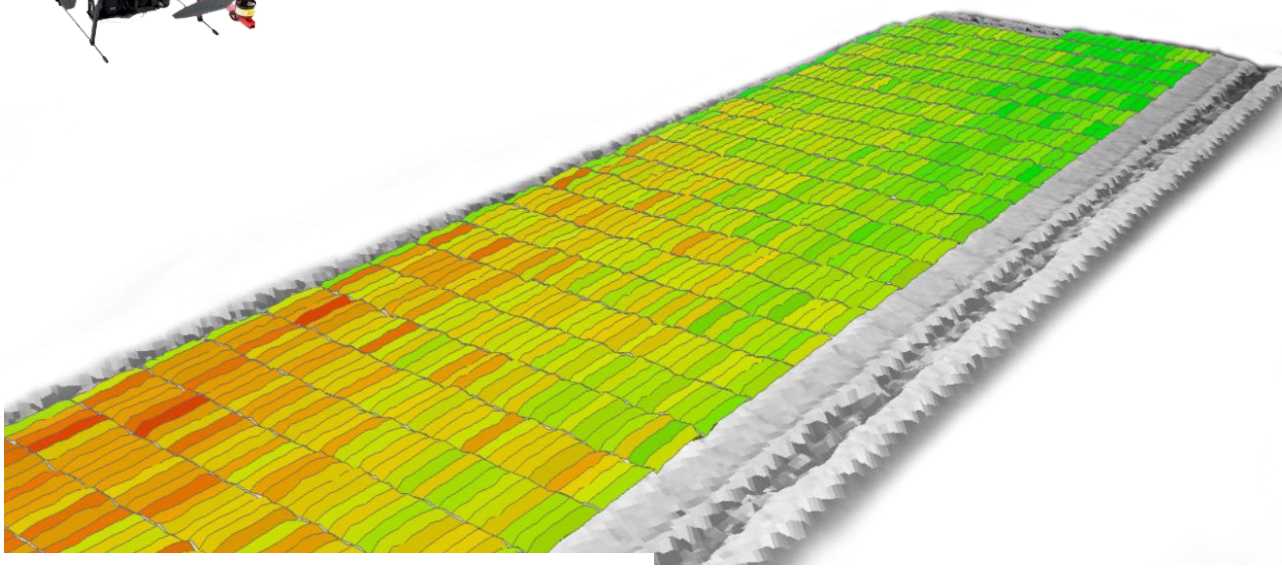
Laboratory of Geo-Information Science and Remote Sensing

Thesis Report GIRS-2016-06

Exploring the usability of Unmanned Aerial Vehicles for non-destructive phenotyping of small-scale maize breeding trials

Darell van der Voort



01-04-2016

WAGENINGEN UNIVERSITY
WAGENINGEN UR

Exploring the usability of Unmanned Aerial Vehicles for non-destructive phenotyping of small-scale maize breeding trials

Darell van der Voort

Registration number: 881113904060

Supervisors:

Dr. Ir. Lammert Kooistra
Dr. Ir. Sander Mücher

A thesis submitted in partial fulfilment of the degree of Master of Science
at Wageningen University and Research Centre,
The Netherlands.

01-04-2016
Wageningen, The Netherlands

Thesis code number: GRS-80436
Thesis Report: GIRS-2016 -06
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing

FOREWORD

This report is the result of seven months of research into the usability of the UAV platform for small-scale maize breeding trials. My combined passion for UAV technology and remote sensing led me to this thesis topic. I thoroughly enjoyed working with these innovative technologies and to explore their capabilities during my thesis. I found the fact that this was a collaborative project between Wageningen UR, Alterra Research Institute and Limagrain BV an added value. Due to this collaboration this research was viewed with three different backgrounds, which made for interesting new and exciting ideas, but also ensured that compromises had to be made. I learned a great deal on UAV flight operations, sensors, regulations, software and processing of data. I see the UAV platform as a very promising technology with legions of applications now, and in the near future. I hope that in my future career I will be able to further improve and work within the field of (UAV) remote sensing and Earth observation.

I would like to thank my supervisor Lammert Kooistra from Wageningen UR for his always enthusiastic guidance, ideas and extensive feedback on this research. My supervisor Sander Mücher from Alterra Research Institute, for his suggestions, feedback, visits to Limagrain and optimism, even when Murphy's Law hit. Furthermore I would like to thank Eva Ampe from Limagrain for her skype session, very enthusiastic input for new ideas and help in R. Also Piet Reijns from Limagrain for his feedback during my presentation. Furthermore from Wageningen UR I would like to thank Henk Kramer for his time to help me with software and reserving the processing PC for me. Also Juha Suomalainen for answering my questions and assistance with some of the pre-processing. Lastly a big thanks to my friends at the University for a great year and half of studying!

ABSTRACT

Quantitative analysis of plant phenotypes in field trials has become a major bottleneck due to the large amount of breeding trials that need to be scored and analyzed. This phenotyping is traditionally done by visual scoring and manual measurements, which is labor intensive, non-systematic and susceptible to human error. Plant breeding would therefore benefit from fast, high-throughput and non-destructive sensing techniques to score and evaluate field-based breeding trials. In controlled circumstances like greenhouses significant progress has been made using image based techniques in high-throughput phenotyping facilities. The next step would be to adopt camera based technology in field experiments, however, a multitude of factors are influencing these observations: light conditions, wind, moisture on leaves, etc. Both ground based vehicles and aerial platforms have been considered for field-based phenomics. Aerial platforms like Unmanned Aerial Vehicles (UAVs) have much potential to be a suitable tool for plant breeders. They can be equipped with multiple sensors and fly at low-altitude in a relative small timespan without affecting field conditions.

In this research we evaluated the usability and accuracy of high-resolution hyperspectral images acquired from an Unmanned Aerial Vehicle (UAV) to quantify the growth parameters biomass and plant height for a large-scale maize breeding trail. Furthermore, this research tested which spectral range and growing stages gave the best relation with biomass and plant height in order to find out which sensors were most suited and what flying periods were essential for phenotyping these traits. The field under research consisted of a phenotyping experiment with 3838 plots with a size of 10.2 m^2 each and planted with different maize (*Zea Mays*) cultivars. Conventional measurements for biomass and height of each cultivar were done at harvest. At four different growth stages, flights with a multi-rotor UAV platform equipped with a combined RGB and hyperspectral camera we carried out, resulting in three geo-rectified products: RGB-orthomosaic, digital surface model and a hyperspectral dataset with 100 bands ranging from 450-950 nm. In the subsequent method development, both individual spatial images and also the fusion (including time) was evaluated using vegetation indices and multivariate statistical approaches.

For the spatial estimation of height at the end of the growing season, a good accuracy could be achieved when correlating UAV derived heights to LiDAR measurements ($R^2 = 0.78$). Accuracy over the whole field differed due to geometrical errors and influence of ground control points ($R^2 = 0.61$). Overall there was a systematic underestimation of -0.30 meters. For the estimation of final biomass, the prediction accuracy was $R^2 = 0.79$ when the spatial layer of the canopy plant model was combined with spectral information from specific parts of the growing season in a multivariate approach. The multispectral indices showed a better performance in estimating biomass than the hyperspectral indices. The beginning and end of the growing season were most important for biomass prediction. The main limitations of this research were the geometrical distortions caused by the hyperspectral push broom scanner and non-linear deformations of the digital surface models due to the lack of sufficient amount of ground control points. The results of this study have shown that crop surface models derived from the UAV based RGB sensor are capable of capturing height variability with high spatial detail. Furthermore, our results suggest that plant height measured at harvest, combined with vegetation indices at the start and end of the growing season may be used to predict maize biomass. Future research should explore the use of more ground control points for improved rectifications of 3D models and the use of multi- or hyperspectral frame cameras in order to minimize geometrical distortions.

Keywords: Unmanned Aerial Vehicle, Remote Sensing, Phenotyping, Breeding Trials, Maize, Structure from Motion, Hyperspectral, Plant Height, Biomass, Multivariate Analysis.

LIST OF ABBREVIATIONS

AHN2	Actueel Hoogtebestand Nederland
CSM	Crop Surface Model
CSMU	Crop Surface Model based on UAV
CSMA	Crop Surface Model based on AHN2
DEM	Digital Elevation Model
DSM	Digital Surface Model
DTM	Digital Terrain Model
GCP	Ground Control Point
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
HYMSY	Hyperspectral Mapping System
LiDAR	Light Detection and Ranging
MTVI2	Modified Triangular Vegetation Index
NDVI	Normalized Difference Vegetation Index
OSAVI	Optimized Soil Adjusted Vegetation Index
RTK	Real Time Kinetic
SfM	Structure from Motion
TOC	Top Of Canopy
UAV	Unmanned Aerial Vehicle
VI	Vegetation Index
WDVI	Weighted Difference Vegetation Index

TABLE OF CONTENTS

1 INTRODUCTION

1.1 CONTEXT AND BACKGROUND

With current predictions on human population growth, a world's population of 8.8 to 10 billion by mid-century is unavoidable (Cleland 2013). To secure a quality of life for current and future generations, sufficient food production is essential. Worldwide there is evidence that food production and the distribution is problematic, leading to 3.7 billion human beings that are malnourished (Pimentel and Pimentel 2006). As diets are not likely to change, much of the future need for food will be driven by increased quantity (Cleland 2013). Crops like maize, wheat and grain are essential for food security as these crops provide the resource for both food and feed for the growing human and animal population. Furthermore plants are becoming a form of renewable energy in the form of biofuels (Grayson 2011) and are increasingly being used as a raw material for products (Boehlje and Broring 2010). Improving plant productivity is therefore a key factor in order to address these major economic, ecological and societal challenges. A way of improving plant productivity is by cross-breeding and refining plant traits in order to produce highly efficient seeds that are adapted for different needs and environmental conditions (Ramming and Fear 1993). Plant breeding can be defined as 'the art and science of changing genetic architecture of plants for the benefit of mankind' (Panguluri and Kumar 2013). It has been a practice since the beginning of agriculture, although the more scientific approaches started to emerge around 1900 (Panguluri and Kumar 2013). Plant breeding companies nowadays are in the forefront of this scientific plant research, introducing traits and creating new plant varieties with high yield potentials.

1.2 PROBLEM STATEMENT

To find new traits with high yield potentials thousands of breeding plots need to be scored and analyzed. This phenotyping is traditionally done by visual scoring and manual measurements, which is labor intensive, not systematic and susceptible to human error. A phenotype is 'any measurable characteristic or trait of a plant and is a result of combination of genes expressing in the plant (referred to as genotype), environmental influence, and their interactions' (Panguluri and Kumar 2013). Selecting and rejecting certain plant traits is based on phenotyping data and thus of great importance for a plant breeder. 'The quantitative analysis of these plant phenotypes has become the major bottleneck due to the large amount of breeding trials that need to be scored and analyzed' (EPPN 2013). Currently this process is destructive, time consuming and expensive. Plant breeding would therefore benefit from fast, high-throughput and non-destructive sensing techniques to score and evaluate breeding trials. Hence, new state-of-the-art technologies must be developed and studied to accelerate breeding through phenomics (Tester and Langridge 2010). Developments are made in the field of non-destructive use of sensors for phenotyping. Although these sensors are already widely used in controlled environments like greenhouses, still a wider range of non-destructive tools are necessary for the automation of field condition phenotyping (Pieruschka and Poorter 2012). A next step would be to adopt this camera based technology in phenotyping field experiments. A multitude of factors however would influence these field based observations: atmospheric conditions, wind, moisture on leaves, etc. Both ground based vehicles and aerial platforms have been considered for field-based phenomics. Aerial platforms like Unmanned Aerial Vehicles (UAVs) have much potential to be a suitable tool for plant breeders. They can be equipped with multiple sensors and fly at low-altitude in a relative small timespan without

affecting field conditions. High resolution imagery from RGB cameras can be used in Structure from Motion (SfM) computer vision algorithms to derive canopy architecture traits. Furthermore, UAVs make it possible to fuse these architectural traits with hyperspectral data to characterize biochemical traits. To improve and develop the UAV platform into a suitable tool for plant breeders, new methods are required to fully utilize the capabilities of these platforms with multiple sensors in large scale field based phenomics. Additional progress is required on experimental protocols (spatial, temporal, spectral settings) which: 1) optimize the required signal to quantify the relevant trait(s); and 2) which reduces the effect of disturbing factors resulting in robust and reproducible retrieval of trait value including the associated uncertainty.

1.3 RESEARCH OBJECTIVES & QUESTIONS

The objective of this research is to evaluate the usability and accuracy of using UAV based optical sensors to derive products for the phenotyping of maize breeding trials, with a main focus on biomass and plant height estimation. Furthermore, this research aims to test which spectral range and growing stage gives the best correlation to biomass in order to find out which sensors are most suited and what flying times are essential for phenotyping these traits. The project is in collaboration with research institute Alterra and plant breeding company Limagrain BV.

1.4 RESEARCH QUESTIONS

The main question of this thesis is '*Are UAV platforms a suitable tool for phenotyping of maize breeding trials and what growing stage gives the best results?'* The main question will be subdivided into three sub-questions:

- *What is the accuracy of estimated plant height using UAV derived digital crop models?*
- *Can the combination of specific vegetation indices measured over time and plant height provide better estimates of above ground biomass for maize breeding trials?*
- *Which growing stages of maize gives the best relation between RGB, hyperspectral and multispectral band vegetation indices and above ground biomass?*

1.5 THESIS OUTLINE

We will first start with a literature review based on the problem description and research questions in chapter 2. We will then describe the study area, set up of plots and data collection procedures in chapter 3. Subsequently, in chapter 4 the methods will be described. Chapter 5 will show all the results based on the methodology and in chapter 6 we will discuss these results. In chapter 7 the conclusion and outlook will be given.

2 LITERATURE REVIEW

2.1 IMPORTANT MAIZE TRAITS

To address the issue of an ever growing world population in combination with a changing climate, we need crops that have high-yielding genotypes and that are adapted to our future climate. These new crops need to have high fundamental and stable yield characteristics, even under abiotic stresses. The genomics revolution and gene technology are believed to provide solutions to these breeding challenges and it is predicted that this technology will become more common in the near future (Furbank and Tester 2011). This will lead to an increase of high-performing genotypes in the agricultural germplasm. To truthfully test the performance of these new genotypes they have to be assessed under field conditions using phenotyping methods, or 'phenomics'.

Phenotyping is a fundamental tool in many industries. Examples are the medical, biology and plant breeding industries. In the plant breeding industry, phenotyping stands for characterizing the performance of the plants for desired trait(s) (Panguluri and Kumar 2013). Fast and high-throughput crop phenomics have increasingly become more discussed as an approach to significantly improve phenotyping efforts for plant breeding (Sankaran et al. 2015). There are numerous traits that can be used for analyzing and scoring trials, such as plant height, stem diameter, vigor, greenness, leaf angle etc. The main traits for maize are; vigor, height, yield, quality and greenness (Figure 1). One of the main interests for a breeder is the estimation of above ground plant biomass at harvest and plant height, as this is an indication for yield potential. It is a widely accepted fact that biomass estimation is a good tool for yield prediction of grain crops like maize (Oerke et al. 2010).

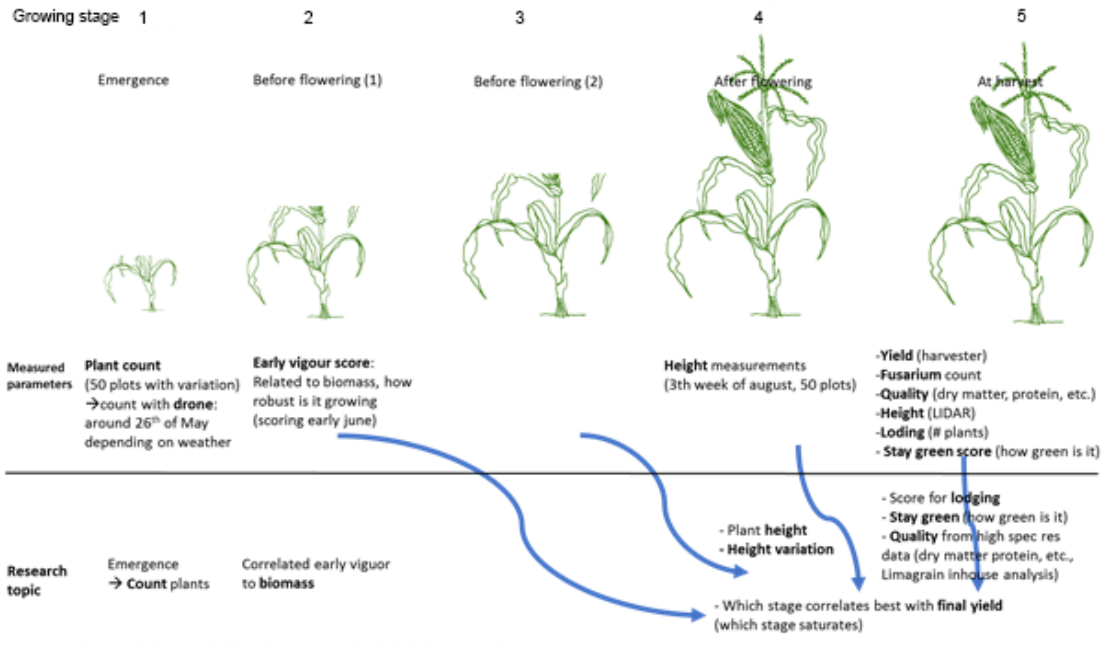


Figure 1 Important maize traits and growing stages (source: Limagrain)

2.2 RELATIONSHIP BETWEEN IMPORTANT MAIZE TRAITS AND YIELD

The relationship between maize yield and plant height has not yet been documented very well. Yin et al. (2011) has shown that the development of plant height of maize is a very good indicator of the

expected yield. He found significant correlations between yield and growing stage 3, 4 and 5 (Figure 1), with the strongest correlations being at growing stage 4 and 5 ($R^2 = 0.54$, $R^2 = 0.81$), although R^2 values differed at different N treatments. Nonetheless Yin et al. (2011) has proven a significant positive correlations between the height of the maize plant at later growing stages and the final yield. Height measurements where in this case done manually with a measuring stick. Similar researches have used Light Detection and Ranging (LiDAR) techniques to indicate the relationship between plant height and yield of other crops (Oilseed rape, Winter rye, Winter wheat and Grassland) and also found strong correlations (Ehlert, Adamek, and Horn 2009; L. Zhang and Grift 2012). The use of LiDAR in dynamic mode, e.g. driven past field edges, proofed to be an accurate measurement method for measuring the height of *Miscanthus giganteus*, a grass with a spikey canopy surface in the same height range as maize (2.70 meters on average). Here an average error of 3,8% was achieved (L. Zhang and Grift 2012). Although these methods proofed to be suitable tools for predicting crop yields, manual height measurements are extremely labor intensive, not objective and susceptible to human error. While LiDAR techniques are more objective, they nowadays mostly require machines driving around a field and would be restricted to measurements at harvest when mounted on a combine. Another way to estimate yield is by remote sensing technology. Vegetation indices calculated from spectral data have been positively correlated to above ground biomass using UAV, aircraft and satellite based imagery (Kross et al. 2015; Jr, Hunt, and Al 1999; Jin et al. 2014). However, these studies did not look at correlation between biomass and VIs during different growing stages (Figure 1). Currently above ground biomass can only be accurately measured at harvest, which is a destructive process and does not provide yield information over the growing season. Phenotyping these traits under field conditions could thus be improved to faster, more accurate non-destructive methods. By combining novel technologies like remote sensing, spectroscopy, image analysis and high-performance computing new field-based phenomics can be developed that address this 'phenotyping bottleneck' (Furbank and Tester 2011).

2.3 PHENOTYPING PLATFORMS

There are both ground based as well as airborne solutions for phenomics under field conditions (White et al. 2012; Andrade-Sanchez et al. 2014; Chapman et al. 2014). These two different platforms have their own pros and cons when it comes to phenotyping under field conditions. Plant breeding trials fields are often large and contain thousands of small-scale plots. Manually measuring ground based field conditions is extremely time consuming and is even more difficult when the measurements are time-sensitive. Atmospheric conditions for example can rapidly change and influence time consuming measurements. Ground based phenomics platforms are only able to measure a small surface area from one plot at a time, while airborne platforms are able to capture hundreds of plots in one image, effectively reducing the effect of changing atmospheric conditions. However ground-based platforms have a huge benefit in accurate Global Navigation Satellite System (GNSS) positioning as they are able to make use of a Real Time Kinematic (RTK) GNSS system and are much more stable in time and space (Andrade-Sanchez et al. 2014). Accurately georeferencing airborne-based imagery proves to be much more of a challenge, although automated methods for georeferencing ultrahigh resolution imagery are being developed (Turner, Lucieer, and Wallace 2014). As there are many crop types with different heights and the set-up of plots can also differ greatly in size and row-spacing, there exists no 'one size fits all' solution for ground-based platforms (Sankaran et al. 2015). Additionally having ground-based platforms drive around the plots can cause soil compactness and potentially influence plant growth. Airborne platforms like Unmanned Aerial Vehicles (UAVs) have the potential to be this 'one size fits

all' tool as they have the capability of capturing ultrahigh centimeter resolution imagery at low-altitude in a relative small timespan without affecting field conditions.

Whereas the more traditional satellite remote sensing often lacks sufficient spatial and temporal resolutions, especially for small breeding trial plots, UAVs provide full control over spatial, temporal and spectral resolutions. One is not limited to satellite revisiting time and flying can be done when needed. Furthermore, different types of sensors can be mounted under the UAV, making it possible to build custom remote sensing tools, especially designed for scoring and evaluating breeding trials. Such aerial monitoring with designated sensors provides detailed information about the spatial variability, which is often invisible to ground observers. Furthermore, hyperspectral sensors provide highly detailed spectral data which is useful for analyzing e.g. chlorophyll, plant quality, and yield. Next to the hyperspectral sensor a high resolution RGB camera can simultaneously be mounted on a UAV, providing detailed imagery as input for photogrammetric Structure from Motion (SfM) computer vision algorithms (Dellaert et al. 2000). This input is used for constructing Digital Surface Models (DSM), which can be used for estimating plant height. Aerial remote sensing techniques have long been used for estimating above ground biomass (Jin et al. 2014; Kross et al. 2015) by calculating vegetation indices (VIs) in the near infrared (NIR) region (Qi et al. 1994). Stereovision techniques have been used for constructing plant 3D models that have the ability to derive plant height (Jay et al. 2015). Studies have shown a 96% accuracy in using stereovision in a controlled environment for deriving plant height (Lati, Filin, and Eizenberg 2013b; Lati, Filin, and Eizenberg 2013a). As UAVs are becoming more accessible and Structure from Motion computer algorithms for building 3D models using just one camera are increasingly getting better, these tools have a huge potential for field-based height and biomass phenomics.

This state-of-the-art unmanned aerial vehicle technology is already been widely used in agricultural practices (Lelong 2008; C. Zhang and Kovacs 2012; Díaz-Varela et al. 2015) and is being tested for various phenotyping applications such as plant count, vigor score and the characterization of yield potential (Sankaran et al. 2015). There are several alternatives for environmental and agricultural practices, namely: a balloon construction (Jensen et al. 2007), an airship without an internal structural framework or a keel (Inoue, Morinaga, and Tomita 2000) or model airplanes and helicopters (Hunt et al. 2005; Sugiura, Noguchi, and Ishii 2005). Though, due to the ease of flying by GPS autopilot options and stable cameras, UAVs are becoming the first choice for these practices (Pudelko, Stuczynski, and Borzecka-Walker 2015).

2.4 Deriving canopy architecture traits using UAVs and computer vision algorithms

SfM techniques can be carried out with any RGB camera an does not need any prior calibration, making it a convenient tool for in-field phenotyping (Jay et al. 2015). This technique has been used for the 3D modeling of plants (Santos and Oliveira 2012), yet testing this technique for field conditions has not much been explored. Due to bi-directional effects and mismatching in ortho-rectification these digital surface models are not always accurate. It is therefore interesting to test the accuracy of these UAV based crop surface models to manual and LiDAR measurements. Díaz-Varela et al. (2015) has used very high resolution imagery acquired from a fixed wing UAV using a consumer-grade camera to construct automatic 3D reconstructions of olive trees. They obtained significant correlations between RTK-GNSS height measurements and heights derived from the 5 centimeter resolution 3D surface model ($R^2 = 0.83$ and RMSE = 35 cm). However, the olive trees were solitary and had a relative large and heterogeneous canopy surface. Therefore crop trees are easier to accurately measure with SfM techniques than smaller,

denser crops like maize. Nevertheless, they showed that a decrease in spatial resolution resulted in a loss of within-crown heterogeneity. A lower spatial resolution of the DSM lead to a loss of detail and had implications for the retrieval of tree height and the errors obtained (R^2 = 0.5 when using 50cm resolution (Díaz-Varela et al. 2015)). Maize has a spikey canopy surface which is difficult to capture due to small area size of canopy and wind effects. This shows that estimating plant height for small-scale plots, centimeter spatial resolutions are needed in order to capture the maize crop phenology. At a coarser resolution parts of one or more plants and their shadows will merge into one pixel signal, losing information on the height variability within a plot. Furthermore Díaz-Varela et al. (2015) did not use ground control points (GCPs), leading to canopy height quantification that was conducted on a relative basis. To improve the absolute accuracy of the DSMs ground control points should have been used. The advantage of this method is that it was fully automated and no data collection in the field was needed. However for accurately estimating plant height for thousands of small-scale maize breeding trials accurate results in absolute means are needed. Therefore the use of ground control points is advised. Lastly Díaz-Varela et al. (2015) focused on open canopies and used visible ground points for height differencing. When looking at the set-up and row spacing of maize fields there is not much open canopy, especially at harvest stage (Figure 1). In order to differentiate between ground pixels and canopy pixels in dense maize fields a digital terrain model (DTM) before harvesting and a digital surface model (DSM) of the crop's canopy is needed. Grenzdorffer (2014) showed that it also possible to estimate the canopy height by statistical analysis of the derived 3D point cloud coming from the SfM algorithms. By classifying points into 'ground' and 'vegetation' points differencing between the heights of these points can be applied to calculate the relative plant height. Although, this method proofed to be less accurate especially with higher leaf area index (LAI) values as there were less reliable ground points. This is especially true for dense crops like maize. The differencing method where the canopy height is determined by the difference between a DTM and a DSM showed to be the fastest and most accurate method. Grenzdorffer (2014) furthermore showed that taking the 95% or 99% percentile of the DCM makes sure that single tall plants do not determine the height of a grid cell. He applied the differencing method to two crops with a homogenous canopy surface (winter wheat and oilseed rape) and to maize, which has a spikier canopy surface. For maize the height results derived from the differencing method deviated most from manual measurements in comparison to the more homogenous crops. Maize height was always underestimated using this method. Some of the underestimation could be corrected by using the 95% percentile or max values. Grenzdorffer (2014) also emphasized the fact that if differences between plants are small, the level of detail and accuracy of both the DTM and DSM need to be sufficient in order to show subtle differences in height variability. This again shows the need of very high resolution imagery which is georeferenced using ground control points to apply this method for small-scale maize trials. A study conducted by Bendig et al. (2014) used these ultrahigh resolution images derived from a RGB camera mounted on a UAV to calculate multi-temporal crop surface models for barley. The test site consisted out of 18 barley cultivars randomized in 54 plots of 3 x 7 m with sufficient spacing between the plots. Bendig et al. (2014) used 15 GCPs measured with a differential global positions system to create the crop surface models. This resulted in a very strong correlation between the derived plant height and reference measurements (R^2 = 0.92) over the whole growing season (six measurements). In general the bias of the derived plant height was about 0.1 m lower than the reference height. This shows that using very high resolution images on a homogenous canopy surface in combination with enough GCPs, very accurate height results can be achieved for smaller plots. However the sample size is rather small; only 54 plots versus thousands of plots used in typical phenotyping trials. Furthermore the plot spacing was wide, which is also not the case in large phenotyping trials. It

is therefore interesting to look at this method applied to these large phenotyping trials with thousands of plots that have a spikey canopy surface.

2.5 FUSING OF PLANT HEIGHT AND HYPERSPECTRAL DATA FOR BIOMASS ESTIMATION

Bendig et al. (2014) furthermore looked at the correlation between this derived plant height and above ground biomass per plot and found good correlations between both dry and fresh biomass ($R^2 = 0.82$ and $R^2 = 81$). The main limitation for this method was the influence of lodging cultivars in the later growth stage, producing irregular plant heights. Also wind effects produced errors, although they found that wind primarily causes a shift in the x-y direction. These lodging and wind phenomenon's also occurs in maize cultivars and could therefore influence the results. Furthermore the height prediction errors became larger at later growing stages and showed an increasing trend. Meaning larger errors in biomass prediction at later growing stages thus indicating that growth stages influence prediction accuracy due to the increasing spatial variability (Bendig et al. 2014). In a later study Bendig et al. (2015) combined these crop surface model with ground-based hyperspectral data to calculate vegetation indices (VIs) for the same study area. Combining the VIs with height data by using multiple linear regression ($R^2 = 0.84$) performed better than the VIs alone ($R^2 = 0.79$) for predicting biomass. However, in this study the relationship between height and biomass produced the most robust results. Bendig et al. (2015) indicated that plant height is competitive with vegetation indices for biomass estimation in summer barley. As the architectural traits of maize differ greatly from barley, it is interesting to test if VIs would also be competitive with height for maize.

Furthermore it is useful to determine what flying times are essential. As flying a UAV needs trained pilots and each flight generates additional costs (C. Zhang and Kovacs 2012), it is beneficial to determine at what growing stage the spectral data correlates best to the maize breeding trials biomass in order to reduce frequency of flights. Furthermore it is beneficial for breeders to know what sensors, RGB, multi or hyperspectral, give the best correlation to biomass in order to develop an optimized integrated platform.

3 STUDY AREA AND DATA DESCRIPTION

3.1 STUDY AREA

The study area was located in the southern part of the Netherlands near the village of Reusel (Lat 51.316500 Long 5.171000). The test fields were located on the Van den Borne farm, but were planted and managed by Limagrain. The test field was approximately 6.2 ha in size (Figure 2).

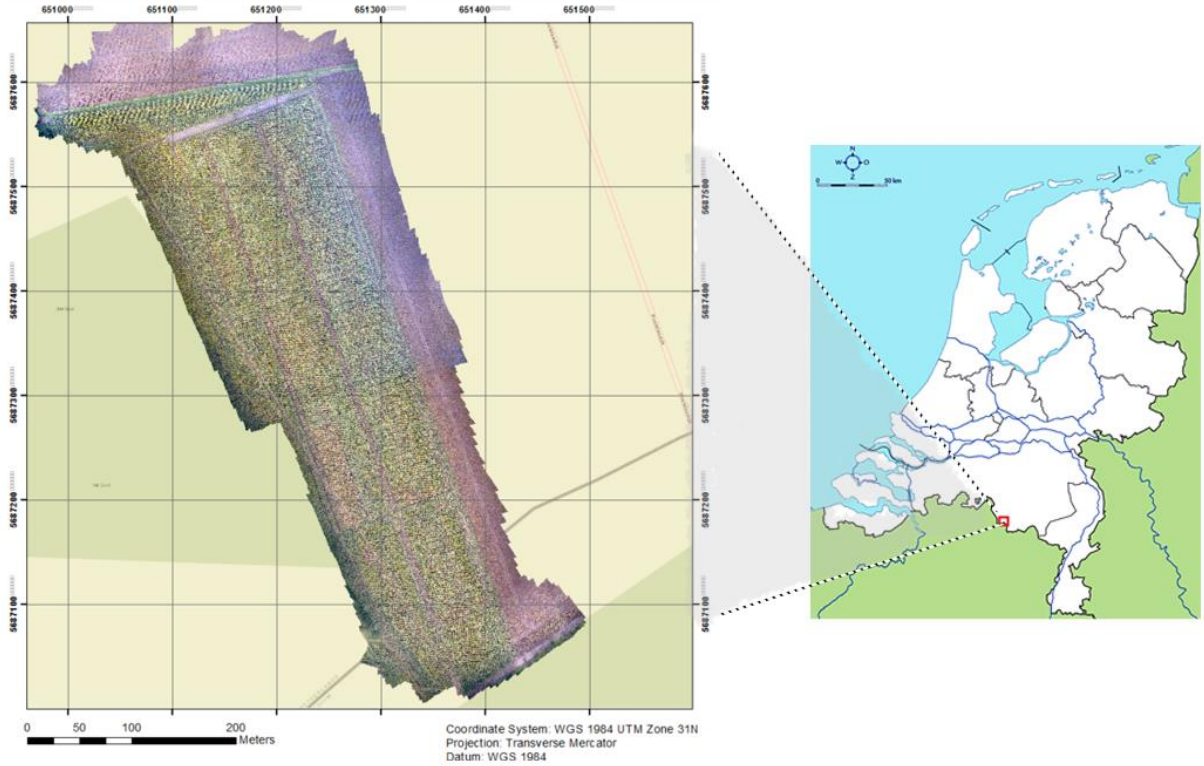


Figure 2 RGB orthomosaic of study area from 22-09-2015 overlaid on Openstreetmap (left). Overview map (right).

3.2 SET UP OF EXPERIMENTAL PLOTS

The test field contained a total of 3838 plots which were planted with 29 different maize (*Zea Mays)* cultivar trial groups. Reference data for 2696 of those plots was made available. The remaining 1142 plots were used as border plots between fields. The reference data consisted out of a cultivar code, size, average plant height per plot as measured by Limagrain with LiDAR (2696 plots) and manual measurements by Limagrain (56 plots). Each plot was 10.2 m^2 in size (6.8 by 1.5 meters) and contained approximately 50 maize plants divided over two equal rows. Biomass and LiDAR measurements of each plot were done at harvest by Limagrain and were provided for this study. The plots had two different set-ups (Figure 3). One was a two row set-up, in which the plots were located directly next to each-other. The other was a four row set-up in which four rows of the same species were planted, but only the inner two rows were used for evaluating.

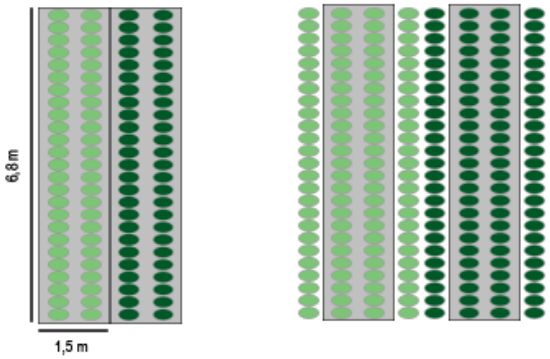


Figure 3 the two different set-ups of experimental plots (grey = plot area)

The four row plots had less influence from surrounding plots, as the plants next to the plot were from the same cultivar. This was different for the two row set up, in which for example overhanging leafs had a potential influence. Also geometrical errors were more apparent in the two row set up, as there was no spacing between the plots. (Figure 4, left).

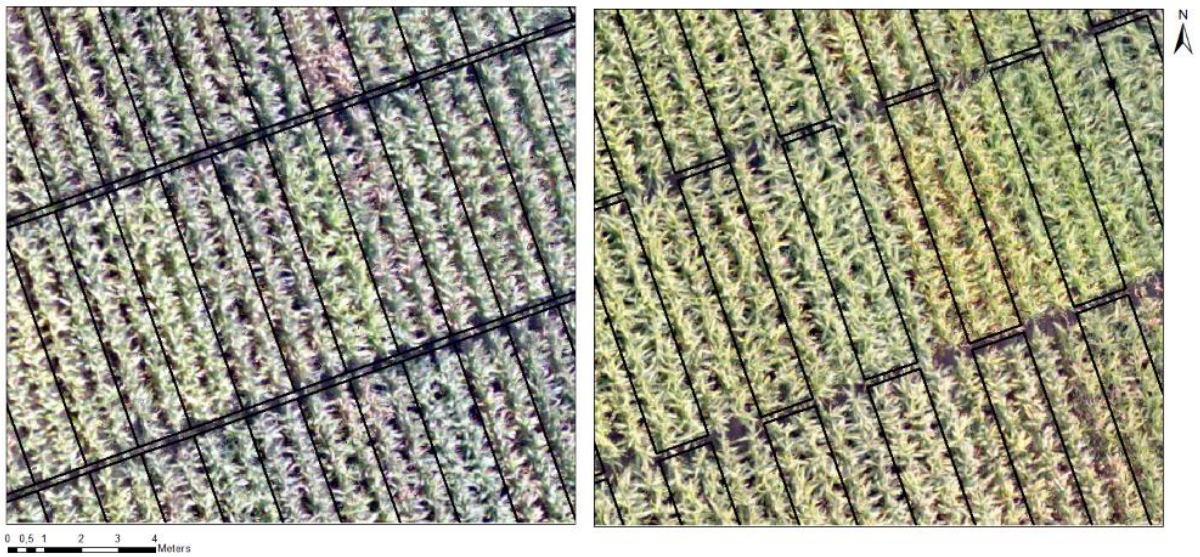


Figure 4 RGB images acquired from UAV on 22-09-2015 showing the difference between set-up of plots on geometrical and bi-directional influence. Left: image with the two row set-up, right: image with the four row set-up.

3.3 UAV IMAGE COLLECTION

3.3.1 Platform and software

We used the Aerialtronics Altura AT8 v1 octocopter (Figure 5, right) as the Unmanned Aerial Vehicle (UAV) platform for this study. The UAV carried a Panasonic GX1 RGB camera with a 14 mm pancake lens. This camera was able to take 16MPix every 2 seconds. Simultaneously the hyperspectral mapping system (HYMSY) was mounted on the UAV (Figure 5, left). The HYMSY is a custom-made push broom spectrometer with a range of 400-950 nm, 9 nm FWHM, 25 lines/s and 328 px/line (Suomalainen et al. 2014). It operated together with the RGB camera and a miniature GPS-Inertial Navigation System. Full specifications on the payload can be found in Annex A.

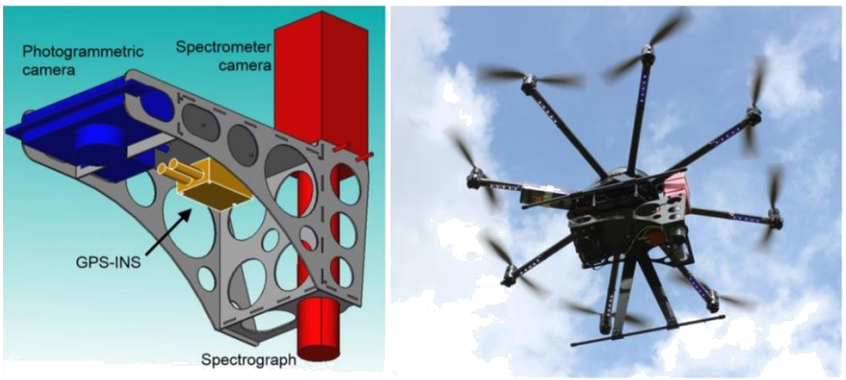


Figure 5 Hyperspectral sensor (left) and UAV used (right) (Suomalainen et al. 2014)

The processing chain of this integrated system was based on a photogrammetric algorithm to derive digital surface models (DSM) from the RGB images and provides a high accuracy orientation of the system of the DSM. The hyperspectral data from the push broom spectrometer was then georectified by projecting it onto the DSM. The photogrammetric orientations and the GPS-INS data were also used to reach an even higher geometrical accuracy. The system was able to produce a RGB orthomosaic between 1 and 5 cm resolution, a DSM at 5 to 10 cm resolution and a hyperspectral data cube between 10 to 50 cm resolution, depending on the flight altitude and speed (Suomalainen et al. 2014).

3.3.2 UAV flights

In total four flight days were planned and carried out at four different growing stages; at plant emerge (26-05-2015), before flowering (25-06-2015), after flowering (30-07-2015) and at harvest (22-09-2015). Each flight date contained five partially overlapping flight lines, as can be seen in Figure 6. The blue areas contain a large amount of overlapping which is beneficial for the quality of the 3D surface models.

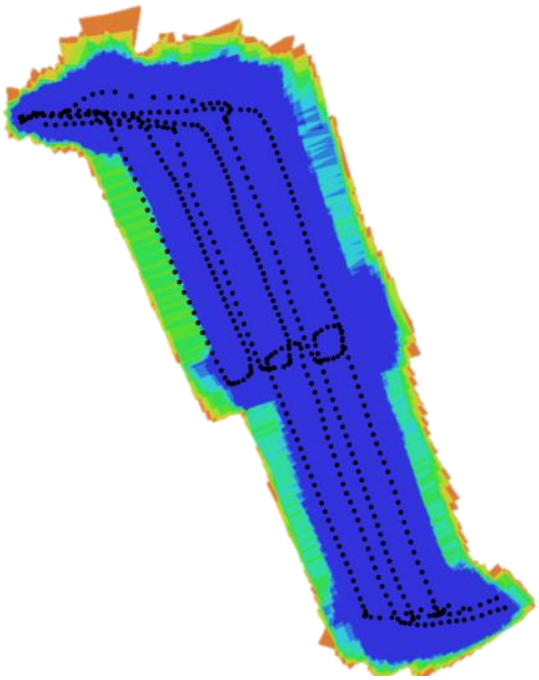


Figure 6 Flight lines per flight date. From blue to red are the amount of overlapping images (high to low)

3.4 GROUND DATA

3.4.1 Limagrain

Plot information and ground data at harvest was provided by Limagrain. Fresh biomass (average in grams per plot) was measured at harvest and averaged per plot. Reference plant height was measured with a LiDAR system (average in meters per plot). This resulted in an average height value per plot somewhere between the plume and the most upper leaf. Height variability within a plot was not provided. The manual height measurements were taken with a static measuring stick for each plant and averaged over the plot.

3.4.2 Ground control points

During the last flight date a total of five ground control points were measured with a RTK-GNSS system (Figure 7, right). The ground control points were made out of wooden plates which were 1 by 1 meter in size. The surface was black with a white circle in the middle (Figure 7, left). Six height points were used for the construction of the digital terrain model (DTM).

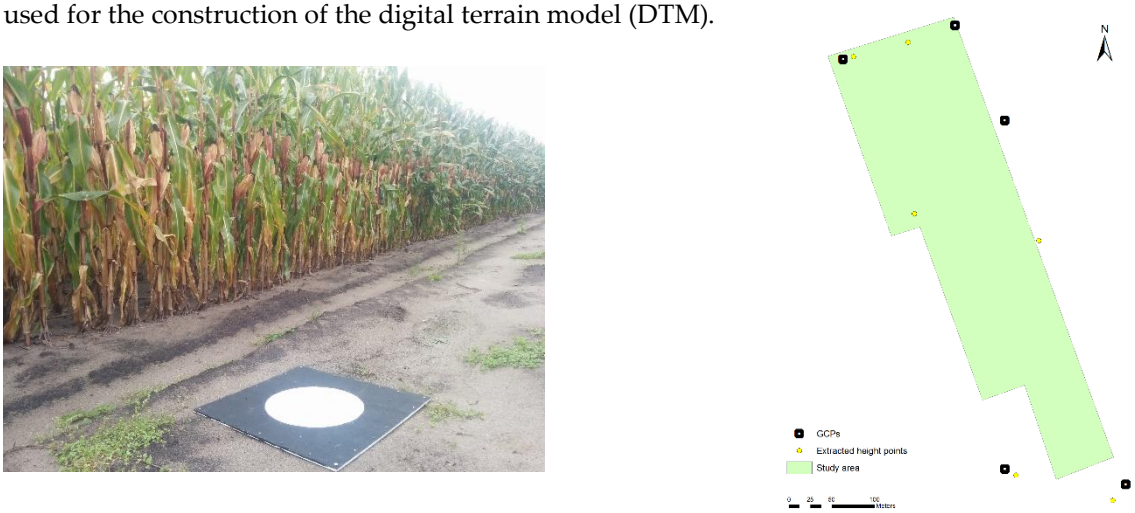


Figure 7 Photo of panels for ground control points (left) and location of GCPs within the experimental field and extracted height points for DTM (right)

Unfortunately GCPs were only made during the last flight. Due to unforeseen technical and weather circumstances the amount of available time to position GCPs during the last flight was very restricted. Therefore the total number of GCPs was limited. From the generated digital surface model (DSM), artificial GCPs were extracted based on known '0' points like tractor tracks (see methods). These GCPs were used as input for absolute calibration of the digital terrain model (DTM). For a map of the GCPs see annex B.

4 METHODOLOGY

We tested three UAV data products of the UARSF facility (www.wageningenur.nl/uarsf) on their capability to quantify plant traits in phenotyping experiments: hyperspectral data, vegetation indices and derivatives from crop surface models (CSM). With these data products we evaluated the accuracy of estimation and prediction of plant height and biomass. We used statistical analysis between the data products and measured reference field data to test the accuracy. High resolution multiple angle RGB images and structure from motion (SfM) software were used to construct 3D point clouds and highly detailed orthophotos. From these 3D point clouds multi temporal digital surface models (DSM) were obtained. We then used these DSMs to construct crop surface models from which plant height, volumes and surface areas were derived. Subsequently the estimated plant height per plot was correlated to the measured plant height per plot using regression analysis on both manual measurements as well as LiDAR measurements. This was done to test the accuracy of UAV derived plant height products compared to traditional manual measurements and LiDAR methods. Lastly we introduced data fusion in both time and products to develop biomass prediction models using crop surface model derivatives and hyperspectral data products. Below the different steps of the developed methodology will be explained in more detail.

4.1 PRE-PROCESSING

In order to generate three georeferenced data products from the raw UAV data, pre-processing was done partly manually. We pre-processed the last flight date first as ground control points were available for this flight. First the RGB images were filtered based on altitude and flight path. Al images that were taken below a flight altitude of 80 meter and around the take-off/landing zone were excluded. The selected raw images were then converted to TIFF files using the ExifTool (Harvey 2010) in the command line.

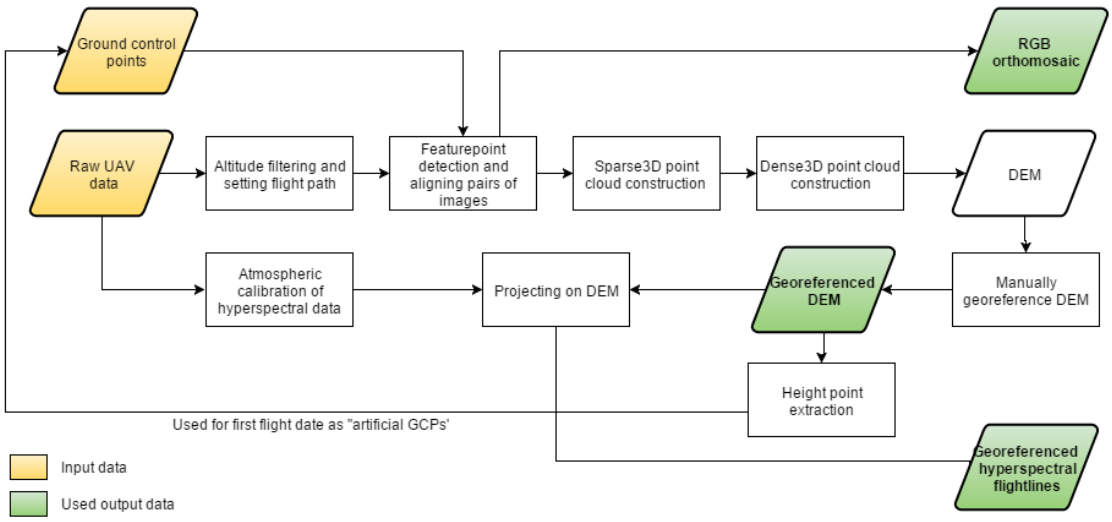


Figure 8 Pre-processing workflow

4.1.1 Ortho-mosaic and digital elevation model processing

From here we loaded all the RGB TIFF files into Agisoft Photoscan Pro 1.16 together with the RTK-GNNS data from the GCPs. Using the photogrammetric SfM computer vision algorithms from the Agisoft Photoscan workflow (Agisoft 2014) and the overlapping, low altitude high resolution RGB

imagery as an input, ortho-mosaics and multi-temporal 3D digital surface models were created. Traditional stereovision uses 2D images from multiple view angles to construct a 3D model. SfM required only a single camera and used its motion for creating multiple angles (Figure 9). This is a fundamentally different approach from traditional stereovision and enables SfM to be used with a single camera mounted on a UAV.

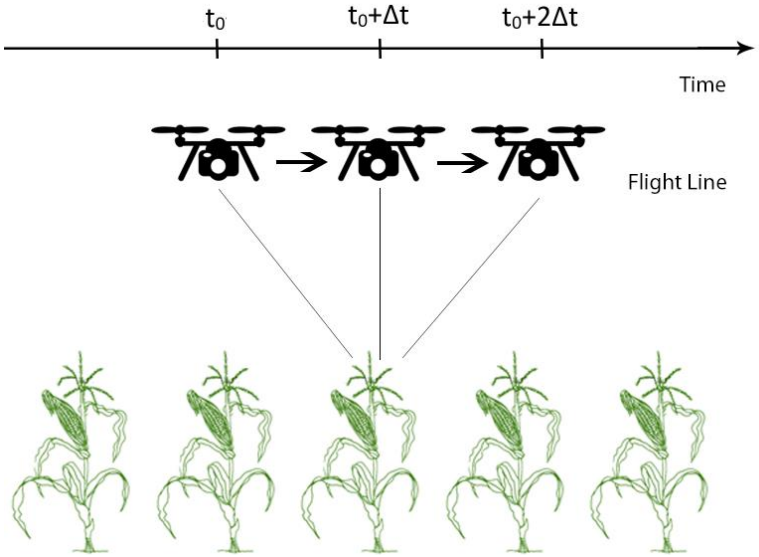


Figure 9 Structure from Motion set-up

The UAV measured the camera speed and acquired images at a set interval Δt. By setting the Δt at a short enough amount of time, the acquired images will be overlapping. A 90% overlap along the flight path is ideal and 70% overlap the minimum. Sideways, a 50% overlap is ideal. Therefore Δt was set at 2 seconds. With lower overlapping, the DSM will become less accurate. Due to this overlapping, every point was observed from different viewing angles (Figure 9). These points were matched with the same point in the nearest image in terms of Euclidean distance. The Agisoft software package has the ability to read the Global Navigation Satellite System (GNSS) data from the imagery's meta-data, assisting in the alignment of the images.

For accurately georeferencing the DSMs in x, y and z dimensions, stereo ground control points were used. Stereo GCPs are GCPs that are observed form multiple angles and is a cross between a regular GCP and a tie point. The point with known ground coordinates had to be identified in two or more images. The points had the same ID and coordinates, however pixel location was different in each image. These points were then used to identify how images are related. Three GCPs is the mathematical necessary minimum, and for this study we used five GCPs. We placed the GCPs so that each point was observed from multiple angles as described in Figure 9. We measured the GCPs location with a Real Time Kinematic (RTK) GNSS system for accurate location and height reference data. The GCPs were then merged in the Agisoft Photoscan workflow. This process was not automated and required the manual marking of each GCP in every image. After alignment of the images based on the photogrammetric orientations, GPS data and GCPs, only the GCPs were used in a re-alignment step in order to 'force' the DSM to use the GCPs as only valid reference. In this way, possible non-linear deformations of the model were removed by optimizing the estimated point cloud on the known reference coordinates. This optimization adjusted the estimated point coordinates and camera parameters, effectively minimizing the sum of reprojection error and reference coordinate misalignment error (Agisoft 2014).

Following these steps, a partly georeferenced RGB orthomosaic was produced at 1.6 centimeter resolution. We manually optimized the georeferenced RGB orthomosaics for better accuracy using tie points on a recent satellite image in ArcMap 10.3. Next to the RGB orthomosaic the Photoscan software created a mesh from the dense point cloud, forming a digital elevation model.

The output DSMs had a resolution of 2.5 centimeters. The DSMs for all flight dates were then manually georeferenced to the RGB orthomosaic using tie points in ArcMap 10.3 for better accuracy. Subsequently, the raw hyperspectral data was then atmospherically calibrated using a reference panel and the pixels projected on the georeferenced DSM. This resulted in georeferenced hyperspectral flight lines, based on the DSM. However due to the push broom scanner, geometrical errors were inevitable (Figure 10, left). Manual re-georeferencing the flight lines improved some of the errors, nevertheless it was not possible to get all the flight lines geometrically correct. Due to this geometrical mismatching only one hyperspectral flight line, which was most geometrically correct, was used over all flight dates (Figure 10, right). This flight line showed the least geometrical distortions in its flight path and could be warped to the ortho-mosaic with considerably more precision than the highly distorted flight lines.

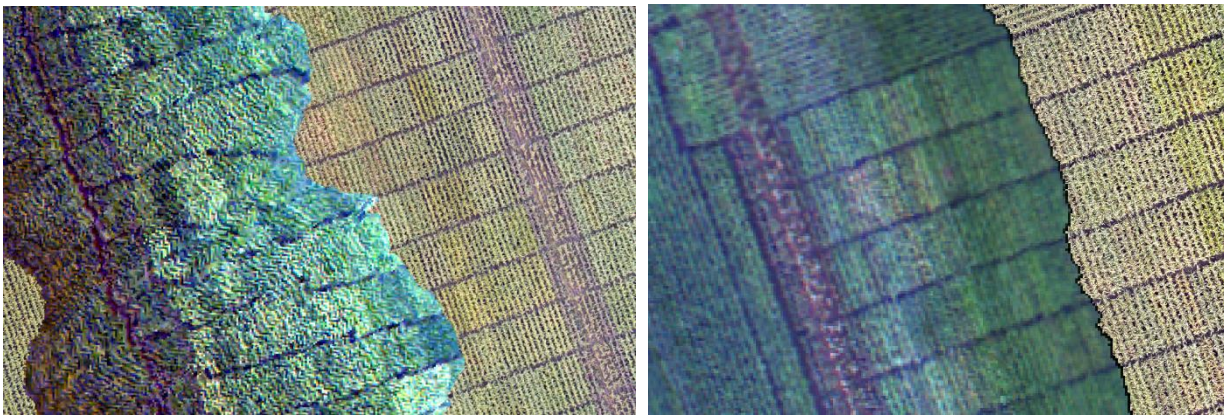


Figure 10 Geometrical distortions caused by push broom scanner (left) and flight line used (right). Both images are from 22-09-2015.

From the georeferenced DSM, six height points with a known '0' value were extracted. 0 value points were points that did not change in z-value over the growing season. These consisted mostly out of tractor tracks in the field. The extracted height points were then used as 'artificial GCPs' for the pre-processing of other flight dates (Figure 8) in order to create an accurate and referenced digital terrain model (DTM). After pre-processing the three data products (RGB ortho-mosaics, DEMs, and hyperspectral data cubes) were fully calibrated and georeferenced for all flight dates.

4.1.2 Shapefile processing

The shapefiles for the 3838 plots were generated with ArcMap 10.3 (Annex C). Their size, spacing and location was based on the seeding pattern data obtained from the GNSS systems mounted on the tractors. This implicated that the plots were drawn in 100% straight lines. However, as the ortho-mosaic was never 100% geometrically straight, manual adjustment of the shapefile containing the plots was needed. The manual adjustment of the geometrical locations of single plots was based on multiple ortho-mosaics and carried out in ArcMap 10.3.

4.2 Crop Surface Models and calculation of crop height

The DSM at harvest was then used to calculate the minimum/maximum, standard deviation and average plant height per plot to show height variability between plots using the method proposed by Bendig et al. (2014). They generated a workflow using ArcGIS (Figure 11, right) for calculating crop height by subtracting the mean bare ground digital elevation model (DTM) from the digital surface model at harvest. This results in a crop surface model (CSM) with absolute crop heights. We used this method and implemented it in the R programming language. Furthermore we derived a Triangular Irregular Network (TIN) model from the CSM using ArcGIS tools. From this we calculated volume and surface area. The plot shapefile was then used to extract plant height, volume and surface area on plot level in the R programming language (Figure 11, left).

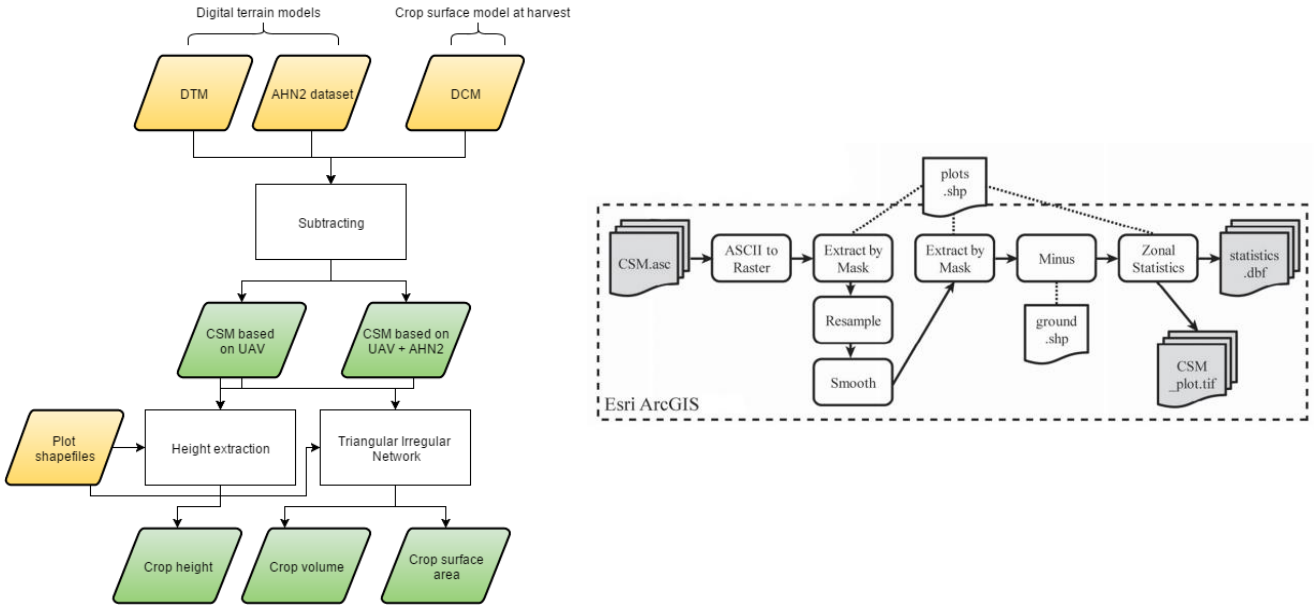


Figure 11 Crop Surface Models workflow (left) and Plant Height (PH) extraction method (right) as proposed by (Bendig, Bolten, and Bareth 2013)

In order to compare the generated DTM, we used the 'Actueel Hoogtebestand Nederland' (AHN2) dataset. This is a height dataset obtained with an airborne LiDAR system in which all vegetation is removed, based on classification of cloud points. For the study area the dataset had a point density of six to ten points per square meter with an error margin of 15 centimeters for 95% of the dataset (van der Zon 2013). We used the AHN2 dataset of 2011 with a resolution of 0.5 meters to test the differences between this dataset and the UAV obtained DTM. Furthermore we evaluated the crop surface models derived from the AHN2 dataset to test if data from another source could be fused with UAV products in a case were the DTM at the beginning of the growing season is not available. The workflow is described in figure 11. In total two CSMs were calculated, one based on the UAV DTM and one based on the AHN DTM (Table 1). We used two different thresholds for each CSM, the mean and top of canopy. We calculated the mean values by averaging the heights per plot and the top of canopy pixels by using a two sigma ($\mu + 2\sigma$) threshold per plot.

Table 1 Calculated datasets based on two different DTM sources and two different DSM values

Dataset	DTM source	DSM source	Resolution	Threshold	Date
UAV DTM	UAV	-	2.5 cm	-	26-05-2015
AHN2 DTM	AHN	-	50 cm	-	2011
DSM	-	UAV	2.5 cm	-	22-09-2015
$CSMA_{mean}$	AHN2	UAV	2.5 cm	μ	22-09-2015
$CSMA_{toc}$	AHN2	UAV	2.5 cm	$\mu + 2\sigma$	22-09-2015
$CSMU_{mean}$	UAV	UAV	2.5 cm	μ	22-09-2015
$CSMU_{toc}$	UAV	UAV	2.5 cm	$\mu + 2\sigma$	22-09-2015
Vegetation indices	-	UAV	25 cm	-	Growing season

Where DTM = digital terrain model, DSM = digital surface model, CSMA = crop surface model based on AHN2 and CSMU = crop surface model based on UAV.

4.3 HYPERSPECTRAL BASED BIOMASS PREDICTION

For biomass estimation we used the UAV derived hyperspectral data products to calculate vegetation indices during four different growing stages; emerge, before flowering, after flowering and at harvest (Figure 1). Calculations of vegetation indices using multi or hyperspectral wavelengths can give estimates of above ground biomass (Qi et al. 1994; Jin et al. 2014; Kross et al. 2015; Bendig et al. 2015). As there are numerous vegetation indices (Huete et al. 1997), a selection was made for this research. We choose the VIs on their characteristics, type and performance with regard to biomass estimation. The indices use in this study are presented in Table 2

Table 2 Vegetation indices used in this study

Name	*Formula*	*Spectral region*	*Reference*
Normalized Difference Vegetation Index	$NDVI = \frac{(\rho NIR - \rho RED)}{(\rho NIR + \rho RED)}$	Multi- and hyperspectral	(Rouse et al. 1974)
Weighted Difference Vegetation Index	$WDVI = NIR - C * R$	Multi- and hyperspectral	(Clevers 1989)
Optimized Soil-Adjusted Vegetation Index	$OSAVI = \frac{(1 + 0.16) * (\rho 800 - \rho 670)}{(\rho 800 - \rho 670 + 0.16)}$	Multi- and Hyperspectral	(Rondeaux et al. 1996)
Modified Triangular Vegetation Index 2	$MTVI2 = \frac{1.5[1.2(\rho_{800} - \rho_{550}) - 2.5(\rho_{670} - \rho_{550})]}{\sqrt{(2 * \rho_{800} + 1)^2 - (6 * \rho_{800} - 5 * \sqrt{\rho_{670}}) - 0.5}}$	Hyperspectral	(Haboudane et al. 2002)

One slope-based VI and three distance-based VIs were used (Table 1). Slope-based VIs are very susceptible for Leaf Area Index (LAI) and do get saturated quite fast (Mroz and Sobieraj 2007), yet the Normalized Difference Vegetation Index (NDVI), a slope-based VI, is widely used and studies shows

good results in using NDVI for dry biomass estimation of maize (Coefficient Variation = 16% and R^2 = 0.95 (Kross et al. 2015)). Therefore, the NDVI was taken into account for this research. Distance-based VIs are less sensitive to LAI and offer a good correction for soil background (Clevers 1991), which is useful in later growing stages. The Weighted Difference Vegetation Index (WDVI) was therefore used with a C value of 1.24 (NIR/R for soil pixels). Furthermore there are indices that show a greater sensitivity to higher LAI and biomass, such as the Modified Triangular Vegetation Index 2 (MTVI2) (Haboudane et al. 2002) which scored good results in estimating dry biomass of maize (Coefficient Variation = 37% and R^2 = 0.73 (Kross et al. 2015)). Furthermore the soil-adjusted OSAVI vegetation index was used for its good results in prediction biomass (Peng and Gitelson 2011). The inputs for NDVI and WDVI can be broadband (multispectral) or narrowband (hyperspectral). We used broadband regions with an average of 640-670nm for the red band and 770-900nm for the NIR band. These regions gave a good representation for the color and are also widely used on satellites like the Landsat series and the Sentinel constellation (USGS 2014). We used narrowband regions of 670nm wavelength for the red band and 800nm wavelength for the NIR band. Reflection of 670 nm is the maximum absorption in the red region and studies have found strong correlations on LAI and chlorophyll content using the 670 and 800 nanometer region ratio (Chaoyang et al. 2008; Sims and Gamon 2002). Using these wavelengths the greenness of the vegetation was calculated. We calculated the indices using the R programming language with the hyperspectral data cubes as an input. Subsequently we averaged the vegetation indices on plot level and extracted the values using the same workflow as figure 11. R^2, probability value (p-value) and root mean square error (RMSE) was calculated to test the correlation between the different indices and biomass measured at harvest.

4.4 COMBINING EXPLANATORY VARIABLES

As a last step, we used Correlated Component Regression (CCR) to introduce fusion of the UAV derived data products over time (Figure 12). As the numbers of predictors was fairly large, especially through time, and were presumably correlated with each other, correlated component regression was introduced. CCR is a suitable tool for predicting outcome variables from a large number of predictors. It is a proven method to prevent overfitting and only takes significant predictors into account (Deal 2011). It also provides information on the significance of each predictor. Studies have shown that CCR outperforms comparable approaches that are based on stepwise regression, penalized regression and PLS regression when the number of predictors is fairly large and are correlated (Magidson 2010). This is due to the fact that CCR is better in capturing important suppressor variables in the final predictors. CCR used K correlated components, which are a linear combination of the predictors, to predict an outcome variable. The first component S_1 captured the prime predictors, while S_2 captured the suppressor variables. If other components improved prediction significantly they were also included. To reduce variables, a step-down algorithm was used were the least important predictors got removed. To determine the number of components and predictors, K-fold cross validation was used (Magidson 2010). The input variables were all four VIs over the growing season, volume, surface area, and the UAV based CSM top of canopy height. We furthermore tested ten models with different predictors, based on the important variables.

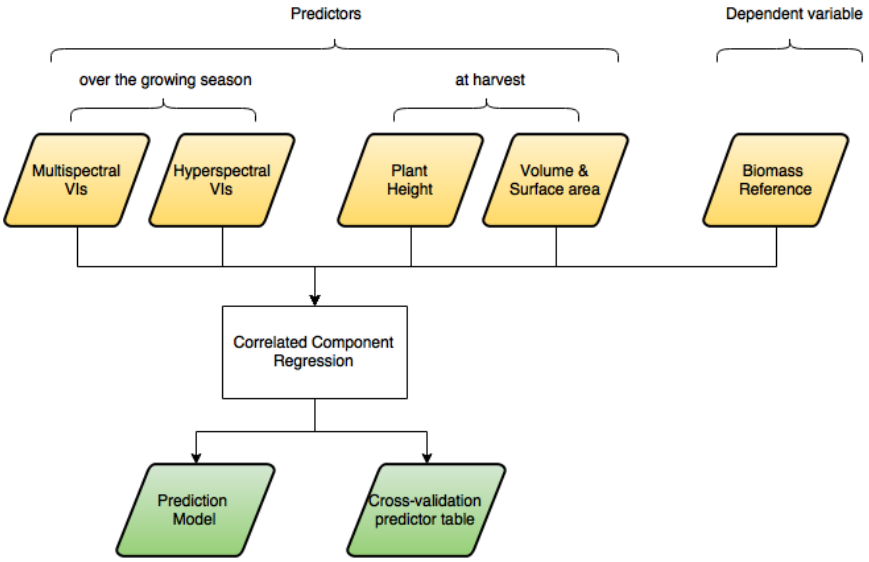


Figure 12 Workflow showing the fusion of data products over time for biomass prediction

4.5 Statistical Analysis

We evaluated the ability of the vegetation indices in predicting above ground biomass as well as the 3D models ability to predict canopy height using the Root Mean Square error (RMSE) calculation. The biomass and height data from Limagrain was used as the observed values (reference) and the UAV derived spatial maps as predicted values. As RMSE can be sensitive to bias, also the bias and standard error of prediction corrected for the bias (SEPc) were calculated. This is a bias-corrected error.

$$\mathrm{RMSE} = \sqrt{[1/n \sum_{i=1}^{n} (\hat{y}_i - y_i)^2]} \quad (1)$$

$$\mathrm{BIAS} = 1/n \sum_{i=1}^{n} (\hat{y}_i - y_i) \quad (2)$$

$$\mathrm{SEPC} = \sqrt{[1/n \sum_{i=1}^{n} (\hat{y}_i - y_i - \mathrm{BIAS})^2]} \quad (3)$$

Where $\hat{y}_i$ is the model predicted value and y_i the measured value.

For testing linearity in the ability of biomass prediction the coefficient of determination (R^2) was used. The R^2 is calculated per VI as a function of the measured biomass using the following formula:

$$R^2 = 1 - \frac{SS_E}{SS_T}. \quad (4)$$

Where SS_E is the total sum of squares of residuals and SS_T is the total sum of squares.

5 RESULTS

The results in this chapter are based on the methods described in the methodology chapter. The results are presented in the order of the methodology and the research questions. We start with the results from the pre-processing, then the analysis of the derived crop height and subsequently the correlation between biomass and vegetation indices including fusion of data products and time.

5.1 UAV DERIVED DATA PRODUCTS

5.1.1 Digital Surface Model

The first data product created from the UAV based RGB imagery was an ortho-mosaic and digital surface model (DSM) at harvest (Annex F). The ortho-mosaic had a high spatial resolution of 1.6 centimeters. For the DSM, a resolution of 2.5 centimeters was obtained. Figure 13 shows the DSM overlaid with the RGB imagery to create a 3D crop model. This is a textured wireframe as is created with Agisoft Photoscan. To the left of the bare soil we see the maize trial plots. Figure 13 furthermore shows the solid wireframe of the same area with bare soil. Here we clearly see an interpolation effect of the SfM algorithms (Figure 13 right). The edges of the maize plots were not straight, but had an oblique angle from the crop canopy to the soil.

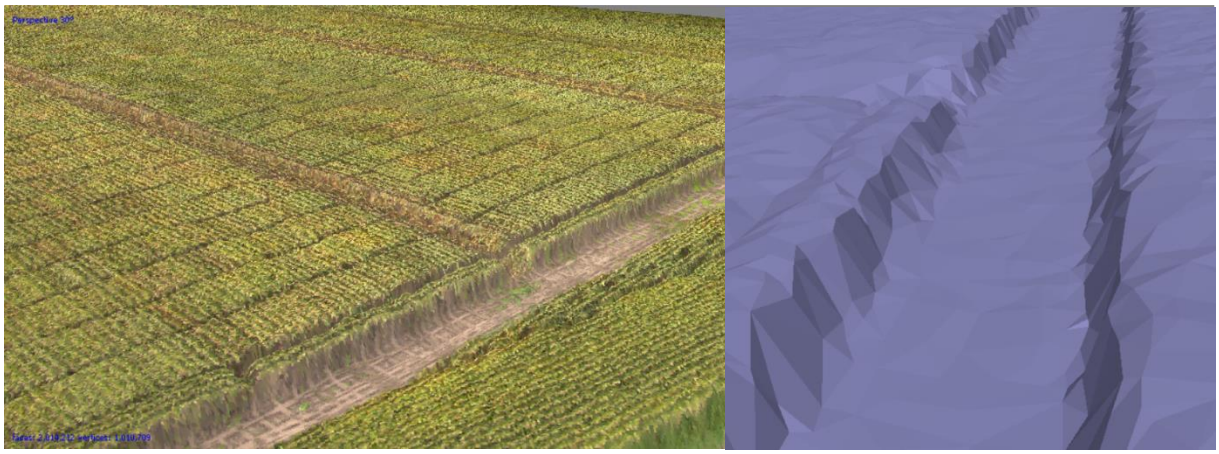


Figure 13 3D textured wireframe (left) and solid wireframe (right) screenshot of the maize field on 22-09-2015 derived from Agisoft Photoscan

To test the accuracy of the DSM compared to the RTK-GNSS points we extracted the height values from the DSM at the same coordinates as the RTK-GNSS measurements. The outcome showed a clear underestimation of the DSM points in the Northern part and an overestimation in the Southern part (Figure 14). The three GCPs in the Northern part had an average underestimation of 30 centimeters, while the two GCPs in the Southern part had an average overestimation of 70 centimeters (for a map of the GCPs see annex A).

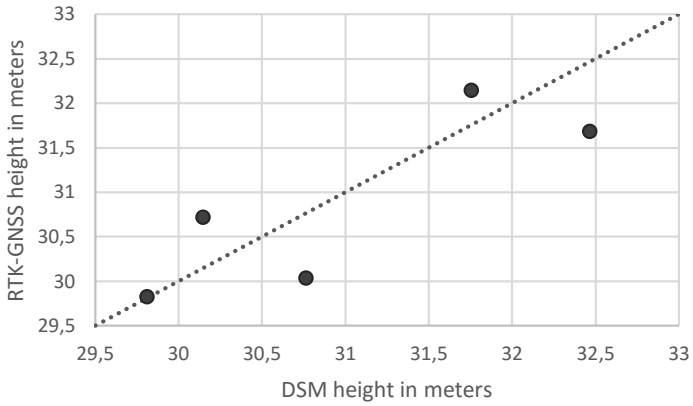


Figure 14 Graph showing the differences between the GCPs height in NAP from the RTK-GNSS and DSM

5.1.2 Digital Terrain Models

The DTM that was created with the help of six extracted height values from the DSM at harvest had a resolution of 2.5 cm and showed a centimeter level spatial accuracy (Figure 15, Annex E). In figure 15 we see a part of the DTM with a 20 meter long profile line (black line). The graph in the lower left corner of the image indicates the depth profile of this line. We see that the soil profile was clearly visible, showing height differences of around 10 centimeter. These were the sowing tracks of the planted maize. The image in the right top corner illustrates a solid 3D wireframe of the terrain. This figure shows the high level of detail that was obtained using ultrahigh resolution RGB imagery.

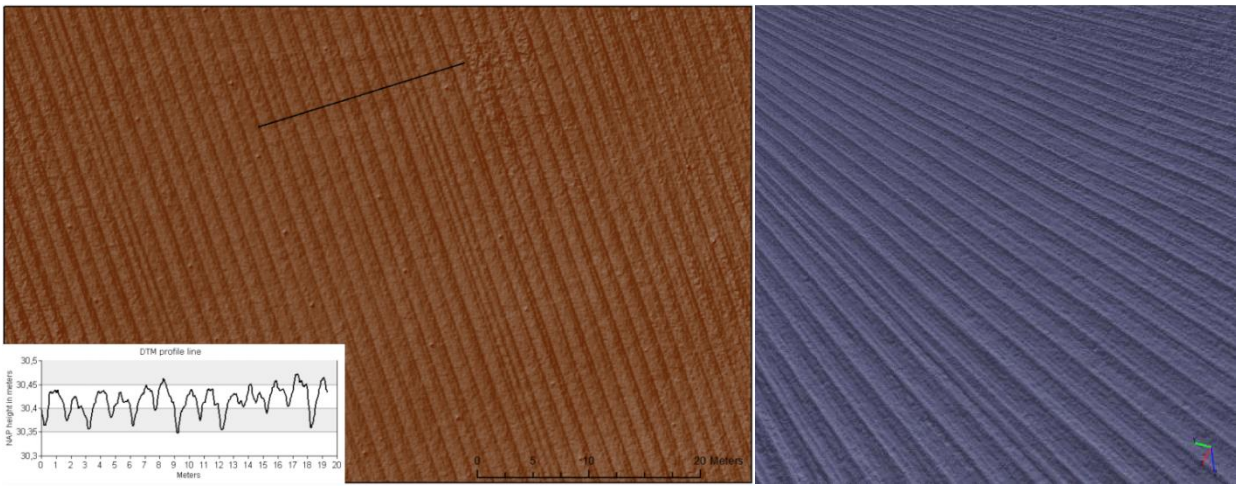


Figure 15 Digital Terrain Model (26-05-2015) with graph of profile line (bottom left) and 3D solid wireframe (top right)

The AHN2 digital terrain model that was used had a resolution of 50 centimeters (Annex D). The grid cells of this DTM were smoothed out using inverse distance weighting. Due to the combination of the lower resolution and the interpolation, the AHN2 DTM showed less detail. Sowing tracks and other corrugations for example were not as clearly visible. Furthermore the AHN2 DTM was obtained 4 years earlier in 2011 and exhibited a slightly different height variability over the field compared to the UAV based DTM. Furthermore this dataset was not recorded around the time of soil preparation, but based on the classification of vegetation and soil points.

5.2 Crop Surface Models

By subtracting the DSM at harvest from the two different DTMs, crop surface models (CSM) were created for the final growing stage at harvest (22-09-2015). The first crop surface model was based on the AHN2 DTM (CSMA), while the second crop surface model was based on the UAV DTM (CSMU). As the GCPs were measured in NAP height, the z-value of both models had the NAP reference system. This resulted in CSMs with absolute crop heights with a resolution of 2.5 cm. The crop heights were calculated for the whole study field. A map of the CSMs can be found in annex G and H.

5.2.1 CSM capabilities and level of detail

In Figure 16 we see a zoomed in section of the CSMU (Annex H) of approximately 3200 m². A gamma stretch of 0.5 was applied to emphasize the height differences. The purple bare soil strip in the middle of the image separates the two and four row plot set-ups. It is easy to distinguish the two and four row plots based on the height variability that can be observed between the plots. On the left we clearly see the height difference between the four rows of the same cultivars, while on the right we can distinguish the two row plots. The graph in figure 16 illustrates the black profile line. It shows the maize canopy rows and the obtained level of detail in height. We can clearly see the four different cultivars, starting with the lowest (left side of the profile line) to the highest plot (right side of the profile line). There was about 60 cm difference between the lowest and highest canopy row. The graph of figure 16 illustrates that the CSM captured both canopy and in-between leaf pixels. The graph depicts all the pixels in the profile line and clearly shows the shape of the maize canopy and lower leaf pixels. We didn't observe any ground pixels when there was a dense canopy, however with a less dense canopy some ground pixels became apparent. This is most visible in the right side of figure 16 in between the width of the plots.

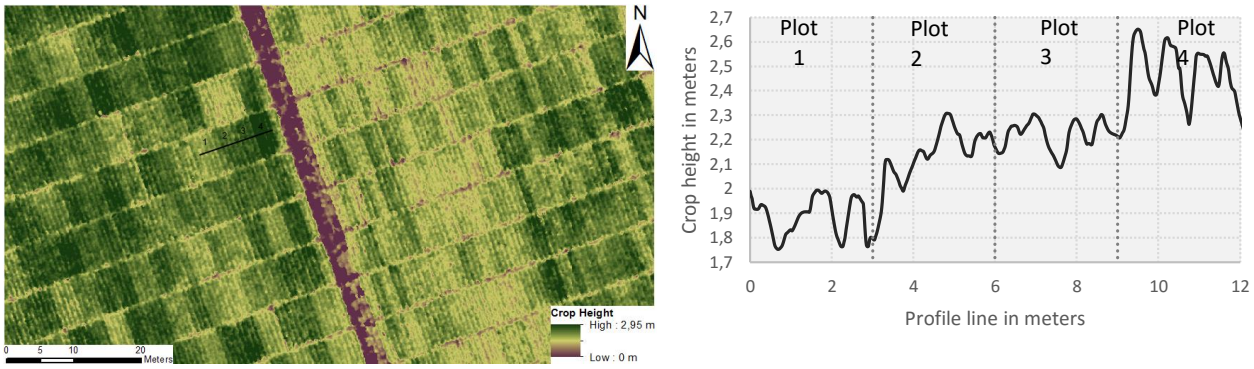


Figure 16 Zoomed in section of the Crop Surface Model (left) and profile line graph (right). The profile line is from a four row plot set-up

When looking at one plot in more detail (Figure 17) we can see that there is significant height variability captured within a plot. One plot contains 16320 height pixels (2.5 cm pixel size) on average. This illustrates the SfM algorithms capabilities of capturing height variability within a small plot when using ultrahigh spatial resolution imagery. When we look at the consistency of the height per row we see some variability. This indicates that the SfM was not always capable of capturing the spikey canopy surface, resulting in inconsistent peaks. This inconsistency is visible in the graph of figure 17 in the lower right corner. Another observed phenomenon were the edge-effects around the plot. There was a clear edge where the height dropped in the front-left side of the plot in figure 17 (left). This edge-effect is also visible in the graph.

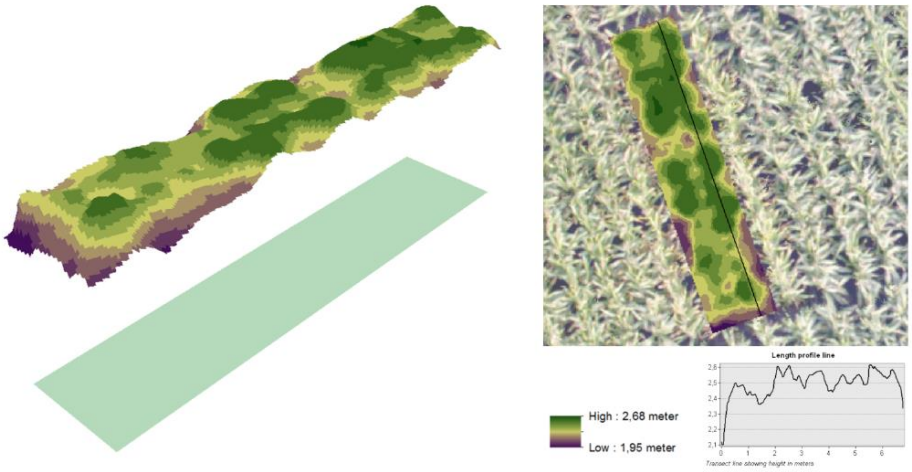


Figure 17 Detailed analysis of a single two row plot

When we look at the cumulative distribution (Figure 18) of all the height pixels for the plot in figure 17 we see that 80% of the height pixels were around 2.3 to 2.5 meters. About 10% of the pixels were the edge pixels of around 1.95 to 2.2 meters. The top 10% was around 2.5 to 2.7 meters. In this upper 10% region the manual height of this plot was also measured (2.65 meters). We see that the average of the CSM derived height was only 2.43 meters as this average was influenced by the edge and in-between pixels (around 90% of the pixels). To exclude all but top of the canopy pixels we took the top 2.3% by using 2 sigma ($\mu + 2\sigma$) threshold. We then see that with 2.67 meters this height was more accurate compared to hand measured height. This clearly illustrates the difficulties with a spikey canopy surface and edge-effects of a plot, but on the other hand shows that SfM algorithms are very capable of capturing detailed height variability which is interesting for a plant breeder. The manual height measurement of this plot was 2.65 meter on average, which is very close to the 2.67 meter derived from the top of the canopy pixels. However, the height measured with the terrestrial LiDAR system was 2.80 meter, meaning there is still a difference of approximately 4.6%, or 13 centimeter in height for this particular plot. From now on we will refer to the mean height derived from a CSM as CSM_{mean} and to the top of the canopy pixels as CSM_{toc}.

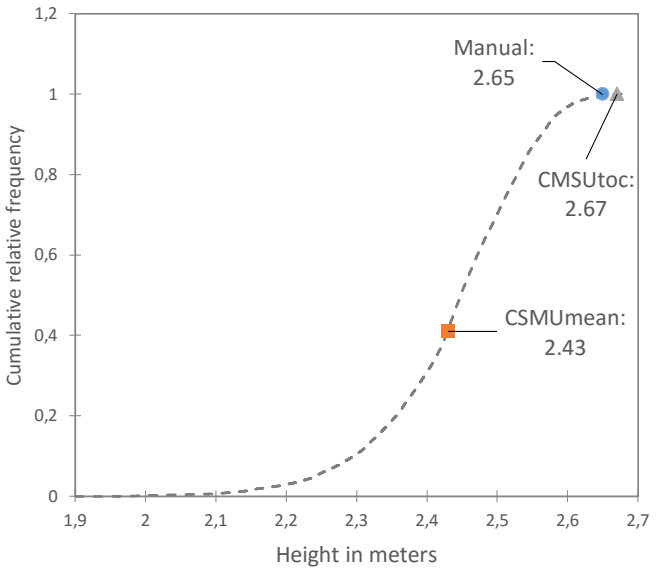


Figure 18 Cumulative relative frequency of 16320 height pixels for a single plot (RIH3EI03014011)

5.3 COMPARING THE DIFFERENT CROP SURFACE MODELS

5.3.1 AHN2 DTM based CSM (CSMA)

The first CSM was based on the AHN2 DTM, which we will refer to as CSMA (Annex G). The terrain of the study area was not flat, but rather showed an upward slope from the Northern part to the Southern part of the field. The AHN2 Digital Terrain Model captured this slope from the North to South, with a height difference of about 1 to 1.5 meters (Figure 19). The slope was steepest till around 2/3 of the test field, roughly the length of block 2 (Figure 21). From there on the slope got less steep, especially for block 1 (Figure 19).

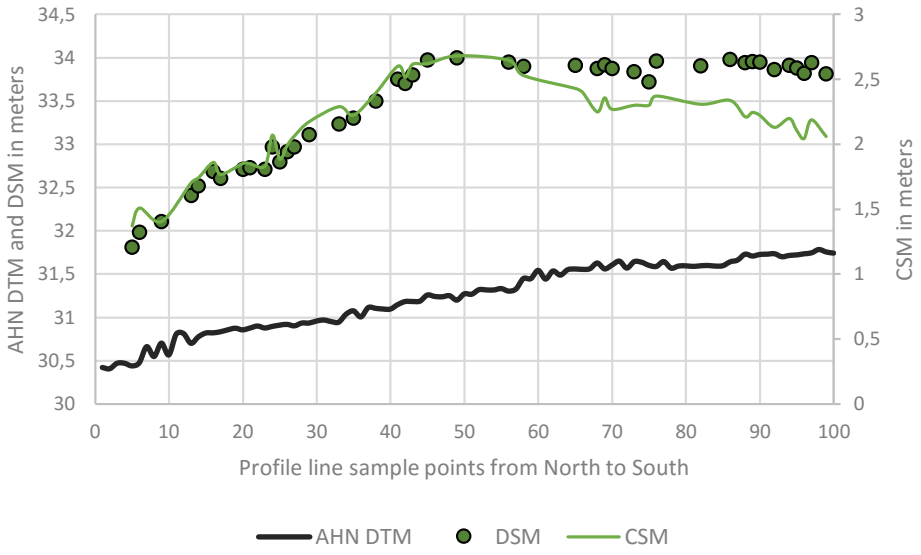


Figure 19 Profile lines from North to South for AHN DTM (black line), DSM sample points (green points) and CSM (green line).

The DSM also captured this slope, although with a slightly different angle. Due to the underestimation of the GCP in the Northern part (Figure 14) and the lack of GCPs inside the field, the DSM started too low and then started to overestimate compared to the AHN2 DTM (Figure 19). Overall the DSM was more skewed than the AHN2 DTM (Figure 20, top). This led to the crop surface model having a slope, similar to the AHN2 DTM slope. This can be seen in both figure 19 and figure 20 (middle). Figure 20 is a schematic representation of the cause of the error in height for block 2. It shows the effect of the difference in slope angle of the DSM compared to the AHN DTM on the CSMA. The effect caused the CSMA plant height to be too low in the Northern part and to have an increase in height for the southern part of block 2.

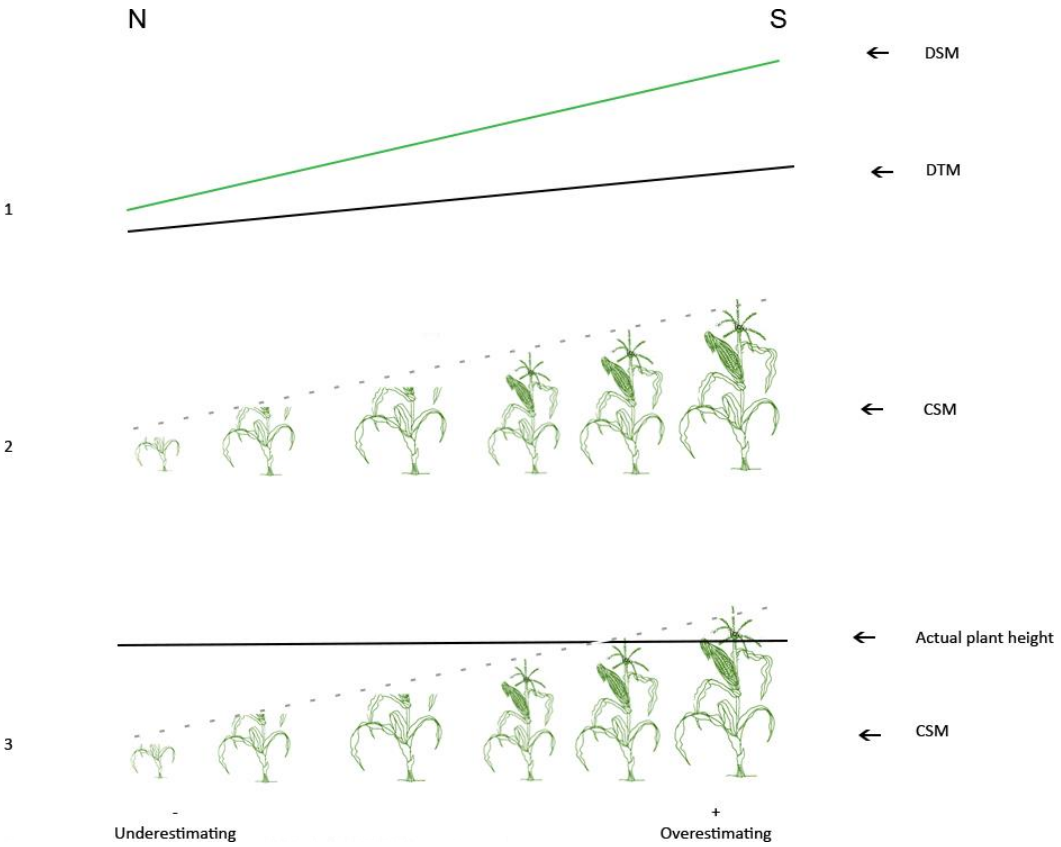
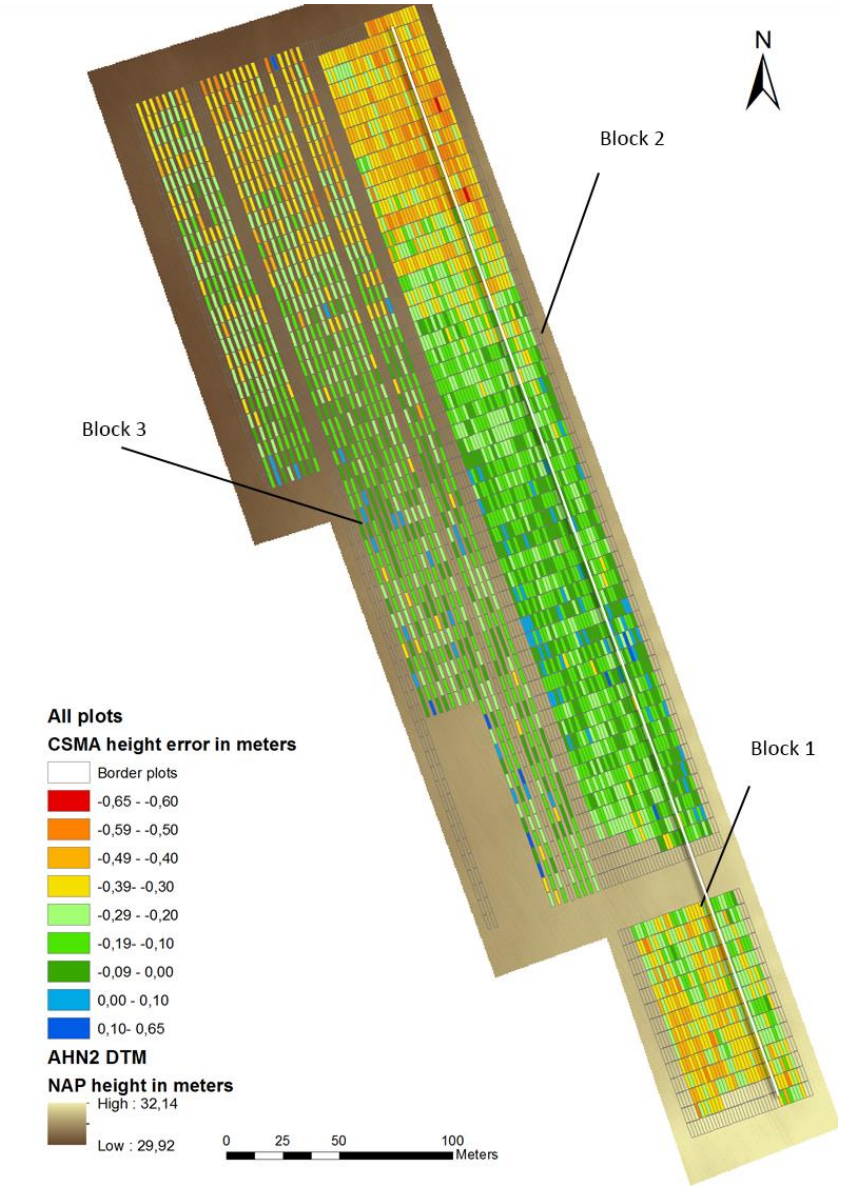
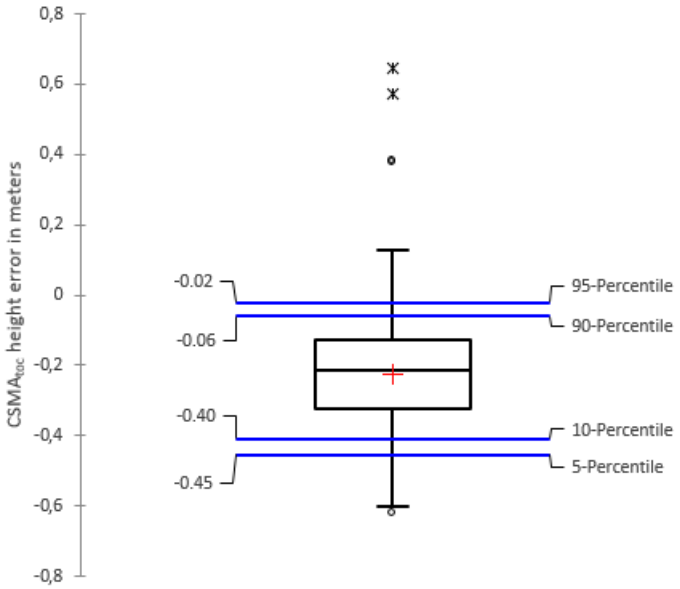


Figure 20 Simple schematic representation for the causes of height errors over the field from North to South for 2/3ths of its length (block 2)

The effect of this can clearly be seen in the map of figure 21. This map shows the AHN2 DTM and the height errors for all plots based on difference between LiDAR and CSMA top of canopy ($CSMA_{toc}$). The pattern of underestimating plant height in the North and increasing in the South is clearly visible in block 2 and 3. As the AHN2 DTM for block 1 did not have such a steep slope (Figure 19) the error caused by different slope angles was less apparent. Therefore the height results from block 1 were more representable. The average height error for this block was -0.30 meters. Also the mean and top of canopy height thresholds correlate strongly to the LiDAR measurements for block 1, with the strongest correlation for the top of canopy threshold ($R^2 = 0.77$).

When using this 30 centimeter height error over the total plots, a significant correlation was found. In total, 79.6% of the 2696 plots had a height error between 0 and -0.30 meters. When using this 80% of the total (2147 plots) a strong correlation was found between the estimated plant height from the $CSMA_{toc}$ and the plant height measured with LiDAR. The correlation showed an R^2 of 0.81. The bias for these plots was -0.17 meters. The mean values of the CSM ($CSMA_{mean}$) showed a lower correlation with an R^2 of 0.75 for the same plots. When looking at the percentiles of the height errors we see that the median for all plots was at -0.22 meters (Figure 21). Furthermore we see that 95% of the data was underestimating, using the top of canopy threshold. There were three outliers which were probably caused by erroneous points in the 3D point cloud. These results indicate that the AHN2 was capable of estimating plant height, but that using AHN2 data as a DTM caused systematic errors over the study field due to height inconsistency between the DTM and DSM over time.





Percentile	*Height error*
Maximum 100%	0,64
99%	0,04
95%	-0,02
90%	-0,06
3rd Quartile 75%	-0,13
Median 50%	-0,22
1st Quartile 25%	-0,32
10%	-0,41
5%	-0,45
1%	-0,53
Minimum 0%	-0,62

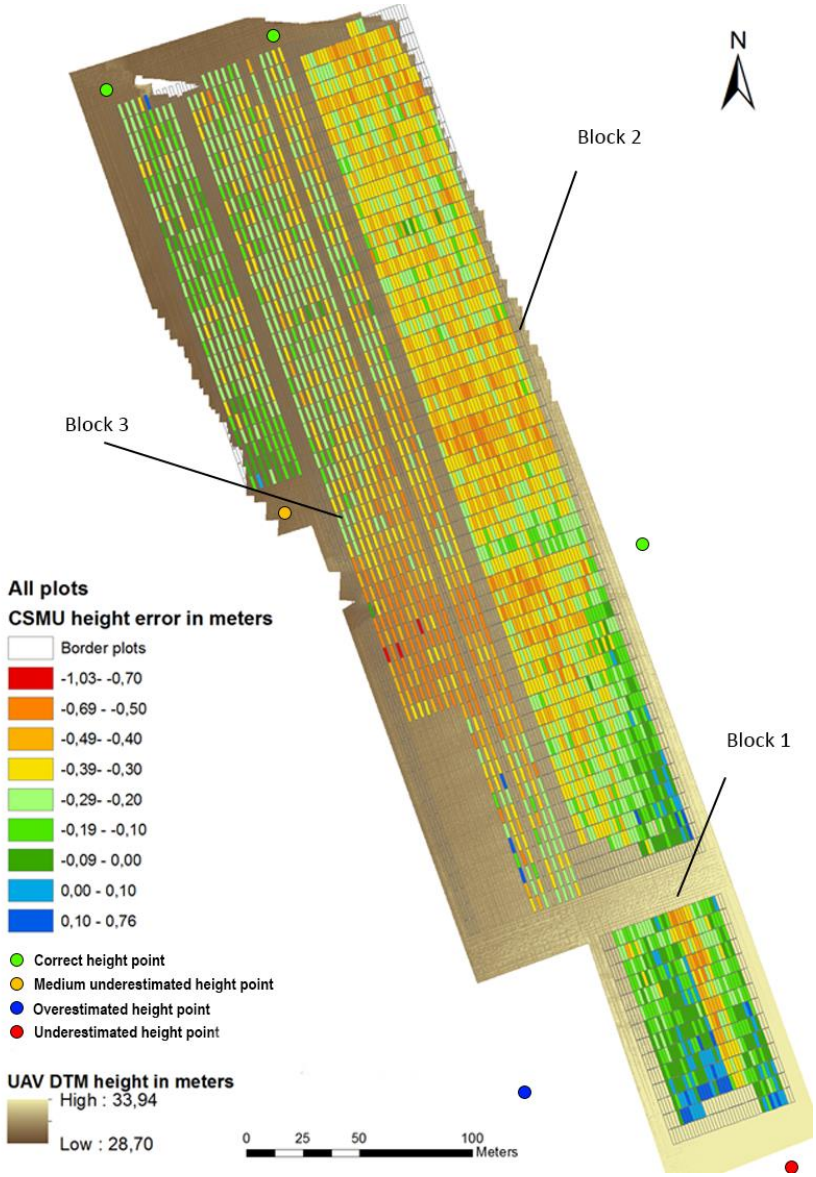
Figure 21 Map of AHN2 DTM and height errors for all plots based on difference between LiDAR and CSMA top of canopy (μ + 2σ). White line represents the profile line for Figure 1 (top). Boxplot and table of error range (bottom).

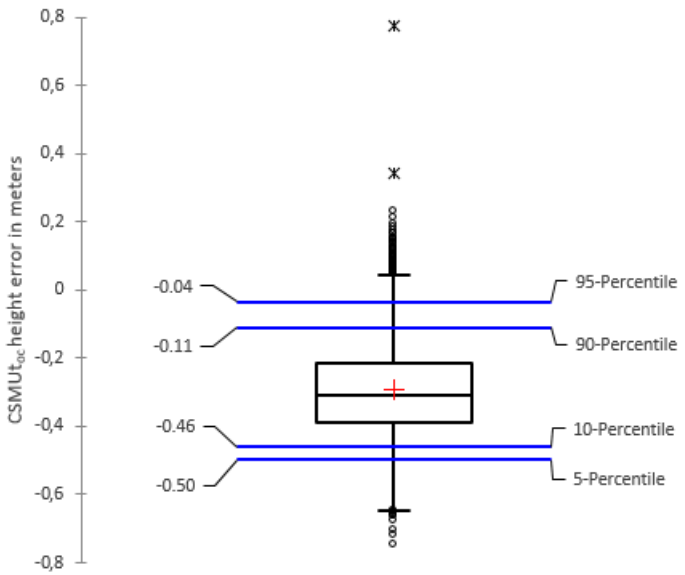
5.3.2 UAV DTM based CSM (CSMU)

To obtain a UAV based CSM (Annex H), a DTM was needed for the first flight date (26-05-2015). As there were no GCPs used during this first flight, the DTM obtained from the UAV based RGB sensor was created with the help of photogrammetric orientations, the GPS-INS data and extracted height points from the DSM of 22-09-2015 (Annex E). In total six of these height points were extracted from the DSM at locations with a known '0' value, such as tractor tracks (Figure 22).

Overall the height points had a low error when comparing the Z values from the DSM with the Z values from the DTM. Especially the three points in the North performed well, with an average error of just 0.003 meters. The point located in the middle left of the map was slightly underestimating in de DTM compared to the DSM values with a value of 0.05 meters. In the Southern part the points showed significantly more deviation from the DSM values with the point located at the bottom left overestimating with 0.30 meters and the point located at the bottom right underestimating with 0.24 meters. The effect of these deviations can be seen in the estimated plant height errors of that area. The plant height errors corresponded with the under or overestimation of the DTM compared to the DSM. Overall the DTM was significantly better in modeling the height differences over the terrain relative to the DSM. This can be seen in the more homogenous spread of errors, with a systematic underestimation of around -0.31 meters. As the extracted height points were coming from the DSM, and most points were located near the edges of the field, the interpolation effect could mean that the DTM was overall slightly overestimating. As we do not have any reference data the precise accuracy of the DTM could not be tested.

The plant height errors of the CSMU contained more outliers that overestimated than underestimated (Figure 22, bottom left). The outliers were located in the areas were the height points had the most uncertainty. Especially block 1 contained a significant amount of outliers. Block 3 with the four row set-up showed less underestimation than the two row set-ups of block 2. The four row plots had less disturbing effects from surrounding plots, which could affect the height estimation. In total 90% of the data underestimated, ranging from -0.50 meters to -0.04 meters, with 50% ranging from -0.39 to -0.22 meters (Figure 22, bottom right).





Percentile	*Height error*
Maximum 100%	0,78
99%	0,08
95%	-0,04
90%	-0,11
3rd Quartile 75%	-0,22
Median 50%	-0,31
1st Quartile 25%	-0,39
10%	-0,46
5%	-0,50
1%	-0,59
Minimum 0%	-0,75

Figure 22 Map of UAV DTM and height errors for all plots based on difference between LiDAR and CSMU top of canopy (μ + 2σ). (top). Boxplot and table of error range (bottom).

5.3.3 Comparing the CSMs

Overall the CSM derived from the UAV (CSMU) performed the best in terms of correlation to the LiDAR reference measurements (Table 3). Both thresholds had higher correlations compared to the AHN2 derived CSM (CSMA). The $CSMU_{toc}$ performed the best with an R^2 of 0.61. The $CSMU_{mean}$ had a slightly lower R^2 of 0.59. As the CSMU had a higher overall underestimation than the CSMA (Figure 23), the biases and RMSE were higher for the CSMU dataset. The $CSMU_{mean}$ dataset underestimated the most with a bias of -0.51 meters and a RMSE of 0.53 meters. The $CSMA_{mean}$ showed a similar underestimation with a bias of -0.43 meters and a RMSE of 0.45 meters. The top of canopy datasets of both CSMA and CSMU had the lowest underestimation with a bias of -0.22 and -0.26 meters respectively. This lower bias of the CMSA dataset can be explained due to the overestimation of a large area of the CSMA due to the slope angle differences (Figure 20). Furthermore the errors were less spread out for the CSMU datasets (Figure 23). This again points to a systematic underestimation for the CSMU, but with more predictable biases. All datasets had outliers, with the biggest outliers caused by overestimation. These outliers were not removed and thus influenced the correlation of the datasets.

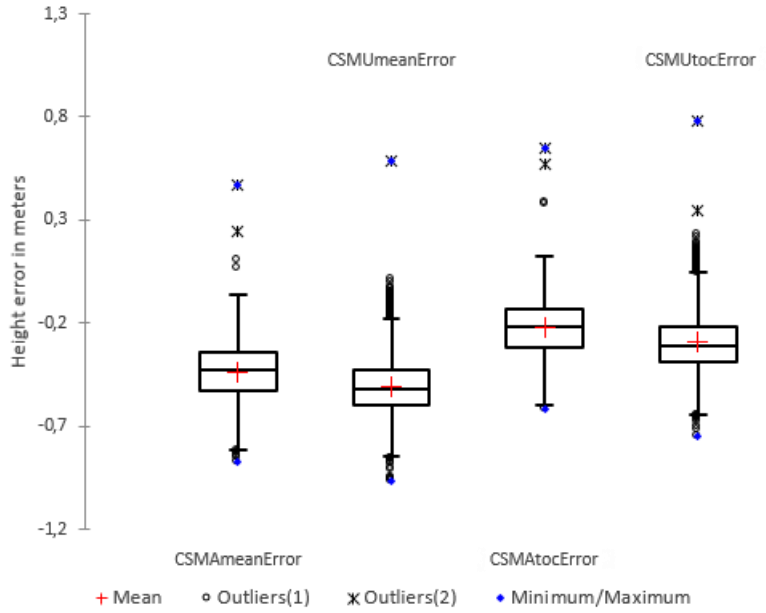


Figure 23 Boxplots showing height errors in meters and outliers of the different CSMs and thresholds (n = 2696)

Table 3 Comparison table of the different CSMs and thresholds compared to LiDAR measurements (n =2696)

Dataset	R^2	RMSE in m	BIAS in m	Threshold
$CSMA_{mean}$	0.53	0.45	-0.43	μ
$CSMA_{toc}$	0.57	0.26	-0.22	$\mu + 2\sigma$
$CSMU_{mean}$	0.59	0.53	-0.51	μ
$CSMU_{toc}$	0.61	0.33	-0.29	$\mu + 2\sigma$

5.4 COMPARING DIFFERENT MEASUREMENT METHODS

Additional to LiDAR, manual height measurements for 56 plots were made. These measurements were taken using a measurement stick and heights averaged for one plot. These plots were located in block 3 and had a four row set-up. For comparing the different measurement methods we used the CSMU.

The CMSU was more strongly correlated to the LiDAR height measurements than to the manual height measurements. The strongest correlation was found between the top of canopy threshold and the LiDAR measurements, with an R^2 of 0.78. The bias and RMSE for this correlation was -0.28 meters and 0.29 meters respectively. Although the correlations between the CSMU and manual height measurement were lower, the bias and RMSE of these correlations were also lower. The overall manual height was less high than similar LiDAR heights. The lowest bias was found between the top of canopy threshold and the manual measurements with -0.12 meters and a RMSE of 0.15 meters. Overall the mean thresholds showed higher biases and root mean square errors. This again shows the effect of in-between leaf and border pixels, which influence the mean height value of a plot. We can also state that the LiDAR techniques are better in capturing the spikey maize canopy, as the bias was always negative. It is notable that the manual methods also had a negative bias compared to the LiDAR measurements of -0.15 meters. In comparison, the correlation between LiDAR and manual measurements (R^2 of 0.73) was lower than the correlation with CSMU and manual measurements (R^2 of 0.78). Therefore, the CSMU performed better in terms of correlation to the manual measurements than the LiDAR dataset.

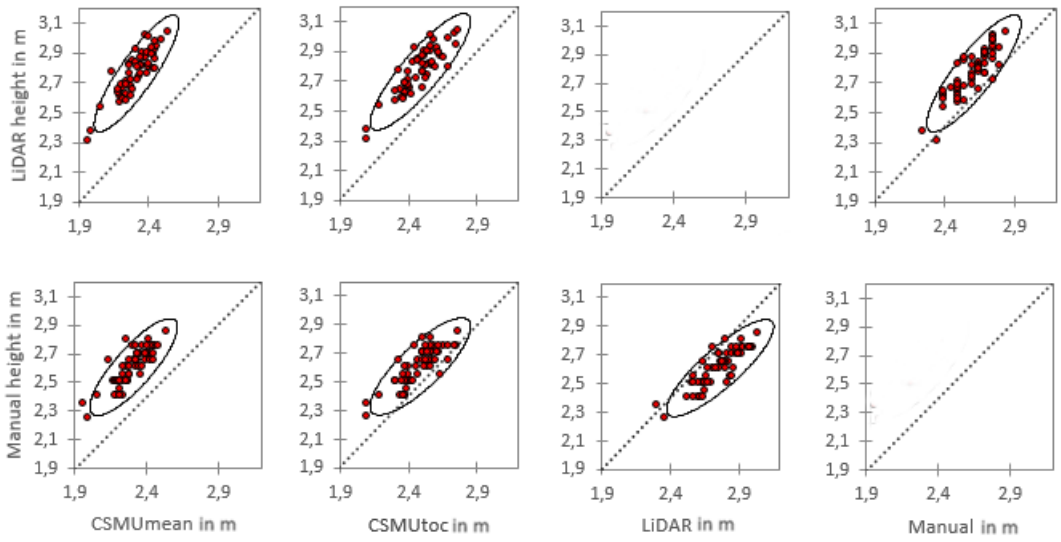


Figure 24 Correlations between manual and LiDAR measurements of maize height and CSM thresholds (n = 56)

Table 4 Coefficients of determinations and error statistics between different maize height measurement methods. RMSE and bias in meters. (n = 56)

Correlation	R^2	RMSE in m	BIAS in m	SEPc	Threshold
$CSMU_{mean}$ ~ LiDAR	0.76	0.46	-0.45	0,076	μ
$CSMU_{toc}$ ~ LiDAR	0.78	0.29	-0.28	0,075	$\mu + 2\sigma$
$CSMU_{mean}$ ~ Manual	0.69	0.30	-0.29	0,074	μ
$CSMU_{toc}$ ~ Manual	0.67	0.15	-0.12	0,084	$\mu + 2\sigma$
Manual ~ LiDAR	0.73	0.17	-0.15	0,079	

5.4 BIOMASS ESTIMATION

5.4.1 Biomass estimation using Crop Surface Model

Previous studies have found good predictions of biomass using crop surface models over the growing season (Bendig et al. 2015; J. Liu et al. 2009; Bendig et al. 2014). In this study we only used a crop surface model at harvest, which is not necessarily the best growing stage for biomass estimation based on crop height. However Bendig et al. (2015) found good correlations between biomass and the last two growing stages for barley. We applied correlation analysis between the measured biomass at harvest and several CSM derivatives such as; volume, surface area and the mean and top of canopy heights from the UAV based crop surface model. To make the results comparable with the estimation of biomass using hyperspectral data, we choose the 72 plots for which the vegetation indices were calculated. Furthermore, these 72 plots were located in block 3 resulting in lower geometrical errors compared to other blocks.

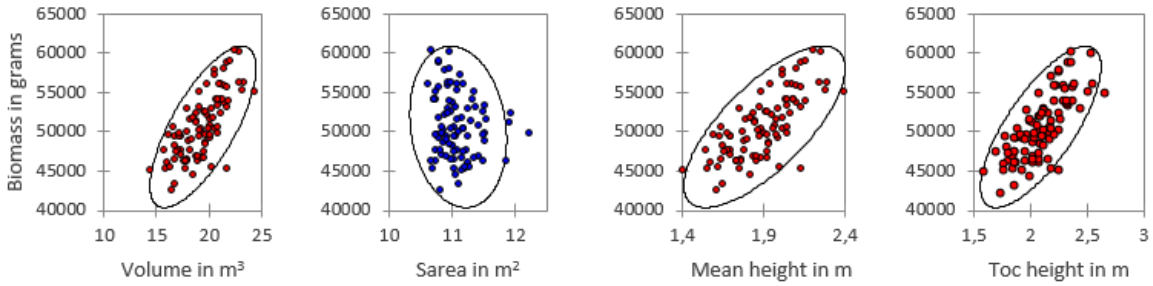


Figure 25 Correlations between biomass (in grams) and CSM derivatives (in meters) (n = 72).

Here we observed some significant correlations between biomass and the CSM derivatives. All but the surface area showed significant R^2 and p-values. The surface area did not have much significant variability in its values, due to the same plot area sizes and relative small differences in height. The p-value of the surface area was 0.14 with a non-significant R^2 of 0.02. The other derivatives scored better and all had significant p-values. The R^2 values for all derivatives where somewhat the same, ranging between 0.53 and 0.56. This shows that the CSM derivatives at harvest have significant prediction power for estimating biomass.

Table 5 Correlation and p-values between biomass and CSM derivatives (n = 72).

	Volume in m^3	*SArea in m^2*	*$CSMU_{mean}$ in m*	*$CSMU_{toc}$ in m*
R^2	0.53	0.02	0.53	0.56
p-values	**< 0,0001**	0,1408	**< 0,0001**	**< 0,0001**

5.4.2 Biomass estimation using vegetation indices

Using the hyperspectral data cubes four vegetation indices were calculated over time for all growing stages. The indices were calculated for June, July and September. Due to the geometrical errors caused by the hyperspectral push broom scanner (Figure 9, Methodology) only one flight line was used over time. From this flight line only plots from block 3 were chosen, as this were the four row set-up plots in order to minimize the effect of geometrical errors on the results. A total of 72 plots from block 3 were chosen, based on the geometrical correctness of the hyperspectral flight line. The hyperspectral sensor was capable of distinguishing the variability between plots. We applied a minimal noise fraction (MNF) analysis to the hyperspectral data cube to show this variability (Figure 26). This showed that it was possible to see spectral differences between the plots.

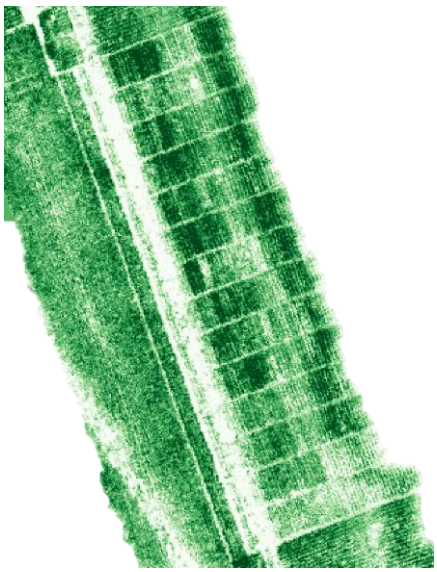


Figure 26 Minimal Noise Fraction of flight 4, flight line 1

5.4.3 Vegetation indices over time

The mean value for each index per growing stage showed a clear greening up stage, which saturated most indices. From July to September the index values lowered again (Figure 27). The hyperspectral indices, especially OSAVI, saturated the most. The multispectral indices showed less saturation. This was especially true for the WDVI due to the correction for soil background. All indices except for MTVI2 had lower index values in September compared to July. The MTVI2 seemed to have stayed relatively saturated with a difference of just 0.03 for that index.

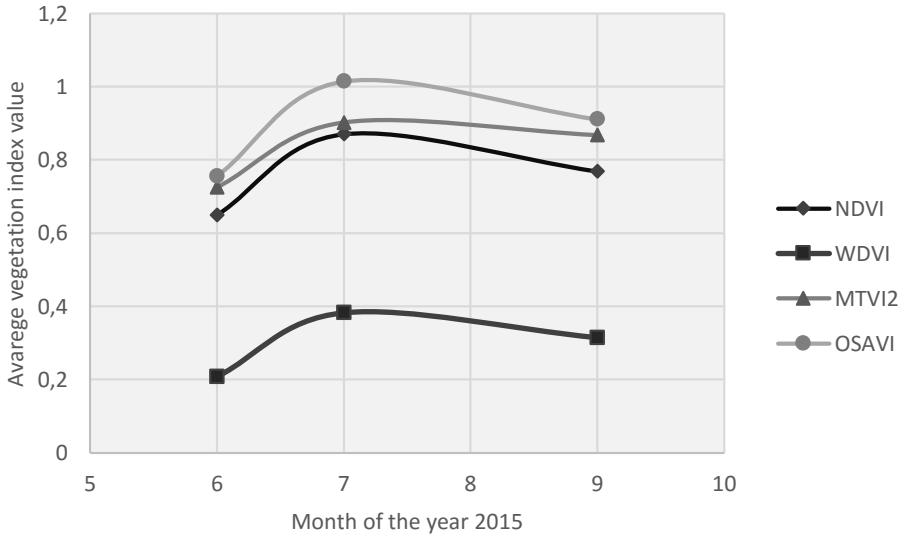


Figure 27 Average vegetation index value for 72 plots for each month (June, July and September)

5.4.4 Sensitivity over time

When looking at the variability for each index per growing stage we see that the most variation was found in June (Figure 28). Secondly September also showed significant variation for all indices. In July we see that almost all indices showed minimum variability due to the saturation effect. The WDVI was the only index that still had significant variability in July. This indicates that in June and September the maize cultivars inhabited more prediction capabilities than July.

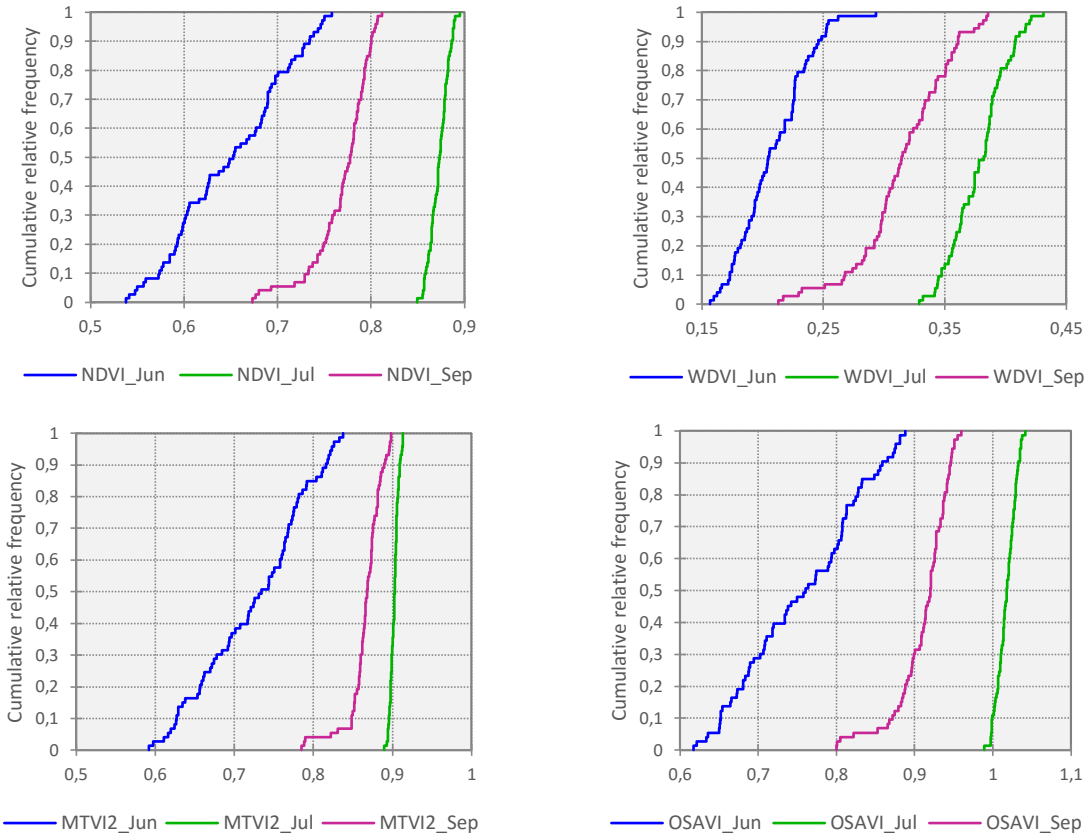


Figure 28 Cumulative relative frequency graphs for all indices showing their sensitivity over time. X-axis represents the vegetation index value

5.4.5 Correlation over time

The correlation for all vegetation indices varied over time, although almost all correlations were very weak (Table 6). In June no strong correlations were found with biomass. The best correlation in June was found using the MTVI2 with a R^2 of 0.08. Although all correlations in June were weak, all but the WDVI were significant with a p-value different from 0 with a significance level alpha=0.05. In June a lot of soil background was still visible and the correction for that in the WDVI could have caused the non-significance. In July the converse was true. Here all indices except for the WDVI were non-significant. The WDVI showed a significant correlation of $R^2 = 0.11$ with biomass in July. The non-significance of the other indices could be explained due to saturation and low sensitivity in this month. Especially slope-based indices like the NDVI are very susceptible for Leaf Area Index (LAI) and do get saturated quite fast. The WDVI is a distance-based index and was less sensitive to LAI. It stayed relative sensitive due to its correctness for soil background. In September all indices were highly significant ($p = < 0.0001$) and showed the strongest correlations. In September the WDVI again had the strongest correlation with

biomass with an R^2 of 0.50. Also the NDVI and OSAVI index performed quite well with an R^2 of 0.34 and 0.33 respectively. In this stage the MTVI2 performed worst with an R^2 of just 0.19. The MTVI2 stayed relative saturated (Figure 27) and did not show much sensitivity in September (Figure 28).

Table 6 Coefficient of determination table showing correlations between vegetation indices over time and biomass (n = 72).

Index	*NDVI Jun*	*WDVI Jun*	*MTVI2 Jun*	*OSAVI Jun*	*NDVI Jul*	*WDVI Jul*	*MTVI2 Jul*	*OSAVI Jul*	*NDVI Sep*	*WDVI Sep*	*MTVI2 Sep*	*OSAVI Sep*
R^2	0.07	0.05	0.08	0.06	0.01	0.11	0.03	0.01	0.34	0.50	0.19	0.33
p-value	**0.0249**	0.0513	**0.0170**	**0.0343**	0.4133	**0.0046**	0.1805	0.3800	**< 0.0001**	**< 0.0001**	**0.0001**	**< 0.0001**

5.4.3 Fusion of CSM derivatives and Vegetation indices for biomass prediction

When combining the CSM derivatives at harvest and the vegetation indices over time, the prediction performance of biomass improved significantly. Volume and height proofed to be good estimates of biomass with a significant correlation of around R^2 0.54, while the vegetation indices alone did not proof to be very good estimators at a single point in time.

The input in the correlated component regression were all 15 predictors over time for a total of 72 plots. 50 plots were used as a training and test set, using k-fold cross validation and 22 plots were used as a validation set. The model performed best with nine predictors in the k-fold cross validation tests. With more than nine predictors the model did not improve significantly anymore. Figure 29 shows how many times a predictor was taken into account in the cross validation steps with a total of 100 rounds with 8 folds. Here we clearly see the importance of the plant height and the vegetation indices in June and September for the prediction of biomass. All vegetation indices except MTVI2 in September and OSAVI in June were taken into account more than 600 times. The lowest prediction capabilities came from all the indices in July except WDVI and from the Surface area. Here we again see the good performance of the WDVI in July as it was taken into account 666 out of 800 times in the k-fold cross validation.

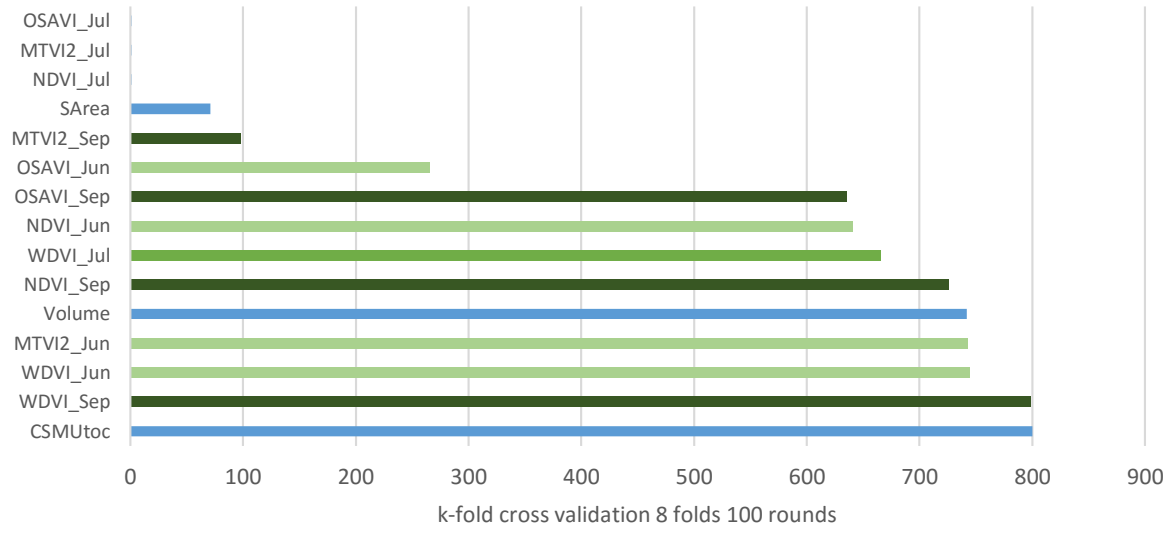


Figure 29 K-fold cross validation predictor count table

The standardized coefficients of the prediction model revealed a similar pattern (Figure 30). The standardized coefficients referred to how many standard deviations the dependent variables changed, per standard deviation increase in the predictor variable. Here we see that the CSM derivatives volume

and height had a greater effect on the dependent variable biomass than the vegetation indices. For the vegetation indices, the WDVI in September had the greatest effect on the dependent variable. This was in line with the k-fold cross validation predictor count table.

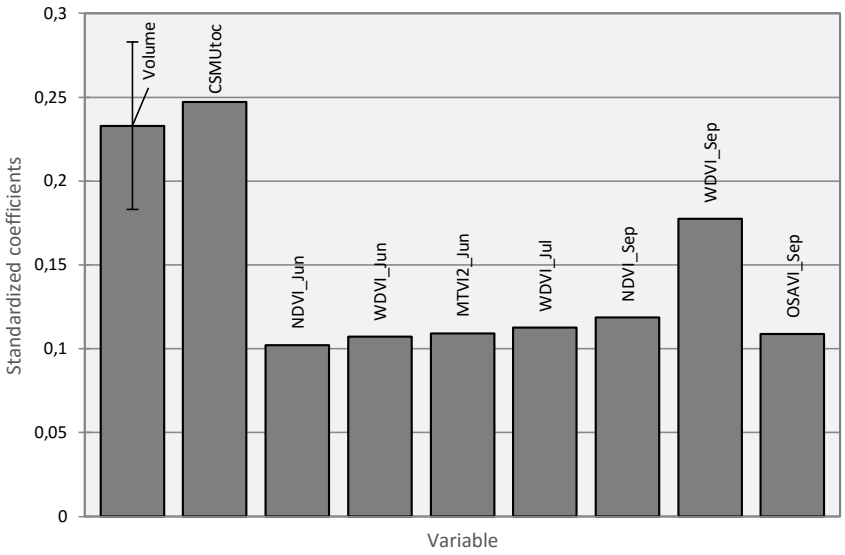


Figure 30 Standardized coefficients of predictors for the dependent variable biomass

For the final prediction model, nine predictors were used. When applying the model to the independent validation set of 22 plots a R^2 of 0.79 was achieved with a RMSE of 2.11 kilograms (Table 7). The average biomass for these plots was 50.62 kilograms per plot. Thus the error for the prediction model was around 4.2%. The prediction model was more accurate with lower biomass values than in the higher regions (Figure 31). When the biomass increases, the predicted biomass starts to deviate more from the 1:1 line. This could be an indication of a complex non-linear relationship. The CCR analysis could not model the complexity of a non-linear relationship that well. Furthermore, the underestimation of the biomass with higher values could point to a saturation effect.

Table 7 Goodness of fit statistics for final biomass prediction model

	Validation	*Cross-validation*	*Std. dev.(CV)*
Number of observations	23	50	
Sum of weights	23	50	
RMSE in grams	2114.12	2403.76	
NMSE	0.2320	0.2638	0,0203
R^2	0.79	0.71	0,0183

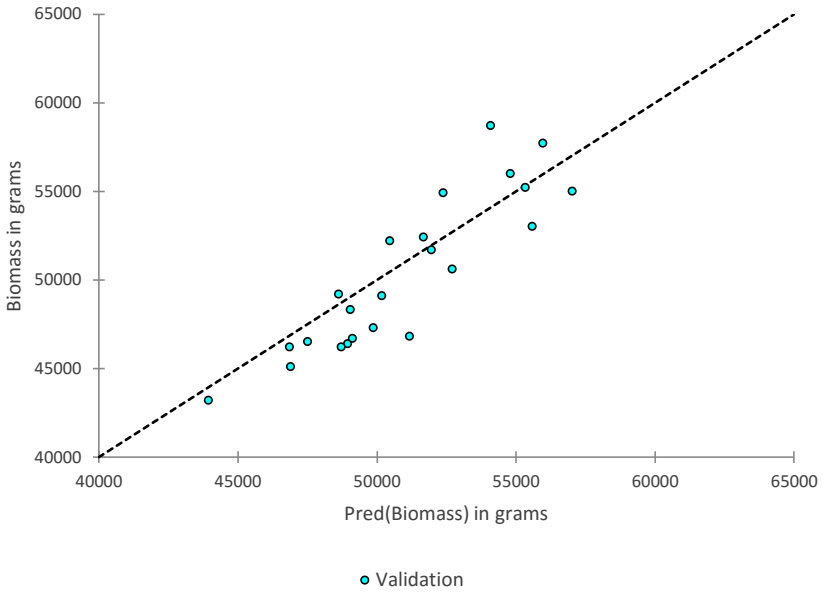


Figure 31 CCR biomass prediction model validation set

The difference in performance between the cross-validation and validation set was due to the higher amount of outliers in the training set (6) than the validation set (2). The average standardized residual for the training set was 0.8, while for the validation set this was 0.73 (Figure 32).

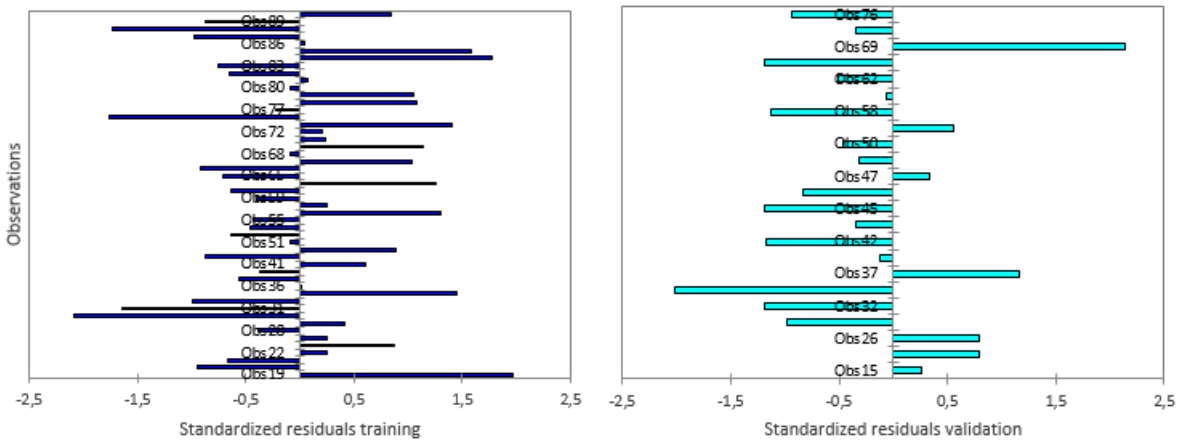


Figure 32 Standardized residuals for training and validation set

To compare the performance of the nine predictors CCR model, we used the same training and validation sets for different prediction models. We also looked at the type of sensor and what growing stage was required for the combination of predictors (Table 8). The CCR model outperformed all the other predictors and models. The CCR model with its predictors over time achieved an R^2 of 0.79 with an RMSE of 2.11 kilo. This is a significant improvement over the height and volume predictors, which achieve an R2 of 0.53 and 0.51 respectively for the same validation set. Also their RMSE errors are higher with an average error of 3.1 kilograms. The vegetation indices summed up over the growing season performed considerably better than when used at a single point in time. However even when summed up over the growing season, the vegetation indices performed the worst in comparison to the other models (Table 10). Notably are the hyperspectral indices which had the lowest correlations, especially the MTVI with an R^2 of just 0.26 and the biggest RMSE of 3.9 kilograms. The OSAVI index performed a

little better with an R^2 of 0.31 and a RMSE of 3.7 kilograms. The NDVI had a similar performance with a R^2 of 0.34 and a slightly lower RMSE of 3.6 kilograms. The best vegetation index predictor was the WDVI which attained a R^2 of 0.58 with a RMSE of 3.1 kilograms, which is in the similar range of height and volume.

When using the WDVI at harvest combined with height, the prediction capabilities improved to an R^2 of 0.74 with a RMSE of 2.4 kilograms. Adding the WDVI at the start of the growing season yielded an R^2 of 0.78 and a RMSE of 2.14, just slightly outperformed by the CCR model with nine predictors.

The Multi- and hyperspectral sensors did not prove to be good biomass estimators on its own, even when fused over time. The multispectral sensors outperform the hyperspectral sensors in this sense. The RGB sensors were capable of producing better products for biomass estimation, at a single point in time. However the best models used a combination of sensors and time. The best performance was achieved with a combination of the RGB and hyperspectral sensor over the growing season. Not far behind was the RGB and multispectral sensor, measured at before flowering and harvest growing stage. Lastly using a combination of RGB and multispectral sensors at harvest also produced significant prediction models.

Table 8 Performance table of different predictors and models for biomass estimation of maize crop

Model name	# Predictors	Validated R^2	RMSE in grams	Formula	Sensor	Growing stage
Height	1	0.53	3142.10	y = ax +b	RGB	At harvest
Volume	1	0.51	3099.92	y = ax +b	RGB	At harvest
Height + Volume	2	0.53	3123.03	y = ax +b	RGB	At harvest
NDVI	1 (sum)	0.34	3615.49	y = ax +b	Multispectral	Over growing season
WDVI	1 (sum)	0.58	3080.40	y = ax +b	Multispectral	Over growing season
MTVI2	1 (sum)	0.26	3927.51	y = ax +b	Hyperspectral	Over growing season
OSAVI	1 (sum)	0.31	3733.47	y = ax +b	Hyperspectral	Over growing season
Height + WDVI sep	2	0.74	2427.38	y = ax +b	RGB + Multispectral	At harvest
Height + WDVI jun + WDVI sep	3	0.78	2140.72	y = ax +b	RGB + Multispectral	Start and harvest
CCR Model	9	0.79	2114.12	y = ax +b	RGB + Hyperspectral	Over growing season

6 DISCUSSION

The purpose of this study was to evaluate the usability and accuracy of using UAV based optical sensors to derive products that could be used in large-scale phenotyping field trials. The main question of this study was to find out if the UAV platform, equipped with a RGB and hyperspectral sensor, is a suitable tool for plant breeders and at what growing stage the best results could be obtained. The focus was on height and biomass estimation at the end of the growing season for maize crops. We used RGB cameras to construct multi-temporal high resolutions ortho-photos and digital elevation models by implementing structure from motion computer vision algorithms. From these canopy architectural models, we derived crop surface models to estimate plant height. Furthermore, this study tested which spectral range in hyperspectral data cubes and growing stage had the strongest significant correlation with fresh biomass at harvest. This was done in order to find out which type of sensors were most suited and what flying times were essential for height and biomass estimation. Lastly, we fused the canopy architectural traits with the hyperspectral data to characterize biomass traits over time. We tested these new novel technologies in order to develop methods to fully utilize the capabilities of the UAV platform with high-resolution spectra cameras to create a state-of-the-art tool for plant breeders. We also analysed some of the disturbing factors that influenced the retrieval of trait values and its associated uncertainty.

6.1 ACCURACY OF PLANT HEIGHT

6.1.1 Assessing the digital surface model

The first research question was to test the accuracy of estimated plant height using UAV derived digital surface models. The RGB sensor that was operated together with a GNSS system proofed to be a robust tool for the 3D modelling of the field (Figure 13 and 15). The flight paths of the UAVs were overlapping around 90% along the flight path and around 50% sideways. The imagery that was obtained from the RGB sensor was a suitable input for the structure form motion algorithms. Highly detailed and georeferenced ortho-photo mosaics were obtained with a high spatial resolution of 1.6 centimetres. The digital surface model at harvest had a spatial resolution of 2.5 centimetres. In total five ground control points, which were measured with a RTK-GNSS system, were used in creating this digital surface model. To test the accuracy of the DSM compared to these GCPs, we extracted the DSM Z values at the same coordinates as the GCPs. Here we observed a deviation between four out of the five of the RTK-GNSS and the DSM height values (Figure 14). This implicates that the surface model deviated from the 'real-world' height at least four places. The Northern part of the study field was modelled lower and the Southern part higher compared to the RTK-GNSS measurements. The 'real-world' terrain also showed an upward slope from the North to the South. The accuracy of the final model depends on many factors, such as overlap between imagery and the shape of the object surface (Agisoft 2014). Possible non-linear deformations and misalignment errors of the model were partly removed by optimizing the estimated point cloud and camera parameters based on the measured GCPs. However the result of this optimization was based on the number and also positioning of the GCPs. Due to the positioning of the GCPs around the fields, and lack of GCPs inside the fields the DSM had no reference inside the field and therefore could not correct for the height variability of the terrain. Hence the skewed digital surface model. Agisoft recommends to use at least 10-15 GCPs that are evenly distributed over the field in order to generate accurate orhophoto's and digital elevation models (Agisoft 2014). Bendig et al. (2014) for example used 15 GCPs for 54 plots, mostly around plot corners. The total study field was just 0.10 ha, which is considerably smaller compared to our study field of 6.2 ha. They achieved an R^2 of 0.92 for estimating height of summer barley crops. This shows the considerably higher accuracy

when using a sufficient amount of GCPs. Even though our test field was larger in size, using more GCPs provides a more solid basis for the georeferencing and corrections of non-linear deformations.

Another method would be to leave out the GCPs entirely. Díaz-Varela et al. (2015) calculated tree heights from a Digital Surface Model without the use of GCPs or a DTM and obtained a relative RMSE from 6% to 20%. They based their heights on the identification of local maxima and the measurement of their difference in height to a reference altitude of the ground in their neighbourhood. For this reference altitude they used a circular neighbourhood of two meter around the trees to identify the minimum DSM values. However this method would be difficult to implement in dense canopy surfaces like maize as there won't be sufficient ground in the neighbourhood. Grenzdorffer (2014) developed a method that classified points obtained from the SfM dense point cloud into 'ground' and 'vegetation'. By differentiating these points, the relative height was obtained. This method proved to be less accurate, especially with higher leaf area index values as there were less reliable ground points. This is especially true for maize. In figure 16 we see that there were not a lot of ground pixels to be identified in between the plots. Grenzdorffer (2014) concluded that the classification of point cloud method was not suitable to determine the crop height for uniform arable crops and that the difference method using a DSM and DTM is simpler, faster and required significantly less expertise. Another problem that could arise when not using GCPs are non-linear deformations of the elevation model. During the georeferencing process in the Agisoft Photoscan workflow, the model is linearly transformed using seven parameter similarity transformation (three parameters for translation, three for rotation and one for scaling) (Agisoft 2014). This transformation can only improve a linear model misalignment. It is not possible to remove the non-linear component without the use of GCPs to optimize the point cloud. On top of that, the georeferencing would only be based on the GPS data from the UAV, which is significantly less accurate (> 5 meter) than the RTK-GNSS system (< 2 centimetres).

For non-linear transformations the amount of GCPs is based on the order of the polynomial transformation based on $((t+1)(t+2))/2$, where t = the order of the transformation. So in order to correct for more complicated image rectifications, a third or fourth order polynomial transformation is needed. Therefore we recommend to optimize the digital elevation models based on 10 to 15 GCPs, especially because high accuracy measurements are to be obtained from these models for thousands of plots. In our particular study area it would be difficult to use ground control points in the field around plots, as the maize vegetation was very dense at harvest and there was little to no spacing between the different plots. Nevertheless, more accurate modelling of the surface would have been achieved with the use of more GCPs. By leaving plots empty, GCPs could be placed inside the field and still be viewed from more than two angles. Another more permanent solution would be to place pole's that reach over the maize canopy surface that have known x, y and z coordinates. As our test field was approximately 500 meter in length and 140 meter wide, three rows with 45 meter width spacing and a placement of GCP at every 100 meter in length would be ideal to spread out 15 GCPs evenly over the field. The placement of the GCPs is most important at harvest, when the final plant height needs to be determined and before planting, to create a digital terrain model. However the DTM could also be based on the DSM by extracting ground points as is done in this study (Figure 8). Our method also yielded good results, even without the use of GCP for the DTM, as the DTM was modelled after ground points derived from the DSM.

6.1.2 Crop Surface Models

To derive a crop surface model with absolute plant heights we extracted digital terrain models from the surface models. We have used two different DTMs from different sources. The first DTM used was the

AHN2 dataset and the second was a DTM derived from UAV based RGB images acquired in the beginning of the growing season.

6.1.3 AHN2 based CSM

The AHN2 dataset is a LiDAR based terrain model. It was obtained in 2011 and has a spatial accuracy of 0.50 meters. The AHN2 datasets are generally seen as good models, but they do however contain some problems (van der Zon 2013). AHN2 datasets are not necessarily obtained when no vegetation is present. Therefore, points are classified into ground and vegetation points, from which subsequently the vegetation points get removed. Due to miss-classification vegetation is not always removed. Furthermore, the point density of the AHN2 dataset is just six to ten points per square meter from which the height grid is obtained. This is significantly less than the hundreds of points per square meter that were obtained with the SfM algorithms. In total the AHN2 dataset has an error margin of 15 centimetres for 95% of the dataset (van der Zon 2013).

The CSM derived from the AHN2 dataset showed a skewness error (Figure 19). The angle of the upward slope in the DSM was slightly higher than the angle for the AHN2 dataset. This resulted in a systematic underestimation of plant height in the Northern part and an increase in height further south (Figure 20 and 21). This could either be explained by a) an error in the AHN2 dataset, due to for example miss-classification, b) a change of the landscape over the time from 2011 to 2015, or c) the incapability of the DSM to accurately model the landscape due to the lack of sufficient amount of GCPs. As we do not have any reference data for 2011 we cannot test the accuracy of the AHN2 dataset. All three causes seem plausible, however the incapability of the DSM to accurately model the terrain seems most probable. This can be linked to the deviation from the DSM height compared to the ground points measured with the RTK-GNSS (Figure 14) and the lack of sufficient GCPs inside the field (Figure 7, left). This shows that when using data from other sources caution is advised as to whether changes have occurred over time or regarding the accuracy difference of both sources. Furthermore, resolutions for the DSM and AHN2 dataset differed greatly (Table 1). The AHN2 dataset had a 20 times courses resolution compared to the DSM. Díaz-Varela et al. (2015) showed that decreasing the spatial resolution of DEMs resulted in a loss of detail and had implications for the retrieval of plant height, even with trees. Grenzdorffer (2014) also emphasized the fact that if differences between plants are small, the level of detail and accuracy of both the DTM and DSM need to be sufficient in order to show subtle differences in height variability. This lower resolution of the AHN2 dataset thus effected the accuracy of the CSM. Due to the small-scale plots and the spikey canopy surface of maize, centimetre spatial resolutions are needed in order to capture this maize crop phenology. Overall the AHN2 based CSM showed a significant correlation to the LiDAR measured plant height (R^2 = 0.57), but due to the spatial differences in height values compared to the DSM and difference in resolution, it proved not to be an accurate method for estimating plant height at this scale.

6.1.4 UAV based CSM

In order to calculate a CSM based on UAV data we created a DTM using RGB imagery from the first flight date (26-05-2015). First of all, this dataset was a problematic due to a) the lack of GCPs, b) low temporal resolution of inflight GPS, c) low overlap of flight lines. Nevertheless, we used the RGB imagery together with extracted height points from the DSM to create a DTM in NAP heights (Annex E). In total six height points were extracted from the DSM with a known '0' value such as tractor tracks and used as markers in the DTM imagery (Figure 7, left). Accurately locating the points in the DTM imagery proofed to be quite a challenge as our only reference was the pattern of tractor tracks. Therefore, the DTM could not be precisely georeferenced, but did show a great level of detail (Figure 15). However,

as the height differences between plants were small, accurately georeferencing the DTM is necessary in order to contain these subtle differences in the CSM. Despite these difficulties only two out of six extracted height points in the DTM deviated from the DSM with around 0.30 meters under and overestimation (Figure 22, top). Again, as we did not have reference data for this date, the real accuracy of the DTM could not be assessed. The resulting CSM was a significantly better model than the AHN2 based CSM (Table 3, Figure 22). The plant height errors were more homogenous, with a systematic underestimation of around -0.31 meters. This indicates that this DTM was significantly better in modeling the height differences over the terrain relative to the DSM. However, as the extracted height points were coming from the DSM, and most points were located near the edges of the field, the interpolation effect could mean that the DTM was overall slightly overestimating. This could partly explain the systematic underestimation of -0.31 meters. Overall the UAV based CSM showed stronger correlations to the LiDAR height measurements (R^2 = 0.61) compared to the AHN2 based CSM (R^2 = 0.57; Table 3). Although the bias was higher, it was more systematic and thus easier to correct for (Figure 23). Our method of creating a DTM based on extracted height points from the DSM is basically a combination of classifying ground points in a DSM as proposed by Díaz-Varela et al. (2015) and subsequently the height differencing method as proposed by Bendig, Bolten, and Bareth (2013). This combination proved to be a good method, even with a small number of GCPs (Figure 22 and Table 3).

6.1.5 Accuracy of the CSM

The CSMs had a spatial resolution of 2.5 centimetres and shows a great level of detail (Figure 16, 17). Each plot contained around 16300 height points that were able to capture height variability within a plot with high detail (Figure 17). Also the spatial outlines of the plots were easily distinguishable from these elevation models. This proofed the capabilities of capturing plant height and its variability of a field using SfM algorithms. Only a small percentage of the 16300 height points per plot contained canopy pixels. The largest part, over 80% were leaf pixels not belonging to the canopy (Figure 18). The lowest pixels were caused by plot borders, especially on the width side of the plot as there was a small gap row between the widths of all plots (Figure 17). These non-top of canopy leafs and border pixels heavily influenced the mean height values for a plot (Figure 18). By taking the max height value, single tall plants or noise could determine the height of a grid cell. Therefore we used the method as proposed by Grenzdorffer (2014) to use the top 95% to 99% of the height data from a plot. Still an underestimation is expected due to the spikey canopy architectural traits of the maize plant. Grenzdorffer (2014) showed that crops with a more homogenous canopy surface like winter wheat and oilseed rape had less to no underestimation compared to spikey canopy surfaces of maize.

If we look at the phenology of the maize plant, we see that in the final growing stage the tassel of the plant is developed. This tassel is about 30 to 40 centimeters in height (Figure 33). The tassel is very susceptible for wind, especially during the last growing stage. These wind effects caused errors in the SfM algorithms as it relies on observing the same object from multiple angles through time. Furthermore, the relative area size of the tassel compared to the lower leaves was rather small and was therefore more difficult to capture in the imagery. The LiDAR datasets had a much higher point density compared to the point cloud generated from the DSM. Therefore, the LiDAR was better in capturing the tassel of the plant. The manual measurements did not take the tassel into account and took the upper leaf as maximum height. This partly explains the systematic underestimation of maize plant height for both UAV and AHN2 datasets compared to LiDAR and why the UAV height was more true to manual measurements.

Figure 33 Development of the tassel of a maize plant during the last growing stage (Ritchie, Hanway, and Benson 1993).

As the UAV based CSM had a more systematic underestimation with an average of -0.31 meters (Figure 22, bottom) this again seemed to be the more correct CSM.

6.1.6 Comparison with manual measurement methods

For 56 plots we were able to compare the CSM, LiDAR and manual height per plot (Figure 24, Table 4). Here it became clear that also the manual measurements were underestimating compared to the LiDAR measurements with an average of -0.15 meters. Although the biases and RMSE from the CSMU were lower for the manual measurements than the LiDAR measurements, the correlations were also lower. This could partly be explained due to the low variability of the manual height measurements, as a lot of plot heights were the same and the precision was only 0.05 meter in contrast to the 4 decimals precision of the CSM heights. The CSM heights had a stronger correlation to the LiDAR measurements than the manual height measurements had with LiDAR (Table 4). This points to either a) an overestimation of the LiDAR measurements, b) an underestimation of the manual height measurements, or c) a combination of both. Nevertheless, it proves that the CSM height values were more true to manual height measurement methods in terms of bias (Table 4) and that the underestimation of the CSM would be significantly lower when compared to manual measurements over the whole field. In similar researches, reference heights were always manual measurements (Bendig et al. 2014; Grenzdorffer 2014; Díaz-Varela et al. 2015; Jay et al. 2015; Yin et al. 2011). Comparison of CSM with LiDAR and manual measurements has not yet been done before. Therefore, this research provided valuable insides in the capabilities and differences between these methods.

6.1.7 Photoscan depth filtering

Depth filtering also explains part of the underestimation. In order to capture more details of the spikey canopy surface, depth filter settings should be evaluated. Photoscan calculates depth maps for every image. Due to factors like poor texture, noisy or badly focused images, there can be some outliers among the points. Depth filtering is applied to remove these outliers. The depth filtering of the 3D point cloud was set to 'mild', in order to obtain as much top of canopy pixels as possible from the spikey maize canopy. Depth filtering could also be turned completely off, however methods should then be developed in order to handle and classify real canopy points from noise. Further research could look into improving this depth filtering and classification for spikey canopy surfaces, however the wind and bi-directional effects will be inevitable.

6.2 BIOMASS PREDICTION USING CSMS AND VEGETATION INDICES OVER TIME

6.2.1 Biomass estimations based on the crop surface model

Our second research question was if the combination of vegetation indices and plant height could predict better estimates of above ground biomass for maize breeding trials compared to using only height. We first started by looking at the plant height, volume and surface area and see the correlation between these CSM derivatives and fresh biomass measured at harvest. Here we found significant correlations of these derivatives and biomass, with an R^2 of around 0.53 (Table 5). Previous studies found good predictions of biomass using crop surface models over the growing season (Bendig et al. 2015; J. Liu et al. 2009; Bendig et al. 2014). As we only had a CSM at harvest, we could not test the correlation between height and volume with biomass over the growing season. Bendig et al. (2014) found good correlations between biomass and height over the growing season for barley crops (R^2 = 0.81), which is significantly better than our correlation with height at harvest and biomass (R^2 = 0.56). The correlation of Bendig et al. (2014) was however exponential with height data at six growing stages during the season. Furthermore Bendig et al. (2014) estimated plant height from the CSM showed a correlation of R^2 = 0.92 to the manual reference height, which was a stronger correlation in comparison to our best findings (R^2 = 0.78). Their main limitation was the influence of lodging cultivars in the later growth stages, which produced irregular plant heights. Wind effects furthermore were a cause of error for irregular plant heights. Due to errors and wind effects in our CSM, there were also some irregular plant heights, which could partly explain the lower correlation. Bendig et al. (2014) also proved that height prediction errors became larger at later growing stages for barley and showed an increasing trend. This indicates that larger errors could occur in biomass prediction at later growing stages due to the increasing spatial variability. Therefore this could also have affected our results. We also calculated volume by generating a triangular irregular network of the CSM. This volume showed a significant correlation to biomass with an R^2 of 0.53 (Table 5). Using volumes from a CSM to predict biomass has not yet been explored much, however this study illustrated that volumes have significant prediction power for biomass. Further research into crop volumes derived either from a triangular irregular network, or the mesh directly from the 3D point cloud over the growing season should be developed in order to test these methods further.

6.2.2 Biomass estimations based on vegetation indices

Secondly we looked at the correlations between four different vegetation indices at three different growing stages and biomass (Table 6). We used two multispectral vegetation indices: the NDVI and WDVI and two hyperspectral vegetation indices: the MTVI2 and OSAVI. In total 12 vegetation index maps were calculated over the growing season. As most of the flight lines had substantial geometrical distortions (Figure 10, left) they were unsuitable as reliable measurements for these small-scale plots. Furthermore the resolution of these hyperspectral data cubes was much coarser at 25 centimeter spatial resolution (Table 1). Therefore small errors in the geometry of the hyperspectral data cubes implicated big errors on the results, especially though time. Next to that it was not possible to mosaic different flight lines due to changing atmospheric and illumination conditions between two flights. Based on these constrains only one flight line was chosen which showed the least distortion over time (Figure 10, right). This flight line was automatically projected on the DEM at pre-processing and later manually georeferenced to the ortho-photo with the highest precession possible. The result was a sufficient geometrical correct hyperspectral flight line for three flight dates. The geometrical accuracy was visually assed by comparing it to the high resolution ortho-photo. To further cancel out geometrical errors we choose to only use plots from the four row set-up. As the surrounding two rows of these plots were of the same cultivar, geometrical errors in the x, y direction had less of an influence (Figure 3). Plots located

at the borders of the hyperspectral flight line were removed, as the border areas were prone to calibration or bi-directional effect errors. In total 72 plots which were geometrically most correct were chosen. These plots are located together in block 3.

Minimal noise fraction (MNF) analysis that was applied to the hyperspectral flight lines proved that the spectral differences from the hyperspectral sensor was capable of distinguishing the variability between plots (Figure 26). These hyperspectral flight lines were used to calculate the four vegetation indices for June, July and September. This resulted in 12 vegetation indices maps. All indices showed a clear greening up stage from June to July and a decline from July to September (Figure 27). Except for the MTVI2, which seemed to stay relative saturated. When normalizing the four indices over time from 0 to 1, the MTVI2 declined 0.20 between July and September. In comparison, the other indices declined with a normalized average of around 0.42. Considering the saturation, the MTVI2 normally has shown low or no saturation to high LAI or biomass (Haboudane 2004) and also in our study did not saturate. However it showed the lowest sensitivity for the months with the highest LAI, July and September (Figure 28). The MTVI2 combines hyperspectral reflectance in the NIR, red and green wavelengths in order to reduce the effect of changing leaf chlorophyll content variation for crop green LAI estimation (Jiangui Liu, Pattey, and Jégo 2012). This low change in index value from July to September could partly be explained by the effects of suppressing chlorophyll.

Due to the saturation effect of the dense green canopy in July (Figure 27), almost all indices lost much of their sensitivity. Especially the two hyperspectral indices MTVI2 and OSAVI lost a lot of sensitivity in July (Figure 28). The NDVI, but especially the WDVI still showed significant variation in this month. This also relates to the correlation for the indices at a single point in time. Here only the WDVI had a significant correlation in July (Table 6). Overall all indices showed very weak correlations to biomass when correlated at a single point in time. VIs measure only the photosynthetically active components such as leafs. As the majority of the maize biomass is coming mostly from the stem and the kernels which are photosynthetically inactive components, vegetation indices are expected to have low correlations to the total biomass of the maize plant. Link, Senner, and Claupein (2013) found similar results when using an airborne spectrometer for prediction maize biomass. In the early growing stages they did not found any significant correlation using 730/740 and 760/730 ratio indices. After flowering of the maize, they found correlations of 0.34 between the reference biomass at harvest and the spectral data. This again illustrates that the relationship between biomass and spectral indices for maize is not very strong.

6.2.3 Fusion of CSM derivatives and Vegetation indices

To answer our last research question on which growing stages of maize are best for the prediction of biomass and which sensors are most suited, we tested several models in which we fused different products over time. To test these models we divided 72 plots for which biomass was measured into 50 plots for training and k-fold cross validation and 22 plots as a validation set. For this multivariate analysis we used the correlated component regression method. Deal (2011) showed that the correlated component regression is a suitable tool when the sample size is rather small and the number of predictors is large. It is a proven method to prevent overfitting and only takes significant predictors into account. Magidson (2010) stated that CCR outperforms comparable approaches when the number of predictors is fairly large.

As our first model, we used all 15 predictors as an input: height, volume, surface area, vegetation indices (12). The CCR model used 9 out of these 15 predictors for its optimal model, which is still rather large and could point to overfitting. However by implementing the k-fold cross validation and step-down

algorithm we were able to determine the most important predictors. In the k-fold cross validation predictor count table (Figure 30) it was evident that height and the WDVI at harvest were the most important predictors as they were taken into account 800 out of 800 times. Furthermore, volume and both the WDVI and MTVI2 in June were important. All indices in July, except the WDVI were not taken into account. This result clearly illustrates the importance of measurements at the beginning of the growing season and at harvest. During the greenest stage, only the multispectral index WDVI proved to be of significance. This outcome seems somewhat contradicting to the individual correlation of the vegetation indices with biomass at a single point in time. For example all the indices showed low correlation in June, but seemed to be important predictors in the correlated component regression. This can be explained that when examining correlations between two variables, only the unconditional (crude) associations between the variables were examined. With the correlated component regression model, we examined the joint associations of the independent variables (the predictors) with the dependent variable (biomass). For every predictor we looked at its association with biomass, conditional on the other predictors. This approach can yield very different, even contrary results to the unconditional correlations. The second output of this CCR regression were the standardized coefficients (Figure 31). These coefficients refer to how many standard deviations the dependent variable changed, per standard deviation increase in the predictor variable. The standardized coefficients for the nine predictors indicated that height and volume had a bigger effect on biomass than the vegetation indices (Figure 31). The WDVI in September was the vegetation index with the biggest effect. There is however some controversy on using standardized coefficients for comparing variables. Greenland et al. (1991) indicated that the standardized regression coefficients appear to simplify the interpretation of regression results as they are unit less quantities that therefore seem to be comparable. However they stated that this comparability is misleading. As distributions between the predictors varied, the meaning of a standard deviation also varied between the predictors. Nonetheless, when combining the correlations, sensitivity and k-fold cross validation results we can clearly state that measurements at harvest and at the beginning of the growing season are most important. The nine predictors' model yielded an R^2 of 0.79 with a RMSE of 2.11 kilograms (Table 7), which was an error of around 4.2%. The model showed more deviations with high biomass values, which could point to a) a complex non-linear relationship of the predictors and biomass or b) a saturation effect with higher values of biomass. The CCR is a linear model, which is not good at handling the complexity of a non-linear relationship (Magidson 2010). Furthermore saturation could have occurred in the predictors in the plots with higher amounts of biomass. The variation in the predictors could have decreased for plots with a higher amount of biomass, leading to less sensitivity. Chen et al. (2010) showed that the NDVI for example saturates at medium to high fresh biomass. This is around 2 kg/m^2 and as our average biomass was around 4.9 kg/m^2 at harvest, saturation effects could have occurred. The MTVI2 should have shown greater sensitivity to higher biomass values (Haboudane 2004), but did not prove to be an important predictor in our models.

Bendig et al. (2015) came to the same conclusion with respect to combining plant height with vegetation indices over time using multiple linear regression. In their study, the indices alone showed lower correlations over the growing season (MSAVI $R^2 = 0.22$) than combined with height (height + MSAVI $R^2 = 0.77$). This was for a barley crop, which has a considerably more homogenous canopy surface than maize. The correlations for plant height to the reference heights were very strong ($R^2 = 0.92$) and were taken over the growing season. Bendig et al. (2015) showed that for barley, height over the growing season was a better predictor for biomass ($R^2 = 0.80$) than combining it with vegetation indices. This indicates that plant height over the growing season is competitive with vegetation indices for biomass

estimation in barley. As the vegetation indices used in this study also had a significantly lower correlation to biomass than height with biomass, it would be interesting to use multiple CSMs over the growing season to test for better biomass prediction models. In this study, multiple CSMs could have been made over the growing season, however as there were no reference data available for the whole growing season, validation of these CSMs would not have been possible.

6.2.4 Comparing multivariate models over time

In order to see the differences in measurement times and sensors, we applied different models to the 50 training and 22 validation sets (Table 8). These models showed the lowest correlations and highest errors for the hyperspectral vegetation indices summed up over the growing season. The multispectral indices, especially the WDVI were significantly better in modeling the prediction of biomass. However due to the fact that most of the maize plants biomass is made up of photosynthetically inactive components, VIs alone are not sufficient as biomass predictors. Height and volume scored average with an average R^2 of 0.52. It was notable that combining height and volume did not improve the model. This indicates that height and volume were competitive. Nonetheless it was interesting to test the capabilities of volume as other similar studies did not look into this predictor. The best models used a combination of the RGB and multi- or hyperspectral sensor. Measurements with the RGB and multispectral sensor at harvest alone yielded good results ($R^2 = 0.74$). However, combing it with data from the beginning of the growing season ($R^2 = 0.78$) or the whole growing season ($R^2 = 0.78$) improved the model. These were expected results and proves that information over the growing season is essential for good biomass prediction models. Furthermore, overly complex models like the CCR model with nine predictors are not needed. The simple model with just three predictors at the start and at harvest using the RGB and multispectral sensor yielded near identical results, with a difference in R^2 of 0.01 and RMSE of 0.27 kilogram.

6.2.5 Limitations

The main limitations for these study were:

- The first flight date was flown for a different purpose than this study. The main focus for the first flight was intentionally to count plants. Therefore, the interval of the UAV miniature GPS-Inertial Navigation System were set to 10 seconds, instead of with every image. Furthermore, there was less overlap in the flight lines which resulted in gaps in de DTM and no GCPs were used.
- The flight of 22-09-2015 was flown at a later time than planned due to unforeseen circumstances with the UAV platform. This meant that the last flight was flown only two hours before harvest. This left little time for preparation and therefore only 5 GCPs could be placed. Also the distribution of these GCPs could have been more evenly.
- The hyperspectral push broom scanner produced too much geometrical distortions and warps for most of the flight lines to be used with small plots. Therefore, this type of scanner is not suited for phenotyping at this scale. Hyperspectral frame cameras like the Rikola camera currently available within UARSF should produce much better and stable results.
- During the flight at 22-09-2015 weather circumstances were not ideal as there were patches of clouds which influenced illumination circumstances during and between flights.
- Conductivity for the plots in block 3 varied due to very wet and dry areas in this block. This could have influenced correlations due to the higher amount of variation.

7 CONCLUSION AND OUTLOOK

7.1 CONCLUSION

In this research we evaluated the use of a UAV platform and different sensors for non-destructive phenotyping of small-scale maize breeding trials, and more specifically for estimating plant height and biomass of 3838 plots of maize (*Zea Mays)* cultivars. The research was designed to fulfill the needs of high-throughput and non-destructive field based phenotyping.

The RGB sensor was capable of creating highly detailed crop surface models with a spatial resolution of 2.5 centimeters. Each plot of 10.2 m^2 contained 16300 height pixels on average. This substantial amount of measurement points demonstrated that the height variability within small-scale plots was captured with high spatial detail. This spatial variability is interesting for plant breeding if the precision of the measurements are accurate enough. The accuracy of the CSM related to the LiDAR height measurements depended greatly on the accuracy of the elevation models and ground control points. For maize, a UAV derived DTM provided the best basis for deriving a CSM because this DTM was significantly better in modeling the height differences over the terrain relative to the DSM. The UAV based CSMU, using a top of canopy threshold proved to be the best estimator for plant height. Furthermore, the CSMU dataset outperformed manual measurements in terms of correlation to the LiDAR method. In comparison with the LiDAR measurements a systematic bias was found with an average spatial homogenous underestimation of -0.31 meters, using the top of canopy pixels per plot. The systematic bias was caused by a combination of wind effects and the incapability of accurately modeling the spikey canopy surface. The LiDAR measurements were better in capturing the spikey maize canopy, as the bias was always negative. Taken the systematic underestimation into account, the CSMU was able to estimate plant height with an average RMSE of 0.30 meters.

For biomass estimations the CSMU at harvest was able to yield a correlation of $R^2 = 0.53$ to biomass with a RMSE of 3.14 kilograms (6.2% error) using just height data. Combining the CSMU with vegetation indices over the growing season using a multivariate approach, significantly improved this correlation to an R^2 of 0.79 with an RMSE of 2.11 kilograms (4.2% error). However, the geometric accuracy of the hyperspectral images influenced the results. Geometrical distortions caused by the push broom scanner were of such nature that correcting for that was difficult, or not possible at al.

The growing stages that showed the highest significant correlations were the start, but mostly the end of the growing season. In September all indices were highly significant ($p = < 0.0001$) and showed the strongest correlations. Due to the saturation effect of a dense green canopy during the middle of the growing season, this stage saturated most indices and gave the lowest and most non-significant correlations. Al together all correlations were rather weak, but did show a significant amount of variance in June and September. Overall the multispectral indices outperformed the hyperspectral indices in both the multivariate approach as the individual correlations. The multispectral WDVI demonstrated to be the most significant VI biomass predictor for maize, especially at harvest.

Overall, our results suggest that plant height measured at harvest, combined with vegetation indices at the start and end of the growing season may be used to predict maize biomass. This prediction can be improved by increasing the spatial accuracy of the crop surface models and using multi- or hyperspectral frame cameras in order to minimize geometrical distortions. To improve spatial accuracy we recommend the use of 10 to 15 GCPs, based on the order of the non-linear polynomial transformation that is needed for the rectifications of 3D models. The use of stable hyperspectral frame cameras should be explored with regard to high precision hyperspectral mapping. For this, we recommend the Rikola hyperspectral frame camera currently available within UARSF as this camera should produce much better and stable results.

7.2 OUTLOOK

To fully implement the UAV platform as a suitable tool for plant breeders depends on the development of the UAV platform regarding stability, control, reliability, positioning, autonomy and improvements of sensors (Sankaran et al. 2015). As remote sensing techniques and UAV platforms are improving rapidly we firmly believe that this improved technology will help to add new methods for aerial field-based phenomics in the near future. Battery capacities and aerodynamics are rapidly improving and allow for longer and safer flying times. These improvements will also lead to an increase of payload capacity, allowing for more heavy and advanced sensors. Furthermore regulations will improve, with reserved airspace for UAV platforms which will be monitored and will have an airplane-like collision avoidance algorithm for UAVs sharing three-dimensional airspace. Such systems are already being implemented (Aerialtronics 2014). Collision avoidance is a key factor in enabling the integration of UAVs into real life use (Pham et al. 2015). Fully implemented collision avoidance systems would give room for programmed and autonomous UAV flights, without the need of being in line-of-sight. This technique would furthermore make it possible for UAVs to operate in swarms in order to cover large fields, or divide multiple tasks.

Technological advances in UAV built-in GNSS systems are going towards real time kinematic systems which provide much higher accuracy of the UAVs location and will also improve the structure from motion outputs. Producing more accurate elevation models, even without the use of ground control points. Experiments are also done with digital ground control points, so called 'beacons', that can automatically communicate with the UAV and send its x, y and z coordinates (Flonew 2015). This technique is still in its early initial phase, but nonetheless shows the potential impact beacon technology could have on UAV navigation in the future.

New state-of-the-art hyperspectral frame cameras are already used at the Unmanned Aerial Facility from Wageningen UR. These hyperspectral frame cameras produce much more stable images with a higher spatial resolution, compared to the more traditional push broom scanner. Geometrical distortion and warps are much less of a factor in frame cameras and stable hyperspectral ortho-photos could be produced of a whole study field. Furthermore, with the advancement of battery capacities and aerodynamics UAVs are now capable of carrying a LiDAR payload. This technology gives room for high precision aerial 3D monitoring, which is an essential aspect of field-based phenomics. Another advantage of these improved battery capacities is the increased duration of a single flight. As of now, five separate flights had to be made in order to capture the whole field. By being able to fly a test field at once, this time span is shortened which is needed in order to compare traits as illumination conditions vary over time.

The research presented in this thesis has opened a number of research opportunities that should be explored in the near-future. The main opportunity would be to produce higher accuracy 3D maps over the growing season using new techniques such as an airborne LiDAR or improved GNSS systems to better predict biomass and plant growth. In summary, the UAV platform combined with the right sensors has great potential for high-throughput field-based phenomics. It can be the 'on size fits all' solution to address the current phenotyping bottleneck. A multitude of sensors can be equipped on these platforms that can fly at low-altitude in a relative small timespan without affecting field conditions. With the current speed of advancement of these technologies we firmly believe in the exponential growth of UAV applications, including field-based phenomics, in the near future.

8 REFERENCES

Aerialtronics. 2014. "Aerialtronics Technology Improves Safety." http://www.aerialtronics.com/2014/08/aerialtronics-revolutionarily-improves-safety/.

Agisoft. 2014. "Agisoft PhotoScan User Manual: Professional Edition, Version 1.1." http://www.agisoft.com/pdf/photoscan-pro_1_1_en.pdf.

Andrade-Sanchez, Pedro, Michael A. Gore, John T. Heun, Kelly R. Thorp, A. Elizabete Carmo-Silva, Andrew N. French, Michael E. Salvucci, and Jeffrey W. White. 2014. "Development and Evaluation of a Field-Based High-Throughput Phenotyping Platform." *Functional Plant Biology* 41 (1): 68–79. doi:10.1071/FP13126.

Bendig, Juliane, Andreas Bolten, and Georg Bareth. 2013. "UAV-Based Imaging for Multi-Temporal, Very High Resolution Crop Surface Models to Monitor Crop Growth Variability." *Photogrammetrie - Fernerkundung - Geoinformation* 2013 (6). E. Schweizerbart'sche Verlagsbuchhandlung: 551–562. doi:10.1127/1432-8364/2013/0200.

Bendig, Juliane, Andreas Bolten, Simon Bennertz, Janis Broscheit, Silas Eichfuss, and Georg Bareth. 2014. "Estimating Biomass of Barley Using Crop Surface Models (CSMs) Derived from UAV-Based RGB Imaging." *Remote Sensing* 6 (11). Multidisciplinary Digital Publishing Institute: 10395–10412. doi:10.3390/rs61110395.

Bendig, Juliane, Kang Yu, Helge Aasen, Andreas Bolten, Simon Bennertz, Janis Broscheit, Martin L. Gnyp, and Georg Bareth. 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 (2015). Elsevier B.V.: 79–87. doi:10.1016/j.jag.2015.02.012.

Boehlje, Michael, and Stefanie Broring. 2010. *On the Increasing Multifunctionality of Agricultural Raw Materials : Three Dilemmas for Innovation and Adoption*. Wageningen University.

Chaoyang, Wu, Niu Zheng, Quan Tang, and Huang Wenjiang. 2008. "Estimating Chlorophyll Content from Hyperspectral Vegetation Indices: Modeling and Validation." *Agriculturl and Forest Meteorlogy* 148: 1230–1241.

Chapman, Scott, Torsten Merz, Amy Chan, Paul Jackway, Stefan Hrabar, M. Dreccer, Edward Holland, Bangyou Zheng, T. Ling, and Jose Jimenez-Berni. 2014. "Pheno-Copter: A Low-Altitude, Autonomous Remote-Sensing Robotic Helicopter for High-Throughput Field-Based Phenotyping." *Agronomy* 4 (2): 279–301. doi:10.3390/agronomy4020279.

Chen, Peng-Fei, Tremblay Nicolas, Ji-Hua Wang, Vigneault Philippe, Wen-Jiang Huang, and Bao-Guo Li. 2010. "New index for crop canopy fresh biomass estimation." *Guang pu xue yu guang pu fen xi = Guang pu* 30 (2). China: 512–517.

Cleland, John. 2013. "World Population Growth; Past, Present and Future." *Environmental and Resource Economics* 55 (4): 543–554. doi:10.1007/s10640-013-9675-6.

Clevers, J.G.P.W. 1989. "The Application of a Weighted Infrared-Red Vegetation Index for Estimating Leaf Area Index by Correcting for Soil Moisture." *Remote Sensing of Environment* 29: 25–37.

Clevers, J.G.P.W. 1991. "Application of the WDVI in Estimating LAI at the Generative Stage of Barley." *ISPRS Journal of Photogrammetry and Remote Sensing* 46: 37–47.

Deal, Ken. 2011. "Conducting Correlated Component Regression Using COREXpress." *Software Review*.

Dellaert, Frank, Steven M Seitz, Charles E Thorpe, and Sebastian Thrun. 2000. "Structure from Motion without Correspondence." *Proceedings of the {IEEE} Conference on Computer Vision and Pattern*

Recognition 2: 557–564. doi:10.1109/CVPR.2000.854916.

Díaz-Varela, Ramón, Raúl de la Rosa, Lorenzo León, and Pablo Zarco-Tejada. 2015. "High-Resolution Airborne UAV Imagery to Assess Olive Tree Crown Parameters Using 3D Photo Reconstruction: Application in Breeding Trials." *Remote Sensing* 7 (4): 4213–4232. doi:10.3390/rs70404213.

Ehlert, Detlef, Rolf Adamek, and Hans Juergen Horn. 2009. "Laser Rangefinder-Based Measuring of Crop Biomass under Field Conditions." *Precision Agriculture* 10 (5): 395–408. doi:10.1007/s11119-009-9114-4.

EPPN. 2013. "Why Do We Need Plant Phenotyping?"

Flonew. 2015. "FloNew." http://flonew.com/.

Furbank, Robert T, and Mark Tester. 2011. "Phenomics--Technologies to Relieve the Phenotyping Bottleneck." *Trends in Plant Science* 16 (12): 635–644. doi:10.1016/j.tplants.2011.09.005.

Grayson, Michelle. 2011. "Biofuels." *Nature* 474 (7352). Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.: S1–S1.

Greenland, Sander, Malcolm Maclure, James J Schlesselman, Charles Poole, and Hal Morgenstern. 1991. "Standardized Regression Coefficients: A Further Critique and Review of Some Alternatives." *Epidemiology* 2 (5). Lippincott Williams & Wilkins: 387–392.

Grenzdorffer, G. J. 2014. "Crop Height Determination with UAS Point Clouds." *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* 40 (1): 135–140. doi:10.5194/isprsarchives-XL-1-135-2014.

Haboudane, D. 2004. "Hyperspectral Vegetation Indices and Novel Algorithms for Predicting Green LAI of Crop Canopies: Modeling and Validation in the Context of Precision Agriculture." *Remote Sensing of Environment* 90 (3): 337–352. doi:10.1016/j.rse.2003.12.013.

Haboudane, D., J. R. Miller, N. Treblay, P. J. Zacro-Tejeda, and L. Dextraze. 2002. "Integrated Narrow-Band Vegetation Indices for Prediction of Crop Chlorophyll Content for Application to Precision Agriculture." *Remote Sensing of Environment* 81: 416–426.

Harvey, Phil. 2010. "ExifTool by Phil Harvey."

Huete, A. R., H. Q. Liu, K. Batchily, and W. van Leeuwen. 1997. "A Comparison of Vegetation Indices over a Global Set of TM Images for EOS-MODIS." *Remote Sensing of Environment* 59: 440–451.

Hunt, E. Raymond, Michel Cavigelli, Craig S. T. Daughtry, James E. Mcmurtrey, and Charles L. Walthall. 2005. "Evaluation of Digital Photography from Model Aircraft for Remote Sensing of Crop Biomass and Nitrogen Status." *Precision Agriculture* 6 (4): 359–378. doi:10.1007/s11119-005-2324-5.

Inoue, Y., S. Morinaga, and A. Tomita. 2000. "A Blimp-Based Remote Sensing System for Low-Altitude Monitoring of Plant Variables: A Preliminary Experiment for Agricultural and Ecological Applications." *International Journal of Remote Sensing* 21 (2). Taylor & Francis Group: 379–385. doi:10.1080/014311600210894.

Jay, Sylvain, Gilles Rabatel, Xavier Hadoux, Daniel Moura, and Nathalie Gorretta. 2015. "In-Field Crop Row Phenotyping from 3D Modeling Performed Using Structure from Motion." *Computers and Electronics in Agriculture* 110 (January): 70–77. doi:10.1016/j.compag.2014.09.021.

Jensen, T., A. Apan, F. Young, and L. Zeller. 2007. "Detecting the Attributes of a Wheat Crop Using Digital Imagery Acquired from a Low-Altitude Platform." *Computers and Electronics in Agriculture* 59 (1-2): 66–77. doi:10.1016/j.compag.2007.05.004.

Jin, Yunxiang, Xiuchun Yang, Jianjun Qiu, Jinya Li, Tian Gao, Qiong Wu, Fen Zhao, Hailong Ma, Haida Yu, and Bin Xu. 2014. "Remote Sensing-Based Biomass Estimation and Its Spatio-Temporal Variations in Temperate Grassland, Northern China." *Remote Sensing* 6 (2): 1496–1513. doi:10.3390/rs6021496.

Jr, E Raymond Hunt, Raymond Hunt, and E T Al. 1999. "Evaluation of Digital Photography from Model Aircraft for Remote Sensing of Crop Biomass and Nitrogen Status," 359–378.

Kross, Angela, Heather McNairn, David Lapen, Mark Sunohara, and Catherine Champagne. 2015. "Assessment of RapidEye Vegetation Indices for Estimation of Leaf Area Index and Biomass in Corn and Soybean Crops." *International Journal of Applied Earth Observation and Geoinformation* 34 (2015). Elsevier B.V.: 235–248. doi:10.1016/j.jag.2014.08.002.

Lati, Ran Nisim, Sagi Filin, and Hanan Eizenberg. 2013a. "Plant Growth Parameter Estimation from Sparse 3D Reconstruction Based on Highly-Textured Feature Points." *Precision Agriculture* 14 (6): 586–605. doi:10.1007/s11119-013-9317-6.

Lati, Ran Nisim, Sagi Filin, and Hanan Eizenberg. 2013b. "Estimating Plant Growth Parameters Using an Energy Minimization-Based Stereovision Model." *Computers and Electronics in Agriculture* 98 (2013). Elsevier B.V.: 260–271. doi:10.1016/j.compag.2013.07.012.

Lelong, Camille C. D. 2008. "Assessment of Unmanned Aerial Vehicles Imagery for Quantitative Monitoring of Wheat Crop in Small Plots." *Sensors* 8 (5): 3557–3585. doi:10.3390/s8053557.

Link, J., D. Senner, and W. Claupein. 2013. "Developing and Evaluating an Aerial Sensor Platform (ASP) to Collect Multispectral Data for Deriving Management Decisions in Precision Farming." *Computers and Electronics in Agriculture* 94 (June): 20–28. doi:10.1016/j.compag.2013.03.003.

Liu, J., E. Pattey, J. Shang, S. Admira, G. Jégo, H. McNairn, A. Smith, B. Hu, F. Zhang, and J. Freementle. 2009. "Quantifying Crop Biomass Accumulation Using Multi-Temporal Optical Remote Sensing Observations." *Proceedings of the 30th Canadian Symposium on Remote Sensing*, 22–25.

Liu, Jiangui, Elizabeth Pattey, and Guillaume Jégo. 2012. "Assessment of Vegetation Indices for Regional Crop Green LAI Estimation from Landsat Images over Multiple Growing Seasons." *Remote Sensing of Environment* 123 (August): 347–358. doi:10.1016/j.rse.2012.04.002.

Magidson, Jay. 2010. "Correlated Component Regression : A Prediction / Classification Methodology for Possibly Many Features." *Training*, 1–19.

Mroz, Marek, and Anna Sobieraj. 2007. "Comparison of Several Vegetation Indices Calculcated on the Basis of a Seasonal SPOT XS Time Series, and Their Suitability for Land Cover." *Technical Sciences* 7: 40–66.

Oerke, Erich-Christian, Roland Gerhards, Gunter Menz, and Richard A. Sikora. 2010. *Precision Crop Protection : The Challenge and Use of Heterogeneity*. Springer.

Panguluri, Siva Kumar, and Are Ashok Kumar. 2013. *Phenotyping for Plant Breeding*. New York, NY: Springer New York.

Peng, Yi, and Anatoly A. Gitelson. 2011. "Application of Chlorophyll-Related Vegetation Indices for Remote Estimation of Maize Productivity." *Agricultural and Forest Meteorology* 151 (9): 1267–1276. doi:10.1016/j.agrformet.2011.05.005.

Pham, Hung, Scott a Smolka, Scott Stoller, Dung Phan, and Junxin Yang. 2015. "A Survey on Unmanned Aerial Vehicle Collision Avoidance Systems." *Department of Computer Science*.

Pieruschka, Roland, and Hendrik Poorter. 2012. "Phenotyping Plants: Genes, Phenes and Machines." *Functional Plant Biology* 39 (11). CSIRO Publishing: 813. doi:10.1071/FPv39n11_IN.

Pimentel, David, and Marcia Pimentel. 2006. "Global Environmental Resources versus World Population Growth" 9: 9–12.

Pudelko, Rafal, Tomasz Stuczynski, and Magdalena Borzecka-Walker. 2015. "The Suitability of an Unmanned Aerial Vehicle (UAV) for the Evaluation of Experimental Fields and Crops." *Zemdirbyste-Agriculture* 99 (4). Lithuanian Research Centre Agriculture & Forestry, Alexandras Stulginskis Univ, Kedainiu r, lt-58344, Lithuania: 431–436. Accessed September 15.

Qi, J, A Chehbouni, A R Huete, Y H Kerr, and S Sorooshian. 1994. "A Modified Soil Adjusted Vegetation Index." *Remote Sensing of Environment* 48 (2): 119–126. doi:http://dx.doi.org/10.1016/0034-4257(94)90134-1.

Ramming, David W, and Carlos D Fear. 1993. "Dedication: John H. Weinberger—Fruit Breeder and Horticulturalist." In *Plant Breeding Reviews*, 1–10. John Wiley & Sons, Inc. doi:10.1002/9780470650035.ch1.

Ritchie, Steven, John Hanway, and Garren Benson. 1993. "How a Corn Plant Develops." *Special Report Iowa State University of Science and Technology*, no. No. 48.

Rondeaux, G., M. Steven, and 1 Baret, F. 1996. "Optimization of Soiladjusted Vegetation Indices." *Remote Sensing of Environment* 55: 95–107.

Rouse, J. W., R. H. Haas, J. A. Schell, and D. W. Deering. 1974. "Monitoring Vegetation Systems in the Great Plains with ERTS." *NASA Goddard Space Flight Center 3d ERTS-1 Symposium*, pp. 309–317.

Sankaran, Sindhuja, Lav R. Khot, Carlos Ziga Espinoza, Sanaz Jarolmasjed, Vidyasagar R. Sathuvalli, George J. Vandemark, Phillip N. Miklas, 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.

Santos, Thiago T, and Alberto A De Oliveira. 2012. "Image-Based 3D Digitizing for Plant Architecture Analysis and Phenotyping." In *Workshop on Industry Applications (WGARI) in SIBGRAPI 2012 (XXV Conference on Graphics, Patterns and Images)*. doi:10.13140/2.1.4576.1608.

Sims, D.A., and J.A. Gamon. 2002. "Relationships between Leaf Pigment Content and Spectral Reflectance across a Wide Range of Species, Leaf Sturucres and Developmental Stages." *Remote Sensing of Environment* 81: 337–354.

Sugiura, Ryo, Noboru Noguchi, and Kazunobu Ishii. 2005. "Remote-Sensing Technology for Vegetation Monitoring Using an Unmanned Helicopter." *Biosystems Engineering* 90 (4): 369–379. doi:10.1016/j.biosystemseng.2004.12.011.

Suomalainen, Juha, Niels Anders, Shahzad Iqbal, Gerbert Roerink, Jappe Franke, Philip Wenting, Dirk Hünniger, Harm Bartholomeus, Rolf Becker, and Lammert Kooistra. 2014. "A Lightweight Hyperspectral Mapping System and Photogrammetric Processing Chain for Unmanned Aerial Vehicles." *Remote Sensing* 6 (11). Multidisciplinary Digital Publishing Institute: 11013–11030. doi:10.3390/rs61111013.

Tester, Mark, and Peter Langridge. 2010. "Breeding Technologies to Increase Crop Production in a Changing World." *Science* 327 (5967): 818–822. doi:10.1126/science.1183700.

Turner, D, A Lucieer, and L Wallace. 2014. "Direct Georeferencing of Ultrahigh-Resolution UAV Imagery." *Geoscience and Remote Sensing, IEEE Transactions on*. doi:10.1109/TGRS.2013.2265295.

USGS. 2014. "What Are the Band Designations for the Landsat Satellites?" http://landsat.usgs.gov/band_designations_landsat_satellites.php.

van der Zon, Niels. 2013. "Kwaliteitsdocument AHN-2," 1–30. http://www.ahn.nl/.

White, Jeffrey W., Pedro Andrade-Sanchez, Michael A. Gore, Kevin F. Bronson, Terry A. Coffelt, Matthew M. Conley, Kenneth A. Feldmann, et al. 2012. "Field-Based Phenomics for Plant Genetics Research." *Field Crops Research* 133 (July): 101–112. doi:10.1016/j.fcr.2012.04.003.

Yin, Xinhua, M. Angela McClure, Ngowari Jaja, Donald D. Tyler, and Robert M. Hayes. 2011. "In-Season Prediction of Corn Yield Using Plant Height under Major Production Systems." *Agronomy Journal* 103 (3): 923–929. doi:10.2134/agronj2010.0450.

Zhang, Chunhua, and John M. Kovacs. 2012. "The Application of Small Unmanned Aerial Systems for Precision Agriculture: A Review." *Precision Agriculture* 13 (6): 693–712. doi:10.1007/s11119-012-9274-5.

Zhang, Lei, and Tony E. Grift. 2012. "A LIDAR-Based Crop Height Measurement System for Miscanthus Giganteus." *Computers and Electronics in Agriculture* 85 (July): 70–76. doi:10.1016/j.compag.2012.04.001.

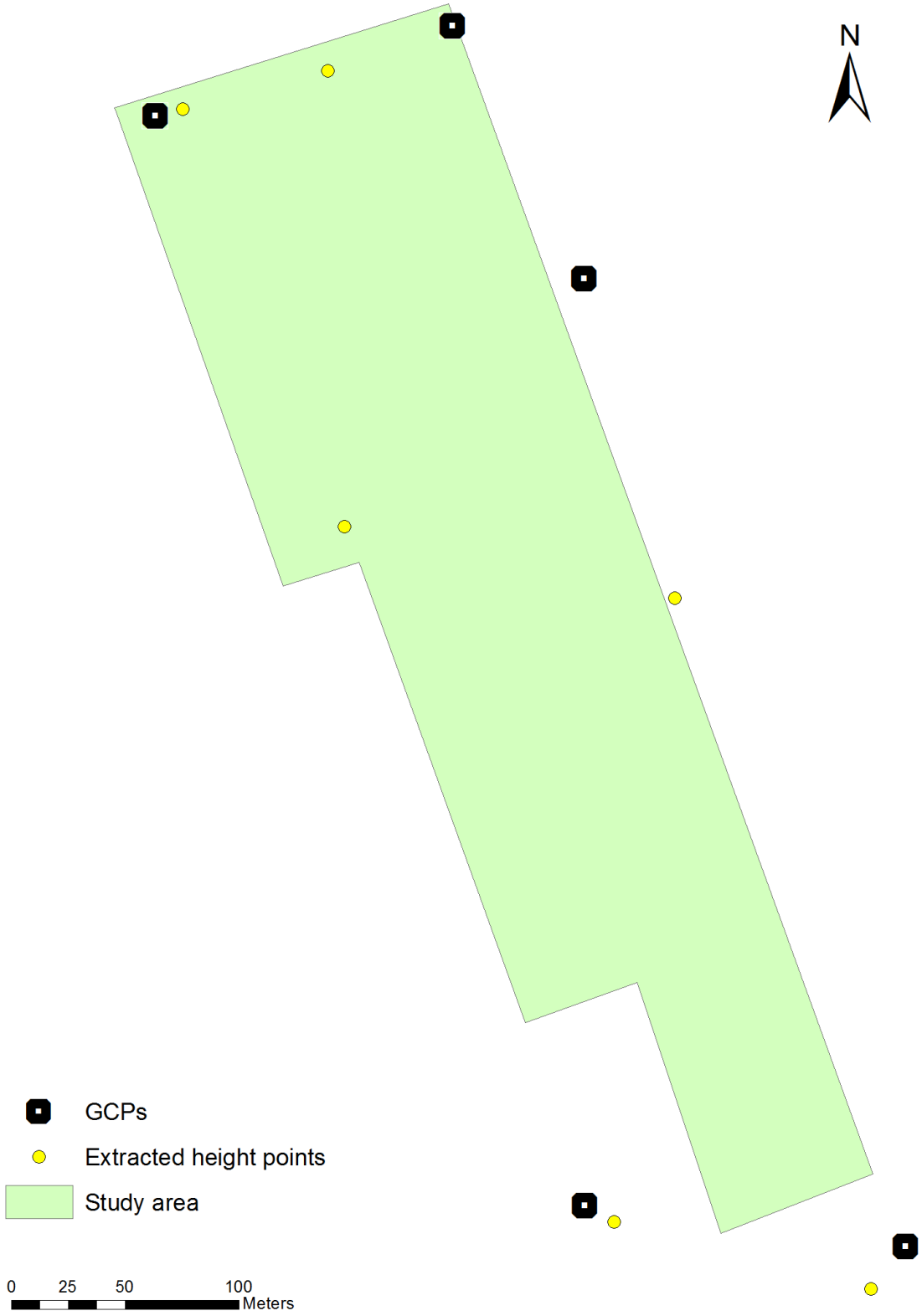
9 APPENDICES

Annex A

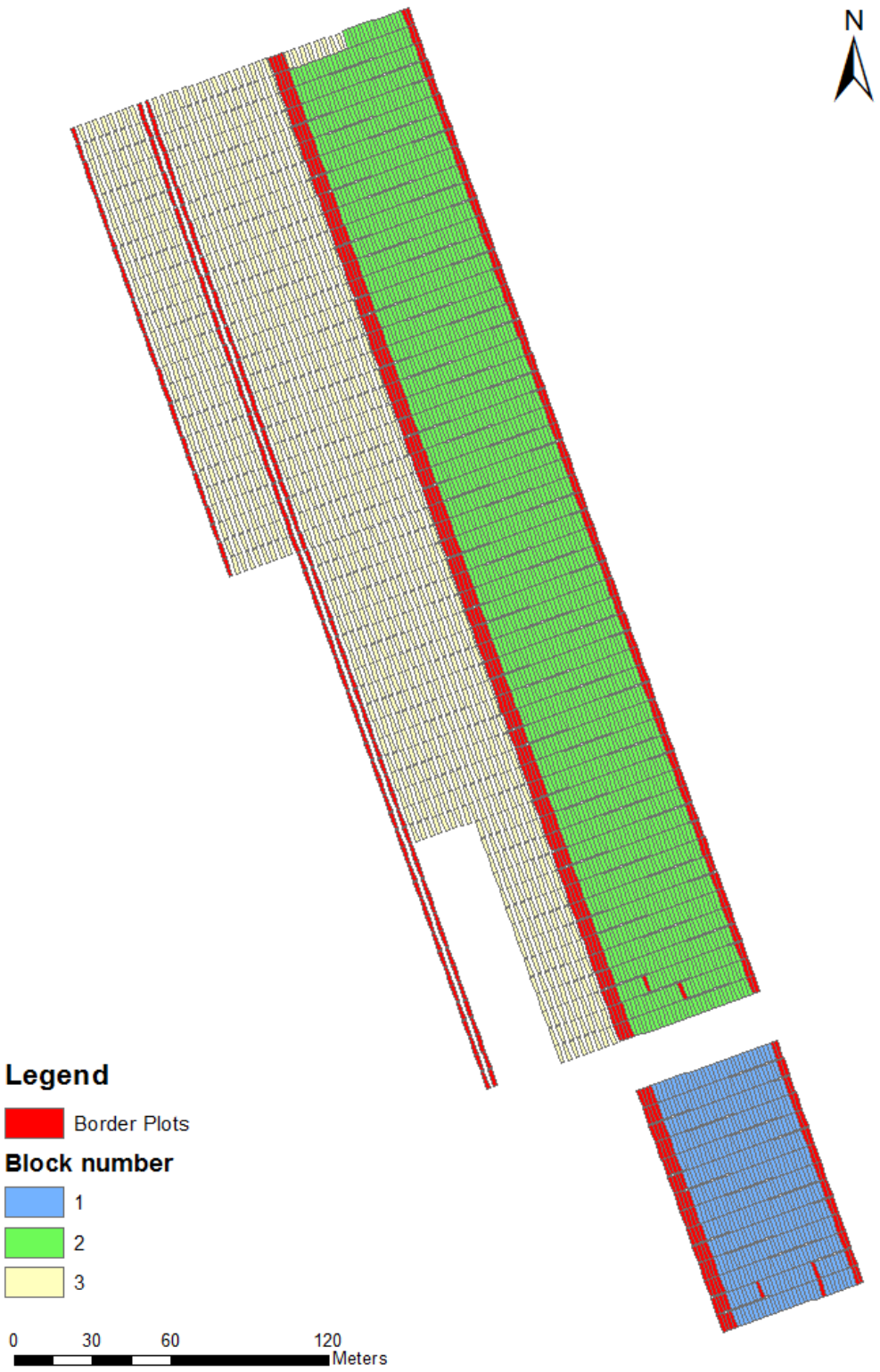
Specifications of UAV payload (Suomalainen et al. 2014)

System	
Components	- Custom spectrometer: o PhotonFocus SM2-D1312, o Specim ImSpector V10 2/3”, o 12mm lens - XSens MTi-G-700 GPS-INS + antenna - Panasonic GX1 + 14 mm pancake lens - Raspberry PI data sink - Synchronization electronics - Carbon fiber frame - 12V Battery
Total weight (ready-to-fly)	2.0 kg
Power intake	10 W @ 12 V
Spectrometer (SM2 + ImSpector)	
Spectral range	400–950 nm
Spectral resolution (FWHM)	9 nm
Cross-track pixels	328 (1312 unbinned)
Swath width	0.8 x flight altitude
Scan rate	~20 lines/s
Signal-to-noise ratio	300:1 (for unbinned pixel at full well)
Camera (*Panasonic* GX1)	
Image size	4608 x 3464 pixels [1.2 x 0.9] x flight altitude
Max frame rate	0.7 Hz @ raw format
GPS-INS (XSens MTi-G-700)	
Roll & pitch accuracy	0.3°
Yaw accuracy	Nominal: 1° Observed: >5°
Spatial accuracy	4 m

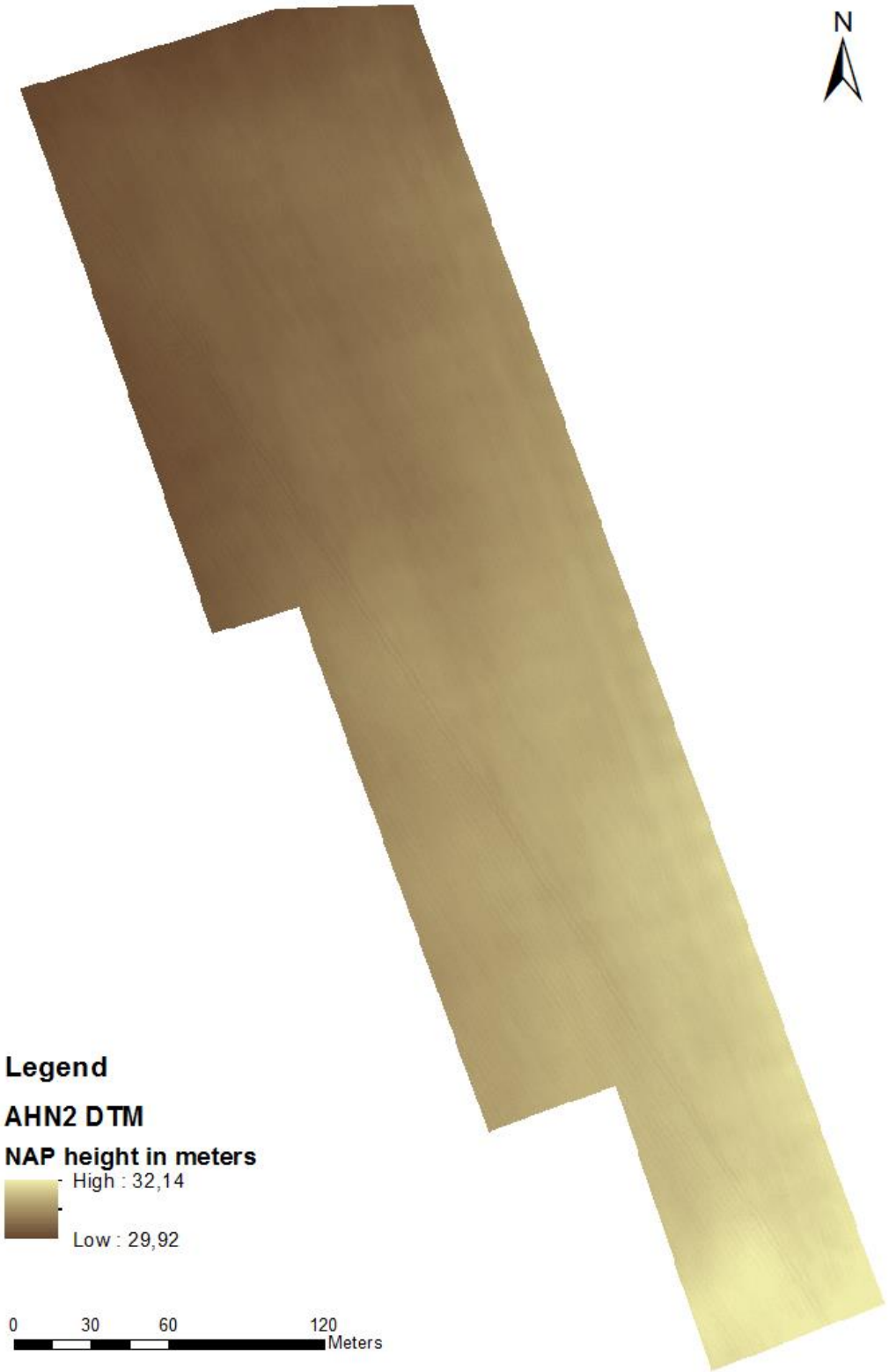
Annex B



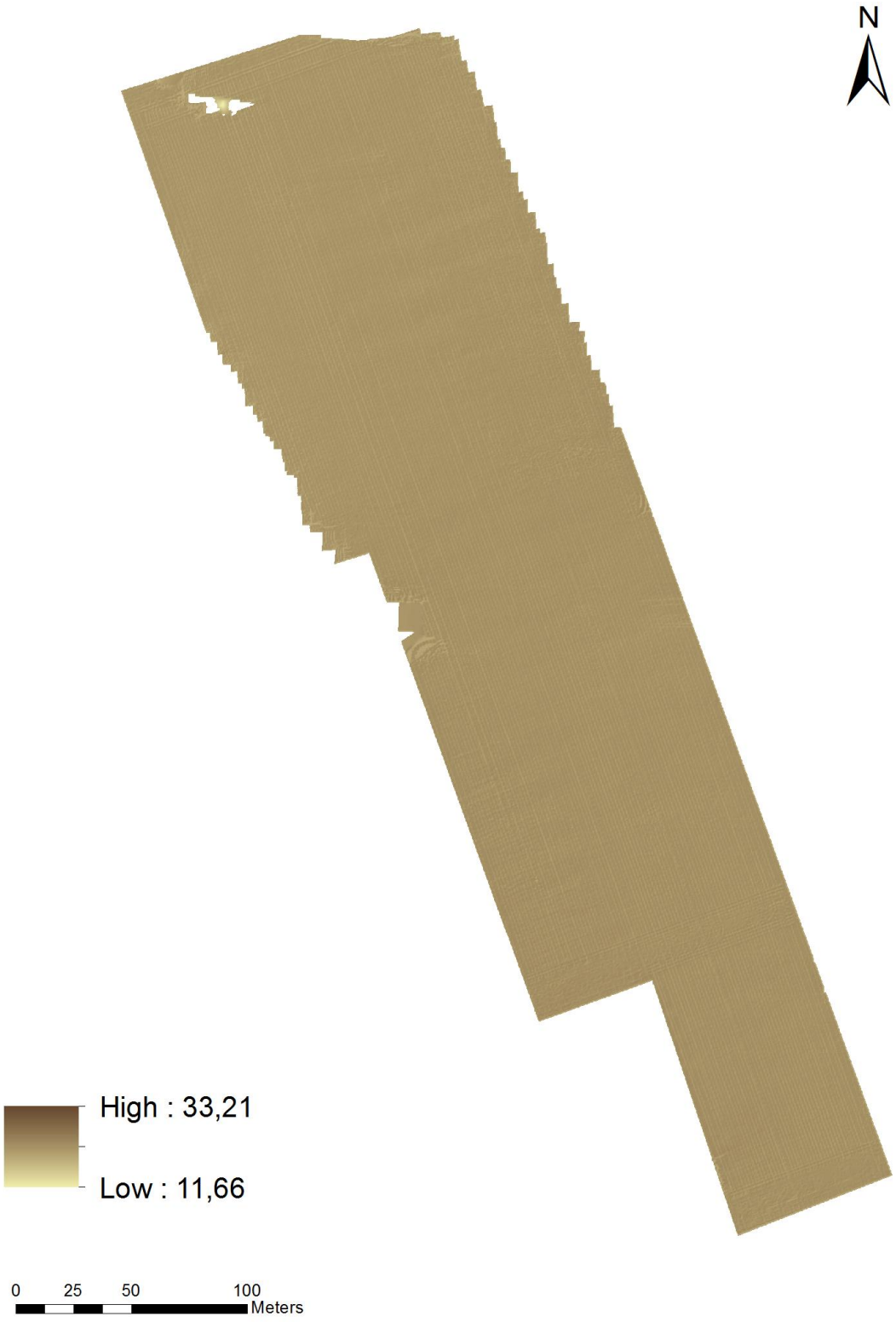
Annex C



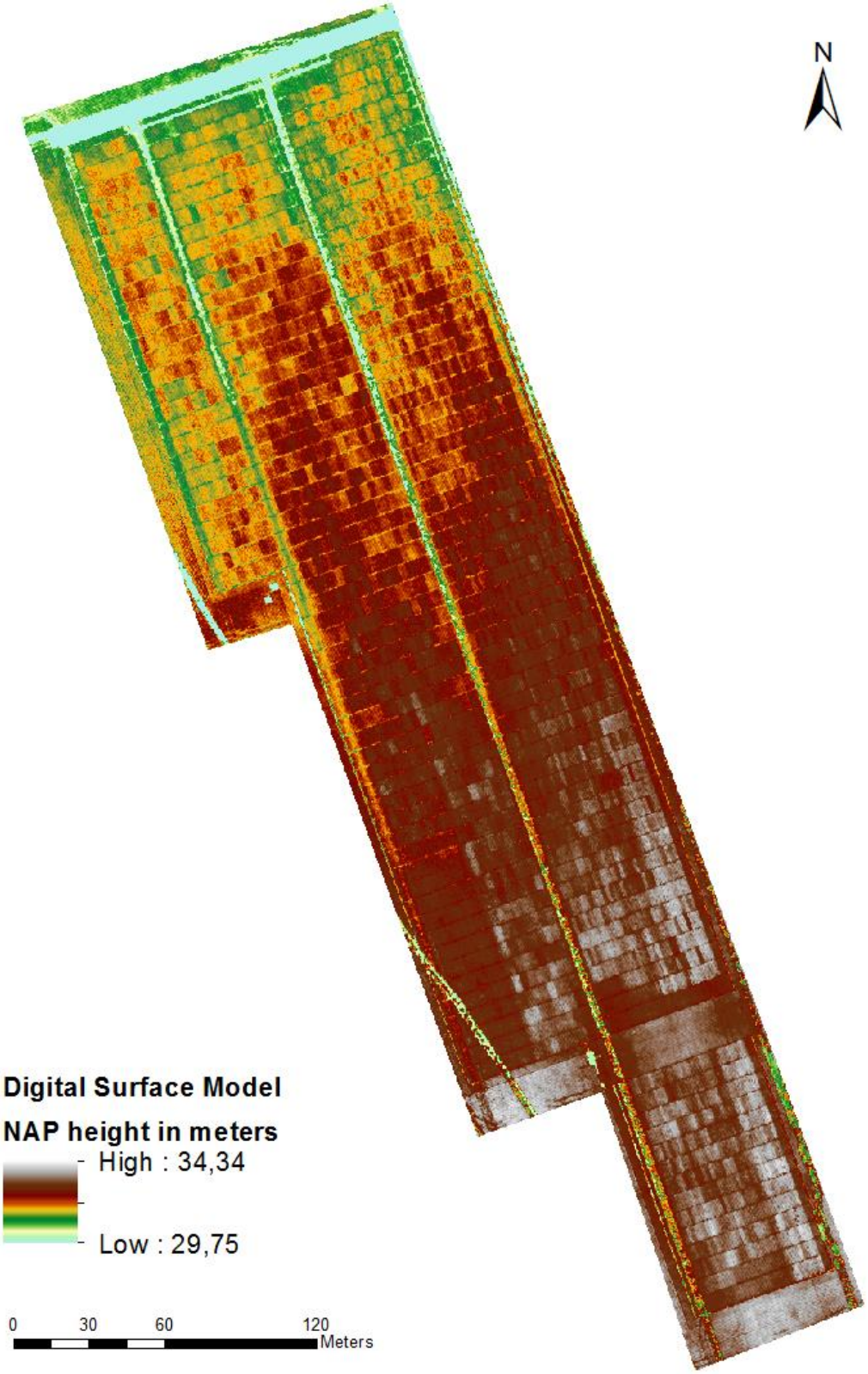
Annex D


N
Legend
AHN2 DTM
NAP height in meters
High : 32,14
Low : 29,92
0
30
60
120
Meters

Annex E



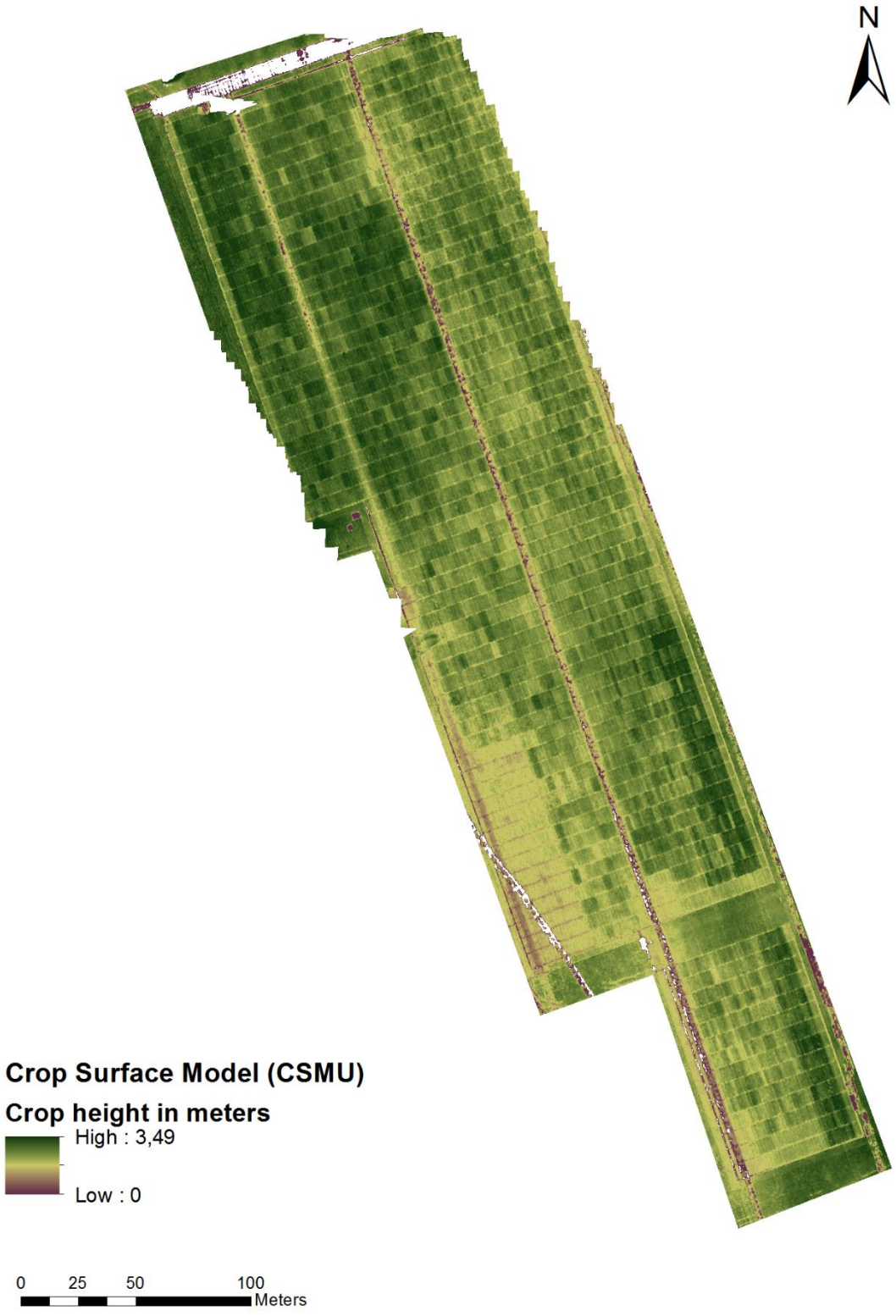
Annex F


N
Digital Surface Model
NAP height in meters
High : 34,34
Low : 29,75
0
30
60
120
Meters

Annex G


N
Crop Surface Model (CSMA)
Plant height in meters
High : 3,18
Low : 0
0
30
60
120
Meters

Annex H


N
Crop Surface Model (CSMU)
Crop height in meters
High : 3,49
Low : 0
0
25
50
100
Meters