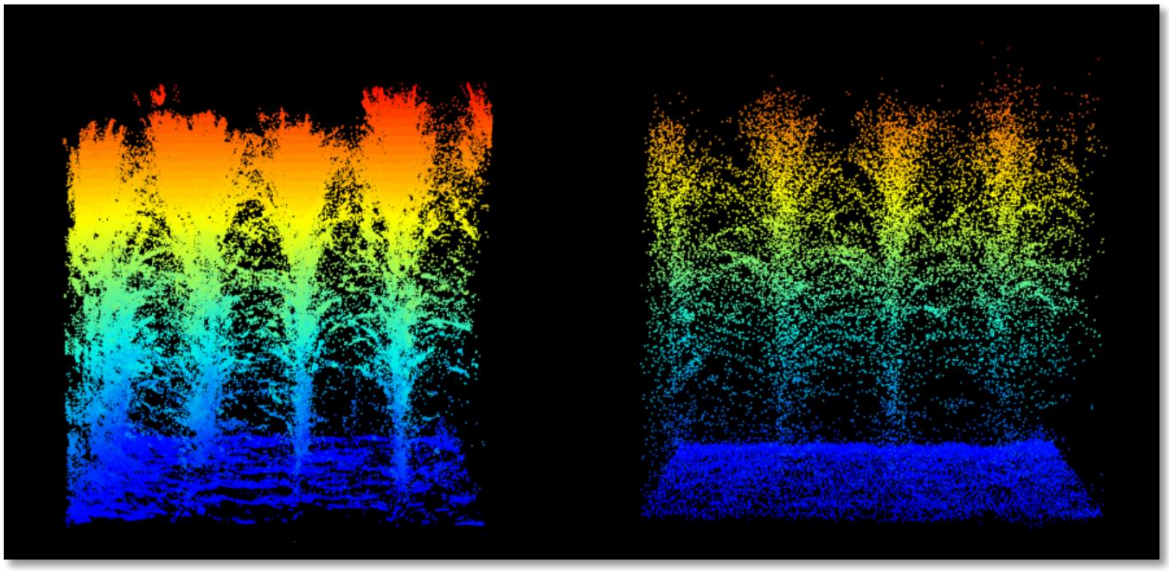
Geo-information Science and Remote Sensing

Thesis Report GIRS-2021-32

Leaf Area Index and Above Ground Biomass Estimation from Unmanned Aerial Vehicle and Terrestrial LiDAR data using Machine Learning Approaches

Gözde COŞKUN



25.05.2021

WAGENINGEN
UNIVERSITY & RESEARCH

Leaf Area Index and Above Ground Biomass Estimation from Unmanned Aerial Vehicle and Terrestrial LiDAR data using Machine Learning Approaches

Gözde Coşkun

Registration number 1029474

Supervisors:
Dr Benjamin Brede
Dr Harm Bartholomeus

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

Thesis code number: GRS-82000_2019_0_
Thesis Report: GIRS-2021-32
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing

1. INTRODUCTION 1
2. METHODS 4
2.1. Study Area 6
2.2. Field Experiment 6
2.3. Data acquisition 7
2.3.1. Non-Destructive Measurements 7
2.3.2. Destructive Measurements 8
2.3.3. Processing LiDAR Data 9
2.4. Feature Extraction 10
2.4.1. Digital Elevation Models(DEMs) 10
2.4.2. Canopy Height Model 12
2.4.3. Canopy Volume 12
2.4.4. Canopy 3D Profile Index 12
2.5. Modelling 13
2.5.1. Multiple Linear Regression Model 14
2.5.2. Random Forest Model 14
2.5.3. Support Vector Regression 15
2.6. Accuracy Assesment 15
3. RESULTS 16
3.1. Relationship Between Predictor Variables and Response 16
3.2. Predictability Performance of AGB and LAI with all crop samples 18
3.3. Predictability Performance of AGB and LAI by Crops 19
4. DISCUSSION 23
5. CONCLUSIONS and RECOMMENDATIONS 27
7. REFERENCES 29
8. ADDITIONAL FIGURES 33

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor's Dr Benjamin Brede and Dr Harm Bartholomeus, for their encouragement, patience, support, enlightening my way and guiding the research**.** I would also like to thank to Dainius Masiliunas for programming support and to Jelle Ten Harkel for sharing all the details of his thesis with me.

Furthermore, armful thanks to my family, who supported me in all circumstances, from my birth to this day, to my friends for motivating me. I would like to express my deepest appreciation to M. Kürşat Coşkun, thanks for taking this journey with me, making me smile, trusting, encouraging, tremendous understanding and believing in me whenever I lose hope and giving up.

Additionally, I greatly appreciate that this master education and thesis study were carried out with the financial support and contribution of the Republic of Turkey.

Abstract

Above-ground biomass (AGB) and leaf area index (LAI) are two critical crop parameters for evaluating crop growth, health, and productivity. Various techniques have been used in the past for the estimation of crop parameters. Development and accessibility of remote sensing technologies by airborne, ground devices, satellites opened a new era for estimating the crop parameter. In this paper, the crop parameters LAI and AGB were estimated using Unmanned Aerial Vehicle LiDAR and ground-based Terrestrial Laser Scanning (TLS) data applying three machine learning methods over four crop types with different canopy properties; wheat, sugarbeet, soybean and maize. The study compared how the predictive performance of three modelling techniques multiple linear regression (MLR), random forest regression (RFR), support vector machine regression (SVR) differ, how the predictability performance differ between crop parameters, crop types and LiDAR sensors. The analysis was carried out using all the crop samples obtained and the crop-specific. Two statistical criteria were used as evaluation metrics, the coefficient of determination (R^2) and Root Mean Square Error (RMSE). The results obtained from the models applied with all crop samples indicated that the predictability performance of AGB was over 0.80 R^2 and the LAI was predicted with the R^2 of 0.76. In crop-specific modelling, the estimation of soybean and maize AGB and LAI parameters with models applied with features obtained from both UAV-LS and TLS data was satisfactory. The prediction was failed on sugar beet AGB estimation (R^2 in 0.54–0.76 range) and wheat LAI prediction (R^2 in 0.46–0.86 range). Based on the model performances, the RFR and SVR models performed better than the linear model MLR in most of the analyses. The models compared to the data of the sensors used in this study, the models applied with the features obtained from UAV-LS and TLS gave similar R^2 and $RMSE$ when made with all crop samples. However, in crop-specific analyses, the models applied with the features obtained from UAV-LS showed a powerful performance in all LAI estimation models. Although the results obtained from models applied using TLS and UAV-LS features to estimate AGB were very close, only sugarbeet AGB estimation gave better results with TLS data than UAV-LS. Given these results, the lacking application in crop studied with LiDAR remote sensing will be a step to close to fill the gap using both TLS and UAV-LS derived data for several crop types.

Keywords: Terrestrial LiDAR; UAV LiDAR; Above Ground Biomass; Leaf Area Index; Machine learning; Precision Agriculture

Makine Öğrenimi Yaklaşımları Kullanılarak İnsansız Hava Aracı ve Karasal LiDAR verilerinden Yaprak Alanı İndeksi ve Yer Üstü Biyokütle Tahmini

Toprak üstü biyokütle (AGB) ve yaprak alanı indeksi (LAI), bitki büyümesini, sağlığını ve üretkenliğini değerlendirmek için iki kritik bitki parametresidir. Bitki parametrelerinin tahmini için geçmişte çeşitli teknikler kullanılmıştır. Hava, yer cihazları, uydular tarafından uzaktan algılama teknolojilerinin geliştirilmesi ve erişilebilirliği, bitki parametresini tahmin etmek için yeni bir dönem açtı. Bu yazıda, bitki parametreleri LAI ve AGB, İnsansız Hava Aracı LiDAR ve yer tabanlı Karasal Lazer Tarama (TLS) verileri kullanılarak, farklı bitki özelliklerine sahip dört ürün türü üzerinde üç makine öğrenimi yöntemi uygulanarak tahmin edilmiştir; buğday, şeker pancarı, soya fasulyesi ve mısır. Çalışma, çoklu doğrusal regresyon (MLR), rastgele orman regresyonu (RFR), destek vektör makine regresyonu (SVR) olacak şekilde, üç modelleme tekniğinin tahmin performansının nasıl farklılık gösterdiğini, öngörülebilirlik performansının bitki parametreleri, bitki türleri ve LiDAR sensörleri arasında nasıl farklılık gösterdiğini karşılaştırdı. Analiz, önce elde edilen tüm bitki numuneleri ile ve her bitki türüne göre ayrı ayrı gerçekleştirilmiştir. Değerlendirme ölçütleri, belirleme katsayısı (R2) ve Ortalama Kare Hatası (RMSE) olmak üzere iki istatistiksel kriter kullanılmıştır. Tüm bitki örnekleriyle uygulanan modellerden elde edilen sonuçlar, AGB'nin öngörülebilirlik performansının 0.80 R^2'nin üzerinde olduğunu ve LAI'nin 0.76 R^2 ile tahmin edildiğini göstermiştir. Ürüne özel modellemede, hem UAV-LS hem de TLS verilerinden elde edilen özelliklerle uygulanan modellerle soya fasulyesi ve mısır AGB ve LAI parametrelerinin tahmini tatmin edici olmuştur. Şeker pancarı AGB tahmininde (0,54-0,76 aralığında R^2) ve buğday LAI tahmininde (0,46-0,86 aralığında R^2) tahmin başarısız oldu. Model performanslarına dayalı olarak, analizlerin çoğunda RFR ve SVR modelleri doğrusal model MLR'den daha iyi performans göstermiştir. Modeller, bu çalışmada kullanılan sensörlerin verileriyle karşılaştırıldığında, UAV-LS ve TLS'den elde edilen özelliklerle uygulanan modeller, tüm Bitki numuneleri ile yapıldığında benzer R^2 ve RMSE vermiştir. Ancak ürüne özgü analizlerde UAV-LS'den elde edilen öznitelikler ile uygulanan modeller tüm LAI tahmin modellerinde güçlü bir performans göstermiştir. AGB tahmini için TLS ve UAV-LS özellikleri kullanılarak uygulanan modellerden elde edilen sonuçlar çok yakın olmasına rağmen, sadece şeker pancarı AGB tahmini TLS verileri ile UAV-LS'ye göre daha iyi sonuçlar vermiştir. Bu sonuçlar göz önüne alındığında, LiDAR uzaktan algılama ile çalışılan bitkideki eksik uygulama, çeşitli bitki türleri için hem TLS hem de UAV-LS'den türetilen verileri kullanarak boşluğu doldurmak için bir adım olacaktır.

Anahtar Kelimeler: Karasal LiDAR; İHA LiDAR; Yer Üstü Biyokütle; Yaprak Alan İndeksi; Makine öğrenme; Hassas tarım

List of Figures

Figure 1. Flow chart of the process of estimating AGB and LAI. 5
Figure 2. Overview of fields within the study area and OpenStreetMap used as background map and the Google maps used as the inset map 6
Figure 3. Pre-processing steps of data by Riegl RiCOPTER.(Brede et al., 2017) 10
Figure 4 The correlation between UAV-LS derived features and AGB and LAI, the colours represent the harvest date and the shapes indicate the crop types 17
Figure 5 The MLR and SVR results with sugarbeet-LAI estimation using UAV-LS features 21

List of Tables

Table 1. The software and implementations 4
Table 2 VZ-400 and VUX R-1UAV main characteristics. (Brede et al., 2017) 7
Table 3 LiDAR data acquisition and related harvest number 8
Table 4. An Overview of field-measured AGB and LAI by crops. 9
Table 5. Machine Learning Models implemented in this study 14
Table 6 The coefficient of determination ($R2$) of the linear relationship between predictor variables and response 16
Table 7. The Result of Predictability Performance of AGB and LAI with all crop samples 18
Table 8 Predictability Performance of Wheat AGB and LAI 20
Table 9 Predictability Performance of Sugar Beet AGB and LAI 21
Table 10 Predictability Performance of Soybean AGB and LAI 22
Table 11 Predictability Performance of Maize AGB and LAI 23

List of Additional Figures

Annex 1 UAV-LS derived Canopy Features Histogram 33
Annex 2 Correlation Between AGB and UAV-LS derived features 34
Annex 3 Correlation Between LAI and UAV-LS derived features 35
Annex 4 UAV-LS - Predicted and Observed values of AGB with all samples 36
Annex 5 UAV-LS - Predicted and Observed values of LAI with all samples 37
Annex 6 UAV-LS - Predicted and Observed values of Wheat AGB 38
Annex 7 UAV-LS - Predicted and Observed values of Wheat LAI 39
Annex 8 UAV-LS - Predicted and Observed values of Sugarbeet AGB 40
Annex 9 UAV-LS - Predicted and Observed values of Sugarbeet LAI 41
Annex 10 UAV-LS - Predicted and Observed values of Soybean AGB 42
Annex 11 UAV-LS - Predicted and Observed values of Soybean LAI 43
Annex 12 UAV-LS - Predicted and Observed values of Maize AGB 44
Annex 13 UAV-LS - Predicted and Observed values of Maize LAI 45
Annex 14 TLS derived Canopy Features Histogram 46
Annex 15 Correlation Between AGB and TLS derived features 47
Annex 16 Correlation Between LAI and TLS derived features 48
Annex 17 TLS-Predicted and Observed values of AGB with all samples 49
Annex 18 TLS-Predicted and Observed values of LAI with all samples 50
Annex 19 TLS-Predicted and Observed values of Wheat AGB 51
Annex 20 TLS-Predicted and Observed values of Wheat LAI 52
Annex 21 TLS-Predicted and Observed values of Sugarbeet AGB 53
Annex 22 TLS-Predicted and Observed values of Sugarbeet LAI 54
Annex 23 TLS-Predicted and Observed values of Soybean AGB 55
Annex 24 TLS-Predicted and Observed values of Soybean LAI 56
Annex 25 TLS-Predicted and Observed values of Maize AGB 57
Annex 26 TLS-Predicted and Observed values of Maize LAI 58

List of Abbreviations

3DPI	Three Dimensional Profile Index
AGB	Above Ground Biomass
ANN	Artificial Neural Network
CHM	Canopy Height Model
CV	Canopy Volume
DEMs	Digital Elevation Models
DSM	Digital Surface Model
DTM	Digital Terrain Model
GCP	Ground Control Panel
k-NN	K nearest Neighbour
LAI	Leaf Area Index
LiDAR	Light Detection and Ranging
LPV	LiDAR Projected Volume
ML	Machine Learning
MLR	Multiple Linear Regression
NA	Not available
PA	Precision Agriculture
RBF	Radial Basis Kernel Function
RFR	Random Forest Regression
SVMs	Support Vector Machines
SVR	Support Vector Machine Regression
TIN	Triangular Irregular Network
TLS	Terrestrial Laser Scanner
UAV	Unmanned Aerial Vehicle
UAV-LS	Unmanned Aerial Vehicle Laser Scanner

1. INTRODUCTION

The human population in the world is increasing day by day. Meeting the food demand relative to the growing population will require a significant increase in crop production. Precision agriculture (PA) which is the science of helping increase crop yields using high-tech sensors and analysis tools, can tackle this demand (Singh et al., 2020). PA is a popular concept adopted worldwide to increase production, shorten labour time and provide efficient fertilizer and irrigation systems (Singh et al., 2020). Accordingly, PA requires accurate and high-efficiency screening of phenotypic traits. Precise and accurate phenotypic variables assessment is crucial for advancing crop yield and crop breeding (Jin et al., 2020). Among these variables, above-ground biomass (AGB) and leaf area index (LAI) are two critical parameters for evaluating crop growth, health, and productivity.

Various techniques have been used in the past for the estimation of crop parameters. LAI and AGB can be measured directly by destructive sampling and non-destructive field measurements. The traditional, destructive methods are time-consuming, require intensive work, and have high costs(Jimenez-Berni et al., 2018). Particularly in the studies performed in large areas, data may be lost or can be misinterpreted, which can lead to a decrease in confidence in accuracy. Development and accessibility of remote sensing technologies by airborne, ground devices, satellites opened a new era for estimating the crop parameter. With these technologies, it is possible to rapidly create reliable and accurate maps of extended cropland by exploiting the different types of remote sensing signals of crops depending on their status (Quebrajo et al., 2018).

The data can be obtained using passive and active remote sensing sensors such as optical cameras, radar sensors, terrestrial or UAV light detection and ranging (LiDAR). However, the data derived using passive optical remote sensing sensors usually can be susceptible to interference from environmental effects and saturation at high biomass and LAI (Jin et al., 2020). In contrast, active remote sensing systems like LIDAR can provide accurate three-dimensional (3D) output and a reliable estimation. Deery et al. (2014) indicated that the images produced by LiDAR point clouds retain image quality and data reliability in biomass estimations by not being affected by changes in ambient light conditions and shadows which can cause overexposure or underexposure in optical sensors, like RGB cameras. Wang et al. (2017) studied the potential of hyperspectral and LiDAR data for a better estimation of the maize biomass. He stated the most robust relationship was observed between LiDAR-derived

metrics with field-observed biomass, while all the vegetation indices derived from hyperspectral images showed weak correlation.Moreover, Walter et al. (2019) estimated wheat biomass with LiDAR-derived vegetation height and volume. He demonstrated the correlation between field-measured AGB, and LiDAR Projected Volume (LPV) and field-measured canopy height and LiDAR Canopy Height (LCH) were generally strong. Nevertheless, LiDAR data acquisition and processing costs are relatively high to monitor large areas (Dan et al., 2018).

The estimation of AGB and LAI is not directly possible using remote sensing derived data. These parameters can be obtained in two ways, either using physical methods like radiative transfer models or statistical models like parametric and non-parametric regression methods (Verrelst et al., 2015). Statistically, parameters are estimated by creating a regression relationship between features, independent variable derived from remote sensing data, and dependent variable ground-truth data collected in the field. Among the most popular statistical methods, machine learning is increasingly being used for crop parameter estimations due to their flexibility and capability of linear and nonlinear data sets (Ali et al., 2015). The most common machine learning methods for estimation of crop parameters are artificial neural network (ANN)(González-Sanchez et al., 2014; Jin et al., 2020; Kucukonder et al., 2016), multiple Linear regression (MLR) (Ali et al., 2015; Comba et al., 2020; González-Sanchez et al., 2014), k-nearest neighbour (k-NN)(Fu et al., 2019; González-Sanchez et al., 2014; Han et al., 2019),random forest regression (RFR)(Breiman, 2001; Fu et al., 2019; Jin et al., 2020) and support vector machine regression (SVR)(Clevers et al., 2007; Durbha et al., 2007; García-Gutiérrez et al., 2013).

LiDAR is an active remote laser-based technology that is well established in its application to derive vegetation features for forest areas (Brede et al., 2017; Jimenez-Berni et al., 2018). Applications in crops still remain underdeveloped. There are a few studies performed on crops. For instance, Harkel et al. (2019) estimated biomass using the 3DPI algorithm and crop height applying the linear regression methods and stated that both 3DPI and height estimation on sugar beet, winter wheat worked well while estimation on potato was less reliable. Han et al. (2019) used structural and spectral information provided by remote sensing from UAV combined with machine learning to estimate maize biomass and concluded that the combination of machine learning and UAV remote sensing is a promising alternative for AGB prediction. Comba et al. (2020) evaluated LAI in vineyards by a multivariate linear regression model using crop canopy descriptors derived from the UAV-LS 3D point clouds, and evaluation results showed a good

correlation between the LiDAR-derived crop canopy descriptors estimated and field-measured LAI.

A comprehensive study of AGB and LAI estimation using LiDAR-derived features with the crops applying several statistical methods have not been fully explored and utilized. This research will be a step to close to fill the gap using both Terrestrial Laser Scanner (TLS) and UAV-LS derived data for several crop types.

The objective of this thesis is to estimate the crop parameters LAI and AGB using Unmanned Aerial Vehicle LiDAR and ground-based Terrestrial Laser Scanning (TLS) data applying machine learning methods over four crop types with different canopy properties; wheat, sugarbeet, soybean and maize. Considering the LiDAR point clouds cannot be directly used for estimation, some features needed to be extracted, such as canopy height (CH)(Harkel et al., 2019; Jimenez-Berni et al., 2018; Walter et al., 2019), canopy volume (CV)(Jin et al., 2020; Verma et al., 2016) and canopy three-dimensional profile index (3DPI)(Jimenez-Berni et al., 2018) derived from LiDAR returns.

The regression analyses will be performed between LiDAR-derived features and field measured AGB and LAI. The research will answer the main question, “How accurately can we obtain the crop parameters LAI and AGB with features derived from LiDAR sensors with machine learning methods?”. In order to understand the overall performance, different factors in the prediction models were analyzed by asking the following questions:

1. How does performance differ between the selected machine learning models?
2. How does performance differ between AGB and LAI?
3. How does performance differ between crop types?
4. How does performance differ between UAV-LS and TLS LiDAR?

2. METHODS

The research method consists of five parts that can be followed in Figure 1: Field Experiment, data acquisition, feature extraction, model training, and validation and accuracy assessment. The software and implementations used in this study can be seen in Table 1.

Table 1. The software and implementations

Software	Package/Tools	Usage
LAStools[1] 201003		Processing LAZ/LAS file, DEMs generation
QGIS 3.16	GDAL, value tool	Quality check of the pixels
CloudCompare 2.11.3		Visualization of LAZ file
Rstudio 1.2.5033	caret, lidR, dplyr,raster	Feature extraction, Modelling, Validation

1- Both the tool itself and scripting by batch were used.

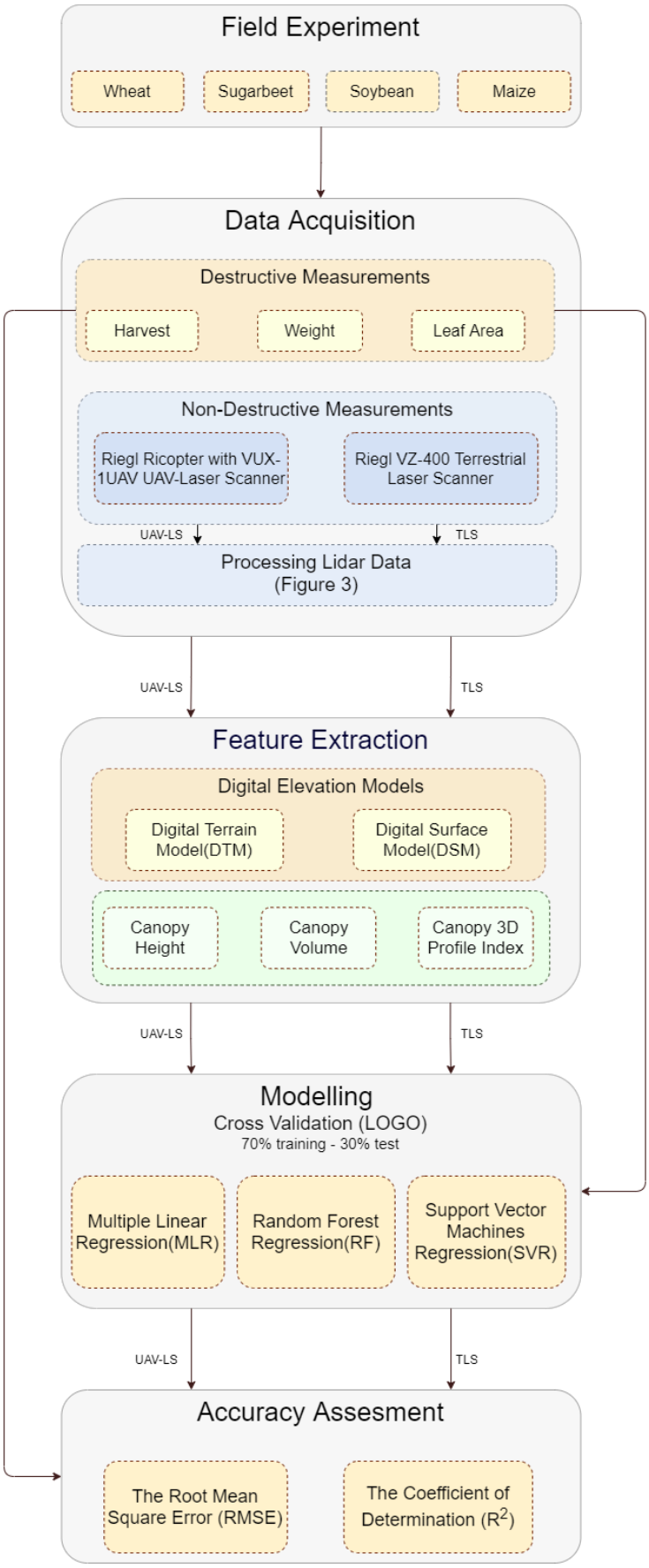


Figure 1. Flow chart of the process of estimating AGB and LAI.

2.1. Study Area

The study site were located north of Wageningen and approximately 1.30 km away from the Wageningen University & Research (WUR) campus and belongs to Unifarm, the WUR experimental farm (Figure 2**).**

Figure 2. Overview of fields within the study area and OpenStreetMap used as background map and the Google maps used as the inset map

2.2. Field Experiment

The total measured area was roughly 1600 m^2. Four types of crops were planted in the study area, which was wheat (*Triticum aestivum L.*), sugar beet (*Beta vulgaris*), soybean (*Glycine max L.*), and maize (*Zea mays L.*). The experimental design aimed at maximizing the differences of the crop development stages to enable parameter estimation over the crop life cycle. For this, each crop type was planted at two sowing dates to increase the variety of development at sampling events. Furthermore, harvest spots were distributed over four repetitions along the length of the field in order to use within-field variation to diversify development stages further. Finally, seven sampling events were spread over the growing season. Harvest plots were

marked with wooden/bamboo sticks and geo-located with Real-time kinematic-global navigation satellite systems (RTK-GNSS) to sub-centimetre accuracy.

2.3. Data acquisition

The LiDAR point clouds and field data were collected to create a regression relationship between features. The field experiment included both destructive and non-destructive measurements, with the non-destructive measurement typically performed 1-3 days prior to the destructive sampling events.

2.3.1. Non-Destructive Measurements

LiDAR data was acquired using a RIEGL RICOPTER with VUX-1UAV UAV-Laser scanner and a RIEGL VZ-400 Terrestrial Laser Scanner (Table 2)**.** Sampling days were scheduled at regular intervals every two weeks when crop growth was judged sufficiently compared to the last sampling event, the weather allowed, and when pilots were available. Final, samplings were performed on the days that can be seen in Table 3.

Table 2 VZ-400 and VUX R-1UAV main characteristics. (Brede et al., 2017)

Characteristic	VZ-400	VUX-1UAV
Maximum Pulse Repition Rate (PRR) (kHz)	300	550
Maximum effective measurement rate (kHz)	120	500
Minimum|Maximum range (m)	1.5|350	3|920
Accuracy|Precision (mm)	5|3	10|5
Laser wavelength (nm)	1550	1550
Beam divergence (mrad)	0.3	0.5
Weight (kg)	9.6	3.75

UAV flight planning was done in a way to maximize lateral overlap and efficiently use airtime. Ground Control Panels (GCP) were placed in the field to support flight line co-registration and co-registration with the TLS. Styrofoam GCPs of 120 cm x 60 cm were set up with the two panes forming a 90˚ ~~o~~ angle between them (Brede et al., 2017). After flying the UAV-Laser scanner, TLS measurements were performed. Cylindrical retro-reflective targets were set up for later coarse co-registration of TLS scans (Brede et al., 2017). Five retro-reflective targets were registered with RTK-GNSS and scanned in a fine-scan mode in order to register the TLS point clouds in a global coordinate reference system, which was WGS84 UTM31N. Further scans were performed at each location that would be harvested after this sampling event.

Table 3 LiDAR data acquisition and related harvest number

Flight Date	Harvest
13.05.2020	H1
26.05.2020	H2
10.06.2020	H3
23.06.2020	H4
15.07.2020	H5
5.08.2020	H6
1.09.2020	H7

2.3.2. Destructive Measurements

Harvesting was done 1-3 days after non-destructive measurements to collect ground truth data for all crops. All above-ground plant material was harvested and, per plot, put into bags. In an indoor laboratory, weighing the whole plot fresh mass was performed with precision scales, usually on the same day as the harvest. Then, two samples per bag were taken from the fresh mass and weighted.

For the first sample, all leaves were separated from other organs and leaves were weighed separately. Then leaf area of these leaves was determined with a LiCOR Leaf Area Meter. The leaf area of the sample was scaled to the whole plot with the ratio of the sample weight to the whole plot weight.

The whole second sample was dried at 105°C for 48h until all the water was vaporated, then re-weighted. The dry-to-fresh weight ratio was used together with the sample to whole plot fresh weight to scale dry weight for the whole plot. Dry weight is interpreted as AGB. Both leaf area and biomass are translated into density parameters by taking the plot surface area into account.

The targeted sample number at plot level decreased from 944 to 909 due to operating errors such as mixing some bags during destructive sampling, incorrect measurement, or incorrect adhering of the bags' labels. Erroneous samples were excluded from the further calculation.

Different plants were harvested at each destructive sampling event. The harvested plants and the summary of the acquired data can be seen in Table 4 where Hi represents the number of harvest events. The calculation was performed at the plot level in the table, and raw plot data was used.

Table 4. An Overview of field-measured AGB and LAI by crops.

	Above Ground Biomass, g			Leaf Area Index, m2/m2		
Harvest	min	mean	max	min	mean	max
			Wheat			
H1	151,4	938,5	1931,4	0,12	0,46	0,85
H2	195,5	477,8	1199,7	0,64	0,99	1,75
H3	307,9	956,9	1637,8	0,64	1,85	3,36
H4	719,5	1298,9	2089,4	0,54	1,69	3,68
			Sugar beet			
H2	116,6	170,6	234,6	0,4	0,61	0,79
H3	5,8	210,4	779,4	0,03	1,1	2,73
H4	78	608,8	1683,5	0,34	1,53	3,66
H5	388,4	1158	1989,6	1,65	3,785	5,996
			Soybean			
H3	1,1	5	13,7	0,007	0,02	0,04
H4	41,1	102,2	273,4	0,21	0,36	0,64
H5	467,6	755	1141,6	2,01	3,7	5,29
H6	1979	2541	3059	4,67	5,99	7,64
H7	1839	2545	3733	3,56	4,81	6,16
			Maize			
H4	5,1	22,3	55,3	0,02	0,09	0,24
H5	18,7	159,9	424,8	0,06	0,66	1,72
H6	114,3	511,9	976,2	0,36	1,85	3,37
H7	357,5	861,7	1622	0,91	2,43	4,84

2.3.3. Processing LiDAR Data

Point clouds are large data sets composed of 3D point data derived from LiDAR laser scanners. The data derived from Riegl RiCOPTER was pre-processed to convert raw data into common format point cloud datasets. The required steps were followed according to Brede et al. (2017) can be seen in Figure 3.

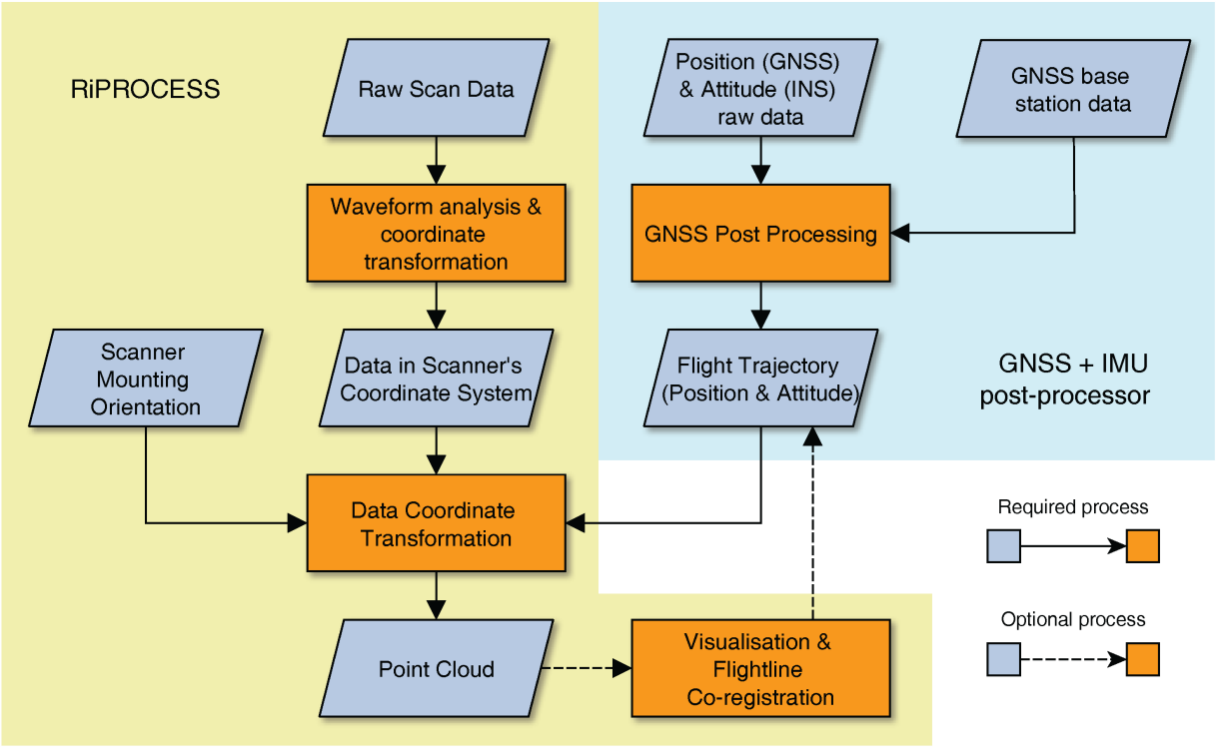


Figure 3. Pre-processing steps of data by Riegl RiCOPTER.(Brede et al., 2017)

2.4. Feature Extraction

2.4.1. Digital Elevation Models(DEMs)

In this section, DEMs generation will be described to perform the extraction of the canopy height model (Wang et al.), canopy volume(CV) and canopy three-dimensional profile index (3DPI). To derive these features from laser data, an estimate of the ground elevation at each location (x, y) with a canopy hit was needed. This was accomplished by creating digital terrain models (DTMs) from ground hits and Digital Surface Models (DSMs) from top of the canopy hits (Magnussen et al., 2011).

The generated raw point clouds cover a much larger area than the studied area. Therefore, the point cloud was clipped to the study area to reduce data size and thus increase processing speed for further operations using LAStools (LAStools, 2020). Both TLS and UAV-LS derived data were tiled with 20m tile and 2m buffer to avoid edge artefacts and interpolation errors while processing LiDAR in tiles (LAStools, 2020).

To generate the digital elevation models (DEMs), we started with cleaning noise points. UAV-LS derived point clouds did not have much noise, but TLS had dense noise points due to wind, rain, insects, or small particles in the air, reflections on water, and other effects(Tilly et al.,

2014). Since the noise was almost at the same height as the maize plants, especially in the 7th harvest, they needed to be cleaned manually, looking at the perspective view at the point clouds using lasview tool.

In the next step, LiDAR points were classified into ground and nonground points with the lasground tool of LAStool. A Triangular Irregular Network (TIN) filter was performed iteratively to dispose of the nonground points. The TIN filter detected the lowest point, which was most likely a terrain surface, and triangulate them to yield a TIN (Irwan Hariyono et al., 2018). Then, we merged and converted the points to a raster with a 10cm x 10cm pixel resolution for calculating the DTM.

In past years, it was suggested to use only the first returns to generate the interpolating surface for DSM generation at high resolution (Khosravipour et al., 2016). Using only the first returns causes to miss some of the LiDAR information and details, which are essential to UAV-LS derived data products. We used a spike-free algorithm presented by Khosravipour et al. (2016), which considers not only first returns but also second and third returns and systematically prevents the returns that could cause sudden spikes. The required freeze constraint number, which was 0.15 for this study, was calculated by three times the averaging the pulse spacing to apply the spike free algorithm (Khosravipour et al., 2016). Generating a DSM with the spike-free algorithm was a very time consuming and challenging process with the available computational resources. Hence, 50% of the points are thinned and flagged and not considered. The thinned points were not included in the calculation to speed up the processing time. Finally, the points were merged and converted to a DSM raster with a 10cm x 10cm pixel resolution.

TLS surveys usually consist of different scanning positions. Even though these positions are well determined, the possibility of shading effects, the fact that the impact angle is not wide enough, or the ground area cannot be entirely observed cause large triangulated artefacts on the DTM rasters (Crespo-Peremarch et al., 2020). As UAV-LS laser beams are almost vertical, these limitations are less prevalent than with TLS. Correspondingly, ground points could be measured better since there is not enough crop growth to cover the ground at the first measurement date (2020-05-13). DTM was calculared from the first flight of UAV-LS and used in combination with DSMs of other harvests to derive CHM, CV and 3DPI.

2.4.2. Canopy Height Model

Most of the published methodologies are based on the determination of canopy height, which is the difference between a digital surface model and a digital terrain model (Loudermilk et al., 2009). This equation gives us the canopy height per plot (Magnussen et al., 1998). In this study, the canopy height was calculated pixel by pixel with a 10 cm x 10 cm pixel size. The pixels were extracted at the plot level and then aggregated by averaging to calculate each plot's mean canopy height. Since harvest and measurement were done for different crops in each destructive sampling, the canopy height obtained with LiDAR was filtered by harvest number.

2.4.3. Canopy Volume

LiDAR-derived plot volume was produced by multiplying the canopy height at pixel level with the 10cm x 10cm pixel size for each pixel in a plot. All calculated prism pixel volumes were extracted at the plot size and then aggregated by summing to calculate each plot's canopy volume. Filtering operation of canopy volume by harvest number was performed due to the different crops measured in each destructive sampling.

2.4.4. Canopy 3D Profile Index

Since LiDAR gives a 3D representation of the canopy, the 3D profile index was calculated as a third feature. The formula of the index is:

$$3DPI = \sum_{i=0}^{i=n} \left(\frac{p_i}{p_t} \, e^{k \frac{P_{cs}}{P_t}} \right) \qquad (1)$$

where n is the maximum number of vertical layers; pi is the number of points located in layer i, pt is the total LiDAR points for all layers; pcs is the cumulative sum of LiDAR points intercepted above a given layer i; k is the exponential of the correction factor (Jimenez-Berni et al., 2018)

3D profile index was applied to the ground classified LAZ file. The vertical layer thickness of each LAZ file formed by the point clouds used in this study ranges from 51 to 55.5m ellipsoid height. This layer must be sliced into vertically thinner layers to calculate the 3DPI index. The slice size was 0.1 m that should not be smaller than the expected co-registration error, which was approximately 0.05m for this study. The zero values were recognized as NA in programming R. To avoid this, all the NA's were replaced with zero in each slice. The replacement does not cause any under or overestimate because zero is a neutral element in a

sum. Since there were approximately more than 4000-pixel zeros, aggregation was applied using the median, which was less sensitive to the outlier than the mean in the extraction section. The filtering operation was performed for 3DPI by harvest number as explained in previous canopy features.

2.5. Modelling

AGB and LAI were estimated using Multiple linear regression (MLR), Support Vector Machine Regression (SVR) and Random Forest Regression (RFR). The estimation was performed at the plot level. During the field study, 909 samples were collected and calculated at the plot level for modelling. The models were trained and tested discretely to all collected samples and by filtering the samples based on the crop types. When all samples are filtered and grouped by crop types, 221 wheat, 224 sugar beets, 256 soybeans, 208 maize samples were procured.

Before building the models, simple linear regression (LR) was applied to assess the relationship between LiDAR-derived features and response variables.

The machine learning models used for estimation were trained using one of the resampling methods called leave-one-group-out (LOGO) cross-validation(CV) to avoid overestimation. All samples were randomly allocated prior to 70% training and 30% validation set. In LOGO method, the percentage of all training samples randomly divided into groups was specified as 70% training and 30% test set during the process. Each time, 70% of all samples were used to fit the model, and the remaining 30% was used as the test set to estimate the best model for prediction.The best training model derived from LOGO CV process was used to predict the response variables using the validation set. Field measured AGB and LAI were the response variable, and LiDAR-derived features CHM, CV, 3DPI were predictor variables. A cross-validated modelling process was done repeatedly, and the number of repetitions was determined as 10. Tuning parameters required for random forest regression and support vector machines regression were adjusted using cross-validation implemented grid search parameter optimization.

Table 5. Machine Learning Models implemented in this study

Model	Strategy	Method	Tuning parameters
LR	Linear regression	Simple Linear Regression	
MLR	Linear regression	Multiple linear regression	–
RFR	Regression trees	Random forest	*mtry*
SVR	Non-linear regression	Radial basis function SVR	*Sigma* and *C*

2.5.1. Multiple Linear Regression Model

Multiple Linear Regression (MLR) is a parametric machine learning approach used to determine linear relationships between dependent and independent variables in cases where there is more than one predictor variable. MLR was selected due to its simple procedure and fitting potential. The MLR formulas are defined as:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \cdots + \beta_k X_{ik} + \varepsilon_i \qquad (2)$$

where Y_i is the model response, the predicted variable, which are AGB and LAI in this case. β_0 is intercepted. $\beta_1, \ldots, \beta_k$ is called slopes or regression coefficients in statistics. $X_{i1}, \ldots, X_{ik}$ are predictor variables which are LiDAR extracted features CHM, CV, 3DPI from the previous section. ε_i is the residual error term that is assumed to be normally distributed with 0 mean and constant variance (Ali et al., 2015; Civelekoglu et al., 2007). MLR was performed in R software caret package with cross-validation implemented.

2.5.2. Random Forest Model

Random forest (RF) is an ensemble learning model that can be applied to both regression and classification problems (Breiman, 2001). For regression, random forests are created with growing trees based on a random vector, with the tree predictor taking numerical values instead of class labels. The output values are numerical, and it is assumed that the training set is drawn independently of the distribution of the random Y, X vector (Breiman, 2001). The random forest model was chosen because it reduces bias, overfitting, and robustness to outliers and noise and gives good results with random samples in estimation-related studies(Breiman, 2001).

A random forest algorithm was implemented using R and the "caret" package in this study. The model was trained and tested on both all samples and filtered samples by crops separately. The number of regression trees ($ntree$) hyper-parameter was set to 100, and $mtry$ parameter, a randomly selected predictor number at each split, was tuned and assigned as 2 or 3.

2.5.3. Support Vector Regression

Support vector machines (SVMs) are a non-parametric approach. They can be used for both classification and regression problems. The SVM algorithm creates a multidimensional hyperplane that separates the two classes by maximizing the margin between the two data sets. However, it is not always possible to separate the given dataset linearly (Zacarias et al., 2013). Support vector regression (SVR) converts the nonlinear regression problem into a linear problem using kernel functions to map the original data into a higher dimensional space (Cristianini et al., 2000). The radial basis kernel function (RBF) produces a robust result and a smaller number of input parameters. (Chen et al., 2011; Cristianini et al., 2000). RBF was chosen to fit the data. The regularization parameter (C) of radial basis kernel shows how well the model can handle the error was tuned and set as 0.25. Sigma(σ) varied between 1 and 6.5 as in all calculation. The SVR model was applied using R programming and the caret package.

2.6. Accuracy Assesment

The modelling application was performed on all samples of response field measured AGB and LAI and LiDAR-derived predictor variables CHM, CV and 3DPI separately to see how the performance differs between models and the response variables. To apply models on the four crops types, wheat, sugar beet, soybean and maize, we filtered the predictors by crop types. Estimation was done with the three machine learning methods, MLR, RFR, and SVR.

Two statistical criteria were used as evaluation metrics, the coefficient of determination (R^2) and Root Mean Square Error (RMSE), to determine how well the model predicts and the model performance on new data (Han et al., 2019). The comparison among the model performance will be the answer to the research questions mentioned in section 1**.** The formula of performance metrics are;

$$R^2 = 1 - \frac{\sum_{i=0}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=0}^{N}(y_i - \bar{y}_i)^2} \qquad (3)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=0}^{N}(y_i - \hat{y}_i)^2} \qquad (4)$$

where N is the total sample size, y_i is the ith measured AGB and LAI the sample, $\hat{y}_i$ is the ith predicted value, and $\bar{y}_i$ is the ith mean measured value.

3. RESULTS

The results obtained from the two sensors will be explained comparatively for the two parameters AGB and LAI and two LiDAR sensors.

3.1. Relationship Between Predictor Variables and Response

Linear regression was applied to see how the features extracted by LiDAR sensor data relate to AGB and LAI. The results shown in Table 6 represents the coefficient of determination (R^2) between LiDAR-derived CHM, CV, 3DPI and field-measured AGB and LAI. As shown in the table, over 70% of the variation in AGB/LAI was explained by the CHM and CV, both LiDAR extracted features and crop parameters. 3DPI and the AGB and LAI do not show a linear relationship seen in Figure 4. Therefore, the R^2 of 3DPI ~ AGB and LAI is lower than the other features which are around 0.50. Additionally, TLS derived 3DPI showed a lower correlation with both AGB and LAI, where other features had a better result than UAV-LS.

Table 6 The coefficient of determination (R^2) of the linear relationship between predictor variables and response

	UAV-LS		TLS	
	AGB	LAI	AGB	LAI
CHM	0.71	0.73	0.72	0.71
CV	0.71	0.73	0.73	0.71
3DPI	0.50	0.52	0.48	0.48

Considering the canopy volume was calculated depending on the canopy height, the correlation ratios were expected to be the same. Only TLS derived CHM and CV were associated with AGB at a different rate of R^2 ($CHM, R^2 = 0.72, CV\ R^2 = 0.73$)

As shown in Figure 4, the slope of some dates causes deviations or clustering around the regression line. Since more than one crop is harvested and measured in each field measurement, their distribution may differ from each other.

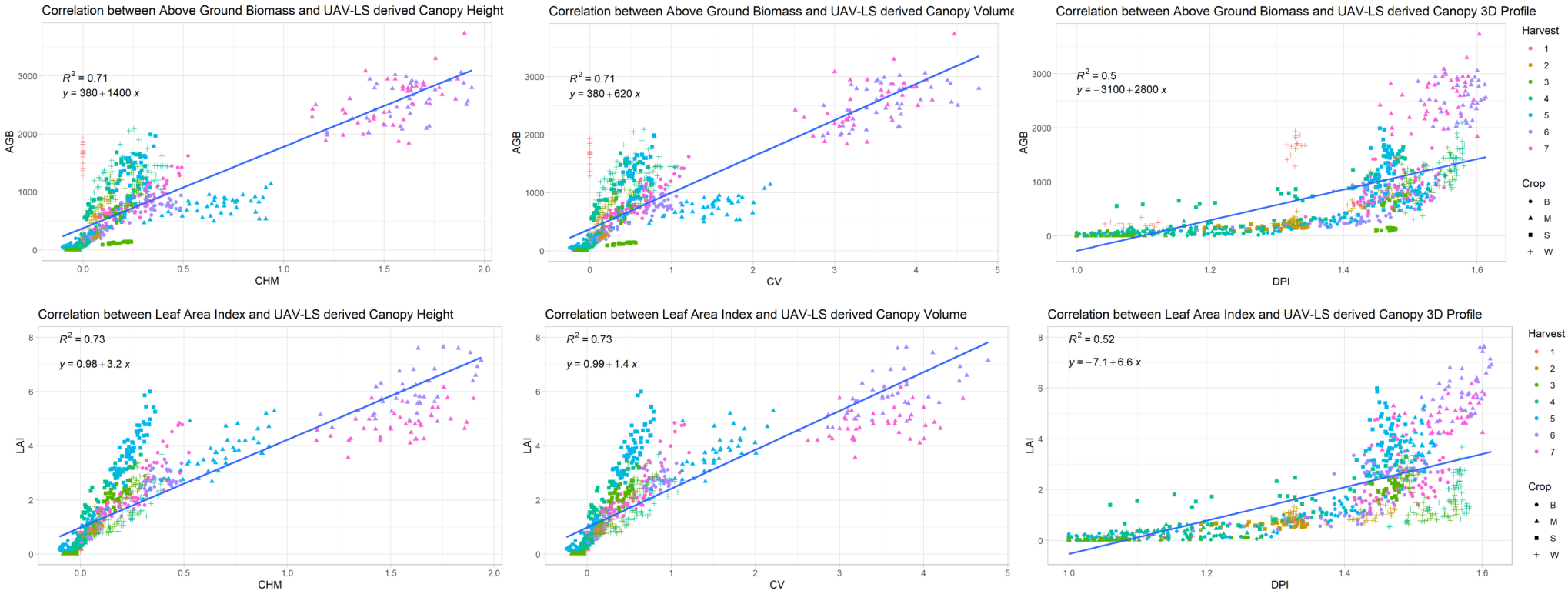


Figure 4 The correlation between UAV-LS derived features and AGB and LAI, the colours represent the harvest date and the shapes indicate the crop types

3.2. Predictability Performance of AGB and LAI with all crop samples

In this section, the result of the predictability performance of AGB and LAI with all crop samples will be presented. Then, the predictability performance of AGB and LAI with which sensor, and the final predictability performance of AGB and LAI with which model will be explained. The results can be seen in Table 7.

Table 7. The Result of Predictability Performance of AGB and LAI with all crop samples

	Above Ground Biomass			Leaf Area Index		
Model	Parameter	RMSE, g	R-squared	Parameter	RMSE, m2	R-squared
			UAV Laser Scanner			
MLR		342	0.80**		0.75**	0.81**
RFR	mtry=3	248	0.89**	mtry=2	0.54**	0.90**
SVR	sigma = 3.55, C = 0.25	307	0.84**	sigma = 3.46, C = 0.25	0.59**	0.88**
			Terrestrial Laser Scanner			
MLR		341	0.80**		0.84**	0.76**
RFR	mtry=3	276	0.87**	mtry=2	0.66**	0.85**
SVR	sigma = 3.47, C = 0.25	299	0.85**	sigma = 4.11, C = 0.25	0.64**	0.86**

AGB and LAI were estimated using three machine learning models using all samples derived from field measured AGB and LAI and LiDAR features. The predictability performance of AGB with features obtained from two LiDAR sensors ranges from 0.80 to 0.89 R^2. The model with the best prediction performance applied with both UAV-LS and TLS features to estimate AGB was RFR (UAV-LS $R^2 = 0.89, RMSE = 248g$, TLS $R^2 = 0.87, RMSE = 276g$). SVR gave a very close result to RFR as well and outperformed MLR. Although the R^2 values of MLR model applied with TLS features are very close to the MLR performed with UAV-LS features; the RMSE value is slightly lower.

The overall predictability performance of LAI estimation has R^2 ranges between 0.76 and 0.90 at the plot scale. As shown in Table 7, the models applied with TLS features showed lower performance than UAV-LS. Among the models, MLR underperformed LAI prediction with a lower R^2 and higher $RMSE$ compared to the RFR and SVR. The best prediction performance was observed RFR ($R^2 = 0.90, RMSE = 0.54m^2/m^2$) with UAV-LS features and SVR $R^2 = 0.86, RMSE = 0.64m^2/m^2$) with TLS derived features.

According to the results, the prediction performed with the features obtained using the UAV-LS gave a slightly better result than the TLS on the LAI estimation. In particular, the R^2 value obtained for LAI estimated by TLS is considerably lower than the value obtained with UAV-LS.. Considering the predictive performance of the machine learning models, overall, RFR and

SVR models that can work with both linear and nonlinear data have performed better than MLR. The models applied with the features derived from UAV-LS and TLS sensors also gave approximately similar results.

Overall, the prediction models applied with all crop samples and LiDAR-derived features can estimate AGB and LAI with the R^2 values that are above 70%. In order to see whether the ability of models to predict AGB and LAI is successful in crop specificity and to investigate whether data collection for one crop can help predict another, each crop sample filtered according to their harvest date and crop types. The predictability of AGB and LAI specific to crops and the predictive performance of models will be discussed in the next section.

3.3. Predictability Performance of AGB and LAI by Crops

The extracted features were filtered by crop types to apply the models separately. The Predictability Performance of AGB and LAI will be explained based on four crop types with different canopy properties wheat, sugar beet, soybean and maize.

Wheat

The results obtained from the wheat crop that can be seen in Table 8. According to wheat AGB estimation results differ considerably between models. The best result was obtained by RFR ($R^2 = 0.83, RMSE = 193g$) applied with UAV-LS derived features and SVR ($R^2 = 78, RMSE = 236g$) applied with TLS derived features. Important to realize that the AGB estimation results obtained with multiple linear regression (MLR) are considerably lower than other models for both sensors (UAV-LS $R^2 = 0.39, RMSE = 364g$, TLS $R^2 = 0.46, RMSE = 371g$). Due to the wheat crops being in early growth and were still tiny in H1 and H2, the sensors could have measured objects at the same height or taller. Another reason for the high $RMSE$ is that the planting dates of the plants are different. In wheat AGB values shown in Table 4; the difference between the minimum and maximum values in harvest number 1 (H1) and 2 (H2) are relatively high. The crops planted at different times may form outlier. Since the linear models are not flexible to the outliers, it could cause a high $RMSE$.

Table 8 Predictability Performance of Wheat AGB and LAI

Wheat						
	Above Ground Biomass			Leaf Area Index		
Model	Parameter	RMSE, g	R-squared	Parameter	RMSE, m2/m2	R-squared
UAV Laser Scanner						
MLR		364	0.39**		0.41**	0.73**
RFR	mtry=2	193	0.83**	mtry=2	0.38**	0.77**
SVR	sigma = 1.98, C = 0.25	204	0.80**	sigma = 1.20, C = 0.25	0.29**	0.86**
Terrestrial Laser Scanner						
MLR		371	0.46**		0.49**	0.46**
RFR	mtry=3	248	0.76**	mtry=2	0.50**	0.61**
SVR	sigma = 1.42, C = 0.25	236	0.78**	sigma = 2.28, C = 0.25	0.52**	0.58**

R^2 results from wheat LAI estimation with three models range between 0.73 and 0.86 for UAV-LS and 0.46 to 0.61 for TLS. SVR best performs the predictability performance of LAI with UAV-LS derived features ($R^2 = 0.86, RMSE = 0.29 m^2/m^2$), RFR outperformed SVR and MLR with TLS features ($R^2 = 0.61, RMSE = 0.50 m^2/m^2$). The prediction performed with the features derived from UAV-LS is reasonably higher than TLS features for LAI estimation.

Sugarbeet

Table 9 represents an overview of the sugar beet AGB and LAI prediction results. The coefficient of determination for estimating sugar beet AGB ranges between 0.54 to 0.76 with the features obtained from the two sensors. Based on the R^2 and RMSE, the best AGB estimation was performed by applying MLR applied with UAV-LS data ($R^2 = 0.60, RMSE = 341 g$), SVR applied with TLS data ($R^2 = 0.76, RMSE = 263 g$). SVR showed the lowest result with UAV-LS data and best performance applied with TLS. R^2 and $RMSE$ values obtained from the models applied with TLS features are higher than those applied with UAV features.

Table 9 Predictability Performance of Sugar Beet AGB and LAI

Sugar Beet						
	Above Ground Biomass			Leaf Area Index		
Model	Parameter	RMSE, g	R-squared	Parameter	RMSE, m2/m2	R-squared
UAV Laser Scanner						
MLR		341	0.60**		0.46**	0.91**
RFR	mtry=2	357	0.56**	mtry=2	0.45**	0.91**
SVR	sigma = 2.30, C = 0.25	366	0.54**	sigma = 2.05, C = 0.25	0.46**	0.91**
Terrestrial Laser Scanner						
MLR		347	0.58**		0.55**	0.83**
RFR	mtry=2	287	0.71**	mtry=3	0.49**	0.86**
SVR	sigma = 1.83, C = 0.25	263	0.76**	sigma = 2.49, C = 0.25	0.53**	0.84**

The results, as shown in the table, indicate that the predictability performance of LAI is better than AGB, where the R^2 value is over 0.83, and $RMSE$ is around 0.50 m^2/m^2. RFR showed the best estimation performance using the features derived from both UAV-LS and TLS. The MLR and SVR results associated with sugarbeet LAI estimation using UAV-LS derived features are seemed to be similar. Still, the decimals are somewhat different and can be seen in Figure 5.

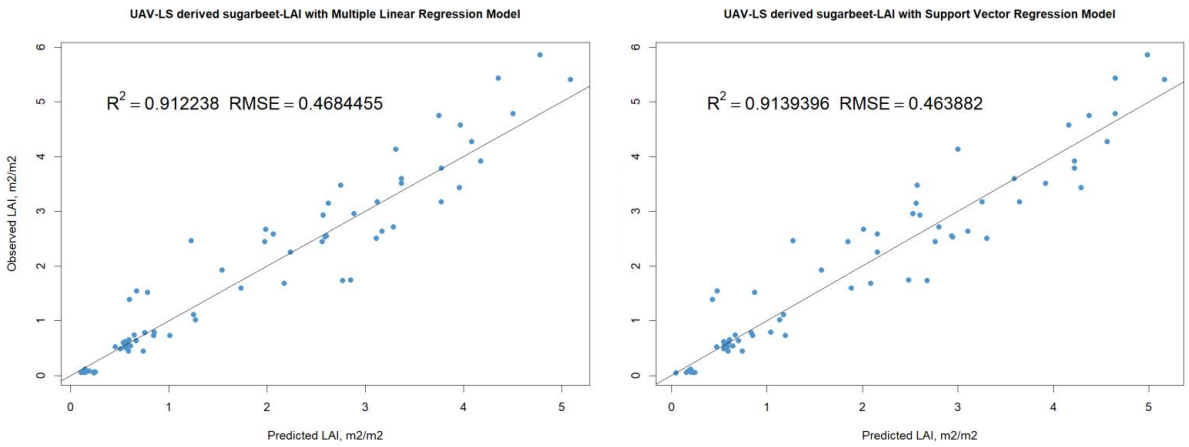


Figure 5 The MLR and SVR results with sugarbeet-LAI estimation using UAV-LS features

Overall, the modelling with TLS derived features gives better results than UAV-LS on sugarbeet AGB estimation based on the R^2 and $RMSE$. Besides, the best LAI prediction was performed with the models applied using UAV-LS features.

Soybean

The results obtained from the soybean crop are shown in Table 10. The coefficient of determination for estimating soybean AGB ranges 0.86 and 0.91, with $RMSE$ of 120g and 158g,

respectively. Looking at the AGB estimation results, the SVR model applied with UAV-LS data ($R^2 = 0.93, RMSE = 96g$), and the RFR model applied with TLS data(($R^2 = 0.90, RMSE = 125g$) showed better performance than other models with a high R^2 and low $RMSE$. The best prediction performance was seen with the SVR model with a considerably low RMSE.

Considering the overall LAI estimates, the models applied with TLS features performed hardly better than the UAV-LS features. MLR model applied with TLS data showed the best performance with a high R^2 and relatively low $RMSE$ ($R^2 = 0.92, RMSE = 0.30\ m^2/m^2$). Likewise, among the models applied with UAV-LS, SVR model ($R^2 = 0.92, RMSE = 0.34\ m^2/m^2$) outperformed the RFR and MLR.

Table 10 Predictability Performance of Soybean AGB and LAI

Soybean						
	Above Ground Biomass			Leaf Area Index		
Model	Parameter	RMSE, g	R-squared	Parameter	RMSE, m2/m2	R-squared
UAV Laser Scanner						
MLR		114	0.91**		0.36**	0.91**
RFR	mtry=3	108	0.92**	mtry=3	0.38**	0.90**
SVR	sigma = 4.82, C = 0.25	96	0.93**	sigma = 1.42, C = 0.25	0.34**	0.92**
Terrestrial Laser Scanner						
MLR		130	0.89**		0.30**	0.92**
RFR	mtry=2	125	0.90**	mtry = 2	0.32**	0.91**
SVR	sigma = 1.76, C = 0.25	138	0.88**	sigma = 1.60, C = 0.25	0.31**	0.92**

In summary, these results show that the models applied with UAV-LS derived features in AGB prediction and models performed with TLS derived features in LAI prediction showed better performance according to the R^2 and the $RMSE$.

Maize

As Table 11 shows, the predictability performance of maize AGB and LAI results of all models applied with the features obtained from the LiDAR sensors are very similar and consistent. Even though the R^2 and RMSE values seem to be the same. The decimal parts differ from each other. According to the prediction performance of AGB, the models applied with TLS data performed better than models applied with UAV-LS features. Among the models applied with both UAV-LS and TLS data, the RFR model outperformed the other models (UAV-LS $R^2 = 0.94, RMSE = 274g$, TLS $R^2 = 0.95, RMSE = 245g$). The SVR model also followed the RFR, outperforming the MLR model.

According to the LAI part of Table 11, the models applied using UAV-LS derived features outperform the TLS. Considering the results derived using UAV-LS features, MLR and RFR model showed somewhat similar performance based on the R^2. However, it can be seen that MLR outperformed RFR model with a lower $RMSE$ ($R^2 = 0.94, RMSE = 0.56 m^2/m^2$). Among the models applied using TLS derived features on maize LAI estimation was observed with, the RFR model (TLS $R^2 = 0.94, RMSE = 0.60 m^2/m^2$).

Table 11 Predictability Performance of Maize AGB and LAI

Maize						
	Above Ground Biomass			Leaf Area Index		
Model	Parameter	RMSE, g	R-squared	Parameter	RMSE, m2/m2	R-squared
UAV Laser Scanner						
MLR		294	0.93**		0.56**	0.94**
RFR	mtry=2	274	0.94**	mtry=2	0.58**	0.94**
SVR	sigma = 5.30, C = 0.25	275	0.94**	sigma = 5.46, C = 0.25	0.62**	0.93**
Terrestrial Laser Scanner						
MLR		267	0.94**		0.65**	0.93**
RFR	mtry=2	245	0.95**	mtry = 2	0.59**	0.94**
SVR	sigma = 4.96, C = 0.25	258	0.95**	sigma = 4.05, C = 0.25	0.69**	0.91**

Based on the information obtained from the predictability performance of maize AGB and LAI results, all models perform almost identical.

4. DISCUSSION

The LiDAR has been demonstrated for the non-destructive phenotyping of AGB and LAI in an increasing number of studies (Deery et al., 2020; Eitel et al., 2016; Harkel et al., 2019; Jimenez-Berni et al., 2018; Li et al., 2015; Walter et al., 2019). In the context of the study, the crop parameters LAI and AGB were estimated using UAV-LS and TLS derived features with the machine learning methods MLR, RFR and SVR over four crops: wheat, sugar beet, soybean, and maize.

The models applied with all crop samples, UAV-LS and TLS derived features, produced very similar results in AGB prediction. Tilly et al. (2014) demonstrated the sufficiency of TLS (Riegl VZ-1000) for AGB estimation. However, Brede et al. (2017) have stated that although the UAV-LS's perspective is vertical, it still reaches the ground level at similar rates as the TLS for all returns. This may have caused the UAV-LS to perform as well as TLS in AGB estimation. In the LAI estimation, UAV-LS performed better than TLS. Wang et al. (2020) and Tang et al. (2014)studied several LiDAR sensors to estimate LAI and pointed out that field-

measured LAI estimates showed the best agreement with UAV-LS derived data at the plot scale. Although the multi-return and side-looking capabilities of the TLS (Brede et al., 2017) ensured as advantageous to detect all part of crops, the fact that the UAV-LS can provide vertical structural information of vegetation (Wang et al., 2017) may have enabled crop leaves and top to be measured more clearly.

The machine learning models differed in the prediction of AGB with all samples and crop-specific modelling. While nonlinear models RFR and SVR performed better in modelling with all crop samples modelling, linear model MLR came to the fore in crop-specific calculations. As shown in Figure 4 The correlation between UAV-LS derived features and AGB and LAI, the colours represent the harvest date and the shapes indicate the crop types, the slope of some dates causes deviations around the main regression line due to more than one crop was measured in each field measurement. As stated before, their distribution may differ from each other. Each crop can show a linear correlation within itself. However, according to Table 7, all crops AGB and LAI show a 0.80 R^2 linear correlation with the LiDAR derived features. RFR and SVR performed quite comparably to each other. Both models can give powerful results with both linear and nonlinear data. In modelling with all samples, RFR yielded better than SVR and MLR.

In the study specific to crops, the results of the models applied with the features derived from UAV-LS and TLS results may differ between crop samples. Looking at the wheat crop AGB and LAI estimation, although there are minor differences between models in the estimation of both AGB and LAI of the wheat crop, it has been observed that UAV-LS data give slightly better results than TLS.

In Sugarbeet AGB estimation, TLS data have a higher R^2 values and lower $RMSE$, while the model results applied with UAV-LS data stand out in LAI estimation.

It can be said that the R^2 and $RMSE$ values obtained from the model applied with UAV-LS data in soybeans gave better results than TLS, but both sensors gave robust results for LAI, and no comparison could be made.

The models applied with TLS data performed slightly better in the maize AGB prediction. The R^2 values in the LAI estimation were very similar in the two sensors, but due to the $RMSE$ values derived from TLS data models were higher than UAV-LS, it can be said that the models applied with UAV-LS outperformed TLS.

When the model performances are evaluated, even if the crops show a linear relationship, MLR's being more sensitive to outliers has sometimes caused it to lag behind RFR for some of the crops. Wheat and maize AGB estimation were performed the with RFR model better than SVR and MLR. In this study, regression models were applied using active sensor LiDAR-derived features. In another study applied with passive optical sensors, Wang et al. (2016) estimated the wheat biomass using appropriate spectral vegetation indices. They indicated that the RFR model produced more accurate estimates of wheat biomass with $R^2 = 0.79$ than the SVR and ANN models. Moreover, Han et al. (2019) conducted a similar study with maize crop and estimated the maize-AGB using optical UAV remote-sensing data with machine learning methods RFR, MLR and ANN. The limitation of optical UAV causes the uncertainties associated with observation angle, illumination conditions, canopy structure, and leaf morphology characteristics estimating the AGB. Additionally, among the statistical models, they applied, due to the data derived by optical UAV and field measured showed a nonlinear relationship, RFR($R^2 = 0.699$)performed better than the ANN ($R^2 = 0.691$) on the test set and outperformed MLR.

In the sugarbeet AGB estimation, MLR showed high R^2 and low RMSE with the UAV-LS features where SVR outperformed the MLR and RFR with TLS features. Harkel et al. (2019) estimated sugarbeet biomass by applying the linear regression model using UAV-Based LiDAR-derived 3DPI and reported the R^2 of 0.68 which was close to this study's result. Additionally, soybean AGB estimation was performed with the SVR model and outperformed MLR and RFR. In this study, one 1D and two 3D features extracted from LiDAR point clouds, which can provide accurate 3D information of plant shapes and canopy structures (Omasa et al., 2006), were used and overall, an R^2 value over 0.88 was obtained from all applied models. Similarly, Maimaitijiang et al. (2019) estimated soybean AGB using unmanned aerial systems derived photogrammetric point clouds applying the regression models with linear performance and indicated that 3D metrics provide better estimates of AGB than 2D structural metrics with the R^2 of 0.849. Based on these results, it can be said that using the 3D features derived from LiDAR point clouds can provide satisfactory accuracy in AGB prediction.

Considering the general results of AGB estimations, except for the model applied with sugar beet data, over 70% R^2 and approximately 250g RMSEs were observed from all other models, both with all samples and crop-specific model.

Looking at the LAI estimation, the models that applied UAV-LS data performed pretty much better than TLS overall in calculations. Tang et al. (2014) studied several LiDAR sensors to estimate LAI and pointed out that field-measured LAI estimates showed the best agreement with UAV-LS derived data. Multiple flights and viewing perspectives (Brede et al., 2017) from the top of the crops are advantageous for UAV-LS for multiscale LAI estimation(Wang et al., 2020) against TLS. It was observed that the RMSE values of the models applied with TLS data were higher than UAV-LS, even if the R^2 values of the models applied with all crop samples using UAV-LS and TLS features were very similar to predict LAI. In the machine learning models for LAI estimation, the nonlinear model performed considerably better than the linear model with applied all samples. According to the R^2 and $RMSE$ while the best LAI estimation was performed with the RFR model with the UAV-LS features, the SVR model achieved the best performance with TLS derived features.

In the crop-specific modelling analyses, according to the R^2 and $RMSE$ values, except soybean crop, LAI estimates performed with the models using UAV-LS features outperformed the estimation made with TLS. R^2values in soybeans are almost the same for two sensors. It can be said by looking at RMSE values that LAI estimation made with TLS features shows better performance than UAV-LS.

Herewith, it can be said that LAI can be estimated by statistical methods using both UAV-LS and TLS derived data. Similarly to this study, Hosseini et al. (2015) used airborne radar remote sensing to estimate LAI for maize and soybean with active remote sensing data. He stated that depending on the radar band; LAI could be estimated successfully (C-band, maize $R^2 = 0.688$ HH–HV and e $R^2 = 0{,}656$ VV–HV, soybean $R^2 = 0{,}64$ for both HH–HV and VV–HV). LAI can also and often be predicted using passive sensors with physical and statistical methods. In statistical LAI estimations made using the optical remote sensing method, the statistical relationships between vegetation indices (VIs) and LAI are explored to estimate LAI (Potithep et al., 2013; Turner et al., 1999; Wang et al., 2005). For instance, Li et al. (2016) evaluated the RFR method for predicting grassland LAI using ground measurements and Landsat images. They demonstrated that if a good choice can be made between vegetation indices derived from optical remote sensing data, LAI can be predicted strongly with the RFR model. In another study of Gupta et al. (2006) in wheat and chickpea LAI estimation using Landsat images, the relationship between hyperspectral vegetation indices and LAI was observed for wheat (R^2 in 0.86–0.97 range) and for chickpea (R^2 in 0.85–0.88 range). Based on the result, wheat LAI

estimation can be done more successfully with Landsat images than LiDAR point clouds derived features.

In addition to all of this, as stated before, the study was carried out with the data collected from all crop samples to see the model performance, and robust results were obtained for both parameter estimation. Crop-specific estimations were also conducted to investigate whether collecting data for one crop could help predict another and to see if the ability of models to predict AGB and LAI was successful in crop specificity. Results show that AGB estimation also yielded very good R^2 values from soybean and maize and even wheat, but sugar beet results were lower than other crops. Likewise, in the LAI estimation, soybean, maize, and sugar beet had good results, while wheat showed lower R^2 values than other crops. Based on this, it can be said that the parameter estimation performed with samples collected from a crop cannot be used to predict another crop parameter for this study. The reason for the high R^2 and low $RMSE$ obtained from the models applied with all samples is due to the fact that some crop samples give remarkable estimation results.

5. CONCLUSIONS and RECOMMENDATIONS

In this context, the crop parameters LAI and AGB estimated using UAV-LS and TLS derived features CHM, CV and 3DPI with machine learning methods over four crop types with different canopy properties; wheat, sugarbeet, soybean and maize. Finally, the research questions mentioned in the introduction section can be answered.

"How accurately can we obtain the crop parameters LAI and AGB with features derived from LiDAR sensors with machine learning methods?"

In models with the features obtained from LiDAR data, according to R^2 and $RMSE$ values of AGB and LAI estimation results obtained from the study with all crop samples were entirely satisfactory. The predictability of AGB was over 0.80 R^2, and the predictability of LAI was above 0.76 R^2. However, in crop-specific regression models of the Wheat LAI estimation and sugarbeet, AGB estimation failed, R^2 was below 0.70.

1. *How does performance differ between the selected machine learning models?*

In the study performed with all crop samples, it was observed that RFR performed better than SVR and MLR and SVR performed relatively better than MLR based on the R^2 and $RMSE$. RFR and SVR gave very similar results to each other with some crops. In the crops

specific study, the performances of the models in AGB and LAI predictions differ. A generalization cannot be made. In conclusion, it can be said that the models RFR and SVR can handle both linear and non-linearly correlated datasets perform better than the linear model MLR.

2. *How does performance differ between AGB and LAI?*

In the models performed with all crop samples' field measured and LiDAR-derived features, the possibility emerged that AGB and LAI could be predicted at a satisfactory rate.

R2 results higher than 0.80 for AGB estimation, 0.76 for LAI were obtained in models applied with all samples. Both parameters could be predicted quite well with all samples. However, a generalization cannot be made for crop-specific modelling.

3. *How does performance differ between crop types?*

In crop-specific calculations, the Wheat AGB prediction was quite successful, when the LAI prediction failed based on the R^2 and $RMSE$ values. While the predictability performance of both crop parameters in soybean and maize was above 0.80 R^2, the R^2 values obtained from the AGB estimation in sugar beet were between 0.54 and 0.76.

4. *How does performance differ between UAV-LS and TLS LiDAR?*

The models applied with the features obtained from UAV-LS and TLS gave similar results when made with all crop samples. However, according to the analysis results made specific to the crops, the models applied with the features obtained from UAV-LS showed a powerful performance in all LAI estimates. Although the results obtained from models applied using TLS and UAV-LS features to estimate AGB were very close, only sugarbeet AGB estimation gave better results with TLS data than UAV-LS.

In this thesis, AGB and LAI were estimated using statistical methods with LiDAR extracted features CHM, CV and 3DP over four crop types. More LiDAR features can be extracted in future studies, and it can be investigated which and how the features affects the crop parameter estimation. Unlike other remote sensing sensors, 3D profile of an object can be generated from the LiDAR sensor. The features can be selected in this direction. As a second recommendation, generating DEMs from LiDAR data is both time-consuming and needs more computational

power. The targeted working time should be planned considering the processing time of LiDAR data.

7. REFERENCES

Ali, I., Greifeneder, F., Stamenkovic, J., Neumann, M., & Notarnicola, C. (2015). Review of Machine Learning Approaches for Biomass and Soil Moisture Retrievals from Remote Sensing Data. *Remote Sensing, 7*. doi:10.3390/rs71215841

Brede, B., Lau, A., Bartholomeus, H. M., & Kooistra, L. (2017). Comparing RIEGL RiCOPTER UAV LiDAR Derived Canopy Height and DBH with Terrestrial LiDAR. *Sensors (Basel), 17*(10). doi:10.3390/s17102371

Breiman, L. (2001). Machine Learning, Volume 45, Number 1 - SpringerLink. *Machine Learning, 45*, 5-32. doi:10.1023/A:1010933404324

Chen, G., & Hay, G. (2011). A Support Vector Regression Approach to Estimate Forest Biophysical Parameters at the Object Level Using Airborne Lidar Transects and QuickBird Data. *Photogrammetric Engineering & Remote Sensing, 77*, 733-741. doi:10.14358/PERS.77.7.733

Civelekoglu, G., Yigit, N. O., Diamadopoulos, E., & Kitis, M. (2007). Prediction of Bromate Formation Using Multi-Linear Regression and Artificial Neural Networks. *Ozone: Science & Engineering, 29*(5), 353-362. doi:10.1080/01919510701549327

Clevers, J. G. P. W., van der Heijden, G., Verzakov, S., & Schaepman, M. E. (2007). Estimating Grassland Biomass Using SVM Band Shaving of Hyperspectral Data. *Photogrammetric Engineering and Remote Sensing 73 (2007) 10, 73*. doi:10.14358/PERS.73.10.1141

Comba, L., Biglia, A., Ricauda Aimonino, D., Tortia, C., Mania, E., Guidoni, S., & Gay, P. (2020). Leaf Area Index evaluation in vineyards using 3D point clouds from UAV imagery. *Precision Agriculture, 21*, 881-896. doi:10.1007/s11119-019-09699-x

Crespo-Peremarch, P., Torralba Pérez, J., Carbonell-Rivera, J. P., & Ruiz, L. (2020). COMPARING THE GENERATION OF DTM IN A FOREST ECOSYSTEM USING TLS, ALS AND UAV-DAP, AND DIFFERENT SOFTWARE TOOLS. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B3-2020*, 575-582. doi:10.5194/isprs-archives-XLIII-B3-2020-575-2020

Cristianini, N., & Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge: Cambridge University Press.

Dan, L., Gu, X., Pang, Y., Chen, B., & Liu, L. (2018). Estimation of Forest Aboveground Biomass and Leaf Area Index Based on Digital Aerial Photograph Data in Northeast China. *Forests, 9*, 275. doi:10.3390/f9050275

Deery, D., Jimenez-Berni, J., Jones, H., Sirault, X., & Furbank, R. (2014). Proximal Remote Sensing Buggies and Potential Applications for Field-Based Phenotyping. *Agronomy, 4*(3), 349-379. Retrieved from https://www.mdpi.com/2073-4395/4/3/349

Deery, D. M., Rebetzke, G. J., Jimenez-Berni, J. A., Condon, A. G., Smith, D. J., Bechaz, K. M., & Bovill, W. D. (2020). Ground-Based LiDAR Improves Phenotypic Repeatability of Above-Ground Biomass and Crop Growth Rate in Wheat. *Plant Phenomics, 2020*, 8329798. doi:10.34133/2020/8329798

Durbha, S., King, R., & Younan, N. H. (2007). Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer. *Remote Sensing of Environment, 107*, 348-361. doi:10.1016/j.rse.2006.09.031

Eitel, J., Magney, T., Vierling, L., Greaves, H., & Zheng, G. (2016). An automated method to quantify crop height and calibrate satellite-derived biomass using hypertemporal lidar. *Remote Sensing of Environment, 187*. doi:10.1016/j.rse.2016.10.044

Fu, Y., He, H. S., Hawbaker, T. J., Henne, P. D., Zhu, Z., & Larsen, D. R. (2019). Evaluating k-Nearest Neighbor (kNN) Imputation Models for Species-Level Aboveground Forest Biomass Mapping in Northeast China. *Remote Sensing, 11*(17), 2005. Retrieved from https://www.mdpi.com/2072-4292/11/17/2005

García-Gutiérrez, J., Martínez-Álvarez, F., Troncoso, A., & Riquelme, J. (2013). *A Comparative Study of Machine Learning Regression Methods on LiDAR Data: A Case Study* (Vol. 239).

González-Sanchez, A., Frausto-Solis, J., & Ojeda, W. (2014). Predictive ability of machine learning methods for massive crop yield prediction. *SPANISH JOURNAL OF AGRICULTURAL RESEARCH*. doi:10.5424/sjar/2014122-4439

Gupta, R. K., Vijayan, D., & Prasad, T. S. (2006). The relationship of hyper-spectral vegetation indices with leaf area index (LAI) over the growth cycle of wheat and chickpea at 3nm spectral resolution. *Advances in Space Research, 38*(10), 2212-2217. doi:https://doi.org/10.1016/j.asr.2003.02.091

Han, L., Guijun, Y., Dai, H.-y., Xu, B., Yang, H., Feng, H., . . . Xiaodong, Y. (2019). Modeling maize above-ground biomass based on machine learning approaches using UAV remote-sensing data. *Plant Methods, 15*. doi:10.1186/s13007-019-0394-z

Harkel, t., Bartholomeus, & Kooistra. (2019). Biomass and Crop Height Estimation of Different Crops Using UAV-Based Lidar. *Remote Sensing, 12*, 17. doi:10.3390/rs12010017

Hosseini, M., McNairn, H., Merzouki, A., & Pacheco, A. (2015). Estimation of Leaf Area Index (LAI) in corn and soybeans using multi-polarization C- and L-band radar data. *Remote Sensing of Environment, 170*, 77-89. doi:https://doi.org/10.1016/j.rse.2015.09.002

Irwan Hariyono, M., & Windiastuti, R. (2018). *Classification of LiDAR Data to Generate Digital Terrain Model*.

Jimenez-Berni, J. A., Deery, D. M., Rozas-Larraondo, P., Condon, A. G., Rebetzke, G. J., James, R. A., . . . Sirault, X. R. R. (2018). High Throughput Determination of Plant Height, Ground Cover, and Above-Ground Biomass in Wheat with LiDAR. *Frontiers in Plant Science, 9*(237). doi:10.3389/fpls.2018.00237

Jin, S., Su, Y., Song, S., Xu, K., Hu, T., Yang, Q., . . . Guo, Q. (2020). Non-destructive estimation of field maize biomass using terrestrial lidar: an evaluation from plot level to individual leaf level. *Plant Methods, 16*(1), 69. doi:10.1186/s13007-020-00613-5

Khosravipour, A., Skidmore, A., & Isenburg, M. (2016). Generating spike-free digital surface models using LiDAR raw point clouds: A new approach for forestry applications. *International Journal of Applied Earth Observation and Geoinformation, 52*, 104-114. doi:10.1016/j.jag.2016.06.005

Kucukonder, H., Boyaci, S., & Akyuz, A. (2016). A modeling study with an artificial neural network: Developing estimation models for the tomato plant leaf area. *TURKISH JOURNAL OF AGRICULTURE AND FORESTRY, 40*, 203-212. doi:10.3906/tar-1408-28

LAStools. (2020). LAStools - efficient LiDAR processing software (version 201003, academic). Retrieved from http://rapidlasso.com/LAStools

Li, W., Niu, Z., Huang, N., Wang, C., Gao, S., & Wu, C. (2015). Airborne LiDAR technique for estimating biomass components of maize: A case study in Zhangye City, Northwest China. *Ecological Indicators, 57*, 486-496. doi:10.1016/j.ecolind.2015.04.016

Li, Z., Xiao-ping, X. I. N., Huan, T., Fan, Y., Bao-rui, C., & Bao-hui, Z. (2016). Estimating grassland LAI using the Random Forests approach and Landsat imagery in the meadow steppe of

Hulunber, China. *Journal of Integrative Agriculture, 15*, 60345-60347. doi:10.1016/S2095-3119(15)61303-X

Loudermilk, E. L., Hiers, J. K., O'Brien, J. J., Mitchell, R. J., Singhania, A., Fernandez, J. C., . . . Slatton, K. C. (2009). Ground-based LIDAR: a novel approach to quantify fine-scale fuelbed characteristics. *International Journal of Wildland Fire, 18*(6), 676-685. doi:https://doi.org/10.1071/WF07138

Magnussen, S., & Boudewyn, P. (1998). Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. *Canadian Journal of Forest Research, 28*(7), 1016-1031. doi:10.1139/x98-078

Magnussen, S., & Boudewyn, P. (2011). Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. *Canadian Journal of Forest Research, 28*, 1016-1031. doi:10.1139/x98-078

Maimaitijiang, M., Sagan, V., Sidike, P., Maimaitiyiming, M., Hartling, S., Peterson, K. T., . . . Fritschi, F. B. (2019). Vegetation Index Weighted Canopy Volume Model (CVMVI) for soybean biomass estimation from Unmanned Aerial System-based RGB imagery. *ISPRS Journal of Photogrammetry and Remote Sensing, 151*, 27-41. doi:https://doi.org/10.1016/j.isprsjprs.2019.03.003

Omasa, K., Hosoi, F., & Konishi, A. (2006). 3D lidar imaging for detecting and understanding plant responses and canopy structure. *Journal of Experimental Botany, 58*(4), 881-898. doi:10.1093/jxb/erl142

Potithep, S., Nagai, S., Nasahara, K. N., Muraoka, H., & Suzuki, R. (2013). Two separate periods of the LAI–VIs relationships using in situ measurements in a deciduous broadleaf forest. *Agricultural and Forest Meteorology, 169*, 148-155. doi:https://doi.org/10.1016/j.agrformet.2012.09.003

Quebrajo, L., Pérez-Ruiz, M., Perez Urrestarazu, L., Martínez, G., & Egea, G. (2018). Linking thermal imaging and soil remote sensing to enhance irrigation management of sugar beet. *Biosystems Engineering*. doi:10.1016/j.biosystemseng.2017.08.013

Singh, P., Pandey, P., Petropoulos, G., Pavlides, A., Srivastava, P., Koutsias, N., . . . Bao, Y. (2020). Hyperspectral remote sensing in precision agriculture: present status, challenges, and future trends. In (pp. 121-146).

Tang, H., Brolly, M., Zhao, F., Strahler, A. H., Schaaf, C. L., Ganguly, S., . . . Dubayah, R. (2014). Deriving and validating Leaf Area Index (LAI) at multiple spatial scales through lidar remote sensing: A case study in Sierra National Forest, CA. *Remote Sensing of Environment, 143*, 131-141. doi:https://doi.org/10.1016/j.rse.2013.12.007

Tang, H., Brolly, M., Zhao, F., Strahlerb, A., Schaafb, C., Gangulyd, S., . . . Dubayah, R. (2014). Deriving and validating Leaf Area Index (LAI) at multiple spatial scales through lidar remote sensing: A case study in Sierra National Forest, CA. *Remote Sensing of Environment, 143*. doi:10.1016/j.rse.2013.12.007

Tilly, N., Hoffmeister, D., Cao, Q., Huang, S., Lenz-Wiedemann, V., Miao, Y., & Bareth, G. (2014). Multitemporal crop surface models: accurate plant height measurement and biomass estimation with terrestrial laser scanning in paddy rice. *Journal of Applied Remote Sensing, 8*(1), 083671. Retrieved from https://doi.org/10.1117/1.JRS.8.083671

Turner, D. P., Cohen, W. B., Kennedy, R. E., Fassnacht, K. S., & Briggs, J. M. (1999). Relationships between Leaf Area Index and Landsat TM Spectral Vegetation Indices across Three Temperate Zone Sites. *Remote Sensing of Environment, 70*(1), 52-68. doi:https://doi.org/10.1016/S0034-4257(99)00057-7

Verma, N., Lamb, D., Reid, N., & Wilson, B. (2016). Comparison of Canopy Volume Measurements of Scattered Eucalypt Farm Trees Derived from High Spatial Resolution Imagery and LiDAR. *Remote Sensing, 8*, 388. doi:10.3390/rs8050388

Verrelst, J., Camps-Valls, G., Muñoz-Marí, J., Rivera, J. P., Veroustraete, F., Clevers, J. G. P. W., & Moreno, J. (2015). Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties - A review. *ISPRS Journal of Photogrammetry and Remote Sensing, 108*, 273-290.

Walter, J. D. C., Edwards, J., McDonald, G., & Kuchel, H. (2019). Estimating Biomass and Canopy Height With LiDAR for Field Crop Breeding. *Frontiers in Plant Science, 10*(1145). doi:10.3389/fpls.2019.01145

Wang, & Fang. (2020). Estimation of LAI with the LiDAR Technology: A Review. *Remote Sensing, 12*(20), 3457. Retrieved from https://www.mdpi.com/2072-4292/12/20/3457

Wang, Nie, Xiaohuan, Luo, & Sun. (2017). Estimating the Biomass of Maize with Hyperspectral and LiDAR Data. *Remote Sensing, 11*, 1-12. doi:10.3390/rs9010011

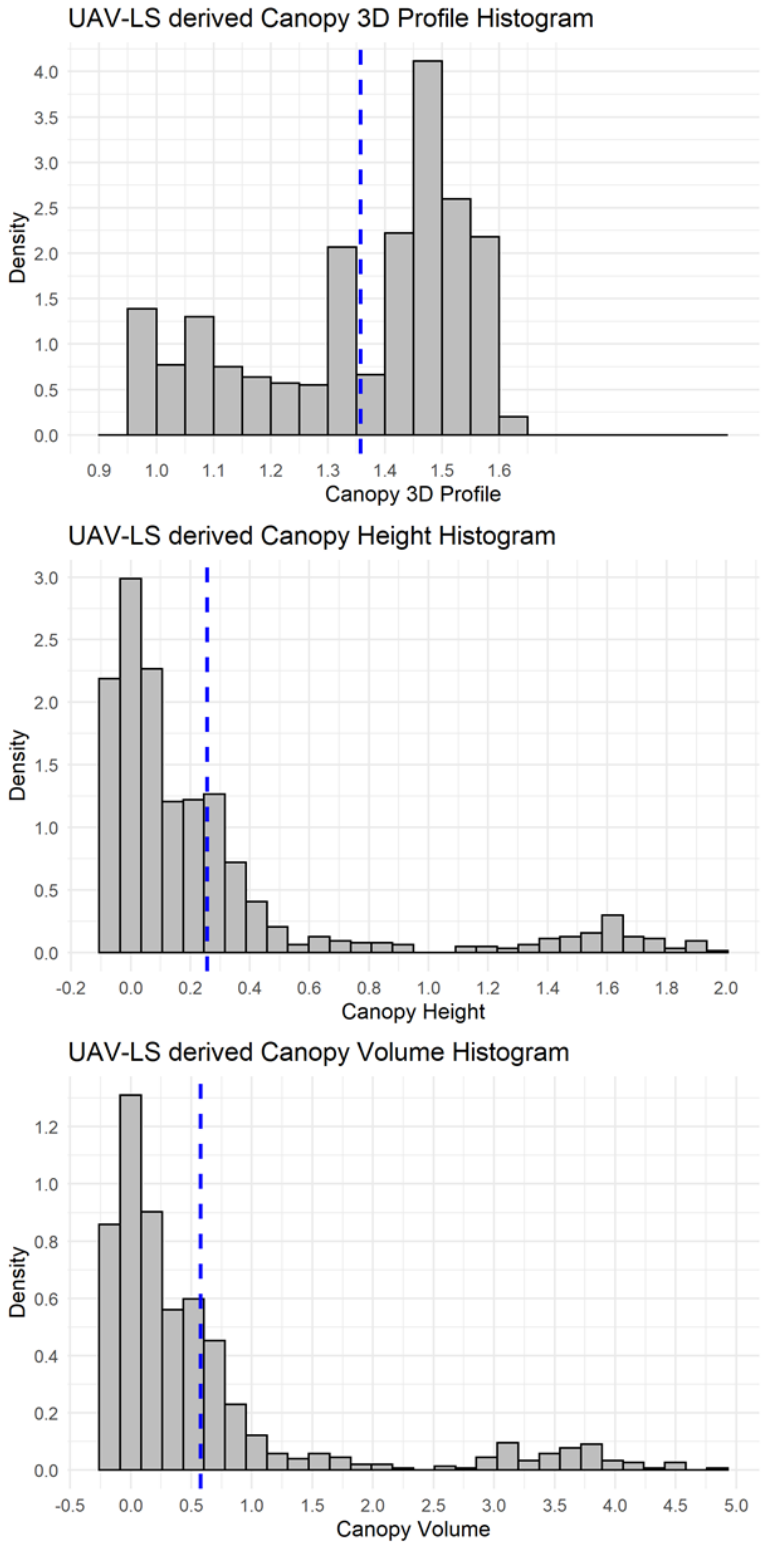
Wang, L. a., Zhou, X., Zhu, X., Dong, Z., & Guo, W. (2016). Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. *The Crop Journal, 4*(3), 212-219. doi:https://doi.org/10.1016/j.cj.2016.01.008

Wang, P., Li, J., Li, J., Li, Z., Schmit, T. J., & Bai, W. (2014). Advanced infrared sounder subpixel cloud detection with imagers and its impact on radiance assimilation in NWP. *Geophysical Research Letters, 41*(5), 1773-1780. doi:https://doi.org/10.1002/2013GL059067

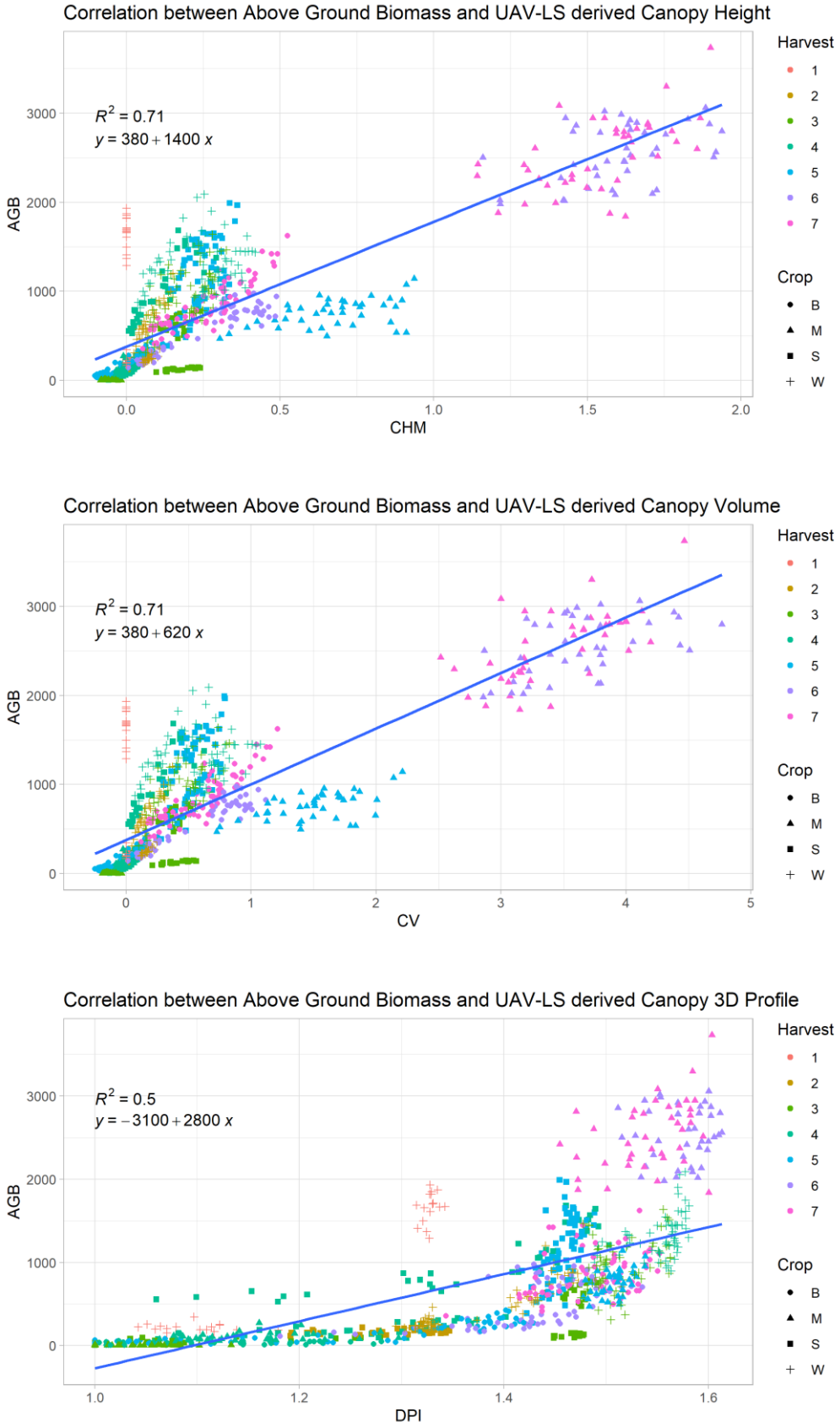
Wang, Q., Adiku, S., Tenhunen, J., & Granier, A. (2005). On the relationship of NDVI with leaf area index in a deciduous forest site. *Remote Sensing of Environment, 94*(2), 244-255. doi:https://doi.org/10.1016/j.rse.2004.10.006

Zacarias, O., & Boström, H. (2013). Comparing Support Vector Regression and Random Forests Modeling for Predicting Malaria Incidence in Mozambique. *International Journal on Advances in ICT for Emerging Regions (ICTer)*. doi:10.1109/ICTer.2013.6761181

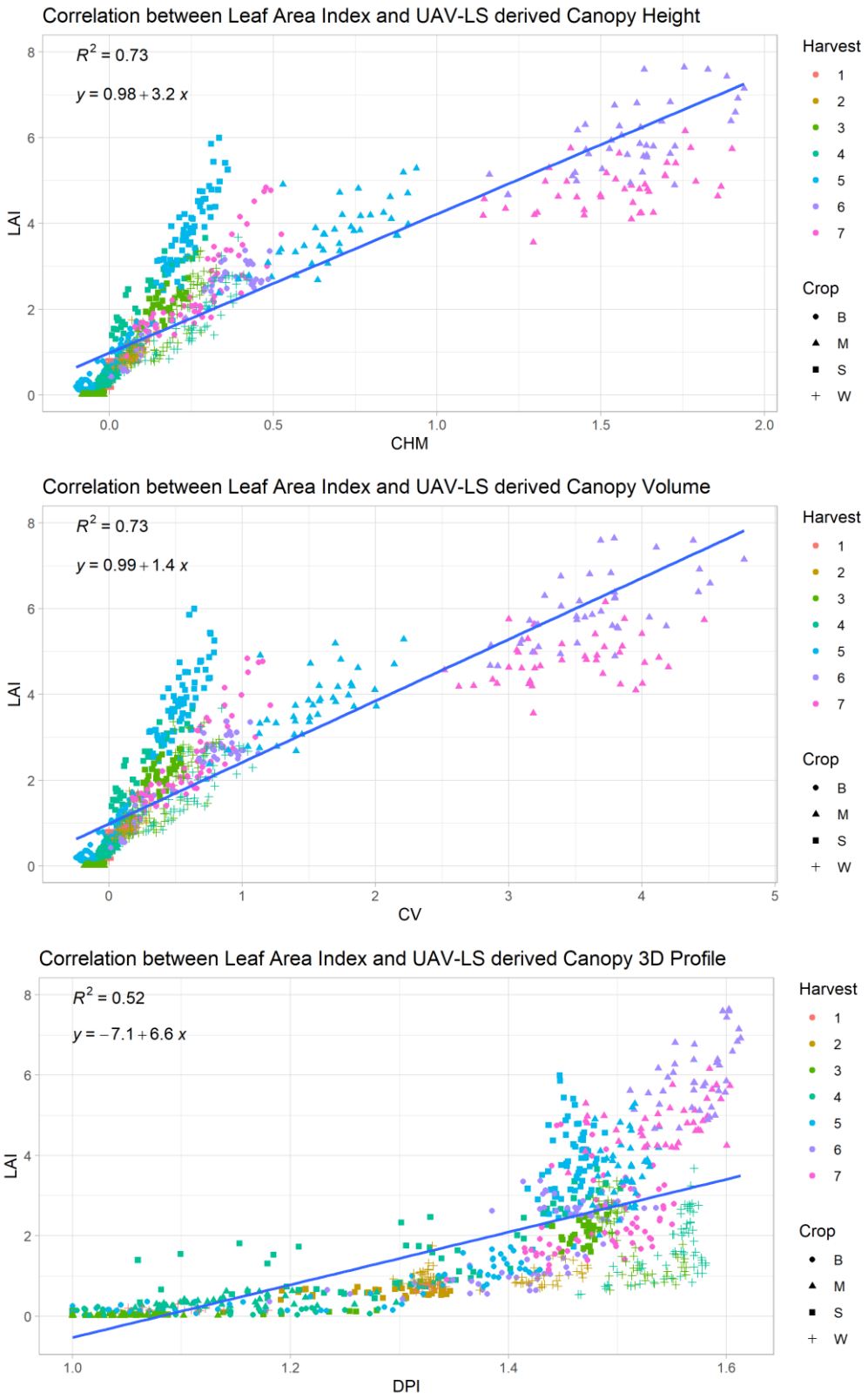
8. ADDITIONAL FIGURES



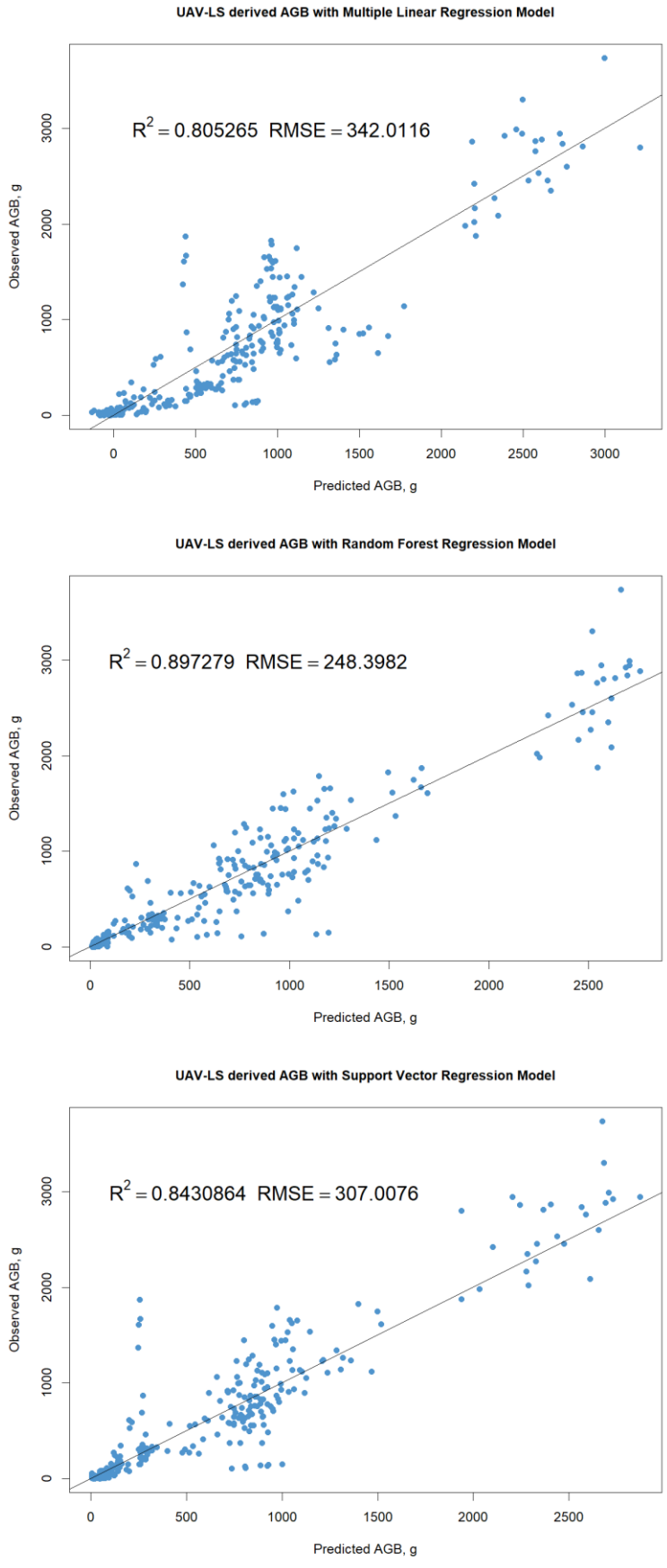
Annex 1 UAV-LS derived Canopy Features Histogram



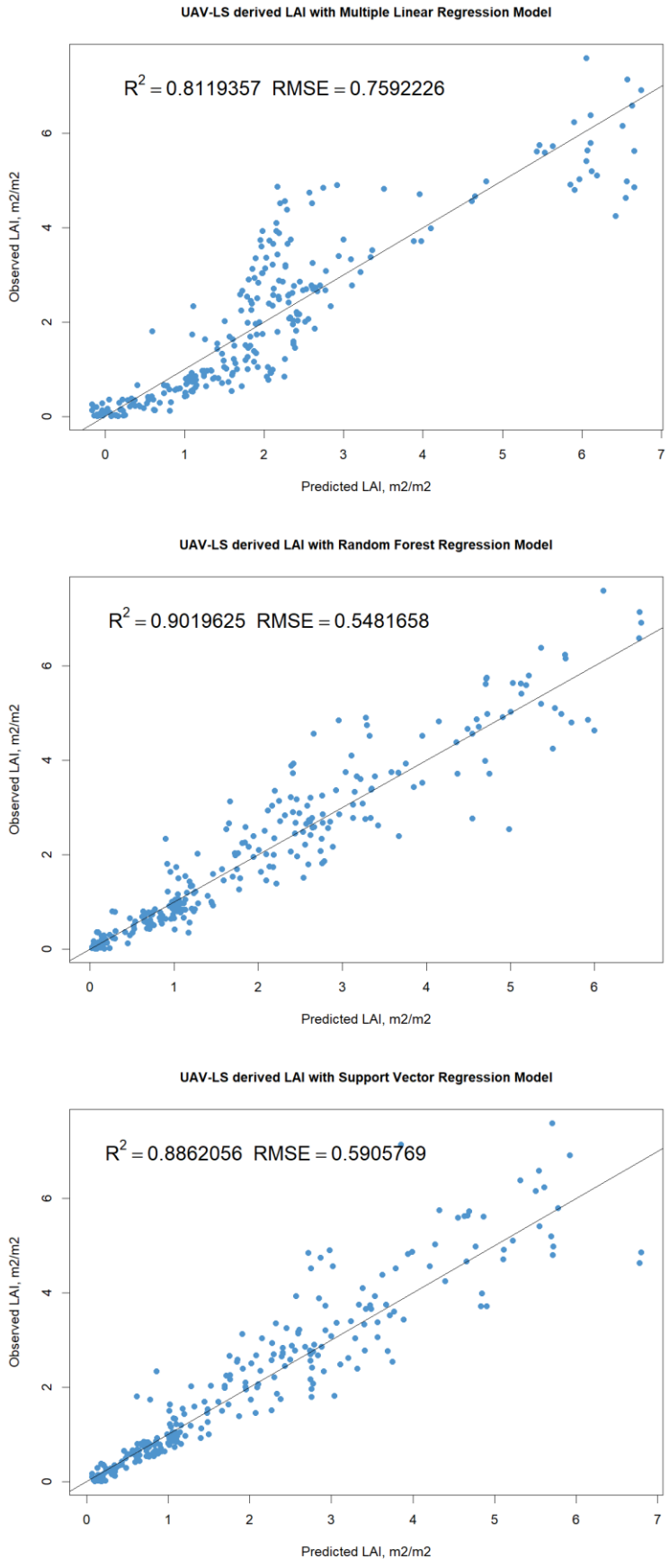
Annex 2 Correlation Between AGB and UAV-LS derived features



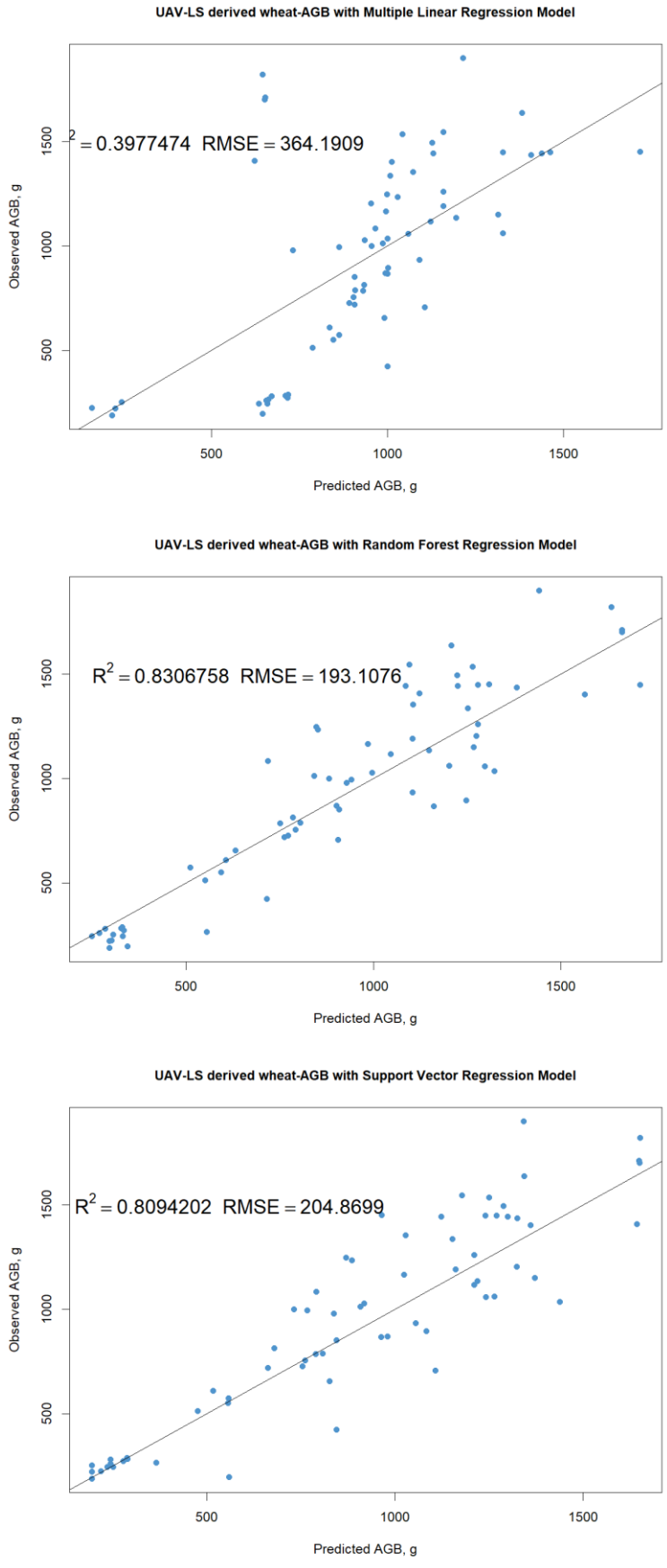
Annex 3 Correlation Between LAI and UAV-LS derived features



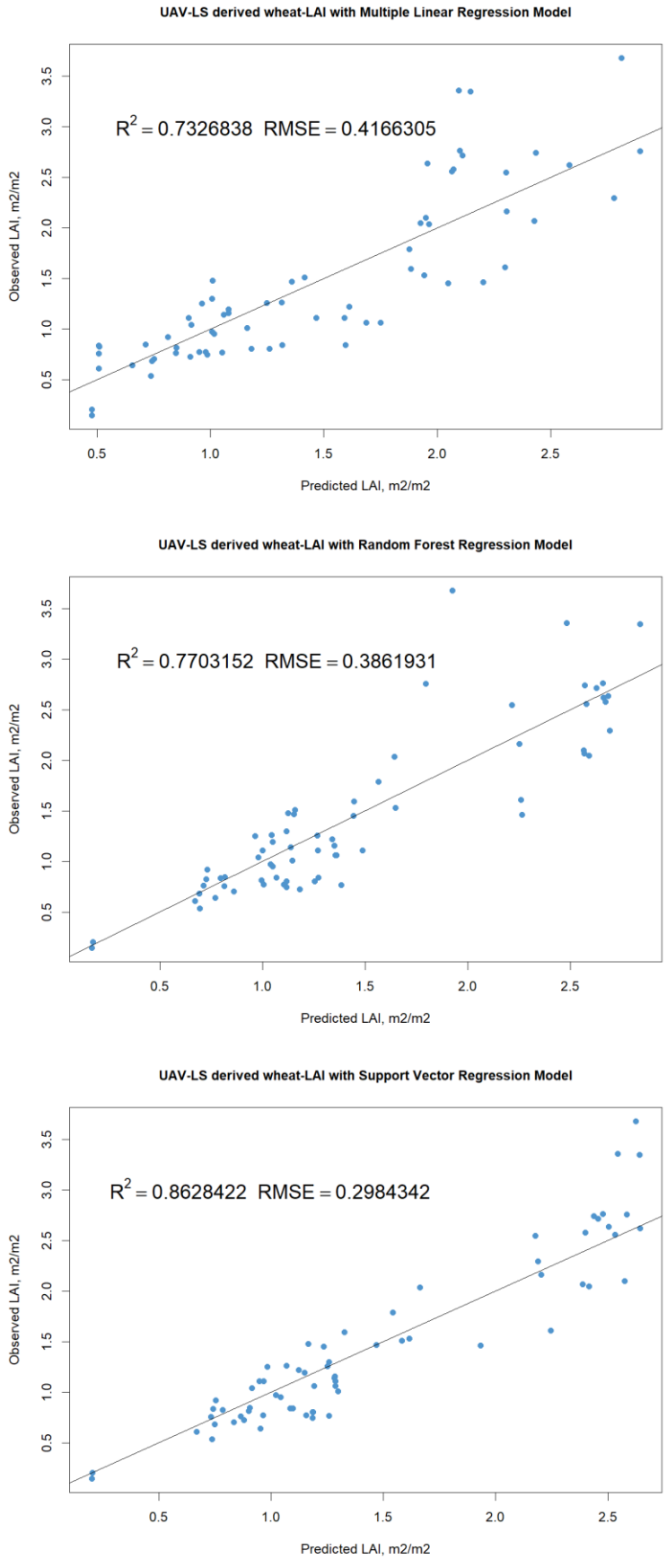
Annex 4 UAV-LS - Predicted and Observed values of AGB with all samples



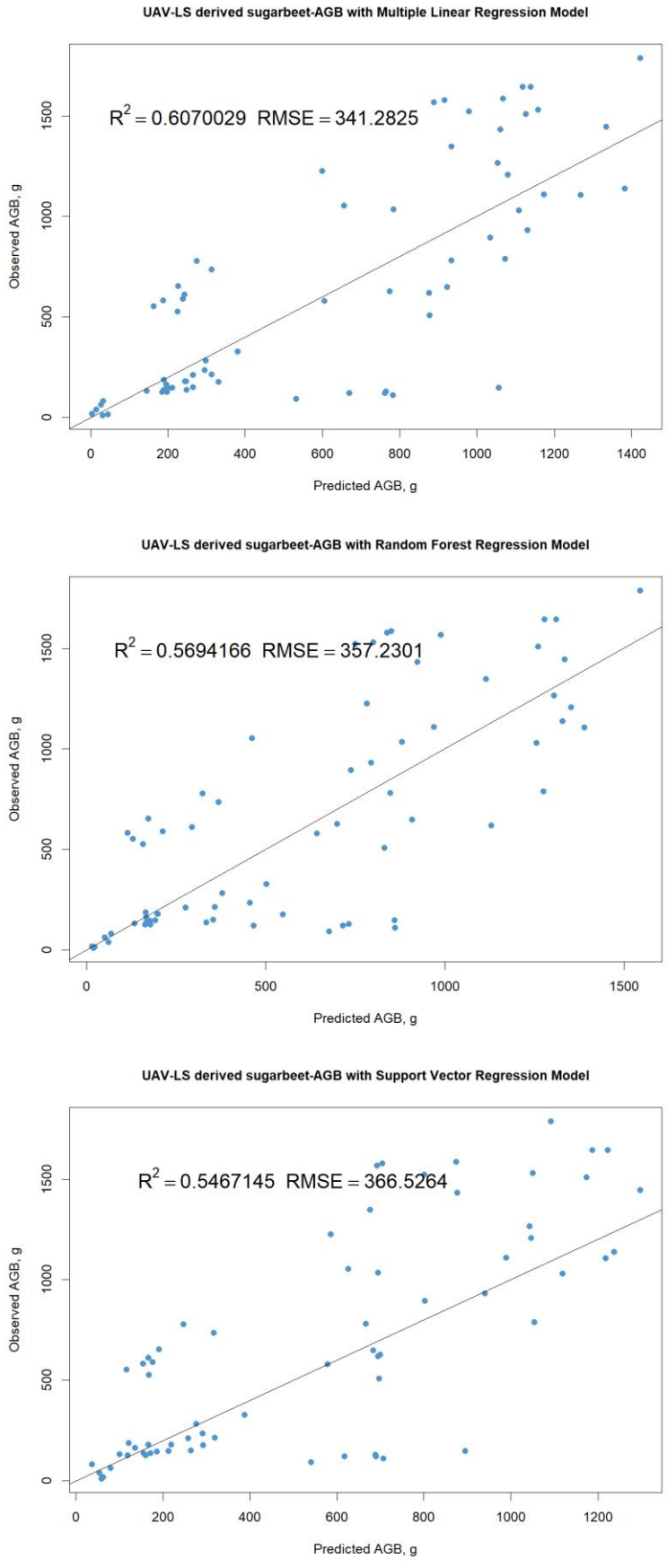
Annex 5 UAV-LS - Predicted and Observed values of LAI with all samples



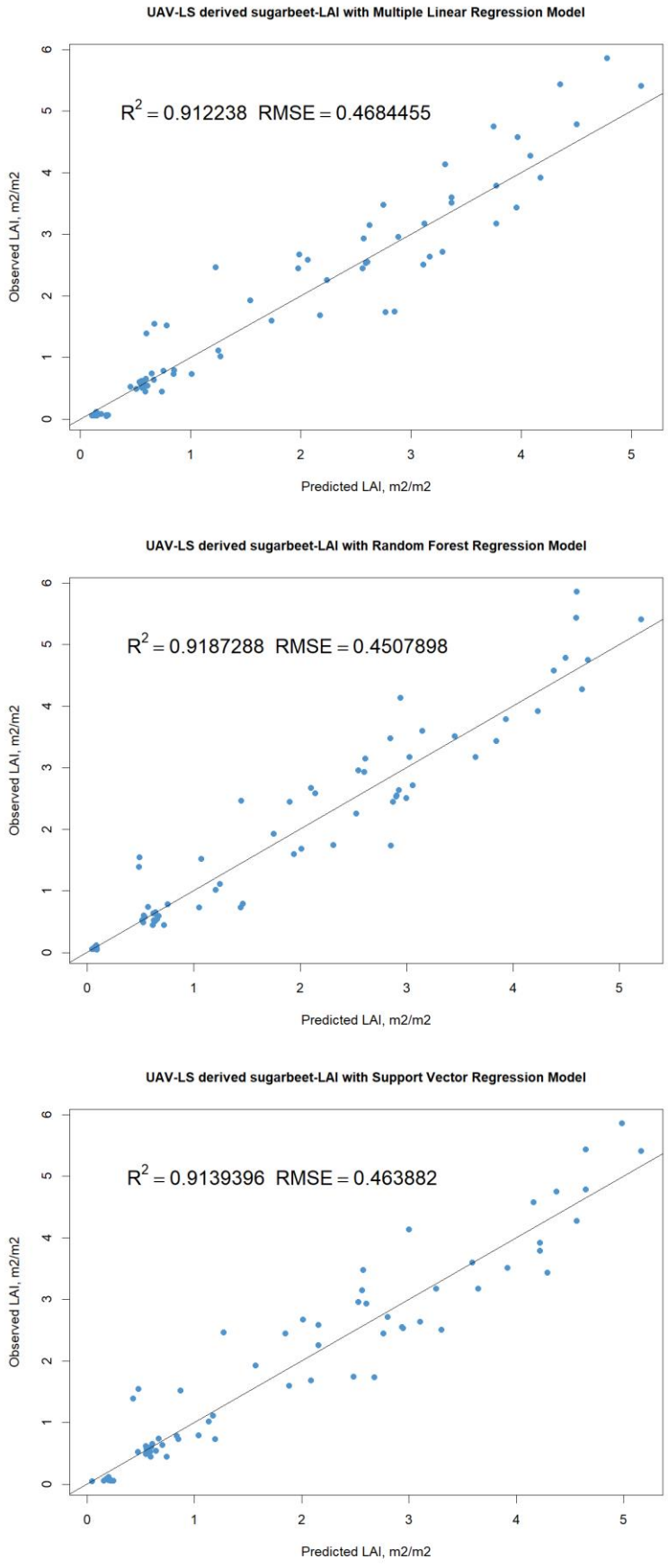
Annex 6 UAV-LS - Predicted and Observed values of Wheat AGB



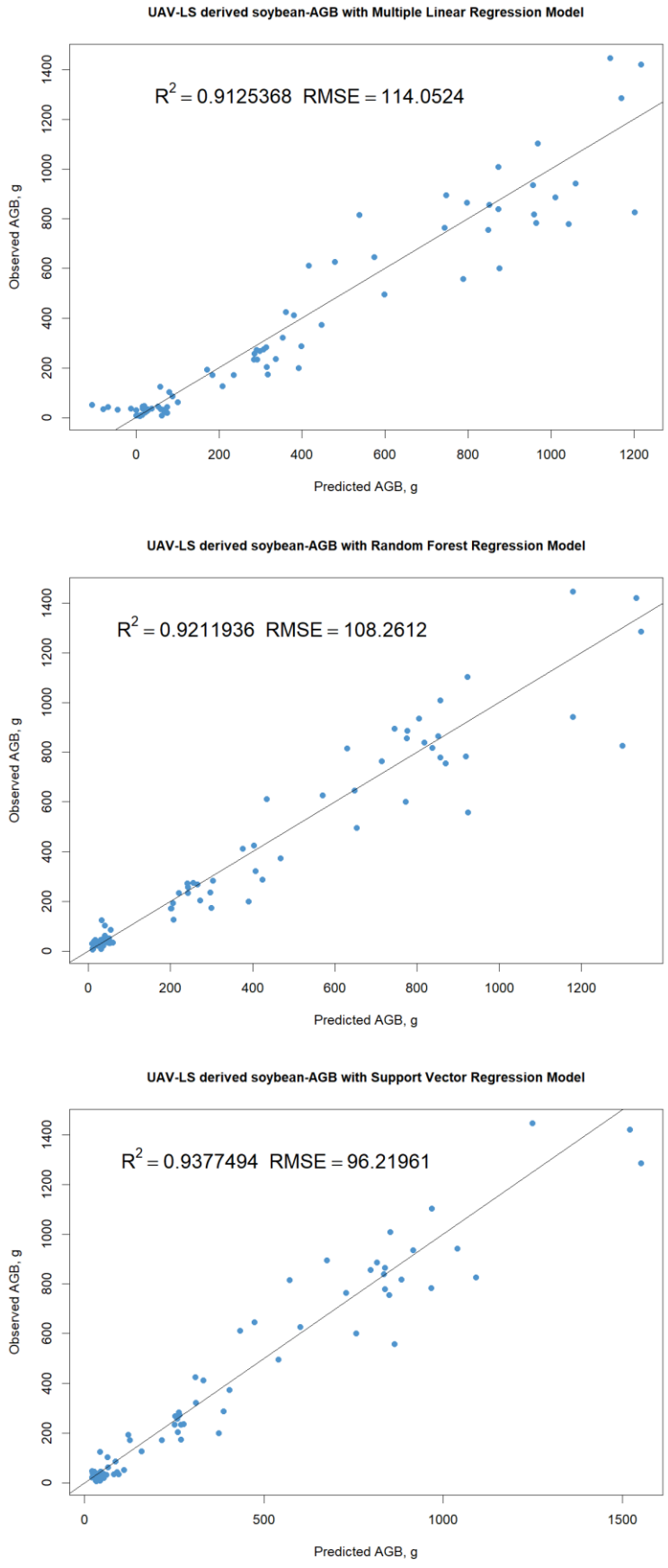
Annex 7 UAV-LS - Predicted and Observed values of Wheat LAI



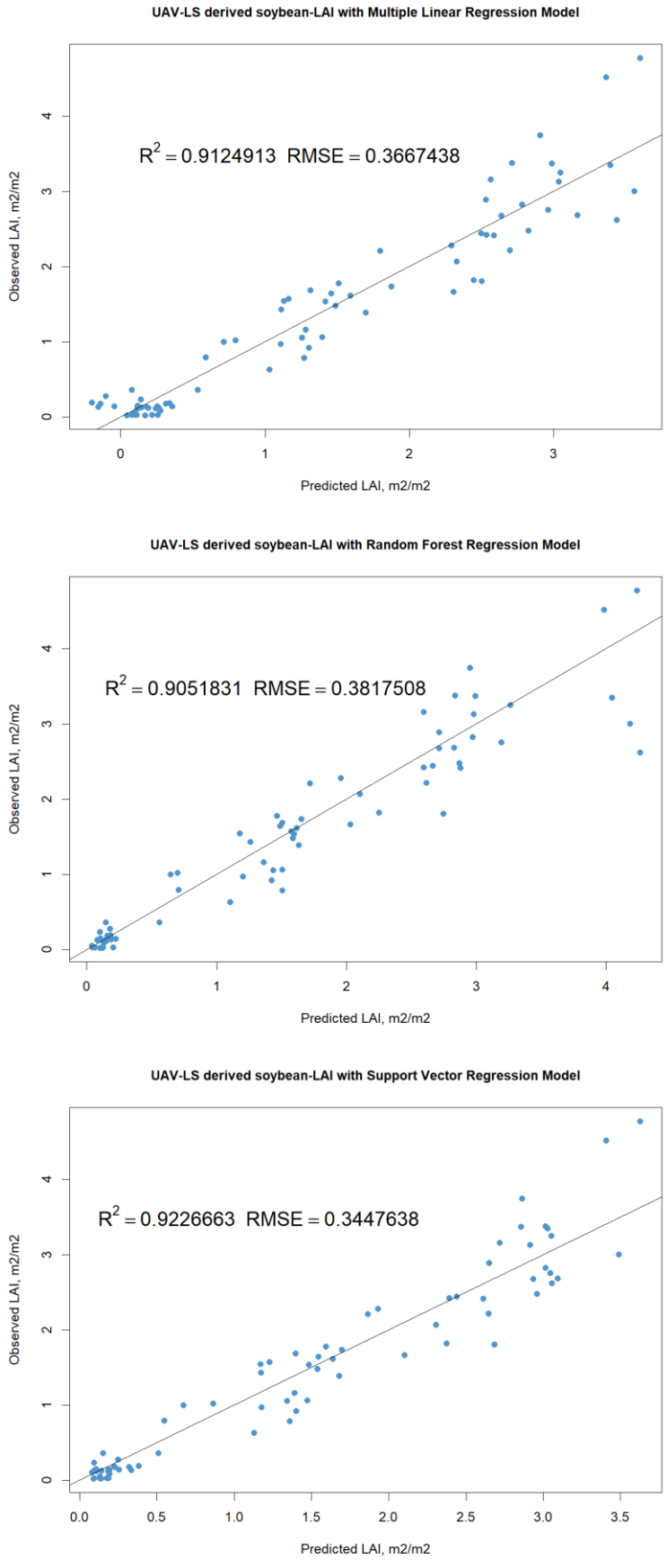
Annex 8 UAV-LS - Predicted and Observed values of Sugarbeet AGB



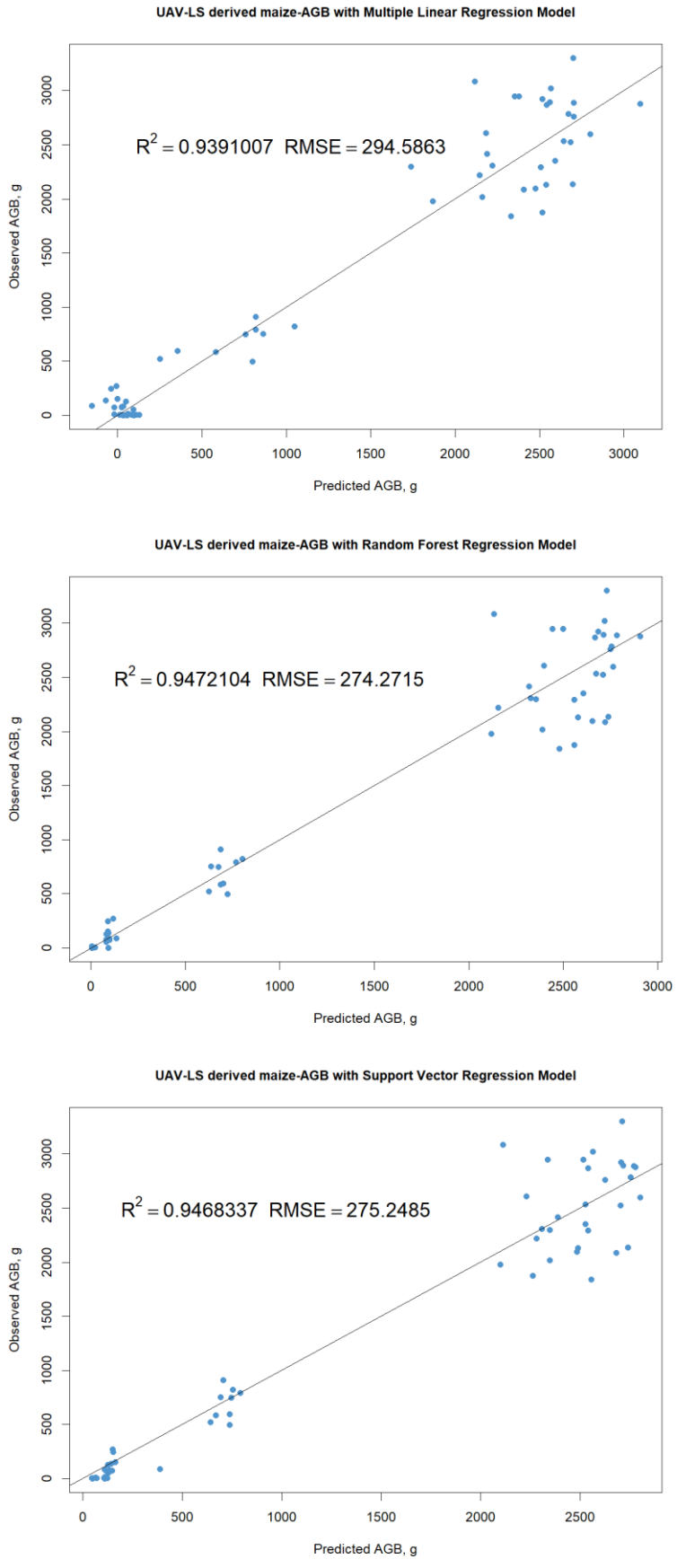
Annex 9 UAV-LS - Predicted and Observed values of Sugarbeet LAI



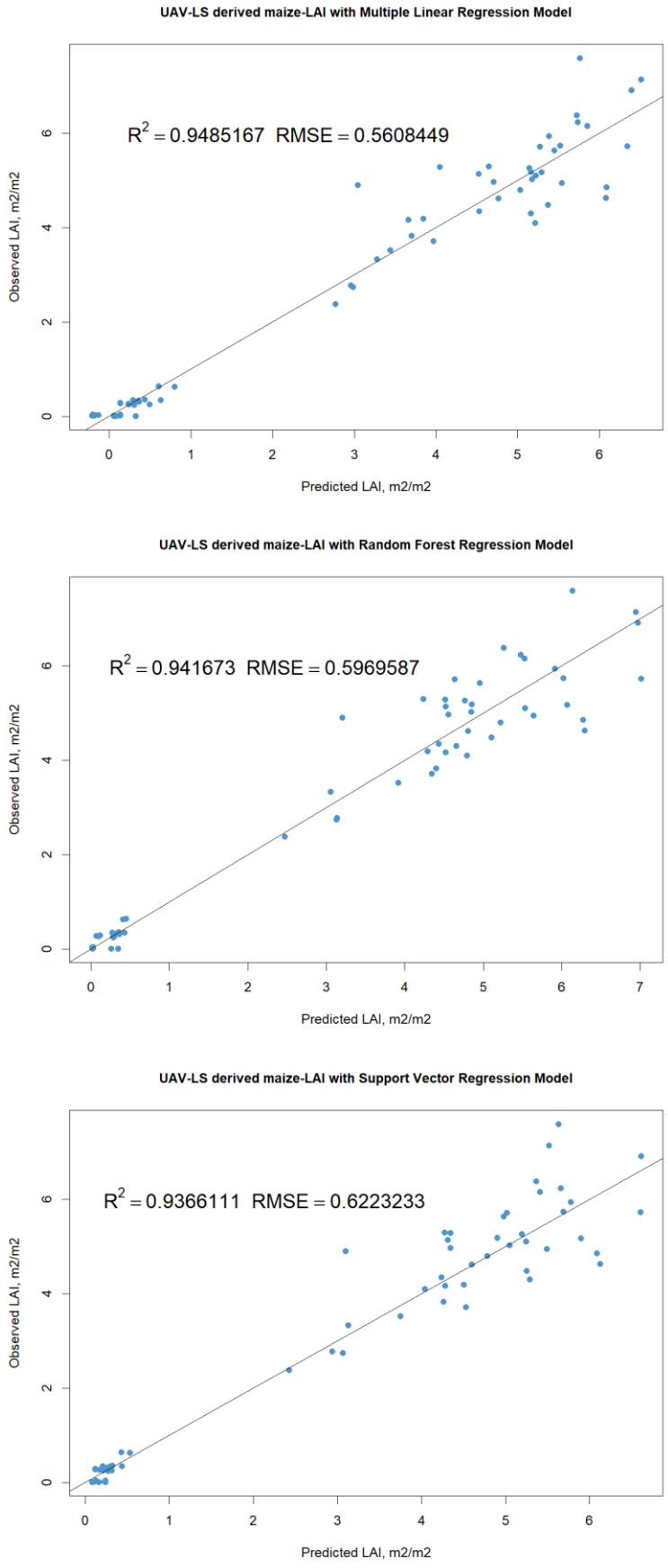
Annex 10 UAV-LS - Predicted and Observed values of Soybean AGB



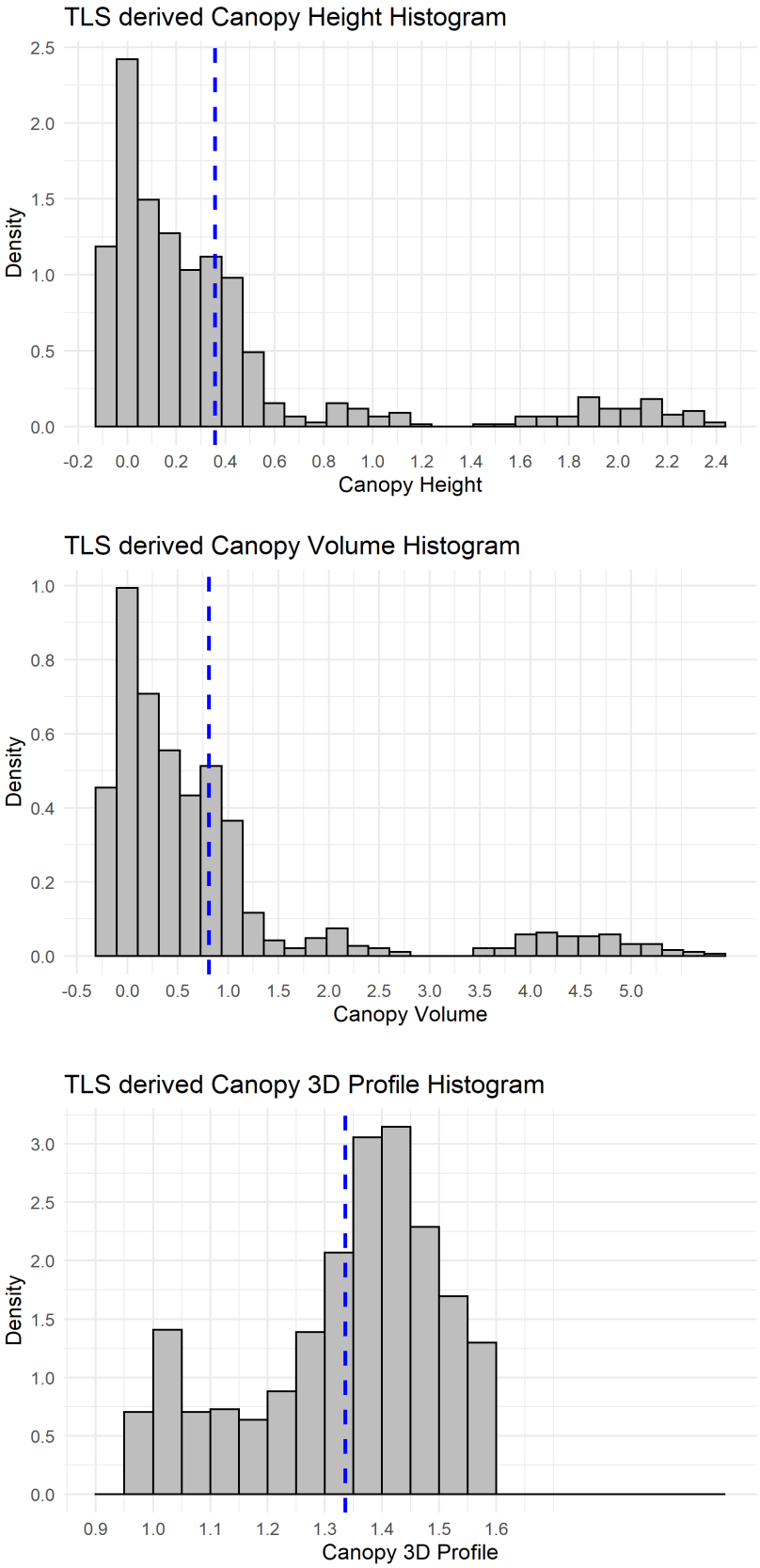
Annex 11 UAV-LS - Predicted and Observed values of Soybean LAI



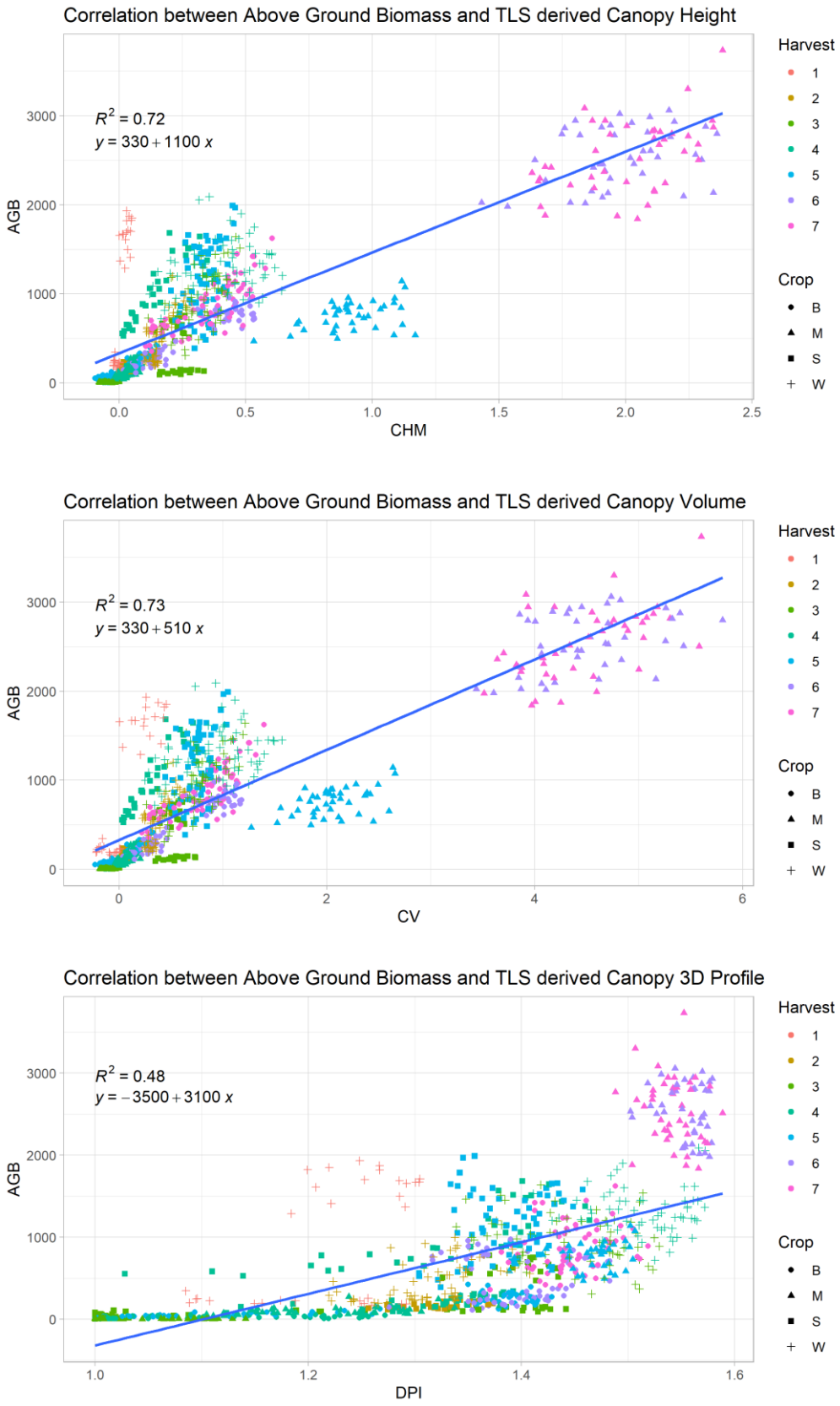
Annex 12 UAV-LS - Predicted and Observed values of Maize AGB



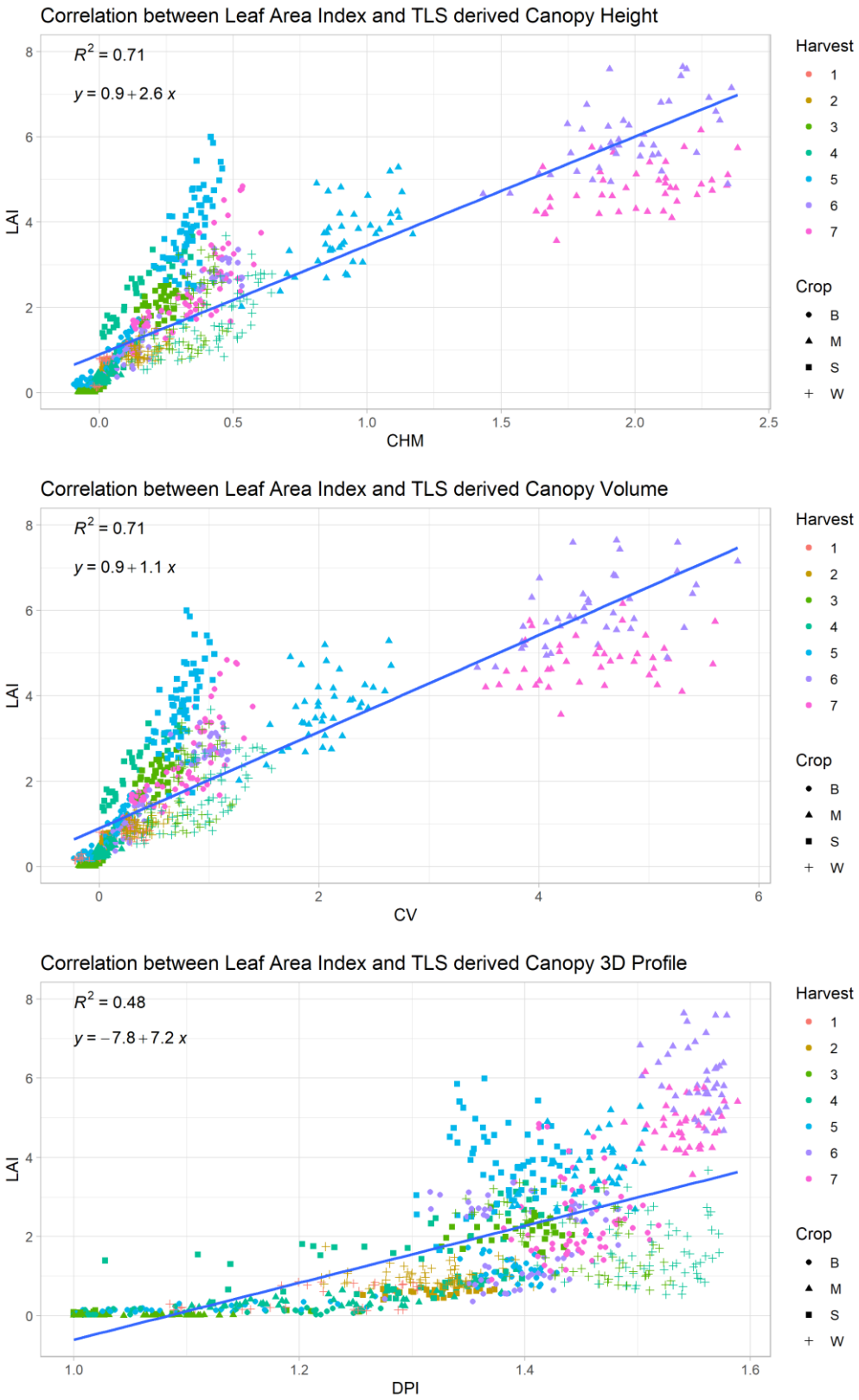
Annex 13 UAV-LS - Predicted and Observed values of Maize LAI



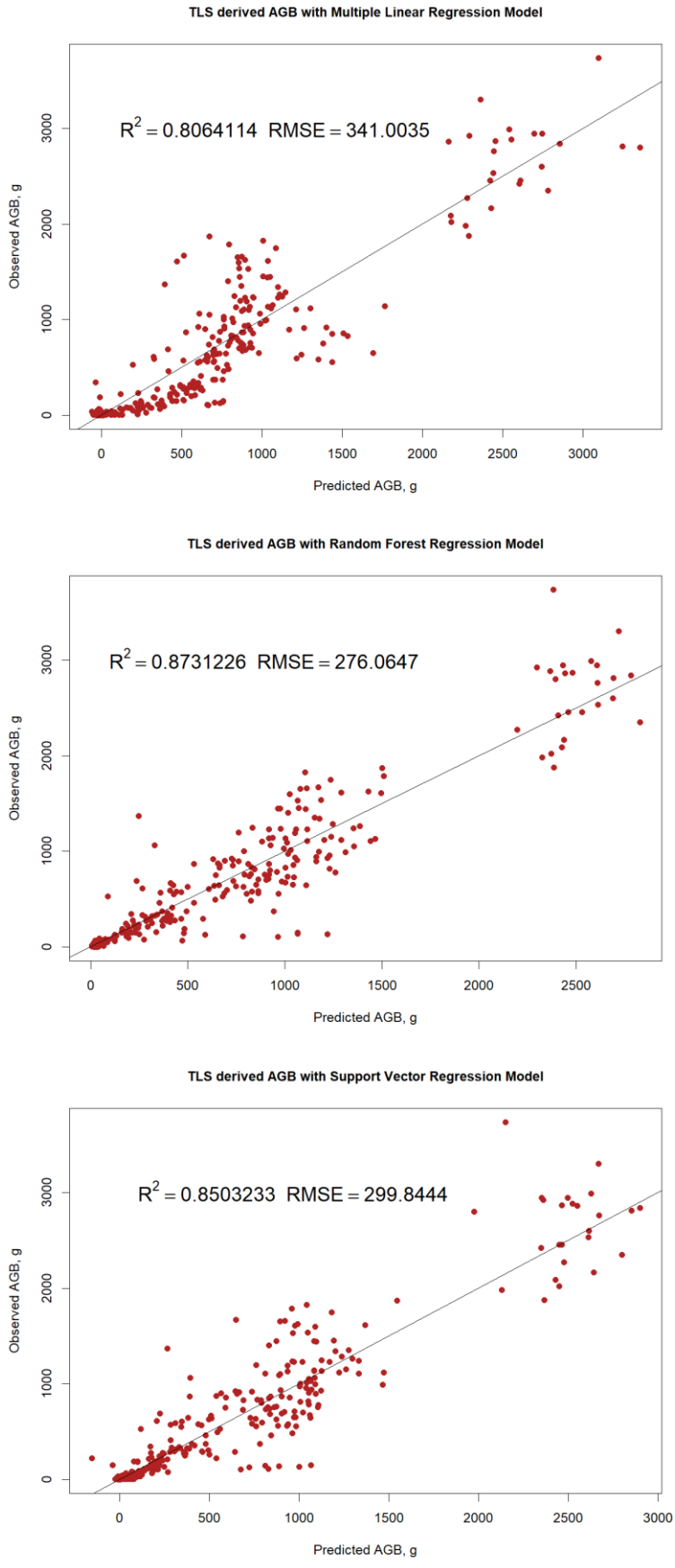
Annex 14 TLS derived Canopy Features Histogram



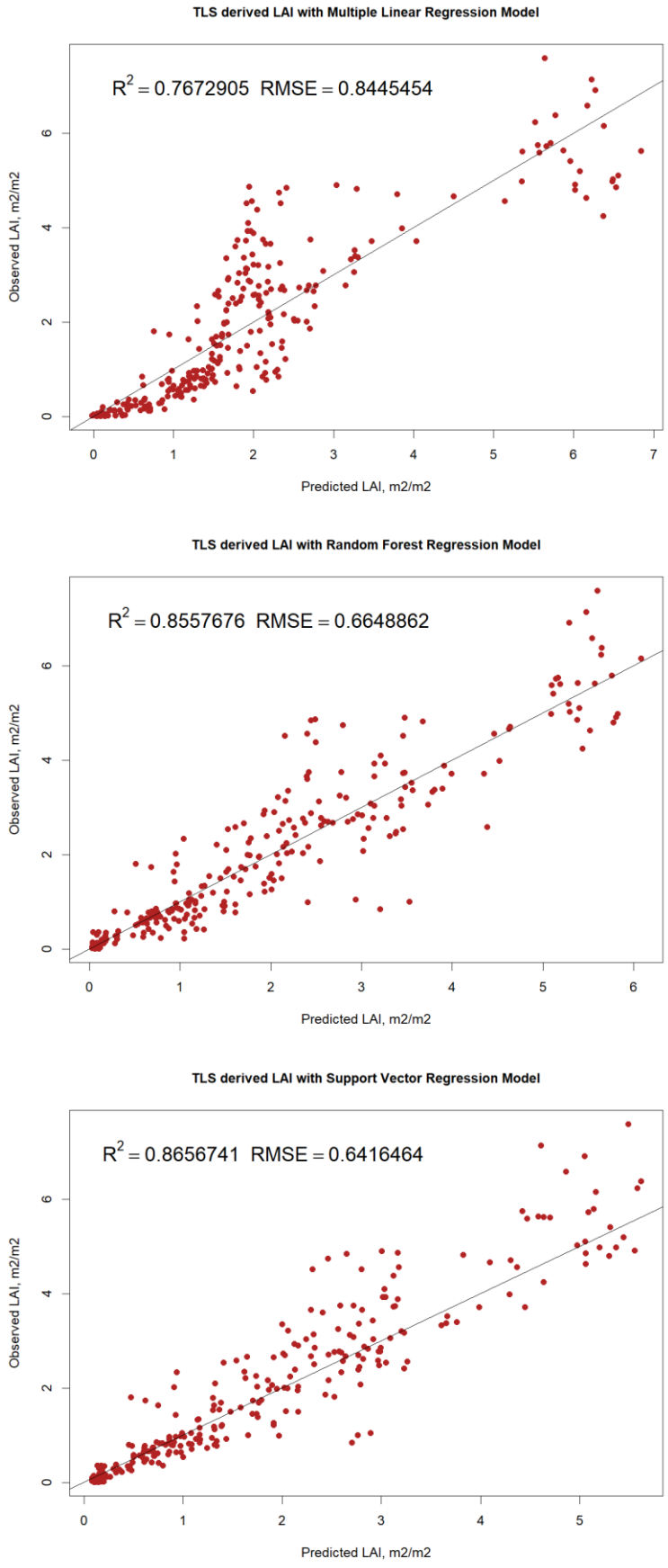
Annex 15 Correlation Between AGB and TLS derived features



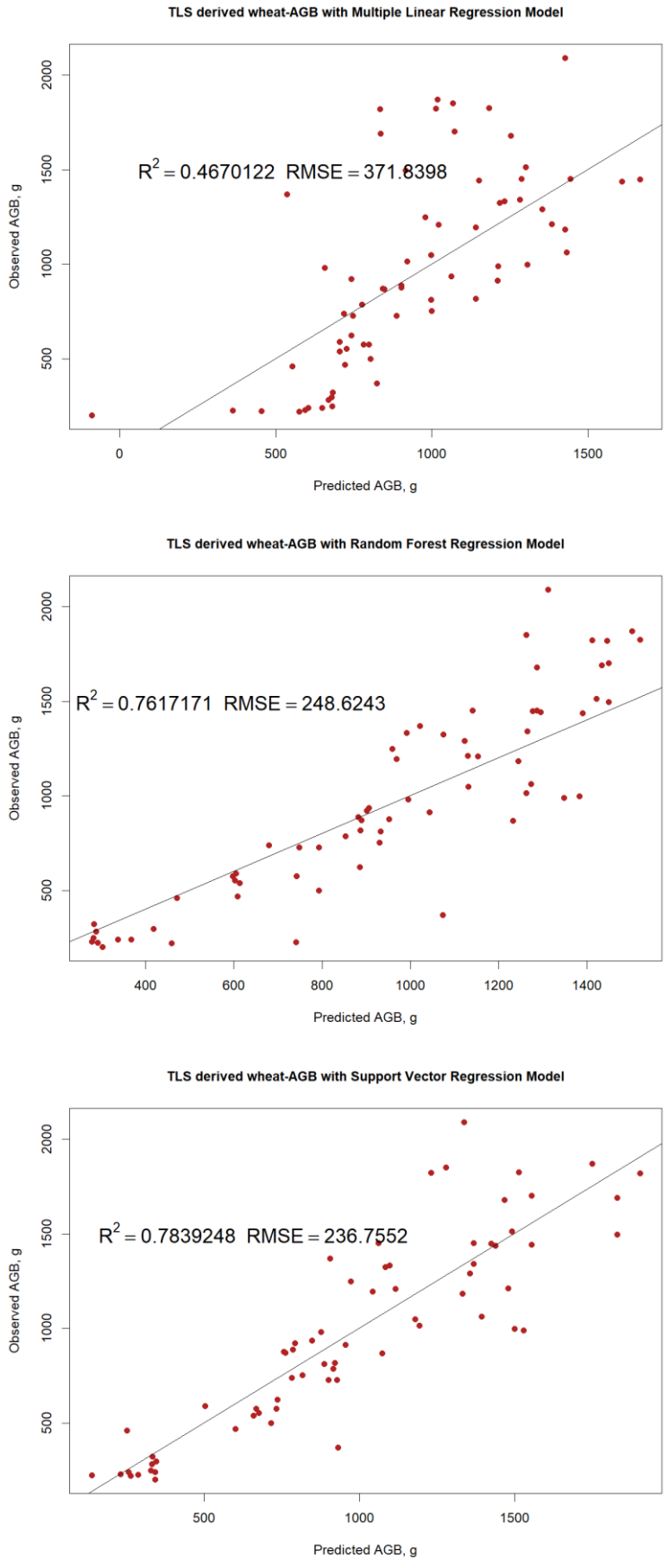
Annex 16 Correlation Between LAI and TLS derived features



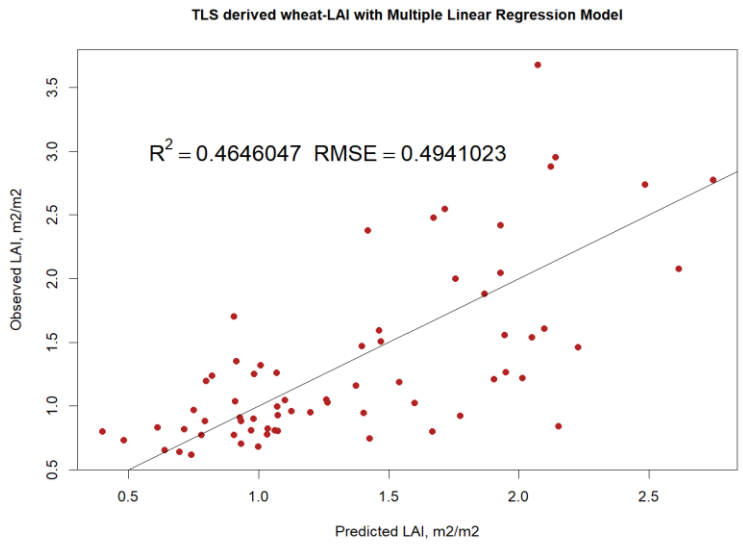
Annex 17 TLS-Predicted and Observed values of AGB with all samples

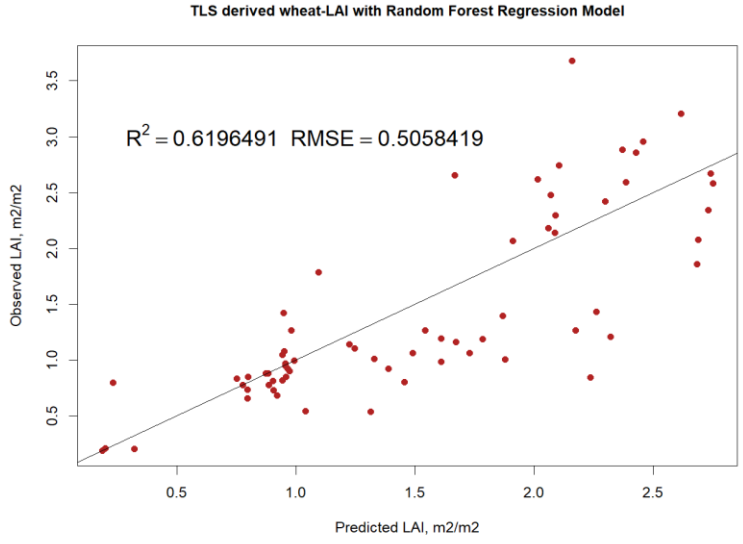


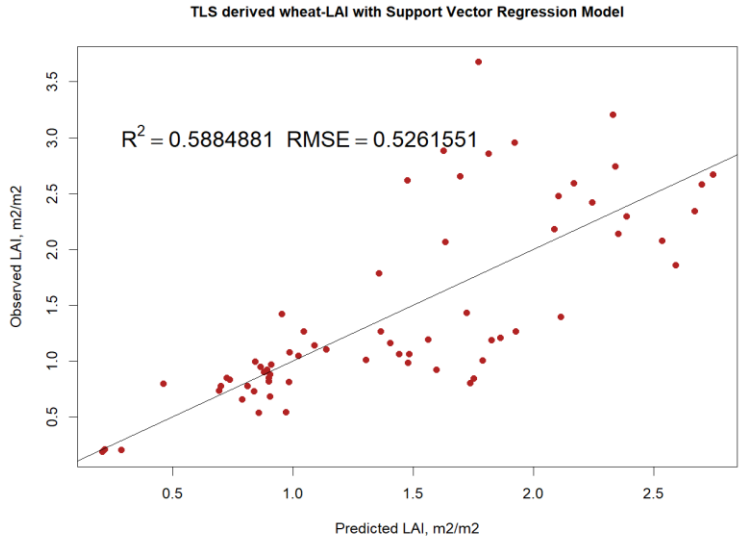
Annex 18 TLS-Predicted and Observed values of LAI with all samples



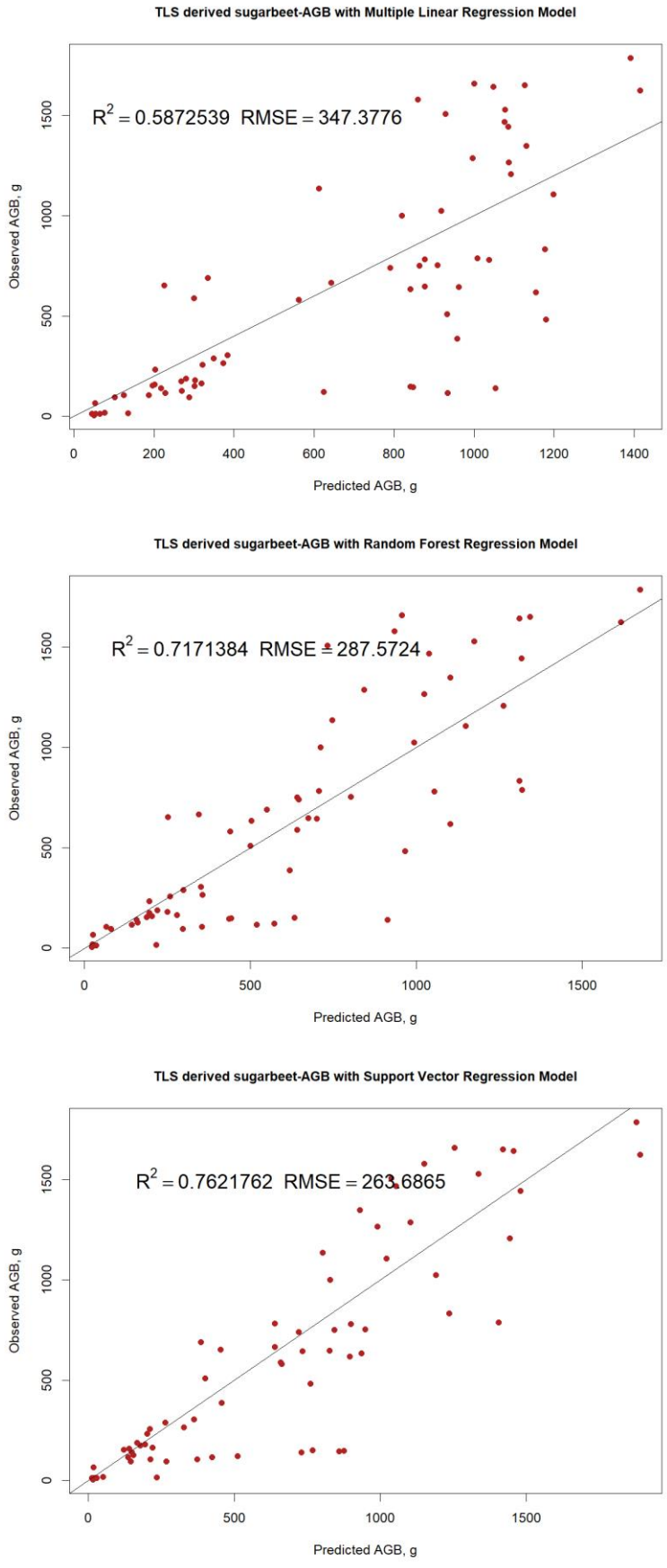
Annex 19 TLS-Predicted and Observed values of Wheat AGB



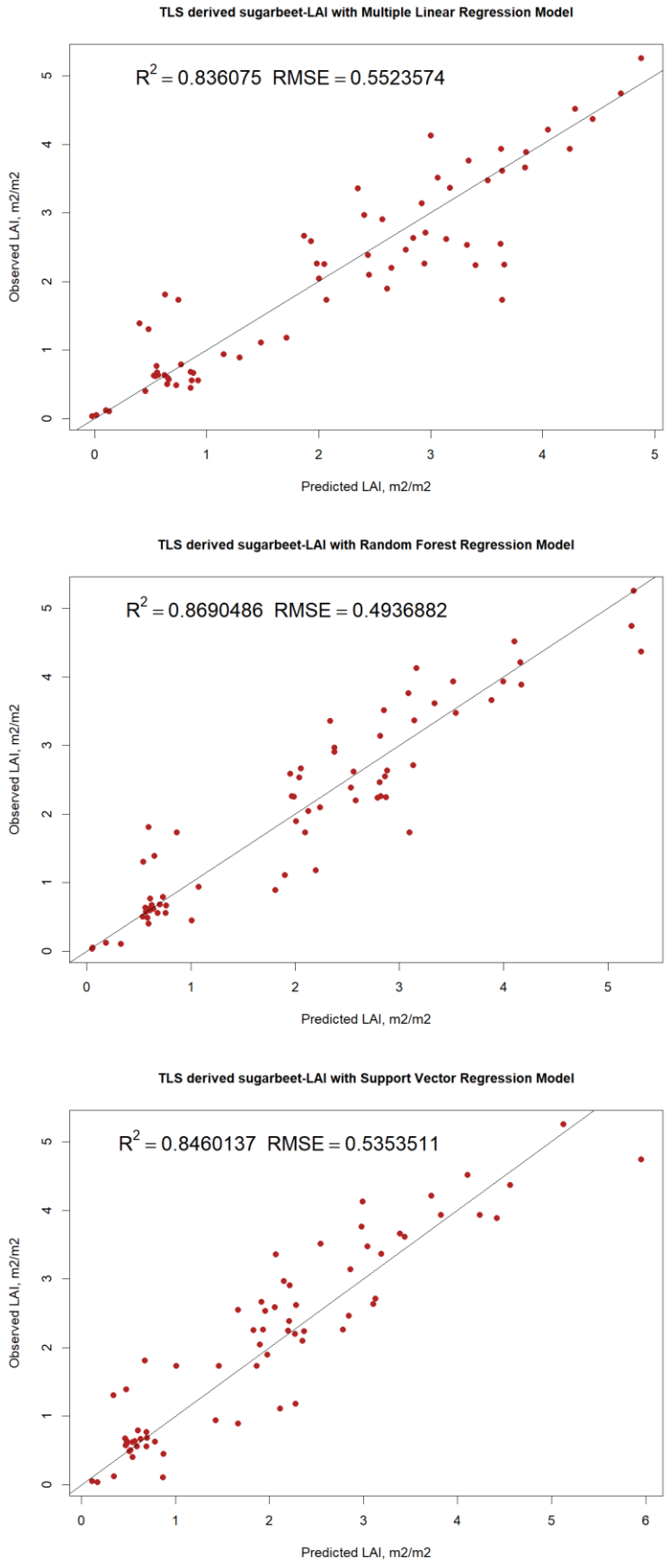




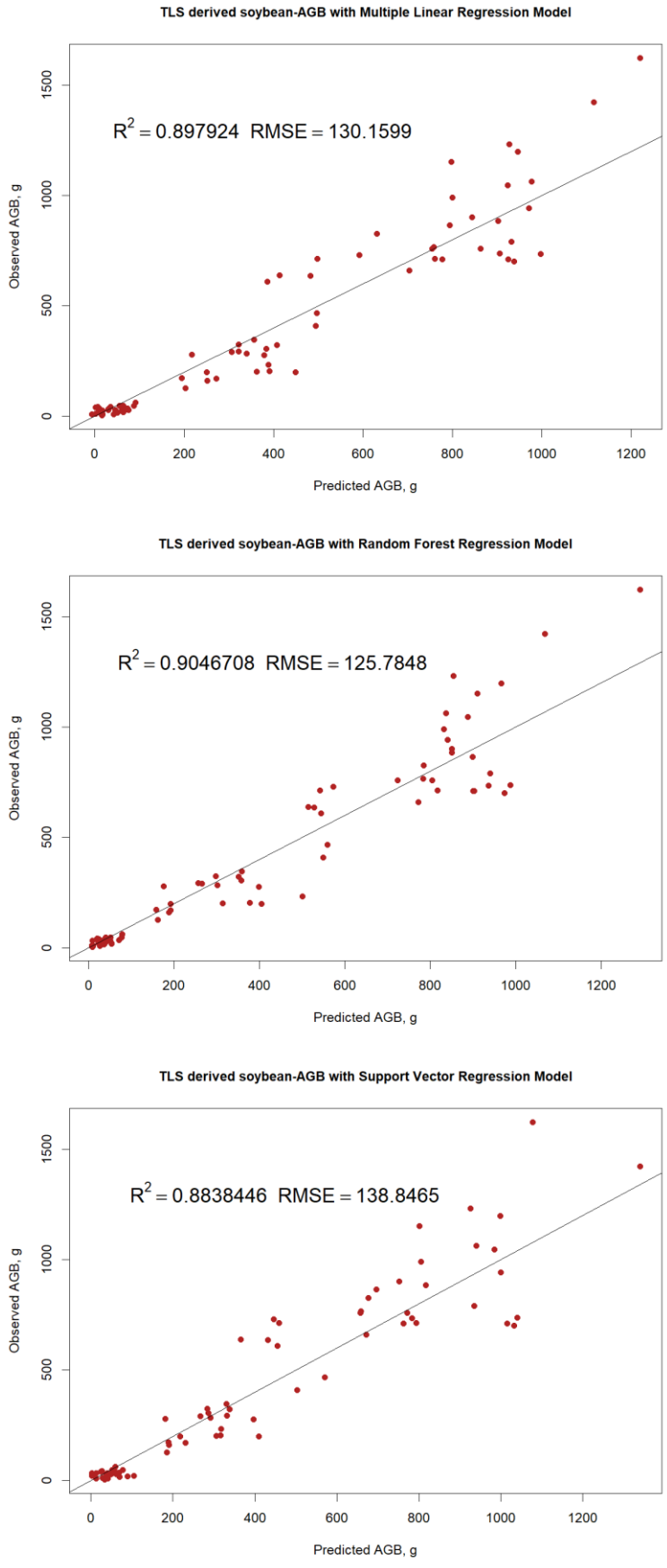
Annex 20 TLS-Predicted and Observed values of Wheat LAI



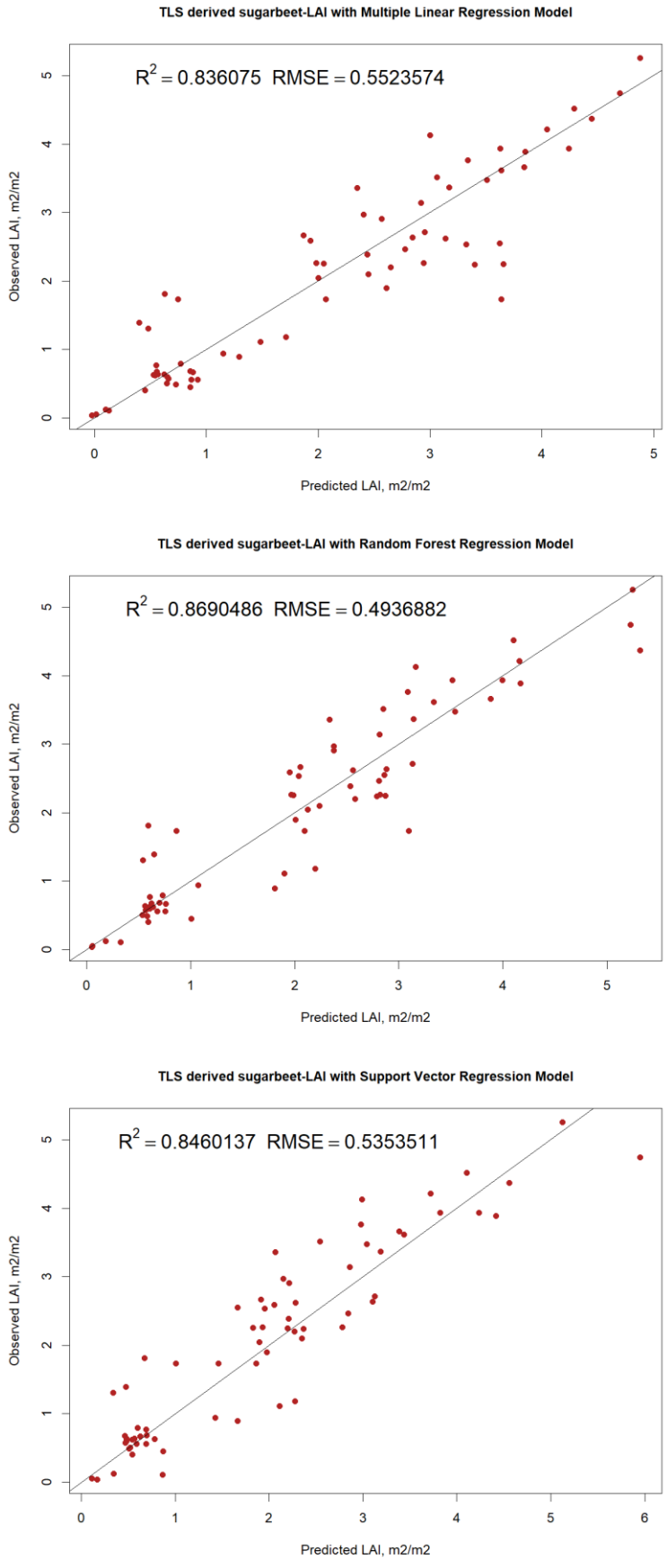
Annex 21 TLS-Predicted and Observed values of Sugarbeet AGB



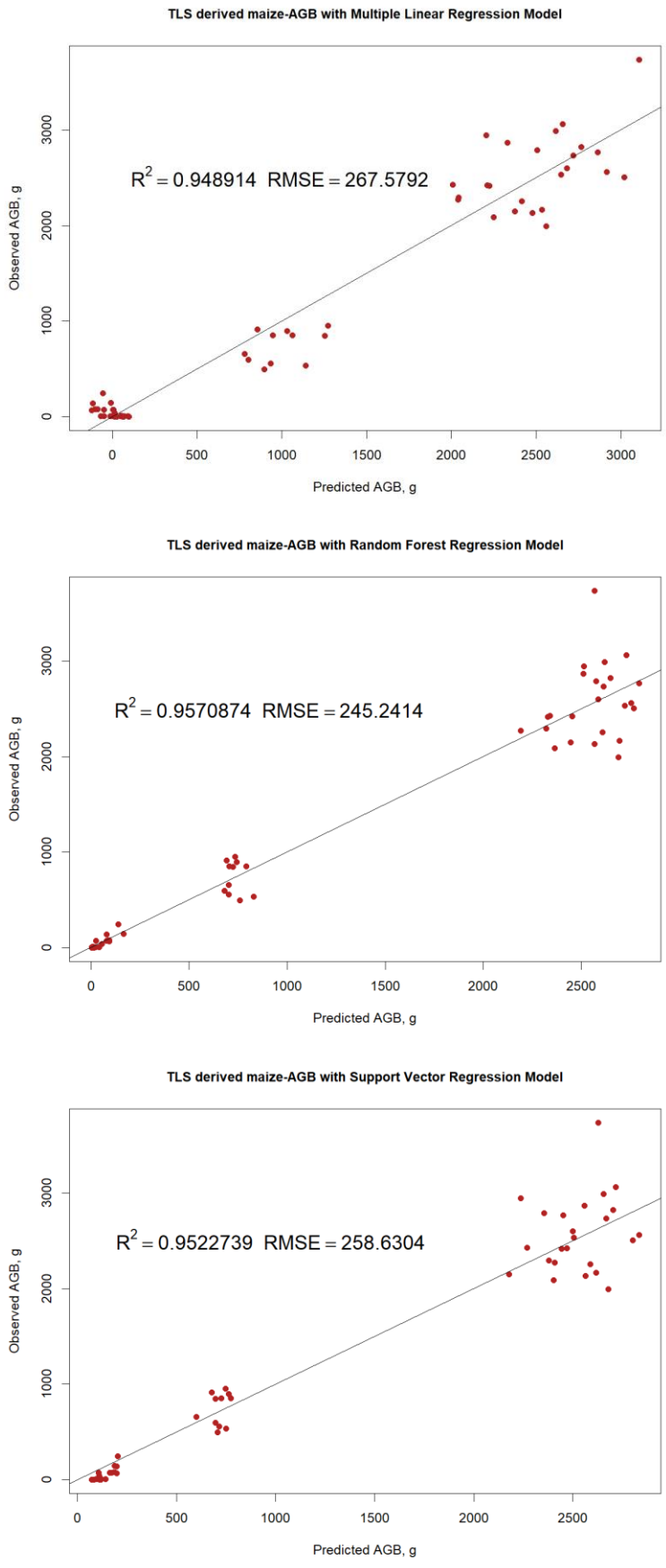
Annex 22 TLS-Predicted and Observed values of Sugarbeet LAI



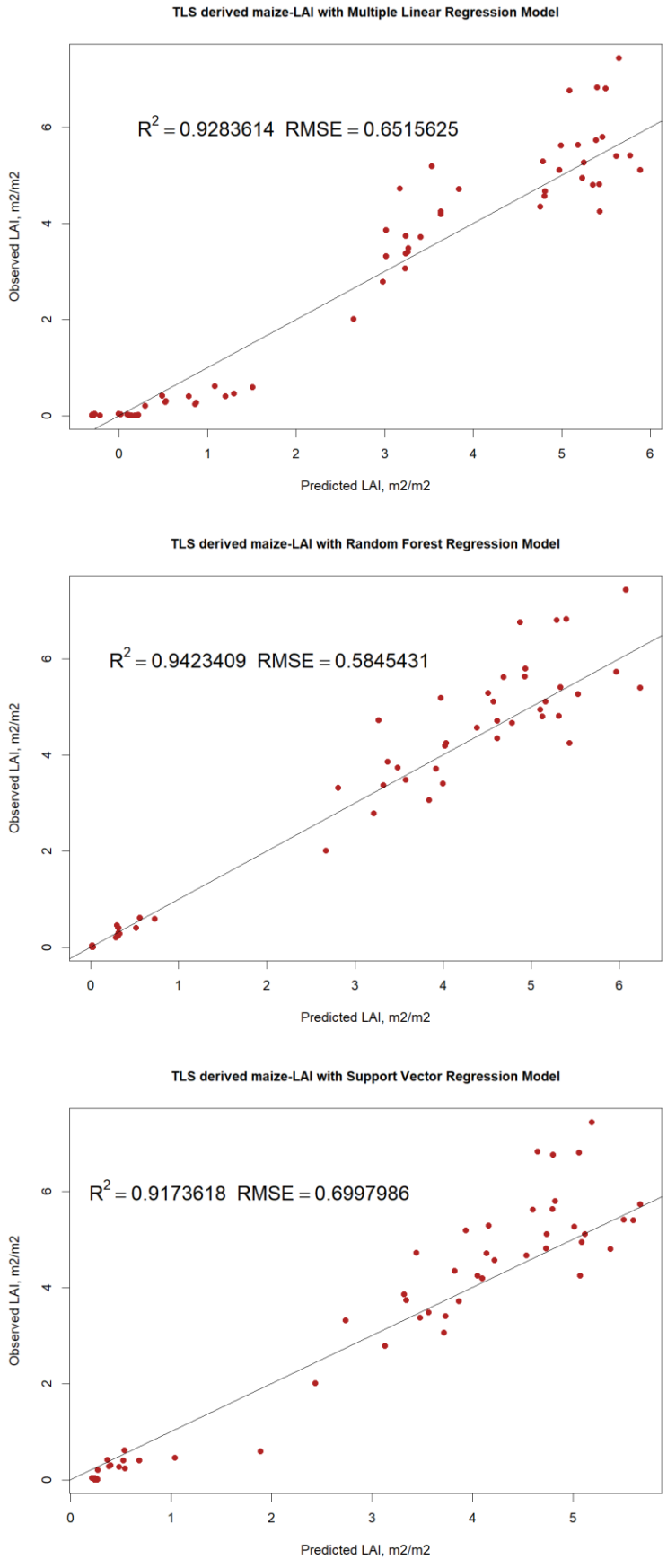
Annex 23 TLS-Predicted and Observed values of Soybean AGB



Annex 24 TLS-Predicted and Observed values of Soybean LAI



Annex 25 TLS-Predicted and Observed values of Maize AGB



Annex 26 TLS-Predicted and Observed values of Maize LAI