University of São Paulo
"Luiz de Queiroz" College of Agriculture

Multispectral aerial images to phenotype yield potential and tree inventory mapping: case studies in dry pea (*Pisum sativum*) and apple (*Malus domestica*) nursery

Juan José Quirós Vargas

Dissertation presented to obtain the degree of Master in Science. Area: Agricultural Systems Engineering

Piracicaba
2017

Juan José Quirós Vargas
Agronomist

Multispectral aerial images to phenotype yield potential and tree inventory mapping: case studies in dry pea (*Pisum sativum*) and apple (*Malus domestica*) nursery

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Advisor:
Prof. Dr. **THIAGO LIBÓRIO ROMANELLI**

Dissertation presented to obtain the degree of Master in Science. Area: Agricultural Systems Engineering

Piracicaba
2017

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP

Quirós Vargas, Juan José

Multispectral aerial images to phenotype yield potential and tree inventory mapping: case studies in dry pea (*Pisum sativum*) and apple (*Malus domestica*) nursery / Juan José Quirós Vargas. - - versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2017.

52 p.

Dissertação (Mestrado) - - USP / Escola Superior de Agricultura "Luiz de Queiroz".

1. Índice de vegetação 2. Sensoriamento remoto 3. Inventário 4. Fenotipagem I. Título

DEDICATION

This work is dedicated to Jesus Christ, the only and real Master, and to my family: Vanessa and Violeta, who are my greatest motivation to keep on fighting for our dreams. With only six months of being born, Violeta has already taught me what true love, happiness, and gratitude in God are.

ACKNOWLEDGEMENT

Foremost, infinite thanks to God, Jesus Christ.

I am grateful with Vanessa Vargas for her support on my life: as beautiful and sweet as strong and brave to face calm and difficult moments in our lives. I thank my family in Costa Rica, my parents and brothers Sonia and Victor, Leonardo, Monica, Ana and Victor Jr., who regardless the distance are always present in my mind and heart; They are my greatest examples to follow. To my nephews: The most difficult proof during this time was staying away from them.

I thank my advisor, Prof. PhD Thiago Romanelli, for all the attention, trust and support he has given me through this academic stage. I am glad to have work with professors Lav Khot and Sindhuja Sankaran from Washington State University (WSU); they trusted on me and, gave me the opportunity of joining his laboratory team, providing the data and invaluable academic orientation for this research.

In the same way, I am grateful with professors PhD. Carlos Vettorazzi (ESALQ/USP) and PhD. João Camargo (Embrapa Informática Agropecuária) for their technical assistance all around this process. I also thank Professor PhD. Ricardo F. Peterson for accepting me in this excellent university. Thanks to Professor PhD. Roberto Fritsche Neto, for his important participation in my final exam, bringing to the interior of an engineering environment a very valuable vision from the perspective of genetics and biology. I also appreciate his help in the last corrections to this document.

I thank my friends Mauricio Martello, Tiago Rodrigues and Rodrigo Muniz, who have always been present to advise me and support me. With his friendship they have make me feel in family, as I was in my contry. I also thank Mr. José Aguero, for been an example to follow in these first years of my professional life. Without the financial assistance of CAPES it would not be possible to do this work.

BIOGRAPHY

I was born in a neighborhood at the south of Costa Rican capital city, in the heart of a family as happy as a hard worker. As a teenager, I learned about computer-aided design at Don Bosco Technical College, which later served to work as topographical designer while pursuing a career in agronomy at the University of Costa Rica (UCR). This has made me want to learn about agricultural mapping, which a few years later has become what is now my passion in agronomy: precision agriculture. The interest in understanding the patterns of spatial distribution in a crop led me to enter into the remote sensing area, which was the central theme of this master I did in this great institution that is ESALQ/USP.

EPIGRAPH

“Have always in mind that you can
recover money, love,
honor, or anything else lost. But you
will never be able recover
the lost time.”.

Roberto Bolaños

CONTENTS

RESUMO

Imagens aéreas multiespectrais para fenotipagem e contagem de plantas: estudos de caso em ervilha (*Pisum sativum*) e viveiro de maçã (*Malus domestica*)

A coleta de dados de campo envolve processos de grande consumo em tempo e dinheiro, ademais de levar o risco de possíveis erros de medição. Com o avanço tecnológico nos últimos anos, surgiram ferramentas de sensoriamento remoto de baixo custo para facilitar procedimentos de medição em campo, sendo uma das técnicas mais conhecidas o uso de câmeras multiespectrales acopladas a um ARP. Essas ferramentas são complementadas pela implementação de procedimentos em programas SIG e de processamento de imagens, a partir dos quais são desenvolvidas metodologias que visam extrair valores alvo desde um determinado conjunto original de dados. Neste trabalho, foram utilizadas imagens multiespectrais no desenvolvimento de dois estudos de caso: (1) para estimativa de produtividade em parcelas para pesquisa de ervilha, e (2) para contagem de plantas em um viveiro de maçã plantado diretamente no solo; ambos os campos localizados no estado de Washington, EUA. No primeiro caso, foi criada uma metodologia confiável e replicável para estimativa de produtividade como técnica de fenotipagem de alto rendimento; enquanto no segundo caso, foi desenvolvido um algoritmo capaz de identificar o número de plantas de maçã com mais de 95% de exatidão. Em ambos os estudos, o sensoriamento remoto é usado como uma ferramenta eficiente e prática na melhora de operações de campo.

Palavras-chave: Índice de vegetação; Sensoriamento remoto; Inventário; Fenotipagem de alto rendimento

ABSTRACT

Multispectral aerial images to phenotype yield potential and tree inventory mapping: case studies in dry pea (*Pisum sativum*) and apple (*Malus domestica*) nursery

Field data collection involves time and money consuming processes, additionally carrying possible measurement errors. With the technological advance in the last years, low cost remote sensing tools have emerged to facilitate procedures for in-field measurements, being one of the most known techniques the use of multispectral cameras coupled to RPA. These tools are complemented by the implementation of procedures in GIS and image-processing software, from which are developed methodologies leading to extract target values from a certain original set of data. In this work, multispectral images were used in two case studies: (1) for yield estimation in pea plots for breeding research, and (2) for plant counting in an apple nursery planted directly on the soil; both fields are located in Washington State, USA. In the first case, a reliable and replicable methodology for yield estimation was created as a high throughput phenotyping technique; while in the second case an algorithm capable of identifying the number of apple plants with more than 95% accuracy was developed. In both studies, remote sensing is used as an efficient and practical way to improve field operations under the specified conditions of each case.

Keywords: Vegetation index; Remote sensing; Inventory; High-throuhput Phenotyping

1. MULTISPECTRAL AERIAL IMAGING TO PHENOTYPE YIELD POTENTIAL IN DRY PEA (*PISUM SATIVUM*)

ABSTRACT

Dry pea (*Pisum sativum*) is an important grain legume and Palouse region of the Pacific Northwest U.S. is one of the major producing area. The USDA-ARS dry pea breeding program focuses on developing varieties of both spring and autumn-sown high yield peas and have high levels of resistance to biotic and abiotic stresses. Yield potential is one of the critical traits influencing the selection of high performing varieties. In this study, low-altitude aerial multispectral imaging was performed to phenotype yield differences, aiming to create a reliable image processing protocol to estimate yield potential in a dry pea breeding program. Three dry pea experiments with 25, 16 and 10 varieties at two locations were used for this purpose. Images were acquired at budding, flowering and pod development stages. Normalized difference vegetation index (NDVI) and green normalized vegetation index (GNDVI) image based features (SUM, MIN, MAX, MEAN) were extracted using image GIS and image processing software. MEAN NDVI and GNDVI was consistently more correlated with yield. The best scenarios for yield estimation were at the beginning of pod development, where linear regressions were generated with R^2 of 0.89 and 0.76 to 0.79 for Pullman and Genesee, respectively. The equations of those regressions were used to estimate yield, reaching accuracies of 89 % (± 9.33) and 93 % (± 6.25), being reliable to perform yield predictions in the same conditions of this study. The methodology here presented can be replicated as a reliable image processing protocol to estimate yield potential for dry pea breeding experiments.

Keywords: Field phenomics; NDVI; GNDVI; Breeding; Image analysis

1.1 INTRODUCTION

Pea (*Pisum sativum*) is a cold season annual grain legume crop that is produced worldwide for human consumption and animal feed (McPhee, 2003; Coyne et al., 2011). Pea is often grown in rotation with cereal grain to break the disease cycle, and its symbiotic relationship with *Rhizobium* leguminosarium bacteria enhances the soil fertility through atmospheric nitrogen fixation. Moreover, dry pea is nutritious source of food, feed and fodder, rich in protein, soluble starch, vitamins, and minerals (Cheng et al., 2015). Dry pea is also

economically important crops with production in more than 87 countries (McPhee, 2003). The Palouse region in Pacific Northwest is the leading producer of dry pea in the USA (USDA-NASS, 2016). The dry pea breeding program of the United States Department of Agriculture (USDA) - Agricultural Research Service (ARS) has released several dry pea varieties over the years (McPhee and Muehlbauer, 2002, 2004; McGee et al., 2012; McGee and McPhee, 2012; McGee et al., 2013) with higher yield potential and stress tolerance.

Phenotyping is a key aspect of plant breeding programs, which refers to evaluation of crop growth factors and performances resulting from interaction between the genetic make-up and the environment. In general, field phenotyping is challenging due to the resource requirements of on-field measurements, such as personnel and time (Araus and Cairns, 2014). Currently, remote sensing (RS) has emerged as a practical alternative for high-throughput crop evaluations in field research plots.

RS works with data taken by sensors at certain distance from the plants, from few centimeters or meters (proximal sensing) up to hundreds of meters (aerial sensing) or kilometers away (orbital sensing) (Florenzano, 2011). Both aerial and satellite images have been used for yield estimation in rice (Tennakoon et al., 1992; Swain et al., 2010), sugarcane (Rao et al., 2002; Murillo and Carbonal, 2012), sunflower (Vega et al, 2015), corn (Shanahan et al., 2001; Fang et al., 2008; Yao et al., 2015), wheat (Serrano et al., 2000; Paiva et al., 2011), and cotton (Brandão et al., 2011).

Remotely Piloted Aircraft (RPA) can be integrated with other sensors to acquire aerial data at desired spatial and temporal resolution to estimate yield potential (Zhang and Kovacs, 2012). Both low (multispectral) and high spectral resolution (hyperspectral) sensors can be utilized for this purposes. Although aerial imaging for yield estimation has been explored, there is very limited literature in validity of the approach in plant breeding programs (Haghighattalab et al., 2016; Shi et al., 2016). Given the sizable scale of plant breeding plots (up to fee hectares), multirotor RPAs integrated with multispectral cameras can acquire useful spectral (visible and near infrared) information for phenotyping applications (Sankaran et al., 2016).

One of the major aspects of RS-based phenotyping application is the image processing. Aerial Images collected using RPA-based multispectral camera needs to be processed accurately in order to extract usable information. For these purpose, it is necessary to have image processing protocols for image analysis and spatial data organization. These processing steps include methods such as supervised and unsupervised classification, segmentation, extraction

of vegetation indexes (VIs), and zonal statistics among others. Moreover, the protocols need to be evaluated for a specific conditions in order to utilize the techniques confidently in the breeding programs and accelerate the image processing time.

Therefore, the major objective of this study was to create a reliable image processing protocol to estimate yield potential for dry pea breeding program. For this purpose, were evaluated four statistical parameters (MEAN, MIN, MAX, SUM) of two VIs (NDVI and GNDVI), at two locations and three data collection moments (budding, flowering and pod development). We differenciate also between the use of individual and mosaic images.

1.2 MATERIAL AND METHODS

1.2.1 Field experimental plot

Data were collected at two locations: Pullman (46.700032 ° N, and -117.138405 ° W) and Genesee (48.600468 ° N, -118.118065 ° W), Washington, USA. These regions are major dry pea growing locations with variable weather and soil type. Pullman and Genesee field sites with pea (Pasivum sativum) breeding trials are shown in Figures 1a and 2a; the dimension of the plots are 1.5 m wide by 7.5 m long. The dry pea entries were planted in Genesee and Pullman on 21 April, 2015 and 26 April, 2015, respectively. A seed treatment was applied to all seeds prior to planting that contains the fungicides fludioxonil (0.56 g kg^{-1}, Syngenta, Greensboro, NC), mefenoxam (0.38 g kg^{-1}, Syngenta, Greensboro, NC), and thiabendazole (1.87 g kg^{-1}, Syngenta, Greensboro, NC), thiamethoxam (0.66 mL kg^{-1}, Syngenta, Greensboro, NC) for insect control, and molybdenum (0.35 g kg^{-1}). Pea entries were planted at a density of 86 seeds m^{-2} (about 860,000 seeds ha^{-1}). A randomized complete block design with three replications was used for all field location trials. Weeds were controlled by a single post-plant/pre-emergence application of metribuzin (0.42 kg ha-1, Bayer Crop Science, Raleigh, NC) and linuron (1.34 kg ha^{-1}, NovaSource, Phoenix, AZ).

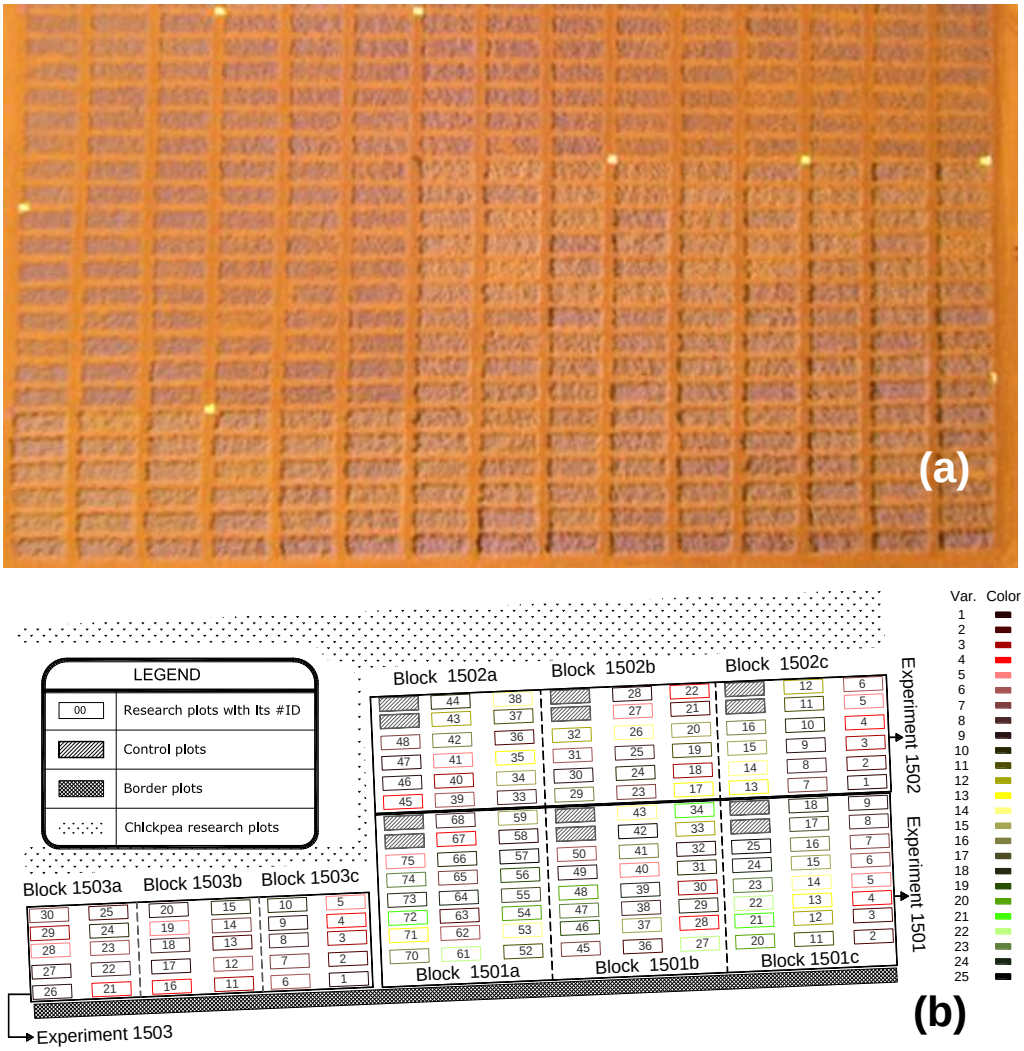


Figure 1 - (a) False color aerial image (R, G, NIR) at budding stage; and (b) distribution of experiments, blocks, and plots in Pullman field site.

There were three experiments in pea trials (labeled as 1501, 1502, and 1503, different panels) in Pullman field site with 75, 48 and 30 plots, respectively (3 replicate plots per entry). In experiment, 1501 and 1503 the varieties planted were green pea, while in experiment 1502 were yellow pea. The number of varieties planted in 1501 was 25 varieties, experiment 1502 was 16 varieties, and experiment 1503 was 10 varieties. Each block had the same number of plots, and the experiment 1503 is the only one without control or check plots (Figure 1b). In Genesee field site, the experiment 1501 and 1502 were planted with the same experimental design, number of plots, and varieties same as those described for Pullman field site (Figure 2b). Plots were harvested using a Zurn® 150 plot combine and seed was further cleaned mechanically with an air blower. Seed yield (kg ha^{-1}) was determined for all entries at both locations.

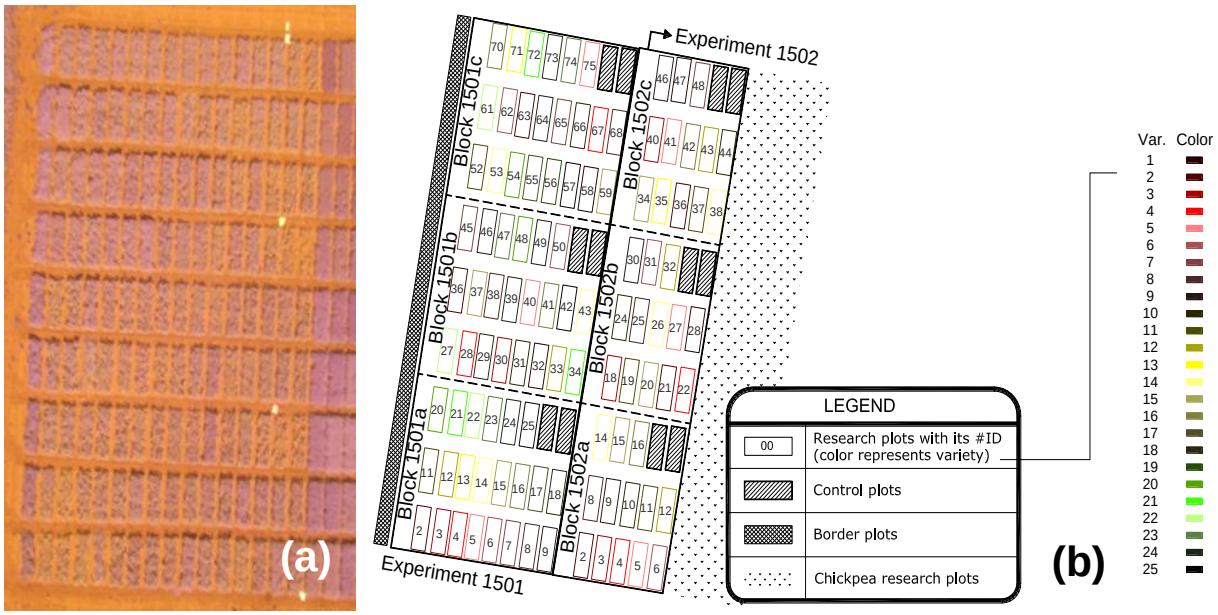


Figure 2 - (a) False color aerial image (R, G, NIR) at flowering; and (b) distribution of experiments, blocks, and plots in Genesee field site.

1.2.2 Remote sensing data collection

The RPA-multirotor OktoXL 6S12 (HiSystems GmbH, Moormerland, Germany) was used for imaging the field plots. It has a maximum load capacity of 2.5 kg, powered by a 6500 mAh Lithium-ion polymer battery, and capable of manual and automated control. This platform has embedded several sensors to maintain stability over the flight such as accelerometer, gyroscope, compass, GPS (Global Positioning System), and pressure sensor. For data acquisition in the present study, the RPA was manually controlled with a 4 km radio transmitter (MX20 Hott, Graupner, Stuttgart, Germany). The flight time was approximately 10 min.

The sensor used was a 3-band modified multispectral camera (Canon Powershot ELPH 340 HS, LDC LLC, Carlstadt, New Jersey, USA). The sensor captures information in the red, green, and near infrared regions of the electromagnetic spectrum; identified as bands 1, 2, and 3 in this study, respectively. Was placed on a gimbal with the ability to adjust the position of the camera during the flight. The sensor firmware has been set to acquire images at intervals of 5 s. The 8-bit, 16-megapixel images were stored on the camera's memory card. The data was acquired at 110 m above ground level (AGL) with a resolution of 5 cm per

pixel at this height; it´s angle of view at 25 mm of focal length is 51.3 °. At Genesee field location, additionally 40 m data was acquired to see the effect of using mosaic high resolution imaging; for this case images parallel and perpendicular overlap were 75 % and 85 %, respectively.

At each location, data was collected at least twice in 2015 during the crop cycle. In Pullman field site, airborne images were taken at budding, flowering, and pod development. At pod development, the dry pea trials 1502 and 1503 were harvested due to early senescence. The first data collection was close to reproductive stage 3 (R3), or budding stage; while the second and third were close to reproductive stages 4 (R4) and 5 (R5), flowering and pod development stages respectively. In Genesee field side, images were taken when the plots were at flowering and pod development stages, or R4 and R5, flowering and pod development. In flowering, leaves had already begun to turn yellow in early maturing varieties. At pod development, the leaves progress from yellowing process to senescence phase (Schwartz and Langham, 2010).

1.2.3 Data processing

ArcGIS (10.2, ESRI, Redlands, California) software was used for all raster and vector data processing. For the mosaics generation, the Agisoft Photoscan imaging platform (Agisoft LLC, Russia) was used. Statistical parameters were extracted from each plot, taking as data population of a variable the digital numbers (DN) of the image pixels of the two VI images extracted in this study: normalized difference vegetation index and green normalized difference vegetation index. The parameters extracted were mean (MEAN), maximum (MAX), minimum (MIN) and summation (SUM). The MEAN refers to the average of the values found within each plot polygon; the MAX and MIN represent the highest and lowest values within each polygon that delimits a plot; and SUM refers to the sum of all values comprised within a polygon or plot. The proposed methodology for obtaining the mentioned values for each plot is composed of two phases: first vector processing and subsequent raster processing (Figure 3).

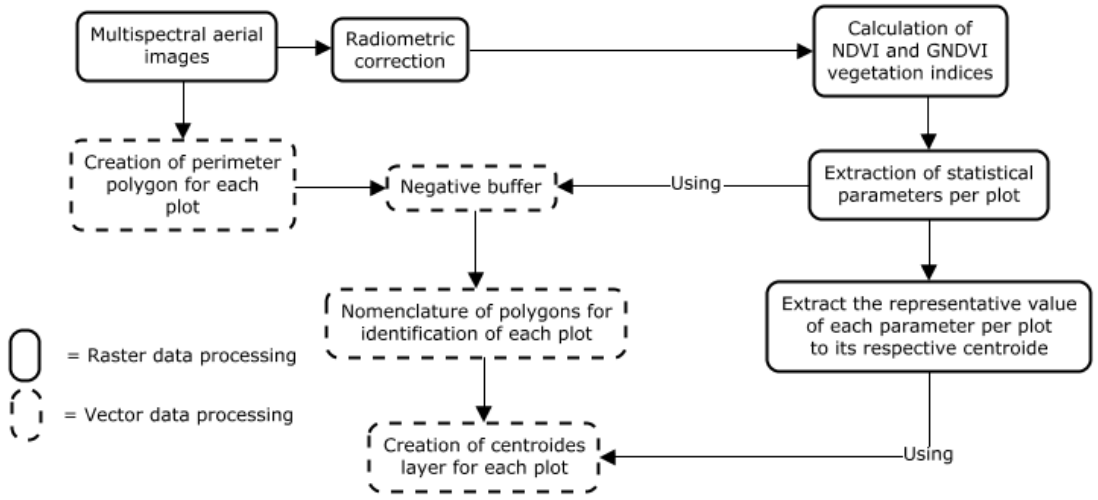


Figure 3 - Raster and vector data processing steps using ArcGIS.

The goal of vector processing was to generate a layer of points in which each plot is represented by a point from which statistical parameters can be extracted (Figure 4). The first step was to create the polygons covering all the area of each plot; then, in order to eliminate any border effect, a negative buffer was created, which refers to a parallel boundary towards the interior of the plot. The distance between the original border and the one created with the negative buffer was 0.3 m (Bareth et al., 2015) (Figure 4f). Each polygon generated in the last step was differentiated with a nomenclature allowing its identification according to the location, time of imaging, and experiment to which it belonged.

The first step in raster data processing was to generate the raster NDVI (Equation 1) and GNDVI (Equation 2) for each image taken at a given time of imaging and location, using the "Raster Calculator" tool of ArcGIS 10.2.

NDVI = (Near Infrared - Red)/(Near Infrared+ Red) **Eq. 1**

GNDVI = (Near Infrared - Green)/(Near Infrared+ Green) **Eq. 2**

For the determination of vegetation indices, radiometrically corrected images were used. This correction was performed using a 25 x 25 cm white reference panel (Spectralon Reflectance Target, CSTM-SRT-99-100, Spectra Vista Cooperation, New York, USA). The panel represents the maximum gray level in the resulting 8-bit image (256 levels). For each image of NDVI and GNDVI, the parameters mentioned above (MEAN, MAX, MIN, SUM) were

extracted using the tool "Zonal Statistics", and lastly, the "Extract Multi Values" command was used to extract each value to the point representing each plot.

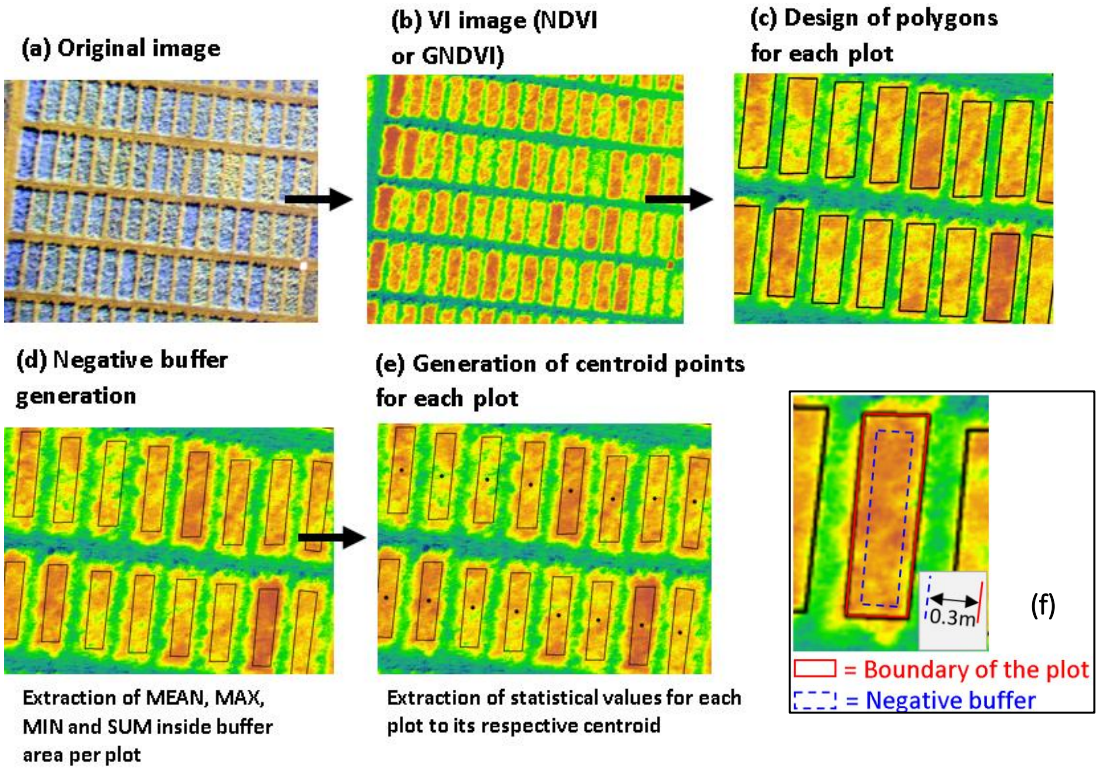


Figure 4 - Sequence of steps followed in the data processing to calculate the parameters MEAN, MAX, MIN and SUM from NDVI or GNDVI images for each plot; being (a) and (b) the original and VI images, (c) the polygons designed per plot, which in (d) are with the negative buffer, and (e)are extracted extracted its centroid points.

1.2.4 Statistical analysis

The attributes table of the resulting point layer was exported to **.txt* format. SAS version 9.3 (SAS Institute, Raleigh, NC) was used to calculate the Pearson's correlation coefficients (r) between extracted parameters and yield based on the average yield and image features of the three repetitions per variety. The results of these correlations were organized by experiment, and separated according to the statistical parameter, time of imaging, location, and vegetation index. For discriminating between high and low correlations, a 0.75 threshold was defined. The scenario of statistical parameter presenting constantly correlations equal or higher to that limit was identified as the best scenario

correlating with yield, and was used for further steps in the methodology; scenarios presenting correlations lower than 0.75 were discarded.

Statistical parameter with higher correlations was identified for each experiment. Linear regression yield equations were generated for those highly correlated scenarios to develop relationships for yield prediction. Each yield equation was validated by calculating the percentage of accuracy for each individual plot using the following calculation (Eq. 3):

$$\text{Accuracy (\%)} = |100-|(EY*100)/AY|| \quad \textbf{Eq. 3}$$

Where, EY represents the estimated yield and AY represents the actual yield. For a better comprehension of the results, the dispersion of the results was studied by calculating coefficient of variation (CV):

$$CV = SDA/AA \quad \textbf{Eq. 4}$$

where, AA refers to average accuracy in yield estimation using the respective yield equation in individual plots, and SDA refers to standard deviation. Similar data processing steps were applied for mosaic images of NDVI and GNDVI in Genesee, aiming to analyze different performance between stitched and non-stitched raster data sets as well as higher resolution images. Yield equations resulting in and lower SDA and CV estimations were defined as representative yield equations for future predictions in breeding plots in identical conditions of crop, locality, and growth stage.

Finally, effect of distortion correction on the image quality was evaluated. The correction was performed using Agisoft's PhotoScan software (Agisoft LLC, St. Peterburg, Russia), which uses a 2-step procedure to generate a rectified image. In the first step, common points are identified between the images with which they are aligned. In this step, an arbitrary coordinate system is generated to position the camera along with a cloud of low-density points, and a self-calibration process of the camera is performed. In the second step, a cloud of points of high density is generated based on the relative position of the camera detected in the previous step, with which it is sought to generate depth information of objects and/or surfaces captured in the image. There is a third step in the workflow of this

program, but it is focused on the generation of 3D models (Moutinho et al., 2015), which was not utilized in this study.

1.3 RESULTS

1.3.1 Correlations between vegetation indices and seed yield using individual images

In most of the cases, the correlation coefficients between the statistical parameter and seed yield were significantly high (p>0.05).

Table 1 - Correlations coefficient (r) for each experiment regarding zonal statistics, growth stage, VI and location.

Zonal statistics	Feature	Location						
		Pullman				Genesee		
		Growth stage	1501	1502	1503	Growth stage	1501	1502
			Correlation coefficient				Correlation coefficient	
SUM	NDVI	Budding	0.62	0.59	0.47	-	-	-
	GNDVI		0.72	0.69	0.76		-	-
	NDVI	Flowering	0.83	0.34	0.93	Flowering	0.76	0.56
	GNDVI		0.81	0.27 *	0.93		0.75	0.58
	NDVI	Pod development	–0.01*	-	-	Pod development	0.57	0.33
	GNDVI		–0.25*	-	-		0.71	0.18*
MIN	NDVI	Budding	0.60	0.54	0.36	-	-	-
	GNDVI		0.57	0.50	0.45		-	-
	NDVI	Flowering	0.68	0.46	0.85	Flowering	0.57	0.30
	GNDVI		0.62	0.42	0.81		0.51	0.43
	NDVI	Pod development	–0.36 *	-	-	Pod development	0.25	0.43
	GNDVI		–0.05*	-	-		0.18*	0.46
MAX	NDVI	Budding	0.62	0.68	0.64	-	-	-
	GNDVI		0.34	0.56	0.51		-	-
	NDVI	Flowering	0.73	0.45	0.82	Flowering	0.71	0.51
	GNDVI		0.65	0.38	0.69		0.63	0.40
	NDVI	Pod development	–0.66	-	-	Pod development	–0.03*	0.20
	GNDVI		–0.56*	-	-		0.47	0.07*
MEAN	NDVI	Budding	**0.80**	0.69	**0.81**	-	-	-
	GNDVI		**0.75**	0.59	**0.76**		-	-
	NDVI	Flowering	**0.86**	0.45	**0.93**	Flowering	**0.77**	0.56
	GNDVI		**0.83**	0.31	**0.93**		**0.75**	0.58
	NDVI	Pod development	–**0.86**	-	-	Pod development	0.57	0.54
	GNDVI		–0.72	-	-		0.71	0.43

*All the correlations are significant at p>0.05 with the exception of those marked with **

Special emphases was given to the data, when r was consistently higher than 0.75. Only MEAN statistics in experiments 1501 and 1503 in Pullman presented regularly correlations higher than 0.75, being selected to generate linear regressions between the yield and VI data per plot in the three imaging times (Figure 5). No constant high correlation was found in Genesee.

1.3.2 Linear regressions between vegetation indices and seed yield using individual images

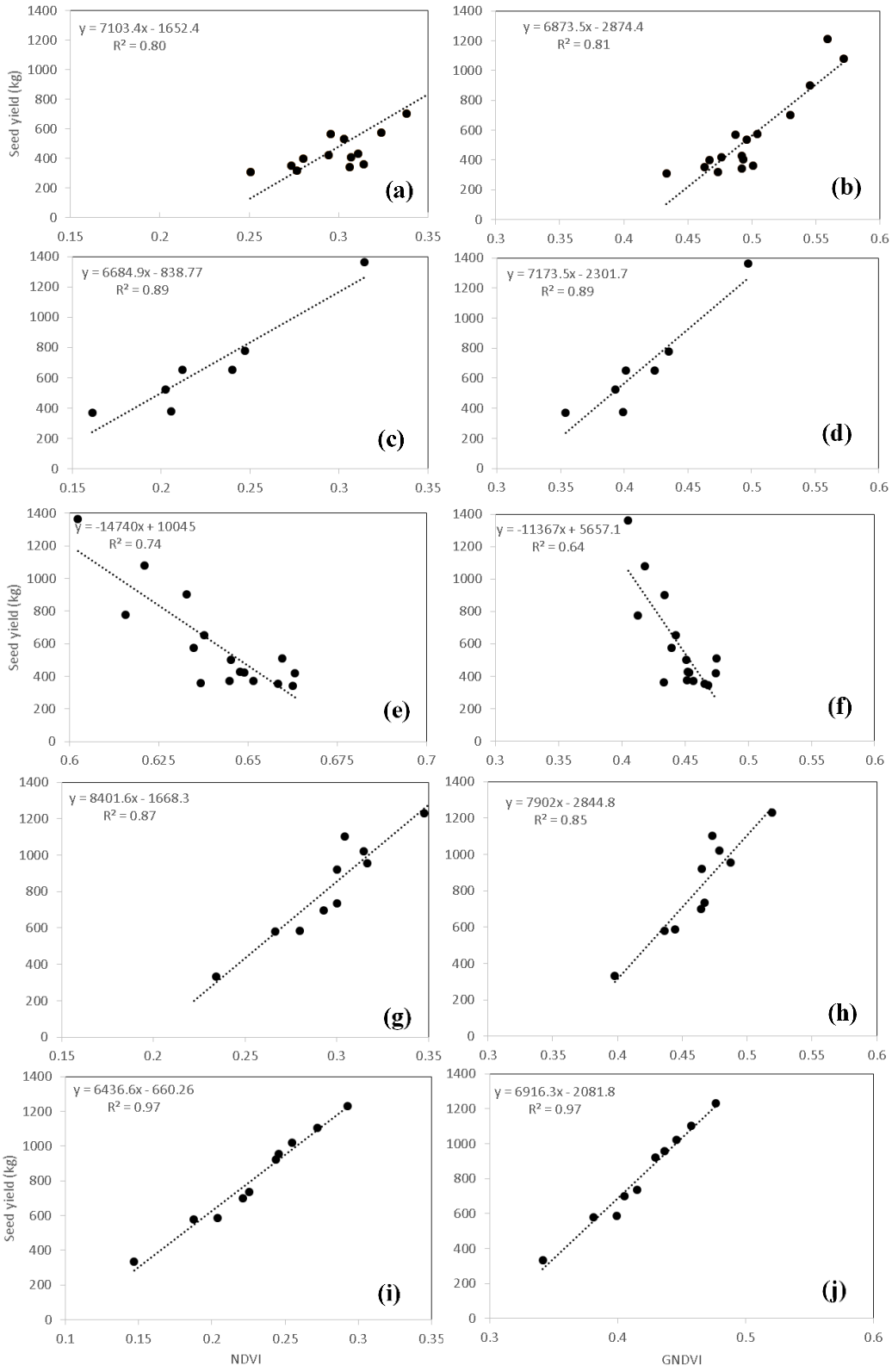


Figure 5 - Linear regressions between yield and MEAN VI (NDVI and GNDVI) for experiments in Pullman 1501 at budding stage be for budding (a and b), flowering (c and d) and pod development (e and f); and 1503 at budding (g and h) and flowering (i and j).

A mosaic image of Genesee was analyzed also obtaining VI-Yield correlations, separating between scenarios with correlations values constantly higher and lower than 0.75 (Table 2).

1.3.3 Correlations between vegetation indices and seed yield using mosaic images

Table 2 - Correlations regarding zonal statistics, phenological stage, VI and location using mosaic images of experiments 1501 and 1502 in Genesee.

Zonal statistics	DAP	VI	Genesee 1501	Genesee 1502
SUM	Flowering	NDVI	0.79	0.53
	Flowering	GNDVI	0.80	0.35
SUM	Pod development	NDVI	0.47	0.16*
MIN	Pod development	GNDVI	0.72	0.35
	Flowering	NDVI	0.63	0.38
	Flowering	GNDVI	0.71	0.04*
MIN	Pod development	NDVI	0.48	0.07*
MAX	Pod development	GNDVI	0.64	(-)0.08*
	Flowering	NDVI	0.71	0.01
	Flowering	GNDVI	0.70	0.11*
MAX	Pod development	NDVI	0.20*	0.43
MEAN	Pod development	GNDVI	0.18*	(-)0.03*
	Flowering	NDVI	**0.79**	0.53
	Flowering	GNDVI	**0.79**	0.37
MEAN	Pod development	NDVI	0.57	0.47
	Pod development	GNDVI	0.73	0.37

*All the correlations are significant at p>0.05 with the exception of those marked with **

As observed with individual images, using mosaic raster regularly correlations higher than 0.75 were obtained with MEAN, in this case, more specifically at budding stage in experiment 1501. This scenario was selected to generate linear regressions between the yield and VI data per plot (Figure 6). No high correlation statistics was found with the mosaic image in experiment 1502.

1.3.4 Linear regressions between vegetation indices and seed yield using mosaic images.

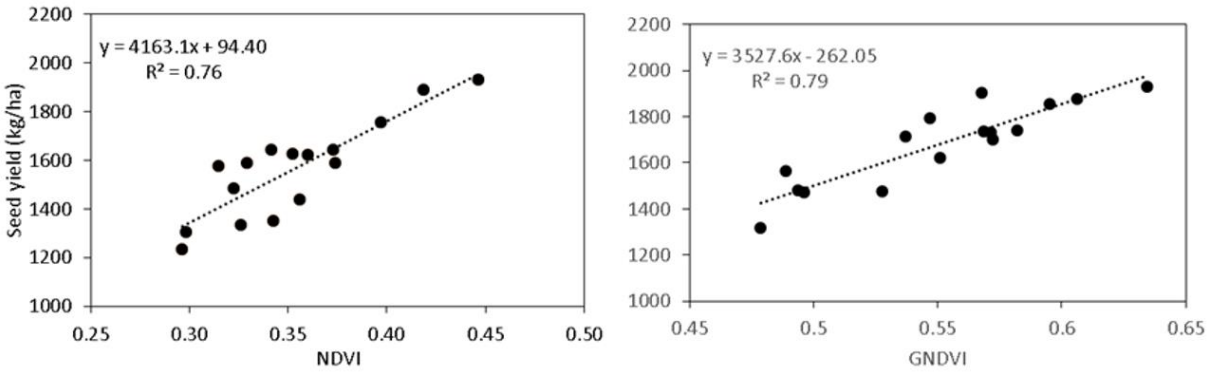


Figure 6 - Linear regressions between MEAN VI (NDVI and GNDVI) and yield using mosaic images of experiment 1501 in Genesee.

1.3.5 Statistical analysis

The percentage of accuracy and CV for the estimations of yield through linear regressions with individual and mosaic images regarding growth stage and VI were evaluated.

Table 3 - Percentage of accuracy for the estimations of yield using linear regressions with individual and mosaic images regarding growth stage and VI.

Type of image	DAP	VI	Experiment			
			1501		1503	
			Acc. (%) (+-SD)	CV	Acc. (%) (+- SD)	CV
Individual (Pullman)	Budding	NDVI	71.04 (+-29.44)	0.40	81.90 (+-15.65)	0.19
		GNDVI	70.12 (+-26.67)	0.38	74.39 (+-18.55)	0.25
	Flowering	NDVI	77.03 (+-19.59)	0.25	**89.00 (+-8.55)**	**0.10**
		GNDVI	75.44 (+-21.93)	0.29	**88.09 (+-9.33)**	**0.11**
	Pod development	NDVI	60.00 (+-31.48)	0.53	-	-
		GNDVI	67.33 (+-31.41)	0.47	-	-
Mosaic (Genesee)	Flowering	NDVI	**93.08 (+-6.25)**	**0.07**	-	-
		GNDVI	**92.44 (+-6.55)**	**0.07**	-	-

For Pullman estimations only experiment 1503 at flowering stage had acceptable percentage of accuracy with low CV (<0.20) with values of 89.00 % (+-8.55) and 88.09 % (+-9.33) when using NDVI and GNDVI, respectively. All the other scenarios in this location presented a combination of low accuracy with high variability (high SD and CV), which means that results will not be replicable in further experiments on same conditions.

Contrary to what is observed in Genesee individual images, where low correlations were found, using mosaic images for this location gave high accuracy results of 93.08 % (+-6.25) and 92.44 % (+- 6.55) when using NDVI and GNDVI images, respectively. These high accuracy values outweigh those obtained in Pullman experiment 1503 at flowering.

Based on those results, equations of regressions presented in Figure 5 (i) and (j), and in Figure 6 can be used for yield prediction in further crops in the same conditions of this study. Equations 5 and 6 are suggested for yield prediction using individual images in the same conditions of Pullman experiment 1503 using NDVI and GNDVI (respectively) images at flowering:

$$y = 6436.6x - 660.26 \quad \textbf{Eq. 5}$$

$$y = 6916.3x - 2081.8 \quad \textbf{Eq. 6}$$

In turn, equations 7 and 8 are suggested for yield prediction using mosaic images in the same conditions of Genesee experiment 1501 using NDVI and GNDVI (respectively) images at flowering:

$$y = 4163.1x + 94.398 \quad \textbf{Eq. 7}$$

$$y = 3527.6x - 262.05 \quad \textbf{Eq.8}$$

1.3.6 Graphical summary

Taking account of all the data colections, the two VI, 4 statistical parameters, 3 experiments and 2 types of image, more than 4500 individual correlations were calculated, where only two scenarios were selected as reliable for future yield estimations. Those scenarios are highlighted in Figure 7, which gives a graphical view of the decantation of results within the complete data pool.

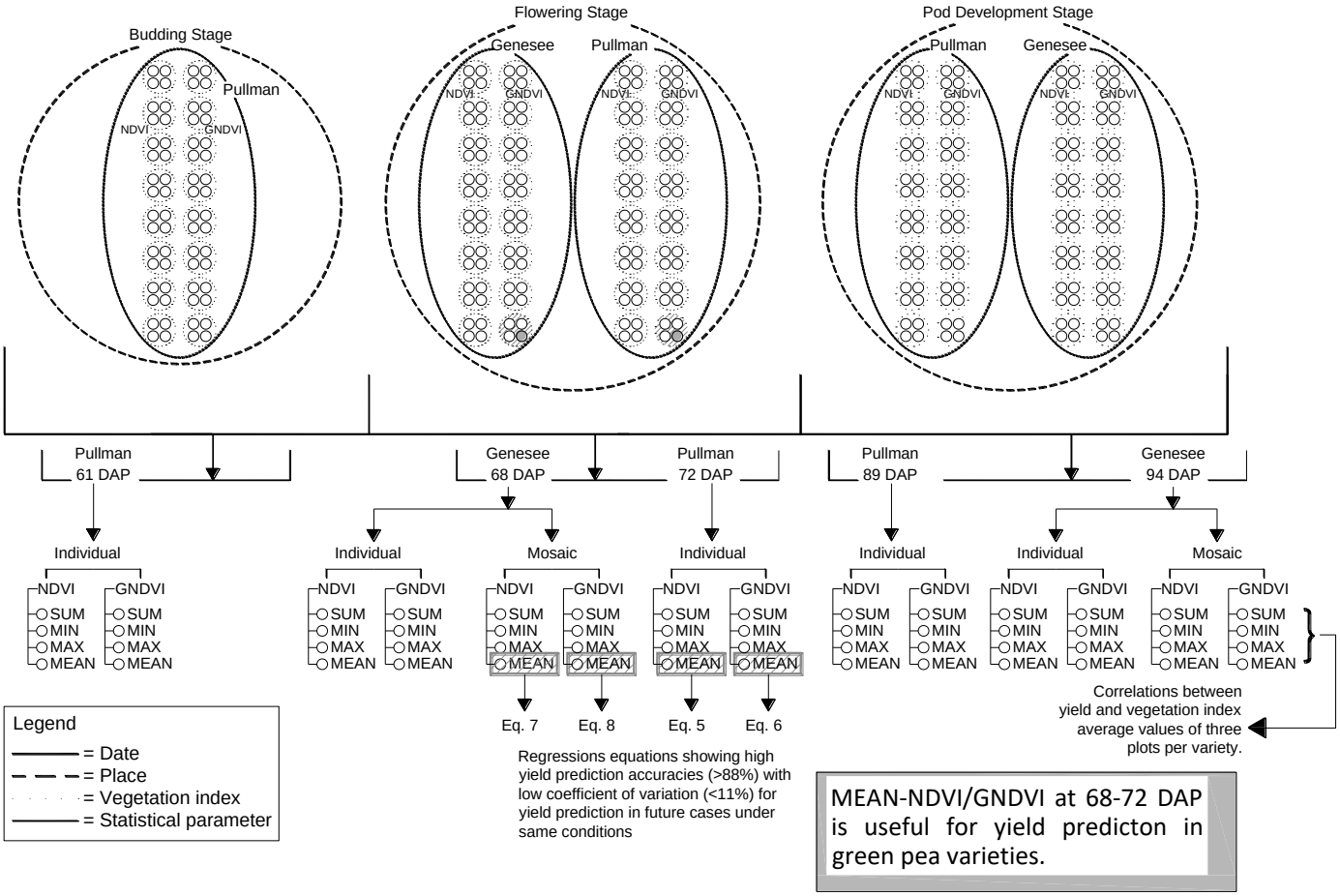


Figure 7 - Graphic summary of the steps followed in the proposed methodology and results obtaining.

1.4 DISCUSSION

Using MEAN, being the sum of all the DN divided by the numbers of cells covering the region, constantly higher correlations were observed since it is a value depending on both the amount of biomass as well as the area covered; this was observed by means of both, individual and mosaic images. With calculation of SUM there is a lack of spatial dimension that lead to results, despite good, not as accurate as the ones found when using MEAN. One other disadRPAage of using SUM is when we have presence of negative DNs of VI, like at pod development stage in Pullman, when adding those negative signs to positive ones loose the sense of the sum. Similar case arise when using MIN and MAX, which values do not guarantee to represent the region analyzed since could occur randomly, for example, for pixels of soil, weeds or plants of abnormal vigor exposed just by chance.

The best scenarios for yield estimation for both localities were at beginning pod filling phenological state (budding and flowering stages respectively), for which four equations (Eq. 5, 6, 7 and 8) were generated for future yield predictions using GNDVI and NDVI data.

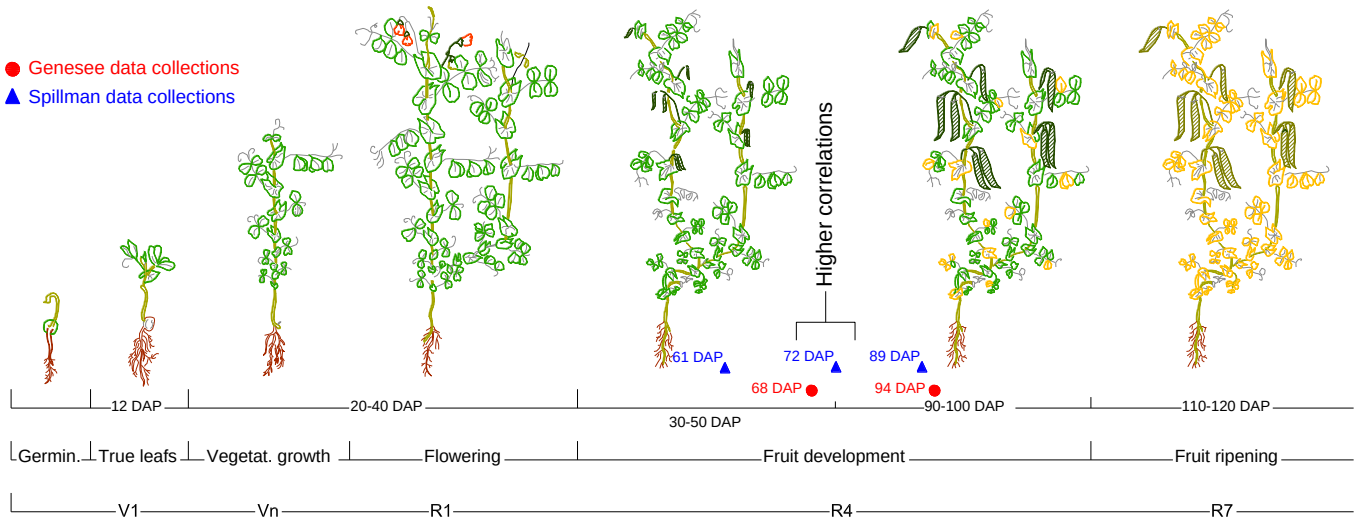


Figure 8 - Imaging moments in Genessee and Pullman according to crop phenological.

In Pullman, no good results are with experiment 1502 since its plots had planted yellow dry pea, different from experiments 1501 and 1503, which have plots with green dry pea. Differences in the growth of both types of pea may explain the lower correlations in experiment 1502.

Results were not affected by VI saturation since all the imaging moments were after the greatest vegetative growth, which according to Zając et al. (2012) is the moment to see this phenomenon, losing its sensitivity to changes in the amount of biomass (Gu et al., 2013). The negative high correlations at pod development stage in Pullman experiment 1501 are attributed to the effect of yellowing leafs (Sankaran et al., 2015) due to the combination of starting senescence and a general stress that affected the crop (direct communication with Rebecca McGee, August 2016), worsened by conditions due to the geographic location (Appendix A). The main conditions differentiating both places are: lower temperature and higher precipitation and humidity in Pullman when compared with Genesee (Table 4).

Table 4 - Weather conditions in June and July in Pullman and Genesee.

	June		July	
	Pullman*	Genesee**	Pullman*	Genesee**
Minimum temp. (°C)	8.70	14.20	11.20	15.70
Average temp. (°C)	18.60	23.60	21.10	24.80
Maximum temp. (°C)	27.30	31.80	29.50	33.10
Relative humidity (°C)	57.50	43.50	45.90	44.10
Precipitation (mm)	5.00	0.00	5.00	0.30
Solar radiation (MJ/m^2)	815.00	822.00	820.00	831.00

**255masl / **768masl*

According to Burkart et al. (2015), depending on it´s magnitude, differences in the angle of the image can cause significant variations in the values of VIs, as it is the case of NDVI. Some low altitude (40 m) images collected in Genesee at pod development, in experiment 1502, presented a certain angle of inclination when they were captured (Appendix B); this may have influenced results with mosaic image processing in this scenario, since there were observed the lower correlations and was the only case where the difference between the correlation found in the mosaic image and the individual differed by 0.05.

Table 5 - Average correlations obtained between the MEAN value of NDVI and GNDVI in 20 randomly selected plots in experiments 1501 and 1502, using individual and mosaic images.

Experiment	VI	Flowering		Pod development	
		Individual	Mosaic	Individual	Mosaic
		Average coorelation			
1501	NDVI	0.77	0.79	0.57	0.57
	GNDVI	0.75	0.7	0.71	0.73
1502	NDVI	0.56	0.53	0.5	0.47
	GNDVI	0.58	0.37*	0.42	0.37*

**Difference higher ore qual than 0.05.*

To evaluate the effect of the inclination of the images mentioned, a test was done calculating the MEAN NDVI of 20 plots with the same image with and without geometric correction (Appendix C). The average percentage error was 2 %, with a standard deviation of 1.6, which means that the inclination in those images did not affect the results of correlations obtained with individual images.

A high value of R2 does not ensure a perfect fitting of the equation of a regression for predictions with different sets of data; that is whys linear regressions were validated. Only four regressions were confirmed as confidence for yield predictions because of it high accuracy and low CV (<0.20): Pullman experiment 1503 at flowering stage with individual images, and Genesee experiment 1501 at budding stage mosaic image. Equations Eq. 5 and 6 can be employed for yield estimation with multispectral individual images of dry green pea breeding plots taken at flowering stage in Pullman; while Eq. 7 and 8 can be employed for yield estimation with multispectral mosaic images of dry green pea breeding plots taken at budding stage in Genesee.

The reader of this study how want to replicate the methodology can use a lower limit of person coefficient. Since yield is the result of the interaction of many genes the threshold set at 0.75 can be very strict, so that it can be lowered down to 0.50 regarding the researcher criteria. Values higher than 0.75 can be used when working with the correlation between VI and crop features resulting from the expression of one or few genes.

1.5 CONCLUSIONS

For both Vis analyzed in this study, NDVI and GNDVI, MEAN was the statistical parameter showing consistently higher correlations with yield.

With images taken at the beginning of pod development we obtained higher correlations compared with the other two dates of data collection.

The use of mosaic images showed only gave different results when images presented certain inclination.

The methodology of this study can be replicated as a reliable image processing protocol to estimate yield potential for dry pea breeding experiments.

REFERENCES

Araus J, Cairns J (2014) Field high-throughput phenotyping: the new crop breeding frontier. **Trends in Plant Science** 19: 152-61

Brandão Z, et al. (2011) Estimativa da produtividade do algodoeiro através de imagens de satélite. **8º Congresso Brasileiro de Algodão & I Cotton Expo** 1602-1609

Bareth G, Bolten A, Hollberg J, Aasen H, Burkart A, Schellberg J (2015) Feasibility study of using non-calibrated RPA-based RGB imagery for grassland monitoring: case study at the Rengen Long-term Grassland Experiment (RGE), Germany. **DGPF Tagungsband** 24: 1-7

Burkart A, et al. (2015) Angular Dependency of Hyperspectral Measurements over Wheat Characterized by a Novel RPA Based Goniometer. **Remote Sensing** 7: 725-746

Cheng P, et al. (2015) Association mapping of agronomic and quality traits in USDA pea single-plant collection. **Molecular breeding** 35: 75

Coyne CJ, McGee RJ, Redden RJ, Ambrose MJ, Furman BJ, Miles CA (2011) **Genetic adjustment to changing climates: Pea. Crop Adaptation to Climate Change** 238-250

Fang H, Liang S, Hoogenboom G, Teasdale J, Cavigelli M (2008) Corn-yield estimation through assimilation of remotely sensed data into the CSM-CERES-Maize model. **International Journal of Remote Sensing** 29: 3011-3032

Florenzano TG (2011) Iniciação em SR. **Oficina de Textos** 128

Gu Y, et al. (2013) NDVI saturation adjustment: A new approach for improving cropland performance estimates in the Greater Platte River Basin, USA. **Ecological Indicators** 30: 1–6

Haghighattalab A, Pérez LG, 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(1): 35

McPhee KE, Muehlbauer FJ (2002) Registration of 'Franklin' green dry pea.(Registrations Of Cultivars). **Crop Science** 42: 1378-1379

McGee RJ, et al. (2012) Registration of pea germplasm lines partially resistant to aphanomyces root rot for breeding fresh or freezer pea and dry pea types. **Journal of Plant Registrations** 6: 203-207

McGee RJ, McPhee KE (2012) Release of autumn-sown pea germplasm ps03101269 with food-quality seed characteristics. **Journal of Plant Registrations** 6: 354-357

McGee RJ, Nelson H, McPhee KE (2013) Registration of 'Lynx' Winter Pea. **Journal of Plant Registrations** 7: 261-264

McPhee K (2003) Dry pea production and breeding: A minireview. **Journal Of Food Agriculture And Environment** 1: 64-69

McPhee KE, Muehlbauer FJ (2004) Registration of 'Stirling' green dry pea. **Crop Science** 44: 1868-1870

Murillo P, Carbonel J (2012) Principios y aplicaciones de la precepción remota en el cultivo de la caña de azúcar en Colombia. **Centro de Investigación de la Caña de Azúcar CENICAÑA** 181

Moutinho O, et al. (2015) Estudo comparativo de software fotogramétrico em diferentes ambientes com RPAS: Comercial vs. Open Source.**VIII Conferencia Nacional de Cartografia e Geodesia**

Paiva C, Tsukahara R, França G (2013) Estimativa da produtividade da cultura do trigo via SR no município de Piraí do Sul, no Estado do Paraná. **Anais XVI Simpósio Brasileiro** de SR – SBSR 158-164

Rao PV, Rao VV, Venkataratnam L (2002) Remote sensing: A technology for assessment of sugarcane crop acreage and yield. **Sugar Tech** 4: 97-101

Sankaran S, et al. (2015) Field-based crop phenotyping: Multispectral aerial imaging for evaluation of winter wheat emergence and spring stand. **Computers and Electronics in Agriculture** 118, 372–379

Sankaran S, Khot LR, Espinoza CZ, Jarolmasjed S, Sathuvalli VR, Vandemark GJ, 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

Schwartz HF, Langham MA (2010) Growth stages of pea (Psivum sativum L.). **Legume ipmPIPE**. Available online: https://www.ndsu.edu/pubweb/pulse-info/growthstages-pdf/peagrowthstagescards.pdf

Serrano L, Filella I, Penuelas J (2000) Remote sensing of biomass and yield of winter wheat under different nitrogen supplies. **Crop Science** 40: 723-731

Shanahan JF, et al. (2001) Use of remote-sensing imagery to estimate corn grain yield. **Agronomy Journal** 93: 583-589

Shi Y, Thomasson JA, Murray SC, Pugh NA, Rooney WL, Shafian S, et al. (2016) Unmanned aerial vehicles for high-throughput phenotyping and agronomic research. **PloS one** 11(7): e0159781

Swain KC, Thomson SJ, Jayasuriya HP (2010) Adoption of an unmanned helicopter for low-altitude remote sensing to estimate yield and total biomass of a rice crop. **Transactions of the ASABE** 53: 21-27

Tennakoon SB, Murty VV, Eiumnoh A (1992) Estimation of cropped area and grain yield of rice using remote sensing data. **International Journal of Remote Sensing** 13: 427-439

Vega F, Carvajal F, Perez M, Orgaz F (2015) Multi-temporal imaging using an unmanned aerial vehicle for monitoring a sunflower crop. **ScienceDirect** 132: 19-27

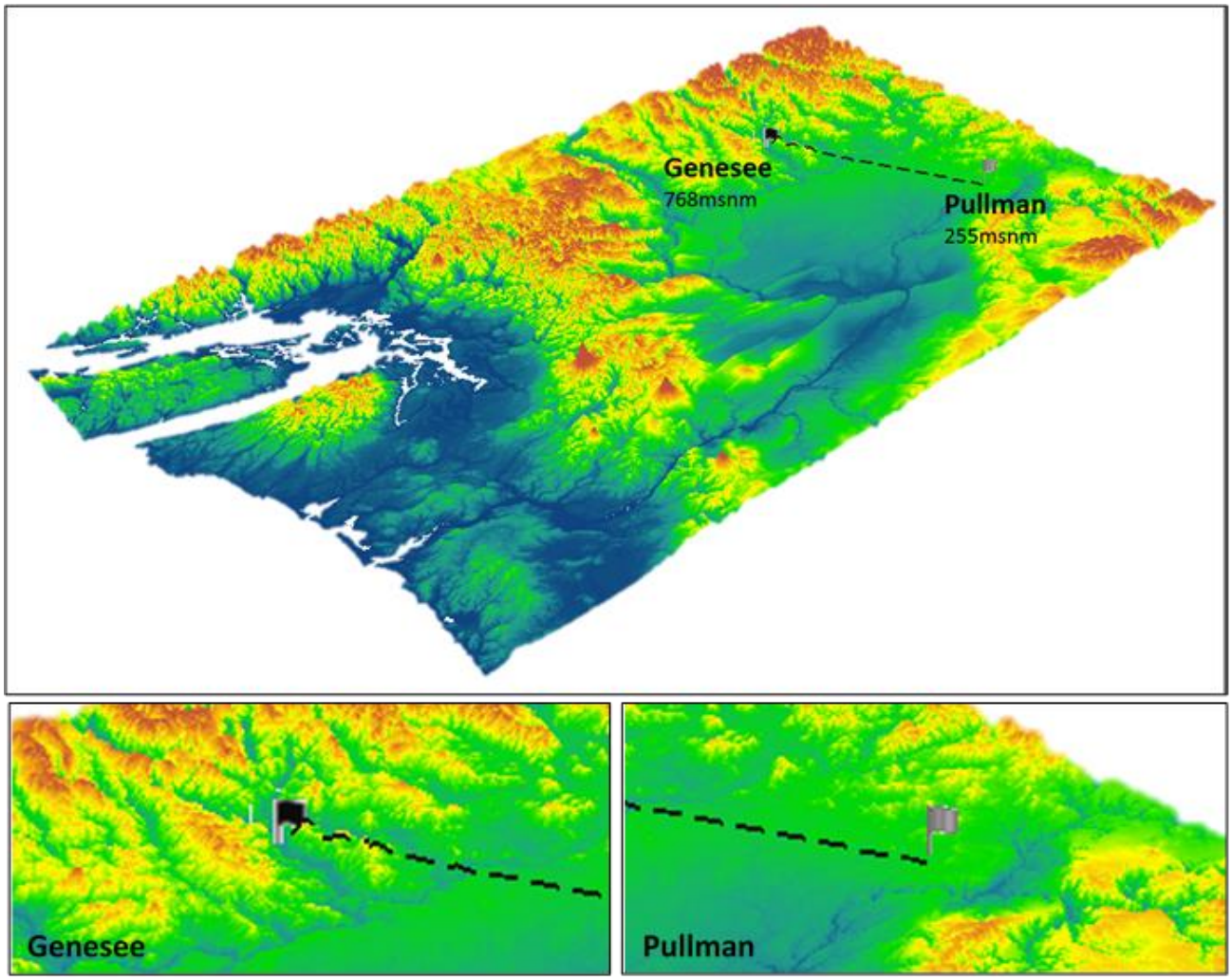
Yao F, Tang Y, Wang P, Zhang J (2015) Estimation of maize yield by using a process-based model and remote sensing data in the Northeast China Plain. **Physics and Chemistry of the Earth** 87: 142-152

Zając et al. (2012) Morphological-developmental reaction and productivity of plants and canopy of semileafl ess pea (*Pisum sativum* L.) after seed vaccination with Rhizobium and foliar micronutrient fertilization. **Journal of Applied Botany and Food Quality** 85: 188 – 197

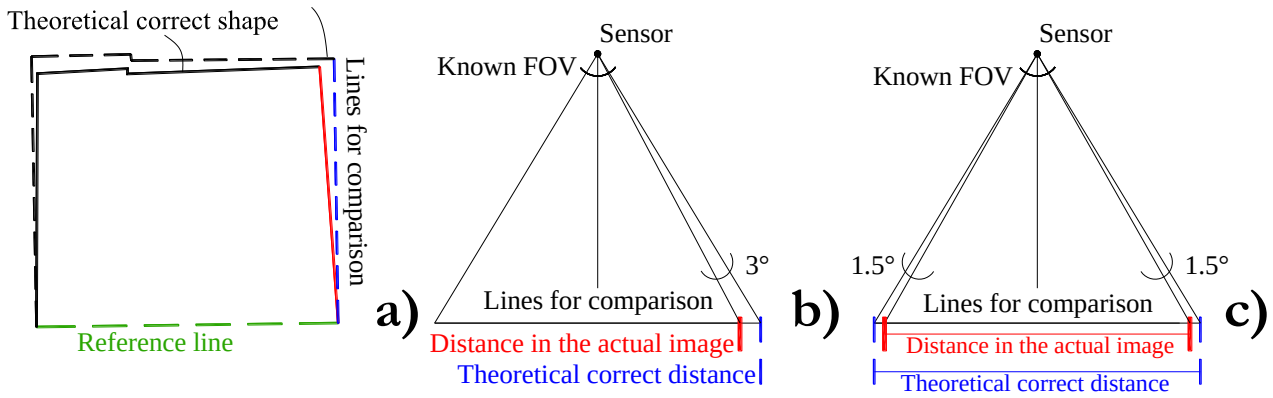
Zhang C, Kovacs M (2012) The application of small unmanned aerial systems for precision agriculture: a review. **Precision Agriculture** 13: 693-712

APPENDICES

APPENDIX A. Comparison of MEAN NDVI values in 20 plots obtained from an orthorectified image, against those obtained from the same image without being rectified.



APPENDIX B. Calculation of angle inclination.



APPENDIX C. Comparison of MEAN NDVI values in 20 plots obtained from an orthorectified image, against those obtained from the same image without being rectified.

MEAN NDVI		% of
With correction	Without correction	error
0.387	0.398	2.80
0.375	0.373	0.50
0.322	0.329	2.20
0.397	0.391	1.50
0.385	0.386	0.30
0.394	0.412	0.46
0.369	0.388	5.10
0.400	0.416	4.00
0.372	0.382	2.70
0.444	0.453	2.00
0.411	0.411	0.00
0.472	0.468	0.80
0.483	0.483	0.00
0.437	0.434	0.70
0.470	0.470	0.00
0.462	0.450	2.60
0.495	0.485	2.00
0.490	0.487	0.60
0.491	0.476	3.10
0.472	0.455	3.60
	Average	2.00
	SD	1.60

2 POTENTIAL OF LOW ALTITUDE MULTISPECTRAL IMAGING FOR IN-FIELD APPLE (*MALUS DOMESTICA*) TREE NURSERY INVENTORY MAPPING

ABSTRACT

Aerial imaging based plant counting is a widely studied topic in remote sensing. Nevertheless, existing methodologies are not applicable for in-field nurseries tree counting due to the complexity in canopy shapes, irregular plant spacing and growth, and diverse textures captured on the images. In this study, a new algorithm has been developed in accordance to the specific requirements of apple nurseries images. Algorithm composed of two key steps, i.e. a raster processing followed by a vector analyses. First step attempt to isolate apple plant pixels using spatial, spectral, and radiometric enhancements. In vector processing, filters based on the size and location of the polygons were applied to isolate the areas resulting from earlier step to represent apple plants. Algorithm was evaluated to estimate number of apple trees in an young nursery imaged with low altitude multispectral imaging system at four altitudes of 10, 25, 40, and 50 m. Multispectral imaging sensor consisted of near-infrared (NIR), green and blue as three bands. For 10- and 25-m images, algorithm performance was evaluated in individual as well as in mosaic images. Low altitude images with ≤ 25 m above ground level were ideally suited for young apple nursery tree count with 5% or less estimation error. Tree count average accuracy was 97% and 95% for 10-m altitude individual and mosaic images, respectively. Similarly, those values were 92% and 88% for 25-m altitude images. Based on the results, images at 40 m are recommended only when the methodology include four extra steps added to the base algorithm to have tree count accuracies of about 88%. Images at 50 m are not recommended in any case due to the low accuracy obtained (~ 75%). Overall, the low altitude multispectral imaging integrated with image processing algorithm developed in this study will aid nursery growers in low-cost and timely tree count needed for inventory mapping and management.

Keywords: Remote sensing; GNDVI; Image filtering; Plant counting

2.1 INTRODUCTION

In Washington State, USA about 60 thousand hectares area is under apple tree fruit production with most of the production sold in fresh market (USDA, 2014). For the ease of orchard mechanization to address pressing labour issues and with goal of being economically viable in competing international fresh fruit market, growers are transitioning from old cultivars

to new apple cultivars that produce better quality fruits. Old tree architectures are also being replaced with modern vertical and y-trellised tree architectures.

Grower needs have translated into increased demand of cultivar specific seedlings (e.g. Honeycrsip) with most local nurseries pre-sold for next few years. Overall, nurseries are generally planted in plastic pots under semi-controlled conditions, where individuals present regular shapes and spacing (Hooijdonk et al., 2015). In the U.S. Pacific Northwest, the panorama changes, with plants growing directly on the ground with somewhat irregular shapes and spacing.

Nursery growers thus use labour-intensive techniques to count every seedling per season, i.e. four times a year, to have cultivar specific inventory. Manual plant counting is time consuming process and accounts for one of the major production expenses. Plant/seedling inventory is also critical for adequate insurance of the seedlings in late fall to avoid economic losses to nursery growers due to winter weather before harvesting the dormant seedlings.

Thus, nursery growers are in need of alternative and rapid inventory mapping techniques. Remotely Piloted Aircrafts (RPAs) is emerging as a versatile technology for range of imaging based agriculture applications. RPA have been employed for crop emergence (Sankaran et al., 2015a; Khot et al., 2016), biotic and abiotic stress scouting (Falkenberg et al., 2007; Sankaran et al., 2015b), irrigation management (Gago et al., 2015) and also for potted nursery tree inventory mapping (Leiva, 2014).

Between the aerial data collection and the generation of inventory maps, are placed the algorithms necessary to convert sensor raw data into information for practical applications (Sankaran et al., 2015b). Some algorithms exist for well-developed trees (Delenne et al., 2010; Recio et al., 2013) and potted nursery trees counting from aerial imaging. However, for nursery trees grown from the ground, new algorithms are needed understanding imagery context, i.e. plant, weed, mulch, irrigation lines, trellis pols, are required. Therefore, key objective of this study was to develop a robust algorithm to process the RPA acquired multi-spectral images for accurately counting number of plants in young 2-year apple nurseries where seedlings are planted directly into the ground at somewhat irregular spacing. Focus was also to evaluate optimal imaging resolution needed to effectively count the nursery plants.

2.2 MATERIAL AND METHODS

2.2.1 Study site and data collection

The field data was collected from an apple tree nursery in second year of the production. Data was collected mid- September under dry/sunny conditions (Figure 9).

Figure 9 - A study site (Lat: 47° 0'29.25"; Long: 119°49'30.29").

Small RPA (OktoXL 6S12, Oktokopter Inc. Germany) integrated with 3-band multispectral imaging sensor (XNiteCanonNDVI, LDP LLC, NJ) was used to acquire the aerial images of study plots. Bands 1, 2 and 3 represents near infrared, green and blue wavelengths in his respective order. RPA can be operated both manually or using an autopilot guided by the GNSS embedded in the hardware with a defined flight plan. In this study, field maps were used for waypoints guided flights. Three-band sensor constitutes NIR, green (G), and blue (B). Sensor has a resolution of 0.35 cm per pixel for images taken at 10-m altitude AGL.

Apple nursery plot of 50 length × 7.5 m width m was imaged at four different flight altitudes of 10, 25, 40 and 50 m AGL. Inside the plot located were five tree lines (rows) spaced at 1.5 m. The accuracy and the standard deviation (SD) of the estimation was calculated for each row inside the images, so that, each crop line represents one repetition. Accuracy was also calculated separately with individual images as well as mosaics for the 10 m and 25 m data altitude scenarios.

Tree spacing was about 0.30 to 0.75 m. In the middle row of study plot, different coloured panels along with reflectance panel (Figure 10) were placed for ease of stitching images from low altitude flights. On the same day, ground-reference data was also collected to have actual count of the apple nursery trees per row.

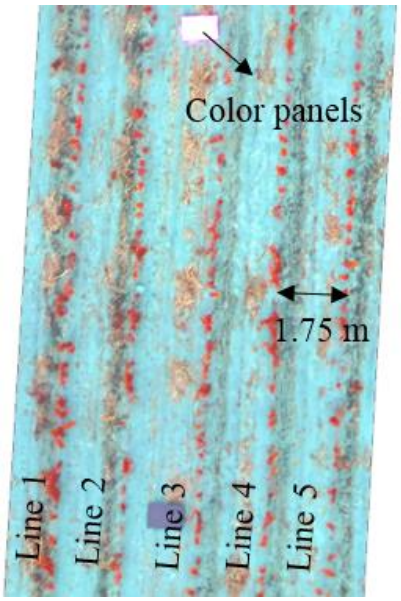


Figure 10 - Spacing of rows and color panels.

2.2.2 Data processing

Image processing algorithm was implemented using ERDAS (2014, Hexagon Geospatial) (for raster analysis) and ArcGIS (10.2, ESRI, Redlands, Ca) (for vector analysis) software. Figure 11 enlist 15 steps performed by the custom developed image processing algorithm, following the order through which they are applied. For 10 m and 25 m images, performance of the algorithm was analysed in individual as well as in mosaic images. Agisoft Photoscan (Agisoft LLC, Russia) was used to generate the mosaics.

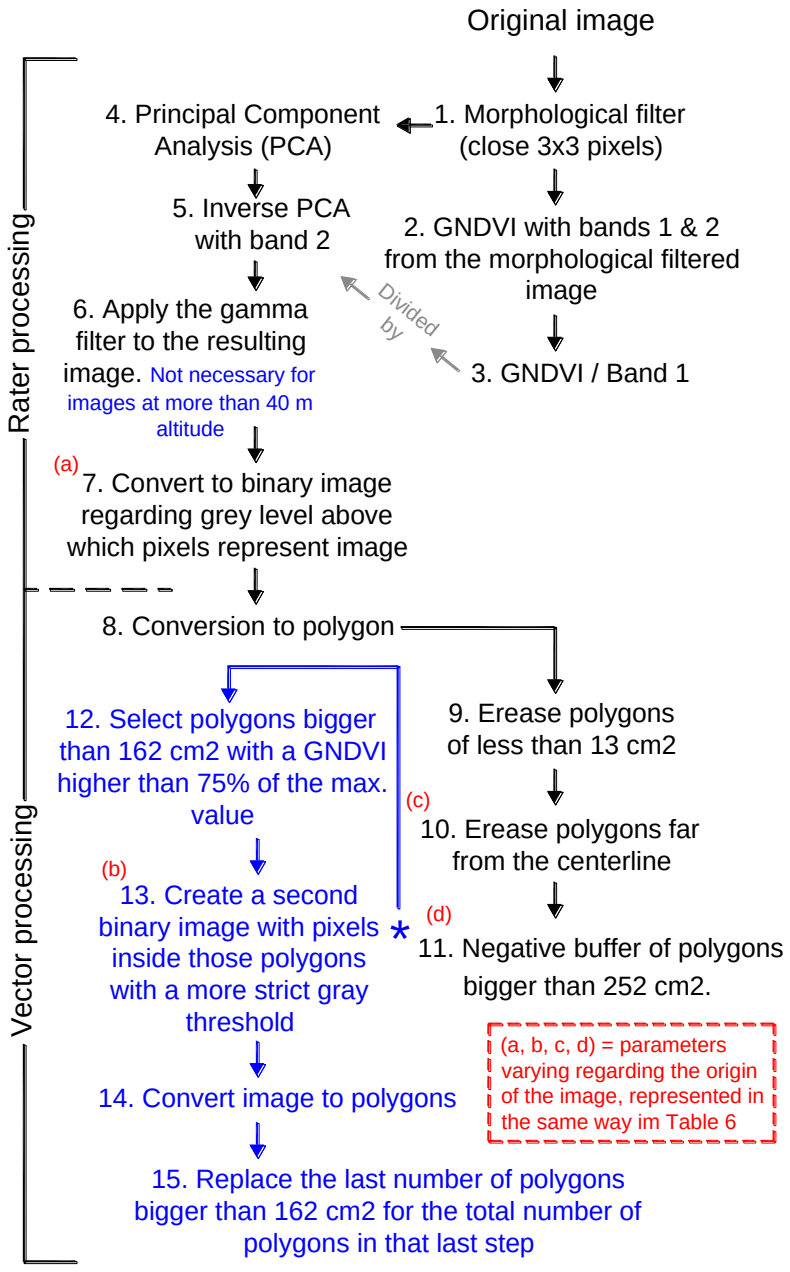


Figure 11 - Apple nursery tree counting algorithm to process RPA based multispectral data.

The nursery tree counting algorithm was thought to be a two-step process as raster and vector processing. The raster processing was done to isolate the apple plant pixels using spatial, spectral, and radiometric enhancements. The vector processing then searched apple plants by applying filters based on the size and location of the polygons resulting in the raster processed data.

In step 1, morphological filter was applied to the original image in submatrices of 3×3 pixels, using the closing function; which tend to smooth sections of contours by filling

holes and blending fine breaks along the boundary of objects (Goyal, 2015). The principle of the morphological closing is to detect the actual shape of an object inside an image in two steps: first enlarging progressively the limits of forefront pixel areas (dilation), and then eroding the boundaries of the dilated regions in that image (Pedrini and Robson, 2008). In the Eq. 9, the expression φB (f) represents the closing operator of an image (f) being first dilated with certain structure element B, and then being eroded with the resulting structure element β (Soille, 1999).

$$\varphi B (f) = \varepsilon\beta \ [\delta \varphi B(f)] \quad \textbf{Eq. 9}$$

Resulting image of the morphological filter was used to generate the Green Normalized Difference Vegetation Index (GNDVI) (step 3). Eq. 10 represents the formula for GNDVI.

$$GNDVI = (Near\ Infrared - Green)/(Near\ Infrared + Green) \quad \textbf{Eq. 10}$$

In step 4, the same morphologically filtered image of step-1 was used to calculate the Principal Component Analyses (PCA) to help in the identification of the patterns of the apple plants in the image, by reducing redundant data and delimiting the image to a portion of the possible 0 – 255 data range. For that, was selected the Principal Component 2 (PC2) to perform the inverse PCA (iPCA) because that PC presented the least inter-band variability (step 5), and so the highest differentiation between apple plants and the rest of the objects.

The result of step 3 was divided by the result of the step 5, and the outcome of that division was filtered using a radiometric adjustment called Gamma Stretch, which is a conversion that utilize a non-linear multiplication generally used in grey-scale image processing, and can be represented as is shown in the Eq. 11 where is described the linear map from P to Ω groups (Shi and Cai, 2011):

$$\phi: P \rightarrow \Omega, \Omega = \{\omega \mid \omega = \phi (X), \phi (X) = \pi x/2xm\} \quad \textbf{Eq. 11}$$

where, P is the digital number (DN), or pixel value, ϕ is the value of angle, x is the grey value and xm the central value in the scale of greys.

Resulting raster was then converted to binary image and pixels were counted as plant above set threshold. Threshold varied with imaged altitude and is reported in Table 6.

Table 6 - Algorithm threshold parameters depending on the imaging altitude.

Altitude (m)	Threshold (pixels)		Buffer from centreline (cm) (c)	Negative buffer in polygons (d)
	Grey 1 (a)	Grey 2 (b)		
10	150	-	20	1
25	160	-	20	3.5
40	110	225	30	7
50	60	126	30	3.5

Notes: a, b, c and d are represented in the same way as in Figure 3.

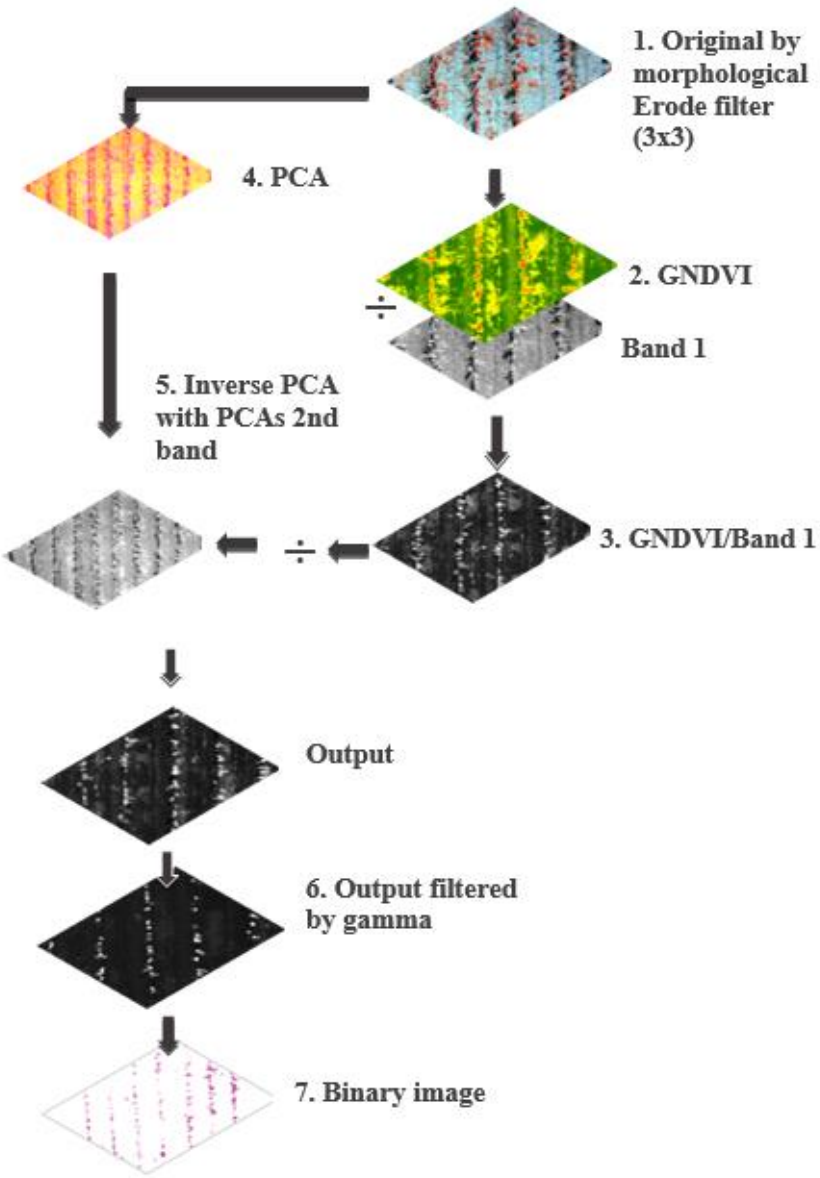


Figure 12 - Graphic representation on steps followed in the raster processing of the algorithm

Moving to the vector processing (Figure 13), the binary image was then converted to polygons and smaller polygons (< 13 cm^2) were deleted (step 9). In step 10, polygons far from the centreline, row line, were deleted. Measure of the centreline buffer depended on the image as well as the value of the negative buffer applied to polygons (> 252 cm^2). The Figure 13 shows graphically each step of the vector processing part of the algorithm.

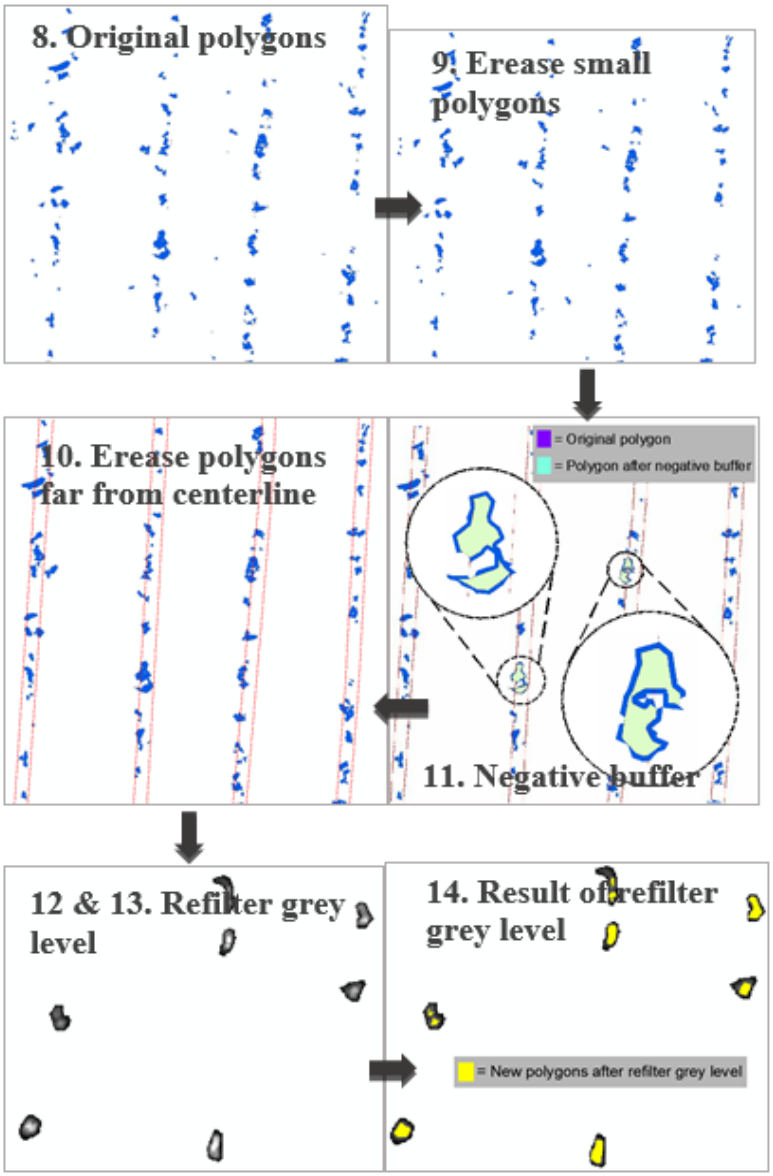


Figure 13 - Graphic representation on steps followed in the vector processing of the algorithm.

The automated algorithm needed four more steps to process images acquired at 40 and 50 m altitude AGL. First of those additional steps was to select polygons bigger than 162 cm^2 presenting a GNDVI higher than the 75% in relation with maximum value presented in

an image. At step 13, the algorithm created a second binary image with the raster resulting in step 12. For that case, a new threshold ('grey 2' in Table 6) is required and it varies for images taken at different altitudes. Finally, the second binary image was converted to polygon (step 14), replacing those previously selected polygons with area > 152 cm^2. The algorithm accuracy (%) towards tree counting from images acquired at different altitudes was then estimated using Eq. 12

$$\% A = | (EV*100)/RV | \quad \textbf{Eq. 12}$$

where, % A = percentage of accuracy in the estimation; EV = estimated value; RV = real value.

2.3 RESULTS

Plant counting accuracies for individual images from each of the altitude are listed in Table 7. Since 40 and 50 m images covered all the area in one image, stitching was only necessary for 10 and 25 m altitude data (Table 8).

Table 7 - Nursery plant counting accuracies for varied attitude individual images processed by customized algorithm.

Alt. (m)	Row	Actual number of plants	Estimated number of plants	% Acc.	% Acc. Average	SD
10	1	31	36	116	97	3
	2	28	26	93		1
	3	30	32	107		1
	4	33	28	85		3
	5	32	27	84		3
25	1	33	31	94	93	1
	2	29	29	100		0
	3	31	34	110		2
	4	32	29	91		2
	5	34	24	71		5
40	1	-	-	-	67	-
	2	121	83	69		19
	3	115	83	72		16
	4	123	85	69		19
	5	126	73	58		27
50	1	-	-	-	56	-
	2	121	75	62		23
	3	115	68	59		24
	4	123	62	50		31
	5	126	64	51		31

Plant counting accuracies for mosaic images for 10 m and 25 m are listed in Table 8.

Table 8 - Nursery plant counting accuracies for mosaic images at 10 m and 25 m, processed by customized algorithm.

Alt. (m)	Row	Actual value	Estimated value	% Acc.	% Acc. Average	SD
10	1	-	-	-	92	-
	2	121	101	83		10
	3	115	112	97		2
	4	123	117	95		3
	5	126	129	102		2
25	1	-	-	-	87	-
	2	121	106	88		8
	3	115	104	90		6
	4	123	103	84		10
	5	126	114	90		6

Employing the algorithm to mosaic images its accuracy still acceptable, but the variation of the data increases; what means that the replication of the results presented cannot be guaranteed as it is the estimations with individual images.

Table 9 presents the results generated after apply the extra steps to 40 and 50 m acquired images.

Table 9 - Nursery plant counting accuracies with modified algorithm.

Alt. (m)	Row	Actual value	Estimated value	% Acc.	% Acc. Aver.	SD
40	1	-	-	-	88	-
	2	121	119	98		1
	3	115	98	85		9
	4	123	99	80		12
	5	126	110	87		8
50	1	-	-	-	63	-
	2	121	79	65		21
	3	115	76	66		20
	4	123	71	58		26
	5	126	73	58		27

By applying those extra steps the performance of the algorithm improved by about 22% for 40 m altitude images whereas it was not possible to significantly improve accuracies in the 50 m images.

2.4 DISCUSSION

Highest accuracy percentage and lower data dispersion was obtained with images acquired at 10 m and 25 m AGL, showed in Table 7, where averages accuracies were 97 % and 95 %, respectively; the lower accuracies and higher variations observed in rows 1, 4 and 5 are related to the distortion in of the images close to its edges; that's the reason why rows 2 and 3, which are in the middle of the image, presented higher accuracies. Bottommost accuracy and greater SD was for 50 m altitude acquired images. High percentage of accuracy at low altitudes are attributable, first, to the adequate isolation of plant pixels due to the correct combination of morphological filter, with GNDVI and iPCA images. Using the iPCA from the PC2 could be highlighted subtle details in the boundaries of apple plants that were

obscured by higher contrast in the original image. Figure 14 shows the example of how the images of PC´s 1, 2 and 3 looks.

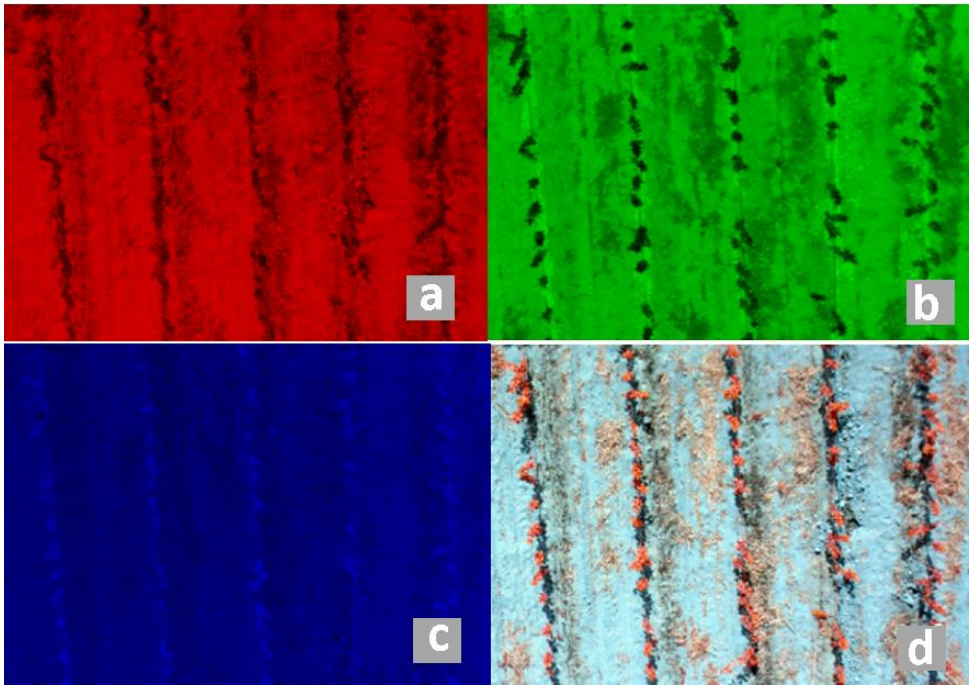


Figure 14 - Images of PC´s 1 (a), 2 (b) and 3 (c); and its respective original view (d).

For 40 m images, the accuracy was 66%, while it was of 55 % for 50 m images. Overall, the higher the flight altitude, lower was the resolution of the images, conducting to the omission of some targets that cannot be identified (Wulder et al., 2001). Therefore, in order to improve the results for higher altitude images, algorithm was modified by adding four more steps (Figure 11).

The extra steps, which results are on Table 9, are focused on the detection of big areas within GNDVI image, which are separated and re-filtered in terms of grey level with expectation to increase the amount of polygons, following the principle that “as bigger and greener the polygon, higher must be the quantity of plants inside”.

Applying each technique separately, or in a different order as the presented here, may not provide the same results. Gamma stretch constitute a critical step as this filter, basically “characterizes the reproduction of tone scale” (Poynton, 1998). Such gamma stretch leads to finely differentiated pixels of each of the apple plants through radiometric adjustment of the intensity projected at each pixel. Thus, gamma stretch would have made the pixel brighter (apple pixels) or opaquer (soil, straw or rock pixels).

Overall, based on the results of the study, we recommend to use 25 m or lower altitude images to have 95% or higher plant count, which is typical error in manual ground count. For general estimations of larger areas one can possibly use images at 25 m AGL but the plant count error might be > 5%. Images at 40 m can be used only in cases where larger areas need to be analysed in a short time, and quick advance in operations is required; with these images the user can expect an error higher than 10%. For the algorithm presented here, we do not recommend images at 50 m in any occasion.

Based on the fact that "the context modulates relationships between objects in an image" (Fazan, 2007), the key limitations in obtaining highest accuracies at any altitude could be: (1) the wide variety of plant shapes; (2) the heterogeneity of textures: soil, straw, apple plants, weeds, rocs, and the combinations; and (3) the overlapping of branches and leaves between neighbour plants (Figure 15).

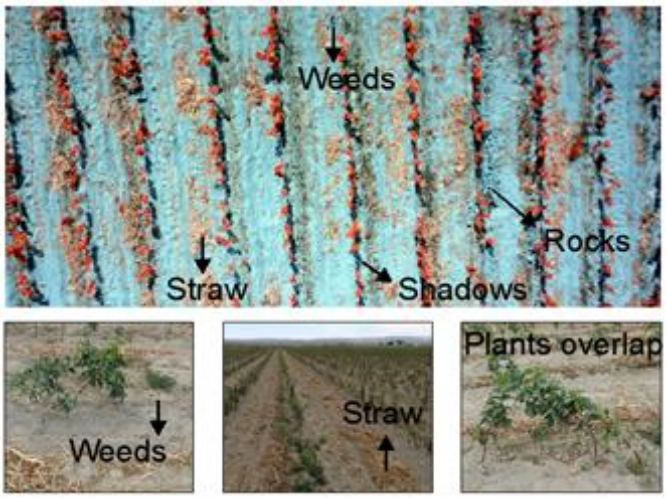


Figure 15 - Examples of factors limiting the improvements in the aerial imaging based plant counting accuracies.

Similar to this study, other pertinent studies also reported similar issues with developed algorithms for similar proposes. Some example include segmentation for crowns delineation (Culvenor, 2002; Jing et al., 2012; Kanda et al., 2004), edge detection (Adnan and Yang, 2005), and other kinds of kernel functions (Dralle and Rudeno, 1996), as well as those that employ simple discrimination options to classify pixels (Nunes et al., 2014). Another limiting factor was the irregular spacing between plants, which blocked the use of the prediction of plant location to facilitate the detection of objects (Barbosa et al., 2015.)

Future steps to improve the performance of the algorithms include the use of elevation data as an extra step inside the process, attempting to approach the height differential between the plants and the other objects captured in the image. Some researchers have shown interesting results for objects detection from 3D surfaces, but all of them worked with either with high density of images or very low altitude data (< 5 m) (Bending et al., 2013; Jay et al., 2015; Uysal et al., 2015). Also, most of above studies have been with big size and well-spaced targets, where it is easier to differentiate objects (Strîmbu and Strîmbu, 2015; Parent et al., 2015).

2.5 CONCLUSIONS

For accurate plant counting using RPA based imaging and the algorithm developed in this study, use of low altitude images in the order of 10 to 25 m is recommended. Images at 40 m are a good option only in cases where larger areas need to be analysed in a short time, and quick advance in operations is required. Further improvements of the algorithm will include the evaluation of the use of elevation data as one more attribute and integration of artificial neural network functions in estimating the tree count.

REFERENCES

Adnan O, Yang C (2005) Detecting citrus tree size, centroid, and health conditions from airborne high-spatial-resolution multispectral imagery. **Geospatial Goes Global: From Your Neighborhood to the Whole Planet** 5 pp

Barbosa A, et al. (2015) Utilização de veículo aéreo não tripulado na detecção de falhas de plantio em um canavial. **XLIV Congresso Brasileiro de Engenharia Agrícola** 9 pp

Culvenor D (2002) TIDA: an algorithm for the delineation of tree crowns in high spatial resolution remotely sensed imagery. **Computers & Geosciences** 28: 33-44

Delenne C, Durrieua S, Rabatel G, Deshayes M (2010) From pixel to vine parcel: A complete methodology for vineyard delineation and characterization using remote-sensing data. **Computer and Electronics in Agriculture** 70: 78-83

Dralle K, Rudemo M (1996) Stem number estimation by kernel smoothing of aerial photos. **Canadian Journal of Forest** 26: 1228-1236

Fazan A (2007) Predição de sombras de edifícios sobre vias urbanas com base em um modelo digital de elevações e dados de imagens aéreas de alta-resolução. **Faculdade de Ciências e Tecnologia / UNESP** 102 pp

Falkenberg N, et al. (2007) Remote sensing of biotic and abiotic stress for irrigation management of cotton. **Agricultural Water Management** 87: 23-31

Gago J, et al. (2015) RPAs challenge to assess water stress for sustainable agriculture. **Agricultural Water Management** 153: 9-19

Goyal M (2011) Morphological Image Processing. **International Journal of Computer Science & Technology** 2: 161-165

Hooijdonk B, Tustin D, Dayatilake D, Oliver M (2015) Nursery tree design modifies annual dry matter production of newly grafted 'Royal Gala' apple trees. **Scientia Horticulturae** 199: 404-410

Jing L, Hu B, Noland T, Li J (2012) An individual tree crown delineation method based on multi-scale segmentation of imagery. **ISPRS Journal of Photogrammetry and Remote Sensing** 70: 88-98

Kanda F, Kubo M, Muramoto K (2004) Watershed segmentation and classification of tree species using high resolution forest imagery. **Proceedings 2004 IEEE International Geoscience and Remote Sensing Simposium** 6: 3822-3825

Khot L, et al. (2016) RPA imaging-based decision tools for arid winter wheat and irrigated potato production management. **International Journal of Remote Sensing** 37: 125-137

Leiva J (2014) Use of Remote Imagery and Object-based Image Methods to Count Plants in an Open-field Container Nursery. **University of Arkansa**n 123 pp

Nunes G, Vieira D, Carvalho S (2014) Preliminated data obtained by RPA into clonal forests stands of Eucalyptus Urograndis H13 in the state of Matto Grosso. **XI Seminário de Atualização em sensoriamento Remoto e Sistemas de Informação Geográfica Aplicados à Engenharia Florestal** 211-218

Pedrini H, Robson W (2008) Análise de Imagens Digitais - Princípios, Algoritmos e Aplicações. **Thomson** 528 pp

Poynton C (1998) The rehabilitation of gamma. **Proceedings of SPIE** 3299: 232-249

Recio J, Hermosilla T, Ruiz L, Palomar J (2013) Automated extraction of tree and plot-based parameters in citrus orchards from aerial images. **Computers and Electronics in Agriculture** 90: 24-34

Sankaran S, Khot L, Carter A (2015) (a) Field-based crop phenotyping: Multispectral aerial imaging for evaluation of winter wheat emergence and spring stand. **Computers and Electronics in Agriculture** 118: 372-379

Sankaran S, et al (2015) (b) Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review. **European Journal of Agronomy** 70: 112-123

Shi J, Cai Y (2011) A Novel Image Enhancement Method Using Local Gamma Correction with Three-level Thresholding. **IEEE** 978: 4244-8625

Soille P (1999) Morphological Image Analysis: Principles and Applications. **Springer** 316 pp

USDA NASS. (2014) Washington Annual Statistical Bulletin. Available online at: https://www.nass.usda.gov/Statistics_by_State/Washington/Publications/Annual_Statistical_Bulletin/2014/bulletin_wa_2014.pdf (accessed 08 February 2016).