KU LEUVEN

ARENBERG DOCTORAL SCHOOL
FACULTY OF BIOSCIENCE ENGINEERING

High-throughput field phenotyping using a drone with RGB imagery

Exploiting the spectral, spatial and temporal dimensions

Manuscript for preliminary defense

Irene Borra Serrano

Supervisor:

Prof. dr. ir. Ben Somers

Co-supervisors:

Dr. ir. Peter Lootens

Prof. dr. ir. Wouter Saeys

Dr. ir. Koen Mertens

Dissertation presented in partial fulfilment of the requirements for the degree of Doctor of Bioscience Engineering (PhD)

October 2020

HIGH-THROUGHPUT FIELD PHENOTYPING USING A DRONE WITH RGB IMAGERY

EXPLOITING THE SPECTRAL, SPATIAL AND TEMPORAL DIMENSIONS

Irene BORRA SERRANO

Supervisor:

Prof. dr. ir. B. Somers

Co-supervisors:

Dr. ir. P. Lootens

Prof. dr.ir. W. Saeys

Dr. ir. K. Mertens

Members of the Examination Committee:

Prof. dr. ir. E. Schrevens

Dr. ir. J. Aper

Dr. A. de Castro

Prof. dr. ir. B. Van de Poel

Dissertation presented in partial fulfilment of the requirements for the degree of Doctor of Bioscience Engineering (PhD)

October 2020

Doctoraatsproefschrift nr. 1667 aan de faculteit Bio-ingenieurswetenschappen van de KU Leuven.

© 2020 Irene Borra Serrano

Uitgegeven in eigen beheer, Irene Borra Serrano, Gent

Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotokopie, microfilm, elektronisch of op welke andere wijze ook zonder voorafgaandelijke schriftelijke toestemming van de uitgever.

All rights reserved. No part of the publication may be reproduced in any form by print, photoprint, microfilm, electronic or any other means without written permission from the publisher.

Table of contents

Summary V

Samenvatting VII

List of abbreviations XI

List of tables XV

List of figures XVII

Chapter 1: Remote sensing-based plant phenotyping 1

1.1 Feeding the world under a changing climate 1

1.2 The role of plant phenotyping in breeding and biological research 2

1.2.1 Definition of traits and their measurement 4

1.2.2 Phenotyping methods 6

1.2.3 The plant phenotyping community 12

1.3 Remote sensing for field phenotyping 13

1.3.1 Sensors 15

1.3.2 Platforms 19

1.3.3 Data gathering and derived products from UAV 21

1.3.4 Data extraction 25

1.4 State of the art in remote sensing based high-throughput field phenotyping 26

1.5 Species studied in this dissertation 28

1.5.1 Perennial ryegrass (*Lolium perenne* L.) 28

1.5.2 Soybean (*Glycine max* (L.) Merr.) 31

1.6 Research objectives and dissertation outline 34

1.6.1 Thesis outline 36

Chapter 2: Towards an objective evaluation of persistency of *Lolium perenne* swards using UAV imagery 39

2.1 Introduction 40
2.2 Materials and methods 42
2.2.1 Field trials and plant material 42
2.2.2 Data acquisition of FT1 42
2.2.3 Application of the methodology to FT2 47
2.3 Results 48
2.3.1 Development of the methodology: FT1 48
2.3.2 Validation on an independent test field 54
2.4 Discussion 56
2.4.1 Three UAV-based VIs deliver robust results under different environmental conditions 56
2.4.2 Three UAV-based VIs provide a good estimation of the area covered by green grass .. 57
2.4.3 The visual scoring system (UAV-based VIs) consistency compared to the visual scoring by breeders 57
Chapter 3: Canopy height measurements and non-destructive biomass estimation of *Lolium perenne* swards using UAV imagery 61
3.1 Introduction 62
3.2 Materials and methods 64
3.2.1 Field trial 64
3.2.2 Data acquisition 65
3.2.3 Data handling and analysis 67
3.3 Results 70
3.3.1 Accuracy of UAV products 70
3.3.2 Comparison between the UAV-derived canopy height and the canopy height measured with the rising plate meter 71
3.3.3 Correlation between variables 72
3.3.4 Dry-matter yield models 73

3.4 Discussion 75
3.4.1 Accuracy of UAV-derived data 75
3.4.2 Canopy height determined using the RPM and CHM 76
3.4.3 Dry-matter yield prediction 77
Chapter 4: Closing the phenotyping gap: high resolution UAV time series for soybean growth analysis provides objective data from field trials 81
4.1 Introduction 82
4.2 Materials and Methods 83
4.2.1 Field Trial 83
4.2.2 Data acquisition 84
4.2.3 Data Handling and Statistical Analysis 85
4.3 Results 89
4.3.1 Model Fitting for Canopy Cover and Canopy Height 89
4.3.2 Extra Phenotypic Parameters Derived from the Fitted Curves 92
4.3.3 Relations between Phenotypic Parameters 94
4.3.4 Linking UAV-Based and Breeder Information 96
4.4 Discussion 99
4.4.1 General Assessment of the Methodology Developed 99
4.4.2 Describing General Growth Patterns in the Different Groups 100
4.4.3 Estimating SY and Maturity Using UAV-Derived Information 101
4.4.4 Perspectives for Practical Applications 102
4.4.5 Practical Lessons Learned and Future Perspectives 102
4.5 Conclusions 102
Chapter 5: General discussion 105
5.1 Technical considerations 105

5.1.1 UAV-RGB as a phenotyping instrument 105
5.1.2 Extraction and interpretation of data from UAV-RGB images 105
5.1.3 Modeling crop growth using multi-temporal UAV data 107
5.2 Practical implementations 108
5.2.1 Persistency and dry matter yield in perennial ryegrass 108
5.2.2 Growth and development in soybean 109
5.3 Lessons learned and recommendations for future use of the results 110
5.3.1 In breeding 110
5.3.2 Link with precision agriculture 111
5.4 Conclusions 112
5.5 Future developments and perspectives 113
5.5.1 Image processing and data extraction 113
5.5.2 Combining information from different sensors and new traits estimation 114
5.5.3 Data standardization and sharing 115
5.5.4 Advanced mathematical methods 116
5.5.5 HTFP in genetic research and for modelers 116
References 119
Appendices 141
Supplementary tables and figures of Chapter 1 141
Supplementary figures of Chapter 4 144
Supplementary explanations 150
Curriculum Vitae 153

Summary

Altered climatic conditions leading to more frequent and intense abiotic events (e.g. drought or heatwaves) and biotic stresses (pathogens and pests) are going to impact on our food system endangering the access to safe, nutritious and sufficient food. To mitigate the negative effects, the adaptation of crops to future climate has to go also hand in hand with an increase in global crop production in order to achieve the estimated demand for the growing world population. Furthermore, this should be achieved requiring less inputs and increasing efficiency of resources used (i.e. more sustainable production).

Different approaches can be taken to accelerate adaptation and increase crop production. This can be done by further optimization of crop management and adjustments of the inputs (precise fertilization, irrigation, …), but especially by the development of improved crop varieties through breeding. Breeding can highly contribute to secure agricultural productivity in this changing environment. However, the current process to adapt a crop to the new growing conditions requires at least a decade of experimental work. The different steps involved in the breeding process are based on the evaluation and selection of genotypes by highly labor intensive measurements and visual assessments of plant characteristics. In order to speed up and improve the breeding process, new technologies are required in different disciplines. In this dissertation the focus is on plant phenotyping, the process of assessing the performance of a genotype for a particular environment. For this purpose, breeding field trials are established to monitor and evaluate the plant performance in the target growth environment.

In those trials, a large number of genotypes are established in small plots and a selection of the best and more appropriate ones (to the targeted environment) is carried out based on desirable characteristics/traits (e.g. drought tolerance, yield, …). Traits are usually assessed visually, by field measurements or destructive sampling. These methods entail limitations: they are to a certain extent subjective and the measurements take a lot of time. How crops are phenotyped determines the efficiency of the breeding program and the resulting genetic gain. This genetic gain can be improved by upscaling and increasing the accuracy and frequency of the phenotypic evaluation.

The breeding process can be accelerated via high-throughput field phenotyping using a sensor mounted on a drone. This allows to collect more objective phenotypic data in a faster manner, even for large breeding trials. Drones allow to fly over the crop at crucial moments in crop development, as frequent as it is needed. Imagery is captured in a short period of time (5’ to 15’ depending on the size of the trial and the resolution required) and therefore under very similar light conditions. To promote integration in breeding programs, the technology should provide high resolution data at an acceptable level of cost and must be user-friendly. Therefore, this PhD research focused on the use of a drone equipped with an RGB camera, which can capture imagery at a high resolution, for quantitative monitoring of plant traits relevant in breeding. From the imagery, vegetation indices and canopy height models were derived, and variability within and between plots was assessed. The potential of this technology was investigated in two model crops, perennial ryegrass and soybean. Both species have economic and ecological importance, but differ in morphology, growth, management, breeding process and final use. Perennial ryegrass is a fodder crop, monocot, and cultivated for the biomass and is harvested several

times during one growing season. Soybean is a grain legume, dicot and grown for the seeds. Here the technical possibilities to perform high-throughput phenotyping in a field environment are investigated. We developed procedures that can easily be integrated and applied in breeding programs which were tested and validated in ongoing breeding trials. Evaluation and discussion of the specific genotypes in terms of physiology was beyond the scope of this PhD research.

Important breeding traits tackled for perennial ryegrass are yield and persistency. Persistency is a crucial trait to ensure productivity along the years a pasture is exploited. This trait is traditionally assessed by visual scoring. This trait is not easy to evaluate and the assessment is not free from subjectivity. Therefore, as an alternative, we have developed a methodology based on vegetation indices (spectral information) derived from drone-based RGB imagery. Different vegetation indices were tested under different light conditions, revealing that it is possible to evaluate persistency of hundreds of plots with UAV imagery in a more consistent way and achieving a better differentiation between genotypes than visual scoring. Biomass yield, another important trait, is typically monitored by destructive sampling which is laborious and time-consuming. A rising plate meter can be used as a non-destructive method for intermediate measurements. The measurements are influenced by canopy height and density, but are limited in evaluating variability within a plot. In this dissertation, we present a non-destructive method capable of capturing field variability and temporal evolution of crop height used to estimate yield. We tested different models for dry matter yield estimation and all models based on drone data resulted in a lower RMSE compared to a rising plate meter. The best one achieved an improvement from an RMSE of 986 (rising plate meter) to 679 kg/ha (drone). In the case of soybean, the main breeding traits are related to seed yield, maturity (ripening of the seeds), growth and development. Multi-temporal canopy cover and height data can provide information about crop growth and performance. By fitting growth curves per plot/genotype to the multi-temporal data, relevant phenotypic characteristics (e.g. canopy closure, growth rate, senescence, …) were derived. Differences between genotypes were depicted more clearly. Based on the phenotypic characteristics seed yield and maturity were estimated by means of a multiple linear regression (adjusted R^2 = 0.49 and 0.80, respectively). This approach resulted in biologically interpretable parameters that were informative for relevant traits.

In conclusion, our results demonstrate the potential of RGB imagery captured with a drone in a breeding context, enabling to obtain objective high-throughput phenotypic data from large field trials with high spatial and temporal resolution. The methodologies developed in this dissertation can support or replace traditional breeding methods.

Samenvatting

Veranderde klimaatomstandigheden zullen leiden tot frequentere en intensievere abiotische (bijv. droogte of hittegolven) en biotische stress (ziekten en plagen). Dit zal verstrekkende gevolgen hebben voor ons voedselsysteem, waardoor de toegang tot voldoende, veilig en voedzaam voedsel in gevaar komt. Om de negatieve effecten te verminderen en te voldoen aan de toenemende vraag van een groeiende wereldbevolking, moet de adaptatie van de gewassen aan het toekomstige klimaat ook gepaard gaan met een toename van de wereldwijde voedselproductie. Bovendien moet dit worden bereikt door gebruik te maken van minder 'inputs' en door de efficiëntie van de gebruikte middelen te verhogen of m.a.w. door een duurzamere productie.

Het versnellen van de klimaatadaptatie en de verhoging van de voedselproductie kan gebeuren op verschillende manieren. Dit kan door de verdere optimalisatie van de teelttechniek en aanpassingen aan inputzijde (plaatsspecifieke bemesting, irrigatie, …), maar vooral ook door de ontwikkeling van verbeterde variëteiten door middel van veredeling. Veredeling kan in hoge mate bijdragen aan het veiligstellen van de productiviteit van de landbouw in een veranderende omgeving. Het huidige proces om een gewas aan te passen aan de nieuwe teeltomstandigheden vereist echter ten minste een decennium van experimenten. Hierbij zijn de verschillende stappen in het veredelingsproces gebaseerd op evaluatie en selectie van vele kruisingen met bijhorende, zeer arbeidsintensieve metingen en visuele beoordelingen van de plantkenmerken. Om het veredelingsproces te versnellen en te verbeteren zijn nieuwe technologieën nodig in verschillende disciplines. In dit proefschrift ligt de focus op plantenfenotypering, het proces van het beoordelen van de prestaties van een genotype in een bepaalde omgeving. Hiervoor worden veredelingsveldproeven uitgevoerd om de prestaties van de plant in de beoogde groeiomgeving op te volgen en te evalueren.

Bij deze experimenten wordt een groot aantal genotypen geteeld in kleine plots en wordt een selectie gemaakt van de beste en meest geschikte genotypen (voor de beoogde omgeving) op basis van de beoogde kenmerken/'traits' (bv. droogtetolerantie, opbrengst, …). Kenmerken worden meestal visueel beoordeeld, door middel van veldmetingen, of via destructieve bemonstering. Deze methoden hebben een aantal beperkingen: ze zijn in zekere mate subjectief en de metingen nemen veel tijd in beslag. De manier waarop de gewassen worden gefenotypeerd bepaalt de efficiëntie van het veredelingsprogramma en de daaruit voortvloeiende genetische winst. Deze genetische winst kan worden verbeterd door het opschalen van de experimenten en het verhogen van de nauwkeurigheid en de frequentie van de fenotypische evaluaties.

Het veredelingsproces kan dus worden versneld door gebruik te maken van '*high-throughput field phenotyping*' met behulp van een sensor gemonteerd op een drone. Dit laat toe om sneller meer objectieve fenotypische gegevens te verzamelen zelfs voor grote veredelingsveldproeven. Drones laten toe om op cruciale momenten in de gewasontwikkeling over het gewas te vliegen met de gewenste frequentie. Hierbij worden alle metingen in een korte periode (5' tot 15' afhankelijk van de grootte van de veldproef en de gewenste resolutie) en dus onder zeer gelijkaardige lichtomstandigheden uitgevoerd. Om de integratie in veredelingsprogramma's te bevorderen moet deze technologie objectieve data met hoge resolutie aanleveren aan een redelijke prijs en moet de techniek

gebruiksvriendelijk zijn. Daarom richtte dit doctoraatsonderzoek zich op het gebruik van een drone uitgerust met een RGB-camera die beelden kan vastleggen met een zeer hoge resolutie voor de kwantitatieve beoordeling van plantkenmerken van belang bij de veredeling. Uit de hoge resolutie beelden kunnen vegetatie-indexen en hoogtemodellen berekend worden en kan de variabiliteit binnen en tussen plots worden beoordeeld. Het potentieel van deze technologie werd onderzocht voor twee modelgewassen, namelijk voedergrassen (raaigras) en soja. Beide soorten zijn economisch en ecologisch belangrijk, maar verschillen in morfologie, groei, teelt, veredelingsproces en eindgebruik. Raaigras is een monocotyl voedergewas, wordt geteeld voor de biomassa en wordt meerdere malen geoogst tijdens één groeiseizoen. Soja is een dicotyl, zaaddragende peulvrucht, en wordt geteeld voor de zaden. Hier worden de technische mogelijkheden onderzocht om high-throughput phenotyping uit te voeren onder veldomstandigheden. We ontwikkelden procedures die gemakkelijk geïntegreerd en toegepast kunnen worden in veredelingsprogramma's. Deze technieken werden daarom dan ook getest en gevalideerd in lopende veredelingsproeven. De plantfysiologische evaluatie van de genotypen viel buiten de scope van dit doctoraatsonderzoek.

Belangrijke kenmerken waarnaar veredeld wordt voor voedergrassen zijn opbrengst en persistentie. Persistentie is cruciaal om de productiviteit gedurende de volledige looptijd dat een weide wordt geëxploiteerd te garanderen. Dit kenmerk wordt traditioneel beoordeeld aan de hand van visuele scores. Dit kenmerk is niet eenvoudig te evalueren en de beoordeling is in zekere mate subjectief. Daarom hebben we als alternatief een methode ontwikkeld op basis van vegetatie-indices (spectrale informatie) afgeleid van RGB-dronebeelden. Verschillende vegetatie-indices werden getest onder verschillende lichtomstandigheden. Hieruit bleek dat het mogelijk is om honderden plots objectief te evalueren op persistentie op een consistentere manier en met een groter onderscheidend vermogen dan mogelijk is via traditionele visuele scoring in het veld. De biomassa, een ander belangrijk kenmerk, wordt traditioneel bepaald door destructieve bemonstering. Dit is een precieze maar omslachtige en tijdrovende meting. Als niet- destructieve methode voor tussentijdse metingen wordt vaak een '*rising plate meter*' gebruikt. De resulterende meetwaarde wordt zowel door de gewashoogte alsook door de gewasdensiteit beïnvloed, maar laat niet toe de variabiliteit binnen een plot goed in te schatten. In dit proefschrift stellen we een niet-destructieve methode voor die in staat is om de veldvariabiliteit en de temporele evolutie van de gewashoogte die gebruikt wordt om de opbrengst te schatten, vast te leggen. We hebben verschillende modellen getest voor de schatting van de droge stof opbrengst en alle modellen gebaseerd op de drone data resulteren in een lagere RMSE in vergelijking met de '*rising plate meter*'. Het beste model resulteert in een verbetering van de RMSE van 986 ('*rising plate meter*') tot 679 kg/ha (drone data). Bij de veredeling van soja zijn de belangrijkste kenmerken de zaadopbrengst, het tijdstip van afrijpen van de zaden en de gewasgroei en -ontwikkeling. Multi-temporele data voor planthoogte en gewasontwikkeling leveren informatie over de groei en de prestaties van het gewas. Door groeicurven te berekenen op basis van de tijdsreeksen, en dit per plot/genotype, werden extra, relevante fenotypische kenmerken (bv. sluiten van de rijen, groeisnelheid, senescentie, …) afgeleid. Verschillen tussen genotypen kunnen duidelijker worden aangegeven. Op basis van deze fenotypische parameters en een meervoudige lineaire regressie wordt de zaadopbrengst en de afrijping voorspeld (respectievelijk: adjusted R^2 = 0,49 en 0,80). Voor deze relevante kenmerken kan aangegeven worden welke interpreteerbare ontwikkelingseigenschappen hier aan de basis liggen.

Samenvattend tonen deze resultaten de mogelijkheden van RGB-beelden genomen met een drone in een veredelingscontext. Hiermee kunnen objectieve '*high-throughput*' fenotypische gegevens bekomen worden van grote veldproeven met een hoge spatiale en temporele resolutie. De in dit proefschrift ontwikkelde methoden kunnen de traditionele veredelingsmethoden ondersteunen of zelf vervangen.

List of abbreviations

A-GG	area covered by green grass
AHR	artificial height references
AIC	Akaike Information Criterion
ANOVA	analysis of variance
APS-C sensor	Advanced Photo System type-C
CC	canopy cover
CC_6W	cover percentage after six weeks emergence
CC_75PCT	thermal time when 75% of cover is achieved
CC_AGRmax	maximum absolute growth rate
CC_Cmax	maximum cover
CC_RsqA	adjusted R2 of the curve fitting for canopy cover
CC_SNC	senescence rate calculated with CC data (CC_Cmax – average cover two last time points)/CC_Cmax
CC_tm	thermal time when maximum growth rate is achieved
cGCPs	control ground control points
CH	canopy height
CH_AGRmax	maximum absolute growth rate
CH_cv	coefficient of variation
CH_DET	determinacy calculated as CH_te – R1
CH_Hmax	maximum height
CH_min	plot minimum canopy height
CH_max	maximum
CH_mean	mean
CH_P50	median
CH_P90	90th percentile
CH_RsqA	adjusted R2 of the curve fitting for canopy height
CH_SNC	senescence rate derived from CH data (CC_Hmax—average height two last time points)/CC_Hmax
CH_stdev	standard deviation
CH_te	thermal time when CC_Hmax occurred
CH_tm	thermal time when maximum growth rate is achieved
CHM	canopy height model
CHMv	canopy height model variables
CIVE	color Index of vegetation
COM1	combination 1
COM2	combination 2
CV	coefficient of variation
CWSI	crop water stress index
DEM	digital elevation model

DMY	dry matter yield
DSM	digital surface model
DTM	digital terrain model
EP	percentage of emergence
EPPN	European Plant Phenotyping Network
ExG	excess green
ExG_cv	coefficient of variation
ExG_mean	plot mean Excess green
ExG_P50	median
ExG_P90	90th percentile
ExG_stdev	standard deviation
ExGR	excess Green minus Exc. Red
ExGR_cv	coefficient of variation
ExGR_mean	plot mean Excess green minus Excess red
ExGR_P50	median
ExGR_P90	90th percentile
ExGR_stdev	standard deviation
ExR	excess Red
FAIR	Findable, Accessible, Interoperable and Reusable
FT1	field trial 1
FT2	Field trial 2
GCC	normalized green intensity
GCP	ground control point
GDD	growing/growth Degree Days
GDDdif	growth degree days difference between cut date and measurement date
GDDv	environmental variables
GIS	geographic information system
GLI	green leaf index
GP	growth periods or Groups
GWAS	Genome Wide Association Mapping
HIS	hue-saturation-intensity
HTFP	High-throughput field phenotyping
IPPN	International Plant Phenotyping Network
ISO	sensitivity to light
LAI	leaf area index
LiDAR	light detection and ranging
Lp	perennial ryegrass
Lp-CPs	control plots
Lp-TPs	target plots
LR	linear regression
ML	machine learning
MLR	multiple linear regression

NDVI	normalized difference vegetation index
NGRDI	norm. green-red difference Index
NIR	near-infrared region
NRMSE	normalized RMSE
OSAVI	Optimized Soil Adjusted Vegetation Index
P50	median
P90	90th percentile
PCA	principal component analysis
PCR	principal component regression
PLS	partial least squares
R1	thermal time when R1 stage starts (beginning bloom)
R8	thermal time when R8 stage is reached (full maturity)
RADAR	radio detection and ranging
RCI	ratio Color Index
RF	random forest
RGB	red, green, blue / visible
RMSE	root mean square error
RPAS	remotely piloted aircraft system
RPM	rising plate meter
RQ	research question
RTK GPS	real time kinematic global positioning system
SfM	structure from motion
stdev or SD	standard deviation
SY	seed yield
SWIR	short-wavelength infrared region
Tb	base temperature
TIN	triangulated irregular network
TPE	target populations of environments
UAV	unmanned aerial vehicle
UAV-HTFP	UAV for high-throughput field phenotyping
UAV-RGB	UAV equipped with a visible sensor
VEG	vegetativen
vGCPs	validation ground control points
VI	vegetation index
VIF	variance inflation factor
VIv	vegetation index variables
WI	Woebbecke Index

List of tables

Table 1-1. Main breeding traits for crops and their conventional trait evaluation method. Adapted from: Araus et al. (2018); Reynolds et al. (2020) and Roitsch et al. (2019). 5

Table 1-2. Overview of sensors and their potential applications, advantages and disadvantages. Adapted from: Sankaran et al. (2015b); Yang et al. (2017) and Yao et al. (2019). 17

Table 2-1. List of RGB vegetation indices used in this study. R= Red, G= Green and B= Blue bands. ... 45

Table 2-2. Comparison of the stability of the vegetation indices (VI) over the different flights for the March 18 and the March 22 using a subset of 90 plots (n=90). Pearson correlation (r) and significance level (p), with bold showing $r > 0.9$ and $p < 0.001$. Intercept and slope of a linear regression, with bold values showing an intercept between -0.1 and 0.1, and a slope between 0.9 and 1.1. An * marks the selected vegetation indices, based on the pre-defined criteria 48

Table 2-3. Example of the results of the supervised classification matrix of an on-ground frame. 49

Table 2-4. Pearson correlation coefficients between the four flights over the two dates (18: March 18; 22: March 22; 1: first flight; 2: second flight) for Excess green 2 (ExG2) (n=90 plots). 51

Table 2-5. Contingency table between breeders for scoring persistency of *Lolium perenne* plots. The number of plots (count) and frequency (%) in each score class (3-9) per breeder is presented. Agreement in the plot scores between both breeders are shown in bold. 52

Table 2-6. Coefficients of variation (CV) and F-values (F varieties) per ploidy and block (trial design) of the visual Breeder score and ExG2 index derived from the four flights carried out over field trial 1 (ExG2 FT1). Statistical significance or not at the 5% level indicated per F-value marked with a plus or minus sign. 53

Table 3-1. Overview of the variables derived from the canopy height models, the orthophotos, and the weather station, that are used to predict dry matter production. 69

Table 3-2. Overview of the models used to estimate dry matter production. Models used were: linear regression (LR), multiple linear regression (MLR), principal component regression (PCR), partial least square regression (PLS), and random forest (RF). Model results are evaluated using the root mean square error of the validation dataset (RMSE, kg ha-1), normalized RMSE (NRMRE, %), Akaike information criterion (AIC), and R^2 or approximated R^2 (for PLS and RF). PLS and RF were optimized in terms of RMSE by selecting 2 (Ncomp3) and 4 components (Mtry4), respectively. 74

Table 3-3. Best MLR to predict DMY with all variables sets included in the model. For all variables mean, standard deviation (SD), effect size and standardized effect size are presented. 75

Table 4-1. Overview of the phenotypic parameters. 88

Table 4-2. Mean, median, standard deviation (Stdev) and coefficient of variation (cv) of the phenotypic parameters derived from fitting the canopy cover data. Data are presented per group (GP1-4). For the abbreviations, see Table 4-1. 93

Table 4-3. Mean, median, standard deviation (Stdev) and coefficient of variation (cv) of the phenotypic parameters derived from fitting the canopy height data. Data are presented per group (GP1-4). For the abbreviations, see Table 4-1. 94

Table 4-4. Statistics and coefficients of the normalized variables for the MLRs for seed yield (SY) and R8 stage based on the phenotypic parameters and the raw flight data. 97

Table A-1. Time required in each step from the UAV mission planning to the data extraction. Flight specifications: RGB 12mm at 30m with 80% overlap. 1.425 ha covered (0.575 trial). 104 images in total. 141

List of figures

Figure 1-1. A) Integrated strategies leading to a common decision support tool to accelerate genetic gain based on the use of new technologies. Each field is described with key advantages (∧) and disadvantages (∨). B) Detail of the discipline tackled in this dissertation. Source: Li et al. (2018). 2

Figure 1-2. Key elements to accelerate genetic gain in which high-throughput phenotyping can contribute (directly or indirectly): increase of the size of the breeding program, breeding cycles acceleration, increase selection accuracy, ensure genetic variation and improve of the decision support tools. TPE = Target populations of environments (where the breeding products would be grown). Source: Araus et al. (2018). 6

Figure 1-3. Controlled phenotyping platforms. Examples of A) top-view imaging cabinets; B) automatic systems for root or top-view imagery and C) side-view imaging conveyor belt system with resulting imagery of barley. Sources: Lemnatec website (https://www.lemnatec.com/), Emphasis website (https://emphasis.plant-phenotyping.eu/) and Chen et al. (2014). 8

Figure 1-4. Phenotyping needs in breeding and precision agriculture scenarios. Source: Chawade et al. (2019). 11

Figure 1-5. Electromagnetic spectrum. Source: Lillesand et al. (2015). 14

Figure 1-6. Spectral signature of vegetation based on structural and biochemical composition. Source: Karlovska et al. (2015). 14

Figure 1-7. Examples of field phenotyping platforms. (A) Phenopole, (B) phenomobile, (C) industrial portal crane fixed platform with elements detailed; (D) cable system fixed platform and (E)UAV. Sources: Deery et al., (2014), Fernandez-Gallego et al. (2018a), Kirchgessner et al. (2017) and Virlet et al. (2017). 21

Figure 1-8. Schematic overview of the steps to capture imagery using an UAV 24

Figure 1-9. (A) Botanical description of *Lolium perenne*; (B) Close-up of a breeding trial; (C) Overview of a breeding trial and (D) Top-view of a plot with canopy and after cut. Source: Flora von Deutschland, Österreich und der Schweiz 1885; ILVO 29

Figure 1-10**.** (A) Soybean developmental stages; (B) Overview of a breeding trial; (C) Top-view canopy and (D) Close-up of a breeding trial. Source part A: Manitoba Pulse & Soybean Growers (https://www.manitobapulse.ca/). 33

Figure 1-11. Schematic overview of the layout of the dissertation. 37

Figure 2-1. (A) Orthophoto of the breeding trial (set images: morning March 18); (B) subset of the 90 plots selected for the second breeder score; (C) on-ground frame imagery capture using tripod; and (D) example of one on-ground frame image. 44

Figure 2-2. Workflow for data collection, processing and analysis, with indication of the data used in each step. 46

Figure 2-3. (A) Example of two on-ground frame images; (B) Results of the supervised pixel classification of green grass (green), old grass (yellow) and bare soil (white); (C) frame area from the orthophoto; and (D) ExG2 grid calculation. 50

Figure 2-4. Linear regression with 95% confidence interval performed with the index (A) ExG2, (B) GLI and (C) GCC to estimate the area covered by green grass (A-GG). 50

Figure 2-5. Theoretical relation between ExG2, GCC and GLI. Dashed lines indicates the range of ExG2 values observed in this study. 50

Figure 2-6. Comparison of the negative selection for six cases: (A) Breeder A versus ExG2 for the flight closest to the scoring (n=102 2x accessions); (B) Breeder A versus ExG2 for the last flight (n=102 2x accessions); (C) Breeder B versus ExG2 for the flight closest to the scoring (n=135 4x accessions); (D) Breeder B versus ExG2 for the last flight (n=135 4x accessions); (E) ExG2 for the two flights showing the highest correlation (n=237 accessions); and (F) ExG2 for the two flights showing the lowest correlation (n=237 accessions). The percentages of agreement in the selections are shown, black dots represent the selections that agree and gray dots the disagreement. The quadrant corresponding to the negative selection is indicated by darker lines corresponding to the median of the axis values represented. 54

Figure 2-7. Comparison negative selection (50% worst accessions) in validation field trial FT2: (A) Breeder A versus ExG2 (n=138 2x accessions); (B) Breeder B versus ExG2 (n=96 4x accessions). The percentages of agreement in the selection are shown, black dots represent the selections that agree and gray dots the disagreement. The quadrant corresponding to the negative selection is indicated by darker lines. 55

Figure 2-8. Orthophoto and derived ExG2 from the flight over field trial 2 (FT2) of two tetraploid accessions, labeled as 4224 (plots A, B and C; n=3) and 6347/2 (plots X and Y; n=2) that showed disagreement between the breeder score (B Score) and the VI (ExG2). The mean value per accession was calculated in Trialwizard taking into account the design of the experiment. 55

Figure 3-1. Schematic overview of the experimental field trial with control and validation ground control points (cGCPs, vGCPs), artificial height references (AHR) located in four zones (numbered 1-4), *Lolium perenne* control and target plots (Lp-CPs, Lp-TPs,) and plots where biomass samples were taken (biomass samples). 65

Figure 3-2. Timelines of (i) the UAV flights, (ii) the rising plate meter measurements (RPM), (iii) the biomass measurements of the control plots (Lp-CPs), (iv) the biomass measurements of the target plots (Lp-TPs) and (v) the field trial cutting regime. Growth periods (GP) are indicated. 67

Figure 3-3. Schematic overview of the workflow followed in this manuscript. DEM: digital elevation model; CHM: canopy height model; DSM: digital surface model; DTM: digital terrain model; vGCPs: validation ground control points; AHR: artificial height references; Lp-CPs: control plots; Lp-TPs: target plots; CHMv: canopy height model variables; VIv: vegetation index variables; GDDv: environmental variables; RPM: rising plate meter; LR: linear regression; MLR: multiple linear regression; PCR: principal component regression; PLS: partial least squares; RF: random forest; RMSE: root mean square error; NRMSE: normalized RMSE; AIC: Akaike Information Criterion. 67

Figure 3-4. (A) Schematic presentation of the calculation of the canopy height model (CHM). (B) Side and top view of a set of artificial height references (AHR). (C-G) Accordance of real height and the height of the twenty eight AHR determined in four different zones of the field trial for 5 flights during growth period 3. Due to adverse weather conditions, any flight was possible between 14/07 and 01/08. The gray band around the line indicate the 95% confidence interval. 71

Figure 3-5. (A) Representation of a plot (dashed lines) and the positions measured with the rising plate meter (RPM, grey squares) in an orthophoto and canopy height model. The positions of the RPM measurements were RTK GPS referenced (black dots, yellow markers were placed in the field) to enable extracting canopy height from the same location from the canopy height models (CHM). (B-E) Comparison of the canopy height using the rising plate meter (RPM) and the mean canopy height (CH_mean) derived from the different canopy height models for the four growth periods (GP). A regression was applied per flight. The correlation of all data for each GP is calculated and shown in the upper left corner of each figure and the (dot dashed) regression line with the 95% confidence interval (gray band). (F) Comparison for the entire growing season with a linear and a loess regression and their 95% confidence interval between CHM and RPM. 72

Figure 3-6. Pearson correlation map of canopy height model variables (CHMv), vegetation index variables (VIv), environmental variables (GGDv), and rising-plate meter measurements (RPM) used for predicting dry matter yield (DMY) (n = 154). All colored correlations are significant at the $p < 0.05$ level. 73

Figure 4-1 . Schematic overview of (A) the trial with group locations (white boxes, GP1-4), (B) zoom of two plots, marked by a red box in A, with the mask applied to the images to select pixels corresponding to vegetation, (C) timeline of the sowing dates (per GP) and UAV flights and (D) workflow followed in the manuscript. 84

Figure 4-2. Schematic representation of the phenotypic parameters derived from the curve fitting for (A) canopy cover using the Gompertz function and (B) canopy height using the Beta function. By default the two last points where not taken into account in the fitting therefore they are represented by outline symbols. For abbreviations and explanations of the parameters, see Table 4-1. 87

Figure 4-3. (A) Adjusted R^2 of the curve fitting for canopy cover (CC_RsqA) and canopy height (CH_RsqA) data of all plots. A color code is used to show the percentage of emergence. (B) Zoom of the plot distribution with an adjusted R^2 higher than 0.70 (dashed lines in (A), n = 291 plots remaining for further analysis). 90

Figure 4-4. (A-F) Boxplots of the phenotypic parameters derived from the curve fitting of canopy cover data. The data is presented per group (GP 1–4) and all groups together (All). For abbreviations see Table 4-1. 91

Figure 4-5. (A-F) Notched boxplots of the phenotypic parameters derived from the curve fitting of canopy height data. The data is presented per group (GP 1–4) and all groups together (All). For abbreviations see Table 4-1. 92

Figure 4-6. PCA of (A) UAV derived phenotypic parameters and field observations (in red) and (B) individual plots (n = 148) with color code related to the group/sowing date. For the abbreviations, see Table 4-1. 95

Figure 4-7. Correlograms between the UAV derived phenotypic parameters and field observations (A-D) per group (GP1-4) and (E) grouped (n = 148). For the abbreviations, see Table 4-1. 95

Figure 4-8. (A) MLR model building overview for SY with RMSE and Adj. R^2 evolution. The red dashed line indicates the selected (best) model with the lowest RMSE. The subset selection regression with normalized UAV phenotypic variables included a 10 k-fold cross validation. (B) In each subsequent step one variable was added, and all possible combinations were evaluated. The model index = 0 included only the groups (GP1-4, GP1 is the reference group) as dummy variables. For the abbreviations, see Table 4-1. Graphs are provided per variable with the value of their respective coefficients. 98

Figure 4-9. (A) Representation of how data could be used to select interesting plots in a breeding program based on phenotypic data using three parameters: CC_6W, CH_Hmax, and CC_SNC. (B) Zoom of the plot distribution based on a applicable values range for selection. For the abbreviations, see Table 4-1. 99

Figure A-1. Breeding cycle ryegrass - Generalized breeding scheme for ryegrasses describing the steps to perform at different scales and years. Source: Humphreys et al. (2010) 142

Figure A-2. Breeding cycle soybean - Generalized breeding scheme for soybean. Source: U. Guelph (dept. of plant agriculture) NAPB web seminars 2015 143

Figure A-3**.** Evolution of thermal time (°C d) during the growing season of the trial. First flight and the two weeks gap are indicated. 144

Figure A-4. Histograms with median values (dashed lines) of the phenotypic parameters (A-F) derived from the curve fitting of canopy cover data. The data is presented per group (GP 1-4) and all groups together (All). For the abbreviations see Table 4-1. 145

Figure A-5. Histograms with median values (dashed lines) of the phenotypic parameters (A-F) derived from the curve fitting of canopy height. The data is presented per group (GP 1-4) and all groups together (All). For the abbreviations see Table 4-1. 146

Figure A-6. MLR model building overview for SY stage with RMSE and Adj.R^2 evolution. The red dashed line indicates the selected (best) model with the lowest RMSE. The subset selection regression with normalized UAV time points included a 10 k-fold cross validation. In each subsequent step one variable was added and all possible combinations were evaluated. The model index = 0 included only the groups (GP1-4, GP1 is the reference group) as dummy variables. Graphs are provided per variable with the value of their respective coefficients. 147

Figure A-7. MLR model building overview for R8 stage with RMSE and Adj.R^2 evolution. The red dashed line indicates the selected (best) model with the lowest RMSE. The subset selection regression with normalized UAV phenotypic variables included a 10 k-fold cross validation. In each subsequent step one variable was added and all possible combinations were evaluated. The model index = 0

included only the groups (GP1-4, GP1 is the reference group) as dummy variables. For the abbreviations see Table 4-1. Graphs are provided per variable with the value of their respective coefficients. 148

Figure A-8. MLR model building overview for R8 stage with RMSE and Adj.R^2 evolution. The red dashed line indicates the selected (best) model with the lowest RMSE. The subset selection regression with normalized UAV time points included a 10 k-fold cross validation. In each subsequent step one variable was added and all possible combinations were evaluated. The model index = 0 included only the groups (GP1-4, GP1 is the reference group) as dummy variables. Graphs are provided per variable with the value of their respective coefficients. 149

Figure A-9. Zoom of the same area in the last three flights carried out in which can be observed a decrease in the height values due to lead drop in the senescence stage. 151

Figure A-10. Rising plate meter used in a grassland field. How it is placed: first the pole is posed on the ground, then a plate of 30x30 cm is gently placed on top of the canopy and based on the distance between the plate and the device on top and the total length of the pole, the height of the canopy is calculated and registered. 152

Chapter 1: Remote sensing-based plant phenotyping

1.1 Feeding the world under a changing climate

Food security comprises not only prevention of hunger, but refers also to access to food and to its quality and nutritive value. According to recent estimations, over two billion people do not have regular access to safe, nutritious and sufficient food (FAO, IFAD, UNICEF, WFP and WHO, 2019). Furthermore, climate change poses a serious threat to food security because, depending on the region, altered climatic conditions can result in an increased prevalence and or intensity of abiotic stresses such as drought, flooding, heatwaves, and biotic stresses with increased attacks of pathogens and pests (i.e. a state where the crop is growing in a non-ideal growth conditions (e.g. lack of water) that lead to deficiencies in growth, yield, permanent damage or death). Therefore, a detrimental effect of climate change on agricultural productivity is expected in large areas of the world. This negative impact on the food system and on rural livelihoods will further increase the challenges related to food security, nutrition and health (FAO, IFAD, UNICEF, WFP and WHO, 2019).

The need to adapt to future climatic situations should be considered in combination with the need to increase global crop productivity to fulfil the demands of approximately 9.5 billion people. As this should be achieved in a sustainable way, identification and implementation of innovative strategies for sustainable crop productivity, increased use efficiency of resources and inputs in addition to the protection of the remaining natural resources and agro-ecosystems is imperative (Shah and Wu, 2019). Considering that the impact of climate change differs in different regions of the world and the high diversity of agricultural systems, tackling these agronomic and climatic challenges requires targeted solutions for different agro-ecological and pedoclimatic zones (Altieri and Nicholls, 2017). Measures to adapt to climate change and to reduce the impact of agricultural production on the natural environment range from technological solutions to adjustments in farm management. Possible measures include, among others, adjusting farm operations and their timing (performing the right application on the right place at the right moment (i.e. precision agriculture)), cultivating crop species adapted to the new environmental conditions or selecting improved crop varieties through breeding (Peng et al., 2020). In this context, crop adaptation is considered a key element to reduce the environmental impact of agriculture whilst securing crop productivity (Chawade et al., 2018). This requires the continuous release of high-yielding varieties of agricultural crops adapted to our future climate and cultivation conditions. Therefore, plant breeding (i.e. the process to develop and release new plant varieties) is key to sustain agricultural productivity while strengthening both ecological and societal resilience, and to achieve six international sustainability targets: food security, safety and quality; food and seed sovereignty; social justice; agrobiodiversity; ecosystem services and climate robustness (Lammerts van Bueren et al., 2018). In general terms, plant breeding involves the evaluation (i.e. phenotyping) of thousands of plants, the creation and evaluation of new candidate varieties, to finally retain a small selection of the best performing ones. The duration of this whole process is variable depending on the crop and the breeding system applied, but in most agricultural crops it typically takes at least a decade to develop and release a new variety (Lenaerts et al., 2019). This process can be accelerated by

implementing recent advances in different scientific disciplines, including plant phenotyping, but also by implementing genomics-based selection, optimizing the breeding strategy or accelerating the generation turnaround (Li et al. 2018, Figure 1-1). In all cases the incorporation of a new method or technology into the breeding program has as objective to increase the genetic gain per generation (see further in this introduction). This PhD dissertation focuses on the development and evaluation of phenotyping methods suitable to phenotype multiple genotypes in the field. In the following sections traditional phenotyping methods, available imaging platforms and the potential of field high-throughput plant phenotyping are discussed in detail.

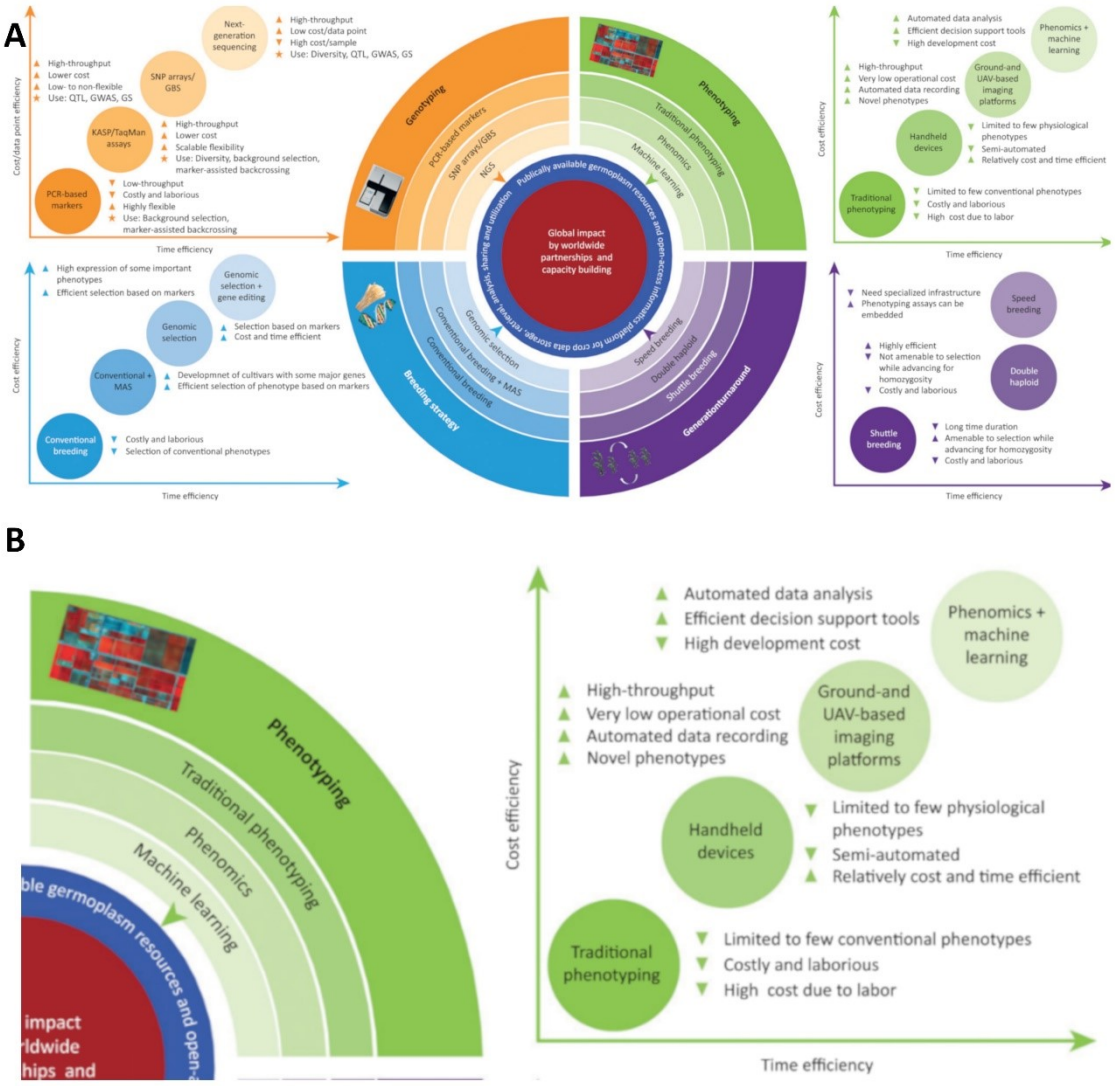


Figure 1-1. A) Integrated strategies leading to a common decision support tool to accelerate genetic gain based on the use of new technologies. Each field is described with key advantages (Λ) and disadvantages (V). B) Detail of the discipline tackled in this dissertation. Source: Li et al. (2018).

1.2 The role of plant phenotyping in breeding and biological research

All breeding programs aim at improving the performance of a particular crop within a specific environment and agricultural context (Casler and van Santen, 2010). The specific objectives are highly

diverse and depend on the crop, the intended use and the target environment. For example, grain protein content is an important objective in soybean, while leaf protein content is a relevant criterion in forage crop species such as perennial ryegrass. Similarly, drought tolerance is only relevant for environments in which water shortage is a risk, while flooding tolerance might be relevant in other regions.

The breeding scheme applied depends to a large extent on the reproduction system of the crop considered (e.g. a self-pollinating crop like soybean versus an obligate cross-pollinating crop like perennial ryegrass), but controlled mating of selected individuals and selection of best performing individuals/populations among the progenies are the essence of each plant breeding program. As a consequence, the breeding process comprises three main types of activities: hybridization, genotype fixation (e.g. line, clone or population) and evaluation of performance in (field) trials. In a typical breeding program (field) evaluation activities imply a large-scale logistical operation involving thousands to hundreds of thousands of plants at the start of the evaluation and selection process, to end with a small number of selected breeding lines or genotypes that are then released as varieties. In general terms, the improvement is accomplished by selecting plants which carry desirable characteristics, which are named **traits** (e.g. plant height, seed protein content, flowering time). Thus, a trait can be defined as a morphological, biochemical, physiological, structural, phenological or behavioral characteristic that may be measured which influences plant performance or status. The process of assessing the performance of a genotype for a given set of traits in a given environment is called **plant phenotyping**.

Plant phenotyping is as old as plant breeding and still largely depends on visual inspections and labor intensive destructive measurements. While these approaches are still highly valuable in breeding context, they have as drawback subjectivity (i.e. in the case of visual inspections) and lack of temporal resolution to describe processes such as plant growth (i.e. when only end-point measurements of biomass yield are performed). As a consequence, these methodologies are increasingly complemented and replaced by more advanced methods based on new technologies. The new phenotyping tools allow to evaluate plant performance in a more objective and efficient manner and can assist the breeding process enormously, as described further in this PhD thesis. The term **phenomics** was recently introduced to refer to the study of crop growth, performance and composition via high-throughput accurate acquisition and analysis of multidimensional phenotypes throughout development (Furbank and Tester, 2011; Tardieu et al., 2017).

Plant breeding programs benefit also from basic biological knowledge in the area of genomics, and the implementation of technological advances in the areas of phenomics and genomics is considered essential to facilitate and accelerate plant breeding (Figure 1-1, Li et al., 2018; Xu et al., 2017). Genomics involves the study of the structure, function and inheritance of all genes of an organism (the genome). The development of new agricultural germplasm that supports future agricultural production can benefit from a good understanding of how genetic instructions from a single gene or the whole genome translate into the full set of phenotypic traits of an organism depending on the environment in which it develops. This requires: (i) genomic information collected for a sample of genotypes, (ii) genotypes exposed to a range of relevant environments, and (iii) phenotypes measured with high spatial and temporal resolution at low cost. Genomic information can then be combined with phenomic information using analytic approaches such as Genome Wide Association Mapping (GWAS) to identify

which genomic regions (and possibly which genes) are responsible for plant performance, in interaction with the environment (see Gupta et al., 2019 for a recent review). Accurate phenotyping approaches that allow the recording of plant responses at relevant spatiotemporal resolution are therefore also essential in GWAS. However, our current capacity to collect useful high-quality phenotypic data lags behind the current capacity to generate high-throughput genomics data, resulting in an underutilization of the progress achieved in genomics, i.e. a "phenotypic gap or bottleneck" (Houle et al., 2010; Mir et al., 2019; Tripodi et al., 2018). Indeed, recent developments in next-generation DNA sequencing technologies have boosted the possibilities to unravel and understand crop genomes tremendously, but the quality of a typical GWAS study depends to an even larger extent on the resolution and quality of the phenotypic data considered.

We can therefore conclude that precise phenotyping using high-throughput tools is needed to identify the genetic and environmental components which influence complex traits (Tripodi et al., 2018) and unravel the interactions between genotype, environment and management in fundamental and applied plant science and in germplasm screening in breeding programs (Roitsch et al., 2019). Consequently, the implementation of efficient phenotyping approaches is considered a key element to improve resilience and to adapt crops to biotic and abiotic stresses related to climate change and finally increase production by 25-70% to meet the food demand by 2050 (Araus and Kefauver, 2018; Hunter et al., 2017).

1.2.1 Definition of traits and their measurement

Breeders have to decide 'what' and 'how' to measure in order to obtain germplasm that fulfils the target, e.g. higher yield (Watt et al., 2020). Most traits of relevance for breeding, such as seed yield, biomass yield, fruit size, etc., are the result of biophysical processes that occurred over time such as an entire crop growth cycle (Weiss et al., 2020). These traits are rather complex as they are an integration of several characteristics/processes over time. For example, if we consider seed yield, a good performance can be achieved through structural optimization (e.g. reduction of plant height to avoid lodging), optimization of photosynthesis (e.g. increase of canopy radiation use efficiency), increase of disease resistance or optimization of flowering time. It might therefore be interesting, next to the evaluation of the performance at the level of the complex trait, to decompose the trait into its components and consider them separately. Analysis of these 'more simple' traits is more straightforward. In addition, this approach allows to investigate the contribution of the different components and their interactions to the overall performance of the crop. If the trait itself is difficult to measure or if it requires expensive procedures, it may be quantified in an indirect manner by measuring traits that are related to it, but easier or cheaper to measure (Prasad et al., 2007). Under some circumstances it might even be relevant to use secondary traits measured in an earlier developmental stage or generation to accelerate selection and shorten the breeding cycle (Richards, 2000).

Traits can be grouped in classes according to different criteria. In Table 1-1 traits have been classified in three groups: 'Growth, development, phenology and morphology', 'Physiology' and 'Biotic interactions'. The way in which the traits are measured can range from methods based on visual inspections performed by experienced operators to the use of sophisticated instruments. Visual scoring, ruler measurements and counts are means to follow-up genotype performance non-destructively. Ruler

measurements and counts can deliver objective information, when standardized protocols are defined and applied. Visual scoring can be performed very quickly by well-trained operators. Visual scoring is, however, not free of constraints as (i) it can be subjective and have low repeatability across observers (Roitsch et al., 2019), (ii) introduce scorer-specific bias even if they are appropriately trained (Bai et al., 2018) and (iii) be influenced by interference from surrounding plants (Wang et al., 2019). Physiological traits can be determined objectively and precisely using devices such as a ceptometer for light interception, an IR gas analyzer to measure gas exchange or a SPAD for measuring chlorophyll content. On the other side, these measurements are (i) time-consuming, and hence (ii) expensive to carry out and (iii) do not cover spatial variability sufficiently. Destructive samples can deliver precise information, but collecting them might be laborious, time consuming and costly, and can only be done at harvest.

Table 1-1. Main breeding traits for crops and their conventional trait evaluation method. Adapted from: Araus et al. (2018); Reynolds et al. (2020) and Roitsch et al. (2019).

Breeding trait	Traditional evaluation method
Growth, development, phenology and morphology	
Heading and maturity	Visual scoring
Plant establishment /emergence	Visual scoring or counting N plants/n seeds
Early vigor	Visual scoring
Winter hardiness	Visual scoring
Yield / biomass	Destructive samples (fresh and dry weight)
Lodging	Visual scoring
Root development	Excavations, soil coring
Plant/canopy height	Ruler, rising plate meter
Green cover/cover fraction	Visual scoring
Persistency	Visual scoring
Growth stage determination	Visual scoring
Senescence	Visual scoring
Stay-green	Visual scoring
Physiology	
Light interception	Ceptometer
Water use efficiency/water content/drought stress	Destructive samples, water potential measurement
Photosynthesis/transpiration rate	Gas exchange, chlorophyll fluorescence
Canopy temperature	Thermography
Chlorophyll content	SPAD measurement
N uptake efficiency	Indirect estimation of N Isotopic tracer ^{15}N tracers, handheld NDVI measurements
Biotic interactions	
Disease resistance	Visual scoring
Weed infestation	Visual assessment

Genetic gain is a central concept in breeding as it reflects the breeding progress. It can be defined as the amount of increase in performance that is achieved annually for a particular trait (Fischer and Edmeades, 2010; Xu et al., 2017). As illustrated in Figure 1-2, the genetic gain can be increased by (i) increasing the size of the breeding program to enable higher selection intensity, (ii) enhancing the

selection accuracy (higher repeatability), (iii) ensuring adequate screening and use of genetic variation, (iv) accelerating the breeding cycles and (v) improving decision support tools (Araus et al., 2018; Masuka et al., 2017). This implies that the performance of a breeding program can be improved by upscaling and increasing the accuracy of the screening of a large(r) number of plants, but this is usually hampered by the inability to do so using the traditional way of working (Wang et al., 2019). Therefore, reliable high-throughput phenotyping can contribute directly or indirectly to higher genetic gains, as it can allow a more efficient, and more detailed and accurate evaluation of a large number of plants. In combination with genomics information, precision phenotyping tools are expected to enhance the prediction accuracy and to increase genetic gain by shortening breeding cycles or by achieving a higher gain per breeding cycle (Cobb et al., 2013; Crossa et al., 2017).

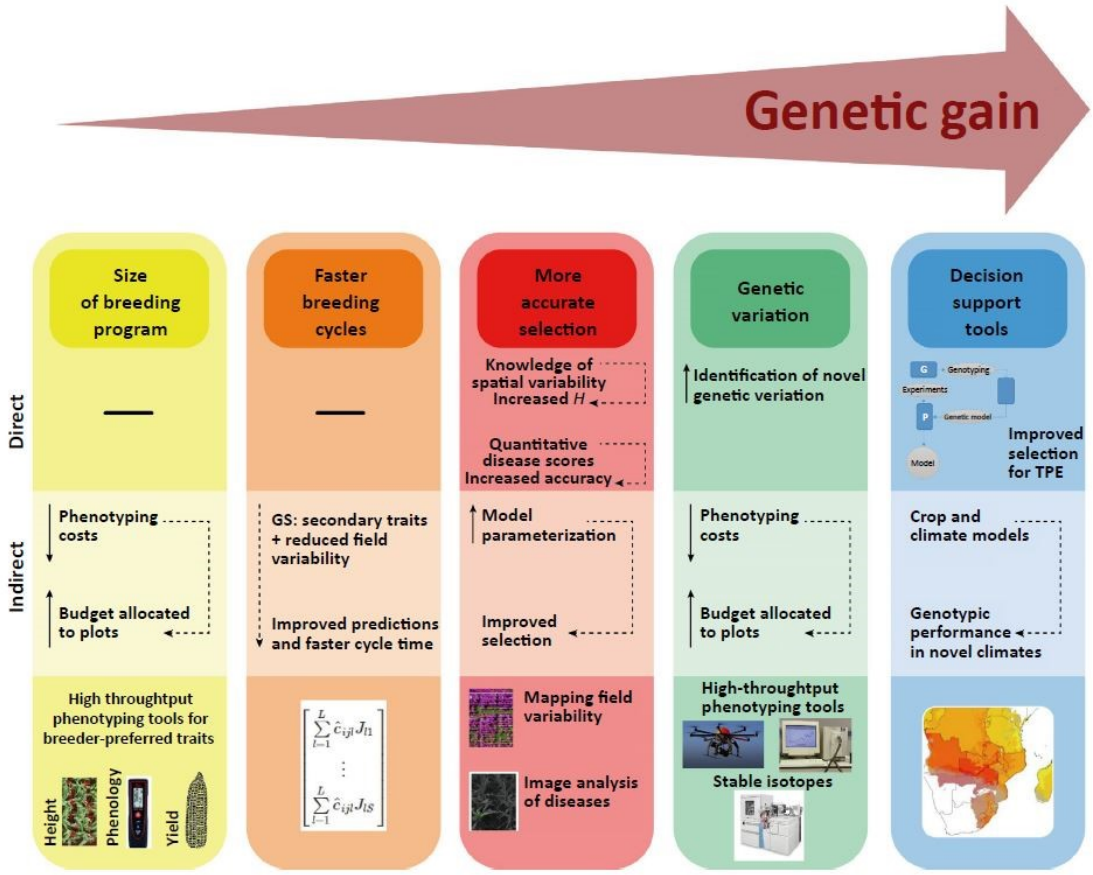


Figure 1-2. Key elements to accelerate genetic gain in which high-throughput phenotyping can contribute (directly or indirectly): increase of the size of the breeding program, breeding cycles acceleration, increase selection accuracy, ensure genetic variation and improve of the decision support tools. TPE = Target populations of environments (where the breeding products would be grown). Source: Araus et al. (2018).

1.2.2 Phenotyping methods

Phenotyping is defined as the set of methods and protocols applied to accurately and precisely determine plant growth, architecture and composition at different scales (from organs to canopies) (Fiorani and Schurr, 2013). Direct quantification of the phenotype involves measurement of structural and functional characteristics (e.g. plant biomass, root morphology, leaf characteristics, photosynthetic

efficiency and biotic or abiotic stress resistance/tolerance) as well as the chemical composition of plants or plant parts (e.g. feed and food related quality, secondary metabolites with roles in plant defense or biosphere-atmosphere interactions). The method chosen to phenotype depends on the crop, the trait to evaluate, the developmental stages and the resources available. As it is typically necessary to evaluate multiple traits, a method that allows to evaluate multiple traits in parallel is recommended to reduce costs, time and possible errors related to revisiting or involvement of many operators in data acquisition (Chawade et al., 2019). In addition, the size of a typical phenotyping trial varies between hundreds to thousands of plants/plots, requiring the implementation of efficient measurement methods and protocols.

Phenotyping can be carried out in different conditions: controlled, semi-controlled or natural. Controlled conditions are achieved in growth-chambers and greenhouses and can be automated using conveyor belts or robots to transport the plants to the measurement unit. In these settings, a broad set of environmental variables can be manipulated to simulate the growth conditions that are relevant for the phenotyping experiment. A semi-controlled environment can be established in the field and allows to alter or control part of the environmental conditions to simulate specific situations (e.g. using rain-out shelters to simulate dry conditions). Natural settings correspond to the field, and have as advantage that crop-sized plants can be grown, but are constrained to seasonal growing conditions (Fahlgren et al., 2015), requiring repeated evaluations over different years or locations. In what follows a non-exhaustive overview is provided of recent advances in the area of plant phenotyping under controlled and natural environments, focusing on image-based approaches.

1.2.2.1 Controlled condition phenotyping

Automatic phenotyping of plants grown under controlled conditions allows to dissect plant responses to different levels of a specific environmental variable, a stress condition or combinations of these, and to describe temporal changes in this response. In such strict controlled conditions, high levels of replicability and reproducibility can be achieved. Furthermore, it is possible to perform several cycles of evaluation per year.

A crucial advance in phenotyping was the introduction of non-invasive methods to capture data at the plant and canopy level. The capacity and detailed data provided by **imaging sensors**, methods to automate the procedure and the computation power required became affordable and made a range of applications in controlled conditions possible (Fiorani and Schurr, 2013). This kind of set-ups typically allow acquisition of multi-temporal data to follow the progression of growth, while increasing the experimental capacity for genotypes, treatments and biological replicates by reducing or even eliminating the number of replicated samples needed when invasive methods are applied (Fahlgren et al., 2015). Recording data for hundreds to thousands of plants in a single day is the ideal situation in controlled condition phenotyping platforms (Li et al., 2014). The phenotyping effort can become rather efficient in controlled environments, as the accuracy, precision and throughput of the measurements can be high, and costs can be minimized through automation, imagery, data integration and optimal experimental design and analysis pipelines (Junker et al., 2015; Mir et al., 2019).

Phenotyping platforms in controlled conditions are diverse: from a simple bench-top cabinet, a growth chamber with a robot to automate the phenotyping process, to fully automated conveyor belt systems with stationary imaging platforms (Figure 1-3). Image acquisition is often executed in a dedicated chamber, under stable light conditions adjusted for optimal imaging and subsequent automatic image analysis and data processing. Examples of high-throughput plant phenotyping platforms are GROWSCREEN (Walter et al., 2007), PHENOPSIS (Granier et al., 2006), Phenoscope (Tisné et al., 2013), InfraPhenoGrid (Pradal et al., 2017) or Scanalyzer HTS (Arvidsson et al., 2011). More examples can be found in Campbell et al. (2018); Mir et al. (2019); Rahaman et al. (2015); Varshney et al. (2018) or Zhao et al. (2019).



Figure 1-3. Controlled phenotyping platforms. Examples of A) top-view imaging cabinets; B) automatic systems for root or top-view imagery and C) side-view imaging conveyor belt system with resulting imagery of barley. Sources: Lemnatec website (https://www.lemnatec.com/), Emphasis website (https://emphasis.plant-phenotyping.eu/) and Chen et al. (2014).

Phenotyping in controlled conditions is typically used to identify genotypes, traits and genes of potential interest for crop improvement, and to develop and test hypotheses concerning biological processes. Main limitations when agricultural crops are evaluated are the fact that plants are grown at low light intensities, temperature profiles do not match target environments, the root growth is limited by the pot size, planting densities are irrelevant for agronomic purposes and in most cases it is impossible to

evaluate a complete crop growth cycle from sowing to harvesting (Junker et al., 2015; Poorter et al., 2016, 2012). Consequently, given the complexity of the interactions across traits and the interactions with the environment, the results obtained in growth chambers or greenhouses should be validated in field environments to test their relevance for breeding (Tardieu et al., 2017). In Poorter et al. (2012) recommendations to design and set up an experiment in controlled conditions are presented such as where to grow the plants, which rooting medium and pot size to use, number of plants required and environmental factors to take into account for plant growth (e.g. light quantity or air humidity). A simple checklist to track and report experimental conditions is also proposed.

Humplík et al. (2015) discuss a series of stress-related studies using automated shoot phenotyping in controlled conditions and indicate that drought is one of the most widely monitored traits in these phenotyping systems. Granier et al. (2006) reported on one of the first studies related to drought tolerance in *Arabidopsis* accessions using a phenotyping platform equipped with an RGB sensor. Based on leaf area measurements and pot weighing to assess transpiration rates, accessions with low sensitivity to soil water deficit were identified. Further developments involved capturing imagery from different positions (not only top-view) allowing to extract plant height and combining plant height with plant width and leaf colors, Neumann et al. (2015) investigated chlorophyll degradation caused by drought in barley. Pereyra-Irujo et al. (2012) used a self-constructed platform with RGB imagery built in the greenhouse to evaluate growth and water-use efficiency in soybean. Growth and water-use efficiency were also studied in barley by Honsdorf et al. (2014) using a commercial platform of LemnaTec (LemnaTec, GmbH, Germany).

Traits related to crop establishment can also be investigated in controlled conditions using measurements of emergence, early vigor, development of leaf area, early biomass accumulation and root system elongation (Fiorani and Schurr, 2013). When the use of RGB sensors is combined with NIR, hyperspectral, fluorescence or thermal cameras, the different sources of data (reflecting emission at different wavelengths of the electromagnetic spectrum) can be combined to phenotype different traits. For example, growth, biomass accumulation or nutritional status in barley (Chen et al., 2014), water use efficiency in maize (Ge et al., 2016), leaf chemical properties in different crops (Pandey et al., 2017), genetic architecture of plant growth in maize (Zhang et al., 2017) and rapeseed (Li et al., 2020), drought tolerance in rice (Guo et al., 2018) and salinity tolerance in chickpea (Atieno et al., 2017). Root phenotyping under controlled conditions has also been demonstrated with automated imaging platforms, enabling to capture images of roots and above-ground plant parts simultaneously (Watt et al., 2020). If images are taken at different points in time, it is even possible to monitor root development.

In a typical experiment in controlled conditions, a single stress or a combination of two stress factors is investigated, and often at a specific development stage. This is ideal to understand detailed responses of crops to these particular conditions. However, in a field environment, i.e. the crop production growing conditions of farmer fields, the situation is much more complex. Mimicking these environmental conditions in controlled conditions is not feasible. Consequently, results obtained under controlled conditions often do not correlate well with the expression in the field (Poorter et al., 2016). Therefore, plant phenomics should not be limited to linking a genotype to a phenotype in a controlled condition, but ideally it should also characterize the plasticity of the plant phenome exposed to a range of environmental conditions (Tardieu et al., 2017).

1.2.2.2 Field plant phenotyping

Typical field phenotyping trials have as purpose to discover novel traits, identify germplasm with relevant traits, validate traits (e.g. proof of concepts), monitor dynamic traits through the season and measure multiple phenotypes (Watt et al., 2020). The goal is to determine the performance of a genotype in an environment that resembles the 'natural or target' environment to a larger extent than controlled condition experiments. This situation allows to investigate the plant/crop response to the environment in a more integrated manner and under more realistic settings. This information can be used to refine hypotheses defined in controlled conditions, raise new research hypotheses or lead to new screening protocols (Cobb et al., 2013). From the perspective of breeding, phenotyping in field conditions will maximize the relevance to breeders as it allows to evaluate the phenotype over a full growing season and estimate the interaction between genotype and environment. Breeding trials usually involve small plots for evaluation of lines or populations for different (agronomic) traits (e.g. yield) to perform a selection (Chawade et al., 2019).

Field phenotyping is an appropriate choice in particular for traits that are strongly influenced by the growth environment or that are expressed only under growth environments that are difficult to reproduce in controlled environment facilities (Tripodi et al., 2018). However, field phenotyping entails difficulties compared to controlled conditions such as environmental heterogeneity (e.g. spatial variation in the soil), uncontrollable weather conditions, and variable light conditions for imaging. In addition, it should be kept in mind that the effect of (or the response to) different environmental factors might be difficult to separate when plants are evaluated in the field. Indeed, in controlled conditions it is possible to control light intensity, temperature and relative humidity independently, but in the field these will be highly correlated (warm – dry – sunny weather vs. cold – wet – cloudy weather). An additional limitation, concerning the use of imagery, is that most field phenotyping methodologies are based on a top-of-canopy view which has restrictions as lower plant parts are hidden from the sensor's field of view (Watt et al., 2020). In Figure 1-4, an overview of the phenotyping needs and challenges in breeding and precision agriculture is presented. At this moment, only the top part of the figure concerning breeding is relevant, further comments about the differences between them will be tackled in the conclusions section.

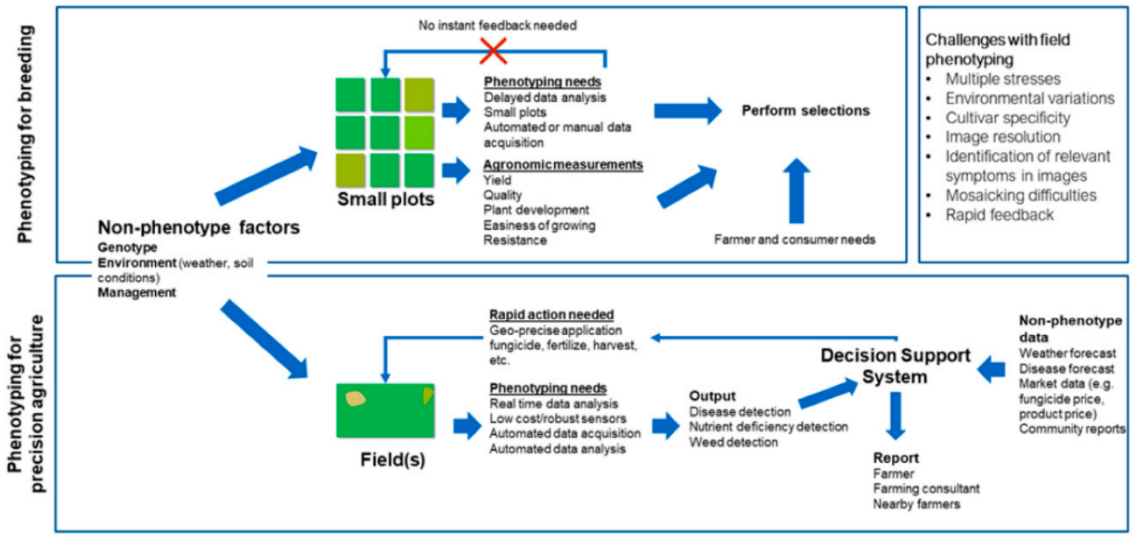


Figure 1-4. Phenotyping needs in breeding and precision agriculture scenarios. Source: Chawade et al. (2019).

Field phenotyping using remote sensing is particularly interesting as it allows efficient and repeated screening of large collections of plants and can replace or complement labor intensive on-ground measurements. In addition, current technologies allow to achieve a high level of resolution and precision, and are suitable to detect spatial variability. This is referred to **high-throughput field phenotyping (HTFP)**. One of the main advantages of these systems is their versatility, as the frequency and resolution of data collection can be chosen and adapted to the trait(s) to measure allowing customization based on the objective of the study. It should not be forgotten that environmental data collection is key to understand differences between genotypes and their response (in the form of growth, development or yield) to the environmental conditions (Reynolds et al., 2020). Although HTFP holds great promise for plant breeding, cost efficiency must also be considered when implementing it in breeding programs (Reynolds et al., 2019). The new phenotyping approach should ideally lead to a reduction of the overall phenotyping cost, allow to achieve a higher level of resolution, deliver complementary information which is difficult to obtain by other means, or bring any other further advantage to the breeding program.

HTFP is also of high value for genetic research, as the availability of detailed (if required, multi-temporal) phenotypic data of plants evaluated in the field allows to link processes that occur at molecular level with morphological characteristics and use linkage mapping or GWAS to dissect the genetic architecture of complex traits (Bazakos et al., 2017). By measuring phenotypes at high resolution in large collections of plants, a high-resolution screening of the genome is possible and the genetic factors that determine relevant plant phenotypes can be identified.

Also for plant and crop modelling, multi-temporal phenotypic data at high resolution is of great value, because estimation of a large number of parameters in an identifiable manner, as required in modelling, requires the use of large quantities of experimental data (De Swaef et al., 2019). Two different approaches can be used to link the (physical) phenotypic measurements to the trait to estimate. One approach can be to use an empirical model which directly relates inputs to outputs by mathematical

relations (e.g. growth curve fitting). The other approach is a mechanistic model focusing on the causality between inputs and outputs by describing the mechanisms involved and relying on assumptions (e.g. functional structural plant model) (Baker et al., 2018). A comparison of both approaches can be found in Feng et al. (2014).

From the high number of publications and reviews related to field phenotyping published, it is clear that there exists a strong interest and that numerous research groups are working on this topic. More information can be found in Araus et al. (2018), Araus and Kefauver (2018), Chawade et al. (2019) and Mir et al. (2019). For further information about phenomics, we refer to Zhao et al. (2019) and Watt et al. (2020). In these reviews, the history of crop phenomics from controlled to field environment is presented and future prospects in this area.

1.2.3 The plant phenotyping community

In Costa et al. (2019) a bibliometric mapping is presented to analyze developments and the evolution of plant phenotyping research. It can be observed that the community has rapidly grown in the last decades and more advances and progress are expected in the coming years.

Challenges that plant phenotyping encounter are the necessity (i) to handle huge amounts of data generated from diverse sources or devices and its heterogeneity, (ii) data management, (iii) storage, (iv) the processing speed and (v) the quality/robustness/veracity of the data. Another challenge is (vi) the need for data exchange in a standard and comprehensive way following the FAIR data principles (Findable, Accessible, Interoperable and Reusable) to enable interoperability between platforms and reproducibility of the results. Moreover, (vii) guidelines and standard protocols are lacking and (viii) validation and usefulness of the data to support decision-making (Araus et al., 2018; Chawade et al., 2019).

To overcome these problems, the exchange of methods and protocols is key. Therefore, websites and online databases are established to share phenotyping know-how (Roitsch et al., 2019). For example, the website Index Database (https://www.indexdatabase.de/) provides information on (remote sensing) indices and their application, while the Quantitative Plant (https://www.quantitative-plant.org/) provides image analysis software tools and plant models. Miappe (https://www.miappe.org/) is an open community set-up related to the minimum information that has to be registered related to plant phenotyping experiments.

In addition, to support the rapid developments in the area of plant phenotyping and to join forces, several international networks have been created such as the International Plant Phenotyping Network (IPPN, https://www.plant-phenotyping.org/) and the European Plant Phenotyping Network (EPPN, https://eppn2020.plant-phenotyping.eu/). The main aims of IPPN and EPPN are to provide all relevant information about plant phenotyping, to increase the visibility and impact of plant phenotyping and enable cooperation by fostering communication between stakeholders in academia, industry and government. Another example of an international collaborative network was the Cost action FA1603 (2014-2018) 'The quest for tolerant varieties – Phenotyping at plant and cellular level' (https://www.plant-phenotyping.org/home_costfa1306), which allowed to organize meetings and trainings to facilitate collaboration. Nowadays EMPHASIS (https://emphasis.plant-phenotyping.eu/), the

successor of EPPN continues to gather plant phenotyping knowledge and tools and is building a robust network among researchers and between researchers and users to exchange experiences (Pieruschka and Schurr, 2019). In 2018 an 'Emphasis - Field Phenomics workshop' was organized at ILVO (Belgium) where both industry and academia from different countries discussed the future of field phenotyping. The number of attendants and the report clearly demonstrate the interest of all parties related to field phenotyping (https://emphasis.plant-phenotyping.eu/meeting_reports).

1.3 Remote sensing for field phenotyping

Remote sensing is the process of detecting and monitoring the physical characteristics of an area or object by measuring its reflected and emitted radiation at a distance. It involves the use of sensors to capture data in a rapid and non-invasive way. These sensors cover specific parts of the electromagnetic spectrum (Figure 1-5) and measure the amount of (incoming solar) radiation which includes radiation from short to very long wavelengths that is reflected back from the objects of interest.

In the case of vegetation, the reflectance is influenced by its physiology, biochemical and structural (leaf) composition and is controlled by three physical mechanisms: absorption, reflection and transmission. Any changes in leaf structure caused by plant genotype, leaf age (e.g. senescence), health or physiological status (e.g. related to light environment, nutrition status or water supply), may affect absorption, transmittance and reflectance of electromagnetic waves and contribute to fluctuations in the spectral signature (Carter, 1993; Guangyu et al., 2012; Mohammed et al., 2000). Henceforth, the vegetation has a specific '**spectral signature**' or pattern which represents the fraction of the reflected radiation as a function of wavelength (Figure 1-6) and that can be used to evaluate plant growth characteristics and status. The visible (VIS, 400–700 nm), near-infrared (NIR, 700–1300 nm) and short-wavelength infrared (SWIR, 1300–2500 nm) regions contain information about leaf physiological status, water content, pigments and structural components of biomass. In the visible region, the amount of light reflected at each wavelength depends on the concentrations of photosynthetic pigments (chlorophyll, carotenoids and anthocyanins) in the leaf palisade cells which absorb the majority of the (incident) radiation (Ustin et al., 2009). The higher absorption by chlorophyll in the wavelength ranges corresponding to the blue and red region results in the green appearance of leaf tissue. Carotenoids absorb in the blue and green region, resulting in orange and red colors. In the case of anthocyanins absorption occurs in the green region, resulting in blue and purple colors. In the near-infrared region (NIR), reflectance is higher and determined by cell and tissue structure of leaves. In addition, canopy structure also plays a role as NIR radiation is transmitted from the upper to the lower leaves of the canopy which can reflect solar radiation back to the upper section of the leaf cover (Gibson, 2013). The red edge part (ascending part of spectrum curve at 680–750 nm, explained by chlorophyll absorbing light in a lower amount in this region than in the visible part of the spectrum) shifts as the plant grows and matures and can be used as an indicator of water stress and plant nutritional status (Filella and Peñuelas, 1994; Gibson, 2013). Components such as lignin and cellulose and especially water absorb the radiation (i.e. reflectance decrease) in the shortwave-IR region (SWIR). Therefore, the particular response in this region can be linked to leaf water content (Kokaly et al., 2009). For example, healthy plants reflect more NIR light than stressed plants. And stressed plants reflect more red wavelength light than healthy plants. This is explained by the loss of tissue structure and the reduced absorption in the

chlorophyll active band in stressed plants, there is a general tendency of stress to change leaf chlorophyll concentration (Karlovska et al., 2015; Mohammed et al., 2000).

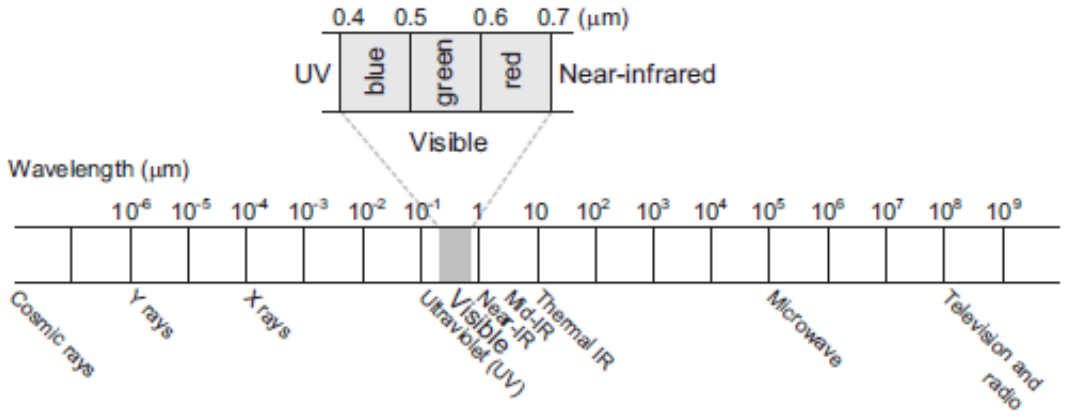


Figure 1-5. Electromagnetic spectrum. Source: Lillesand et al. (2015).

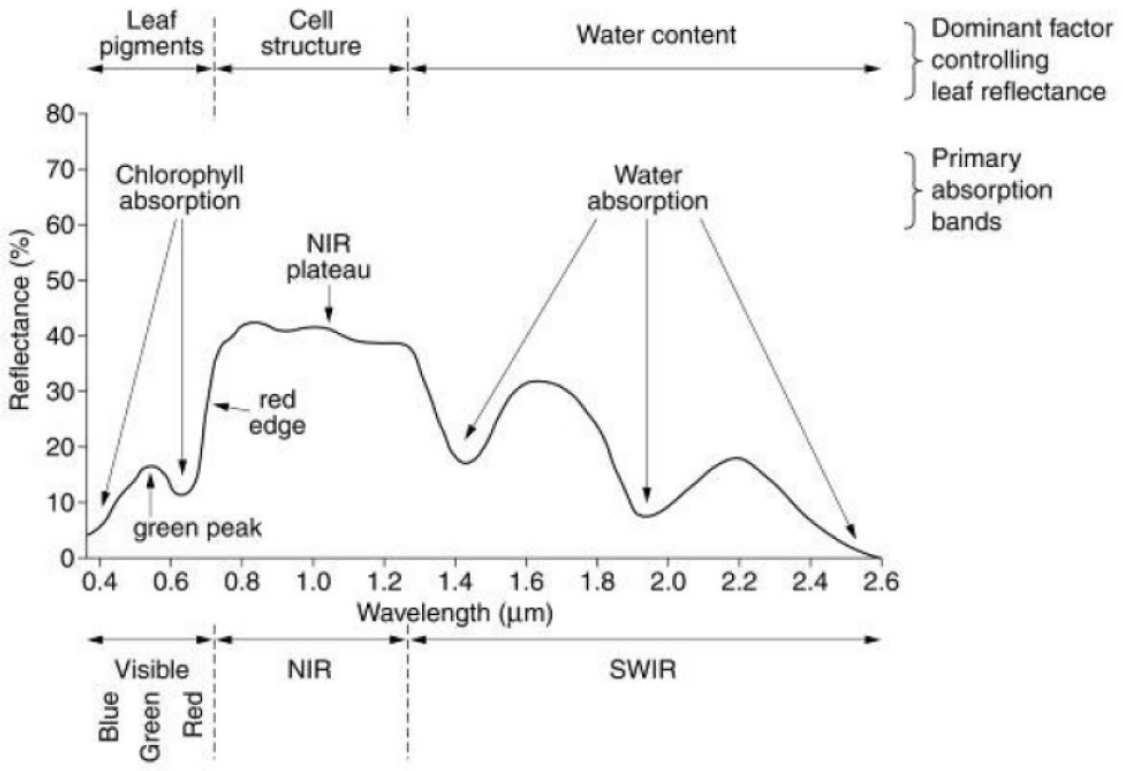


Figure 1-6. Spectral signature of vegetation based on structural and biochemical composition. Source: Karlovska et al. (2015).

1.3.1 Sensors

In Table 1-2 a non-exhaustive overview of the different kinds of sensors used in remote sensing is given with some of their advantages and disadvantages. Two main types of sensors can be identified: **passive sensors** rely on solar radiation as light source, while **active sensors** include an artificial light source (e.g. LiDAR – Light Detection and Ranging or SAR – synthetic Aperture Radar). The discussion will be limited to the passive sensors as it is the type of sensor used in this dissertation because they provide spectral information, have enough spatial resolution to apply Structure from Motion (explained further) to the images and are easier to mount on platforms and to process the data. Active sensors, such as LiDAR, require more complicated processing, are a heavy equipment (improved since the start of this PhD), more expensive and do not provide spectral information. Passive sensors are the most used for plant phenotyping.

When a camera system is used, one has to choose the number of wavebands to include from three to dozens and the minimum necessary spatial resolution. Even if RGB sensors (400-700 nm, visual range, three broad bands: red, green and blue) are more limited in spectral information than multispectral and hyperspectral sensors, they provide high spatial resolution data. Currently, none of the single multispectral (400-1000 nm) sensors available achieves the spatial resolution of RGB sensors. Easier data handling and lower price are additional advantages of RGB sensors. The high number of studies published in which RGB sensors are used to generate RGB-based products (orthophotos or digital elevation models (DEMs)) (Yao et al., 2019), demonstrates that RGB cameras play a key role in non-invasive field phenotyping (Guo et al., 2015). Nevertheless, the extra spectral information provided by multispectral sensors, in the red-edge and near-infrared region, is more powerful to detect changes in the spectral reflectance related to abiotic and biotic stress detection and other physiological traits. RGB only "sees" visual damages/symptoms and therefore the spectral sensitivity to the chlorophyll level of the vegetation is limited for more sophisticated analyses. In recent years, modified RGB sensors have been developed that are able to capture a portion of the visible (two of the three bands) and infrared region (by removing a filter that blocks the infrared region) to achieve higher spectral resolution imagery. However, extensive calibration is needed to extract consistent spectral data as the images are subject to distortions, vignetting (reduction of an image's brightness or saturation toward the periphery compared to the image center), and chromatic aberrations (Berra et al., 2017; Holman et al., 2019). Software such as Lightroom (Adobe Systems Incorporated, USA) could correct for vignetting. Hyperspectral sensors cover a wider spectral range (from 400-2500 nm) and allow 'to see' in a higher number of narrow spectral bands than RGB but are not free of limitations either. Their high spectral resolution typically comes at the expense of a lower spatial resolution, data redundancy and collinearity between bands. In addition, the high data volumes generated are associated to higher photogrammetry processing costs and difficulties to geo-reference and stitch the images.

Thermal sensors allow to monitor the canopy temperature through the temperature dependent emission of long wavelength infrared radiation (7.5-13 μm) by the canopy. This canopy temperature can be linked to plant water status and is therefore suitable for phenotyping in the context of drought stress tolerance. If water supply is lower than the quantity required by the plant, stomatal closure can occur (due to the stress). This leads to a reduction of the transpiration rate, which is the plant cooling system, and to a reduction of photosynthesis and growth (Sankaran et al., 2015b). However, the processing and

interpretation of thermal data is not simple, as canopy temperature depends not only on the level of stomatal closure, but also on canopy structure/density and might be affected by rapid changes in environmental conditions in the field (e.g. wind speed or radiation) (Maes and Steppe, 2012; White et al., 2012). Moreover, these infrared cameras typically have lower capture rates (speed at which the camera can collect images, which limits fly speed and area cover) and lower spatial resolution than RGB cameras. Therefore, thermal images obtained with this kind of sensors contain fewer texture/details and the photogrammetry process is harder due to the lack of common points detected over images (Yao et al., 2019).

Table 1-2. Overview of sensors and their potential applications, advantages and disadvantages. Adapted from: Sankaran et al. (2015b); Yang et al. (2017) and Yao et al. (2019).

SENSOR Spectral bands Wavelength range Sensor pixels	Potential applications	Advantages	Disadvantages
Passive sensors			
Visual or RGB camera 3 bands (red, green and blue) 400-700 nm (VIS) > 20 Mpix	Plant height, lodging, canopy cover, intercepted radiation, LAI, 3D structure, leaf color, leaf angle, visual analysis, biomass.	- High resolution (sub-centimeter up to millimeter) - Easy to operate and to install on the platform - Well-modeled camera geometry with a large number of software solutions - Simple data processing - Low cost and weight	- Lack of radiometric/geometric calibration - Low radiometric resolution (lack of spectral information)
Multispectral camera 5-10 bands 400-1000 nm (VIS, NIR) ~ 1.2 Mpix	See above + leaf nitrogen content, yield, chlorophyll content, weed emergence, health status	- Wider spectrum range and narrower bandwidth - Often come with means of radiometric calibration - Imagery from most of the sensors can be well-processed for geometric reconstruction - Allow sub-decimeter to centimeter mapping	- Data format compatibility (sometimes 12 or 16-bit) for software packages - Higher cost than RGB - Discontinuous spectra - Lower spatial resolution than RGB
Hyperspectral camera > 30 bands 400-2500 nm (VIS, NIR, SWIR) ~ variable depending on how data is capture (line or area)	See above + net photosynthesis, nitrogen, chlorophyll content, disease detection, biophysical, physiological or biochemical parameters	- Higher spectral information (more bands) than RGB and multispectral - 10 nm-level bandwidth	- High cost - Complex data processing (require specialized/dedicated software and the users need to take care of the data format and geometric corrections) - Dimension reduction is needed for typical classification tasks - Sensor compatibility to drones limited - Lower spatial resolution than RGB - More sensitive to weather/light conditions - No 3D models

Thermal sensor 1 band 7.5-13 µm (LWIR) ~ 0.3 Mpix	Canopy temperature, stomatal conductance, water potential	- Surface temperature measurement - Indirect determination of crop growth status under abiotic and biotic stress - Relatively easy to be processed	- Lack of texture information brings difficulties in 3D reconstruction tasks - For direct temperature measurement, it needs careful and frequent calibration - Cost is relatively high - Difficult to eliminate the influence of soil - Sensitive to weather (wind) - lower resolution than RGB
Active sensors			
LiDAR NIR 1000-1550 nm -	Plant height, biomass	- Rich point cloud information - Acquisition of high precision 3D canopy structure - Direct geometric measurement - Data under thin canopies (multiple returns of the signals)	- High equipment cost - Highly dependent on expensive onboard GPS/IMU measurement - Increased payload - May not work in GPS-denied regions - Complex data processing
Synthetic Aperture Radar - 10^6-10^8 nm (microwave-radio range)	Crop identification, crop acreage monitoring, key crop trait estimation, yield prediction	Collects data even in cloudy weather	- High cost - Complex data processing - Mainly limited to satellites (for large plots)
Ground Penetrating Radar Ultrawideband 0-1000 MHz	Detection of root bulk root biomass or large diameter tree roots	High throughput	- Cannot detect fine roots - Limited ability to detect genotypic differences

1.3.2 Platforms

To survey (large) breeding or experimental fields in a remote way, different systems/platforms are available (Figure 1-7). Below, we discuss their advantages and constraints, and to which extent they are suitable for HTFP.

In general, three types of **ground-based systems** can be distinguished: handheld sensors, mobile and (semi-)fixed or stationary systems (Araus et al., 2018). Handheld instruments or phenopoles (sensors mounted on a pole or tripod) can be used by one person. This option is often chosen as a cheap alternative to the more expensive platforms or to be able to measure closer to the canopy to overcome a low sensor resolution and still enable high resolution imaging at ground level (e.g. hyperspectral and thermal sensors). They allow to measure e.g. canopy cover, greenness, % plant emergence or counting the number of emerging plants, flowers or ears but at a relatively low throughput. Because the measurements are time consuming, they are less suited for large field trials. For thermal measurements, this is not the best option if many plots need to be monitored as the time needed per measurement can result in noisy data. The main reason is that the thermal response strongly depends on the environmental conditions, and these can change quickly over short periods of time. Therefore, poles and tripods are not considered adequate for HTFP (Reynolds et al., 2019). Nonetheless, they can be very valuable to validate measurements performed with a UAV or airplane (ground truth reference).

Phenomobiles or proximal sensing carts (i.e. carts manually or automatically moving across the trial) can be used to generate large amounts of data with high spatial and spectral resolution as multiple sensors can be mounted for simultaneous measurements. They are often used in combination with hyperspectral cameras as the resolution of the sensors is limited, and because they can be covered to create stable illumination settings or equipped with calibration panels if external light is used (Crain et al., 2016; Deery et al., 2014; Van De Vijver et al., 2020). Limitations of phenomobiles are the number of plots that can be screened simultaneously, the total number of plots that can be screened over a given period of time and soil conditions (i.e. mud) (Zhang and Kovacs, 2012). In addition, the plots could be damaged and soil compacted by the platform itself (Reynolds et al., 2020). As a consequence, data could be difficult to handle and combine, and data collection can be influenced by environmental short-term changes (illumination or temperature), the limitations already mentioned for phenopoles. In addition, the data is often not georeferenced which may make it difficult to link measurements of the same plant/plot at different moments if no measures are taken to label the imagery/data properly. However, there are also examples in which a GPS receiver was incorporated (Thompson et al., 2018).

Fixed and semi-fixed platforms can provide high spatial and temporal resolution with different types of sensors, but they are limited to one specific field or area (e.g. Kirchgessner et al., 2017; Virlet et al., 2017). A cable suspended platform or a camera bay mounted on an industrial portal crane system are moved around to monitor all plots in a dedicated research area. The main advantages are that these systems allow to position the sensor close to the plants in an automated way without causing any disturbance. In this way, high resolution imagery (millimeter resolution) can be obtained at a high temporal resolution. Disadvantages are the need for extensive automation to control the entire system, the investment cost, the fact that equal data acquisition is not ensured (different viewing angles for

border plots may happen), limited area covered, and the fixed position of the system to a specific field (limited in design and number of plots) (Kirchgessner et al., 2017). Especially when the soil is to be kept in good condition, crop rotation is needed, reducing the area that can be used for experimentation. Cameras mounted on fixed poles are used to frequently phenotype a trial resulting in very high temporal resolution (e.g. Aasen et al., 2020). Extra disadvantages are the need to implement complicated mathematical corrections to extract relevant data from side view images to avoid the interaction of the viewing angle in terms of spatial resolution, light environment and changing position of the regions of interest, as canopies develop (Aasen et al., 2020).

A second group of platforms are **air-based platforms** such as airplanes and satellites. The higher the platform is positioned, the larger the area that can be screened, but the lower the spatial resolution and revisiting time. In addition, the presence of clouds might limit the use of platforms imaging from high altitude. Satellites are the main remote sensing platform to monitor natural vegetation and agricultural fields, but for breeding purposes and for biological research they present a limitation that is difficult to overcome: a low spatial resolution of 0.7 up to 30 m depending on the spatial program and spectral bands (e.g. Pleiades 0.70 m, Sentinel-2 10-20 m (https://earth.esa.int/web/eoportal/satellite-missions/) or Landsat 8-30 m (https://www.usgs.gov/land-resources/nli/landsat)). However, as they are able to oversee large areas frequently, they are ideal systems to monitor crops in terms of precision crop farming.

A versatile platform that allows to combine sensors, to register high resolution imagery, is flexible to monitor several trials regularly and screen a large number of plots simultaneously is an Unmanned aerial vehicle (UAV), remotely piloted aircraft system (RPAS) or drone. Two types can be distinguished: copters and fixed wings. The first are more flexible in terms of flying closer to the ground. The minimum altitude at which these systems can fly depends on the downdraft of airstream generated by the rotors. Small copters, e.g. DJI Mavic or Phantom (DJI Sciences and Technologies Ltd., China), can fly as low as 3-5 m without any interference with the canopy, but can only use fixed sensor systems which entails limitations in spatial resolution and spectral range. They are able to stay in the air for approx. 30 minutes. Heavier copters are more flexible in payload and can carry multiple, heavier sensors at a time mounted on a gimbal, but cause more downdraft (the closest flight altitude is 18-20 m without affecting the canopy through turbulence). Typically, their flight can last approx. 20-30 minutes. Fixed wings on the other hand can fly at a higher altitude (up to 100-120 m depending on restrictions by local flight laws and regulations) and at a higher speed. Consequently, they can cover a larger area but at lower spatial resolution. They are also less flexible in terms of sensors that can be mounted (e.g. eBee (Parrot Group, France)). Fixed wings are used rather for precision crop farming than for HTFP. One of the main difficulties associated to the use of UAVs (either copters or fixed wings) is that they are dependent on weather conditions: the use is hampered by strong winds (≤ 4 Beaufort), rain and Kp-index (describes the disturbance of the Earth's magnetic field caused by the solar wind). Detailed information related to UAVs can be found in Colomina and Molina (2014), Pádua et al. (2017), Pajares (2015), Salamí et al. (2014), Sankaran et al. (2015b), Shi et al. (2016) and Yang et al. (2017) and a review of the current status of UAV regulations in Stöcker et al. (2017).

The choice of platform therefore depends on the phenotyping requirements, and it usually implies a trade-off between spatial resolution, precision, throughput and cost. According to Mir et al. (2019),

UAVs are possibly the best suited platforms for HTFP in the context of breeding and biological research as they allow imaging of a large number of plots within a short timeframe, which reduces the risk of changes in illumination conditions, with different sensors and at a high spatial resolution (up to mm). Their flexibility allows frequent revisits (high temporal resolution) or planning of flights at critical moments in crop development for the evaluation of the trait of interest and to investigate the biological response of plant/crops to environmental clues.

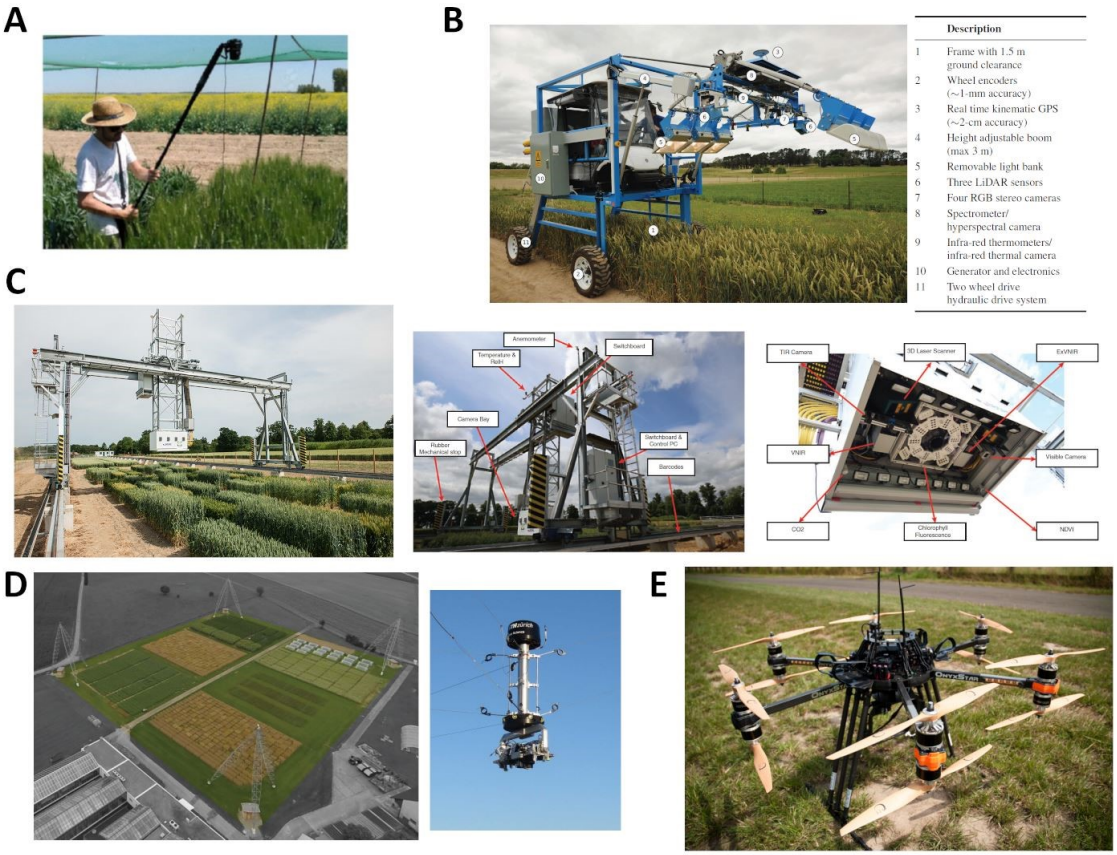


Figure 1-7. Examples of field phenotyping platforms. (A) Phenopole, (B) phenomobile, (C) industrial portal crane fixed platform with elements detailed; (D) cable system fixed platform and (E)UAV. Sources: Deery et al., (2014), Fernandez-Gallego et al. (2018a), Kirchgessner et al. (2017) and Virlet et al. (2017).

1.3.3 Data gathering and derived products from UAV

For the measurement of some traits, such as seedling emergence, single flights at specific developmental stages are needed while for others, such as canopy growth, a series of flights at different moments during the season is required. The trait(s) evaluated will therefore determine the flight frequency (one time evaluation or a multi-temporal dataset) and the timing in the growing season. Once these aspects have been clarified, a flight calendar can be planned for the trial and other technical aspects can be considered.

A schematic overview of the overall process of gathering images using a drone and calculating derived products is shown in Figure 1-8. This process consist of the following steps: (i) making a flight plan that defines the route for the drone while overlapping images are taken, (ii) the flight over the area of

interest is carried out (based on the flight plan) and imagery is captured, (iii) photogrammetry software is used to build a 3D model based on all images, after calibration and correction of the images, and using ground control points for correct positioning of the model, and (iv) based on the 3D model two products can be constructed i.e. an orthomosaic or orthophoto (one single image of the area flown), and a digital elevation model (DEM) or height model (image that contains altitude data of the area). The latter may not be possible if the resolution of the sensor used it too low (e.g. thermal sensor).

Mission planning comprises decisions on the sensor to use, the flight altitude, the timing, etc. and the development of a flight plan. These aspects determine to a large extent the final quality of the products obtained. Detailed information related to the flight planning can be found in Roth et al. (2018b). The following aspects should be considered:

I. The flight area. Width and length of the trial plus borders should be considered to ensure full coverage of the trial. For a good DEM calculation, a wider area needs to be considered than for the calculation of an orthomosaic as each location of the trial needs coverage with the same number of images to obtain the same precision at all locations in the final product.
II. The front and side overlap (common area between consecutive and collateral images). This determines the number of images and the capture rate. It is recommended to achieve at least 75% frontal overlap and a minimum of 60% side overlap (Pix4D support, https://support.pix4d.com/) or even more depending on the specifications of the camera, to ensure the building of the 3D model. To create one single image enough common points between images are required to place them next to each other.
III. The camera specifications. The focal length (distance between the lens and the image sensor) and sensor size (area that determines how much light is used to create the image) will determine the angle of view or area captured. The shorter the focal length, the wider the angle and the greater area captured and vice versa.
IV. The flight speed and flight altitude. These factors influence the quality (sharpness) and the spatial resolution of the orthophoto and the DEM. The higher the altitude, the lower the spatial resolution (larger pixel size).

To be able to perform a spectral inspection and calibration of the images, color neutral references (calibration panels) should be placed in the field, e.g. grey cards or Spectralon (Labsphere, Inc., USA) with known reflectance percentages for white balance, color and spatial distortion effects (Berra et al., 2017). There is no standard procedure to position them, but it is recommended to place at least one spectral reference in the middle of the field or at the beginning and end of the flight plan. Once the set of images has been obtained, spectral inspection and calibration should be performed prior to the stitching of the images. This involves verifying that the white balance is stable, thus that no or limited changes in the light condition occurred during the flight, using the spectral references. Improved processing of the calibrated images enables more consistent and accurate results (Murchie et al., 2018). The calibration process includes the following steps: a relative radiometric calibration (to uniform output across the pixels and time based in dark signal), a spectral calibration (to determine the spectral response of the bands) and an absolute radiometric calibration (to transform from pixel values to radiance based on coefficients derived from an integrating sphere or cross-calibration with an already radiometrically calibrated device). The most common procedure to generate reflectance maps from

spectral UAV data used is the sensor calibration which provides sensor-specific calibration factors to apply to the captured data. In some cases, researchers implement their own calibration procedures if the calibration files or suitable calibration procedures are not provided by the sensor manufacturer (Aasen et al., 2018). Multispectral sensors such as mini-MCA-6 (Tetracam, Inc., USA) or RedEdge-MX (MicaSense, Inc., USA) are equipped with their own calibration software (e.g. Pixel Wrench) or light sensor to calibrate the imagery. The transformation to reflectance could be carried out i) based on incident irradiance, estimated using atmospheric radiative transfer models or measured with an irradiance spectrometer, or ii) using reference targets with a known spectral reflectance to calculate reflectance factors and apply an empirical line method (Aasen et al., 2018). Software's such as ENVI (Research Systems Inc., USA) are often used to carry out this procedure. In the case of RGB imagery simpler procedures could be used such as histogram equalization (it distributes the pixels intensity over the full intensity range) to enhance contrast and white balance correction based on known reflectance of the reference panels. More information related to calibration process can be found in Aasen et al. (2018).

Another important element to decide upon is the number and location of ground control points (GCPs), which are used in the photogrammetric processing to georeference the imagery (e.g. to locate the images precisely into a coordinate system). Their number and placement determines their presence frequency in the imagery and will differ depending on the purpose of the flight (Roth et al., 2018b). Mesas-Carrascosa et al. (2015) and Torres-Sánchez et al., (2018b) tested the influence of different forward overlap values, flight altitudes, flight mode (stop or cruising mode) and number of GCPs on the DEM and orthophoto building process. They recommend (i) to maximize spatial resolution and spectral discrimination to cover a broader area than area of interest, thus a flight planning with 10% of security margin, (ii) to place the GCPs in that margin to have a lower error range in the study area than in the adjacent areas close to the field perimeter, and (iii) to fly in cruising mode and with 95% forward overlap and 60% side overlap.

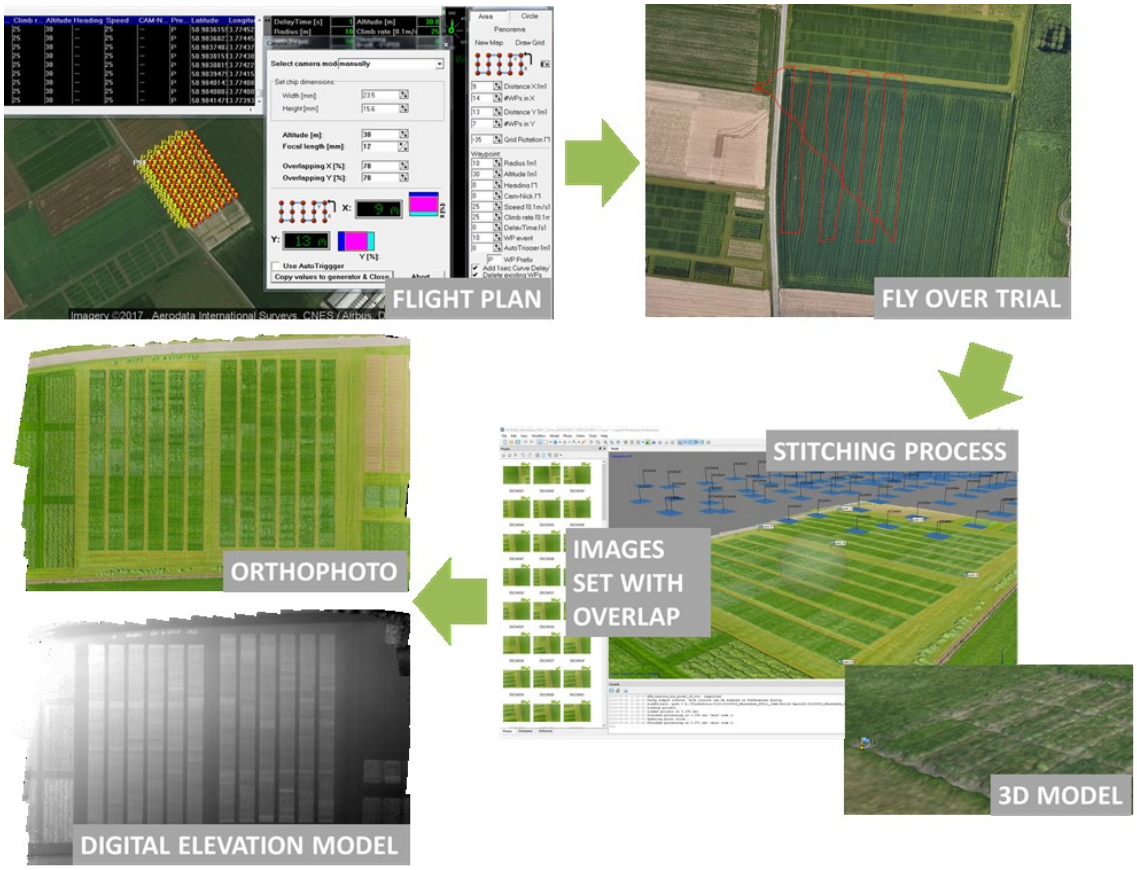


Figure 1-8. Schematic overview of the steps to capture imagery using an UAV

Once the images have been collected, inspected and corrected, photogrammetry software is used to stitch, i.e. to connect and place the images correctly next to each other to form one single image. There are a multitude of options for this, but two broadly used software packages are Metashape (Agisoft LLC, Russia) and Pix4Dmapper (Pix4D, Switzerland). The quality of the reconstructed scene depends on factors such as the geometry and physical properties of the objects, illumination (diffuse and direct light, sun position), weather conditions and sensor and software settings (Probst et al., 2018), being vegetation objects one of the most challenging cases (Gatziolis et al., 2015). Stitching errors in the form of spatial discontinuities could happen such as misalignment of rows, crop structure not correctly built within plots, discontinuities in color or sharpness. To avoid these errors, attention must be paid to the flight planning.

The photogrammetry process (or bundle adjustment) consists of two parts. First, a sparse model is created (a low-density cloud where key scene features are identified in multiple images) and camera/image positions are calculated based on the scene features. Second, based on the sparse model, the point cloud (dense model) is generated. The Structure from Motion (SfM) algorithms/techniques behind these two steps are not freely accessible and detailed documentation is not available, as it is part of the IP of the software developers. Nevertheless, the principles and theory of this technique can be consulted in James and Robson (2012), Smith et al. (2016) and Ullman (1979). The photogrammetry process is fully automated and the only information that needs to be provided to the software (apart

from the images) are the coordinates of the GCPs (i.e. their geographical position). The software automatically estimates camera positions, aligns the images by location and matches the images using common points to establish the relative positions of the sensor during the flight. The output is a three-dimensional georeferenced reconstruction in the form of a point cloud (dense 3D model, Figure 1-8) that represents the area flown and the targeted objects (plants, plots, etc.). Based on the 3D model, the orthophoto and the DEM of the experiment trial are generated. The presented procedure refers in particular to RGB images, but most aspects are transferable to other sensors types (multispectral, hyperspectral and thermal) with some specific adaptations. The main differences are related to the calibration process, including for example different reference panels for thermal sensors.

1.3.4 Data extraction

Regions of interest or polygons with the size of an individual plant or plot (or any other object) need to be defined and correctly located (georeferenced) to extract the data corresponding to each object present in the trial. The creation and labelling of the polygons and the data extraction can be performed manually or automatically using a GIS software (e.g. QGIS (open source) or ArcGIS Pro) or using scripts in R or Python. Data can be derived from the orthophoto and DEM in the form of e.g. vegetation indices or canopy height models at different scales (individual plant, plot or field).

Vegetation indices (VIs) are combinations of spectral information of two or more wavebands to enhance sensitivity to a specific trait, e.g. ground cover or stress response. Many VIs have been proposed to quantify pigment content, structural properties and physiological properties at leaf, plant and canopy level. An overview of frequently used VIs can be found in Ihuoma and Madramootoo (2017); Kefauver et al. (2017); Salamí et al. (2014); Xue and Su (2017) and Yang et al. (2017). As VIs are easy to apply (low computational cost and easy to extract), they are widely used in breeding, precision agriculture and environmental monitoring. However, they also have limitations. For example, the most commonly used NDVI (Normalized difference vegetation index) is limited to a leaf area index (LAI, projected area of leaves over a unit ground surface) < 3 (low canopy density), above which it saturates (Royo et al., 2003). Thus, it can only be used in early stages of canopy development or for vegetation under abiotic stress in which growth is reduced (Reynolds et al., 2020). Moreover, VIs mainly contain information on the top of the canopy and are less suitable to assess very dense canopies.

Changes in illumination conditions or shadows can be removed or minimized in images and sensitivity to variations in greenness enhanced by transformations from the RGB to another color space such as HSI (hue-saturation-intensity) or CIELab (Casadesús and Villegas, 2014; Philipp and Rath, 2002). As an example, hue corresponds to the color tint; saturation is an indicator of how much the pure color is diluted with white color (S = 0% (white) to S = 100% (complete saturation)); and intensity is an achromatic measurement of the reflected light. For CIELab, dimension L^* represents lightness or luminosity whereas a^* represents the red–green spectrum of chromaticity, and b^* represents the yellow–blue color spectrum (Gnädinger and Schmidhalter, 2017; Gracia-Romero et al., 2019; Pointer, 2009).

Thanks to the high resolution that imaging sensors can attain, when overlap between images is foreseen in the flight scheme, the SfM approach allows to build DEMs with a resolution up to 1-2 cm. This allows

to estimate plant/canopy height via a canopy height model when the digital terrain model (DTM) is subtracted from the DEM (Bendig et al., 2015). Usually, the DTM is built based on a flight before or right after sowing, by interpolation of ground pixels (i.e. pixels classified as ground based on spectral reflectance) or by interpolation of GPS points. Both spectral (VIs) and spatial (height) data can be combined to estimate biovolume, which can be used as an indicator of biomass (Bendig et al., 2015), and used to estimate yield.

1.4 State of the art in remote sensing based high-throughput field phenotyping

The majority of studies on field plant phenotyping published to date were carried out in cereals such as maize, rice, wheat or barley. Mainly plant height, biomass and grain yield have been studied using different techniques and sensors.

At early stages in development, single plant counting, germination rate or plant density reflect the final field emergence and are related to the potential yield. Their estimation is possible, for instance, using RGB imagery captured with a phenomobile and implementing artificial neural networks (interconnected neurons characterized by a transfer function). For example, Liu et al. (2017b) obtained an average relative error of 12% in wheat. In maize, Gnädinger and Schmidhalter (2017) obtained an average error ≤5% between in-situ and image-based plant counting applying an enhancement procedure to RGB imagery taken using an UAV. Emergence has also been determined in wheat using UAV and multispectral imaging (Sankaran et al., 2015a). In this study VIs displayed a high correlation with ground-truth emergence data (r^2 = 0.86).

Several other studies have developed methods to estimate plant height. For example, Madec et al. (2017) estimated plant height in wheat using RGB imagery captured with an UAV and compared it with LiDAR obtaining a correlation of r^2 = 0.98 and a RMSE = 8.4cm between them. SfM techniques showed a slight underestimation in both, but the data obtained with the two methodologies were consistent. Bai et al. (2016) used a phenomobile equipped with five sensing modules (ultrasonic distance sensor, infrared thermal radiometer, NDVI sensor, portable spectrometer and RGB camera) to estimate the following crop canopy traits in soybean and wheat: height, temperature, NDVI and green pixel fraction. Canopy height measurements based on UAV imagery have been used to estimate biomass in for example barley (Bendig et al., 2015, 2014; Näsi et al., 2018), maize (Geipel et al., 2014; Li et al., 2016; Niu et al., 2019) and winter wheat (Yue et al., 2017). The accuracy of the estimations was rather high in these studies with r^2 values between 0.71 and 0.82. Tilly et al. (2014) used a Terrestrial Laser Sensor to measure plant height and estimate biomass in rice obtaining $r^2 > 0.86$. Based on plant height data derived from oblique RGB images captured from a pole, Brocks and Bareth (2018) estimated fresh and dry biomass in barley using exponential and simple linear regressions with r^2 = 0.79 for the best model. Tattaris et al. (2016) compared UAV data with ground-based and satellite data in terms of canopy temperature and NDVI to estimate spring wheat biomass obtaining the best result (r^2 = 0.79) with UAV imagery.

Using RGB, multispectral and thermal imagery captured with a ground-based platform and an UAV, Gracia-Romero et al. (2019) estimated grain yield in durum wheat based on traits such as crop vigor or green biomass derived from VIs and compared both platforms. The performance of the step-wise linear

regression models built with ground-based and UAV data to estimate grain yield was similar (r^2 = 0.63). UAV was, however, superior for the screening of large trials and to carry out thermal measurements. While the studies discussed above derived biomass and grain yield based on height data or vegetation indices, other studies have focused on the direct counting of the ears. For example, grain yield was successfully estimated in wheat using ground-based RGB imagery captured from a pole in combination with automatic ear-counting algorithms, correctly detecting >90% of the ears (Fernandez-Gallego et al., 2018b). More recently, they also developed a system for automatic wheat ear counting using RGB images obtained from a UAV (Fernandez-Gallego et al., 2020). The methodology is based on the use of machine learning and accuracies higher than 90% were demonstrated in this case, with the advantage of covering a larger area than using ground-based imagery.

Multi-temporal data has also been used to estimate growth rate or to determine the developmental stages in wheat (Du and Noguchi, 2017; Holman et al., 2016), sorghum (Pugh et al., 2018) and maize (Han et al., 2019; Pugh et al., 2018). Another use of multi-temporal data is to fit growth curves (Chang et al., 2017; Pugh et al., 2018). These examples demonstrate the application, value and potential of multi-temporal data to describe growth and derive specific traits.

Remote sensing has also been used to estimate the response of plants to abiotic stress. For example, Zarco-Tejada et al. (2013), proposed a new formulation of the Photochemical Reflectance Index (PRI) as an indicator for drought stress. In this case, UAV multispectral imagery was used. Data derived from multispectral imagery were compared to stomatal conductance, leaf water potential and CWSI (Crop Water Stress Index, derived from thermal sensor), and r^2 values in the 0.75-0.79 range were obtained. Chapman et al. (2014) demonstrated that canopy temperature can be used for the selection of resistance to water stress. Nitrogen deficiency affects photosynthesis and growth. Therefore, it is relevant to estimate nitrogen content, e.g. using the red-edge inflection point derived from UAV multispectral imagery (best result of the regression analysis r^2 = 0.89) (Geipel et al., 2016). To tackle disease detection hyperspectral data have been used in combination with advanced spectral classifiers such as support vector machines in potato (Franceschini et al., 2019; Van De Vijver et al., 2020) and barley (Thomas et al., 2018).

Besides the breeding programs for the agricultural crops previously mentioned, there are also programs for woody crops. Here, the objectives are to obtain more suitable cultivars for mechanization and use in intensive orchards, to improve quality production and increase tolerance to biotic and abiotic stresses. As mentioned above, the progress in such breeding programs is also hampered by the lack of efficient phenotyping methods (de Castro et al., 2019). Moreover, it is hampered by the time required to establish the crop (some years are needed until fruit production starts and even more to reach its peak production). The ability to quantify traits as early as possible will result in time, labor and cost savings (Rallo et al., 2018). The selection of desirable genotypes is usually based on traits such as tree vigor, vegetative growth habits, tree height and crown diameter/area and volume. HTFP could be used in the form of 3D models and has been applied to assess growth traits rapidly and accurately in vineyards (de Castro et al., 2018; Matese et al., 2017, 2015), olive (Díaz-Varela et al., 2015; Jiménez-Brenes et al., 2017), almond (López-Granados et al., 2019; Torres-Sánchez et al., 2018a) and poplar (Peña et al., 2018) trees.

More detailed information on high-throughput phenotyping and related studies can be found in the recent reviews by Jin et al. (2020), Yang et al. (2017) and Yang et al. (2020).

1.5 Species studied in this dissertation

Two contrasting model crops were used for methodology development and testing in this dissertation. They were chosen because of their high economic and environmental relevance, and because by the start of this research they had been underrepresented in studies regarding remote sensing and development of high-throughput methodologies. Furthermore, they represent two different groups of agricultural crops with different reproduction systems, what implies a different level of diversity in the selection units used for evaluation breeding (lower level of genetic diversity in soybean than in perennial ryegrass) and also very different breeding purposes (perennial ryegrass is a forage crop while soybean is a grain legume). In what follows the main characteristics of the two crops are described.

1.5.1 Perennial ryegrass (*Lolium perenne* L.)

1.5.1.1 Main use and breeding objectives

From an agricultural point of view, grasslands are an important source of feed for ruminants. Moreover, they also provide important ecosystem services, including erosion control, water management and water purification, support biodiversity and cultural services and are an important stock of carbon (Lugato et al., 2015; O'Mara, 2012; Wilkins and Humphreys, 2003). *Lolium* species account for approx. 23% of the area of grassland in Europe and are also dominant in New Zealand and in temperate regions such as Japan and Australia, with perennial ryegrass taking up the largest share (Humphreys et al., 2010).

In North Western Europe, perennial ryegrass (*Lolium perenne* L., family: Poeaceae) is one of the dominant forage grasses thanks to its rapid establishment and growth, high yield potential and excellent feed quality. The species is perennial and productive over long growing seasons (from March to November in Western Europe) and can be either mown (mainly for silage making) or grazed. The large amount of genetic variation and range of heading dates available provide many adaptive phenotypes and genetic resources (Humphreys et al., 2010).

Perennial ryegrass can be cultivated in pure swards or in combination with other species. Frequent companion species are legumes such as white and red clover (*Trifolium repens* and *T. pretense*) and Lucerne (*Medicago sativa*) (Huyghe et al., 2014). These swards are mown several times during each growing season depending on the cultivation environment from 3 to 5 times, being the spring cut the most important one as it contributes to more than half of the annual yield (Sampoux et al., 2011). Ideally, a good total dry matter yield (ton DMY y^{-1}) should be combined with a good distribution over the growing season.

One of the factors that complicate the breeding process in perennial ryegrass is its obligate outcrossing nature. This means that a genotype cannot be self-fertilized, which prevents the development of homozygous lines by repeated self-fertilization. Therefore, perennial ryegrass varieties are populations of related individuals, each of them carrying a different combination of genes. This has clear

consequences for the development of phenotyping methods, as the level of phenotypic diversity that is expected within an 'element' in a ryegrass breeding program (usually a family created by crossing two genotypes or a population created by intercrossing several genotypes) is much higher than what is expected in an 'element' of the breeding program of a self-pollinating species such as soybean (see below). This is due to the fact that each individual plant in a ryegrass family or population carries a different genotype that, in interaction with the environment and the management, can result in a different phenotype. Thus, the level of intra-plot heterogeneity is expected to be higher in field plots (and farmers' fields) of perennial ryegrass than in those of self-pollinating crops (Figure 1-9B).



Figure 1-9. (A) Botanical description of *Lolium perenne*; (B) Close-up of a breeding trial; (C) Overview of a breeding trial and (D) Top-view of a plot with canopy and after cut. Source: Flora von Deutschland, Österreich und der Schweiz 1885; ILVO

The breeding goals in perennial ryegrass are related to biomass yield and nutritional quality. Regarding biomass, given the perennial character of the species, a high level of persistency over years is required, in combination with a high dry matter yield (DMY) and a good spread of the DMY over the whole growing season (March-October). High digestibility, high content of water-soluble carbohydrates with a good

balance of fibers, high level of proteins and lipids, tolerance to abiotic stress conditions and resistance to diseases and pests are also important targets. Breeding perennial ryegrass is a lengthy process, and a genetic gain per decade of 4-5% for dry matter yield was demonstrated over the period 1963-2007 in North-Western Europe (Chaves et al., 2009; Wilkins and Humphreys, 2003), while for other crops the average gain was 8-12% (Li et al., 2018). For example, Wilkins and Humphreys (2003) were successful in improving dry matter yield and water-soluble carbohydrate content. However, they pointed out that the main limitation was the number of plots that could be evaluated at each generation. It is therefore anticipated that implementation of advanced phenotyping tools based on remote sensing imagery, which allow to monitor a high number of plots, will speed up the breeding process of *L. perenne*, allowing higher genetic gains.

1.5.1.2 Traits considered in this dissertation

Swards containing perennial ryegrass are exploited for multiple years (from two to five years depending on the country) (Sampoux et al., 2011) and reseeding is expensive and affects production (lower production during the establishment year). A high level of persistency is therefore an important breeding goal. Persistency can be defined as the capacity to survive and produce optimal yields over several growing seasons (Herrmann et al., 2008; Taylor and Quesenberry, 1996) and it is more critical in the second and subsequent years after establishment to avoid yield losses (Wilkins and Humphreys, 2003). Persistency depends on innate perenniality (also related to the capacity to regrow after mowing and grazing) and on the ability to cope with various abiotic and biotic stresses, which can vary in different cultivation regions. Abiotic stresses can be freezing temperatures, low light intensities in winter, heat or drought, while biotic stresses are pests and diseases (for example rusts) (Humphreys et al., 2010). In current breeding programs, persistency is scored visually by experienced operators. While the visual evaluation of persistency is quick, it could be biased by the observer (Luscier et al., 2006). Observer bias associated to visual estimations have been estimated from 13 to 90% for cover and density (Hatton et al., 1986; Horst et al., 1984; Kennedy and Addison, 1987). By the start of this study a handful of studies had reported on the use of digital images for quantifying canopy cover (Liu et al., 2007; Luscier et al., 2006; Richardson et al., 2001), but neither of these had estimated persistency.

As described above, dry matter yield is also a very relevant selection criterion in perennial ryegrass. Currently, this trait is monitored mainly using destructive samplings that follow the mowing schema that is used by farmers (e.g. in North-Western Europe 4-5 cuts per year). In some cases, plant/canopy height measured using a ruler or a rising plate meter is used to provide a non-destructive estimation of DMY. These approaches present constraints. For example, destructive measurements can only be taken at specific moments and do not allow to analyze temporal changes between cuts. The rising plate meter provides poor estimations of DMY in sparse swards or non-uniform canopies (Viljanen et al., 2018), is relatively reliable only for a sward with a height of 20-25 cm, and needs calibration with destructive samples. Regarding imaging different methods have been proposed to estimate DMY (Holshof et al., 2015; Näsi et al., 2018; Possoch et al., 2016; Rueda-Ayala et al., 2019; Viljanen et al., 2018), but none of them considered perennial ryegrass. In addition, a limited number of samples were considered in these studies, what might compromise the applicability of the methodology developed to other trials and situations.

It is therefore clear that HTPF methods to screen persistency and biomass accumulation would be very useful items in the toolbox of the perennial ryegrass breeder. By the start of this research this option had not been explored yet. Recently, the gap is also being filled by other research groups. For example, Bareth et al. (2019), who studied the potential of UAV-based imagery for forage crop monitoring. Grüner et al. (2019) estimated biomass in two heterogeneous grasslands. Interestingly, Grüner et al. (2019) concluded that the transfer of the models between grassland types is difficult because of differences in morphological characteristics of the component species, and that the prediction accuracies obtained for grasslands were lower than those obtained for arable crops due to their canopy variability. Gebremedhin et al. (2019a) combined NDVI and plant height data to estimate biomass of individual perennial ryegrass plants at early stages of the breeding program. Finally, Lussem et al. (2019) estimated biomass for an entire growing season (4 cuts) using height and vegetation indices. Further background and motivation will be given in the Chapters 2 and 3.

1.5.2 Soybean (*Glycine max* (L.) Merr.)

1.5.2.1 Main use and breeding objectives

Soybean (*Glycine max* L., family: Fabaceae) is an annual legume crop cultivated for the grain. It is used for the production of food products and is one of the main protein sources for animal feed (Sugiyama et al., 2015). It has about 40% protein and 20% oil on weight basis in its seeds (Singh and Hymowitz, 1999). Soybean was originally domesticated in China and then introduced in the US and Brazil (López-López et al., 2010). Masuda and Goldsmith (2009) analyzed the soybean production from 1961 to 2007 and concluded that the cultivated soybean field area and production in that period increased by 36%. Currently, 87.6% of the total world soybean production (>378MT, FAOSTAT 2018, http://www.fao.org/faostat/en/) is located on the American continent, followed by Asia (8.1%) and Europe (3.2%). Europe is one of the largest consumers of soybeans. While the "local" production is starting to develop, it is far from sufficient and massive import remains necessary. Therefore, the interest in soybean cultivation in Europe is increasing. Local production of protein crops such as soybean could help Europe to tackle the following environmental challenges: soil quality, nitrogen management, agro-biodiversity and greenhouse gas emissions (Schreuder and De Visser, 2014).

Soybean is grown under warm conditions in the tropics, subtropics and temperate regions. It is relatively resistant to low and high temperatures, but growth is limited under 18°C and above 35°C. Soybean is photoperiode sensitive for flowering i.e. day length determines flowering and plant maturity. Because it is a short-day species, the flowering occurs when the nights begin to lengthen. Based on maturity, soybean is classified in 13 maturity groups from earliest (000) to latest (X) (Boerma and Specht, 2004), with the early maturing types being less daylength sensitive than late maturing ones. Moreover, the response to daylength is also influenced by temperature. The total growing period length of a soybean crop is 100 to 130 days or more (FAO, http://www.fao.org/home/en/).

Soybean is a self-pollinating crop, and soybean cultivars are pure lines. New varieties are usually developed from bi-parental crosses. After a successful cross and development of heterozygous plants, the breeding material should be returned to the near homozygous state. This usually achieved with the single-seed descent method (i.e. rapid fixing of genes) according to Cahoon et al. (2009), Cober et al.

(2009) and Miladinović et al. (2015). In a representative soybean breeding program, the common procedure is to evaluate the progeny of a selected number of individual plants in progeny row trials. In these trials the small number of seeds available limits the experiment to un-replicated one-row plots at one location (Sun et al., 2015). Consequently, yield measurements are not very accurate at this stage. Hence, there is an opportunity for close remote sensing tools to be used to determine traits correlated to yield at early stages in the selection process in a non-destructive way, improving data quality and reducing time and cost for selection (Moreira et al., 2019).

The breeding goals in soybean are related to seed yield, seed quality and composition, maturity, plant height and lodging resistance, adaptability, disease resistance and stress tolerance (Burton, 1997). Specific goals depend on the target environment, as it is indispensable to adapt the variety to the area where it will be produced. Because soybean genotypes are adapted to rather narrow latitude ranges, the breeding of genotypes with different maturation periods (i.e. maturity groups) allows optimal production in a wide range of latitudes.

As mentioned above, there is an increasing interest in the cultivation of soybean in Europe, and more specifically in North-West Europe including Belgium. Considering the fact that this is a completely new cultivation region for soybean, the selection of varieties with a high production potential in our climate is crucial. In the case of Belgium, only cultivars of the earliest maturity groups (MG000) are suitable, but these cultivars have a low yield potential (Aper et al., 2016). The optimal planting time is end of April – beginning of May to minimize the risk of spring frost and the optimal harvesting is around the end of September – beginning of October. Later on in the season, maturation and thus harvest become more difficult and seed quality is reduced. Therefore, selecting varieties that are well-adapted to our region (i.e. high yielding and producing grains of good quality) involves optimization of multiple aspects related to the growth and the development of the crop throughout the whole growing season (Figure 1-10A).

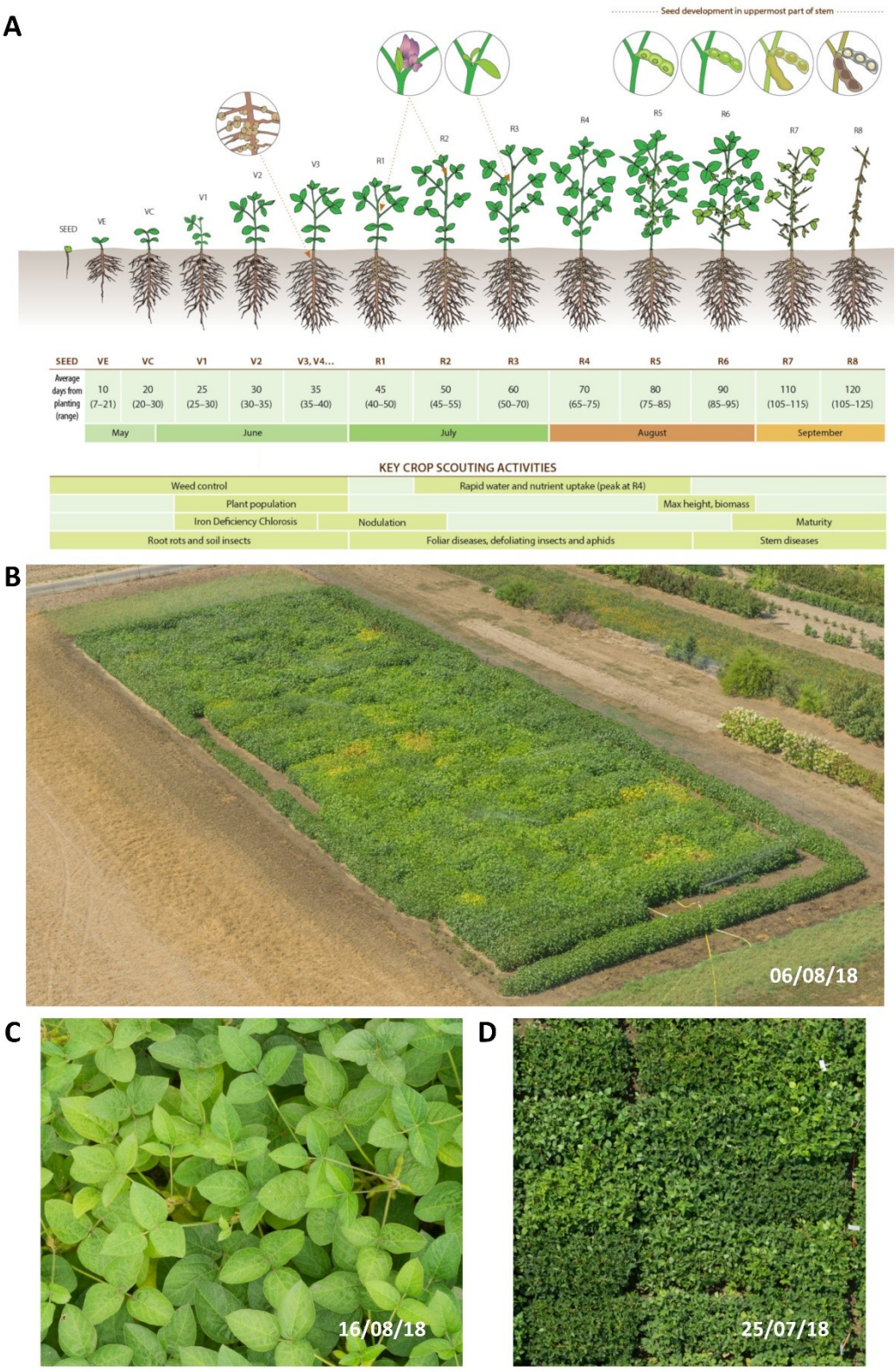


Figure 1-10. (A) Soybean developmental stages; (B) Overview of a breeding trial; (C) Top-view canopy and (D) Close-up of a breeding trial. Source part A: Manitoba Pulse & Soybean Growers (https://www.manitobapulse.ca/).

1.5.2.2 Traits considered in this dissertation

Rapid establishment and full canopy coverage during vegetative and early reproductive stages improve solar radiation interception, which is directly linked to crop growth and yield (Edwards et al., 2005; Edwards and Purcell, 2005), reinforce weed suppression and decrease soil evaporation (Jannink et al., 2001). Plant height is a trait related to yield potential and lodging resistance (Herrero-Huerta et al., 2020; Tilly et al., 2015). Maturity information (95% of the pods with mature brown color), collected by frequent visual assessments, is used to exclude lines with inadequate maturities (too long or too short) for the adaptation to the target growing environment (Gbikpi and Crookson, 1981; Lindsey et al., 2020).

Some examples where multi-temporal datasets are used to estimate the main traits in soybean breeding are presented. Using RGB UAV imagery, Moreira et al. (2019) measured canopy cover to predict yield. Hoyos-Villegas et al. (2014) monitored canopy cover via ground-based imagery to estimate yield . De la Casa et al. (2018) calculated the potential yield based on canopy cover estimations derived from UAV imagery captured with a modified digital camera. Vegetation indices derived from hyperspectral UAV data have also been used for yield prediction (Zhang et al., 2019). Herrero-Huerta et al. (2020) tested the capability of machine learning for grain yield prediction using geometric (e.g. canopy cover, height or canopy volume) and radiometric parameters (e.g. median of the near infrared values) estimated from RGB and multispectral UAV imagery. In another study, three indices (canopy temperature depression, transpiration and canopy greenness) derived from a handheld infrared thermal imager were used to predict yield variation under different water stress levels (Hou et al., 2019). Multispectral UAV imagery has also been used to develop prediction models for yield and maturity (Yu et al., 2016). Estimation of the maturity date using multispectral UAV imagery was investigated by Zhou et al. (2019). Relative maturity and seed yield were estimated with canopy reflectance data measured with a spectroradiometer (ground-based approach) (Christenson et al., 2016). Height, volume and vegetation indices from RGB UAV imagery were used for aboveground biomass estimation (Maimaitijiang et al., 2019). Based on ground-based NDVI values, senescence was estimated to compare soybean genotypes (Craft et al., 2019; Lindsey et al., 2020).

These examples demonstrate the application, value and potential of imaging data to derive specific traits in soybean. However, they only investigated a few time points in crop development and none of them investigated crop development throughout an entire growing season. Indeed, up to now, little attention has been paid to approaches based on the use of multi-temporal data to describe growth and crop performance from sowing to harvest (e.g. by fitting growth curves) and that allow to extract biologically meaningful parameters such as growth rate, senescence or seed yield. More details with respect to this is provided in the Chapter 4.

1.6 Research objectives and dissertation outline

The general objective of this PhD dissertation is to contribute to the advancement of plant phenotyping using remote sensing approaches (UAV-HTFP) in crops for which at the start of this work these tools had been underutilized. In order to broadly explore the potential of this relative new technology, we investigate two species with contrasting characteristics from botanical and agronomic point of view, and bred for different purposes: whereas perennial ryegrass (*Lolium perenne* L.) is an obligate

outcrossing, perennial monocot cultivated for the vegetative parts and harvested several times during one growing season, soybean (*Glycine max* (L.) Merr) is an annual, self-pollinating, dicot cultivated for the grains. As discussed above, these characteristics have consequences for the kind of traits that are relevant to investigate but also regarding the level of intra-plot phenotypic diversity that can expected. In both species we develop and demonstrate novel, easy to implement methods for crop phenotyping suited for the screening of a large number of plots in the context of breeding and/or biological research for relevant traits. This investigation is valuable since it can contribute to a more efficient breeding process in the two crops studied. Furthermore, as these two crops can be considered 'models' for other forage grass species (e.g. tall fescue – *Festuca arundinacea*) and for grain legumes (e.g. faba bean – *Vicia faba*), the outreach of this work is much broader.

More specifically, we develop and test methods to generate relevant and objective phenotypic information from UAV imagery captured with a high resolution RGB camera. In this way, vegetation indices and canopy height models are generated at very high spatial resolution using fairly low-cost technology. The choice of using a relatively accessible RGB sensor, both in terms of cost and technical skills required for operation and data processing is not arbitrary, as this is probably one of the basic requirements to facilitate the adoption of the methods developed (and the technology in general) by the plant breeding community. This manuscript and research are intended to present the possibilities from a technical point of view related to the drone flight, data extraction and data analysis, but we will not engage in genotype modeling.

We do not only develop methods that provide information about a particular relevant stage in crop development and that can replace traditional methods based on visual scoring, tedious measurements or destructive measurements, but investigate also whether the extraction of extra information can complement existing methodologies – putatively accelerating the selection process and increasing the genetic gain. In some applications, we make use of multi-temporal data to derive traits related to crop growth and development and illustrate the added value of multi-point evaluations, not only for breeding but also for biological research in the context of quantitative genetics or modelling.

To ensure that the methods developed are valid and useful in real-world applications, trials that had been set-up in the context of breeding (perennial ryegrass) or biological research (perennial ryegrass and soybean) are investigated. Furthermore, the investigation is carried out in close collaboration with practical breeders who evaluate the same plots using traditional methods in parallel to illustrate the possible added value of the methodology developed. This is in contrast with several of the research reports published in the past, which in many cases focused on a small number of genotypes or which did not make the link with more traditional evaluation methods.

Three cases are considered:

- Persistency in perennial ryegrass
- Dry-matter-yield in perennial ryegrass
- Growth and development in soybean

More specifically, the following research questions (RQ) are tackled thoroughly in this dissertation:

RQ 1. Can persistency of perennial ryegrass be evaluated using UAV imagery? Can this approach replace or complement the traditional way of scoring?

RQ 2. Which is the level of precision that can be attained when estimating canopy height (and its within plot variability) using UAV imagery and can this value be used as proxy for dry-matter yield in perennial ryegrass?

RQ 3. Can UAV-HTFP in multi-temporal series replace labor-intensive measurements to monitor growth and development in soybean? Do these data allow to predict grain yield?

RQ 4. Can UAV-HTFP contribute complementary phenotypic information regarding traits that are difficult to determine by other means?

1.6.1 Thesis outline

After this introductory chapter reviewing the state of the art, presenting the crops investigated and defining the objectives, the research work will be discussed in three chapters addressing the research questions formulated above. In a final chapter, the main results are discussed and general conclusions and future perspectives are presented (Figure 1-11).

In chapter 2, RQ 1 is tackled by developing an alternative methodology for visual scoring to asses persistency in perennial ryegrass. We present a methodology based on the use of vegetation indices derived from UAV imagery captured with an RGB sensor. The performance of multiple vegetation indices is tested to select one that delivers the right information while being relatively insensitive to environmental conditions. The methodology is then applied to two independent field trials and a comparison is made of results obtained using UAV imagery and visual scores recorded by experienced breeders. This work has been published in *Euphytica*.

In chapter 3, we tackle RQ 2 and RQ 3 by developing a phenotyping methodology based on the use of an RGB camera mounted on an UAV to estimate canopy height and its spatial (between and within the plots) and temporal (flights over different dates) variation. We evaluate the capacity of UAV imagery to estimate canopy height, in comparison to the use of a rising plate meter, and its accuracy. Finally, we develop models that allow to monitor the dry-matter accumulation in perennial ryegrass plots over time in a non-destructive way. This work has been published in *Grass and Forage Science*.

Chapter 4 features a different approach related to multi-temporal data analysis of RGB-derived orthomosaics and DEM models. Here, canopy cover and height data are obtained throughout the growing season by UAV imagery in a soybean field trial. The data linked to the growth and development of the crop are used to fit models and derive phenotypic parameters that are related directly or indirectly to breeding traits. These parameters are also used to estimate important characteristics such as grain yield and maturity. So, RQ 3 and RQ 4 are addressed in this chapter. This work has been published in *Remote Sensing*.

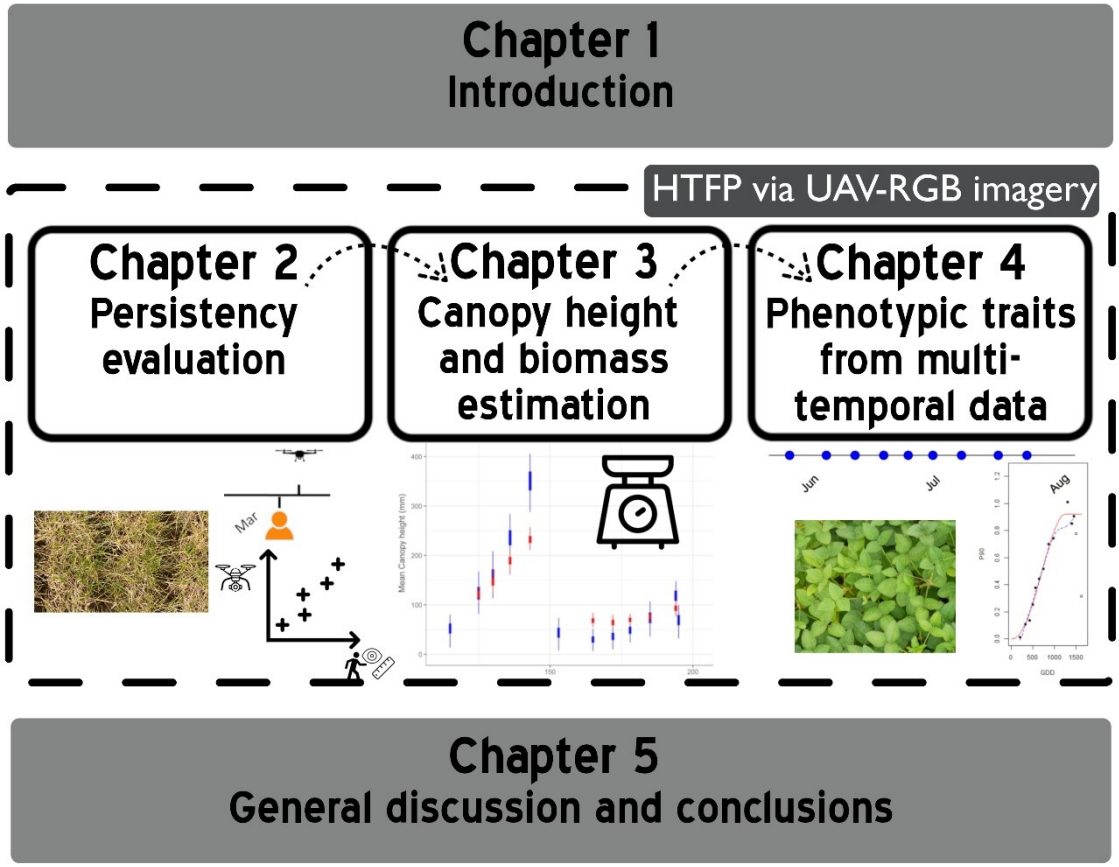


Figure 1-11. Schematic overview of the layout of the dissertation.

Chapter 2: Towards an objective evaluation of persistency of *Lolium perenne* swards using UAV imagery

Adapted from:

Borra-Serrano, I., De Swaef, T., Aper, J. , Ghesquiere, A., Mertens, K., Nuyttens, D., Saeys, W., Somers, B., Vangeyte, J., Roldán-Ruiz, I., Lootens, P. (2018). Towards an objective evaluation of persistency of *Lolium perenne* swards using UAV imagery. *Euphytica* 214: 142.

2.1 Introduction

Perennial ryegrass (*Lolium perenne* L.) is one of the dominant forage grasses in livestock farming in Northwestern Europe in terms of seed sales (Burgon et al., 1998) and land area covered (Wilkins and Humphreys, 2003). Pastures and hay fields containing *L. perenne* are established and subsequently exploited for multiple years. Excellent persistency is desired to avoid reseeding of grasslands, which is expensive and sacrifices grass production during the establishment period, and to avoid substantial nutrient losses through drainage water due to a lower soil coverage (Wilkins and Humphreys, 2003). Therefore, persistency is an important evaluation criterion in perennial ryegrass breeding programs along with forage yield, quality, and disease resistance (Wilkins and Humphreys, 2003; Popay and Hume, 2011).

Although a formal definition of persistency is not straightforward, in a perennial forage crop it can be defined as the capacity to survive and produce optimal yields over several growing seasons (Herrmann et al., 2008; Taylor and Quesenberry, 1996). In perennial ryegrass, persistency predominantly affects the yield in the second and subsequent harvest years (Wilkins and Humphreys, 2003). Plant mortality creates open spaces and allows undesirable species to become established in the sward. This results in losses in productivity and nutritive value (Popay and Hume, 2011). Persistency is a complex trait influenced by intrinsic biological properties determining the degree of perenniality together with biotic and abiotic interactions (Casler and van Santen, 2010; Humphreys et al., 2010). Biotic stresses can include pathogens and pests that can cause loss of vigor and mortality in forage plants. Abiotic stresses such as low temperature, heat, drought and flooding can also affect persistency. In general terms, diploid perennial ryegrass varieties are considered to be more persistent than tetraploids. Diploids are longer-lived and form swards of good density (Humphreys et al., 2010), while tetraploid ryegrass plants produce larger leaves and tillers than diploids, resulting in more open swards. However, persistency displays a high variation within each of these groups. In recent years, tetraploid varieties are catching up with diploid varieties in terms of persistency (Chaves et al., 2009).

In perennial ryegrass breeding programs, enhanced persistence or survivability over the long term is essential. Persistency is usually scored visually through field observations: ground cover is visually evaluated seven days after a cut in the second or third year after sowing, at the end of the production trials, it could be done after the last autumn cut or after the equalization cut after winter time. The breeder rates the vegetation cover with scores ranging from 1 to 5 or 1 to 9 (Luscier et al., 2006), with the highest value assigned to the most persistent swards according to greater soil surface coverage and vigor of the plants. The purpose is to select plant material that demonstrates greater production stability and adaptation to the environment. The persistency score is often used as a negative selection criterion: populations that fall below a certain threshold are discarded. Visual scores are characterized by their low cost in terms of equipment and ease of performance. However, visual ratings have a low repeatability, could be biased by the observer (Luscier et al., 2006), and can be time consuming especially in large breeding trials. Therefore, the availability of a high throughput, objective methodology to estimate persistency in perennial ryegrass swards would offer clear advantages for the breeder, potentially allowing a quicker breeding progress for this trait. Digital imaging can remove the

bias related to the observer and is also a quick and non-invasive method for collecting data on vegetation cover (Liu et al., 2012).

In combination with recent advances in digital photography and image analysis software, remote or proximal digital imaging provides opportunities for improved vegetation data collection and analysis (Luscier et al., 2006; Walter et al., 2012). This is reflected in the multiple studies in which analysis systems have been developed to quantify grass cover based on digital images collected on the ground using tripods (for examples see: Liu et al., 2007; Lootens et al., 2016; Luscier et al., 2006; Lynch et al., 2015; and Richardson et al., 2001). Besides, a large number of studies demonstrate the potential of Unmanned Aerial Vehicles or "drones" (UAVs) for vegetation mapping and monitoring (Araus and Cairns, 2014; Caturegli et al., 2016; Salamí et al., 2014). This remote sensing technology exhibits high spatial and temporal resolution, flexibility in terms of flight planning and image acquisition scheduling, versatility, reduced cost and immediate accessibility of the information to start processing in contrast to other remote sensing platforms such as satellites or airplanes (Xiang and Tian, 2011). Furthermore, while on-ground imagery is limited by scale at the (sub)plot level, UAVs allow simultaneous imaging and characterization of a large number of plots (Araus and Cairns, 2014). On the other hand, UAV images might have a lower spatial resolution (although <0.5cm Ground Sampling Distance can be achieved with appropriate flight settings), and might require more time for processing, which can result in a delay in the decision-making. The UAV procedure consists in a series of steps from mission planning, flight, orthophoto building (which requires most of the time and depend on the computational power) to data extraction in comparison to on-ground imagery capture that requires less preparation and computational power (any orthophoto is built). The main disadvantage of UAVs is their restricted use due to poor flight conditions or regulatory restrictions.

Depending on the application, different approaches for digital image analysis can be followed. For applications in which explicit spatial information on vegetation cover has to be determined (without the necessity to recognize individual plants or plant parts) spectral vegetation indices (VIs) are commonly used (Guijarro et al., 2015; Liu et al., 2007; Meyer and Neto, 2008; Woebbecke et al., 1995). VIs are scalars calculated based on the information from two or more spectral bands, or in other words, calculations using light reflectance at different wavelengths (Rasmussen et al., 2016). VIs constitute a simple and convenient means to extract information from the remotely sensed data contained in spectral reflectance. Their ease of use facilitates the processing and analysis of large amounts of data (Govaerts et al., 1999). Vegetation indices can use bands from different spectral regions. Indices such as NDVI or OSAVI that include the near-infrared region might be more informative but are less human intuitive as human eye is not sensitive to that region of the spectrum (Meyer and Neto, 2008). The indices that involve red, green and blue bands (visual region) emphasize colors such as plant greenness. Authors such as Rasmussen et al. (2016), Meyer and Neto (2008) or Torres-Sánchez et al. (2014) demonstrate different uses for which those indices are suitable related to vegetation cover. Advantages of using RGB cameras are their easier and quicker handling being more intuitive for their use and processing and lower price.

The aim of this work was to develop a reliable methodology based on UAV-derived VIs to estimate persistency of *L. perenne* accessions in a breeding context. We used a field trial to estimate persistency for a set of diploid and tetraploid accessions from the *L. perenne* breeding program at the Flanders

Research Institute for Agriculture, Fisheries and Food (ILVO; Melle, Belgium) according to standard procedures (Humphreys et al., 2010). To this end, the following steps will be taken: 1) VIs choice and data stability study; 2) validation with on-ground imagery; and 3) evaluation of agreement with traditional visual persistency scoring. The developed methodology was further validated by applying it to an independent field trial. Finally, the advantages and limitations of the developed methodology are discussed in relation to the state of the art evaluation methods.

2.2 Materials and methods

2.2.1 Field trials and plant material

For the development of the methodology, we used two experimental fields, one for training (Field Trial 1; FT1) and one for validation (Field Trial 2; FT2), both located in Melle (Belgium). The geographic coordinates (WGS-84 / UTM zone 31N (EPSG:32631)) of the bottom right corner of FT1 were X = 554310 m and Y = 5648345 m and of FT2 were X = 554216 m and Y = 5648109 m. The two fields are divided in diploid and tetraploid block trials, eight (four for each ploidy) in FT1 and seven (four diploid and three tetraploid) trials in FT2. Each trial is a randomized complete block design composed of plots of 1.35 x 5.8 m (7.83 m^2) where each *L. perenne* accession was represented by two or three replicates depending on the block trial. Both fields were managed under a cutting regime of five cuts per year at a mowing height of 5 cm and fertilized each year after sowing. Trial FT1 consisted of 102 diploid and 135 tetraploid accessions (237 in total) and was sown on April 25, 2013. The second independent trial FT2, to investigate the general applicability of the methodology, consisted of 138 diploid and 96 tetraploid accessions (234 in total) and was sown on April 14, 2014. On FT1 and FT2 *L. perenne* accessions were evaluated for persistency in the context of breeding activities.

2.2.2 Data acquisition of FT1

2.2.2.1 UAV-derived orthophotos

On March 18 and March 22 2016 (7 and 11 days after mowing, respectively) a set of nadir overlapping images (60% forward-longitudinal-lap and 60% side-lap) was captured by a sensor mounted on a 12-rotor UAV (model Onyxstar Hydra-12, Altigator, Belgium) at 50 m flight altitude. For imagery, we used a conventional RGB (red-green-blue or visible spectral range) camera (model Sony α6000, Sony Corporation, Tokyo, Japan) equipped with a 35 mm lens (f/1.8, Sony Corporation, Tokyo, Japan). The APS-C sensor was 6000x4000 pixels, resulting in a field of view of 34.3x22.9 m at 50 m flight altitude. The camera was set to operate in manual mode to avoid different settings in successive images. Exposure time (shutter time), F-stop (aperture) and sensitivity to light (ISO) were adjusted and fixed in the field just before the start of each flight. Focus was set to automatic mode. Image trigger was activated by the UAV according to the flight route. The flight route was programmed using the software MikroKopter v2.14b (HiSystems GmbH, Germany) and automated to ensure that the whole experimental field was covered with the necessary overlap to build the orthophoto.

Per date two flights were carried out, one in the morning around 11:30 (local time) and a second in the afternoon around 14:45. Different weather conditions were experienced: the first date was cloudy and the second date was clear sky and sunny, resulting in different light conditions and solar radiation levels during the acquisition of the images. The solar radiation recorded during each of the flights (chronologically) was 175.4 W/m^2, 75.6 W/m^2, 552.7 W/m^2 and 505.5 W/m^2.

These four flights rendered four sets of images. Raw images were used as they enable image corrections without data loss. The raw images were pre-processed per flight in Lightroom 6.5 (Adobe Systems Incorporated, San Jose, CA, USA) to correct for white balance and exposure based on gray cards (15x20 cm Gray card, 18% reference gray, Novoflex Präzisionstechnik GmbH, Germany) used as a color neutral reference. Three gray cards were placed on the ground of the field trial at the start, middle and end of the flight route. The cards were visible in 15% of the images captured and flights were carried out under constant weather conditions (in 7' the image acquisition was completed, duration of the flight was 10' from take-off to landing). The invariant references where used to verify that the white balance was stable and the same correction could be applied to the entire set of images. After corrections were implemented, the raw images were converted to lossless tif files. Agisoft Photoscan Professional Edition (Agisoft LLC, St. Petersburg, Russia) was used for mosaicking and building orthophotos (example provided in Figure 2-1A). To create geo-referenced ortho-mosaicked images, five artificial Ground Control Points (GCPs) were geo-referenced using a RTK GPS (model Stonex S10 GNSS, Stonex SRL, Lissone, Italy) and identified in the images. The four resulting ortho-mosaicked images had a spatial resolution of 0.52 cm defined automatically by the software based on the camera parameters (sensor size and focal length) and flight altitude. The geo-referencing error of the orthophoto ranged from 0.04 to 0.09 m. This reference error represents the distance between the coordinates of the GCP measured by the GPS and the estimated position in the orthophoto (root mean square error). It was used to decide if an additional refinement of the marker location was required or if a GCP should be excluded. The error obtained in this experiment was less than 0.10 m and was thus considered acceptable.

2.2.2.2 Ground-truth assessment

To estimate and compare the performance to assess persistency by different VIs derived from UAV images (see section 2.2.2.3 below), a ground-truth reference dataset was generated for a subset of 30 plots distributed randomly across the trial (Figure 2-1B). In each of these 30 plots on-ground images of 0.30 m x 0.50 m frames, which were geo-referenced with RTK GPS, were captured on March 18, 2016 using a Nikon D90 camera (Nikon Corporation, Tokyo, Japan) mounted on a tripod at a height of 1.2 m in nadir position. The frame was placed in the middle of the plot to avoid border effects and at approximately the same location for all the plots to avoid bias due to positioning of the frame (Figure 2-1C). To avoid and ensure that the shade of the tripod was beside the frame area and not on the image, the tripod was placed opposite to the sun. In Figure 2-1C and D the positioning and an example of a frame-image are shown.

These 30 on-ground images were processed using the supervised classification method 'Spectral Angle Mapper' (SAM) (Kruse et al., 1993) in ENVI/IDL 5.0 (Exelis Inc., Boulder, CO, USA). We defined three classes: 'green grass', 'bare soil' and 'old/dead grass' (not green). Training and validation pixels, called regions of interest (ROIs), were randomly selected for each class in the frame images ensuring that the

set of pixels used for validation was different from that used for training. Training ROIs were used as the basis to conduct the classification. Validation ROIs were used to assess the accuracy of the classification by computation of a confusion matrix and Kappa coefficient. To this end, each validation ROI was located in the classification image and the class assigned was compared to the one given by the operator during the selection of pixels for this analysis. Classification match (or coincidence) was determined through the confusion matrix. Once the procedure had been validated, the class 'green grass' was used to estimate the percentage of area covered by green grass (A-GG) in each frame. In this study, only 'green grass' was considered, as we only used vegetation indices that highlight and differentiate green vegetation from other features (bare soil or old grass).

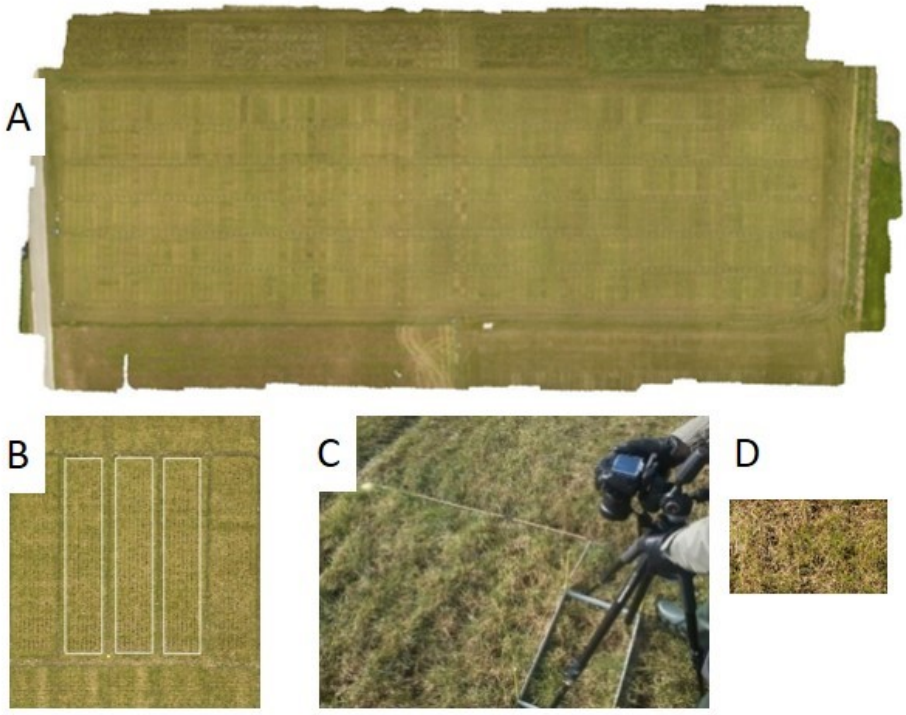


Figure 2-1. (A) Orthophoto of the breeding trial (set images: morning March 18); (B) subset of the 90 plots selected for the second breeder score; (C) on-ground frame imagery capture using tripod; and (D) example of one on-ground frame image.

2.2.2.3 Vegetation indices

A set of 13 vegetation indices (VIs) proposed by different researchers working on RGB imagery were selected as candidates for the persistency evaluation (Table 2-1). For each individual plot of the experiment an average value of each of these indices was determined from UAV orthophotos, excluding the border (15 cm buffer) to avoid influence from pixels that do not belong to the plots.

In addition, we located the 30 georeferenced on-ground frames used as ground-truth on the orthophoto, defined a polygon with the same dimensions and calculated the average value of the different VIs for the selected frame area. The location of the frames and the values of the VIs for plots and frames were determined in QGIS 2.8.7-Wien (QGIS Geographic Information System. Open Source

Geospatial Foundation Project). The VI values and green grass fractions (A-GG) estimated for each frame were used for regression analysis to check whether the VIs were delivering correct information on the green cover fraction.

Table 2-1. List of RGB vegetation indices used in this study. R= Red, G= Green and B= Blue bands.

Index	Abbreviation	Reference	Formula
Non-normalized			
Excess Green 1	ExG1	Woebbecke et al.(1995)	$ExG(1) = 2 * G - R - B$
Normalized			
Color Index of vegetation	CIVE	Kataoka et al.(2003)	$CIVE = {(0{,}441R - 0{,}881G + 0{,}385B + 18{,}78745)}/{(R + G + B)}$
Excess Green 2	ExG2	Woebbecke et al.(1995)	$ExG(2) = {(2 * G - R - B)}/{(R + G + B)}$
Excess Red	ExR	Meyer et al. (1999)	$ExR = {(1{,}4 * R - G)}/{(R + G + B)}$
Green Leaf Index	GLI	Gobron et al.(2000)	$GLI = {(2 * G - R - B)}/{(2 * G + R + B)}$
Normalized Green Intensity	GCC	Woebbecke et al.(1995)	$GCC = {G}/{(R + G + B)}$
Norm. Green-red Difference Index	NGRDI	Gitelson et al.(2002)	$NGRDI = {(G - R)}/{(G + R)}$
Ratio Color Index	RCI		$RCI = {G}/{R}$
Vegetativen	VEG	Hague et al.(2006)	$VEG = {G}/{(R^{0.667} * B^{0.333})}$
Woebbecke Index	WI	Woebbecke et al.(1995)	$WI = {(G - B)}/{(R - G)}$
Combination			
Combination 1	COM1	Guijarro et al.(2011)	$COM(1) = 0.25ExG(2) + 0.3ExR + 0.33CIVE + 0{,}12VEG$
Combination 2	COM2	Guerrero et al.(2012)	$COM(2) = 0.36ExG(2) + 0.47CIVE + 0{,}17VEG$
Excess Green minus Exc. Red	ExGR	Camargo Neto (2004)	$ExGR = ExG(2) - ExR$

2.2.2.4 Visual scoring of persistency

Two experienced breeders visually scored the plots on March 17, 2016 (six days after the equalization cut at the end of the winter before the spring growth starts, and one day before the first UAV flight) using the 1-9 persistency scoring system applied routinely in breeding, with a score of 1 representing low persistency and 9 high persistency. Breeder A scored the diploid plots and Breeder B scored the tetraploid plots, according to standard practice in the ILVO breeding program. In addition, they were requested to score in parallel a subset of 90 plots randomly chosen across the experiment (including both diploid and tetraploids). The scoring was done blind as no labels were present in the field, only a yellow plastic dot on the ground was present to reduce the time searching for the 90 plots.

2.2.2.5 Negative selection procedure

The practical implementation of the developed procedure was tested by comparing it with the current methodology to estimate persistency in perennial ryegrass. In practice, ILVO breeders perform a negative selection by eliminating the 50% accessions with the lowest persistency scores. We simulated

this type of selection using the best performing vegetation index related to persistency per accession and results were compared to accession selection based on breeder scores. The comparison in the negative selection was carried out for diploids and tetraploids separately with the full set of plots per accessions, as is usually done in practical breeding. This allowed us to determine the potential impact of shifting from a system based on visual scoring to a system based on UAV derived images and VIs. For the selection based on breeder scores we followed the standard procedure in the ILVO breeding program. This is based on the processing of the data in Trialwizard (Beerepoot Consultancy BV, The Netherlands), which performs corrections based on 'moving mean' and 'nearest-neighbor' analysis (taking into account the experimental design and the replicates present in the field), providing one single persistency score per accession. For the UAV-derived data, the averages of the VI values estimated per plot were processed in a similar way as the persistency scores, using Trialwizard to obtain a single index value per accession. To establish the limits of the 50% negative selection, the median values were determined for the scores and indices.

2.2.2.6 Data handling and statistical analysis

To facilitate the understanding of the different steps done to develop the methodology, the workflow is schematically illustrated in Figure 2-2, where the data used in each part are indicated.

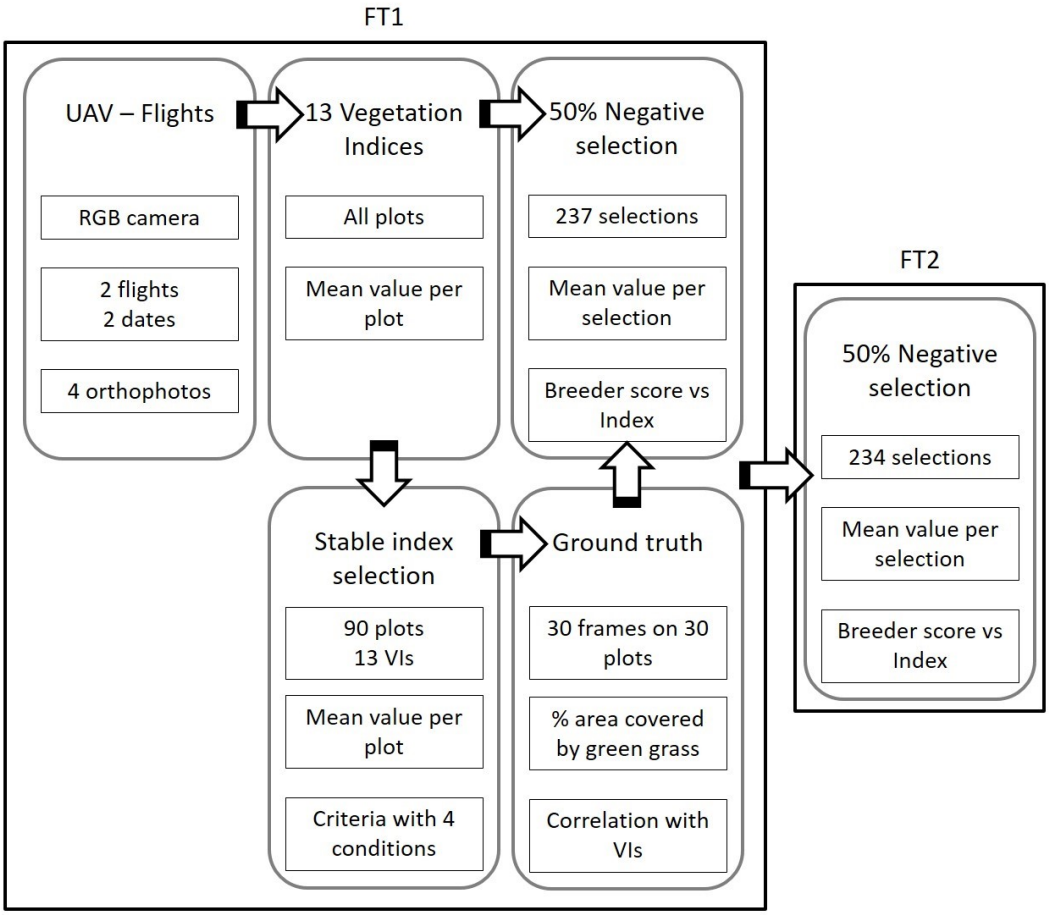


Figure 2-2. Workflow for data collection, processing and analysis, with indication of the data used in each step.

In the first step, we selected VIs that rendered consistent and stable results for different flights (and consequently under different solar radiation conditions) on the same day for the subset of 90 plots (same set described in section 2.2.2.4 above). Consequently, a suitable index has a value as stable as possible minimally influenced by the time of flight or flight conditions. We only considered comparisons of UAV flights within one day (morning and afternoon), as the plants' size may have changed between the two flight dates, which could interfere with the analysis. Secondly, we determined whether the VI values (of the VIs selected in the first step) derived from the UAV flights were representative for the ground-truth values obtained using 30 on-ground images of small frames (0.3 m x 0.5 m) within the plots. The ground-truth is the percentage of area covered by green grass (A-GG) in the on-ground photo frames and the VI values were derived from the frame area in the orthophoto. In the third step, we assessed the effect of the time interval between mowing and UAV flight using all pair-wise combinations of flights for the subset of 90 plots, since UAVs cannot operate under certain weather conditions. This makes that the chosen date is not always viable. In the fourth step, we estimated the consistency of scores assigned to the same plot by each breeder at different times and evaluated the agreement of scores assigned by different breeders for the same subset of 90 plots (n_{2x}=34, n_{4x}=56).

To fulfill the four steps mentioned, Pearson correlation and simple linear regression analyses were carried out between (1) VI's of flights per day, (2) selected VIs and ground-truth, and (3) selected VIs for all flights combinations. To compare the persistency assessment within breeder and between breeders of the subset of plots scored two times, a contingency table was calculated. All the statistical analyses were carried out in RStudio 0.99.879 (RStudio: Integrated Development Environment for R, RStudio Inc., Boston, MA, USA). Additionally, per block trial (eight in total) the coefficient of variation and F-value with significance test were derived from an analysis of variance of the scores and VI calculated using Trialwizard.

As a final step, the percentage of agreement in the negative selection of accessions derived from the VI value and the visual scoring were calculated.

2.2.3 Application of the methodology to FT2

FT2 was used to test the general applicability of the developed methodology. The two abovementioned experienced breeders were responsible for the persistency scoring in this field. The breeder's persistence rating and the UAV flight were performed on October 10th 2016, six days after mowing. A set of nadir images was collected with the same sensor and procedure described in section 2.2.2 at 55 m altitude around 14h00 (80% forward-longitudinal-lap and 60% side-lap). In this case, we only considered the VI selected in previous steps. Negative selection and data analysis was as described in sections 2.2.2.5 and the final part of 2.2.2.6.

2.3 Results

2.3.1 Development of the methodology: FT1

2.3.1.1 Selection of robust UAV-based VIs

Based on the VIs obtained for a subset of 90 plots, the stability of the index values between the two flights carried out on the same day was calculated (Table 2-2). Only three of the investigated VIs fulfilled the pre-defined criteria: ExG2, GCC and GLI. The following criteria were applied for selection of candidate VIs: correlation higher than 0.9 and significant ($p < 0.001$), intercept between -0.1 and 0.1, and slope between 0.9 and 1.1 to avoid offsets. For selection, the VI had to fulfill the criteria for both dates investigated. These VIs showed highly significant correlations between flights carried out on the same day, an intercept close to 0 and a slope around 1 (Table 2-2). These results demonstrate that values obtained under different light conditions for these three UAV-based VIs were highly stable.

Table 2-2. Comparison of the stability of the vegetation indices (VI) over the different flights for the March 18 and the March 22 using a subset of 90 plots (n=90). Pearson correlation (r) and significance level (p), with bold showing $r > 0.9$ and $p < 0.001$. Intercept and slope of a linear regression, with bold values showing an intercept between -0.1 and 0.1, and a slope between 0.9 and 1.1. An * marks the selected vegetation indices, based on the pre-defined criteria .

Index	March 18				March 22				Selection
	r	p	intercept	slope	r	p	intercept	slope	
Non-normalized									
ExG1	**0.924**	**0.001**	3.482	0.817	0.722	**0.001**	16.974	0.580	
Normalized									
CIVE	**0.956**	**0.001**	-0.931	**1.050**	**0.917**	**0.001**	-0.849	**1.045**	
ExG2	**0.957**	**0.001**	**-0.002**	**1.050**	**0.919**	**0.001**	**0.000**	**1.045**	*
ExR	**0.928**	**0.001**	**0.019**	**0.978**	0.859	**0.001**	**-0.007**	**1.040**	
GLI	**0.957**	**0.001**	**-0.001**	**1.047**	**0.920**	**0.001**	**0.000**	**1.040**	*
GCC	**0.957**	**0.001**	**-0.017**	**1.049**	**0.919**	**0.001**	**-0.015**	**1.045**	*
NGRDI	**0.937**	**0.001**	**-0.011**	**0.993**	0.872	**0.001**	**0.002**	**1.041**	
RCI	**0.938**	**0.001**	**0.007**	**0.971**	0.868	**0.001**	**-0.045**	**1.052**	
VEG	**0.945**	**0.001**	**-0.014**	0.776	**0.922**	**0.001**	**-0.061**	**1.063**	
WI	0.892	**0.001**	-0.832	**1.015**	0.600	**0.001**	2.869	0.364	
Combination									
COM1	0.535	**0.001**	5.477	0.143	0.735	**0.001**	-9.966	2.541	
COM2	**0.953**	**0.001**	1.089	0.875	**0.922**	**0.001**	-0.507	**1.056**	
ExGR	**0.950**	**0.001**	**-0.006**	**1.044**	0.898	**0.001**	**0.009**	**1.050**	

2.3.1.2 Correspondence between UAV-based VIs and ground-truth

The next step was to investigate whether the three UAV-based VIs selected in the previous step delivered accurate information regarding green cover. For this comparison, we used pictures of a subset of 30 on-ground frames to generate a ground-truth dataset. Three items were studied to evaluate the

correspondence: 1) frame area in relation to the entire plot area; 2) supervised pixel classification accuracy of on-ground imagery; and 3) correlations between VI and on-ground imagery.

The area imaged in the on-ground image (frame) was considered representative for the state of the corresponding plots, as the VI of the corresponding plot-area extracted from the UAV orthophoto was highly correlated to the VI of the entire plot ($r > 0.845$). Subsequently, the on-ground images of the frames were suitable to relate and validate the use of vegetation indices to estimate persistency. The supervised classification of the pixels in the 30 on-ground frames was performed with a mean overall accuracy of 85% and Kappa coefficient of 0.78. The confusion matrix and the Kappa coefficient were calculated for each frame separately, providing different accuracies for each frame. The classification matrix of an average frame shows that 95.9% of the green grass is classified as green grass (Table 2-3). The classification of old grass and bare soil was less precise, although 78.9 and 88.6% of the bare soil and old grass, respectively, were still correctly classified. Especially bare soil and old grass were less easy to identify separately, but they were confused less frequently with green grass (Table 2-3). As the classification gave a high mean overall accuracy, the estimation of the percentage of area covered by green grass (class 1), bare soil (class 2) and old grass (class 3) was executed. In Figure 2-3A-D, two frames are shown together with their classifications and percentage of area covered by green grass (A-GG) considered as ground-truth. Correlations on the same areas between the VI derived from the frame area and A-GG were calculated for the 13 VIs (results not shown). For the three selected stable VIs (ExG2, GCC, GLI) high correlations were observed between the VI and ground cover ($r > 0.885$). Which supported their use to assess vegetation cover and hence persistency.

Table 2-3. Example of the results of the supervised classification matrix of an on-ground frame.

Classification result	Ground-truth (%)		
	Class 1	Class 2	Class 3
Unclassified	0.00	5.69	0.00
Class 1: 'green grass'	95.85	0.05	4.51
Class 2: 'bare soil'	2.39	78.86	6.92
Class 3: 'old grass'	1.75	15.39	88.57
Overall accuracy: 85.57% Kappa coefficient: 0.779			

2.3.1.3 Choice of one representative UAV-based VI

Multiple linear regression models were built using all or a selection of the 13 VIs. However, this did not lead to improved correlations with A-GG. Therefore, the three UAV-based VIs were considered the most appropriate to fit linear regressions for our purposes. In Figure 2-4, the linear regression is illustrated for the morning flight on the March 18 for ExG2, GLI, GCC and the area of the frame covered by grass (A-GG). Because the scatter plots and regression lines for the three selected UAV-based VIs show high similarities (Figure 2-4) and their values are strongly correlated ($r= 0.981$), one can question if these are actually different or if they capture the same information. Therefore, the relation between ExG2, GCC and GLI was evaluated using a random sample of 1000 RGB values for which the indices were calculated (Figure 2-5). ExG2 and GCC are related linearly and ExG2 and GLI curvilinearly. However, as the range of the values derived from the orthophotos (indicated by the dashed lines in Figure 2-5) was much smaller than in the random sample, all pair-wise relations became approximately linear. Therefore, we can

conclude that the three selected VIs delivered approximately the same information in this experiment. For that reason and for the sake of simplicity, further analyses were carried out using ExG2 exclusively, as this index is the most widely used of the three.

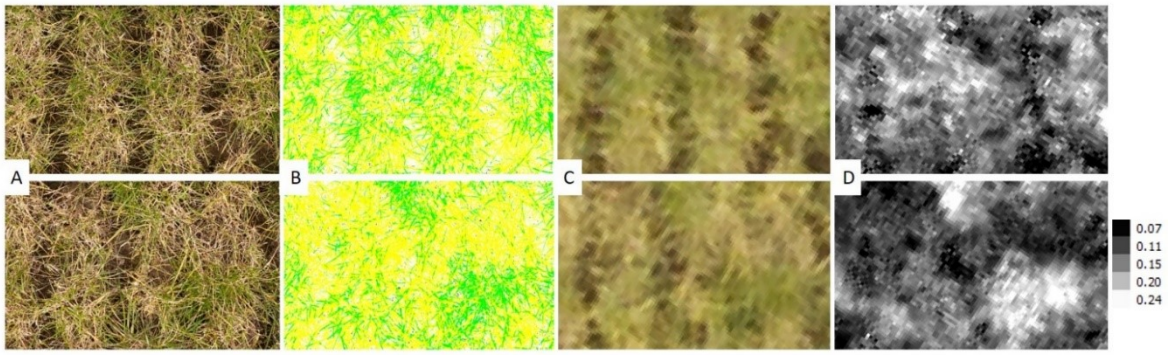


Figure 2-3. (A) Example of two on-ground frame images; (B) Results of the supervised pixel classification of green grass (green), old grass (yellow) and bare soil (white); (C) frame area from the orthophoto; and (D) ExG2 grid calculation.

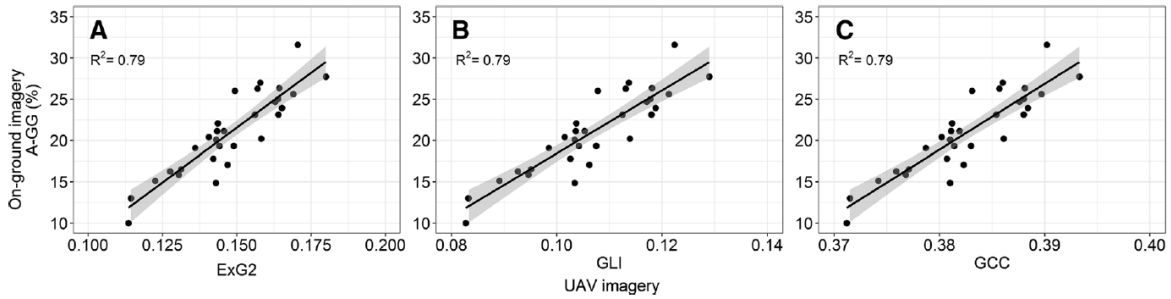


Figure 2-4. Linear regression with 95% confidence interval performed with the index (A) ExG2, (B) GLI and (C) GCC to estimate the area covered by green grass (A-GG).

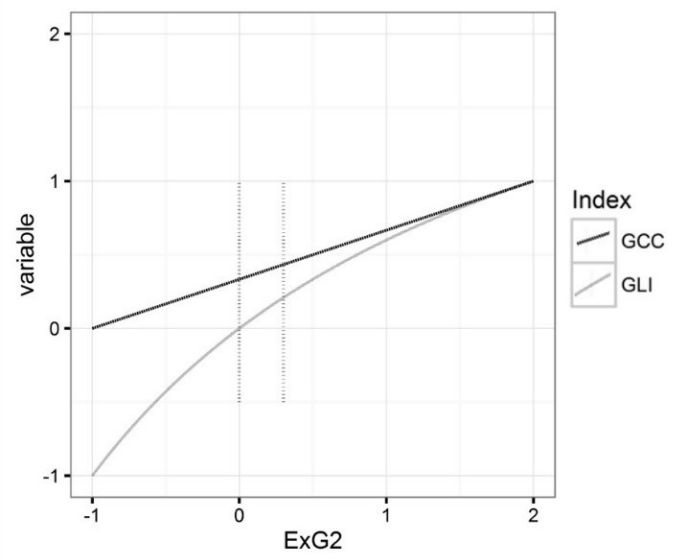


Figure 2-5. Theoretical relation between ExG2, GCC and GLI. Dashed lines indicates the range of ExG2 values observed in this study.

2.3.1.4 Inter-day variability

The next step was to calculate correlations between the four ExG2 datasets (Table 2-4) to estimate if flights performed at a five-day interval deliver similar information. In all cases, the correlations were ≥ 0.871, even though between the dates, the plants had grown and the weather conditions were different. This shows that during this period of the year, flights separated by five days delivered similar results in this experiment, adding flexibility to the implementation of a methodology based on the use of visual imaging from UAVs in case that weather conditions could restrict the use of drones.

Table 2-4. Pearson correlation coefficients between the four flights over the two dates (18: March 18; 22: March 22; 1: first flight; 2: second flight) for Excess green 2 (ExG2) (n=90 plots).

Index	ExG2 18-1	ExG2 18-2	ExG2 22-1
ExG2 18-2	0.957		
ExG2 22-1	0.948	0.930	
ExG2 22-2	0.871	0.876	0.919

2.3.1.5 Comparison of the persistency assessment though UAV-based VIs and breeder scores

The results of the usual procedure of negative selection were compared to the decisions that would have been taken if UAV-based ExG2 values had been used (Figure 2-6). First, based on the additional scores given by the two breeders to the set of 90 plots, we estimated the within-breeder consistency (repeatability of scores assigned at different times by the same breeder) and the agreement of scores assigned by the different breeders. Both breeders assigned consistent scores when they evaluated the same plot two times, with agreements of 61.1% (n=36 diploids plots) and 72.4% (n=54 tetraploids plots) for Breeder A and B, respectively (data not shown). However, the agreement between the scoring by different breeders was only 26.6% which correspond to 25 plots (Table 2-5).

Table 2-5. Contingency table between breeders for scoring persistency of *Lolium perenne* plots. The number of plots (count) and frequency (%) in each score class (3-9) per breeder is presented. Agreement in the plot scores between both breeders are shown in bold.

	Count	Breeder B				
	%	6	7	8	9	Sum
Breeder A	3	0	1 *1.06*	0	0	1 *1.06*
	4	1 *1.06*	3 *3.19*	1 *1.06*	0	5 *5.32*
	5	9 *9.57*	18 *10.64*	10 *10.64*	0	37 *39.36*
	6	**9** ***9.57***	10 *7.45*	7 *7.45*	0	26 *27.66*
	7	1 *1.06*	**11** ***11.70***	4 *4.26*	1 *1.06*	17 *18.09*
	8	0	2 *2.13*	**4** ***4.26***	0	6 *6.38*
	9	0	1 *1.06*	0	**1** ***1.06***	2 *2.13*
	Sum	20 *21.24*	46 *48.94*	26 *27.66*	2 *2.13*	

Furthermore, an analysis of variance for the scores and the index ExG2 is presented in Table 2-6 to study the size of the coefficient of variation (CV) and whether there are significant entry effects (F Varieties). The values of both parameters are shown per ploidy, split by block-design, for the breeder scores and the four flights carried out over FT1. On the whole, the scores had greater CV and lower F-values than the VI.

Table 2-6. Coefficients of variation (CV) and F-values (F varieties) per ploidy and block (trial design) of the visual Breeder score and ExG2 index derived from the four flights carried out over field trial 1 (ExG2 FT1). Statistical significance or not at the 5% level indicated per F-value marked with a plus or minus sign.

				March 18		March 22	
Ploidy	Block	Parameter	Breeder Score	ExG2 FT1	ExG2 FT1	ExG2 FT1	ExG2 FT1
2n	Lp1301	CV	11.3	4.7	4.7	4.6	3.7
		F Varieties	3.3+	7.4+	8.3+	9+	11.4+
	Lp1302	CV	13.5	5.9	5.9	6	4.7
		F Varieties	2.9+	4.4+	4.4+	6.3+	6.9+
	Lp1303	CV	13	7.1	7.7	9	7.6
		F Varieties	4.1+	3.2+	2.9+	2.1+	2.4+
	Lp1304	CV	15.8	7.6	6.4	9.6	7.3
		F Varieties	1.2-	2.2+	2.6+	1.1-	1.3-
4n	Lp1305	CV	11	4.40	3.7	7.4	5.5
		F Varieties	1.4-	3+	5.1+	1.6-	2.1-
	Lp1306	CV	7.5	4.40	4.6	5.8	4.4
		F Varieties	2.1+	3.4+	4.1+	4.4+	5.2+
	Lp1307	CV	7.1	3.90	4.3	7.4	5.9
		F Varieties	4.6+	7.7+	6.9+	3.4+	4.1+
	Lp1308	CV	7.4	5.30	4.7	6.1	5.1
		F Varieties	3.2+	2.5+	2.5+	2.4+	2.8+

Second, we compared the decisions of Breeder A and Breeder B (separately) to ExG2 data from the flight closest to the scoring (morning of March 18) and of the flight that was separated by the longest interval from the scoring (afternoon of March 22). Note that in these comparisons we took the scores of Breeder A for diploid populations and of Breeder B for tetraploid populations. For the selection of the 50% worst populations (negative selection) the scores of Breeder A and the ExG2 values of the flight closest to the scoring displayed a degree of agreement of 74.6% (Figure 2-6A). Correspondingly, for 25.4% of the populations a different decision would have been made using the criteria of Breeder A and the ExG2 values. Corresponding values for Breeder B were lower: 54% for the negative selection (Figure 2-6C). Results obtained taking into consideration the afternoon flight of March 22 were comparable (Figure 2-6B and Figure 2-6D for breeders A and B, respectively). Finally, to complete the analysis, we also compared the consistency of the decisions that would be taken using information from different flights (Figure 2-6 and Table 2-4). As expected, a high degree of agreement was found between flights, with higher values for comparisons involving flights on the same date (91.6%, Figure 2-6E) than between the flights that were furthest apart in time (85.6%, Figure 2-6F).

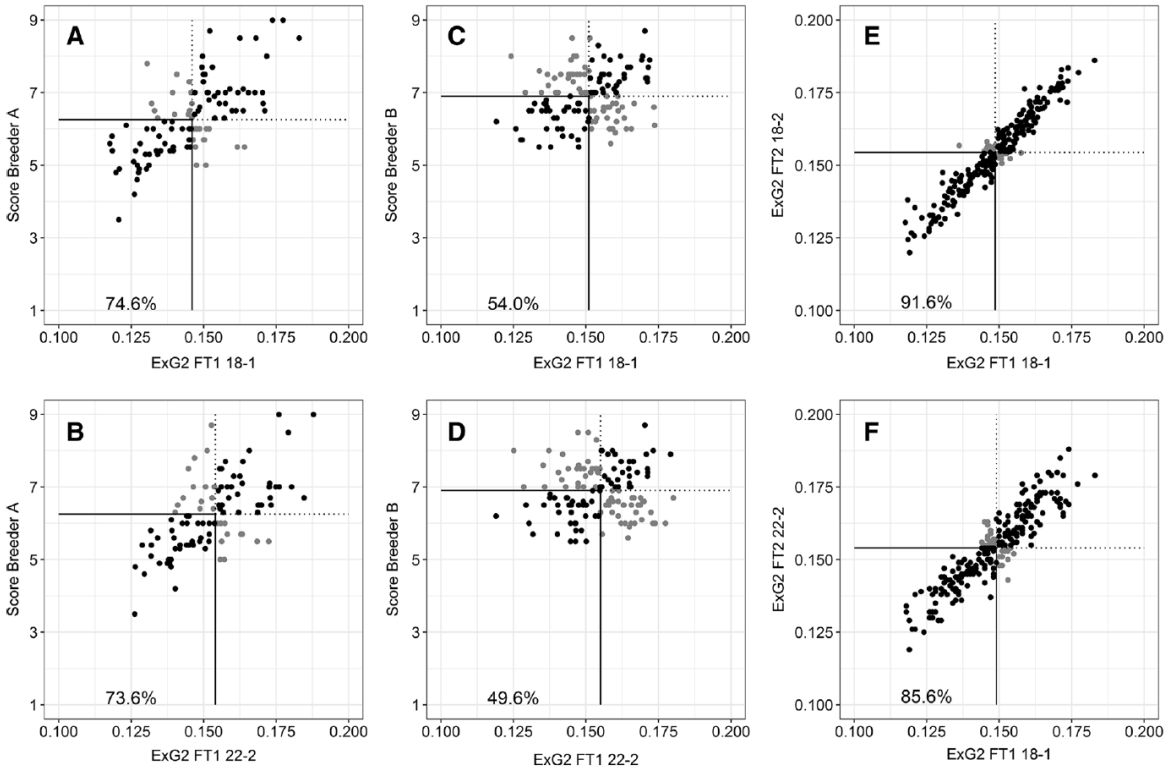


Figure 2-6. Comparison of the negative selection for six cases: (A) Breeder A versus ExG2 for the flight closest to the scoring (n=102 2x accessions); (B) Breeder A versus ExG2 for the last flight (n=102 2x accessions); (C) Breeder B versus ExG2 for the flight closest to the scoring (n=135 4x accessions); (D) Breeder B versus ExG2 for the last flight (n=135 4x accessions); (E) ExG2 for the two flights showing the highest correlation (n=237 accessions); and (F) ExG2 for the two flights showing the lowest correlation (n=237 accessions). The percentages of agreement in the selections are shown, black dots represent the selections that agree and gray dots the disagreement. The quadrant corresponding to the negative selection is indicated by darker lines corresponding to the median of the axis values represented.

2.3.2 Validation on an independent test field

In FT2, the negative selection of the 50% least persistent accessions showed an agreement of 64.5% between the ExG2 values (UAV-flight derived) and the score of Breeder A (Figure 2-7A). The degree of agreement with the scores of Breeder B was 60.4% (Figure 2-7B). These values are similar to those obtained for FT1, except that the agreement value for Breeder A was slightly lower in this case. The range of scores assigned by Breeder A to the accessions investigated in FT2 was lower than for FT1, whereas the opposite was true for Breeder B. This probably reflects a different range of variation in the accessions investigated in FT1 and FT2. The persistency score of an accession is based on two or three repetitions present in the field. The plots corresponding to two accessions which showed disagreement between the visual score and the VI are shown in Figure 2-8. The VIs values are influenced by the green biomass present in the image, as can be seen by comparing plot C which has less green biomass and results in a lower ExG2 value than plot X which has the highest ExG2 value of the plots shown (Figure 2-8).

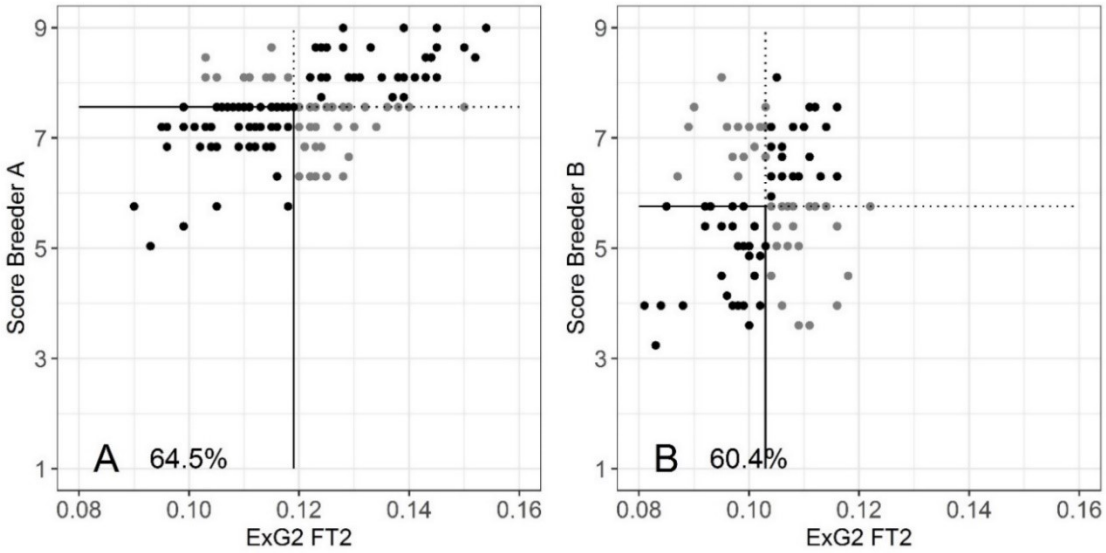
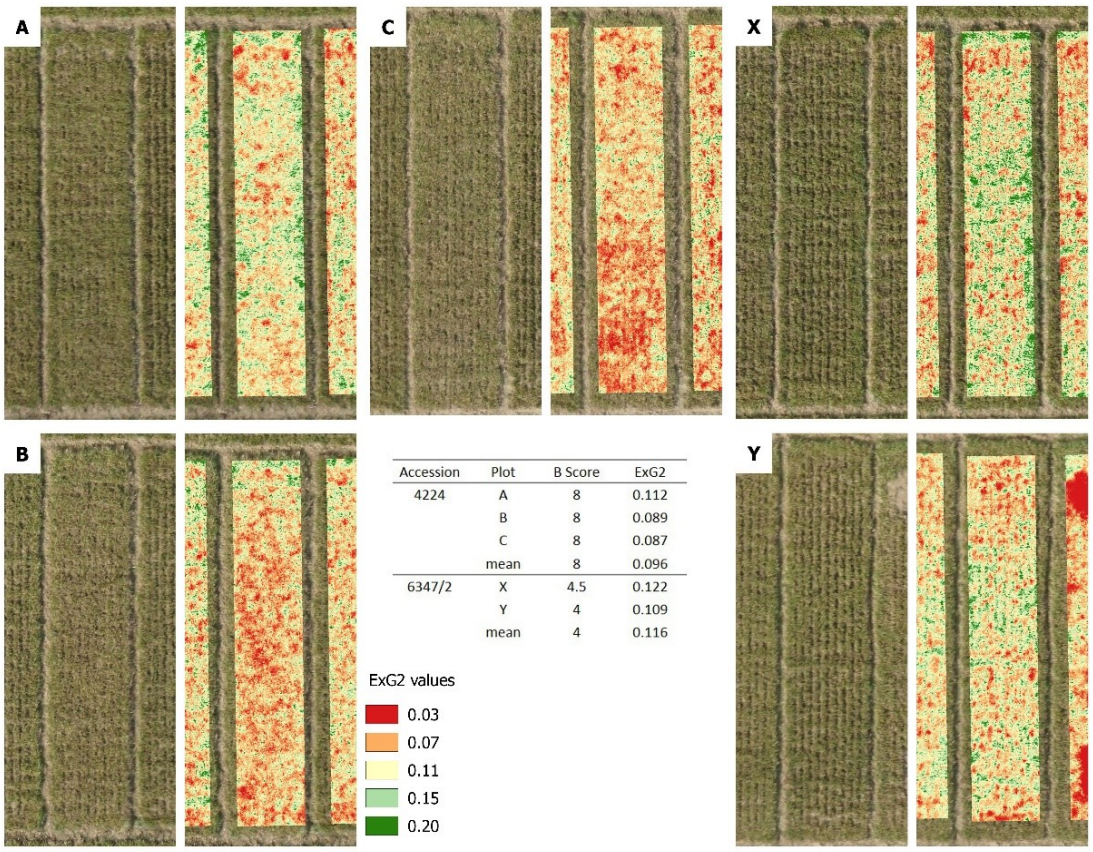


Figure 2-7. Comparison negative selection (50% worst accessions) in validation field trial FT2: (A) Breeder A versus ExG2 (n=138 2x accessions); (B) Breeder B versus ExG2 (n=96 4x accessions). The percentages of agreement in the selection are shown, black dots represent the selections that agree and gray dots the disagreement. The quadrant corresponding to the negative selection is indicated by darker lines.



Accession	Plot	B Score	ExG2
4224	A	8	0.112
	B	8	0.089
	C	8	0.087
	mean	8	0.096
6347/2	X	4.5	0.122
	Y	4	0.109
	mean	4	0.116

Figure 2-8. Orthophoto and derived ExG2 from the flight over field trial 2 (FT2) of two tetraploid accessions, labeled as 4224 (plots A, B and C; n=3) and 6347/2 (plots X and Y; n=2) that showed disagreement between the breeder score (B Score) and the VI (ExG2). The mean value per accession was calculated in Trialwizard taking into account the design of the experiment.

2.4 Discussion

In this study, we propose a methodology based on the use of vegetation indices derived from visual UAV imagery to estimate persistency in diploid and tetraploid perennial ryegrass in the context of breeding activities. The developed methodology has resulted in an accurate method to measure ground cover as a tool to determine the persistency of different accessions.

2.4.1 Three UAV-based VIs deliver robust results under different environmental conditions

It should be clarified that we consider variation in weather conditions between the four flights, but not during the course of a flight. In our case study, the imagery was captured under constant weather conditions, during the seven minutes required to cover the field, that allowed us to use the sensor with fixed settings and pre-processed the entire set of images for one flight at once using the gray cards. Our recommendation is to place two or three references along the flight route to verify whether substantial variation in white balance is present and to decide whether the same pre-processing parameters can be applied to all images or not. In case of a partial cloudy day, changes in the illumination conditions could happen during the flight and result in a set of images that require specific pre-processing per image (or per small group of consecutive images) using multiple invariant targets (e.g. gray cards) located across the entire field to ensure their presence in numerous images.

The use of vegetation indices is widespread (Guijarro et al., 2015; Liu et al., 2007; Meyer and Neto, 2008; and others). As outdoor light conditions can differ greatly, it is important to use indices that are as insensitive as possible to these differences. Atmospheric and lighting conditions change throughout a day, resulting in significant differences between the light intensities at which different images are acquired (Guijarro et al., 2015). In this study, the first criterion to select suitable VIs was therefore their robustness against changes in lighting conditions. Three of the 13 tested VIs tested (EXG2, GLI and GCC) delivered robust data over flights performed at different times on the same day and thus under different light conditions (from 75.6 W/m^2 to 552.7 W/m^2). Not surprisingly, these are all normalized indices; the non-normalized ExG1 investigated did not pass the stability test and was one of the VIs with the lowest performance. This is in agreement with the results reported by Woebbecke et al. (1995) who tested the use of non-normalized and normalized RGB VIs in shaded and unshaded surfaces and concluded that the normalized ones were less variable than the non-normalized ones. The effect of the change in illumination over time on information derived from on-ground digital imagery in a forest scenario was studied by Sonnentag et al. (2012). They concluded that it was not possible to remove the effect entirely on ExG1 and GCC. The effect of spatial and temporal factors on the classification accuracy of vegetation in wheat fields using VIs from UAV imagery was also investigated by Torres-Sánchez et al. (2014), who reported that ExG2 and VEG were the most suitable indices.

Furthermore, the three selected VIs delivered consistent data even when flights were separated by five days and were done in weather conditions with different light intensities. The fact of delivering consistent data even after five days is important, as adverse climatological conditions can prevent flying at pre-defined days after mowing. In this experiment, the plants had resumed growth after mowing, but flights executed within a window of five days still delivered consistent VI results and consistent breeding decisions in terms of populations being selected or eliminated. This finding indicates the practical

applicability of our procedure to evaluate persistency in realistic situations for breeding. It should, however, be noted that although the consistency mentioned here refers to relative values, as the ExG2 values were slightly different for the two dates investigated, this is sufficient for breeders' interests. We cannot determine whether this was due to plant growth or incomplete correction for the light environment despite the application of white balance and exposure correction.

2.4.2 Three UAV-based VIs provide a good estimation of the area covered by green grass

Results presented in section 2.3.1.2 demonstrate that UAV-based EXG2, GLI and GCC values represent reliable estimates of ground cover assessed by on-ground imagery (r^2= 0.791), and hence could be used to estimate persistency. In an attempt to further improve these results, different multiple linear regressions were carried out involving some or all of the indices, but no significant improvement was achieved. Even if other authors (Camargo Neto, 2004; Guijarro et al., 2011) have reported new indices (e.g. ExGR or COM1) that are combinations of existing indices, it should be noted that the indices used in this study were only based on 3 independent variables (R, G, B) and could result in overfitted models. For the three UAV-based VIs, it is not surprising that a combination was not relevant given the linear relationship demonstrated for the range of RGB values considered in this study in pair-wise comparisons. Therefore, a linear regression with any of the three selected indices fit our purposes. As EXG2 is one of the most widely cited and used vegetation indices (Guijarro et al., 2015; Mao et al., 2003; Meyer and Neto, 2008; Sonnentag et al., 2012; Torres-Sánchez et al., 2014) we based further developments and testing of the procedure on this index. Results for GLI or GCC are expected to be similar.

2.4.3 The visual scoring system (UAV-based VIs) consistency compared to the visual scoring by breeders

All pairwise correlations between the four ExG2 datasets were ≥ 0.871 (Table 2-4), meaning that similar and consistent information were delivered by the repeated flights. This creates room for flexibility when implementing the methodology.

In this experiment, two experienced breeders performed the visual scoring of the plots. Both breeders involved in this experiment were trained to assign persistency scores. Their training is reflected in the high 'within-breeder' consistency estimated here (>60% agreement between consecutive scorings by the same breeder). On the other hand, it was found that they probably use slightly different criteria for visual evaluation, as the between-breeder agreement was low (26.6%). When interviewing the breeders about their criteria, clear differences emerged. Breeder B scored the plots using a narrower range of values than Breeder A and took the amount of bare soil present in the plots more into account. This difference may be related to the established thought that tetraploid accessions have worse soil coverage, as Breeder B was responsible for the tetraploid section. This assumption was confirmed by the higher correlation between the percentage of bare soil (determined through a supervised classification) and the visual scores for Breeder B (r= -0.331; p<0.001) than for Breeder A (r = -0.025; non-significant). The two breeders involved in this experiment probably assess persistency using

different criteria and differ in the range of scores (results not shown), which may affect the inter-breeder consistency.

Visual scoring has the advantage of being cheap, quick, and easy to perform, but contains a risk of delivering biased information. Furthermore, even experienced breeders within the same organization do not always agree on a consistent system for assigning scores. Indeed, visual scores are affected by the observer: the assessments require experience and a certain level of skill, but a certain level of subjectivity cannot be avoided (Milberg et al., 2008). Studies to estimate the observer bias affecting visual estimations (Hatton et al., 1986; Horst et al., 1984; Kennedy and Addison, 1987) have shown variances from 13% to 90%.

Taken together, the results indicate that the agreement between decisions taken based on the UAV-based EXG2 data and the visual scoring by a breeder is comparable to the agreement between the scoring done by different breeders. Our results demonstrate that using UAV-based VIs to estimate persistency would allow a higher level of objectivity and flexibility by eliminating the human influences on the evaluation. Moreover, when an analysis of variance is carried out for both type of data, the results for ExG2 showed a smaller coefficient of variation and a greater F-value compared to the breeder scores, which results in an easier differentiation between the accessions.

ExG2 data were compared to scoring data in terms of selection as is done in a real breeding situation. Negative (worst 50%) selection was carried out based on the visual scores and the UAV based vegetation index values. The results of the selection were then compared. The highest levels of agreement between visual scoring and ExG2 data were found for Breeder A with values of 74.6% and 60.9% for FT1 and FT2, respectively. The level of agreement was lower for the visual scorings by Breeder B (54.0% and 58.4%, respectively). We can therefore conclude that the UAV-based approach using ExG2 can reproduce the visual scoring system used by Breeder A for diploid accessions quite well, while it was less successful in reproducing the visual scoring system of Breeder B for tetraploid accessions. The latter might be improved by incorporating information on the percentage bare soil, as Breeder B also included this information in his scoring system.

Regarding the objectivity and reproducibility, even when selections based on the two flights with the lowest correlation were compared, levels of agreement of 85.6% for negative selections was found. Consequently, the UAV-based VI procedure presented here would provide the same end result: the same selection of populations, independent of the flight used. Thus, our results clearly demonstrate that an evaluation system based on UAV-based VIs would significantly increase the reproducibility and objectivity of evaluations of persistency in perennial ryegrass.

In summary, we have developed a procedure using a UAV-based VI that provides useful information on the persistency of perennial ryegrass accessions sown in plots. This system is ready for implementation in practical breeding. In future research, extra variables should be considered to further refine the system, e.g. the amount of bare soil. Moreover, during the breeding experiment, each generation of selection for all-round performance takes at least three years (Wilkins and Humphreys, 2003). Breeders typically score multiple parameters, not only persistency, to develop improved varieties. Consequently, their decisions to select an accession or not is not only based on persistency, but also on other parameters such as yield. Therefore, it could be interesting to perform a follow-up breeding field trial

carrying out flights at critical moments over the three years to support scores involved in the selection, such as persistency or rust, and contribute with extra information to benefit the selection process through monitoring based on UAV imagery.

In conclusion, we have demonstrated that UAV-based vegetation indices can assist breeders to evaluate persistency of *L. perenne* accessions at plot level using a UAV equipped with a visual camera. ExG2 was revealed to be a stable normalized vegetation index, it corresponded very well with the true value of persistency, and it allowed assessment of the populations in a way that was consistent with the breeder evaluations based on a negative selection procedure.

Chapter 3: Canopy height measurements and non-destructive biomass estimation of *Lolium perenne* swards using UAV imagery

Adapted from:

Borra-Serrano, I., De Swaef, T., Muylle, H., Nuyttens, D., Vangeyte, J., Mertens, K., Saeys, W., Somers, B., Roldán-Ruiz, I., Lootens, P. (2019). Canopy height measurements and non-destructive biomass estimation of *Lolium perenne* swards using UAV imagery. *Grass and Forage Science*. 74: 356–369.

3.1 Introduction

Grasslands deliver important economic and ecological services. They provide a substantial fraction of the ruminant feed requirements, prevent soil erosion, build soil fertility, require minimal pesticide use, and are important for carbon sequestration (Lugato et al., 2015; O'Mara, 2012; Wilkins and Humphreys, 2003). In northwest Europe, *Lolium perenne* L. (perennial ryegrass, Lp) is one of the most frequently sown forage grass species because of its rapid establishment and growth, high yield potential and excellent feed quality (Humphreys et al., 2010). These valuable (agro)services must be safeguarded. Breeding goals in perennial ryegrass are therefore to improve the yield of cultivated grasslands as well as their resilience to abiotic and biotic stresses and sustained production of high quality forage over multiple years. In the last decades an average yearly improvement of 0.3-0.4% of dry matter yield (DMY) has been achieved in this crop (Chaves et al., 2009; Laidig et al., 2014; McDonagh et al., 2016) while for other crops an average annual gain of 0.8-1.2% is reached (Li et al., 2018). These values are still far from the estimated necessary annual gain of 2% in crop productivity that are required to achieve the food and bioresource demand for the projected global population in 2050 (Li et al., 2018). Genomics-enabled breeding approaches in combination with precision phenotyping tools are expected to contribute substantially to enhance the prediction accuracy and to accelerate genetic gains by shortening breeding cycles (Crossa et al., 2017). Indeed, the use of precision phenotyping tools to obtain a detailed and accurate characterization of the genetic value of an accession can result in a higher selection accuracy and a higher gain per breeding cycle (Cobb et al., 2013). In a recent review by Gebremedhin et al. (2019b) phenotyping is presented as the bottleneck to implement genomic strategies in forage species. Furthermore, the adoption of precision phenotyping tools by breeders is frequently lagging behind scientific-technological advances in this area due to the lack of versatile and affordable approaches to screen large collections of breeding strains in a high-throughput manner (Araus et al., 2018; Araus and Cairns, 2014; Desta and Ortiz, 2014).

Dry Matter Yield (DMY) is one of the most important traits in forage grass breeding (Barre et al., 2017). Biomass yield data can be obtained by visual assessments, measurements of canopy height (on its own or in combination with estimates of ground cover) or from destructive samplings. Visual assessments are carried out by experienced operators. Height measurements are non-destructive and can be done using rulers or devices such as rising plate meters (RPM), which measure plant height and canopy density. When using an RPM, the grass canopy is (gently) compressed by the plate weight, and the device measures the distance between the plate and the ground. This value is then converted into a value of DMY using a linear relationship (Holshof et al., 2015). Destructive measurements involve cutting, drying and weighing of aboveground samples, and are related to sward management (typically 4-5 times per year depending on management and the characteristics of the growing season). Despite their widespread use, all three approaches present constraints. Visual assessments are relatively cheap but are subjected to subjectivity of the operator. In addition, they can perform poorly when differences among the accessions are limited. Destructive measurements have limited temporal resolution. Using RPM, a high temporal resolution can be achieved but RPM measurements are not accurate in sparse swards or in tall, non-uniform canopies (Viljanen et al., 2018). In addition, with the exception of RPM when several measurements are carried out in the same plot, none of the three approaches are suitable to ascertain intra-plot spatial variability in an accurate manner, rendering a low spatial resolution.

As summarized by Ali et al. (2016), remote sensing from satellites equipped with optical, radio detection and ranging (radar) or light detection and ranging (LiDAR) sensors can be used to estimate temperate grassland crop biomass (Wachendorf et al., 2017). These methods are non-invasive and address some of the issues encountered by the ground-based monitoring approaches described in the previous paragraph, such as the labor intensity, the temporal limitations (and also spatial, depending on the required resolution) and human bias (Ali et al., 2016). However, satellite remote sensing has limitations regarding grassland monitoring and management such as: (i) imagery acquisition and delivery in a timely manner, (ii) low spatial resolution (ranging from 5-100 m depending on the spectral band, sensor and altitude of the orbit), (iii) image interpretation and data extraction, and (iv) dependence on weather conditions (cloud cover obstructing free sight) (Zhang and Kovacs, 2012). Although these limitations of satellite remote sensing have been formulated in the context of grassland management and monitoring, they are also applicable in a breeding context, especially if we consider the typical small plot sizes handled in breeding programs (on the order of 1-10 m^2). There are several options of phenotypic platforms and the choice is dependent on the scale of the work and sensor chosen (Araus et al., 2018). For the application considered in this manuscript, close remote sensing using Unmanned Aerial Vehicles (UAV) was the more appropriate choice.

Indeed, sensors mounted on a UAV allow flights to be scheduled in function of the particular needs of the breeding program considered and they are relatively easy to perform. It is possible to acquire data throughout the whole growing season to obtain a time series of images at high temporal resolution, delivering dynamic information of crop growth and status (Zhang and Kovacs, 2012), while allowing the accurate screening of spatial heterogeneity in even relatively small plots. UAVs equipped with a high resolution consumer digital camera can capture the visible part of the reflected electromagnetic energy in a red, green and blue (RGB) band via imagery with cm-resolution (Bendig et al., 2015). Processing of these images using Structure from Motion (SfM) based software (Dandois and Ellis, 2010), enables the generation of digital elevation models (DEM) derived from 3D point clouds and highly detailed orthophotos that deliver centimeter resolution (Bendig et al., 2014; Tilly et al., 2014). Compared to multispectral, hyperspectral or thermal sensors, RGB sensors have limitations in spectral resolution, but they have the advantages of lower cost and higher spatial resolution, and it is possible to calculate vegetation indices and estimate plant height from the same set of photographs. Spectral and spatial information can then be combined to estimate crop biomass (Bendig et al., 2015).

RGB sensors mounted on UAV (drones) have been used to estimate DMY based on DEM in maize (Li et al., 2016), winter wheat (Yue et al., 2017), barley (Näsi et al., 2018), corn (Geipel et al., 2014), soybean (Raymond et al., 2005), and grasslands, among others. Specifically for grasslands, Possoch et al. (2016), Näsi et al. (2018), Viljanen et al. (2018) and Rueda-Ayala et al. (2019) have demonstrated the use of regression models involving plant height and/or vegetation indices derived from data obtained using RGB or multispectral sensors in combination with UAV to estimate DMY. Possoch et al. (2016) used a simple linear regression to relate canopy height and biomass yield of 18 plots on 11 dates (in total 196 samples) obtaining an r^2 of 0.63. Näsi et al. (2018), used simple linear regression and random forest (RF) to estimate DMY in a mixture of timothy and meadow fescue, but they considered only one flight and a relatively small number of samples (n = 36) which may limit the applicability of their conclusions to other situations. Viljanen et al. (2018) used data from four harvesting dates (all in June) with 96 samples of a

mixture of timothy and meadow fescue to generate multiple linear regression (MLR) and RF models. Finally, Rueda-Ayala et al. (2019) compared an on-ground system with an UAV in a limited number of plots (n=10) in two different species mixtures using simple linear regression models to estimate DMY and the UAV-based system showed a lower capability in the estimation ($R^2 < 0.6$). While these studies illustrate the potential of RGB and multispectral sensors to provide estimations of DMY of crops in general and forage grasses in particular, they were rather limited in the number of plots, conditions tested and/or growth stages considered, which may limit the general applicability of the results found. Therefore, to the best of our knowledge, estimation of DMY from canopy height over an entire growing season for perennial ryegrass using high resolution UAV-RGB images has not been reported yet. This is especially challenging as the height of perennial ryegrass canopies are typically lower (0.2 and 0.4 m) than those of other forage crop species such as timothy-meadow fescue mixtures (0.6-0.7 m) (Viljanen et al., 2018) and bermudagrass/alfalfa mixtures (0.8-1.1 m) (Pittman et al., 2015).

The aim of this study was to develop a methodology for monitoring the spatial and temporal dynamics of biomass accumulation of perennial ryegrass plots throughout the growing season in an affordable, easy-to-use, reliable and non-destructive way using an RGB camera mounted on a UAV. First, we evaluated the accuracy of canopy height models (based on the Structure from Motion (SfM) and the derived digital elevation model) to determine canopy height, using measurements taken using a rising plate meter as reference. Second, correlations between different variable sets (canopy height, vegetation indices, and temperature as environmental driver) were evaluated. Third, different regression models (MLR, PCR, PLS and RF) were built to estimate DMY. Finally, the performance of the different models was compared.

3.2 Materials and methods

3.2.1 Field trial

The data was acquired in 2017 in an experimental field (Figure 3-1) of 0.5 ha located in Melle (Belgium, lat. 50.98 N, long 3.78 E). The trial was sown on a sandy loam soil at the beginning of October 2015 at a density of 2 g seeds/m² and comprised a total of 420 perennial ryegrass accessions from Europe, North Africa and West Asia, as part of the FACCE-JPI ERA-NET+ Grasslandscape project (https://www.faccejpi.com/Research-Themes-and-Achievements/Climate-Change-Adaptation/ERA-NET-Plus-on-Climate-Smart-Agriculture/Grasslandscape). This set of accessions displays a very broad range of genetic and phenotypic diversity regarding growth characteristics. To facilitate the management of the trial, the accessions were grouped according to heading date (124 late, 132 intermediate, 123 early and 41 very early accessions). Within each group the accessions were organized in a randomized block design with three repetitions, rendering a total of 1350 plots (420 accessions x 3 repetitions) of 0.6 x 1.75 m (1.05 m^2). The trial was managed under a cutting regime of four or five cuts per year at a mowing height of 5 cm. We select 39 representative plots of accessions from the different groups. In the following text, we refer to these plots as 'Target plots' (Lp-TPs). In addition, 90 plots (0.6 x 1.75 m, 1.05 m^2), named 'Control plots' (Lp-CPs) below, were sown with cultivar Olano in the border areas of the different groups. The Lp-CPs, organized in three blocks (Figure 3-1), were used for additional measurements and for destructive sampling between cuts to provide extra data for method

development. The field trials were mown five times in 2017 (Figure 3-2), but the flight campaign included only four growth periods (GP1-GP4), as described in the following sections.



Figure 3-1. Schematic overview of the experimental field trial with control and validation ground control points (cGCPs, vGCPs), artificial height references (AHR) located in four zones (numbered 1-4), *Lolium perenne* control and target plots (Lp-CPs, Lp-TPs,) and plots where biomass samples were taken (biomass samples).

3.2.2 Data acquisition

3.2.2.1 UAV flights

In 2017, between April 25 and October 3, a weekly flight was performed when weather conditions allowed (Figure 3-2) using a dodecacopter UAV model Onyxstar HYDRA-12 (AltiGator, Belgium) equipped with a RGB (visible spectral range) camera (α6000, Sony Corporation, Japan), with 6000 x 4000 pixels, equipped with a 12 mm wide-angle lens (ZEISS Touit 2.8/12, Carl Zeiss AG, Germany). Per flight a set of nadir images of the experimental field was collected from a height of 30 m above ground level, ensuring an overlap of 80% in both directions (forward-longitudinal-lap and side-lap). The gimbal on which the sensor was fixed compensated for UAV movements during the flight and guaranteed nadir image collection. The flight route was programmed in MikroKopter-Tool V2.14b (HiSystems GmbH, Germany) and was automated to ensure whole coverage of the trial. The camera was triggered automatically according to the designed flight route. The APS-C sensor was set to record images in manual mode to avoid different settings in successive images. Shutter time (exposure), aperture (F-stop) and sensitivity to light (ISO) were adjusted in the field before the start of each flight based on the light conditions at that moment and auto-focus activated. All 22 flights were executed around 13 h (local time). Each flight delivered 100 images. Seven minutes were needed to capture them (duration of the flight was less than 10 minutes from take-off to landing) which allowed light conditions to remain unchanged. During each flight a gray card of 15 cm x 20 cm (18% reference grey, Novoflex Präzisionstechnik GmbH, Germany) was placed on the ground in the middle of the field, to be used as a color neutral reference to correct for white balance and exposure (it was visible in 15% of the images) and to check if differences in light conditions occurred. The raw images were converted to lossless tiff files in Lightroom v6.7 (Adobe Systems Incorporated, USA). During each flight, artificial height references (AHR) of known height were placed in the field. They consisted of four sets of seven surfaces

of 0.04 m^2 (20 cm x 20 cm) of different heights ranging from 5 to 40 cm (5, 10, 15, 20, 25, 30 and 40 cm). The sets were placed close to the borders and in the middle part of the field (Figure 3-1) and were used to estimate the altimetric accuracy.

The SfM software Agisoft Photoscan v1.2.6 Professional Edition (Agisoft LLC, Russia) was used to process the 22 sets of images and build geo-referenced orthophotos and digital elevation models (DEMs). Thirty-five Ground Control Points (GCPs) geo-referenced using a RTK GPS (Stonex S10 GNSS, Stonex SRL, Italy) were given in the software (Figure 3-1). Twenty-seven of these GCPs were used in the Agisoft workflow (calibration GCP, cGCP) and the remaining 8 GCPs were used for method validation (validation GCPs, vGCPs). The resulting ortho-mosaicked images had a mean planimetric spatial resolution of 0.6 cm and the DEMs had a mean altimetric resolution of 1.5 cm. Both resolutions were defined automatically by the software based on the camera parameters and flight altitude.

Digital elevation models (DEMs) were used to build a Canopy Height Model (CHM) of the field for each date. For this purpose, we built two types of DEMs: 1) Digital Terrain Model (DTM), corresponding to the ground at the time of mowing, and 2) a Digital Surface Model (DSM) derived from the imagery collected in the consecutive weeks with the presence of canopy on the terrain. Two types of DTMs were built: (i) DTM derived from the flight carried out as close as possible to each mowing event (with a correction of 5 cm due to the remaining stubble after the mowing) or (ii) DTM based on a triangulated irregular network (TIN) interpolation of GCPs distributed all over the trial. The CHM was derived by subtracting the DTM from the DSM. In total 40 CHMs were built, (i) 18 of which were produced using the DTM based on the stubble and (ii) 22 using the TIN interpolation.

3.2.2.2 Ground measurements

With the same frequency as the flights (Figure 3-2), a single non-destructive measurement of canopy height (CH) was taken in the central part of each Lp-TP using a rising plate meter (RPM, HerboMETRE, ARVALIS-Institut du Végétal, France) consisting of a square rising plate of 0.09 m^2 (30 x 30 cm). In the 90 Lp-CPs, three CH measurements per plot were recorded, consistently at the same RTK GPS georeferenced locations. These locations were marked on the ground with plastic labels.

In addition, three Lp-CP plots were harvested weekly to determine the DMY. Each biomass sample was weighed fresh and a subsample of approximately 250 g was dried at 70°C for 48 hours to determine the dry matter weight. To obtain the DMY, this weight was extrapolated to kg ha^{-1}. The harvests were done immediately after the UAV flight or the subsequent day. On three occasions (May 31, August 24 and October 4) DMY was determined in a similar way on the same set of 39 Lp-TP plots.

Air temperature during the experiment was recorded using a weather station located at a distance of 2.25 km from the trial and was used to calculate Growing Degree Days (GDD) considering a base temperature of 0°C.

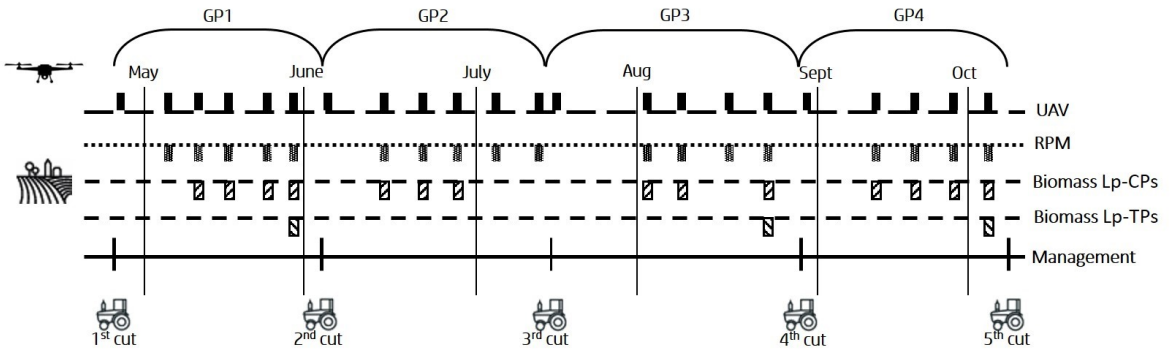


Figure 3-2. Timelines of (i) the UAV flights, (ii) the rising plate meter measurements (RPM), (iii) the biomass measurements of the control plots (Lp-CPs), (iv) the biomass measurements of the target plots (Lp-TPs) and (v) the field trial cutting regime. Growth periods (GP) are indicated.

3.2.3 Data handling and analysis

The workflow followed in this study is presented schematically in Figure 3-3. First, the accuracy of the UAV-products was evaluated. Second, the canopy heights estimated from the UAV and measured with an RPM were compared. Third, DMY was estimated using different models and results were compared. QGIS v2.14.16-Essen (QGIS Geographic Information System; Open Source Geospatial Foundation Project) was used to extract information from the images and to build canopy height models. The statistical analyses were carried out in R v3.5.1 using RStudio v1.1.456 (RStudio: Integrated Development Environment for R, RStudio Inc., USA) with the packages *FactoMineR*, *pls* and *randomForest*.

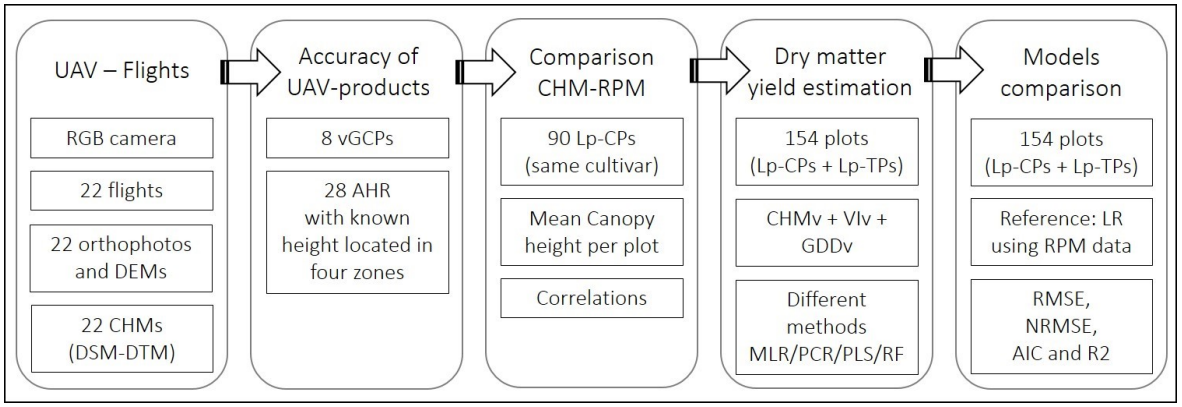


Figure 3-3. Schematic overview of the workflow followed in this manuscript. DEM: digital elevation model; CHM: canopy height model; DSM: digital surface model; DTM: digital terrain model; vGCPs: validation ground control points; AHR: artificial height references; Lp-CPs: control plots; Lp-TPs: target plots; CHMv: canopy height model variables; VIv: vegetation index variables; GDDv: environmental variables; RPM: rising plate meter; LR: linear regression; MLR: multiple linear regression; PCR: principal component regression; PLS: partial least squares; RF: random forest; RMSE: root mean square error; NRMSE: normalized RMSE; AIC: Akaike Information Criterion.

3.2.3.1 UAV product quality

As one of the aims of this study was to measure variables such as canopy height linked to planimetric and altimetric spatial resolution, the quality of the products derived from the flights was evaluated. First, the 8 vGCP described in 3.2.2.1 were used. Variance and root mean square error (RMSE) were calculated between the coordinates (X-easting, Y-northing and Z-altitude) measured with the RTK GPS and the coordinates derived from the orthophoto and the DEM. RMSE is then equivalent to the distance between the measured and the estimated coordinates. To investigate further the altimetric resolution, AHR in the expected range of the canopy considered in this experiment (0-40 cm) described in 3.2.2.1 were used. We calculated the height of the central part (5 x 5 cm) of each object using the CHM of each UAV flight (see Figure 3-4), and compared this value with its known height.

3.2.3.2 Canopy height and vegetation indices

We compared the RPM measurements and canopy height data extracted from the CHM for 18 of the 22 flights; the other four flights were executed just after a cut and no canopy height was measured. We used the Lp-CPs and considered only the RTK GPS locations at which the RPM measurements had been carried out (3 per plot). Georeferenced polygons with the dimension of the plate were defined in QGIS to ensure that data extraction was at the exact same location. In these polygons the mean height values were calculated. Correlations between the CHM values and the RPM data were determined on three levels: per flight, per growth period (all flights between two cuts) and for the entire growing season. The mean CH per plot, calculated from three measurements carried out in each control plot, was used for correlation analysis. Thus, linear regressions were built per flight, growth period and growing season.

Finally, for all whole plots considered for DMY estimation (45 Lp-CPs and 39 Lp-TPs) minimum, maximum, mean, P50 (median), P90, standard deviation and coefficient of variation values were derived for the canopy height from the CHMs and for two vegetation indices: Excess Green (ExG = (2G-(R+B))/(R+G+B)) (Woebbecke et al., 1995) and Excess Green minus Excess Red (ExGR = ExG - ExR = ExG - (1.4R-G)/(R+G+B)) (Camargo Neto, 2004). These indices were selected because they are normalized indices and reported as robust (less variable) with changes in lighting conditions (Woebbecke et al., 1995). In chapter 2, 13 vegetation indices were tested and these two indices rendered consistent and stable results.

3.2.3.3 Prediction of dry-matter yield from UAV and ground-derived information

Different types of regression models were built to estimate DMY for the Lp-CPs and Lp-TPs selected using different sets of variables. An overview of all variables used in the modeling step is presented in Table 3-1. Previous to the model building, Pearson correlation was calculated among the variables. First, we built models incorporating canopy height variables (CHMv) only. Vegetation index variables (VIv) and environmental information (GDDv) were incorporated in two further steps. Because the RPM is a mixed measure combining both the height of the canopy and its density, we also tested the performance of VIv extracted from the UAV-RGB images. By combining CHMv and VIv it was investigated whether the VIs describing the greenness of the canopy could contribute complementary information. A linear regression (LR) was used as reference. Backward multiple linear regression (MLR) using different

variables from the three sets and a principal component regression (PCR) with all variable sets were carried out. In addition, partial least squares regression (PLS) and the machine learning technique random forest (RF) were tested. All variables except RPM were used, since we aimed to estimate DMY based on UAV data and replace RPM measurements. Only in one MLR model, RPM data was added to evaluate the complementarity with other variables.

Table 3-1. Overview of the variables derived from the canopy height models, the orthophotos, and the weather station, that are used to predict dry matter production.

Data source	Variable set		Variable abbreviation	Description
DEM	Canopy Height Model variables (CHMv)		CH_min	Plot minimum canopy height
			CH_max	maximum
			CH_mean	mean
			CH_P50	median
			CH_P90	90th percentile
			CH_stdev	Standard deviation
			CH_cv	Coefficient of variation
Orthophoto	Vegetation Indices variables (VIv)	ExG	ExG_mean	Plot mean Excess green
			ExG_P50	median
			ExG_P90	90th percentile
			ExG_stdev	Standard deviation
			ExG_cv	Coefficient of variation
		ExGR	ExGR_mean	Plot mean Excess green minus Excess red
			ExGR_P50	median
			ExGR_P90	90th percentile
			ExGR_stdev	Standard deviation
			ExGR_cv	Coefficient of variation
Weather station	Environment variables (GDDv)		GDD	Growth degree days
			GDDdif	Growth degree days difference between cut date and measurement date

To estimate the performance of the different models, the complete plot dataset was divided into a training (70%) and validation (30%) set. The derived regression models were applied to the validation dataset and results (Table 3-2) were analyzed and compared based on the root mean squared error (RMSE) and normalized RMSE (NRMSE) between the observed and predicted DMY, using Akaike's Information Criterion (AIC) and the coefficient of determination (R^2). For the MLR models, after the

backward selection procedure (using the training dataset), the results were evaluated in terms of significance of the predictor variables and multicollinearity using the variance inflation factor (VIF). The model was manually adjusted by removing predictor variables with $p > 0.10$ or with a VIF > 10. Since the MLR models were hierarchical, the significance of the variables was evaluated at each stage using ANOVA.

3.3 Results

3.3.1 Accuracy of UAV products

The accuracy test explained in section 3.2.3.1 (performed for both types of CHMs) demonstrated that better results were achieved with the CHMs built using the TIN interpolated DTM. Only data derived from those CHMs are presented.

3.3.1.1 Accuracy of the ground level references

Based on the eight validation GCPs (vGCPs), variance and RMSE were estimated for X, Y and Z coordinates. The variance values ranged from -1.10 cm to 0.20 cm, from 0.04 cm to 0.60 cm and from -2.25 cm to -0.41 cm for the three axes (X, Y and Z respectively). RMSE values ranged from 0.63 cm to 1.49 cm (for X and Y) and from 0.94 cm to 1.89 cm for Z.

3.3.1.2 Accuracy of the height references

The variance was calculated for each AHR, field zone and flight (data not shown), rendering an average variance value of -0.56 cm. In Figure 3-4, the height estimates for the AHR are shown for five representative CHM subsets.

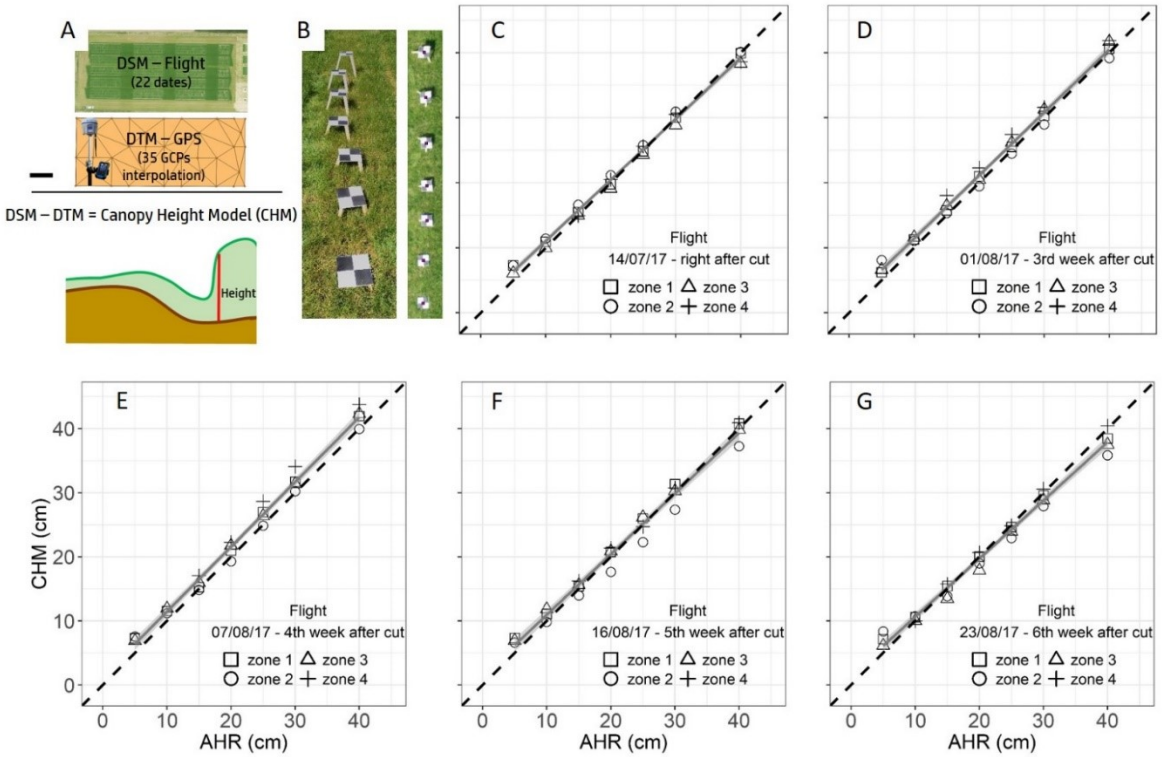


Figure 3-4. (A) Schematic presentation of the calculation of the canopy height model (CHM). (B) Side and top view of a set of artificial height references (AHR). (C-G) Accordance of real height and the height of the twenty eight AHR determined in four different zones of the field trial for 5 flights during growth period 3. Due to adverse weather conditions, any flight was possible between 14/07 and 01/08. The gray band around the line indicate the 95% confidence interval.

3.3.2 Comparison between the UAV-derived canopy height and the canopy height measured with the rising plate meter

Correlations were established between the CHM and RPM per flight and per growth period (4 to 5 flights/GP, GP1-GP4, Figure 3-5). The correlation between canopy height values derived from the CHM (CH_mean, n=3 per plot) and the RPM values were 0.96, 0.80, 0.81 and 0.80 ($p < 0.001$, Figure 3-5B-E) respectively. The correlation for the complete set of flights (17 in total, as one flight had to be discarded due to an error in the measurement with the RPM) was 0.92 ($p < 0.001$ Figure 3-5F). In general, higher canopy height values were obtained from the CHM than from the RPM. The lower canopy height values (underestimation) in the RPM data were expected as the RPM flattens the canopy during the measurement. The fact that this was not observed for the first three weeks of the second GP suggests that the sparse canopy (consequence of a severe drought period) and the open areas in which the ground was visible influenced the mean value calculated for the area covered by the polygon in QGIS, and that the compression effect by the RPM was reduced as the canopy was low. The compression by the plate was higher at later weeks of the regrowth period. A linear and a loess (non-parametric locally weighted) regression were carried out for the entire set of flights (Figure 3-5F) between CHM and RPM to study their relation. It was observed that the relation between both variables changed depending on the mean canopy height. Clearly three zones could be determined with a different relation: from 5 to 13 cm, 13-25 cm, and 25-40 cm.

3.3.3 Correlation between variables

Correlation results are presented in a correlation map (Figure 3-6). Dry matter yield displayed a good correlation with the RPM values (r= 0.79). Correlation between DMY and CHMv such as CH_min, CH_mean, CH_P50, and CH_P90 was also in the same range. The highest correlation was obtained between DMY and CH_P50 (r = 0.80). Low r-values were obtained between DMY and VIv (0.28 - 0.53). A negative correlation was found between GDD and DMY (r = -0.42).

RPM values were highly correlated with CHM data, especially with CH_mean and CH_P50 (r = 0.88). RPM values were also significantly correlated with VIv, but the correlation values were lower compared to CHMv. The highest correlations were found for both ExG_stdev and ExGR_stdev (r = 0.67). The correlation is related to variation in greenness of the canopy. No clear preferred correlation could be found for ExG or ExGR derived variables in relation to RPM. This was related to the high correlations between ExG and ExGR variables (r-values for mean, P50 and P90, ranging between 0.94 and 0.99).

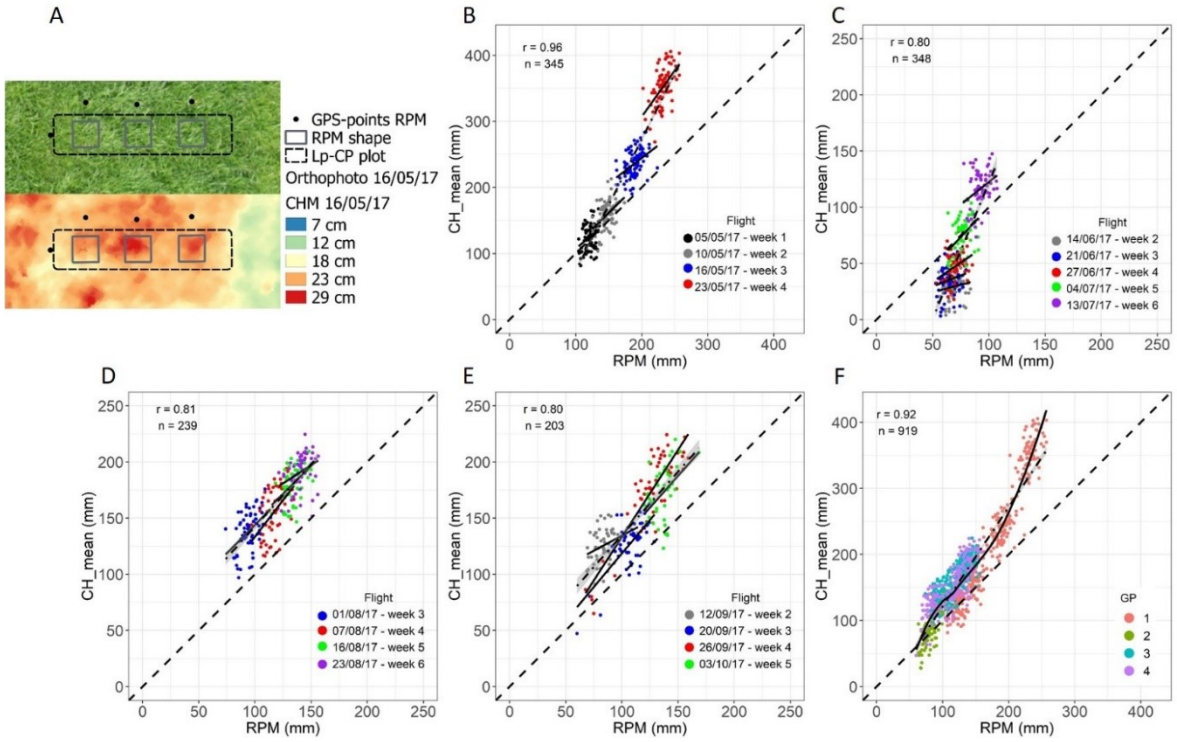


Figure 3-5. (A) Representation of a plot (dashed lines) and the positions measured with the rising plate meter (RPM, grey squares) in an orthophoto and canopy height model. The positions of the RPM measurements were RTK GPS referenced (black dots, yellow markers were placed in the field) to enable extracting canopy height from the same location from the canopy height models (CHM). (B-E) Comparison of the canopy height using the rising plate meter (RPM) and the mean canopy height (CH_mean) derived from the different canopy height models for the four growth periods (GP). A regression was applied per flight. The correlation of all data for each GP is calculated and shown in the upper left corner of each figure and the (dot dashed) regression line with the 95% confidence interval (gray band). (F) Comparison for the entire growing season with a linear and a loess regression and their 95% confidence interval between CHM and RPM.

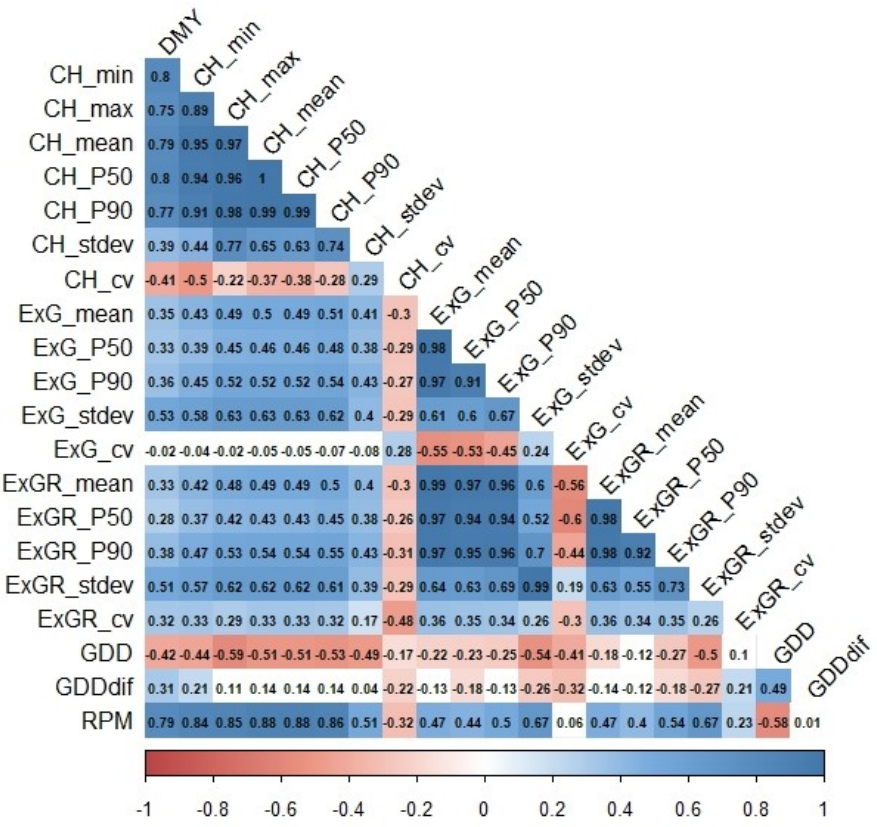


Figure 3-6. Pearson correlation map of canopy height model variables (CHMv), vegetation index variables (VIv), environmental variables (GGDv), and rising-plate meter measurements (RPM) used for predicting dry matter yield (DMY) (n = 154). All colored correlations are significant at the $p < 0.05$ level.

3.3.4 Dry-matter yield models

Different models were built to estimate DMY. Results are presented based on the validation dataset (Table 3-2; see Materials and Methods above). The first model was a linear regression with the RPM variable, which served as reference for other models. For our trial, this linear regression resulted in an RMSE of 986 kg ha^{-1} and a NRMSE of 31.0%. The linear regression based on the best performing UAV-derived variable, CH_P50, performed better than the RPM, with RMSE equal to 876 kg ha^{-1} (NRMSE = 27.6%), which is an improvement of 11%.

Using multiple linear regressions (MLR) we evaluated if a further improvement in the DMY prediction could be achieved. ANOVA was used to compare the different MLR model performances and their significance. In a first multiple linear regression model (MLR), only CHMv were considered in the model. The resulting RMSE was 825 kg ha^{-1} (NRMSE = 26.0%) and thus slightly lower than the linear model with CH_P50 and a 16% improvement compared to the linear RPM model. This result was paired with similar AIC and a slightly higher R^2 and achieved by including the CH_stdev. However, the observed improvement compared to the linear model incorporating CH_P50 was just not large enough to be significant ($p = 0.056$).

CHMv and Viv were combined and the resulting model hardly improved the prediction, and resulted in an RMSE of 822 kg ha^{-1} (NRMSE = 25.9%). Indeed, no significant improvement was found compared to

the first MLR (p = 0.100). To investigate if the RPM contained further complementary information compared to the CHMv and VIv, a third MLR was calculated including also RPM. A higher RMSE was achieved in this case: 857 kg ha^{-1} (NRMSE = 27%) compared to the previous MLRs, showing no clear complementarity. In a fourth MLR model GGDv were included in combination with CHMV and VIv. By incorporating GGDv we added information on growth drivers and on the different GP. In this case, a significant ($p < 0.001$) improvement in the DMY estimation of 31% compared to the linear RPM model was obtained, resulting in an RMSE = 679 kg ha^{-1} (NRMSE = 21.3%).

In the fifth model, a principal component regression (PCR) was used for which five components were selected based on the eigen values, resulting in an RMSE of 714 kg ha^{-1} (NRMSE = 22.5%) or a 28% improvement compared to the linear RPM model. Finally partial least square (PLS) and random forest (RF) methods were evaluated to predict DMY. After optimization, the PLS and RF resulted in RMSEs of 739 (NRMSE = 23.3%) and 769 kg ha^{-1} (NRMSE = 24.2%), respectively. However, these models did not contribute any improvement compared to the best MLR model (MLR with CHMv+VIv+GDDv). In Table 3-3, the summary of the best MLRs is presented. The variable with the highest standard effect size in the model is CH_P50 followed by the GDDdif, which was calculated as the GDD accumulated after each cut.

Table 3-2. Overview of the models used to estimate dry matter production. Models used were: linear regression (LR), multiple linear regression (MLR), principal component regression (PCR), partial least square regression (PLS), and random forest (RF). Model results are evaluated using the root mean square error of the validation dataset (RMSE, kg ha-1), normalized RMSE (NRMRE, %), Akaike information criterion (AIC), and R^2 or approximated R^2 (for PLS and RF). PLS and RF were optimized in terms of RMSE by selecting 2 (Ncomp3) and 4 components (Mtry4), respectively.

Model type	Variables	RMSE (kg ha^{-1})	NRMSE (%)	AIC	R^2
LR	RPM	986	31.0	1707	0.60
	CHMv – CH_P50	876	27.6	1709	0.67
MLR	CHMv	825	26.0	1707	0.71
	CHMv + VIv	822	25.9	1707	0.71
	CHMv + VIv + RPM	857	27.0	1693	0.69
	CHMv + VIv + GDDv	679	21.3	1677	0.81
PCR	CHMv + VIv + GDDv Ncomp 5	714	22.5	-	-
PLS	CHMv + VIv + GDDv Ncomp 3	739	23.3	-	(0.58)
RF	CHMv + VIv + GDDv Mtry 4	769	24.2	-	(0.70)

Table 3-3. Best MLR to predict DMY with all variables sets included in the model. For all variables mean, standard deviation (SD), effect size and standardized effect size are presented.

Variable set	Variable	Mean ± SD	MLR	
			Effect Size (βi)	Std effect size
	DM	3233.38 ± 1613.8		
Intercept (β0)			-959.3	
CHMv	CH_P50	205.42 ± 89.9	10.14 ***	911.3
	CH_stdev	20.15 ± 10.9	-31.34 **	-341.0
VIv	ExG_stdev	0.08 ± 0.02	14057.15 **	281.1
GDDv	GDD	2331.10 ± 936.3	-0.63 ***	-308.7
	GDDdif	561.85 ± 145.9	5.46 ***	796.4

Significance levels: *** p < 0.001 ** p < 0.01 * p < 0.05

3.4 Discussion

3.4.1 Accuracy of UAV-derived data

Prior to the use of DEMs and orthomosaics, an evaluation whether these products do provide correctly georeferenced information is needed and is especially relevant in time series. Here we used two methods: a set of validation ground control points (vGCPs) and artificial height references (AHR). Similar RMSE values were obtained for the eight vGCPs used. However, the flights just before a cut tended to show a slightly higher RMSE (from about 1.24 cm up to 1.89 cm). The higher RMSE was caused by the partial covering of cGCPs by the taller surrounding grass, resulting in fewer images where the cGCPs were visible and could thus be used to calculate the 3D model. This effect was accentuated when the cGCP differed more from the nadir position in the image (results not shown). In our case, the flight plan, overlap and number of images was sufficient and correct, as demonstrated by a mean calculated error in the Z-axis below 2 cm. The mean error calculated in the Z-axis was comparable or even better than shown in other studies in grasslands. For example, Näsi et al. (2018) or Viljanen et al. (2018) presented RMSE values in the Z-axis of 6.9 cm and 1.0 – 4.8 cm, respectively. The size of the area 'free of vegetation' needed around a GCP depends on the canopy height and the lens used. For a taller crop, more open area around the GCP is needed. Therefore, this is an important point of attention when planning the experiment. Furthermore, the stitching software needs sufficient overlap between images for the edges of the field trial to provide the same height accuracy as the central area of the DEM. Thus, the flight plan of the drone should be taken wide enough. In this study, every point of the field trial appeared in a minimum of 15 images, and the GCPs located in the central part of the field were present in 20 images. The number of images was higher than in other studies where the area of interest was captured in a minimum of 9 images (Bendig et al., 2015; Viljanen et al., 2018).

Regarding the AHR, slightly higher variances were obtained for the set located closer to the border (zone 4). This result was expected as they appear in a lower number of images and were covered by a smaller angle of view to reconstruct them in the photogrammetry processing compared to the references located in the middle of the field. The lowest AHR (5 cm) showed higher variances in height. This outcome can be explained by the higher difficulty involved in the detection of less tall objects and their modeling due to (i) their closer distance to the soil that increases the difficulty for the SfM software and (ii) the taller surrounding grass, especially at the end of the growth period that reduces its visibility in the extreme oblique images. To obtain the DTM needed to build the CHM, it is possible to use a flight without vegetation (Bendig et al., 2015), interpolation of points (manually or automatically) classified as ground using the borders of the plots or visual bare soil areas (Viljanen et al., 2018; Yue et al., 2017), to use a ultrasonic or laser scanner (Pittman et al., 2015) or interpolating X, Y and Z coordinates of the plot corners measured with a GPS (Chang et al., 2017). In our case, very good results were obtained using interpolated coordinates measured with the RTK GPS. Using that information, it was possible to determine the slope of the field and to detect bumps or irregularities present. Consequently, a higher precision in AHR height was achieved.

3.4.2 Canopy height determined using the RPM and CHM

The rising plate meter is often used as a reference for non-destructive assessment of biomass in pastures (e.g. Holshof et al., 2015; Mathieu and Fiorelli, 1985). In this study, it was also used as a reference even though we were aware that the RPM compresses the canopy, which makes the height measurement dependent on the canopy density, resulting in lower canopy heights in comparison to the CHM (information acquisition without physical contact). Furthermore, the RPM encounters problems when the sward is sparse, non-uniform, presents low growth rate or is too high. Therefore, it has been reported to be more reliable when the sward is dense and has reached a height of 20-30 cm (Viljanen et al., 2018). Having ruler-based height measurements as an undisturbed reference measurement was not an option as this measurement can only be done very locally and questions would remain with respect to the height of the canopy to be measured (e.g. average canopy height, tallest leaves,…), the number of measurements needed for a good estimation or how to avoid human bias (Sanderson et al., 2001).

The correlation obtained between the canopy height measured with the RPM and the variables derived from the CHM for the Lp-CPs was 0.92 (Figure 3-5B). The correlation is similar to the values reported by Viljanen et al. (2018) for a timothy and meadow fescue mixture (r = 0.91-0.94), similar to the values reported in a 3-year study by Bareth and Schellberg (2018) (r = 0.83-0.91) and higher than the value of 0.55 reported by Possoch et al. (2016). It was observed that the CHM-CH was underestimated for the earliest measurement dates of a growth period and during a growth period with reduced growth due to abiotic stress (i.e. drought stress). This observation may be attributed to the ground level being visible in the imagery, which lowers the DSM as a result of the point cloud interpolation carried out by the photogrammetry software (Brocks and Bareth, 2018). A binomial distribution in the height values could be expected showing a peak for the canopy height and another corresponding to the soil level, but the software may not always be able to detect that difference such that this information may be lost during processing. As PhotoScan does not provide full details on its algorithms and their limitations, it is

recommended to investigate if the systematic error can be mitigated through the inclusion of oblique images (James and Robson, 2014) or a higher number of GCPs (Roth and Streit, 2017). However, this would come at the price of an increase in the processing time to obtain the UAV products.

The RPM merely provides information related to the area covered by the plate. Therefore, many measurements are required within the field to obtain a representative biomass estimate of large plots. Moreover, the spatial resolution of this method is considerably lower than the 1.5 cm resolution of the CHMs. Indeed, UAV imagery allows to capture in-field variability faster with a dense spatial coverage (Bendig et al., 2015). Moreover, the RPM consistency was evaluated by measuring three times the Lp-CPs on the same date (August 16; results not shown). An average Pearson correlation of 0.76 was obtained between the measurements, which indicates that the consistence of the RPM is limited. In the case of CHM, the correlation between two consecutive flights on the same date was much higher (r = 0.97). These findings together with the studies mentioned previously support the use of CHMs to measure canopy height and replace RPM measurements.

3.4.3 Dry-matter yield prediction

In total, 19 variables related to the canopy height model (CHMv), the vegetation indices (VIs) and the environment (GDDv) were used to build models that predict DMY. We tested if the combination of different types of data resulted in a better DMY prediction and investigated the influence of the methodology selected for the regression. As demonstrated by Bendig et al. (2015) and Brocks and Bareth (2018) for barley, Yue et al. (2017) for winter wheat, Näsi et al. (2018) for a timothy and meadow fescue mixture or for maize in Geipel et al. (2014), and as is shown in our regression analysis, the canopy height is (well) correlated to yield. Slightly lower correlation values (r = 0.79) were obtained in comparison to those reported by Viljanen et al. (2018) for a timothy and meadow fescue mixture (r ranges from 0.79 to 0.98). This can probably be explained by (i) the taller crop (up to 0.7 m) in their study, and (ii) the much more diverse set of accessions in the present study, implying differences in canopy density which cannot be captured by CH_P50. Here we used the model for DMY estimation based on RPM measurements as reference. Even if the RPM measured a weighted or averaged canopy height when the linear regression was built, it presented a higher RMSE than the model using the CH_P50, which achieved an 11% reduction of the RMSE.

To increase the performance of the prediction, the MLR model with more CHMv resulted in a slight improvement of 16% compared to the RPM model when CH_stev was included. Variables such as mean, median or 90th percentile described the vertical distribution of the canopy and the standard deviation and coefficient of variation, the vertical complexity and heterogeneity (variation). In a second MLR, spectral information was included in the form of vegetation indices (VIs). As VIs provide information on soil coverage and canopy density, they showed significant correlations with the RPM which is related to sward density. Nevertheless, they presented a lower correlation with DMY. Hence, it was difficult to achieve a relevant decrease in the RMSE compared to the first MLR model when ExG_P90, ExG_stdev and ExG_cv were included in the model. Indeed, this model did not show a significant improvement compared to the MLR with only CHMv. From this non-significant result, it can be concluded that the visual VIs are not crucial to achieve a better DMY estimation. The inclusion of other VIs such as NDVI or other wavelength ranges might add relevant information related to vegetation status and canopy cover

in the earlier weeks after a cut. However, NDVI may reach saturation in dense canopies or when LAI values are about 2.5-3 (Tilly et al., 2014) and thus may not be relevant for DMY estimation, because four or five weeks after the cut, grasslands can reach LAI values of 4 (Lambert et al., 1999). As we focused on the use of a (low-cost) RGB camera with high resolution to obtain accurate height estimations, the performance/contribution of NDVI measurements could not be evaluated in our study. Finally, as growth is strongly driven by temperature (e.g. Voorend et al., 2014), two variables that incorporate environmental data were selected (GDD and GDDdif) and showed to improve the prediction. Indeed, a seasonal growth pattern of high DMY in the spring and to a lesser extent in the autumn has been reported for *Lolium perenne* (Barre et al., 2017). Furthermore, the growth of plants has been shown to be related to temperature sum (Yang et al., 1995) taking into account a species dependent base temperature (Tb for *Lolium perenne* = 0°C) (Lootens et al., 2016).

With each step of the incorporation of a new group of variables, the RMSE decreased without increasing AIC, reaching a DMY model prediction improvement of 31% compared to the reference model based on RPM. Including the RPM data in the MLR resulted in no complementary effects. As can be seen in

Table 3-3, the standardized effect size (calculated as standard deviation multiplied by effect size) was estimated for the significant variables for the best MLR model to explain the variability of the DMY. As expected, the median canopy (CH_P50) explained the most variation followed by the temperature sum per growth period (GDDdif).

As an alternative for MLR, a PCR was executed to safeguard the regression against the correlations between the explanatory variables. The 5 principal components included in the regression model captured 96% of the variance in the X data, but this model did not result in an improvement of the prediction of DMY compared to the best MLR model. Because of the typically large data size, the use of machine learning techniques is increasing in remote sensing. Random Forest (RF) is one of the most popular algorithms because it requires no feature selection and is less sensitive to overfitting than MLR. Viljanen et al. (2018) incorporated this method for DMY estimation and they compared its performance with that of MLR. Similar to our study, one of their best results was obtained with RF, but it did not perform better than the best MLR that incorporated all the variable sets available. The use of PLS in our case did also not result in a further reduction of the RMSE.

To the best of our knowledge, no studies have been performed for perennial ryegrass plots related to DMY estimation with UAV imagery using an RGB camera to compare our results with. This might be due to the problems that could be encountered such as lower canopy than in, for example, cereal or the high variability between the different accessions or cultivars. Compared to previous UAV imagery research in other crops (Geipel et al., 2014, Raymond et al., 2005, Li et al. 2016, Yue et al., 2017, Näsi et al., 2018), in this study a higher number of flights (and thus dates), samples and accessions were involved in the estimation of DMY though different regression techniques, from a simple linear regression to more advanced techniques such as partial least squares regression and random forest. We have shown that it is possible to estimate canopy height of perennial ryegrass plots using a cheap, high resolution RGB camera mounted on a drone and predict DMY in a better way (lower RMSE) than with a rising plate meter. Indeed, a high spatial resolution (1.5 cm), and accurate CHMs in the Z-axis (1-1.9 cm) were achieved. By performing frequent flights throughout the growing season also a high temporal

resolution was achieved, which allowed to monitor all plots in a large field trial in detail in a non-destructive way.

Chapter 4: Closing the phenotyping gap: high resolution UAV time series for soybean growth analysis provides objective data from field trials

Adapted from:

Borra-Serrano, I., De Swaef, T., Quataert, P., Aper, J., Saleem, A., Saeys, W., Somers, B., Roldán-Ruiz, I., Lootens, P. (2020). Closing the Phenotyping Gap: High Resolution UAV Time Series for Soybean Growth Analysis Provides Objective Data from Field Trials. Remote Sensing 12, 1644.

4.1 Introduction

It has been demonstrated that close remote sensing approaches in combination with appropriate experimental designs and data integration can increase accuracy, precision, and throughput of on-field phenotyping experiments while also reducing cost and labor requirements (Mir et al., 2019). These high throughput phenotyping approaches can also be implemented in crop breeding to determine architectural/morphological and physiological traits to aid early detection of desirable genotypes (Shakoor et al., 2017). In these areas, the use of unmanned aerial vehicles (UAVs) is increasing steadily, as UAVs are suitable for use in complex field environments (e.g., muddy soil), are non-invasive, flexible in use (e.g., not restricted in use to particular field plots), and are available at a low cost (Yang et al., 2017). UAV flights can be scheduled in function of the needs of the breeding program or experiment, using combinations of sensors that are suitable to investigate the traits of interest. In one of the most straightforward applications, canopy cover (CC) and canopy height (CH) data can be derived from images obtained using an RGB camera mounted on an UAV (Araus and Kefauver, 2018). Canopy cover, defined as the ratio of the projected plant area to the total area considered, goes through changes from crop establishment to senescence. CC yields information regarding seedling emergence, early vigor, or senescence (Walter et al., 2015) and is therefore an important criterion in a breeding context (Yu et al., 2017). Furthermore, CC data can be used to estimate canopy photosynthesis and transpiration (Mullan and Reynolds, 2010) or to evaluate crop responses to abiotic stresses (e.g., drought stress; (Sayed et al., 2012)). Canopy height (CH) is also dynamic and linked to growth, crop development, and yield (Viljanen et al., 2018 and see chapter 3).

One of the main advantages of UAV-based imaging approaches is the high precision and throughput for data acquisition and processing in comparison to the manual ruler measurements or scoring systems that have been traditionally used to analyze crop growth and development. By performing UAV flights throughout the entire growing season, accurate crop data can be obtained at a high spatio-temporal resolution, delivering dynamic information on crop growth and performance (Zhang and Kovacs, 2012). For example, in Yu et al. (2016), data of five time points (using RGB and modified RGB-NIR imaging) were combined to classify soybean populations in different maturity groups and to estimate final yield. In Holman et al. (2016), a method for deriving crop height and growth rate from multi-temporal 3D digital surface models of winter wheat field trials was developed. In Pugh et al. (2018), UAV-based approaches to estimate height in sorghum and maize were determined and growth curves were fitted for different groups of breeding material. Similarly, in Chang et al. (2017) sigmoidal curves were fitted to UAV-derived multi-temporal height data of sorghum. Multi-temporal height data (four flights) have also been used in combination with clustering analysis methods to detect different growth patterns in maize varieties (Han et al., 2018).

These examples illustrate the value of multi-temporal data to describe crop growth and to predict specific parameters such as maturity, growth rate (in the above examples this was restricted to estimations considering two consecutive measurements) or yield. Up to now, however, little attention was paid to the biological interpretation of the growth curves built using UAV-based information. Nevertheless, the possibility for such interpretation represents one of the main advantages of this integrative approach. Multi-temporal data at high resolution can be of great value for sophisticated

plant and crop modeling approaches. This is because crop models typically contain a large number of parameters that need to be estimated in an identifiable manner (De Swaef et al., 2019) based on preferentially large quantities of objective experimental (field) data. Furthermore, mathematical models can be fitted to multi-temporal data to link growth patterns to relevant environmental and developmental clues for a precise description of crop performance from sowing to harvest. This might also enable derivation of 'hidden' variables as phenome descriptors that are useful in a breeding context (e.g., determinacy).

In this study, we demonstrate this concept using a soybean (*Glycine max* (L.) Merr.) experiment with 454 field plots comprising a diverse set of genotypes. More specifically, we fit mathematical models to multi-temporal data of CC and CH obtained for each plot and explore the possibilities of a global description of crop growth and development throughout the season in the context of breeding applications. First, CC and CH from each flight are estimated. Second, curves are fitted considering multi-temporal data per plot to extract biologically relevant information. Third, the correlation/complementarity of the model-based parameters obtained are evaluated. Fourth, parameters are examined for a possible relation to seed yield and maturity. Finally, an example to illustrate how the data could be used in a breeding context is developed to select interesting plots or to depict variability in the collection of genotypes evaluated.

4.2 Materials and Methods

4.2.1 Field Trial

An experimental field located in Melle (ILVO, Belgium, latitude 50.59 N, longitude 3.47 E) was used. The trial was established in 2018 in the context of the EU project EUCLEG (http://www.eucleg.eu/) and comprised 359 soybean genotypes randomized in a row-column-block augmented design in which three reference genotypes were replicated nine times, nine reference genotypes were replicated six times, and 13 reference genotypes were replicated three times. The remaining 334 genotypes were not replicated, rendering a total of 454 plots of 0.50 x 1.20 m (0.60 m^2) or 0.80 x 0.75 m (0.60 m^2), depending on the group to which the genotype belongs (Figure 4-1A). Within the 454 plots, there were 12 border plots (located between groups sections in Figure 4-1A) that were not considered in our analysis. This design was used to achieve the highest possible genotypic variation for an association mapping study in which genomic and phenotypic data will be combined. The genotypes used were classified beforehand in four groups that were sown on four different dates in an attempt to synchronize flowering of early maturing and late maturing types: April 20 (GP1), May 3 (GP2), May 8 (GP3), and May 11 (GP4) (Figure 4-1C). Each plot consisted of three rows; only the middle row was considered for the measurements. The seeds were inoculated with rhizobium bacteria (BiodozTM, De Sangosse, France) and the field was fertilized with 36 kg N ha^{-1}, 20 kg P2O5 ha^{-1}, 140 kg K2O ha^{-1}, and 20 kg MgO ha^{-1}. For weed control, 910 g ha^{-1} pendimethaline and 72 g ha^{-1} clomazone was applied directly after sowing (on the four sowing dates). A manual correction on June 13 was applied to ensure that no weeds developed that could hamper the growth of the crop or influence UAV-based measurements.

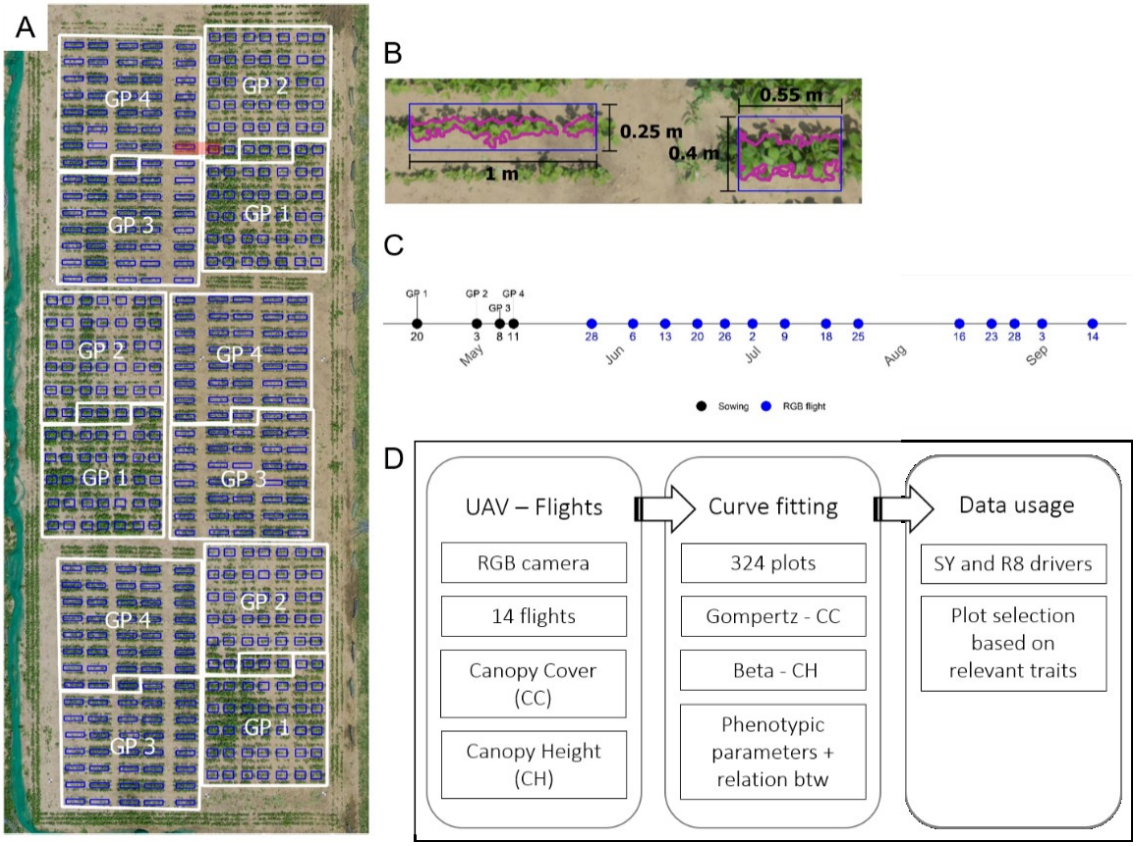


Figure 4-1 . Schematic overview of (A) the trial with group locations (white boxes, GP1-4), (B) zoom of two plots, marked by a red box in A, with the mask applied to the images to select pixels corresponding to vegetation, (C) timeline of the sowing dates (per GP) and UAV flights and (D) workflow followed in the manuscript.

4.2.2 Data acquisition

4.2.2.1 UAV Flights

Between May 28 and September 14, UAV flights were performed once per week when weather conditions allowed (Figure 4-1C), using a dodecacopter model Onyxstar HYDRA-12 (AltiGator, Belgium) equipped with a RGB (visible spectral range) camera (α6000, Sony Corporation, Japan), with 6000 x 4000 pixels and a 35 mm lens (f/1.8 Sony Corporation, Japan). Per flight, a set of nadir images was collected from a height of 30 m above ground level (raw image resolution of 3.35 mm/pix), ensuring an overlap of 70% in both directions (forward-longitudinal-lap and side-lap) at a flight speed of 2.5 m/s. The RGB camera settings (shutter speed, aperture, and ISO) were adjusted before each trial, set in manual mode, and triggered automatically according to the designed flight route. Depending on the weather conditions (light intensity and wind), shutter speed was ≤1/1000 s and ISO was ≤800.

In total, 14 flights were executed around 13:00h (local time) and the required images were captured in less than 20 minutes. A graph thermal time (°C d) versus time for the growing season is included in the supplementary materials (Figure A-3). Correction for white balance and exposure was applied to the images. Processing of the 14 sets of images and building of geo-referenced orthophotos and digital

elevation models (DEMs) were done using Agisoft Photoscan v1.2.6 Professional Edition (Agisoft LLC, Russia) based on 10 Ground Control Points (GCPs) geo-referenced with an RTK GPS (Stonex S10 GNSS, Stonex SRL, Italy). The resulting RGB ortho-mosaicked images had a planimetric spatial resolution of 0.5 cm; the DEMs had an altimetric resolution of 1.25 cm. Spatial resolutions were defined when both products (orthophotos and DEMs) were exported to ensure that products from different flights had the same resolution. Digital elevation models were used to build a Canopy Height Model (CHM) for each date. A triangulated irregular network (TIN) interpolation of 25 GCPs distributed all over the trial was used as Digital Terrain Model (DTM). The CHM was obtained by subtracting the DTM from the DEM.

4.2.2.2 On-Ground Measurements

Traits commonly recorded by breeders to track growth and reproductive development were monitored during the growing season by experienced UAV operators. The percentage of emergence was calculated by dividing the number of emerged plants by the total seed density. Plots with a percentage of emergence lower than 50% were discarded. The reproductive stage (from R1 to R6) was recorded at five fixed dates for each plot, as described in Fehr and Caviness (1977). The date of maturity (R8) was recorded separately by frequent visits to the field trial at the end of the growing season. R1 corresponds to beginning bloom (one open flower at any node on the main stem) and R8 is full maturity (95% of the pods have reached their mature pod color). As the R1 stage of each of the plots did not always coincide with one of the five time points considered, we estimated it by linear regression of the reproductive stage (five time points) and thermal time at time of scoring. Thermal time was calculated as the daily accumulation of the difference between the mean daily temperature and a base temperature (6 °C) over a period of time (°C d). Plots for which R^2 was lower than 0.7 were discarded, as the determination of R1 stage was considered inaccurate. In total, 324 plots remained for further analysis after the discards related to emergence rate and R1 stage. At maturity, five plants of the middle row of each plot were harvested by hand and were threshed together to determine average seed yield and number of seeds per plant.

4.2.3 Data Handling and Statistical Analysis

The different analysis steps are presented schematically in Figure 4-1D. First, canopy cover (CC) and canopy height (CH) data were extracted per plot. Second, sigmoid growth curves were fitted to CC and CH data using the 14 time points available and parameters were estimated (see Section 4.2.3.2). Third, a principal component analysis (PCA) was carried out to investigate the complementarity of parameters describing the genotypes. Last, predictive models were built for seed yield (SY) and R8 using multiple linear regression (MLR) models.

QGIS v2.18.26-Las Palmas (QGIS Geographic Information System; Open Source Geospatial Foundation Project) was used to extract information from the images. Georeferenced polygons with the dimension of the plot middle row were defined in QGIS to ensure that data were extracted from the same location for each and every flight and that only the information of the middle row was captured (Figure 4-1). The curve fitting and the statistical analyses were carried out in R v3.5.1 using RStudio v1.1.456 (RStudio: Integrated Development Environment for R, RStudio Inc., USA) coding customized functions for fitting the curves and using packages such as *FactoMineR* and *optimx* for the PCA and optimization of the

functions, respectively. These customized functions can be shared, upon request to the corresponding author.

4.2.3.1 Canopy Cover and Canopy Height

The canopy cover was calculated per time point and plot using the Excess Green vegetation index (ExG, (Woebbecke et al., 1995)). Pixels corresponding to vegetation were differentiated from pixels corresponding to soil (Figure 4-1B) using a simple thresholding, and the proportion of pixels classified as vegetation was then calculated. CHMs were built for each date (described in Section 4.2.2.1) and used to estimate the 90th percentile (P90) of the canopy height for each plot. Only pixels corresponding to vegetation (the pixel classification described in the previous paragraph was used to build masks) were considered for this calculation.

4.2.3.2 Model Fitting and Parameter Estimation

Canopy cover and canopy height data extracted for the 14 time points considered (timeline Figure 4-1C) were used to fit curves describing their temporal evolution in 324 plots. After reaching the maximum value, subsequent time points were only considered for fitting if the difference in value was not greater than 10% from the maximum canopy cover, and 20% from the maximum plant height. When plants get older and start to senesce, they gradually lose leaves and the CC and/or CH is reduced. These criteria prevented any excess influence of these changes on the CC and/or CH fits, as our interest was to fit the sigmoid part of canopy development. The last two time points were excluded by default, because senescence was already observed in a majority of plots. This implies that a maximum of 12 time points were included in each fit. For the fits, we chose curves that have parameters that are biologically interpretable (Zeide, 1993) and that reach a "plateau" that defines growth stop.

For the canopy cover (CC) data the Gompertz function (Gompertz, 1825) was used. In this function, an initial value (at the beginning of the growing season) is assumed. In the present experiment, the first drone flight was carried out when the plants had already emerged (canopy cover >0%). From the Gompertz fits, two parameters were extracted (Figure 4-2A, formula 1): maximum cover (CC_Cmax) and thermal time when growth rate was maximal (CC_tm). The Gompertz function used for CC has the following equation:

$$CC = CC_C_{max} e^{e^{-k(t-CC_tm)}}, \quad (1)$$

where CC_Cmax [%] is the maximum CC value attained, t is the thermal time [°C d], k [°C-1 d-1] is a constant that determines the curvature of the growth pattern, and CC_tm [°C d] is the inflection point at which the growth rate reaches its maximum value.

For the canopy height (CH) data, the Beta function defined in Yin et al. (2003) was chosen because it has more flexibility than Gompertz at the beginning of the growing season. In addition, for the Beta function the limit is not known in advance, while the CC data values are restricted by definition to the range 0–100. The Beta function was used to estimate the following parameters (Figure 4-2B, formula 2): maximum height (CH_Hmax), thermal time when maximum canopy height was achieved (CH_te), and

thermal time when the growth rate was maximal (CH_tm). The Beta function used for CH has the following equation:

$$CH = CH_H_{max}(1 + \frac{CH_t_e - t}{CH_t_e - CH_t_m})(\frac{t}{CH_t_e})^{\frac{CH_t_e}{CH_t_e - CH_t_m}}, \quad (2)$$

where CH_Hmax [m] is the maximum value of CH reached at CH_te [°C d], t is the thermal time [°C d], and CH_tm [°C d] is the inflection point at which the growth rate reaches its maximum value.

The accuracy of the fitting was evaluated by adjusted R^2 (CC_RsqA and CH_RsqA). Plots with an adjusted R^2 lower than 0.70 were discarded, resulting in 291 remaining plots (when both datasets were merged) for further analysis.

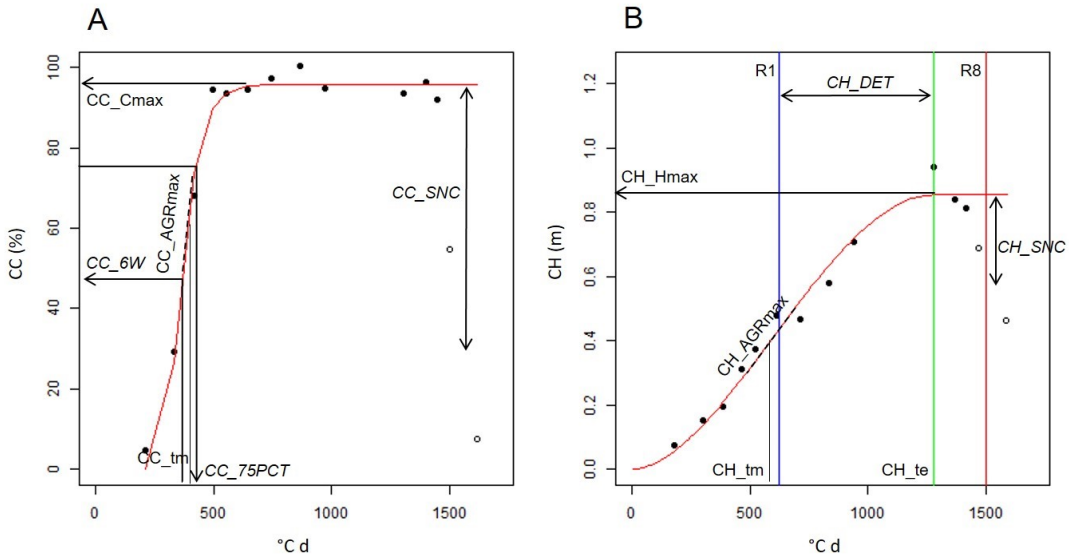


Figure 4-2. Schematic representation of the phenotypic parameters derived from the curve fitting for (A) canopy cover using the Gompertz function and (B) canopy height using the Beta function. By default the two last points where not taken into account in the fitting therefore they are represented by outline symbols. For abbreviations and explanations of the parameters, see Table 4-1.

4.2.3.3 Derived Phenotypic Parameters

Using the fitted growth curves, additional phenotypic parameters were derived (Figure 4-2). These phenotypes are more difficult to measure in the field. From the CC growth function, we calculated the maximum absolute growth rate (CC_AGRmax) as the first derivative of the function (which is the slope), the canopy cover six weeks after emergence (CC_6W; related to the capacity of the genotype to suppress weeds) and the thermal time needed to reach 75% of full canopy cover (CC_75PCT; related to canopy closure). From the CH growth function, maximum absolute growth rate (CH_AGRmax) was calculated and the degree of determinacy (CH_DET) was estimated by calculating the thermal time difference between the CH_te and the R1 stage (based on visual scores). According to this calculation, a high CH_DET value indicates a longer period of growth and a lower level of determinacy. Finally, the senescence rate was calculated as the difference between the maximum value of the fitted curve and

the average value of the last two time points. Calculations were performed for CC and CH (CC_SNC and CH_SNC, respectively). To determine whether a significant difference is present between the groups, a multiple pairwise comparison using Wilcoxon rank sum was carried out for all phenotypic variables.

Table 4-1. Overview of the phenotypic parameters.

Data source		Abbreviation	Description
UAV Orthophoto	Canopy Cover (CC)	CC_tm	Thermal time when maximum growth rate is achieved
		CC_Cmax	Maximum cover
		CC_AGRmax	Maximum absolute growth rate
		CC_6W	Cover percentage after six weeks emergence
		CC_75PCT	Thermal time when 75% of cover is achieved
		CC_SNC	Senescence rate calculated with CC data (CC_Cmax – average cover two last time points)/CC_Cmax
Digital Elevation Model	Canopy Height (CH)	CH_tm	Thermal time when maximum growth rate is achieved
		CH_Hmax	Maximum height
		CH_te	Thermal time when CC_Hmax occurred
		CH_AGRmax	Maximum absolute growth rate
		CH_DET	Determinacy calculated as CH_te – R1
		CH_SNC	Senescence rate derived from CH data (CC_Hmax—average height two last time points)/CC_Hmax
Breeders	Field observations	EP	Percentage of emergence
		SY	Seed yield
		R1	Thermal time when R1 stage starts (beginning bloom)
		R8	Thermal time when R8 stage is reached (full maturity)

4.2.3.4 Relations Among Phenotypic Parameters

A PCA analysis was performed to investigate the relation between the different phenotypic parameters and to find out which ones were correlated or complementary. Twelve phenotypic parameters obtained from the UAV imagery were included (Table 4-1) and three on-ground determined variables (as

quantitative supplementary variables). The four groups (GP1 to GP4) were considered together. To prevent a specific GP from determining the results, an equal number of plots per group were chosen at random (avoiding plots that corresponded to replicates of the same genotype) to ensure a balanced dataset. This resulted in a set of 148 plots. Pairwise correlations were also calculated between the parameters for the same set of plots. In addition, pairwise correlations per GP were estimated separately. When one genotype was represented by more than one plot (see 4.2.1), only one was retained to prevent the replicates from influencing the correlation output.

4.2.3.5 Seed Yield and R8 Stage Prediction from UAV Data

From a breeding perspective it is important to know which variables act as driver for seed yield (SY) or influence the thermal time needed to reach maturity (R8). Therefore, two backward multiple linear regression (MLR) models were built for SY and R8 with different sets of variables. The first MLR model was built using the values of the 12 phenotypic parameters that were determined from the curve fitting, whereas the second MLR model was built on the raw data of the 14 time points (flights). In both cases, the group classification was included as a dummy variable for the basic model to be used as starting point. To avoid collinearity between the phenotypic variables in building the first MLR model, variables were selected based on expert knowledge. Consequently, the variables CC_6W and CH_DET were discarded as they are highly correlated with CC_75PCT and CH_te. For the second MLR model, no expert variable selection was done since the raw data set is difficult to interpret in terms of redundancy. The MLR models were built using subset selection regression including a 10-fold cross validation. The variables for both datasets were normalized. In each subsequent step, the number of variables included was increased by one, while considering all possible combinations of variables and the best model was chosen based on the lowest RMSE.

4.3 Results

4.3.1 Model Fitting for Canopy Cover and Canopy Height

In an exploratory data analysis of CC and CH data, after discarding plots with emergence <50% or R^2 < 0.7 in the estimation of R1 stage, different growth patterns per group were observed (data not shown), which was confirmed in the subsequent analysis.

Fitting of CC data using the Gompertz function resulted in 90.4% fits with an adjusted $R^2 > 0.70$. For the CH data to which the Beta model was fitted, 99.4% fits had an adjusted $R^2 > 0.70$. Combining the 0.7 limits for both fits, rendered 291 plots that fulfilled both criteria (89.8%), with a median CC_RsqA of 0.95 and CH_RsqA of 0.98 (Figure 4-3).

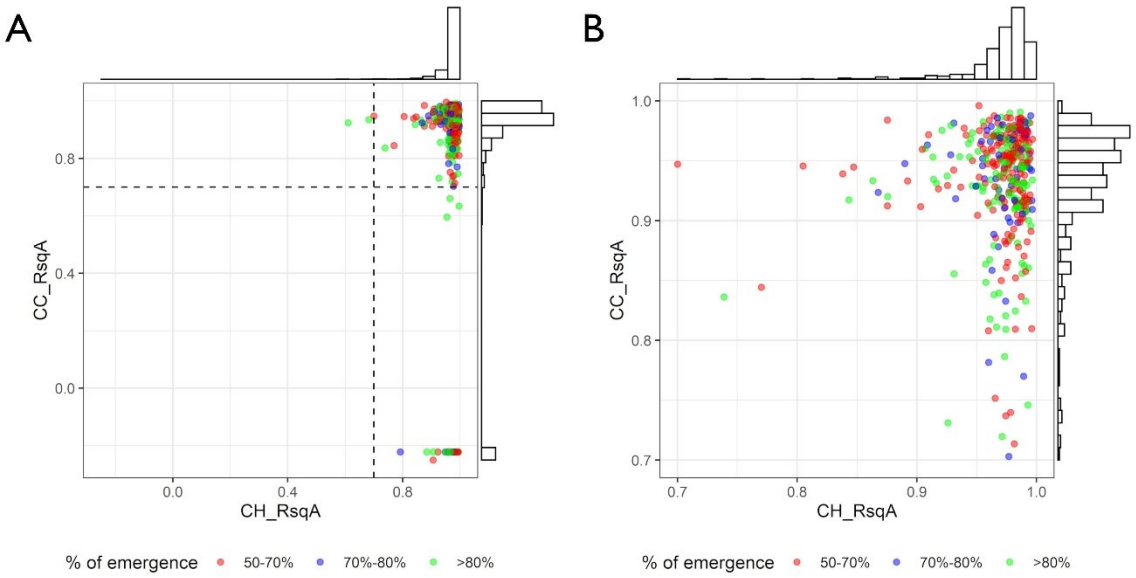


Figure 4-3. (A) Adjusted R^2 of the curve fitting for canopy cover (CC_RsqA) and canopy height (CH_RsqA) data of all plots. A color code is used to show the percentage of emergence. (B) Zoom of the plot distribution with an adjusted R^2 higher than 0.70 (dashed lines in (A), n = 291 plots remaining for further analysis).

The parameter values for CC are presented in Table 4-2. For CC_tm (°C d) the median values ranged from 299 to 446; for CC_Cmax (%) the values ranged from 94.2 to 95.6. On average, CC_tm was lower for the later-sown groups GP3 and GP4 than for the groups GP1 and GP2 (Figure 4-4A). For CH, the parameter values are presented in Table 4-3. For CH_tm (°C d) this resulted in a range from 547 to 735, for CH_Hmax (m) from 0.76 to 0.96 and for CH_te (°C d) from 901 to 1312. To complete the results presented in Figure 4-4; Figure 4-5, we checked whether the differences between groups were significant. For CC_Cmax, the only pair with a significant difference was GP2-GP3. All pairwise comparisons were significant for CC_tm. In the case of CH_Hmax only GP4 showed significant differences with the other three groups. All six pairwise combinations were significant for CH_tm. Almost the same results were found for CH_te, except GP1-GP2, which was not significantly different.

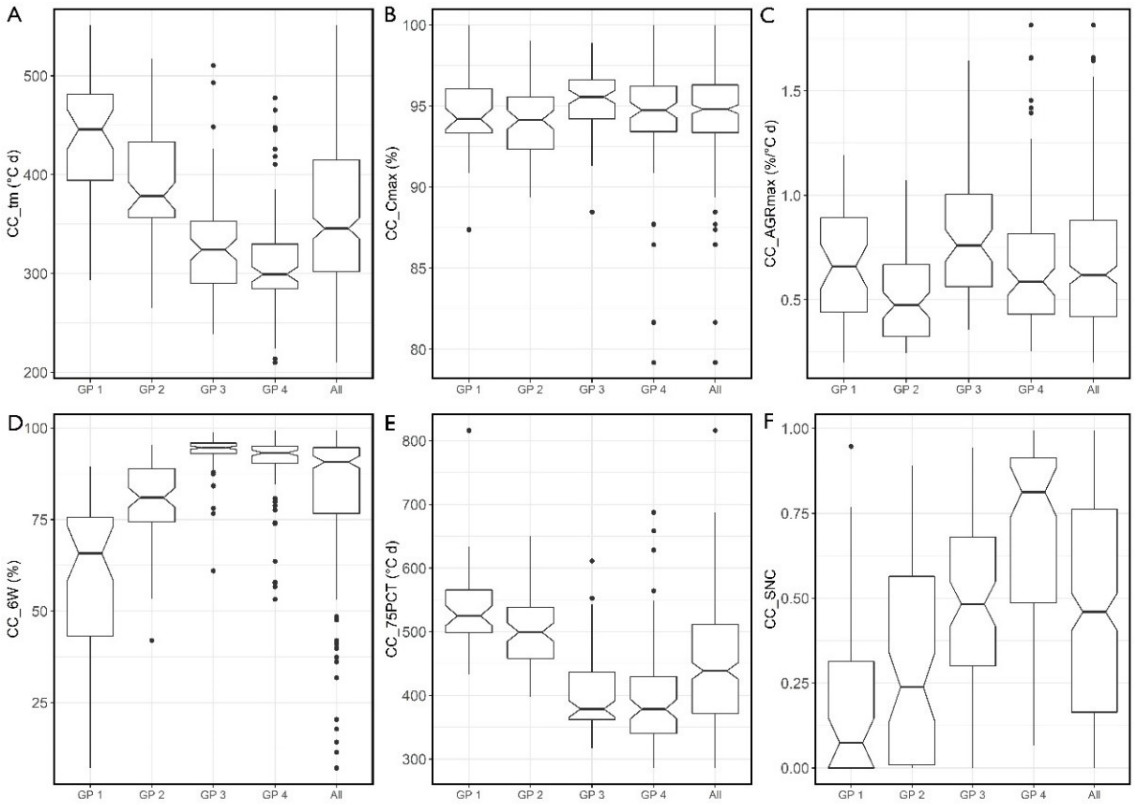


Figure 4-4. (A-F) Boxplots of the phenotypic parameters derived from the curve fitting of canopy cover data. The data is presented per group (GP 1–4) and all groups together (All). For abbreviations see Table 4-1.

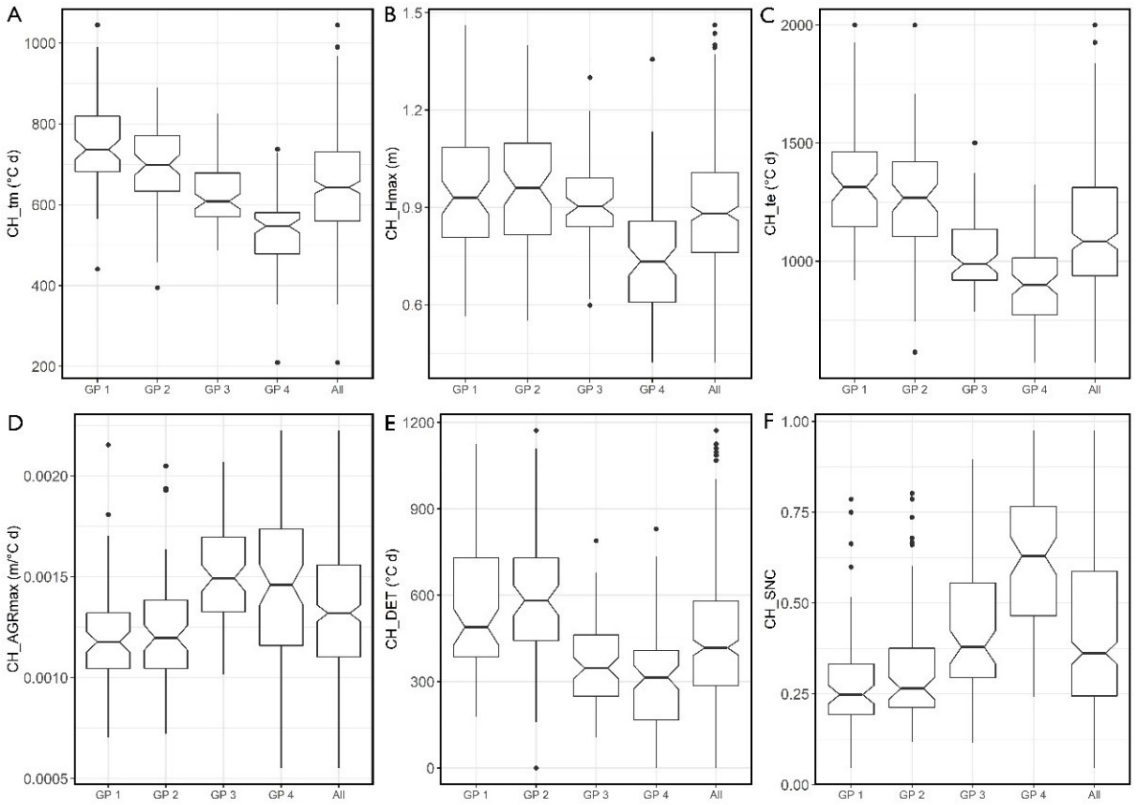


Figure 4-5. (A-F) Notched boxplots of the phenotypic parameters derived from the curve fitting of canopy height data. The data is presented per group (GP 1–4) and all groups together (All). For abbreviations see Table 4-1.

4.3.2 Extra Phenotypic Parameters Derived from the Fitted Curves

As models were fitted to the time series, we could derive phenotypic parameters that correspond to traits that are hard to observe or measure in the field such as growth rate, weed suppression, determinacy or senescence.

The median of the CC_AGRmax (% $°C^{-1}$ d^{-1}) ranged from 0.475 to 0.761 when different GPs were considered. Although the four groups differed in CC_AGRmax (Figure 4-4C) almost all their plots reached high CC_Cmax > 90–95% (Figure 4-4B) as can be expected because standard deviations are <3.1%. Based on the canopy cover growth curve, a value related to weed suppression (Table 4-2 and Figure 4-4D) was derived by estimating the percentage of canopy cover reached six weeks after emergence (CC_6W (%)). The median values varied between 65.8 and 94.6 for the four groups considered. Strikingly, a large spread of values in GP1 and GP2 was observed, indicating a slow canopy development for some of the genotypes contained in these groups. Thermal time values when 75% of canopy coverage was reached (CC_75PCT; °C d) ranged from 372 to 525, meaning that GP1 and GP2 needed more time (were slower) to close canopy. A group effect was also present here, with considerable variation found within each group (Figure 4-4D,E).

The median values for CH_AGRmax (m $°C^{-1}$ d^{-1}) ranged from 0.00118 to 0.00155. The two groups with the highest CH_AGRmax were GP3 and GP4; they reached their maximum height earlier (lower CH_te), but the other two groups showed a higher CH_Hmax (Figure 4-5B–D). The determinacy values are

summarized in Table 4-3 and Figure 4-5E. This variable indicates how long the plants kept growing after reaching R1. The median CH_DET values ranged from 339 to 644 (°C d). Figure 4-5E shows that GP3 and GP4 contained more determinate genotypes than GP1 and GP2, but again with a large variation.

Finally, based both on the canopy cover and the height fits, senescence was estimated (Table 4-2; Table 4-3 and Figure 4-4F and Figure 4-5F). These senescence parameters describe the ratio of decrease for CC and CH between the maximum value derived from the model fitting and the average value of the two last flights. In other words, the higher the values, the more senescence was observed. For CC_SNC median values ranged from 0.074 to 0.815, and in the case of CH_SNC, the median values ranged from 0.245 to 0.639. GP4 showed the highest CC_SNC and CH_SNC values followed by GP3 (Figure 4-4F and Figure 4-5F). Taken together with results for determinacy, we can conclude that GP3 and GP4 contained more determinate genotypes that senesce earlier in the environment tested. The histograms of all parameters are available in the supplementary materials (Figure A-4 and Figure A-5).

The analysis to identify significant differences between groups resulted in the following. In the case of CC_AGRmax, all six pairwise comparisons were significant, except GP1-GP4. For CC_6W and CC_SNC, all pairwise comparisons were significant. For CC_75PCT, all pairwise comparisons significant except for the comparison GP3-GP4. In the case of CH_AGRmax the two pairs that were not significantly different were GP1-GP2 and GP3-GP4. For CH_DET and CH_SNC all comparisons were significant except for GP1-GP2.

Table 4-2. Mean, median, standard deviation (Stdev) and coefficient of variation (cv) of the phenotypic parameters derived from fitting the canopy cover data. Data are presented per group (GP1-4). For the abbreviations, see Table 4-1.

		UAV Parameters					
		Estimated Parameters		Derived Parameters			
		CC_tm (°C d)	CC_Cmax (%)	CC_AGRmax (% °C^{-1} d^{-1})	CC_6W (%)	CC_75PCT (°C d)	CC_SNC (-)
GP1	Mean	390	94.1	0.518	79.5	509	0.315
	Median	446	94.2	0.660	65.8	525	0.074
	Stdev	69	2.6	0.300	21.8	64	0.230
	cv	15.9	2.8	45.3	36.7	11.9	129.5
GP2	Mean	390	94.1	0.518	79.5	509	0.315
	Median	378	94.2	0.475	81.0	499	0.238
	Stdev	51	2.3	0.217	11.5	65	0.300
	cv	13.1	2.4	41.9	14.5	12.8	95.1
GP3	Mean	329	95.3	0.807	93.3	401	0.493
	Median	324	95.6	0.761	94.6	372	0.483
	Stdev	59	1.8	0.293	5.2	57	0.262
	cv	17.9	1.9	36.6	5.5	14.3	53.1
GP4	Mean	306	94.6	0.689	90.6	399	0.715
	Median	299	94.8	0.577	93.4	379	0.815
	Stdev	56	3.1	0.348	9.1	81	0.235
	cv	18.2	3.3	50.4	10.1	20.3	32.9

Table 4-3. Mean, median, standard deviation (Stdev) and coefficient of variation (cv) of the phenotypic parameters derived from fitting the canopy height data. Data are presented per group (GP1-4). For the abbreviations, see Table 4-1.

		UAV Parameters					
		Estimated Parameters			Derived Parameters		
		CH_tm (°C d)	CH_Hmax (m)	CH_te (°C d)	CH_AGRmax (m °C^{-1} d^{-1})	CH_DET (°C d)	CH_SNC (-)
GP1	Mean	753	0.96	1333	0.00118	636	0.274
	Median	735	0.93	1312	0.00118	595	0.245
	Stdev	113	0.21	258	0.00025	293	0.139
	cv	15.1	21.5	19.3	20.9	46.0	50.9
GP2	Mean	695	0.95	1280	0.00121	675	0.336
	Median	699	0.96	1272	0.00119	644	0.261
	Stdev	112	0.18	248	0.00025	262	0.181
	cv	16.1	19.2	19.4	21.0	38.8	53.9
GP3	Mean	627	0.91	1029	0.00151	402	0.427
	Median	609	0.91	989	0.00149	389	0.379
	Stdev	72	0.14	150	0.00026	156	0.173
	cv	11.6	15.0	14.6	16.9	38.7	40.4
GP4	Mean	535	0.77	897	0.00149	340	0.637
	Median	547	0.76	901	0.00155	339	0.639
	Stdev	92	0.18	170	0.00035	181	0.186
	cv	17.2	23.3	19.0	23.5	53.2	29.1

4.3.3 Relations between Phenotypic Parameters

To visualize the relations and complementarity among all phenotypic data, a PCA was carried out (Figure 4-6) and pair-wise correlations were calculated (Figure 4-7). The PCA (for all GP together) explained 81.7% of the variance present using the first four dimensions. Positively correlated variables were CC_75PCT with CC_tm, CH_DET with CH_tm and CH_SNC with CC_SNC. Examples of negatively correlated variables were CC_75PCT and CC_6W (Figure 4-6A and Figure 4-7E). Variables that show an orthogonal relation in the PCA graphs provide complementary data, for example CH_Hmax and CC_6W, indicating that the maximum height achieved is not correlated with the canopy development in early stages. Based on the contributions of the variables to the two first principal components (Figure 4-6A), the most important variables to explain the variability in the data set were CH_tm, CH_te and CC_75PCT for the first principal component and CH_Hmax and CC_AGRmax for the second principal component. The total variance in the PCA is further explained by the CH derived parameters than the CC ones (57.9% versus 42.1%, data not shown). A large overlap between GPs was observed when the plots were plotted in the two first dimensions of the PCA (Figure 4-6B) and similar patterns were observed when correlations were calculated per GP (Figure 4-7), indicating no clear-cut differentiation among GPs.

SY, R1, and R8 values obtained using ground-based methods were plotted on top of the PCA carried out with the phenotypic variables derived from UAV-based data (Figure 4-6, red arrows). SY and R8 are positively correlated, both with each other as well as largely with CH_te, CH_tm, and CH_DET. SY and R8 are negatively correlated with CC_SNC and CH_SNC. This is consistent with the results shown in Figure 4-7. The low correlation of CC_max with all variables was expected, as all plots reached a high maximum canopy cover, resulting in a small variance. Lower correlations between SY, R1, and R8 with all the other variables were observed for GP4 compared to the other groups (Figure 4-7D).

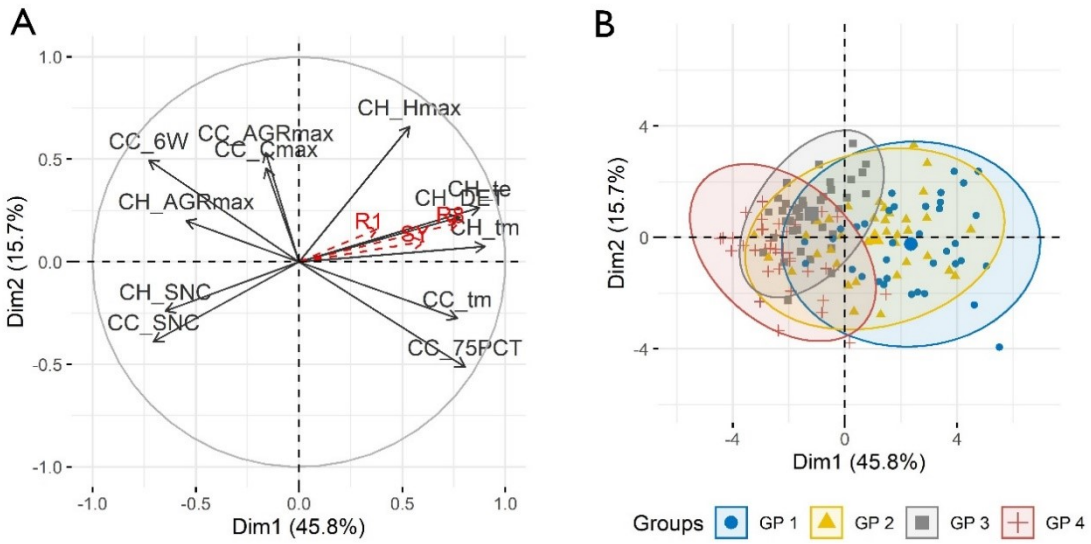


Figure 4-6. PCA of (A) UAV derived phenotypic parameters and field observations (in red) and (B) individual plots (n = 148) with color code related to the group/sowing date. For the abbreviations, see Table 4-1.

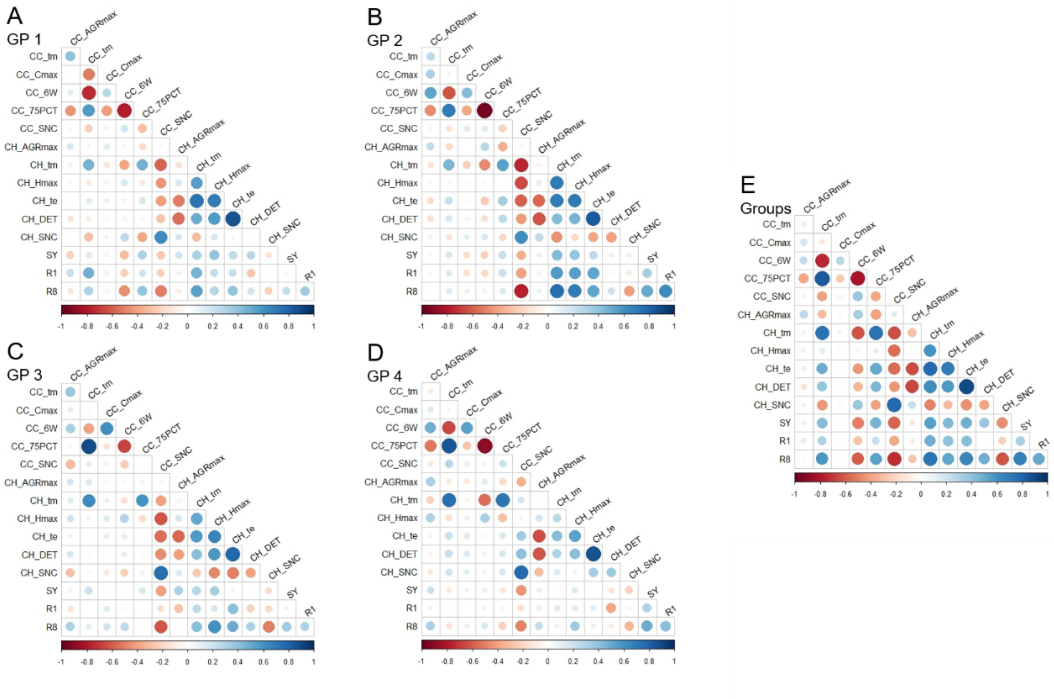


Figure 4-7. Correlograms between the UAV derived phenotypic parameters and field observations (A-D) per group (GP1-4) and (E) grouped (n = 148). For the abbreviations, see Table 4-1.

4.3.4 Linking UAV-Based and Breeder Information

In Table 4-4, the results of the four MLR models built for SY and R8 with different sets of variables are presented. Higher adjusted R^2 value and lower RMSE were obtained for R8 compared to SY. The similar adjusted R^2 between the MLR models based on different datasets indicated that SY and R8 could be estimated with similar performance either using the phenotypic parameters (from curve fitting) or the 14 time points (flights), with a trend towards a better prediction based on the raw data, but resulting in models that are less interpretable.

When UAV phenotypic variables were used in the MLR building, the best result for SY was achieved when using three variables: CC_SNC (senescence ratio), CH_te (thermal time when maximum height is reached) and CH_AGRmax (maximum growth rate) with an RMSE of 4.65 g (9% improvement compared to the basic model). In Figure 4-8A, the RMSE evolution in response to adding a variable is visualized, whereas Figure 4-8B shows the intercept, the coefficients for each GP and the variables included in each model. The red dashed line indicates the best model with three extra variables compared to the basic model. Figure 4-8 and Table 4-4 demonstrate that the models for SY show stable coefficients for the groups, independently of the number of variables added. This indicates a separate effect explained by the dummy variables. The coefficients for CC_SNC, CH_te and CH_AGRmax in the final model have the same magnitude and have consequently comparable importance. The negative coefficient for CC_SNC indicates that an earlier senescence reduces SY. A positive coefficient for CH_te and CH_AGRmax suggests that a later date to achieve the maximum canopy height and a higher AGRmax are linked with a higher SY. In the case of R8, a similar figure is presented and a best model with four phenotypic variables resulted in a RMSE of 63.46 (23% improvement compared to the basic model). The variables included in the model were CC_SNC, CH_Hmax, CH_tm and CC_Cmax. Figures for the other three MLR models are included in the supplementary materials (Figure A-6, Figure A-7 and Figure A-8).

In the case of SY, using the 14 time points, five variables were selected as the best model obtaining a RMSE of 4.58 g (11% improvement compared to the basic model). Four of these variables corresponded to CC data. For R8, eight variables were selected including CC and CH data and RMSE was 56.94 (improvement of 31% compared to the basic model).

In a next step, we explored the information obtained in a breeding context to detect any plots displaying interesting traits. For example, by plotting genotypes/plots in terms of relevant complementary trait values (e.g., CH_Hmax, CC_6W and CC_SNC; maximum height reached, fast soil coverage and the rate of senescence, respectively) (Figure 4-9A). As apparent in Figure 4-9, taller genotypes have a tendency to senesce later. Based on the values obtained, breeders could select a plot with a high CC_6W (good weed suppression), not too tall (low sensitivity to lodging), not too small (strong correlation with seed yield) and a fast senescence. As indicated in Figure 4-9B, plot number 30 fulfils these criteria (plots 110, 131, 177, and 392 are also not too tall and display a good weed suppression, but they senesce later).

Table 4-4. Statistics and coefficients of the normalized variables for the MLRs for seed yield (SY) and R8 stage based on the phenotypic parameters and the raw flight data.

	RMSE	Adj R^2	Int.	GP2	GP3	GP4	Variables Included in Model		
SY	4.650	0.486	17.68	-2.94	-7.55	-6.13	CC_SNC -1.70	CH_te 1.74	CH_AGRmax 1.77
	4.579	0.501	18.31	-3.96	-9.10	-6.09	CC_246 2.05 CC_157 -0.92	CC_228 0.93 CC_199 -0.59	CH_199 1.05
R8	63.463	0.804	1777.51	-112.73	-169.99	-208.54	CC_SNC -44.05 CC_Cmax 10.29	CH_Hmax 14.69	CH_tm 17.60
	56.944	0.842	1790.79	-122.79	-200.07	-221.49	CC_246 43.234 CC_228 11.07 CH_164 -28.45	CH_235 30.626 CC_177 -24.76 CH_177 16.67	CC_206 14.227 CC_171 30.40

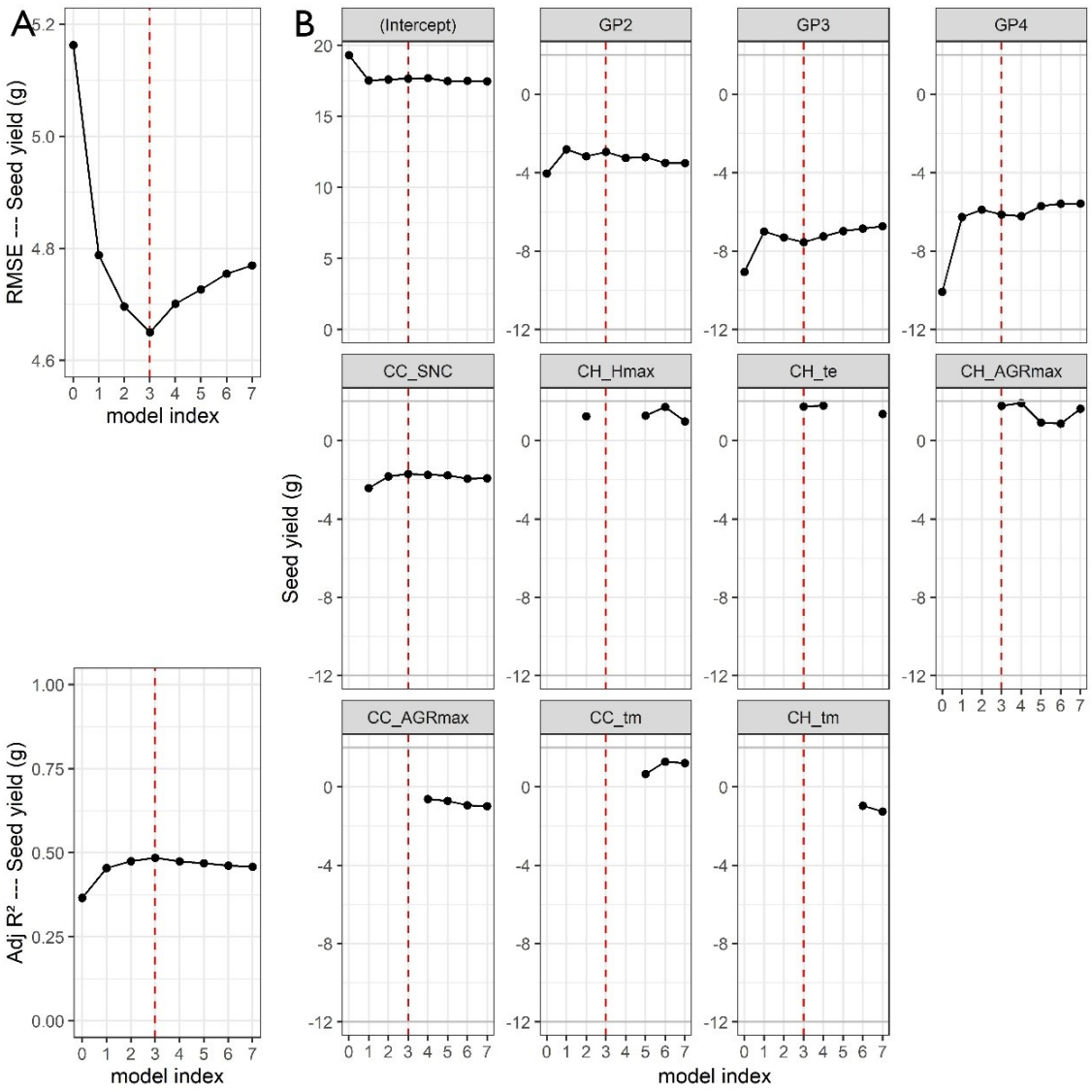


Figure 4-8. (A) MLR model building overview for SY with RMSE and Adj. R^2 evolution. The red dashed line indicates the selected (best) model with the lowest RMSE. The subset selection regression with normalized UAV phenotypic variables included a 10 k-fold cross validation. (B) In each subsequent step one variable was added, and all possible combinations were evaluated. The model index = 0 included only the groups (GP1-4, GP1 is the reference group) as dummy variables. For the abbreviations, see Table 4-1. Graphs are provided per variable with the value of their respective coefficients.

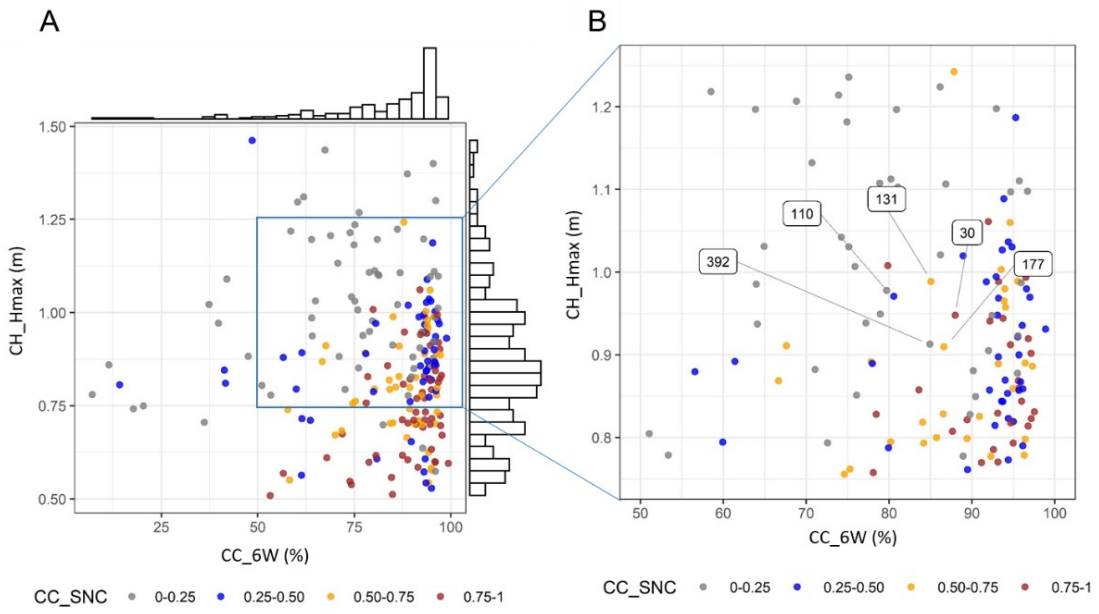


Figure 4-9. (A) Representation of how data could be used to select interesting plots in a breeding program based on phenotypic data using three parameters: CC_6W, CH_Hmax, and CC_SNC. (B) Zoom of the plot distribution based on a applicable values range for selection. For the abbreviations, see Table 4-1.

4.4 Discussion

4.4.1 General Assessment of the Methodology Developed

Weekly flights (n = 14) of a UAV equipped with an RGB sensor allowed to phenotype a large field trial during the entire growing season in terms of the dynamics of canopy cover based on the calculation of the vegetation index Excess Green (ExG), and canopy height based on the Structure from Motion (SfM) photogrammetry with high temporal and spatial resolution. These data were used to fit sigmoid curves (Gompertz and Beta function) in an automated way using R scripts. Model parameters to objectively describe the different soybean plots/genotypes were extracted. The methodology presented here is suitable for the efficient phenotyping of a large number of genotypes under field conditions and can provide breeders with trait information that may be difficult to obtain using other methods, e.g., maximum absolute growth rate, weed suppressive ability, onset and evolution of senescence or determinacy. Large differences were observed among the four groups considered, and high variability was present within the groups (~genotypes). The full canopy cover was reached starting from day of the year (DOY) 165 (mid-June). A decrease in cover and height was observed starting from DOY 200 (mid/end July) and was related to the onset of senescence.

By fitting a sigmoid growth function such as the Gompertz or the Beta function to multi-temporal data, it is possible to determine reliable parameters with a straightforward biological meaning enabling characterization of between-plot differences in growth and developmental processes (Yin et al., 2003). An extra advantage of curve fitting to multi-temporal data is that possible small errors at one time point are smoothed out using the entire time series. However, model choice is critical because use of the

wrong model can introduce bias (Thorney, 1990). The success rate reached for model fitting was 91.0% for canopy cover and 99.6% for canopy height, hence the models chosen were appropriate for fitting these data. The majority of the cases that did not reach a good fit, originated from CC data mainly from genotypes of GP1, the group first sown. Thus, a bad fit was most likely a combination of having only two data points in the increasing phase of the Gompertz model fit for CC data, the first data point having already a canopy cover of >15%, and the reduced flexibility of the Gompertz function to properly fit the first time points. These first time points are important for the model fit, as these present bigger differences because of rapid growth. This suggests that for UAV based phenotyping using time series, flights should start early enough, at plant emergence, and should last long enough to enable full senescence of all treatments. This sometimes conflicts with traditional experimental methods. In this case, bird netting was used to cover the field trial at the time of emergence to prevent bird damage. This of course interferes with UAV imaging, requiring removal and replacement of large areas of netting at the starting phase of the experiment. Near the end of the experiment, full senescence is not always required in a breeding trial before harvests are performed. Therefore, before starting the experiment, these issues should be discussed carefully between the breeders and the phenotyping researchers. A gap in the dataset can happen too—in this study, no flights were possible between July 25 and August 16. This was solved by the curve fitting, as the values during the two weeks gap were estimated. On July 25 the canopy was fully closed, and on August 16 senescence has been initiated only for some plots. Potentially, this could affect the parameters CC_SNC and CH_SNC, but not all other parameters, as full canopy and maximum height were reached before that gap. Scripting the models in R proved appropriate to automatize the fitting step for large sets of plots. Nevertheless, care should be taken to define appropriate initialization values for the Gompertz model as an optimization was not always achieved with the optimx function in R. In our case different values needed to be tested. Regarding canopy height, some datasets contained close to zero or even negative values for some plots in the first and the second flight. This can be explained by the intrinsic altimetric error and the difficulty of detecting small objects in the photogrammetry process of a flight (see chapter 3). This “noise” cannot be corrected or removed, but is in our opinion acceptable (x-y-z average error is 0.07 m).

4.4.2 Describing General Growth Patterns in the Different Groups

Using the derived and estimated phenotypic parameters (Figure 4-4; Figure 4-5), it was possible to detect and quantify differences among the groups. A relevant trait is the canopy cover at six weeks after emergence (CC_6W) as this is related to weed suppression capacity and row closure. In this study, most of the plots of GP3 and GP4 had CC_6W values >90% and closed the row faster than the other two groups. Senescence was also different among groups, with fast senescence for GP4. For CH_AGRmax comparable values were observed for GP3 and GP4 and for GP1 and GP2, but at different thermal times (CH_tm). By combining the results obtained for CH_Hmax and CH_te, we can conclude that plots of GP4 did not grow tall (relatively lower CH_Hmax) but grew faster (lower CH_te in combination with higher CH_AGRmax) than plots in other groups. This confirms our expectations, as GP4 comprised genotypes expected to develop quickly and mature early, according to the information available beforehand. We can thus conclude that the pipeline developed to characterize the seasonal growth and development of soybean is functional and suitable for implementation in real-time breeding applications.

Determinacy describes the extent to which a plant continues to produce extra internodes and consequently extra possibilities to produce pods and seeds after entering the generative phase (R1). Indeterminacy or semi-determinacy can be a favorable trait in soybean for cultivation in Northwest Europe as it represents an 'insurance policy' if a period of unfavorable weather conditions occurs, such as low temperature during flowering or drought in the summer months. Based on experience, determinate genotypes perform less well under Northwestern European conditions (Aper et al., 2016; Schori et al., 2003). In contrast, a high level of determinacy contributes to a synchronous seed maturation, avoiding variable moisture content in the harvested seed that could cause quality problems. In soybean, determinate cultivars are used in areas with long growing seasons while indeterminate cultivars are more suited to short growing seasons (Ping et al., 2014). Detailed information on this trait is therefore very useful for soybean breeders. Here, we present a method to estimate determinacy in a straightforward way that could be implemented routinely in soybean breeding. Overall, genotypes included in groups 1 and 2 had a higher level of indeterminacy and kept growing during a longer thermal time.

4.4.3 Estimating SY and Maturity Using UAV-Derived Information

In Northwest Europe, early maturity (quick progress to R8) in combination with high seed yield and a high protein content are important breeding targets. We evaluated if seed yield (SY) and the R8 stage could be predicted by an MLR with the individual time series data points or with the estimated and derived phenotypic parameters as variables. Similar adjusted R^2 values were obtained using the data of the different flights and using the derived phenotypic parameters. This demonstrates that the phenotypic parameters captured all variance present. Seed yield is influenced by canopy cover because a longer lasting canopy cover most likely results in more photosynthesis for optimal SY at the end of the growing season. This is supported by the presence of CC_SNC in the first MLR model and because four out of five time points included were CC data for the second model. Canopy height also plays a role in terms of growth rate and the moment when maximum height is reached. An early senescence decreases SY and the later the maximum height is reached, the higher the chances for a higher SY.

Maturity (R8) was determined by more CC time points (five dates, DOY range 164–246) than CH time points (three dates were included). The phenotypic variable CC_SNC has the most influence on R8 followed by CH_tm. Similar results related to maturity prediction classification (93% accuracy) were achieved in Yu et al. 2016 using RGB and NIR drone data. In the present study we obtained similar results using higher resolution RGB imagery but by using simpler statistical methods (MLR). Soybean agronomic traits such as yield, maturity, and height were predicted in Yuan et al. (2019) using six flights over four locations by color and texture features. Resulting R^2 values were 0.68, 0.76, and 0.27, respectively, using five regression modeling techniques, i.e., Partial Least Squares Regression, Random Forests, Cubist, Artificial Neural Networks, and Support Vector Regression. In our case, seed yield and R8 estimated by MLR achieved an adjusted R^2 of 0.51 and 0.85, respectively, which is slightly lower for SY but higher for maturity.

4.4.4 Perspectives for Practical Applications

The UAV-estimated and derived phenotypic data can be used directly in a breeding context, as they are related to relevant traits, e.g., weed suppression capacity, fast canopy establishment, early vigor, growth rates, maximum plant height, risk for lodging, determinacy, etc., for each genotype. For example, maximum plant height contributes negatively to lodging resistance, while it is positively correlated with the number of internodes, which in turn is positively correlated with the number of pods. Therefore, breeders should find a compromise considering different traits. For example, the combination of CC_6W, CH_Hmax, and CC_SNC makes it possible to identify genotypes that combine appropriate values for the different traits.

Besides the use in a breeding context, multi-temporal data can be highly valuable for plant and crop modeling. Time series of different traits have been found to be essential for proper and identifiable parameter estimation (De Swaef et al., 2019). Large amounts of data can be gathered using UAV with a high temporal resolution.

4.4.5 Practical Lessons Learned and Future Perspectives

The high spatial resolution achieved from the UAV flights (0.5 cm pixel size for the orthophotos and 1.25 cm for the DEM) revealed the variability present within the plot and yielded a higher number of measurements in comparison to manual data-gathering. High resolution imagery does require a longer processing time, however. To reduce the processing time, a lower spatial resolution could be defined if the aim of the study does not require highly detailed information. Future research could investigate which resolution is sufficient without influencing the predictability of traits.

In this study, we considered 14 flights spread over the entire growing season at regular intervals. This time interval was short in comparison to similar studies, which included an average of only four to 10 dates (e.g., Han et al. (2019); Holman et al. (2016)). During data processing and interpretation, we noticed that a different flight scheme may have been more appropriate. For example, in some cases, as senescence occurred in a short period of time, it could have been appropriate to add to extra flights at the end of the growing season. Similarly, when rapid changes (e.g., fast growth) are expected, extra time points are welcome to ensure a good fit. In this regard, the date of the first flight should also be carefully chosen to capture the entire growing season. In conclusion, planning the flights according to the expected development of the crop may yield more useful data than regular intervals between flights.

4.5 Conclusions

In this study we collected UAV data of a soybean field trial at an unprecedented temporal resolution. The time series data allowed us to fit growth curves with high accuracy (>90%) and to derive relevant traits. The trait data that became available in this study by combining multi-temporal UAV with curve fitting will allow for a better biological interpretation of the differential response of a large set of soybean genotypes. In addition, we developed a methodology to predict seed yield and to estimate the moment at which each accession reached the R8 stage.

Furthermore, this developed strategy may be applied in the future to the analysis of multi-temporal data obtained using other sensors, e.g., multispectral, hyperspectral and thermal, to evaluate responses to abiotic and biotic stresses. This would enable the development of more resilient plant varieties at a faster pace.

Chapter 5: General discussion

High-throughput field phenotyping of large trials is key to increase crop resilience and achieve a more sustainable crop production towards the upcoming climate change. This should result in the development of new varieties (via breeding) and also in an environment-friendly crop management and production systems. Together with the evolutions in crop genomics this could enable the next generation of crop research and breeding (Furbank et al., 2019). Indeed, a good understanding of biological processes linked to growth and development is relevant in biological research too to be able to predict the phenotype based on the interaction between genotype, environment and the management applied (G*E*M concept). The "phenotyping gap or bottleneck" or the lack of quantitative and high throughput plant phenotype data, especially in relevant environments (i.e. the farmers' and breeders' fields), became more and more obvious compared to the fast developments in genotyping and sequencing technology (Araus et al., 2018; Junker et al., 2015) at the start of this research.

Therefore, this PhD dissertation has been looking into ways to support breeding activities and close the phenotyping gap (a) by investigating the potential of UAV-based RGB imagery to provide high throughput phenotypic data in field conditions, (b) by developing simple and accessible methods to facilitate the use of UAV imagery in a breeding context providing relevant phenotypic data/traits, (c) by validating the derived phenotypic data against current standard measurement techniques, (d) by proving the applicability of the methods developed, and (e) by investigating the use of modelling approaches to extract additional traits derived from multi-temporal data. For this investigation, an RGB sensor was chosen, because (i) a high spatial resolution can be achieved avoiding mixed pixels compared to lower resolution sensors, (ii) orthomosaics and detailed height models can be constructed based on the images taken, and (iii) the technology is highly accessible thanks to the lower cost of RGB sensors.

5.1 Technical considerations

5.1.1 UAV-RGB as a phenotyping instrument

UAV-based remote sensing brings a full range of new applications supported by the high numbers of publications and reviews available (Yang et al., 2017). More specific for high throughput phenotyping, UAV data can complement or even replace traditional methods to evaluate and monitor field trials. It allows to monitor plant traits in space and time, increase the number of samplings, and depict inter- and intra-plot variability in a non-destructive way. Via the assessment of traits in field conditions and a better understanding of the processes, breeding programs could better adapt crops to biotic and abiotic stresses through selection for resilience to climate change.

5.1.2 Extraction and interpretation of data from UAV-RGB images

Canopy cover (CC) is the ratio of the projected plant area to the total area considered. CC is dynamic, changing from crop establishment to senescence, and can be linked to seedling emergence, early vigor and senescence (Walter et al., 2015). CC can also be used to evaluate crop responses to drought (Sayed et al., 2012). A rapid establishment and full coverage at early stages enhance solar radiation interception

associated to growth, yield, weed suppression and decreases soil evaporation (Edwards and Purcell, 2005; Jannink et al., 2001). We observed that in the case of soybean CC was also related to seed yield. A longer lasting canopy at the end of the growing season means a longer photosynthetic activity for seed production. In this work, we have estimated CC using a vegetation index (VI), excess green, which enhances green and can be used to differentiate vegetation from soil. This VI can also be used to build a mask and only consider the pixels that correspond to vegetation in the data extraction. Using VIs is a way to transform the spectral values into more interpretable data linked to traits. Although these require low computational cost, their limitations cannot be neglected such as saturation beyond a certain biomass threshold (Royo et al., 2003; Thenkabail et al., 2002) and the fact that they should be tested and validated for each application developed (Togeiro de Alckmin et al., 2020).

Plant height and density are related to light capture and ultimately to yield (Bendig et al., 2014). The estimation of plant/canopy height using UAV-RGB images has been demonstrated repeatedly in the past for several crops (e.g. Bendig et al., 2015; Geipel et al., 2014). Here, we have evaluated the potential to measure canopy height in grass plots ranging in height from 10 to 40 cm, which is significantly lower than the height of crops in which this methodology was demonstrated in the past, such as maize or wheat. Essential in this case was attaining a level of resolution at cm level to be able to differentiate the performance of different grass plots. In this dissertation, height was measured throughout the growing season. With an RGB sensor mounted on a UAV, we obtained detailed imagery with a spatial resolution of 0.5 cm for orthophotos and 1.25 cm for DEM. To be able to measure crop height in an accurate way via UAV imagery, based on Structure from Motion (SfM) photogrammetry, is a great improvement. Nevertheless, it requires a sensor with a sufficiently high number of pixels (e.g. 24Mpix) and a sufficient front and size overlap in the flight lines (Torres-Sánchez et al., 2018b). The relation between sensor resolution, sensor size, focal length of the lens used and flight altitude will determine the spatial resolution. As well, canopy structure (closed or open canopies) and the algorithms used in the stitching software determine the results (Roth et al., 2018b; Torres-Sánchez et al., 2018b). In our cases, the flight plan was correctly design as a high quality 3D model and reconstruction of the canopy was achieved. RGB sensors with a very high resolution (e.g. 42 Mpix) are now commercially available, but have the disadvantage of increasing the data to stitch dramatically, thus increasing the processing time or even preventing the stitching due to lack of sufficient processing resources.

Strikingly, none of the studies on the estimation of crop height available before we initiated this research reported on the use of artificial references to calculate the intrinsic error of the methodology developed. The evaluation carried out in Chapter 3 using references at ground level showed a higher error for the flights done before a cut (when the canopy was well developed) and some ground control points were partially covered. It is important to ensure their visibility in a sufficient number of images to achieve a high quality DEM. We achieved a high precision in the reconstruction of the artificial height references as the average variance was limited. Studies where passive sensors were compared to active sensors have shown that height measurements were similar. Height obtained from RGB images through SfM was highly correlated to LiDAR measurements (r^2 = 0.97 (Madec et al., 2017), r^2 = 0.6-0.77 (Malambo et al., 2018), r^2 = 0.65-0.99 (Ziliani et al., 2018)). When the height measurements from RGB imagery were less accurate this could mainly be attributed to the lower spatial resolution and the limited penetration capacity, and the limitations of SfM in reconstructing the top of the canopy (Bareth et al., 2016). In our

case, the use of an RGB sensor to measure height was found to be suitable, while it requires simpler processing than LiDAR.

When determining the usefulness of a new methodology, it is important to compare it to standard methodologies. For example, a comparison with a traditional height measurement procedure in ryegrass, rising plate meter (RPM), is presented in Chapter 3. The average correlation of the height measurements between the UAV and RPM was 0.92, similar to the values reported by Bareth and Schellberg (2018) and Viljanen et al. (2018). The difference observed between both methods was attributed to the gentle compression of the canopy by the RPM, which will be more or less pronounced depending on the moment of the year, as the canopy density varies throughout the year as we observed in Chapter 3 (Figure 3-5). The relation between the RPM and canopy height can be limited by plant development stages, canopy architecture and density (Nakagami, 2016). In addition, as discussed by Viljanen et al. (2018), the rising plate meter presents limitations in non-uniform, tall or sparse swards. Furthermore, we have tested the RPM for consistency by measuring at repeated occasions the same perennial ryegrass plot and this resulted in an r^2 of only 0.76 between the measurements. Of course, the UAV method is not free from constraints either. For example, if the plants are really small or not covered by enough pixels and images, the reconstruction of the volume would present difficulties and be incomplete. Therefore, the flight planning is key to capture high quality imagery.

Measuring the height of hundreds of field plots is time-consuming when done manually. We performed weekly flights to follow up the evolution of the height in soybean of more than 450 plots and in ryegrass of more than 1350 plots which is hardly feasible manually. In addition, the resolution of the UAV products provided more detailed height data than field measurements, e.g. using a RPM or ruler. The high spatial resolution attained made it possible to study the variability within the plot, providing detailed data about the canopy uniformity. This would have been practically impossible using an RPM, as each measurement would provide an average value for an area of 30x30 cm at a limited number of points in the canopy. Height, plant density and canopy structure would then be averaged using RPM data, resulting in low power to study intra-field variability compared to UAV imagery. To generate reliable data with the RPM, a substantial number of repetitions combined with a spatially uniform distribution would be needed which requires time and effort for large trials (Lu, 2006).

5.1.3 Modeling crop growth using multi-temporal UAV data

UAVs allow to capture imagery at crucial moments or at multiple moments throughout the growing season (temporal resolution) to derive information about a specific trait or to monitor crop development. In this dissertation, depending on the purpose, we carried out flights at specific time point to estimate a trait at crucial moments or we flew weekly to track the development of canopy cover and height throughout the growing season. In the case of soybean with the weekly flights (multi-temporal data), we monitored the growth and development of more than 450 plots. The acquisition of data with such a high frequency allowed to fit growth curves to study the crop development and derive relevant traits. As discussed in Chapter 4, the parameters derived from multi-temporal imagery, through multiple linear regression models or curve fitting, provided extra information on relevant traits which are difficult to obtain for such a high number of plots using other approaches (early vigor, growth rate, the onset of senescence or the time at which a specific phenological stage is reached). A great advantage is that the

phenotypic variables are biologically interpretable and can be used to support breeders' decisions. For example, they allow to compare genotypes in terms of early vigor and weed suppression capacity, important traits related to establishment and canopy closure rate.

5.2 Practical implementations

5.2.1 Persistency and dry matter yield in perennial ryegrass

Vegetation indices have been used for diverse applications and crops. In this dissertation, vegetation indices derived from RGB drone imagery were used to determine ground cover as a proxy for persistency in ryegrass. To prove the applicability of different indices, they were tested in terms of changes in light conditions, data consistency and reproducibility when measured at different moments. A significant higher F-value and lower CV were obtained for the VIs compared to the visual persistency scores, which indicates a better differentiation between accessions to carry out the selection based on this trait. Visual assessments are relatively cheap, but are prone to subjectivity of the operator and can perform poorly when only subtle differences among genotypes are present or if illumination conditions fluctuate. In addition, scores are limited by what a human can visualize at ground level (Bai et al., 2018; Yu et al., 2016). Potentially, access to accurate characterization of the genetic value will result in a higher genetic gain per breeding cycle (Cobb et al., 2013).

Moreover, we have also developed a methodology to estimate the dry matter yield from RGB drone imagery. Different models to estimate DMY based on UAV and RPM height data were compared for ryegrass. It was concluded that including environmental data provided a great improvement and all models built using UAV data obtained a lower RMSE than the one based on RPM data. This was to be expected as clear dynamics in yield are present depending on the season in forage crops (Behaeghe, T.J. 1979). The best model achieved an improvement of 31% with an RMSE = 679 kg/ha (NRMSE = 21.3%) compared to the model based on the rising plate meter. Our findings are comparable to those reported by Grüner et al. (2019) and Lussem et al. (2019). In a recent article, Togeiro de Alckmin et al. (2020) achieved a lower RMSE (= 430 kg/ha) for DMY estimation using vegetation indices derived from spectroradiometer measurements. They used non-parametric models (such as random forest or support vector machine) with four VIs from the red-edge and NIR regions. They concluded that part of the error might be due to the destructive biomass samples and their drying process, an aspect that we did not consider in our study but that requires further attention, and that model selection and tuning are key. However, only 30 plots were monitored in this study and their measurements took each time around 2h. Comparison of the results of our study with those of Togeiro de Alckmin et al. (2020) is however difficult. Our best model resulted in a higher RMSE but this might be due to the high diversity in genotypes included in our trial, with differential responses to the environmental conditions and differing heading date (versus one genotype in that other study). A higher number of destructive samples could be a way to improve the performance of our best model. Alternatively, a specific model could be developed per group. Based on the results obtained, the variability present in our trials and the comparison with other studies, we agree with Schellberg et al. (2008) and Schellberg and Verbruggen (2014) that most approaches based on UAV-imagery have been developed to monitor arable crops (where spatial heterogeneity is lower). And it is important to understand the

spatiotemporal dynamics of grasslands to adapt management decisions as they will differ depending on the environmental conditions. We agree with Lussem et al. (2019) that challenges remain to achieve a global model for biomass estimation, because the heterogeneity in structure and spatiotemporal variability and the predictors/parameters choice and the technique used to build the model led to varying performance of the models.

5.2.2 Growth and development in soybean

The combination of relevant traits provides a complete view on the performance of each genotype and enables selection by the breeders of the most appropriate ones to achieve crop adaptation or to use them as future parents. Via flights done at a high frequency (e.g. weekly) from emergence to full senescence, additional crop development data are available for selection. As we can measure maximum height and estimate when it is reached using curve fitting, comparisons related to rates of change (fast or slow development) become available.

The model choice for the growth curve fitting is important. Our success rate was higher than 90% for both canopy cover and canopy height, so the Gompertz and Beta function were adequate choices. Nevertheless, there were some unsuccessful cases in the group sown first as not enough points were available to cover the increasing phase (fast) development at the beginning of the growing season making it difficult to fit the curve. A higher number of flights at that crucial period of time could be beneficial to increase the success rate of the fitting, which will be taken into account in future fight planning.

The R8 stage which corresponds to maturity is traditionally scored visually. However, this requires frequent revisits to the trial as not all genotypes reach that stage the same day. As many genotypes need evaluation, to score all of them is therefore time consuming. We estimated when R8 was reached via a MLR model obtaining an adjusted R^2 of 0.80 and we found that the two main parameters that are related to R8 are canopy cover senescence and the moment when height maximum growth rate happens. Maturity is related to harvest and it is relevant to know when it is reached precisely for a large number of plots to depict the early maturity genotypes, which are the most suitable to our target environment. Five to ten days of good drying weather are needed after reaching maturity to harvest soybean. From mid-October, maturation becomes more difficult if there are insufficient drying hours and mold can infect the seeds (Aper et al., 2016). To measure seed yield is also quite time consuming as it is done manually and only a few plants are harvested per plot. We also attempted to estimate SY obtaining an adjusted R^2 of 0. 49 (medium accuracy). Nonetheless, via the model we got insight in some main parameters that determine seed yield such as canopy cover senescence and the moment when maximum height is achieved. This extra information can be taken into account by the breeders in future breeding selections to improve seed yield in an indirect manner.

5.3 Lessons learned and recommendations for future use of the results

5.3.1 In breeding

When a UAV mission is planned, several aspects need to be considered. An important parameter in the flight planning is the overlap between images and lines (longitudinal-lap and side-lap) and coverage of the entire trial (≥ 75% overlap) to ensure that the canopy is captured in enough images and at different viewing angles (perspectives) to build its volume during the photogrammetry process. Another key parameter is the (acceptable) error obtained in our height quality evaluation, which is partly explained by the partial covering of the references by the surrounding taller canopy (grass in our study) which reduced visibility of the objects. Therefore, the location, visibility and number of references are important to obtain high quality georeferenced orthophotos and have to be taken into account in the flight planning. A flight planning tool which ensures imagery acquisition with sufficient quality using an UAV was presented in Roth et al. (2018b).

To facilitate the adoption and the use of the methodologies by the breeders, versatility and cost to screen large collections are crucial. The costs associated to the implementation of the methodologies developed in this work (in comparison to standard methods) were not investigated in detail. There are some first estimations of the costs taking into account not only the equipment needed for the data collection, but also the processing and analysis steps to obtain phenotypic data. Reynolds et al. (2019) evaluated the share of data analysis to be up to 20% of the total costs. Nevertheless, the costs are highly variable due to the varied options available in terms of UAV platform, sensors and processing software. In his PhD dissertation, F. Holman (2020) performed a more detailed and specific evaluation of the costs and comparison between manual and UAV methods, resulting in a 40% cost saving using an UAV (11.40 €/plot vs 6.9 €/plot). Of course, the effective cost saving will depend on the size and the number of plots to evaluate in the trials, with a higher expected saving for larger trials. The calculation becomes even more complicated if aspects such as possible increases in genetic gain are also considered. It is therefore clear that a honest comparison of the reward that can be attained through the implementation of HTFP methods will only be possible within some years, once these methodologies are implemented broadly in practical breeding programs.

The comparisons carried out by Holman et al. (2016) and White et al. (2012) indicated that measuring traits manually (in the field or laboratory) is more time-consuming than UAV data collection. However, the data collection is not sufficient, additional steps such as processing and analysis are needed to produce phenotypic data and need to be taken into account. Manual methods imply slow data collection and minimal data processing times compared to the reduced data collection time and high processing requirements of the UAV imagery. The asset of the imagery is the option to semi-automate the processing workflow to 24/7. Holman (2020) compared both methods in terms of time and concluded that the manual methodology takes three times longer. The processing time depends on the number of images and their overlap (rather than the area covered) plus the time needed for the analysis to extract the data (e.g. to obtain maps or traits). It is evident that a trade-off exists between throughput in data acquisition and precision. If required for fast decision making, the UAV mission can be modified to reduce the processing time (e.g. flying at a higher altitude or reducing overlap). However, the final

resolution and quality will be affected. With this premise in mind, the most efficient and effective method for field phenotyping towards measuring the target trait has to be chosen (Murchie et al., 2018).

There are multiple options in terms of sensors and platforms to be used and implemented in breeding programs. Depending on the phenotyping requirements a trade-off has to be made between spatial resolution, cost and processing time, among others. The methodologies developed in this manuscript are based on UAV imagery as it allows to cover a high number of plots to acquire data at a high speed (reduced period of time) and high spatial resolution. The greatest advantage here is the fast data acquisition for a high number of plots. Therefore, the methodologies presented here could be used in the row and plot trials of a breeding program when a high number of plots have to be evaluated. Moreover, UAVs could be used along the entire breeding program as the spatial resolution is very high (up to 2 mm is possible) and could even be used to monitor individual plants in the first years.

5.3.2 Link with precision agriculture

Remote sensing tools are used for different purposes in phenotyping and precision agriculture. In phenotyping, the aim is to carry out a selection based on detailed data (traits such as yield, quality or plant development) acquired and analyzed in a (semi-)automated way. In the case of precision agriculture, the aim is to provide decision support based on imagery at a low cost, e.g. to maximize yield, detect deficient areas or prevent the spread of a disease. The knowledge acquired in phenotyping research such as traits that are linked to yield, or vegetation indices that inform about e.g. health status can be transferred and used in precision agriculture, where in most applications a lower spatial resolution is needed. In this case, it is enough to be able to quantify the spatial and temporal variation in the agricultural areas (variation present in the field and its evolution) at the relevant scale. Based on the method developed for persistency evaluation, areas with lower grass coverage can be detected to remedy the situation as fast as possible by reseeding to avoid weed infestation and yield losses. Moreover, a targeted solution applied to a specific area and not to the entire field saves costs. Transferring methodologies and technological progress from research to practice is crucial. For a widespread use of UAV-based remote sensing in agriculture applications, the requirements are the same as for breeding: low-cost, easily adaptable, user-friendliness (simplicity) and proven valuable. Not all agricultural fields are extensive areas, smallholder farmers grow different crops resulting in a heterogeneous land, which is difficult to monitor with satellite remote sensing, while UAVs could be more suitable for the task (Cucho-Padin et al., 2019). An appropriate sensor for agricultural practices could be a multispectral sensor embedded in a UAV. It would be a flexible system with a sufficient time and spatial resolution. Moreover, the imagery calibration has become more accessible for the users (Togeiro de Alckmin et al., 2020). Multispectral UAV data could also be used to calibrate Sentinel-2 data imagery for the coverage of large areas, as was demonstrated for the quantification of damages in rice (Pla et al., 2019).

Maes and Steppe (2019) reviewed the progress of the use of UAVs in precision agriculture for applications such as weed detection or nutrient status, among others. They identified complementarity and fusion of data from different sensor types and the development of growth models as the next steps to be taken in this field. In chapter 4 we have contributed to the second step identified via the curve fitting. Recently, Segarra et al. (2020) reviewed the advancements and Sentinel-2 applications in

precision agriculture. They indicated that monitoring for biotic stresses (e.g. pests) is difficult due to the pixel resolution as damage is only sensed when it is significant. Hence, combination with data derived from ground-based or UAV platforms might be the way to go. They concluded that as Sentinel-2 imagery is publicly available and accessible, precision agriculture can benefit from it and a more productive and sustainable crop production could be achieved.

5.4 Conclusions

Regarding the questions raised in chapter 1, we may conclude for the first research question (RQ1) that the methodology tested and validated in chapter 2 delivers robust and objective data to evaluate persistency and can be implemented in breeding programs. The method based on the RGB spectral information in the form of vegetation indices provided consistent results even for flights carried out at different moments, obtained a high correlation with the ground based imagery used as ground-truth and differentiates better between the genotypes. The second research question (RQ2) was related to the accuracy of canopy height measurements derived from UAV-RGB imagery. The precision of UAV-RGB to estimate height of (less tall) canopies was tested in chapter 3 using artificial height references. At ground level, the average error for height ranged from 0.94 to 1.89 cm, and using references from 5 to 40 cm resulted in an average underestimation of the height with 0.56 cm. Canopy height derived from UAV imagery for ryegrass was compared to the traditional method, a rising plate meter (RPM) obtaining a correlation of 0.92. For these reasons, the height measurement could be replaced by RGB imagery getting more detailed data and catching the variability present in the field at a higher resolution. Moreover, RGB extracted data (vegetation indices, canopy cover and height data) was used to estimate dry matter yield in grass, achieving a 30% improvement (r^2 = 0.81) compared to the traditional RPM method.

In this dissertation, two crops were monitored in terms of canopy cover and height throughout the growing season in a remote and non-destructive manner with an RGB sensor. For soybean, the evolution of the crop was captured with high spatial and temporal resolution, i.e. with detailed information within plots and in a high number of time points. As demonstrated in chapter 3, UAV can be used to replace ground based measurements as it allows to measure more than 450 plots in a short period of time. The third research question (RQ3) was related to the prediction of seed yield. An approach for soybean seed yield estimation was carried out resulting in a medium accuracy (adjusted R^2 = 0.49), although we obtained information about the parameters that determine it. The great advantage of UAV-RGB imagery at that high resolution and fitting curves, as done in chapter 4, is the provision of additional phenotypic parameters linked to growth rate, senescence, determinacy, phenological stage and early vigor. These parameters are difficult to estimate in the field, but are informative and relevant for selection as they can be used to (biologically) interpret and compare the responses of a large set of genotypes. As we demonstrated the advantages of a multi-temporal RGB dataset, also the fourth research question (RQ4) can be answered positively. In addition, the strategy developed in chapter 4 is transferable to other sensors such as multispectral cameras.

5.5 Future developments and perspectives

The focus of this PhD research was on the evaluation, development and validation of methodologies to produce accurate phenotypic data from an RGB sensor embedded in an UAV in two model crops derived for use in biological understanding and breeding. Despite the great progress and the high number of applications of UAVs to phenotype crops in a high throughput manner, research gaps remain to be addressed. The lack of standard protocols, the actual use in a "real" breeding context, or the link with important research fields such as genomics for association mapping and modelling, are only a few of the outstanding knowledge gaps and future research tracks that show potential to improve the understanding of plant behavior and to improve crop production in a changing climate through more resilient varieties.

5.5.1 Image processing and data extraction

In this dissertation, the imagery was captured and processed with the highest spatial resolution possible in terms of flight planning and computational power, i.e. flight altitude as low as possible without disturbing the crop due to downdraft of the rotors of the UAV, high overlap between the images, coverage of the entire trial in a single flight, etc., to explore the potential of the sensor. This had its impact on the processing time needed to produce the orthomosaic and the height model. It would be interesting to investigate which resolutions, spatial for orthomosaics and altimetric for height models, are sufficient to be able to extract the needed phenotypic data and to reduce the processing time. This would allow users to have the data at their disposal faster and at a lower cost. The effect of lower resolution could result in more mixed pixels and consequently more averaged data preventing to discriminate between extremes. However, a trade-off will exist and can be evaluated by downsampling orthomosaics and height models.

When a licensed software such as Agisoft Photoscan or Pix4D is used, we have limited knowledge about the processing algorithms and the influence or importance of the parameters chosen. This "black box" determines the accuracy and quality of the orthophoto and DEM. Thus, to get the best possible results the impact of the SfM algorithm used should be studied. New open source tools have recently become available for flight planning (e.g. PhenoFly (Roth et al., 2018b)) and data analysis (e.g. FIELDimageR (Matias et al., 2020)). These tools should be tested for more applications and scenarios to strengthen them and spread their use. One of the principles of open source software is transparency. Therefore, detailed information on the algorithms and code are accessible for a more flexible use compared to a privative/proprietary software.

To automate the data extraction and have fully operable workflows to calculate indices and extract data will also reduce the processing time. The delineation of the regions of interest (plots) can be time-consuming when this is done manually. If the design of the experiment is regular (sowing controlled by GPS or equidistant plots), the task can be simplified by using a software such as MiniGis (geo-konzept GmbH) or programing language to define the plots. Indices and statistics can be calculated in a pipeline that can be created with the different steps in a GIS software such as QGIS or ArcGIS (Graphical Modeler or Model Builder) or using a script.

The different options mentioned to decrease the time required from capturing the imagery to the data analysis depend on the studied crop. The morphology of the crop, e.g. leaves size, height, canopy structure and even the sown/plant density, has to be taken into account to plan the UAV mission. For example, a lower spatial resolution will not affect equally to a woody crop than to a cereal crop as they differ in structure and height. In our cases, a lower spatial resolution in perennial ryegrass could result in mixed pixels (if pixel resolution > leave width) resulting in a reduction of the contrast and a lower quality of the digital elevation which could decrease the differentiation between plots. The influence of height and overlap on computation time and DEM accuracy in olive trees has been investigated in Torres-Sánchez et al. (2018b) and a spatial quality evaluation of resampled UAV imagery for weed mapping was carried out in Borra-Serrano et al. (2015).

5.5.2 Combining information from different sensors and new traits estimation

In the majority of the reported studies the imagery set has been captured with the camera perpendicular to the field (nadir position). However, this has limitations as only data from the top part of the canopy is included. If we are interested in estimating leaf area index (LAI), indicative for the plant canopy performance or yield potential, or crop morphology, it is an option to include oblique images as tested by Roth et al. (2018a). However, this could increase the processing time to obtain the UAV products as a higher number of images would have to be processed.

In this dissertation, one sensor was used to develop the methodologies but there are other sensors available to complement the spectral data that allow to evaluate multiple traits simultaneously together with their complex and dynamic interaction. Being able to measure multiple traits at multiple times while combining sensors will increase the data available to generate phenotype information (phenotyping capabilities) and increase efficiency in the selection (more data in a shorter time) as a better characterization is achieved (Rebetzke et al., 2019). In our case, it would be interesting to incorporate multispectral sensors which offer extra spectral information without losing too much resolution. Another option could be a modified commercial 'off the shelf' (COTS) camera which has a higher spatial resolution, but requires a more complex radiometric calibration than the commercial multispectral sensors (Holman et al., 2019). Incorporating the near infrared and red edge region would give access to a full range of vegetation indices linked to traits such as stress or diseases. For example, the Green Normalized Difference Vegetation Index (GNDVI) has been used among other indices for plant disease detection (Pourazar et al., 2019), while near-infrared bands were used to detect water stress in ornamental plants (Freeman et al., 2019). Those cameras tend to have a higher price and lower spatial resolution than RGB cameras, but the added value in terms of spectral information could justify the investment. We are aware that the traits considered in this dissertation are not the only ones taken into account in breeding selection, but that several traits are combined. Therefore, other traits should be evaluated and investigated to provide a more complete picture of the genotypes evaluated in the breeding trials as a basis for selection and development of improved varieties. Not all traits are scored at the same moment, so drone flights at different moments during the growing season are required to cover the critical stages. The advantage of a trait-based approach is that heritability of a yield component may be much higher than the yield itself, thus maximizing genetic gain over those achievable components compared to selection for yield alone.

It is not always straightforward to scale a trait measured at leaf level to canopy level. For example, photosynthetic capacity or light interception are influenced by leaf age and the different layers of the canopy profile, or biomass accumulation and radiation use efficiency require an analysis of the canopy at different developmental stages. Therefore, indirect estimates/approaches could be a solution, such as canopy ground cover as an alternative to light interception. Fractional ground cover estimated with UAV-RGB imagery as a proxy for canopy light interception has been used in maize combined with incident radiation from a weather station and against biomass samples at different moments to derive radiation use efficiency at different stages of crop growth (Tewes and Schellberg, 2018). Another example is the combination of LiDAR and 3D modelling to estimate the green area index (Liu et al., 2017a). As LiDAR provides data on the vertical distribution of the canopy, it can be used to model light penetration and interception at different heights and times of the day. Most efforts have been made towards direct selection for traits such as yield, but indirect selection could also be beneficial. It gives the opportunity to introduce new alleles to maintain genetic variability and a path to early selection of one or more adaptation drivers (Fischer and Rebetzke, 2018), e.g. in a water-limited scenario (Richards et al., 2010). A review of the possibilities of indirect selection is presented in Fischer and Rebetzke (2018).

Murchie et al. (2018) indicated the need to understand photosynthetic dynamics from a mechanistic and biological point of view in phenomics, as it is an important target for improving yield. Although progress in the measurement techniques has been made, the dynamic processes involved in photosynthesis have not yet been described in detail. Photosynthesis integrates major primary processes (e.g. light harvesting or gas exchange), is sensitive to events at the whole-plant level and highly responsive to fluctuations in the environmental conditions. Therefore, it is not a single linear process. To measure properties under steady-state conditions is convenient and relevant, but does not allow to predict the dynamic response (Poorter et al., 2016). A set of characteristics which include static properties and dynamic responses are required to describe the photosynthetic apparatus. Inclusion of dynamic responses such as stomatal responses is important. Although it is challenging to capture photosynthetic responses in the field, this is where the appropriate phenotype is expressed and not in controlled environments (Poorter et al., 2016). At present, there is no fast and reliable method available to detect photosynthetic carbon-uptake rates for large-scale field phenotyping (Watt et al., 2020). Sun-induced fluorescence, stomatal responses through thermography, plant water use efficiency or a more detailed canopy structure (e.g. leaf angle) are informative parameters related to photosynthesis to be addressed.

The adaptability of some ground sensors (e.g. hyperspectral sensors) to UAV is also a challenge, not only in terms of miniaturization and costs, but also in the preprocessing required to use the data (radiometric correction to turn images to reflectance or geometric corrections for georeferenced images) and the spatial resolution (mixed pixels). Integration of data from different platforms and sensor types can be used to target specific traits such as growth (Potgieter et al., 2018).

5.5.3 Data standardization and sharing

As mentioned earlier in this dissertation, there are initiatives to implement the FAIR principles (Findable, Accessible, Interoperable and Reusable) to phenotypic data. To reach this ambitious aim, standards

should be defined in collaboration between the different groups which carry out HTFP. Especially within EPPN2020 and Emphasis (https://eppn2020.plant-phenotyping.eu/) necessary steps are being undertaken to improve the interoperability between platforms and theirs datasets. To achieve this, an exchange of experiences, methodologies and datasets with detailed information (data acquisition, weather conditions, etc.) is needed. Therefore, a revision of our protocols and flight missions and establishing a structure to make our data accessible has to be discussed and implemented. The huge amount of data generated from diverse sources (scores, imagery, etc.) could limit the exchange. Websites and online databases are becoming available to share phenotyping know-how (Roitsch et al., 2019) such as the Quantitative Plant (https://www.quantitative-plant.org/) where it is possible to find image analysis software tools and plant models available for use.

5.5.4 Advanced mathematical methods

Machine learning (ML) consists of a group of computerized modeling approaches that learn patterns from the data for automatic decision making without explicit rules programmed (Singh et al., 2016). The learning process can be supervised (labeled data) or unsupervised (no initial classification), and the modeling objective could be discriminative (to distinguish between different data patterns) or generative (to learn and generate similar patterns synthetically). As major advantages, ML facilitates to discover patterns and to look simultaneously at a combination of factors (instead of each one individually) in large datasets (Singh et al., 2016). ML approaches require a pivotal pre-processing step and an appropriate selection of methods and model choice. There are a large number of choices and it is important to pick the appropriate one based on the type and amount of data available to tackle the specific problem (Singh et al., 2016). They are considered the most viable approach to analyze the growing (in size, scope and complexity) image datasets (Shakoor et al., 2017).

ML has been used to make progress in quantifying traits such as plant stresses, yield or root traits. Plant traits can be modelled using ML, as it uses advanced statistical techniques able to approximate complex non-linear functions between image features and biophysical parameters (Herrero-Huerta et al., 2020). Recent examples of soybean yield prediction with UAV imagery using ML resulted in a RMSE of 15.9% (Maimaitijiang et al., 2020) and in an accuracy of 90.7% (Herrero-Huerta et al., 2020). In Singh et al. (2016) detailed explanation of machine learning tools for stress phenotyping can be found.

5.5.5 HTFP in genetic research and for modelers

The availability of accurate and precise high-throughput phenotyping data is crucial to correlate with genomics data and bridge the phenotyping gap. For example, genomic selection requires accurate phenotyping measurements on the training population that will be used in the genetic selection model for predictions in the test set. The phenotypic data is not used to select individual plants, but rather to train models and predict the performance of non-phenotyped individuals from their DNA-marker scores. However, this requires a large training population when working with low heritability traits (Cobb et al., 2013). Most of the genomic selection models are single trait, but as measurement of multiple (correlated) traits is possible using a phenotyping platform, multiple-trait genetic selection models could be developed (Crain et al., 2018). Rutkoski et al. (2016) compared a multiple-trait model to a single trait

one and observed an increase in the prediction accuracy, reflecting the advantage of combining genomic selection and high-throughput phenotyping platform measurements.

In addition, HTFP will permit to link processes that occur at molecular level with morphological characteristics and use linkage mapping and association mapping to dissect the genetic architecture of complex traits (Bazakos et al., 2017). Accurate and precise phenotyping has to be designed to interpret and dissect the genetic architecture of complex traits and to understand how specific DNA variants (genes or alleles) determine the inheritance of diverse phenotypes (Cobb et al., 2013).

Multi-temporal phenotypic data can also be relevant for plant and crop modelling. Every model needs a dataset measured in the field to be fitted to build/train the model. HTFP can be the source to provide robust data for 'physiological breeding' or to construct a model or ideal plant (i.e. 'ideotype') with combined traits. As to phenotype genetically complex quantitative physiological traits has become more technologically tractable, and not only simple visible phenotypes are available (Reynolds 2016), the approach to construct an ideal plant is feasible.

References

Aasen, H., Honkavaara, E., Lucieer, A., Zarco-Tejada, P., 2018. Quantitative remote sensing at ultra-high resolution with UAV spectroscopy: a review of sensor technology, measurement procedures, and data correction workflows. Remote Sens. 10, 1091. https://doi.org/10.3390/rs10071091

Aasen, H., Kirchgessner, N., Walter, A., Liebisch, F., 2020. PhenoCams for field phenotyping: using very high temporal resolution digital repeated photography to investigate interactions of growth, phenology, and harvest traits. Front. Plant Sci. 11, 593. https://doi.org/10.3389/fpls.2020.00593

Ali, I., Cawkwell, F., Dwyer, E., Barrett, B., Green, S., 2016. Satellite remote sensing of grasslands: from observation to management. J. Plant Ecol. 9, 649–671. https://doi.org/10.1093/jpe/rtw005

Altieri, M.A., Nicholls, C.I., 2017. The adaptation and mitigation potential of traditional agriculture in a changing climate. Clim. Change 140, 33–45. https://doi.org/10.1007/s10584-013-0909-y

Aper, J., De Clercq, H., Baert, J., 2016. Agronomic characteristics of early-maturing soybean and implications for breeding in Belgium. Plant Genet. Resour. Characterisation Util. 14, 142–148. https://doi.org/10.1017/S1479262115000180

Araus, J.L., Cairns, J.E., 2014. Field high-throughput phenotyping: the new crop breeding frontier. Trends Plant Sci. https://doi.org/10.1016/j.tplants.2013.09.008

Araus, J.L., Kefauver, S.C., 2018. Breeding to adapt agriculture to climate change: affordable phenotyping solutions. Curr. Opin. Plant Biol. 45, 237–247. https://doi.org/10.1016/j.pbi.2018.05.003

Araus, J.L., Kefauver, S.C., Zaman-Allah, M., Olsen, M.S., Cairns, J.E., 2018. Translating high-throughput phenotyping into genetic gain. Trends Plant Sci. https://doi.org/10.1016/j.tplants.2018.02.001

Arvidsson, S., Pérez-Rodríguez, P., Mueller-Roeber, B., 2011. A growth phenotyping pipeline for *Arabidopsis thaliana* integrating image analysis and rosette area modeling for robust quantification of genotype effects. New Phytol. 191, 895–907. https://doi.org/10.1111/j.1469-8137.2011.03756.x

Atieno, J., Li, Y., Langridge, P., Dowling, K., Brien, C., Berger, B., Varshney, R.K., Sutton, T., 2017. Exploring genetic variation for salinity tolerance in chickpea using image-based phenotyping. Sci. Rep. 7, 1–11. https://doi.org/10.1038/s41598-017-01211-7

Bai, G., Ge, Y., Hussain, W., Baenziger, P.S., Graef, G., 2016. A multi-sensor system for high throughput field phenotyping in soybean and wheat breeding. Comput. Electron. Agric. 128, 181–192. https://doi.org/10.1016/j.compag.2016.08.021

Bai, G., Jenkins, S., Yuan, W., Graef, G.L., Ge, Y., 2018. Field-based scoring of soybean iron deficiency chlorosis using RGB imaging and statistical learning. Front. Plant Sci. 9, 1–12. https://doi.org/10.3389/fpls.2018.01002

Baker, R.E., Peña, J.M., Jayamohan, J., Jérusalem, A., 2018. Mechanistic models versus machine learning, a fight worth fighting for the biological community? Biol. Lett. 14. https://doi.org/10.1098/rsbl.2017.0660

Bareth, G., Bendig, J., Tilly, N., Hoffmeister, D., Aasen, H., Bolten, A., 2016. A comparison of UAV- and TLS-derived plant height for crop monitoring: using polygon grids for the analysis of crop surface models (CSMs). Photogramm. - Fernerkundung - Geoinf. 2016, 85–94. https://doi.org/10.1127/pfg/2016/0289

Bareth, G., Lussem, U., Menne, J., Hollberg, J., Schellberg, J., 2019. Potential of non-calibrated UAV-based rgb imagery for forage monitoring: Case study at the rengen long-term grassland experiment (rge), Germany. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch. 42, 203–206. https://doi.org/10.5194/isprs-archives-XLII-2-W13-203-2019

Bareth, G., Schellberg, J., 2018. Replacing manual rising plate meter measurements with low-cost UAV-derived sward height data in grasslands for spatial monitoring. PFG - J. Photogramm. Remote Sens. Geoinf. Sci. 86, 157–168. https://doi.org/10.1007/s41064-018-0055-2

Barre, P., Ruttink, T., Muylle, H., Lootens, P., Sampoux, J.P., Rohde, A., Combes, D., Roldán-Ruiz, I., 2017. Natural diversity in vegetative and reproductive investments of perennial ryegrass is shaped by the climate at the place of origin. Grass Forage Sci. 73, 193–205. https://doi.org/10.1111/gfs.12304

Bazakos, C., Hanemian, M., Trontin, C., Jiménez-Gómez, J.M., Loudet, O., 2017. New strategies and tools in quantitative genetics: how to go from the phenotype to the genotype. Annu. Rev. Plant Biol. 68, 435–455. https://doi.org/10.1146/annurev-arplant-042916-040820

Behaeghe, T. J., 1979. De seizoenvariatie in de grasgroei. Thesis, Gent, 272 pp.

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

Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., Gnyp, M.L., Bareth, G., 2015. Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. Int. J. Appl. Earth Obs. Geoinf. 39, 79–87. https://doi.org/10.1016/j.jag.2015.02.012

Berra, E.F., Gaulton, R., Barr, S., 2017. Commercial off-the-shelf digital cameras on Unmanned Aerial Vehicles for multitemporal monitoring of vegetation reflectance and NDVI. IEEE Trans. Geosci. Remote Sens. 55, 4878–4886. https://doi.org/10.1109/TGRS.2017.2655365

Boerma, H.R., Specht, J.E., 2004. Soybeans: improvement, production and uses. Soybeans Improv. Prod. uses.

Borra-Serrano, I., Peña, J.M., Torres-Sánchez, J., Mesas-Carrascosa, F.J., López-Granados, F., 2015. Spatial quality evaluation of resampled unmanned aerial vehicle-imagery for weed mapping. Sensors (Switzerland) 15. https://doi.org/10.3390/s150819688

Brocks, S., Bareth, G., 2018. Estimating barley biomass with crop surface models from oblique RGB imagery. Remote Sens. 10. https://doi.org/10.3390/rs10020268

Burgon, A., Bondesen, O.B., Verburgt, W.H., Hall, A.G., Bark, N.S., Robinson, M., Timm, G., Fairey, D.T., Hampton, J.G., 1998. The forage seed trade. Forage seed Prod. Vol. 1 Temp. species. 271–286.

Burton, J.W., 1997. Soyabean (*Glycine max* (L.) Merr.). F. Crop. Res. 53, 171–186. https://doi.org/10.1016/S0378-4290(97)00030-0

Cahoon, E.B., Clemente, T.E., Damude, H.G., Kinney, A.J., 2009. Modifying vegetable oils for food and non-food purposes, in: oil Crops. Springer New York, pp. 31–56. https://doi.org/10.1007/978-0-387-77594-4_2

Camargo Neto, J., 2004. A combined statistical-soft computing approach for classification and mapping weed species in minimum-tillage systems. ETD Collect. Univ. Nebraska - Lincoln.

Campbell, Z.C., Acosta-Gamboa, L.M., Nepal, N., Lorence, A., 2018. Engineering plants for tomorrow: how high-throughput phenotyping is contributing to the development of better crops. Phytochem. Rev. 17, 1329–1343. https://doi.org/10.1007/s11101-018-9585-x

Carter, G.A., 1993. Responses of leaf spectral reflectance to plant stress. Am. J. Bot. 80, 239–243. https://doi.org/10.1002/j.1537-2197.1993.tb13796.x

Casadesús, J., Villegas, D., 2014. Conventional digital cameras as a tool for assessing leaf area index and biomass for cereal breeding. J. Integr. Plant Biol. 56, 7–14. https://doi.org/10.1111/jipb.12117

Casler, M.D., van Santen, E., 2010. Breeding objectives in forages, in: Boller, B., Posselt, U.K., Veronesi, F. (Eds.), Fodder Crops and Amenity Grasses. Springer New York, New York, NY, pp. 115–136. https://doi.org/10.1007/978-1-4419-0760-8_5

Caturegli, L., Corniglia, M., Gaetani, M., Grossi, N., Magni, S., Migliazzi, M., Angelini, L., Mazzoncini, M., Silvestri, N., Fontanelli, M., Raffaelli, M., Peruzzi, A., Volterrani, M., 2016. Unmanned aerial vehicle to estimate nitrogen status of turfgrasses. PLoS One 11, 1–13. https://doi.org/10.1371/journal.pone.0158268

Chang, A., Jung, J., Maeda, M.M., Landivar, J., 2017. Crop height monitoring with digital imagery from Unmanned Aerial System (UAS). Comput. Electron. Agric. 141, 232–237. https://doi.org/10.1016/j.compag.2017.07.008

Chapman, S.C., Merz, T., Chan, A., Jackway, P., Hrabar, S., Dreccer, M.F., Holland, E., Zheng, B., Ling, T.J., Jimenez-Berni, J., 2014. Pheno-copter: A low-altitude, autonomous remote-sensing robotic helicopter for high-throughput field-based phenotyping. Agronomy 4, 279–301. https://doi.org/10.3390/agronomy4020279

Chaves, B., De Vliegher, A., Van Waes, J., Carlier, L., Marynissen, B., 2009. Change in agronomic performance of *Lolium perenne* and *Lolium multiflorum* varieties in the past 40 years based on data from Belgian VCU trials. Plant Breed. 128, 680–690. https://doi.org/10.1111/j.1439-0523.2009.01621.x

Chawade, A., Armoniené, R., Berg, G., Brazauskas, G., Frostgård, G., Geleta, M., Gorash, A., Henriksson, T., Himanen, K., Ingver, A., Johansson, E., Jørgensen, L.N., Koppel, M., Koppel, R., Makela, P., Ortiz, R., Podyma, W., Roitsch, T., Ronis, A., Svensson, J.T., Vallenback, P., Weih, M., 2018. A transnational and holistic breeding approach is needed for sustainable wheat production in the Baltic Sea region. Physiol. Plant. 164, 442–451. https://doi.org/10.1111/ppl.12726

Chawade, A., Van Ham, J., Blomquist, H., Bagge, O., Alexandersson, E., Ortiz, R., 2019. High-throughput field-phenotyping tools for plant breeding and precision agriculture. Agronomy 9, 1–18. https://doi.org/10.3390/agronomy9050258

Chen, D., Neumann, K., Friedel, S., Kilian, B., Chen, M., Altmann, T., Klukas, C., 2014. Dissecting the phenotypic components of crop plant growth and drought responses based on high-throughput image analysis. Plant Cell Online 26, 4636–4655. https://doi.org/10.1105/tpc.114.129601

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

Cobb, J.N., DeClerck, G., Greenberg, A., Clark, R., McCouch, S., 2013. Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype-phenotype relationships and its relevance to crop improvement. Theor. Appl. Genet. 126, 867–887. https://doi.org/10.1007/s00122-013-2066-0

Cober, E.R., Cianzio, S.R., Pantalone, V.R., Rajcan, I., 2009. Soybean, in: Oil Crops. Springer New York, New York, NY, pp. 57–90. https://doi.org/10.1007/978-0-387-77594-4_3

Colomina, I., Molina, P., 2014. Unmanned aerial systems for photogrammetry and remote sensing: a review. ISPRS J. Photogramm. Remote Sens. 92, 79–97. https://doi.org/10.1016/j.isprsjprs.2014.02.013

Costa, C., Schurr, U., Loreto, F., Menesatti, P., Carpentier, S., 2019. Plant phenotyping research trends, a science mapping approach. Front. Plant Sci. 9, 1–11. https://doi.org/10.3389/fpls.2018.01933

Craft, J.C., Lindsey, L.E., Barker, D.J., Lindsey, A.J., 2019. Quantification of soybean leaf senescence and maturation as impacted by soil- and foliar-applied nitrogen. Crop. Forage Turfgrass Manag. 5, 180051. https://doi.org/10.2134/cftm2018.07.0051

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

Crain, J.L., Wei, Y., Barker, J., Thompson, S.M., Alderman, P.D., Reynolds, M., Zhang, N., Poland, J., 2016. Development and deployment of a portable field phenotyping platform. Crop Sci. 56, 965–975. https://doi.org/10.2135/cropsci2015.05.0290

Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., de los Campos, G., Burgueño, J., González-Camacho, J.M., Pérez-Elizalde, S., Beyene, Y., Dreisigacker, S., Singh, R., Zhang, X., Gowda, M., Roorkiwal, M., Rutkoski, J., Varshney, R.K., 2017. Genomic selection in plant breeding: methods, models, and perspectives. Trends Plant Sci. 22, 961–975. https://doi.org/10.1016/j.tplants.2017.08.011

Cucho-Padin, G., Loayza, H., Palacios, S., Balcazar, M., Carbajal, M., Quiroz, R., 2019. Development of low-cost remote sensing tools and methods for supporting smallholder agriculture. Appl. Geomatics 1–17. https://doi.org/10.1007/s12518-019-00292-5

Dandois, J.P., Ellis, E.C., 2010. Remote sensing of vegetation structure using computer vision. Remote Sens. 2, 1157–1176. https://doi.org/10.3390/rs2041157

de Castro, A., Jiménez-Brenes, F., Torres-Sánchez, J., Peña, J., Borra-Serrano, I., López-Granados, F., 2018. 3-D Characterization of Vineyards Using a Novel UAV Imagery-Based OBIA Procedure for Precision Viticulture Applications. Remote Sens. 10, 584. https://doi.org/10.3390/rs10040584

de Castro, A.I., Rallo, P., Suárez, M.P., Torres-Sánchez, J., Casanova, L., Jiménez-Brenes, F.M., Morales-Sillero, A., Jiménez, M.R., López-Granados, F., 2019. High-Throughput System for the Early Quantification of Major Architectural Traits in Olive Breeding Trials Using UAV Images and OBIA Techniques. Front. Plant Sci. 10. https://doi.org/10.3389/fpls.2019.01472

de la Casa, A., Ovando, G., Bressanini, L., Martínez, J., Díaz, G., Miranda, C., 2018. Soybean crop coverage estimation from NDVI images with different spatial resolution to evaluate yield variability in a plot. ISPRS J. Photogramm. Remote Sens. 146, 531–547. https://doi.org/10.1016/j.isprsjprs.2018.10.018

De Swaef, T., Bellocchi, G., Aper, J., Lootens, P., Roldán-Ruiz, I., 2019. Use of identifiability analysis in designing phenotyping experiments for modelling forage production and quality. J. Exp. Bot. 70, 2587–2604. https://doi.org/10.1093/jxb/erz049

Deery, D., Jimenez-Berni, J., Jones, H., Sirault, X., Furbank, R., 2014. Proximal remote sensing buggies and potential applications for field-based phenotyping, Agronomy. https://doi.org/10.3390/agronomy4030349

Desta, Z.A., Ortiz, R., 2014. Genomic selection: genome-wide prediction in plant improvement. Trends Plant Sci. 19, 592–601. https://doi.org/10.1016/j.tplants.2014.05.006

Díaz-Varela, R.A., De La Rosa, R., León, L., Zarco-Tejada, P.J., Lucieer, A., Rascher, U., Bareth, G., Baghdadi, N., Thenkabail, P.S., 2015. High-Resolution Airborne UAV Imagery to Assess Olive Tree Crown Parameters Using 3D Photo Reconstruction: Application in Breeding Trials 7, 4213–4232. https://doi.org/10.3390/rs70404213

Du, M., Noguchi, N., 2017. Monitoring of wheat growth status and mapping of wheat yield's within-field spatial variations using color images acquired from UAV-camera System. Remote Sens. 9. https://doi.org/10.3390/rs9030289

Edwards, J.T., Purcell, L.C., 2005. Soybean yield and biomass responses to increasing plant population among diverse maturity groups: I. Agronomic Characteristics. Crop Sci. 45, 1770–1777. https://doi.org/10.2135/cropsci2004.0564

Edwards, J.T., Purcell, L.C., Karcher, D.E., 2005. Soybean yield and biomass responses to increasing plant population among diverse maturity groups: II. Light Interception and Utilization. Crop Sci. 45, 1778–1785. https://doi.org/10.2135/cropsci2004.0570

Fahlgren, N., Gehan, M.A., Baxter, I., 2015. Lights, camera, action: high-throughput plant phenotyping is ready for a close-up. Curr. Opin. Plant Biol. 24, 93–99. https://doi.org/10.1016/j.pbi.2015.02.006

FAO, IFAD, UNICEF, WFP and WHO, 2019. The State of Food Security and Nutrition in the World 2019. Safeguarding against economic slowdowns and downturns. Rome, FAO.

Fehr, W., Caviness, C., 1977. Stages of soybean development. Spec. Rep. 87.

Feng, L., Mailhol, J.C., Rey, H., Griffon, S., Auclair, D., De Reffye, P., 2014. Comparing an empirical crop model with a functional structural plant model to account for individual variability. Eur. J. Agron. 53, 16–27. https://doi.org/10.1016/j.eja.2013.11.005

Fernandez-Gallego, J.A., Kefauver, S.C., Gutiérrez, N.A., Nieto-Taladriz, M.T., Araus, J.L., 2018a. Automatic wheat ear counting in-field conditions: simulation and implication of lower resolution images, in: Neale, C.M., Maltese, A. (Eds.), Remote Sensing for Agriculture, Ecosystems, and Hydrology XX. SPIE, p. 23. https://doi.org/10.1117/12.2500083

Fernandez-Gallego, J.A., Kefauver, S.C., Gutiérrez, N.A., Nieto-Taladriz, M.T., Araus, J.L., 2018b. Wheat ear counting in-field conditions: High throughput and low-cost approach using RGB images. Plant Methods 14, 1–12. https://doi.org/10.1186/s13007-018-0289-4

Fernandez-Gallego, J.A., Lootens, P., Borra-Serrano, I., Derycke, V., Haesaert, G., Roldán-Ruiz, I., Araus, J.L., Kefauver, S.C., 2020. Automatic wheat ear counting using machine learning based on RGB UAV imagery. Plant J. tpj.14799. https://doi.org/10.1111/tpj.14799

Filella, I., Peñuelas, J., 1994. The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. Int. J. Remote Sens. 15, 1459–1470. https://doi.org/10.1080/01431169408954177

Fiorani, F., Schurr, U., 2013. Future scenarios for plant phenotyping. Annu. Rev. Plant Biol. 64, 267–291. https://doi.org/10.1146/annurev-arplant-050312-120137

Fischer, R.A., Edmeades, G.O., 2010. Breeding and cereal yield progress. Crop Sci. 50, S-85-S-98. https://doi.org/10.2135/cropsci2009.10.0564

Fischer, R.A., Rebetzke, G.J., 2018. Indirect selection for potential yield in early-generation, spaced plantings of wheat and other small-grain cereals: A review. Crop Pasture Sci. https://doi.org/10.1071/CP17409

Franceschini, M.H.D., Bartholomeus, H., van Apeldoorn, D.F., Suomalainen, J., Kooistra, L., 2019. Feasibility of Unmanned Aerial Vehicle optical imagery for early detection and severity assessment of late blight in potato. Remote Sens. 11, 224. https://doi.org/10.3390/rs11030224

Freeman, D., Gupta, S., Hudson Smith, D., Maja, J.M., Robbins, J., Owen, J.S., Peña, J.M., de Castro, A.I., 2019. Watson on the farm: Using cloud-based artificial intelligence to identify early indicators of water stress. Remote Sens. 11. https://doi.org/10.3390/rs11222645

Furbank, R.T., Jimenez-Berni, J.A., George-Jaeggli, B., Potgieter, A.B., Deery, D.M., 2019. Field crop phenomics: enabling breeding for radiation use efficiency and biomass in cereal crops. New Phytol. 223, 1714–1727. https://doi.org/10.1111/nph.15817

Furbank, R.T., Tester, M., 2011. Phenomics - technologies to relieve the phenotyping bottleneck. Trends Plant Sci. 16, 635–644. https://doi.org/10.1016/j.tplants.2011.09.005

Gatziolis, D., Lienard, J.F., Vogs, A., Strigul, N.S., 2015. 3D Tree dimensionality assessment using photogrammetry and small Unmanned Aerial Vehicles. PLoS One 10, e0137765. https://doi.org/10.1371/journal.pone.0137765

Gbikpi, P.J., Crookson, R.K., 1981. A whole-plant indicator of soybean physiological maturity. Crop Sci. 21, 469–472. https://doi.org/10.2135/cropsci1981.0011183x002100030030x

Ge, Y., Bai, G., Stoerger, V., Schnable, J.C., 2016. Temporal dynamics of maize plant growth, water use, and leaf water content using automated high throughput RGB and hyperspectral imaging. Comput. Electron. Agric. 127, 625–632. https://doi.org/10.1016/j.compag.2016.07.028

Gebremedhin, A., Badenhorst, P., Wang, J., Giri, K., Spangenberg, G., Smith, K., 2019a. Development and validation of a model to combine NDVI and plant height for high-throughput phenotyping of herbage yield in a perennial ryegrass breeding program. Remote Sens. 11. https://doi.org/10.3390/rs11212494

Gebremedhin, A., Badenhorst, P.E., Wang, J., Spangenberg, G.C., Smith, K.F., 2019b. Prospects for measurement of dry matter yield in forage breeding programs using sensor technologies. Agronomy 9, 65. https://doi.org/10.3390/agronomy9020065

Geipel, J., Link, J., Claupein, W., 2014. Combined spectral and spatial modeling of corn yield based on aerial images and crop surface models acquired with an Unmanned Aircraft System. Remote Sens. 6, 10335–10355. https://doi.org/10.3390/rs61110335

Geipel, J., Link, J., Wirwahn, J., Claupein, W., 2016. A programmable aerial multispectral camera system for in-season crop biomass and nitrogen content estimation. Agriculture 6, 4. https://doi.org/10.3390/agriculture6010004

Gibson, P., 2013. Introductory remote sensing principles and concepts, introductory remote sensing principles and concepts. Routledge. https://doi.org/10.4324/9780203714522

Gitelson, A.A., Kaufman, Y.J., Stark, R., Rundquist, D., 2002. Novel algorithms for remote estimation of vegetation fraction. Remote Sens. Environ. 80, 76–87. https://doi.org/10.1016/S0034-4257(01)00289-9

Gnädinger, F., Schmidhalter, U., 2017. Digital counts of maize plants by Unmanned Aerial Vehicles (UAVs). Remote Sens. 9. https://doi.org/10.3390/rs9060544

Gobron, N., Pinty, B., Verstraete, M.M., Widlowski, J.L., 2000. Advanced vegetation indices optimized for up-coming sensors: design, performance, and applications. IEEE Trans. Geosci. Remote Sens. 38, 2489–2505. https://doi.org/10.1109/36.885197

Gompertz B., 1825. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. In a letter to Francis Baily, Esq. F. R. S. &c. Philos. Trans. R. Soc. London 182, 513–583. https://doi.org/10.1098/rstl.1825.0026

Govaerts, Y.M., Verstraete, M.M., Pinty, B., Gobron, N., 1999. Designing optimal spectral indices: a feasibility and proof of concept study. Int. J. Remote Sens. 20, 1853–1873. https://doi.org/10.1080/014311699212524

Gracia-Romero, A., Kefauver, S.C., Fernandez-Gallego, J.A., Vergara-Díaz, O., Nieto-Taladriz, M.T., Araus, J.L., 2019. UAV and ground image-based phenotyping: a proof of concept with durum wheat. Remote Sens. 11. https://doi.org/10.3390/rs11101244

Granier, C., Aguirrezabal, L., Chenu, K., Cookson, S.J., Dauzat, M., Hamard, P., Thioux, J.J., Rolland, G., Bouchier-Combaud, S., Lebaudy, A., Muller, B., Simonneau, T., Tardieu, F., 2006. PHENOPSIS, an automated platform for reproducible phenotyping of plant responses to soil water deficit in Arabidopsis thaliana permitted the identification of an accession with low sensitivity to soil water deficit. New Phytol. 169, 623–635. https://doi.org/10.1111/j.1469-8137.2005.01609.x

Grüner, E., Astor, T., Wachendorf, M., 2019. Biomass prediction of heterogeneous temperate grasslands using an SFM approach based on UAV imaging. Agronomy 9. https://doi.org/10.3390/agronomy9020054

Guangyu, C., Taihui, Z., Jian, M., Yi, S., Xin, C., 2012. Changes in spectral reflectance of vegetation in response to specific nutrient supply, in: Lecture Notes in Electrical Engineering. Springer, Berlin, Heidelberg, pp. 671–675. https://doi.org/10.1007/978-3-642-28744-2_87

Guerrero, J.M., Pajares, G., Montalvo, M., Romeo, J., Guijarro, M., 2012. Support Vector Machines for crop/weeds identification in maize fields. Expert Syst. Appl. 39, 11149–11155. https://doi.org/10.1016/j.eswa.2012.03.040

Guijarro, M., Pajares, G., Riomoros, I., Herrera, P.J., Burgos-Artizzu, X.P., Ribeiro, A., 2011. Automatic segmentation of relevant textures in agricultural images. Comput. Electron. Agric. 75, 75–83. https://doi.org/10.1016/J.COMPAG.2010.09.013

Guijarro, M., Riomoros, I., Pajares, G., Zitinski, P., 2015. Discrete wavelets transform for improving greenness image segmentation in agricultural images. Comput. Electron. Agric. 118, 396–407. https://doi.org/10.1016/j.compag.2015.09.011

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

Guo, Z., Yang, W., Chang, Y., Ma, X., Tu, H., Xiong, F., Jiang, N., Feng, H., Huang, C., Yang, P., Zhao, H., Chen, G., Liu, H., Luo, L., Hu, H., Liu, Q., Xiong, L., 2018. Genome-Wide Association Studies of image traits reveal genetic architecture of drought resistance in rice. Mol. Plant 11, 789–805. https://doi.org/10.1016/j.molp.2018.03.018

Gupta, P.K., Kulwal, P.L., Jaiswal, V., 2019. Association mapping in plants in the post-GWAS genomics era, in: Advances in Genetics. Academic Press Inc., pp. 75–154. https://doi.org/10.1016/bs.adgen.2018.12.001

Hague, T., Tillett, N.D., Wheeler, H., 2006. Automated crop and weed monitoring in widely spaced cereals. Precis. Agric. 7, 21–32. https://doi.org/10.1007/s11119-005-6787-1

Han, L., Yang, G., Dai, H., Yang, H., Xu, B., Feng, H., Li, Z., Yang, X., 2019. Fuzzy clustering of maize plant-height patterns using time series of UAV remote-sensing images and variety traits. Front. Plant Sci. 10, 1–17. https://doi.org/10.3389/fpls.2019.00926

Han, L., Yang, G., Yang, H., Xu, B., Li, Z., Yang, X., 2018. Clustering field-based maize phenotyping of plant-height growth and canopy spectral dynamics using a UAV remote-sensing approach. Front. Plant Sci. 9, 1–18. https://doi.org/10.3389/fpls.2018.01638

Hatton, T.J., West, N.E., Johnson, P.S., 1986. Relationships of the error associated with ocular estimation and actual total cover. J. Range Manag. 39, 91–92. https://doi.org/Doi 10.2307/3899697

Herrero-Huerta, M., Rodriguez-Gonzalvez, P., Rainey, K.M., 2020. Yield prediction by machine learning from UAS-based multi-sensor data fusion in soybean. Plant Methods 16, 78. https://doi.org/10.1186/s13007-020-00620-6

Herrmann, D., Boller, B., Studer, B., Widmer, F., Kölliker, R., 2008. Improving persistence in red clover: insights from QTL analysis and comparative phenotypic evaluation. Crop Sci. 48, 269–277. https://doi.org/10.2135/cropsci2007.03.0143

Holman, F.H., Riche, A.B., Castle, M., Wooster, M.J., Hawkesford, M.J., 2019. Radiometric calibration of 'commercial off the shelf' cameras for UAV-based high-resolution temporal crop phenotyping of reflectance and NDVI. Remote Sens. 11, 1657. https://doi.org/10.3390/rs11141657

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

Holman, F.H., 2020. Development and evaluation of Unmanned Aerial Vehicles for high throughput phenotyping of field-based wheat trials. Doctoral Thesis, London, 271 pp.

Holshof, G., Stienezen, M.W.J., Galama, P.G., 2015. Calibration of five rising plate meters in the Netherlands. Grassl. Sci. Eur. 20, 233–235.

Honsdorf, N., March, T.J., Berger, B., Tester, M., Pillen, K., 2014. High-throughput phenotyping to detect drought tolerance QTL in wild barley introgression lines. PLoS One 9, e97047. https://doi.org/10.1371/journal.pone.0097047

Horst, G.L., Engelke, M.C., Meyers, W., 1984. Assessment of visual evaluation techniques. Agron. J. 76, 619–622. https://doi.org/10.2134/agronj1984.00021962007600040027x

Hou, M., Tian, F., Zhang, T., Huang, M., 2019. Evaluation of canopy temperature depression, transpiration, and canopy greenness in relation to yield of soybean at reproductive stage based on remote sensing imagery. Agric. Water Manag. 222, 182–192. https://doi.org/10.1016/j.agwat.2019.06.005

Houle, D., Govindaraju, D.R., Omholt, S., 2010. Phenomics: the next challenge. Nat. Rev. Genet. 11, 855–866. https://doi.org/10.1038/nrg2897

Hoyos-Villegas, V., Houx, J.H., Singh, S.K., Fritschi, F.B., 2014. Ground-based digital imaging as a tool to assess soybean growth and yield. Crop Sci. 54, 1756–1768. https://doi.org/10.2135/cropsci2013.08.0540

Humphreys, M., Feuerstein, U., Vandewalle, M., Baert, J., 2010. Ryegrasses, in: Boller, B., Posselt, U.K., Veronesi, F. (Eds.), Fodder Crops and Amenity Grasses. Springer New York, New York, NY, pp. 211–260. https://doi.org/10.1007/978-1-4419-0760-8_10

Humplík, J.F., Lazár, D., Husičková, A., Spíchal, L., 2015. Automated phenotyping of plant shoots using imaging methods for analysis of plant stress responses - a review. Plant Methods 11, 1–10. https://doi.org/10.1186/s13007-015-0072-8

Hunter, M.C., Smith, R.G., Schipanski, M.E., Atwood, L.W., Mortensen, D.A., 2017. Agriculture in 2050: recalibrating targets for sustainable intensification. Bioscience 67, 386–391. https://doi.org/10.1093/biosci/bix010

Huyghe, C., De Vliegher, A., Van Gils, B., Peeters, A., 2014. Grasslands and herbivore production in Europe and effects of common policies. Editions Quae, pp.320.

Ihuoma, S.O., Madramootoo, C.A., 2017. Recent advances in crop water stress detection. Comput. Electron. Agric. 141, 267–275. https://doi.org/10.1016/j.compag.2017.07.026

James, M.R., Robson, S., 2014. Mitigating systematic error in topographic models derived from UAV and ground-based image networks. Earth Surf. Process. Landforms 39, 1413–1420. https://doi.org/10.1002/esp.3609

James, M.R., Robson, S., 2012. Straightforward reconstruction of 3D surfaces and topography with a camera: accuracy and geoscience application. J. Geophys. Res. Earth Surf. 117. https://doi.org/10.1029/2011JF002289

Jannink, J.L., Jordan, N.R., Orf, J.H., 2001. Feasibility of selection for high weed suppressive ability in soybean: absence of tradeoffs between rapid initial growth and sustained later growth. Euphytica 120, 291–300. https://doi.org/10.1023/A:1017540800854

Jiménez-Brenes, F.M., López-Granados, F., Castro, A.I., Torres-Sánchez, J., Serrano, N., Peña, J.M., 2017. Quantifying pruning impacts on olive tree architecture and annual canopy growth by using UAV-based 3D modelling. Plant Methods 13, 1–15. https://doi.org/10.1186/s13007-017-0205-3

Jin, X., Zarco-Tejada, P., Schmidhalter, U., Reynolds, M.P., Hawkesford, M.J., Varshney, R.K., Yang, T., Nie, C., Li, Z., Ming, B., Xiao, Y., Xie, Y., Li, S., 2020. High-throughput estimation of crop traits: a review of ground and aerial phenotyping platforms. IEEE Geosci. Remote Sens. Mag. 0–0. https://doi.org/10.1109/MGRS.2020.2998816

Junker, A., Muraya, M.M., Weigelt-Fischer, K., Arana-Ceballos, F., Klukas, C., Melchinger, A.E., Meyer, R.C., Riewe, D., Altmann, T., 2015. Optimizing experimental procedures for quantitative evaluation of crop plant performance in high throughput phenotyping systems. Front. Plant Sci. 5, 1–21. https://doi.org/10.3389/fpls.2014.00770

Karlovska, A., Grinfelde, I., Alsina, I., Prieditis, G., Roze, D., 2015. Plant reflected spectra depending on biological characteristics and growth conditions, in: Proceedings of the 7th International Scientific Conference Rural Development 2015. Aleksandras Stulginskis University. https://doi.org/10.15544/rd.2015.045

Kataoka, T., Kaneko, T., Okamoto, H., Hata, S., 2003. Crop growth estimation system using machine vision. Proc. 2003 IEEE/ASME Int. Conf. Adv. Intell. Mechatronics (AIM 2003) 2, 1079–1083. https://doi.org/10.1109/AIM.2003.1225492

Kefauver, S.C., Vicente, R., Vergara-Díaz, O., Fernandez-Gallego, J.A., Kerfal, S., Lopez, A., Melichar, J.P.E., Serret Molins, M.D., Araus, J.L., 2017. Comparative UAV and field phenotyping to assess yield and nitrogen use efficiency in hybrid and conventional barley. Front. Plant Sci. 8, 1–15. https://doi.org/10.3389/fpls.2017.01733

Kennedy, K.A., Addison, P.A., 1987. Some considerations for the use of visual estimates of plant cover in biomonitoring. J. Ecol. 75, 151. https://doi.org/10.2307/2260541

Kirchgessner, N., Liebisch, F., Yu, K., Pfeifer, J., Friedli, M., Hund, A., Walter, A., 2017. The ETH field phenotyping platform FIP: a cable-suspended multi-sensor system. Funct. Plant Biol. 44, 154–168. https://doi.org/10.1071/FP16165

Kokaly, R.F., Asner, G.P., Ollinger, S. V., Martin, M.E., Wessman, C.A., 2009. Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies. Remote Sens. Environ. 113, S78–S91. https://doi.org/10.1016/j.rse.2008.10.018

Kruse, F.A., Lefkoff, A.B., Boardman, J.W., Heidebrecht, K.B., Shapiro, A.T., Barloon, P.J., Goetz, A.F.H., 1993. The spectral image processing system (SIPS)-interactive visualization and analysis of imaging spectrometer data. Remote Sens. Environ. 44, 145–163. https://doi.org/10.1016/0034-4257(93)90013-N

Laidig, F., Piepho, H.-P., Drobek, T., Meyer, U., 2014. Genetic and non-genetic long-term trends of 12 different crops in German official variety performance trials and on-farm yield trends. Theor. Appl. Genet. 127, 2599–2617. https://doi.org/10.1007/s00122-014-2402-z

Lambert, R., Peeters, A., Toussaint, B., 1999. LAI evolution of a perennial ryegrass crop estimated from the sum of temperatures in spring time. Agric. For. Meteorol. 97, 1–8. https://doi.org/10.1016/S0168-1923(99)00059-3

Lammerts van Bueren, E.T., Struik, P.C., van Eekeren, N., Nuijten, E., 2018. Towards resilience through systems-based plant breeding. A review. Agron. Sustain. Dev. https://doi.org/10.1007/s13593-018-0522-6

Lenaerts, B., Collard, B.C.Y., Demont, M., 2019. Review: Improving global food security through accelerated plant breeding. Plant Sci. 287, 110207. https://doi.org/10.1016/j.plantsci.2019.110207

Li, H., Feng, H., Guo, C., Yang, S., Huang, W., Xiong, X., Liu, J., Chen, G., Liu, Q., Xiong, L., Liu, K., Yang, W., 2020. High-throughput phenotyping accelerates the dissection of the dynamic genetic architecture of plant growth and yield improvement in rapeseed. Plant Biotechnol. J. 1–9. https://doi.org/10.1111/pbi.13396

Li, H., Rasheed, A., Hickey, L.T., He, Z., 2018. Fast-forwarding genetic gain. Trends Plant Sci. 23, 184–186. https://doi.org/10.1016/j.tplants.2018.01.007

Li, L., Zhang, Q., Huang, D., 2014. A review of imaging techniques for plant phenotyping. Sensors 14, 20078–20111. https://doi.org/10.3390/s141120078

Li, W., Niu, Z., Chen, H., Li, D., Wu, M., Zhao, W., 2016. Remote estimation of canopy height and aboveground biomass of maize using high-resolution stereo images from a low-cost unmanned aerial vehicle system. Ecol. Indic. 67, 637–648. https://doi.org/10.1016/j.ecolind.2016.03.036

Lillesand, T., Kiefer, R., Chipman, J., 2015. Remote sensing and image interpretation. Seventh Ed., John Wiley and Sons, Inc., New York.

Lindsey, A.J., Craft, J.C., Barker, D.J., 2020. Modeling canopy senescence to calculate soybean maturity date using NDVI. Crop Sci. 60, 172–180. https://doi.org/10.1002/csc2.20079

Liu, S., Baret, F., Abichou, M., Boudon, F., Thomas, S., Zhao, K., Fournier, C., Andrieu, B., Irfan, K., Hemmerlé, M., Solan, B. de, 2017a. Estimating wheat green area index from ground-based LiDAR measurement using a 3D canopy structure model. Agric. For. Meteorol. 247, 12–20. https://doi.org/10.1016/j.agrformet.2017.07.007

Liu, T., Li, R., Jin, X., Ding, J., Zhu, X., Sun, C., Guo, W., 2017b. Evaluation of seed emergence uniformity of mechanically sown wheat with UAV RGB imagery. Remote Sens. 9, 1241. https://doi.org/10.3390/rs9121241

Liu, Y., Mu, X., Wang, H., Yan, G., 2012. A novel method for extracting green fractional vegetation cover from digital images. J. Veg. Sci. 23, 406–418. https://doi.org/10.1111/j.1654-1103.2011.01373.x

Liu, Z.Y., Huang, J.F., Wu, X.H., Dong, Y.P., 2007. Comparison of vegetation indices and red-edge parameters for estimating grassland cover from canopy reflectance data. J. Integr. Plant Biol. 49, 299–306. https://doi.org/10.1111/j.1744-7909.2007.00401.x

Lootens, P., Ruttink, T., Rohde, A., Combes, D., Barre, P., Roldán-Ruiz, I., 2016. High-throughput phenotyping of lateral expansion and regrowth of spaced *Lolium perenne* plants using on-field image analysis. Plant Methods 12, 1–14. https://doi.org/10.1186/s13007-016-0132-8

López-Granados, F., Torres-Sánchez, J., Jiménez-Brenes, F.M., Arquero, O., Lovera, M., de Castro, A.I., 2019. An efficient RGB-UAV-based platform for field almond tree phenotyping: 3-D architecture and flowering traits. Plant Methods 15, 160. https://doi.org/10.1186/s13007-019-0547-0

López-López, A., Rosenblueth, M., Martínez, J., Martínez-Romero, E., 2010. Rhizobial symbioses in tropical legumes and non-legumes. Springer, Berlin, Heidelberg, pp. 163–184. https://doi.org/10.1007/978-3-642-05076-3_8

Lu, D., 2006. The potential and challenge of remote sensing-based biomass estimation. Int. J. Remote Sens. https://doi.org/10.1080/01431160500486732

Lugato, E., Bampa, F., Panagos, P., Montanarella, L., Jones, A., 2015. Potential carbon sequestration of European arable soils estimated by modelling a comprehensive set of management practices. Glob. Chang. Biol. 20, 3557–3567. https://doi.org/10.1111/gcb.12551

Luscier, J.D., Thompson, W.L., Wilson, J.M., Gorham, B.E., Dragut, L.D., 2006. Using digital photographs and object-based image analysis to estimate percent ground cover in vegetation plots. Front. Ecol. Environ. 4, 408–413. https://doi.org/10.1890/1540-9295(2006)4[408:UDPAOI]2.0.CO;2

Lussem, U., Bolten, A., Menne, J., Gnyp, M.L., Schellberg, J., Bareth, G., 2019. Estimating biomass in temperate grassland with high resolution canopy surface models from UAV-based RGB images and vegetation indices. J. Appl. Remote Sens. 13, 1. https://doi.org/10.1117/1.jrs.13.034525

Lynch, T.M.H., Barth, S., Dix, P.J., Grogan, D., Grant, J., Grant, O.M., 2015. Ground cover assessment of perennial ryegrass using digital imaging. Agron. J. 107, 2347. https://doi.org/10.2134/agronj15.0185

Madec, S., Baret, F., de Solan, B., Thomas, S., Dutartre, D., Jezequel, S., Hemmerlé, M., Colombeau, G., Comar, A., 2017. High-throughput phenotyping of plant height: comparing Unmanned Aerial Vehicles and ground LiDAR estimates. Front. Plant Sci. 8, 1–14. https://doi.org/10.3389/fpls.2017.02002

Maes, W.H., Steppe, K., 2019. Perspectives for remote sensing with Unmanned Aerial Vehicles in precision agriculture. Trends Plant Sci. 24, 152–164. https://doi.org/10.1016/j.tplants.2018.11.007

Maes, W.H., Steppe, K., 2012. Estimating evapotranspiration and drought stress with ground-based thermal remote sensing in agriculture: a review. J. Exp. Bot. 63, 4671–4712. https://doi.org/10.1093/jxb/ers165

Maimaitijiang, M., Sagan, V., Sidike, P., Hartling, S., Esposito, F., Fritschi, F.B., 2020. Soybean yield prediction from UAV using multimodal data fusion and deep learning. Remote Sens. Environ. 237, 111599. https://doi.org/10.1016/j.rse.2019.111599

Maimaitijiang, M., Sagan, V., Sidike, P., Maimaitiyiming, M., Hartling, S., Peterson, K.T., Maw, M.J.W., Shakoor, N., Mockler, T., Fritschi, F.B., 2019. Vegetation Index weighted Canopy Volume Model (CVM VI) for soybean biomass

estimation from Unmanned Aerial System-based RGB imagery. ISPRS J. Photogramm. Remote Sens. 151, 27–41. https://doi.org/10.1016/j.isprsjprs.2019.03.003

Malambo, L., Popescu, S.C., Murray, S.C., Putman, E., Pugh, N.A., Horne, D.W., Richardson, G., Sheridan, R., Rooney, W.L., Avant, R., Vidrine, M., McCutchen, B., Baltensperger, D., Bishop, M., 2018. Multitemporal field-based plant height estimation using 3D point clouds generated from small Unmanned Aerial Systems high-resolution imagery. Int. J. Appl. Earth Obs. Geoinf. 64, 31–42. https://doi.org/10.1016/j.jag.2017.08.014

Mao, W., Wang, Yiming, Wang, Yueqing, 2003. Real-time detection of between-row weeds using machine vision. 2003 ASAE Annu. Meet. 0300, 1.

Masuda, T., Goldsmith, P.D., 2009. World soybean production: area harvested, yield, and long-term projections. Int. Food Agribus. Manag. Rev. 12, 143–162. https://doi.org/10.22004/ag.econ.92573

Masuka, B., Magorokosho, C., Olsen, M., Atlin, G.N., Bänziger, M., Pixley, K. V., Vivek, B.S., Labuschagne, M., Matemba-Mutasa, R., Burgueño, J., Macrobert, J., Prasanna, B.M., Das, B., Makumbi, D., Tarekegne, A., Crossa, J., Zaman-Allah, M., van Biljon, A., Cairns, J.E., 2017. Gains in maize genetic improvement in Eastern and Southern Africa: II. CIMMYT Open-Pollinated Variety Breeding Pipeline. Crop Sci. 57, 180–191. https://doi.org/10.2135/cropsci2016.05.0408

Matese, A., Di Gennaro, S.F., Berton, A., 2017. Assessment of a canopy height model (CHM) in a vineyard using UAV-based multispectral imaging. Int. J. Remote Sens. 38, 2150–2160. https://doi.org/10.1080/01431161.2016.1226002

Matese, A., Toscano, P., Di Gennaro, S.F., Genesio, L., Vaccari, F.P., Primicerio, J., Belli, C., Zaldei, A., Bianconi, R., Gioli, B., 2015. Intercomparison of UAV, aircraft and satellite remote sensing platforms for precision viticulture. Remote Sens. 7, 2971–2990. https://doi.org/10.3390/rs7030297

Mathieu, A., Fiorelli, J.-L., 1985. Utilisation d'un herbomètre pour l'interprétation du déroulement d'un pâturage. Possibilités d'estimation de la production de matière sèche ou de caractérisation de l'état de l'herbe. Fourrages 105, 3–29.

Matias, F.I., Caraza-Harter, M. V., Endelman, J.B., 2020. FIELDimageR: An R package to analyze orthomosaic images from agricultural field trials. Plant Phenome J. 3. https://doi.org/10.1002/ppj2.20005

McDonagh, J., O'Donovan, M., McEvoy, M., Gilliland, T.J., 2016. Genetic gain in perennial ryegrass (*Lolium perenne*) varieties 1973 to 2013. Euphytica 212, 187–199. https://doi.org/10.1007/s10681-016-1754-7

Mesas-Carrascosa, F.J., Torres-Sánchez, J., Clavero-Rumbao, I., García-Ferrer, A., Peña, J.M., Borra-Serrano, I., López-Granados, F., 2015. Assessing optimal flight parameters for generating accurate multispectral orthomosaicks by UAV to support site-specific crop management. Remote Sens. 7, 12793–12814. https://doi.org/10.3390/rs71012793

Meyer, G.E., Hindman, T.W., Laksmi, K., 1999. Machine vision detection parameters for plant species identification, in: Meyer, G.E., DeShazer, J.A. (Eds.), . International Society for Optics and Photonics, pp. 327–335. https://doi.org/10.1117/12.336896

Meyer, G.E., Neto, J.C., 2008. Verification of color vegetation indices for automated crop imaging applications. Comput. Electron. Agric. 63, 282–293. https://doi.org/10.1016/j.compag.2008.03.009

Miladinović, J., Vidić, M., Dordević, V., Balešević-tubić, S., 2015. New trends in plant breeding - example of soybean. Genetika 47, 131–142. https://doi.org/10.2298/GENSR1501131M

Milberg, P., Bergstedt, J., Fridman, J., Odell, G., Westerberg, L., 2008. Observer bias and random variation in vegetation monitoring data. J. Veg. Sci. 19, 633–644. https://doi.org/10.3170/2008-8-18423

Mir, R.R., Reynolds, M., Pinto, F., Khan, M.A., Bhat, M.A., 2019. High-throughput phenotyping for crop improvement in the genomics era. Plant Sci. 282, 60–72. https://doi.org/10.1016/j.plantsci.2019.01.007

Mohammed, G.H., Noland, T.L., Irving, D., Sampson, P.H., Zarco-Tejada, P.J., Miller, J.R., 2000. Natural and stress-induced effects on leaf spectral reflectance in Ontario species. Science Development and Transfer, Ontario Ministry of Natural Resources.

Moreira, F.F., Hearst, A.A., Cherkauer, K.A., Rainey, K.M., 2019. Improving the efficiency of soybean breeding with high-throughput canopy phenotyping. Plant Methods 15, 1–9. https://doi.org/10.1186/s13007-019-0519-4

Mullan, D.J., Reynolds, M.P., 2010. Quantifying genetic effects of ground cover on soil water evaporation using digital imaging. Funct. Plant Biol. 37, 703. https://doi.org/10.1071/FP09277

Murchie, E.H., Kefauver, S., Araus, J.L., Muller, O., Rascher, U., Flood, P.J., Lawson, T., 2018. Measuring the dynamic photosynthome. Ann. Bot. 122, 207–220. https://doi.org/10.1093/aob/mcy087

Nakagami, K., 2016. Effects of sites and years on the coefficients of rising plate meter calibration under varying coefficient models. Grassl. Sci. 62, 128–132. https://doi.org/10.1111/grs.12117

Näsi, R., Viljanen, N., Kaivosoja, J., Alhonoja, K., Hakala, T., Markelin, L., Honkavaara, E., 2018. Estimating biomass and nitrogen amount of barley and grass using UAV and aircraft based spectral and photogrammetric 3D features. Remote Sens. 10, 1–32. https://doi.org/10.3390/rs10071082

Neumann, K., Klukas, C., Friedel, S., Rischbeck, P., Chen, D., Entzian, A., Stein, N., Graner, A., Kilian, B., 2015. Dissecting spatiotemporal biomass accumulation in barley under different water regimes using high-throughput image analysis. Plant, Cell Environ. 38, 1980–1996. https://doi.org/10.1111/pce.12516

Niu, Y., Zhang, L., Zhang, H., Han, W., Peng, X., 2019. Estimating above-ground biomass of maize using features derived from UAV-based RGB imagery. Remote Sens. 11, 1–21. https://doi.org/10.3390/rs11111261

O'Mara, F.P., 2012. The role of grasslands in food security and climate change. Ann. Bot. 110, 1263–1270. https://doi.org/10.1093/aob/mcs209

Pádua, L., Vanko, J., Hruška, J., Adão, T., Sousa, J.J., Peres, E., Morais, R., 2017. UAS, sensors, and data processing in agroforestry: a review towards practical applications. Int. J. Remote Sens. 38, 2349–2391. https://doi.org/10.1080/01431161.2017.1297548

Pajares, G., 2015. Overview and current status of remote sensing applications based on Unmanned Aerial Vehicles (UAVs). Photogramm. Eng. Remote Sensing 81, 281–329. https://doi.org/10.14358/PERS.81.4.281

Pandey, P., Ge, Y., Stoerger, V., Schnable, J.C., 2017. High throughput in vivo analysis of plant leaf chemical properties using hyperspectral imaging. Front. Plant Sci. 8, 1–12. https://doi.org/10.3389/fpls.2017.01348

Peña, J.M., de Castro, A.I., Torres-Sánchez, J., Andújar, D., Martín, C.S., Dorado, J., Fernández-Quintanilla, C., López-Granados, F., 2018. Estimating tree height and biomass of a poplar plantation with image-based UAV technology. AIMS Agric. Food 3, 313–323. https://doi.org/10.3934/AGRFOOD.2018.3.313

Peng, B., Guan, K., Tang, J., Ainsworth, E.A., Asseng, S., Bernacchi, C.J., Cooper, M., Delucia, E.H., Elliott, J.W., Ewert, F., Grant, R.F., Gustafson, D.I., Hammer, G.L., Jin, Z., Jones, J.W., Kimm, H., Lawrence, D.M., Li, Y., Lombardozzi, D.L., Marshall-Colon, A., Messina, C.D., Ort, D.R., Schnable, J.C., Vallejos, C.E., Wu, A., Yin, X., Zhou, W., 2020.

Towards a multiscale crop modelling framework for climate change adaptation assessment. Nat. Plants. https://doi.org/10.1038/s41477-020-0625-3

Pereyra-Irujo, G.A., Gasco, E.D., Peirone, L.S., Aguirrezábal, L.A.N., 2012. GlyPh: a low-cost platform for phenotyping plant growth and water use. Funct. Plant Biol. 39, 905. https://doi.org/10.1071/FP12052

Philipp, I., Rath, T., 2002. Improving plant discrimination in image processing by use of different colour space transformations. Comput. Electron. Agric. 35, 1–15. https://doi.org/10.1016/S0168-1699(02)00050-9

Pieruschka, R., Schurr, U., 2019. Plant phenotyping: past, present, and future. Plant Phenomics 2019, 7507131. https://doi.org/10.34133/2019/7507131

Ping, J., Liu, Y., Sun, L., Zhao, M., Li, Y., She, M., Sui, Y., Lin, F., Liu, X., Tang, Z., Nguyen, H., Qiu, L., Tian, Z., Nelson, R.L., Clemente, T.E., Specht, J.E., Ma, J., 2014. Dt2 is a gain-of-function MADS-domain factor gene that specifies semideterminacy in soybean. Plant Cell 26, 2831–2842. https://doi.org/10.1105/tpc.114.126938

Pittman, J.J., Arnall, D.B., Interrante, S.M., Moffet, C.A., Butler, T.J., 2015. Estimation of biomass and canopy height in bermudagrass, alfalfa, and wheat using ultrasonic, laser, and spectral sensors. Sensors 15, 2920–2943. https://doi.org/10.3390/s150202920

Pla, M., Bota, G., Duane, A., Balagué, J., Curcó, A., Gutiérrez, R., Brotons, L., 2019. Calibrating Sentinel-2 imagery with multispectral UAV derived information to quantify damages in Mediterranean rice crops caused by Western Swamphen (*Porphyrio porphyrio*). Drones 3, 45. https://doi.org/10.3390/drones3020045

Pointer, M.R., 2009. A comparison of the CIE 1976 colour spaces. Color Res. Appl. 6, 108–118. https://doi.org/10.1002/col.5080060212

Poorter, H., Fiorani, F., Pieruschka, R., Wojciechowski, T., van der Putten, W.H., Kleyer, M., Schurr, U., Postma, J., 2016. Pampered inside, pestered outside? Differences and similarities between plants growing in controlled conditions and in the field. New Phytol. 212, 838–855. https://doi.org/10.1111/nph.14243

Poorter, H., Fiorani, F., Stitt, M., Schurr, U., Finck, A., Gibon, Y., Usadel, B., Munns, R., Atkin, O.K., Tardieu, F., Pons, T.L., 2012. The art of growing plants for experimental purposes: a practical guide for the plant biologist. Funct. Plant Biol. 39, 821–838. https://doi.org/10.1071/FP12028

Popay, A.J., Hume, D.E., 2011. Endophytes improve ryegrass persistence by controlling insects. Pasture Persistence – Grassl. Res. Pract. Ser. 15, 149–156.

Possoch, M., Bieker, S., Hoffmeister, D., Bolten, A.A., Schellberg, J., Bareth, G., 2016. Multi-Temporal crop surface models combined with the RGB vegetation index from UAV-based images for forage monitoring in grassland. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch. 2016-Janua, 991–998. https://doi.org/10.5194/isprsarchives-XLI-B1-991-2016

Potgieter, A.B., Watson, J., Eldridge, M., Laws, K., George-Jaeggli, B., Hunt, C., Borrell, A., Mace, E., Chapman, S.C., Jordan, D.R., Hammer, G.L., 2018. Determining crop growth dynamics in sorghum breeding trials through remote and proximal sensing technologies, in: International Geoscience and Remote Sensing Symposium (IGARSS). Institute of Electrical and Electronics Engineers Inc., pp. 8244–8247. https://doi.org/10.1109/IGARSS.2018.8519296

Pourazar, H., Samadzadegan, F., Dadrass Javan, F., 2019. Aerial multispectral imagery for plant disease detection: radiometric calibration necessity assessment. Eur. J. Remote Sens. 52, 17–31. https://doi.org/10.1080/22797254.2019.1642143

Pradal, C., Artzet, S., Chopard, J., Dupuis, D., Fournier, C., Mielewczik, M., Nègre, V., Neveu, P., Parigot, D., Valduriez, P., Cohen-Boulakia, S., 2017. InfraPhenoGrid: a scientific workflow infrastructure for plant phenomics on the Grid. Futur. Gener. Comput. Syst. 67, 341–353. https://doi.org/10.1016/j.future.2016.06.002

Prasad, B., Carver, B.F., Stone, M.L., Babar, M.A., Raun, W.R., Klatt, A.R., 2007. Genetic analysis of indirect selection for winter wheat grain yield using spectral reflectance indices. Crop Sci. 47, 1416–1425. https://doi.org/10.2135/cropsci2006.08.0546

Probst, A., Gatziolis, D., Strigul, N., 2018. Intercomparison of photogrammetry software for three-dimensional vegetation modelling. R. Soc. Open Sci. 5. https://doi.org/10.1098/rsos.172192

Pugh, N.A., Horne, D.W., Murray, S.C., Carvalho, G., Malambo, L., Jung, J., Chang, A., Maeda, M., Popescu, S., Chu, T., Starek, M.J., Brewer, M.J., Richardson, G., Rooney, W.L., 2018. Temporal estimates of crop growth in sorghum and maize breeding enabled by Unmanned Aerial Systems. Tppj 1, 0. https://doi.org/10.2135/tppj2017.08.0006

Rahaman, M.M., Chen, D., Gillani, Z., Klukas, C., Chen, M., 2015. Advanced phenotyping and phenotype data analysis for the study of plant growth and development. Front. Plant Sci. 6, 1–15. https://doi.org/10.3389/fpls.2015.00619

Rallo, L., Barranco, D., Díez, C.M., Rallo, P., Suárez, M.P., Trapero, C., Alfaro, F.P., 2018. Strategies for olive (Olea europaea l.) breeding: Cultivated genetic resources and crossbreeding, in: Advances in Plant Breeding Strategies: Fruits. Springer International Publishing, pp. 555–600. https://doi.org/10.1007/978-3-319-91944-7_14

Rasmussen, J., Ntakos, G., Nielsen, J., Svensgaard, J., Poulsen, R.N., Christensen, S., 2016. Are vegetation indices derived from consumer-grade cameras mounted on UAVs sufficiently reliable for assessing experimental plots? Eur. J. Agron. 74, 75–92. https://doi.org/10.1016/j.eja.2015.11.026

Raymond, H.E., Cavigelli, M., Daughtry, C.S.T., E., M.J., Walthall, C.L., 2005. Evaluation of digital photography from model aircraft for remote sensing of crop biomass and nitrogen status. Precis. Agric. 6, 359–378. https://doi.org/10.1007/s11119-005-2324-5

Rebetzke, G.J., Jimenez-Berni, J., Fischer, R.A., Deery, D.M., Smith, D.J., 2019. Review: high-throughput phenotyping to enhance the use of crop genetic resources. Plant Sci. 282, 40–48. https://doi.org/10.1016/j.plantsci.2018.06.017

Reynolds, D., Baret, F., Welcker, C., Bostrom, A., Ball, J., Cellini, F., Lorence, A., Chawade, A., Khafif, M., Noshita, K., Mueller-Linow, M., Zhou, J., Tardieu, F., 2019. What is cost-efficient phenotyping? Optimizing costs for different scenarios. Plant Sci. 282, 14–22. https://doi.org/10.1016/j.plantsci.2018.06.015

Reynolds, M., Chapman, S., Crespo-Herrera, L., Molero, G., Mondal, S., Pequeno, D.N.L., Pinto, F., Pinera-Chavez, F.J., Poland, J., Rivera-Amado, C., Saint Pierre, C., Sukumaran, S., 2020. Breeder friendly phenotyping. Plant Sci. https://doi.org/10.1016/j.plantsci.2019.110396

Richards, R.A., 2000. Selectable traits to increase crop photosynthesis and yield of grain crops. J. Exp. Bot. 51, 447–458. https://doi.org/10.1093/jexbot/51.suppl_1.447

Richards, R.A., Rebetzke, G.J., Watt, M., Condon, A.G. (Tony), Spielmeyer, W., Dolferus, R., 2010. Breeding for improved water productivity in temperate cereals: phenotyping, quantitative trait loci, markers and the selection environment. Funct. Plant Biol. 37, 85. https://doi.org/10.1071/FP09219

Richardson, M.D., Karcher, D.E., Purcell, L.C., 2001. Quantifying turfgrass cover using digital image analaysis. Crop Sci. 41, 1884–1888. https://doi.org/10.2135/cropsci2001.1884

Roitsch, T., Cabrera-Bosquet, L., Fournier, A., Ghamkhar, K., Jiménez-Berni, J., Pinto, F., Ober, E.S., 2019. Review: new sensors and data-driven approaches—A path to next generation phenomics. Plant Sci. 282, 2–10. https://doi.org/10.1016/j.plantsci.2019.01.011

Roth, L., Aasen, H., Walter, A., Liebisch, F., 2018a. Extracting leaf area index using viewing geometry effects—A new perspective on high-resolution unmanned aerial system photography. ISPRS J. Photogramm. Remote Sens. 141, 161–175. https://doi.org/10.1016/j.isprsjprs.2018.04.012

Roth, L., Hund, A., Aasen, H., 2018b. PhenoFly planning tool: flight planning for high-resolution optical remote sensing with Unmanned Aerial Systems. Plant Methods 14, 1–21. https://doi.org/10.1186/s13007-018-0376-6

Roth, L., Streit, B., 2017. Predicting cover crop biomass by lightweight UAS-based RGB and NIR photography: an applied photogrammetric approach. Precis. Agric. 1–22. https://doi.org/10.1007/s11119-017-9501-1

Royo, C., Aparicio, N., Villegas, D., Casadesus, J., Monneveux, P., Araus, J.L., 2003. Usefulness of spectral reflectance indices as durum wheat yield predictors under contrasting Mediterranean conditions. Int. J. Remote Sens. 24, 4403–4419. https://doi.org/10.1080/0143116031000150059

Rueda-Ayala, V.P.; Peña, J.M.; Höglind, M.; Bengochea-Guevara, J.M.; Andújar, D., 2019. Comparing UAV-Based Technologies and RGB-D Reconstruction Methods for Plant Height and Biomass Monitoring on Grass Ley. Sensors 19, 535. https://doi.org/10.3390/s19030535

Rutkoski, J., Poland, J., Mondal, S., Autrique, E., Pérez, L.G., Crossa, J., Reynolds, M., Singh, R., 2016. Canopy temperature and vegetation indices from high-throughput phenotyping improve accuracy of pedigree and genomic selection for grain yield in wheat. G3: Genes, Genomes, Genetics 6, 2799–2808. https://doi.org/10.1534/g3.116.032888

Salamí, E., Barrado, C., Pastor, E., 2014. UAV flight experiments applied to the remote sensing of vegetated areas. Remote Sens. 6, 11051–11081. https://doi.org/10.3390/rs61111051

Sampoux, J.P., Baudouin, P., Bayle, B., Béguier, V., Bourdon, P., Chosson, J.F., Deneufbourg, F., Galbrun, C., Ghesquière, M., Noël, D., Pietraszek, W., Tharel, B., Viguié, A., 2011. Breeding perennial grasses for forage usage: An experimental assessment of trait changes in diploid perennial ryegrass (*Lolium perenne* L.) cultivars released in the last four decades. F. Crop. Res. 123, 117–129. https://doi.org/10.1016/j.fcr.2011.05.007

Sanderson, M.A., Rotz, C.A., Fultz, S.W., Rayburn, E.B., 2001. Estimating forage mass with a commercial capacitance meter, rising plate meter, and pasture ruler. Agron. J. 93, 1281–1286. https://doi.org/10.2134/agronj2001.1281

Sankaran, S., Khot, L.R., Carter, A.H., 2015a. Field-based crop phenotyping: multispectral aerial imaging for evaluation of winter wheat emergence and spring stand. Comput. Electron. Agric. 118, 372–379. https://doi.org/10.1016/j.compag.2015.09.001

Sankaran, S., Khot, L.R., Espinoza, C.Z., Jarolmasjed, S., Sathuvalli, V.R., Vandemark, G.J., Miklas, P.N., Carter, A.H., Pumphrey, M.O., Knowles, R.R.N., Pavek, M.J., 2015b. Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: a review. Eur. J. Agron. 70, 112–123. https://doi.org/10.1016/j.eja.2015.07.004

Sayed, M.A., Schumann, H., Pillen, K., Naz, A.A., Léon, J., 2012. AB-QTL analysis reveals new alleles associated to proline accumulation and leaf wilting under drought stress conditions in barley (*Hordeum vulgare* L.). BMC Genet. 13, 61. https://doi.org/10.1186/1471-2156-13-61

Schellberg, J., Hill, M.J., Gerhards, R., Rothmund, M., Braun, M., 2008. Precision agriculture on grassland: applications, perspectives and constraints. Eur. J. Agron. 29, 59–71. https://doi.org/10.1016/j.eja.2008.05.005

Schellberg, J., Verbruggen, E., 2014. Frontiers and perspectives on research strategies in grassland technology. Crop Pasture Sci. 65, 508–523. https://doi.org/10.1071/CP13429

Schori, A., Charles, R., Peter, D., 2003. Soja: sélection, agronomie et production en Suisse. Rev. Suisse d'Agronomie 35, 69–75.

Schreuder, R., De Visser, C., 2014. EIP-AGRI Focus Group Protein Crops : final report. Eur. Innov. Partnersh. Agric. Product. Sustain. 48.

Segarra, J., Buchaillot, M.L., Araus, J.L., Kefauver, S.C., 2020. Remote sensing for precision agriculture: Sentinel-2 improved features and applications. Agronomy 10, 641. https://doi.org/10.3390/agronomy10050641

Shah, F., Wu, W., 2019. Soil and crop management strategies to ensure higher crop productivity within sustainable environments. Sustainability 11, 1485. https://doi.org/10.3390/su11051485

Shakoor, N., Lee, S., Mockler, T.C., 2017. High throughput phenotyping to accelerate crop breeding and monitoring of diseases in the field. Curr. Opin. Plant Biol. 38, 184–192. https://doi.org/10.1016/j.pbi.2017.05.006

Shi, Y., Thomasson, J.A., Murray, S.C., Pugh, N.A., Rooney, W.L., Shafian, S., Rajan, N., Rouze, G., Morgan, C.L.S., Neely, H.L., Rana, A., Bagavathiannan, M. V., Henrickson, J., Bowden, E., Valasek, J., Olsenholler, J., Bishop, M.P., Sheridan, R., Putman, E.B., Popescu, S., Burks, T., Cope, D., Ibrahim, A., McCutchen, B.F., Baltensperger, D.D., Avant, R. V., Vidrine, M., Yang, C., 2016. Unmanned Aerial Vehicles for high-throughput phenotyping and agronomic research. PLoS One 11, e0159781. https://doi.org/10.1371/journal.pone.0159781

Singh, A., Ganapathysubramanian, B., Singh, A.K., Sarkar, S., 2016. Machine learning for high-throughput stress phenotyping in plants. Trends Plant Sci. 21, 110–124. https://doi.org/10.1016/j.tplants.2015.10.015

Singh, R.J., Hymowitz, T., 1999. Soybean genetic resources and crop improvement, in: Genome. National Research Council of Canada, pp. 605–616. https://doi.org/10.1139/g99-039

Smith, M.W., Carrivick, J.L., Quincey, D.J., 2016. Structure from motion photogrammetry in physical geography. Prog. Phys. Geogr. Earth Environ. 40, 247–275. https://doi.org/10.1177/0309133315615805

Sonnentag, O., Hufkens, K., Teshera-Sterne, C., Young, A.M., Friedl, M., Braswell, B.H., Milliman, T., O'Keefe, J., Richardson, A.D., 2012. Digital repeat photography for phenological research in forest ecosystems. Agric. For. Meteorol. 152, 159–177. https://doi.org/10.1016/j.agrformet.2011.09.009

Stöcker, C., Bennett, R., Nex, F., Gerke, M., Zevenbergen, J., 2017. Review of the current state of UAV regulations. Remote Sens. 9, 459. https://doi.org/10.3390/rs9050459

Sugiyama, A., Ueda, Y., Takase, H., Yazaki, K., 2015. Do soybeans select specific species of *Bradyrhizobium* during growth? Commun. Integr. Biol. 8, e992734. https://doi.org/10.4161/19420889.2014.992734

Sun, M., Goggi, S.A., Matson, K., Palmer, R.G., Moore, K., Cianzio, S.R., 2015. Thin plate spline regression model used at early stages of soybean breeding to control field spatial variation. J. Crop Improv. 29, 333–352. https://doi.org/10.1080/15427528.2015.1026623

Tardieu, F., Cabrera-Bosquet, L., Pridmore, T., Bennett, M., 2017. Plant phenomics, from sensors to knowledge. Curr. Biol. 27, R770–R783. https://doi.org/10.1016/j.cub.2017.05.055

Tattaris, M., Reynolds, M.P., Chapman, S.C., 2016. A direct comparison of remote sensing approaches for high-throughput phenotyping in plant breeding. Front. Plant Sci. 7, 1–9. https://doi.org/10.3389/fpls.2016.01131

Taylor, N.L., Quesenberry, K.H., 1996. Red clover science. Kluwer Academic Publishers, Dordrecht.

Tewes, A., Schellberg, J., 2018. Towards remote estimation of radiation use efficiency in maize using UAV-based low-cost camera imagery. Agronomy 8, 16. https://doi.org/10.3390/agronomy8020016

Thenkabail, P.S., Smith, R.B., De Pauw, E., 2002. Evaluation of narrowband and broadband vegetation indices for determining optimal hyperspectral wavebands for agricultural crop characterization. Photogramm. Eng. Remote Sensing 68, 607-622.

Thomas, S., Kuska, M.T., Bohnenkamp, D., Brugger, A., Alisaac, E., Wahabzada, M., Behmann, J., Mahlein, A.-K., 2018. Benefits of hyperspectral imaging for plant disease detection and plant protection: a technical perspective. J. Plant Dis. Prot. 125, 5–20. https://doi.org/10.1007/s41348-017-0124-6

Thompson, A.L., Thorp, K.R., Conley, M., Andrade-Sanchez, P., Heun, J.T., Dyer, J.M., White, J.W., 2018. Deploying a proximal sensing cart to identify drought-adaptive traits in upland cotton for high-throughput phenotyping. Front. Plant Sci. 9, 507. https://doi.org/10.3389/fpls.2018.00507

Thorney, J.H.M., 1990. A new formulation of the logistic growth equation and its application to leaf area growth. Ann. Bot. 66, 309–311. https://doi.org/10.1093/oxfordjournals.aob.a088029

Tilly, N., Aasen, H., Bareth, G., 2015. Fusion of plant height and vegetation indices for the estimation of barley biomass. Remote Sens. 7, 11449–11480. https://doi.org/10.3390/rs70911449

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. J. Appl. Remote Sens. 8, 083671. https://doi.org/10.1117/1.JRS.8.083671

Tisné, S., Serrand, Y., Bach, L., Gilbault, E., Ben Ameur, R., Balasse, H., Voisin, R., Bouchez, D., Durand-Tardif, M., Guerche, P., Chareyron, G., Da Rugna, J., Camilleri, C., Loudet, O., 2013. Phenoscope: an automated large-scale phenotyping platform offering high spatial homogeneity. Plant J. 74, 534–544. https://doi.org/10.1111/tpj.12131

Togeiro de Alckmin, G., Kooistra, L., Rawnsley, R., Lucieer, A., 2020. Comparing methods to estimate perennial ryegrass biomass: canopy height and spectral vegetation indices. Precis. Agric. 1–21. https://doi.org/10.1007/s11119-020-09737-z

Torres-Sánchez, J., de Castro, A.I., Peña, J.M., Jiménez-Brenes, F.M., Arquero, O., Lovera, M., López-Granados, F., 2018a. Mapping the 3D structure of almond trees using UAV acquired photogrammetric point clouds and object-based image analysis. Biosyst. Eng. 176, 172–184. https://doi.org/10.1016/j.biosystemseng.2018.10.018

Torres-Sánchez, J., López-Granados, F., Borra-Serrano, I., Peña, J.M., 2018b. Assessing UAV-collected image overlap influence on computation time and digital surface model accuracy in olive orchards. Precis. Agric. 19, 115–133. https://doi.org/10.1007/s11119-017-9502-0

Torres-Sánchez, J., Peña, J.M., de Castro, A.I., López-Granados, F., 2014. Multi-temporal mapping of the vegetation fraction in early-season wheat fields using images from UAV. Comput. Electron. Agric. 103, 104–113. https://doi.org/10.1016/j.compag.2014.02.009

Tripodi, P., Massa, D., Venezia, A., Cardi, T., 2018. Sensing technologies for precision phenotyping in vegetable crops: current status and future challenges. Agronomy 8. https://doi.org/10.3390/agronomy8040057

Ullman, S., 1979. The interpretation of structure from motion. Proc. R. Soc. London. Ser. B. Biol. Sci. 203, 405–426. https://doi.org/10.1098/rspb.1979.0006

Ustin, S.L., Gitelson, A.A., Jacquemoud, S., Schaepman, M., Asner, G.P., Gamon, J.A., Zarco-Tejada, P., 2009. Retrieval of foliar information about plant pigment systems from high resolution spectroscopy. Remote Sens. Environ. 113, S67–S77. https://doi.org/10.1016/j.rse.2008.10.019

Van De Vijver, R., Mertens, K., Heungens, K., Somers, B., Nuyttens, D., Borra-Serrano, I., Lootens, P., Roldán-Ruiz, I., Vangeyte, J., Saeys, W., 2020. In-field detection of *Alternaria solani* in potato crops using hyperspectral imaging. Comput. Electron. Agric. 168, 105106. https://doi.org/10.1016/j.compag.2019.105106

Varshney, R.K., Thudi, M., Pandey, M.K., Tardieu, F., Ojiewo, C., Vadez, V., Whitbread, A.M., Siddique, K.H.M., Nguyen, H.T., Carberry, P.S., Bergvinson, D., 2018. Accelerating genetic gains in legumes for the development of prosperous smallholder agriculture: integrating genomics, phenotyping, systems modelling and agronomy. J. Exp. Bot. 69, 3293–3312. https://doi.org/10.1093/jxb/ery088

Viljanen, N., Honkavaara, E., Näsi, R., Hakala, T., Niemeläinen, O., Kaivosoja, J., 2018. A novel machine learning method for estimating biomass of grass swards using a photogrammetric canopy height model, images and vegetation indices captured by a drone. Agriculture 8, 70. https://doi.org/10.3390/agriculture8050070

Virlet, N., Sabermanesh, K., Sadeghi-Tehran, P., Hawkesford, M.J., 2017. Field Scanalyzer: an automated robotic field phenotyping platform for detailed crop monitoring. Funct. Plant Biol. 44, 143–153. https://doi.org/10.1071/FP16163

Voorend, W., Lootens, P., Nelissen, H., Roldán-Ruiz, I., Inzé, D., Muylle, H., 2014. LEAF-E: a tool to analyze grass leaf growth using function fitting. Plant Methods 10, 1–13. https://doi.org/10.1186/1746-4811-10-37

Wachendorf, M., Fricke, T., Möckel, T., 2017. Remote sensing as a tool to assess botanical composition, structure, quantity and quality of temperate grasslands. Grass Forage Sci. 1–14. https://doi.org/10.1111/gfs.12312

Walter, A., Liebisch, F., Hund, A., 2015. Plant phenotyping: from bean weighing to image analysis. Plant Methods 11, 1–11. https://doi.org/10.1186/s13007-015-0056-8

Walter, A., Scharr, H., Gilmer, F., Zierer, R., Nagel, K.A., Ernst, M., Wiese, A., Virnich, O., Christ, M.M., Uhlig, B., Jünger, S., Schurr, U., 2007. Dynamics of seedling growth acclimation towards altered light conditions can be quantified via GROWSCREEN: a setup and procedure designed for rapid optical phenotyping of different plant species. New Phytol. 174, 447–455. https://doi.org/10.1111/j.1469-8137.2007.02002.x

Wang, J., Badenhorst, P., Phelan, A., Pembleton, L., Shi, F., Cogan, N., Spangenberg, G., Smith, K., 2019. Using sensors and Unmanned Aircraft Systems for high-throughput phenotyping of biomass in perennial ryegrass breeding trials. Front. Plant Sci. 10, 1–9. https://doi.org/10.3389/fpls.2019.01381

Watt, M., Fiorani, F., Usadel, B., Rascher, U., Muller, O., Schurr, U., 2020. Phenotyping: new windows into the plant for breeders. Annu. Rev. Plant Biol. 71, 689–712. https://doi.org/10.1146/annurev-arplant-042916-041124

Weiss, M., Jacob, F., Duveiller, G., 2020. Remote sensing for agricultural applications: a meta-review. Remote Sens. Environ. 236, 111402. https://doi.org/10.1016/j.rse.2019.111402

White, J.W., Andrade-Sanchez, P., Gore, M.A., Bronson, K.F., Coffelt, T.A., Conley, M.M., Feldmann, K.A., French, A.N., Heun, J.T., Hunsaker, D.J., Jenks, M.A., Kimball, B.A., Roth, R.L., Strand, R.J., Thorp, K.R., Wall, G.W., Wang, G., 2012. Field-based phenomics for plant genetics research. F. Crop. Res. 133, 101–112. https://doi.org/10.1016/j.fcr.2012.04.003

Wilkins, P.W., Humphreys, M.O., 2003. Progress in breeding perennial clovers for temperate agriculture. J. Agric. Sci. 140, 129–150. https://doi.org/10.1017/S0021859605005101

Woebbecke, D.M., Meyer, G.E., Von Bargen, K., Mortensen, D.A., 1995. Color indices for weed identification under various soil, residue, and lighting conditions. Trans. ASAE 38, 259–269. https://doi.org/10.13031/2013.27838

Xiang, H., Tian, L., 2011. Method for automatic georeferencing aerial remote sensing (RS) images from an Unmanned Aerial Vehicle (UAV) platform. Biosyst. Eng. 108, 104–113. https://doi.org/10.1016/j.biosystemseng.2010.11.003

Xu, Y., Li, P., Zou, C., Lu, Y., Xie, C., Zhang, X., Prasanna, B.M., Olsen, M.S., 2017. Enhancing genetic gain in the era of molecular breeding. J. Exp. Bot. https://doi.org/10.1093/jxb/erx135

Xue, J., Su, B., 2017. Significant remote sensing vegetation indices: a review of developments and applications. J. Sensors. https://doi.org/10.1155/2017/1353691

Yang, G., Liu, J., Zhao, C., Li, Zhenhong, Huang, Y., Yu, H., Xu, B., Yang, X., Zhu, D., Zhang, X., Zhang, R., Feng, H., Zhao, X., Li, Zhenhai, Li, H., Yang, H., 2017. Unmanned Aerial Vehicle remote sensing for field-based crop phenotyping: current status and perspectives. Front. Plant Sci. 8. https://doi.org/10.3389/fpls.2017.01111

Yang, S., Logan, J., Coffey, D.L., 1995. Mathematical formulae for calculating the base temperature for growing degree days. Agric. For. Meteorol. 74, 61–74. https://doi.org/10.1016/0168-1923(94)02185-M

Yang, W., Feng, H., Zhang, X., Zhang, J., Doonan, J.H., Batchelor, W.D., Xiong, L., Yan, J., 2020. Crop phenomics and high-throughput phenotyping: past decades, current challenges, and future perspectives. Mol. Plant 13, 187–214. https://doi.org/10.1016/j.molp.2020.01.008

Yao, H., Qin, R., Chen, X., 2019. Unmanned Aerial Vehicle for remote sensing applications - A review. Remote Sens. 11, 1–22. https://doi.org/10.3390/rs11121443

Yin, X., Goudriaan, J., Lantinga, E.A., Vos, J., Spiertz, H.J., 2003. A flexible sigmoid function of determinate growth. Ann. Bot. 91, 361–371. https://doi.org/10.1093/aob/mcg029

Yu, K., Kirchgessner, N., Grieder, C., Walter, A., Hund, A., 2017. An image analysis pipeline for automated classification of imaging light conditions and for quantification of wheat canopy cover time series in field phenotyping. Plant Methods 13. https://doi.org/10.1186/s13007-017-0168-4

Yu, N., Li, L., Schmitz, N., Tian, L.F., Greenberg, J.A., Diers, B.W., 2016. Development of methods to improve soybean yield estimation and predict plant maturity with an Unmanned Aerial Vehicle based platform. Remote Sens. Environ. 187, 91–101. https://doi.org/10.1016/j.rse.2016.10.005

Yuan, W., Wijewardane, N.K., Jenkins, S., Bai, G., Ge, Y., Graef, G.L., 2019. Early prediction of soybean traits through color and texture features of canopy RGB imagery. Sci. Rep. 9, 1–18. https://doi.org/10.1038/s41598-019-50480-x

Yue, J., Yang, G., Li, C., Li, Z., Wang, Y., Feng, H., Xu, B., 2017. Estimation of winter wheat above-ground biomass using Unmanned Aerial Vehicle-based snapshot hyperspectral sensor and crop height improved models. Remote Sens. 9. https://doi.org/10.3390/rs9070708

Zarco-Tejada, P.J., González-Dugo, V., Williams, L.E., Suárez, L., Berni, J.A.J., Goldhamer, D., Fereres, E., 2013. A PRI-based water stress index combining structural and chlorophyll effects: assessment using diurnal narrow-band

airborne imagery and the CWSI thermal index. Remote Sens. Environ. 138, 38–50. https://doi.org/10.1016/j.rse.2013.07.024

Zeide, B., 1993. Analysis of Growth Equations. For. Sci. 39, 594–616. https://doi.org/10.1093/forestscience/39.3.594

Zhang, C., Kovacs, J.M., 2012. The application of small Unmanned Aerial Systems for precision agriculture: a review. Precis. Agric. 13, 693–712. https://doi.org/10.1007/s11119-012-9274-5

Zhang, X., Huang, C., Wu, D., Qiao, F., Li, W., Duan, L., Wang, K., Xiao, Y., Chen, G., Liu, Q., Xiong, L., Yang, W., Yan, J., 2017. High-throughput phenotyping and QTL mapping reveals the genetic architecture of maize plant growth. Plant Physiol. 173, 1554–1564. https://doi.org/10.1104/pp.16.01516

Zhang, X., Zhao, J., Yang, G., Liu, J., Cao, J., Li, C., Zhao, X., Gai, J., 2019. establishment of plot-yield prediction models in soybean breeding programs using UAV-based hyperspectral remote sensing. Remote Sens. 11, 2752. https://doi.org/10.3390/rs11232752

Zhao, C., Zhang, Y., Du, J., Guo, X., Wen, W., Gu, S., Wang, J., Fan, J., 2019. Crop phenomics: current status and perspectives. Front. Plant Sci. 10. https://doi.org/10.3389/fpls.2019.00714

Zhou, Jing, Yungbluth, D., Vong, C.N., Scaboo, A., Zhou, Jianfeng, 2019. Estimation of maturity date of soybean breeding lines using UAV-based imagery. 2019 ASABE Annu. Int. Meet. https://doi.org/10.13031/aim.201900427

Ziliani, M., Parkes, S., Hoteit, I., McCabe, M., 2018. Intra-season crop height variability at commercial farm scales using a fixed-wing UAV. Remote Sens. 10, 2007. https://doi.org/10.3390/rs10122007

Appendices

Supplementary tables and figures of Chapter 1

Table A-1. Time required in each step from the UAV mission planning to the data extraction. Flight specifications: RGB 12mm at 30m with 80% overlap. 1.425 ha covered (0.575 trial). 104 images in total.

Task description	Person intervention (h:min)	Extra computer time (h:min)
Mission planning	0:10	
In field activities*		
Preparation of the flight	0:30	
GCPs location and measurement	0:30	
Flight operation	0:10	
Pre-processing of the images		
Uploading data	0:05	0:20
Image calibration	0:20	2:00
Image mosaicking	0:30**	6:30
Export Orthomosaic and DEM plus report generation	0:10	0:30
Quality inspection	0:10	
Spatial data extraction		
Build polygons	1:00***	
Calculate VIs	0:30	
Masking and segmentation	2:00	
Data extraction	0:20	
	6:15	9:20
Total (h:min)	**15:35**	

*two persons are needed

**preparation of the project + checking GCPs location

***only once, reusable when multiple flights are done

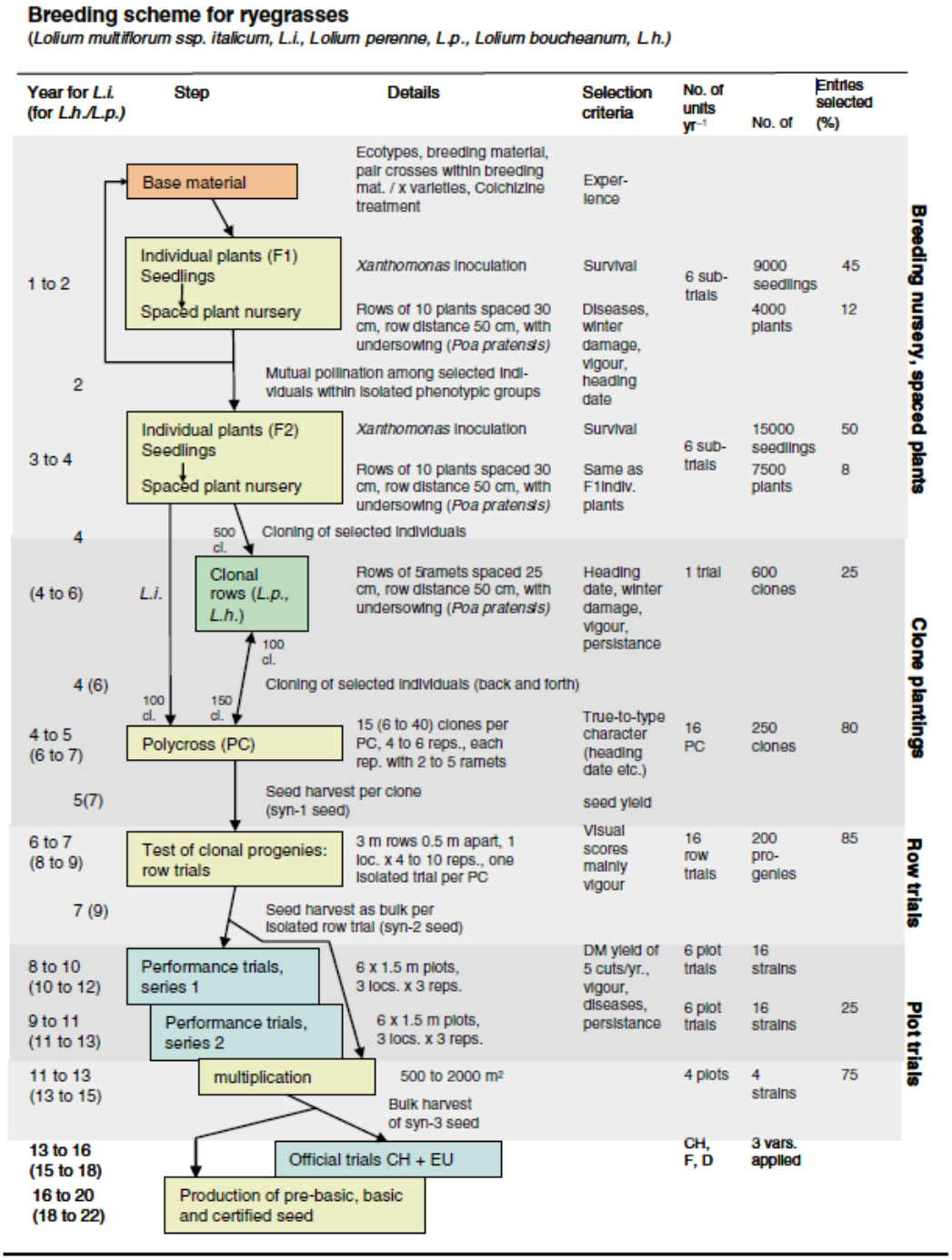


Figure A-1. Breeding cycle ryegrass - Generalized breeding scheme for ryegrasses describing the steps to perform at different scales and years. Source: Humphreys et al. (2010)

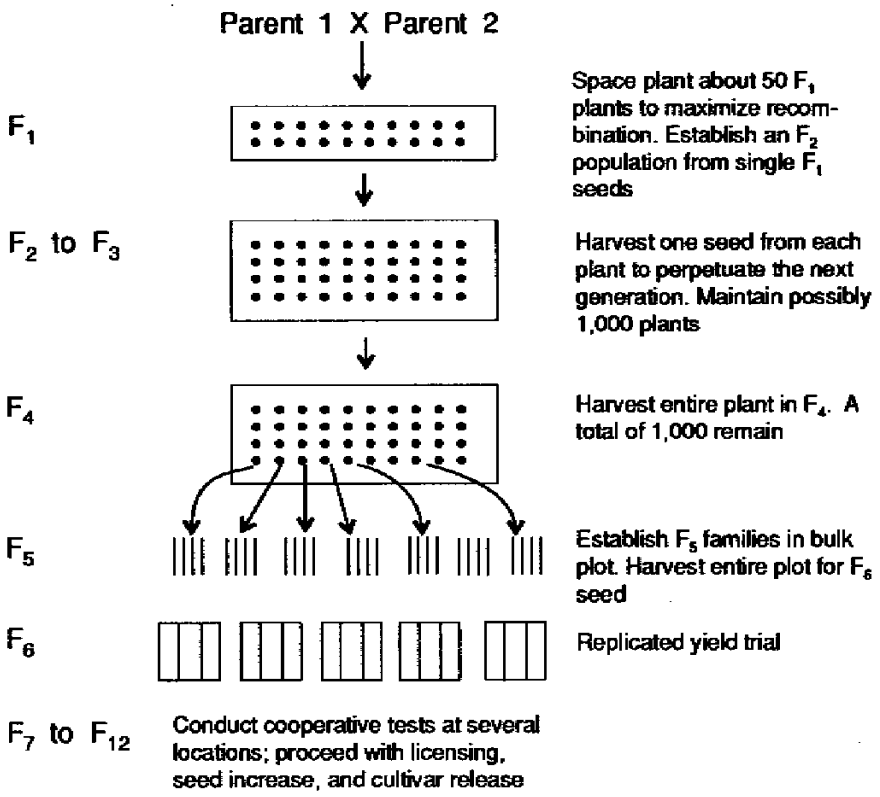


Figure A-2. Breeding cycle soybean - Generalized breeding scheme for soybean. Source: U. Guelph (dept. of plant agriculture) NAPB web seminars 2015

Supplementary figures of Chapter 4

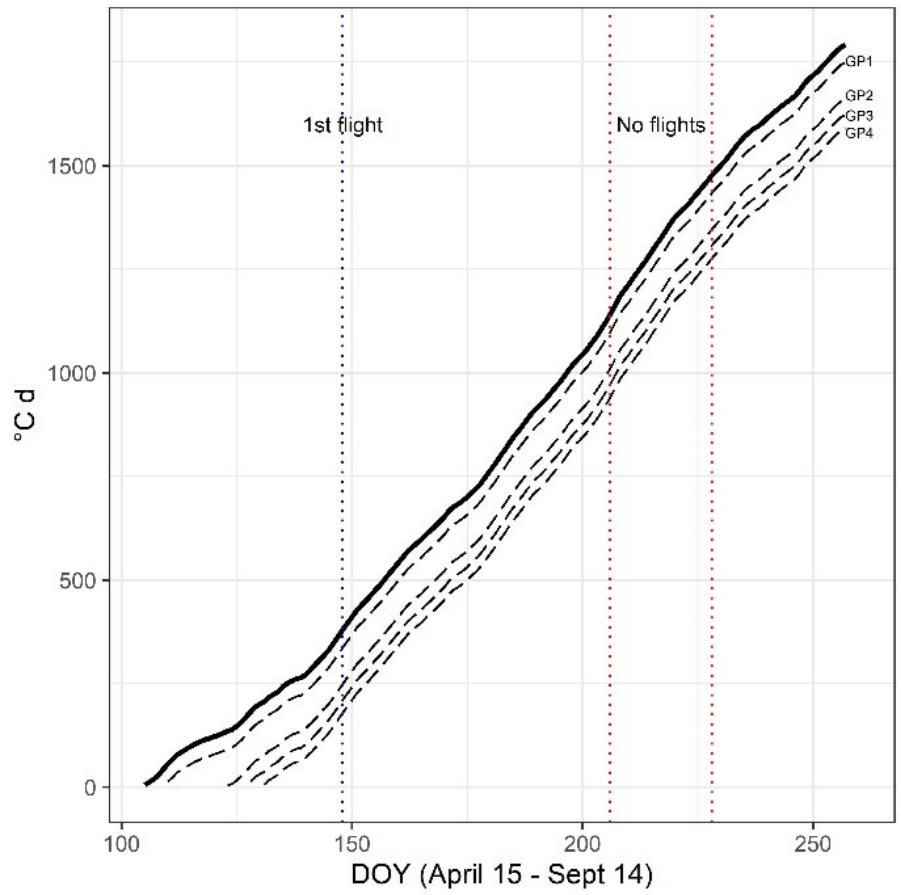


Figure A-3**.** Evolution of thermal time (°C d) during the growing season of the trial. First flight and the two weeks gap are indicated.

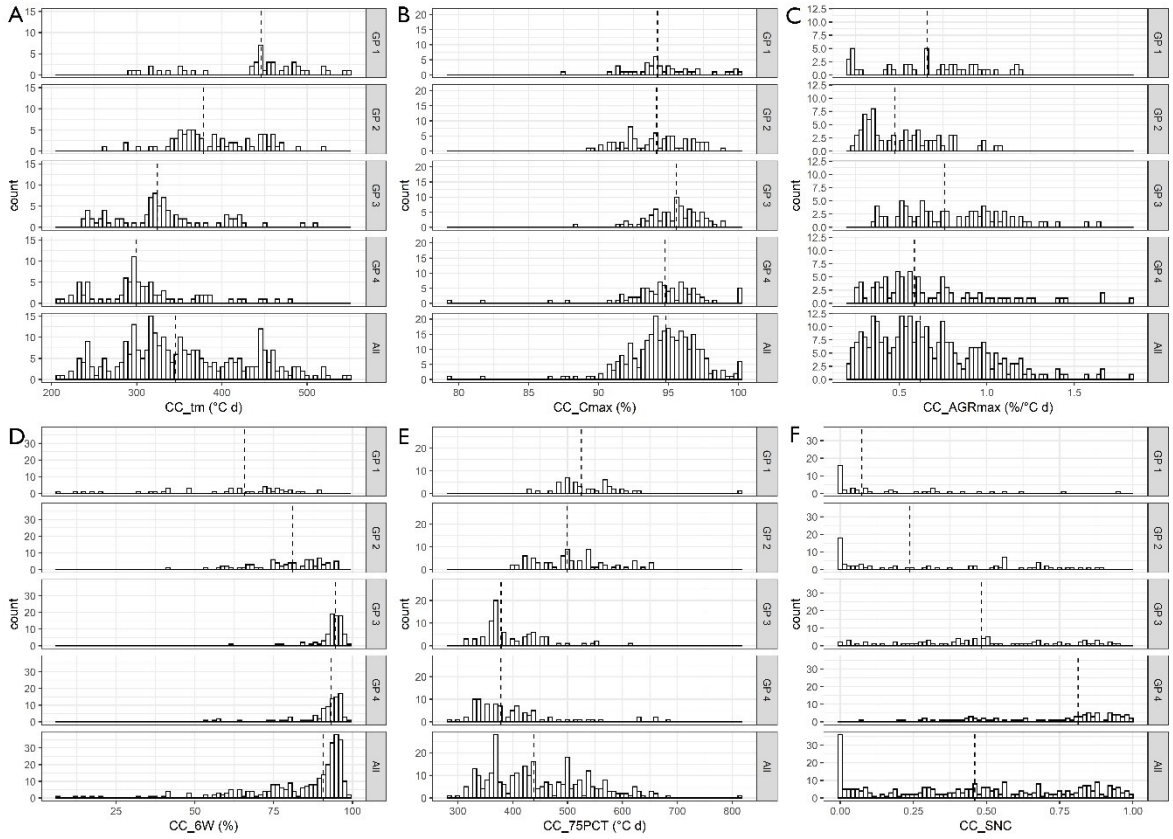


Figure A-4. Histograms with median values (dashed lines) of the phenotypic parameters (A-F) derived from the curve fitting of canopy cover data. The data is presented per group (GP 1-4) and all groups together (All). For the abbreviations see Table 4-1.

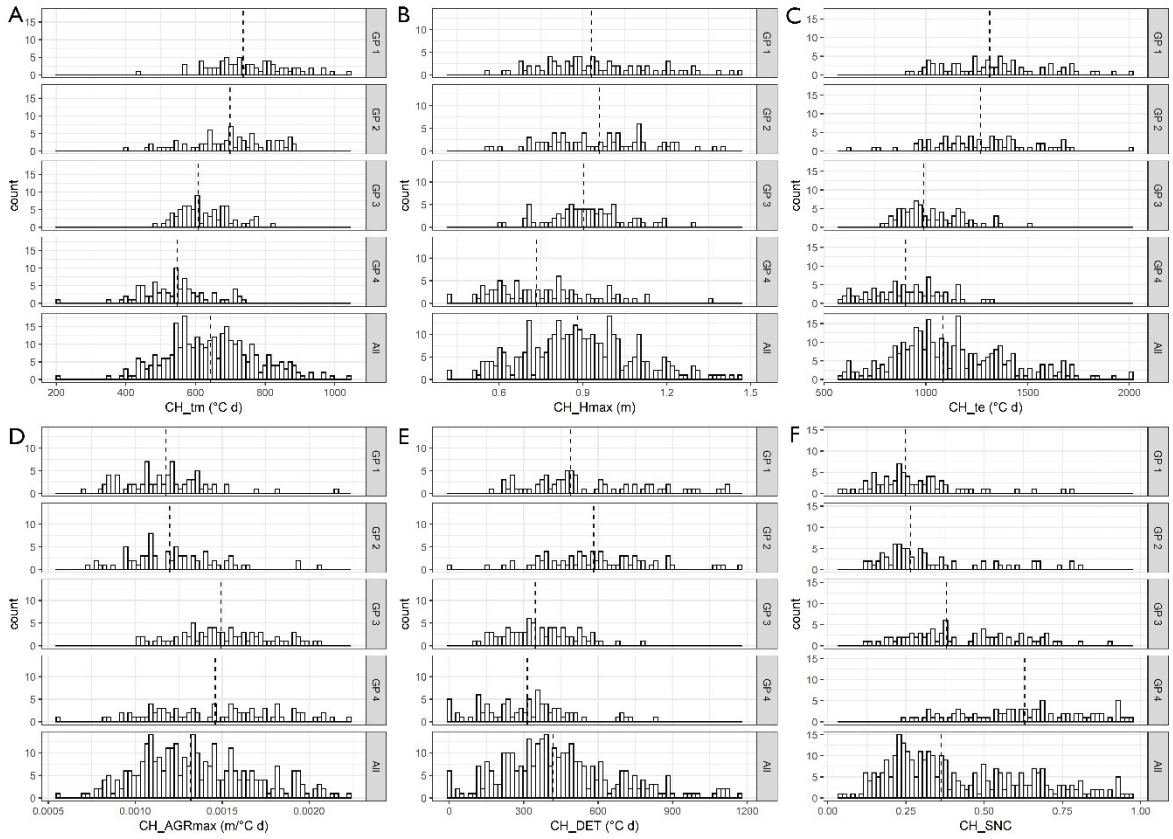


Figure A-5. Histograms with median values (dashed lines) of the phenotypic parameters (A-F) derived from the curve fitting of canopy height. The data is presented per group (GP 1-4) and all groups together (All). For the abbreviations see Table 4-1.

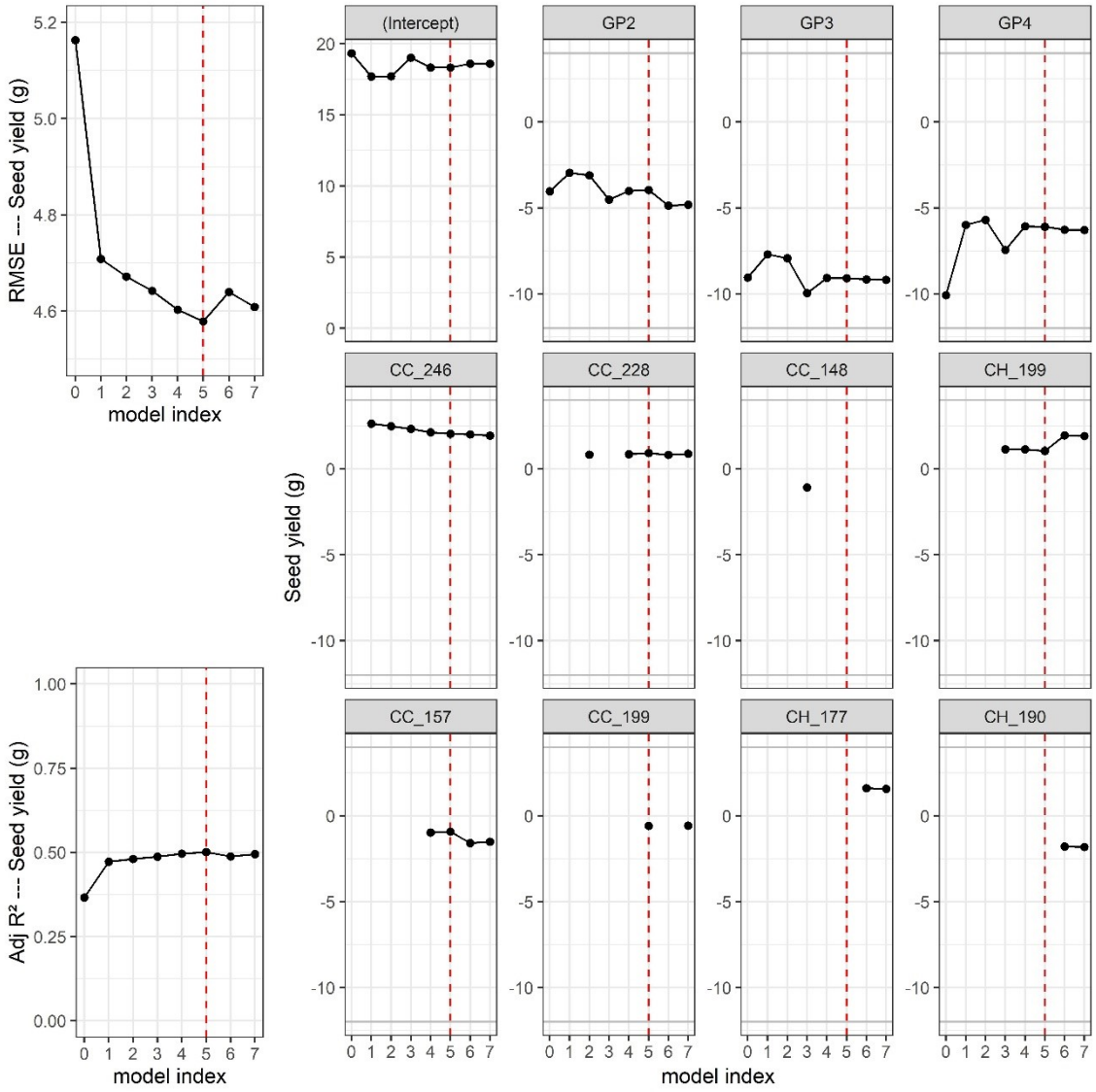


Figure A-6. MLR model building overview for SY stage with RMSE and Adj.R² evolution. The red dashed line indicates the selected (best) model with the lowest RMSE. The subset selection regression with normalized UAV time points included a 10 k-fold cross validation. In each subsequent step one variable was added and all possible combinations were evaluated. The model index = 0 included only the groups (GP1-4, GP1 is the reference group) as dummy variables. Graphs are provided per variable with the value of their respective coefficients.

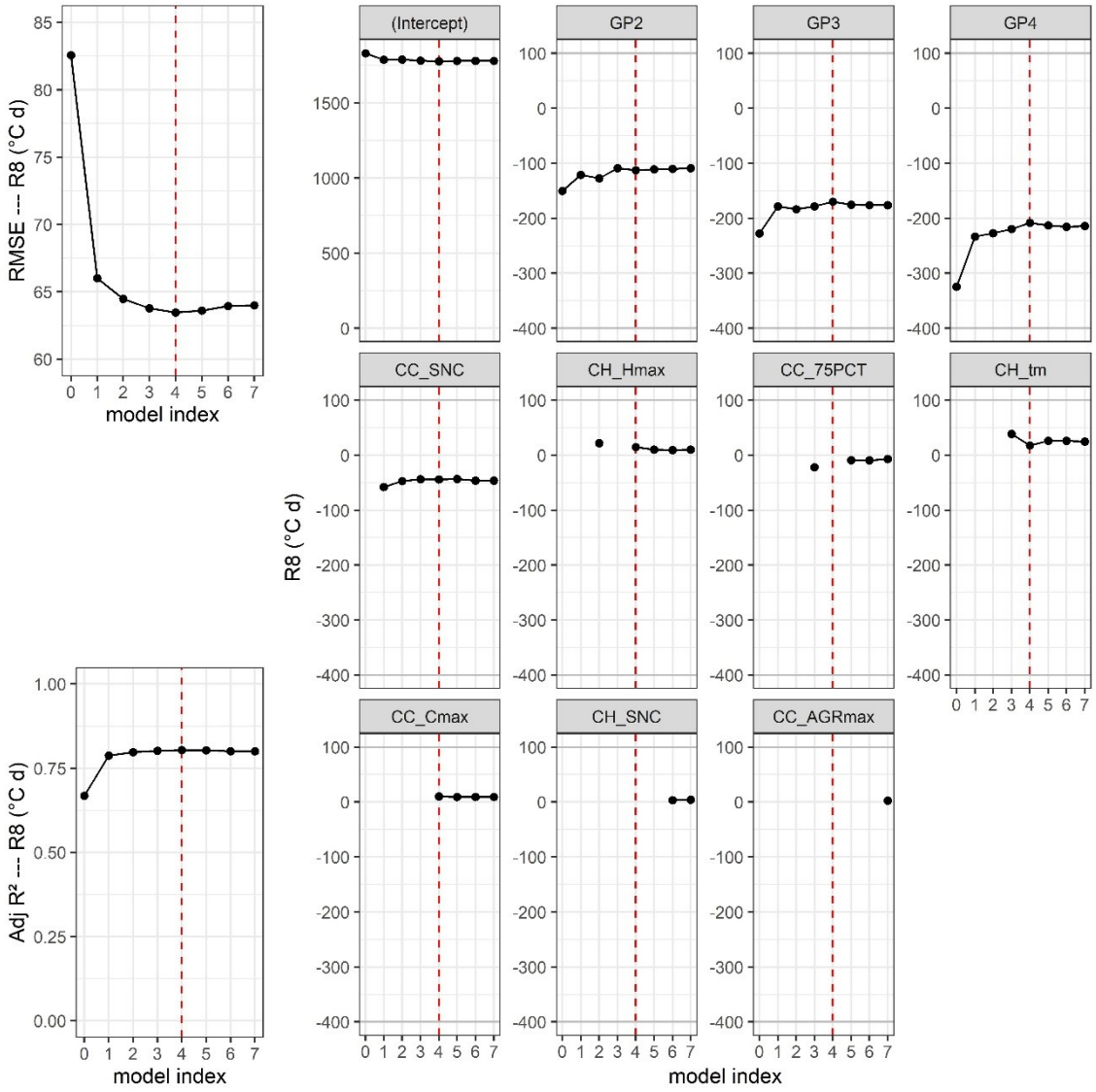


Figure A-7. MLR model building overview for R8 stage with RMSE and Adj.R² evolution. The red dashed line indicates the selected (best) model with the lowest RMSE. The subset selection regression with normalized UAV phenotypic variables included a 10 k-fold cross validation. In each subsequent step one variable was added and all possible combinations were evaluated. The model index = 0 included only the groups (GP1-4, GP1 is the reference group) as dummy variables. For the abbreviations see Table 4-1. Graphs are provided per variable with the value of their respective coefficients.

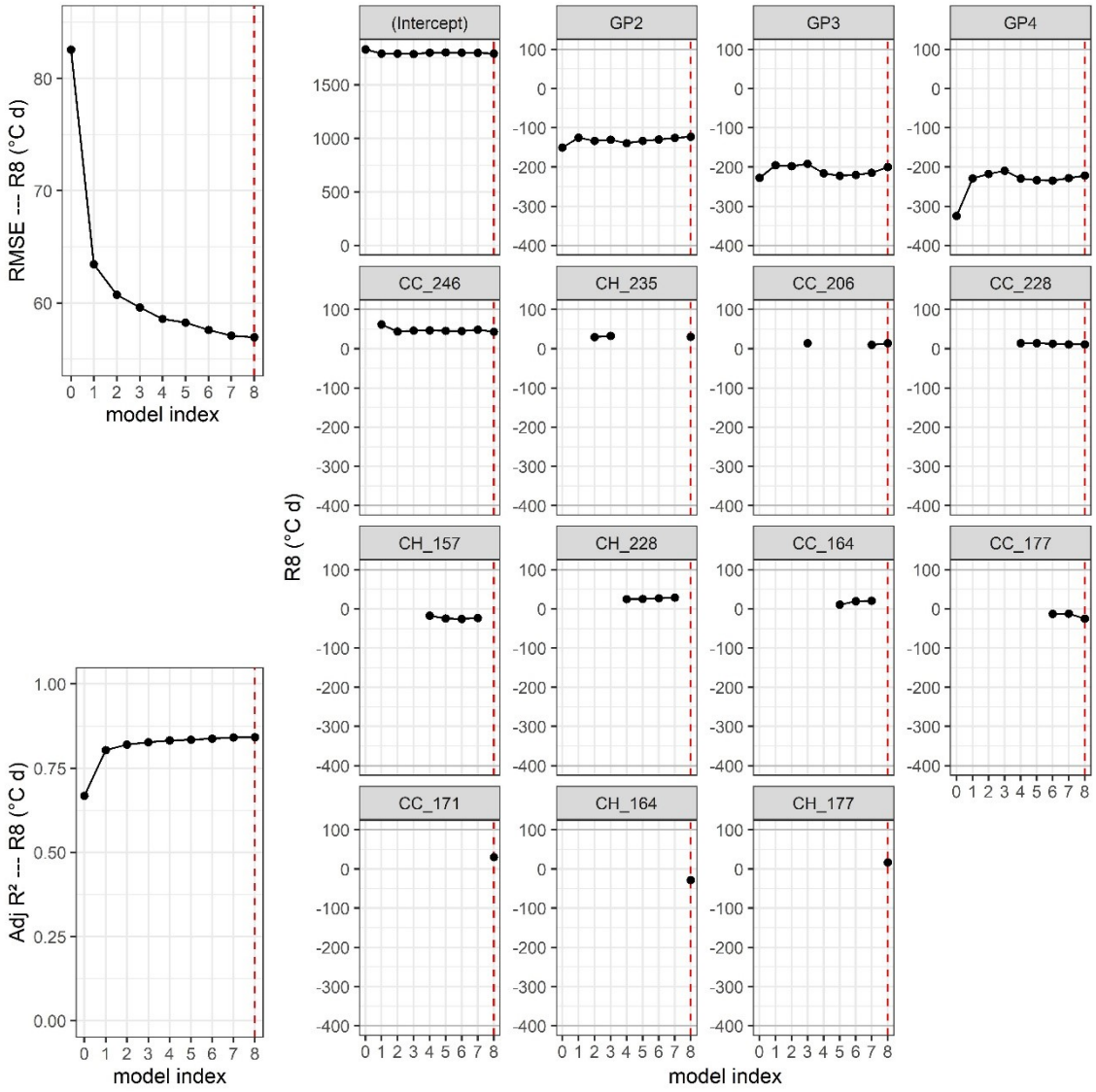


Figure A-8. MLR model building overview for R8 stage with RMSE and Adj.R² evolution. The red dashed line indicates the selected (best) model with the lowest RMSE. The subset selection regression with normalized UAV time points included a 10 k-fold cross validation. In each subsequent step one variable was added and all possible combinations were evaluated. The model index = 0 included only the groups (GP1-4, GP1 is the reference group) as dummy variables. Graphs are provided per variable with the value of their respective coefficients.

Supplementary explanations

- Selection of plots in CH3

The 39 plots corresponded to 13 cultivars x 3 replicates. The cultivars were selected according to heading data, with the intention to be used in an extrapolation of the height to yield in the respective heading date group in a later analysis carried out by one of the partners (not ILVO) inside the project in which this trial was established (FACCE-JPI ERA-NET+ Grasslandscape project, https://www.faccejpi.com/Research-Themes-and-Achievements/Climate-Change-Adaptation/ERA-NET-Plus-on-Climate-Smart-Agriculture/Grasslandscape).

- Selection of plots in CH4

The 148 plots used in the MLR models correspond to a balanced set with an equal number of plots per group to avoid a higher weight/importance/influence of one group compared to the others. Moreover plots with an emergence of <50% and a bad fitting (i.e. ajdR^2 < 0.7 for CH and for CC) were discarded to avoid noise or low quality data in the analysis

- Thresholding approach

A simple thresholding was used when mask were built to only extract data of the pixels corresponding to vegetation in the images (i.e. remove soil). This method corresponds to a manual selected threshold value based on the ExG histograms to differentiate both classes. For this specific index, the lower values and closer to 0 correspond to the soil and the higher values (max. value 1) are vegetation pixels. In the case of multi-temporal data, to establish the value between those two classes multiple orthomosaics (approx. five) were inspected if a time series.

- Effect of lodging on CH_SNC scores

CH_SNC was estimated based on the two last flights carried out. Senescence in soybean get visible as leaves turn yellow/brown and leaves drop as can be seen in Figure A-9. Leaf drop will affect the height values. An extra figure was added in the appendices to illustrate this.

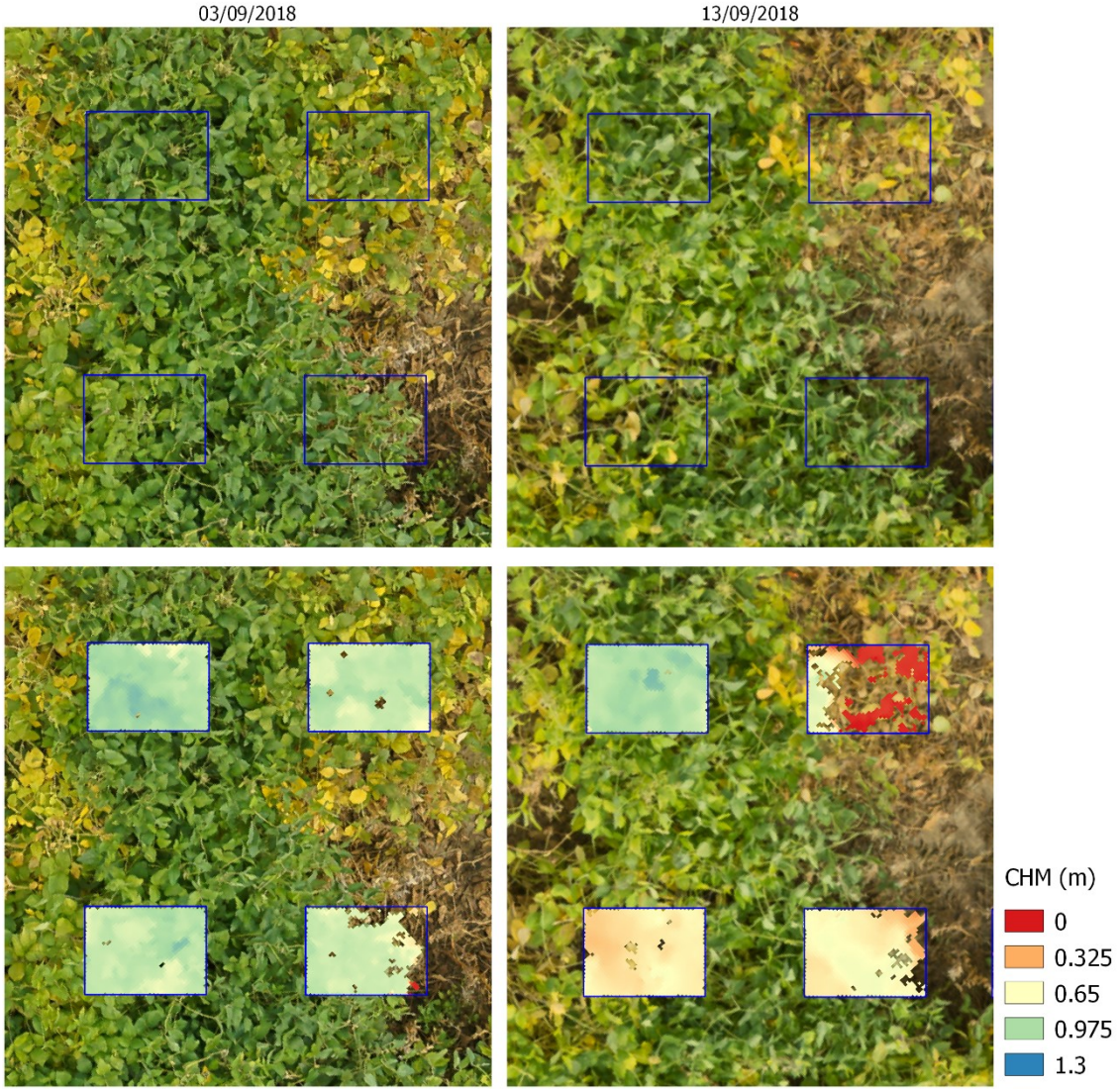


Figure A-9. Zoom of the same area in the last three flights carried out in which can be observed a decrease in the height values due to lead drop in the senescence stage.

- F varieties

The F-value for varieties indicates the variation between varieties relative to the random error in the trial. Higher F-values indicate better quality observations. A "+" or "-" sign indicates whether F-values are statistically significant or not (at the 5% level) (TrialWizard manual, https://www.beerepootconsultancy.com/wp-content/uploads/Manual_TrialWizard.pdf).

- Random Forest details

  Number of trees: 400

  Number of variables tried at each split: 19

  % var explained: 68.5

- Rising plate meter

A picture of a rising plate meter was added in Appendices (Figure A-10) with a short description of the operating principle.



Figure A-10. Rising plate meter used in a grassland field. How it is placed: first the pole is posed on the ground, then a plate of 30x30 cm is gently placed on top of the canopy and based on the distance between the plate and the device on top and the total length of the pole, the height of the canopy is calculated and registered.

Curriculum Vitae

Professional experience

01/09/19 – Present	Scientific attaché **Institute for Agricultural and Fisheries Research** (ILVO) – Melle (Belgium) Involved in different projects that use UAV imagery to derive phenotypic data
01/02/16 – 31/08/19	PhD Student in the Research Project iSense **Institute for Agricultural and Fisheries Research** (ILVO) – Melle (Belgium) UAV-RGB imagery as a phenotyping tool to derive breeding traits for genotypes selection
21/04/14 – 31/12/15	Forestry engineer in the Research Project RECUPERA2020 **Institute of Sustainable Agriculture** (IAS-CSIC) – Córdoba (Spain) Identification of species, field sampling with GPS and creation of thematic maps. Process and analysis of images captured with UAV.

Education

Master in Geomatics, Remote Sensing and Space Models Applied to Forestry Management
University of Córdoba (Spain)
Research MSc's thesis → Using images from an Unmanned Aerial Vehicle for trees height determination in forest area

Forestry engineer
University of Córdoba (Spain)
BSc's final project → Evaluation and comparison of digital classification techniques: combination of airborne sensors

Lists of publications

Peer-reviewed papers

Borra-Serrano, I., Swaef, T. De, Quataert, P., Aper, J., Saleem, A., Saeys, W., Somers, B., Roldán-Ruiz, I., Lootens, P., 2020. Closing the phenotyping gap: high resolution UAV time series for soybean growth analysis provides objective data from field trials. Remote Sens. 12, 1644. https://doi.org/10.3390/rs12101644

Van De Vijver, R., Mertens, K., Heungens, K., Somers, B., Nuyttens, D., Borra-Serrano, I., Lootens, P., Roldán-Ruiz, I., Vangeyte, J., Saeys, W., 2020. In-field detection of *Alternaria solani* in potato crops using hyperspectral imaging. Comput. Electron. Agric. 168, 105106. https://doi.org/10.1016/j.compag.2019.105106

Fernandez-Gallego, J.A., Lootens, P., Borra-Serrano, I., Derycke, V., Haesaert, G., Roldán-Ruiz, I., Araus, J.L., Kefauver, S.C., 2020. Automatic wheat ear counting using machine learning based on RGB UAV imagery. Plant J. tpj.14799. https://doi.org/10.1111/tpj.14799

Borra-Serrano, I., De Swaef, T., Muylle, H., Nuyttens, D., Vangeyte, J., Mertens, K., Saeys, W., Somers, B., Roldán-Ruiz, I., Lootens, P., 2019. Canopy height measurements and non-destructive biomass estimation of *Lolium perenne* swards using UAV imagery. Grass Forage Sci. 74, 356–369. https://doi.org/10.1111/gfs.12439

Borra-Serrano, I., De Swaef, T., Aper, J., Ghesquiere, A., Mertens, K., Nuyttens, D., Saeys, W., Somers, B., Vangeyte, J., Roldán-Ruiz, I., Lootens, P., 2018. Towards an objective evaluation of persistency of *Lolium perenne* swards using UAV imagery. Euphytica 214, 1–18. https://doi.org/10.1007/s10681-018-2208-1

de Castro, A., Jiménez-Brenes, F., Torres-Sánchez, J., Peña, J.M., Borra-Serrano, I., López-Granados, F., 2018. 3-D characterization of vineyards using a novel UAV imagery-based OBIA procedure for precision viticulture applications. Remote Sens. 10, 584. https://doi.org/10.3390/rs10040584

Torres-Sánchez, J., López-Granados, F., Borra-Serrano, I., Peña, J.M., 2018. Assessing UAV-collected image overlap influence on computation time and digital surface model accuracy in olive orchards. Precis. Agric. 19, 115–133. https://doi.org/10.1007/s11119-017-9502-0

Mesas-Carrascosa, F.J., Torres-Sánchez, J., Clavero-Rumbao, I., García-Ferrer, A., Peña, J.M., Borra-Serrano, I., López-Granados, F., 2015. Assessing optimal flight parameters for generating accurate multispectral orthomosaicks by UAV to support site-specific crop management. Remote Sens. 7, 12793–12814. https://doi.org/10.3390/rs71012793

Borra-Serrano, I., Peña, J.M., Torres-Sánchez, J., Mesas-Carrascosa, F.J., López-Granados, F., 2015. Spatial quality evaluation of resampled unmanned aerial vehicle-imagery for weed mapping. Sensors 15, 19688–19708. https://doi.org/10.3390/s150819688

Conference contributions (during PhD)

High throughput phenotyping of winter wheat growth using an UAV. Borra-Serrano, I.; De Swaef, T.; Derycke, V.; Haesaert, G.; Inzé, D.; Roldán-Ruiz, I.; Lootens, P., 2019. II Spanish Symposium on Physiology and Breeding of Cereals, Córdoba, Spain. 6-7/03/2019.

UAV based field phenotyping of winter wheat: can a proxy be found to measure chlorophyll content? Lootens, P.; De Swaef, T.; Borra-Serrano, I.; Derycke, V.; Haesaert, G.; Inzé, D.; Roldán-Ruiz, I., 2019. II Spanish Symposium on Physiology and Breeding of Cereals, Córdoba, Spain. 6-7/03/2019.

Yield estimation of perennial ryegrass plots in breeding trials using drone images. Aper, J.; Borra-Serrano, I; Ghesquiere, A.; De Swaef, T.; Lootens, P.; Baert, J., 2019. 20th Symposium of the European Grassland Federation and the 33rd Meeting of the EUCARPIA Section 'Fodder Crops and Amenity Grasses', Zürich, Switzerland. 24-27/06/2019.

Grassland monitoring and phenotyping based on UAV-imagery. Borra-Serrano, I.; De Swaef, T.; Aper, J.; Muylle, H.; Mertens, K.; Nuyttens, D.; Vangeyte, J.; Saeys, W.; Somers, B.; Roldán-Ruiz, I.; Lootens, P., 2018. 17th meeting of the EUCARPIA Section Biometrics in Plant Breeding, Gent, Belgium. 03/09-05/09/2018.

UAV based field phenotyping of winter wheat: what do vegetation indices derived from RGB and multispectral data offer to the breeder? Lootens, P.; De Swaef, T.; Borra-Serrano, I.; Derycke, V.; Louwers, M.; Muylle, H.; Haesaert, G.; Inzé, D; Roldán-Ruiz, I., 2018. 17th meeting of the EUCARPIA Section Biometrics in Plant Breeding, Gent, Belgium. 03/09-05/09/2018.

Field based phenotyping of forage crops using close remote sensing tools. Lootens, P.; Borra-Serrano, I.; De Swaef, T.; Aper, J.; Roldán-Ruiz, I., 2018. Plant Biology Europe 2018, Copenhagen, Denmark. 18/06-21/06/2018.

Phenotyping grass canopy height evolution based on SfM and DEM estimation using a drone. Lootens, P.; Borra-Serrano, I.; De Swaef, T.; Muylle, H.; Mertens, K.; Nuyttens, D.; Vangeyte, J.; Saeys, W.; Somers, B.; Roldán-Ruiz, I., 2018. COSTFA1306 Meeting 2018 Leuven, Belgium. 20/03/2018.

Non-destructive monitoring of grassland canopy height using a UAV. Borra-Serrano, I.; De Swaef, T.; Muylle, H. ; Mertens, K.; Nuyttens, D.; Vangeyte, J.; Willner, E.; Hegarty, M. J.; Barre, P.; Sampoux, J.P.; Saeys, W.; Somers, B.; Roldán-Ruiz, I.; Lootens, P., 2017. 11th European Conference on Precision Agriculture, Edinburgh, United Kingdom. 17/07-20/07/2017.

Non-destructive analysis of biomass accumulation in grasslands using UAV. Lootens, P.; Borra-Serrano, I.; De Swaef, T.; Muylle, H.; Mertens, K.; Nuyttens, D.; Vangeyte, J.; Willner, E.; Hegarty, M.J.; Philippe, B.; Sampoux, J.P.; Roldán-Ruiz, I., 2017. 3rd meeting COST FA1306, Oeiras, Portugal. 27/03-28/03/2017.

Persistency evaluation in a breeding context using UAV imagery. Borra-Serrano, I.; De Swaef, T.; Aper, J.; Ghesquiere, A.; Mertens, K.; Nuyttens, D.; Saeys, W.; Ben, S.; Vangeyte, J.; Roldán-Ruiz, I.; Lootens, P., 2017. 3rd meeting COST FA1306, Oeiras, Portugal. 27/03-28/03/2017.

Contribution to education

Bachelor or Master thesis co-supervision.

Bachelor Vastgoed en Landmeten. AP Hogeschool Antwerpen. Academic year: 2016-2017. Fotogrammetrische verwerking van UAV foto's tot ruimtelijke 3D informatie.

MSc in de biowetenschappen: land- en tuinbouwkunde. UGent. Academic year: 2019-2020. UAV gebaseerde fenotypering ter ondersteuning van veredeling van tuinrozen

Bachelor in de Agro- en biotechnologie. Odisee Hogeschool. Academic year: 2019-2020. Wintertarwe opvolgen met dronetechnologie.

Practical session in Master education

POC Master Bio-ingenieurswetenschappen: Remote Sensing of Vegetative Systems (KULeuven). Academic year: 2018-2019 (26/04/19). Can UAV be used as a phenotyping tool?

Seminars

ISENSE – Precision Crop Farming. UCL 31/05/2017 (Drone and Research at universities)

Grassland monitoring for precision management using UAV remote sensing. ILVO 02/06/2017

Grassland monitoring. Rothamsted Research 03/10/2017

Grassland sensing using airborne sensors. Noordzee Drones 14/06/2018

Phenotyping breeding trial with an UAV. GFPI Sommertagung der Abteilung Futterpflanzen 11/04/2019

Can UAV be used as a phenotyping tool?. ILVO 04/10/2019