PHENOTYPING CANOLA (*Brassica napus* L.) AGRONOMIC TRAITS AND ESTIMATING SEED YIELD USING UNOCCUPIED AERIAL VEHICLE (UAV)-BASED MULTISPECTRAL IMAGERY

A Thesis Submitted to the
College of Graduate and Postdoctoral Studies
In Partial Fulfillment of the Requirements
For the Degree of Doctor of Philosophy
In the Department of Plant Sciences
University of Saskatchewan
Saskatoon, Saskatchewan, Canada

By

Ti Zhang

PERMISSION TO USE

In presenting this thesis/dissertation in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis/dissertation in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis/dissertation work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis/dissertation or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis/dissertation.

Requests for permission to copy or to make other uses of materials in this thesis in whole or part should be addressed to:

Head of the Department of Plant Sciences
College of Agriculture and Bioresources
University of Saskatchewan
51 Campus Drive
Saskatoon, Saskatchewan S7N 5A8, Canada

OR

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan S7N 5C9, Canada

ABSTRACT

Canola (*Brassica napus* L.), as an important oilseed crop, is widely grown in Canada. In plant breeding, there has been great improvement in genetic techniques, but traditional field phenotyping methods are a limitation in breeding genotype selection for new cultivar development. These conventional phenotyping methods are labour-intensive, time-consuming, and can be subjective and destructive. With the development of phenotyping technologies such as unoccupied aerial vehicles (UAVs) and multispectral sensors, desirable phenotypic traits and seed yield can be estimated digitally. In this thesis, the main objective was to develop an efficient and non-destructive method to estimate canola flowering number, flowering layer depth, canopy height, and seed yield using UAV-based multispectral imagery collected during the entire crop season. Canola field experiments were conducted using 56 diverse Brassica genotypes under diverse environments from 2016 to 2018 in central Saskatchewan. A UAV mounted with a multispectral sensor was used for imagery collection. In the flowering number estimation study, the normalized difference yellowness index (NDYI)-based pixels were significantly correlated with actual canola flower numbers with coefficient of determination (R^2) ranging from 0.54 to 0.95 ($p < 0.05$). Moreover, seed yield could be estimated using cumulative NDYI-based pixels extracted from multi-temporal imagery collected during the flowering stage with R^2 up to 0.42 ($p < 0.05$). In the flowering layer depth and canopy height estimation study, canopy height and flowering layer depth could be quantified using a crop surface model generated from UAV-based imagery with R^2 up to 0.90 ($p < 0.05$) and 0.42 ($p < 0.05$), respectively. Furthermore, the cumulative UAV-derived canopy height at the flowering stage and the cumulative UAV-derived flowering layer depth were significantly correlated with seed yield (R^2 up to 0.46 and 0.34, respectively; $p < 0.05$). In the last seed yield estimation study, a machine learning method (i.e., random forest regression model) was used to investigate 30 digitalized input variables including 28 cumulative vegetation indices and 2 cumulative canopy structural phenotypes (i.e., cumulative UAV-derived canopy height at the flowering stage and cumulative flowering layer depth) for yield estimation. The ranking of variable importance by the random forest model indicated that the cumulative blue normalized difference vegetation index at the flowering stage, the cumulative UAV-derived canopy height at the flowering stage, and the cumulative normalized difference vegetation index at the vegetative stage were the most important indicators to estimate seed yield in canola. Generally, all results demonstrated that flowering number and canopy structural feature (i.e., canopy

height and flowering layer depth) could be efficiently assessed by UAV-based imagery. The digital cumulative phenotypes at the vegetative and flowering stages selected by a random forest regression model can assist crop researchers and farmers in early yield estimation and crop management decisions.

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to my supervisor Dr. Steve Shirtliffe, for his continuous encouragement, valuable advice, and guidance through the project. I would like to acknowledge and thank my advisory committee members: Dr. Sally Vail, Dr. Xulin Guo, Dr. Rosalind A. Bueckert, and Dr. Yuguang Bai for their support, advice, and inspiration. I would like to say special thanks to my external examiner Dr. Malcolm Morrison. I would like to acknowledge Mr. Eric N. Johnson for his support and help in editing my thesis and manuscripts. A special acknowledgement goes to Dr. Hema S.N. Duddu, Ms. Menglu Wang, and Mr. Ryu Seungbum for their friendship and assistance in imagery and ground data collection.

I appreciate all technician staff at Kernen Research Farm, Agriculture and Agri-Food Canada Research Farm at the Saskatoon, Melfort, and Scott locations, who provided technical assistance in field operations and taught me practical knowledge of agriculture. Many thanks go to all summer students in the Weed Science lab who helped me with data collection.

I would like to acknowledge the financial support from a Canada First Research Excellence Fund to the Plant Phenotyping and Imaging Research Center (P2IRC), and Global Institute for Food Security (GIFS) at the University of Saskatchewan, SK, Canada. I am grateful for the scholarships including the 2016-2019 Western Grains Research Foundation Endowment Fund Graduate Scholarship, the 2016-2017 Rene Vandeveld Postgraduate Scholarship, and 2015-2016 Chinese Scholarship Council Scholarship.

Finally, I would like to thank my family and friends for their motivation and support. I thank my husband Yichao Shen for his support in taking care of me and our baby. In the end, I thank my mother for coming to Canada to help and support my family.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2D	Two-dimensional
3D	Three-dimensional
AUC	Area under curve
AUDPC	Area under the disease progress curve
AUFPC	Area under the flowering progress curve
BNDVI	Blue normalized difference vegetation index
$CI_{red\ edge}$	Red edge chlorophyll index
CSMs	Crop surface models
DEMs	Digital elevation models
DSMs	Digital surface models
FL	Flowering layer depth
GCP	Ground control point
GNDVI	Green normalized difference vegetation index
GPS	Global positioning system
GSD	Ground sampling distance
HFI	High resolution flowering index
mtry	The number of variables for each split
MYI	Modified yellowness index
NA	No data available
NAM	Nested association mapping
NDRE	Normalized difference red edge index
NDVI	Normalized difference vegetation index
NDYI	Normalized difference yellowness index
NIR	Near IR
ntree	The number of decision trees
r	Correlation coefficient
R^2	Coefficient of determination
RBNI	Red-blue normalizing index
RF	Random forest
RMSE	Root mean square error
RVI	Ratio vegetation index

SfM	Structure-from-motion
SK	Saskatchewan
UAV	Unoccupied aerial vehicles
VARI	Visible atmospherically resistant index
VIP	Variable importance in the projection
VIs	Vegetation indices

CHAPTER 1 INTRODUCTION

1.1 Background

Food security is becoming a global challenge with increasing human population, variable climate conditions, and limited natural resources, necessitating an increase in crop productivity (Furbank and Tester, 2011; Tilman et al., 2011). Although there has been great progress in genotyping technology, plant breeding is currently limited by challenges in field-based phenotyping.

Plant phenotype is defined as observable plant characteristics related to structure, morphology, and physiology (Granier and Vile, 2014). These crop traits are expressed as the interactions between genes and environmental factors (Granier and Vile, 2014). Phenotyping is the process of assessing various plant traits (i.e., plant performance under specific environmental conditions) (Granier and Vile, 2014). Conventional field phenotyping is difficult because it is time consuming and labor intensive in addition to often being destructive and subjective. To efficiently measure association between gene or genome and phenotypes, high throughput phenotyping using remote sensing techniques and advanced mathematical algorithms is a useful tool to qualify and quantify plant traits in crop research (Haghighattalab et al., 2016).

Current improvements in aerial-based platforms and sensors make it possible to efficiently collect phenotypes *via* analysis of digital imagery. Low-cost unoccupied aerial vehicles (UAVs) equipped with various sensors (i.e., cameras) can quickly collect large quantities of field data, which enables plant breeders to efficiently detect crop traits of numerous genotypes in a large-scale field (White et al., 2012; Sankaran et al., 2015). In addition to the efficiency of data collection, UAV-based imagery can provide temporal and novel digital phenotypes related to crop performance over the growing season, which has less damage on the crop canopy and less bias in data collection compared with manual measurement or visual rating.

Aerial-based phenotyping platforms enable higher temporal resolution due to ease of accessing field trials and less requirements for labor and time. Sensors mounted to these platforms allow

crop researchers to gather high spatial and spectral resolution imagery. Application of UAVs with imaging sensors such as commercial RGB cameras and multispectral cameras in agriculture has gained the attentions of many researchers over the last decade (Haghighattalab et al., 2016).

Canola is an important oilseed crop grown in Canada, contributing $29.9 billion to the Canadian economy annually and is primarily grown in western Canada including Saskatchewan, Alberta, and Manitoba (Canola Council of Canada, 2021). Seed yield is one of the most important plant phenotypes in breeding programs, which is controlled by multiple genes contributing small effects and correlated with many growth parameters. Yield prediction models based on the different growth parameters estimated digitally may provide efficient selection criteria in breeding projects by associating yield with its related digital phenotypes under different environmental conditions (Sabaghnia et al., 2010). Field phenotypes, including flowering number, flowering layer depth, and canopy height are potential indicators for yield estimation, which could be assessed by two-dimensional (2D) vegetation indices or canopy structural data extracted from three-dimensional (3D) digital surface models (DSMs) at plot level (Sulik and Long, 2015, 2016; Bendig et al., 2014, 2015).

1.2 Hypotheses and Broad Objectives

1.2.1 Hypothesis:

The general hypothesis is that agronomic traits of canola including flowering number, flowering layer depth (i.e., vertical distance from the bottom flower to the top flower of canopy), canopy height, and seed yield can be quantified or estimated by UAV-based multispectral imagery at reproductive stage. More specifically, one hypothesis is that yellow flower number can be estimated by spectral vegetation index and used for yield estimation. Second, canopy structural information such as flowering layer depth and canopy height can be quantified using 3D DSMs extracted from UAV-based images. The two digitalized canopy structural phenotypes can be indicators for yield estimation. Lastly, seed yield can be estimated using the combination of 2D and 3D imagery information collected at reproductive stage.

1.2.2 Broad Objectives:

In this study, the main objective was to develop an efficient method to phenotype flower number, canopy structural information, and seed yield in canola using UAV-based imagery, which can enhance the genotype selection process in plant breeding programs. More specifically, the first objective was to investigate the relationship between the actual and UAV-based flowering numbers, and to evaluate the potential of the digitalized flowering trait to estimate yield (Chapter 3). The second objective was to quantify crop canopy structural information including flowering layer depth and canopy height using DSMs derived from UAV-based imagery (Chapter 4). The third objective was to develop a descriptive yield model using UAV-based digital imagery with diverse breeding genotypes under various environmental conditions (Chapter 5).

References

Araus, J. L., and Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. Trends in Plant Science, 19, 52-61. doi: 10.1016/j.tplants.2013.09.008

Bendig, J., Bolten, A., Bennertz, S., Broscheit, J., Eichfuss, S., and Bareth, G. (2014). Estimating biomass of barley using crop surface models (CSMs) derived from UAV-based RGB imaging. Remote Sensing, 6, 10395-10412. doi: 10.3390/rs61110395

Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., et al. (2015). Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. International Journal of Applied Earth Observation and Geoinformation, 39, 79-87. doi: 10.1016/j.jag.2015.02.012

Canola Council of Canada. (2021). Growing opportunity for all Canadians. Available online at: https://www.canolacouncil.org/about-canola/economic-impact/ (Accessed December 18, 2021).

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

Granier, C., and Vile, D. (2014). Phenotyping and beyond: modelling the relationships between traits. Current Opinion in Plant Biology, 18, 96-102. doi: 10.1016/j.pbi.2014.02.009

Haghighattalab, A., Pérez, L. G., Mondal, S., Singh, D., Schinstock, D., Rutkoski, J., et al. (2016). Application of unmanned aerial systems for high throughput phenotyping of large wheat breeding nurseries. Plant Methods, 12, 35. doi: 10.1186/s13007-016-0134-6

Sabaghnia, N., Dehghani, H., Alizadeh, B., and Mohghaddam, M. (2010). Interrelationships between seed yield and 20 related traits of 49 canola (*Brassica napus* L.) genotypes in non-stressed and water-stressed environments. Spanish Journal of Agricultural Research, 2, 356-370. doi: 10.5424/sjar/2010082-1195

Sankaran, S., Khot, L. R., Espinoza, C. Z., Jarolmasjed, S., Sathuvalli, V. R., Vandemark, G. J., et al. (2015). Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: a review. European Journal of Agronomy, 70, 112-123. doi: 10.1016/j.eja.2015.07.004

Sulik, J. J., and Long, D. S. (2015). Spectral indices for yellow canola flowers. International Journal of Remote Sensing, 36, 2751-2765. doi: 10.1080/01431161.2015.1047994

Sulik, J. J., and Long, D. S. (2016). Spectral considerations for modeling yield of canola. Remote Sensing of Environment, 184, 161-174. doi: 10.1016/j.rse.2016.06.016

Tilman, D., Balzer, C., Hill, J., and Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. Proceedings of the National Academy of Sciences, 108, 20260-20264. doi: 10.1073/pnas.1116437108

White, J. W., Andrade-Sanchez, P., Gore, M. A., Bronson, K. F., Coffelt, T. A., Conley, M. M., et al. (2012). Field-based phenomics for plant genetics research. Field Crops Research, 133, 101-112. doi: 10.1016/j.fcr.2012.04.003

CHAPTER 2 LITERATURE REVIEW

2.1 Canola Production in Canada and its Use

Canola (*Brassica napus* L., genome AACC, 2n = 38) is an allopolyploid species, deriving from natural hybridization between *Brassica rapa* (genome AA, 2n = 20) and *Brassica oleracea* (genome CC, 2n = 18) (Rana et al., 2004; Li et al., 2014). Canola cultivars commonly grown in Canada belongs to three Brassica species (i.e., *Brassica napus*, *Brassica rapa*, and *Brassica juncea*). The generic term "canola" is derived from "Canadian oil" and refers to rapeseed varieties with low erucic acid (less than 2% of its fatty acid profile) and low glucosinolates (less than 30 micromoles per gram of air-dried, oil-free meal) (Canola Council of Canada, 2021). The nutritional desirability of a high level of erucic acid content in the rapeseed oil was questioned. In addition, glucosinolates of rapeseed meal used in rations led to the pungent odour and taste for livestock consumption, which could reduce feed efficacy (Canola Council of Canada, 2021). The first canola cultivars with low erucic acid and glucosinolates were developed from common rapeseed by Canadian breeders during the1960's and 1970's through traditional plant breeding methods (Stefansson and Gold, 1995).

The advantages of canola as edible oil on human health have been well known for decades. Canola oil, as a healthy vegetable oil, is low in saturated fat and consists of two essential fatty acids (i.e., alpha-linolenic acid and linoleic acid) which cannot be produced by human (Scarth and McVetty, 1999; Lin et al., 2013). Previous studies have shown the health benefits of canola oil in reducing blood cholesterol levels and heart disease risk (Lin et al., 2013). In addition, serving as source for livestock feed, canola meal contains high quality protein (Kalscheur and Moore, 2018). Canola meal has shown advantages in dairy rations with increased milk yield, milk protein yield, and dry matter intake compared with other vegetable protein source such as soybean meal (Martineau et al., 2013; Broderick et al., 2016). Canola is the predominant oilseed crop grown in Canada which remains the world leader in production and exports (Clayton et al., 2000; Morrison et al., 2016). With the growing global demand for canola, it is critical for Canada to maintain and improve canola yield and seed quality to meet the market demands.

2.2 Phenotyping in Plant Breeding

The plant phenotype is defined as the ultimate plant expression of the structure, morphology, and physiology affected by genetic factors, environmental conditions, and their interactions (White et al., 2012; Araus and Cairns, 2014; Granier and Vile, 2014). Phenotying in a breeding program is defined as a process to measure or estimate various plant expressions such as canopy volume, growth rate, physiological characteristics (e.g., leaf area index and flower number, etc.), susceptibility to biotic and abiotic stress, and grain yield (White et al., 2012; Sankaran et al., 2015). Phenomics refers to the research field of improved phenotyping technologies for studying plant growth, performance, and composition (Furbank and Tester, 2011; White et al., 2012; Sankaran et al., 2015)

To solve the global food security problem, crop production must double by 2050 to satisfy the food demand of increasing populations, unstable climate conditions, and limited natural resources (Furbank and Tester, 2011; Tilman et al., 2011). Over the last decade, many efforts have been made to improve production practices to increase crop productivity (Sankaran et al., 2015). Additionally, there has been great progress in genetic and genomic technologies for breeding programs, which rapidly reduced the costs and time of genotyping (White et al., 2012). During the breeding process, field-based phenotyping plays an important role in evaluating plant performance, which contributes to selection of ideal genotypes that are high yielding and stress tolerant by associating genotype with corresponding phenotype (Montes et al., 2007; Furbank and Tester, 2011; Sankaran et al., 2015). To select for better crop genotypes and develop new varieties with desired traits, plant breeders need to assess many distinct genotypes grown in multiple environments to detect interactions between genotype and environment (White et al., 2012; Araus and Cairns, 2014). However, the current methods of phenotyping crop samples are very laborious requiring manual collection of phenotypic data from numerous genotypes in the field such as flowering traits and plant height (Araus and Cairns, 2014; Shi et al., 2016). In general, the conventional methods are labor intensive, time consuming, expensive, destructive, and subjective (e.g., scoring scale of crop growth conditions) in large breeding trials (Araus and Cairns, 2014; Sankaran et al., 2015).

By recognizing the limitation in existing phenotyping approaches, developing cost and time efficient techniques must be pursued to accelerate the breeding program. High-throughput

phenotyping uses various sensors integrated with autonomous or semi-autonomous platforms to collect phenotypic data across different genotypes in field trials (Shi et al., 2016). This novel data collection technique can provide more rapid and accurate phenotypic data collection with less labor efforts compared with conventional phenotyping methods, which allow more genotypes entering the breeding workflow (Araus and Cairns, 2014; Shi et al., 2016).

2.3 Remote Sensing Techniques in High-throughput Phenotyping

2.3.1 Remote Sensing Techniques

Remote sensing is defined as a method of collecting data or information *via* platforms equipped with sensors without physically touching subjects (Araus and Cairns, 2014; Wójtowicz et al., 2016). The information carrier in remote sensing is electromagnetic radiation reflected or emitted by objects, which can be collected by remote sensing sensors (Araus and Cairns, 2014; Wójtowicz et al., 2016). Regarding plants, the amount of radiation reflected or emitted is influenced by leaf pigments, cell structure, and water content (Kumar and Silva, 1973). The most widely used wavelengths in remote sensing are visible light (400 to 700 nm), near infrared (700 to 1,100 nm), shortwave infrared (1,100 to 3,000 nm), thermal infrared (7,000 to 14,000 nm), and microwave bands (1 mm to 1 m) (Araus and Cairns, 2014; Wójtowicz et al., 2016).

By correlating the radiation information with plant phenotypes, many categories of plant traits can be measured such as flowering traits, plant height, leaf area index, green biomass, and yield prediction through non-invasive approaches. Therefore, the remote sensing techniques can be used for field crop phenotyping, which are repeatable, accurate, non-destructive and time-efficient methods to gathering the information generated through radiation reflected or emitted by crop canopy (Araus and Cairns, 2014).

2.3.2 Platforms in Remote Sensing for Field Phenotyping

Available remote sensing platforms for field phenotyping range from ground-based to aerial-based systems including ground vehicles, hand-held instruments, satellites, aircraft, helicopters,

and unoccupied aerial vehicles (UAVs) (White et al., 2012; Sankaran et al., 2015; Haghighattalab et al., 2016; Wójtowicz et al., 2016).

Ground-based remote sensing has been applied in field experiments to investigate plant responses to biotic and abiotic stress (Jackson, 1986; White et al., 2012; Wójtowicz et al., 2016). Sensors mounted on the ground-based platforms are expected to provide higher spatial and spectral resolutions with all measurements at a shorter distance compared to aerial-based systems, which is useful in relatively small-scale field trials (Jackson, 1986; White et al., 2012; Wójtowicz et al., 2016). However, low portability and long transporting time for the ground-based platforms can preclude the utilization at multiple experimental locations (Haghighattalab et al., 2016).

Compared with ground-based platforms, aerial-based remote sensing can rapidly record imagery of entire field and have no requirements for good soil conditions to enter. It is more practical for multiple large experimental areas and various environmental conditions. In recent years, small, unoccupied aerial vehicles (UAVs) with the aircraft remotely piloted from the ground station, have been developed for use in agriculture research (Shi et al., 2016; Wójtowicz et al., 2016). Compared with other aerial-based platforms such as satellite, small UAVs have gained more attention in agriculture projects because of the relatively low operation cost, user-friendly and accurate navigation planning flight control, flexibility in flight height, opportunity for more frequent digital data collection, and high-resolution images (Sankaran et al., 2015; Haghighattalab et al., 2016; Shi et al., 2016; Wójtowicz et al., 2016).

UAVs can be classified into two groups: fixed-wing and rotary-wing UAVs (Sankaran et al., 2015; Shi et al., 2016; Wójtowicz et al., 2016). Selection of UAVs mainly depends on experimental objectives (Sankaran et al., 2015; Wójtowicz et al., 2016). Fixed-wing UAVs have higher flight speed and height, which enables the capture of a wider range of field areas compared to rotary-wing UAVs (Shi et al., 2016; Wójtowicz et al., 2016). In addition, fixed-wing UAVs allow heavier sensors and longer flight duration (Shi et al., 2016; Wójtowicz et al., 2016). On the other hand, rotary-wing UAVs can take off and land vertically without runways and have better hovering capability over the target field areas (Sankaran et al., 2015; Wójtowicz et al., 2016).

2.3.3 Sensors Used for Small UAV Platforms

Different sensors have been developed for field phenotyping mounted to the small UAV platforms (Araus and Cairns, 2014; Sankaran et al., 2015; Shi et al., 2016; Wójtowicz et al., 2016). These sensors are lightweight to meet the limited payload lift capability of small UAVs (Sankaran et al., 2015; Shi et al., 2016). In agriculture research, the most widely used sensors for small UAVs are RGB (red, green and blue wavelength bands from 400 to 700 nm), multispectral (i.e., 3 to 6 spectral bands from 400 to 1,000 nm), and thermal sensors (commonly from 7,000 to 14,000 nm), which are able to detect the wavelength of spectral radiation and the amount of energy (Zhang and Kovacs, 2012; Sankaran et al., 2015; Shi et al., 2016; Wójtowicz et al., 2016). Apart from these sensors, hyperspectral sensors (10s to 100s of continuous spectral bands) and LiDAR (light detection and ranging) emitting its own pulsed light to detect objects, have promising potential to be used for small UAVs (Sankaran et al., 2015; Shi et al., 2016). Hyperspectral sensors can provide a wide spectral range but require development in lowering its weight to meet the maximum payload lift capabilities of UAVs (Shi et al., 2016). LiDAR can be used to collect morphological data by penetrating through crop canopy and collecting internal canopy structure and canopy density (Sankaran et al., 2015; Shi et al., 2016).

Spatial, spectral and temporal resolutions are important factors to be considered when applying UAVs for research. Spatial resolution is narrowly defined as pixel size while broadly defined as the total area size that aerial images can cover (Wójtowicz et al., 2016). Spectral resolution means the ability to define distinct wavelength intervals or the width of wavelength bands (i.e., the higher spectral resolution, the narrower the wavelength range for a specific band) (Wójtowicz et al., 2016). Usually, temporal resolution indicates the frequency of data collection while the broad sense of temporal resolution refers to the earliest date when you can find historical data from saved images taken by imaging sensors.

2.4 Application of UAV in High-throughput Phenotyping

2.4.1 Mechanisms of Crop Information (Canopy Reflectance and Canopy Structure) Extraction from UAV-based Images

UAV-based phenotyping is mainly based on the characteristics of spectral radiation reflected and the amount of light energy emitted by plants, which are recorded by remote sensors (Christenson et al., 2016; Wójtowicz et al., 2016). Generally, the amount of spectral reflectance of plants is correlated to morphological and physiological traits due to leaf pigments, cell structure, and water content (Kumar and Silva, 1973); thus, ratios of different bands in the visual light, near and shortwave infrared wavelengths, named as vegetation indices (VIs), were used to characterize plant traits (Sankaran et al., 2015; Christenson et al., 2016; Wójtowicz et al., 2016). The most popular and widely used vegetation index is the normalized difference vegetation index (NDVI) (Equation 2.1) calculated as a ratio of difference over sum of the near infrared and red bands, which was first used by Rouse et al. (1974) in Great Plains study. NDVI is commonly used to identify vegetation area, plant health status, and biomass for agriculture research.

$$\boldsymbol{NDVI} = \left(\frac{R_{NIR} - R_{red}}{R_{NIR} + R_{red}}\right) \quad [2.1]$$

Where R_{NIR} represents the reflectance value of near infrared wavelength (700 - 1,100 nm), and R_{red} represents the reflectance value of red wavelength (600 - 700 nm).

Aside from VIs calculated based on canopy reflectance, canopy structure information can be captured *via* building 3-dimensional (3D) crop surface models (CSMs) using structure-from-motion (SfM) photogrammetry, which allows the measurements of canopy height, plant biomass, and lodging condition (Westoby et al., 2012; Shi et al., 2016; Malambo et al., 2018). SfM is a method providing high-resolution digital evaluation models (DEMs) from overlapping images (Westoby et al., 2012; Malambo et al., 2018). This technique applies bundle adjustment based on image matching features, which has been used for characterizing canopy height in various research such as barley (Bendig et al., 2014), wheat (Holman et al., 2016; Panday et

al., 2020), corn (Shi et al., 2016; Malambo et al., 2018), and sorghum (Shi et al., 2016; Malambo et al., 2018).

2.4.2 Crop Traits Evaluation Using Imagery

Many crop traits can be evaluated by high-throughput phenotyping method such as emergence rate, canopy cover, flowering number, canopy height, chlorophyll content, leaf area index, biomass, plant maturity, and yield (White et al., 2012; Araus and Cairns, 2014; Wójtowicz et al., 2016). Among those phenotypes, flowering number is an important phenotype positively correlated with seed yield in canola (Tayo and Morgan, 1975; Diepenbrock, 2000; Kirkegaard et al., 2018). Previous research has investigated flowering using imagery-based method in many crops (Guo et al., 2015; Sulik and Long, 2015, 2016; Fang et al., 2016; Wan et al., 2018; Xu et al., 2018; d'Andrimont et al., 2020; Han et al., 2021; Zhang et al., 2021b). For example, in canola, Sulik and Long (2016) reported the normalized difference yellowness index (NDYI) could effectively capture the yellowness of canola flower and estimate county-level yield with a coefficient of determination (R^2) value of 0.76. Multiple authors (d'Andrimont et al., 2020; Han et al., 2021; Zhang et al., 2021b) found NDYI or NDYI-based pixel could successfully detect canola flower and determine the peak flowering timing using satellite or multispectral imagery with moderate to high estimation accuracy. Wan et al. (2018) and Gong et al. (2018) found the combination of vegetation index and classification methods can improve canola flower detection. In addition, Guo et al. (2015) estimated flowering panicles in rice using the machine learning method (i.e., support vector machine) and found a significant correlation between the visually measured and image-based flowering panicles with correlation coefficient (r = 0.82). Xu et al. (2018) applied convolutional neural network to count cotton bloom from UAV-based color imagery with less than 10% estimation error.

Canopy height is another important crop trait in agriculture research, which is usually related to crop lodging, mechanical harvest, and yield estimation. This phenotype has been studied using imagery for biomass or yield estimation in various crops such as barley (Bendig et al., 2014), corn (Geipel et al., 2014), poppy crop (*Papaver somniferum* L.) (Iqbal et al., 2017), cotton (Feng et al., 2019), wheat (Panday et al., 2020), and cool season crops including chickpea, dry pea, camelina, and canola (Zhang et al., 2021a). In a barley field trial conducted in Germany, estimated canopy height extracted from the CSMs generated by UAV-based 3D

geometry information was significantly regressed with ground reference height and dry biomass (R^2 = 0.92 and 0.82, respectively) (Bendig et al., 2014). A poppy crop study demonstrated the digital canopy height could be reliably estimated using CSMs developed from RGB imagery with R^2 ranging from 0.93 to 0.97 (Iqbal et al., 2017). They also stated that the digital canopy height is a promising indicator of poppy capsule volume (R^2 = 0.74). Feng et al. (2019) applied UAV-based imaging method to estimate canopy height in cotton using the difference in DEMs between crop canopy and soil surface with less than 10% estimation error. This study reported a significant correlation between the digitalized canopy height and cotton yield with *r* ranging from 0.66 to 0.96. Similarly, low estimation error (5% to 11.9%) was observed when comparing the actual and estimated canopy heights in wheat using CSMs derived from RGB images (Panday et al., 2020). There were significant relationships between the estimated canopy height and above ground biomass and grain weights (R^2 = 0.66 and 0.70, respectively). Zhang et al. (2021a) reported significant correlations between actual and estimated canopy heights using CSMs extracted from UAV-based RGB imagery in wheat, chickpea, dry pea, camelina, and canola.

2.4.3 Yield Estimation Using UAV-based Imagery and The Machine Learning Method

UAVs with various sensors have been used for yield or biomass estimation in many crops such as canola, rice, wheat, barley, corn, and cotton (Swain and Zaman, 2012; Bendig et al., 2014, 2015; Johnson et al., 2016; Sulik and Long, 2016; Gong et al., 2018; Peng et al., 2019; Feng et al., 2020; Ramos et al., 2020; Zhang et al., 2021b). For example, canola seed yield was predicted by NDYI in a linear regression model, proposed by Sulik and Long (2016). They stated that there were significant correlations between NDYI and seed yield with higher R^2 values than NDVI at the peak flowering time. NDVI performed better in yield estimation at the vegetative stage by capturing greenness. But NDYI could catch more yellowness of flowering for yield prediction during the flowering period. They suggested that NDYI collected at the peak flowering time could provide more informative data for yield prediction because early flowering contributes more pods to final yield than late flowering. A field study conducted by Swain and Zaman (2012) investigated the application of canopy reflectance for estimating rice yield. The study illustrated that the NDVI extracted from UAV-based multispectral images collected at 65 days after planting (i.e., booting stage) had a linear relationship with yield (R^2 = 0.76) (Swain and Zaman, 2012).

Compared with using crop information obtained from a single image date to estimate biomass or yield, multi-temporal imagery over a growing season may provide more details of the dynamics of crop growth status and progress (Xue et al., 2007; Wang et al., 2014; Zhou et al., 2017; Malambo et al., 2018; Zhang et al., 2021b). Xue et al. (2007) found the cumulative vegetation indices, accumulated from heading to maturity, had the most robust relationship with grain yield in wheat. Similar results were found by Wang et al. (2014) who stated that the cumulative NDVI and ratio vegetation index (RVI) observed from jointing to filling stages improved the yield estimation in wheat compared to the indices obtained from a single image date. Zhou et al. (2017) reported the cumulative NDVI and visible atmospherically resistant index (VARI) from booting to the heading stage had better performance for rice yield estimation with a R^2 value of 0.75 and 0.72, respectively than when comparing with a single image.

Crop yield is a complex crop trait affected by genotype, environments, and their interactions. Therefore, simply applying a single input variable in a yield or biomass prediction model may result in inaccurate results because of the potential saturation of VIs at a specific growth stage, which neglects important growth condition and information during the remaining growing season (Geipel et al., 2014; Liu et al., 2019; Feng et al., 2020; Wan et al., 2020). Geipel et al. (2014) applied a multiple linear regression model including several VIs representing crop coverage and digital crop heights computed from CSMs to estimate corn yield. They reported that the multiple linear regression model improved the accuracy of yield estimation compared with a simple linear regression model (Geipel et al., 2014). Liu et al. (2019) stated that the combination of NDVI with a texture metrics (i.e., contrast) improved the accuracy of biomass prediction compared with a single VI, and this combination had a great potential for biomass estimation in canola. Feng et al. (2020) found that yield prediction accuracy in cotton increased when DSM-based plant height, cotton fibre index, canopy temperature and canopy colour space were considered in the yield estimation model compared to one image feature. Wan et al. (2020) indicated the combination of NDVI, NDYI, canopy height and canopy coverage greatly improved the prediction results of yield in rice compared to using a single variable.

Within the last decade, various modeling algorithms have been applied to estimate yield or biomass using imagery-based phenotypes as independent variables (Wang et al., 2016; Liu et

al., 2019; Feng et al., 2020; Wan et al., 2020). Random forest (RF) is one of the most used models to build a descriptive yield model. It is an ensemble machine learning method for classification and regression determination. For the regression purposes, it can efficiently handle many independent variables, has less sensitivity to outliers and less over-fitting issues (Wang et al., 2016; Liu et al., 2019). In addition, the RF regression model needs fewer model hyperparameters (i.e., number of decision trees and number of variables for each split) compared with other machine learning methods such as artificial neural network and support vector regression (Wang et al., 2016; Liu et al., 2019). A RF regression model creates many randomized and independent regression trees (i.e., decision trees) which forms the "forest". Each decision tree is trained by randomly selected a subset of training samples and variables using a bootstrap aggregating algorithm. Bootstrapping aggregating is an ensemble technique which chooses a random sample from the original dataset. Then the RF regression model combines the multiple decision trees and takes the average of all the outputs to develop an optimal ensemble model with accurate and stable prediction. It is important to determine two model hyperparameters including the number of decision trees (*ntree*) and the number of variables at each tree (*mtry*). For the RF regression model, the default values for the *ntree* and *mtry* are 500 trees and 1/3 of total number of variables, respectively (Wang et al., 2016).

2.5 Pre-processing UAV-based Images (Geometric and Radiometric Calibration)

Large quantities of overlapped images covering the whole experimental field can be obtained by small UAVs. Prior to analyzing the image of study area, a critical task is combing all individual overlapped images into a complete field map of the field using a process of mosaicking in professional image software (e.g., Pix4D pro mapper and PhotoScan) (Torres-Sánchez et al., 2014). The image software will search for common points in the overlapped images and match the overflown area. However, commonly used image sensors mounted to UAV platforms have limitations in image geometry because of geometric distortion problems of consumer-grade camera, which may result in inaccurate size of object (Haghighattalab et al., 2016). Therefore, geometric calibration prior to analyzing images is critical (Torres-Sánchez et al., 2014; Haghighattalab et al., 2016). According to the positions of ground geo-reference points (i.e., ground control points), image geometry can be constructed, and the individual aerial images are projected over it for orthophoto process (Torres-Sánchez et al., 2014;

Haghighattable et al., 2016). Global positioning system (GPS) measures geo-location for each ground control point (GCP) (Haghighattable et al., 2016). Although lens distortion is another internal problem, this issue is expected to be compensated by the image sensor itself.

Ortho-mosaic images also need radiometric calibration because of unstable illumination intensity in the field resulting from shading from cloud as well as atmospheric absorption differing within and among imaging dates (Haghighattable et al., 2016). Gray-scale or white calibration panels can be used for calibration to correct for this (Wang et al., 2015; Haghighattable et al., 2016). Some software (e.g., Pix4D pro mapper) can semi-automatically calibrate the radiometric and geometric issues of images, which improve the precision of following image analysis (Torres-Sánchez et al., 2014; Haghighattalab et al., 2016).

2.6 Challenges in Field-based High-throughput Phenotyping Implementation

Collecting raw phenotyping data is not difficult with advances in UAVs platforms and sensors; however, subsequent processing and mining "big phenotyping data" sets remain challenging (White et al. 2012; Araus and Cairns, 2014). More specifically, combining and comparing the imagery collected by different sensors requires imagery processing standardization by applying efficient image software tools (White et al. 2012; Araus and Cairns, 2014). In addition, properly understanding and effectively using large volumes of phenotyping data needs advanced algorithms and modeling for incorporating genotypic, phenotypic, and environmental data (White et al. 2012; Araus and Cairns, 2014). Otherwise, extensive calculations or inaccurate modeling may limit real-time information of plants, which may hinder breeders and farmers from predicting or making accurate decisions for genotypes selection and crop management (White et al. 2012; Araus and Cairns, 2014; Haghighattable et al., 2016). In conclusion, automatic and rapid imaging processing and advanced statistical tools need to be improved to drive forward the application of UAV-based imagery in agriculture field (White et al. 2012; Araus and Cairns, 2014; Haghighattable et al., 2016).

References

Araus, J. L., and Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. Trends in Plant Science, 19, 52-61. doi: 10.1016/j.tplants.2013.09.008

Bendig, J., Bolten, A., Bennertz, S., Broscheit, J., Eichfuss, S., and Bareth, G. (2014). Estimating biomass of barley using crop surface models (CSMs) derived from UAV-based RGB imaging. Remote Sensing, 6, 10395-10412. doi: 10.3390/rs61110395

Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., et al. (2015). Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. International Journal of Applied Earth Observation and Geoinformation, 39, 79-87. doi: 10.1016/j.jag.2015.02.012

Broderick, G. A., Colombini, S., Costa, S., Karsli, M. A., and Faciola, A. P. (2016). Chemical and ruminal in vitro evaluation of Canadian canola meals produced over 4 years. Journal of Dairy Science, 99, 7956-7970. doi: 10.3168/jds.2016-11000

Canola Council of Canada. (2021). History of canola seed development. Available online at: https://www.canolacouncil.org/canola-encyclopedia/history-of-canola-seed-development/#footnote_1_2301 (Accessed February 22, 2021).

Christenson, B. S., Schapaugh Jr, W. T., An, N., Price, K. P., Prasad, V., and Fritz, A. K. (2016). Predicting soybean relative maturity and seed yield using canopy reflectance. Crop Science, 56, 625-643. doi: 10.2135/cropsci2015.04.0237

Clayton, G., Turkington, K., Harker, N., O'Donovan, J., and Johnston, A. (2000). High yielding canola production. Better Crops, 84. 26-27. Available online at: http://www.ipni.net/publication/bettercrops.nsf/0/405DA9F381C7D59F852579800081FDE7/$FILE/Better%20Crops%202000-1%20p26.pdf (accessed January 05, 2020).

d'Andrimont, R., Taymans, M., Lemoine, G., Ceglar, A., Yordanov, M., and van der Velde, M. (2020). Detecting flowering phenology in oil seed rape parcels with Sentinel-1 and-2 time series. Remote Sensing of Environment, 239, 1-14. doi: 10.1016/j.rse.2020.111660

Diepenbrock, W. (2000). Yield analysis of winter oilseed rape (*Brassica napus* L.): a review. Field Crops Research, 67, 35-49. doi: 10.1016/S0378-4290(00)00082-4

Fang, S., Tang, W., Peng, Y., Gong, Y., Dai, C., Chai, R., et al. (2016). Remote estimation of vegetation fraction and flower fraction in oilseed rape with unmanned aerial vehicle data. Remote Sensing, 8, 416. doi: 10.3390/rs8050416

Feng, A., Zhang, M., Sudduth, K. A., Vories, E. D., and Zhou, J. (2019). Cotton yield estimation from UAV-based plant height. Transactions of the ASABE, 62, 393-404. doi: 10.13031/trans.13067

Feng, A., Zhou, J., Vories, E. D., Sudduth, K. A., and Zhang, M. (2020). Yield estimation in cotton using UAV-based multi-sensor imagery. Biosystems Engineering, 193, 101-114. doi: 10.1016/j.biosystemseng.2020.02.014

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

Geipel, J., Link, J., and Claupein, W. (2014). Combined spectral and spatial modeling of corn yield based on aerial images and crop surface models acquired with an unmanned aircraft system. Remote Sensing, 6, 10335-10355. doi: 10.3390/rs61110335

Gong, Y., Duan, B., Fang, S., Zhu, R., Wu, X., Ma, Y., et al. (2018). Remote estimation of rapeseed yield with unmanned aerial vehicle (UAV) imaging and spectral mixture analysis. Plant Methods, 14, 1-14. doi: 10.1186/s13007-018-0338-z

Granier, C., and Vile, D. (2014). Phenotyping and beyond: modelling the relationships between traits. Current Opinion in Plant Biology, 18, 96-102. doi: 10.1016/j.pbi.2014.02.009

Guo, W., Fukatsu, T., and Ninomiya, S. (2015). Automated characterization of flowering dynamics in rice using field-acquired time-series RGB images. Plant Methods, 11, 1-15. doi: 10.1186/s13007-015-0047-9

Haghighattalab, A., Pérez, L. G., Mondal, S., Singh, D., Schinstock, D., Rutkoski, J., et al. (2016). Application of unmanned aerial systems for high throughput phenotyping of large wheat breeding nurseries. Plant Methods, 12, 35. doi: 10.1186/s13007-016-0134-6

Han, J., Zhang, Z., and Cao, J. (2021). Developing a new method to identify flowering dynamics of rapeseed using landsat 8 and sentinel-1/2. Remote Sensing, 13, 105. doi: 10.3390/rs13010105

Holman, F. H., Riche, A. B., Michalski, A., Castle, M., Wooster, M. J., and Hawkesford, M. J. (2016). High throughput field phenotyping of wheat plant height and growth rate in field plot trials using UAV based remote sensing. Remote Sensing, 8, 1031. doi: 10.3390/rs8121031

Iqbal, F., Lucieer, A., Barry, K., and Wells, R. (2017). Poppy crop height and capsule volume estimation from a single UAS flight. Remote Sensing, 9, 647. doi: 10.3390/rs9070647

Jackson, R. D. (1986). Remote sensing of biotic and abiotic plant stress. Annual Review of Phytopathology, 24, 265-287. doi: 10.1146/annurev.py.24.090186.001405

Johnson, J. B. and Jimmerson, J., (2003). Canola and Rapeseed. Briefing No. 60, Agricultural Marketing Policy Center, Montana State University, Bozeman, MT.

Johnson, M. D., Hsieh, W. W., Cannon, A. J., Davidson, A., and Bédard, F. (2016). Crop yield forecasting on the Canadian Prairies by remotely sensed vegetation indices and machine learning methods. Agricultural and Forest Meteorology, 218, 74-84. doi: 10.1016/j.agrformet.2015.11.003

Kalscheur, K. F., and Moore, S. A. (2018). Canola meal, a proven advantage in various diet formulations. In Symposium Proceedings, 145-148.

Kirkegaard, J. A., Lilley, J. M., Brill, R. D., Ware, A. H., and Walela, C. K. (2018). The critical period for yield and quality determination in canola (*Brassica napus* L.). Field Crops Research, 222, 180-188. doi: 10.1016/j.fcr.2018.03.018

Kumar, R., and Silva, L. (1973). Light ray tracing through a leaf cross section. Applied Optics, 12, 2950-2954. doi: 10.1364/AO.12.002950

Li, F., Chen, B., Xu, K., Wu, J., Song, W., Bancroft, I., et al. (2014). Genome-wide association study dissects the genetic architecture of seed weight and seed quality in rapeseed (*Brassica napus* L.). DNA Research, 21, 355-367. doi: 10.1093/dnares/dsu002

Lin, L., Allemekinders, H., Dansby, A., Campbell, L., Durance-Tod, S., Berger, A., et al. (2013). Evidence of health benefits of canola oil. Nutrition Reviews. 71, 370–385. doi: 10.1111/nure.12033

Liu, Y., Liu, S., Li, J., Guo, X., Wang, S., and Lu, J. (2019). Estimating biomass of winter oilseed rape using vegetation indices and texture metrics derived from UAV multispectral images. Computers and Electronics in Agriculture, 166, 105026. doi: 10.1016/j.compag.2019.105026

Malambo, L., Popescu, S. C., Murray, S. C., Putman, E., Pugh, N. A., Horne, D. W., et al. (2018). Multitemporal field-based plant height estimation using 3D point clouds generated from small unmanned aerial systems high-resolution imagery. International Journal of Applied Earth Observation and Geoinformation, 64, 31-42. doi: 10.1016/j.jag.2017.08.014

Martineau, R., Ouellet, D. R., and Lapierre, H. (2013). Feeding canola meal to dairy cows: a meta-analysis on lactational responses. Journal of Dairy Science. 96, 1701- 1714. doi: 10.3168/jds.2012-5740

Montes, J. M., Melchinger, A. E., and Reif, J. C. (2007). Novel throughput phenotyping platforms in plant genetic studies. Trends in Plant Science, 12, 433-436. doi: 10.1016/j.tplants.2007.08.006

Morrison, M. J., Harker, K. N., Blackshaw, R. E., Holzapfel, C. J., and O'donovan, J. T. (2016). Canola yield improvement on the Canadian Prairies from 2000 to 2013. Crop and Pasture Science, 67, 245-252. doi: https://doi.org/10.1071/CP15348

Panday, U. S., Shrestha, N., Maharjan, S., Pratihast, A. K., Shrestha, K. L., and Aryal, J. (2020). Correlating the plant height of wheat with above-ground biomass and crop yield using drone imagery and crop surface model, a case study from Nepal. Drones, 4, 28. doi: 10.3390/drones4030028

Peng, Y., Li, Y., Dai, C., Fang, S., Gong, Y., Wu, X., et al. (2019). Remote prediction of yield based on LAI estimation in oilseed rape under different planting methods and nitrogen fertilizer applications. Agricultural and Forest Meteorology, 271, 116-125. doi: 10.1016/j.agrformet.2019.02.032

Ramos, A. P. M., Osco, L. P., Furuya, D. E. G., Gonçalves, W. N., Santana, D. C., Teodoro, L. P. R., et al. (2020). A random forest ranking approach to predict yield in maize with UAV-based vegetation spectral indices. Computers and Electronics in Agriculture, 178, 10579. doi: 10.1016/j.compag.2020.105791

Rana, D., Boogaart, T., O'Neill, C. M., Hynes, L., Bent, E., Macpherson, L., et al. (2004). Conservation of the microstructure of genome segments in *Brassica napus* and its diploid relatives. The Plant Journal, 40, 725-733. doi: 10.1111/j.1365-313X.2004.02244.x

Rouse, J.W, Haas, R.H., Scheel, J.A., and Deering, D.W. (1974) Monitoring vegetation systems in the Great Plains with ERTS. NASA/GSFC Type III Final Report. Greenbelt, MD, 317.

Sankaran, S., Khot, L. R., Espinoza, C. Z., Jarolmasjed, S., Sathuvalli, V. R., Vandemark, G. J., et al. (2015). Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: a review. European Journal of Agronomy, 70, 112-123. doi: 10.1016/j.eja.2015.07.004

Scarth, R., and McVetty, P. B. (1999). Designer oil canola–a review of new food-grade Brassica oils with focus on high oleic, low linolenic types. Processing of the 10th International Rapeseed Congress, Canberra, Australia. Available online at: http://www.regional.org.au/au/gcirc/4/57.htm (accessed January 10, 2020).

Shi, Y., Thomasson, J. A., Murray, S. C., Pugh, N. A., Rooney, W. L., Shafian, S., et al. (2016). Unmanned aerial vehicles for high-throughput phenotyping and agronomic research. PLOS One, 11, e0159781. doi: 10.1371/journal.pone.0159781

Stefansson, B., and Gold, R. D. (1995). “Rapeseed,” in Harvest of gold: the history of field crop breeding in Canada, eds. A. E. Slinkard and D. R. Knott (University of Saskatchewan, Saskatoon, Canada), 140–152.

Sulik, J. J., and Long, D. S. (2015). Spectral indices for yellow canola flowers. International Journal of Remote Sensing, 36, 2751-2765. doi: 10.1080/01431161.2015.1047994

Sulik, J. J., and Long, D. S. (2016). Spectral considerations for modeling yield of canola. Remote Sensing of Environment, 184, 161-174. doi: 10.1016/j.rse.2016.06.016

Swain, K. C., and Zaman, Q. U. (2012). “Rice crop monitoring with unmanned helicopter remote sensing images,” In Remote sensing of biomass-principles and applications, eds. L. Fatoyinbo (InTech: Rijeka), 253-272.

Tayo, T. O., and Morgan, D. G. (1975). Quantitative analysis of the growth, development and distribution of flowers and pods in oil seed rape (*Brassica napus* L.). The Journal of Agricultural Science, 85, 103-110. doi: 10.1017/S0021859600053466

Tilman, D., Balzer, C., Hill, J., and Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. Proceedings of the National Academy of Sciences, 108, 20260-20264. doi: 10.1073/pnas.1116437108

Torres-Sánchez, J., Peña, J. M., De Castro, A. I., and López-Granados, F. (2014). Multi-temporal mapping of the vegetation fraction in early-season wheat fields using images from UAV. Computers and Electronics in Agriculture, 103, 104-113. doi: 10.1016/j.compag.2014.02.009

Wan, L., Cen, H., Zhu, J., Zhang, J., Zhu, Y., Sun, D., et al. (2020). Grain yield prediction of rice using multi-temporal UAV-based RGB and multispectral images and model transfer–a case study of small farmlands in the South of China. Agricultural and Forest Meteorology, 291, 108096. doi: 10.1016/j.agrformet.2020.108096

Wan, L., Li, Y., Cen, H., Zhu, J., Yin, W., Wu, W., et al. (2018). Combining UAV-based vegetation indices and image classification to estimate flower number in oilseed rape. Remote Sensing, 10, 1484. doi: 10.3390/rs10091484

Wang, C., and Myint, S. W. (2015). A simplified empirical line method of radiometric calibration for small unmanned aircraft systems-based remote sensing. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8, 1876-1885. doi: 10.1109/JSTARS.2015.2422716

Wang, L., Tian, Y., Yao, X., Zhu, Y., and Cao, W. (2014). Predicting grain yield and protein content in wheat by fusing multi-sensor and multi-temporal remote-sensing images. Field Crops Research, 164, 178-188. doi: 10.1016/j.fcr.2014.05.001

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

Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., and Reynolds, J. M. (2012). 'Structure-from-Motion'photogrammetry: a low-cost, effective tool for geoscience applications. Geomorphology, 179, 300-314. doi: 10.1016/j.geomorph.2012.08.021

White, J. W., Andrade-Sanchez, P., Gore, M. A., Bronson, K. F., Coffelt, T. A., Conley, M. M., et al. (2012). Field-based phenomics for plant genetics research. Field Crops Research, 133, 101-112. doi: 10.1016/j.fcr.2012.04.003

Wójtowicz, M., Wójtowicz, A., and Piekarczyk, J. (2016). Application of remote sensing methods in agriculture. Communications in Biometry and Crop Science, 11, 31-50.

Xu, R., Li, C., Paterson, A. H., Jiang, Y., Sun, S., and Robertson, J. S. (2018). Aerial images and convolutional neural network for cotton bloom detection. Frontiers in Plant Science, 8, 2235. doi: 10.3389/fpls.2017.02235

Xue, L., Cao, W., and Yang, L. (2007). Predicting grain yield and protein content in winter wheat at different N supply levels using canopy reflectance spectra. Pedosphere, 17, 646–653. doi: 10.1016/S1002-0160(07)60077-0

Zhang, C., and Kovacs, J. M. (2012). The application of small unmanned aerial systems for precision agriculture: a review. Precision Agriculture, 13, 693-712. doi: 10.1007/s11119-012-9274-5

Zhang, C., Craine, W. A., McGee, R. J., Vandemark, G. J., Davis, J. B., Brown, J., et al. (2021a). High-throughput phenotyping of canopy height in cool-season crops using sensing techniques. Agronomy Journal, 113, 3269-3280. doi: 10.1002/agj2.20632

Zhang, T., Vail, S., Duddu, H. S., Parkin, I. A., Guo, X., Johnson, E. N., et al. (2021b). Phenotyping flowering in canola (*Brassica napus* L.) and estimating seed yield using an unmanned aerial vehicle-based imagery. Frontiers in Plant Science, 12, 1178. doi: 10.3389/fpls.2021.686332

Zhou, X., Zheng, H. B., Xu, X. Q., He, J. Y., Ge, X. K., Yao, X., et al. (2017). Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 130, 246-255. doi: 10.1016/j.isprsjprs.2017.05.003

CHAPTER 3 PHENOTYPING FLOWERING IN CANOLA (*Brassica napus* L.) AND ESTIMATING SEED YIELD USING AN UNOCCUPIED AERIAL VEHICLE (UAV)-BASED IMAGERY

The content of this chapter has been published in Frontiers in Plant Science: Zhang, T., Vail, S., Duddu, H. S., Parkin, I. A., Guo, X., Johnson, E. N., and Shirtliffe, S. J. (2021). Phenotyping flowering in canola (*Brassica napus* L.) and estimating seed yield using an unmanned aerial vehicle-based imagery. Frontiers in Plant Science, 12, 1178. doi: 10.3389/fpls.2021.686332

This chapter has been reformatted from the published version of the manuscript based on the requirements of manuscript-style thesis and dissertations.

Abstract

Phenotyping traits related to crop performance is critical for genotype selection and variety development in plant breeding. Canola (*Brassica napus* L.) flowers semi-indeterminately and the bright yellow flowers accumulate over a relatively long period. Flower production of canola plays an important role in yield determination. Yellowness of canola petals may have a critical reflectance signal and can be a good predictor of pod number and, therefore, seed yield. However, estimating flowering based on traditional visual scales is subjective, time-consuming, and labor-intensive. Recent developments in phenotyping technologies using unoccupied aerial vehicles (UAVs) make it possible to effectively capture crop information and to predict crop yield *via* imagery. The objectives were to investigate the application of vegetation indices in estimating canola flower numbers and to develop a descriptive model of canola seed yield. Fifty-six diverse Brassica genotypes including 54 *B. napus* genotypes, a *Brassica carinata* genotype, and a *Brassica juncea* genotype were grown near Saskatoon, SK, Canada from 2016 to 2018, and near Melfort and Scott, SK, Canada in 2017. Aerial imagery with geometric and radiometric corrections was collected through the flowering stage using a UAV mounted with a multispectral camera. The normalized difference yellowness index (NDYI) was a useful vegetation index for representing canola yellowness, which was related to canola flowering number during the full flowering stage. However, the flowering pixel number estimated by the thresholding method improved the ability of NDYI to detect yellow flowers with coefficient of determination (R^2) values ranging from 0.54 to 0.95. Moreover, compared with using a single image date, the NDYI-based flowering pixel numbers accumulated over time covers more growth information and can be a good predictor of pod number and thus canola yield with R^2 ranging from 0.12 to 0.42. These results indicate that NDYI-based flowering pixel number can

estimate flowering number. The cumulative flowering number extracted from imagery over time can be a potential phenotype associated with canola seed yield.

3.1 Introduction

Canola (*Brassica napus* L.) is the predominant oilseed crop grown in Canada (Clayton et al., 2000). Canada has the largest area of canola production in the world (Statistics Canada, 2018). With the growing global demand for canola, Canada needs to maintain and improve canola yield and seed quality to meet the market demands. Yield components of canola consist of the number of pods, the seeds per pod, and the weight per seed (Tayo and Morgan, 1975; McGregor, 1981; Diepenbrock, 2000; Ivanovska et al, 2007; Faraji, 2012). Among these components, pod number retained at maturity is the most important factor as it is influenced most by environment constraints (Tayo and Morgan, 1975; McGregor, 1981; Diepenbrock, 2000; Ivanovska et al, 2007; Faraji, 2012; Gan et al., 2016; Kirkegaard et al., 2018). The flowering stage of canola is important for yield estimation as flowers produced in the first 2-3 weeks from anthesis contribute to 75% of the pods at maturation (Tayo and Morgan, 1975). Additionally, the flowering period can last from 2 to 6 weeks, which is a major portion of the crop growth cycle (Gan et al., 2016; Kirkegaard et al., 2018). Thus, flower production is one of the most important factors in determining final seed yield (Tayo and Morgan, 1975; Diepenbrock, 2000; Faraji et al., 2008; Fang et al., 2016; Gong et al., 2018; Kirkegaard et al., 2018; Zhang and Flottmann, 2018).

During the plant breeding process, field-based phenotyping plays an important role in evaluating plant performance. It contributes to the selection of ideal genotypes that are high-yielding by associating genotype with the corresponding phenotype (Montes et al., 2007; Sankaran et al., 2015). To select for better canola genotypes and eventually develop better varieties, breeders need to assess many distinct genotypes grown in multiple environments to detect interactions between genotype and environment (White et al., 2012; Araus and Cairns, 2014).

The estimation of flowering number based on traditional visual scales is subjective, labor-consuming, and is often destructive (Sulik and Long, 2015; Fang et al., 2016; Wan et al., 2018).

Although ground-based platforms such as Greenseeker (Trimble Inc., Westminster, CO, United States), Crop Circle (Holland Scientific Inc., Lincoln, NE, United States), or time-lapse RGB imaging can provide adequate spectral data, these platforms still require a prohibitive amount of time and labor (Xu et al., 2017; Hassan et al., 2019). Additionally, data collection using these ground-based platforms may cause crop canopy damage, soil compaction, and is also subject to environmental conditions that could limit field access. Therefore, it is necessary to develop an objective, non-destructive, and efficient method to estimate flower numbers. With this, one can model seed yield by assessing radiometric data of the crop canopy, which has the potential to accelerate breeding methods for yield improvement. Current improvements in aerial-based platforms and imaging sensors make it possible to effectively collect phenotypes *via* analyzing digital imagery (Kim et al., 2019). Unoccupied aerial vehicles (UAVs) equipped with various sensors can quickly provide large quantities of field data enabling plant breeders to efficiently detect traits of numerous plots in large-scale field trials (Kefauver et al., 2017).

Spectral reflectance of the crop canopy is correlated with morphological and physiological traits. Leaf pigment composition and cell structure can affect reflectance of crop; thus, ratios or differences of different bands in the visual light, near IR (NIR) and shortwave IR wavelengths (i.e., vegetation indices) can be a tool to characterize plant traits (Sankaran et al., 2015; Wójtowicz et al., 2016). Previous studies have shown that multispectral reflectance profiles of visible bands (i.e., blue, green, and red), and NIR bands could estimate canopy features such as nitrogen use efficiency (Kefauver et al., 2017; Prey et al., 2020), leaf area index (Tunca et al., 2018; Blancon et al., 2019), and flower numbers (Guo et al., 2015; Sulik and Long, 2015, 2016; Carl et al., 2017; Gong et al., 2018; Wan et al., 2018; Xu et al., 2018). These plant traits investigated remotely have the potential to improve yield estimates. Flower number, as an important factor in determining crop yield, is correlated with optical properties in various crops, such as rice (Guo et al., 2015), cotton (Xu et al., 2017, 2018), and canola (Sulik and Long, 2015, 2016; Gong et al., 2018; Wan et al., 2018). Guo et al. (2015) applied a machine learning model, the support vector machine, for rice flowering identification using RGB images, which resulted in a significant correlation between the visually counted rice flowering panicles and identified flowering (correlation coefficients ranging from 0.64 to 0.82). In canola, there are three different canopy morphologies during the growing season, namely, the vegetative phase (green canopy dominated by leaves), the flowering phase (yellow canopy dominated by the yellowness of flower petals), and the mature phase (green or brown canopy

because of pods and branches) (Sulik and Long, 2016). During the flowering phase, the yellowness of canola petals is due to carotenoid absorption of blue and reflectance of a mixture of green and red wavelengths (Sulik and Long, 2015, 2016), but the yellow color has little impact on red edge and NIR reflectance unlike a green vegetative canopy (Shen et al., 2009; Migdall et al., 2010; Sulik and Long, 2015, 2016). Thus, the contributed red light decreases the normalized difference vegetation index (NDVI) values (Equation 3.1) and adversely impacts the ability of NDVI to monitor crop growth condition and estimate yield during the flowering phase (Shen et al., 2009, 2010; Sulik and Long, 2015, 2016). Sulik and Long (2015) found that the ratio of green and blue was strongly correlated with the actual flower numbers with a coefficient of determination (R^2) of 0.87, and they proposed the normalized difference yellowness index (NDYI) (Equation 3.2) could be a potential yield predictor ($R^2 = 0.76$) (Sulik and Long, 2016). d'Andrimont et al. (2020) and Han et al. (2020) reported that NDYI successfully captured canola yellowness and detected the peak flowering dates using Sentinel-2 satellite imagery (European Space Agency, United Nations). Fang et al. (2016) found that reflectance at 550 nm was the most sensitive band to estimate flowering coverage with an estimation error below 6% when compared with wavelengths at 490, 670, 720, 800, and 900 nm. Wan et al. (2018) and Gong et al. (2018) reported that combining vegetation index and image classification methods (i.e., k-means clustering method by CIE L*a*b space and pixel-level spectral mixture analysis) improved the accuracy of flower numbers and yield estimation in canola with R^2 values of 0.89 and 0.75, respectively.

Although several studies have detected canola flowering number and predicted yield, most of these field experiments were conducted with relatively few canola genotypes and environments, which may neglect the effect of genotype and environmental fluctuations on yellowness of flower (Ohmiya, 2011) and petal size (Jiang and Becker, 2003; Qian et al., 2021). In addition, yield estimation models used in those studies were based on only one image date (Sulik and Long, 2016; Gong et al., 2018), which ignores the effect of time and duration of flowering (Tayo and Morgan, 1975). Thus, the cumulative reflectance data of flowering throughout the flowering period may provide a better estimate of crop yield. Therefore, the objectives in this study were to use UAV multispectral data to detect flowers within a wide range of canola genotypes (Table A.1, Appendix A) and to estimate seed yield in canola using time series imagery collected during the flowering period.

3.2 Materials and Methods

3.2.1 Experimental Sites and Plant Materials

The experiment was conducted at the Agriculture and Agri-Food Canada Research Farm near Saskatoon, SK, Canada from 2016 to 2018 (52° 10' 52.9" N, 106° 30' 10.6" W in 2016; 52° 10' 59.3" N, 106° 30' 53.7" W in 2017; 52° 10' 57.7" N, 106° 30' 01.4" W in 2018), and near Melfort (52° 49' 9.6" N, 104° 35' 46.9" W) and Scott (52° 21' 55.3" N, 108° 52' 32.6" W), SK, Canada in 2017 (Table 3.1). The soil texture at Saskatoon was a clay loam with a pH of 7.3 and an organic matter content of 5.5%. The field plots were set up in a randomized incomplete block design (i.e., rectangular lattice design) with three replications (Figure 3.1). A rectangular lattice design was used to reduce spatial variability within each block. Individual plot size was 6.0 m long x 1.2 m wide in 2016 and 2018, and 6.0 m long x 1.5m wide in 2017. Fifty-six genotypes (Table A.1, Appendix A; Saskatoon Research and Development Center, Agriculture and Agri-Food Canada) including 54 diverse *B. napus* genotypes, a *B. carinata* genotype, and a *B. juncea* genotype were selected and planted with a row spacing of 0.3 m. Fifty of the diverse genotypes were used as founders to form a spring *B. napus* Nested Association Mapping (NAM) population from crossing founders to a common reference genotype (Ebersbach et al., 2022). This panel, which represents diverse germplasm resources and the historical basis of canola breeding programs, differs in geographic origin, pedigree, phenotypes, and genotype (Ebersbach et al., 2022). Seeding occurred on May 27, 28, and 21 in 2016, 2017, and 2018, respectively, at a seeding rate of 108 seeds m^{-2} (Table 3.1). Out of the 56 genotypes, 16 were selected and planted in two adjacent but separate plots as double plots. The criteria of genotype selection for the double plots were based on contrasting seed quality (i.e., seed color, acid detergent lignin, seed glucosinolates, and seed erucic acid) and phenological similarities (Taye et al., 2020). The reason for setting double plots was to preserve one plot for imaging and yield determination without any subsamples being removed. The 16 *B. napus* genotypes planted in double plots were YN04-C1213sp013, DH27298, NAM-0, 5, 13, 14, 17, 23, 30, 37, 32, 43, 46, 72, 76, and 79.

Table 3.1 Summary of canola trials and data collection (imagery acquisition and manual flower count) at Saskatoon, SK, Canada from 2016 to 2018, and at Melfort and Scott, SK, Canada in 2017.

Site	Year	Seeding date	Number of genotypes	Flight altitude (m)	Image acquisition dates	Manual flower count dates
Saskatoon	2016	May 27	56	20	July 14; 19; 26 August 06	July 15; 22; 29 August 05
	2017	May 28	56	20	July 07; 11; 15; 19; 22; 26 August 01; 09; 16; 22	July 10;18; 25 August 01
	2018	May 21	56	25	June 28 July 06; 09; 16; 20; 24; 27; 30 August 03; 07	July 10; 17; 24; 31
Melfort	2017	May 18	16	15	July 05; 13; 20; 26	July 05; 20; 26
Scott	2017	June 20	16	20	August 09; 16; 29	August 09; 16

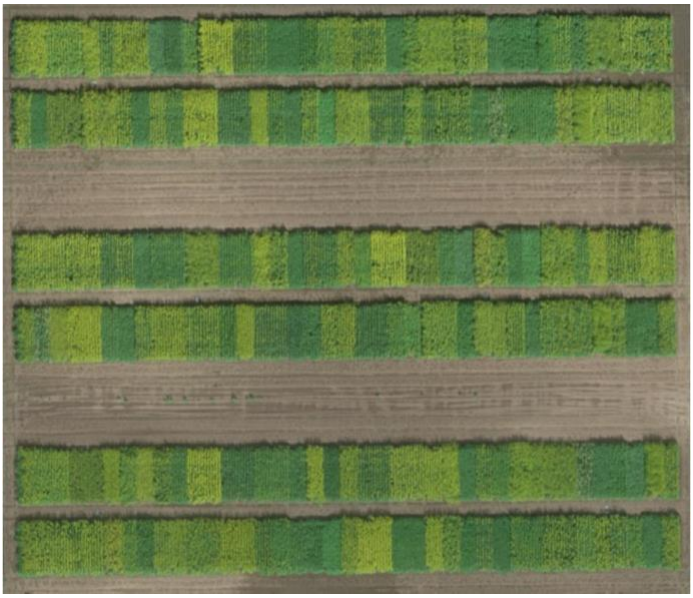


Figure 3.1 The overview of experimental plot layout at the Agriculture and Agri-Food Canada Research Farm (52° 10' 52.9" N, 106° 30' 10.6" W) near Saskatoon, SK, Canada on July 14, 2016.

The selected 16 *B. napus* genotypes were planted in a randomized complete block design with three replications at the Melfort and Scott locations in 2017. The 16 genotypes were seeded in double-plots and subsampling occurred in destructive plots. The neighbouring non-destructive plots were used for imaging and final yield determination. All genotypes were planted in 5 m long x 1.2 m wide plots at Melfort and in 5m long x 1 m wide plots at Scott. Canola was seeded

on May 18 at Melfort and re-seeded on June 20 at Scott due to hail damage, at a desired seeding rate of 108 seeds m^{-2} with a row spacing of 0.3 m (Table 3.1). Edge® (ethalfluralin) was applied as a pre-emergence herbicide at a rate of 19.1 kg ha^{-1} to control weeds. Any weeds not controlled by the herbicides were removed by hand.

3.2.2 Image Acquisition

3.2.2.1 Platform and sensor

The UAV used in this study was a Draganflyer X4-P model in 2016 and 2017 (DraganFly Inc., Saskatoon, SK, Canada). It is a rotary-wing platform with a maximum payload of 800 g. It can semi-automatically depart and land based on GPS navigation mode and optional Surveyor software. Flight mission was planned in Surveyor software (DraganFly Inc., Saskatoon, SK, Canada) by importing ground coordinates of the field boundaries. The other rotary-wing platform was a Draganflyer Commander model (DraganFly, Inc., Saskatoon, SK, Canada), used in 2018, which differs from the X4-P model in its maximum payload capacity (1,000 g).

A multi-spectral camera (RedEdge, MicaSense Inc., Seattle, WA, United States) was used to acquire images (12-bit image) with an image resolution of 1.2 megapixels (1,280 x 960 pixels) for each of five spectral bands (blue: 475 ± 10 nm; green: 560 ± 10 nm; red: 668 ± 5 nm; red edge: 717 ± 5 nm; near infrared: 840 ± 20 nm) (Table 3.2). The focal length of the camera is 5.5 mm and ground sampling distance at 15, 20, and 25 m above ground level was 1.02, 1.36, and 1.70 cm per pixel, respectively (Table 3.2). Images of a MicaSense reflectance panel (RedEdge, MicaSense Inc., Seattle, WA, United States) were taken before and after each UAV flight for radiometric calibration. To geo-reference aerial images, six ground control points (GCPs) were distributed across the experimental area during the whole crop season in 2016 at Saskatoon. The size of the GCPs was 60 cm x 60 cm, which were geolocated by Trimble GeoExplorer 2008 GPS (Trimble Inc., Westminster, CO, United States). GCPs were manually placed at the same location when phenotyping canola by UAV, which provided an overlay of images taken from various dates and reduced workload by using the same geolocation information for each GCP. For the four locations in 2017 and 2018, GCPs were permanently mounted within guard plots to avoid manually carrying GCPs to the field.

Table 3.2 Basic specifications for the multispectral camera (RedEdge) equipped on the unoccupied aerial vehicle (UAV) platforms.

GSD[a] (cm per pixel) (per band)	Flight altitude (m)	Sensor resolution per band (MP)[b]	Focal length (mm)	Full width at half maximum (nm)	Peak wavelength (nm)
				Blue: 465 - 485	Blue: 475
1.02	15			Green: 550 - 570	Green: 560
1.36	20	1.2[c]	5.5	Red: 663 - 673	Red: 668
1.70	25			Red edge: 712 - 722	Red edge: 717
				NIR: 820 - 860	NIR: 840

[a]GSD: ground sampling distance
[b]MP: megapixel
[c]image resolution: 1.2 MP = 1,280 x 960 pixels

3.2.2.2 UAV flight schedule

The UAV, equipped with a multispectral camera, captured the images of the fields weekly during the flowering stage at Saskatoon in 2016 and at Melfort and Scott in 2017 (Table 3.1). The imagery was collected semi-weekly in 2017 and 2018 at Saskatoon for the duration of canola flowering (Table 3.1). For the Saskatoon location, although weather conditions such as rain, clouds, and heavy wind limited the flight schedule, image timing interval was achieved as close to 7 days in 2016, and 4 days in 2017 and 2018. For the Melfort and Scott locations in 2017, image collection was carried out at a 7-day interval.

3.2.3 Image Process and Data Extraction

3.2.3.1 Image pre-process

Multispectral images were processed, stitched, and calibrated in Pix4Dmapper Pro (Pix4D Inc., San Francisco, CA, United States). Individual images were aligned based on common points from the overlapped images to generate a geo-referenced image that matched the overflown study area. Geometric calibration was done by importing the geo-location of GCPs to reduce geometric distortion problems of the camera. A system coordinate, World Geodetic System 1984, was applied to generate geo-referenced images. The images of the MicaSense reflectance panel were used in the radiometric calibration to enhance spectral consistency between different flight dates. Then the five generated reflectance maps corresponding to each spectral band were exported and used for further analysis.

3.2.3.2 Vegetation index calculation, thresholding, and accumulation of flowering progress

ArcGIS software version 10.4.1 (ESRI Canada, Toronto, ON, Canada) was applied for plot segmentation, vegetation indices calculation, and thresholding. In this study, the middle three rows for each plot were segmented using polygon shapes with assigned plot numbers. The polygon shapes were generated using the "Create Feature" tool. Vegetation index maps were derived *via* calculation of the reflectance maps using the "Raster Calculator" tool. Commonly used vegetation indices, NDVI (Rouse et al., 1974), NDYI (Sulik and Long, 2016), green normalized difference vegetation index (GNDVI) (Gitelson et al., 1996), and normalized difference red edge index (NDRE) (Gitelson and Merzlyak, 1997) were calculated as following equations to compare with the actual flower number counts:

$$\boldsymbol{NDVI} = \left(\frac{R_{NIR} - R_{red}}{R_{NIR} + R_{red}}\right) \quad [3.1]$$

$$\boldsymbol{NDYI} = \left(\frac{R_{green} - R_{blue}}{R_{green} + R_{blue}}\right) \quad [3.2]$$

$$\boldsymbol{GNDVI} = \left(\frac{R_{NIR} - R_{green}}{R_{NIR} + R_{green}}\right) \quad [3.3]$$

$$\boldsymbol{NDRE} = \left(\frac{R_{NIR} - R_{red\ edge}}{R_{NIR} + R_{red\ edge}}\right) \quad [3.4]$$

where R_{NIR}, R_{red}, R_{green}, R_{blue}, and $R_{red\ edge}$ are the reflectance values at bands centered on 840, 668, 560, 475, and 717 nm, respectively (Table 3.2). NDVI is the most commonly used vegetation index to identify crop growth condition and yield estimation (Rouse et al., 1974). NDYI has previously shown strong correlation with seed yield (Sulik and Long, 2016). GNDVI (Gitelson et al., 1996) and NDRE (Gitelson and Merzlyak, 1997) are related to photosynthesis and have been reported in previous research.

Canola flowers and leaf organs co-existed within each plot during flowering; thus, the "Conditional Function" [Con (index map > threshold value, index map, "")] in the "Raster Calculator" tool was used to separate flowering pixels from non-flowering pixels by applying threshold values on vegetation index maps. Threshold values were manually determined by comparing the composited RGB images with calculated index maps so that most flowering pixels could be selected and segmented. All pixels in the index map that have values larger

than the threshold values were kept in a threshold index map, otherwise, pixels were discarded. Then, the "Zonal Statistics" tool was used to extract the summary statistics of the threshold index map, which included the number of flowering pixels per plot.

This study involved 56 diverse genotypes with a high flowering density gradient and diverse flowering timing. In most cases, the 56 genotypes showed different development stages within an image. It was difficult to determine which image date was the perfect timing for yield estimation. For this reason, the area under the flowering progress curve (AUFPC) was used to calculate accumulation of flowering progress during the flowering season using the following equation:

$$AUFPC = \left(\frac{F_1+F_2}{2} - F1\right)(t_2 - t_1) + \left(\frac{F_2+F_3}{2} - F1\right)(t_3 - t_2) + \cdots + \left(\frac{F_{n-1}+F_n}{2} - F1\right)(t_n - t_{n-1}) \quad [3.5]$$

where F_1, F_2, F_3, F_{n-1}, and F_n represent the flowering pixel numbers at each image date and t_1, t_2, t_3, t_{n-1}, and t_n represent Julian date at each image timing. The flower accumulation refers to the overall flower progress over the flowering stage. The AUFPC is an adjusted accumulation equation based on the area under the disease progress curve (AUDPC), which is used in general in pathology studies for estimating the effect of disease progression on crop yield (Jeger and Viljanen-Rollison, 2001; Simko and Piepho, 2012). Similar to the AUDPC, the advantage of the AUFPC is providing a baseline for each genotype to adjust flowering progress, which can reduce the effect of diverse initial flowering pixel numbers of each genotype on the calculated area. The AUFPC equation converted several flowering pixel numbers at a series of image timings into a single value for reporting. The higher AUFPC number indicates greater overall flowering number. Figure 3.2 displays an example of flowering progress over time for a genotype (NAM-23). Seven data points on the curve line represent NDYI-based pixel numbers for each image date. Pictures under the seven points are corresponding threshold index maps. Then, the area under the curve line was calculated using the AUFPC equation (Equation 3.5) for NAM-23. The same mathematical method was used to calculate flowering progress for all other genotypes across 5 site years.

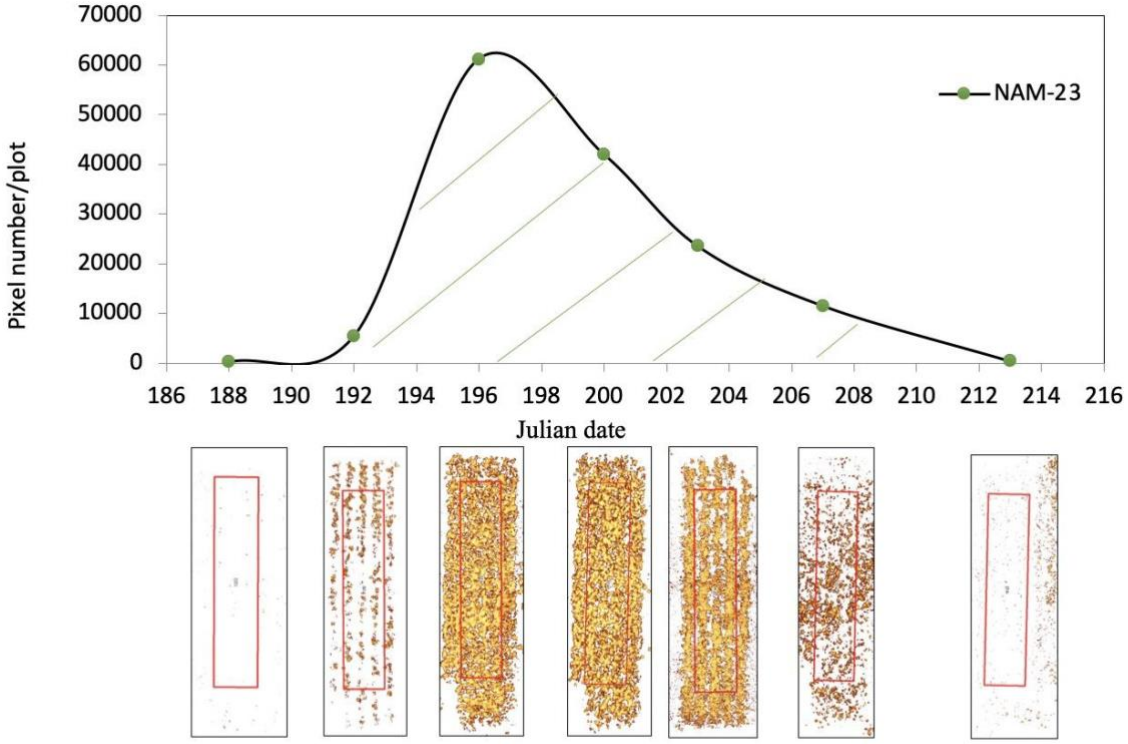


Figure 3.2 The growth pattern of flowering progress for a *Brassica napus* genotype (NAM-23) during the flowering stage at the Agriculture and Agri-Food Canada Research Farm (52° 10' 59.3" N, 106° 30' 53.7" W) near Saskatoon, SK, Canada in 2017. X-axis is imagery acquisition date (Julian date) in 2017. The y-axis is normalized difference yellowness index (NDYI)-based pixel number per plot. A solid curve line is the flowering progress trend of NAM-23. Seven points on the progress curve line represent NDYI-based pixel number per plot at seven imagery acquisition dates. Seven pictures under each point are corresponding false-color images after thresholding with flowers highlighted in yellow. The region of interest was highlighted in red.

3.2.4 Ground Reference Data Collection

Canola flowering started in early July and ended in early August. The first row of each plot was manually sampled to estimate flowering at a 7-day interval from July to August. Flower numbers on the main stem and branches was counted within a 0.075 m^2 area. Then, the counted flower number was converted to per-plot value (i.e., actual flower number per plot area). The manual flower count for total 168 experimental plots required approximately eight hours with the help of nine summer students per sampling date. Grain yield was straight combined by a small plot combine harvester when the crop was mature and dry. This occurred multiple times due to differing maturity dates of the *B. napus* genotypes. To reduce the border effect, the

middle four rows of each plot were harvested. All harvested seeds were air-dried to 10% seed moisture. Final yields were weighed after seed cleaning.

3.2.5 Statistical Analysis

The PROC LATTICE procedure of SAS version 9.4 (SAS Institute, Cary, NC, United States) was used to analyze the data. The LATTICE procedure reduced variations within blocks. After data adjustment, PROC REG in SAS version 9.4 was used as the statistical tool to investigate the simple linear regressions between ground reference data and imagery. Both ground reference data (i.e., actual flower number and seed yield) and imagery (i.e., vegetation indices) were averaged for each genotype. In this study, a zero-intercept linear regression model was used as there was no flowering pixel prior to the commencement of flowering when comparing the actual flower number and imagery. Furthermore, the fitted intercept values were close to zero in most cases. Scatterplots of variables were observed to determine whether data could be combined for analysis. In the case where data could not be combined, data were analyzed within site years (Tables B.1, B.2, Appendix B).

3.3 Results and Discussion

3.3.1 Regression Between Flowering Pixel Number and Actual Flower Numbers

Initial regression results showed that GNDVI and NDRE did not demonstrate significant correlations with the actual flower count ($p > 0.05$, data not shown). Meanwhile, regression results showed NDYI had greater coefficients of determination (R^2) than NDVI with actual flower numbers over 3 years of study. Increased red light from the yellow petals can reduce NDVI values and affect its ability to detect canola growth conditions. Noise from soil background and green vegetation within a plot at the early flowering stage may result in an adverse effect on the ability of NDYI to detect yellow flowers. For this reason, NDYI maps were used to extract flowering pixels and remove non-flowering pixels by the thresholding method. Flowering pixels were detected and counted when pixel values were greater than NDYI-based threshold levels. Threshold values were 0.59, 0.52, and 0.45 in 2016, 2017, and 2018, respectively.

Across 5 site years, the R^2 values between NDYI-based flowering pixel numbers and actual flower numbers ranged from 0.54 to 0.95 during flowering duration (Figures 3.3–3.7) depending on the date of sampling. There were significant relationships between flowering pixel numbers and actual flower numbers in 2016 at Saskatoon (Figure 3.3). Not surprisingly, the early flowering stage (July 15) had a significant regression relationship with actual flower numbers with R^2 of 0.85 (Figure 3.3A). Developing flowers were on the upper part of a plant at the early flowering stage so sensors could easily detect these early-blooming flowers. Whereas, the late flowering stage (August 05) showed the weakest regression (R^2 = 0.54) (Figure 3.3D) compared with other image dates, which might be a result of lower sensitivity of NDYI to differentiate yellow flowers and dark green pods. Dark green pods impart more green reflectance, which can make NDYI less sensitive to yellow flowers, as yellow is a composite color of green and red (Yates and Steven, 1987; Sulik and Long, 2015, 2016). Additionally, the potential reason why it had the smallest R^2 was the occlusion of late-developing flowers growing on the lower branches.

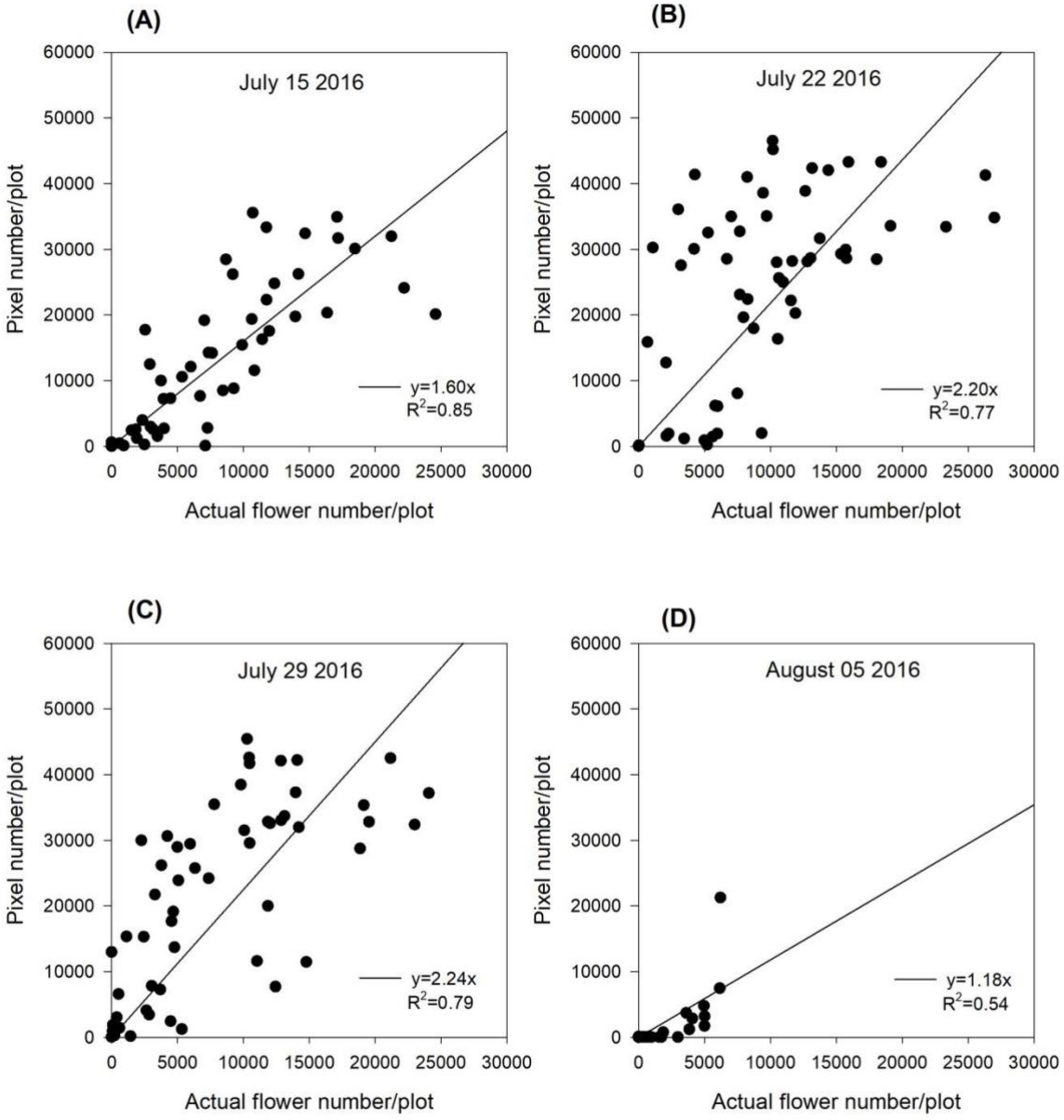


Figure 3.3 The relationship between actual flower numbers per plot and pixel numbers extracted from aerial images during the flowering stage at Saskatoon, SK, Canada in 2016. Actual flower numbers per plot were manually measured. Pixel number per plot was detected by the thresholding method. (A) Regression equation for July 15, 2016: y=1.60x, R^2=0.85. (B) Regression equation for July 22, 2016: y=2.20x, R^2=0.77. (C) Regression equation for July 29, 2016: y=2.24x, R^2=0.79. (D) Regression equation for August 05, 2016: y=1.18x, R^2=0.54.

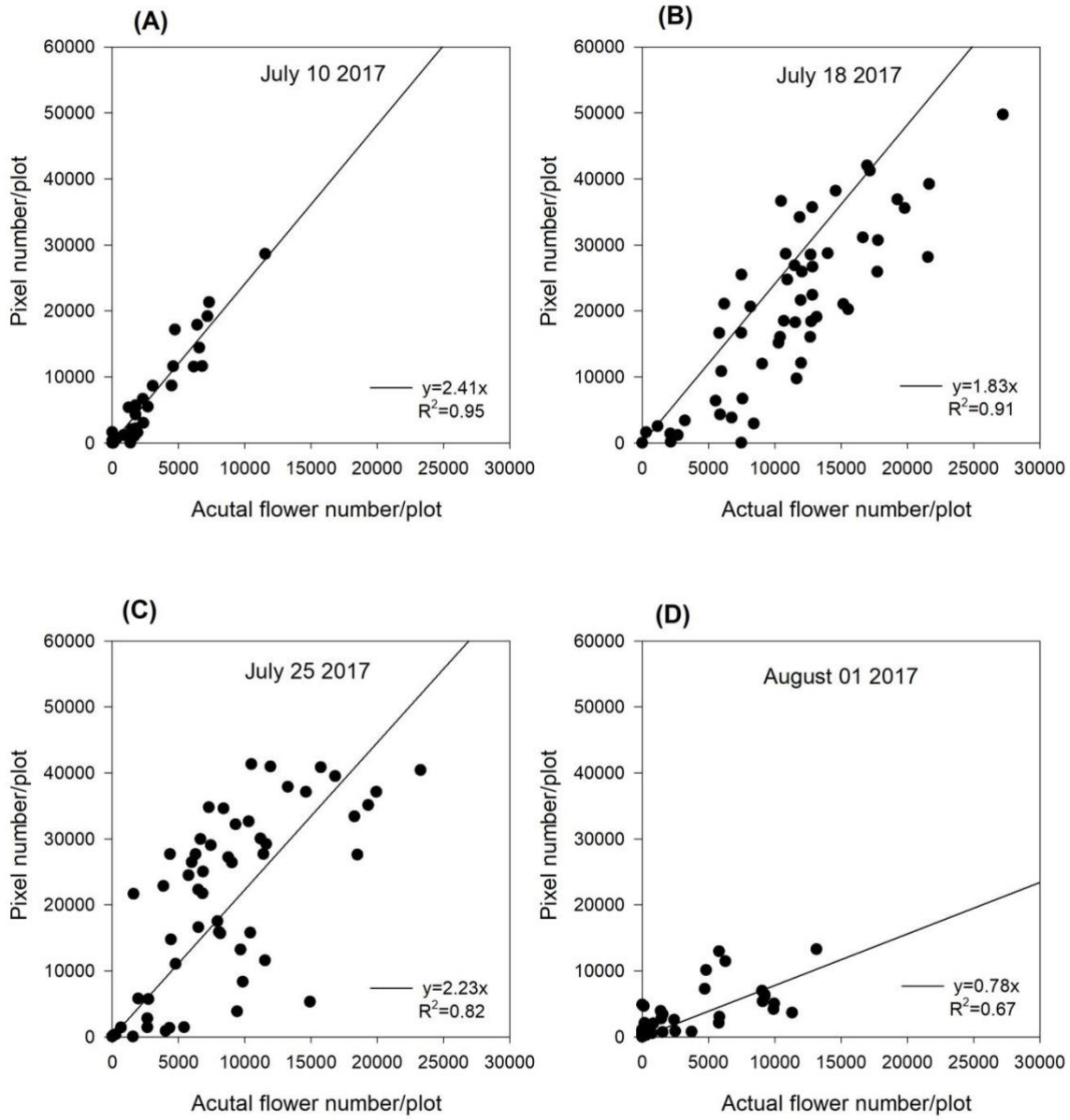


Figure 3.4 The relationship between actual flower numbers per plot and pixel numbers extracted from aerial images during the flowering stage at Saskatoon, SK, Canada in 2017. Actual flower numbers per plot were manually measured. Pixel number per plot was detected by the thresholding method. (A) Regression equation for July 10, 2017: y=2.41x, R^2=0.95. (B) Regression equation for July 18, 2017: y=1.83x, R^2=0.91. (C) Regression equation for July 25, 2017: y=2.23x, R^2=0.82. (D) Regression equation for August 01, 2017: y=0.78x, R^2=0.67.

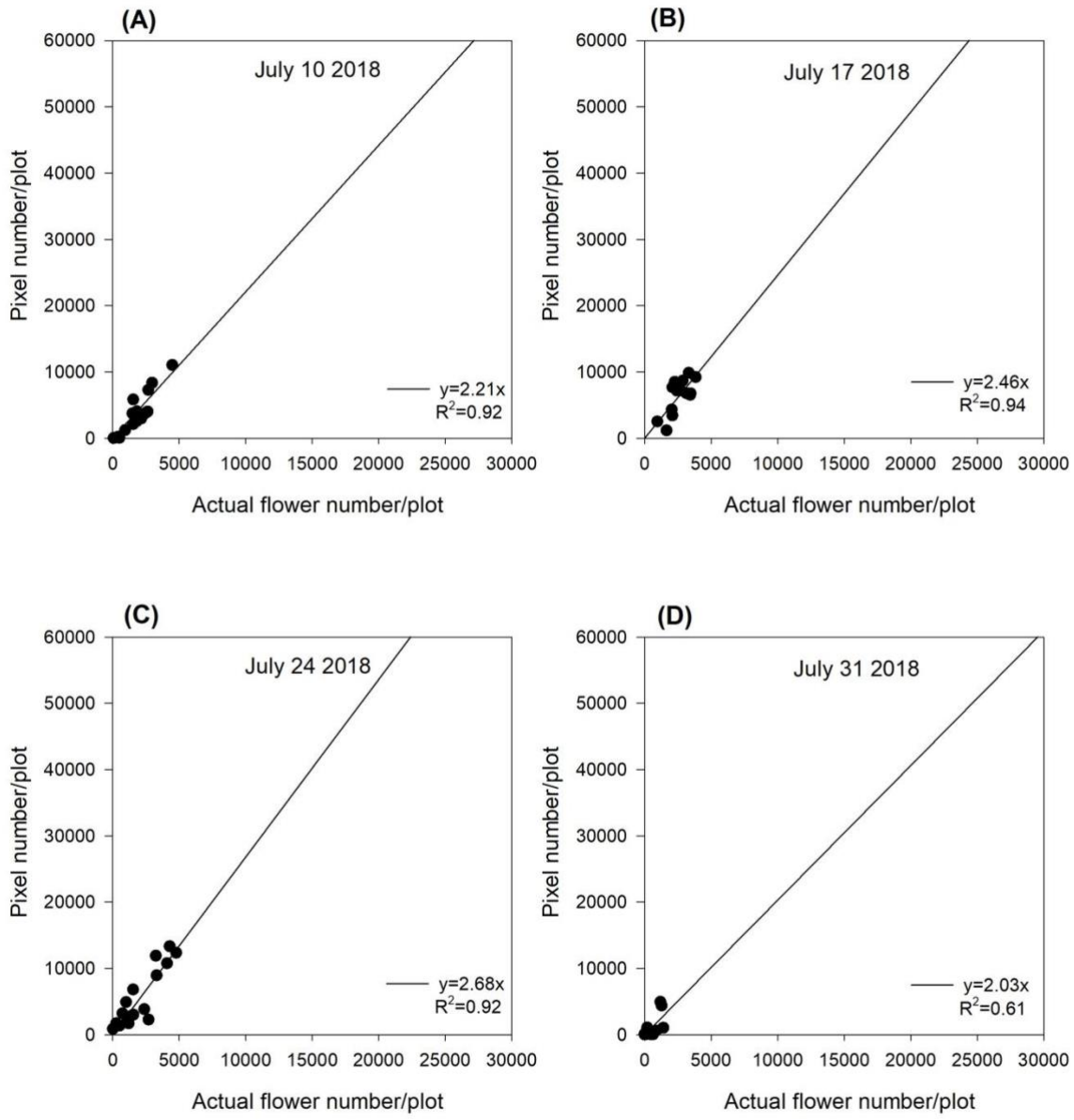


Figure 3.5 The relationship between actual flower numbers per plot and pixel numbers extracted from aerial images during the flowering stage at Saskatoon, SK, Canada in 2018. Actual flower numbers per plot were manually measured. Pixel number per plot was detected by the thresholding method. (A) Regression equation for July 10, 2018: y=2.21x, R^2=0.92. (B) Regression equation for July 17, 2018: y=2.46x, R^2=0.94. (C) Regression equation for July 24, 2018: y=2.68x, R^2=0.92. (D) Regression equation for July 31, 2018: y=2.03x, R^2=0.61.

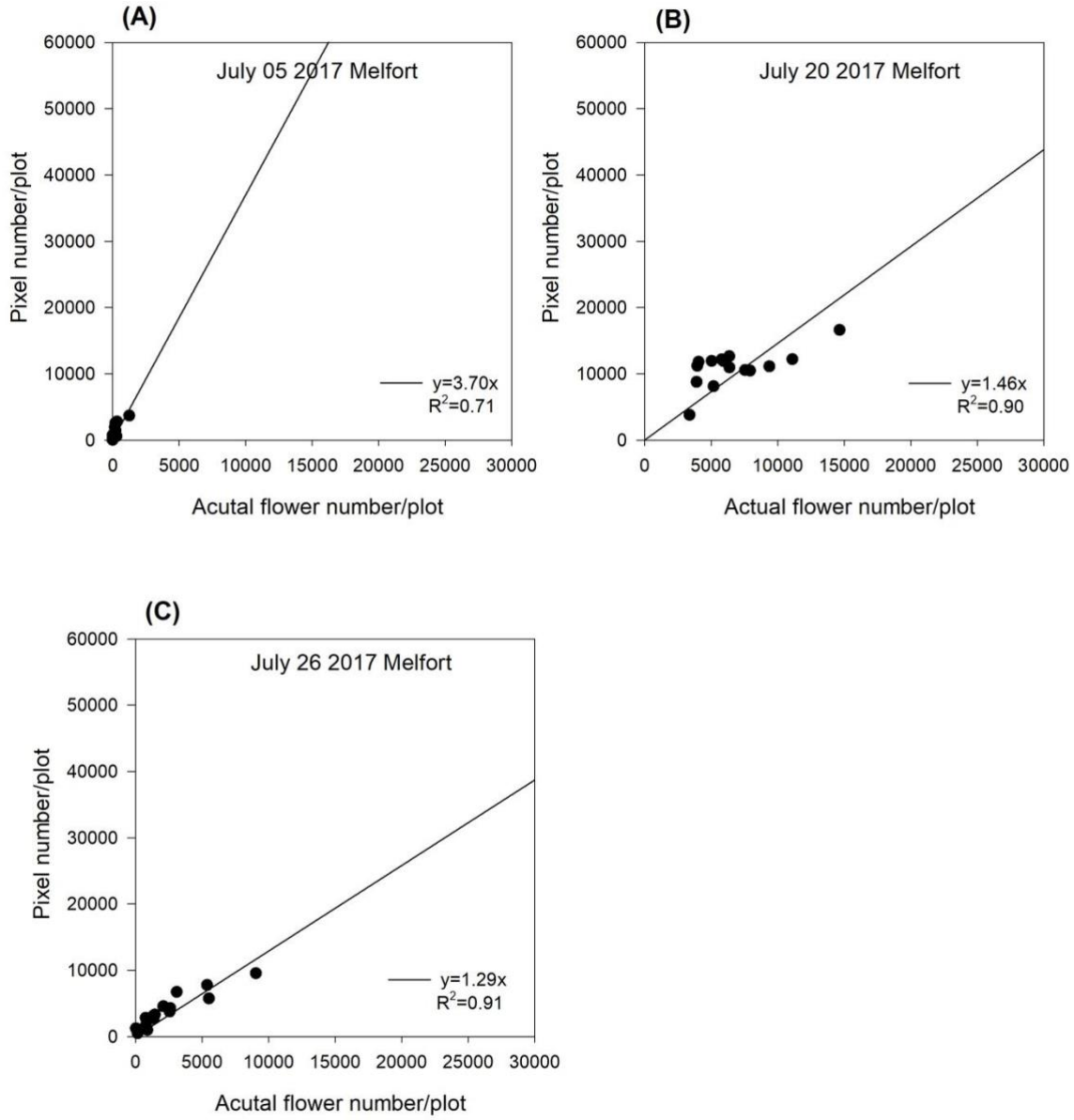


Figure 3.6 The relationship between actual flower numbers per plot and pixel numbers extracted from aerial images during the flowering stage at Melfort, SK, Canada in 2017. Actual flower numbers per plot were manually measured. Pixel number per plot was detected by the thresholding method. (A) Regression equation for July 05, 2017: y=3.70x, R^2=0.71. (B) Regression equation for July 20, 2017: y=1.46x, R^2=0.90. (C) Regression equation for July 26, 2017: y=1.29x, R^2=0.91.

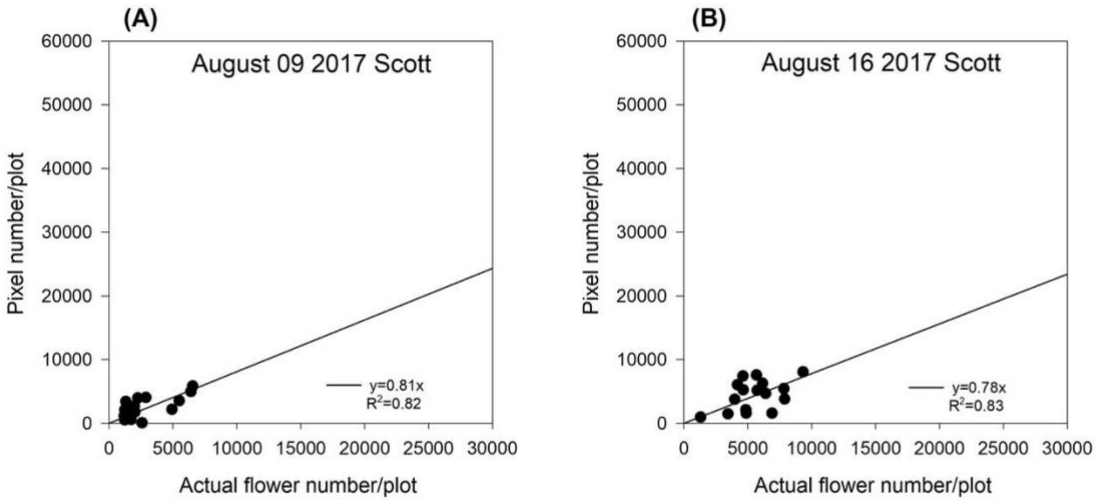


Figure 3.7 The relationship between actual flower numbers per plot and pixel number extracted from aerial images during the flowering stage at Scott, SK, Canada in 2017. Actual flower numbers per plot were manually measured. Pixel number per plot was detected by the thresholding method. (A) Regression equation for August 09, 2017: y=0.81x, R^2=0.82. (B) Regression equation for August 16, 2017: y=0.78x, R^2=0.83.

The Saskatoon location in 2017 and 2018 had similar regression patterns between flowering pixel numbers and actual flower numbers (Figures 3.4, 3.5). There were significant relationships at the early flowering stages (July 10, 2017 and July 17, 2018) (Figures 3.4, 3.5). Similar to 2016, the relationships became weaker with the late flowering stage (August 01, 2017 and July 31, 2018) (Figures 3.4, 3.5). Although the late flowering stages had weaker regressions compared with the early flowering timing, the regressions at the peak flowering dates (July 25, 2017 and July 24, 2018) had high R^2 ranging from 0.82 to 0.92 (Figures 3.4, 3.5).

For the Melfort location in 2017, the first image date (July 05) had the weakest regression (R^2 = 0.71) (Figure 3.6A). Ground measurement errors from subsampling plants could be a potential reason for decreased regressions at the early flowering stage. Moreover, another

potential reason could be the actual flower number counted within a small sampling area might not be representative of entire flower number for each plot. However, the peak flowering time (July 20) and late flowering stage (July 26) showed significant relationships with higher R^2 up to 0.91 compared with the early flowering stage (Figure 3.6). The potential reason why this site year had a greater R^2 at the late flowering stage is that flight altitude (15m) at Melfort in 2017 was lower than the other site years (Table 3.1). The higher resolution may have increased the ability of the sensor to detect flowers growing lower in the canopy. Although the flight altitude was relatively low compared with other locations, there was no significant canopy movement due to the UAV platform. The re-seeding date at Scott was June 22, 2017. Flowering started relatively late with a shorter duration compared with other site years. There was no imagery collected at the end of the flowering stage, and thus, those relationships are unknown. At Scott, the R^2 values for the regressions between flowering pixel numbers and actual flower numbers followed similar patterns as the Saskatoon location. The early flowering stage (August 09) and the peak flowering time (August 16) had significant relationships (Figure 3.7).

For the Saskatoon location over 3 years, slopes were relatively consistent at the early flowering stages (Figures 3.3–3.5). Slope values became smaller with the delayed flowering stage. There was a smaller slope value at the late flowering stage (slope = 1.18) compared with the peak flowering time (slope = 2.20) at Saskatoon in 2016. The Saskatoon location in 2017 and 2018 had similar patterns (Figures 3.4, 3.5). The Melfort location had similar patterns with a smaller slope at the late flowering stage (Figure 3.6) but the slope of the first image date (slope = 3.70) was greater than the other image dates. This indicated that early flowering imagery overestimated the actual flower number. Experimental plots at this location showed non-uniform flowering with fewer flowers at the front of each plot which might be caused by the edge effect. Thus, manual flower count based on subsampling plants at the front row of a plot may not accurately represent the average flower numbers. In 2017 at Scott, slopes were consistent at the early and the peak flowering times (Figure 3.7). The slope values at this location were smaller than the other site years. A potential reason for this underestimation of flower number was that the plots had a more condensed canopy and there were more branches at this site year than other site years (data not shown) due to poor emergence percentage. Thus, for the Scott location, there were more flowers produced on the lower branches resulting in more occlusion issues. As mentioned above, there was no available data collected at the end of flowering; thus, the relationship at this stage is unknown.

In general, although the linear regression slopes varied across site years, the high R^2 values indicated that the flowering pixel numbers extracted from the threshold NDYI map performed well to predict actual flower numbers at the early and peak flowering stages in canola (R^2 up to 0.95). Moreover, the early flowering stages across 5 site years showed greater slope values, as most flowers at this early stage were visible and had less occlusion. In contrast, flowers growing on lower branches were likely to be underestimated at the late flowering stages. Subsampling variability might make the actual flower count non-representative for a plot, which might reduce the accuracy of flower estimation. These results were consistent with Sulik and Long (2015), wherein the ratio of blue and green was significantly correlated to the yellow flowers in canola with a significant R^2 value of 0.87 at the full flowering stage. Wan et al. (2018) reported a significant correlation (R^2 = 0.89) between manually counted flower number and classified flower coverage area using k-means clustering method based on the CIE L*a*b space model during the full flowering period. Xu et al. (2018) found that white cotton flowers had higher prediction accuracy at the early flowering stage. The lower classification accuracy at the later growth stage may have resulted from coverage of leaves which increased misclassified non-flowers when using a convolutional neural network (Xu et al., 2018). They recommended that using one raw image and side-view cameras might solve this issue, as more cotton flowers would be detected from different perspectives.

3.3.2 Yield Estimation Using UAV-derived Flowering Accumulation During The Flowering Period

Flowering pixel numbers derived from the threshold NDYI map were able to estimate actual flower numbers across 5 experimental site years. Initially, regression analysis was used to determine the relationship between yield and flowering pixel numbers at each image date. Among the 5 site years, in most cases, there were no significant relationships until the middle of July when most varieties started blooming (Table 3.3). In addition, it is difficult to determine a single well-defined image time for crop yield estimation because of various environmental fluctuations and various flowering timings in large-scale breeding programs, especially involving many diverse genotypes. Furthermore, important flowering progress information may be missed if yield estimation is only based on imagery from a single date (Haynes and Weingartner, 2004; Gan et al., 2016). Although flower formation at the later stage may

contribute less than early timing points, they may still have the potential to increase final grain yield. Therefore, integrating all aspects of the entire flowering duration using AUFPC can reflect flowering accumulation progress and improve the accuracy of crop yield estimation.

Table 3.3 The coefficient of determination (R^2) between flowering pixel numbers from a single image date and yield at Saskatoon, SK, Canada from 2016 to 2018 and at Melfort and Scott, SK, Canada in 2017.

Site	**2016**	R^2	**2017**	R^2	**2018**	R^2
Saskatoon	July 14	0.04	July 07	0.02	June 28	< 0.01
	July 19	< 0.01	July 11	< 0.01	July 06	0.02
	July 26	0.02	July 15	0.04	July 09	0.06
	August 06	0.04	July 19	0.29***	July 16	0.36***
			July 22	0.33***	July 20	0.22***
			July 26	0.06	July 24	0.07
			August 01	0.06	July 27	< 0.01
			August 09	0.02	July 30	0.03
			August 16	0.05	August 03	0.03
			August 22	0.05	August 07	0.02
Melfort			July 05	< 0.01		
			July 13	0.23*		
			July 20	0.02		
			July 26	0.14		
Scott			August 09	0.46**		
			August 16	0.32*		
			August 29	0.01		

* Significant at the 0.05 probability level.
** Significant at the 0.01 probability level.
***Significant at the 0.001 probability level.

There were significant relationships between the UAV-derived flower accumulation and yield during the flowering period (Figures 3.8, 3.9). In 2016, at Saskatoon, UAV-derived flower accumulation had a significant relationship with yield ($R^2 = 0.12$, $p < 0.05$) (Figure 3.8A). The calculation of flower accumulation progress was done by integrating the flowering pixel numbers over four image dates at a 7-day interval, which missed the starting point of the flowering period. There was no adequate imagery data for the entire flowering period, so it may be the reason for the low accuracy of yield estimation. In both 2017 and 2018 at Saskatoon, imagery were collected semi-weekly (Table 3.1). For the 2 site years, the relationships between the UAV-derived flower accumulation and seed yield were relatively stronger compared to the 1st experimental year ($R^2 = 0.30$, $p < 0.05$ in 2017; $R^2 = 0.34$, $p < 0.05$ in 2018) (Figures 3.8B,C). At the Melfort and Scott locations in 2017, there were more consistent and stronger regressions

(Figure 3.9) using the UAV-derived flower accumulation, when compared with a single image date (Table 3.3).

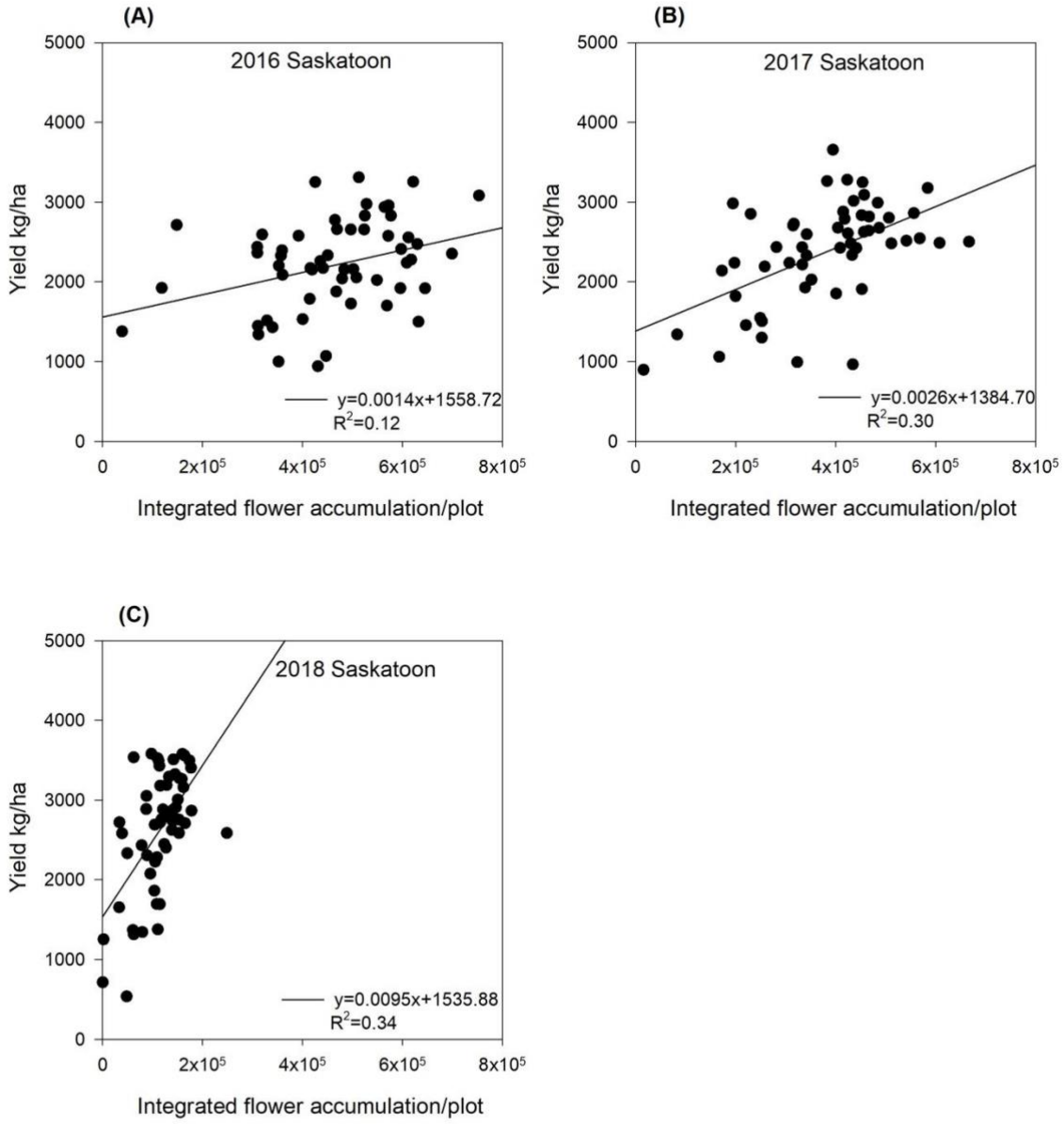


Figure 3.8 The relationship between seed yield and UAV-derived flower accumulation at Saskatoon, SK, Canada from 2016 to 2018. The UAV-derived flower accumulation was calculated using the area under flowering progress curve function. (A) Regression for the Saskatoon location in 2016: y=0.0014x+1558.72, R^2=0.12. (B) Regression for the Saskatoon location in 2017: y=0.0026x+1384.70, R^2=0.30. (C) Regression for the Saskatoon location in 2018: y=0.0095x+1535.88, R^2=0.34.

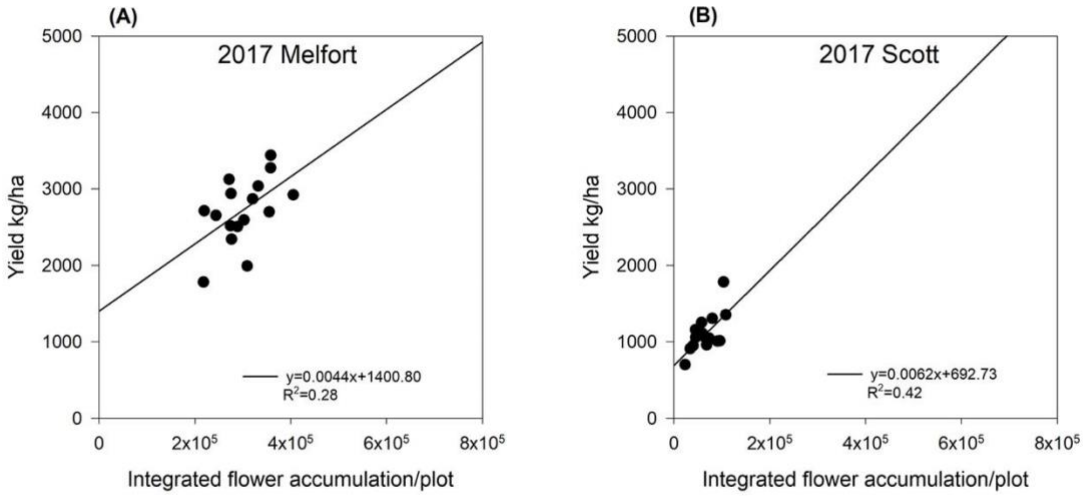


Figure 3.9 The relationship between seed yield and UAV-derived flower accumulation at Melfort and Scott, SK, Canada in 2017. The UAV-derived flower accumulation was calculated using the area under flowering progress curve function. (A) Regression for the Melfort location in 2017: y=0.0044x+1400.80, R^2=0.28. (B) Regression for the Scott location in 2017: y=0.0062x+692.73, R^2=0.42

In general, compared with using a single image, applying the integration of flowering progress to estimate yield includes more information to provide consistent accuracy (Figures 3.8, 3.9). Although the R^2 values for yield estimation were not high, the results still demonstrate the potential of AUFPC to predict yield, especially for crops producing bright flowers (e.g., canola and cotton) under different environmental conditions.

Several studies have reported similar results (Sulik and Long, 2016; Gong et al., 2018; Xu et al., 2017; Hassan et al., 2019). Sulik and Long (2016) reported that the NDYI values during flowering had high accordance with yield observations (R^2 = 0.76), which showed a better correlation with seed yield than NDVI at the peak flowering time in canola. Gong et al. (2018) found that NDVI multiplied by leaf-related canopy fraction had the strongest relationship with canola yield with low estimation errors (coefficient of variation < 13%) at the early flowering

stages. Some research also investigated yield estimation using canopy reflectance data in other crops including cotton and wheat (Xu et al., 2017; Hassan et al., 2019). Xu et al. (2017) reported that the estimated cotton flower number derived from aerial images using a convolutional neural network significantly correlated with cotton yield (R^2 = 0.36). Hassan et al. (2019) reported that UAV-based NDVI measured at the grain filling stage could be a promising tool for wheat yield prediction with R^2 ranging from 0.83 to 0.89 in field conditions.

Regression results had smaller R^2 values compared with the previous studies. This is probably associated with many diverse genotypes (i.e., 56 diverse genotypes) estimated in this study (Tables A.1, Appendix A). Most previous research only planted one or few genotypes. Stability of pigments in rapeseed flowers for each genotype may change under different developmental stages (Ohmiya, 2011). These factors can impact yellow to some degree (Ohmiya, 2011). The inconsistent yellowness may explain that the less model variation could be explained by the flower accumulation when more varieties were involved in regression analysis. Furthermore, flowering pixels extracted based on threshold values may not have been consistent over the flowering stage, as each threshold value was determined manually. In addition, canola yield components include pod number, seed number per pod, and seed weight. Although pod number per plant is highly correlated with seed yield (Tayo and Morgan, 1975; McGregor 1981; Ivanovska et al., 2007), only 45% of flowers produce pods (McGregor 1981). Seed weight per pod and thousand seed weight also significantly correlated with seed yield (Ivanovska et al, 2007). The simple regression analysis of flower numbers could not fully explain yield variation. Additional yield components considered in the yield estimation model would improve the accuracy of seed yield estimation. Finally, severe flower abortion and poor pod formation can happen under drought and heat stress during the crop season (Faraji et al, 2008). Flowering progress only reveals part of crop growth stages, so some varieties even with high AUFPC may end up with low yield under stress, which may result in a weaker relationship between the UAV-derived flower accumulation and seed yield. Combining UAV-based reflectance data at both flowering and pod stages may enhance yield estimation accuracy.

Usually, breeding programs need to assess a large number of varieties or breeding genotypes across multiple environmental conditions. Therefore, from a practical perspective, these results revealed a more realistic yield estimation trend for large-scale breeding programs. Moreover, most previous research used one image date or selected the largest reflectance index value for

each plot across all sampling dates to estimate crop yield. In fact, it is difficult to determine the best image date for yield estimation using multiple crop varieties grown in differing environmental conditions. Fluctuating environments can influence flowering progress; therefore, UAV-derived flower accumulation is a promising and predictable variable in the descriptive yield model.

3.4 Conclusions

In this study, a simple and effective approach was proposed to estimate relative flower numbers and model seed yield based on the cumulative flowering pixel. This study results showed that flowering pixel numbers estimated by the thresholding method significantly regressed with manual flower count during the flowering stage with R^2 value up to 0.95, indicating that flowering pixel numbers can be used as a good indicator of flower number in the field. Additionally, the integration of flowering progress from consecutive images *via* AUFPC math function was more consistently and strongly related to yield compared with using a single image date because cumulative flowering pixel over time utilizes more information of crop growth. Therefore, the UAV-derived flower accumulation can be a promising indicator for yield estimation. These tools do not require extra coding or strong computer science background, can be used for calculating thresholds and vegetation indices, and is a convenient tool for agronomists and breeders. Future studies need to consider and test a multivariate model including multiple vegetation indices related to other yield components and more reflectance information from the pod stage to improve yield estimation accuracy.

References

Araus, J. L., and Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. Trends in Plant Science, 19, 52-61. doi: 10.1016/j.tplants.2013.09.008

Blancon, J., Dutartre, D., Tixier, M. H., Weiss, M., Comar, A., Praud, S., et al. (2019). A high-throughput model-assisted method for phenotyping maize green leaf area index dynamics using unmanned aerial vehicle imagery. Frontiers in Plant Science, 10, 685. doi: 10.3389/fpls.2019.00685

Carl, C., Landgraf, D., der Maaten-Theunissen, V., Biber, P., and Pretzsch, H. (2017). *Robinia pseudoacacia* L. flower analyzed by using an unmanned aerial vehicle (UAV). Remote Sensing, 9, 1091. doi: 10.3390/rs9111091

Clayton, G., Turkington, K., Harker, N., O'Donovan, J., and Johnston, A. (2000). High yielding canola production. Better Crops, 84. 26-27. Available online at: http://www.ipni.net/publication/bettercrops.nsf/0/405DA9F381C7D59F852579800081FDE7/$FILE/Better%20Crops%202000-1%20p26.pdf (accessed January 05, 2020).

d'Andrimont, R., Taymans, M., Lemoine, G., Ceglar, A., Yordanov, M., and van der Velde, M. (2020). Detecting flowering phenology in oil seed rape parcels with Sentinel-1 and-2 time series. Remote Sensing of Environment, 239, 1-14. doi: 10.1016/j.rse.2020.111660

Diepenbrock, W. (2000). Yield analysis of winter oilseed rape (*Brassica napus* L.): a review. Field Crops Research, 67, 35-49. doi: 10.1016/S0378-4290(00)00082-4

Ebersbach, J., Khan, N. A., McQuillan, I., Higgins, E. E., Horner, K., Bandi, V., et al. (2021). Exploiting high-throughput indoor phenotyping to characterize the founders of a structured *B. napus* breeding population. Frontiers in Plant Science, 12, 780250-780250. doi: 10.3389/fpls.2021.780250

Fang, S., Tang, W., Peng, Y., Gong, Y., Dai, C., Chai, R. et al. (2016). Remote estimation of vegetation fraction and flower fraction in oilseed rape with unmanned aerial vehicle data. Remote Sensing, 8, 416. doi: 10.3390/rs8050416

Faraji, A. (2012). Flower formation and pod/flower ratio in canola (*Brassica napus* L.) affected by assimilates supply around flowering. International Journal of Plant Production, 4, 271-280. doi: 10.22069/ijpp.2012.710

Faraji, A., Latifi, N., Soltani, A., and Rad, A. H. S. (2008). Effect of high temperature stress and supplemental irrigation on flower and pod formation in two canola (*Brassica napus* L.) cultivars at Mediterranean climate. Asian Journal of Plant Sciences, 7, 343-351. doi: 10.3923/ajps.2008.343.351

Gan Y.T., Harker, K. N., Kutcher, H. R., Gulden, R. H., Irvine, B., May, W. E., et al. (2016). Canola seed yield and phenological responses to plant density. Canadian Journal of Plant Science, 96, 151-159. doi: 10.1139/cjps-2015-0093

Gitelson, A. A., and Merzlyak, M. N. (1997). Remote estimation of chlorophyll content in higher plant leaves. International Journal of Remote Sensing, 18, 2691-2697. doi: 10.1080/014311697217558

Gitelson, A. A., Merzlyak, M. N., and Lichtenthaler, H. K. (1996). Detection of red edge position and chlorophyll content by reflectance measurements near 700 nm. Journal of Plant Physiology, 148, 501-508. doi: 10.1016/S0176-1617(96)80285-9Get

Gong, Y., Duan, B., Fang, S., Zhu, R., Wu, X., Ma, Y., et al. (2018). Remote estimation of rapeseed yield with unmanned aerial vehicle (UAV) imaging and spectral mixture analysis. Plant Methods, 14, 1-14. doi: 10.1186/s13007-018-0338-z

Guo, W., Fukatsu, T., and Ninomiya, S. (2015). Automated characterization of flowering dynamics in rice using field-acquired time-series RGB images. Plant Methods, 11, 1-15. doi: 10.1186/s13007-015-0047-9

Han, J., Zhang, Z., and Cao, J. (2021). Developing a new method to identify flowering dynamics of rapeseed using landsat 8 and sentinel-1/2. Remote Sensing, 13, 105. doi: 10.3390/rs13010105

Hassan, M. A., Yang, M., Rasheed, A., Yang, G., Reynolds, M., Xia, X., et al. (2019). A rapid monitoring of NDVI across the wheat growth cycle for grain yield prediction using a multi-spectral UAV platform. Plant Science, 282, 95-103. doi: 10.1016/j.plantsci.2018.10.022

Haynes, K. G., and Weingartner, D. P. (2004). The use of area under the disease progress curve to assess resistance to late blight in potato germplasm. American Journal of Potato Research, 81, 137-141. doi: 10.1007/BF02853611

Ivanovska, S., Stojkovski, C., Dimov, Z., Marjanović-Jeromela, A., Jankulovska, M., and Jankuloski, L. (2007). Interrelationship between yield and yield related traits of spring canola (*Brassica napus* L.) genotypes. Genetika, 39, 325-332. doi: 10.2298/GENSR0703325I

Jeger, M. J., and Viljanen-Rollinson, S. L. H. (2001). The use of the area under the disease-progress curve (AUDPC) to assess quantitative disease resistance in crop cultivars. Theoretical and Applied Genetics, 102, 32-40. doi: 10.1007/s001220051615

Jiang, L., and Becker, H. C. (2003). Inheritance of apetalous flowers in a mutant of oilseed rape. Crop Science, 43, 508-510. doi: 10.2135/cropsci2003.5080

Kefauver, S. C., Vicente, R., Vergara-Díaz, O., Fernandez-Gallego, J. A., Kerfal, S., Lopez, A., et al. (2017). Comparative UAV and field phenotyping to assess yield and nitrogen use efficiency in hybrid and conventional barley. Frontiers in Plant Science, 8: 1733. doi: 10.3389/fpls.2017.01733

Kim, J., Kim, S., Ju, C., and Son, H. I. (2019). Unmanned aerial vehicles in agriculture: a review of perspective of platform, control, and applications. IEEE Access, 7, 105100-105115. doi: 10.1109/ACCESS.2019.2932119

Kirkegaard, J. A., Lilley, J. M., Brill, R. D., Ware, A. H., and Walela, C. K. (2018). The critical period for yield and quality determination in canola (*Brassica napus* L.). Field Crops Research, 222, 180-188. doi: 10.1016/j.fcr.2018.03.018

McGregor, D. I. (1981). Pattern of flower and pod development in rapeseed. Canadian Journal of Plant Science, 61, 275-282. doi: 10.4141/cjps81-040

Migdall, S., Ohl, N., and Bach, H. (2010). "Parameterisation of the land surface reflectance model SLC for winter rape using spaceborne hyperspectral CHRIS data," in Hyperspectral Workshop. (Vol. ESA SP-683). Frascati.

Montes, J. M., Melchinger, A. E., and Reif, J. C. (2007). Novel throughput phenotyping platforms in plant genetic studies. Trends in Plant Science, 12, 433-436. doi: 10.1016/j.tplants.2007.08.006

Ohmiya, A. (2011). Diversity of carotenoid composition in flower petals. Japan Agricultural Research Quarterly: JARQ, 45, 163-171. doi: 10.1016/j.tplants.2007.08.006

Prey, L., Hu, Y., and Schmidhalter, U. (2020). High-throughput field phenotyping traits of grain yield formation and nitrogen use efficiency: optimizing the selection of vegetation indices and growth stages. Frontiers in Plant Science, 10, 1672. doi: 10.3389/fpls.2019.01672

Qian, M., Fan, Y., Li, Y., Liu, M., Sun, W., Duan, H., et al. (2021). Genome-wide association study and transcriptome comparison reveal novel QTL and candidate genes that control petal size in rapeseed. Journal of Experimental Botany, 72, 3597-3610. doi: 10.1093/jxb/erab105

Rouse, J. W., Haas, R. H., Schell, J. A., Deering, D. W., and Harlan, J. C. (1974). Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. NASA/GSFC Type III Final Report, Greenbelt, Md, 371.

Sankaran, S., Khot, L. R., Espinoza, C. Z., Jarolmasjed, S., Sathuvalli, V. R., Vandemark, G. J., et al. (2015). Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: a review. European Journal of Agronomy, 70, 112-123. doi: 10.1016/j.eja.2015.07.004

Shen, M., Chen, J., Zhu, X., and Tang, Y. (2009). Yellow flowers can decrease NDVI and EVI values: evidence from a field experiment in an alpine meadow. Canadian Journal of Remote Sensing, 35, 99-106. doi: 10.5589/m09-003

Shen, M., Chen, J., Zhu, X., Tang, Y., and Chen, X. (2010). Do flowers affect biomass estimate accuracy from NDVI and EVI?. International Journal of Remote Sensing, 31, 2139-2149. doi: 10.1080/01431160903578812

Simko, I., and Piepho, H. P. (2012). The area under the disease progress stairs: calculation, advantage, and application. Phytopathology, 102, 381-389. doi: 10.1094/PHYTO-07-11-0216

Statistics Canada. (2018). Seeding decisions harvest opportunities for Canadian farmers. Available online at: https://www150.statcan.gc.ca/n1/pub/95-640-x/2016001/article/14813-eng.htm (accessed May 23, 2018).

Sulik, J. J., and Long, D. S. (2015). Spectral indices for yellow canola flowers. International Journal of Remote Sensing, 36, 2751-2765. doi: 10.1080/01431161.2015.1047994

Sulik, J. J., and Long, D. S. (2016). Spectral considerations for modeling yield of canola. Remote Sensing of Environment, 184, 161-174. doi: 10.1016/j.rse.2016.06.016

Taye, Z. M., Helgason, B. L., Bell, J. K., Norris, C. E., Vail, S., Robinson, S. J., et al. (2020). Core and differentially abundant bacterial taxa in the rhizosphere of field grown Brassica napus genotypes: implications for canola breeding. Frontiers in Microbiology, 10, 3007. doi: 10.3389/fmicb.2019.03007

Tayo, T. O., and Morgan, D. G. (1975). Quantitative analysis of the growth, development and distribution of flowers and pods in oil seed rape (*Brassica napus* L.). The Journal of Agricultural Science, 85, 103-110. doi: 10.1017/S0021859600053466

Tunca, E., Köksal, E. S., Çetin, S., Ekiz, N. M., and Balde, H. (2018). Yield and leaf area index estimations for sunflower plants using unmanned aerial vehicle images. Environmental Monitoring and Assessment, 190, 1-12. doi: 10.1007/s10661-018-7064-x

Wan, L., Li, Y., Cen, H., Zhu, J., Yin, W., Wu, W., et al. (2018). Combining UAV-based vegetation indices and image classification to estimate flower number in oilseed rape. Remote Sensing, 10, 1484. doi: 10.3390/rs10091484

White, J. W., Andrade-Sanchez, P., Gore, M. A., Bronson, K. F., Coffelt, T. A., Conley, M. M., et al. (2012). Field-based phenomics for plant genetics research. Field Crops Research, 133, 101-112. doi: 10.1016/j.fcr.2012.04.003

Wójtowicz, M., Wójtowicz, A., and Piekarczyk, J. (2016). Application of remote sensing methods in agriculture. Communications in Biometry and Crop Science, 11, 31-50.

Xu, R., Li, C., Paterson, A. H., Jiang, Y., Sun, S., and Robertson, J. S. (2018). Aerial images and convolutional neural network for cotton bloom detection. Frontiers in Plant Science, 8, 2235. doi: 10.3389/fpls.2017.02235

Xu, R., Paterson, A. H., and Li, C. (2017). Cotton flower detection using aerial color images. ASABE. In 2017 ASABE Annual International Meeting. American Society of Agriculture and Biological Engineers. doi: 10.13031/aim.201701080

Yates, D. J., and Steven, M. D. (1987). Reflexion and absorption of solar radiation by flowering canopies of oil-seed rape (*Brassica napus* L.). The Journal of Agricultural Science, 109, 495-502. doi: 10.1017/S0021859600081703

Zhang, H., and Flottmann, S. (2018). Source-sink manipulations indicate seed yield in canola is limited by source availability. European Journal of Agronomy, 96, 70-76. doi: 10.1016/j.eja.2018.03.005

Contributions by others to Chapter 3

Dr. Steve J. Shirtliffe and Dr. Sally Vail designed the field experiments. Dr. Sally Vail and Dr. Isobel A. P. Parkin provided and prepared the plant materials. Dr. Steve J. Shirtliffe contributed to funding acquisition. Ti Zhang, Dr. Hema S. N. Duddu, and Dr. Steve J. Shirtliffe collected the images and performed imagery pre-processing. Ti Zhang conducted imagery processing and ground data collection, performed the statistical analysis, and wrote the chapter under the supervision of Dr. Steve J. Shirtliffe and with contributions from Dr. Sally Vail, Dr. Hema S. N. Duddu, Mr. Eric N. Johnson, Dr. Xulin Guo, and Dr. Rosalind A. Bueckert.

Transition section between Chapter 3 and Chapter 4

Chapter 3 investigated the potential of two-dimensional vegetation indices extracted from UAV-based imagery to detect canola flower number and estimate seed yield. The normalized difference yellowness index-based pixel number could detect yellow flowers, and the accumulation of flowering pixels over time was a potential digital phenotype for yield estimation. The objective of the next study in Chapter 4 was to evaluate whether canopy structural information (i.e., flowering layer depth and canopy height) throughout the crop season can be estimated using three-dimensional digital surface models extracted from UAV-based imagery and has the potential to estimate seed yield in canola.

CHAPTER 4 QUANTIFYING CANOPY HEIGHT AND FLOWERING LAYER DEPTH TO ESTIMATE SEED YIELD IN CANOLA (*Brassica napus* L.) USING UNOCCUPIED AERIAL VEHICLE-BASED MULTI-TEMPORAL IMAGERY

The content of this chapter will be submitted to Remote Sensing.

Abstract

Unoccupied aerial vehicles (UAVs) mounted with imaging sensors have become a useful phenotyping approach for assessing crop traits *in situ* because of their rapid, low-cost, and non-destructive manner. Plant height and flowering traits are common field measurements for analyzing plant growth. In tradition, manual plant height measurement and visual rating for estimating flowering number are labor-intensive, destructive, and subjective. To overcome the limitations, the objectives were to evaluate the potential of UAV-based multispectral imagery to quantify canopy height and flowering layer depth (i.e., vertical distance from the bottom flower to the top flower), and to investigate whether the digitalized crop traits can be seed yield predictors. This study was conducted using fifty-six diverse Brassica genotypes from 2016 to 2018, near Saskatoon, SK, Canada. The results demonstrated that there were significant relationships between actual and UAV-derived canopy heights during reproductive stage with coefficient of determination (R^2) ranging from 0.67 to 0.90. The peak flowering time and pod stage were the optimum growth stages for canopy height estimation with the highest R^2 values. Significant relationships between the manually measured and UAV-derived flowering layer depth were observed with R^2 ranging from 0.13 to 0.42. Additionally, compared with using the crop information extracted from a single image date, there were more consistent and significant regressions between the digitalized cumulative canopy height or cumulative flowering layer depth over the reproductive period and seed yield (R^2 up to 0.46 and 0.34, respectively). This study indicates that UAV-based multispectral imagery can estimate canopy height and flowering layer depth with high to moderate accuracy. The digitalized cumulative canopy height and flowering layer depth at the flowering stage can be potential indicators for seed yield estimation.

4.1 Introduction

Canola (*Brassica napus* L.) is an important vegetable oil source grown worldwide. Crop yield is one of the most crucial targets for canola growers. Accurate crop yield forecasts enable farmers to take effective crop management decisions during the growing season to optimize pest control and crop nutrient supply (Peng et al., 2019). In addition, early field evaluation of crops depending on secondary traits is now common for high-yielding genotype selection in plant breeding programs (Hassan et al., 2019).

Canola leaves are the major plant structures for photosynthesis. Stem also plays an important role in photosynthesis through pod and seed development stages. Plant height is an important indicator of crop stand and yield estimation (Ma et al., 2015; Ivanovska et al., 2007; Bendig et al., 2014; Iqbal et al., 2017; Assefa et al., 2018; Zhang et al., 2021a). Plant height was found positively correlated with seed yield in Brassica genotypes (Ivanovska et al., 2007; Assefa et al., 2018). Usually, maximum stem length overlaps the peak flowering time (Canola Council of Canada, 2021). Crop lodging is determined as the aboveground part of plant falling over, which is a key factor affecting canola yield potential (Wu and Ma, 2016, 2018). Normally, lodged plants make the canopy closer to the ground, which can impede harvest and lower yield as it can cause uneven pod maturity and increase diseases severity (Wu and Ma, 2016, 2018). Thus, crop canopy height measurement during the whole growing season can be an indicator of plant growth status and may assist in crop management decision-making and yield estimation (Ivanovska et al., 2007; Bendig et al., 2014; Iqbal et al., 2017; Assefa et al., 2018).

Flowering number in canola is an important phenotype that determines seed yield potential (Tayo and Morgan, 1975; Diepenbrock, 2000; Faraji et al., 2008; Faraji, 2012; Fang et al., 2016; Sulik and Long, 2016; Gong et al., 2018; Kirkegaard et al., 2018; Zhang and Flottmann, 2018; Zhang et al., 2021b). Accurate yield estimation using flowering traits may allow canola breeders to select the high-yielding genotypes at the early reproductive growth stages. It could permit selection for a yield proxy in years when harvest environments are unfavorable.

Plant height is normally measured manually from the bottom of the crop to the top of the highest point of plant using a meter-ruler, which is time-consuming and labor intensive. Flowering traits are visually assessed which is time-consuming and subjective, especially when

assessing many genotypes under multiple field conditions (Sulik and Long, 2015; Fang et al., 2016; Wan et al., 2018; Zhang et al., 2021b). In breeding programs, flowering is typically limited to being recorded as the date of commencement, and may not be recorded in large nursery trials.

With the advancement of UAVs, crop information obtained by remote sensors at high spatial and temporal resolutions has attracted more attention. A commonly used approach to estimate yield is applying a vegetation index calculated using spectral reflectance bands in the visual region (blue, green, and red) and near IR (NIR). Studies have shown significant relationships between the vegetation indices and seed yield in corn (Ramos et al., 2020), rice (Zhou et al., 2017), sunflower (Tunca et al., 2018), and canola (Sulik and Long, 2016; Zhang et al., 2021b).

An alternative method to estimate yield is to model it using canopy height information, extracted from crop surface models (CSMs) (Geipel et al., 2014; Bendig et al., 2014 and 2015; Iqbal et al., 2017; Panday et al., 2020). Structure from motion (SfM) and multi-view stereopsis techniques has enabled efficient production of dense point clouds (Bendig et al., 2015), which can be used to develop CSMs and detailed orthoimages (Bendig et al., 2015). This detailed three-dimensional (3D) model recreation enables a new and realistic view of crop growth and health. In addition, it may provide more crop information (i.e., dynamics of plant height throughout the crop season) when there is no significant spectral difference in two-dimensional (2D) reflectance maps (Bendig et al., 2015). The method of using CSMs derived from 3D point clouds has shown the suitability to estimate plant height for biomass and seed yield estimation in various crops such as corn, barley, poppy crop (*Papaver somniferum* L.), and wheat (Geipel et al., 2014; Bendig et al., 2014 and 2015; Iqbal et al., 2017; Panday et al., 2020). Geipel et al. (2014) studied yield modeling using plant height information extracted from UAV-based imagery in corn. They found yield could be estimated using digital height information after completion of canopy closure. Bendig et al. (2014 and 2015) found that plant height derived from UAV-based CSMs was significantly correlated with dry biomass in barley. Iqbal et al. (2017) reported significant correlations between UAV-extracted plant height and ground-measured plant height (R^2 = 0.92 - 0.97) in poppy crop. Their study also indicated the UAV-based plant height was a reliable indicator for capsule volume of poppy crop at the productive stage. Panday et al. (2020) estimated a linear regression between plant height derived from

CSMs and grain yield of wheat. This study showed a significant relationship between grain yield and the plant height obtained from CSMs with R^2 of 0.70.

Many researchers have evaluated canola flowering traits by analyzing canopy reflectance from 2D orthoimages (Sulik and Long, 2015, 2016; Gong et al., 2018; Wan et al., 2018; Zhang et al., 2021b). Sulik and Long (2015, 2016) and Zhang et al. (2021b) found that there were significant linear relationships between the normalized difference yellowness index (NDYI) and flower numbers, and the NDYI could be a potential yield indicator. Similarly, Fang et al. (2016), Gong et al. (2018), and Wan et al. (2018) reported that the combination of vegetation index and sophisticated machine learning methods or pixel-level spectral mixture analysis improved the estimates of flower numbers and crop yield. Although the 2D reflectance maps provide information for flowering number estimation, the measurement of flowering layer depth using 3D information may be an alternative approach to estimate flowering volume (i.e., flowering volume = plot area x flowering layer depth), which may have a similar growth pattern as flowering number. Flowering layer depth refers to the difference between the top flowering height and the lowest flowering height (Figure 4.1). Instead of using the top view of 2D canopy reflectance information, 3D image reconstruction using the SfM and multi-view stereo technique can measure crop canopy structure using high vertical resolution. Tayo and Morgan (1975) reported the lowest flower buds on the main stem were the first to become yellow. Flowering then continued acropetally. Usually, the flowers on the main stem and the upper four axillary racemes will develop into productive pods at maturity. Flowering layer depth could be estimated by using 3D data. Vanbrabant et al. (2020) found the 3D method showed a higher accuracy of flower cluster estimation in pear ($R^2 > 0.70$) than 2D method with R^2 of 0.53.

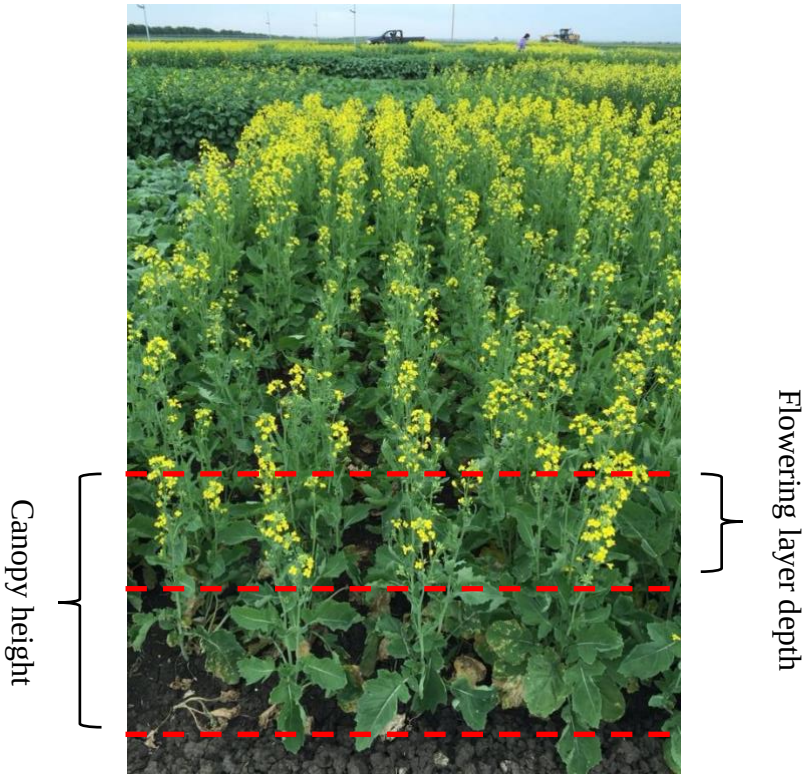


Figure 4.1 Canopy height refers to the vertical distance from the base of a plant to the maximum point of crop canopy. Flowering layer depth is the difference between the top flowering height and lowest flowering height.

Some studies have investigated the estimation accuracy of plant height using UAV-based imagery or LiDAR. However, little research has studied the application of canopy height information in estimating canola yield using UAV-based imagery. Additionally, the measurement of flowering layer depth, a potential indicator of plant growth condition, is a new phenotype combining 2D and 3D information which may assist flowering number estimation, and therefore yield prediction. Thus, the objectives of this study were to investigate the ability of UAV-based imagery to estimate canola canopy height and flowering layer depth, and to assess the potential of digital canopy height information and flowering layer depth to estimate yield.

4.2 Materials and Methods

4.2.1 Study Area

The field study was conducted near Saskatoon from May 2016 to September 2018 (52° 10' 52.9" N, 106° 30' 10.6" W in 2016; 52° 10' 59.3" N, 106° 30' 53.7" W in 2017; and 52° 10' 57.7" N, 106° 30' 01.4" W in 2018). Fifty-six diverse genotypes (Ebersbach et al., 2022; Table A.1, Appendix A) were planted in rectangular lattice design with three replicates at a desired rate of 108 seeds m^{-2}. Out of the 56 genotypes, 16 were selected and planted in two adjacent but separate plots as double plots, as described in Chapter 3. The dimension of plot sizes was 1.2 m wide by 6.0 m long in 2016 and 2018, and 1.5 m wide by 6.0 m long in 2017. Canola plots were seeded on May 27, 28, and 21 in 2016, 2017 and 2018, respectively (Table 4.1). Pesticides were applied based on requirement. Weeds not controlled by the pesticides in field were manually removed.

Table 4.1 Details of canola trials and data acquisition (image collection, actual canopy height and flowering layer depth measurement dates) near Saskatoon, SK, Canada from 2016 to 2018.

Year	Seeding date	Flight altitude (m)	Image collection date	Height measurement date	Flowering layer depth measurement date
2016	May 27	20	June 13; 30		
			July 14; 19; 26	July 19; 27	July 19; 27
			August 06; 22	August 11	
			September 08	September 11	
2017	May 28	20	June 28		
			July 07; 15; 19; 22; 26	July 04	July 25
			August 09; 16; 22; 29	August 14	
2018	May 21	25	June 28		
			July 06; 09; 16; 20; 24; 27; 30	July 10; 17; 25; 31	July 10; 17; 25
			August 03; 07; 14; 20; 28	August 07; 14; 20; 28	

4.2.2 Image Acquisition

A Draganflyer X4-P multirotor UAV (DraganFly Inc., Saskatoon, SK, Canada) was used for imagery collection in 2016 and 2017. The other UAV named Draganflyer Commander (DraganFly Inc., Saskatoon, SK, Canada) was used to acquire imagery in 2018. A multi-

spectral camera (RedEdge, MicaSense Inc., Seattle, WA, United States) was mounted on the UAVs to collect aerial imagery (12-bit image) of blue (475 ± 10 nm), green (560 ± 10 nm), red (668 ± 5 nm); red edge (717 ± 5 nm), and near-infrared (840 ± 20 nm) bands with an image resolution for each band of 1.2 megapixels (1,280 x 960 pixels). The UAV flight mission was programmed using Draganfly Surveyor software (DraganFly Inc., Saskatoon, SK, Canada). The UAVs were flown over the canola trials at a ground speed of 1.7 meter per second and at 20 and 25 m above ground level which resulted in a ground sampling distance of 1.36 and 1.70 cm per pixel, respectively. During the flight mission, all aerial images were collected at 80% front and 80% side overlaps. A MicaSense calibration target was imaged before and after each flight mission for radiometric calibration. Ground control points (GCPs) were evenly placed in the experiment area for georeferencing throughout the crop season. GCPs were geolocated by Trimble GeoExplorer 2008 GPS (Trimble Inc., Westminster, CO, United States).

The UAV-based multispectral imagery was collected at a 7-day interval during the field season in 2016, at 3-day interval in 2017 and 2018 during the flowering period, and at 7-day interval at the vegetative and pod stages from 2017 to 2018 (Table 4.1).

4.2.3 Image Processing and Data Extraction

4.2.3.1 Image pre-processing

The UAV-based multispectral imagery was pre-processed in Pix4Dmapper Pro (Pix4D Inc., San Francisco, CA, United States). The basic processing steps include (1) initial processing for the photo alignment, (2) generation of a dense point cloud and mesh, and (3) generation of digital surface models (DSMs) and orthomosaic (Duddu et al., 2019; Zhang et al., 2021b).

(1) In the initial processing step, images were imported and stitched using features extraction from the overlapped images. Sparse point cloud was created after the key point extraction and image matching. Radiometric calibration was done by using the images of the MicaSense calibration target. To georeference the point cloud, the geo-location of GCPs was imported. Bundle block adjustment was applied to adjust the position and orientation of photos.

(2) In the second step, a dense point cloud and mesh were created based on the sparse point cloud. To generate suitable 3D point cloud, important parameters were established (i.e., image scale: half image size; point density: optimal; the minimum number of matches: 3). The image scale parameter in Pix4D is keypoints image scale, which refers to the image size where additional 3D points are computed. The point density is a parameter used to determine the point cloud density, and the minimum number of matches per 3D point means the minimum number of images that each 3D point needs to be correctly re-projected in (Maturbongs et al., 2019). In this study, the default processing parameters in Pix4d were used to generate a high-quality point cloud with less noise and artifacts.

(3) In the last step, the DSMs of the experiment area was generated, which was exported for further canopy height extraction and flowering layer depth determination. As the second product, calibrated orthomosaic maps (i.e., calibrated reflectance maps) were generated and exported for vegetation index calculations. The spatial resolution for the two products (i.e., DSM and reflectance map) was 1.36 cm per pixel at 20m and a 1.70 cm per pixel resolution at 25m.

4.2.3.2 Canopy height extraction and accumulation of canopy height dynamics

After image pre-processing, the generated DSMs and calibrated reflectance maps were imported into the ArcGIS version 10.4.1 (ESRI Canada, Toronto, ON, Canada) for canopy height estimation.

The calibrated reflectance maps of red, green, and blue were used to generate composited RGB image using the “Composite Band” tool for identifying plant and bare soil area. To segment the area of interest for each experimental plot and the adjacent bare soil surface, polygon shapes were generated using the “Create Feature” tool. Then, the generated plant and soil polygons were applied to the DSMs. The mean height value of plant area for each plot and corresponding adjacent soil surface were extracted using the “Zonal Statistics” tool. The UAV-derived canopy height was determined by subtracting the DSM of the bare soil surface from the DSM of the corresponding plot area:

$$UAV\text{-}derived\ canopy\ height = DSM_{crop} - DSM_{soil} \quad [4.1]$$

Where DSM_{crop} and DSM_{soil} represent crop canopy and neighboring bare soil surface heights, respectively.

Cumulative UAV-derived canopy height refers to an overview of canopy height dynamics throughout the entire crop season. It reveals both the magnitude of canopy height and duration of height change. To combine dynamics of canopy height over the whole crop season, cumulative UAV-derived canopy height was computed using the area under the flowering curve formula (Equation 3.5), which has been used to improve the accuracy of yield estimation in canola (Zhang et al., 2021b).

4.2.3.3 Flowering layer depth determination

4.2.3.3.1 Flowering layer depth determination using the thresholding method

To exclude the flowering area from non-flowering area, the NDYI was calculated using the "Raster Calculator" tool in the ArcGIS version 10.4.1. NDYI has shown its ability to efficiently separate canola flowers from vegetative or soil surface area (Sulik and Long, 2016; d'Andrimont et al., 2020; Han et al., 2021; Zhang et al., 2021b). It was calculated as following formula:

$$\boldsymbol{NDYI} = \left(\frac{R_{green} - R_{blue}}{R_{green} + R_{blue}}\right) \quad [4.2]$$

Where R_{green} and R_{blue} represent green and blue bands on 560 nm and 475 nm, respectively.

The output NDYI maps were grouped into flowering pixel and non-flowering pixels by the thresholding method. The details of selecting and segmenting flowering pixels for each plot by applying threshold values to the NDYI maps were described by Zhang et al. (2021b). The segmented flowering pixels with pixel values larger than threshold values were converted into a value of "1", and the non-flowering pixels with pixel values below threshold values were converted into "no data" using the "Reclassify" tool. The threshold NDYI maps became a

binary layer with “1” and “no data”, which were combined with DSMs to extract the height of flowering area. Similarly, to estimate the height of non-flowering area, the flowering pixels were converted into “no data”, and non-flowering pixels were converted into “1”. The non-flowering binary layer was combined with DSMs. Finally, flowering pixel height and non-flowering pixel height were subtracted to determine the flowering layer depth. The mean values of flowering height and corresponding non-flowering height per plot were extracted using the “Zonal Statistics” tool. The flowering layer depth using the threshold method was $FL_{threshold}$, calculated using the following formula:

$$FL_{threshold} = DSM_{flowering\ area} - DSM_{non\text{-}flowering\ area} \quad [4.3]$$

where $DSM_{flowering\ area}$ and $DSM_{non\text{-}flowering\ area}$ represent the mean height of flowering area and non-flowering area, respectively.

4.2.3.3.2 Flowering layer depth determination using the UAV-derived canopy height difference between the bolting and flowering stages

The bolting stage in canola begins when canola flower buds are formed and are first visible in the rosette of leaves. The flower buds quickly become yellow after bolting. Flowering duration is the period from the first opening bud on the main raceme until all flowers developed into pods. Canopy heights at the bolting and flowering stages were calculated using the overall plot mean height of the DSMs at the bolting and flowering stages. The flowering layer depth using the mean height difference was FL_{mean}, and it was calculated using the following formula:

$$FL_{mean} = DSM_{canopy\ height\ at\ flowering\ stage} - DSM_{canopy\ height\ at\ bolting\ stage} \quad [4.4]$$

where $DSM_{canopy\ height\ at\ flowering\ stage}$ represents UAV-derived canopy height at the flowering stage, and $DSM_{canopoy\ height\ at\ bolting\ stage}$ represents UAV-derived canopy height at the bolting stage.

4.2.3.3.3 Flowering layer depth determination using the maximum height values at the bolting and flowering stages

Flowering layer depth refers to the vertical distance from the top flower to the bottom flower of a plant. Usually, the top flower for each plant is the highest point. In this method, instead of using the mean height value at plot-level, canopy heights were estimated using the maximum height value at the bolting and flowering stages, which is $FL_{maximum}$. It was calculated using the following formula:

$$FL_{maximum} = DSM_{maximum\ height\ at\ flowering\ stage} - DSM_{maximum\ height\ at\ bolting\ stage} \quad [4.5]$$

where $DSM_{maximum\ height\ at\ flowering\ stage}$ and $DSM_{maximum\ height\ at\ bolting\ stage}$ represent the maximum height at the flowering and bolting stages.

4.2.4 Ground Reference Measurement

Canopy height refers to the vertical distance from the base of a plant to the highest point of crop canopy (Figure 4.1). Canola flowering layer depth refers to the vertical distance from the bottom flower to the top flower of a plant (Figure 4.1). Three plants were randomly selected from the front, middle, and back of each plot, and manually measured *in situ* with a meter-ruler for canopy height and flowering layer depth determination. Canopy height was measured from July to early September for 56 diverse genotypes in 2016 and 2017, and the selected 16 genotypes in 2018 (Table 4.1). Flowering layer depth was measured in July for 56 diverse genotypes in 2016, and the selected 16 genotypes in 2017 and 2018 (Table 4.1). There were no adequate summer students to manually measure canopy height and flowering layer depth for all 56 genotypes; thus, the ground reference data were collected for the selected 16 genotypes in the 2017 and 2018 seasons. The criteria of genotype selection was described in Chapter 3.

To reduce border effect, the middle four rows of each experimental plot were straight combined by a small plot combine harvester when the crop was mature and dry. This occurred multiple times due to differing maturity dates of the Brassica genotypes. All harvested seeds were air-dried to 10% seed moisture. Final seed yields were determined after air-drying and seed cleaning.

4.2.5 Statistical Analysis

PROC LATTICE in SAS 9.4 (SAS institute Inc., Cary, NC) was applied to investigate and adjust data to reduce variations within blocks (SAS version 9.4, SAS institute Inc., Cary, NC). Mean values per Brassica genotype were computed across three replications for ground reference data (i.e., seed yield and manual canopy height and flowering layer depth) and image-based features (i.e., UAV-derived canopy height and flowering layer depth) within each year. These mean values were used to evaluate relationships between ground reference data and image-based features by PROC REG (SAS 9.4).

4.3 Results and Discussion

4.3.1 Canopy Height Estimation Using UAV-based Imagery

Across 3 site years, there were significant relationships between the ground measured and UAV-derived canopy heights with coefficient of determination (R^2) ranging from 0.24 to 0.90. The peak flowering time and pod stage achieved high R^2 values between the manually measured and UAV-derived canopy heights when canola usually reached its tallest height.

In 2016, there were significant relationships between the actual and UAV-derived canopy heights at reproductive stage with R^2 values ranging from 0.74 to 0.90 (Figure 4.2). Compared with the early flowering and late pod stages, the peak flowering time and the early pod stage had the strongest relationships with R^2 value of 0.80 and 0.90, respectively (Figures 4.2B,C). Not surprisingly, canola had not reached the maximum plant height at the early flowering stage. The relatively short height might result in less accurate estimation *via* UAV-imagery (Figure 4.2A). The R^2 value (0.76) between the actual and UAV-derived canopy heights decreased at the late pod stage (Figure 4.2D). For this site year, plant lodging occurred at late development stage, which resulted in shorter and uneven canopy height. The reduced canopy height caused by lodging might have resulted in decreased estimation accuracy because of relatively lower spatial resolution compared with taller plants. In addition, it was difficult to accurately determine the average canopy height for non-uniform plots by manually measuring limited samples (i.e., three plant samples from the front, middle and back of each plot) (Wilke et al.,

2019; Morrison et al., 2021). In contrast, canopy height estimation *via* imagery could cover larger sampling area than manual measurement, which indicates imagery may reveal more realistic data than manual measurement (Wilke et al., 2019; Morrison et al., 2021). During the crop season in 2016, UAV-based imagery overestimated canopy height at the early development stage (slope > 1) (Figures 4.2A,B), and underestimated canopy height when crop grew taller at the late growth stage with slope less than one (Figures 4.2C,D). Although the slopes over the season were not consistently larger or less than one, the absolute values of the slopes were close to one with intercepts close to zero, which indicates UAV-derived canopy height was close to the actual canopy height.

Similar results were observed in 2017 and 2018. The flowering and pod stages showed significant relationships between the actual and UAV-derived canopy heights with R^2 ranging from 0.67 and 0.87 (Figures 4.3, 4.4). Usually, a canola plant has reached maximum height at the peak flowering stage. As mentioned above, taller plants could be easily detected *via* imagery. Therefore, a relationship with higher R^2 can be expected after the full flowering stage compared with the early development stage. The bolting stage showed low R^2 values between the actual and UAV-derived canopy heights in 2017 and 2018 (Figures 4.3A, 4.4A). Moreover, the actual canopy height tended to be underestimated at the early development stage in 2017 and 2018 (slop = 0.85 and 0.68, respectively) (Figures 4.3A, 4.4A), which could be resulted from less accurate estimation of smaller plant and incompletely closed canopies at the early development stage compared with the flowering and pod stages. Canopy height was overestimated by the UAV-based imagery at the flowering and pod stages with a slope ranging from 1.14 to 3.17 in 2018, which was different from the height estimation in 2016 and 2017 (Figures 4.2-4.4). It could be resulted from the shorter canopy height in 2018 over the growing season compared with the heights in the first two study years (Figures 4.2-4.4).

In general, there were significant relationships between the actual and UAV-derived canopy heights at both peak flowering and early pod stages when canola had reached full plant height at peak flowering. Although significant relationships between the ground measured and UAV-derived canopy heights were observed at the early flowering and late pod stages, the R^2 values were less compared with the peak flowering and early pod stages. Prior to the peak flowering stage, the raceme continues to elongate and plant height is relatively short. The results indicate that taller plants could be more easily and accurately estimated by UAV-based imagery. These

findings were consistent with a study by Iqbal et al. (2017) who found that the relationship between the ground measured and estimated heights had a greater R^2 for taller plants. In addition, the plant height is also very dynamic before peak flowering. The time gap between the dates of ground reference data collection and taking imagery might also reduce the accuracy of height estimation (Table 4.1). Apart from plant size, crop lodging may also affect the accuracy of height estimation due to non-uniform canopy (Iqbal et al., 2017; Wilke et al., 2019; Morrison et al., 2021). Iqbal et al. (2017) reported that the correlation between yield and estimated plant height was weak when poppy crop development was not uniform within a plot. UAV-based imagery tended to overestimate the canopy height of short plots and underestimate the tall plots. Morrison et al. (2021) also reported overestimation for the short canopy and underestimation for the tall canopy at the early growing season in soybean and wheat.

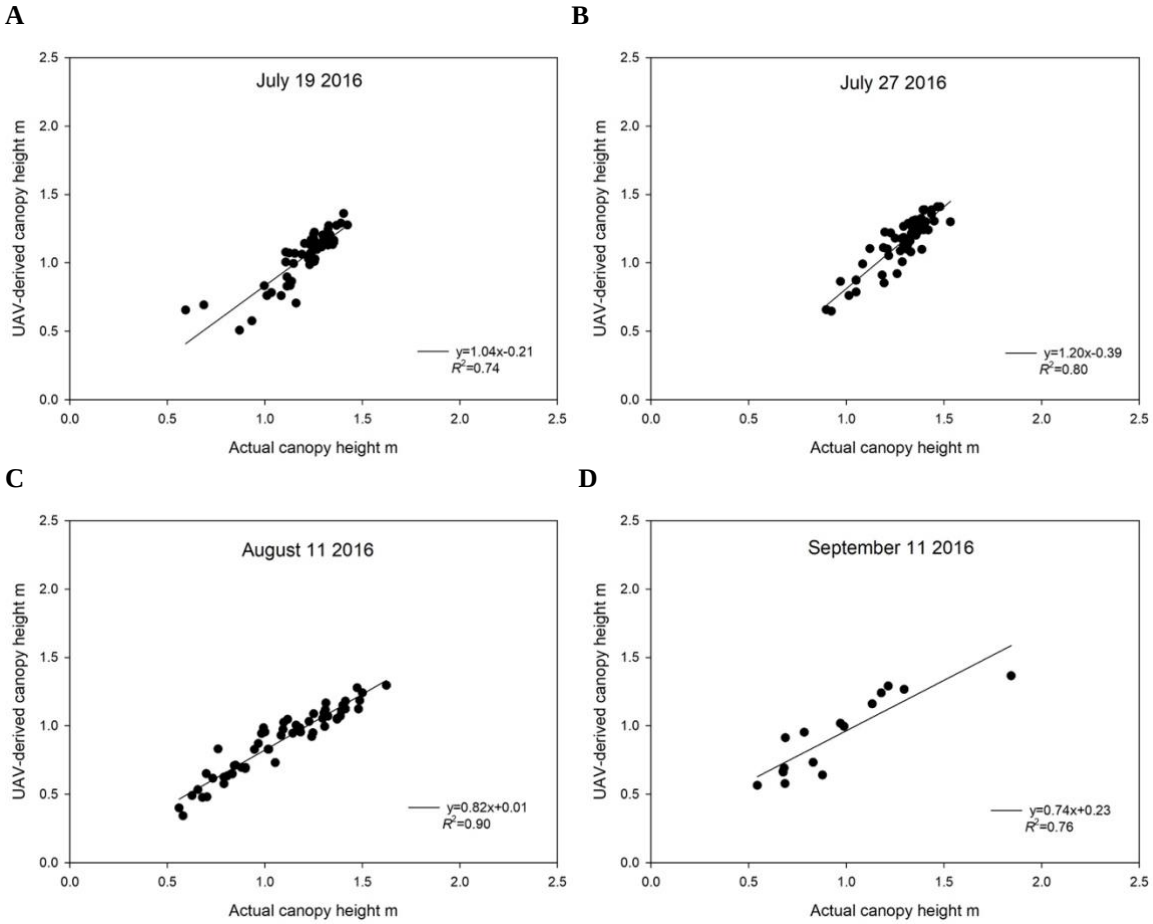


Figure 4.2 Relationships between the UAV-derived canopy height and actual canopy height during the reproductive stage at Saskatoon, SK, Canada in 2016. Actual canopy height per plot were manually measured by averaging the height of three randomly selected plants from the front, middle, and back of each plot. UAV-derived canopy height was extracted from the overall plot mean height of middle three rows for each plot. (A) Regression equation for July 19, 2016: y=1.04x-0.21, R^2=0.74. (B) Regression equation for July 27, 2016: y=1.20x-0.39, R^2=0.80. (C) Regression equation for August 11, 2016: y=0.82x+0.01, R^2=0.90. (D) Regression equation for September 11, 2016: y=0.74x+0.23, R^2=0.76.

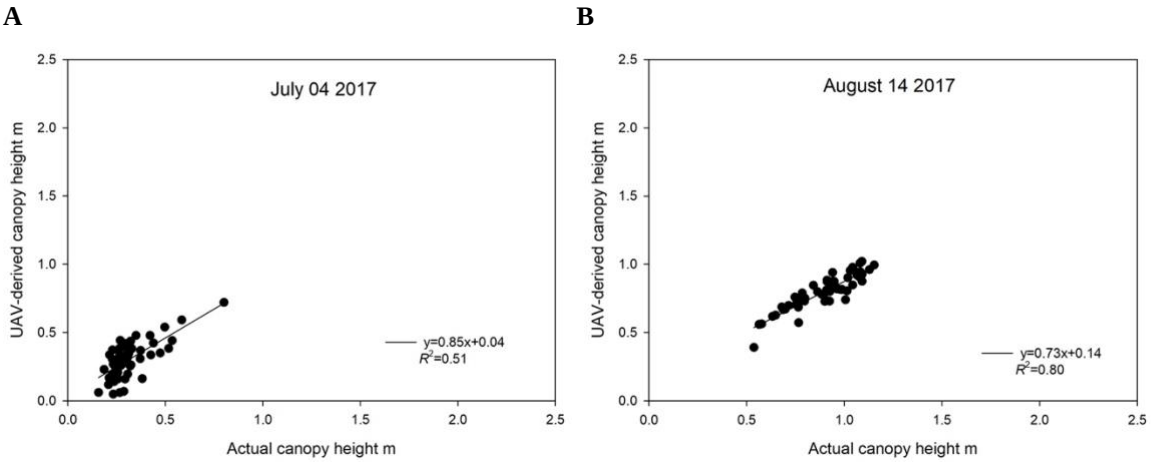
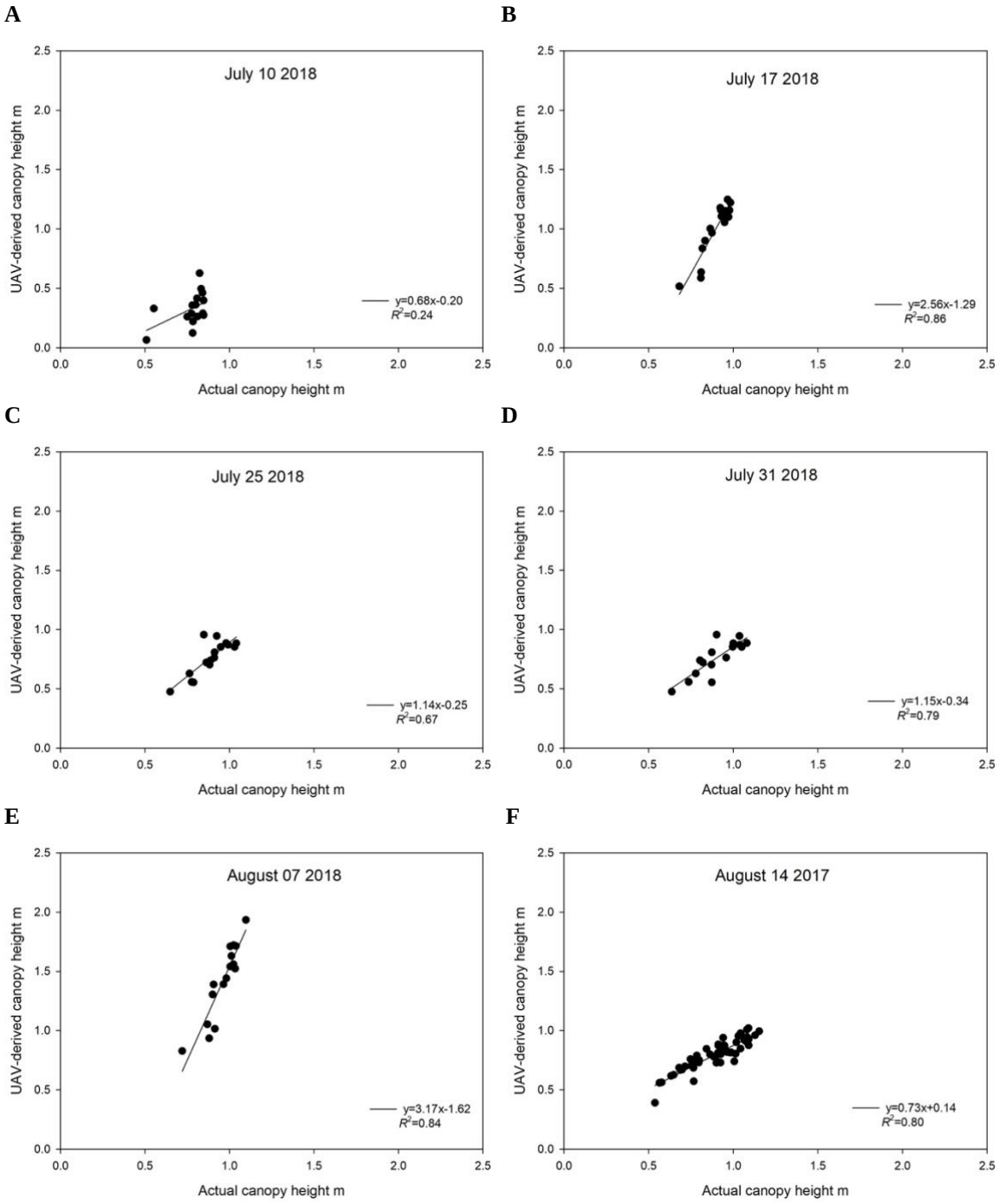


Figure 4.3 Relationships between the UAV-derived canopy height and actual canopy height at Saskatoon, SK, Canada in 2017. Actual canopy height per plot were manually measured by averaging the height of three randomly selected plants from the front, middle, and back of each plot. UAV-derived canopy height was extracted from the overall plot mean height of middle three rows for each plot. (A) Regression equation for July 04, 2017: y=0.85x+0.04, R^2=0.51. (B) Regression equation for August 14, 2017: y=0.73x+0.14, R^2=0.80.


A
July 10 2018
y=0.68x-0.20
R2=0.24
B
July 17 2018
y=2.56x-1.29
R2=0.86
C
July 25 2018
y=1.14x-0.25
R2=0.67
D
July 31 2018
y=1.15x-0.34
R2=0.79
E
August 07 2018
y=3.17x-1.62
R2=0.84
F
August 14 2017
y=0.73x+0.14
R2=0.80
UAV-derived canopy height m
Actual canopy height m
0.0
0.5
1.0
1.5
2.0
2.5

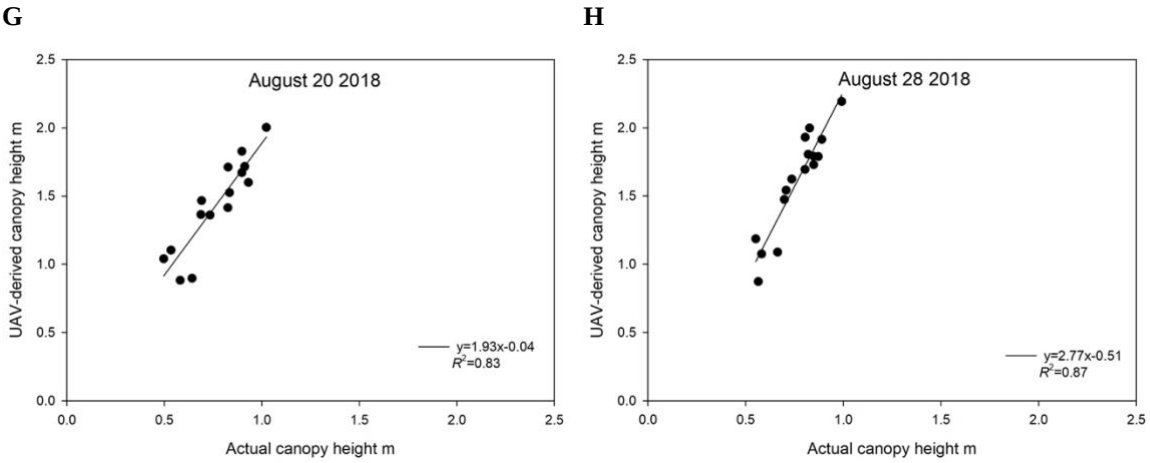


Figure 4.4 Relationships between the UAV-derived canopy height and actual canopy height at Saskatoon, SK, Canada in 2018. Actual canopy height per plot were manually measured by averaging the height of three randomly selected plants from the front, middle, and back of each plot. UAV-derived canopy height was extracted from the overall plot mean height of middle three rows for each plot. (A) Regression equation for July 10, 2018: y=0.68x-0.02, R^2=0.24. (B) Regression equation for July 17, 2018: y=2.56x-1.29, R^2=0.86. (C) Regression equation for July 25, 2018: y=1.14x-0.25, R^2=0.67. (D) Regression equation for July 31, 2018: y=1.15x-0.34, R^2=0.79. (E) Regression equation for August 07, 2018: y=3.17x-1.62, R^2=0.84. (F) Regression equation for August 14, 2018: y=2.13x-0.54, R^2=0.69. (G) Regression equation for August 20, 2018: y=1.93x-0.04, R^2=0.83. (H) Regression equation for August 28: y= 2.77x-0.51, R^2=0.87.

4.3.2 Relationships Between Ground Measured and UAV-derived Flowering Layer Depth

Simple linear and quadratic regression models were tested between the actual flowering layer depth (FL_{actual}) and flowering layer depth extracted from UAV-based imagery. According to the coefficient of determination, there were significant relationships between the FL_{actual} and FL_{mean} in most cases (Table 4.2). However, $FL_{threshold}$ was not significantly regressed with the FL_{actual}. Strong relationships between the FL_{actual} and the $FL_{maximum}$ were not observed except at the early flowering stage (July 19, 2016 and July 10, 2018) (Table 4.2). In addition, compared with the linear regression model, a quadratic regression model fits the dataset better (Table 4.2). Therefore, quadratic regressions between the FL_{actual} and the FL_{mean} would be further discussed.

Table 4.2 Coefficient of determination (R^2) between the actual and UAV-derived flowering layer depth at Saskatoon from 2016 to 2018.

Year	Date	$FL_{threshold}$		FL_{mean}		$FL_{maximum}$	
		Regression	R^2	**Regression**	R^2	**Regression**	R^2
2016	July 19	Linear	0.04	Linear	0.09*	Linear	0.06
		Quadratic	0.05	Quadratic	0.40***	Quadratic	0.47***
	July 27	Linear	0.002	Linear	0.09*	Linear	0.09*
		Quadratic	0.01	Quadratic	0.13*	Quadratic	0.16*
2017	July 25	Linear	0.07	Linear	0.23	Linear	0.18
		Quadratic	0.02	Quadratic	0.36*	Quadratic	0.34
2018	July 10	Linear	0.11	Linear	0.13	Linear	0.07
		Quadratic	0.14	Quadratic	0.13	Quadratic	0.37*
	July 17	Linear	0.05	Linear	0.25*	Linear	0.10
		Quadratic	0.05	Quadratic	0.34*	Quadratic	0.13
	July 25	Linear	0.21	Linear	0.39*	Linear	0.10
		Quadratic	0.22	Quadratic	0.42*	Quadratic	0.23

* Significant at the 0.05 probability level.
** Significant at the 0.01 probability level.
***Significant at the 0.001 probability level.

In 2016, there was a significant relationship between FL_{actual} and FL_{mean} at the early flowering stage (July 19) (Figure 4.5). The relationship had a decreased R^2 at the peak flowering stage (July 27) (Figure 4.5). Moreover, the results showed that Brassica genotypes with shorter flowering layer depth were overestimated. In 2017, the peak flowering stage had a significant regression with R^2 of 0.36 (Figure 4.6). Similarly, the flowering layer depth was overestimated for the genotypes with short layer depth. For the year of 2018, the FL_{mean} was significantly

correlated with the FL_{actual} at full flowering (R^2 = 0.34 and 0.42) (Figures 4.7B,C). Although the regression at the peak flowering (July 17) was not statistically significant, the *p* value was close to the significance level (0.05). Like 2016 and 2017, there was an overestimation for the FL_{actual} using UAV-based imagery. There was no significant relationship between the FL_{actual} and FL_{mean} at the beginning of flowering stage (Figure 4.7A). Ground reference data were not collected at the end of flowering stage for three site years. Therefore, the relationship at this stage was unknown.

In general, results from 3 site years, indicate that imagery consistently overestimated canola FL_{actual}. The UAV-derived FL_{mean} was calculated by subtracting UAV-derived canopy height at the bolting stage from the UAV-derived canopy height at the flowering stage (Equation 4.4). Crop canopy was not completely closed and level at the bolting stage. The bolting flower buds were taller than the leaves, but the canopy surface was discontinuous as the buds and racemes only filled part of the uppermost canopy level. As a result canopy height extracted from the DSMs were shorter than the actual canopy height measured at the bolting stage. But there might be less impact on the UAV-derived canopy height when crop canopy was close, more even, denser, and taller at the flowering stage, which made UAV-derived canopy height was closer to actual height. Although FL_{actual} was overestimated using UAV-based imagery, the consistence of overestimation may not affect the ability to predict crop growth trend, which indicates that the UAV-derived FL_{mean} has the potential to estimate FL_{actual} under field conditions.

The non-significant relationship between the FL_{actual} and FL_{mean} at the beginning of flowering stage in 2018 might be due to the short layer depth. Although flowers were presumably not occluded by the top branches at this stage, the yellow flowering layer were too short to be accurately extracted from imagery taken at 25-meter flight altitude. Flowers with pale yellow petals or flowers with small size petals might be neglected when extracting data from imagery with relatively low spatial resolution in which the size of the organ was lower than the GSD. Improving the GSD with high resolution RGB imagery may improve the possibility of flower detection at the early flowering stage. Moreover, the short flowering layer depth at the early flowering stage might introduce more manual measurement errors and decrease estimation accuracy compared with the late development stages (Morrison et al., 2021).

A **B**

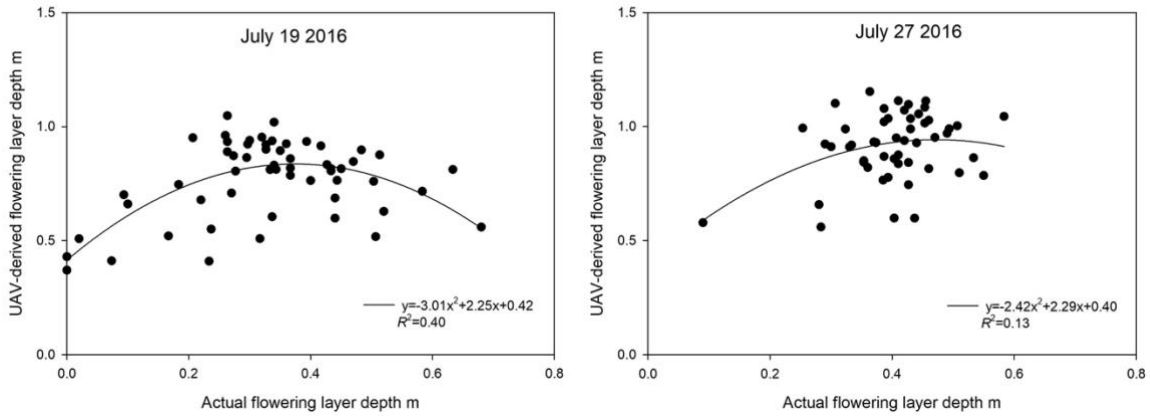


Figure 4.5 Relationships between the UAV-derived flowering layer depth and actual flowering layer depth at Saskatoon, SK, Canada in 2016. Fifty-six genotypes including 54 diverse *B. napus* genotypes, a *B. carinata* genotype, and a *B. juncea* genotype were manually measured for the flowering layer depth. (A) Regression equation for July 19, 2016: $y=-3.01x^2+2.25x+0.42$, $R^2=0.40$. (B) Regression equation for July 27, 2016: $y=-2.42x^2+2.29x+0.40$, $R^2=0.13$.

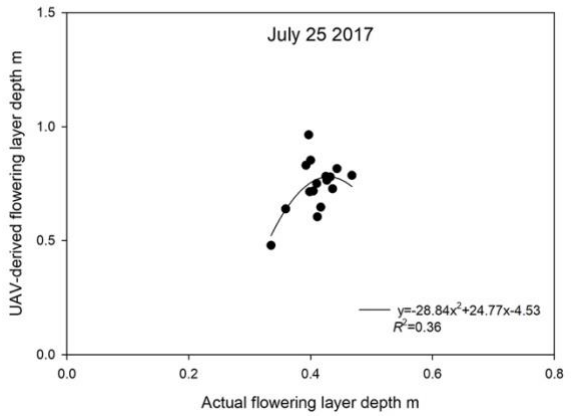


Figure 4.6 Relationships between the UAV-derived flowering layer depth and actual flowering layer depth at Saskatoon, SK, Canada in 2017. Sixteen *B. napus* genotypes were manually measured for the flowering layer depth. Regression equation for July 25, 2017: $y=-28.84x^2+24.77x-4.53$, $R^2=0.36$.

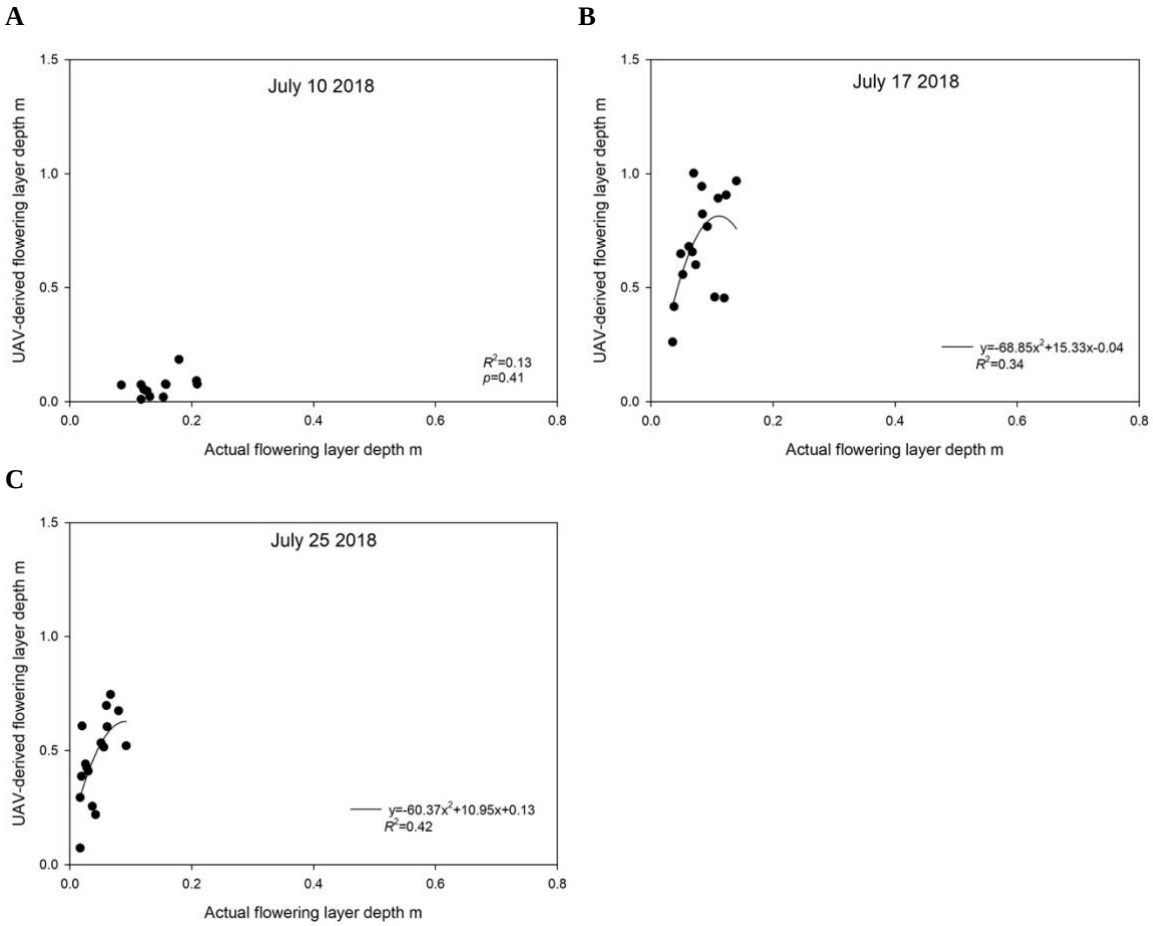


Figure 4.7 Relationships between the UAV-derived flowering layer depth and actual flowering layer depth at Saskatoon, SK, Canada in 2018. Sixteen *B. napus* genotypes were manually measured for the flowering layer depth. (A) There was no significant regression for July 10, 2018: R^2=0.13, p=0.41. (B) Regression equation for July 17, 2018: y=-68.85x^2+15.33x-0.04, R^2=0.34. (C) Regression equation for July 25, 2018: y=-60.37x^2+10.95x+0.13, R^2=0.42.

4.3.3 Seed Yield Estimation Using UAV-derived Canopy Height Extracted from a Single Image Date

Among all site years, there were significant relationships between UAV-derived canopy height and seed yield at the peak flowering stage (R^2 ranging from 0.23 to 0.42, $p < 0.05$) (Figure 4.8). In 2016, the flowering period demonstrated significant relationships between the UAV-derived canopy height and seed yield except at the late flowering stage (August 06). Usually, plant height reached the maximum height at the peak flowering. By peak flowering, the plant stem is the major photosynthetic area providing food source for crop development (Canola Council of Canada, 2021). In addition, taller plants can be easily detected by the sensor with a higher assessment accuracy. The relationships had small R^2 values ranging from 0.10 to 0.12 at the pod stage (Figure 4.8A). These relationships might be due to severe crop lodging at the late growth stage in 2016, which resulted in uneven maturity and smaller plant size with lower accuracy of height estimation. There were no significant relationships between the UAV-derived canopy height and seed yield at the vegetative stage. At the early development stage, canola plants were still short and might require higher spatial resolution for accurate height estimation. Additionally, height information at the vegetative stage might not be a key factor contributing to seed yield. Other factors such as environmental conditions during the remaining growth stage and crop diseases, can affect final seed yield. Therefore, UAV-derived canopy height at the early vegetative stage may contribute little in modeling yield.

In 2017, there were similar patterns in the relationships between the UAV-derived canopy height and seed yield compared with the 2016 assessment. The peak flowering stage (July 26) showed a significant relationship with seed yield with R^2 of 0.28 (Figure 4.8B). Both vegetative (June 28 – July 07) and pod stages (August 16 – August 29) had relatively less R^2 compared with the flowering stage (Figure 4.8B). Not surprisingly, in 2018, the flowering stage showed significant relationships with seed yield. The weakest relationship between the UAV-derived canopy height and seed yield was found at the vegetative stage (R^2 = 0.14 to 0.25) (Figure 4.8C). The pod stage showed significant relationships with seed yield with greater R^2 up to 0.48 compared with the regressions in 2016 and 2017 (Figure 4.8). These relatively stronger relationships at the pod stage might be resulted from less plant lodging and more uniform canopy in 2018.

Generally, consistently significant relationships between the UAV-derived canopy height and seed yield were observed during the flowering stage. The peak flowering time showed the greatest R^2 when canola plants reached the maximum stem length. Canola plant stem played an important role in photosynthesis and became the major sources of food for seed growth by peak flowering when lower pods had started elongating. In addition, a larger plant made it easy for sensor detection during the flowering period. There were no consistently strong relationships with seed yield at the pod stage. Normally, crop lodging can result in non-uniform and shorter canopy height, which can reduce the accuracy of canopy height estimation (Iqbal et al., 2017; Wilke et al., 2019; Morrison et al., 2021). Thus, digital canopy height information at the pod stage may not be a stable indicator for yield estimation when severe lodging occurs. Across 3 site years, the weakest relationships were found at the vegetative stage, indicating that this growth stage might not be a critical growth stage for yield estimation using canopy height information. Canola plants might be too small to be detected by imagery, which could lower the accuracy of height estimation, and therefore yield prediction. The results are in accordance with the study of Geipel et al. (2014). They stated that high resolution imagery was required for height estimation at the early growth stages of corn. In addition, canopy height information at the vegetative stage may not be a key factor contributing to final seed yield.

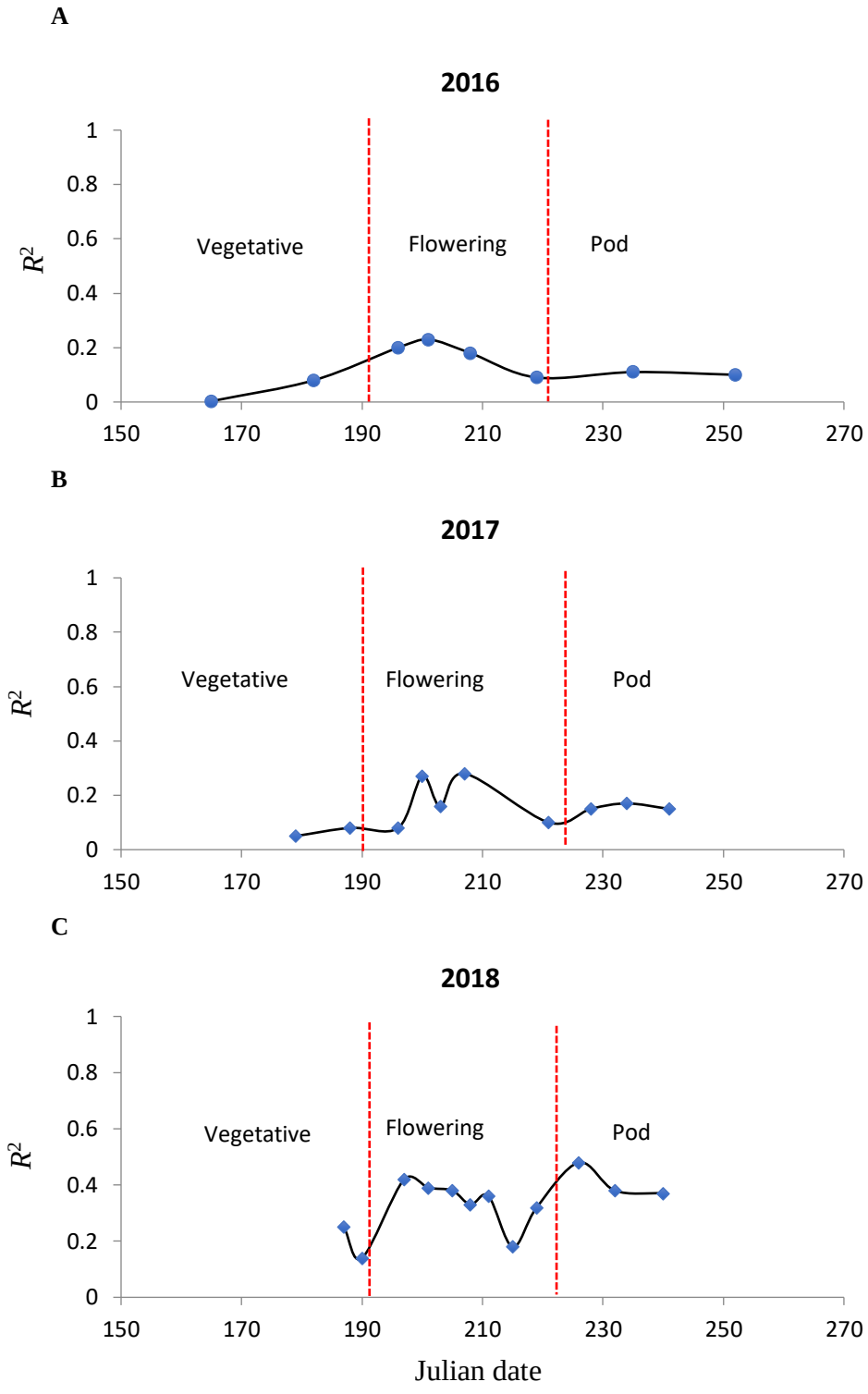


Figure 4.8 Trend lines of coefficient of determination (R^2) between seed yield and the UAV-derived canopy height at Saskatoon, Saskatchewan, Canada in 2016 (A), 2017 (B), and 2018 (C).

4.3.4 Seed Yield Estimation Using The Cumulative Canopy Height Calculated from Multi-Temporal Imagery

Although there were significant relationships between the UAV-derived canopy height and yield during the reproductive stage, lodging at the late development stage such as the pod stage might have an adverse effect on the accuracy of canopy height and yield estimation. Therefore, the dynamics of canopy height during the entire crop season may demonstrate a more realistic plant growth condition such as rapid stem elongation or lodging.

Compared with the canopy height extracted from a single image date, the cumulative UAV-derived canopy height across the flowering stage and entire crop season showed more consistent and significant relationships with seed yield (R^2 ranging from 0.25 to 0.46) (Figures 4.9-4.11). More specifically, in 2016, the UAV-derived canopy height accumulated during the flowering stage and whole crop season showed significant relationships with seed yield with greater R^2 than using a single image (Figures 4.8A, 4.9A,C). A weak relationship between the cumulative UAV-derived canopy height and yield was found at the pod stage (Figure 4.9B). The potential reason might be an adverse impact of severe lodging on canopy height and uniformity, which makes digital height information unreliable for yield estimation.

In 2017, the accumulation of canopy height dynamics throughout the flowering stage and entire crop season had the highest yield estimation accuracy, compared with using the single image date (Figures 4.8B, 4.10A,C). The cumulative UAV-derived canopy height at the pod stage had low yield estimation accuracy ($R^2 = 0.17$) (Figure 4.10B). Similarly, in 2018, the flowering stage and entire crop season showed significant relationships between the cumulative UAV-derived canopy height and yield (Figures 4.11A,C). The accumulation of UAV-derived canopy height at the pod stage had relatively higher yield estimation accuracy compared with the pod stage in 2016 and 2017. It might be due to less lodging in 2018 compared with the first 2 site years. Previous studies have shown the improvement of yield estimation by accumulating multi-temporal vegetation index in rice (Zhou et al., 2017), wheat (Xue et al., 2007; Wang et al., 2014), and canola (Zhang et al., 2021b). They reported that the accumulative imagery could better represent the general crop growth status than using a single image.

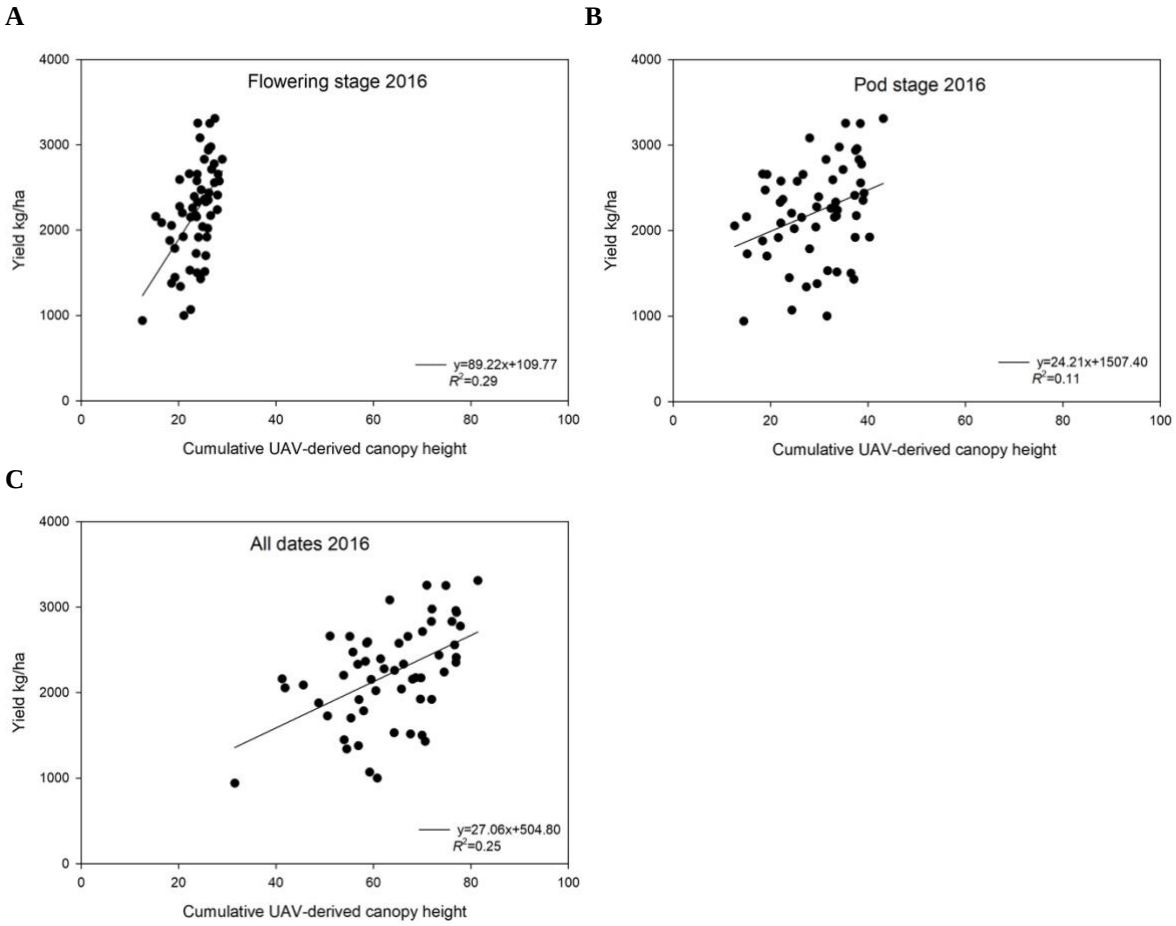


Figure 4.9 Relationships between seed yield and cumulative UAV-derived canopy height at the flowering stage, the pod stage, and throughout the crop season in 2016. (A) Regression equation at the flowering stage: y=89.22x+109.77, R^2=0.29. (B) Regression equation at the pod stage: y=24.21x+1507.40, R^2=0.11. (C) Regression equation throughout the crop season: y=27.06x+504.80, R^2=0.25.

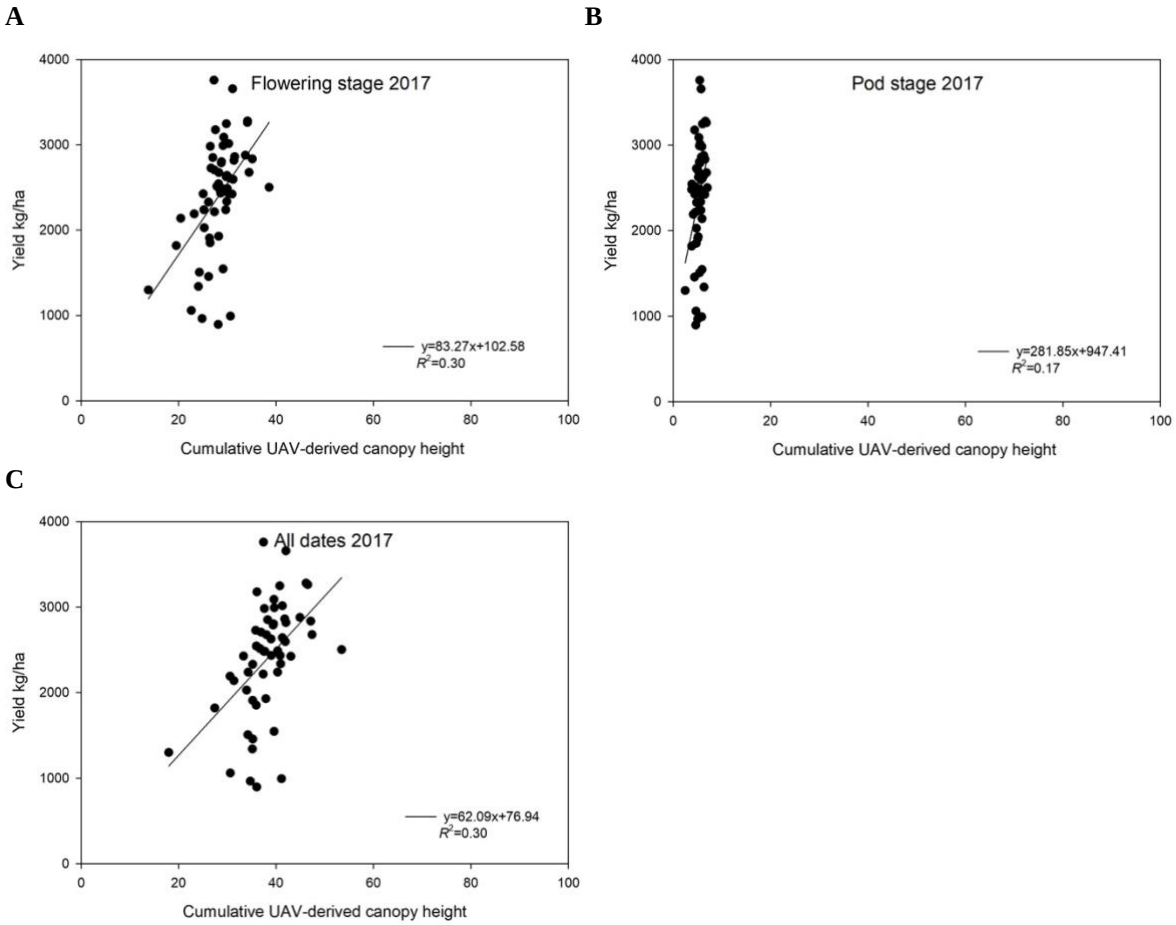


Figure 4.10 Relationships between seed yield and cumulative UAV-derived canopy height at the flowering stage, the pod stage, and throughout the crop season in 2017. (A) Regression equation at the flowering stage: y=83.27x+102.58, R^2=0.30. (B) Regression equation at the pod stage: y=281.85x+947.41, R^2=0.17. (C) Regression equation throughout the crop season: y=62.09x+76.94, R^2=0.30.

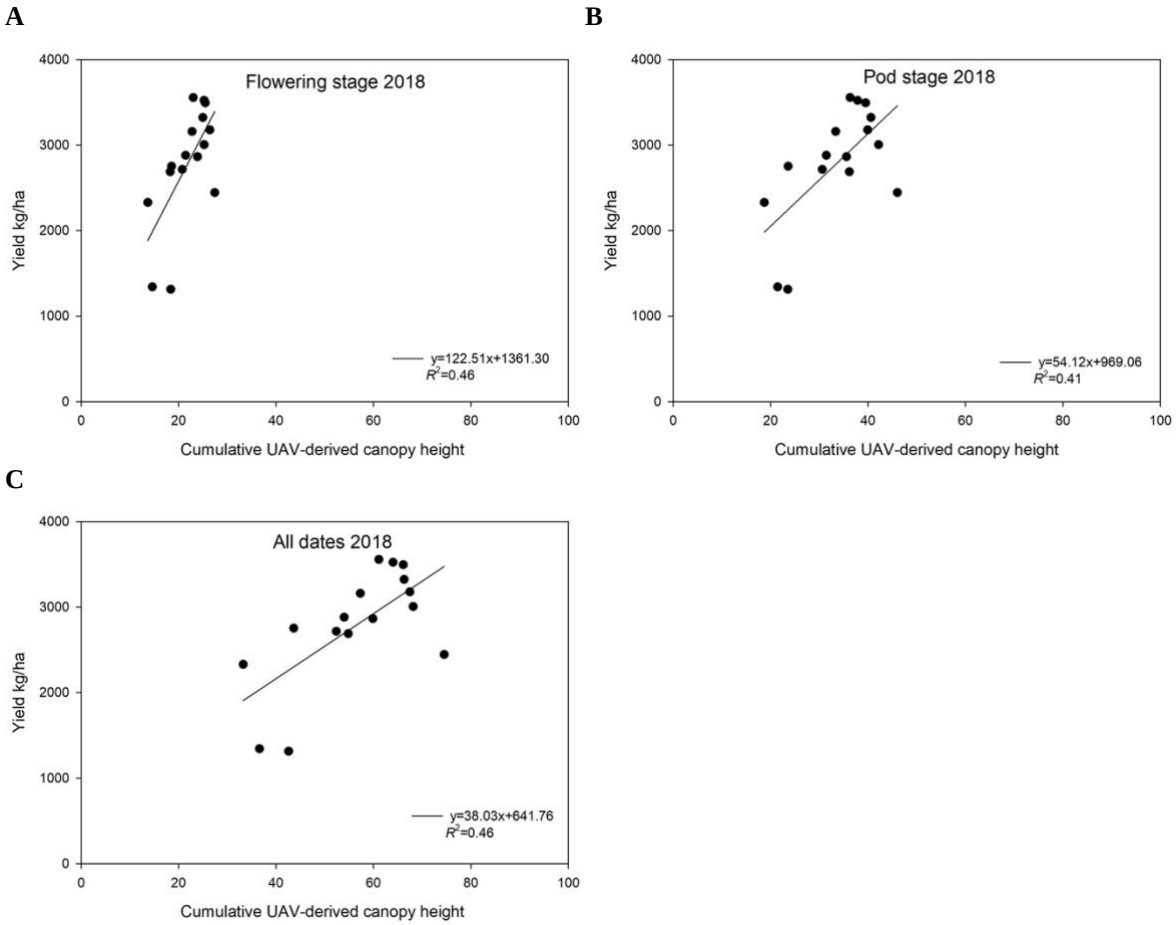


Figure 4.11 Relationships between seed yield and cumulative UAV-derived canopy height at the flowering stage, the pod stage, and throughout the crop season in 2018. (A) Regression equation at the flowering stage: y=122.51x+1361.30, R^2=0.46. (B) Regression equation at the pod stage: y=54.12x+969.06, R^2=0.41. (C) Regression equation throughout the crop season: y=38.03x+641.76, R^2=0.46.

4.3.5 Seed Yield Estimation Using Cumulative UAV-derived Flowering Layer Depth

Flowering timing differed among the diverse 56 Brassica genotypes. To cover the entire flowering layer development progress for each genotype, the FL_{mean} was accumulated over time using the flowering progress curve (AUFPC) (Equation 3.5), as described by Chapter 3. The cumulative FL_{mean} is defined as the overall magnitude and the duration of flowering layer depth change.

Seed yield was significantly and positively correlated with the cumulative FL_{mean} in 2016 (R^2 = 0.20) and in 2017 (R^2 = 0.34) (Figures 4.12A,B). There was no significant relationship between the cumulative FL_{mean} and seed yield in 2018 (Figure 4.12C). The yellow layer of canola flower can reflect or absorb 60% - 65% incoming solar radiation that could have been

used by the photosynthetic plant tissue (Diepenbrock 2000). It seems that shorter yellow FL_{actual} may increase radiation absorption by the leaves and seed development (Diepenbrock 2000). However, less flower number per unit ground can result in fewer pods and seed (Yates and Steven, 1987). In this study, it indicated that the increase in radiation with fewer pods and seed might not be a significant factor influencing seed yield, which is consistent with results found by Yates and Steven (1987). The non-significant regression result in 2018 might be because the relatively shorter FL_{actual} could not be accurately estimated *via* the imagery.

Compared with the cumulative FL_{mean}, there were no significant relationships between the manually measured cumulative FL_{actual} and seed yield over 3 site years (Figure 4.13). The non-significant relationships might be due to the limited sampling size (i.e., 3 plants per plot for each manual measurement) in the field at the flowering stage. Smaller sample size might make the manual measurement not representative for an entire plot (Morrison et al., 2021). In addition, manual measurement was influenced by lacking labor and limited access to field due to wet field conditions. Therefore, the frequency of the manual measurement was lower than UAV imagery collection, resulting in inadequate plant growth information for yield estimation (Table 4.1). Moreover, missing data due to human error could be another factor affecting regression. The results indicate that adequate flowering layer depth progress information might be necessary to improve the accuracy of seed yield estimation.

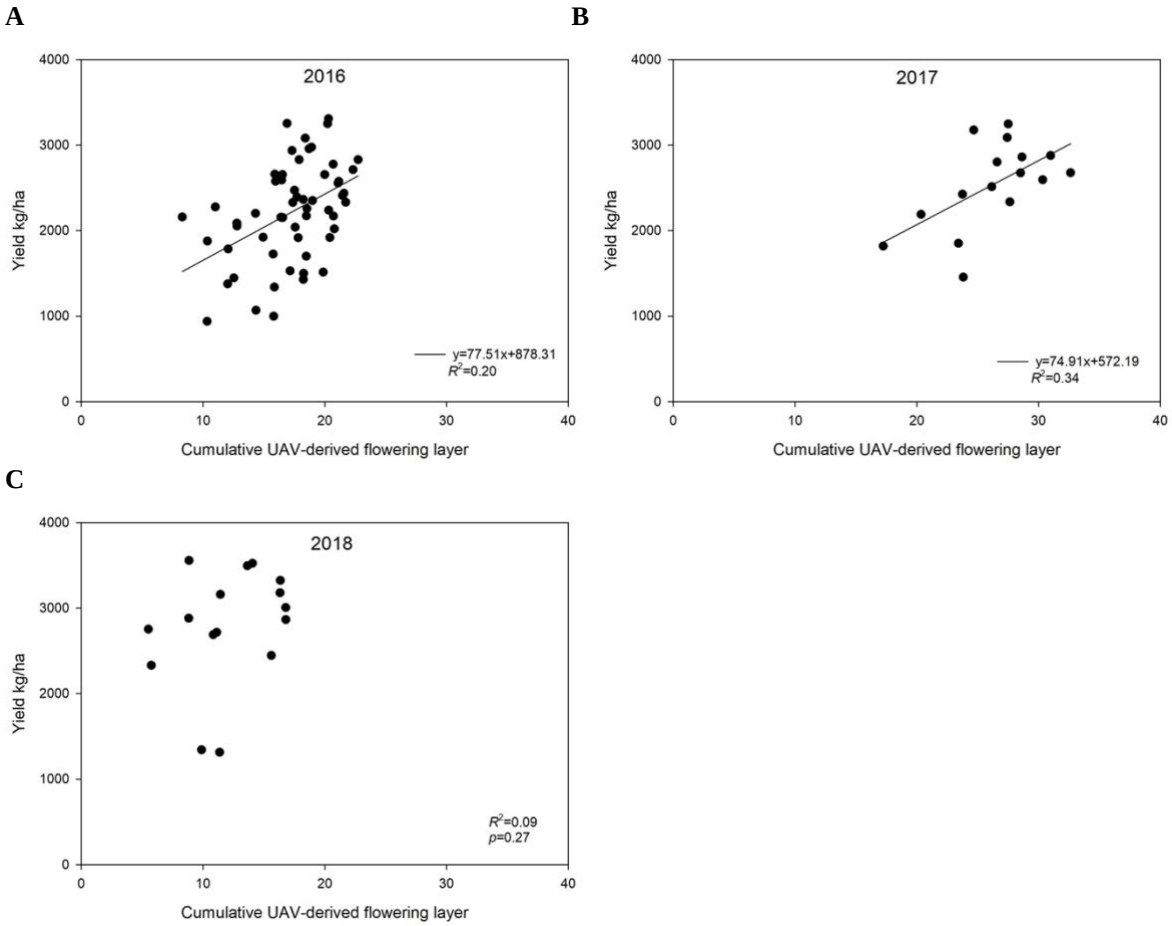


Figure 4.12 Relationships between seed yield and cumulative UAV-derived flowering layer depth from 2016 to 2018 at Saskatoon, SK, Canada. (A) Regression equation in 2016: y=77.51x+878.31, R^2=0.20. (B) Regression equation in 2017: y=74.91x+572.19, R^2=0.34. (C) There was no significant regression in 2018: R^2=0.09, p=0.27.

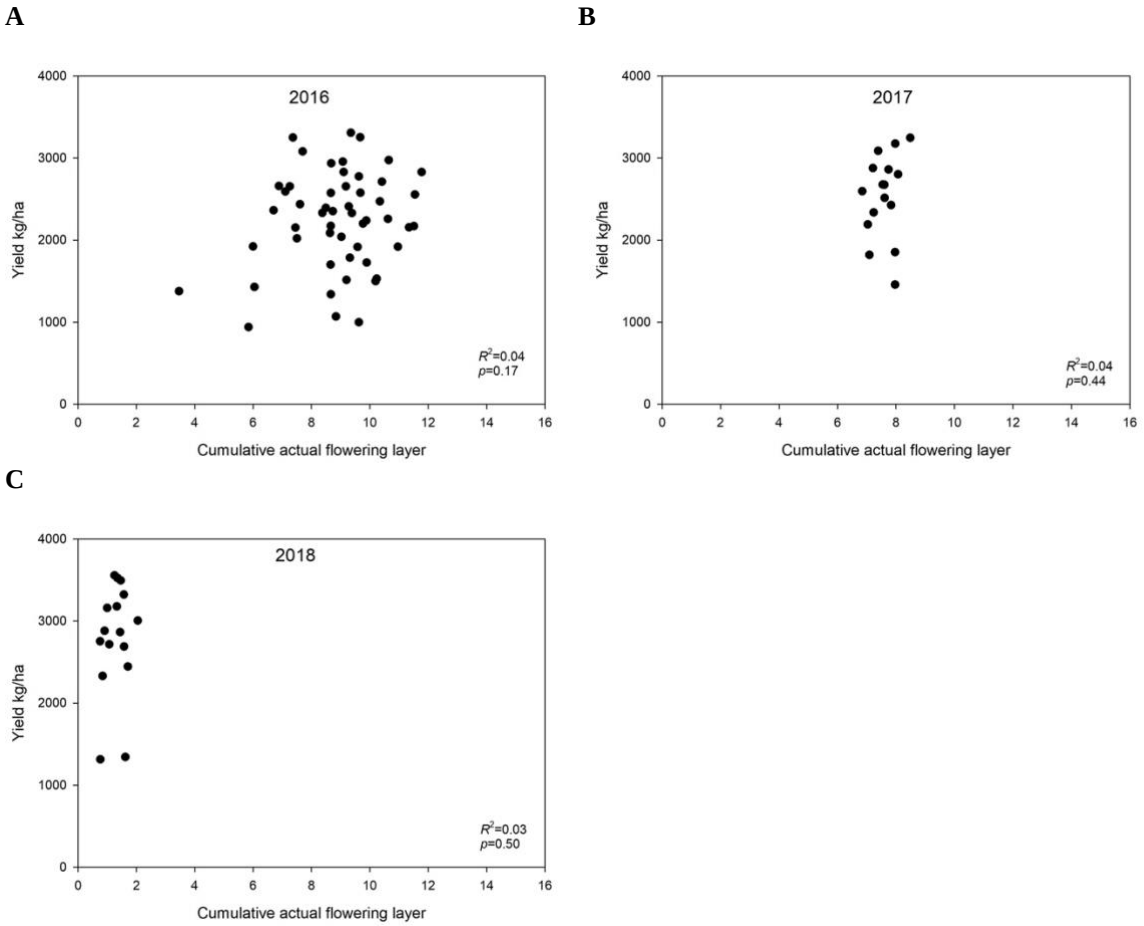


Figure 4.13 Relationships between seed yield and cumulative actual flowering layer depth from 2016 to 2018 at Saskatoon, SK, Canada. There were no significant relationships in 2016 (A), 2017 (B), and 2018 (C).

4.4 Conclusions

In conclusion, canopy height can be estimated using UAV-based imagery during reproductive stage. The peak flowering and pod stages showed the highest accuracy of height estimation using the UAV-based DSMs when plant reached the maximum height. In addition, shorter plants at the early development stage may require sensors with high spatial resolution to increase estimation accuracy. The results also implied the UAV-derived canopy height had the potential to monitor the dynamics of crop canopy height during the growing season, which can be used for crop lodging evaluation. Moreover, this study indicated that canopy height information extracted from a single image date during the flowering and pod stages had the potential for seed yield estimation, but the accuracy of yield estimation using height information at the pod stage was not consistent, and it can be influenced by severe crop lodging. Similar results were reported by Geipel et al. (2014) and Bendig et al. (2015) in barley and

corn, respectively. Their studies indicated that the application of digital height information in yield estimation was limited by certain growth stage. Compared with using the single image date, the accumulation of UAV-derived canopy height dynamics at the flowering period or the entire crop season increased yield estimation accuracy. The cumulative UAV-derived canopy height during the flowering stage may be a more promising phenotype than the entire crop season as the accumulation from the flowering period requires less imagery collection and provides earlier crop evaluation. Moreover, the cumulative FL_{mean} can be a potential predictor variable to explain some variations in seed yield model. This digitalized phenotype can be estimated using UAV-based DSMs.

References

Assefa, Y., Prasad, P. V., Foster, C., Wright, Y., Young, S., Bradley, P., et al. (2018). Major management factors determining spring and winter canola yield in North America. Crop Science, 58, 1-16. doi: 10.2135/cropsci2017.02.0079

Bendig, J., Bolten, A., Bennertz, S., Broscheit, J., Eichfuss, S., and Bareth, G. (2014). Estimating biomass of barley using crop surface models (CSMs) derived from UAV-based RGB imaging. Remote Sensing, 6, 10395-10412. doi: 10.3390/rs61110395

Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., et al. (2015). Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. International Journal of Applied Earth Observation and Geoinformation, 39, 79-87. doi: 10.1016/j.jag.2015.02.012

Canola Council of Canada. (2021). Canola growth stages. Available online at: https://www.canolacouncil.org/canola-encyclopedia/growth-stages/#canola-growth-stage-1-leaf-development (Accessed February 12, 2022).

d'Andrimont, R., Taymans, M., Lemoine, G., Ceglar, A., Yordanov, M., and van der Velde, M. (2020). Detecting flowering phenology in oil seed rape parcels with Sentinel-1 and-2 time series. Remote Sensing of Environment, 239, 1-14. doi: 10.1016/j.rse.2020.111660

Diepenbrock, W. (2000). Yield analysis of winter oilseed rape (*Brassica napus* L.): a review. Field Crops Research, 67, 35-49. doi: 10.1016/S0378-4290(00)00082-4

Duddu, H. S., Johnson, E. N., Willenborg, C. J., and Shirtliffe, S. J. (2019). High-throughput UAV image-based method is more precise than manual rating of herbicide tolerance. Plant Phenomics, 2019, 1-9. doi: 10.34133/2019/6036453

Ebersbach, J., Khan, N. A., McQuillan, I., Higgins, E. E., Horner, K., Bandi, V., et al. (2021). Exploiting high-throughput indoor phenotyping to characterize the founders of a structured *B. napus* breeding population. Frontiers in Plant Science, 12, 780250-780250. doi: 10.3389/fpls.2021.780250

Fang, S., Tang, W., Peng, Y., Gong, Y., Dai, C., Chai, R., et al. (2016). Remote estimation of vegetation fraction and flower fraction in oilseed rape with unmanned aerial vehicle data. Remote Sensing, 8, 416. doi: 10.3390/rs8050416

Faraji, A. (2012). Flower formation and pod/flower ratio in canola (*Brassica napus* L.) affected by assimilates supply around flowering. International Journal of Plant Production, 4, 271-280. doi: 10.22069/ijpp.2012.710

Faraji, A., Latifi, N., Soltani, A., and Rad, A. H. S. (2008). Effect of high temperature stress and supplemental irrigation on flower and pod formation in two canola (*Brassica napus* L.) cultivars at Mediterranean climate. Asian Journal of Plant Sciences, 7, 343-351. doi: 10.3923/ajps.2008.343.351

Geipel, J., Link, J., and Claupein, W. (2014). Combined spectral and spatial modeling of corn yield based on aerial images and crop surface models acquired with an unmanned aircraft system. Remote Sensing, 6, 10335-10355. doi: 10.3390/rs61110335

Gong, Y., Duan, B., Fang, S., Zhu, R., Wu, X., Ma, Y., et al. (2018). Remote estimation of rapeseed yield with unmanned aerial vehicle (UAV) imaging and spectral mixture analysis. Plant Methods, 14, 1-14. doi: 10.1186/s13007-018-0338-z

Hassan, M. A., Yang, M., Rasheed, A., Yang, G., Reynolds, M., Xia, X., et al. (2019). A rapid monitoring of NDVI across the wheat growth cycle for grain yield prediction using a multi-spectral UAV platform. Plant Science, 282, 95-103. doi: 10.1016/j.plantsci.2018.10.022

Iqbal, F., Lucieer, A., Barry, K., and Wells, R. (2017). Poppy crop height and capsule volume estimation from a single UAS flight. Remote Sensing, 9, 647. doi: 10.3390/rs9070647

Ivanovska, S., Stojkovski, C., Dimov, Z., Marjanović-Jeromela, A., Jankulovska, M., and Jankuloski, L. (2007). Interrelationship between yield and yield related traits of spring canola (*Brassica napus* L.) genotypes. Genetika, 39, 325-332. doi: 10.2298/GENSR0703325I

Kirkegaard, J. A., Lilley, J. M., Brill, R. D., Ware, A. H., and Walela, C. K. (2018). The critical period for yield and quality determination in canola (*Brassica napus* L.). Field Crops Research, 222, 180-188. doi: 10.1016/j.fcr.2018.03.018

Ma, B. L., Biswas, D. K., Herath, A. W., Whalen, J. K., Ruan, S. Q., Caldwell, C., et al. (2015). Growth, yield, and yield components of canola as affected by nitrogen, sulfur, and boron application. Journal of Plant Nutrition and Soil Science, 178, 658-670. doi: 10.1002/jpln.201400280

Maturbongs, B., Wing, M. G., Strimbu, B., and Burnett, J. (2019). Forest inventory sensitivity to UAS-based image processing algorithms. Annals of Forest Research, 62, 87-108. doi: 10.15287/afr.2018.1282

Morrison, M. J., Gahagan, A. C., and Lefebvre, M. B. (2021). Measuring canopy height in soybean and wheat using a low-cost depth camera. The Plant Phenome Journal, 4, e20019. doi: 10.1002/ppj2.20019

Panday, U. S., Shrestha, N., Maharjan, S., Pratihast, A. K., Shrestha, K. L., and Aryal, J. (2020). Correlating the plant height of wheat with above-ground biomass and crop yield using drone imagery and crop surface model, a case study from Nepal. Drones, 4, 28. doi: 10.3390/drones4030028

Peng, Y., Li, Y., Dai, C., Fang, S., Gong, Y., Wu, X., et al. (2019). Remote prediction of yield based on LAI estimation in oilseed rape under different planting methods and nitrogen fertilizer applications. Agricultural and Forest Meteorology, 271, 116-125. doi: 10.1016/j.agrformet.2019.02.032

Ramos, A. P. M., Osco, L. P., Furuya, D. E. G., Gonçalves, W. N., Santana, D. C., Teodoro, L. P. R., et al. (2020). A random forest ranking approach to predict yield in maize with uav-based vegetation spectral indices. Computers and Electronics in Agriculture, 178, 105791. doi: 10.1016/j.compag.2020.105791

Sulik, J. J., and Long, D. S. (2015). Spectral indices for yellow canola flowers. International Journal of Remote Sensing, 36, 2751-2765. doi: 10.1080/01431161.2015.1047994

Sulik, J. J., and Long, D. S. (2016). Spectral considerations for modeling yield of canola. Remote Sensing of Environment, 184, 161-174. doi: 10.1016/j.rse.2016.06.016

Tayo, T. O., and Morgan, D. G. (1975). Quantitative analysis of the growth, development and distribution of flowers and pods in oil seed rape (*Brassica napus* L.). The Journal of Agricultural Science, 85, 103-110. doi: 10.1017/S0021859600053466

Tunca, E., Köksal, E. S., Çetin, S., Ekiz, N. M., and Balde, H. (2018). Yield and leaf area index estimations for sunflower plants using unmanned aerial vehicle images. Environmental Monitoring and Assessment, 190, 1-12. doi: 10.1007/s10661-018-7064-x

Vanbrabant, Y., Delalieux, S., Tits, L., Pauly, K., Vandermaesen, J., and Somers, B. (2020). Pear flower cluster quantification using RGB drone imagery. Agronomy, 10, 407. doi: 10.3390/agronomy10030407

Wan, L., Li, Y., Cen, H., Zhu, J., Yin, W., Wu, W., et al. (2018). Combining UAV-based vegetation indices and image classification to estimate flower number in oilseed rape. Remote Sensing, 10, 1484. doi: 10.3390/rs10091484

Wang, L., Tian, Y., Yao, X., Zhu, Y., and Cao, W. (2014). Predicting grain yield and protein content in wheat by fusing multi-sensor and multi-temporal remote-sensing images. Field Crops Research, 164, 178-188. doi: 10.1016/j.fcr.2014.05.001

Wilke, N., Siegmann, B., Klingbeil, L., Burkart, A., Kraska, T., Muller, O., et al. (2019). Quantifying lodging percentage and lodging severity using a UAV-based canopy height model combined with an objective threshold approach. Remote Sensing, 11, 515. doi: 10.3390/rs11050515

Wu, W., and Ma, B. L. (2016). A new method for assessing plant lodging and the impact of management options on lodging in canola crop production. Scientific Reports, 6, 1-17. doi: 10.1038/srep31890

Wu, W., and Ma, B. L. (2018). Assessment of canola crop lodging under elevated temperatures for adaptation to climate change. Agricultural and Forest Meteorology, 248, 329-338. doi: 10.1016/j.agrformet.2017.09.017

Xue, L., Cao, W., and Yang, L. (2007). Predicting grain yield and protein content in winter wheat at different N supply levels using canopy reflectance spectra. Pedosphere, 17, 646–653. doi: 10.1016/S1002-0160(07)60077-0

Yates, D. J., and Steven, M. D. (1987). Reflexion and absorption of solar radiation by flowering canopies of oil-seed rape (*Brassica napus* L.). The Journal of Agricultural Science, 109, 495-502. doi: 10.1017/S0021859600081703

Zhang, C., Craine, W. A., McGee, R. J., Vandemark, G. J., Davis, J. B., Brown, J., et al. (2021a). High-throughput phenotyping of canopy height in cool-season crops using sensing techniques. Agronomy Journal, 113, 3269-3280. doi: 10.1002/agj2.20632

Zhang, H., and Flottmann, S. (2018). Source-sink manipulations indicate seed yield in canola is limited by source availability. European Journal of Agronomy, 96, 70-76. doi: 10.1016/j.eja.2018.03.005

Zhang, T., Vail, S., Duddu, H. S., Parkin, I. A., Guo, X., Johnson, E. N., et al. (2021b). Phenotyping flowering in canola (*Brassica napus* L.) and estimating seed yield using an unmanned aerial vehicle-based imagery. Frontiers in Plant Science, 12, 1178. doi: 10.3389/fpls.2021.686332

Zhou, X., Zheng, H. B., Xu, X. Q., He, J. Y., Ge, X. K., Yao, X., et al. (2017). Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 130, 246-255. doi: 10.1016/j.isprsjprs.2017.05.003

Contributions by others to Chapter 4

Dr. Steve J. Shirtliffe and Dr. Sally Vail designed the field experiments. Dr. Sally Vail and Dr. Isobel A. P. Parkin provided and prepared the plant materials. Dr. Steve J. Shirtliffe contributed to funding acquisition. Ti Zhang and Dr. Hema S. N. Duddu collected the images and performed imagery pre-processing. Ti Zhang conducted imagery processing and ground data collection, performed the statistical analysis, and wrote this chapter under the supervision of Dr. Steve J. Shirtliffe and with contributions from Dr. Sally Vail, Dr. Hema S. N. Duddu, Mr. Eric N. Johnson, Dr. Xulin Guo, and Dr. Rosalind A. Bueckert.

Transition section between Chapter 4 and Chapter 5

Chapter 3 and Chapter 4 showed that flowering number, flowering layer depth, and canopy height could be estimated or quantified using two-dimensional (2D) or three-dimensional (3D) imagery information. Additionally, accumulating these digitalized phenotypes using UAV-based multi-temporal imagery improved the ability to estimate seed yield compared with using a single image. However, applying a single predictor variable in a yield model may not explain adequate variations. Chapter 5 would estimate seed yield using the combination of 2D and 3D imagery with a machine learning method.

CHAPTER 5 SEED YIELD ESTIMATION USING UNOCCUPIED AERIAL VEHICLE (UAV)-BASED MULTISPECTRAL IMAGES AND THE RANDOM FOREST MODEL IN DIVERSE BRASSICA GENOTYPES

The content of this chapter will be submitted to Remote Sensing.

Abstract

Early yield estimation in canola (*Brassica napus* L.) plays an important role in selecting breeding genotypes and developing cultivars through crop breeding that optimize field productivity. Traditional methods for yield estimation based on visual rating or manual sampling before harvest stage are subjective, destructive, labor-intensive, and time-consuming for large-scale fields. In the most recent decade, advances in unoccupied aerial vehicle (UAV) mounted with imaging sensors have demonstrated an ability to acquire high temporal and spatial resolution images for yield estimation without damaging the crop canopy. The objective of this study was to estimate seed yield in canola using multi-temporal UAV-based imagery collected during the entire crop season. Fifty-six Brassica breeding genotypes were investigated in canola field experiments conducted near Saskatoon, SK, Canada from 2016 to 2018. Multispectral imagery was collected at 3-day or 7-day intervals during the crop season. To estimate seed yield, instead of using vegetation indices (VIs) extracted from a single image date, VIs were accumulated at the vegetative (i.e., seedling and bolting stages), the flowering, and the pod stages. These cumulative VIs and two cumulative canopy structural phenotypes (i.e., crop canopy height and flowering layer depth) were evaluated as potential predictor variables using simple linear regression and random forest (RF) regression models. The linear regression results demonstrated that the cumulative UAV-derived canopy height at the flowering stage was an acceptable indicator for yield estimation with coefficient of determination (R^2) value of 0.18. The RF regression model combining 9 cumulative VIs and a cumulative canopy structural phenotype improved the accuracy of yield estimation (R^2 = 0.32; root mean square error = 490.54 kg/ha for the validation dataset) compared with the linear regression model. According to the ranking of variable importance by the RF regression model, the cumulative blue normalized difference vegetation index at the flowering stage, the cumulative UAV-derived canopy height at the flowering stage, and the cumulative normalized

difference vegetation index at the vegetative stage were the most important traits for yield estimation. Moreover, the vegetative (i.e., from seedling to bolting stage) and flowering stages were the optimal early crop growth stage for yield estimation in canola. The results indicate that combining the cumulative VIs and cumulative canopy structural phenotype derived from UAV-based imagery has the potential to estimate seed yield in canola which can assist breeding programs and improve farm productivity.

5.1 Introduction

Canola (*Brassica napus* L.), an important oilseed crop, is cultivated worldwide. Canada is among the top five exporting countries, with 20.8 million seed acres and a production of 18.7 million tonnes in 2020 (Statistics Canada, 2020). Timely and accurate crop yield estimation before harvest is important in high-yielding genotype selection and cultivar development in breeding programs (Araus and Cairns, 2014; Sankaran et al., 2015). Yield estimation plays a vital role in optimizing field management for farmers (Peng et al., 2019) and decision making for policy makers and marketing agencies (Johnson et al., 2016). However, the conventional methods to estimate crop yield before maturity stage using visual field evaluation or straight harvest require extensive field experience and are labor-intensive and destructive. In addition, these methods are subjective due to limited sampling size which may not be representative for the entire plot-level or farm-scale field. Furthermore, traditional crop yield models usually require quantitative information such as crop type, growth stages, environmental conditions, and management operations, which is sometimes difficult to obtain and quantify (Johnson et al., 2016; Peng et al., 2019). To overcome the challenges of the conventional methods to estimate crop yield *in situ* and the complexity of traditional crop growth models, it is necessary to establish an efficient and accurate descriptive yield model for plant breeders, farmers, and policy makers.

In the recent decade, advances in remote sensing techniques have shown potential as an efficient and non-destructive approach for crop yield prediction. Among the remote phenotyping platforms, unoccupied aerial vehicles (UAVs) equipped with multi-spectral sensors have become a relatively low-cost high-throughput phenotyping tool that can provide high spatial and temporal resolution images for large-scale areas. Application of vegetation

indices (VIs) extracted from the multispectral imagery have been widely studied in many crops for yield or biomass estimation such as canola (Sulik and Long, 2016; Gong et al., 2018; Peng et al., 2019; Zhang et al., 2021b), wheat (Han et al., 2021b), barley (Bendig et al., 2015; Johnson et al., 2016), rice (Noureldin et al., 2013; Zhou et al., 2017; Wan et al., 2020), corn (Ramos et al., 2020), and cotton (Feng et al., 2020). In canola, Sulik and Long (2016) determined that normalized difference yellow index (NDYI) at the flowering stage was significantly correlated with county-level seed yield with R^2 up to 0.76. Gong et al. (2018) showed that a set of greenness related VIs such as the normalized difference vegetation index (NDVI) multiplied by leaf-related abundance at the early flowering stage were able to estimate yield under different nitrogen treatments with low estimation error (< 13%). Peng et al. (2019) found that red edge vegetation index ($CI_{red\ edge}$) during the vegetative stage was able to predict seed yield with low estimation error (< 15%). Han et al. (2021b) applied NDVI and soil adjusted vegetation index (SAVI) to estimate seed yield in wheat with R^2 of 0.673 and 0.616, respectively.

Apart from spectral information extracted from 2-dimensional (2D) orthomosaic, canopy structural phenotype (e.g., crop canopy height) can be a potential indicator for monitoring crop health status and yield estimation (Zhang et al., 2021a). The canopy structural phenotypes are defined as the characteristics of crop canopy which can be quantified using 3-dimensional (3D) imagery such as canopy height and flowering layer depth (Chapter 4). Previous research has proven that digital crop height extracted from 3D models such as crop surface models (CSMs) or digital elevation models (DEMs) was strongly correlated with growth parameters and seed yield or biomass in various crops such as barley (Bendig et al., 2015), cotton (Feng et al., 2019), and cool season crops including chickpea, dry pea, camelina, and canola (Zhang et al., 2021a). For example, Bendig et al. (2015) observed that plant height derived from CSMs was the most robust digital phenotype for biomass estimation in barley (R^2 = 0.80 - 0.82) compared with a series of VIs. A linear regression using plant height extracted from DEMs was able to predict yield in cotton with *r* value ranging from 0.66 to 0.96 (Feng et al., 2019). Zhang et al. (2021a) demonstrated that plant height derived from UAV-based RGB images were significantly correlated with actual plant height in four cool-season crops (i.e., chickpea, dry pea, camelina, and canola).

In addition, compared with using UAV-based imagery obtained from a single image date, accumulative or multi-temporal imagery may improve the accuracy of yield estimation (Xue et al., 2007; Wang et al., 2014; Zhou et al., 2017; Wan et al., 2020; Zhang et al., 2021b). Xue et al. (2007) and Wang et al. (2014) found that the cumulative NDVI and ratio vegetation index (RVI) obtained from the jointing to grain filling stage improved yield estimation accuracy in wheat. Zhou et al. (2017) showed that the multi-temporal NDVI derived from multispectral images had a significant relationship with seed yield in rice using a multiple linear regression model (r = 0.73). Zhang et al. (2021b) reported that accumulative NDYI-based pixel number during the flowering stage improved the yield estimation accuracy among 56 Brassica genotypes across 5 site years, compared with using imagery from a single image date.

During the entire crop season, canola has three distinct canopy morphologies with different spectral reflectance including the vegetative stage dominated by leafy vegetation, the flowering stage displaying a yellowish canopy, and the mature stage with green to dark green or brown pods and branches (Domínguez et al., 2015; Pan et al., 2013; Sulik and Long, 2016; Han et al., 2021a; Singh et al., 2021). In addition, canola yield is related to many direct and indirect factors during the entire crop season. Simply inputting a single predictor variable in a descriptive yield model may result in unsatisfactory and inaccurate results because of the potential saturation of VIs and neglecting other important factors which may contribute to final seed yield (Geipel et al., 2014; Liu et al., 2019; Fu et al., 2020). Geipel et al. (2014) set up a multiple linear regression model including crop coverage related VIs and digital crop heights computed from CSMs. They reported that the multiple linear regression model improved the accuracy of yield estimation compared with the simple linear regression model using a single predictor. Liu et al. (2019) also reported that combining VIs and texture metrics computed by a gray level co-occurrence matrix texture algorithm improved above ground biomass estimation in winter canola.

In the contest to build a descriptive model, numerous types of regression models can be used to explore the relationships between remote sensed imagery and crop yield. In this study a simple linear regression model and a random forest (RF) model were applied. These models were used to determine the contribution of each digital phenotype to seed yield and to develop a descriptive yield prediction model. The linear regression model has commonly been used to determine simple relationships between digital phenotypes and yield (Zhou et al., 2017; Fu et

al., 2020). The RF is a non-linear regression model which has low sensitivity to outliers and minimized overfitting risk by the bagging algorithm (Liu et al., 2019; Fu et al., 2020; Ramos et al., 2020). To fit an RF, two model hyperparameters (i.e., *ntree* and *mtry*) need to be set. The hyperparameter *ntree* determines the number of decision trees grown in the RF model, while *mtry* refers to the number of variables for each split (Liu et al., 2019; Fu et al., 2020).

Many researchers have shown that seed yield prediction model using UAV-based digital phenotypes could be a selection tool in breeding programs and could assist in field management. However, most research on canola yield estimation models only applied limited imagery from a single image date or a single growth stage, which neglected the progression of crop growth and might reduce yield prediction accuracy. Combining comprehensive growth information over the entire crop season may improve the accuracy of yield estimation. In this context, time-series imagery during the entire crop season should be considered and evaluated to estimate yield. In addition, diverse genotypes, uncertain environmental conditions, and the interaction between genotypes and environments may make it more difficult to develop a robust yield model. In this study, the main objective was to investigate the potential of cumulative phenotypes calculated from multi-temporal images for yield estimation using 56 highly diverse Brassica genotypes across three site years.

5.2 Materials and Methods

5.2.1 Experimental Field

A three-year field study was conducted at Agriculture and Agri-Food Canada Research Farm near Saskatoon (latitude: 52.181577, longitude: -106.499884) in Saskatchewan from 2016 to 2018. Fifty-six diverse Brassica breeding genotypes (Ebersbach et al., 2022; Table A.1, Appendix A) were grown in a rectangular lattice design with three replicates, resulting in 168 plots in total with plot sizes of 6.0 m long x 1.2 m wide in 2016 and 2018, and 6.0 m long x 1.5 m wide in 2017. Canola was seeded at a recommended rate of 108 seeds m^{-2}. Pesticides were applied based on field conditions. More details of study area and field management operations were descripted by Chapter 3. Seed yield of each experimental plot was determined at maturity

stage by harvesting the four central rows to reduce border effect using a small plot combine harvester.

5.2.2 Multispectral Imagery Collection

Across three site years, UAV platforms (i.e., Draganflyer X4-P model and Draganflyer Commander model) (DraganFly Inc., Saskatoon, SK, Canada) equipped with a multispectral sensor (RedEdge, MicaSense Inc., Seattle, WA, United States) were used to take imagery over the study area. The sensor has five channels consisting of blue (R_{blue}, 475 ± 10 nm), green (R_{green}, 560 ± 10 nm), red (R_{red}, 668 ± 5 nm), red edge ($R_{red\ edge}$, 717 ± 5 nm), and near-infrared (R_{nir}, 840 ± 20 nm) with an image resolution of 1.2 megapixels and radiometric resolution of 12-bit. Flight altitudes were 20 m and 25 m above ground level, leading to ground sampling distance of 1.36 cm per pixel and 1.70 cm per pixel, respectively.

Images of a calibration panel were taken before and after each flight for radiometric calibration; and geo-locations of ground control points were recorded by GeoExplorer 2008 GPS (Trimble Inc., Westminster, CO, United States) for geometric calibration. To cover the entire crop growth progress, imagery collection was arranged weekly in 2016, semi-weekly during the flowering period in 2017 and 2018, and weekly during the vegetative and pod stages in 2017 and 2018 (Table 5.1).

Table 5.1 Details of canola trials and imagery acquisition at Saskatoon, SK, Canada from 2016 to 2018.

Year	Seeding date	Number of genotypes	Flight altitude (m)	Image acquisition
2016	May 27	56	20	June 13; 24; 30 July 14; 19; 26 August 06; 22 September 08
2017	May 28	56	20	June 28 July 07; 11; 15; 19; 22; 26 August 01; 09; 16; 22; 29
2018	May 21	56	25	June 07; 20; 28 July 06; 09; 16; 20; 24; 27; 30 August 03; 07; 10; 14; 17; 20; 23; 28; 31

5.2.3 Imagery Processing and Data Extraction

5.2.3.1 Pre-processing

Multispectral imagery was initially processed in a commercial software; Pix4Dmapper Pro (Pix4D Inc., San Francisco, CA, United States) for mosaicking, radiometric and geometric calibration. Five calibrated reflectance maps were obtained and exported for further vegetation indices (VIs) calculation. Meanwhile, digital surface models (DSM) were generated using the Structure from Motion algorithm for height-related traits extraction. More details were descripted in Chapter 4.

5.2.3.2 Vegetation indices, canopy height, flowering layer depth, and cumulative image-based features calculation during the crop season

The generated orthomosaic images and DSMs were imported into ArcGIS 10.4.1 (ESRI Canada, Toronto, ON, Canada) for further data extraction. The "Raster Calculator" tool computed VIs using different combinations of the five calibrated reflectance maps. According to previous research, several commonly used VIs were tested in this study including normalized difference vegetation index (NDVI), red edge ratio vegetation index (RVI), chlorophyll index ($CI_{red\ edge}$), blue normalized difference vegetation index (BNDVI), green normalized difference vegetation index (GNDVI), and visible atmospherically resistant index (VARI) (Rouse et al.,

1974; Gitelson et al., 1996 and 2002; Xue et al., 2004; Vina et al., 2004; Sulik and Long, 2016; Hussain et al., 2020). These VIs have shown the possibility to predict crop biomass or yield. Additionally, high resolution flowering index (HFI), red-blue normalizing flowering index (RBFI) (Fernando and Ha, 2021; personal communication), and modified yellowness index (MYI) (Ha, 2021; personal communication) have been proposed and recommended for yellowish flower detection during the flowering stage. Details of these VIs are shown in Table 5.2.

Table 5.2 Details of vegetation indices applied in this study.

Indices	Name	Formula	Reference
NDVI	Normalized difference vegetation index	$(R_{nir}-R_{red})/(R_{nir}+R_{red})$	Rouse et al. (1974)
RVI	Ratio vegetation index	R_{nir}/R_{red}	Xue et al. (2004)
$CI_{red\ edge}$	Red edge chlorophyll index	$(R_{nir}/R_{red\ edge})-1$	Gitelson et al. (2002)
BNDVI	Blue normalized difference vegetation index	$(R_{nir}-R_{blue})/(R_{nir}+R_{blue})$	Sulik and Long (2016)
GNDVI	Green normalized difference vegetation index	$(R_{nir}-R_{green})/(R_{nir}+R_{green})$	Gitelson et al. (1996)
VARI	Visible atmospherically resistant index	$(R_{green}-R_{red})/(R_{green}+R_{red}+R_{blue})$	Vina et al. (2004)
HFI	High resolution flowering index	$(R_{red}-R_{blue})\ x(R_{green}-R_{blue})$	Fernando and Ha (2021)
RBNI	Red-blue normalizing index	$(R_{red}-R_{blue})/(R_{red}+R_{blue})$	Fernando and Ha (2021)
MYI	Modified yellowness index	$(R_{red}\ x\ R_{green})/R_{blue}$	Ha (2021)

In this experiment, to cover the dynamics of crop growth over time, the selected VIs were accumulated at the vegetative, flowering, and pod stages, resulting in 27 cumulative VIs. The cumulative VIs were calculated using the area under curve (AUC) formula:

$$AUC = \left(\frac{X_1+X_2}{2}\right)(t_2 - t_1) + \left(\frac{X_2+X_3}{2}\right)(t_3 - t_2) + \cdots + \left(\frac{X_{n-1}+X_n}{2}\right)(t_n - t_{n-1}) \quad [5.1]$$

where X_1, X_2, X_3, X_{n-1}, and X_n represent the VIs values at each image date, and t_1, t_2, t_3, t_{n-1}, and t_n represent Julian date at each image timing.

Chapter 3 showed that a cumulative VI extracted from multi-temporal imagery (i.e., cumulative NDYI-based flowering pixel) had the potential as a promising indicator of crop growth status and yield production; thus, the cumulative NDYI-based pixel number is included for further regression analysis. Details about the calculation of the cumulative NDYI-based flowering pixel number were described by Chapter 3.

In addition to the 2D image-based features, two cumulative canopy structural phenotype extracted from 3D DSMs (i.e., the cumulative UAV-derived canopy height and flowering layer depth at the flowering stage) has been proved to be potential indicators to estimate seed yield (Chapter 4). The two digital cumulative phenotypes were also tested in Chapter 5. The details of extracting and calculating the cumulative UAV-derived canopy structural phenotypes were described in Chapter 4.

5.2.4 Data Analysis and Model Development

For each cumulative image-based feature, mean values per Brassica genotype were calculated across three replications within each year. In addition, mean yield per genotype was averaged across three replicates within each year, resulting in 168 data points (i.e., 168 = 56 genotype x 3 site years). These mean values (i.e., cumulative image-based features and seed yield) per genotype were used in further regression analysis.

The simple linear regression model was calculated using PROC REG in SAS 9.4 (SAS Institute, Cary, NC, United States) to estimate the linear relationships between seed yield and single cumulative image-based features (i.e., VIs and canopy structural phenotypes).

Prior to applying random forest (RF) regression model, all cumulative VIs and canopy structural phenotypes were pre-screened to avoid covariance. To select suitable independent variables within these cumulative VIs and canopy structural phenotypes, Pearson correlation was applied. The independent variables with Pearson correlation coefficient (r) larger than 0.9 were considered as highly correlated input variables and removed before further RF regression analysis. Pearson correlation analysis was done by the package 'corrplot' in R (Wei and Simko, 2021). According to the initial correlation results, twelve highly correlated variables were eliminated (Figure 5.1). The remaining variables were input as predictors in a random forest regression model (Figure 5.2).

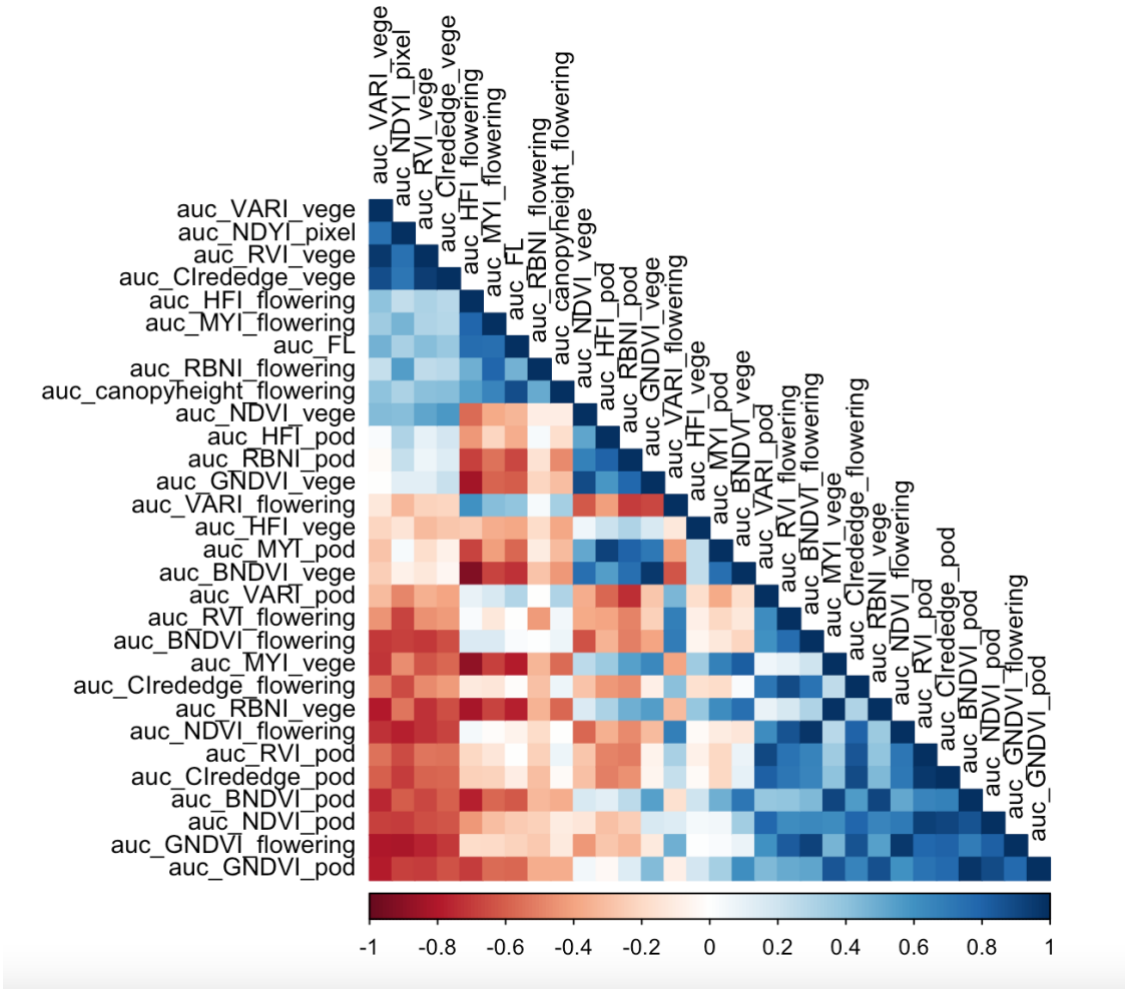


Figure 5.1 Heatmap of correlation between each cumulative vegetation indices and canopy structural phenotypes. auc_VARI represents cumulative VARI; auc_NDYI_pixel represents cumulative NDYI-based pixel number; auc_RVI represents cumulative RVI; auc_CIrededge represents cumulative $CI_{red\ edge}$; auc_HFI represents cumulative HFI; auc_MYI represents cumulative MYI; auc_FL represents cumulative flowering layer depth; auc_RBNI represents cumulative RBNI; auc_canopyheight represents cumulative UAV-derived canopy height; auc_NDVI represents cumulative NDVI; auc_GNDVI represents cumulative GNDVI; auc_BNDVI represents cumulative BNDVI; "_Vege" represents at the vegetative stage; "_flowering" represents at the flowering stage; "_pod" represents at the pod stage.

The RF regression model was constructed with the package 'randomForest' in R (Liaw and Wiener, 2002). The RF regression model is a non-linear ensemble approach for regression determination which is widely used for crop yield or biomass prediction because of its non-sensitivity to data skewness and less overfitting problems (Liu et al., 2019; Wan et al., 2020). Two basic model hyperparameters need to be determined to fit RF model including the number of variables for each split (*mtry*) and the number of trees (*ntree*) to grow. To determine the suitable hyperparameters, the hyperparameter *ntree* was evaluated from 100 to 1000 at a 100-interval. The hyperparameter mtry was determined from 3 to 18 at a 1-interval. Finally, in this study, the default hyperparameters (i.e., *mtry* = 6; *ntree* = 500) were applied in the RF

regression model because there was no obvious improvement after model tuning. After the initial RF modeling, the least important variables were gradually removed based on the variable importance in the projection (VIP). The lower VIP value (i.e., mean decrease accuracy %) represents less contribution to the seed yield estimation. Seventy percent data from 2016 to 2018 were randomly selected as the training dataset for the regression model (Wan et al., 2020). Then, the remaining 30% of the dataset obtained over the three site years were used to validate the generated regression model (Wan et al., 2020). Yield estimation model was evaluated based on two statistical indexes including the coefficient of determination (R^2) and root mean square error (RMSE).

5.3 Results and Discussion

5.3.1 Seed Yield Estimation Using Single Cumulative Vegetation Indices and Canopy Structural Phenotypes

To combine and standardize datasets across three site years, the selected vegetation indices (VIs) were accumulated over three growth stages (i.e., vegetative stage, flowering stage, and pod stage) into one value for further regression analysis using the area under curve function (Equation 5.1). In addition, the cumulative VIs calculated using multi-temporal images usually contain more details of crop growth progress compared with using a single image date (Zhou et al., 2017; Liu et al., 2019; Zhang et al., 2021b). Therefore, in the linear regression analysis, the selected VIs accumulated at the vegetative, flowering, and pod stages, the cumulative NDYI-based flowering pixel number, the cumulative flowering layer depth, and the cumulative UAV-derived canopy height at the flowering stage were applied as inputs to estimate seed yield (Table 5.3). Unfortunately, there were no significant relationships between the single cumulative VIs and seed yield except the cumulative UAV-derived canopy height at the flowering stage (R^2 = 0.18) (Table 5.3). Although the R^2 value for the relationship between the cumulative UAV-derived canopy height at the flowering stage and seed yield was not as high as previous studies (Sulik and Long, 2016; Gong et al., 2018; Liu et al., 2019), the significant regression result is still acceptable because the datasets were based on the responses of 56 diverse Brassica genotypes over three site years (Table B.1, Appendix B). Unlike this study, most previous research collected datasets from only one or few varieties and often only one or

two environments (Sulik and Long, 2016; Gong et al., 2018; Liu et al., 2019). The crop growth patterns with less diverse varieties and environments may be easier to be explained by a regression model (Zhang et al., 2021b). The results indicated that the cumulative UAV-derived canopy height at the flowering stage could be a promising yield indicator in canola when using a single predictor for yield estimation.

Table 5.3 Coefficient of determination (R^2) and root mean square error (RMSE) for the linear regression model between seed yield and the cumulative vegetation indices and canopy structural phenotypes at Saskatoon, SK, Canada.

	Growth stages					
	***R*²**			**RMSE (kg/ha)**		
Cumulative Indices	**Vegetative**	**Flowering**	**Pod**	**Vegetative**	**Flowering**	**Pod**
auc_NDVI	0.02	0.07	0.06	698.98	682.82	686.80
auc_RVI	< 0.01	0.07	0.06	706.80	681.40	685.50
auc_$CI_{red\ edge}$	0.02	0.07	0.02	700.40	681.62	701.64
auc_BNDVI	0.02	0.09	0.04	704.22	676.47	692.12
auc_GNDVI	0.03	0.06	0.02	696.80	1180.51	699.84
auc_VARI	< 0.01	0.08	0.09	706.67	726.49	676.10
auc_HFI	0.04	< 0.01	0.03	692.64	707.22	695.07
auc_RBNI	< 0.01	0.07	< 0.01	706.98	680.54	881.50
auc_MYI	< 0.01	0.07	0.06	706.49	720.77	685.97
auc_ndyi_ pixel	NA	< 0.01	NA	NA	706.04	NA
auc_FL	NA	0.03	NA	NA	695.01	NA
auc_canopyheight	NA	0.18**	NA	NA	642.04	NA

* Significant at the 0.05 probability level.
** Significant at the 0.01 probability level.
***Significant at the 0.001 probability level.
NA represents no data available.
auc_NDVI represents cumulative NDVI.
auc_RVI represents cumulative RVI.
auc_$CI_{red\ edge}$ represents cumulative $CI_{red\ edge.}$
auc_BNDVI represents cumulative BNDVI.
auc_GNDVI represents cumulative GNDVI.
auc_VARI represents cumulative VARI.
auc_HFI represents cumulative HFI.
auc_RBNI represents cumulative RBNI.
auc_MYI represents cumulative MYI.
auc_NDYI_pixel represents cumulative NDYI-based pixel number.
auc_FL represents cumulative flowering layer depth.
auc_canopyheight represents cumulative UAV-derived canopy height.

5.3.2 Seed Yield Estimation Using the Random Forest Model

After the initial RF modeling, the least important variables were removed gradually based on the ranking of variable importance (Figure 5.2). Then, ten cumulative variables including the cumulative BNDVI, the cumulative UAV-derived canopy height, the cumulative VARI, the cumulative RVI, the cumulative NDYI-based pixel at the flowering stage, the cumulative NDVI, the cumulative GNDVI, the cumulative $CI_{red\ edge}$, the cumulative HFI at the vegetative stage and the cumulative MYI at the pod stage were selected and used as inputs in the RF regression model, which provided more accurate seed yield estimation (R^2 = 0.92 ; RMSE = 237.15 kg/ha for the training data; R^2 = 0.32 ; RMSE = 490.54 kg/ha for the validation data) compared with the linear regression model using a single predictor variable (i.e., cumulative UAV-derived canopy height at the flowering stage) (R^2 = 0.18; RMSE = 642.04 kg/ha). The R^2 value for the validation dataset was increased by 78% and RMSE value reduced 24%. Similar results were found by Liu et al. (2019) who reported that the $CI_{red\ edge}$ and RVI were the most important indicators based on the ranking of variable importance in the RF regression model for biomass estimation in winter canola. Wan et al. (2020) found that combining multi-temporal NDVI, NDYI, canopy height, and canopy coverage resulted in robust yield prediction in rice (r = 0.83 - 0.85) with a RF model. Ramos et al. (2020) stated that NDVI and GNDVI were the top variables in the RF for yield prediction in corn.

In this yield estimation study, the results showed that the flowering and vegetative stages were more important than the pod stage for yield estimation in canola when using UAV-based imagery. Although the cumulative MYI at the pod stage provided some contributions to the yield estimation (Figure 5.2), the other 9 cumulative variables at the early growth stages (i.e., vegetative or flowering stages) could provide satisfying yield estimation for on-time field management strategy to maintain final yield.

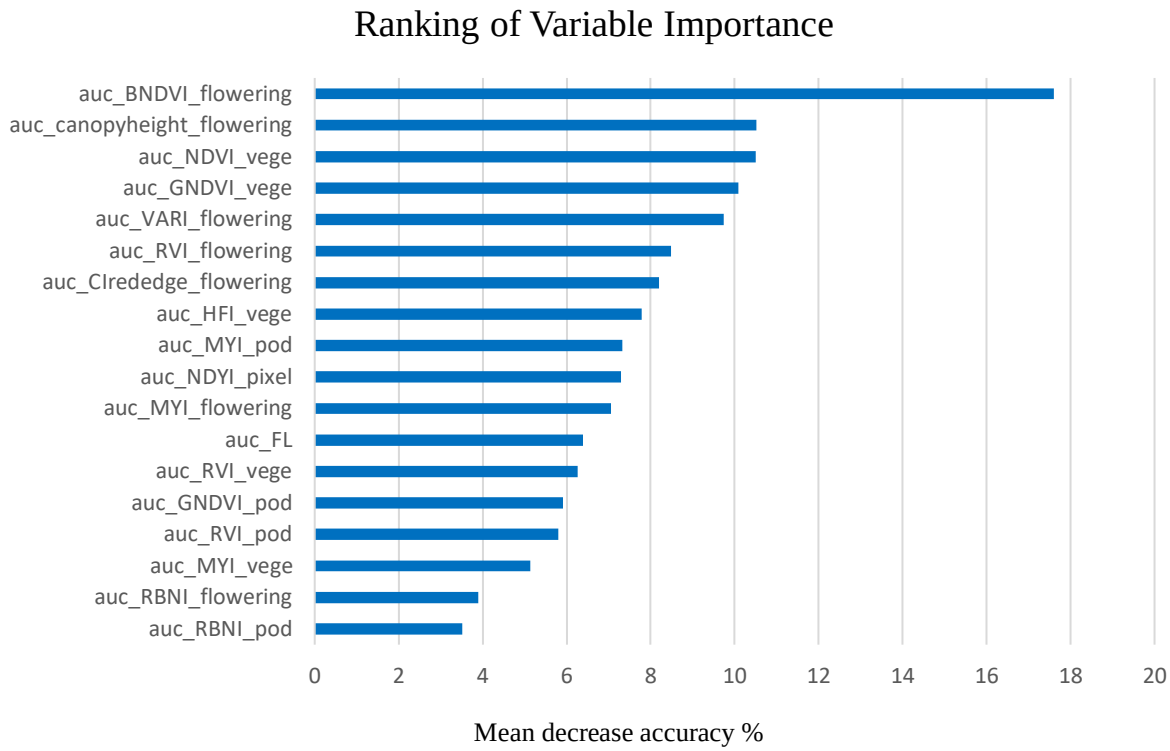


Figure 5.2 Ranking of the variable importance of various cumulative vegetation indices in the random forest regression model. auc_BNDVI represents cumulative BNDVI; auc_canopyheight represents cumulative UAV-derived canopy height; auc_NDVI represents cumulative NDVI; auc_GNDVI represents cumulative GNDVI; auc_VARI represents cumulative VARI; auc_RVI represents cumulative RVI; auc_CIrededge represents cumulative $CI_{red\ edge}$; auc_HFI represents cumulative HFI; auc_MYI represents cumulative MYI; auc_NDYI_pixel represents cumulative NDYI-based pixel number at the flowering stage; auc_FL represents cumulative flowering layer depth at the flowering stage; auc_RBNI represents cumulative RBNI; "_Vege" represents at the vegetative stage; "_flowering" represents at the flowering stage; "_pod" represents at the pod stage.

The ranking of variable importance in the RF regression model revealed that the cumulative BNDVI and the cumulative UAV-derived canopy height at the flowering stage, and the cumulative NDVI at the vegetative stage were the most important inputs for seed yield estimation in canola (Figure 5.2). Although Chapter 3 indicated that the cumulative NDYI-based pixel number during the flowering period was a significant indicator for yield estimation, the RF results demonstrated that the cumulative BNDVI at the flowering stage was more sensitive for seed yield estimation. During the flowering stage, both green vegetation and yellow flowers were visible before and after the peak flowering. The cumulative NDYI-based pixel number detected and represented the flowering production progress (Sulik and Long, 2016; Zhang et al., 2021b). The BNDVI value of each plot was the average reflectance of both green vegetation and yellow flowers. Chlorophyll absorbs at blue wavelength band and reflects NIR (Sulik and Long, 2016). Meanwhile, the yellowness of flowers due to carotenoids also

absorbs blue reflectance with few effects on NIR (Sulik and Long, 2015; Zhang et al., 2021b). Therefore, the greater BNDVI value could be expected with healthier green vegetation and more flowers and pods. This digital phenotype (i.e., cumulative BNDVI at the flowering stage) included more crop growth information compared with the cumulative NDYI-based pixel number. Moreover, the cumulative UAV-derived canopy height at the flowering stage derived from 3D DSMs was the second most important input variable in the RF. The accumulation of UAV-derived canopy height during the flowering stage included not only the canopy height information, but also the dynamic change of canopy height over time such as lodging or uniformity which had effects on the final seed yield (Chapter 4). Furthermore, as a common VI, NDVI extracted from a single image date could not explain many variations in yield estimation (Table 5.3), but the cumulative NDVI during the vegetative stage could be a promising yield indicator because it could capture the dynamic change of crop photosynthetic capacity at this stage using NIR and red reflectance. As expected, similar to NDVI, the cumulative NDVI became non-sensitive at the flowering stage because of the effect of yellow flowers (Sulik and Long, 2015, 2016; Zhang et al., 2021b).

Not surprisingly, the RF regression model underestimated seed yield in the validation dataset (slope = 0.38) (data not shown). This was probably because the vegetative parts (i.e., leaves, stem, and branches) and yellow flowering might be occluded by the upper part of crop canopy. These lower plant parts could not be captured by UAV-based images, and the missing crop growth information might reduce the accuracy of yield estimation. Similar results have been reported by Zhang et al. (2021b) who stated the missing flowering accumulation progress might be a reason of reduced accuracy of yield estimation using UAV-based imagery.

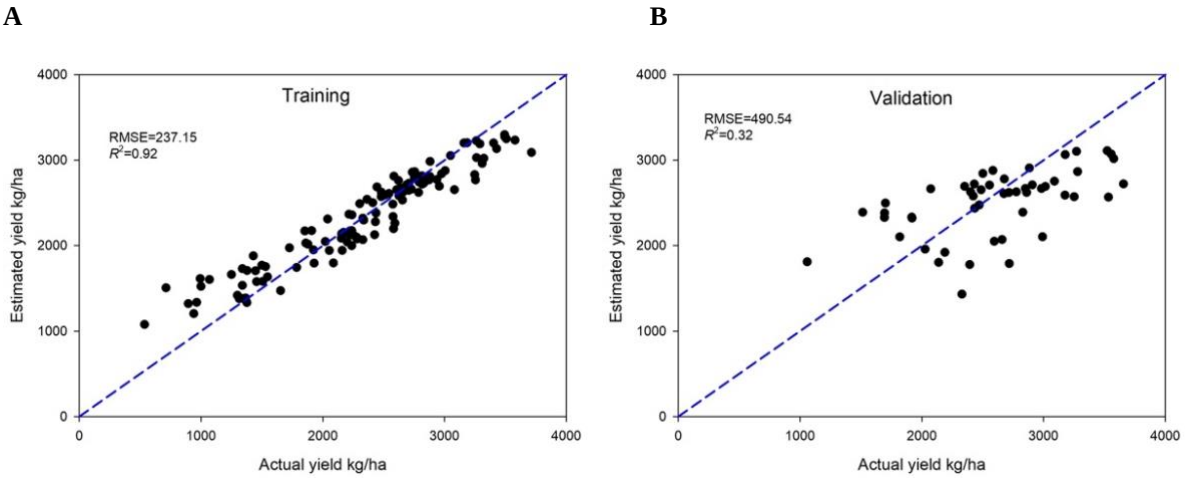


Figure 5.3 Actual seed yield compared with estimated seed yield with the random forest model using the cumulative BNDVI, the cumulative UAV-derived canopy height, the cumulative VARI, the cumulative RVI, the cumulative NDYI-based flowering pixel at the flowering stage, the cumulative NDVI, the cumulative GNDVI, the cumulative $CI_{red\ edge}$, the cumulative HFI at the vegetative stage, and the cumulative MYI at the pod stage for the training dataset (A) and the validation dataset (B). 70% of raw data were used as the training dataset and the remaining data were considered as the validation dataset. The blue dash line represents 1:1 reference line.

5.4 Conclusions

In this study, seed yield was regressed with the cumulative VIs and canopy structural phenotypes using a simple linear regression model and a random forest (RF) regression model with 56 genotypes grown under three site years. The RF regression model selected 9 cumulative VIs and a canopy structural phenotype (i.e., the cumulative BNDVI, the cumulative UAV-derived canopy height, the cumulative VARI, the cumulative RVI, the cumulative NDYI-based pixel at the flowering stage, the cumulative NDVI, the cumulative GNDVI, the cumulative $CI_{red\ edge}$, the cumulative HFI at the vegetative stage, and the cumulative MYI at the pod stage) as input indicators for yield estimation.

According to the ranking of variable importance, the RF model indicated that the cumulative BNDVI at the flowering stage, the cumulative UAV-derived canopy height at the flowering stage, and the cumulative NDVI at the vegetative stage were the most important variables for yield estimation, which combined multi-temporal crop spectral and canopy structural information. In addition, the vegetative and flowering stages could be the optimal growth stages for early yield estimation as only one cumulative VI was calculated at the pod stage within the 10 selected input variables. Moreover, compared with the simple linear regression model, the

RF regression model improved the accuracy of seed yield estimation in canola (R^2 = 0.32; RMSE = 490.54 kg/ha for the validation dataset). In the simple linear regression analysis, there was a significant relationship between the cumulative UAV-derived canopy height at the flowering stage and seed yield. Although the R^2 value of 0.18 was not as high as previous research, considering 56 various Brassica genotypes and three different site years (Table B.1, Appendix B), the cumulative UAV-derived canopy height at the flowering stage with a significant relationship was still acceptable as an individual indicator for canola yield estimation.

These results indicate that combining the cumulative VIs and a canopy structural phenotype over the vegetative and flowering stages extracted from UAV-based imagery have the potential to estimate seed yield under field conditions. The ten selected variables, especially the cumulative BNDVI and the cumulative UAV-derived canopy height at the flowering stage, and the cumulative NDVI at the vegetative stage were promising digital phenotypes which could assist plant breeders in yield estimation of diverse Brassica genotypes under multiple field conditions.

The future studies need to consider RGB imagery with higher resolution which may capture more crop growth information. Moreover, thermal sensors can be included in the future studies to gain new insights into field phenotyping and improve yield estimation accuracy. In addition, the selected input variables by the RF can be applied to satellite imagery which may allow regional-scale yield estimation and field scouting.

References

Araus, J. L., and Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. Trends in Plant Science, 19, 52-61. doi: 10.1016/j.tplants.2013.09.008

Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., et al. (2015). Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. International Journal of Applied Earth Observation and Geoinformation, 39, 79-87. doi: 10.1016/j.jag.2015.02.012

Domínguez, J. A., Kumhálová, J., and Novák, P. (2015). Winter oilseed rape and winter wheat growth prediction using remote sensing methods. Plant, Soil and Environment, 61, 410-416. doi: 10.17221/412/2015-PSE

Ebersbach, J., Khan, N. A., McQuillan, I., Higgins, E. E., Horner, K., Bandi, V., et al. (2021). Exploiting high-throughput indoor phenotyping to characterize the founders of a structured *B. napus* breeding population. Frontiers in Plant Science, 12, 780250-780250. doi: 10.3389/fpls.2021.780250

Feng, A., Zhang, M., Sudduth, K. A., Vories, E. D., and Zhou, J. (2019). Cotton yield estimation from UAV-based plant height. Transactions of the ASABE, 62, 393-404. doi: 10.13031/trans.13067

Feng, A., Zhou, J., Vories, E. D., Sudduth, K. A., and Zhang, M. (2020). Yield estimation in cotton using UAV-based multi-sensor imagery. Biosystems Engineering, 193, 101-114. doi: 10.1016/j.biosystemseng.2020.02.014

Fu, Z., Jiang, J., Gao, Y., Krienke, B., Wang, M., Zhong, K., et al. (2020). Wheat growth monitoring and yield estimation based on multi-rotor unmanned aerial vehicle. Remote Sensing, 12, 508. doi: 10.3390/rs12030508

Geipel, J., Link, J., and Claupein, W. (2014). Combined spectral and spatial modeling of corn yield based on aerial images and crop surface models acquired with an unmanned aircraft system. Remote Sensing, 6, 10335-10355. doi: 10.3390/rs61110335

Gitelson, A. A., Merzlyak, M. N., and Lichtenthaler, H. K. (1996). Detection of red edge position and chlorophyll content by reflectance measurements near 700 nm. Journal of Plant Physiology, 148, 501-508. doi: 10.1016/S0176-1617(96)80285-9

Gitelson, A. A., Zur, Y., Chivkunova, O. B., and Merzlyak, M. N. (2002). Assessing carotenoid content in plant leaves with reflectance spectroscopy. Photochemistry and Photobiology, 75, 272-281. doi: 10.1562/0031-8655(2002)0750272ACCIPL2.0.CO2

Gong, Y., Duan, B., Fang, S., Zhu, R., Wu, X., Ma, Y., et al. (2018). Remote estimation of rapeseed yield with unmanned aerial vehicle (UAV) imaging and spectral mixture analysis. Plant Methods, 14, 1-14. doi: 10.1186/s13007-018-0338-z

Han, J., Zhang, Z., and Cao, J. (2021a). Developing a new method to identify flowering dynamics of rapeseed using Landsat 8 and Sentinel-1/2. Remote Sensing, 13, 105. doi: 10.3390/rs13010105

Han, X., Wei, Z., Chen, H., Zhang, B., Li, Y., and Du, T. (2021b). Inversion of winter wheat growth parameters and yield under different water treatments based on UAV multispectral remote sensing. Frontiers in Plant Science, 12, 639. doi: 10.3389/fpls.2021.609876

Hussain, S., Gao, K., Din, M., Gao, Y., Shi, Z., and Wang, S. (2020). Assessment of UAV-onboard multispectral sensor for non-destructive site-specific rapeseed crop phenotype variable at different phenological stages and resolutions. Remote Sensing, 12, 397. doi: 10.3390/rs12030397

Johnson, M. D., Hsieh, W. W., Cannon, A. J., Davidson, A., and Bédard, F. (2016). Crop yield forecasting on the Canadian Prairies by remotely sensed vegetation indices and machine learning methods. Agricultural and Forest Meteorology, 218, 74-84. doi: 10.1016/j.agrformet.2015.11.003

Liaw, A., and Wiener, M. (2002). Classification and regression by randomForest. R News, 2, 18-22. Available online at: https://CRAN.R-project.org/doc/Rnews/. (Accessed December 01, 2021).

Liu, Y., Liu, S., Li, J., Guo, X., Wang, S., and Lu, J. (2019). Estimating biomass of winter oilseed rape using vegetation indices and texture metrics derived from UAV multispectral images. Computers and Electronics in Agriculture, 166, 105026. doi: 10.1016/j.compag.2019.105026

Noureldin, N. A., Aboelghar, M. A., Saudy, H. S., and Ali, A. M. (2013). Rice yield forecasting models using satellite imagery in Egypt. The Egyptian Journal of Remote Sensing and Space Science, 16, 125-131. doi: 10.1016/j.ejrs.2013.04.005

Pan, Z., Huang, J., and Wang, F. (2013). Multi range spectral feature fitting for hyperspectral imagery in extracting oilseed rape planting area. International Journal of Applied Earth Observation and Geoinformation, 25, 21-29. doi: 10.1016/j.jag.2013.03.002

Peng, Y., Li, Y., Dai, C., Fang, S., Gong, Y., Wu, X., et al. (2019). Remote prediction of yield based on LAI estimation in oilseed rape under different planting methods and nitrogen fertilizer applications. Agricultural and Forest Meteorology, 271, 116-125. doi: 10.1016/j.agrformet.2019.02.032

Ramos, A. P. M., Osco, L. P., Furuya, D. E. G., Gonçalves, W. N., Santana, D. C., Teodoro, L. P. R., et al. (2020). A random forest ranking approach to predict yield in maize with UAV-based vegetation spectral indices. Computers and Electronics in Agriculture, 178, 105791. doi: 10.1016/j.compag.2020.105791

Rouse, J. W., Haas, R. H., Schell, J. A., Deering, D. W., and Harlan, J. C. (1974). Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. NASA/GSFC Type III Final Report, Greenbelt, Md, 371.

Sankaran, S., Khot, L. R., Espinoza, C. Z., Jarolmasjed, S., Sathuvalli, V. R., Vandemark, G. J., et al. (2015). Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: a review. European Journal of Agronomy, 70, 112-123. doi: 10.1016/j.eja.2015.07.004

Singh, K. D., Duddu, H. S., Vail, S., Parkin, I., and Shirtliffe, S. J. (2021). UAV-Based hyperspectral imaging technique to estimate canola (*Brassica napus* L.) seedpods maturity. Canadian Journal of Remote Sensing, 47, 33-47. doi: 10.1080/07038992.2021.1881464

Statistics Canada. (2020). Production of principal field crops, November 2020. Available online at: https://www150.statcan.gc.ca/n1/daily-quotidien/201203/dq201203b-eng.htm (accessed January 05, 2021).

Sulik, J. J., and Long, D. S. (2016). Spectral considerations for modeling yield of canola. Remote Sensing of Environment, 184, 161-174. doi: 10.1016/j.rse.2016.06.016

Vina, A., Gitelson, A. A., Rundquist, D. C., Keydan, G. P., Leavitt, B., and Schepers, J. (2004). Monitoring maize (*Zea mays* L.) phenology with remote sensing. Agronomy Journal, 96, 1139-1147. doi: 10.2134/agronj2004.1139

Wan, L., Cen, H., Zhu, J., Zhang, J., Zhu, Y., Sun, D., et al. (2020). Grain yield prediction of rice using multi-temporal UAV-based RGB and multispectral images and model transfer–a case study of small farmlands in the South of China. Agricultural and Forest Meteorology, 291, 108096. doi: 10.1016/j.agrformet.2020.108096

Wang, L., Tian, Y., Yao, X., Zhu, Y., and Cao, W. (2014). Predicting grain yield and protein content in wheat by fusing multi-sensor and multi-temporal remote-sensing images. Field Crops Research, 164, 178-188. doi: 10.1016/j.fcr.2014.05.001

Wei, T., and Simko, V. (2021). R package 'corrplot': visualization of a correlation matrix. (version 0.92). Available online at: https://github.com/taiyun/corrplot. (Accessed December 01, 2021).

Xue, L., Cao, W., and Yang, L. (2007). Predicting grain yield and protein content in winter wheat at different N supply levels using canopy reflectance spectra. Pedosphere, 17, 646–653. doi: 10.1016/S1002-0160(07)60077-0

Xue, L., Cao, W., Luo, W., Dai, T., and Zhu, Y. (2004). Monitoring leaf nitrogen status in rice with canopy spectral reflectance. Agronomy Journal, 96, 135–142. doi: 10.2134/ agronj2004.0135.

Zhang, C., Craine, W. A., McGee, R. J., Vandemark, G. J., Davis, J. B., Brown, J., et al. (2021a). High-throughput phenotyping of canopy height in cool-season crops using sensing techniques. Agronomy Journal, 113, 3269-3280. doi: 10.1002/agj2.20632

Zhang, T., Vail, S., Duddu, H. S., Parkin, I. A., Guo, X., Johnson, E. N., et al. (2021b). Phenotyping flowering in canola (*Brassica napus* L.) and estimating seed yield using an unmanned aerial vehicle-based imagery. Frontiers in Plant Science, 12, 1178. doi: 10.3389/fpls.2021.686332

Zhou, X., Zheng, H. B., Xu, X. Q., He, J. Y., Ge, X. K., Yao, X., et al. (2017). Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 130, 246-255. doi: 10.1016/j.isprsjprs.2017.05.003

Contributions by others to Chapter 5

Dr. Steve J. Shirtliffe and Dr. Sally Vail designed the field experiments. Dr. Sally Vail and Dr. Isobel A. P. Parkin provided and prepared the plant materials. Dr. Steve J. Shirtliffe contributed to funding acquisition. Ti Zhang, Dr. Hema S. N. Duddu, and Dr. Steve J. Shirtliffe collected the images and performed imagery pre-processing. Ti Zhang performed the statistical analysis and wrote this chapter under the supervision of Dr. Steve J. Shirtliffe and with contributions from Dr. Sally Vail, Mr. Eric N. Johnson, Hansanee Fernado, Dr. Anjika Attanayake, Dr. Thuan Ha, Dr. Xulin Guo, and Dr. Rosalind A. Bueckert.

CHAPTER 6 GENERAL DISCUSSION AND CONCLUSION

6.1 General Discussion

Conventional phenotyping methods such as visual rating and manual measurement under field conditions are laborious, time-consuming, subjective, and destructive, which has become the bottleneck affecting genotype selection and new cultivar development with desirable crop performance (Araus and Cairns, 2014; Sankaran et al., 2015). With the advancement of aerial platforms mounted with imaging sensors, crop traits could be rapidly evaluated with relatively high accuracy (Araus and Cairns, 2014; Sankaran et al., 2015). In this thesis, the main hypothesis is agronomic traits of canola including flowering number, flowering layer depth, canopy height, and seed yield can be quantified or estimated by unoccupied aerial vehicle (UAV)-based multispectral imagery at reproductive stage.

Canola flowering number is an important and interesting crop trait related to pod number and final seed yield (Tayo and Morgan, 1975; Diepenbrock, 2000; Faraji et al., 2008; Kirkegaard et al., 2018; Zhang and Flottmann, 2018). With the bright yellow flowering and relatively long flowering period (Gan et al., 2016; Kirkegaard et al., 2018), the dynamics of flowering progress in the field could be captured by the UAV-based imagery and have the potential to estimate yield.

In the first study, it was hypothesized that yellow flower number could be estimated by spectral vegetation index and used for yield estimation (Chapter 3). Linear regression analysis was applied to investigate the relationship between the manually measured flower number and digitalized flowering number extracted from UAV-based imagery. The results indicated that normalized difference yellowness index (NDYI)-based pixel number had the potential to detect flower yellowness and estimate flower number with coefficient of determination (R^2) value up to 0.95. Several researchers found the ratio of green and blue and NDYI could successfully detect canola flower (Sulik and Long, 2016; d'Andrimont et al., 2020; Han et al., 2021). NDYI is a ratio of sum and difference between blue and green wavelength bands (Sulik and Long, 2016). Yellowness of flower petals decreases blue and increases red and green reflectance with little impact on near IR (NIR), therefore, NDYI is more sensitive to yellow flowers than other

greenness-related vegetation indices (VIs) using NIR (Sulik and Long, 2016). The average of NDYI for each plot contained both yellow flowering and green vegetation. Thresholding method was applied to extract flowering pixels and exclude non-flowering ones, which improved the sensitivity of NDYI to detect yellow flowers.

Chapter 3 showed that the NDYI-based pixel extracted from a single image was not consistently correlated with yield. The results contrast with Sulik and Long (2016) who found NDYI extracted at the peak flowering stage was consistently and significantly correlated with seed yield in canola with R^2 up to 0.76. The inconsistent relationships in my study might be due to more diverse genotypes involved and various environments (Table B.1, Appendix B) compared with the study by Sulik and Long (2016) who investigated fewer varieties. My field study evaluated 56 diverse breeding genotypes showing various flowering timing and growth habits. In addition, the field experiments were conducted under various environmental conditions (i.e., five site years), which made flowering timing and progress pattern more unpredictable. Therefore, it was difficult to determine an optimal image timing representing a specific development stage for all genotypes under various environmental conditions for yield estimation.

The area under the flowering progress curve math function was used to accumulate flowering progress over the stage. This new digital cumulative flowering phenotype was significantly and consistently correlated with seed yield with R^2 up to 0.42. Compared with using a single image, the cumulative NDYI-based pixels, as a new phenotype, captured the complete flowering progress, which provided more crop growth information and was significantly correlated with seed yield with higher R^2. Similar results were reported in wheat (Xue et al., 2007; Wang et al., 2014) and rice (Zhou et al., 2017). The authors stated that the cumulative vegetation indices improved yield estimation when comparing with using a single image (Xue et al., 2007; Wang et al., 2014; Zhou et al., 2017). These findings proved the first hypothesis that canola flower number was related to canopy reflectance and could be estimated by spectral VI (i.e., NDYI-based pixel number), and the accumulation of the NDYI-pixels over the flowering stage can be an indicator of seed yield in canola.

Crop growth status and seed yield can be estimated by two-dimensional (2D) canopy reflectance (Chapter 3). Canopy height, as a canopy structural feature, can be a potential

indicator of plant growth status and may assist in yield estimation (Ivanovska et al., 2007; Bendig et al., 2014; Iqbal et al., 2017; Assefa et al., 2018; Zhang et al., 2021). Additionally, flowering layer depth may show similar growth pattern as flowering number correlated with seed yield. These two canopy structural phenotypes extracted from three-dimensional (3D) digital surface models (DSMs) may be an alternative method for yield estimation.

The second hypothesis of my study is canopy structural phenotypes (i.e., flowering layer depth and canopy height) can be quantified using 3D DSMs extracted from UAV-based images and the digitalized canopy structural phenotypes can be indicators for yield estimation (Chapter 4). The results of Chapter 4 demonstrated that canopy height could be quantified using DSMs derived from UAV-based imagery. The optimum image timing was the peak flowering and early pod stages achieving the highest R^2 values between manually measured and UAV-derived canopy heights. Usually canola has reached its maximum height at the peak flowering and early pod stages. The taller plants are more visible and easier to be captured by imaging sensor due to relatively high spatial resolution. Iqbal et al. (2017) reported similar results that the accuracy of canopy height estimation using UAV imagery was higher for taller plants in poppy crop. Moreover, plant height changes dynamically before peak flowering, thus, time gap between the dates of imaging and ground reference data collection might be another reason for the weaker correlations between the manually measured and UAV-derived canopy heights before peak flowering. Furthermore, the non-uniform lodging condition at the late pod stage might also reduce the estimation accuracy because the limited manual measurements might not be representative for the whole plot (Iqbal et al., 2017; Morrison et al., 2021).

The manually measured flowering layer depth was significantly correlated with UAV-derived flowering layer depth at the peak flowering stage. The flowers at the early flowering stage were not occluded by the upper part of canopy, but the shorter depth might reduce the estimation accuracy due to relatively low image resolution at 25 m flight altitude or more errors from manual measurement (Morrison et al., 2021). Flowering layer depth was generally overestimated by UAV-based images, but the consistence of overestimation can still represent the trend of dynamic change of this flowering trait. The overestimation might be resulted from the underestimation of canopy height at the bolting stage due to thin and incompletely closed canopy.

In terms of yield estimation, like NDYI-based flowering pixel accumulation (Chapter 3), both UAV-derived canopy height and flowering layer depth accumulated over the flowering stage showed more consistent and significant relationships with seed yield compared with using a single image. The two digitalized cumulative phenotypes over the flowering stage included more crop growth information and might reveal more realistic change of crop status than a single image (Zhou et al., 2017). These findings proved the second hypothesis that flowering layer depth and canopy height could be quantified using UAV-based images, and the accumulation of these two phenotypes at reproductive stage (i.e., flowering stage) could be applied as new input variables for yield estimation.

Chapters 3 and 4 indicated that the digitalized cumulative flowering number and canopy structural traits could be applied as a single predictor variable for yield estimation in canola. Seed yield is an important and complex crop trait. A single predictor variable in a yield model may not adequately explain variations. The third hypothesis is that seed yield can be estimated using the combination of 2D and 3D imagery information collected at reproductive stage (Chapter 5).

In the third study, VIs were accumulated at the vegetative, flowering, and pod stages resulting in 27 cumulative VIs. According to Chapters 3 and 4, the cumulative NDYI-based pixel number, the cumulative UAV-derived canopy height, and the cumulative flowering layer at the flowering stage also demonstrated the potential to estimate yield; thus, they were included as input variables in the yield model. A commonly used machine learning method (i.e., random forest model) was applied in this study and it selected 10 variables as the inputs with the highest estimation accuracy. Compared with a simple linear regression model, a random forest (RF) regression model improved estimation accuracy and lowered error. The higher estimation accuracy was also observed in wheat yield (Fu et al., 2020) and canola biomass when using the RF (Liu et al., 2019).

The ranking of variable importance by the RF regression model showed three digital cumulative phenotypes including the cumulative blue normalized difference vegetation index (BNDVI) at the flowering stage, the cumulative UAV-derived canopy height at the flowering stage, and the cumulative normalized vegetation index (NDVI) at the vegetative stage were the most promising (i.e., top three) indicators for yield estimation. BDNVI consists of blue and

NIR which can be sensitive to both chlorophyll and yellowness of flower petals as chlorophyll absorbs blue and reflects NIR, and yellow petals absorb blue but have no or little impact on NIR (Sulik and long, 2016). The heathier green vegetation and more yellow flowers are expected to be correlated with greater BNDVI value. In this study, the cumulative BNDVI at the flowering stage correlated with the vegetation health status and the progress of flower production contributed most to the yield estimation. The cumulative UAV-derived canopy height at the flowering stage and the cumulative NDVI at the vegetative stage reflected the dynamic change of canopy height at the flowering stage and plant vigor at the early growth stage, respectively, demonstrating significant contributions to yield. Furthermore, the vegetative and flowering stages were the optimum period for early yield estimation using the cumulative UAV-derived phenotypes. These findings agreed the third hypothesis that the combination of 2D and 3D imagery information improved yield estimation with higher R^2 value and smaller root mean square error compared with applying a single predictor variable. However, different from the hypothesis, among the 10 cumulative phenotypes selected by the RF regression model, most of them were accumulated at the vegetative and flowering stages which indicated that both vegetative and flowering stages were the critical period for early yield estimation in canola using UAV-based imagery.

6.2 Implications

This research supports the introduction of UAV-based imagery to provide plant breeders and farmers with more reliable phenotypic data without damaging crop canopy. This new phenotyping method is less labour intensive, time consuming, destructive, and subjective compared to the conventional phenotyping methods. Plant breeders and farmers can get earlier crop traits evaluation for a large-scale field without harvesting, which may accelerate breeding process and assist field management. More specifically, the flowering number estimation study implied that canola flower number could be estimated more efficiently and objectively using UAV-based imagery. The canopy structure study indicated canopy height and flowering layer depth could be quantified using images with high to moderate estimation accuracy. The yield estimation study using a machine learning method implied that the accumulation of digitalized crop traits at the early development stages could monitor crop growth status and be applied in yield estimation.

6.3 Future Study

This dissertation evaluated four crop traits including flowering number, canopy height, flowering layer depth, and seed yield in canola across three years and three locations using 56 diverse Brassica genotypes (Table A.1, Appendix A). Future studies can investigate more crop traits such as days to flowering, flowering duration, plant lodging, biomass, crop growth rate, and nitrogen use efficiency. In breeding programs, plant breeders can use these digitalized phenotypes to evaluate and select diverse breeding genotypes with desirable traits more efficiently and frequently. For example, flowering timing and flowering duration derived from UAV-based imagery can be used to quickly identify genotypes with the optimal flowering traits to avoid local environmental stress such as heat, drought, and frost. In addition, to develop a more robust yield model, more digital crop traits can be included and validated in a real breeding project. Moreover, sensors with higher spatial resolution such as RGB camera can be applied to capture more details of crop traits. Thermal camera is another sensor to be considered in the future as it can provide canopy temperature to evaluate crop responses to biotic and abiotic stress. Furthermore, satellite images could be investigated for regional-scale field scouting and yield estimation.

References

Araus, J. L., and Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. Trends in Plant Science, 19, 52-61. doi: 10.1016/j.tplants.2013.09.008

Assefa, Y., Prasad, P. V., Foster, C., Wright, Y., Young, S., Bradley, P., et al. (2018). Major management factors determining spring and winter canola yield in North America. Crop Science, 58, 1-16. doi: 10.2135/cropsci2017.02.0079

Bendig, J., Bolten, A., Bennertz, S., Broscheit, J., Eichfuss, S., and Bareth, G. (2014). Estimating biomass of barley using crop surface models (CSMs) derived from UAV-based RGB imaging. Remote Sensing, 6, 10395-10412. doi: 10.3390/rs61110395

d'Andrimont, R., Taymans, M., Lemoine, G., Ceglar, A., Yordanov, M., and van der Velde, M. (2020). Detecting flowering phenology in oil seed rape parcels with Sentinel-1 and-2 time series. Remote Sensing of Environment, 239, 1-14. doi: 10.1016/j.rse.2020.111660

Diepenbrock, W. (2000). Yield analysis of winter oilseed rape (*Brassica napus* L.): a review. Field Crops Research, 67, 35-49. doi: 10.1016/S0378-4290(00)00082-4

Faraji, A., Latifi, N., Soltani, A., and Rad, A. H. S. (2008). Effect of high temperature stress and supplemental irrigation on flower and pod formation in two canola (*Brassica napus* L.) cultivars at Mediterranean climate. Asian Journal of Plant Sciences, 7, 343-351. doi: 10.3923/ajps.2008.343.351

Fu, Z., Jiang, J., Gao, Y., Krienke, B., Wang, M., Zhong, K., et al. (2020). Wheat growth monitoring and yield estimation based on multi-rotor unmanned aerial vehicle. Remote Sensing, 12, 508. doi: 10.3390/rs12030508

Gan Y.T., Harker, K. N., Kutcher, H. R., Gulden, R. H., Irvine, B., May, W. E., et al. (2016). Canola seed yield and phenological responses to plant density. Canadian Journal of Plant Science, 96, 151-159. doi: 10.1139/cjps-2015-0093

Han, J., Zhang, Z., and Cao, J. (2021). Developing a new method to identify flowering dynamics of rapeseed using landsat 8 and sentinel-1/2. Remote Sensing, 13, 105. doi: 10.3390/rs13010105

Iqbal, F., Lucieer, A., Barry, K., and Wells, R. (2017). Poppy crop height and capsule volume estimation from a single UAS flight. Remote Sensing, 9, 647. doi: 10.3390/rs9070647

Ivanovska, S., Stojkovski, C., Dimov, Z., Marjanović-Jeromela, A., Jankulovska, M., and Jankuloski, L. (2007). Interrelationship between yield and yield related traits of spring canola (*Brassica napus* L.) genotypes. Genetika, 39, 325-332. doi: 10.2298/GENSR0703325I

Kirkegaard, J. A., Lilley, J. M., Brill, R. D., Ware, A. H., and Walela, C. K. (2018). The critical period for yield and quality determination in canola (*Brassica napus* L.). Field Crops Research, 222, 180-188. doi: 10.1016/j.fcr.2018.03.018

Liu, Y., Liu, S., Li, J., Guo, X., Wang, S., and Lu, J. (2019). Estimating biomass of winter oilseed rape using vegetation indices and texture metrics derived from UAV multispectral images. Computers and Electronics in Agriculture, 166, 105026. doi: 10.1016/j.compag.2019.105026

Morrison, M. J., Gahagan, A. C., and Lefebvre, M. B. (2021). Measuring canopy height in soybean and wheat using a low-cost depth camera. The Plant Phenome Journal, 4, e20019. doi: 10.1002/ppj2.20019

Sankaran, S., Khot, L. R., Espinoza, C. Z., Jarolmasjed, S., Sathuvalli, V. R., Vandemark, G. J., et al. (2015). Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: a review. European Journal of Agronomy, 70, 112-123. doi: 10.1016/j.eja.2015.07.004

Sulik, J. J., and Long, D. S. (2016). Spectral considerations for modeling yield of canola. Remote Sensing of Environment, 184, 161-174. doi: 10.1016/j.rse.2016.06.016

Tayo, T. O., and Morgan, D. G. (1975). Quantitative analysis of the growth, development and distribution of flowers and pods in oil seed rape (*Brassica napus* L.). The Journal of Agricultural Science, 85, 103-110. doi: 10.1017/S0021859600053466

Wang, L., Tian, Y., Yao, X., Zhu, Y., and Cao, W. (2014). Predicting grain yield and protein content in wheat by fusing multi-sensor and multi-temporal remote-sensing images. Field Crops Research, 164, 178-188. doi: 10.1016/j.fcr.2014.05.001

Xue, L., Cao, W., and Yang, L. (2007). Predicting grain yield and protein content in winter wheat at different N supply levels using canopy reflectance spectra. Pedosphere, 17, 646–653. doi: 10.1016/S1002-0160(07)60077-0

Zhang, C., Craine, W. A., McGee, R. J., Vandemark, G. J., Davis, J. B., Brown, J., et al. (2021). High-throughput phenotyping of canopy height in cool-season crops using sensing techniques. Agronomy Journal, 113, 3269-3280. doi: 10.1002/agj2.20632

Zhang, H., and Flottmann, S. (2018). Source-sink manipulations indicate seed yield in canola is limited by source availability. European Journal of Agronomy, 96, 70-76. doi: 10.1016/j.eja.2018.03.005

Zhou, X., Zheng, H. B., Xu, X. Q., He, J. Y., Ge, X. K., Yao, X., et al. (2017). Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 130, 246-255. doi: 10.1016/j.isprsjprs.2017.05.003

APPENDIX A

Table A.1 Details of 56 Brassica genotypes involved in the canola trials.

NAM CODE	Name	Origin	Group	Canola Quality
NAM-23	Daichousen (mizuyasu)	North Korea	Asian	Yes
NAM-31	Wase Chousen	Korea	Asian	Yes
NAM-32	Dong Hae 2	South Korea	Asian	No
NAM-33	Dong Hae 3	South Korea	Asian	No
NAM-34	Dong Hae 11	South Korea	Asian	No
NAM-36	81N058-5	Unknown	Asian	No
NAM-39	Nakate Chousen	Korea	Asian	No
NAM-45	Dong Hae 6	South Korea	Asian	No
NAM-46	Dong Hae 21	South Korea	Asian	No
NAM-47	Dong Hae 23	South Korea	Asian	No
NAM-53	Buk Wuk 27	South Korea	Asian	No
NAM-66	N12-C11609	Unknown	Asian	No
NAM-8	86004	Unknown	Asian	Yes
NAM-83	SWU Chinese 9	China	Asian	No
NAM-85	DC21	South Korea	Asian	No
NAM-86	Kinki 22	South Korea	Asian	No
NAM-15	N01D-1330	Australia	Australian	Yes
NAM-37	Wesroona	Australia	Australian	No
NAM-38	Wesreo	Australia	Australian	No
NAM-4	Wesway	Australia	Australian	No
NAM-82	Tribune	Australia	Australian	Yes
NAM-0	N99-508	Canada	Canadian Adapted	Yes
NAM-12	46A65	Canada	Canadian Adapted	Yes
NAM-17	N00-C3661sp2	Canada	Canadian Adapted	Yes
DH27298	DH27298	Canada	Canadian Adapted	Yes
DH12075	DH12075	Canada	Canadian Adapted	Yes
NAM-51	AC Elect	Canada	Canadian Adapted	Yes
NAM-71	ACS N22	Canada	Canadian Adapted	Yes
NAM-72	ACS YN03-C656	Canada,winter background	Canadian Adapted	Yes
NAM-75	Magnum	Canada	Canadian Adapted	Yes
NAM-76	Ebony	Canada	Canadian Adapted	Yes
NAM-78	Unknown	Canada	Canadian Adapted	Yes
YN04-C1213sp013	YN04-C1213sp013	Canada	Canadian Adapted	Yes
NAM-1	Czyzowska	Poland	European	No
NAM-10	Global	Sweden or Denmark	European	Yes
NAM-13	Campino	Germany	European	Yes
NAM-14	Svalöf's Gulle	Sweden	European	No
NAM-28	Topas	Sweden	European	Yes
NAM-30	Egra	Unknown	European	Yes
NAM-40	Mlochowski	Poland	European	No
NAM-56	MAZOWIECKI	Poland	European	No

NAM-57	Mozart	Denmark	European	Yes
NAM-65	Puma	Sweden or Denmark	European	Yes
NAM-68	Lirawell	Germany	European	Yes
NAM-73	Optima	Denmark	European	No
NAM-87	Tanto	France	European	Yes
NAM-88	SRS1632	Poland	European	No
NAM-25	Nolza 541	Argentina	Other	Yes
NAM-26	Nolza 531	Argentina	Other	Yes
NAM-29	PSA12	RSYN	Other	No
NAM-42	PI432392	Bangladesh	South Asian	No
NAM-43	PI432395	Bangladesh	South Asian	No
NAM-5	BN-1	India	South Asian	No
NAM-79	PAK 85912	Pakistan	South Asian	No
B.carinata	GID-5461	NA	NA	NA
B.juncea	Centennial Brown	NA	NA	No

NA represents no data available.

APPENDIX B

Table B.1 Analysis of variance and covariance table of fixed and random effects on seed yield of 56 Brassica genotypes at Saskatoon, SK, Canada from 2016 to 2018 (i.e., three site years). The fixed effect was genotype, and the random effects were site year and the interaction between site year and genotype.

Effect	Z value	F value	*p* value
Site year	0.77	NA	0.22
Site year x genotype	5.47	NA	< 0.0001***
Genotype	NA	9.25	< 0.0001***

*, **, ***, significant at the 0.05, 0.01, and 0.001 probability, respectively.
NA represents no data available.

Table B.2 Analysis of variance and covariance table of fixed and random effects on seed yield of the selected 16 Brassica genotypes at Saskatoon, Melfort, and Scott, SK, Canada from 2016 to 2018 (i.e., five site years). The fixed effect was genotype, and the random effects were site year and the interaction between site year and genotype.

Effect	Z value	F value	*p* value
Site year	1.39	NA	0.08
Site year x genotype	3.81	NA	< 0.0001***
Genotype	NA	8.12	< 0.0001***

*, **, ***, significant at the 0.05, 0.01, and 0.001 probability, respectively.
NA represents no data available.