Multi-temporal and and multispectral modeling of wheat crop parameters : Estimating within-field variability using UAV imagery

Auteur : Taconet, Julien
Promoteur(s) : Brostaux, Yves
Faculté : Gembloux Agro-Bio Tech (GxABT)
Diplôme : Master en bioingénieur : sciences et technologies de l'environnement, à finalité spécialisée
Année académique : 2019-2020
URI/URL : http://hdl.handle.net/2268.2/10543

LIÈGE université
Gembloux
Agro-Bio Tech

MULTI-TEMPORAL AND MULTISPECTRAL MODELING OF WHEAT CROP PARAMETERS: ESTIMATING WITHIN-FIELD VARIABILITY USING UAV IMAGERY

JULIEN TACONET

TRAVAIL DE FIN D'ETUDES PRESENTE EN VUE DE L'OBTENTION DU DIPLÔME DE MASTER BIOINGENIEUR EN SCIENCES ET TECHNOLOGIES DE L'ENVIRONNEMENT

ANNEE ACADEMIQUE 2019-2020

(CO)-PROMOTEUR(S): YVES BROSTAUX, BENJAMIN DUMONT

Remerciements

Je tiens à remercier mon promoteur de Travail de Fin d'études, Yves Brostaux, de m'avoir inclus dans ce projet et de m'avoir suivi et soutenu tout au long de ce travail. Ce soutien s'est manifesté à plusieurs reprises le long de mon parcours universitaire, dans des moments de questionnements quant à mon orientation. Je me retrouve aujourd'hui pleinement satisfait de mes choix d'études, vous avez entretenu mon goût de la gestion de données à travers votre pédagogie et vos conseils avisés. Merci !

A mon co-promoteur Benjamin Dumont, qui malgré la crise sanitaire a su suivre le travail de fin d'études de plusieurs étudiants. Un grand merci pour les nombreux conseils, et pour le temps consacré à mon suivi et à la réalisation de ce travail.

Je voudrais exprimer ma reconnaissance à Samuel Quevilliers, Cédric Geerts, Adrien Michez et Jérôme Pierreux pour m'avoir régulièrement renseigné et répondu à mes nombreux questionnements. Merci aussi de m'avoir formé à l'utilisation des équipements nécessaires à mon travail.

A ma compagne Romane Quintin, qui a été à mes côtés tout au long de ce travail. Merci d'avoir su apporter un regard extérieur à mon mémoire et pour le soutien lors des moments les plus difficiles.

A mes parents, merci d'avoir fait de moi la personne que je suis alors que mon parcours universitaire se termine. Vous m'avez soutenu économiquement et moralement tout au long de mes études, avec de précieux conseils lors de mes questionnements. Je pense sincèrement avoir dirigé mon parcours dans la bonne direction, et c'est en grande partie grâce à vous.

A mes frères et sœurs qui pendant ce travail, étaient dispersés en France, en Italie, en Espagne et au Burkina-Faso, merci d'avoir été présents à travers ces appels et messages pendant le confinement, qui nous ont permis malgré la distance, de garder ce contact et entretenir ces liens si précieux. Merci pour tous vos conseils et votre soutien, qui m'ont inspiré et permis d'aller au bout de mon parcours étudiant.

Résumé

Étant donné l'amplitude des défis dont l'agriculture doit faire face au 21ème siècle, il est crucial de développer des stratégies agronomiques compatibles avec le climat afin de nourrir une population mondiale croissante, tout en fournissant une alimentation diversifiée dans le respect de la santé humaine. En réponse à l'impact de l'agriculture sur le changement climatique à travers l'élevage intensif et son implication dans le processus de la dégradation des terres, dû à des pratiques agricoles non-durables et à l'utilisation excessive d'intrants chimiques, une expérimentation à long-terme s'initie en 2020. Co-construite par Gembloux Agro-Bio Tech et des agriculteurs, celle-ci vise à étudier deux systèmes agricoles innovants, visant à éliminer l'utilisation de moyens de contrôle chimiques des ravageurs, produire une grande diversité de cultures et questionner le rôle de l'animal dans l'agriculture. Un suivi rapproché de plusieurs paramètres est prévu, partiellement effectué à travers l'utilisation d'imagerie drones. En effet, certains paramètres culturaux (particulièrement le Leaf Area Index, ou Index de Surface Foliaire) sont considérés comme étant de bons indicateurs de l'état physiologique de la plante, qui est conditionné par l'état du sol. Ce travail vise à fournir des lignes directrices pour le suivi cultural de cette parcelle expérimentale à travers de l'imagerie drone. Un modèle d'Index de Surface Foliaire (LAI) et de matière sèche d'épi ont été créés avec succès pour le froment d'hiver. Ces modèles ont été créés en utilisant des images RGB et multispectrales, ainsi que des données de vérité terrain collectées à l'occasion de quatre vols drones, du gonflement d'épi à la maturité. Cet outil a permis de cartographier ces paramètres sur la totalité de la parcelle, d'étudier leur évolution à travers la saison et seront possiblement utilisés pour suivre l'évolution des conditions de la parcelle pendant l'expérience à venir. De plus, l'implémentation de cette expérience nécessite d'examiner l'homogénéité des conditions du sol dans les blocs expérimentaux. Cela a été effectué dans ce travail à travers la carte de LAI, et un repositionnement réfléchi des blocs a été proposé afin de réduire l'hétérogénéité des conditions du sol, au sein des blocs et entre ceux-ci.

Abstract

Given the amplitude of challenges agriculture has to face in the 21st century, it is crucial to develop climate-compatible agronomical strategies in order to feed a growing population, while respecting human health by providing the necessary diversity of nutritional products. In response to the impact of livestock on climate change and to the crying alarm of land degradation progress, that occurs through the excessive use of chemical inputs and non-durable agricultural practices, an experiment aiming to study two innovative cropping systems is to be set in Gembloux Agro Bio-Tech. These systems are aligned with the principles of removing chemical pest control, providing a higher crop diversity and will question the role of livestock in agriculture. Close-monitoring of various parameters is to be set up during the experiment, and will partly be carried out through Unmanned Aerial Vehicle (UAV or drone) imagery. In fact, some crop parameters (especially Leaf Area Index) are considered to be an accurate indicator of plant status, which is highly driven by soil conditions. This work aims to set up guidelines for crop monitoring through drone imagery in the specific context of this parcel. A Leaf Area Index (LAI) model and an ear dry matter model have been successfully created for winter wheat. These models were built using both RGB and multispectral UAV imagery, and ground truth collected on four drone flights, from booting to ripening. This tool allowed to map these parameters on the whole parcel, study their evolution through the season and will possibly be used to monitor the evolution of soil conditions in the future experiment. Furthermore, the implementation of the experiment requires to survey the homogeneity of the conditions within the experimental blocs. This was done through the LAI map, and careful repositioning of the experimental blocs was carried out in order to reduce heterogeneity within and between these blocs.

Table of Contents

Table of Figures

Table of Tables

Table of Maps

1. Introduction

a. Global context

According to the FAO (2017), one of the main challenges for agriculture in the 21st century is to achieve sustainable food production to meet increasing demand, while ensuring a sustainable resource base. As climate change's growing threat to food supply increases, it is crucial to "develop climate-compatible growth strategies that do not conflict with mitigation goals required to minimize further warming" (FAO, 2010).

At the dawn of the 21st century, the world had witnessed an extraordinary crop production growth through an intensification of favorable areas and a slight increase in cropland surfaces (Pingali, 2012; Foley, 2005). In response to the alarming picture that 1 billion people were hungry every day in the 1950s, industrialized and developing countries together with development organizations decided to take action through a series of interventions in agriculture (Spielman & Pandya-lorch, 2010). The world grain harvests in 2005 had doubled in the past 40 years due to the "Green Revolution" (GR), which was characterized by an increase of 700% in fertilizer use, 70% of irrigated cropland areas, the use of high-yield cultivars and mechanization (Matson,1997; Gleick,2003). The increase in production that characterized the GR was followed by a drop of the food prices and an increased consumption of high-calorie and protein food of the poor (Spielman & Pandya-lorch, 2010). This resulted in a reduction of global hunger rates, with the share of undernourished people dropping between 1970 and 1990 (FAO, 2016). The tendency remained unchanged after 1990 thanks to international global agendas such as the Millenium Development Goals, with this share decreasing to 11 percent in 2016 from 19 percent in 1990 (Willet & al., 2019; IFRPI, 2017). Yet the GR came at the cost of severe environmental degradation. In fact, appropriate policies and research aimed towards a judicious use of inputs was largely lacking (Pingali, 2012). The conversion of natural ecosystems to croplands, increased water use, soil degradation and chemical runoffs have had unintended consequences beyond the areas cultivated. To cite a few, some irrigated land has become heavily salinized, causing a loss of 1.5 million hectares of arable land per year (Wood & Sherr, 2000). The loss of native habitats, especially for pollinators, have had negative impacts on agricultural productivity (Kremen *et al*., 2002). Overuse of nitrogen and phosphorus-rich fertilizers created dead zones in lakes and coastal ecosystems (Diaz & Rosenberg, 2008). These are examples of consequences that partly explain the slowdown in yield growth from the mid-80's (Pingali, 2012).

The tight correlation between agriculture and climate is verified (Paustian *et al.*, 1997; Aydinalp & Cresser, 2008). Agriculture is the major cause of global environmental changes, being one of the leading GHG (Green House Gases) emitters to an extent of 19-29% (Vermeulen, 2012) and responsible for 20% of the annual increase of total GHG emissions (IPCC, 1996). Agriculture is also highly sensitive to climate fluctuations such as precipitation irregularity, temperature changes and extreme weather events (Singh R.S.G.S., 2017), which are a direct consequence of anthropogenic GHG emissions.

Land use activities are defined by Foley *et al.* (2005) as "whether converting natural landscapes for human use or changing management practices on human-dominated lands". Throughout most of history, the increase in food production was achieved by expanding the total area dedicated to agriculture, through conversion of forests and natural grasslands (Wood & Sherr, 2000). According to the FAO (2011), roughly 40 percent of non-ice-covered land surface was dedicated to croplands and pastures in the 2000s, and had been increasing by 12 percent in the past four decades. The

FAO (2011) states this extension was achieved mainly at the expense of forest, grasslands and wetlands habitats. All of the net increase in agricultural surface between the 1960s and 2010 is attributable to a net increase of irrigated croplands, while rainfed cropland total area slightly declined (FAO, 2011). Yet land well-suited for crop production is limited and in competition with other land-uses, for example biofuel production (World Energy Council, 2011). Furthermore, the patterns of human concentration are increasing, which has greatly limited the opportunity for further geographic expansion of land for agriculture (Wood *et al.*, 2000; FAO, 2011). In addition, the second part of the twentieth century witnessed a historical degree of land degradation through intensification of agriculture. Land degradation is defined in the IPCC's Special Report on Climate Change and Land (2019) as: "a negative trend in land condition, caused by direct or indirect human-induced processes including anthropogenic climate change, expressed as long-term reduction or loss of at least one of the following: biological productivity, ecological integrity or value to humans". The IPCC points out unsustainable agriculture and forestry, in combination with socio-economic pressures such as urbanization and population growth as majorly responsible for land degradation. Global estimates of total degraded area vary between 10 million km^2 and 60 million km^2 (Gibbs and Salmon, 2016). The FAO TerraStat interpretation of the Global Assessment of Soil Degradation (GLASOD) commissioned by the United Nations Environment Program, estimated that 66 percent of the world's arable land is affected by a form of land degradation, with 26 percent severely or very severely degraded (FAO, 2000). A direct consequence of this degradation is the loss of ecosystem services: an estimated 9.2% decrease in the global annual value of ecosystem services is due to land degradation (Gibbs and Salmon, 2016). These ecosystem services include soil fertility, agricultural products, climate regulation and recreational activities. The economic value of the losses due to the degradation of ecosystem services have been estimated to 10% of the world's GDP (Sutton *et al.,* 2016). Another direct consequence of land degradation is yield reduction: worldwide, 75 billion tons of soil are lost every year, corresponding to a loss of approximately 400 billion dollars per year (FAO, 2006). The geographical limits of further expansion of agriculture and the loss of arable land due to land degradation makes feeding the world in the 21st century a huge challenge.

Not only has the GR come at the cost of environmental degradation, it has also contributed to a substantial change in dietary patterns. Although global calorie intake increased, the effects on diets were unevenly distributed around the globe. In some regions, the GR allowed for a broader diversification of diets through an increased accessibility to micro-nutrient dense foods, thanks to the drop in food prices (Pingali, 2012). On the other hand, in other regions, accessibility to micronutrient-rich crops has diminished, due to their replacement by higher value staple crops or to a rise in their price relative to staples (Pingali, 2012). This accessibility issue towards micronutrient rich food is one of the driving processes of a global shift to unhealthy diets, along with rapid urbanization and income increases (Popkin, 2006; Global Panel on Agriculture and Food Systems for Nutrition, 2016; Pingali, 2012). The shift towards high-calory, heavily processed and animal sourced food led to an increase in obesity and diet-related diseases, co-existing with wide-scale undernutrition. While 820 million people remain undernourished, 2.1 billion adults face an overweight or obesity issue, and the prevalence of diabetes has doubled in the past 30 years (Willet *et al*., 2019).

Human health and environmental sustainability are inextricably linked by diets (Willet *et al*., 2019). As previously stated, the food industry has a major role in anthropogenic GHG emissions. Most importantly, up to 80% of these emissions are directly associated with livestock production (FAO, 2006), and this production is mainly driven by demand for livestock products (Delgado, 1999). Diets consisting of high consumption of red and processed meats and low consumption of fruits and vegetables are major diet-related risk factors, which partly explain substantial early mortality rates in most regions (Lim S.S. *et al*., 2012). Willet *et al*. (2019) define "Lose-lose" diets as unhealthy and

environmentally unsustainable, commonly characterized as being “high in calories, added sugars, saturated fats, processed foods, and red meats.” As the world population increases, with an estimation of 9.8 billion people by 2050 (United Nations, 2018), demand for agricultural products is projected to increase by 50 percent in 2050 compared to 2012 (FAO, 2017). Furthermore, the rate of increase of the middle-class is approaching its peak: around 140 million people annually add up to the 3.2 billon people in the middle class by the end of 2016, and their numbers are expected to rise up to 5 billion by 2050 (Homi, 2017). As a growing and wealthier population adopt unhealthy and environmentally unsustainable diets, the situation is expected to get worse without targeted dietary changes (Popkin, 2006). The negative effects on human health will be exacerbated indirectly through negative impacts on the environment. These include respiratory problems due to poor air quality resulting from the burning of biomass and land clearing for agriculture, food insecurity due to reduced yields resulting from changing climatic conditions and extreme weather events, and diminished nutrient content in crops resulting from higher atmospheric CO2 concentrations (Willet *et al*., 2019). Switching to a reduced animal-sourced food diet now appears as a necessity to mitigate climate change, positively impacting human health and allowing future access to safe and affordable food (Tilman D. & Clark M., 2014). Such a dietary change would allow for environmental benefits such as reduced pressure on land (Stehfest E, 2009) and reduction in GHG emissions (Hallström, 2014). Changing diets may even be more effective than a technological option to mitigate climate change (Popp, 2010). Yet, poor nutrition still has no dedicated Millenium Development Goal and no Sustainable Development Goal (except for SDG 2 zero hunger, which addresses only part of the problem). Actions aiming to improve nutrition have failed to gain international commitment. In response to the lack of global traction to this issue, *The Lancet* has decided to address nutrition issues through two commissions published in January 2019, followed by a series of papers. *Food in the Anthropocene: the EAT-Lancet Commission on Healthy Diets for Sustainable Food Systems* (2019) approaches nutrition through several perspectives, highlighting the link between nutritional targets and environmental stability.

b. “AgricultureIsLife” platform: experimenting innovative food systems

The present work is one local research response to the above introduced global challenges and context. It takes place under “AgricultureIsLife”, the TERRA Teaching and Research Centre’s experimental platform in Gembloux, which has been host to long-term experiments aiming to anticipate changes in the agricultural systems in order to provide practical and sustainable solutions. A new co-constructed experiment involving scientists from the TERRA Centre and farmers is to be initiated in 2020. Its goal is to evaluate the long-term stability of innovative rotations in crop systems in a system approach, with the aim of getting rid of synthetic pesticides and herbicides. Two innovative 8-year rotation cropping systems are to be set up for the experiment on a 30 hectares parcel. Both are planned to answer locally to the optimal diet defined in the Eat-Lancet Commission on healthy diets for sustainable food systems, adapted to the loamy region of Hesbaye (Belgium).

In this commission, Willet *et al*. (2019) defined a “safe operating space” for food systems that sets quantitative scientific targets to achieve food production for a healthy diet while remaining within the planetary boundaries. The “optimal diet” is defined in this commission through scientific targets in terms of ranges of intakes for food groups (e.g. 100-300g of fruits per day), and its healthiness is defined on the basis of nutrient intake and mortality rates. The planetary boundaries framework is a set of environmental boundaries for food production to ensure a stable earth system, and they are defined on the basis of six processes that regulate the state of the earth system, including “the total global amount of cropland use, biodiversity loss, water use, greenhouse-gas emissions, and nitrogen and phosphorus pollution that can be due to food production” (Willet *et al*., 2019). This commission

concludes that it is possible to switch to an environmentally sustainable food system to ensure a healthy diet to 10 billion people by 2050, thus providing a framework for a “Great Food Transformation”. In fact, they warn this goal can only be achieved through “widespread, multisector, multilevel action that includes a substantial global shift towards healthy dietary patterns, large reductions in food loss and waste, and major improvements in food production practices.” (Willet *et al*., 20019). The Eat-Lancet commission lists five main strategies to achieve the Great Food Transformation, targeting national and international commitment to this goal, the food production system, land governance and waste reduction. Noticeably, adaptation of the food production system to local conditions is promoted. In fact, matching production practices to local constraints can sustainably increase food production (Seufert *et al*., 2012). The definition of the healthy dietary pattern, which consists of ranges of intakes for each food group allows flexibility to apply it locally through its tailoring to preferences and cultures. Additionally, while the production and consumption of meat and dairy products has to decrease to achieve this goal, adaptation to local environmental contexts is required to determine at what extent it should decrease (Willet *et al*., 2019). In fact, animal production can be beneficial in some contexts, providing an essential income base for poverty alleviation, being beneficial for grassland ecosystem services and providing for particular nutritional needs (especially children and vulnerable populations) (Smith, 2012).

AgricultureIsLife’s new experiment bans the use of pesticides and herbicides, but operations on the soil will be allowed. The first cropping system integrates the animal in its management through its presence in the rotation for grazing in meadows, as a tool for weeds and pest-control, allowing animal-sourced input fertilizers and exportation of by-products. The second does not take livestock into account in any way, with no organic fertilization (utilizing only mineral fertilizers) and production fully destined for human use. A third rotation is planned to account for current practices, with two variants, one allowing the use of herbicides, the other one banning all pesticide usage. Additionally, according to the third strategy described in the Eat-Lancet commission for a Great Food Transformation, “sustainably intensify food production, generating high-quality output », a number of agroecological levers are planned. Long term 8-year rotations including resistant/tolerant crop varieties and temporary meadows will be implemented, combined with the use of inter-cropping. An ecological network promoting biodiversity composed of hedges, messicoles, agroforestry and flowery strips are also to be set up. The parameters to be measured across the experiment are plant growth, total aerial and root biomass, monitoring of beetles and earthworms, and soil parameters such as diversity and dynamics of microbiota, pH, Total Organic Carbon or bioavailability of elements. Final yield and its sanitary and nutritional value are also to be evaluated, along with economical, energetical and workforce-related aspects. Drone technology is to be used as a monitoring tool through the entirety of the experiment, with regular flights planned for image acquisition which will be used as a basis for modeling applications.

This new experiment is to be ready by September 2020, date of sowing of the first crop. Since November 2019, a homogenization crop consisting of winter wheat (*Triticum aestivum*) has been sowed and will be harvested in august 2020. Since the area covered by the experimental platform is composed of a grouping of different fields, this pre-treatment aims to homogenize in-field conditions. A universal soil property is that they are variable across space and time. Heterogeneity of soils is often distinguished by its origin: environmentally induced vs. plant-induced. The first one can be induced by geological processes such as soil formation and weathering (Ricklefs, 1977), while the latter is mainly due to biotic processes of plants such as root activity, plant colonization and decomposition (Gibson, 1988; Jackson and Caldwell, 1993). Plant-soil feedback mechanisms are significant sources of heterogeneity in soils, because plants often have species-specific effects on their environments (Brandt *et al*., 2014). A third heterogeneity source is agricultural practices, if those

differ within the area of interest. Soil heterogeneity is a major source of variation in crop yield (Li H. *et al*., 2001). An experiment conducted to determine the effect of different treatments must be carefully planned so that the results reflects the effect of treatments, not of an uncontrolled source of variation. A significant method to control error is blocking: this technique consists in assigning experimental units that are as similar as possible together in the same group (called block), then applying all treatments separately and independently in each block (Gomez & Wiley, 1983). Treatment replications allows for a better estimation of experimental error and increase the range of conditions under which the treatment's effects are evaluated, while randomization of treatments application on plots inside blocks allow to remove bias from error estimation (Steel and Torrie, 1980). Yet, reduction of experimental error relies on separating the heterogeneous field into blocks that are nearly homogeneous, which requires careful dimensioning and positioning (Mulla *et al*., 1990). In AgricultureIsLife's future experiment, 4 cropping systems are to be compared. Considering the limited amount of space on the parcel where it has been implemented, 4 blocks have been set up, each one consisting of 8 plots. Each block is divided in two temporalities, so each cropping system is represented twice in each block, with a 4-year shift. This will allow to have the entirety of the 8-year rotation's diversity of crops covered in 4 years and will provide results earlier, even though the experiment will be ongoing on the long-term.

c. Remote Sensing for Precision Agriculture

According to the third main strategy defined in the Eat-Lancet commission to achieve the Great Food Transformation, "sustainably intensify food production, generating high-quality output », Precision Agriculture (PA) techniques are recommended to successfully intensify food production per area unit. Since the mid-1980s, PA has been growing in interest, using sensors and information systems to account for "within field variability", allowing informed management thus increased productivity (Robert, 2002). In-field spatial variability in crop yield can be due to soil condition (which is itself determined by the parent material, regional variations in climate and previous land management), pests or diseases and weeds. In addition, there is a temporal component in this variability resulting from weather patterns and management practices (Gebbers, 2010). Temporal variation tends to be insufficiently recognized in PA, especially within-season variations (Mcbratney, 2005). A timed input management guided by feedback from crop monitoring allowed Van Alphen and Stoorvogel (2000) to save 17-25% of fertilizer inputs compared to the up-to-date fertilization advices given by extension services. The main drivers of PA lie in the rise of GPS, sensor, GNSS, GIS and microcomputer technologies, combined with increasing awareness of soil in-field variability (Robert, 2002). This has made it possible for farmers to access to this variability at an unprecedented level of precision. Amongst data acquisition technologies, Remote Sensing's (RS) potential in agriculture has been recognized as immense through multi-scale monitoring, which is a valuable asset for future challenges (European Space Agency., 2019). In fact, while the international nature of agricultural markets and the global impacts of temperature rising require large-scale management, the effect of environmental impacts and food insecurity are felt locally (FAO, 2010). A coordination between large-scale policies and local adaptations is required to provide appropriate answers to global warming.

Common applications of RS include water and nutrient management, pest and disease detection and yield prediction (Pinter, 2003). RS can be categorized into satellite, aerial and near-ground RS. Though satellite sensors can acquire images at a very large scale, they are tailored by their spatial resolution (typically 500m to 0.5m). While imagery at 10m-20m spatial resolution is available in open-data (ex. Copernicus's SENTINEL program), the use of high-resolution satellite imagery is restrained by its cost. Furthermore, satellites have a fixed orbit period (16 days for Landsat and 26 days for SPOT) which can be a constraint when studying plant phenology because of the fixed timing

acquisition, and passive sensors cannot penetrate clouds, which makes data acquired under cloudy conditions stale. Aerial RS is also hampered by economic considerations (Holman *et al*., 2016). On the other hand, Unmanned Aerial Vehicle's (UAV) interest in remote sensing applications for PA have been increasing recently (FAO and ITU, 2018). Drones are remotely-controlled small aircraft, with no pilot on-board. They offer a flexible, reliable and low-cost solution for data acquisition at ultra-high spatial resolution (from millimeter scale to meter). They can be customized with a variety of sensors and can be flown at high frequencies to match study-specific requirements, which allows for endless possibilities of applications. Such applications range from soil erosion monitoring (Marzolff *et al*., 2012) to hazards and disaster risk monitoring (Gomez C. & Purdie H., 2016). Agriculture applications include wheat yellow rust monitoring from multispectral imagery (Su *et al*., 2018), assessing water stress (Gago, 2015), or plant monitoring and phenotyping (Sagan *et al*., 2019).

The next section is a literature review in the domains of interest in this work. First, UAV technology will be described in its globality, from its advantages and drawbacks to the methods used to obtain quality output data. Wheat crop parameters that will be monitored in this work will then be described, along with relevant information that must be collected to carry out this work. Vegetation Indexes will be defined, their use justified in the context of remote sensing, and finally, the role and diversity of machine learning algorithms for crop parameter modeling will be described.

2. Literature review

a. UAV

i. Structure-from-Motion photogrammetry

An important asset of UAV-acquired imagery is its ability to be used in a Structure-from-Motion (SfM) photogrammetry workflow. This emerging method offers high-resolution 3D structural or topographic reconstruction of the scene, calculating the position, orientation and scene geometry through the set of overlapping images (Westoby *et al*., 2012). This allows to generate useful 3D models (e.g. Crop Surface Models which can be combined with a Digital Elevation Model (DEM) to estimate crop height). SfM photogrammetry relies on an algorithm that detects ties points (similar pixels) between pictures. This requires the pictures to be overlapping, between the pictures (front overlap) and between the flight lines (side overlap). Most studies recommend a front overlap of 80% and a 60-75% side overlap, for these values usually lead to the production of high-quality 3D models and orthoimages (Tmuši *et al*., 2020).

ii. UAV technical limitations

On the other hand, UAV imagery presents its set of challenges. The relatively low flying height implies a small footprint, which means longer flights to cover the whole area, and they are often limited by battery capacity (typically 20 minutes per flight). The UAV's flight speed is also tailored by the camera's sensor shutter speed, measured in fractions of a second, which controls how long it is exposed to light. A camera mounted on a UAV flying at an excessive speed compared to the sensor's minimal shutter speed will result in motion blur, reducing the image's quality (Sieberth *et al.,* 2014) There is often a compromise to be made between flight altitude, UAV cruise velocity, spatial resolution and) area covered. Additionally, UAV imagery often suffers from occlusion or changes in illumination conditions during the flight, while variations in topography may result in variations in spatial resolution (depending on the ability of the flight-programming software to allow for terrain altitude following, thus keeping a constant spatial resolution across the whole flight).

iii. UAV imagery pre-processing

A series of processing steps has to be carried out once the UAV images are acquired. Schott (2007) calls the entire workflow the "image chain approach". Usually, the initial data is Digital Numbers (DNs) and the final data is reflectance. Sensor-related calibration, scene reflectance generation and correction need to be carried out in case reflectance is the desired output. Sensor-related calibration starts with relative radiometric calibration, which aims to transform the DNs into normalized DNs, which have a uniform response across space and time during the time of operation. These transformations include dark-signal correction (influence of sensor temperature on readings) and vignetting (influence of the optical path of the camera on the incoming radiant flux) (Aasen *et al*., 2018). The instrument recordings then need to be transformed into reflectance, which can be done using a reflectance panel through the Empirical Line Method (ELM). A reflectance panel is a surface of known reflectance, which will allow to calibrate the sensor to the conditions at the time of flight adapting the model to the atmospheric conditions at the time of flight. Multiple reflectance panels are recommended to fit a linear model through the least-squared method. Mamaghani *et al* (2018) compared an ELM fitted with 2 panels and with 1 panel, following the MicaSense in-built procedure. Their results show that the use of 2 panels allow to reduce the mean reflectance errors by a factor of 2 compared to the use of only 1 panel. Another weakness of the ELM procedure comes from the fact that the images of the panel are taken by holding the UAV above the panel. There is a large

(invisibly) shaded part of the hemisphere by the operator, the UAV and the sensor. This can induce a severe bias that can account for up to 15% of the readings under cloudy conditions (Aasen *et al*., 2018), and it will affect the retrieval of vegetation parameters. Another bias comes from the variation of illumination conditions during the flight. Usually, an image of the panel before and after the flight allows to take continuous variations of the the atmospheric conditions in account through a simple interpolation of the model. Miura and Huete (2009) state that this method is suitable for flights shorter than 30 minutes and under stable illumination conditions. Since the panel is only recorded twice and is not visible in between, this procedure is not suitable for longer flights or flights with highly variable illumination conditions.

iv. Georeferencing

There are multiple methods for UAV image georeferencing. Two main methods can be distinguished: direct and indirect georeferencing (Tmuši *et al*., 2020). Each image's camera orientation and location are defined by six exterior orientation parameters (EOP). Direct georeferencing implies the use of additional positioning and orientation sensors: onboard differential global navigation satellite system techniques are currently being adapted for UAS photogrammetry. Such techniques require two GNSS receivers: a base reference station and a rover GNSS receiver mounted on the UAV. Indirect georeferencing require the use of accurately georeferenced GCPs to estimate the EOPs through the aerial triangulation method. Therefore, georeferencing's accuracy is dependent on the accuracy of the GCP measurements and the accuracy of the onboard sensors. Padró *et al* (2019) compared four georeferencing methods for environmental monitoring, and concluded that the use of geo-referenced GCPs is still the most affordable solution and still delivers accurate results. Martínez-carricondo *et al*. (2018) suggest that in order to maximize horizontal accuracy, GCPs must be positioned on the edge of the study area, and vertical error is minimized by placing them in a stratified way inside the study area.

Finally, the evaluation of the accuracy of geospatial products can be done by using independent checkpoints.

b. Wheat crop parameter monitoring

Wheat crop parameter monitoring through UAV-acquired imagery has been subject to many studies. In the following section, we will focus on some of the parameters that have been monitored through UAV imagery.

i. Crop height

Crop height is defined by Pérez-Harguindeguy *et al* (2016) as "the shortest distance between the upper boundary of the main photosynthetic tissues (excluding inflorescences) on a plant and the ground level, expressed in meters". Flying the UAV before crop growth allows to produce a 3D-reconstructed scenery of bare soil, thus furnishing a surface reference (DEM). Subtracting this reference to later Crop Surface Models (CSM) developed during crop growth results in a Crop Height Model (CSM). This method has been evaluated for a diversity of crops: sorghum (Chang *et al*., 2017; Malambo *et al*., 2018), maize (Li W. *et al*., 2016; Malambo *et al*., 2018), vine (Matese *et al*., 2017), sugarcane (Henrique *et al*., 2017), wheat (Holman *et al*., 2016) and rice and summer barley (Bendig, 2013). Holman *et al* (2016) have compared this method on wheat to standard ground measurements and Terrestrial LiDAR derived CSM. Results indicated the UAV-derived height measurements were more accurate than ground measurements. These results must be taken with care, since their

image's spatial resolution was very fine (0.9-2cm/pixel) and Grenzdörffer (2014) demonstrated that a degradation of spatial resolution implied a degradation of the correlation between UAV and ground-based measurements. Error sources are multiple: the UAV survey (ex. camera settings, illumination conditions, spatial resolution), data processing, and crop phenology.

Data processing choices will affect the resulting surface model precision. Haala (2013) proved the underlying feature matching algorithm relative to the software used for 3D-reconstruction will affect it. Choices made for the processing steps should also be optimized. For example., according to Holman *et al* (2016), "There was a clear advantage to using "mild depth filtering" during processing in order to achieve greater crop height accuracy, likely due to the reduction in "smoothing" of the model allowing for more of the small plant features of interest to remain during the depth filtering step of model processing".

Crop phenology is also a factor influencing the output precision of the CSM. For example, during the early growth stages, the sparse and low plants do not have a closed canopy which induces a high error rate. There is a minimum detectable height set between 5-15cm depending on crop variety (Grenzdörffer, 2014). On the other hand, in later development stages, canopy grows in density and the CSM becomes increasingly accurate. Grenzdörffer (2014) suggests the crop's canopy structure also determines the model's precision. A dense and flat canopy (ex. wheat) will deliver more accurate results than a sparse or spiky canopy (ex. corn), where the CSM tends to under-estimate crop height.

ii. Above Ground Biomass and potential yield

Wheat potential yield has been defined by Acevedo *et al* (2002) as the product of grain density and grain weight. Grain density is tailored by the number of grains per spikelet, spikelets per spike, spikes per plant, and plants per m^2. Figure 1 shows a simple diagram summarizing this hierarchy.

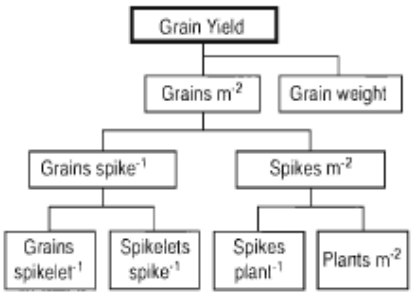


Figure 1: Components of grain yield. Source: Slafer et al., 1996.

These yield components are determined by environmental and/or genetic factors during specific development stages. Figure 2 is a visual representation of wheat development stages and the yield components that are determined during these periods. It shows grain density (grains.m^{-2}), one of the major yield components, is mainly determined before anthesis. The grain is filled after anthesis, determining the second major yield component, grain weight.

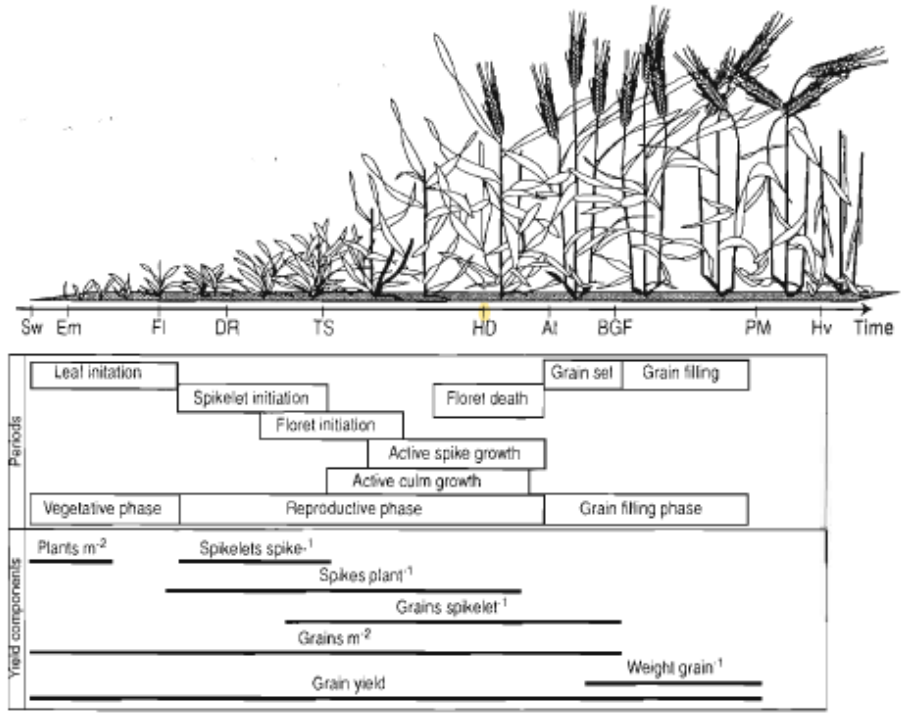


Figure 2: Yield components in relationship with wheat development stages. Source: Slafer et al., 1996.

These two components cannot be studied independently, since it has been proved (Slafer *et al*., 1996) an increase in grain density does not imply a proportional increase in yield. In fact, they observed that with the increase of grain number, the numbers of lighter weighted grains increased. Grain weight has been reported by Acevedo *et al* (2002) to be mostly determined by the plant's genotype, while grain density results of both genotype and environmental factors which can trigger a diversity of compensation mechanisms. Richards (1996) supports the idea that it is the partitioning of Carbon into the reproductive structures that determines yield rather than increases in biomass. The main environmental factors driving grain density in the absence of biotic or abiotic stresses is solar radiation and temperature (Fischer, 1985). Nitrogen availability has also been proven to be of major importance driving the sub-components of grain density. In fact, N deficiencies can impact the number of plants that develop in the early phases of growth, their capacity to grow into a spike, the number of spikelets per spikes and the number of fertile grains in each spikelet, depending on the deficiency's timing and intensity (Oscarson, 2000).

It has also been proven through several studies that the sink force, determined by the number of grains, is critical in yield potential and has a greater influence than the source force, resulting from post-anthesis photosynthetic products (Reynolds *et al*., 2009; Richards, 1996). For example, Richards (1996) trimmed green leaves to half their area after anthesis and observed the final yield was not affected: the plant compensated by increasing stomatal conductance. It appeared wheat plants have a photosynthetic system operating below it's potential, and the driving process is the number of grains: changes in leaf conductance, photosynthesis or temperature are a response to the increased demand. Post-anthesis grain filling is also host of a well-known phenomenon: mobilization and transfer of reserves to the grain. In fact, adverse conditions for photosynthesis are compensated by mobilizing Carbon reserves deposited in vegetative parts before anthesis (Saeedipour & Foad 2010). The relative contribution of the different parts of the plants in post-anthesis C mobilization has been studied by Gebbing *et al*., (1998). It has been reported that leaf blade, sheaths and the stem contribute to an extent of approx. 30% of their total carbon content at

anthesis for mobilization. Glumes and roots contributed to mobilization to an extent of respectively 18% and 44% of their total C content at anthesis.

For an accurate assessment of potential yield through RS at anthesis, it is important to weight plant organs separately. The pre-anthesis development phases are crucial since they determine grain density and the vegetative C stock destined to be allocated to grain filling. Nitrogen availability is also a critical factor that drives the sub-components of grain density. Thus, monitoring the weight evolution of plant organs separately would allow to evaluate the relative contribution of pre-anthesis C mobilization and post-anthesis photosynthetic products to grain filling, and their impact on final yield. RS-derived parameters could allow to estimate final yield through imagery acquired at anthesis since grain density is already determined at this stage, and post-anthesis grain-filling can compensate adverse conditions for photosynthesis.

Above-Ground Biomass (AGB) has been previously modeled in a number of studies through the use of spectral bands and Vegetation Indices, but a common limitation of such models is the loss of sensitivity of near infra-red and red band-based VIs, resulting in the loss of accuracy of the model at middle to late development stages. More recently, wheat monitoring studies focusing on the addition of plant height derived from 3D reconstruction in the model resulted in increased accuracy, overcoming the saturation problem (Bendig *et al*., 2015; Yue J *et al*., 2017).

AGB and height are particularly significant in assessing life activities of vegetation, monitoring crop growth and estimating potential yield (Yao *et al*., 2017). As stated before, crop height can be easily derived from SfM workflow and is a crucial parameter to improve the accuracy of AGB estimation through UAV imagery.

iii. Leaf Area Index

LAI ($m^2.m^{-2}$) is one of the most important biophysical parameters, defined as the total one-sided green leaf or needles area per unit horizontal soil (Chen & Black, 1992). It governs photosynthesis, respiration and transpiration of plants (Bonan, 1993). LAI is crucial in agronomical modelling, since it reflects the influence of biotic and abiotic conditions, especially yield-limiting and -reducing factors, such as mismanagement or plant diseases (Boegh *et al*., 2002).

There are various methods to estimate LAI. Destructive measurement is time-consuming, expensive, and therefore limited to small areas, but it is usually the most accurate method to this day (Bréda, 2003). Multispectral remote-sensing techniques have been increasingly used in the recent decades for LAI modelling, improving the spatial and non-destructive assessment of this parameter. Two main approaches have been developed in recent years using RS: the use of bio-physical radiative transfer models, or calibration of empirical-statistical models, calibrated with *in situ* LAI measurements, in order to establish a relationship between LAI and the canopy's reflectance spectra and/or derived spectral indices (Siegmann & Jarmer, 2015). The first approach does not require much calibration data, and by definition generalizes well. Compared to empirical approaches, their drawbacks include the strong computational requirements and the need of expert knowledge for model parametrization in order to permit successful outcomes (Berger, 2018). Empirical models, on the other hand, are well adapted to local context but require more calibration data, and are often limited to cross-validation for quality assessment. The major drawback of such models is the risk of overfitting: a phenomenon that can be defined as an error occurring when the model is too closely fit to a limited number of samples (Hawkins, 2004).

Empirical linear relationships can also be fitted between leaf area and leaf biomass. A relationship between leaf area and leaf dry matter (r^2 = 0.951) has been established by Aase (1978) for winter wheat. Aase (1978) also found that leaf area could also be linearly related to total plant dry matter (r^2 = 0.948) though the fifth growth stage (Feeke's scale). Such relationships are site- and cultivar-dependent but can be used to estimate LAI on larger areas where leaf biomass has been measured.

iv. Wheat phenology

Crop parameter monitoring should be associated with the plant's phenology. In fact, some models may only be valid between specific crop development stages. For example, Lu *et al* (2019) modeled wheat AGB through UAV imagery from jointing to anthesis using three regression techniques, and observed their accuracies degrading with growing development stages with the highest degradation between booting to heading. Liu Y. *et al* (2016) also noticed a relatively low accuracy for their leaf nitrogen content estimation model during late growth stages of wheat. On the other hand, Siegmann & Jarmer (2015) showed that an extended data range through the inclusion of data at multiple wheat development stages was essential in improving their LAI regression algorithms. They suggest models calibrated on datasets derived from multiple development stages, covering a wide range of LAI values could be used for LAI estimation of other wheat fields within the region.

In order to record the plant's development stage, a widely-used scale is Zadok's scale, which consists of a series of numbers from 0 to 99. The first digit accounts for the main growth stage (for ex. 1 for Germination, 4 for booting) but those are not sufficiently detailed for precise tracking of the plant. Its complementation with a secondary scale consisting of the second digit allows for more precise recordings (Zadoks *et al*., 1974).

c. Vegetation indexes

These various crop parameters have been studied through the use of hyperspectral Vegetation Indexes (VIs), themselves being combinations of different absorption bands. Most VIs are computed using the contrast between the red band, of maximum absorption due to chlorophyll pigments, and the Near-InfraRed (NIR) band, of maximum reflection caused by leaf cellular structure (Zhao *et al*., 2012). The use of VIs to estimate crop parameters at canopy scale is recommended thanks to their capacity to attenuate variation caused by canopy geometry, soil background, atmospheric conditions and sun view angles (Mutanga & Skidmore, 2004). As described by Basso *et al*. (2013) some VIs are a priori better correlated to some crop parameters, for example the Visible Atmospherically Resistant Index (VARI) is linked to LAI. One of the most utilized VIs is Normalized Differential Vegetation Index (NDVI). NDVI, like most VIs, is computed through a combination of red and NIR wavelengths (Treitz and Howarts, 1999) and reflects vegetation cover, vigor and growth dynamics amongst others. NDVI's main disadvantage is its saturation at high densities of vegetation (LAI > 2 m^2m^{-2}) (Zhao *et al*., 2012). Yao *et al.* (2017) successfully assessed the saturation problem of a number of VIs, and concluded that the $MTVI_2$ has higher sensitivity for high LAI values (2 to 7 $m^2\,m^{-2}$). Many VIs have been created and used empirically for different applications, and the choice of a VI has to be done cautiously, taking in account their respective advantages and limitations, and tailoring the choice to the sensors and platform used (Xue & Su, 2017). Most VIs have been developed for satellite or airborne imagery such as the Enhanced Vegetation Index (EVI), the modified triangular vegetation index (MTVI) or the soil adjusted vegetation index (SAVI). However, Yao *et al.* (2017) successfully modelled wheat LAI using 10 commonly used satellite-adapted VIs with UAV imagery, while Yue *et al.* (2017) and Lu *et al.* (2019) obtained high accuracies for wheat AGB estimation using a combination of VIs and canopy height metrics.

d. Machine Learning algorithms diversity in crop parameter modeling

The hyperspectral information in the form of reflectance values or VIs can be used in powerful machine learning regression algorithms such as Support Vector Machine (SVM), Partial Least Squares Regression (PLSR), Random Forest regression (RF) or Artificial Neural Network (ANN) for the assessment of specific crop parameters. For example, Siegman & Jarmer (2015) compared SVM, PLSR and RF algorithms for LAI assessment on wheat satellite reflectance data. They concluded "the regression model robustness largely depends on the applied validation approach and the data range of the LAI used for model building". Liu *et al*., (2016) compared 14 common regression algorithms to assess wheat leaf nitrogen content through UAV imagery, and concluded Random Forest regression was the best performing method. In their estimation of wheat Above Ground Biomass (AGB) using UAV-derived hyperspectral data and crop surface model, Yue *et al.* (2017) used PLSR to achieve high accuracies, while concluding on many studies indicating PLSR, ANN, SVM, RF, PCR or SMLR as robust algorithms to make full-use of hyperspectral information. Considering these conclusions, Random Forest models will be generated and tested in this work in order to assess the various crop parameters, and their relative accuracy and robustness will be tested. It is important to keep in mind that the applied validation method (cross-validation vs. independent validation) has great influence on the resulting accuracy

3. Present work objectives

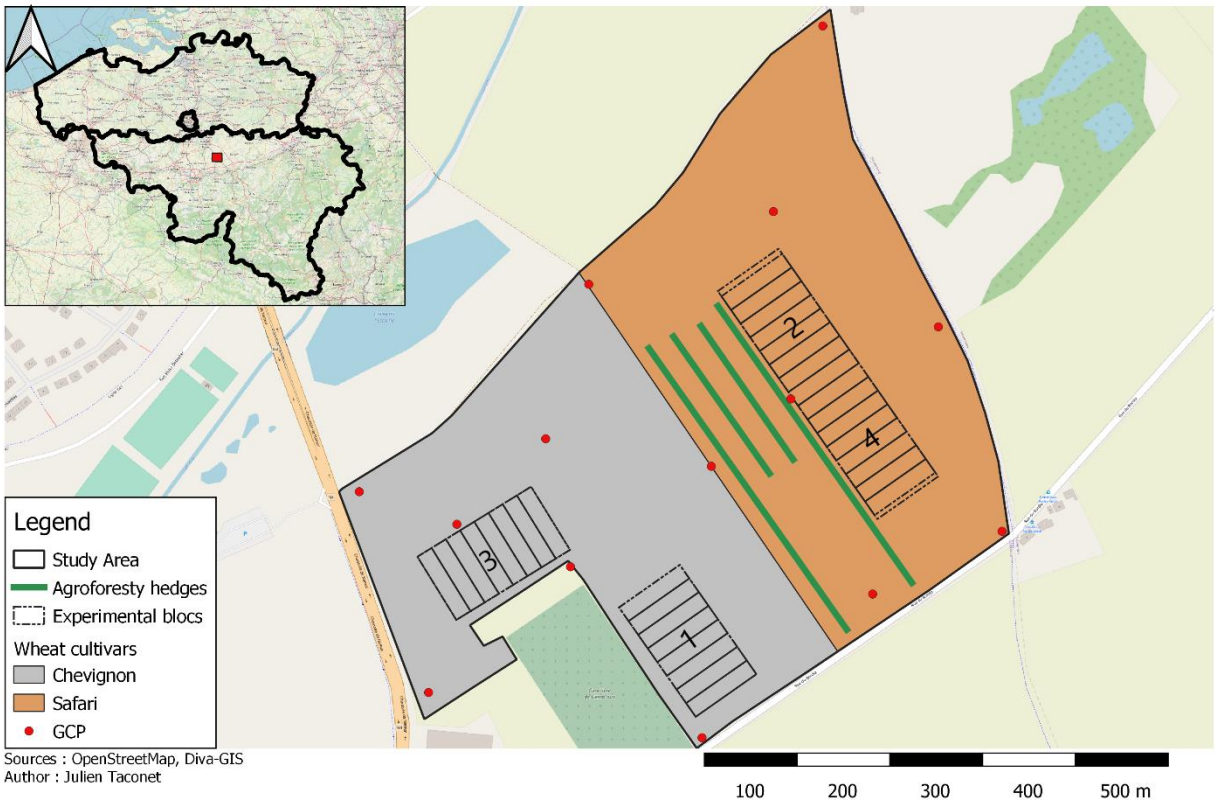
As stated earlier, AgricultureIsLife's experiment will start after the harvest of the homogenization crop. Growing the same crop during one year across the whole area as pre-treatment aims to reduce the effect of plant-induced in-field spatial variability. This operation will not fully erase soil heterogeneity, since environmentally-induced heterogeneity and the influence of past practices will still have a major influence on soils. Furthermore, one year is not enough to completely remove plant-induced in-field spatial variability. Thus, it is important to quantify the innate soil fertility differences due to field spatial variability before the beginning of the experiment. The primary objective of this end of studies work is to detect prospective variability within and between the experimental blocks through UAV-acquired imagery of winter wheat, by modeling and mapping of relevant crop parameters over the whole parcel. Careful repositioning of experimental blocks could then be performed in regards of the results of this work to ensure that the conditions inside each block are as homogeneous as possible. The homogenization crop of AgricultureIsLife's experiment consists of two different varieties of winter wheat, "*Chevignon*" and "*Safari*". The use of two varieties is justified by economic reasons and risk minimization.

Winter wheat will serve as an internal reference crop and is planned to be grown regularly across the 8-year rotations: twice in the rotation that takes the animal in account and four times in the rotation without animals. In order to reach the primary objective, regression models will be created to estimate relevant wheat crop parameters on the experimental area through UAV acquired imagery, while evaluating minimal requirements in terms of sensors and data collection frequency to obtain robust models. Heterogeneity of such crop parameters is driven by several intrinsic factors such as soil and past or present management techniques. The developed models could then be used to monitor wheat crop parameters during growth, attesting for their evolution across the whole experiment. Thus, they would account for the relative effects of the different crop management techniques (which includes amount and nature of inputs, pest and disease control and many more such as the agroecological levers that are to be set up) in regards of their initial spatial distribution.

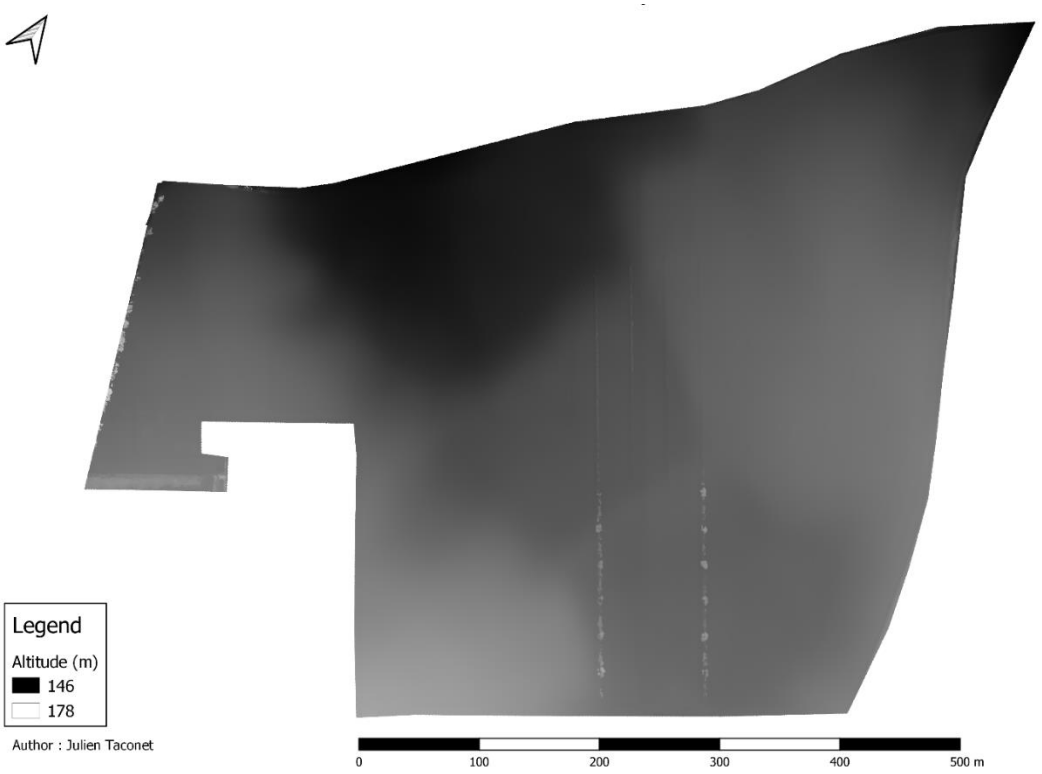
4. Material and Methods

a. Study Area

The study area is located in Gembloux, Belgium. It covers 27.7 hectares, and a UAV flight on bare-soil in November 2019 indicates an altitude variation of 20 meters (between 146 and 166 meters). The location and shape of the study area, as well as the partitioning of both varieties, the theoretical positions of the experimental blocs, their identification numbers, and other constitutive elements of the study area are shown in Map 1. The two varieties of winter wheat were sowed in November 2019, with a 12cm inter-row spacing. This area was formerly used for a trial on spring crops (oat, barley ...), which was located at the right of the agroforestry hedges, roughly covering bloc number 2 and 4. This experiment implied nitrogen and fungicides inputs. Another trial on spelt was carried out at its right. The absolute altitude on the whole parcel is presented through a DEM, extracted from the November 2019 flight, in Map 2.



Map 1: Location of the study area and it's features. GCP = Ground Control Points. The numbers in the experimental blocs correspond to their bloc number.



Map 2: Digital Elevation Model of the study area (extracted from the November 2019 flight).

b. UAS survey description

Currently, UAS publications do not always provide a detailed description of the specific data collection and processing steps (Tmuši *et al.,* 2020). In order to be in accordance with the rise of UAS data collection for environmental studies, the following section was built, as far as the required data was available, following Tmuši *et al.* (2020) recommendations to "ensure reproducibility but also to guarantee a certain quality, while also advancing and educating the broader field".

i. Study design

1. Platform characteristics

The UAS system used in this work is composed of a quadracopter rotary wing- the DJI Phantom 4 PRO - it's integrated RGB camera and a 3D-printed frame was integrated to add a MicaSense multispectral camera to the system. The UAV is capable of vertical take-off and landing, and weighs 1388 g (with an additional 316 g due to the micasense and its frame). The battery allows it to reach a maximum autonomy of 15 minutes per flight under optimal weather conditions, at a maximum speed of 6m.s^{-1}.

2. Sensor characteristics and Camera settings

Two cameras were mounted on the UAV. The first one is the integrated RGB camera from the DJI phantom Pro, the SONY IMX183CLK. The second is the MicaSense RedEdge 3 for multispectral data acquisition. Their technical specifications are reported in Table 1 and their spectral specifications are informed in Table 2 and Table 3 for the RGB and the MicaSense RedEdge respectively.

Table 1: RGB and Multispectral sensor's characteristics.

Sensor	Weight (g)	Sensor size (mm)	Sensor size (pixel)	Focal length (mm)	Min shutter speed (s)	Radiometric resolution (bit)	Pixel size (pitch) (mm)
P4PRO	138	8.8 x 13.1	5472 x 3648	8.8	1/2000	12	0.0024
MicaSense RedEdge 3	150	3.6 x 4.8	1280 x 960	5.5	1/500	16	0.00375

Table 2: SONY IMX183CLK spectral characteristics.

Band Name	Center wavelength (nm)	Band width (nm)
Blue	*460*	*Unknown*
Green	*525*	*Unknown*
Red	*600*	*Unknown*

Table 3: Micasense Rededge 3 spectral characteristics.

Band Name	Center wavelength (nm)	Band width (nm)
Blue	475	20
Green	560	20
Red	668	10
RedEdge	717	10
NIR	842	40

3. Flight plan

The flight altitude was set at 100 meters in order to get a fine spatial resolution, tailored by the area to survey, UAV's flight autonomy (15 min) and cruise speed (6 $m.s^{-1}$). This resulted in a spatial resolution of approximately 2.75 cm/pixel for the drone's integrated RGB camera, and 7cm/pixel for the Micasense RedEdge 3. The minimum shutter speed relative to both sensors was respected (see Table 1). The RGB camera was set on variable mode (Mode A – Aperture priority), so that the shutter speed varies following the instant's brightness, ensuring the image's quality not to be affected by motion. The cameras were oriented nadir at 90°. The overlap between flight lines (sidelap) was set at 80% and the overlap between photographs (frontlap) at 80%. They were set at these values through a compromise between the mosaicking accuracy and the flight time, tailored by the UAVs autonomy, and in accordance with literature (Tmuši *et al*., 2020).

4. UAS control software

The flight is fully automatic and the path was planned on DJI P4PRO's *dronedeploy* flight planning software. This software does not allow for terrain-follow functionality in order to keep a fixed altitude relatively to ground elevation, thus it does not automatically keep a fixed spatial resolution on the whole study area.

5. *Georeferencing*

In the present work, indirect geo-referencing of the images was achieved through the use of Ground Control Points (GCPs). These are 50cm * 50cm targets positioned in the field. 10 GCPs were positioned as recommended by Martínez-carricondo *et al*. (2018). They were previously accurately geo-referenced with the Emlid REACH RS Real Time Kinematic (RTK) GPS. This RTK GPS uses the Walcors network with satellite constellations GPS/QZSS L1, Glonass G1 and Galileo E1. The GCP spatial distribution is shown in Map 1.

ii. Flight mission

In order to cover the whole area, the flight planning software indicated a flight time of 38:59 minutes. However, the additional weight induced by the Micasense and its frame extended the flights duration to 45-60 minutes, corresponding to 4 flights. A screenshot of the flight's waypoints is shown Figure 3. There are 26 flight lines, 27 meters apart. The image file format is of great importance: the RAW file format (unprocessed) is considered by many researchers to be superior to other compressed formats such as JPEG or PNG, since the latter may introduce unnecessary noise (Tmuši *et al*., 2020). However, in this work, the output for the RGB camera was JPEG since it is unchangeable, and TIFF for the multispectral camera.

4 flights were carried out during wheat growth. The dates of flights, sky conditions and the associated wheat development stages are reported in Table 4.

Figure 3: Screenshot of the flight plan (dronedeploy app).

Table 4: Sky conditions and wheat development stages during each flight.

Date of flight	Sky conditions	Chevignon development stage	Safari development stage
25/05/2020	Sparse clouds & direct sunlight	55	47
09/06/2020	Cloudy	69	65
22/06/2020	Medium clouds & direct sunlight	75	71
13/07/2020	Direct sunlight	91	87

iii. Processing of aerial data

1. Geometric processing

The geometric processing of the flight images was carried out on Agisoft Metashape 1.6.3. After the photogrammetric processing workflow, the final products are for each camera, a Digital Elevation Model and an orthomosaic map of the study area. The general workflow for such extraction starts with the generation of a low-density point cloud based on the calculation of camera locations. These locations are calculated through the GCPs visible on the images. The second step is the generation of a high-density point cloud and finally, using the latter to build a georeferenced orthomosaic. The parameters specified during these steps for the creation of the final products are reported in Table 5.

Table 5: Parameters specified in Agisoft Metashape for the processing workflow.

Parameter	RGB outputs	Multispectral outputs
Accuracy	High	High
Tie point limit	40.000	40.000
Key point limit	10.000	10.000
Quality Dense cloud	High	Low
Filtering mode	Agressive	Aggressive

2. Radiometric processing

A combination of the Empirical Line Method (ELM) and an illumination sensor was used on the Micasense Rededge camera in order to generate reflectance images from the images. For the ELM, a picture of the built-in Micasense grey panel of known reflectance was taken before and after the flight, in order to interpolate reflectance values between both pictures. The picture was taken by the drone operator, by flying the drone 2-3 meters over the panel, in order to reduce the shaded hemisphere due to the operator's body. However, technical difficulties occurred on some flights (connection problems between the UAV and the laptop), for which the calibration pictures were taken by holding the drone approximately 2 meters over the panel, taking care not to directly shade the panel. The illumination sensor is mounted at the top of a rod over the drone. It continuously measures solar irradiance, and is used to correct global changes in irradiance, such as a when the sky is completely overcast. Its limitation comes when clouds obstruct the sun which casts a shadow on the ground, but not on the sensor. The sensor will not be able to correct for these situations.

These two methods are used in combination: while the illumination sensor account for global changes of irradiance at high-altitude, the panel is used to correct measurements on the ground. The photogrammetry software integrates both information for radiometric correction, to create accurate reflectance maps.

On the other hand, the RGB camera uses none of these methods for radiometric correction. Reflectance maps thus suffer greatly from changes in illumination during flights.

iv. Accuracy assessment

The global errors, and for each axis, related to each flight and camera were extracted from Metashape and are given in Table 6.

Table 6: SfM photogrammetry products accuracy assessment.

	RGB				Multispectral			
	X (m)	**Y (m)**	**Z (m)**	**Global (m)**	**X (m)**	**Y (m)**	**Z (m)**	**Global (m)**
25/05	0.006	0.009	0.001	0.01	0.009	0.007	0.001	0.01
09/06	0.005	0.01	0.006	0.01	0.005	0.007	0.007	0.008
22/06	0.009	0.01	0.01	0.02	0.006	0.007	0.001	0.01
13/07	0.005	0.01	0.017	0.02	0.008	0.006	0.007	0.011

c. Ground-truth data sampling

Ground-truth data acquisition started as soon as the UAV flight ended to be as synchronous as possible with the imagery and minimize the effect of plant growth. The ground-truth sampling protocol is provided in the Annex 1: Ground-truth sampling protocol.

A 50cm-long stick was used to delimitate the sampling surface, and 3 plant rows were collected on this length. Considering the 12cm inter-row spacing, the sampling surface was 1800cm^2, and was determined through a compromise between time consumption for sampling and the number of pixels for the sample to be representative. Considering the spatial resolution of the imagery, this surface provided more or less 240 and 35 pixels per sample for the RGB and multispectral orthophotos respectively. To calculate canopy height metrics from the CSMs, the sampling surface provided approximately 60 pixels for the RGB-derived CSMs. The spatial resolution of the multispectral-derived CSMs are coarse and higher than the sampling surface, thus the multispectral CSM values are based on the few pixels that intersect the sampling polygon.

The aim of the sampling campaign was to cover at best the in-field variability of biomass and LAI at each date, while taking the same number of samples in each wheat cultivar. To achieve this, the whole study area was covered, and the samples were selected on the basis of the operator's visual inspection. 20 biomass samples were collected per sampling campaign. Out of these 20 samples, 10 were assigned to build a LAI-leaf dry biomass relationship, on which a fourth row of plants was collected. Technical considerations in terms of maneuverability and ease of positioning during sampling were also taken in account. The spatial distribution of the samples is shown in Map 3. Each sample's location was precisely geo-referenced using the RTK-GPS.

Map 3: Location of the samples for each date.

i. Crop height

Crop height over the whole study area will be derived from the CSM built through SfM photogrammetry. The CSM for each UAV flight was built by subtracting the flight's Digital Surface Model (DSM) to the bare-ground Digital Terrain Model (DTM). Considering the sources of uncertainties discussed in Section 2.b.i, the accuracy of the CSM derived from UAV imagery was not evaluated by comparison to ground truth measurements since the latter also has high sources of variability.

ii. Above-Ground Biomass

Ground-truth data was acquired through destructive above-ground biomass sampling. Three crop rows of 50cm were cut at ground level, put in paper bags, brought to the laboratory and organs (stems, leaves, flag leaf and ears) were separated (see Figure 4). The samples were then ceiled in a microperforated bag, dried in the oven at 60 degrees for 7 days before being weighted to obtain dry biomass. AGB was calculated by rescaling the values to kg per m^2, using the following formula:

$$AGB = \frac{dry\ biomass\ sample}{1000} * \frac{10000}{1800}$$

Equation 1: AGB rescaling method, with AGB expressed in [kg.m^{-2}] and dry biomass samples in [g].

The 1800 factor in Equation 1 corresponds to the sampling area, equal to three times 12cm (inter-row spacing) times 50cm (sample length).

Figure 4: Ground-truth sampling protocol: plant organs separated.

iii. Leaf Area

Leaf area (LA) is an extremely time-consuming parameter to measure. It is collected through destructive sampling of a unit area. The leaves are separated from the stem, fresh biomass is weighted and they are spread out on a sticky paper. They are then applied on a standard A3 or A4 white paper for scanning. A 400dpi resolution and PDF files were chosen as output scanning parameters. Finally, these files were submitted to an image analysis algorithm that determines the leaf area, based on the RGB pixel values. Depending on the growth stage of the plant, the necessary time to prepare the sample, spread out the leaves, scan and treat the images can reach up to 1h per 600cm^2 sample.

Due to these considerations, LAI was measured on half of the biomass-destined samples. An empirical linear relationship between leaf area and leaf fresh biomass was created for each sampling date and used to extrapolate the LAI to all samples.

Figure 5 shows an example of a scanned A3 paper. This operation resulted in the collection of 10 samples per date for the creation of the leaf area – dry biomass relationship. Two methods were used in order to extract the LAI sample's dry leaf weight. The first method consists in hypothesizing that the fourth cut row destined to LAI is representative of the 3 other rows destined to AGB. The LAI sample's leaf dry matter is considered equal to the biomass sample's leaf dry matter divided by 3. This approximation has been evaluated by extracting the stem's and/or ear dry biomass from the LAI-destined samples and comparing them to the biomass-destined sample. The second method

consists in rescaling the LAI sample's leaf dry matter by a factor, equal to ratio of the biomass stem's dry matter by the LAI stem's dry matter. This method relies on the hypothesis that the ratio between leaf dry biomass and stem dry biomass in a sample does not vary between two adjacent biomass and LAI samples.

Figure 5: Example of an A3 paper with wheat leaves spread out for LAI measurements.

d. Data pre-treatments

The raw data read in Rstudio consisted in the RGB- and multispectral-derived reflectance orthophotos, and the RGB- and multispectral-derived DSMs. A summary of the pre-treatments they have been submitted to is shown in Figure 6.

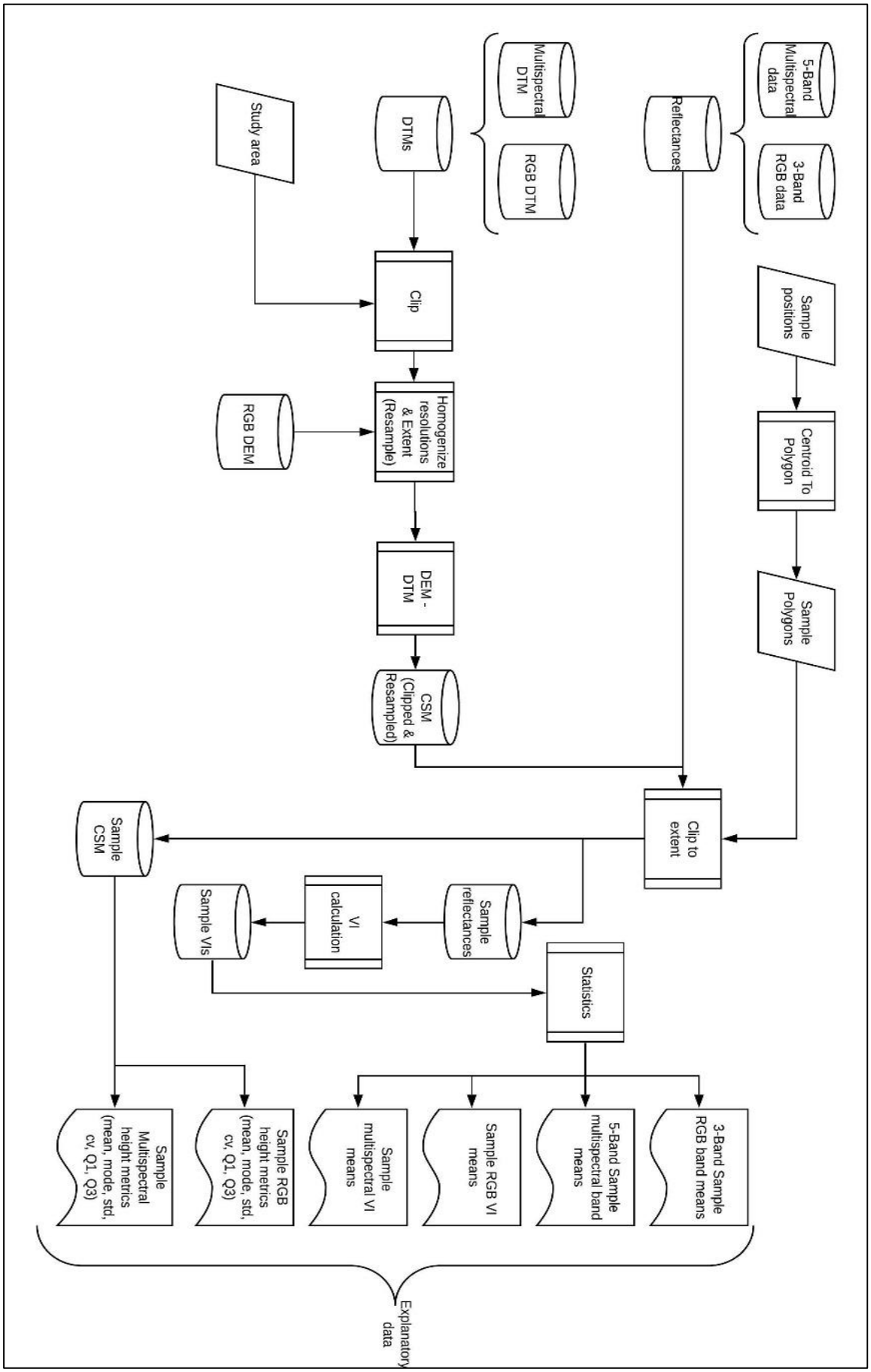


Figure 6: Workflow of the data pre-treatments.

i. Polygonising samples

The samples geo-referenced center positions were imported in the Rstudio environment as shapefiles. 50cm*36cm Regions of Interest (ROI) were delineated and extracted around the sample centers using Rstudio. These are rectangle polygons, representing the sampling surface and oriented in the same direction as the crop lines. An R function was created in order to accomplish this task, findable in the Annex 2: « CentroidToPolygon.R ».

An example of the result of this operation is visible in Figure 7, which shows a polygon built around a geo-referenced center of sample collected on the first flight date, with the second flight's RGB orthophoto as background. Bare-ground is visible on this figure under the sample center, since the orthophoto originates from the flight date following the sampling. Of course, the imagery variables used in the models are extracted on the vegetation cover within each sample polygon, as the orthophotos used in this operation were created using the flight.

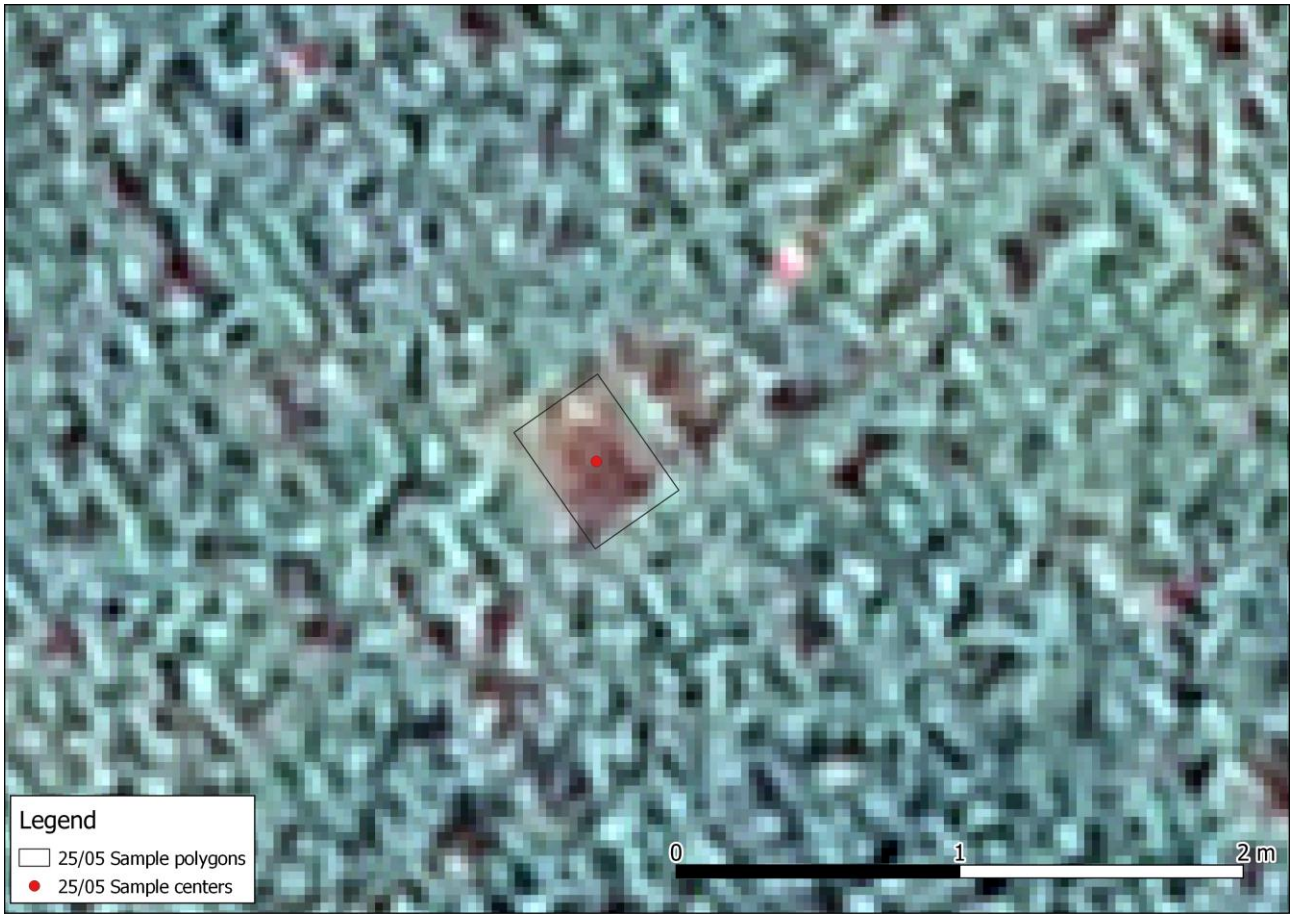


Figure 7: Example of a polygon built around a sample centroid.

ii. Canopy height metrics

The DSM was subtracted to the DTM for each flight, which results in a Crop Surface Model (CSM). In order to achieve this operation, the DTM and the DSMs had to be homogenized in terms of resolution and spatial extent. First, they were all clipped to the study area's polygon. Then, the RGB DSMs were all homogenized to the original DTM's spatial resolution (5.5cm/pixel). The multispectral DSMs were homogenized to the first flight's spatial resolution (57cm/pixel, 25/05). Since the UAV was not equipped with the multispectral camera during the flight in November 2019, which allowed

to extract the DTM, the multispectral DTM is a resampled version of the RGB DTM, with a 57 cm/pixel spatial resolution. All resampling treatments were done using the "bilinear interpolation" method in Qgis.

Canopy height metrics were then derived from the RGB CSM at each UAV flight date: zonal statistics were extracted for each ROI. Mean height, median height, percentile heights (25%, 75%), minimum and maximum values, standard deviation of height and coefficient of variation were used as input variables for the different models. Only the mean of the multispectral CSM was extracted, since they had a very coarse spatial resolution roughly equal to the size of the sampled areas.

iii. Spectral bands and Vegetation Indices

The significance of the raw reflectance of each band (475nm, 560nm and 668nm, 717nm and 842nm) was assessed to build the crop parameter models. In addition to these variables, Vegetation Indices were used in this work and selected on the basis of former wheat crop parameter modeling studies (through UAV imagery). A wide variety of VIs were tested to model the various crop parameters assessed in this work and are summarized in Table 7. The mean of the VIs and of the spectral bands within the ROI were extracted with the extract() function (raster package), and used as explanatory variables for the models. In order to avoid any pixel-value distortion through resampling, the VI calculation and their extraction in the ROI were done using the raw orthophotos.

Table 7: Summary of VIs derived from UAV imagery for crop parameter estimation. R,G,B,Re and NIR stand for red, green, blue, Red-Edge and Near Infra-Red bands respectively.

Index	Name	Formulation	References
RVI	Ratio Vegetation Index	$\frac{NIR}{Re}$	Yue *et al.*, 2017
NDVI	Normalized Difference Vegetation Index	$\frac{NIR - R}{NIR + R}$	Yue *et al.*, 2017
TVI	Triangular VI	$0.5(120(Re - G) - 200(R - G))$	Yue *et al.*, 2017
MTVI1	Modified Triangular VI	$1.2(1.2(NIR - G) - 2.5(R - G))$	Yue *et al.*, 2017
MTVI2	Modified Triangular VI	$\frac{1.5(1.2(NIR - G) - 2.5(R - G))}{((2NIR + 1)^2 - \left(6NIR - 5R^{\frac{1}{2}}\right) - 0.5)^{1/2}}$	Yue *et al.*, 2017
SAVI	Soil-Adjusted VI	$\frac{NIR - R}{NIR + R} * (1 + L)$	Huete *et al.*, 1988
MSAVI	Modified Soil-Adjusted VI	$0.5(2NIR + 1 - \left((2NIR + 1)^2 - 8(NIR - R)\right)^{\frac{1}{2}})$	Yue *et al.*, 2017
GBNDVI	Green-Blue NDVI	$NIR - NIR(G + B) + (G + B)$	Yao *et al.*, 2017
RDVI	Renormalized Differential VI	$(NIR - Re) * (NIR + Re)^{1/2}$	Yao *et al.*, 2017
GNDVI	Green Normalized VI	$\frac{NIR - G}{NIR + G}$	Yao *et al.*, 2017
MSR	Modified Simple Ratio	$\frac{NIR/_{Re} - 1}{(NIR/_{Re} + 1)^{0.5}}$	Yao *et al.*, 2017
EVI	Enhanced VI	$\frac{2.5 * (NIR - R)}{(NIR + 6R - 7.5G + 1)}$	Yue *et al.*, 2017
EVI2	Enhanced VI	$\frac{2.5 * (NIR - R)}{(NIR + 2.4R + 1)}$	Yue *et al.*, 2017
GI	Greenness Index	$\frac{G}{R}$	Yue *et al.*, 2017
OSAVI	Optimized Soil-Adjusted VI	$\frac{1.16(NIR - R)}{NIR + R + 0.16}$	Yue *et al.*, 2017

WDRVI	Wide Dynamic Range VI	$\frac{0.1NIR - R}{0.1NIR + R}$	Yue *et al*., 2017
VARI	Visible Atmospherically Resistant Index	$\frac{G - R}{G + R - B}$	Lu *et al*., 2019
ExG	Excess Green VI	$2 * G - R - B$	Lu *et al*., 2019
ExB	Excess Blue VI	$\frac{1.4 * B - G}{G + R + B}$	Lu *et al*., 2019
ExR	Excess Red VI	$\frac{1.4 * R - G}{G + R + B}$	Lu *et al*., 2019
ExGR	Excess Green minus Excess Red	ExG - ExR	Lu *et al*., 2019
GRVI	Green Red VI	$\frac{G - R}{G + R}$	Lu *et al*., 2019
MGRVI	Modified Green Red VI	$\frac{G^2 - R^2}{G^2 + R^2}$	Lu *et al*., 2019
RGBVI	Red Green Blue VI	$\frac{G^2 - B * R}{G^2 + B * R}$	Lu *et al*., 2019
IKAW	Kawashima Index	$\frac{R - B}{R + B}$	Lu *et al*., 2019
GLI	Green Leaf Index	$\frac{2 * G - R - B}{-R - B}$	Lu *et al*., 2019
DVI	Difference Vegetation Index	$NIR - Re$	Yue *et al*., 2017
DVI2	Difference Vegetation Index	$Re - R$	Yue *et al*., 2017
DVI3	Difference Vegetation Index	$G - R$	Yue *et al*., 2017

e. Model creation

i. Random Forest algorithm

1. Algorithm description

The Random Forest (RF) algorithm, as described by Breiman (2001), builds a high number of decision trees based on bootstrapped datasets from the original dataset, and uses a random subset of variables at each node of a tree (also called feature bagging). A bootstrapped (or bagged) dataset is a random selection of observations, without replacement. Feature bagging allows to avoid the generation of correlated trees: if a few variables are strong predictors of the dataset, they will be systematically selected in each tree, making them correlated (Ho T.K, 2002). In addition, bootstrapping allows to calculate a specific error rate, the Out-Of-Bag (OOB) error rate. Each sample is used as input in the trees that do not contain it in their bootstrap sample, and the proportion of incorrectly classified samples this way gives the OOB error. In regression, the accuracy measurement is based on the OOB predictions, defined as the pseudo R-squared value. The formula to calculate it is shown in Equation 2.

$$\text{pseudo R} - \text{squared} \;=\; 1 \;-\; \text{MSE} \,/\, \text{Var(y)}$$

Equation 2: Random Forest's Pseudo R-squared formula, where MSE = Mean Square Error, Var(y) = variance of targets.

Three parameters describe the RF algorithm, and are susceptible to fine-tuning: the minimum size of terminal nodes, the number of trees grown, and the mtry parameter, which defines the number of variables randomly selected at each node. For regression purposes, the default value of the minimum nodesize is 5, the number of trees grown is 500 and mtry is p/3, where p is the total number

of explanatory variables. Amongst these 3 parameters, mtry is the most sensitive for prediction quality and optimizable (Brostaux, 2017).

2. Error sources

The variance error is related to the small fluctuations in the dataset, and a high-variance algorithm causes it to model the noise in the dataset. This is known as overfitting. The bias error, on the other hand, is an error caused by erroneous assumptions on the dataset, thus missing the important relationships between predictor and response variables. Machine learning algorithms have to trade-off between these two errors, which is known as the bias-variance dilemma (Kohavi *et al*., 1996). Single decision trees are usually built on a high variance and low bias trade-off. In a RF, the averaging of decision trees aims to reduce the model's variance, while keeping the bias unchanged. This means that they are almost insensitive to overfitting, even when increasing the number of trees (Breiman, 2001). The risks in overfitting come from a dataset with very simple relationships, where the predictions are very close to each other, thus the variance reduction effect is limited. The mtry parameter can also be a source of over-fitting: if this parameter is set too high, the trees will look very similar and be correlated.

3. Variable importance

Variable importance can be computed in a RF and is a measure of the prediction strength of the variables. Two variable importance measures have been implemented. In a regression algorithm, the first one is called "percentage of increase in Mean Square Error", in which the accuracies when the OOB samples are passed down each tree are recorded. Then, the values of the studied variable are permuted through the OOB samples: the decrease in accuracy due to the permutation is a measure of the variable importance (Trevor *et al*., 2017). The second one, called "mean decrease in impurity", is based on averaging, across all trees, the reduction in the sum of squared errors whenever a variable is chosen to split a node.

Careful attention has to be paid when analyzing the results of the mean decrease in impurity graphs. Strobl *et al*., in 2007, pointed out that "the variable importance measures of Breiman's original Random Forest method [...] are not reliable in situations where potential predictor variables vary in their scale of measurement or their number of categories.".

4. Conclusions on the RF algorithm

Compared to other machine learning algorithms, the random forest is less susceptible to overfitting through the averaging of multiple trees, hands a measure of variable importance, has an in-built validation measure (OOB score) and only has a few hyper-parameters susceptible to tuning. On the other hand, it acts as a non-interpretable "black box" and its complexity makes it computationally expensive.

ii. Present work models

1. Creation of different input datasets

Many different random Forests were created and assessed in this work. In order to evaluate the requirements in terms of equipment to obtain a robust model, 3 major datasets were created. The first one was built using the RGB camera-derived variables exclusively. A second dataset was created with the multispectral derived variables. The last one contains both RGB and multispectral derived variables (see Figure 8 for a workflow of the model creation steps).

Two subsets were made within each of these major datasets, to evaluate the impact of wheat cultivars on the built models (one subset with both cultivars and one for each cultivar separately). In total, 9 input variables datasets were created. Three crop parameters were modeled through a Random Forest algorithm on each dataset: LAI, total dry AGB and ear dry biomass.

Another subset was created within the major datasets, removing the data coming from the last flight, where wheat had a very different spectral response due to leaf senescence, in order to evaluate the impact of the addition of the ground truth data coming from this flight on the models. The separation of cultivars was not made after this subset, considering that 20 to 30 samples per model is not sufficient to draw conclusions.

2. Model parametrization

Each dataset was submitted to outlier detection with the outForest function (outForest package): this recent algorithm uses the Random Forest model to regress a given variable using all the others in the dataset. If the scaled absolute difference between the predicted value and the observed value is above a certain threshold, the value is considered as an outlier. The default value of threshold (3) were left untouched in this work, and the replacement method used was “replace”, which imputes the outlier with the predicted value.

After the outlier detection, a variable selection algorithm was applied to the models. The VSURF algorithm (VSURF package) was used, which works in three steps. The first step (thresholding) allows to eliminate the irrelevant variables from the dataset, by computing a high number of Random Forests, and computing each variable’s mean importance to eliminate the ones under the given threshold. The second step consists in building random forests with a growing number of variables selected in step 1. Each model’s Out-Of-Bag (OOB) error estimation is stored and the model with the lowest OOB error is selected. Finally, step 3 starts from the same point as in step 2, but each variable is included in the final model if it’s impact on the decrease of the OOB error when added to the model is higher than a given threshold.

Finally, the mtry parameter (number of variables randomly sampled as candidates at each node) of the resulting Random Forest was fine-tuned, using the tuneRF function (randomForest package). This function computes the optimal mtry parameter value with respect to the OOB error estimate.

Each model was created using the same seed, set at 97.

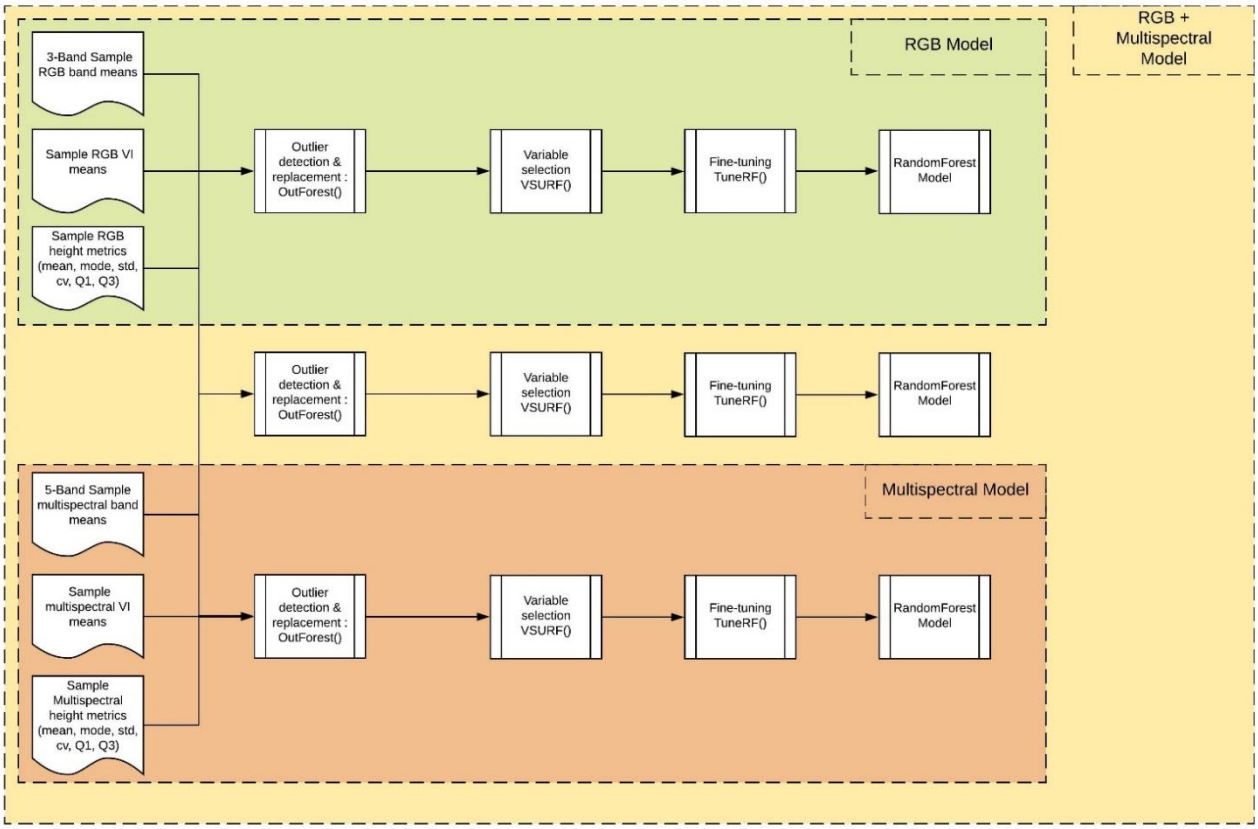


Figure 8: Workflow of the random forest optimization and creation of the major datasets.

f. Model predictions

The most accurate model between the mixed-cultivar models, for each crop parameter assessed in this work, was used to predict their values on the whole study area. This choice was made in a matter of generalization, and considering the fact that the mixed-cultivar models were built using twice the data as the single-cultivar models. In order to use the predict function from the random forest package on a raster object (a rasterstack in this case), they must all be on the same spatial resolution and extent. For each date and for the models using mixed sources of variables, RGB and multispectral orthophotos, as well as RGB and multispectral CSMs were resampled to the RGB's orthophoto's spatial resolution. The relevant VIs for each distinct model were calculated. Finally, a rasterstack object was created for each model with its relevant variables, and used as input for the predict function.

Finally, the output rasters were used to generate maps of the crop parameter variations through the flight dates. A workflow of the steps used to generate the predicted crop parameter maps is shown in Figure 9.

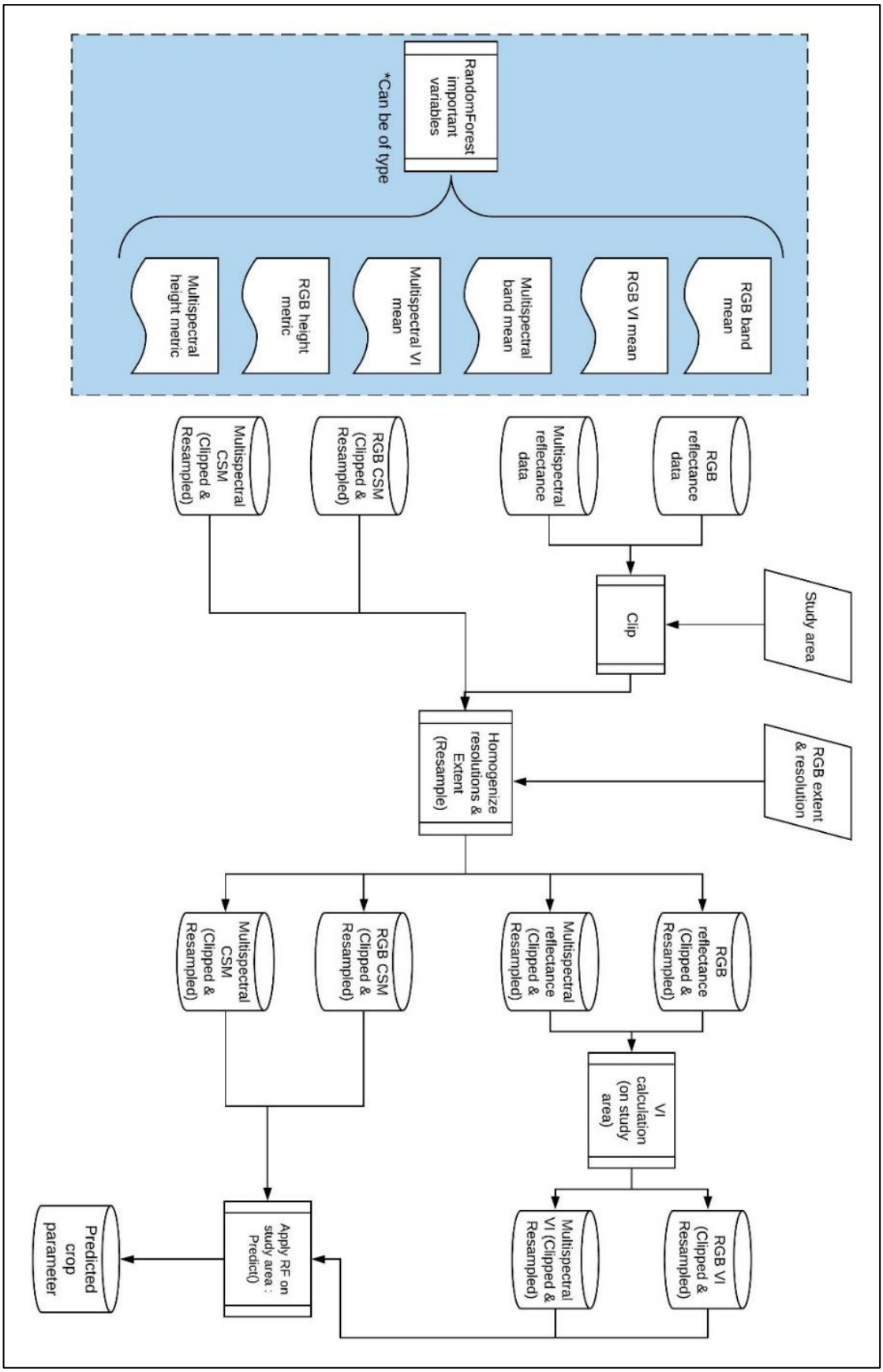


Figure 9: Workflow of data preparation for the prediction step of the random forest.

g. Analysis of Variance of modeled parameters on experimental blocs

The LAI parameter predictions were used as basis to study the in-field heterogeneity, as LAI is a good indicator of plant health and soil conditions. Another parameter that could have been chosen for this task is the ear dry biomass, but in the case of wheat, compensation of this parameter by carbon remobilization homogenizes its response to heterogeneous growth conditions. On the other hand, low LAI values detected on wheat are a good indicator of lower yields for some of the following crops. The 25/05 flight was used for the study, as the LAI predicted values are the most variable at this date, since the leaf apparatus is completely developed. A LAI relative deviation map was created using Equation 3.

$$Relative\ LAI\ deviation = \frac{LAI_i - \overline{LAI}}{\overline{LAI}}$$

Equation 3: Relative LAI deviation formula, with the relative LAI deviation expressed in [%], LAI_i is the LAI value of the i_{dth} pixel, and $\overline{LAI}$ is the mean of the LAI values on the whole study area.

These values (set between -100% and 100%) were then regrouped in 5 classes.

On the basis of the LAI relative deviation map, a proposition of new positions for the experimental blocs was proposed, by setting them in order to cover lower values of relative deviation. This modification was saved as a new shapefile.

Finally, the pixel values within the original and the new experimental blocs were extracted, along with their bloc and unit number. One-way Analysis of Variance tests of means (ANOVA) were run between the major blocks within each block in order to check if the proposed displacement decreased, as planned, the within-bloc variability. This can be done by observing the evolution of the F-statistic, which in this situation is proportional to the ratio of the variation between groups by the variation within groups. The variation within groups is quantified as the sum of squares of the differences between each pixel and its experimental group mean (SSE, error sum of squares), and the variation between groups is quantified as the sum of squares of the differences between each groups mean and the grand mean, mean of all values in all units (SSEmean, total sum of squares for the mean). Studying these output ANOVA parameters allow to give evidence for the evolution of the homogeneity between the experimental blocs, and within the experimental blocs and units.

5. Results

a. LAI – leaf dry biomass relationship

The LAI – leaf dry biomass relationship obtained through the procedure described in section 2.b.iii is presented in Figure 10.

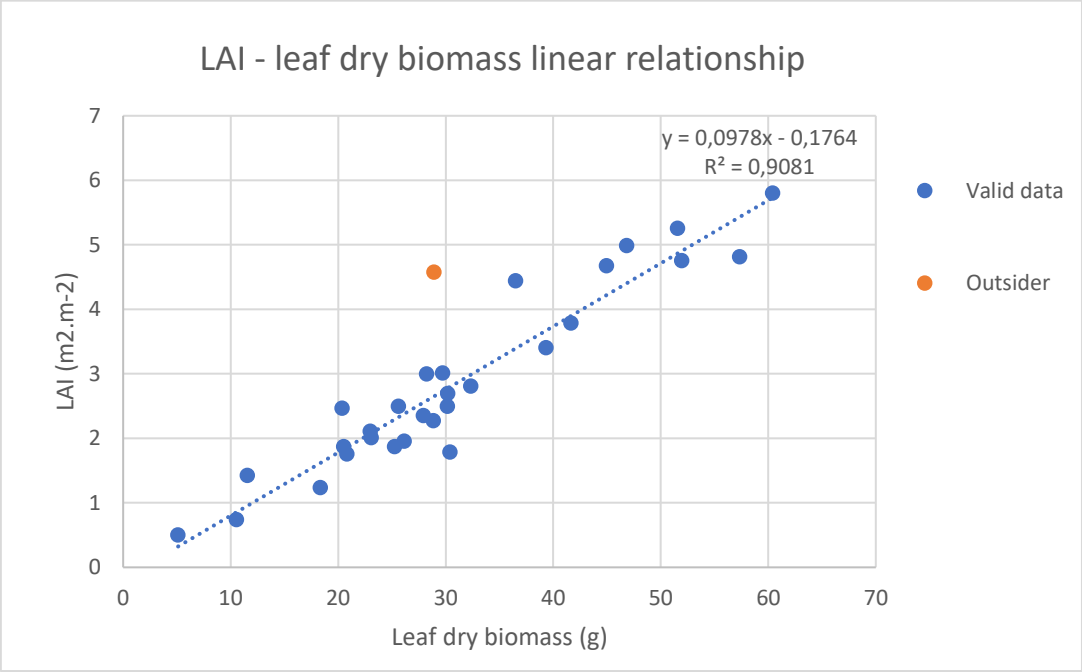


Figure 10: LAI - leaf dry biomass relationship.

This relationship is almost perfectly linear, with an R^2 value of 0.908. It has been built using 30 measures of LAI, and by removing one outsider value. It was built in order to extrapolate LAI values, which were measured on only half of the samples, to all samples.

b. Model accuracies

The pseudo R-squared value (or percentage of variance explained) of each model built with all the ground-truth data is reported in Table 8. These values originate from the final models that were used for predictions, after variable selection and outsider management. In addition, Table 9 shows the number of samples used as input for each model. These tables are reported in Table 10 and Table 11 for the models built with the 3 first flights data only.

Table 8: Summary of model accuracies, represented by the Pseudo R-squared value (%), based on the OOB predictions. S = Safari, C = Chevignon.

		RGB + Multi	RGB only	Multi only
Total plant dry matter model	S + C	46,34	41,83	39,62
	S	44,63	24,76	35,72
	C	47,58	45,94	29,32
Ear dry matter model	S + C	80,76	70,09	79,56
	S	68,9	70,9	63,08
	C	85,67	70,89	83,13
LAI model	S + C	87,14	79,63	84,88
	S	85,54	74,08	82,42
	C	86,22	81	84,98

Table 9: Number of couples of observations used in each model. S = Safari, C = Chevignon.

		RGB + Multi	RGB only	Multi only
Total plant dry matter model	S + C	79	77	80
	S	40	40	40
	C	37	37	40
Ear dry matter model	S + C	68	68	71
	S	31	31	31
	C	37	37	40
LAI model	S + C	79	77	80
	S	40	40	40
	C	37	37	40

By analyzing Table 8, we can identify that the total plant dry matter models do not hand satisfactory results (the best model, for the Safari cultivar and using both RGB and multispectral data, barely reaches a Pseudo R-squared value of 50%). These models can't be used as prediction tools. The results of the two other models, for ear dry matter and LAI prediction, are encouraging.

The ear dry matter model's accuracies are comprised between 63% and 86%. While the models with multispectral data only and with RGB + multispectral data achieve similar accuracies, with a maximum of 3% difference, the RGB only models tend to imply a loss of accuracy (except with the Safari cultivar). In fact, this loss reaches 9% for the model with both cultivars, while there is an 15% drop for the Chevignon model. On the other hand, for the Safari cultivar, the RGB only model is the most accurate one, being 2% higher than the RGB + multispectral model, and 8% higher than the multispectral only model. Comparing model accuracies between cultivars yields interesting results. The Chevignon cultivar seems to achieve better results than Safari, but only through the addition of multispectral data. In fact, the accuracies of models with RGB data only are almost identical between cultivars, and there is a loss of 17% and 20% respectively for the RGB + multispectral and multispectral only models between Chevignon and Safari. Finally, the combination of both cultivars yields a good accuracy of 80%, with a slight decrease in respect of the Chevignon cultivar only model (-5%), but with a greater increase compared to the Safari only model (+12%).

The LAI model's accuracies start at 74% to reach up to 87%, which is an evidence that this model works well. The same conclusions as the ear dry matter model can be drawn concerning the decrease in accuracy due to the exclusion of the multispectral data, with a 5% to 8% accuracy loss. Meanwhile, the inclusion of the RGB data in the model with both cultivars achieves a very slight increase of 1%. The accuracy loss from the Chevignon to the Safari model that was raised for the ear dry matter model is also observable, but at a lower extent (2% to 7%). The combination of both cultivars in the same LAI model does not imply a loss in accuracy, the latter being comparable to the Chevignon only models.

Table 9 shows the number of samples differ between multispectral only and RGB only models. After exploring the datasets, the GI vegetation index means has been detected to have "NA" values for 3 samples in the Chevignon cultivar, which were removed before running the model. One of these samples had another missing value for the IKAW VI.

Table 10: Summary of model accuracies represented by the Pseudo R-squared value (%), based on the OOB predictions (without last flight data).

	RGB + Multi	RGB only	Multi only
Total plant dry matter model	17,29	7,28	20,7
Ear dry matter model	54,27	10,86	59,81
LAI model	65,9	51,81	63,61

Table 11: Number of couples of observations used in each model (without last flight data).

	RGB + Multi	RGB only	Multi only
Total plant dry matter model	59	57	60
Ear dry matter model	48	48	51
LAI model	59	57	60

The removal of the last flight data implies a great reduction in all models, as seen in Table 10. The LAI model could still be used for prediction purposes, with a Pseudo R-squared value of 66%, but the ear dry matter model is no longer accurate enough. It is interesting to note that with the 3 first flight data too, there is almost no reduction in the accuracy of the models with the Multispectral data only models compared to the models using RGB and multispectral data combined.

c. Variable importance

Variable importance plots (percentage of increase in mean square errors, and mean decrease in node impurity) are reported in Figure 11 to Figure 16, for the selected LAI and ear dry matter models with all data. The 15 most important variables are shown from decreasing order of importance. Table 12 presents the variables that were selected by the VSURF algorithm, which were used for model predictions.

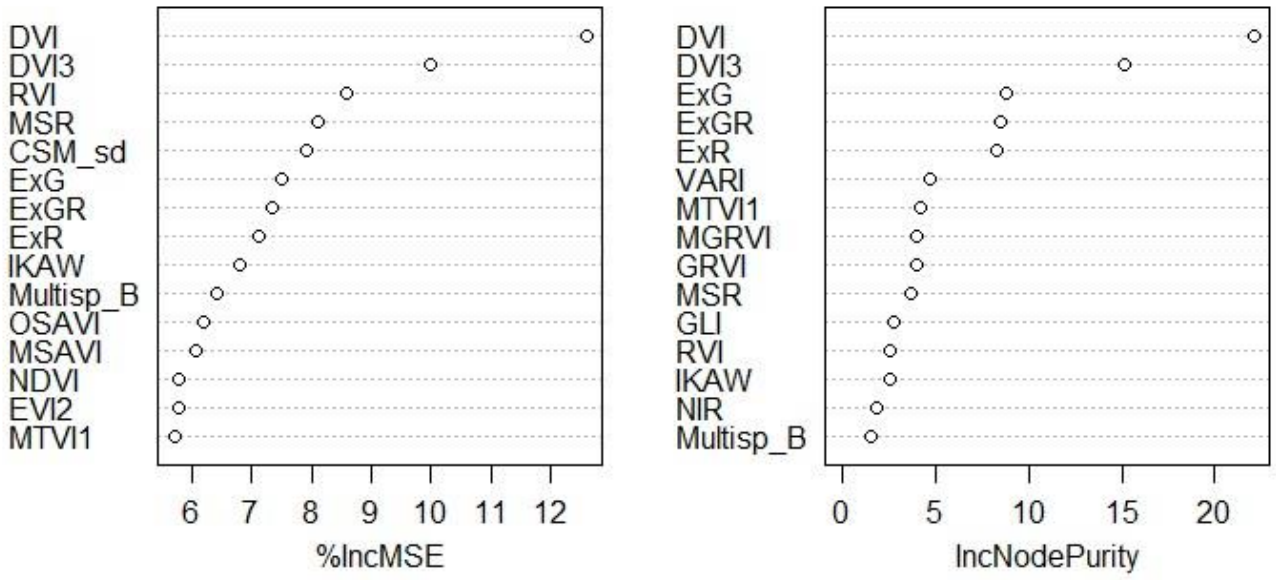


Figure 11: Variable importance plots of the LAI model built on all variables (%IncMSE = Percentage of increase in mean square error, IncNodePurity = mean decrease in node impurity).

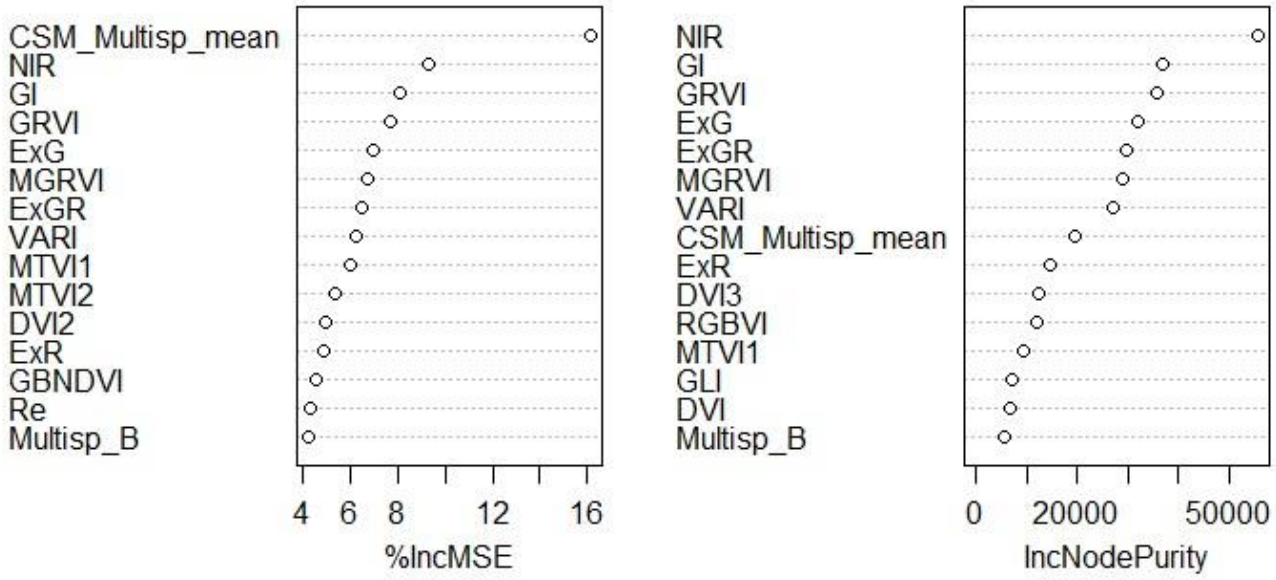


Figure 12: Variable importance plots of the ear dry matter model built on all variables (%IncMSE = Percentage of increase in mean square error, IncNodePurity = mean decrease in node impurity).

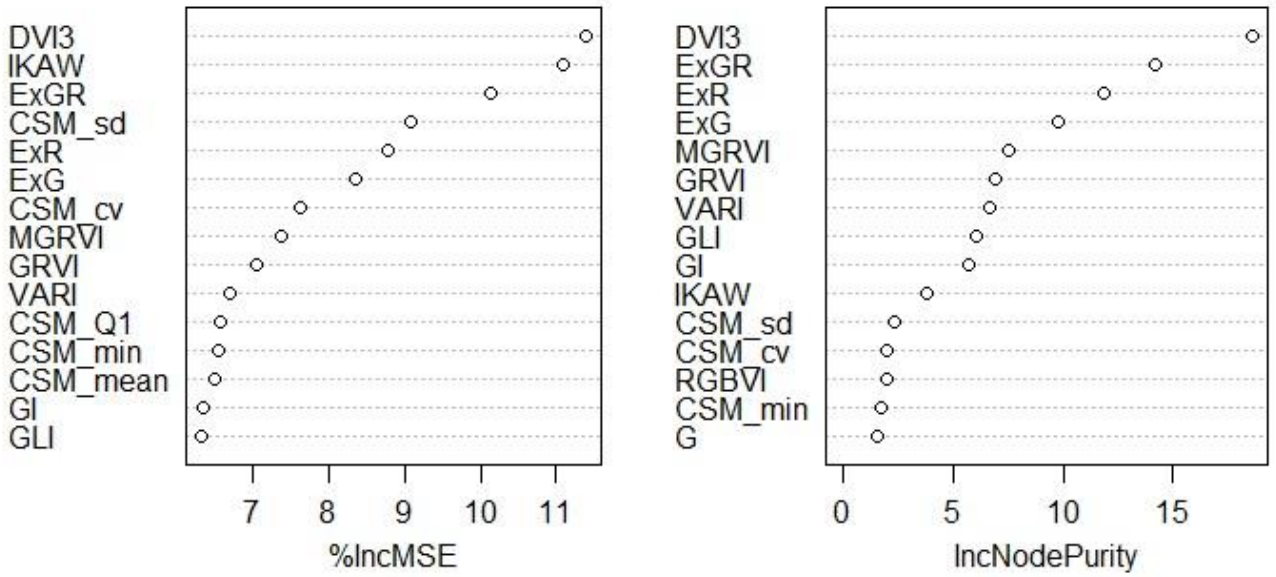


Figure 13: Variable importance plots of the LAI model built on RGB variables (%IncMSE = Percentage of increase in mean square error, IncNodePurity = mean decrease in node impurity).

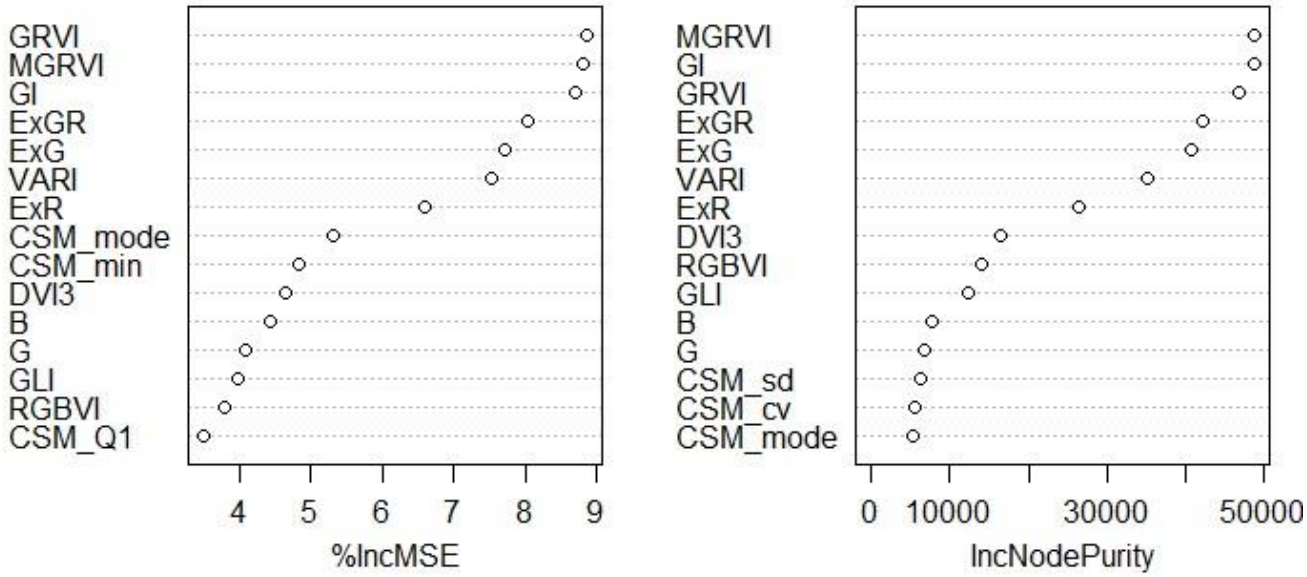


Figure 14: Variable importance plots of the ear dry matter model built on RGB variables (%IncMSE = Percentage of increase in mean square error, IncNodePurity = mean decrease in node impurity).

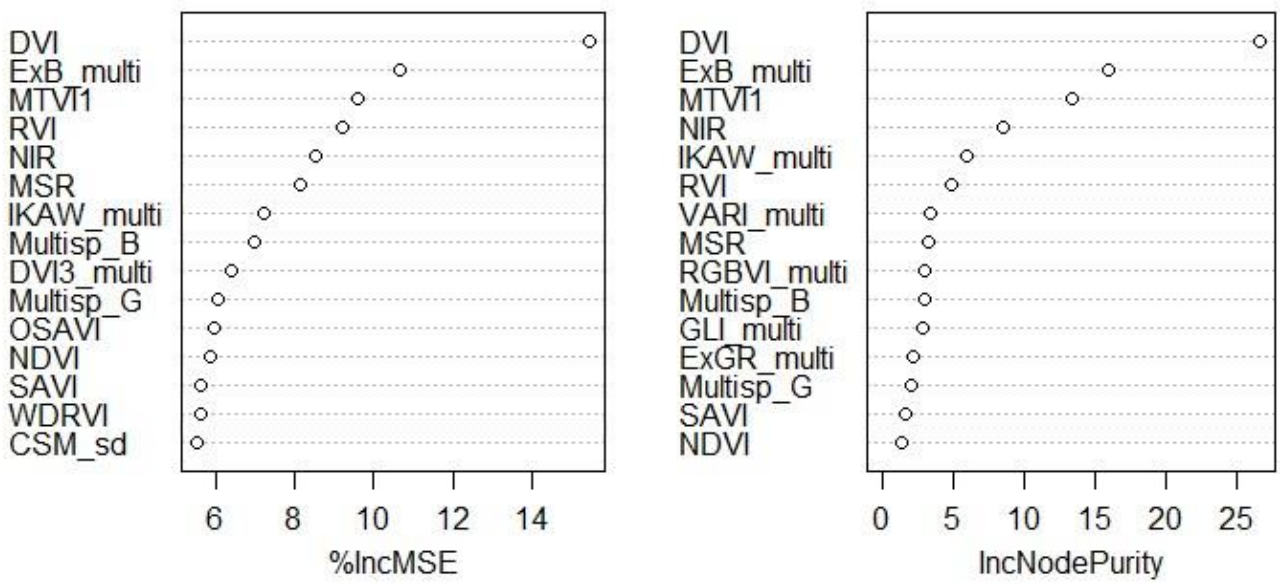


Figure 15: Variable importance plots of the LAI model built on multispectral variables (%IncMSE = Percentage of increase in mean square error, IncNodePurity = mean decrease in node impurity).

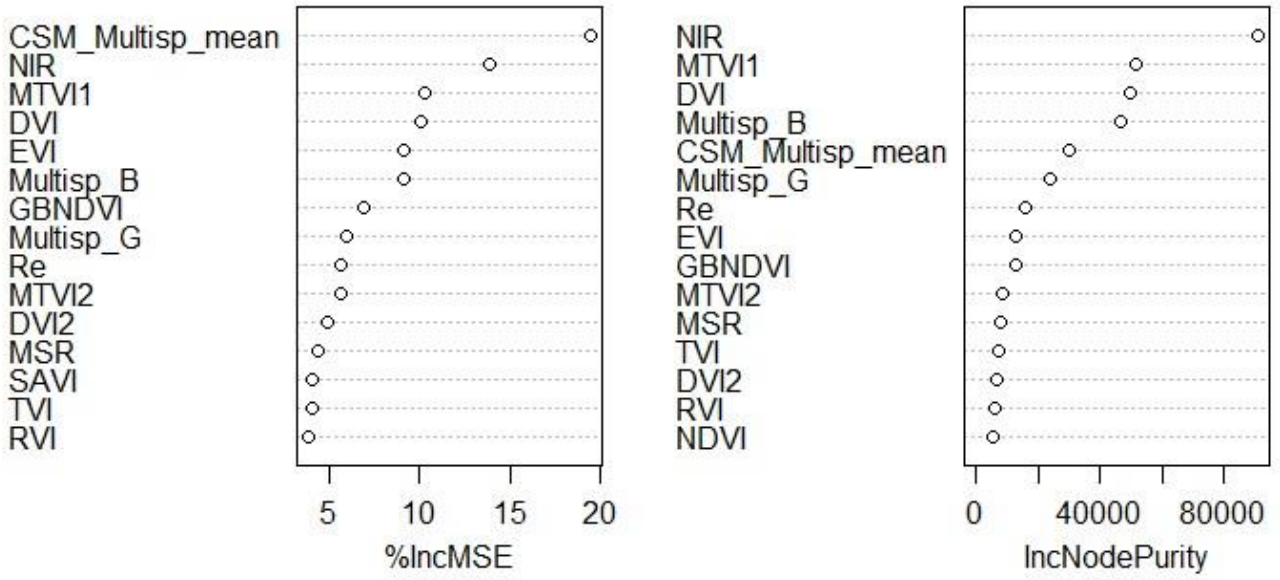


Figure 16: Variable importance plots of the ear dry matter model built on multispectral variables (%IncMSE = Percentage of increase in mean square error, IncNodePurity = mean decrease in node impurity).

Table 12: Selected variables by the VSURF algorithm with the different datasets for both models, by decreasing order of importance. See Table 7 for VI acronyms.

	RGB + Multi	RGB only	Multi only
Ear dry matter model	NIR, GRVI, ExGR, Multispectral CSM mean, MTVI1, Multisp_B, GBNDVI	MGRVI, ExGR	NIR, MTVI1, Multispectral Blue bland, Multispectral CSM mean, EVI, GBNDVI
LAI model	DVI, DVI3, ExGR, VARI, MTVI1, RVI, MSR, SAVI	DVI3, ExR, ExGR, GI, VARI, CSM_cv	DVI, ExB, MTVI1, IKAW, MSR, RVI

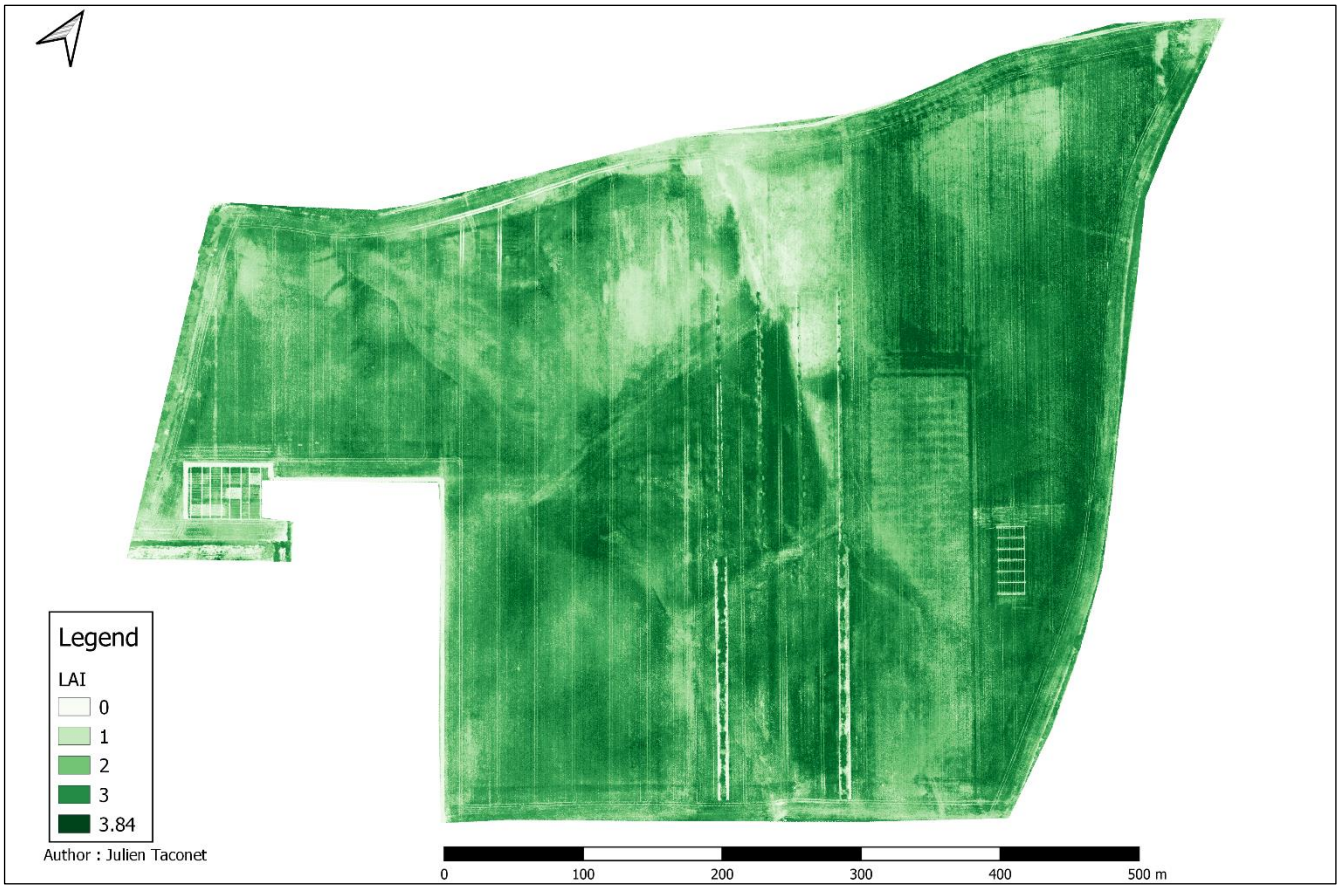
Figure 11 indicates the LAI model is mostly based on two VIs: the DVI (uses NIR and Re bands) and the DVI3 (uses green and red bands). RVI, MSR (derived from the multispectral camera) and excess VIs (ExG, ExR, ExGR) derived from RGB variables are the next most important variables, but at a lower importance scale. Standard deviation of Crop Surface Model appears as an important variable in the %Increase in MSE graph, while it does not appear in the IncNodePurity graph: it is an indicator that this variable wrongly interpreted as important. In the end, the VSURF algorithm eliminated the ExG and ExR variables from its selection since they are part of the ExGR variable, and added the MTVI1, SAVI and RVI indexes, three multispectral-derived VIs and the VARI RGB VI. The VSURF algorithm also eliminated the CSM standard deviation. Figure 13, showing the variable importance for RGB data only, confirms the important RGB variables for LAI modeling are DVI3 and the excess VIs. The CSM standard deviation still appears as an important variable in Figure 13, with the addition of the CSM coefficient of variation, which we can see in Table 12 has been selected by the VSURF algorithm (as last important variable).

Both importance graphs in Figure 12 indicate the NIR band and the multispectral CSM are the most important multispectral-derived variables for the ear dry matter model. The RGB-derived important variables for this model are GI, GRVI, ExG, ExGR, VARI and MGRVI. These appear as the most important VIs in the RGB-only models too, Figure 14, while the VSURF algorithm only selected MRGVI and ExGR in the RGB + multispectral model, as seen in Table 12. It is interesting to note that the multispectral CSM is considered as an important variable, while the RGB-derived CSM does not appear in the 15 most important variables. The multispectral blue band also comes out as important in the models with all variables, and has been selected by the VSURF algorithm, while it's RGB equivalent was not.

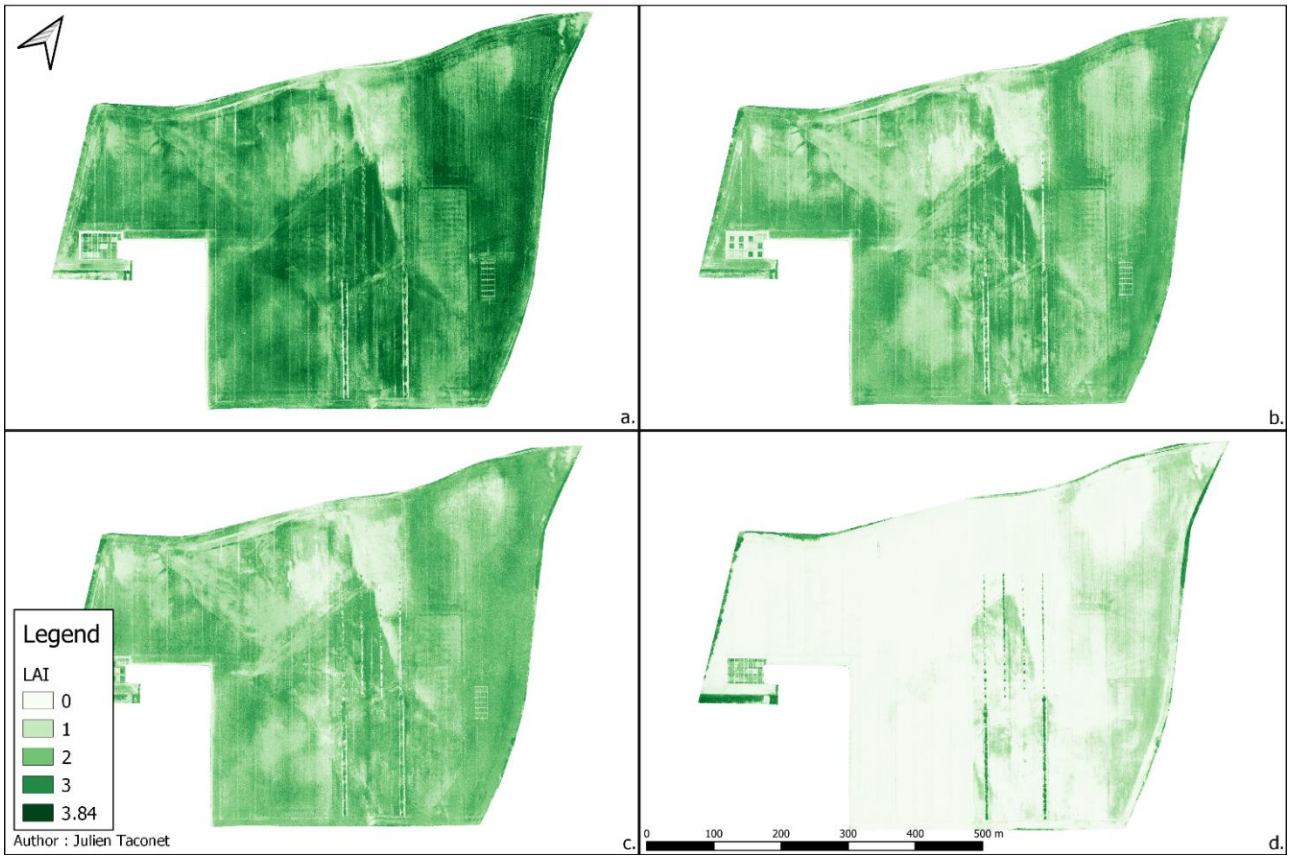
d. Using the created models to predict crop parameters on the whole parcel

Two models were run for prediction, one for the LAI and another for ear dry biomass. Both models use mixed-cultivar data, variables derived from both RGB and multispectral cameras, and all flight data.

Map 4 shows the predictions of the LAI model for the 25/05 flight. Map 5 presents the predicted LAI for each of the flights from the same model. The ear dry matter predictions are presented in Map 6 for the 25/05 flight, and Map 7 presents the results for all flights.



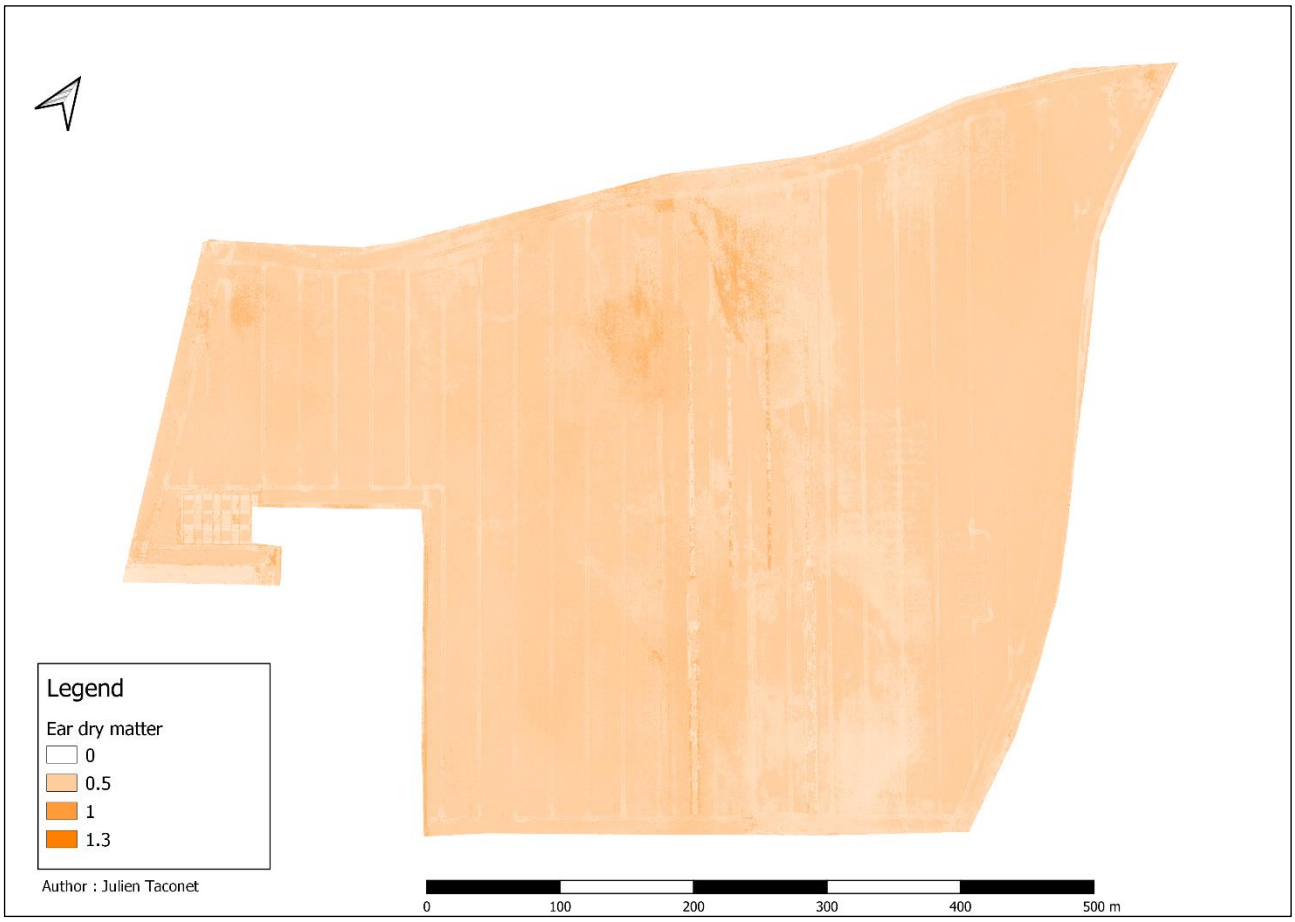
Map 4: LAI ($m^2.m^{-2}$) predicted from the model with all data and all variables on the 25/05.



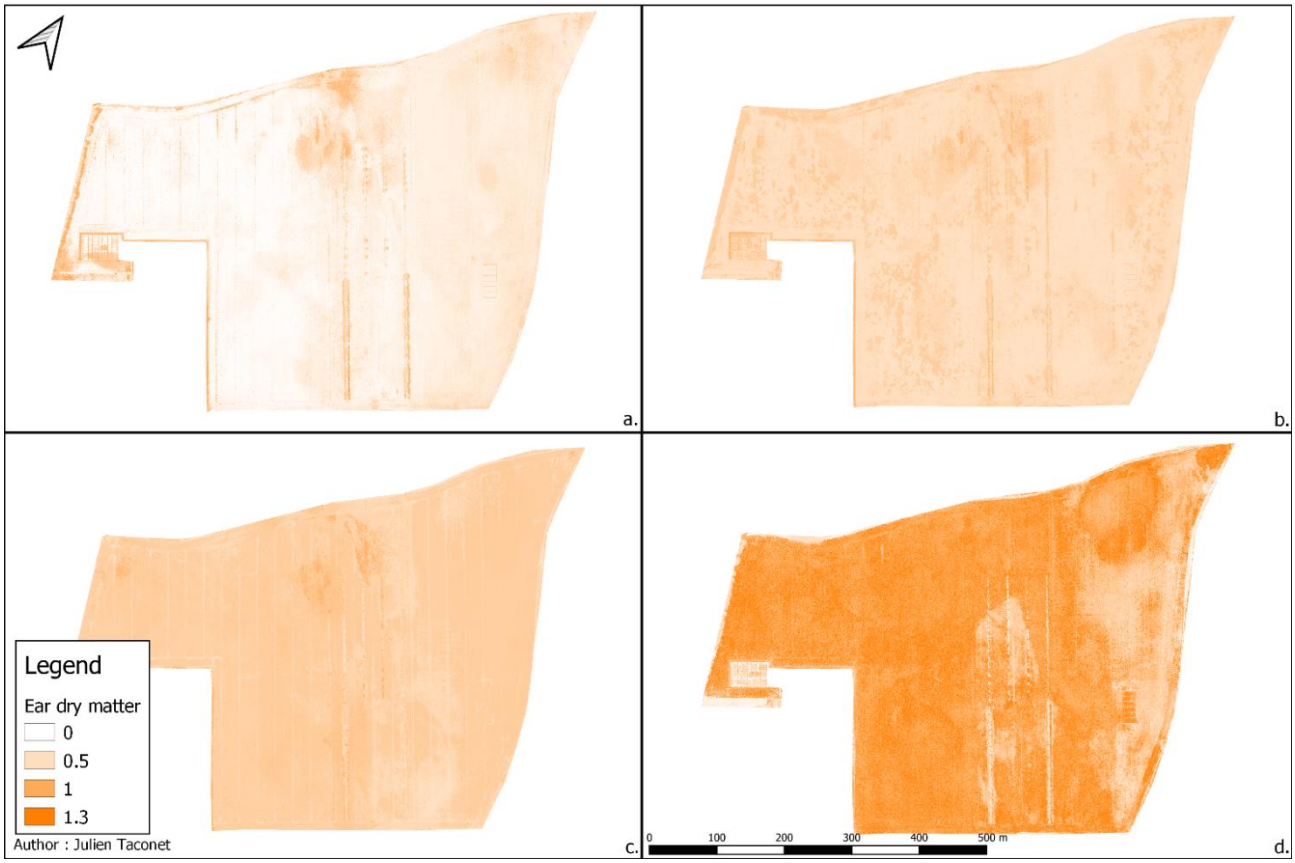
Map 5: Predicted LAI ($m^2.m^{-2}$) from the model with all data and all variables for all dates. a = 25/05, b = 09/06, c = 22/06, d = 13/07.

We can see in Map 4 the LAI varies from 0 to 3.84 on the parcel. A value of 0 corresponds to the road surrounding the tests at the far left of the picture or along the cemetery, and the maximum value corresponds to the hedges. Higher values (between 3.5 and 3.8) are found in the middle of the agroforestry hedges. The Safari cultivar seems to have higher LAI values than Chevignon. The former spring crop trials from 2019 are clearly visible at the right of the map, next to the agroforestry hedge, and the brome trials too, at the right of the spring crop trials. A pattern is clearly visible, running from the lower-right corner to the upper left corner of the map: this pattern follows the general altitude variations of the parcel, visible in the DEM Map 2. Along this pattern on the right, the lowest LAI values of the area can be found. Finally, at the upper right corner, two zones with lower LAI values are visible.

Map 5 seems to indicate LAI values steadily decrease between the three first flights. The LAI then drops down between 0 and 1 on the last flight, a few days before wheat arrives at maturity. This last flight emphasizes on the delay that the Safari cultivar (to the right) has on Chevignon, through the remaining greenness of its leaves, while they have all reached senescence and yellowed in the Chevignon cultivar. This delay can also be seen in the middle of the agroforestry hedges on the last map.



Map 6: Ear dry matter (kg.m²) predicted from the model with all data and all variables on the 22/06 flight.



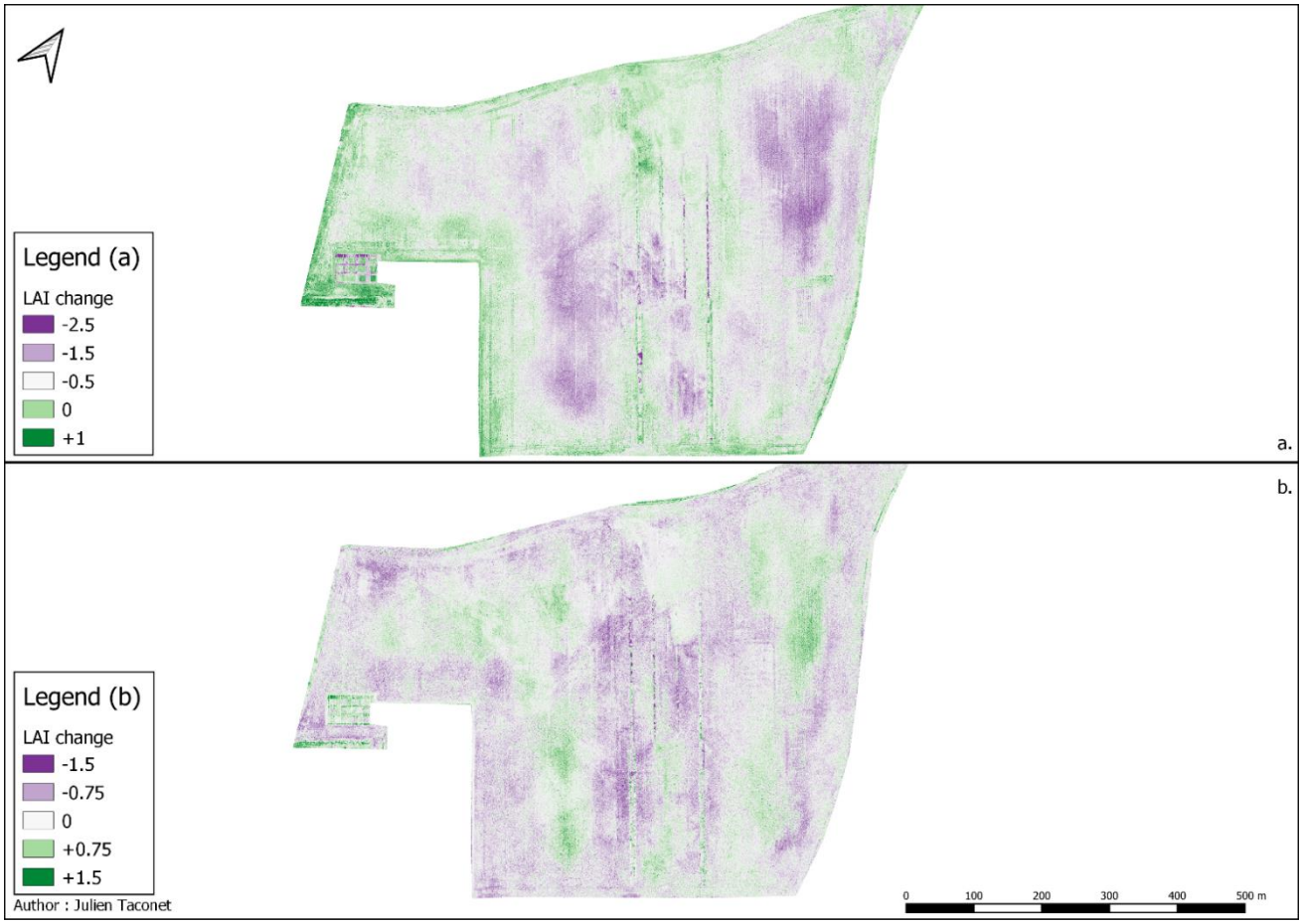
Map 7: Predicted Ear dry matter (kg.m^2) from the model with all variables and all data for all dates. a = 25/05, b = 09/06, c = 22/06, d = 13/07.

Map 6 shows the ear dry matter on the 22/06 seems almost homogeneously distributed along the whole field. Only few areas with lower values appear at the bottom-right and top-right of the agroforestry hedges, and a few higher values at the top and top-left of the latter.

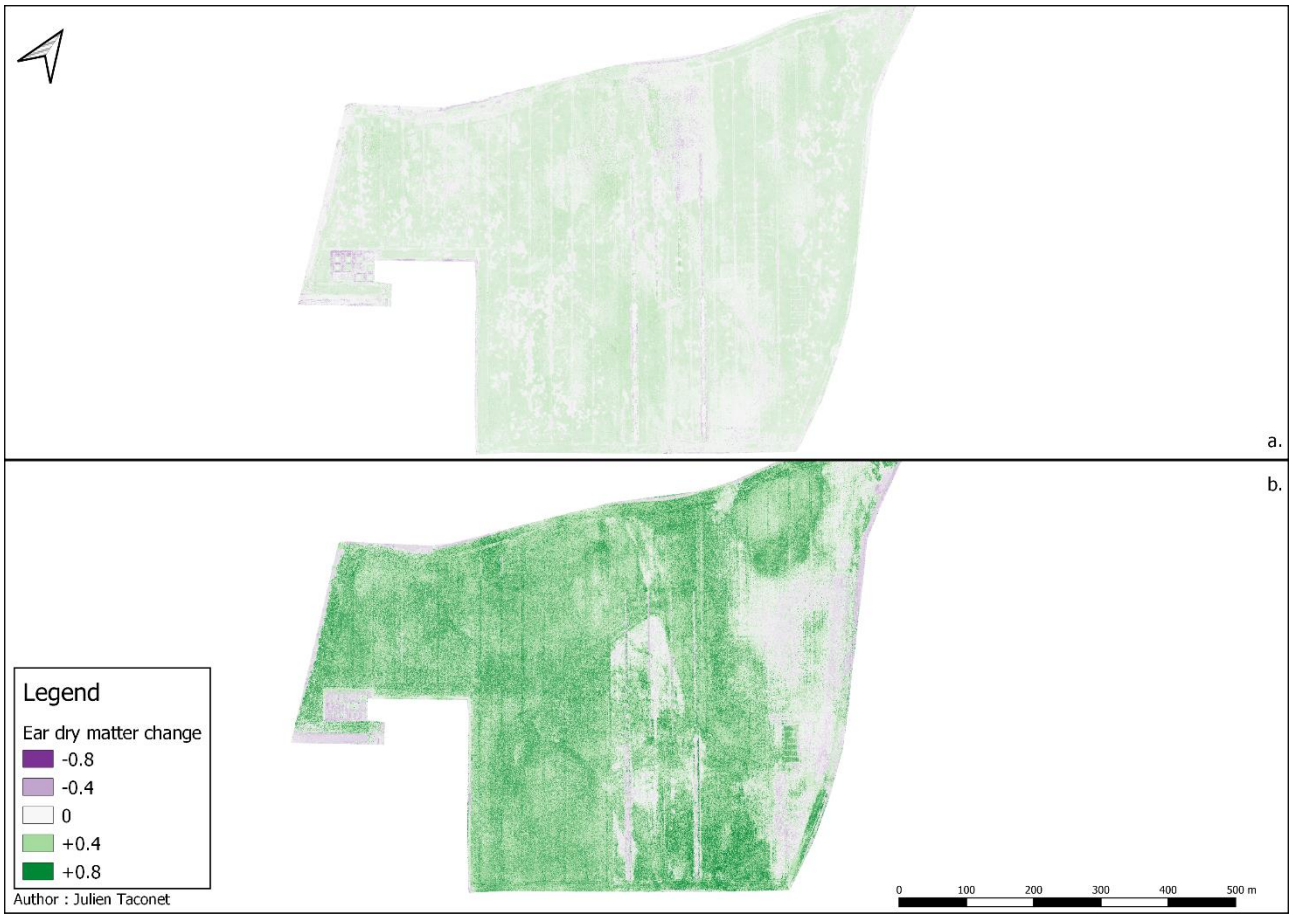
At first look on Map 7, panel a, the Safari (right side of the parcel) cultivar seems to have higher ear dry biomass values (between 0.4 and 0.6 kg/m^2) than the Chevignon cultivar (between 0.2 and 0.4 kg/m^2) on the 25/05. A few high values also appear at the top of the map, between the agroforestry hedges. The highest increase in ear dry matter in Map 7 is visible between the 22/06 (Zadok's stage 71 for Safari, 75 for Chevignon) and the 13/07 (Zadok's stage 87 for Safari, 91 for Chevignon). Between the 25/05 and the 09/06, the ear dry matter increases globally on the parcel, with higher increases in a spot pattern. On the 22/06, the ear dry matter seems homogeneous on the whole parcel, with both cultivars having reached their flowering stage. Areas with slightly lower ear dry matter values are visible in the middle-right of the map, between the agroforestry hedges. The delay that was observed on LAI for the Safari cultivar in two areas, at the right of the map and between the agroforestry hedges is also observable on the 13/07. This is a sign that the grain is not completely filled yet in these areas.

e. Crop parameter temporal changes through flights based on model predictions

In order to monitor the evolution of the studied crop parameters, two change maps sets were created. The first set (Map 8, Map 9) represents the LAI and ear dry matter evolution with a 2-week interval, between the 25/05 and the 09/06, and between the 09/06 and the 22/06. The second set (Map 10) represents the evolution of these parameters with a 1-month interval, between the first flight and the third flight.



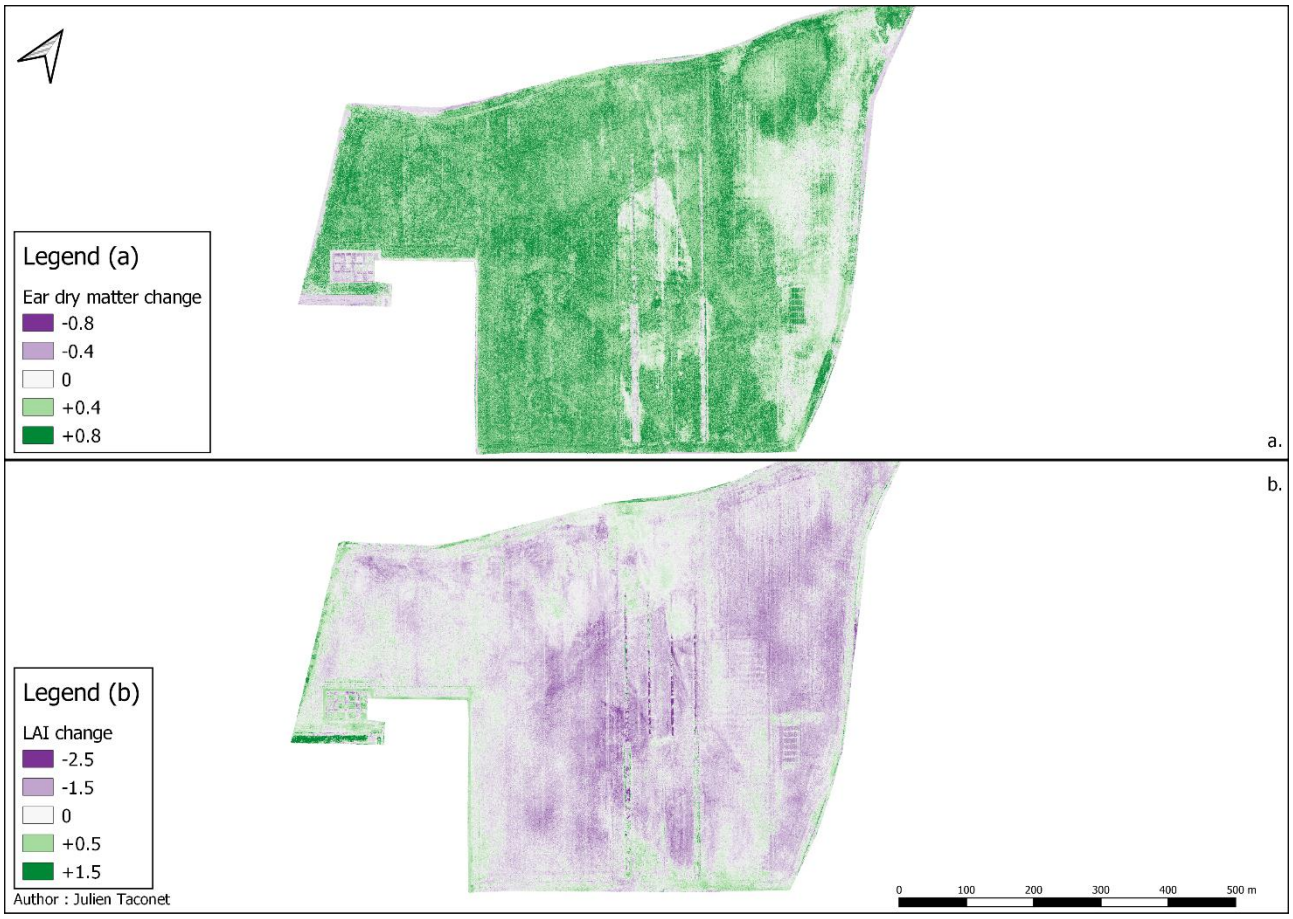
Map 8: LAI ($m^2.m^{-2}$) variation, (a) between the 25/05 and the 09/06 and (b) between the 09/06 and the 22/06.



Map 9: Ear dry matter variation (kg.m²), between (a) the 09/06 and the 22/06 and (b) between the 22/06 and the 13/07.

Map 8 and Map 9 show the changes in crop parameters between the three first flights. We can see on Map 8 a global tendency for predicted LAI values to decrease through the flights. It does not evolve homogeneously on the whole study area: between the 25/05 and the 09/06, LAI in the inside areas seem to decrease at a high extent, while the borders increase a little. A small zone at the top-middle, between the agroforestry hedges, increases too. Between the 09/06 and the 22/06, the previous patterns seems to reverse, with a high LAI decrease on the borders of the areas and within the agroforestry hedges, and a much lighter increase on the inside.

The predicted ear dry matter change values are visible in Map 9. Globally, the ear dry matter increases at a much higher extent between the 22/06 and the 13/07 than between 09/06 and the 22/06. Panel b. shows that the ear dry matter increases at a much higher extent for the Chevignon cultivar than for Safari, where values stagnate between the agroforestry hedges and on the right of the field.



Map 10: Ear dry matter change (kg.m^2) (a) between the 13/07 and the 22/06 and LAI change (m^2.m^{-2}) (b) between the 22/06 and the 25/05.

Map 10 shows the ear dry matter changes between the second flight (09/06) and the last flight (13/07), and the LAI changes between the first flight (25/05) and the third flight (22/06). As previously noted, ear dry matter values on the Safari cultivar increase at a lower extent than the Chevignon cultivar.

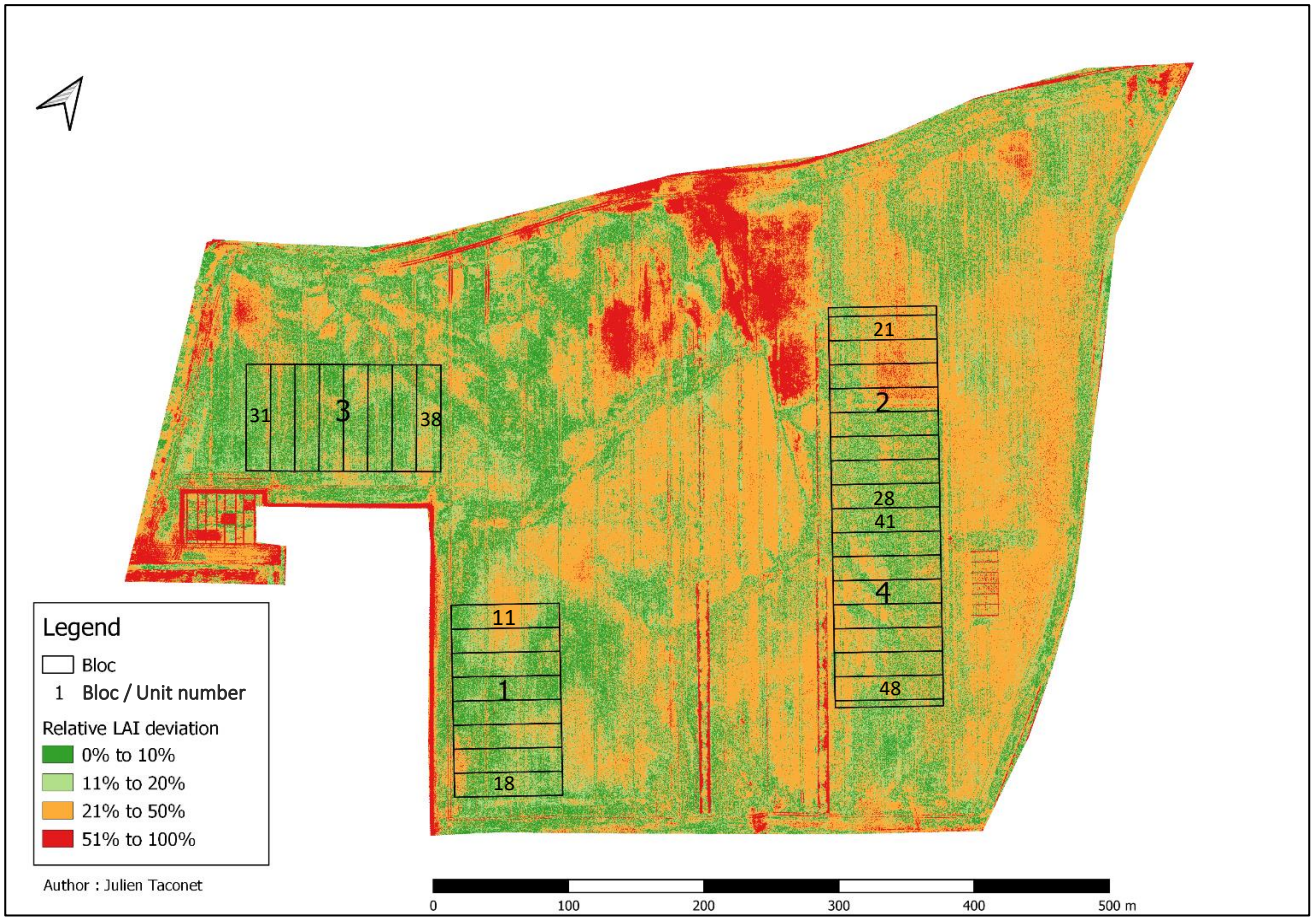
On the other hand, LAI decreases on most of the parcel, with the highest decrease on the Safari cultivar, and between and on the left of the agroforestry hedges. The left part of the map in the Chevignon cultivar does not evolve much, only slight decreases are noted, same as at the top-right and top of the agroforestry hedges.

It is interesting to note that, in the Safari cultivar, the LAI values are more variable outside the clearly-visible spring crop trials. This phenomenon is not as clear on the ear dry matter change map.

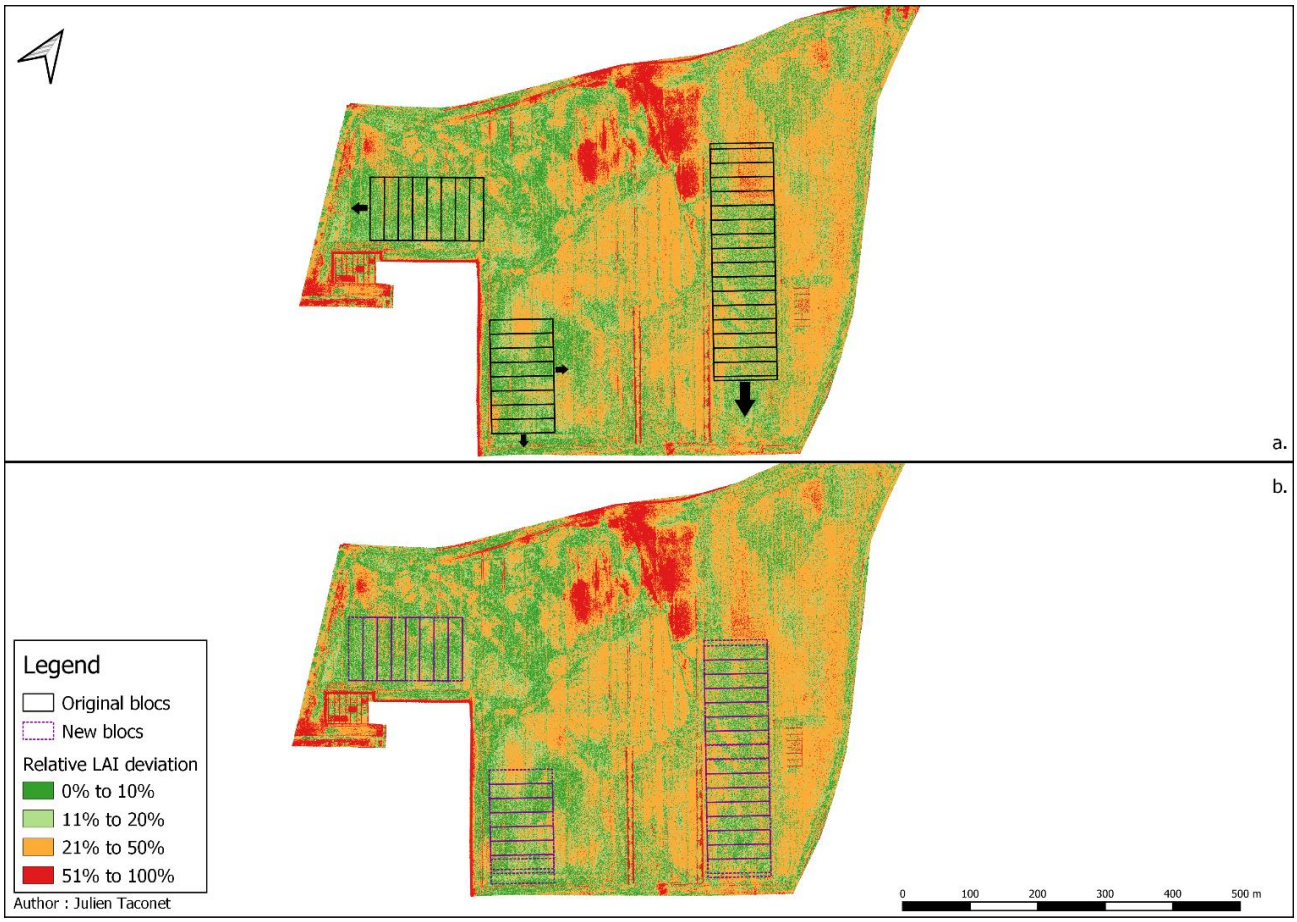
f. Analysis of Variance results

Map 11 shows the relative LAI deviation map of the 25/05 flight with the experimental blocs, and Map 12 shows the blocs original locations, their number and the proposition of displacement in regards of the relative LAI deviations.

Table 13 summarizes the F-values of the ANOVA tests of means run between and within the experimental blocs, for their original location and the new location.



Map 11: Relative LAI deviation on the 25/05.



Map 12: (a) Original experimental blocs location and (b) modification proposition of the experimental bloc's location in regards of the 25/05 LAI relative deviation map.

Map 11 indicates bloc 3's unit number 38 covers an area with highly variable LAI values compared to the LAI mean on the whole parcel. Since there are lower values of relative variability on the left, it has been proposed to be moved to the left, as shown in Map 12. Table 13 confirms that the new position reduces the inter-unit variability, with a noticeable reduction in SSE_M, and Figure 20 shows the new unit 38 LAI values closer to the other unit means than the former 38.

Bloc number 1 covers an extremely-variable, rather small area at its bottom-left corner on units 18, 17 and 16, with a high number of red pixels. In order to avoid it, it was proposed to displace the bloc to the right, having as side-effect to cover slightly more orange pixels with medium deviation values. Additionally, there is some space for the bloc to be moved downwards, allowing the medium-valued LAI relative deviations covering unit 11 to be replaced by lower values at the bottom. Table 13 shows this displacement has a high impact on this bloc's inter-unit variability, roughly dividing the SSE_M by 4.

Bloc number 2 also covers a highly variable LAI area at its top, on units 21 to 24. This is why it has been proposed to be moved down towards lower variable LAI values, as seen in Map 12, and Table 13 shows a drastic reduction in SSE_M as consequence. This automatically displaces bloc 4, since it is joint to bloc 2, but the values in SSE, SSE_M and F-score are left almost unchanged from the displacement. This proposition has a huge impact on bloc 2's inter-unit variability, while leaving bloc 4's almost unchanged.

Finally, we can see the overall effect of the displacements on the Inter-Bloc ANOVA results in Table 13. A great reduction in SSE_M is noted, combined with a lower reduction in SS_E. This means the inter-bloc variability has greatly reduced, accompanied by a slighter reduction in within-bloc variability.

Table 13: Comparison of ANOVA outputs between the original blocs and their new location. SSE_M = sum of squares for the mean, SS_E = Error sum of squares.

Bloc number	Original SSE_M	New SSE_M	Original SS_E	New SS_E	Original F-value	New F-value
Bloc 1	225 797	48 065	2 858 483	2 542 729	171 898	41 135
Bloc 2	857 772	256 663	3 868 168	2 994 046	483 190	186 531
Bloc 3	431 216	299 626	2 730 326	2 287 855	343 692	284 997
Bloc 4	357 844	364 527	3 535 271	3 548 652	220 273	223 541
Inter-Bloc	987 898	226 466	14 489 879	12 313 582	1 350 276	373 548

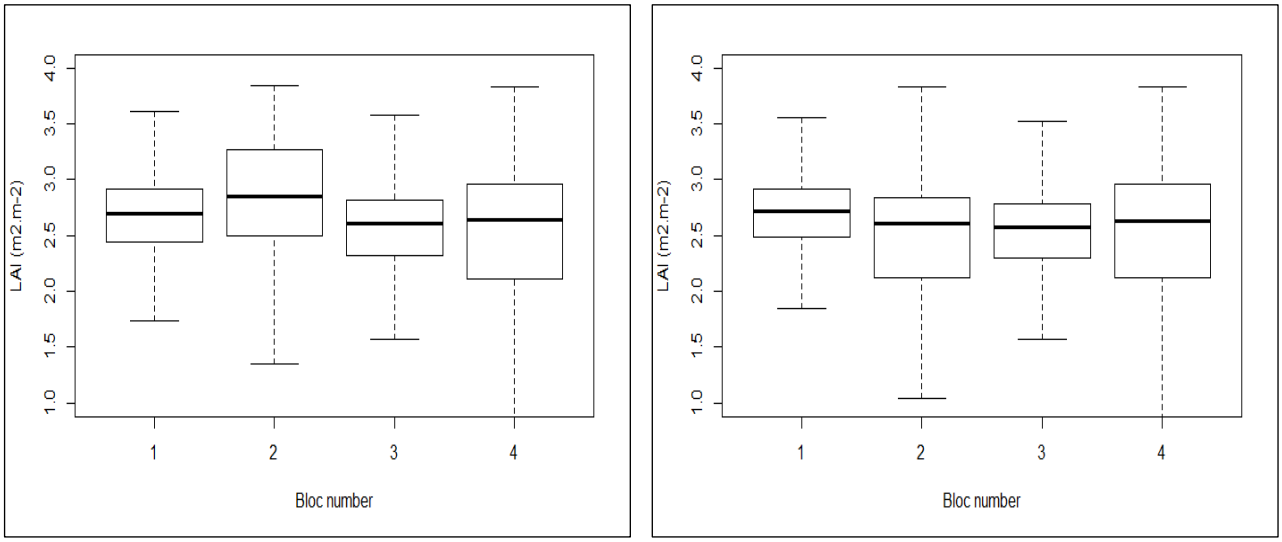


Figure 17: LAI boxplots within the original experimental blocs (left) and their new positions (right).

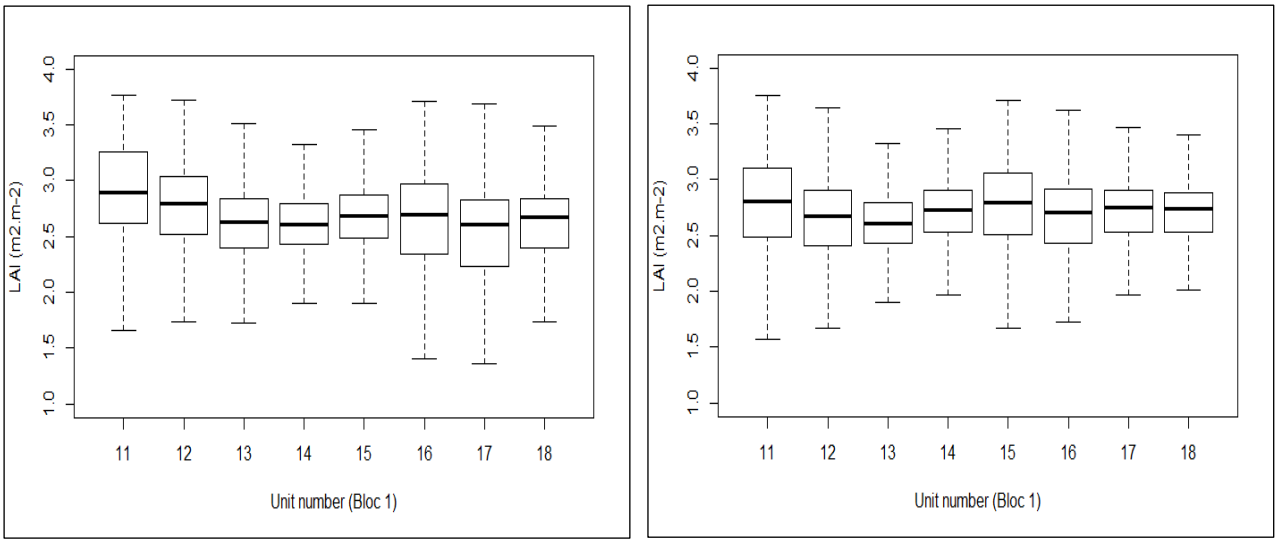


Figure 18: LAI boxplots within the units of original experimental bloc 1 (left) and its new position (right).

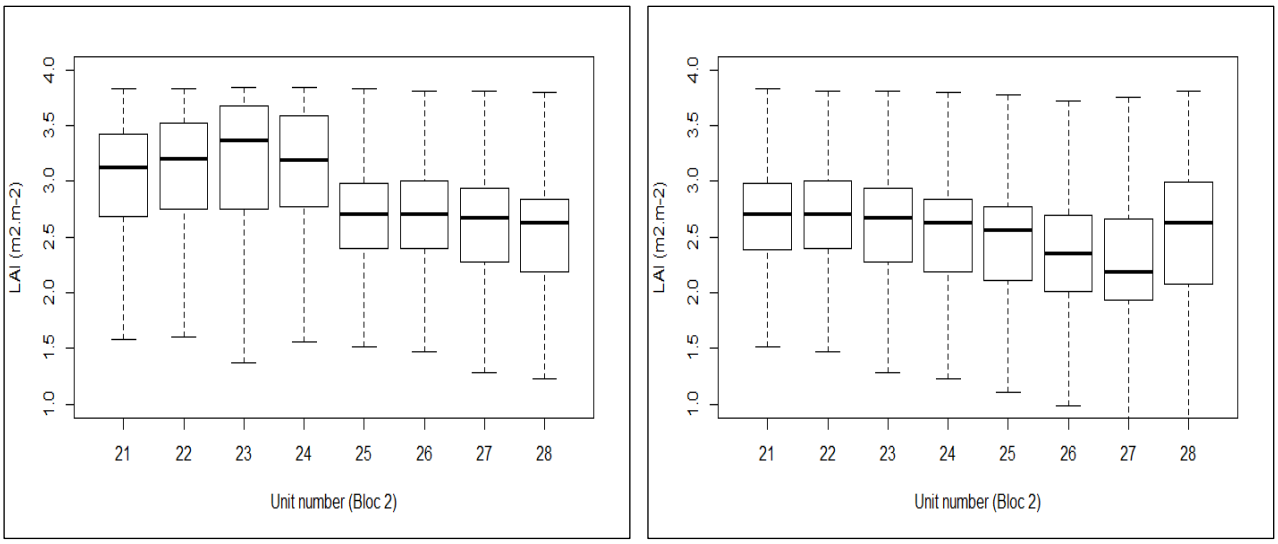


Figure 19: LAI boxplots within the units of original experimental bloc 2 (left) and its new position (right).

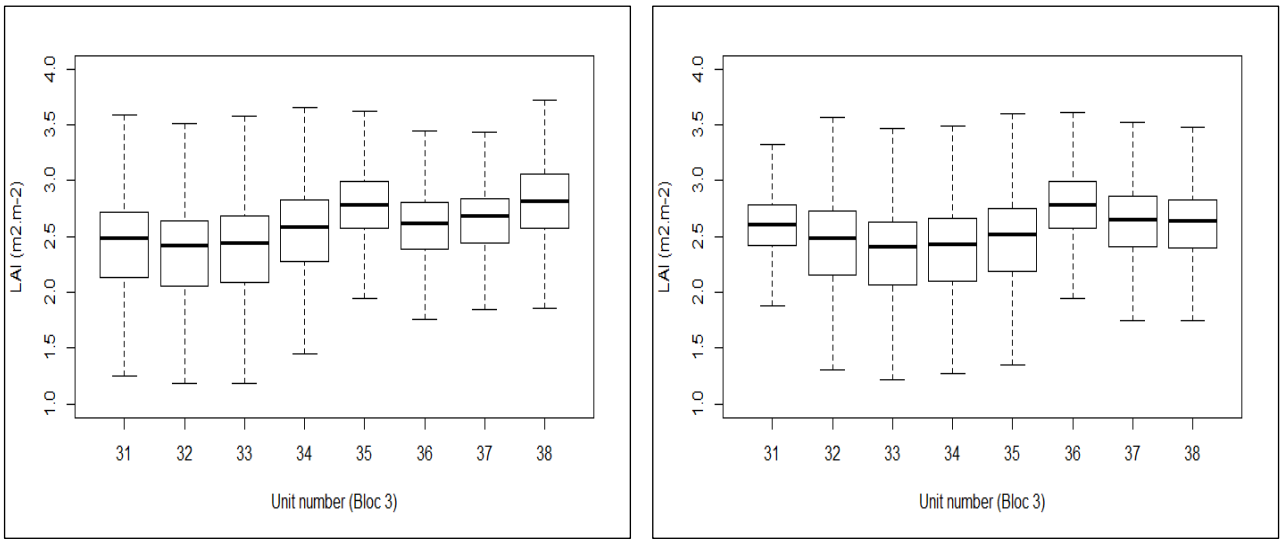


Figure 20: LAI boxplots within the units of original experimental bloc 3 (left) and its new position (right).

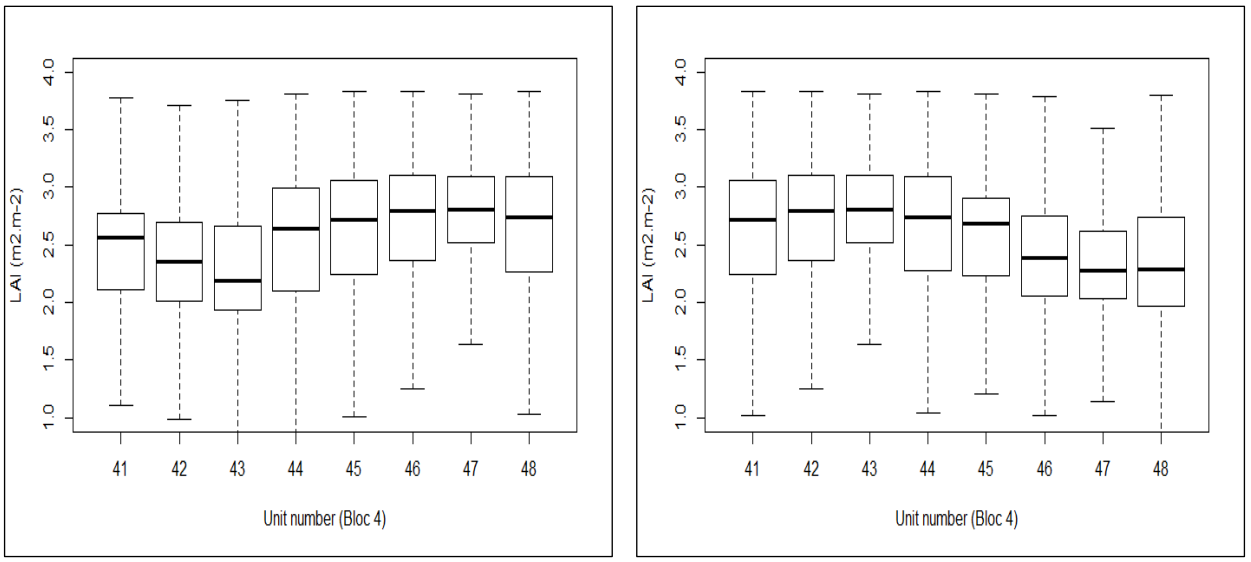


Figure 21: LAI boxplots within the units of original experimental bloc 4 (left) and its new position (right).

6. Discussion

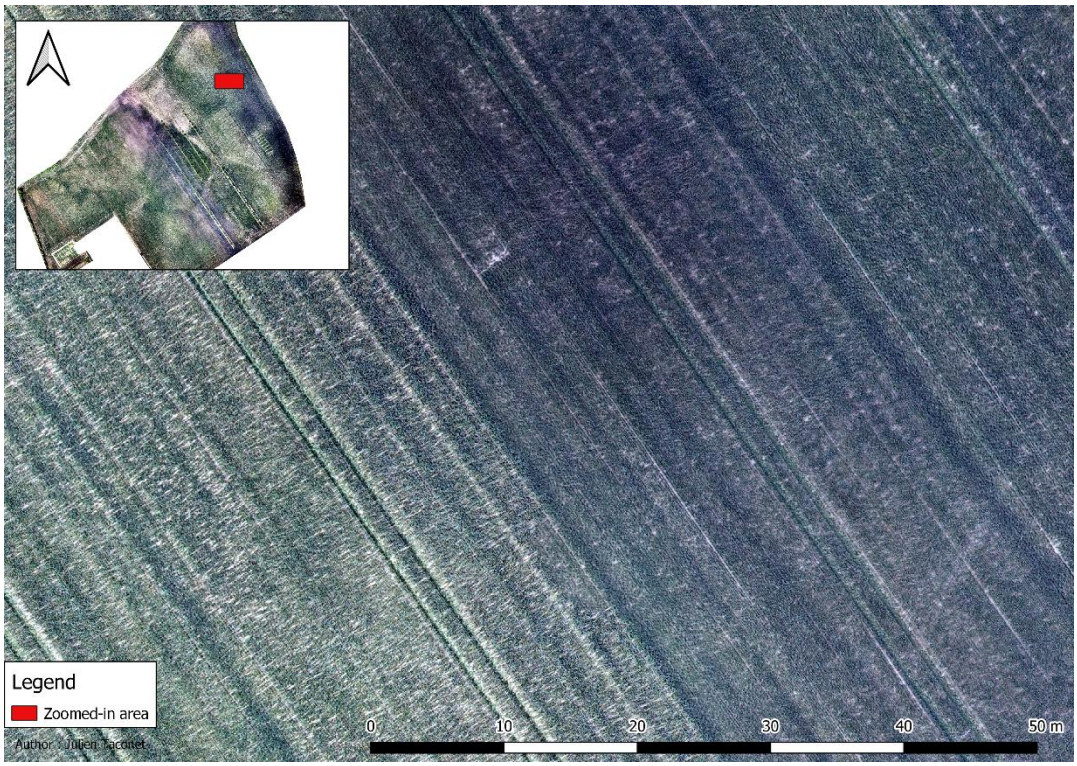
a. Model accuracies and error sources

i. UAV data error sources

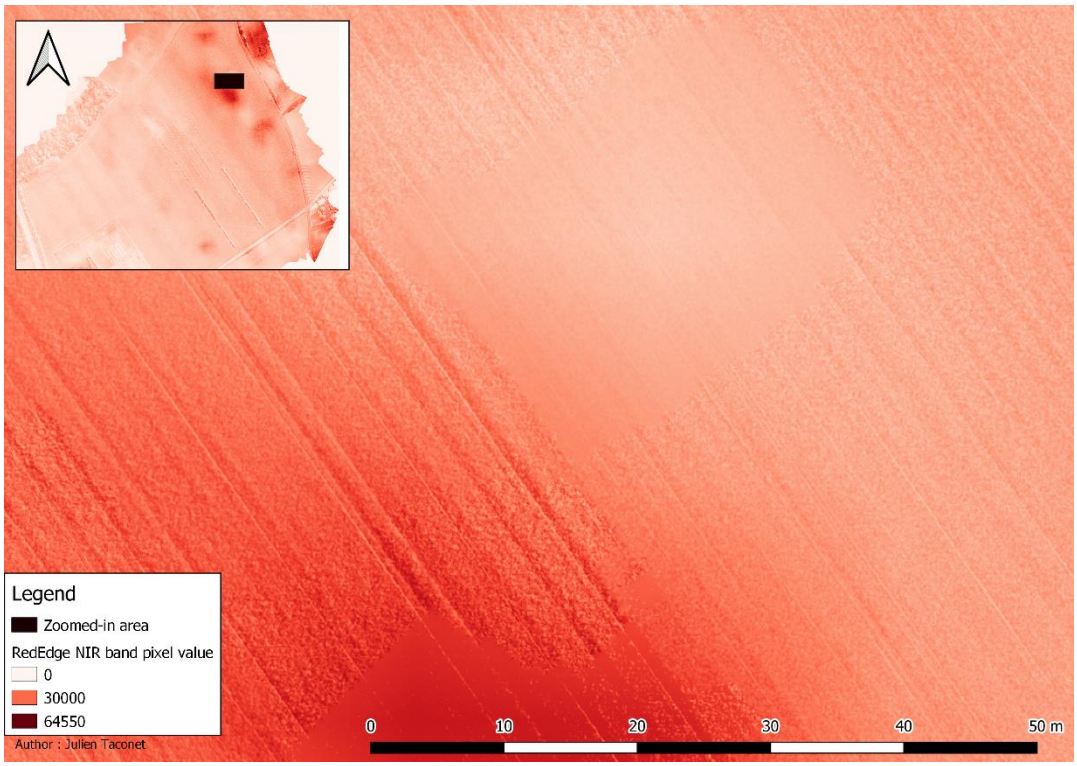
As stated in section 2.a, UAV technology has many advantages. It's flexibility to be equipped with sensors of all types and its low flight altitude allowing high spatial resolution outputs make it an ideal tool for crop monitoring. However, a successful flight with quality output data depends on the weather conditions during the flight. It's potential flexibility to be flown at the exact desired frequency, unlike satellite's fixed-frequency orbits, is highly tailored by weather conditions. Rain, wind, or extreme temperatures and humidity can impact the communication between the operator and the drone, ultimately causing its loss of control (Ranquist *et al*., 2016). On the other hand, the range of conditions where the UAV is able to be flown with no risks is high, but only specific conditions allow the output data to be suitable for crop parameter modeling.

Through visual interpretation of the output orthophotos, it has been noted that the flights run during changing illumination conditions (due to clouds obstructing the sun) are source of undesired noise. Map 13, Map 14 and Map 15 respectively show the impact on the RGB orthophoto, the Multispectral orthophoto and the LAI predictions of a cloud obstructing the sun on the 25/05 flight. Photo 1 shows the sky conditions during the flight at this date. An oblique separation is clearly visible in the middle of the 3 maps between two flight lines, translated by a darker (on the RGB orthophoto) or blurred (on the NIR band) area at the right. The fact that the multispectral image was also affected, despite the luminosity sensor, is explained by the cloud casting a shadow on the ground, but not the UAV. This ultimately impacts the LAI model predictions, as seen on Map 15. Samples extracted in those areas, as it is the case in this present work, can cause high biases in the output models. The Empirical Line Method used to transform pixel absolute values into reflectance values does not take these changes in account. The ELM method should also be carried out with care, by taking the pictures of the calibration panel with the same illumination conditions before and after the flight (with the sun obstructed or with direct sunlight, depending on the dominant conditions).

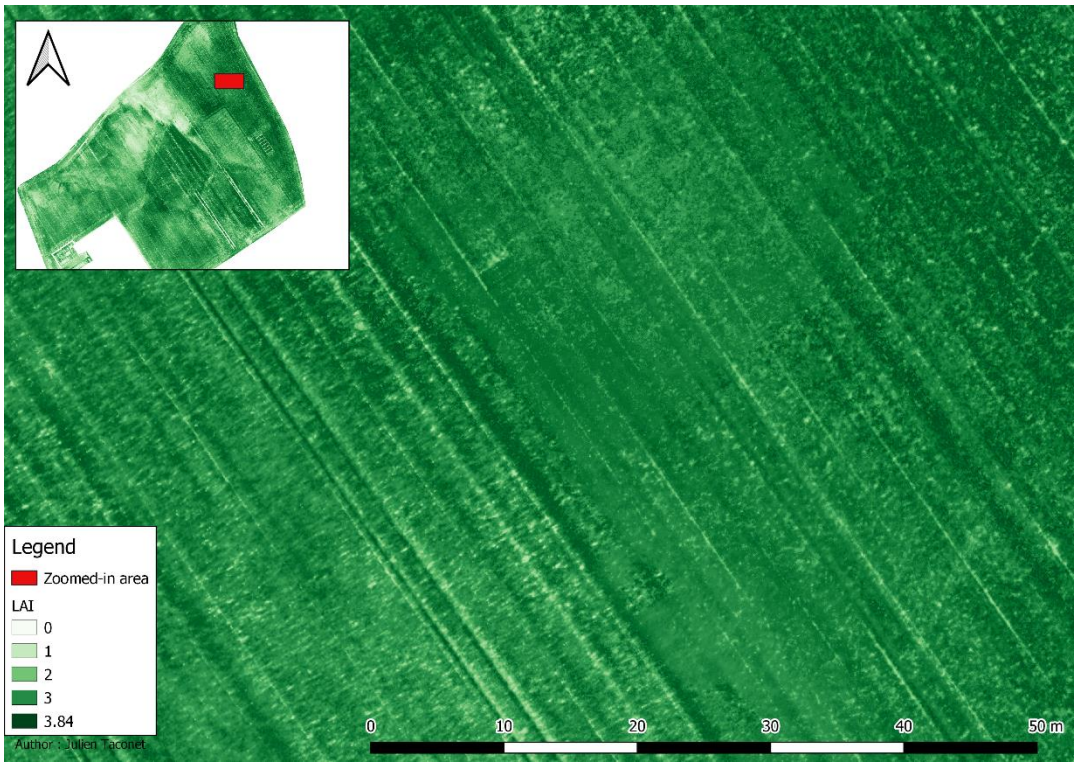
Photo 1: Sky conditions over the parcel during the 25/05 flight.



Map 13: Effect of sunlight obstruction by clouds on RGB orthophotos (25/05 flight).



Map 14: Effect of sunlight obstruction by clouds on the NIR band from the multispectral camera (25/05 flight).



Map 15: Effect of sunlight obstruction on the output LAI map (25/05 flight).

While changing illumination conditions are undesirable, direct sunlight can also be a source of error on reflectance and elevation UAV products (Ortega-terol *et al*., 2017), due to what is called the "sun glint effect" and the "hotspot effect". The sun glint effect (when the sun angle of incidence on a specular surface equals the reflection angle) was detected on the 25/05 flight (see Figure 25a.), and a first inspection of the RGB CSM within the sampled surfaces revealed the related errors: many CSM values were negative. The origin of this inaccuracy was detected in Agisoft Metashape as being a tie-point matching problem. After recalculating tie-points and extracting new products from Agisoft, the negative CSM values had passed positive, but some were still as low as 3cm. All samples were made where wheat reached at least 20cm, and a poor correlation of 0.42 was calculated between dry stem biomass and CSM means. The accuracy of the CSM is disappointing compared to the work of Holman *et al.* (2016) (see section 2.b.i). This can be explained mostly by the illumination conditions, and by the fact that the spatial resolution of outputs is roughly twice as higher than in their study.

Ortega-terol *et al.,* in 2017, proposed an SfM strategy to minimize the impact of sun glint or hotspot effects, which could be used in future flights. The main consequences of their strategy are the removal of the areas where sun problems are detected, which leads to an improvement of SfM photogrammetry cartographic products.

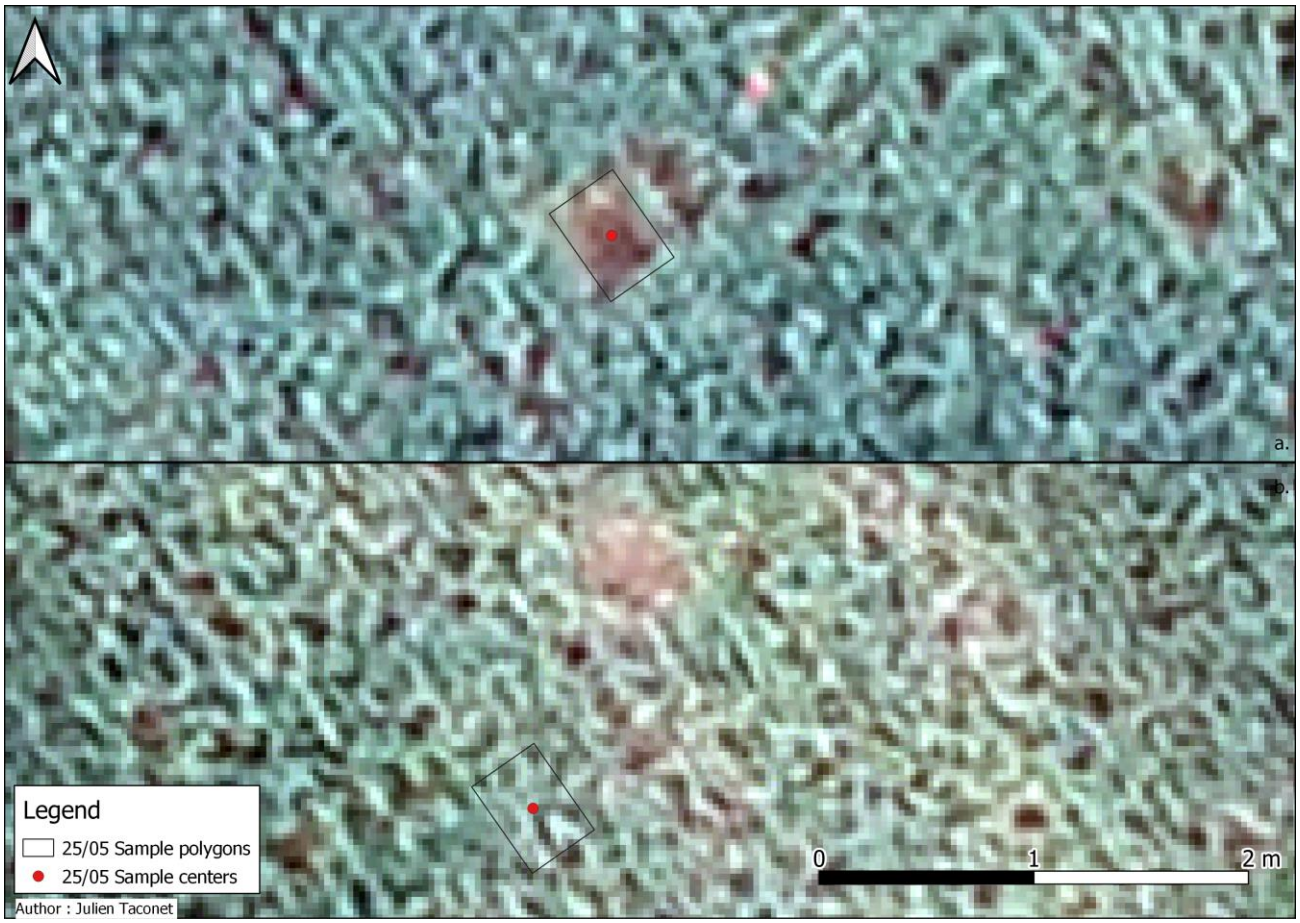
The previous considerations necessarily had negative consequences on the accuracy of the models, though they cannot be quantified. While weather conditions during the flights are a major source of error, SfM photogrammetry parameter choices can also affect the general quality of the outputs. It is also important to keep in mind that the output format of the flight pictures should be, if possible, set to RAW to be used in the SfM photogrammetry software, since the JPEG and TIFF formats used in this work can be a first source of unwanted variability as pointed out in section 4.b.ii.

ii. Ground-truth protocol error sources

The ground-truth data acquisition also brought its lot of uncertainties in the model.

Firstly, an explanation for the total dry biomass poor accuracy can be found in the sampling protocol. In fact, in order to use the LAI-leaf dry biomass relationship, the yellow leaves were removed from the pool of weighted leaves at every date. This implies a growing bias in the total dry biomass data through the flight dates, as more leaves reached senescence. This protocol error was detected before the last flight, and yellow and green leaves were weighted on the 13/07. This allowed to quantify the mean under-estimation of total dry biomass through the 3 first dates, from 5% to 10%.

A second accuracy degradation source comes from the positions of sample centers recorded with the RTK-GPS. Visual inspection of the orthophotos from the 09/06 to the 13/07 revealed that most points are accurately geo-referenced. In fact, the bare-ground, left after destructive sampling, is visible on the next flight. Some points, on the other hand, were recorded with less visible satellites due to vegetation disturbance of their signals (for example close to the agroforestry hedges) or to an unknown source (the 13/07 date was marked by a lower amount of satellites in reach of the GPS). An example of an accurately and an inaccurately geo-referenced point is shown in Map 16. In this map, we can see, part (a): the bare-ground is located right under the sample center record, and the polygon delimitates it perfectly, while in part (b): the bare-ground is located 1.5m North-East of the sample center record. This inaccuracy could be a source of model degradation starting at a few decimeters error in positioning, since the micro-variability on the parcel is high. A fix for this problem could reside in modifying the sample center point records manually through visual inspection of the orthophotos. This solution is not suitable for the last data records, except if another flight is programmed after the sampling campaign.



Map 16: Accurately (a) and inaccurately (b) geo-referenced sample centers on the 25/05 flight.

Another model source of inaccuracy in the sampling protocol resides in leaf scanning for LAI determination. As seen in Figure 22, the scans can show areas with light reflections over green leaves. These are due to the roughness of the leaf, which deforms the transparent cover sheet. As a consequence, green pixels are considered white by the image-recognition algorithm, thus leading to an under-estimation of LAI values. A possible solution could be to stack less leaves per scanned paper, or to use a tougher and stickier cover sheet.

Figure 22: Example of leaf scan for LAI determination.

An interesting point to reveal on the sampling protocol is the fact that only leaves were used for the LAI model. On the other hand, the VIs used for the LAI model are linked to the reflectance of the whole plant, mostly of the canopy. What is actually interpreted by the model is a Green Area Index (GAI): there is a remaining part of greenness that was not taken in account in the sampling protocol, which is the photosynthetic activity of the stem and of the ears (through the glumes and lemmas). These sources of greenness were not measured; the LAI models were not calibrated on the entirety of the sources of green reflectance. These ear and stem photosynthetic activities increase at and after flowering, and could partly explain LAI variations Map 8.

To conclude on ground-truth data error sources, it is important to remind that the LAI values used in the models are extracted from a linear relationship and are not direct measurements, and that the leaf dry biomasses were corrected by a factor. This relationship has been created to alleviate the ground-truth sampling work, and can have an impact on the LAI model. It could be positive if the relationship is accurate enough and its “smoothing effect” allows to compensate for random error sources, or it could lead to a systematic error if the relationship was built on biased data (as the LAI-scanning method could have brought).

It is also important to remind that the ear dry matter model was built on less data than the LAI model: on the 25/05, the Safari cultivar was still at stage Z47, and almost all sample’s ears and stems were weighted together due to the difficulty to separate them.

iii. Data processing error sources

The last set of error sources on models that will be discussed comes from the data processing choices in raster resampling and outlier detection.

Raster resampling was carried out on Qgis. This choice was made after noticing that one or both R functions resample() or writeraster(), from the raster package, randomly modified the multispectral camera NIR band in some specific areas over the parcel. The output datatype parameter default value is suspected to be at the origin of this phenomenon. Special care must be taken when using R base raster package, as it can lead to erroneous conclusions.

The bilinear interpolation resampling method inevitably induces error since it modifies pixel values of a given raster object, but this method is known to have the best trade-off between accuracy and computation time.

Finally, outlier detection through the outForest() function detected 4 outliers in the RGB band means, 6 in the multispectral band means, 18 in the RGB CSM metrics and 35 in the various VIs. In total, 22 samples are affected by at least 1 outlier, 17 if the RGB CSM outliers are not accounted for (since they have no impact on the models used for prediction). It has been verified that these outliers are not related spatially or temporally. It has been chosen to replace them by the predicting mean matching from the non-outliers. This can lead to a great degradation in the model's accuracy, but removing the observations with missing values was not an option considering the high proportion of samples affected. A first alternative to this choice would have been to choose the other replacement option, which replaces the outlier value by the OOB predictions based on the other variables, but it could also lead to missing important information and to a virtual increase in accuracy. Another alternative would have been to remove variables that exceed a certain outlier percentage threshold, reducing the dataset variables and observations at a certain extent, but keeping the original data untouched. On the other hand, this could have removed important variables as side-effect.

iv. Impact of different wheat cultivars on the models

The two different cultivars, Chevignon and Safari, undeniably have different reflectances in the visible domain. This is shown in Figure 23, where Safari seems to have more reflectance in the blue band than Chevignon. Figure 24, on the other side, shows their response in the IR domain through a false-color infrared image. No noticeable difference in this domain seems visible between the two cultivars.

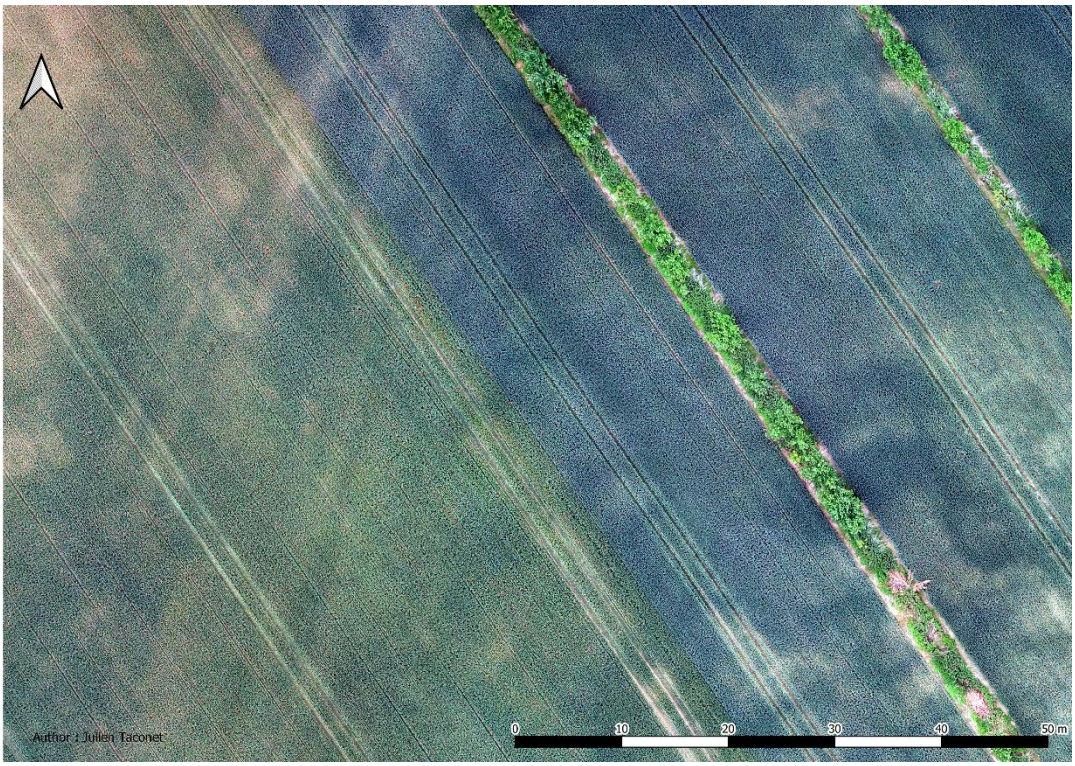


Figure 23: RGB orthophoto of Chevignon (right) and Safari (left) cultivars.

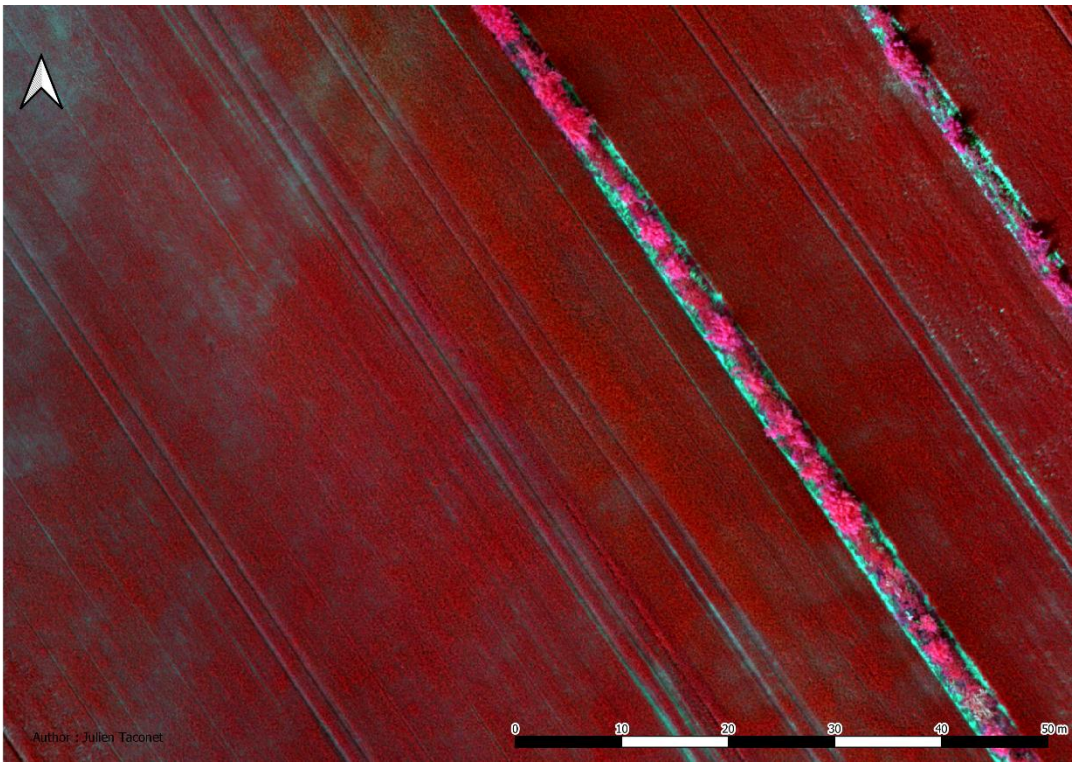


Figure 24: False-color infrared orthophoto of Chevignon (right) and Safari (left) cultivars.

As we can see in Table 8, the models have very different accuracies depending on the selected cultivar(s). This proves the cultivars can have an impact on the models, especially for ear dry matter, where an important degradation of model accuracy is observed from the Chevignon model to the Safari model. The generalized model with both cultivars results in an intermediate accuracy, that will not yield the best results for any of the cultivars, but has the advantage of being more robust (with

twice as many samples as the cultivar-specific models used to create it) and applicable to both cultivars.

If wheat crop parameters are to be monitored on this parcel in the future, it is important to adapt the selected model the studied cultivar. First of all, this work proved that cultivar-specific models are not transposable to another cultivar. If the same cultivars are studied, the generalized model will yield the best and safer results on the whole parcel. The generalized model could also be used on cultivars that were not comprised in this work, though it must be taken with care, and a test set should be sampled to check its accuracy. If the Chevignon cultivar is studied alone, the models based only on Chevignon would give the best results, as their accuracies overcome the generalized one's. However, if the Safari cultivar is studied and the ear dry matter must be monitored, it would be preferred to use the generalized model. As a matter of fact, the model for the Safari cultivar is based on fewer (-6) samples than the Chevignon cultivar, which makes it less robust. This lack of data, originating from the early development stages of the ear (during ear swelling, Z47, on the 25/05), is also be the reason why the Safari cultivar model's accuracies are degraded compared to the Chevignon cultivar or the generalized model (see Table 8). In fact, the collected ear dry matter data on the 25/05 increases the range of the response data, with low ear dry matter values, corresponding to an increased range in the explanatory data (photosynthetic activity at its maximum).

This is also the reason why the ear dry matter shows higher values for the Safari cultivar on the 25/05 (Map 6) : no ground truth data for this cultivar on this date was collected, thus the model wrongly interpreted it's reflectance by assigning it higher ear dry matter values. This underlines a particular weakness of the ear dry matter model: special attention must be paid when using it outside the range of development stages it was built on, because random forests work poorly outside the range of values they were trained on. Blindly following the model's results will lead to erroneous conclusions, for example in Map 9, the first conclusion that could come in one's mind is that there is a problem with the Safari cultivar. It seems that compared to Chevignon cultivar, the ear dry matter increases at a lower extent, but it is only due to the Safari predictions on the 25/05, where reflectance was wrongly interpreted since it was not linked to any ground-truth data. A first solution to this problem could have been to assign values of 0 to ear dry matter on this date, though it is not exactly true since ears were swelling in the flag leaf's sheath at this moment. A more accurate model could have been built by weighing the ear dry matters even if they were not out of their sheath, but that would have required a great amount of time.

v. Impact of the last flight data on the models

The last flight (13/07) is a very particular source of data. In fact, it greatly increased the range of response data: ground-truth LAI values were equal to 0 on all the Chevignon part of the study area, set between 0 and 1 on the Safari part, and a great increase in ear dry matter occurred between this date and the previous. The range of the explanatory data also reached extremes: with leaf senescence, the photosynthetic activity was almost null, greatly affecting both RGB and multispectral responses. The consequence of this enlargement of data ranges, due to maturation and leaf senescence, is a huge increase in model accuracies, as the differences between Table 8 and Table 10 relate.

Table 10 is basically an indicator of the model's ability to track changes between the early heading stages (Z47) and the early grain development stages (Z71). It informs us that the models are less effective in tracking changes in ear dry matter and LAI between these stages, with pseudo R-squared values respectively equal to 60% and 65%. In fact, the variations in both explained variables between

these stages are less extreme, thus the models are more affected by noise in the images. This is the reason why, despite the very high accuracies reported in Table 8, the results of this work must still be taken with care, especially the change maps. On the other hand, the high data-range collected at each date in this work allows to yield accurate relative maps for a single flight, in order to study the distribution of a given parameter on the parcel, provided that the illumination conditions within the flight stay constant.

As an example of wrong interpretation risks, it was pointed out in the results section that in Map 6, there is an area with high ear-dry matter between the agroforestry hedges, at the top. This area was actually a dryer area than the rest of the parcel, as we can see in Map 4, where LAI is very low. As the addition of 20 samples in the model from the 13/07 brought in these new, extreme values of ear dry matter and reflectance, this area was interpreted as having high ear dry matter values, since its reflectance approached mature wheat reflectance.

vi. Relative LAI deviation map error source

An error source in the analysis of variance maps was detected in this work. In fact, the LAI mean, calculated to be used in Equation 3, is based on the whole study area polygon, which also covers a few areas with bare-ground, trials with other crops, agroforestry hedges, and a few other artifacts. This should have had a small impact on the LAI mean, but could account for a few percentages in relative LAI deviation. The same operations should have been carried out on a new polygon, built by removing every artifact, except for the wheat parcel.

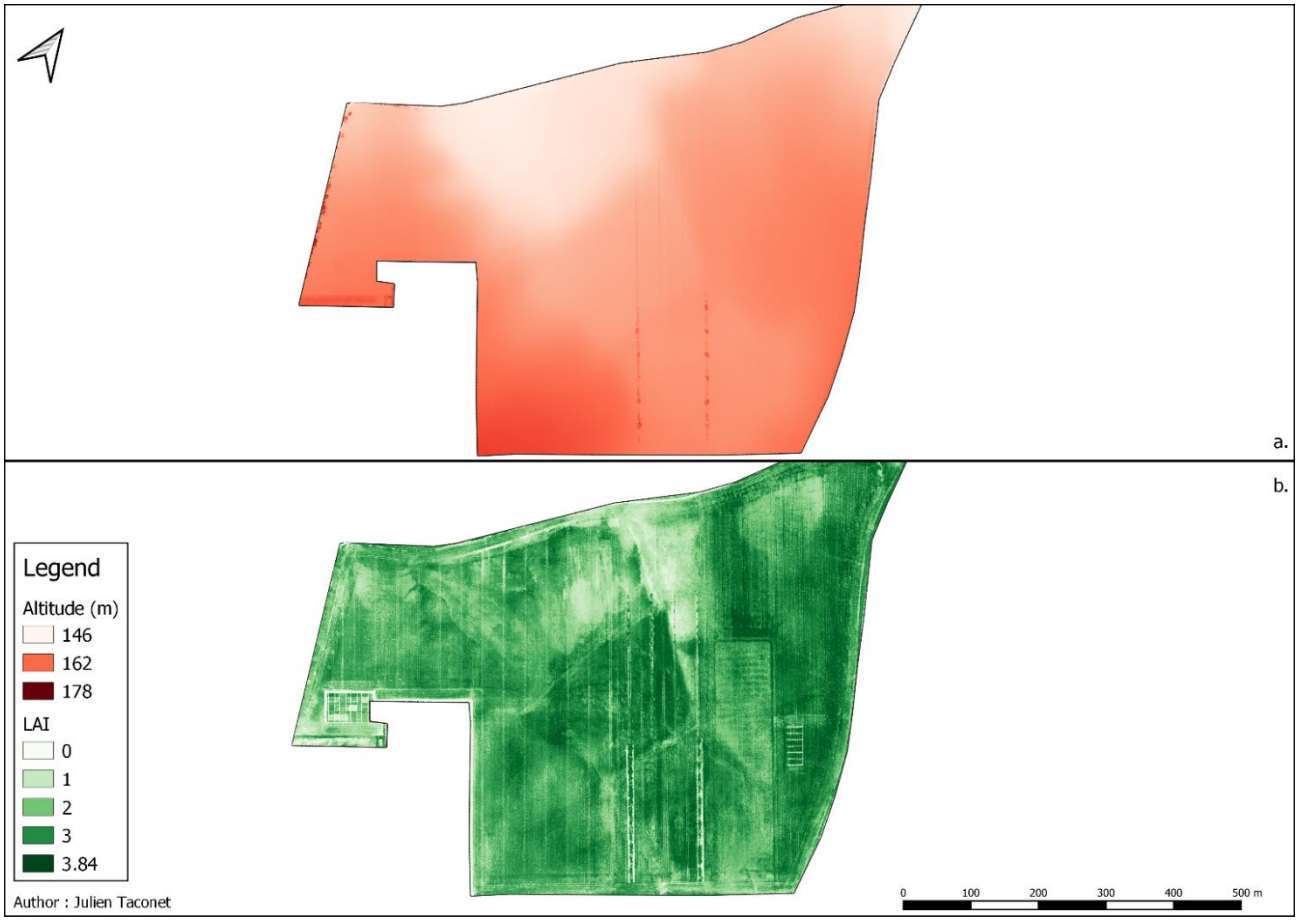
b. Model results

In this section, an analysis of the ability of the models and predicted maps to reveal information that was not used as input will be carried out. The impact of elevation and slope, soil type and former experiments on predicted LAI will be discussed. This is an important step in this study, since the relocation of experimental blocs towards more suitable areas depends on the LAI prediction maps, and on the hypothesis that LAI is a mirror of soil conditions.

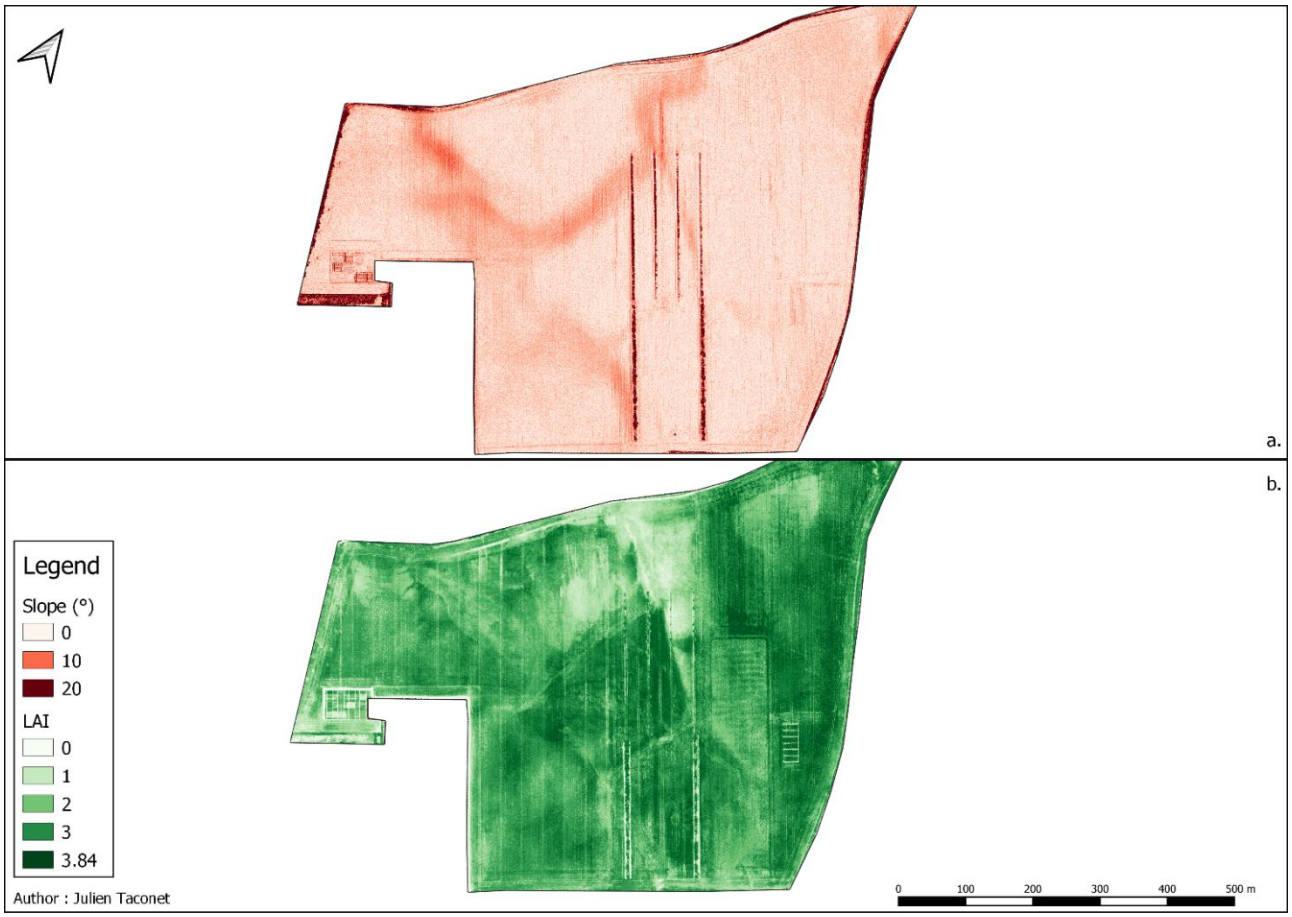
i. Impact of elevation and slope on the crop parameters

Map 17 and Map 18 allow to visualize the impact of elevation and slope on the LAI predictions for the 25/05 flight. It is interesting to study these parameters in parallel, since the developed models are not based on slope or absolute elevation metrics, only on Crop Surface Model metrics.

These maps clearly highlight that the source of the pattern on the LAI maps that is observed from the bottom right to the top left of the map is slope and elevation. Map 17 shows that at the top left of the map, where the absolute altitude is the lowest, the LAI values are also low. This translates a response of the plant to different soil conditions, since this area is a settling pond, where finer soil particles and water tend to accumulate, thus reducing soil ventilation. The slope Map 18 also points out the effect slope has on the crop: all areas with higher slopes (varying between 10° and 20°) are subject to a diminished LAI. This can be explained by the runoff phenomenon, which erodes the finer soil particles, thus reducing its capacity to provide important nutritional elements to the crop.



Map 17: DEM (a) and LAI predictions on the 25/05 flight (b).



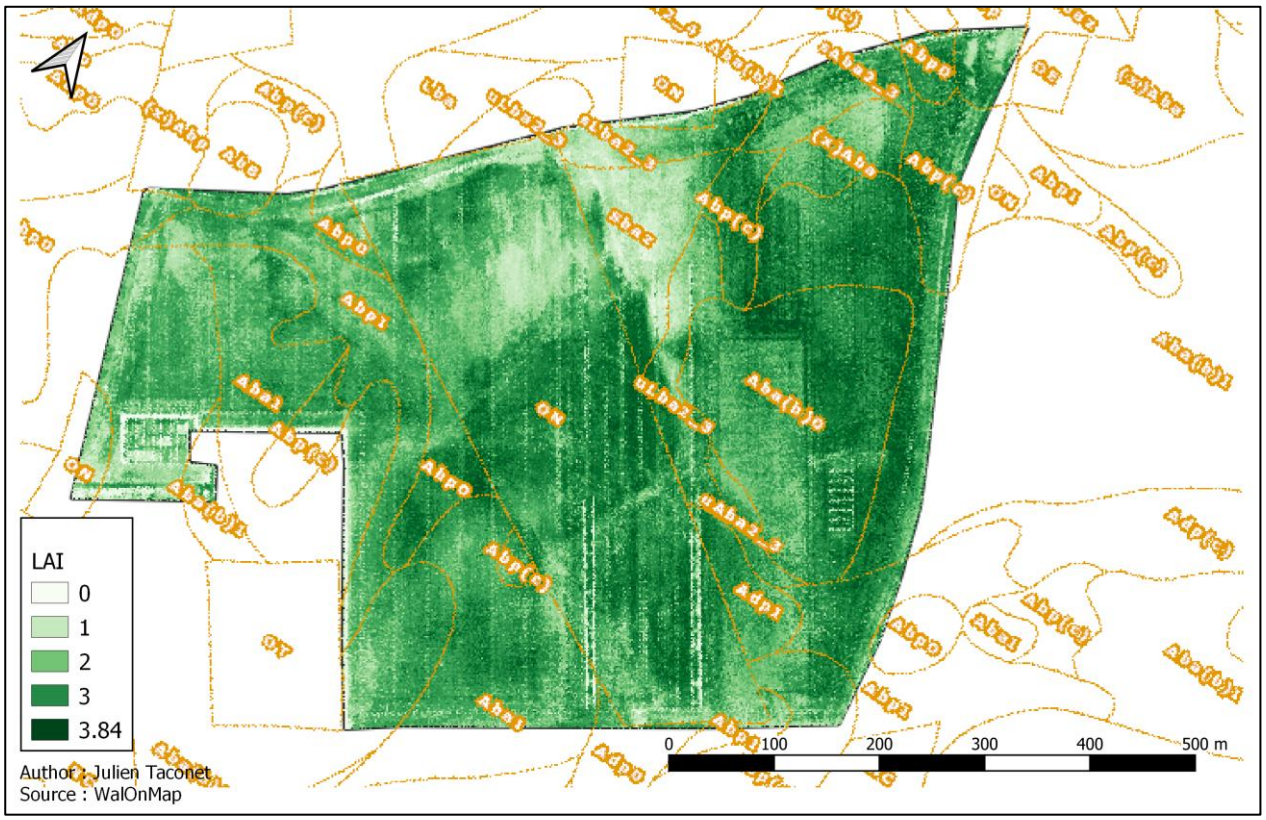
Map 18: Slope (a) and LAI predictions on the 25/05 flight (b).

ii. Impact of the soil type on the crop parameters

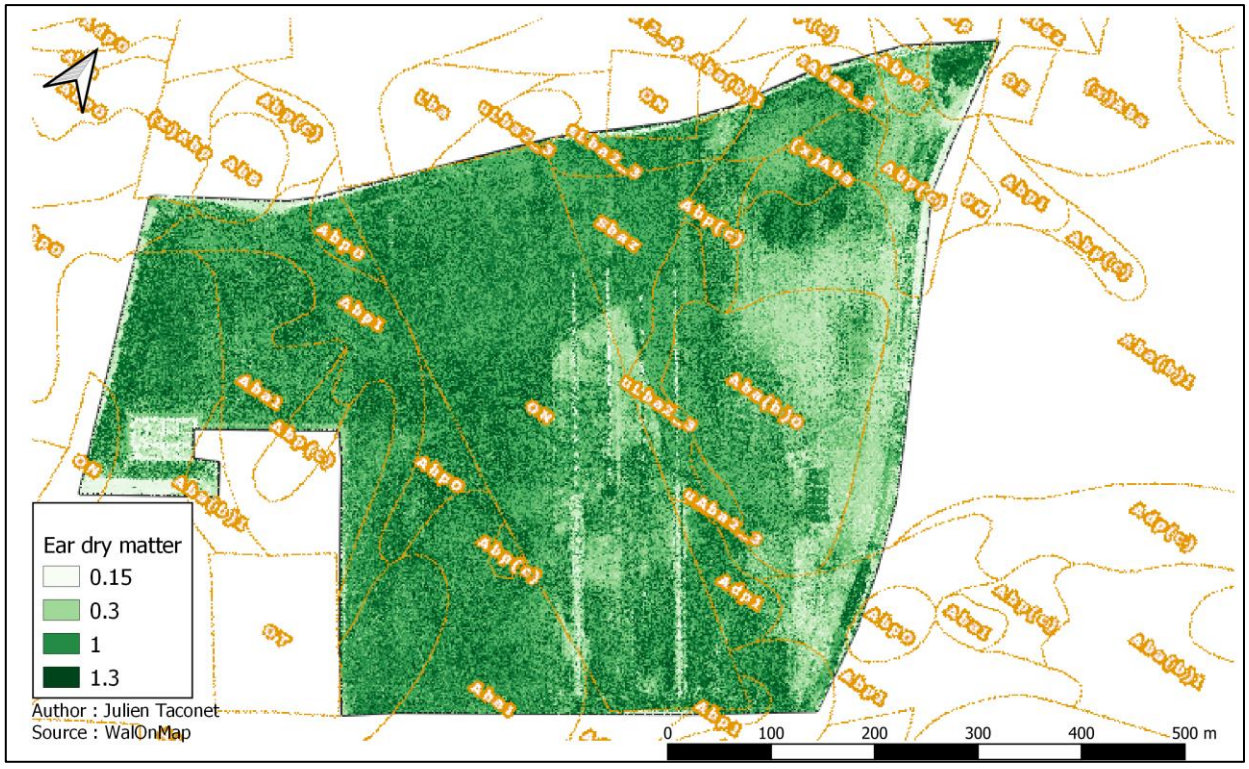
Soil type also has a major impact on both LAI and ear dry matter predictions. The WalonMap legend is furnished in Annex 3: WalonMap's Numeric Soils Map legend.

Map 19 shows the soil map, extracted from the WalonMap database, covering the LAI prediction map. Most of the parcel is composed of Aba and Abp soils, which are respectively silty soils with favorable drainage and a textural B horizon, and silty soils with favorable drainage, a textural B horizon and no real profile development. The diagonal slope pattern is registered as a backfill soil. The two points of interest are the negative impacts on LAI of the Sbaz and (x)Aba soil types. The Sbaz soil type (Sand-silty soil, with favorable drainage, textural B horizon and increasingly coarse with depth) can be found covering the area with the lowest LAI on the whole parcel, at the top of the agroforestry hedges. This is an area subject to runoff since it has an important slope on part of it, but the triangle formed by this soil type covers a larger area and is clearly visible through the LAI prediction maps. It is important to remember soil type is closely linked to elevation and slope, and this part of the parcel is located right next to the settling pond: water inevitably runs off this area following the slope. The other point of interest is located at the top-right of the map. The (x)Aba soil type (silty soil with favorable drainage, a B textural horizon, and a substrate starting at medium depth) is the only possible explanation for the slight drop in LAI in this region since it is not a steep nor deep region, and underlines the capacity of the developed LAI model to reflect the soil conditions on this parcel.

Map 20, on the other hand, shows that the differences in soil types are not visible through the ear dry matter model predictions at wheat maturity. The patterns of lower ear dry matter are clearly linked to the development stage delay between Chevignon and Safari, visible in the LAI predictions Map 5.



Map 19: Soil types and LAI predictions on the 25/05 flight.



Map 20: Soil types and ear dry matter predictions on the 13/07 flight.

iii. Impact of former experiments on the crop parameters

The impact of the former trial on spring crops is clearly visible on all LAI prediction maps, at the right of the agroforestry hedges. It is now covered with wheat, but the trials left a trace on the soil's state, which is reflected through lower LAI values for all dates, Map 5. A darker line running around the trial is visible, and corresponds to the former tractor tracks: the remaining nitrogen from the former experiment was assimilated by wheat during the season, which is reflected through higher LAI values. The brome trial is visible on this map at the right of the first one, but is explained by the fact that winter wheat was not sown over this area: it was covered with grass during the whole season.

c. Equipment and flight frequency requirements

In this section, an analysis of the advantages and drawbacks of the three tested systems (RGB, Multispectral and the combination of both) will be carried out. The frequency of the flights in relation with their ability to accurately monitor the parcel will also be addressed.

i. Equipment considerations

The main differences in the RGB equipment and the multispectral equipment are the spatial and radiometric resolution of the output products, and their price.

The analysis of the summary of model accuracies in Table 8 shows that the combination of both cameras yields the most accurate models. The multispectral-only models come next, with a very slight degradation in accuracy, and finally the RGB-only models, with a noticeable degradation.

The models in this work were built on the combination of both cameras to obtain the most accurate crop parameter maps, in order to study the in-field variability. In terms of equipment, this required to print a 3D-frame to integrate the Micasense camera to the DJI-P4 Pro drone, adding additional weight thus reducing its autonomy. A 3D-printer was needed as well as the expertise to create the frame on the specialized software and to successfully integrate it to the drone. The Micasense Rededge kit, which includes the camera, a calibration panel and a light sensor for real-time calibration sums up to a total of 6,500$[1]. Mount kits are also available to acquisition, though they are only adaptable to specific UAVs and cost 600$. This underlines the sufficient amount of resources one should dispose of to use multispectral equipment. On the other hand, the DJI-P4 pro UAS (includes drone, camera, flight-planning software, operator remote-controller) prices start at 1700$.

The combination of both cameras leads to the highest precision crop parameter maps. On the other hand, applying such models based on rasters with different resolutions implies a resampling of raster values, which is a computationally-extensive operation, given the size of the study area and the fine spatial resolution of RGB products. In this work, the resampling operations were based on the finest RGB camera resolution. A high amount of computing time could have been saved during model prediction by resampling to the multispectral spatial resolution. The choice of which resolution the alignment should be done given the desired application and computation power. Considering the spatial resolution of the outputs, computation time for the creation or usage for prediction of RGB-only models is roughly 20 times higher than for multispectral-only models. Finally, using both cameras allows the outForest() outlier detection algorithm to detect errors in one camera based on the other, acting as a validation method for red, green and blue pixel values.

A significant amount of resources and expertise can be saved by using the RGB system only. But at what impact on the model precisions? While the spatial resolution of the RGB camera is finer, its spectral resolution is coarser. The RGB camera is not subject to a calibration protocol: models built on RGB-only data can lead to inaccurate models. In the worst case, a flight made during changing illumination conditions, combined to a sampling campaign where all samples were made on equal ground illumination conditions (this can occur by chance) can build apparently accurate models, leading to erroneous conclusions when analyzing the outputs.

[1] https://micasense.com/shop/Sensors-c26039403 (visited on the 01/08/2020)

Table 8 shows the RGB-only models are less accurate than the multispectral-only models, especially for the ear dry matter model. Considering the LAI, there is a slight increase of 5% in accuracy in the multispectral-only model compared to the RGB-only model. Another 2% are gained by combining both data sources. Apparently, the total accuracy gain when comparing an RGB-only model to a model with all variables hardly outweighs the additional resources required to build the latter. Considering the impact of the last flight data in section 6.a.v, it is important to compare the accuracy of the models built without the last flight data. Table 10 reveals that the accuracy of the LAI model is greatly affected when the multispectral information is removed, and the RGB-only model becomes hardly exploitable with a 52% pseudo R-squared value. All things considered, an RGB-only model for LAI monitoring can be worth creating and using if limited resources are available. Multispectral equipment will achieve higher accuracies and lower computation time, but is disposable if extreme accuracy is not required.

Now to the ear dry matter model: Figure 12, the variable importance plot of the ear dry matter model for all variables, reveals that the multispectral CSM and the NIR band are of upmost importance. This is interesting since the multispectral CSM has a very coarse spatial resolution (57cm/pixel) compared to the RGB CSM (5.5cm/pixel). The RGB CSM metrics do not appear at all in the variable importance graphs, underlining the fact that they are irrelevant, and that errors pointed out in section 6.a.i play an important role. This also means accurate height metrics are crucial when creating an ear dry matter model, especially if an accurate monitoring of ear dry matter from heading to early grain development is desired, as seen in Table 10 where the multispectral-only model achieves a pseudo R-squared value 60% compared to 10% for the RGB-only model.

Comparing the multispectral-only and the RGB + multispectral ear dry matter model accuracies in Table 8 reveals that the addition of the RGB variables improves the pseudo R-squared value of only 1%. This brings to think that for the creation of an ear dry matter model, a choice should be made between the multispectral equipment and the RGB equipment in regards of the objective, the available funds and computation power, since the accuracy gained by the combination of both data sources is negligible compared to the increase in resources and computation requirements. Again, the usage of multispectral equipment is justified if higher accuracies are required, but depending on the application, RGB equipment can be sufficient.

ii. Flight frequency considerations

Map 8, Map 9 and Map 10 allow to study the ability of the created models to track changes in the crop parameters, based on 2-week intervals and on a 1-month interval.

The ear dry matter change maps can hand useful information at a 2-week interval. Map 9 shows the ear dry matter increases at a much higher extent between the 3rd and the 4th flight than between the 2nd and the 3rd flight. Furthermore, the ear dry matter change in Map 10 is very similar to Map 9 (panel b) except for the scale of variation, which is a sign that the most intense phase of grain-filling, between the two last flights, dominates the final distribution of this parameter. This corresponds well to the grain-filling curve, as drawn by Girousse *et al.* (2018). On the other hand, the 1-month interval temporal resolution is too coarse to study this evolution. Accurate monitoring of the wheat ear grain-filling rates at field-scale through frequent UAV flights is a possible future application of the ear dry matter model created in this work. Care must be taken when using it, since it is only valid between the growth stages it was trained on.

The interpretation of the LAI change maps is not as crystal-clear as the ear dry biomass change maps. An analysis of the LAI change maps reveals that most of the parcel is subject to a decrease or stagnation of LAI values between the two first flights with a few areas where it increases. Between the second and the third flight, LAI values decrease in the areas where it had increased or stagnated between the past flights, and slightly increase where it had previously strongly decreased. The theoretical wheat leaf development pattern is an increase in leaf area until approximately 10 days after anthesis (Acevedo *et al*., 2002), followed by a tendency to decrease. Anthesis was reached on the second flight. Thus, the areas with a great decrease in LAI between the two first flights, followed by an increase between the next flights can be interpreted as an evidence of the impact of noise induced by major illumination condition changes conditions within and between the flights. An area has been inspected on the basis of this hypothesis: Figure 25 and Figure 26 show the RGB orthophotos and LAI change values over this area between the three first flights. Figure 25 shows an area with lower plant density and spots of apparent soil, which causes sun reflection on the 25/05. The 09/06 and 22/06 do not show sun reflection problems, but the lighting conditions are very different between these two flights, as it can be seen through the difference in the working tracks reflectance. This is interpreted by the LAI model by a decrease in values between the two first flights, and an increase between the two next flights. This observation was done in all inspected areas where these illogical LAI variation tendencies have been noticed. The fact that this model was calibrated on the LAI, but actually reflects the GAI was discussed section 6.a.ii, may also explain some elements in the LAI change maps.

LAI monitoring is not an easy task when flights are not run with the same sky conditions, especially if the flights are temporally close to each other. After wheat booting, LAI values have almost reached their peak, and the slight variations this parameter could show in a 2-week interval seem to fade out in noise due to changing illumination conditions between flights.

It was reported that the two last weeks of march were particularly dry weeks, before a few rainfalls on the beginning of June, and that this lack of rain heavily impacted the area at the top-right of the agroforestry hedges. This is visible on the LAI Map 4, and the 2-week interval change Map 8 shows an increase in LAI on this area between the two first flights. It was interesting to note that LAI changes in some areas could be explained through weather variations, but the overall impression for LAI is that noise overcomes reality when it comes to track fine changes in LAI over the whole parcel.

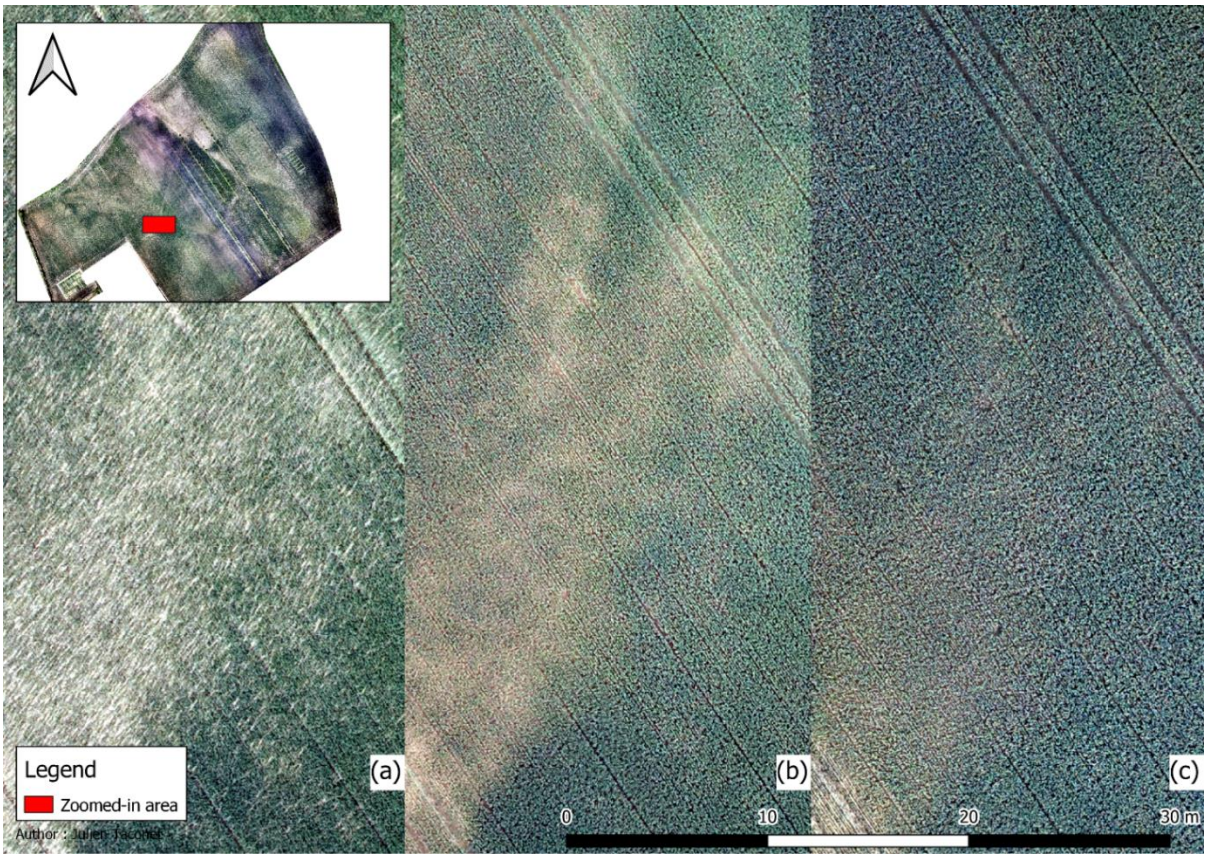


Figure 25: Inspection of a noisy area: RGB orthophotos on the (a) 25/05, (b) 09/06, and (c) 22/06.

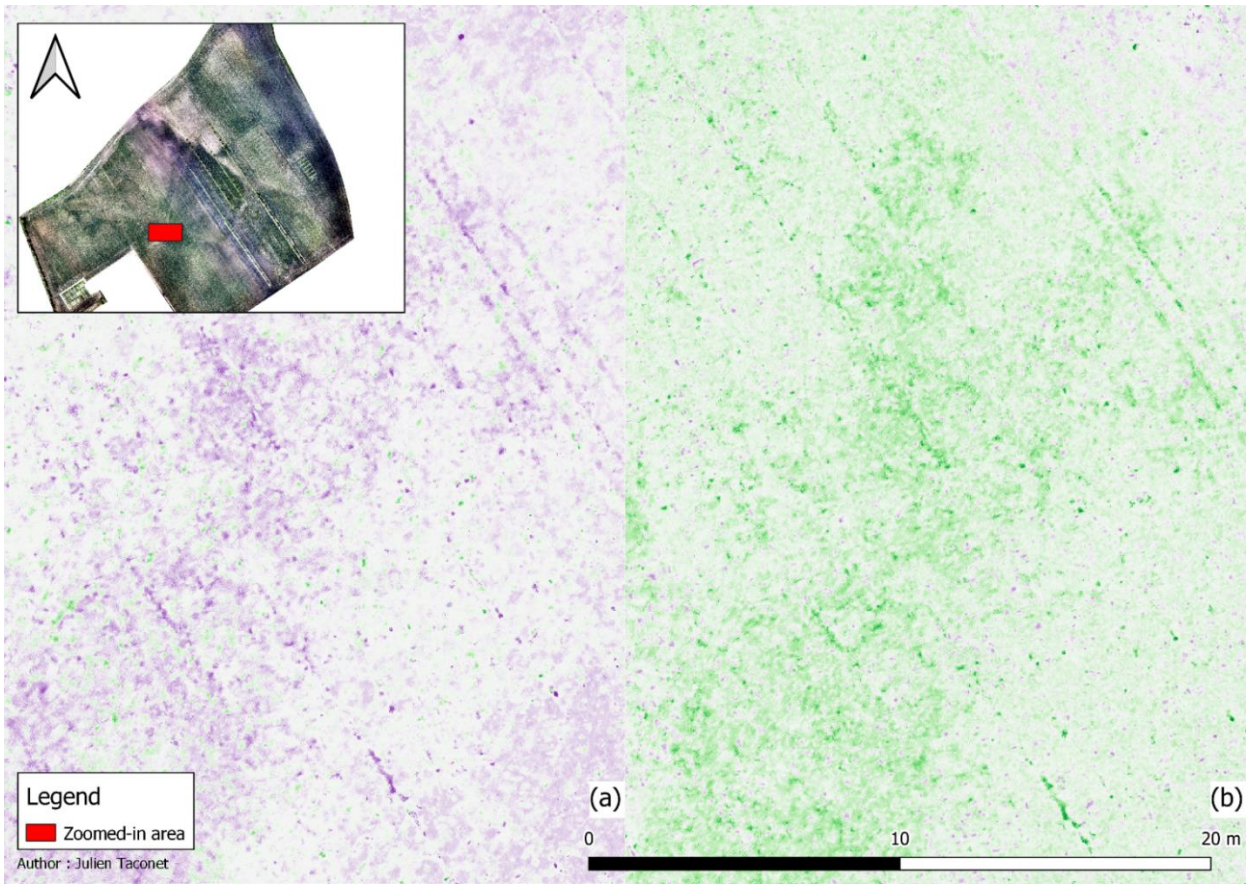


Figure 26: Inspection of a noisy area: LAI change between the (a) 25/05 and the 09/06, and (b) between the 09/06 and the 22/06.

As more flights imply higher investments, it is important to discuss flight frequency in relation with the information gain. In fact, flight frequency for model building should be thought depending on the crop parameter to monitor, in function of the variations of this parameter. An increased number of flights should be carried out when it varies most, and resources can be saved with fewer flights when it varies less. A fixed-frequency flight calendar, as it has been done in this work, has the advantage of being less organizationally expensive, but its inflexibility makes it prone to the appearance of noise because of non-essential additional flights, and to missing important variations in the studied parameter. In fact, between the 22/06 and the 13/07, huge variations in ear dry matter and LAI were noted: intermediate values would have been welcome in order to increase the robustness of the model. On the other hand, two carefully-selected flight dates between the 22/05 and the 22/06, could have been enough (instead of 3) to cover the variations in these parameters, while reducing the impact of noise in the models.

Once an robust model was built, if it is less affected by noise than in this work (by paying careful attention to having the same weather conditions between flights especially if using RGB equipment, or using only multispectral equipment, but still paying attention to have homogeneous illumination conditions within each flight), accurate tracking of LAI could be realized by flying the drone at any given date. Ear dry matter could also be accurately tracked this way, but only within the development stages the model was built on. Final yield prediction could also be accomplished by flying the UAV when the whole parcel reached anthesis (since grain density is fixed at this stage) and studying the ear dry matter values partitioning within the studied area. Anticipatory corrections, if desired, could also be carried out through the analysis of the model outputs from a flight before anthesis, and compensating for adverse conditions in the affected areas.

7. Conclusion

The present work is the first step of a larger study, AgricultureIsLife's experiment on new agricultural methods. It has focused on creating relevant crop parameter models to answer to a double objective: map the spatial variability of soil conditions in the parcel and monitor the evolution of the crop parameters over time during the experiment. Many models were created, but only the most accurate were used for interpretation purposes. The studied parameters (LAI and ear dry matter) seem to be modeled accurately enough to answer to these objectives, thought the sources of inaccuracies were discussed. Their ability to track crop parameter variations was assessed, and it appears that the LAI model is strongly affected by noise due to changing illumination conditions between flights, while the ear dry matter model is less prone to it and could be used for monitoring purposes. Recommendations were made for future works in crop parameter modeling through UAV imagery, in order to minimize the impact of noise. Flight frequency was discussed, and it was concluded that the flight schedule should be adapted to the temporal variations of the studied crop parameters, with a particular focus on careful selection of flight dates in function of weather conditions and the type of sensors the drone is equipped with. The necessary equipment for such monitoring was also questioned: an accurate monitoring should be based on multispectral equipment, and RGB sensors are disposable of. It has been proved that a monitoring with limited funds could still be done based solely on RGB equipment, but with lower accuracy, and this equipment is especially affected by the weather conditions during the flights.

The models were used to map the spatial variability of soil conditions in the parcel. Presence of variability in a parcel as large as this is inevitable, and it is important, before launching a trial of this magnitude, to quantify this variability. The experimental blocs must be set up in the most homogeneous areas possible, to reduce the impact of the in-field variability on the experiment results, which must be known from researchers. The Leaf Area Index parameter was used for this objective, since it is a reliable parameter to reflect the soil conditions. The maps created were reliable enough to be submitted statistical analyses, and new experimental bloc positions were proposed in order to reduce their within- and between-variability.

The created models will be eventually used to monitor the evolution of the crop parameters over time during the experiment. For each rotation, winter wheat is planned to be grown every four years. A UAV flight over the parcel will allow to run the models, visualize LAI or ear dry matter and attend for their evolution. This allows to obtain information on the evolution of soil conditions in relation with their management technique, information which can be an element when conclusions will be drawn on the efficiency and the impact on soils of the tested systems.

8. References

Aase J.K., 1978. Relationship between leaf area and dry matter in winter wheat. *Agron. J.*

Aasen H., Honkavaara E., Lucieer A. & Zarco-Tejada P.J., 2018. Quantitative Remote Sensing at Ultra-High Resolution with UAV Spectroscopy: A Review of Sensor Technology, Measurement Procedures , and Data Correction Workflows. *Remote Sens.* 1–42.

Acevedo E., Silva P. & Silva H., 2002. Wheat Growth and Physiology. *In*: Curtis, B.C., Rajaram, S., Gomez Macpherson, H. eds. *Bread Wheat – Improvement and Production*. FAO Plant Production and Protection Series, No. 30. Food and Agriculture Organization of the United Nations, Rome.

Alphen B.J.V.A.N. & Stoorvogel J.J., 2001. A Methodology for Precision Nitrogen Fertilization in High-Input Farming Systems. *Precis. Agric.* 319–332.

Aydinalp C. & Cresser M.S., 2008. The Effects of Global Climate Change on Agriculture 1. *Am. J. Agric. Environ. Sci.* **3**(5), 672–676.

Basso B., Cammarano D. & Carfagna E., n.d. Review of Crop Yield Forecasting Methods and Early Warning systems. *Dep. Geol. Sci. Michigan State Univ.* 1–56.

Bendig J., Willkomm M., Tilly N., Gnyp M.L., Bennertz S., Qiang C., Miao Y. & Bareth G., 2013. Very high resolution crop surface models (CSMs) from UAV-based stereo images for rice growth monitoring In Northeast China. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **XL**(September), 4–6.

Bendig J.V., 2015. Unmanned aerial vehicles (UAVs) for multi-temporal crop surface modelling - A new method for plant height and biomass estimation based on RGB-imaging.

Berger K., Atzberger C., Danner M., Vuolo F. & Hank T., 2018. Evaluation of the PROSAIL Model Capabilities for Future Hyperspectral Model Environments : A Review Study. *Remote Sens.*

Boegh E., Soegaard H., Broge N., Hasager C.B., Jensen N.O., Schelde K. & Thomsen A., 2002. Airborne multispectral data for quantifying leaf area index , nitrogen concentration , and photosynthetic efficiency in agriculture. *Remote Sens. Environ.* **81**, 179–193.

Brandt A.J., Pino G.A. & Burns J.H., 2014. Experimental Protocol for Manipulating Plant-induced Soil Heterogeneity. *J. Vis. Exp.* (March), 1–9.

Bréda N.J.J., 2003. Ground-based measurements of leaf area index : a review of methods , instruments and current controversies. *J. Exp. Bot.* **54**(392), 2403–2417.

Breiman L.E.O., 2001. Random Forests. *Mach. Learn.* 5–32.

Brostaux Y., 2017. Course on Multivariate analysis: Data Mining and Machine Learning.

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.

Chen J.M. & Black T.A., 1992. Defining leaf area index for non-flat leaves. *Plant, Cell Environ.* 421–429.

D.J Mulla., A.U B. & R. K., 1990. Methods for Removing Spatial Variability from Field Research Trials. *In*: *Advances in Soil Sciences, Vol.13*.

Delgado C., 1999. Live stock to 2020 The Next Food Revolution. *Outlook Agric.* (May 1999).

Diaz R.J. & Rosenberg R., 2008. Spreading Dead Zones and Consequences for Marine Ecosystems. *Science (80-.).* (August), 926–930.

European Space Agency, EO4SD Agriculture and Rural Development cluster (2019). Final Report on Earth Observation for Sustainable Development in Agriculture and Rural Development, 2019., 1–29.

FAO, 2000. Land resource potential and constraints at regional and country levels. *World Soil Resour. Reports*.

FAO, 2006. Livestock's long shadow: environmental issues and options.

FAO, 2006. *The State of Food Insecurity in the World*.

FAO, 2010. "Climate-Smart" Agriculture: Policies, Practices and Financing for Food Security, Adaptation and Mitigation.

FAO, 2011. The State of the World's Land and Water Resources for Food and Agriculture. Managing Systems at Risk.

Fischer R.., 1985. Physiological limitation to producing wheat in semitropical and tropical environments and possible selection criteria. *In*: *Proc. Int. Symp. Wheats for More Tropical Environments*. DF, CIMMYT., Mexico, 209–230.

Foley J.A., Defries R., Asner G.P., Barford C., Bonan G., Carpenter S.R., Chapin F.S., Coe M.T., Daily G.C., Gibbs H.K., Helkowski J.H., Holloway T., Howard E.A., Kucharik C.J., Monfreda C., Patz J.A., Prentice I.C., Ramankutty N. & Snyder P.K., 2005. Global consequences of Land Use. *Science (80-.).* (July), 570–575.

Gago J., Douthe C., Coopman R.E., Gallego P.P., Ribas-carbo M., Flexas J., Escalona J. & Medrano H., 2015. UAVs challenge to assess water stress for sustainable agriculture. *Agric. Water Manag.* **153**, 9–19.

Gebbing T., Schnyder H. & Kühbauch W., 1998. Carbon mobilization in shoot parts and roots of wheat during grain filling : assessment by 13 C / 12 C steady-state labelling , growth analysis and balance sheets of reserves. *Plant, Cell Environ.* 301–313.

Gibbs H.K. & Salmon J.M., 2016. Mapping the world ' s degraded lands. *Appl. Geogr.* **57**, 12–21.

Gibson D.J., 1988. The Maintenance of Plant and Soil Heterogeneity in Dune Grassland. *J. Ecol.* **76**(2), 497–508

Girousse C., Allard V., Piquet-pissaloux A., Le J. & Id G., 2018. Different grain-filling rates explain grain- weight differences along the wheat ear. *PLoS One* 1–15..

Gleick P.H., 2003. Water Use. *Annu. Rev. Environ. Resour* 275–314.

Global panel on Agriculture and Food Systems for Nutrition, 2016. Food systems and diets: Facing the challenges of the 21st century.

Gomez A.A. & Wiley J., 1983. *Statistical Procedures For Agricultural Research*.

Gomez C. & Purdie H., 2016. UAV- based Photogrammetry and Geocomputing for Hazards and Disaster Risk Monitoring – A Review. *Geoenvironmental Disasters*.

Grenzdörffer G.J., 2014. Crop Height Determination With UAS Point Clouds. *Int. Soc. Photogramm. Remote Sens.* **XL**(November), 17–20.

Haala N., 2013. The Landscape of Dense Image Matching Algorithms. *Photogramm. week* 271–284.

Hallström E., Carlsson-Kanyama A. & Borjesson P., 2014. Environmental impact of dietary change : a systematic review. *Environ. Energy Syst. Stud.* **91**.

Hawkins D.M., 2004. The Problem of Overfitting. *J. Chem. Inf. Model.* 1–12.

Henrique C., Souza W. De, Augusto R., Lamparelli C., Rocha J.V., Sergio P. & Magalhães G., 2017. Height estimation of sugarcane using an unmanned aerial system (UAS) based on structure from motion (SfM) point clouds. *Int. J. Remote Sens.* **38**(8–10), 2218–2230.

Ho T.K., 2002. A Data Complexity Analysis of Comparative Advantages of Decision Forest Constructors. *Pattern Anal. Appl.* 102–112.

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.*

Homi K., 2017. The Unprecedented Expansion Of The Global Middle Class.

Huete A.R., 1988. A Soil-Adjusted Vegetation Index (SAVI). *Remote Sens. Environ.* **309**, 295–309.

IFRPI, 2017. *Global Food Policy Report*.

Jackson R.B. & Caldwell M.M., 1993. Geostatistical Patterns of Soil Heterogeneity Around Individual Perennial Plants. *J. Ecol.* **81**(4), 683–692.

Kohavi R. & Wolpert D.H., 1996. Bias Plus Variance Decomposition for Zero-One Loss Functions 3 Bias Plus Variance for Zero-One. *Mach. Learn.*

Kremen C., Williams N.M. & Thorp R.W., 2002. Crop pollination from native bees at risk from agricultural intensification. *PNAS* **99**(26).

Li H., Lascano R.J., Booker J., Wilson L.T. & Bronson K.F., 2001. Cotton lint yield variability in a heterogeneous soil at a landscape scale. *Soil Tillage Res.* **58**, 245–258.

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.

Lim S.S., Bull F., Burnett R.T., Byers T.E., Calabria B., Carapetis J., Carnahan E., Chafe Z., Charlson F. & Chen H., 2012. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions , 1990 – 2010 : a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* (November), 1990–2010.

Liu Y., Cheng T., Zhu Y., Tian Y., Cao W., Yao X. & Wang N., 2016. Comparative analysis of vegetation indices, non-parametric and physical retrieval methods for monitoring nitrogen in wheat

using UAV-based multispectral imagery. *In*: *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. 7362–7365.

Lu N., Zhou J., Han Z., Li D., Cao Q., Yao X., Tian Y., Zhu Y. & Cao W., 2019. Improved estimation of aboveground biomass in wheat from RGB imagery and point cloud data acquired with a low - cost unmanned aerial vehicle system. *Plant Methods* 1–16.

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. Int J Appl Earth Obs Geoinformation Multitemporal fi eld-based plant height estimation using 3D point clouds generated from small unmanned aerial systems high-resolution imagery. *Int J Appl Earth Obs Geoinf.* **64**(September 2017), 31–42.

Mamaghani B.G., Sasaki G. V, Connal R.J., Kha K., Knappen J.S., Hartzell R.A., Marcellus E.D., Bauch T.D. & Raque N.G., 2018. An initial exploration of vicarious and in-scene calibration techniques for small unmanned aircraft systems, 1–19.

Martínez-carricondo P., Agüera-vega F., Carvajal-ramírez F., Mesas-carrascosa F., García-ferrer A. & Pérez-porras F., 2018. Geoinformation Assessment of UAV-photogrammetric mapping accuracy based on variation of ground control points. *Int J Appl Earth Obs Geoinf.* **72**(May), 1–10.

Marzolff I., Peter K.D., Ries J.B. & Sensing R., 2012. Unmanned Aerial Vehicle (UAV) for Monitoring Soil Erosion in Morocco. *Remote Sens.* 3390–3416.

Matese A., Filippo S., Gennaro D., Berton A., Matese A., Filippo S., Gennaro D. & Berton A., 2017. Assessment of a canopy height model (CSM) in a vineyard using UAV-based multispectral imaging. *Int. J. Remote Sens.* **38**(8–10), 2150–2160.

Matson P.A., Parton W.J., Power A.G. & Swift M.J., 1997. Agricultural Intensification and Ecosystem Properties. *Science (80-.).* **277**(July).

Mcbratney A.B., Whelan B.M., Ancev T. & Bouma J., 2005. Future Directions of Precision Agriculture. *Precis. Agric.* (February).

Miura T. & Huete A.R., 2009. Performance of Three Reflectance Calibration Methods for Airborne Hyperspectral Spectrometer Data. *Sensors* 794–813.

Mutanga O. & Skidmore A.K., 2004. Narrow band vegetation indices overcome the saturation problem in biomass estimation. *Int. J. Remote Sens.* **1161**.

Ortega-terol D., Hernandez-lopez D. & Gonzalez-aguilera D., 2017. Automatic Hotspot and Sun Glint Detection in UAV Multispectral Images. *Sensors* 1–16.

Oscarson P., 2000. The strategy of the wheat plant in acclimating growth and grain production to nitrogen availability. *J. Exp. Bot.* (December 2000).

Padró J., Muñoz F., Planas J. & Pons X., 2019. Int J Appl Earth Obs Geoinformation Comparison of four UAV georeferencing methods for environmental monitoring purposes focusing on the combined use with airborne and satellite remote sensing platforms. *Int J Appl Earth Obs Geoinf.* **75**(November 2018), 130–140.

Paustian K., Andrén O., H.H. Janzen, R. Lal, P. Smith, G. Tian, H. Tiessen M.V.N. & Woomer P., 1997. Agricultural soils as a sink to mitigate CO2 emissions. *Soil Use Manag.* 230–244.

Pérez-Harguindeguy N., Díaz S., Garnier E., Lavorel S., Poorter H., Jaureguiberry P., Cornwell W.K., Craine J.M., Gurvich D.E., Urcelay C., Veneklaas E.J., Reich P.B., Poorter L., Wright I.J., Ray P., Enrico L., Pausas J.G., Vos A.C. De, BuCSMann N., Funes G., Quétier F., Hodgson J.G., Thompson K., Morgan H.D., Steege H., Heijden M.G.A. Van Der, Sack L., Blonder B., Poschlod P., Vaieretti M. V, Conti G., Staver A.C., Aquino S. & Cornelissen J.H.C., 2016. New handbook for standardised measurement of plant functional traits worldwide. *Aust. J. Bot.* **20**, 715–716.

Pingali P.L. & Improvement W., 2012. Green Revolution: Impacts, limits, and the path ahead. *PNAS* **109**(31), 12302–12308.

Pinter P.J., Hatfield J.L., Schepers J.S., Barnes E.M., Moran M.S., Daughtry C.S.T. & Upchurch D.R., 2003. Remote Sensing for Crop Management. *Photogramm. Eng. Remote Sensing* **69**(6), 647–664.

Popp A., Lotze-campen H. & Bodirsky B., 2010. Food consumption , diet shifts and associated non-CO 2 greenhouse gases from agricultural production. *Glob. Environ. Chang.*

Ranquist E.A., Steiner M. & Argrow B., 2016. Exploring the range of weather impacts on UAS operations.

Reynolds M., Foulkes M.J., Slafer G.A., Berry P., Parry M.A.J., Snape J.W. & Angus W.J., 2009. Raising yield potential in wheat. *J. Exp. Bot.* **60**(7), 1899–1918.

Richards R.A., 1996. Increasing Yield Potential in Wheat : manipulating sources and sinks. *In*: M.P. Reynolds, S.R.& A.M. ed. *Increasing Yield Potential in Wheat: Breaking the Barriers*. DF, CIMMYT, Mexico, 134–149.

Robert P.C., 2002. Precision agriculture : a challenge for crop nutrition management. *Plant Soil* (Figure 2), 143–149.

Saeedipour S. & Foad M., 2010. Effect of Drought at the Post-anthesis Stage on Remobilization of Carbon Reserves and Some Physiological Changes in the Flag Leaf of Two Wheat Cultivars Differing in Drought Resistance. *J. Agric. Sci.* **3**(3), 81–92.

Sagan V., Maimaitijiang M., Sidike P., Eblimit K., Peterson K.T., Hartling S., Esposito F., Khanal K., Newcomb M., Pauli D., Ward R., Fritschi F., Shakoor N. & Mockler T., 2019. UAV-Based High Resolution Thermal Imaging for Vegetation Monitoring , and Plant Phenotyping Using ICI 8640 P , FLIR Vue Pro R 640 ,. *Remote Sens.* (2).

Schott R.J., 2007. *Remote Sensing The Image Chain Approach 2nd Edition*, Oxford University Press, Inc.

Seufert V., Ramankutty N. & Foley J.A., 2012. Comparing the yields of organic and conventional agriculture. *Nature* 2–7.

Sieberth T., Wackrow R. & Chandler J.H., 2014. Motion blur disturbs – The influence of motion-blurred images in photogrammetry. *Photogramm. Rec.* **29**(December), 434–453.

Siegmann B. & Jarmer T., 2015. Comparison of different regression models and validation techniques for the assessment of wheat leaf area index from hyperspectral data. *Int. J. Remote Sens.* **36**(18), 4519–4534.

Singh R.S.G.S., 2017. Traditional agriculture : a climate-smart approach for sustainable food production. *Energy, Ecol. Environ.* **2**(5), 296–316.

Slafer G.A., Calderini D.F. & Miralles D.J., 1996. Yield components and compensation in wheat: opportunities for further increasing yield potential. *In*: Reynolds, M.P. Rajaram, S., Mcnab, A. eds. *Increasing Yield Potential in Wheat: Breaking the Barriers*. DF, CIMMYT, Mexico, 101–133.

Smith J., Sones K., Grace D., Macmillan S., Tarawali S. & Herrero M., 2012. Beyond milk , meat , and eggs : Livestock ' s role in food and nutrition security. *Anim. Front.* 6–13.

Spielman D.J. & Pandya-lorch R., 2010. *Proven Successes in Agricultural Development*.

Stehfest E., Bouwman L., Vuuren D.P. Van, Elzen M.G.J. Den, Eickhout B. & Kabat P., 2009. Climate benefits of changing diet. *Clim. Change* 83–102.

Strobl C., Boulesteix A., Zeileis A. & Hothorn T., 2007. Bias in random forest variable importance measures : Illustrations , sources and a solution. *BMC Bioinformatics* **21**.

Su J., Liu C., Coombes M., Hu X., Wang C., Xu X., Li Q., Guo L. & Chen W., 2018. Wheat yellow rust monitoring by learning from multispectral UAV aerial imagery. *Comput. Electron. Agric.* **155**(August), 157–166.

Sutton P.C., Anderson S.J., Costanza R. & Kubiszewski I., 2016. The ecological economics of land degradation : Impacts on ecosystem service values. *Ecol. Econ.* **129**, 182–192.

Tilman D. & Clark M., 2014. Global diets link environmental sustainability and human health. *Nature* **515**(7528), 518–522.

Tmuši G., Manfreda S., Aasen H., James M.R., Gonçalves G., Ben-dor E., Brook A., Polinova M. & Arranz J.J., 2020. Current Practices in UAS-based Environmental Monitoring. *Remote Sens.*

Treitz P.M. & Howarth P.J., 1999. Hyperspectral remote sensing for estimating biophysical parameters of forest ecosystems. *Prog. Phys. Geogr.* **3**, 359–390.

Trevor H., Tibshirani R. & Friedman J., 2017. *The Elements of Statistical Learning, Data Mining, Inference, and Prediction*, Springer Series.

Vermeulen S.J., Campbell B.M. & Ingram J.S.I., 2012. Climate Change and Food Systems. *Annu. Rev. Environ. Resour.*

Westoby M.J., Brasington J., Glasser N.F., Hambrey M.J. & Reynolds J.M., 2012. Geomorphology ' Structure-from-Motion ' photogrammetry : A low-cost , effective tool for geoscience applications. *Geomorphology* **179**, 300–314.

Willett W., Rockström J., Loken B., Springmann M., Lang T., Vermeulen S., Garnett T., Tilman D. & Declerck F., 2019. The Lancet Commissions Food in the Anthropocene : the EAT – Lancet Commission on healthy diets from sustainable food systems. *Lancet* **6736**(February).

Wood S., Sebastian K. & Scherr S.J., 2000. *Pilot Analysis Of Global Ecosystems*.

World Energy Council, 2011. Global Transport Scenarios 2050.

Xue J. & Su B., 2017. Significant Remote Sensing Vegetation Indices : A Review of Developments and Applications. *J. Sensors* **2017**.

Yao X., Wang N., Liu Y. & Cheng T., 2017. Estimation of Wheat LAI at Middle to High Levels Using Unmanned Aerial Vehicle Narrowband Multispectral Imagery. *Remote Sens.*

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.*

Zhao J., Li J., Liu Q. & Yang L., 2012. A preliminary study on mechanism of lai inversion saturation. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **XXXIX**(September), 77–81.

9. Annexes

Annex 1: Ground-truth sampling protocol

- Sélectionner, sur base d'un examen visuel, une zone où la densité de froment semble homogène
- Placer la latte en bois de 50cm le long d'une ligne de semis
- Acquérir la position précise de l'échantillon au point central de la latte avec le GPS-RTK. Attendre au minimum 30 secondes pour l'acquisition de la position.
- **Biomasse (20 échantillons/date)**
    - Arracher les plants sur les 50cm en longueur et sur 3 lignes de semis, une de chaque coté de la ligne centrale et à l'aide d'un petit trident si besoin.
    - Placer les plants dans la pochette nommée d'après son numéro d'ID : « BIOM – xx »
    - **Une fois tous les échantillons acquis :** déterminer le stade de croissance de chaque échantillon sur l'échelle Zadok.
    - **Pour chaque échantillon :** Découper les racines au niveau du plateau de tallage
    - Enlever les feuilles complètements sèches/brulées au niveau du plateau de tallage
    - En fonction du stade de croissance, séparer les différents organes (épi, tige, feuilles et flag leaf). La flag leaf est séparée du reste lorsqu'elle est complètement étalée, et l'épi lorsque la floraison est initiée.
    - Placer dans l'étuve pour séchage à 60 degrés pendant au moins 3 jours.
    - Peser et noter les masses sèches des organes séparément

- **LAI (10 échantillons/date)**
    - Découper une 4ème rangée de plants parallèlement à l'échantillon de biomasse. Porter attention à ce que cette rangée soit au plus ressemblante aux 3 rangées collectées pour la biomasse.
    - Placer les plants dans la pochette nommée d'après son numéro d'ID : « BIOM – LAI xx »
    - **Une fois tous les échantillons acquis :** placer les échantillons de LAI au frais et entourer les racines de papier essuie-tout et l'humidifier. Bien garder le papier humide pendant toute la durée du travail à suivre.
    - **Pour chaque échantillon :** Séparer les feuilles et la flag leaf, et placer les tiges et épis dans des pochettes microperforées nommées en fonction. Placer ces derniers à l'étuve pour séchage, puis les peser et noter leurs masses sèches.
    - Etaler les feuilles de froment sur du papier adhésif au format A3 ou A4, flag leaf et autres feuilles séparément. Seules les feuilles complètement jaunes ne sont pas étalées.
    - Nommer une feuille A3/A4 (attention à ne pas recouvrir de feuille avec l'étiquette) : « Date_ *numéro d'ID* »
    - Recouvrir le papier adhésif avec la feuille A3/A4
    - Scanner la feuille en pleine couleur et bien noter la résolution choisie.
    - Grace à un algorithme de reconnaissance d'image, déterminer la surface foliaire.

Annex 2: « CentroidToPolygon.R »

```
CentroidToPolygon2 <- function(CRS, centroid, height, width, angle) {

coordinates <- st_coordinates(centroid)
y <- coordinates[,2]
x <- coordinates[,1]

# calculate polygon coordinates for each plot centroid.

square=cbind(x+width*cos(angle)-height*sin(angle),  y+width*sin(angle)+heigh
t*cos(angle),  # NW corner
             x+width*cos(angle)+height*sin(angle),  y+width*sin(angle)-heigh
t*cos(angle),  # NE corner
             x-width*cos(angle)+height*sin(angle),  y-width*sin(angle)-heigh
t*cos(angle),  # SE corner
             x-width*cos(angle)-height*sin(angle),  y-width*sin(angle)+heigh
t*cos(angle), # SW corner
             x+width*cos(angle)-height*sin(angle),  y+width*sin(angle)+heigh
t*cos(angle))  # NW corner again - close ploygon

# Extract the plot ID information
ID=centroid$name

# create spatial polygons from coordinates
polys <- SpatialPolygons(mapply(function(poly, id)
{
  xy <- matrix(poly, ncol=2, byrow=TRUE)
  Polygons(list(Polygon(xy)), ID=id)
},
split(square, row(square)), ID),
proj4string=CRS(as.character(CRS)))

return(polys)
}
```

Annex 3: WalonMap's Numeric Soils Map legend