Eesti Maaülikool
Estonian University of Life Sciences

UNMANNED AIRCRAFT SYSTEMS AND IMAGE ANALYSIS IN YIELD ESTIMATION AND AGRICULTURAL MANAGEMENT

MEHITAMATA ÕHUSÕIDUKI RAKENDAMINE PÕLLUKULTUURIDE SAAGIKUSE JA MAA HARIMISVIISIDE TUVASTAMISEL

KAI-YUN LI

A Thesis
for applying for the degree of Doctor of Philosophy
in Environmental Protection

Väitekiri
filosoofiadoktori kraadi taotlemiseks
keskkonnakaitse erialal

Tartu 2022

Eesti Maaülikooli doktoritööd

Doctoral Theses of the Estonian University of Life Sciences

Eesti Maaülikool
Estonian University of Life Sciences

UNMANNED AIRCRAFT SYSTEMS AND IMAGE ANALYSIS IN YIELD ESTIMATION AND AGRICULTURAL MANAGEMENT

MEHITAMATA ÕHUSÕIDUKI RAKENDAMINE PÕLLUKULTUURIDE SAAGIKUSE JA MAA HARIMISVIISIDE TUVASTAMISEL

KAI-YUN LI

A Thesis
for applying for the degree of Doctor of Philosophy
in Environmental Protection

Väitekiri
filosoofiadoktori kraadi taotlemiseks
keskkonnakaitse erialal

Tartu 2022

Institute of Agricultural and Environmental Sciences Estonian University of Life Sciences

According to the verdict 6-14/17-1 of July 8, 2022, the Doctoral Committee of Environmental Sciences and Applied Biology of the Estonian University of Life Sciences has accepted the thesis for the defence of the degree of Doctor of Philosophy in Environmental Protection.

Opponent: Associate Prof. Dr. **Hui-Ping Tsai**
Dept. Civil Engineering
National Chung Hsing University,
Taichung, Taiwan

Pre-Opponent: Prof. **Tiit Kutser**
Estonian Marine Institute
University of Tartu
Tartu, Estonia

Supervisors: Prof. **Kalev Sepp**
Dr. **Ants Vain**
Institute of Agricultural and Environmental Sciences
Estonian University of Life Sciences
Tartu, Estonia
Prof. **Niall Burnside**
School of Applied Sciences
University of Brighton
Brighton, United Kingdom

Defence of the Thesis: Estonian University of Life Sciences, Fr. R.Kreutzwaldi 5, Tartu on Oct 5th, 2022 at 12.15. The English language was edited by Prof. Niall Burnside. Estonian language was edited by Ülle Sihver (Estonian University of Life Sciences) and by Kaupo Kokamägi. Publication of this thesis is supported by the Estonian University of Life Scieces and by the Doctoral School of Earth Sciences and Ecology created under the auspices of the European Social Fund.

ISSN 2382-7076
ISBN 978-9916-669-63-1 (trükis)
ISBN 978-9916-669-64-8 (pdf)

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

The present thesis is based on the following research papers, which are referred to by their Roman numerals (I- III in the text).

I **Li, K. Y.**, Burnside, N. G., de Lima, R. S., Villoslada, M., Sepp, K., Yang, M. D., Raet, J. Vain, A., Selge, A. & Sepp, K. (2021). The Application of an Unmanned Aerial System and Machine Learning Techniques for Red Clover-Grass Mixture Yield Estimation under Variety Performance Trials. Remote Sensing, 13(10), 1994.

II **Li, K. Y.**, Burnside, N. G., de Lima, R. S., Villoslada, M., Sepp, K., Cabral Pinheiro, V. H., de Lima, B.R.C.A., Yang, M. D., Vain, A., & Sepp, K. (2021). An Automated Machine Learning Framework in Unmanned Aircraft Systems: New Insights into Agricultural Management Practices Recognition Approaches. Remote Sensing, 13(16), 3190.

III **Li, K. Y.**, de Lima, R. S., R., Burnside, N. G., Vahtmäe, E., Kutser, T., Sepp, K., Cabral Pinheiro, V.H., Yang, M.-D., Vain, A. & Sepp, K. (2022). Toward Automated Machine Learning-Based Hyperspectral Image Analysis in Crop Yield and Biomass Estimation. Remote Sensing, 14(5), 1114.

Papers are reproduced by kind permission of the journals concerned.

Author's contributions to the articles:

	Article		
	I	II	III
Idea and study design	*	*	*
Data collection	*	*	*
Data analysis	*	*	*
Manuscript preparation	*	*	*

LIST OF ACRONYMS

AGL:	Above Ground Level
AMP:	Agricultural Management Practices
ANN:	Artificial Neural Network
ARC:	Agricultural Research Centre
Ard:	Automatic Relevance Determination
AUC:	Area Under the Curve
AutoML:	Automated Machine Learning
CASH:	Combined Algorithm Selection and Hyperparameter Optimization
CHM:	Canopy Height Models
CM:	Cultivation Method
CMin+:	Conventional Farming with Mineral Fertilizer Application
CORS:	Continuously Operating Reference Station
DM:	Dry Matter
DN:	Digital Number
DP:	Disking and Ploughing
EDA:	Exploratory Data Analysis
GDI:	Green Difference Index
GDVI:	Green Difference Vegetation Index
GIPVI:	Green Infrared Percentage Vegetation Index
GNDVI:	Green Normalized Difference Vegetation Index
GNSS:	Global Navigation Satellite System
GRDI:	Green Red Difference Index
GRVI:	Green Ration Vegetation Index
GSD:	Ground Sampling Distance
INS:	Inertial Navigation System
LAI:	Leaf Area Index
LDA:	Linear Discriminant Analysis
LOOCV:	Leave One Out Cross Validation
MA:	Manure Application
ML:	Machine Learning
MSR:	Modified Simple Ratio
MSRre:	Red-edge Modified Simple Ratio
MTVI:	Modified Triangular Vegetation Index
NDVI:	Normalized Difference Vegetation Index
NDVIre:	Red-edge Normalized Difference Vegetation Index

NIR:	Near-infrared Region
NRMSE:	Normalized Root Means Square Error
OMin+/-:	Organic farming with/without mineral fertilizer
OSAVI:	Optimized Soil-Adjusted Vegetation Index
P:	Ploughing
P+O:	Pea and Oat Mixture
PA:	Precision Agriculture
PCA:	Principal Component Analysis
PLS:	Partial Least Square
PPK:	Post Processed Kinematic
QP:	Quadratic Programming
R:	Reduced Tillage
R^2:	Coefficient of Determination
RC+G:	Red Clover with Grass Misture
RDVI:	Renormalized Difference Vegetation Index
REGVI:	Red-edge Greenness Vegetation Index
RFR:	Random Forest Regression
RINEX:	Receiver Independent Exchange
ROC:	Receiver Operating Characteristic Curve
RS:	Remote Sensing
RTVIcore:	Red-edge Triangular Vegetation Index
RVSI:	Red-edge Vegetation Stress Index
SB+RC:	Spring Barley with under-sowing Red Clover
SMAC:	Sequential Model-based Algorithm Configuration
SR:	Simple Ratio
SRre:	The Red-Edge Simple Ratio
STM:	Soil Tillage Method
SVR:	Support Vector Regression
SW:	Spring Wheat
UAS:	Unmanned Aerial Systems
UV:	Ultraviolet Area
VI:	Vegetation Indices
VPT:	Variety Performance Trials
WDRVI:	Wide Dynamic Range Vegetation Index

1. INTRODUCTION

My study began in Estonia in September 2018. Previously, as an agronomic in East Asia, the study was focused on the production and quality of rice (*Oryza sativa* L.), maize (*Zea mays* L.), and other tropical crops. However, the composition of agricultural activities and species in Estonia and other Nordic nations is somewhat different from that in Asia. As of the beginning of 2018, Estonia had not constructed a platform for comparing field-based phenotypes or crop breeding using the remote sensing (RS) platform. The supply of agriculture decision support systems is important as the precise and efficient interpretation of phenotypic data is crucial to the future development of precision agriculture. As well as enhancing crop fertilization, disease, pest, and weed identification in Estonia, all of which are in their infancy of development. The major objective of this research, from an agricultural standpoint, is to aid Estonia's agricultural research institutions in developing a reliable remote-sensed platform and delivering innovative real-time analytical methodologies to deal with climate change and food scarcity shocks in the future. In 2018, The study began by contacting Kuusiku Agricultural Center in Estonia to ascertain which comparison tests and crop categories were presently being conducted. According to Statistics Estonia of crop farming area (Ministry of Rural Affairs, 2021), the top three cultivated areas between 2014 and 2019 were mostly wheat, forage crops, and barley. Additionally, the Kuusiku Agricultural Center is conducting comparative experiments on these crops. As a result, the study concentrated on these critical crops first. Thus, under the major axis of sustainable agriculture development, the relevance of selecting these crop cultivation methods, monitoring techniques, and breeding systems are essential.

Legume-based systems, particularly red clover practices, are economically appealing to dairy producers in northern Europe and are critical for organic systems to compete with more conventional or artificially modified systems in terms of profitability. (Doyle & Topp, 2004). For example, in the majority of Northern Europe's nations, including Estonia, red clover (*Trifolium pratense* L.) is one of the dominant perennial forage crop legume species (Annicchiarico et al., 2015; Bender & Tamm, 2018). Legumes can boost grass pasture production by fixing atmospheric nitrogen into the soil through symbiotic rhizobia

located in their root nodules (Thilakarathna et al., 2016). The ability to fix atmospheric nitrogen makes red clover a suitable rotating crop, especially in organic farming systems that do not utilize synthetic nitrogen fertilizers (Vleugels et al., 2019). Numerous studies have shown that when red clover is seeded in combinations with grass species rather than in pure, monocultural stands, it is more effective at growing (Arturi et al., 2012; Zarza et al., 2020). On the other hand, according to data from the European statistical system – EUROSTAT (The European Commission 2017), the total area under cereal cultivation in the eight Baltic Sea-bordering European countries (Denmark, Estonia, Finland, Germany, Latvia, Lithuania, Poland, and Sweden) was 19 million hectares (ha) in 2016, with wheat (Triticum aestivum L.) accounting for 8 million ha, making it the region's most significant cereal (Chawade et al., 2018). Thus, breeding cultivars with well-characterized physiological characteristics and the potential to maximize grain output while allowing for modification of the nitrogen rate and agricultural approaches would be advantageous. This would therefore boost the ability to predict real nitrogen demand using a variety of techniques based on grain production and quality requirements (Chawade et al., 2018; Muñoz-Huerta et al., 2013).

As far as was known, trials and phenotyping methodologies, have come under greater scrutiny in recent years. One of the study's aims is to investigate and broaden the monitoring range beyond controlled conditions such as laboratories and greenhouses to circumstances involving bare soil fields (Araus & Cairns, 2014; G. Yang et al., 2017). As an alternative, variety performance trials (VPT) are a randomized controlled field-based experimental design used to enhance recommendations for environmental management scenarios for variety comparison and breeding selection (Laidig et al., 2014; Lollato et al., 2020). The recognition of various crop genotypes and their reaction to management approaches is a common strategy of VPT. With increased awareness of environmental preservation and the notion of sustainable agriculture, the applicability of eco-friendly farming practices such as decreased tillage and the use of a variety of minerals and organic fertilizers is becoming noticeable (Zhu-Barker & Steenwerth, 2018b).

Regardless of the weather, soil, or management circumstances in current experiments with rigorous model simulation, the sampling and model building challenges are increased by landscape variability (Frazier, 2015).

and a variety of patterns of spatial distribution for geographical objects (Ge et al., 2016). To address these issues, remote sensing (RS) technology enables the measurement of biophysical parameters at research places. Unmanned Aerial Systems (UAS) equipped with multifunctional sensors are regarded as a critical technology for the advancement of precision agriculture (PA) (Mulla, 2013) and sustainable smart farming (Tripicchio et al., 2015). They are frequently used to monitor cultivated areas, offering efficient solutions for accurate decision support, boosting agricultural efficiency, and profitability, minimizing environmental impact and stimulating future technological innovation (Herwitz et al., 2004; Mulla, 2013; C. Zhang & Kovacs, 2012a). UAS outfitted with a variety of unique sensor types have the potential to significantly increase the agreement and synergy between imaging and field reference data. Additionally, these systems may highlight regional monitoring needs, including disease detection, growth observation, yield calculation, and weed control (Tsouros et al., 2019; Xiang & Tian, 2011). Figure 1 illustrates how environmentally friendly sustainable agriculture principles work in conjunction with emerging remote sensing technology to affect agricultural management, productivity, and decision-making. On the other hand, a high spectral resolution imaging system (i.e., hyperspectral imaging) also creates the opportunity to enable increasingly sophisticated agricultural applications. The necessity for research in identifying optimum wavebands to predict crop biophysical characteristics is vital as hyperspectral remote sensing data becomes ever more available and significant (Monteiro et al., 2012; Xavier et al., 2006).

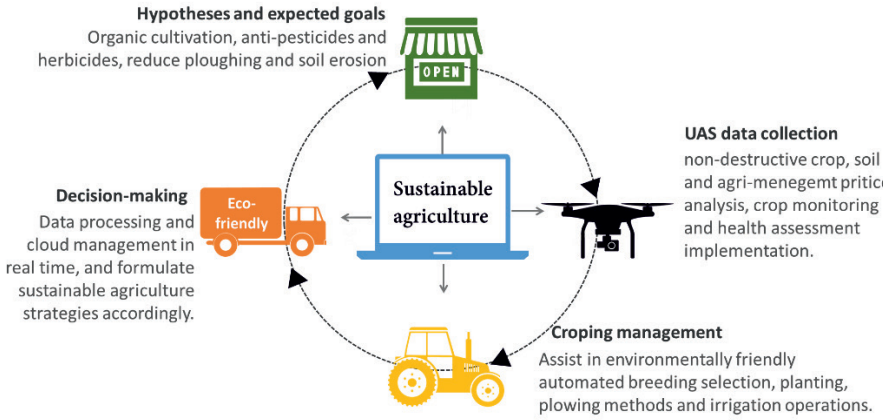

Figure 1. The application of remote sensing technologies under the concept of sustainable agriculture.

This thesis aims to examine how machine learning (ML) technologies have aided significant advancements in image analysis in the area of precision agriculture. These multimodal computing technologies extend the use of machine learning to a broader spectrum of data collecting and selection for the advancement of agricultural practices (Nawar et al., 2017) These techniques will assist complicated cropping systems with more informed decisions with less human intervention, and provide a scalable framework for incorporating expert knowledge of the precision agriculture (PA) system (Chlingaryan et al., 2018). Complexity, on the other hand, can be seen as a disadvantage in crop trials, as machine learning models require training/testing databases, limited areas with insignificant sampling sizes, time and space-specificity, and environmental factor interventions, all of which complicate parameter selection and make using a single empirical model for an entire region impractical. During the early stages of writing this thesis, this study used a relatively traditional machine learning method to address the regression problem of crop yield and biomass prediction [(i.e., random forest regression (RFR), support vector regression (SVR), and artificial neural network (ANN)] to predicted dry matter (DM) yields of red clover. It obtained favourable results, however, the selection of hyperparameters, the lengthy algorithms selection process, data cleaning, and redundant collinearity issues significantly limited the application of the machine learning techniques.

The thesis discusses the recent trend of automated machine learning (AutoML) that has been driving further significant technological innovation in the application of artificial intelligence from its automated algorithm selection and hyperparameter optimization of the deployable pipeline model for unravelling substance problems. However, a present knowledge gap exists in the integration of machine learning (ML) technology with UAS-multispectral and airborne-hyperspectral imaging data categorization and regression applications. In this thesis, a state-of-the-art (SOTA) and entirely open-source AutoML framework, Auto-sklearn was explored, which was built on one of the most frequently used machine learning systems, Scikit-learn. It was integrated with two unique AutoML visualization tools to examine the recognition and acceptance of multispectral vegetation indices (VI) data collected from UAS-multispectral and airborne-hyperspectral narrow-band VIs across a varied spectrum of agricultural management practices (AMP). These procedures incorporate soil tillage method (STM), cultivation method

(CM), and manure application (MA), and are classified as four-crop combination fields (i.e., red clover-grass mixture, spring wheat, pea-oat mixture, and spring barley). Additionally, they have not been thoroughly evaluated and lack characteristics that are accessible in agriculture remote sensing applications.

The aim of this thesis is to study existing gaps in the knowledge base for several critical crop categories and cultivation management methods referring to biomass and yield analysis, as well as to gain a better understanding of the potential for remotely sensed solutions to field-based and multifunctional platforms to meet precision agriculture demands. To overcome these knowledge gaps, this research introduces a rapid, non-destructive, and low-cost framework for field-based biomass and grain yield modelling, as well as the identification of agricultural management practices. The results may aid agronomists and farmers in establishing more accurate agricultural methods and in monitoring environmental conditions more effectively.

Paper **I** examined the variation in DM yields of a red clover-grass mixture across temporal periods (one- and two-year cultivated) and farming operations, utilizing three ML techniques RFR, SVR, and ANN and six multispectral VIs to predict DM yields. Paper **II** employed a SOTA and completely open-source AutoML framework, combined with two novel AutoML visualization tools to focus particularly on the recognition and adoption of UAS-derived multispectral vegetation indices (VI) data across a diverse range of agricultural management practices. Paper **III** explores the airborne hyperspectral system's extensive coverage, high spectral resolution, and varied narrow-band selection. It integrates open-sourced systems (R and Python) combined with automated hyperspectral narrowband vegetation index calculation and the robust AI-based AutoML technology to estimate yield and biomass for three crop categories (spring wheat, pea and oat mixture, and spring barley with red clover) in Estonia.

2. LITERATURE REVIEW

2.1. Development and challenges of forage and grain crops in Nordic countries and Estonia

In Estonia, experiments with red clover result in a mixed-species strategy, with other grass species blended to boost its commercial application value, where estimating complexity may be greater than in monocropping systems. Red clover's performance in pasture farming systems establishes it as a critical economically viable crop regardless of whether it is incorporated into conventional or organic farming operations. Despite the beneficial effects on agricultural productivity, legumes can support reducing greenhouse gas emissions by reducing the use of inorganic nitrogen fertilizers and substituting symbiotically fixed nitrogen, as well as by utilizing perennial grass species, which is a common practice, to reduce carbon loss in cultivated soil (Hanson & Ellis, 2020). This strategy increases the agricultural ecosystem's sustainability compared to monocropping systems and adds to the conservation value of vulnerable bumblebee species (Carvell et al., 2006). In Estonia, the cultivation of clover-grass mixtures has served significant agronomic purposes in co-cultivation increasing the feed value of the mixture, sequestering nitrogen and thereby reducing the amount of fertilizer (X. M. Yang et al., 2019), and achieving C-balance and carbon sequestration (Y. Yang et al., 2019) in crop rotation through perennial grassland.

It is important to note, however, that with the current trend toward global trade, the increasing importation of grain legumes into Europe has resulted in decreased domestic output in a number of countries (Godfray et al., 2010). This trend has prompted concerns about the long-term sustainability and security of protein supplies (Lüscher et al., 2014). Perhaps contrary to this, the main objective of red clover cultivation has been forage yields and persistence (Boelt et al., 2015), which might have a direct impact on company competitiveness and agricultural adaptability. This also emphasizes the critical significance of estimating and quantifying high-yield clover and grass combinations, in particular extending from the laboratory to field-based performance trials and studies. Traditional destructive silage and forage biomass sampling and measurements on-site give precise reference data for developing and evaluating yield models. However, it is time demanding,

labour intensive, and constrained by the gathering of large-scale special quantity parameters (Wachendorf et al., 2018).

2.2. Agricultural management strategies in variety performance trials

A common approach when identifying multiple crop management procedures and their interaction with the environment involves a well-conducted randomized experimental design, in which different agricultural management practices (AMP) are imposed on crops (Andrade et al., 2019). Variety performance trials (VPT) are a valuable method to address this issue. VPTs are regularly implemented in AMP research activities to improve the understanding of diverse systems and develop environmental management recommendations for variety selection (Laidig et al., 2014; Lollato et al., 2020). Concerning the AMPs trial criteria chosen and the recent growth of environmental protection awareness under the concepts of sustainable agriculture, the flexibility of environmental-friendly cultivation methods, such as reduced tillage and the application of various minerals and organic fertilizers, are being developed (Zhu-Barker & Steenwerth, 2018a). For example, tillage reduction is an essential characteristic of agricultural management that changes the soil either physically, chemically, mechanically, or biologically to create the appropriate conditions for seedling sprouting and healthy plant growth (DeLonge et al., 2014; Zhu-Barker & Steenwerth, 2018b), whereas organic additions such as manure or organic fertilizers are widely used methods to enhance soil fertility (Crews & Peoples, 2004a). Moreover, AMP is based on the concept of sustainable cropping ideas (such as reduced tillage intensity (Ashapure et al., 2019; Desta et al., 2021; Fanigliulo et al., 2020; Karlen et al., 2013; Telles et al., 2018; Triplett & Dick, 2008), fertilizer input (Crews & Peoples, 2004b), and organic farming (Crews & Peoples, 2004a; X. M. Yang et al., 2019; Zikeli et al., 2013) combined with mixed cropping systems, can effectively diminish greenhouse gas emissions by reducing the use of inorganic nitrogen fertilizers and replacing them with symbiotically fixed nitrogen, as well as carbon loss (Gianelle et al., 2009; Loide, 2019; Mandal et al., 2020) and soil erosion (Seitz et al., 2019) in cultivated soil. Studying VPT datasets, however, provides unique analysis problems due to their structure, nature, and husbandry variations in each trial. The evaluation of differences in management practices could potentially be confounded due to their nested structure (e.g. as opposed to controlled replicated

treatments) (Munaro et al., 2020). These AMPs have been increasingly proposed as an ecological method involving nutrient management, increased water holding capacity, and recoupled C and N cycling in agricultural ecosystems to improve sustainability (Drinkwater & Snapp, 2007; Gardner & Drinkwater, 2009). Although the specification of weather, soil, and management practices in current cropping systems are vital for robust model simulation and evaluation, these data are usually inaccessible for most cropping systems with adequate geospatial detail and lack of ability to replicate measured yields of field crops that received the best possible AMPs across a broad range of environments (van Ittersum et al., 2013).

In recent years, new developments in precision agriculture (PA) and the development of automated systems for agricultural resource management have been extensively studied and implemented (Pavón-Pulido et al., 2017). The emergence of these techniques seeks to boost crop growth and production, maximize profitability through empirical models and data assimilation, and make a substantial contribution to food security (Karthikeyan et al., 2020; Wen et al., 2021), agricultural disasters risk management (M. Der Yang et al., 2017), and more importantly, address concerns relating to climate change mitigation (Mandal et al., 2020). These challenges and opportunities have pushed remote sensing technology to the forefront. The use of RS for crop monitoring and trait decision-making provides cost-effective, non-destructive, and geographically comprehensive techniques. Additionally, when integrated with phenotypic feature modeling, RS methods may assist in yield prediction and support the determination and assessment of a diverse variety of plant characteristics (Costa et al., 2019). Its use in agriculture is especially critical for increasing the understanding of plant-environment interactions that occur during crop management. (Pieruschka & Schurr, 2019). Therefore, it is required to widen the scope of rapid and accurate RS methodologies for evaluating red clover-grass combination trials in response to a variety of agricultural practices and activities.

2.3. The significance of UAS-multispectral imaging and vegetation indexes in precision agricultural

Among the many remote sensing technologies, the UAS platform is considered one of the most significant technologies for the further development of PA (Mulla, 2013) and sustainable smart farming

(Tripicchio et al., 2015). UAS equipped with multifunctional sensors are frequently employed for the surveillance of cultivated lands, offering efficient solutions for accurate decision support, boosting agricultural efficiency, and profitability, minimizing environmental impact and stimulating future technological innovation (Herwitz et al., 2004; Mulla, 2013; C. Zhang & Kovacs, 2012a). They have developed into a low-cost remote sensing platform capable of acquiring high-resolution images (Mozgeris et al., 2018). The recent proliferation of sensors and cameras that can be incorporated into and mounted on these systems allows the identification and monitoring of plant changes both geographically and temporally for the detection and differentiation of local agricultural practices (Yeom et al., 2019a). For instance, the reflectance of vegetation data captured by UAS-borne sensors is verified by the biological and morphological characteristics of the tissues or leaves' surfaces (C. Zhang & Kovacs, 2012b). Depending upon the sensor types, the vegetation light spectra that may be captured can range from the ultraviolet area (UV), the visible region (RGB), down to the near-infrared region (NIR). UAS equipped with various novel sensor types can be exploited to improve agreement and synergy between imagery and field reference data. In addition, these systems can also identify the regional monitoring requirements, such as disease detection, growth observation, yield estimation, and weed management (Tsouros et al., 2019; Xiang & Tian, 2011).

From the application of UAS platform in precision agriculture, vegetation indices (VI) are one of the most often utilized outputs from UAS imaging applications. They aid in the supply of reliable spatial and temporal information for a wide range of agricultural operations with the ability to reduce soil or environmental noise and enhance their sensitivity for target characteristics (Wachendorf et al., 2018). VIs are typically mathematical combinations of individual or groups of electromagnetic spectrum bands and are meant to reduce the influence of external confounding variables while increasing the detectability of vegetative features (Raeva et al., 2019; Tsouros et al., 2019). Currently, UAS-based remote sensing techniques offer a notable contribution to field-based crop phenotyping investigations (Sankaran, Khot, & Carter, 2015). One of the most applied multispectral VI is the Normalized Difference Vegetation Index (NDVI) with its ratio between the red and near-infrared bands (Rouse et al., 1974). However, NDVI is not only sensitive to soil and atmospheric effects but also certain spectrum

ranges were found to have asymptotic relations as applicability is limited for higher biomass levels (Chao et al., 2019; Xue & Su, 2017). Also, the ability of the reflectance sensor in biomass prediction could be limited by ageing crop materials and diverse canopy structures caused by mixed species (Wachendorf et al., 2018). Therefore, an alternative for increasing the accuracy of various crop modeling tasks is by increasing the varieties and combinations of adjusted and optimized VIs (Osco et al., 2019).

Several UAS-based studies have been conducted in recent years. UAS-RGB-based vegetation indexes and linear regression models were utilized in estimating the red clover DM yield with the best performance R^2 value 0.62 (Lussem et al., 2018). UAS-RGB-based point cloud data generated into photogrammetric canopy height models (CHM) can also be utilized in forage legumes DM prediction; clover-grass canopies showed better performance than lucerne-grass mixtures for DM prediction (Grüner et al., 2019). Concerning another study combining CHM, RGB, and VIs with ML techniques for grass swards silage prediction, the Pearson correlation coefficients reached 0.98 (Viljanen et al., 2018a). Equally, clover related phenotypic research has also received much attention in recent years: It included clover-grass pasture coverage and spatial dynamics monitoring (Abuleil et al., 2015; Bonesmo et al., 2004), and quality parameters, such as the digestibility of organic matter, water-soluble carbohydrates, the nitrogen concentration, and uptake (Oliveira et al., 2020).

2.4. Application of hyperspectral imagery to agricultural yield and biomass estimation

On the other hand, multi-spectral, broadband-based remote sensing has had longstanding success in establishing correlations between conventional indices with yield and crop status. However, due to saturation in dense vegetation at larger leaf area index (LAI) values, multilayered canopies, and various farming systems, the calculated indices can occasionally produce inaccurate measurements and pose limits for quantitative estimation of biochemical properties owing to lower spectral resolution (Haboudane et al., 2004; Mutanga & Skidmore, 2004b; Sahoo et al., 2015; Zarco-Tejada et al., 2005). As an alternative technology, a high spectral resolution imaging system (i.e., hyperspectral imaging) creates the opportunity to enable increasingly sophisticated agricultural applications. The necessity for research in identifying

optimum wavebands to predict crop biophysical characteristics is vital as hyperspectral remote sensing data becomes ever more available and significant (Monteiro et al., 2012; Xavier et al., 2006). With the use of narrow spectral channels of less than 10 nm, hyperspectral remote sensing data has the potential to identify more nuanced differences in vegetation than multispectral data (Stagakis et al., 2010). It has been suggested that hyperspectral data analysis may present a format to provide a deeper understanding of the mechanisms governing spectral reflectance from field scales and canopy levels (Zarco-Tejada, 2000; Zarco-Tejada et al., 2001). These reduced-range channels allow for the detection of detailed plant and crop characteristics that would typically be obscured by broader-band multispectral channels. Innovative approaches for analysing spectral reflectance data are being established as a result of advances within hyperspectral remote sensing technology (Monteiro et al., 2012; Schmidt & Skidmore, 2003). Whilst hyperspectral sensors provide a more detailed depiction of plant canopy reflectance than more traditional multispectral sensors, they come with concerns regarding data redundancy and spectral autocorrelation (J. Feng et al., 2016; Thorp et al., 2017; M. der Yang, Huang, et al., 2020). In an attempt to redress and resolve these challenges, the reduction of data dimensionality is proposed, which can often be achieved via feature extraction, i.e., translating the spectra to a lower-dimensional representation, or selecting only a subset of essential bands or spectral characteristics for analysis (Bajcsy & Groves, 2004). One proposed technique to investigate imaging spectroscopy via spectral characteristics is to use application-specific optimal bands' combination, i.e., narrowband VIs. These narrowband VIs have significantly improved crop characteristics and deliver substantially advanced variability information with a superior dynamic range and considerable improvements over broad bands (Sahoo et al., 2015). There is mounting evidence that narrowband VIs can improve biomass estimations for many land-cover types (Heiskanen et al., 2013). Recently, a study regarding wheat grain yields also revealed that when compared to broadband VIs, hyperspectral indices provided greater estimation ability of grain production and biophysical factors (Xavier et al., 2006). As a result of the emergence of hyperspectral systems, there exists now the possibility to both refine previous spectral indices and build novel approaches that make use of the increased spectral resolution of hyperspectral data. Alternatively, the analysis might suggest that narrow-band, continuous reflectance data from a hyperspectral sensor

is preferred and potentially more accurate for certain remote sensing applications (Thorp et al., 2017).

2.5. The evolution and challenges of machine learning in agricultural remote sensing

In the subject of agricultural RS, machine learning techniques are commonly applied. In general, machine-learning systems are capable of modeling complicated class signatures, accepting a range of predictor data as input, and making no assumptions about the data distribution (i.e., nonparametric) (Maxwell et al., 2018). These incorporating multisensory computing science approaches provide a wide range of valuable information for the expansion of precision farming practices (Nawar et al., 2017). ML techniques may not provide a universal solution in precision farming; however, these approaches enable better determination in verisimilitude scenarios with minimum human intervention. They provide not only a powerful and flexible framework for decision-making but also facilitate the integration of expert knowledge into the PA system (Chlingaryan et al., 2018). For instance, when combined with hyperspectral imaging, ML has significantly improved crop biomass and yield estimation.(Changchun Li et al., 2020; Choudhury et al., 2021; Näsi et al., 2018).

Complexity, however, can be seen as a disadvantage in crop trials since the ML modelling includes training/testing databases, limited areas with insignificant sampling sizes, time and space-specificity, which raises problems in parameter selection and makes use of a single empirical model for an entire region impractical (Colombo et al., 2003; W. Zhang et al., 2019). Likewise, environmental factor interventions also enhance obstacles in parameter selection in ML systems owing to the differences in climate, and soil properties (W. Zhang et al., 2019). Occasionally, even the same crop genotypes may not express similar spectral characteristics in RS imaging, which renders the models invalid. If the reference parameters exist to formulate relationship functions, the genuine implementation results are frequently unsatisfactory owing to mismatches between concepts and realities.

Instead, the robust artificial intelligence-based notion of automated machine learning (AutoML) has emerged to minimize such data-driven expenses and enables experts to build self-regulating machine

learning applications (X. He et al., 2021; Mendoza et al., 2016). AutoML is characterized as a combination of selecting an algorithm and hyperparameter optimization based on the Bayesian optimization method that seeks to identify the optimum (cross-validated) combination of algorithm components by encompassing data from raw datasets to a deployable pipeline ML model, which greatly simplifies these stages for people with limited expertise (Feurer, Klein, et al., 2015a; Thornton et al., 2013; Yao et al., 2018). Recent advancements in AutoML systems such as Auto-WEKA (Thornton et al., 2013), and Auto-sklearn (Feurer, Klein, et al., 2015b) are recommended as an artificial intelligence-based solution for the expanding challenge of ML applications by combining a highly parametric ML framework with a Bayesian optimization method for a given dataset, significantly streamlines these steps for non-experts (Feurer, Klein, et al., 2015b). The standard procedure of ML modelling involves data pre-processing, feature engineering, feature extraction, feature selection, algorithm selection and hyperparameter optimization to increase the model's predictive performance (Remeseiro & Bolon-Canedo, 2019).

2.6. Knowledge gaps and prospects for research

A current gap persists in the knowledge base for multispectral-based AMP analysis and agriculture land use studies in addition to the further understanding of the potential for remotely-sensed solutions to field-based and multifunctional platforms for the demands of plant phenotyping and smart farming management. To address this knowledge gap, this study presents a rapid, non-destructive, low-cost framework for field-based crop yield and biomass modeling. To address this gap in knowledge, this thesis further employed a SOTA and completely open-source AutoML system, Auto-sklearn, which is constructed based on one of the most widely used ML system Scikit-learn in the scientific Python community (Komer et al., 2014), combined with two novel AutoML visualization tools to explore UAS and airborne -derived vegetation indices (VI) as an example for handling the AMPs classification tasks. Finally, an AutoML framework was constructed for hyperspectral imaging regression tasks, and used to explore the applicability of the AutoML models to estimate spring wheat, spring barley, pea and oat mixture grain yields and straw mass in regular mono- or mixed cropping systems in Northern Europe and Estonia.

3. HYPOTHESIS AND AIMS OF THE STUDY

Due to existing agricultural challenges, no platform for comparing field-based phenotypes or crop breeding had been developed using the RS platform. The absence of precise and effective interpretation of phenotypic data necessitates the development of agriculture decision support systems. The principle objective of this research, from an agricultural standpoint, is to assist a dependable remote-sensing platform and delivering innovative real-time analytical methodologies to deal with future climate change and food scarcity shocks. According to spatial resolution observability, UAS and airborne are suited as RS carriers in this investigation. More precisely, the central objective is to employ UAS-multispectral and airborne-hyperspectral imaging, with further employed a SOTA and entirely open-source AutoML system in conjunction with two innovative visualization tools to explore the yield prediction and cultivation management categorization skills of common crops in Estonia.

The following general objectives and hypotheses motivated this study:

1. To explore the relationship between crop period, location, and technique of cultivation, as well as the feasibility of using multispectral-UAS to estimate forage crop production and biomass.

Hypothesis 1: *UAS-multispectral is applicable to achieve accurate predictions in a variety of seasons, periods, and cultivation management*

2. To investigate the use of developing automated machine learning approaches to crop image analysis in developing effective regression and classification models.

Hypothesis 2: *Emerging automated learning approaches that combine classic machine learning methods can effectively solve regression and classification tasks of captured agricultural images.*

3. To investigate the capability of UAS-multispectral imaging to discover agricultural cultivation management strategies.

Hypothesis 3: *Applicable to recognize multiple agri-management categories of common crops in Estonia.*

4. To study the predictive potential of hyperspectral imaging and the synergistic benefits of combining it with AI-based auto-learning algorithms for crop production and monitoring.

Hypothesis 4: *By combining airborne hyperspectral images with a novel automated learning system, it is feasible to perform precise predictions.*

4. METHODOLOGY

4.1. The fundamental methodology of the thesis

The objective of this study was to establish a non-destructive and low-cost framework for biomass and yield modelling in several crop categories and cultivation management methods, as well as to gain a better understanding of the potential for remotely sensed solutions to field-based and multifunctional platforms to meet precision agriculture demands. As agriculture trial fields were often no larger than 3 by 9 meters. We conclude that the UAV equipped with a multispectral sensor andirborne sensor were best suited for the crop yield estimation and agricultural management practices (AMPs) recognition due to its superior velocity and spatial resolution over handheld sensors and satellite imagery. In particular, an unmanned aerial vehicle equipped with multispectral sensors was employed, as well as airborne hyperspectral imagery, to estimate yield and categorize fields by image processing, extraction, and various machine learning calculation procedures.

4.2. Study area and experiment layout

The study area of this thesis was undertaken at the Agricultural Research Centre (ARC) in Kuusiku (58°58'52.7"N 24°42'59.1"E), Estonia (Figure 2). The ARC was established in 1924 by an official institution under the governance of the Ministry of Agriculture and consists of consolidated laboratories and field testing centres. The experimental area covers 226 hectares, of which the 2.87-hectares variety performance trial (VPT) area was selected to consist of two soil types: Calcaric Cambisol and Calcari-Leptic Regosol (FAO, 2006). This experimental design was developed to facilitate the understanding of the physiological conditions and yield performance capabilities of the chosen varieties and their combinations under three types of AMPs. To assess the UAS-based AMP detection capacity, the experiment was put together with three principal experimental factors which include: (1) soil tillage methods (STM), considering reduced tillage (R) (8-10 cm), ploughing (P) at a depth traditionally used in conventional tillage (18-20 cm), and disking (DP) (8-10 cm) as treatments; (2) cultivation methods (CM), considering conventional farming with mineral fertilizer application (CMin+),

organic farming with mineral fertilizer application (OMin+), and organic farming without mineral fertilizer (OMin-); and (3) manure applications (MA) (Figure 2A and 2B). The ARC experimental area has a temperate climate with an average annual temperature of 5.3 °C, where the average daytime temperature was 9.5 °C, and 0.8 °C as the night temperature. The annual precipitation was 75 cm.

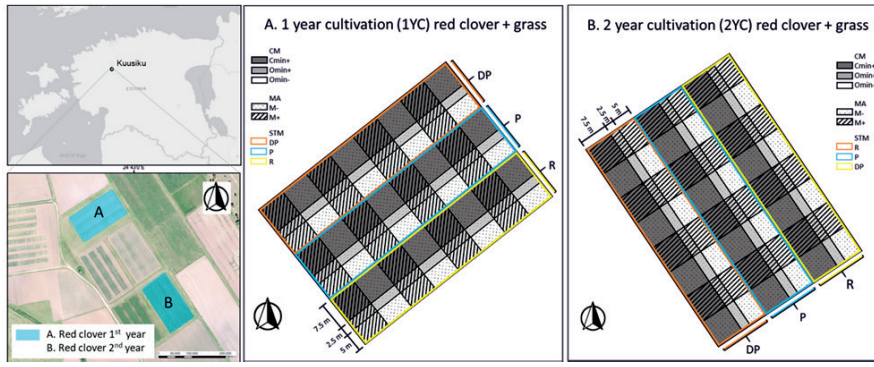

Figure 2. (a) The Agricultural Research Centre (ARC) is situated in Kuusiku, Estonia. A. The experiment layout contains three treatments, 1. Soil tillage method (STM) 2. cultivation method (CM), and manure application (MA) for one-year cultivation (1YC) in Field A with a total of 72 observation plots, and two-year cultivation (2YC) in Field B equally with a total of 72 plots. For caption descriptions, see Table 1. (b) A visual demonstration of the different CM treatments within the 2YC DP area.

Table 1. The farming operation and treatment of the red clover experiment fields.

Farming Operation	Treatment	Description
Soil tillage methods (STM)	Reduced tillage (R)	R (8-10 cm)
	Ploughing (P)	P (18-20 cm)
	Disking and ploughing (DP)	D (8-10 cm) & P (18-20 cm)
Cultivation methods (CM)	Conventional framing with fertilizer (Cmin+)	NPK 5-10-25 [1]
	Organic farming with mineral fertilizer (Omin+)	Patentkali [2]
	Organic farming without mineral fertilizer (Omin-)	N/A
Manure application (MA)	With manure application (M+)	M (30000 kg ha^{-1}) [3]
	Without manure application (M-)	N/A

[1] NPK 5-10-25 (chemical fertilizer) 291 kg ha^{-1} (N-14 kg ha^{-1}, P-13 kg ha^{-1}, and K-60 kg ha^{-1}),
[2] Patentkali (mineral fertilizer) 240 kg ha^{-1} (K-60 kg ha^{-1}, S- 41 kg ha^{-1}, M-14 kg ha^{-1})
[3] Manure 30000 kg ha-1 (N-234 kg ha^{-1}, P-20 kg ha^{-1}, and K-216 kg ha^{-1})

Although the same ARC experiment region was selected, the three papers covered distinct fields, which will be detailed as follows:

4.2.1 Red clover-grass mixture yield estimation

The variability, particularly in dry matter yields of a red clover-grass combination over temporal periods [(one- and two-year farmed (1YC and 2YC)] and farming practices was examined. The mixture's fresh aboveground biomass was cut twice in two fields. Each field contains 72 plots, which means in a total 144 plots were sampled; the first cut took place on 10/06/2019, and the second took place on 16/08/2019. The fresh biomass was weighed by plot and dried to verify its DM yield measured in kilograms per hectare.

4.2.2 Agricultural management practices recognition

The primary focus was particularly on the recognition and adoption of UAS-derived multispectral vegetation indices (VI) data across a diverse range of agricultural management practices. Four different crop categories were selected to enrich the diversity of crop identification. The experimental layout consists of four types of common crop and their regular combinations in Estonia, i.e., Field 1: red clover 75% (Trifolium pratense L.) with grass 25% (Festuca pratensis) (RC+G). Field 2: spring wheat (SW), Field 3: pea and oat mixture (P+O), and Field 4: spring barley with under-sowing red clover (SB+RC) in 2019. Each field comprises 72 plots, which amounts to a total of 288 plots sampled within the study area (Figure 3).

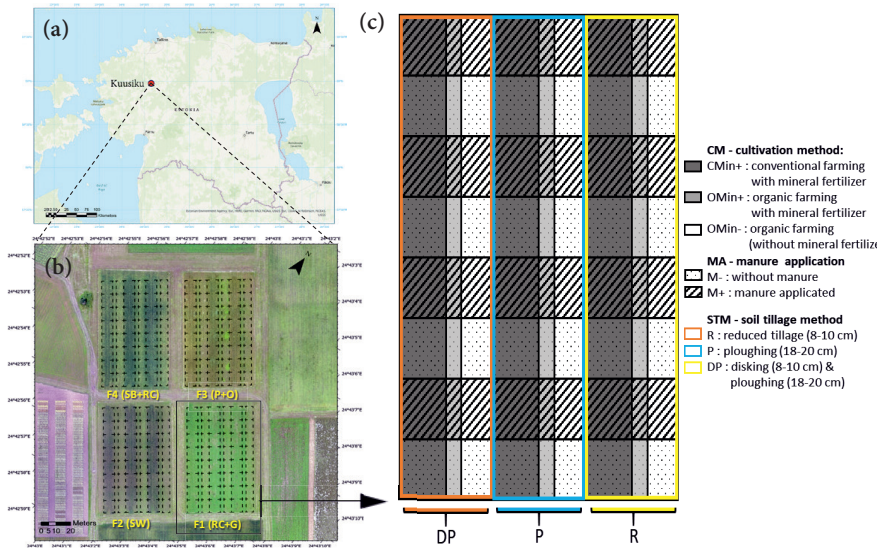

Figure 3. (a) The study area is located at Kuusiku agriculture centre, Estonia. (b) The RGB orthomosaic image from May 30th of the experimental layout fields with four crop types i.e. [F1. (RC+G), F2. (SW), Field 3. (PO), and Field 4. (SB+RC)]

4.2.3 Hyperspectral image analysis in crop yield and biomass estimation

The airborne hyperspectral system's extensive coverage, high spectral resolution, and varied narrow-band selection to estimate yield and biomass for three crop categories (spring wheat, pea and oat mixture, and spring barley with red clover) in Estonia were explored. The experimental fields consist of three commonly cultivated crop categories and their regular cropping combinations in Estonia (Figure 4b), i.e., Field 1: spring wheat (SW) (Figure 4c), as representative of the uniform variety planting field; Field 2: pea and oat mixture (P+O), and Field 3: spring barley with under-sowing red clover (SB+RC) (Figure 4d) as representative of the mixed planting fields. All three fields are part of common crop rotation with a spatial and temporal arrangement (Figure 4).

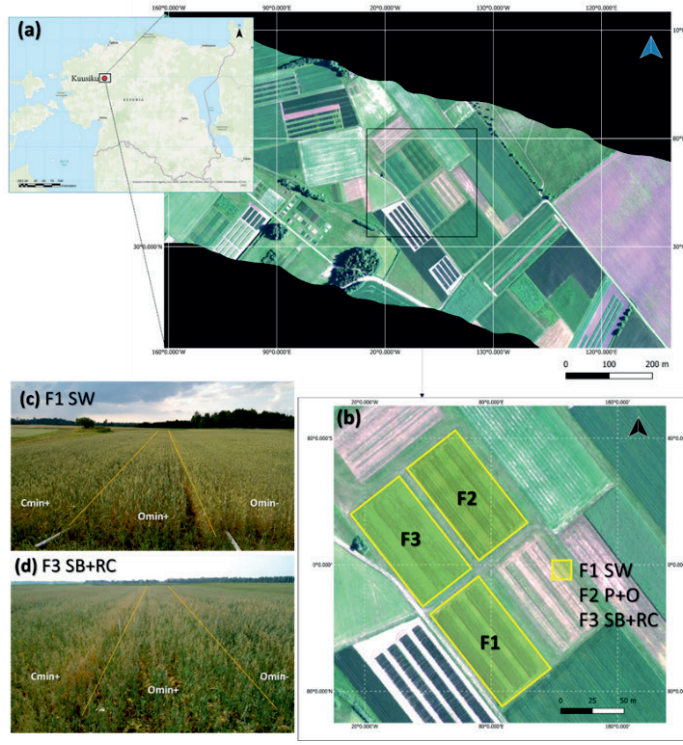

Figure 4. Airborne push-broom hyperspectral image in the Agricultural Research Centre (ARC), Kuusiku, Estonia. (a) Hyperspectral image with the band combination: band 83 (630 nm), band 47 (532 nm), and band 22 (465 nm) light in. (b) The experiment fields of this study, where Field 1 (F1): spring wheat (SW), Field 2 (F2): pea and oat mixture (P+O), and Field 3 (F3): spring barley with under-sowing red clover (SB+RC). The interpretation diagrams represent on-site (c) single variety planting SW, and (d) mixed planting SB+RC.

Figure 5 shows the AMPs and their specific arrangement in SW, P+O, and SB+RC fields. Every field comprised 72 plots, with a total of 216 plots. Based on considerations of budget limitations, labour shortages, excessive scope, and repetitiveness, the sampling of grain yield was taken from 56 out of 72 plots (n = 56), and straw biomass were sampled from 24 out of 72 plots (n = 24) specific from the disking and ploughing (DP) area (Figure 5). The harvesting took place on 5 August 2019 in field SB+RC and on 16 August 2019 in fields SW and P+O. The fresh grain and biomass were weighed by plot and dried to verify its dry grain yield and fresh straw mass measured in kilograms per hectare. However, regarding the mixture P+O field, the total weight of the two crops was calculated, while in the SB+RC field only the SB grain yield and straw mass.

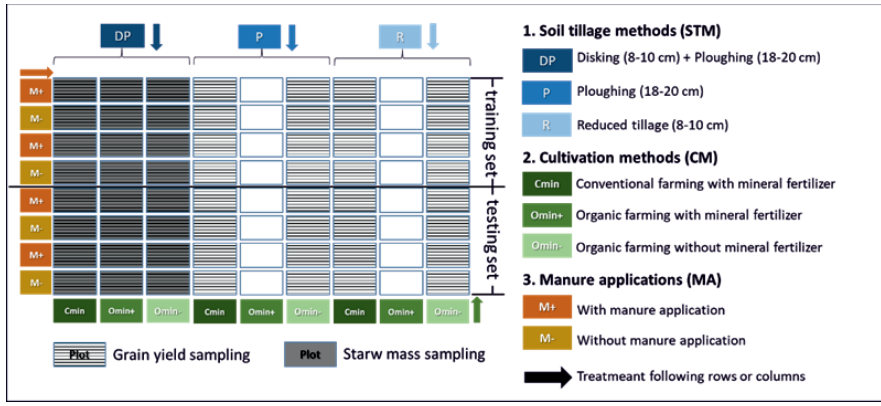

Figure 5. The structure of agriculture management practices (AMPs) and the sampling method of grain yield and straw mass in the SW, P+O, and SB+RC fields. The AMPs contain three treatments: 1. soil tillage method (STM), 2. cultivation method (CM), and manure application (MA), where the grain yield (n = 56) (black striped rectangle box) and straw mass (n = 24) (grey rectangle box). To guarantee that the training area contained all combinations of AMPs, each field was split into training and testing areas equally from the centre. The special arrangements of AMP categories and the sampling method were the same in the 3 fields.

4.3. UAS-multispectral and hyperspectral image acquisition

4.3.1 UAS-multispectral Image acquisition for red clover-grass Mixture fields

Figure 6 presents a workflow of the methodology used to combine the UAS-based image collection, processing, and biomass sampling. To capture data for both image processing and biomass evaluations, the UAS imaging was conducted twice [i.e., 11 days before 1st cut (11DB) and 38 days before 2nd cut (38DB) harvesting] in the summer of 2019. Due to the needs of the other experimental areas, data of 80 hectares were collected, of which 2.4 hectares were used in this study. An eBee Plus device (senseFly, 2016), with onboard GNSS post-processed kinematic (PPK) capabilities, was deployed and equipped with a Parrot Sequoia multispectral sensor. The Parrot Sequoia© (Parrot S.A., Paris, France) sensor captured imagery across four spectral bands: near-infrared (770–810 nm); red-edge (730–740 nm); red (640–680 nm); and green (530–570 nm). The flight lines overlap was set with a frontal image overlap of 80% and lateral image overlap of 75%. All the operations took place between 10 a.m. to 2 p.m. to ensure consistency with the sun's angle,

and to reduce lateral shading within the experimental fields. The images were captured from a height of 120 meters, and the resulting images had ground sampling distance (GSD) of 10 cm per pixel. Prior to each flight mission, an Airinov radiometric calibration target and one-point calibration method (Poncet et al., 2019) was used to facilitate post-flight radiometric correction of the multispectral imagery.

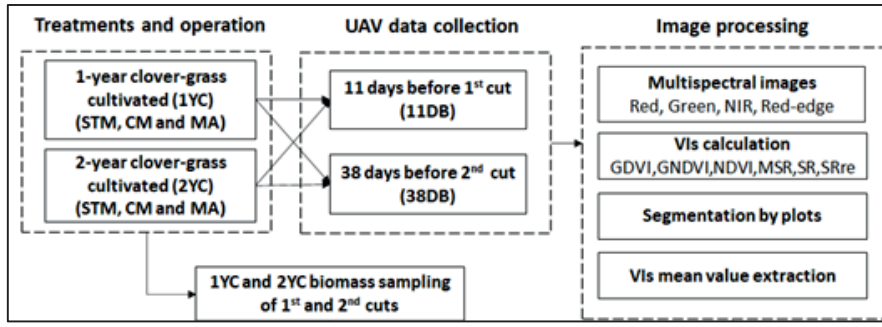

Figure 6. The methodology flowchart of red clover-grass mixture UAV data collection and processing. The image processing rectangular dotted box contains all predictors extracted from the UAV images.

4.3.2 UAS -multispectral Image Processing and Analysis

The UAS-multispectral data was post-processed in SenseFly eMotion 3 (senseFly, 2016) using receiver independent exchange (RINEX) format data provided by the GNSS CORS (Continuously Operating Reference Station) of Estonia (Land Board, Republic of Estonia, 2018) for post-processing kinematics (PPK) corrections. This post-process provided an increase in the geotagging accuracy (Mokroš et al., 2019) of the UAS images from 5 m error to under 0.06 m, where the method and accuracy obtained are similar to (Villoslada et al., 2020) ; and thus less than the one-pixel size in this study. Pix4D v.4.3.31® (Pix4D SA, 1015 Lausanne, Switzerland) software was utilized to process and radiometrically correct (default in Pix4D) the imagery and generate the multispectral orthomosaics. These images were subsequently clipped to represent only the extent of the experimental area.

4.3.3 UAS-multispectral Image acquisition for Agricultural Management Practices Recognition

Figure 7 shows the workflow utilized to combine the UAS-based image collection, processing, sampling, and AutoML framework modified from (Feurer, Klein, et al., 2015b). A fixed-wing UAS eBee Plus (Sensefly Inc., Cheseaux–Lausane, Switzerland) equipped with GNSS PPK capabilities was deployed with a Parrot Sequoia multispectral sensor (version 1.2.1, Parrot, Paris, France). To facilitate seasonal image processing and AMP recognition, UAS images were captured over three timeslots in 2019 at the Kuusiku Research Center: April 23rd (temperature: 16°C, wind speed: 11 km h^{-1} S, sunny), May 30th (temperature: 19°C, wind speed: 12 km h^{-1} WSW, overcast), and July 10th (temperature 20°C, wind speed: 3.6 km h-1 NW, sun with minor cloud cover). The originally designed flight time was 37 minutes and 30 seconds per task over an area of 65.8 hectares (with areas of interest 2.87 hectares in this study). However, depending on the weather conditions and wind speed of the day, the eBee flight time might be slightly different from the number of battery replacements (the endurance of one battery was approximately 20-30 minutes). This data capture protocol was designed to represent the reflectance spectrum characteristics of crops during different growth stages. Flight-line overlap was set using a frontal image overlap of 80% and a lateral overlap of 75% with a target altitude of 120 m above ground level (AGL), resulting in a GSD of 10 cm per pixel. All image data capture procedures were undertaken between the hours of 10 a.m. to 2 p.m. to guarantee the consistency of photo collection quality, and to minimise lateral shading of crops within the VPT fields. An Airinov radiometric calibration target (Airinov, Paris, France) and a one-point calibration method (Poncet et al., 2019) were used to enable post-flight radiometric correction of the multispectral imagery before each flight to remove dark current and lens vignetting effects while postprocessing the image (Kelcey & Lucieer, 2012).

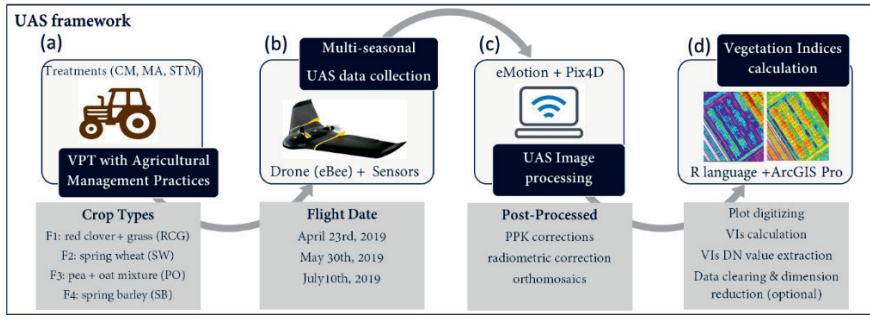

Figure 7. The flowchart of UAS framework in the classification task, where (a) Three types of AMPs are processed for four crop categories. (b) The eBee plus with Parrot Sequoia multispectral sensor with the time series flight (April, May, and July) to collect spectral information from different crop periods. (c) UAS image was post-processed in SenseFly eMotion with PPK corrections and orthomosaics in Pix4D. (d) 19 VIs calculation, segmentation and corresponding plot digital number (DN) extraction for AutoML modelling.

4.3.4 UAS-multispectral Image Processing and Analysis

For pre-processing UAS-multispectral images, SenseFly eMotion 3 applying differential correction data (RINEX) provided by the GNSS CORS (Continuously Operating Reference Station) of Estonia for post-PPK corrections (Land Board, Republic of Estonia, 2018) was used. PPK was reported to increase the higher horizontal and vertical geotagging accuracy when compared to ground control points (GCP) (Mokroš et al., 2019). In this study, the UAS image corrections were decreased from 5 m error to under 0.06 m (less than one-pixel size). Pix4D v.4.3.31® (Pix4D SA, 1015 Lausanne, Switzerland) software was utilized to process and radiometrically correct (calibrated according to the variances between the measured value and target actual reflectance (Poncet et al., 2019)) the imagery, as well as to generate the multispectral orthomosaics. These images were subsequently clipped with a one-metre inward buffer zone from each plot to represent only the extent of the area of the VPTs.

4.3.5 Hyperspectral image data collection

Airborne measurements were carried out in Kuusiku Agricultural Research Centre on 18 June 2019 using hyperspectral imager HySpex [Norsk Elektro Optikk AS (NEO), Norway] owned by Estonian Marine

Institute and operated by the Estonian Land Board. HySpex was flown at an altitude of 900 m which resulted in a spatial resolution of 40 cm (Figure 4a). The spectral resolution of HySpex is approximately 2.69 nm (216 spectral bands ranging from visible to near-infrared with centres between 409 nm and 989 nm). The day was sunny with a wind speed of 2.6 m/s, average air temperature of 10°C. Regarding the growth stages of the main crops on the flight date, spring wheat, spring barley, and oat were approximately in booting to heading stage. For the mixed crops, i.e., field pea and red clover were in the reproductive growth stages, and the flowing stage, respectively.

Raw HySpex image data were converted into units of spectral radiance ($W\ m^{-2}\ nm^{-1}\ sr^{-1}$) using Rad software developed by the NEO. PARGE (Parametric Geocoding, ReSe Applications Schäpfler, University of Zurich) geo-coding software was used for geo correction of the flight lines utilizing accurate altitude and location measurements provided by the GNSS/INS unit. The captured Hyspex flight line used in this study is shown in Figure 4a. Atmospheric influence at such a low altitude was considered minimal and therefore atmospheric correction was not applied to the imagery.

4.3.6 Hyperspectral Image Processing

Most hyperspectral processing techniques now employ commercial software such as Erdas Imagine, ENVI, or the MATLAB hyperspectral toolbox (The Mathworks Inc., 2019). These technologies are often expensive and can have limited statistical analysis capabilities. Therefore, a new package that was built on the open-source software R in 2019 was employed. The hyperspectral data analysis (*Hsdar*) package incorporates several important hyperspectral capabilities from the HyperSpec package (Beleites & Sergo, 2012), with an emphasis on the analysis of large data sets collected in the field for vegetation remote sensing. It is available at https://CRAN.R-project.org/package=hsdar on the Comprehensive R Archive Network (CRAN).

Hyperspectral data was reconstructed into a class named 'Speclib' to offer a framework for handling huge sets in R. This allows the user to store three-dimensional (3D) cube data together with extra adding information into a matrix. This matrix, together with the wavelength information can then be utilized in the *Hsdar* software and used to manage subsequent

calculations. A Savitz-ky-Golay filter (method „*sgolay*") with a length of 15 nm was used in the initial preprocessing stage to reduce noise from the spectra. By fitting a polynomial function to the reflectance data, the filter minimizes noise and removes minor discrepancies between adjacent bands. These noise-reduced hyperspectral data were calculated zonal statistics and converted to a [216 (wavelength bands) multiplied by 216 (plot Shapefile)] table. This table was then subsequently used for preliminary correlation analysis between grain yield and straw mass with the mean wavelength reflectance value into plot level. The correlation analysis results of each narrowband band can be utilized as a consideration in the following selection of narrowband vegetation indexes.

4.4. Vegetation Indices selection and calculation

4.4.1 Vegetation Indices selection and calculation in Red clover-grass mixture yield estimation

Six VIs were calculated using R version 4.0.2 (R Core Team, 2020) (Table 2). The normalized difference vegetation index (NDVI) utilizes the reflectance (ϱ) in the NIR and Red wavelengths, and the outputs range from −1.0 to 1.0. This index was selected for this study as it has a sensitive response to tracking physiological dynamics and biomass (Hassan et al., 2019). However, NDVI reaches saturation when leaf area index (LAI) values are about 2.5–3 or in dense crop canopies [42,43]. The green normalized difference vegetation index (GNDVI) was also calculated and outputs values range from 0 to 1. Previous studies have shown GNDVI to be linearly correlated with LAI and biomass, with the ability to reduce the effects of soil reflectance and estimate nitrogen conditions (Hunt et al., 2008). Similarly, the Simple Ratio (SR), a normalization of ϱ NIR against ϱ Red, was calculated as it has been previously shown that this index can better indicate the strength of canopy photosynthetic material and yield prediction than NDVI under different nitrogen supplies (Serrano et al., 2000). The Red-Edge Simple Ratio (SRre) formula was calculated by replacing the ϱ Red band with the ϱ Red-edge. Its inclusion in the assessment was due to previous studies indicating a higher correlation with plant nitrogen concentration compared to ϱ Red based VIs. This can lessen the soil background influence on crop reflectance (Walsh et al., 2018). Finally, the Modified Simple Ratio (MSR), as a potentially improved version of

the Renormalized Difference Vegetation Index (RDVI), was calculated to linearize the relationship between biophysical parameters (J. M. Chen, 1996) and enhance the sensitivity of vegetation occurrences which can be observed in other VIs. The two experimental fields [i.e., 1YC (n = 72) and 2YC (n = 72)], with a total of 144 plots were digitized in ArcGIS Pro 2.6.3 (ESRI, 2016). The average VIs within each plot were extracted and calculated as the VIs of each plot at the experiment site. To avoid potential edge effects in the fertilizer treatment, a one-meter buffer zone was extended inwards from each plot boundary, and data sampled within this target region (Figure 8). These extracted values were further used in this study when building ML algorithms for clover-grass mixture DM yield estimation and evaluation.

Table 2. Descriptions and formulas of NIR related VIs used in this study.

Vegetation Index	Description	Equation	Reference
NDVI	Normalized Difference Vegetation Index	(ϱ NIR – ϱ R 1) / (ϱ NIR + ϱ R)	(Rouse et al., 1974)
GNDVI	Green Normalized Difference Vegetation Index	(ϱ NIR - ϱ G 2)/ (ϱ NIR + ϱ G)	(A. A. Gitelson et al., 1996)
GDVI	Green Difference Vegetation Index	ϱ NIR 3 - ϱ G	(Sripada et al., 2006)
SR	Simple Ratio	ϱ NIR / ϱ R	(Jordan, 1969)
SRre	Red-edge simple ratio	ϱ NIR / ϱ REG 4	(A. Gitelson & Merzlyak, 1994b)
MSR	Modified simple ratio	((ϱ NIR- ϱ R)-1)/(((ϱ NIR+ ϱ R)*(.5))+1)	(J. M. Chen, 1996)

[1] ϱ R refers to red band, [2] ϱ G refers to green band, [3] ϱ NIR refers to near-infrared, and [4] ϱ REG refers to the red edge.

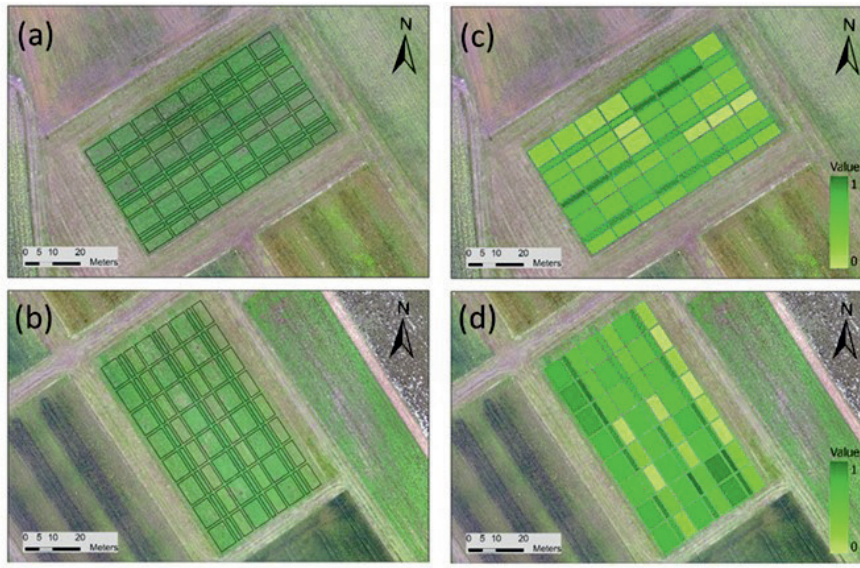

Figure 8. A demonstration of VIs (GDVI as an example) zonal statistics in 1YC and 2YC fields. (a) RGB image with 1-meter buffer zone plot polygons from 1YC11DB, (b) RGB image with 1-meter buffer zone plot polygons from 2YC11DB, (C) GDVI zonal statistic with ROI in 1YC11DB, and (d) GDVI zonal statistics with ROI from 2YC11DB.

4.4.2 Vegetation Indices selection and calculation in AMP Recognition

Nineteen VIs were chosen and calculated to address the issues of heterogeneous crop classes, soil types, and the current absence of valuable referenced parameters in AMPs (see Table 3). More specifically, Datt4, SRre, NDVIre were selected due to their positive correlation with chlorophyll content (Dong et al., 2015; A. Gitelson & Merzlyak, 1994b; J. Zhang et al., 2014); MTVI, MSR, MSRre, RVIS, WDRVI (J. M. Chen, 1996; Haboudane et al., 2004; Henebry et al., 2004; Merton & Huntington, 1999; C. Wu et al., 2008) are known to be sensitive to variations in leaf area index (LAI); GDVI was used for better lower vegetal land cover estimates and characterization (W. Wu, 2014); GIPVI was calculated for its potential in grassland communities detection (Strong et al., 2017); GNDVI, NDVI, RTVIcore were utilized due to their high performance in crop above-ground biomass (AGB) estimation (P. F. Chen et al., 2010; Kross et al., 2015); GDI, GRDI, and RDVI were included due to their ability to compensate for NDVI saturation problems and the potential effects of soil and sun viewing geometry

(Mutanga & Skidmore, 2004b; Vasudevan et al., 2016); GRVI was applied for its sensitivity to soil moisture (Ballester et al., 2019), SR for strongly correlated with comprehensive growth index (CGI) (H. Feng et al., n.d.) and REGVI was included for its sensitivity to deviations in senescence and vegetation stress (Cross et al., 2019).

Table 3. Descriptions and formulas of multispectral UAS derived VIs used in this study. The *ϱ R* refers to the reflectance of the red band, *ϱ G* refers to the reflectance of the green band, *ϱ REG* refers to the reflectance of the red edge, and *ϱ NIR* refers to the reflectance of the near-infrared.

Vegetation Index	Equation	Reference
Datt4	ϱ R /(ϱ G * ϱ REG)	(Datt, 1998)
Green Infrared Percentage Vegetation Index (GIPVI)	ϱ NIR /(ϱ NIR + ϱ G)	(Crippen, 1990)
Green Normalized Difference Vegetation Index (GNDVI)	(ϱ NIR - ϱ G)/(NIR+ ϱ G)	(A. A. Gitelson et al., 1996)
Green Difference Vegetation Index (GDVI)	ϱ NIR - ϱ G	(Sripada et al., 2006)
Green Ration Vegetation Index (GRVI)	ϱ NIR / ϱ G	(Sripada et al., 2006)
Green Difference Index (GDI)	ϱ NIR - ϱ R + ϱ G	(Gianelle & Vescovo, 2007)
Green Red Difference Index (GRDI)	(ϱ G - ϱ R)/(ϱ G + ϱ R)	(Gianelle & Vescovo, 2007)
Normalized Difference Vegetation Index (NDVI)	(ϱ NIR - ϱ R)/(ϱ NIR + ϱ R)	(Rouse et al., 1974)
Red-edge Normalized Difference Vegetation Index (NDVIre)	(ϱ NIR - ϱ REG)/(ϱ NIR + ϱ REG)	(A. Gitelson & Merzlyak, 1994b)
Red-edge Simple Ratio (SRre)	ϱ NIR / ϱ REG	(A. Gitelson & Merzlyak, 1994b)
Renormalized Difference Vegetation Index (RDVI)	((ϱ NIR - ϱ R)/((ϱ NIR + ϱ R)**(.5)))	(Roujean & Breon, 1995)
Red-edge Modified Simple Ratio (MSRre)	((ϱ NIR - ϱ REG)-1)/(((ϱ NIR + ϱ REG)**(0.5))+1)	(C. Wu et al., 2008)
Red-edge Triangular Vegetation Index (RTVIcore)	(100*(ϱ NIR - ϱ REG))-(10*(ϱ NIR - ϱ G))	(P. F. Chen et al., 2010)
Red-edge Vegetation Stress Index (RVSI)	((ϱ R + ϱ NIR)/2)- ϱ REG	(Merton & Huntington, 1999)

Red-edge Greenness Vegetation Index (REGVI)	(ϱ REG - ϱ G)/(ϱ REG + ϱ G)	(Sims & Gamon, 2002)
Simple Ratio (SR)	ϱ NIR / ϱ R	(Jordan, 1969)
Modified Simple Ratio (MSR)	((ϱ NIR - ϱ R)-1)/(((NIR+ ϱ R)**(.5))+1)	(J. M. Chen, 1996)]
Modified Triangular Vegetation Index (MTVI)	1.2*((1.2*(ϱ NIR - ϱ G))-(2.5*(ϱ R - ϱ G)))	(Haboudane et al., 2004)
Wide Dynamic Range Vegetation Index (WDRVI)	(((0.2* ϱ NIR)- ϱ R)/((0.2* ϱ NIR)+ ϱ R))	(A. A. Gitelson, 2004)

4.4.3 VI extraction and principal component analysis

Principal component analysis (PCA) (Kambhatla & Leen, 1997) was employed as an exploratory data analysis (EDA) technique to describe the relationship between three different agricultural management types (CM, MA, and STM) and multispectral UAS-VIs. The PCA was used for testing whether or not it could improve the classification efficiency of AMPs. PCA was conducted using R version 4.0.2 (R Core Team (2020), 2020) and the *FactoMine*R package (Lê et al., 2008). For extraction of the digital number (DN) values from each VIs of four experimental fields (72 plots in each field), a total of 288 plots were digitized in ArcGIS Pro 2.6.3 (ESRI, 2016). As stated previously, a one-meter buffer zone was extended inwards from each plot boundary, to address potential edge effects from agricultural management, the average VIs were isolated and calculated. These extracted values were further used in this study when building ML algorithms and for AutoML assessment and evaluation.

4.4.4 Narrowband vegetation indices selection and calculation from hyperspectral imagery

Optical indices for chlorophyll estimation studies have focused on analyzing reflectance in specific narrow bands, ratios, combinations, and the properties of derivative spectra to minimize extraneous factor changes and increase sensitivity to chlorophyll content (Haboudane et al., 2002). In this study, VIs that were sensitive to canopy structure, biochemistry, and physiology, and those that might potentially indicate variance in grain yields and biomass were targeted. Pigments (i.e., chlorophyll a,

chlorophyll b, and carotenoids) exhibit varied spectral behaviour from an optical standpoint, with specific absorption properties at different wavelengths (Blackburn, 1998). Therefore, pre-defined indices in the *Hsdar* R package were deployed to automatically fit provided wavelength positions and compute corresponding VIs to reduce the intricacy of computation and boost the repeatability of this research (Table 4).

The Normalized Difference Vegetation Index (NDVI) was adopted based on it is sensitivity to green leaf area or green leaf biomass, and it can be used to monitor photosynthetically active vegetation biomass distribution using linear combinations of red and infrared radiances (Tucker, 1979). However, it is crucial to note that NDVI has a saturation effect at richer vegetation covers (Fernández-Manso et al., 2016). To solve the probable saturation problem, NDVI2 was applied with its ability to adequately determine chlorophyll in the presence of a high pigment concentration background (A. Gitelson & Merzlyak, 1994a). The renormalized difference vegetation index (RDVI) narrow band, was employed in this study due to its capacity in identifying mixture phytomass in grassland (Vescovo et al., 2012). The prospect for using the Transformed Chlorophyll Absorption in Reflectance Index (TCARI) in an operational remote sensing situation in the context of precision agriculture was investigated. The R_{700}/R_{670} ratio was chosen to reduce the combined impacts of underlying soil reflectance and non-photosynthetic materials. The changes in reflectance characteristics of background materials (soil and non-photosynthetic components) and the R_{700}/R_{550} ratio are strongly connected to differences in background materials (Haboudane et al., 2002; M. S. Kim et al., 1994). Soil-Adjusted Vegetation Index (SAVI) was conducted to reduce soil-induced fluctuations in vegetations using a transformation approach to decrease soil brightness impacts by counting red and near-infrared wavelengths from spectral data (Huete, 1988). Where Optimized Soil-Adjusted Vegetation Index (OSAVI) with two types of reflectance combinations (OSAVI and OSAVI2) was selected for its simplicity of use in the context of deployable observations on agricultural landscapes, as its estimation requires no knowledge of soil optical properties, and it also provided the best results for most crops (Rondeaux et al., 1996), as well as the distinction of tillage effects in an economically RGB UAV application (Yeom et al., 2019b). In addition, the choice of Simple Ratio (SR) narrow-band indices (R_{515}/R_{550}), different from chlorophyll pigment content detection, is based on its feasibility to predict carotene content on

hyperspectral imagery in heterogeneous canopies (Hernández-Clemente et al., 2012). Carotenoid concentrations reveal important information about plant physiological state (Demmig-Adams & Adams, 1992), and offering a heterogeneous VI source may improve model predictability and minimize collinearity.

Table 4. Descriptions and formulae of narrowband VIs were utilized in this study. Narrowband VIs were calculated which were closest to the wavelengths given in the original *Hsdar* R package references.

Vegetation Index	Description	Equation	Reference
NDVI	Normalized Difference Vegetation Index	$(R_{800}-R_{680}) / (R_{800}+R_{680})$	(Tucker, 1979)
NDVI2	Normalized Difference Vegetation Index 2	$(R_{750}-R_{705})/(R_{750}+R_{705})$	(A. Gitelson & Merzlyak, 1994a)
OSAVI	Optimized Soil Adjusted Vegetation Index	$(1+0.16)*(R_{800}-R_{670})/(R_{800}+R_{670}+0.16)$	(Rondeaux et al., 1996)
OSAVI2	Optimized Soil Adjusted Vegetation Index 2	$(1+0.16)*(R_{750}-R_{705})/(R_{750}+R_{705}+0.16)$	(C. Wu et al., 2008)
RDVI	Renormalized Difference Vegetation Index	$(R_{800}-R_{670})/\sqrt{(R_{800}+R_{670})}$	(Roujean & Breon, 1995)
SR	Simple Ratio	R_{515}/R_{550}	(Hernández-Clemente et al., 2012)
SAVI	Soil-Adjusted Vegetation Index	$(1+L^{1})*(R_{800}-R_{670})/(R_{800}+R_{670}+L)$	(Huete, 1988)
TCARI	Transformed Chlorophyll Absorption Reflectance Index	$((R_{700}-R_{670})-0.2*(R_{700}-R_{550}))*$	(Haboudane et al., 2002)

[1] L, a soil brightness adjustment factor (L) established as 0.5 to suit the majority of land cover types for the SAVI index.

These narrowband VIs were computed and saved in TIFF file format (https://www.adobe.io/open/standards/TIFF.html), which were then utilized to extract spatial information in the SW, P+O, and SB+RC experimental fields. For extraction, a total of 216 plots were digitized in

ArcGIS Pro (ESRI, 2016). Average VIs across every plot were extracted and determined at each plot at the research location, while one-meter buffer zones were calculated inwards from each plot boundary to eliminate unexpected boundary effects. Considering the potential variances in the treatment of each AMP, the field from the centre of the area into training and testing areas were equally divided, ensuring that the training area contained all combinations of AMPs. These collected parameters were then utilized in this study to create AutoML algorithms for estimating and evaluating grain production and straw mass.

4.5. Machine Learning Techniques

4.5.1 Machine learning techniques in red-clover biomass estimation

Parametric regression models may lead to multicollinearity between covariates and overfitting, which renders them impractical when dealing with highly dimensional remotely sensed data (Zheng et al., 2019). Conversely, machine learning algorithms can handle high volumes of predictor variables that are interrelated and have a non-linear relationship with response variables (Poley & McDermid, 2020). A recent remote sensing-based ML study collected data from 220 related articles and found that random forest (RF), support vector machine (SVM), and artificial neural network (ANN) algorithms were amongst the most used ML techniques (L. Ma et al., 2017a). Therefore, these derived ML regressions [random forest regression (RFR), support vector regression (SVR), and artificial neural network (ANN)] were chosen for modeling DM in this study. These algorithms were programmed in Python (Pilgrim, 2009) (version 3.8). The VI values presented in Table 3 were used as continuous predictor variables of the DM regression models, which were divided into training sites and prediction sites, and the parameters of each algorithm were adapted to ensure the performance as effectively as possible for the training and testing dataset.

4.5.2 Random Forest regression

An adaptation of the RF algorithm (Breiman, 2001) was conducted for DM regression models (i.e. RFR). The RFR algorithm fits an ensemble of decision tree models to a set of data. The regression tree algorithm

creates individual decision trees automatically based on randomly chosen samples and subsets of the training data. For random forest construction, the best split is selected among a random subset of the predictors at each node. Calculations were conducted with 100 trees, the minimum number of samples required to split an internal node was set to 2, and the minimum number of samples required to be at a leaf node was set to 1. Tests were run to confirm regression accuracy by using different amounts of trees ranging from 100 to 500, and it was noted that accuracy did not vary substantially with this parameter. Similar results have also been observed in other RF study (Liu et al., 2018). In terms of variable importance, the feature importance values were extracted using the *feature_importances* object located in the *sklearn. ensemble. RandomForestRegressor* class. The algorithm calculates these percentage values based on how every feature decreases the impurity of the split (mean decrease impurity) in each decision tree. The average across all trees in the forest represents the feature importance.

4.5.3 Support Vector regression

Support vector regression (SVR), which is a Kernel-based machine learning method, was used for its low dimensional and quadratic programming (QP) problem converted ability with usually only a scarce training data set needed (Shin et al., 2010). For this study, a linear kernel was used. Three extra parameters were set for the algorithm. The first included the regularization parameter (*C*, cost) set at 500. This parameter controls the trade-off between achieving a low error on the training data and minimizing the norm of the weights. The second parameter, gamma was set at 0.5. It defines how far the influence of a single training example reaches. The third parameter, epsilon, gives a margin of tolerance and was set at 0.01. In terms of variable importance, the coefficients of all six predictors estimated by the inner sklearn algorithm were extracted from the created SVR model using the *coef_ value* located in the *sklearn. svm.SVC* class, and then rescaled to be in terms of percentage.

4.5.4 Artificial Neural Network regression

The gradient-based artificial neural network (ANN), which is also called multi-layer perceptron, is a supervised algorithm that can learn nonparametric and nonlinear features that simulate human brain neural network spreading between layers and receivers and information

processing (Q. He, 1999) for classification or regression tasks. Execution of the ANN algorithm required fine-tuning of certain parameters. In this study, *lbfgs*, which stands for *Limited-memory Broyden–Fletcher–Goldfarb–Shanno*, was used as the solver since it was most optimal in saving memory. The *MLPRegressor* algorithm was executed using one layer with fifteen hidden units, with the regularization parameter (*alpha*) set at 0.00005. The maximum number of iterations allowed for this algorithm was set to 100,000. In terms of variable importance, the weights of all six predictors assigned by the inner *MLPRegressor algorithm* were extracted from the created model using the *coefs_* object located in the *sklearn.linear_module.Perceptron* class, and then rescaled to be in terms of percentage. The SVR and ANN importance scores were similarly extracted but were rescaled to also be in terms of percentage.

4.5.5 Automated machine learning (AutoML) classification with Auto-sklearn

Auto-sklearn (Feurer, Klein, et al., 2015b), a robust and efficient AutoML system first introduced in 2015 and upgraded in 2020 (Feurer et al., 2020), was utilized in this study. Auto-sklearn is developed on the Python scikit-learn machine learning package. It uses 15 classifiers, 14 feature pre-processing methods, and four data pre-processing methods, giving rise to a structured hypothesis space with 110 hyperparameters (Pedregosa et al., 2011). It improves on existing AutoML methods by automatically considering the previous performance on similar datasets, and by constructing ensembles from the models evaluated during the optimization process. At its core, this method combines the highly parametric ML framework with automatically constructed ML pipelines suggested by the Bayesian optimization method sequential model-based algorithm configuration (SMAC) (Hutter et al., 2011). SMAC can automatically construct ML pipelines that include feature selection (i.e. removing insignificant features), transformation (i.e. dimensionality reduction), classifier selection comprising SVM (Suykens & Vandewalle, 1999), RF (Breiman, 2001) and other algorithms, hyperparameter optimization, etc. Subsequently, it then utilizes a random forest technique for swift cross-validation by evaluating one-fold at a time, while at the same time discarding poor-performing hyperparameter settings during early stages. It achieves competitive classification accuracy, in addition to novel pipeline operators that significantly increase classification accuracy on the datasets (Olson et al., 2016). During the feature selection stage,

any highly correlated VIs were removed to eradicate the influence of collinearity. This step was omitted here since Auto-sklearn deals with the low dimensional optimization problems (Feurer, Springenberg, et al., 2015).

In this study, all calculations were done in the open-source operating system LINUX with Intel Core i5-1035G1 CPU (1.00 GHz) and 16 GB RAM. For the AutoML framework, the steps were described in (Feurer, Klein, et al., 2015b). Firstly, the system uses a supplementary approach of extensively applied meta-learning methods to train machine learning models over statistical attributes of datasets and estimate the parameter of models that yield the best precision (Franceschi et al., 2018). Secondly, the system automatically built ensembles of the models considered by Bayesian optimization. Thirdly, the system constructed a highly parameterized ML framework from high-performing classifiers and pre-processors implemented within the ML framework. Finally, the system performed broad empirical analysis using a diverse collection of datasets to demonstrate the resulting Auto-sklearn system outperformed preceding AutoML methods. The major AutoML parameter settings of this study are described in Table 5. Due to computational resource constraints and to test the efficiency of AutoML, CPU time for each run was limited to 60 seconds and the running time for evaluating a single model to 10 seconds as an example of rapid model selection. Subsequently, a total of 1200 seconds with a 10-second single model computing time was used as a representative of the better processing of AutoML models. The data were analysed separately according to the four crop fields (F1-F4), with each field containing 72 plots (n=72) with splitting in the training site and validation site (0.6/ 0.4) for classification modelling.

Table 5. The AutoML main parameters and descriptions that were used in this study.

Parameter Name	Range Value	Description
time_left_for_this_task	60- 1200 sec	The time limit for the search of appropriate models.
per_run_time_limit	10 sec	The time limit for a single call to the machine learning model.
ensemble_size	50 (default)	The number of models added to the ensemble built by *Ensemble selection from libraries of models*.

ensemble_nbest	50 (default)	The number of best models for building an ensemble model.
resampling_strategy	CV; folds = 3	(CV= cross-validation); to handle overfitting
seed	47	Used to seed SMAC.
training/ testing split	(0.6; 0.4)	Data partitioning way

* The other parameters that are not listed on the table were run in default mode.

A recent review study of supervised ML methods applied in land-cover image classification disclosed that RF, SVM, and ANN classifiers were amongst the most commonly used ML techniques from 220 related articles (L. Ma et al., 2017a). Therefore, in this study, these popular ML classifiers were selected for comparison against the accuracy performance of AutoML (with 60-sec run, and 1200-sec run of Auto-sklearn). These algorithms were programmed in Python by the robust ML library Scikit-learn (0.24.2) (Pedregosa et al., 2011) with the perimeter setting as following: *sklearn.ensemble*.RandomForestClassifier [100 trees; *min_samples_split* (2); *leaf_node* (1)]; *sklearn.svm*.SVC [cost (*C*=500); gamma (0.5); epsilon (0,01)], and *sklearn.neural_network*.MLPClassifier [alpha (0.00005); the maximum number of iterations (100,000)]

4.5.6 Automated Machine Learning (AutoML) regression with Auto-sklearn

This study employed the robust and frequently updated AutoML system, Auto-sklearn, based on the scikit-learn ML library in Python (Pedregosa et al., 2011). It employs 15 classifiers, 14 feature processing, and four data pre-processing methods, yielding a 110-hyperparameter structured hypothesis space (Feurer et al., 2020; Feurer, Klein, et al., 2015a). It offers an advancement on existing AutoML approaches by incorporating prior performance on comparable datasets and generates ensembles from the models that were examined throughout the optimization procedure. This technique involves the largely configurable ML prototype with the automatically generated ML pipelines, i.e. feature selection (deleting trivial features), transformation (reducing dimensionality), hyperparameter optimization based on Bayesian optimization strategy SMAC (Hutter et al., 2011). Following that, a Random Forest (Breiman, 2001) approach was utilized for fast cross-validation, assessing one-fold at a time and eliminating poor-performing hyperparameter configurations during the

initial phases. The Random Forest approach delivers a superior accuracy rate, as well as alternative pipeline operators that boost regression performance within the datasets (Feurer, Klein, et al., 2015a; Olson et al., 2016).

All computations in this study were performed on an Intel Core i5-1035G1 CPU (1.00 GHz) with 16 GB RAM utilizing the LINUX open-source operating system. The processes outlined in (Feurer, Klein, et al., 2015a) were executed for the AutoML framework. To begin with, the system employs a supplemental technique based on widely used meta-learning procedures to train machine learning models over the statistical features of datasets and evaluates the model parameters that produce the greatest performance (Franceschi et al., 2018). Second, the system creates ensembles of the models that Bayesian optimization examined, using high-performing regressors and pre-processors employed within the ML framework. Finally, the program works a wide range of empirical examinations on a diverse set of data to determine whether the AutoML regression offers better outcomes than previous regressions. However, any strongly correlated VIs should be eliminated during the feature selection step to avoid the effects of collinearity. Since Auto-sklearn works with low-dimensional optimization issues (Feurer, Springenberg, et al., 2015), this step was bypassed in this stage. Table 6 lists the principal AutoML regression parameters employed in this study. To perform tests, as a demonstration of the practicability and efficiency of AutoML model selection, CPU timing for each task was restricted to 30 seconds, and the runtime for assessing a single model to 10 seconds. The analyses were performed separately for each of the crop fields, with grain yield consisting of 56 plots (n=56) and straw mass divided in the training and test sites (0.5/0.5) for regression modelling (see flowchart Figure 9).

Table 6. The AutoML regression parameters and descriptions were employed in this study.

Parameter Name	Range Value	Description
time_left_for_this_task	30 sec	The time restriction for seeking suitable models.
per_run_time_limit	10 sec	The maximum amount of time a single call to the ML model could perform.

ensemble_size	50 (default)	Several models were added to the ensemble from *Ensemble libraries*.
ensemble_nbest	50 (default)	The amount of best models for building an ensemble model.
resampling_strategy	CV; folds = 3	(CV= cross-validation); to deal with overfitting
seed	47	Used to seed SMAC.
training/ testing split	(0.5; 0.5)	Data partitioning way

* Other options and parameters that aren't shown in the table were set to default.

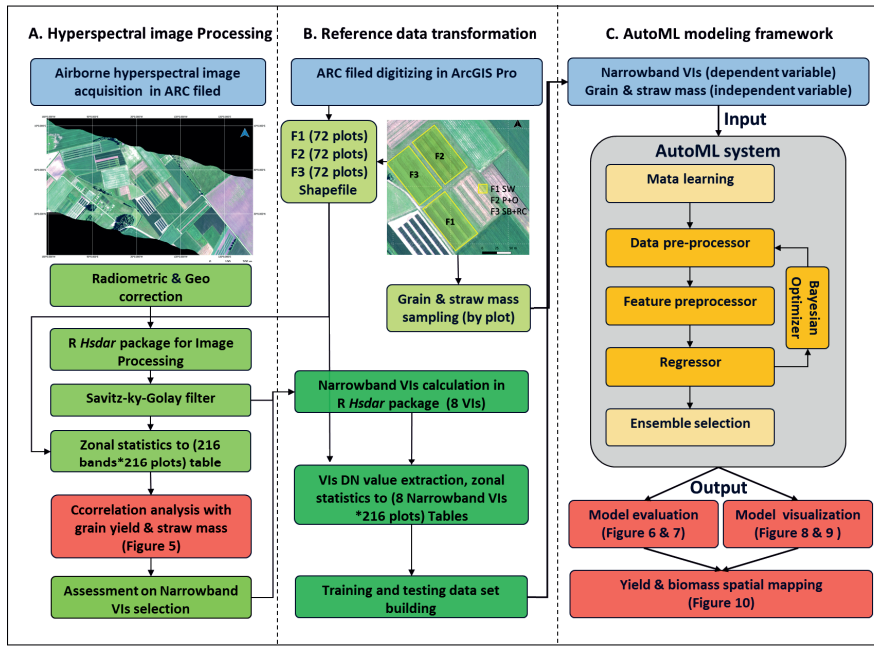

Figure 9. The flowchart of the hyperspectral image processing and AutoML framework was utilized in this study. (A) The hyperspectral image processing framework where hyperspectral imager HySpex was conducted and R *Hsdar* package was employed in the processing steps. (B) Field reference data transformation, ARC field were digitized based on each field and following AMP treatments. The grain yield and straw mass data were collected according to plots. Eight narrowband VIs were selected and calculated and segmented into corresponding plot digital numbers (DN) for AutoML modelling. (C) To achieve robust performance, the Auto-sklearn framework automatically built ML pipelines that were provided by the Bayesian optimization method with warm-started meta-learning and combined with a post hoc ensemble building strategy (adapted from (Feurer, Klein, et al., 2015a)).

4.6. Model evaluation

4.6.1 Regression Model Evaluation

In red-clover biomass estimation, to reduce the potential over-fitting problem of the model, a leave-one-out cross-validation (LOOCV) procedure (Kearns & Ron, 1999) was conducted to validate the three ML techniques. The LOOCV procedure involves creating a model by separating one sample for testing and the rest (n = 36) for validation in every iteration (Figure 10a). Second, all training sites were used to model, evaluate and predict the three ML methods (Figure 10b). As the training and testing dataset comprised two repetitions of results from each treatment, a comprehensive range of crop conditions was covered by the modelling. The variable importance of the VIs was calculated for each ML technique differently and was listed per each model's VIs importance scores, and the suitable models for different periods for DM yield spatial mapping were demonstrated. Finally, experimental treatments (i.e. STM, CM, and MA) were used to explore the relationship between different experiment factors and models (Figure 10c). The testing sites were sampled following a stratified approach based on the three different farming operations, 3 strata with 12 samples in each one in STM (DP, P, and R) and CM (Cmin, Omin+, and Omin-) groups, and 2 strata with 18 samples in MA (M+, and M-). The ML methods, selected parameters, model evaluation techniques, and variable importance calculations are described below.

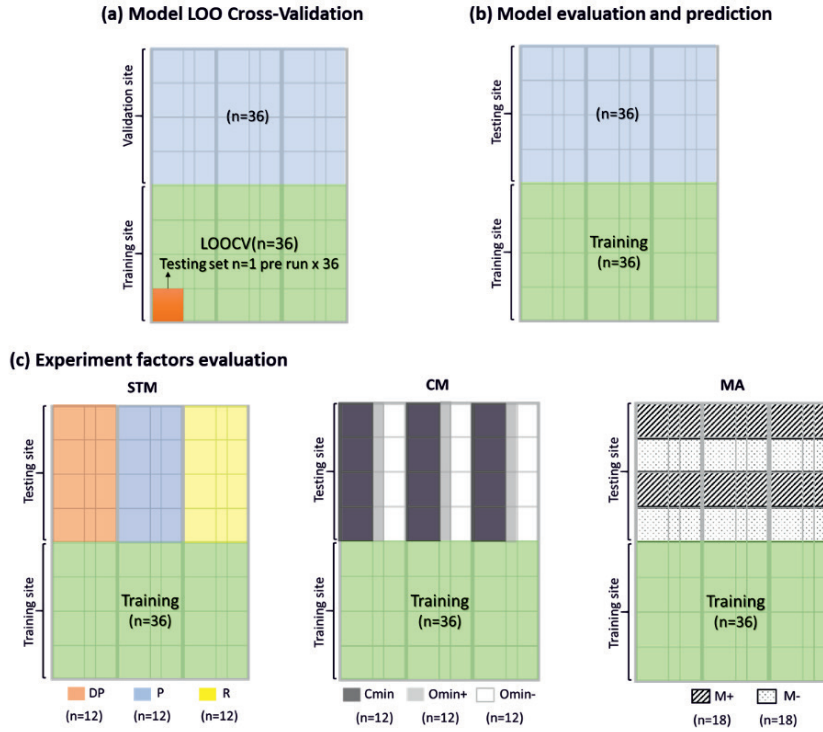

Figure 10. Examples of different sampling methods and regions in field 2YC, the training site (1/2, n = 36) contains two repeated trial plots. (a) The evaluation of RFR, SVR, and ANN using the cross-validation method LOOCV in the training site (1/2, n = 36) and validation site (1/2, n = 36) (b) Model construction including training site (1/2, n = 36) and testing site (1/2, n = 36) (c) Evaluation of the model efficiency across three different treatments: STM (n = 12 for each subset), CM (n = 12 for each subset), and MA (n = 18 for each subset).

For the model prediction evaluation of each model, the accuracy evaluation method described by (Yue et al., 2017), was used. The models' accuracies were measured by the coefficient of determination (R^2) (Equation 1) and normalized root means square error (NRMSE) (Equation 2). The equations used are as follows:

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

$$NRMSE = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}}{\bar{y}} \quad (2)$$

where N represents the total sample size, yi is the ith DM yield value of the sample, ŷi is the ith predicted value, ȳi is the ith computed mean value, and ȳ represents the difference between the maximum and minimum values of the dataset.

4.6.2. Classification model evaluation

The parameters not mentioned were computed as default settings from Scikit-learn, and for accuracy, the calculation refers to Table 7.

Table 7. The confusion matrix-based accuracy evaluation equations used throughout this study.

Indices	Equations
Recall	TP/ (TP+FN)
Precision	TP/ (TP+FP)
Specificity	TN/(TN+FP)
Accuracy	TP/ (TP+TN+FP +FN)
F1-score	2 * Precision * Recall / (Precision + Recall)
False Positive Rate (FPR)	1 - Specificity = FP / (FP + TN)
True Positive Rate (TPR)	Sensitivity = TP / (TP + FN)

For the visualization and evaluation of the Auto-sklearn model, the workflow includes, in general, multiple iterations through feature engineering, algorithm selection, and hyperparameter tuning (Kumar et al., 2016). In this study, an open-source visual steering tool Yellowbrick visualization package (essentially a wrapper for the Sklearn documentation) was conducted for AutoML evaluation (Bengfort & Bilbro, 2019). Yellowbrick contributes to assessing the stability and predictive values of ML models and delivers visualizations for the AutoML classification models. The accuracy evaluation based on the confusion matrix system of the AutoML classification parameters were defined as follows: true positive (TP), false positive (FP), true negative (TN), and false-negative (FN), which have been well described in (Anwar et al., 2019). The equations used in this study are described in Table 7. The derived receiver operating characteristic curve (ROC) graph with the x-axis showing FPR and the y-axis showing TPR was used in this study to show the relationship among specificity and sensitivity for each possible cut-off (Fawcett, 2006) and the area under the curve (AUC) ranges from 0 to 1 to visualize the trade-off between the classifier's

sensitivity and specificity (Fawcett, 2006; Pencina et al., 2008). Macro and micro-averaging ROC were calculated to evaluate overall classifier performance in multi-class problems. In this approach, the ROC curve was calculated anew, based upon the true positive and false-positive rates for all datasets (by weighting curves by the relative frequencies of the dataset and then averaging them) (Gunčar et al., 2018; Sokolova & Lapalme, 2009). In addition, the precision-recall curve (PR) was calculated for different probability thresholds. PR curves were conducted in cases where there was an imbalance in the observations between the classes (Boyd et al., 2013) as another classification evaluation standard to assist with the ROC curve. The prediction errors (confusion matrix) and classification report that displays precision, recall, and F1-score (Chicco & Jurman, 2020) (Table 7) per class as a heatmap in this study.

Alternatively, even though the AutoML framework facilitates the construction of models, given their black-box nature, the complication of the underlying algorithms, and the large number of pipelines they derive leads to reduced trust in AutoML pipelines systems (Q. Wang et al., 2019). Therefore, in this study, PipelineProfiler (Ono et al., 2021) was conducted for AutoML pipelines visualization. PipelineProfiler is a SOTA in visual analytics for AutoML interactive visualization tool that allows the examination of the solution space of end-to-end ML pipelines. It offers a recovering understanding of how the AutoML algorithms are generated and the perceptions of how they can be optimized. As the outcome of the interactive AutoML pipeline matrix plots, where illustrated Pipeline flowchart, primitives used by the pipelines; one-hot-encoded hyperparameters for the primitive across pipelines; the accuracy ranking; primitive contribution view; and the class balancing of correlation score with accuracy. These calculations and expressions are clearly detailed described in the (Ono et al., 2021) article.

5. RESULTS

5.1. The Multispectral-VIs Application for Red Clover-grass Mixture Yield Estimation

5.1.1 The red clover- grass mixture DM modeling and LOOCV

For better evaluation evaluating of the model performance, LOOCV was firstly conducted in this research to assess three machine learning methods (RFR, SVR, and ANN) abilities to predict DM yield using six VI. The distributions of R^2 and NRMSE values under thirty-six LOOCV iterations of the three models are shown in box plots (Figure 11). The results showed that in terms of flight dates, 11DB generally performed better than 38DB. In terms of model performance, ANN had the best performance in 11DB, with the highest R^2 values (1YC = 0.84, 2YC = 0.85), and the lowest median NRMSE with a stable distribution of outliers. RFR's accuracy was slightly smaller than for ANN. In terms of regional prediction, 2YC was on average better than 1YC. Overall, the results of LOOCV showed that the three models have moderate to high accuracies in two locations and different flight dates, and the best R^2 values were observed in 2YC11D (0.80 to 0.85), while the worst were observed in 1YC38DB (0.64 to 0.70).

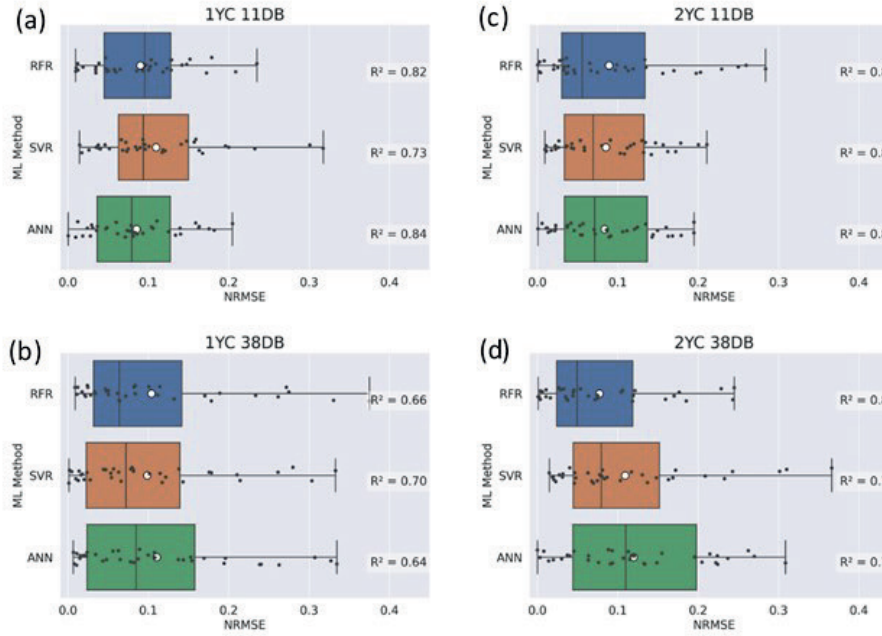

Figure 11. Comparison of the NRMSE and R^2 values resulting from 3 different ML methods (RFR, SVR, and ANN) of LOOCV in (a) 1YC11DB (b) 1YC38Db (C) 2YC11DB (d) 2YC38DB. Each model performed 36 times LOOCV calculations. The R^2 for LOOCV was calculated using the average variance between the actual and prediction value for every iteration of the cross-validation. The black dots showed the NRMSE results of each cross-validation, and the white dots represent its average value. The median line in the box shows the middle value, and the interquartile range of the box (shown in blue, red, and green) represent the 25th to the 75th percentile.

5.1.2 The red clover- grass mixture model evaluation and variable importance

After the cross-validation, the training dataset (n = 36) was used for the calculation of models separately (Figure 10). During the modeling phase, the appropriate combinations of the parameters of the data set were tested. The scatter plots with model predictions and observed DM values compared to the 1:1 line, and their corresponding variable importance are shown in Figures 12 and 13.

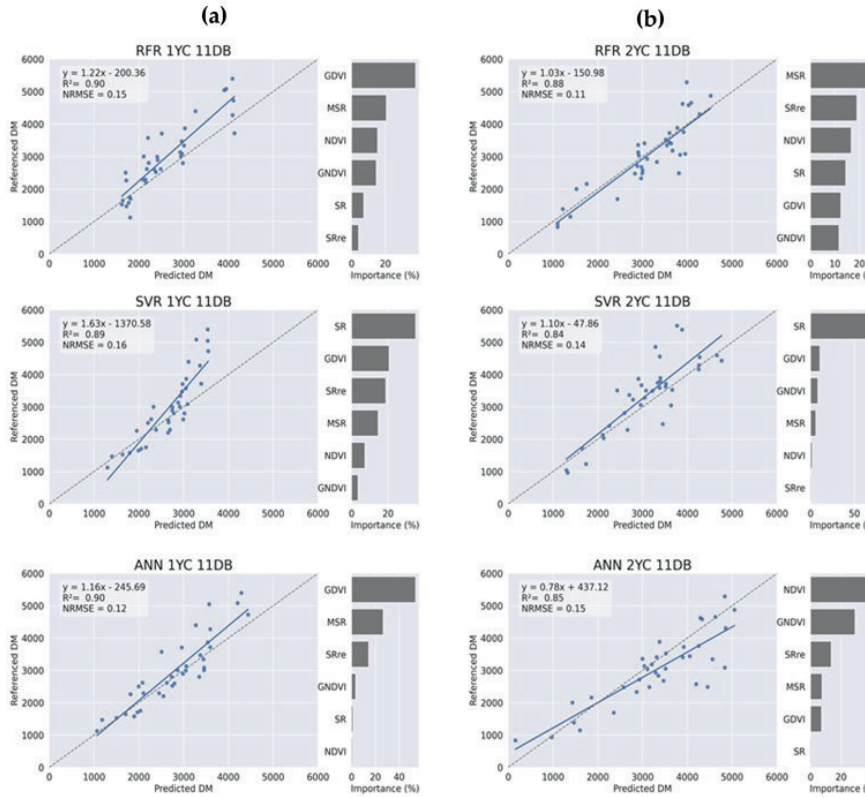

Figure 12. Regression plots of 1YC and 2YC fields based on RFR, SVR, and ANN methods in 11DB flight. The plots correspond to (a) 1YC11DB and (b) 2YC11DB. The horizontal bar plots on the right side of each graph shows the variable importance estimation based on the models. The horizontal axis in the scatter plots describes the predicted DM yield acquired from the model, and the vertical axis stands for the field-observed DM yield. The R^2 = coefficient of determination, NRMSE = normalized root means squared error, and the black dotted line exemplifies the 1:1 slope.

Figure 12 shows three red clover-grass mixture DM models across the 1YC and 2YC fields where the images were captured eleven days before harvesting (11DB). The results indicate that, in 1YC11DB (Figure 12a), the ANN model had the lowest prediction errors (NRMSE= 0.12) and the highest R^2 value (0.90). RFR had a similar performance, but with higher NRMSE. Although the three models performed well, a slight uniform underestimation of DM yield appeared in both the RFR and ANN models. On the other hand, a non-uniform bias appeared in the SVR model, which overestimated small DM values and underestimated large DM values. According to the ranking of variable importance, GDVI and MSR provided higher contributions to the RFR and ANN

models. Concerning the SVM model, larger contributions were found for SR and GDVI. The results of 1YC38DB (Figure 12b) show that the three models performed relatively well (R^2 from 0.84 to 0.88). The slope of RFR was closest to the 1:1 line, with the smallest NRMSE (0.11) and highest R^2 value (0.88); ANN and SVR had similar predictive capabilities. With regard to variable importance ranking, SVR showed similar results compared to 11DB with the highest contribution of SR, while in the results of RFR and ANN, the contribution of VIs did not show obvious similarity with the highest ranking in MSR and NDVI, respectively.

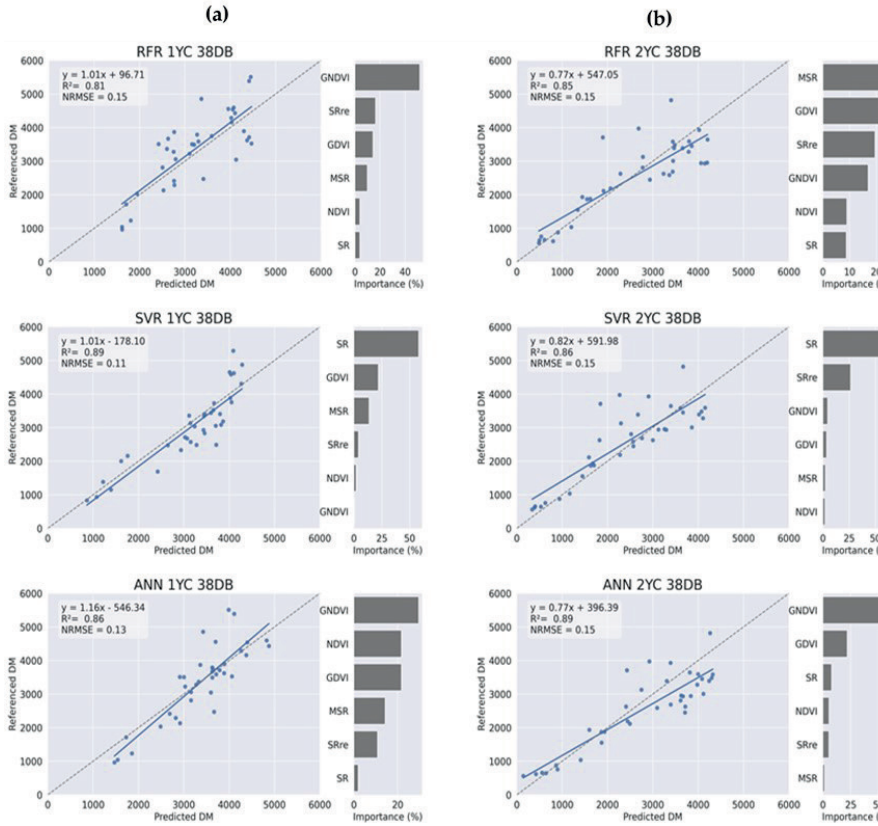

Figure 13. Regression plots of 1YC and 2YC fields based on RFR, SVR, and ANN methods in 38DB flight. The plots correspond to (a) 1YC38DB and (b) 2YC38DB. The horizontal bar plots on the right side of each graph shows the variable importance estimation based on the models. The horizontal axis in the scatter plots describes the predicted DM yield acquired from the model, and the vertical axis stands for the field observed DM yield. The R^2 = coefficient of determination, NRMSE = normalized root means squared error, and the black dotted line exemplifies the 1:1 slope.

Figure 13 shows the behavior of predictive models using the thirty-eight days before harvest (38DB) datasets. The 1YC38DB results (Figure 13a) showed that SVR had the highest R^2 (0.89) and the smallest NRMSE (0.11), where the slope was close to the 1:1 line. In contrast, ANN and RFR relatively had weaker performances. However, the overall performance of the models for the 2YC dataset was slightly inferior to the result of 1YC, showing a higher bias of the slopes. In terms of 2YC38DB results (Figure 13b), ANN had the best performance among the three algorithms (R^2 = 0.89, NRMSE= 0.15). In the ranking of predictor variables in 1YC and 2YC, the GNDVI, MSR, NDVI, and SRre played crucial roles in both RFR and ANN models. Besides, the GNDVI was the most important variable in both 1YC and the model of ANN in the 2YC field. In contrast, SR stably ranked as the most important in the models of SVR, regardless of flight dates or regions.
Based on the evaluation of models (Figures 12 and 13) and their suitability for different periods, prediction maps (Figure 14) of DM yield for both experimental sites were generated.

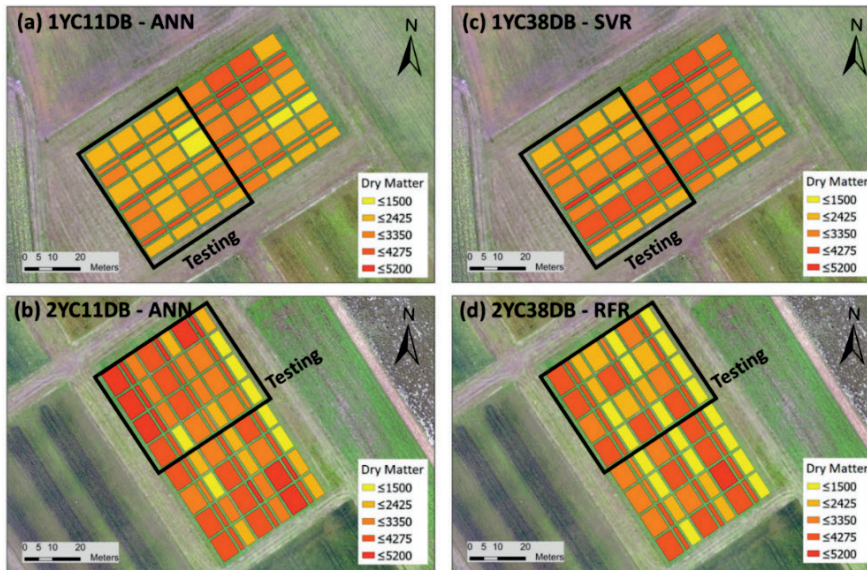

Figure 14. The spatial testing (prediction) site mapping output of DM yields (kg ha^{-1}) in 1YC and 2YC fields based on 11DB and 38DB flights by machine learning techniques at the plot level: (a) 1YC11DB, (b) 2YC11DB, (c) 1YC38DB, and (d) 2YC38DB. The best performing algorithm was chosen for each of the four categories, as shown in the previous results.

5.2. The Multispectral-VIs and AutoML framework into agricultural management practices recognition.

5.2.1 The AMPs observation in VPTs and VIs calculation

Figure 15 displays the observation of onsite crop VPTs [i.e., Field 1 (F1) (Figure 15a) and Field 2 (F2) (Figure 15b) with CM treatments] and one of the VIs (NDVI; Figure 15c) captured on July 10th from F1 and F2. It can be observed from the onsite AMPs treatment photographs of F1 and F2 in July that is not readily distinguishable. In addition, it can be seen from the NDVI image that the heterogeneity within the plot may be caused by edge effects or uneven fertilization. For this reason, the plot average value considering the pixels' inward boundary clipping to decrease the noise was used.

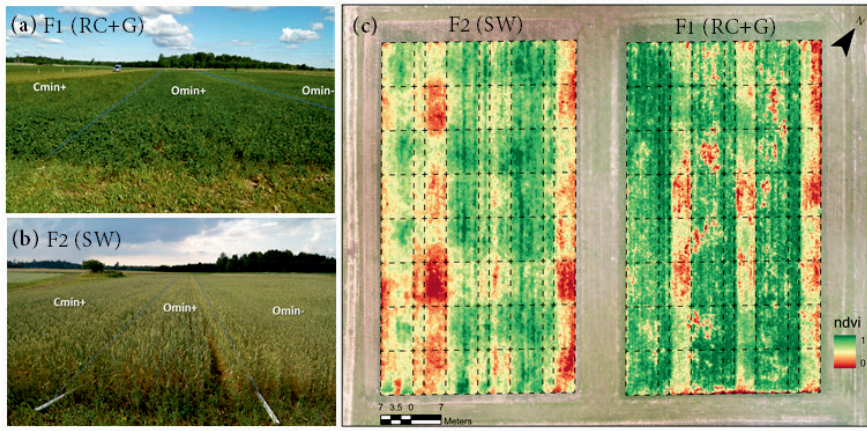

Figure 15. Interpretation diagrams representing onsite crop VPTs and the calculation of VIs per the image captured on July 10th (a) Field 1: red clove +grass (RC+G) with CM treatment. (b) Field 2: Spring wheat (SW) field with CM treatment. (a) Normalized Difference Vegetation Index (NDVI) image captured of F1(RC+G) and F2 (SW) VPT.

5.2.2 Monthly PCA Analysis in Various Crop Growth Periods

The PCA results (Figure 16) was conducted as the first step of data exploration in this study to gain an understanding of the relationship between VIs and different AMP categories during the three flight periods (April, May, and July) with their corresponding growing stages. The results show that on May 30th and July 10th, the PC1 and PC2 captured most of the variation from the F1 to F4 fields with 98.3%, 98.7%, 97.3%, and 97.6%, respectively, on May 30th (Figure 16b), and

with 98.7%, 94.0%, 95.4%, and 95.4%, respectively, on July 10th (Figure 16c); followed by April 23rd (Figure 16a). In addition, during the three flight periods, the PCA results in May and July provide better separation of the three AMP categories throughout the four crop cultivation areas based on the coloured concentration ellipses where the sizes determined by a 0.95-probability level. In terms of the AMP category, the subclasses of CM (Cmin+ Cmin+ and the other two categories) and MA (M+ and M-) seems easier for non-overlapping AMP clustering, followed by STM. In terms of crop types, F1 (SW) were better clustered in April, while F2 (SW), F3 (P+O), and F4 (SB+RC) were better clustered in May or July. Given the better clustering performance in May, follow-up AutoML analysis was conducted on the UAS multispectral-VIs data of this month. In general, feature selection (finding the most relevant spectral bands) and extraction (reduced set of new significant variables) are commonly used to solve the collinearity and overfitting problems in the dimensionality reduction process (Serpico et al., 2003). However, after test results, using PCA 95% feature extraction in the preliminary experiments could not significantly improve the classification efficiency. Therefore, these PCA results were simply used as a reference basis for AutoML classification.

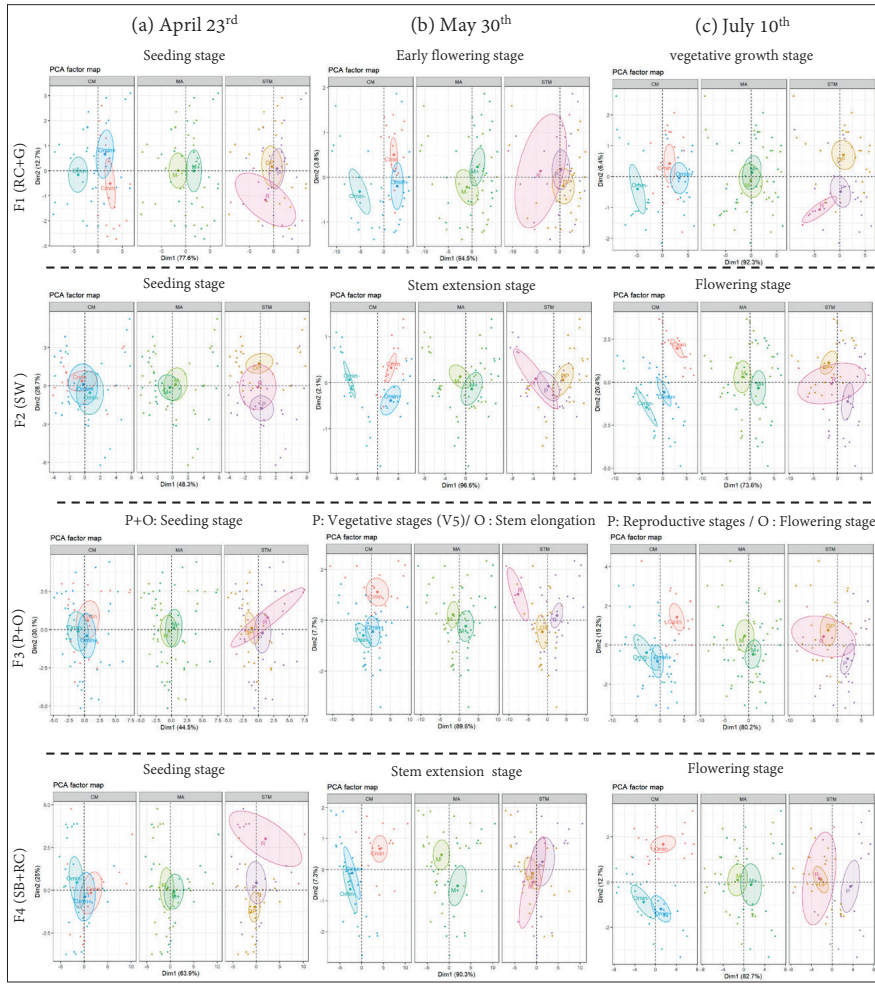

Figure 16. PCA biplot of 19 VI variables (n = 72) of each crop field at April, May, and July. Each biplot shows the PCA individuals (3 AMPs) [i.e., CM (Cmin, Omin+, Omin-), MA (M+, M-), STM (DP, P, R)] of the first (x-axis : PC1 score) and second (y-axis : PC2 score) principal components (the variation explained by the dimensions are shown on the axes); four crop categories (F1-F4) and its corresponding growing stage from top to bottom. Coloured concentration ellipses (size determined by a 0.95-probability level) show the observations grouped by marked AMP sub-classes.

5.2.3 AutoML ROC and AUC Evaluation of AMP Recognition in May

The different subclasses and average results of ROC/AUC were calculated for evaluation of the AutoML performance for the AMP classification ability in UAS multispectral-VIs that were captured in May

(Figure 17), where AUC values were categorized in this study as AUC = 0.5: no discrimination; 0.7≦AUC≦0.8 (acceptable discrimination); 0.8≦AUC≦0.9 (excellent discrimination); 0.9≦AUC≦1.0 (outstanding discrimination) (Fawcett, 2006).

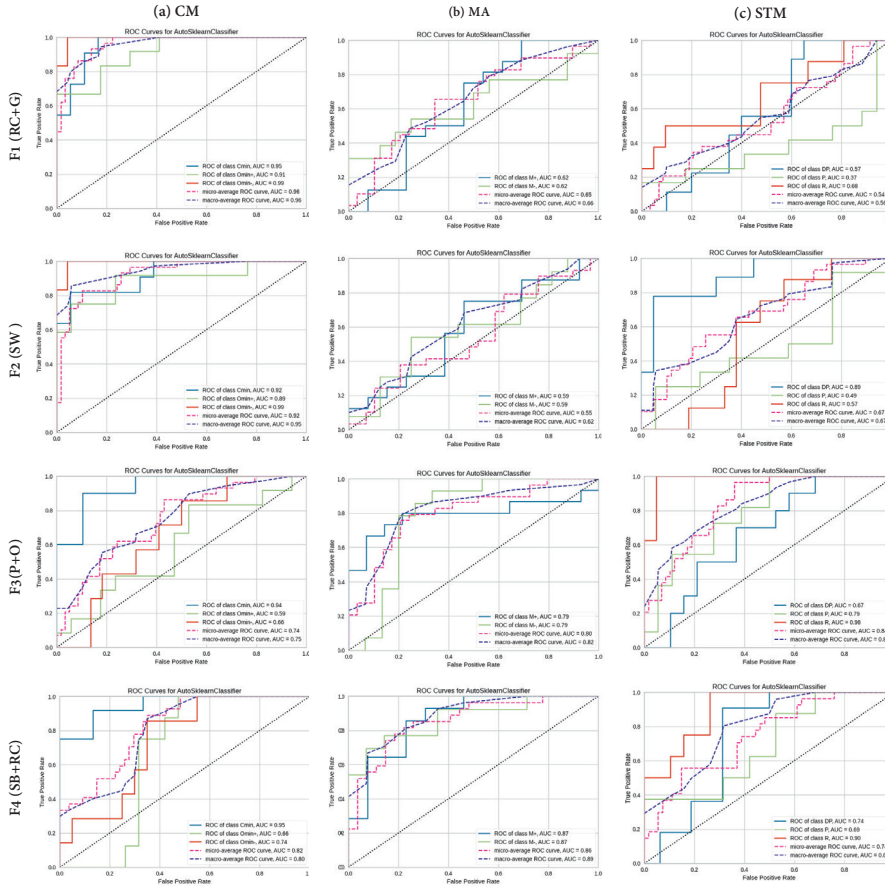

Figure 17. ROC curves and AUC of the AutoML classification corresponding to the subclasses within the AMPs for the acquisition of the UAS multispectral-VIs DN in May. From left to right, the ROC curves computed on (a) CM [(Cmin (blue lines), Omin+ (green lines), Omin- (red lines)]; (b) MA[M+ (blue lines), M- (green lines)]; (c) STM [DP (blue lines), P (green lines), R (red lines)]; and their micro (pink dotted line) and macro (dark blue dotted line) average performance. Four crop categories (F1-F4) from top to bottom.

The AutoML results showed that the micro-average ROC of CM's classification results in F1(RC+G) and F2 (SW) were higher (AUC = 0.95, and 0,92, respectively). Especially in the subclass Omin-, the AUC both reached 0.99, the micro-average ROC followed by F4, and F3

(P+O), with 0.86 and 0.75, respectively) (Figure 17a). On the contrary, MA classification results show that the micro-average AUC in F3 and F4 were higher (AUC = 0.83, and 0,89, respectively), followed by F1 (AUC = 0.71). F2 performance for MA was the worst (AUC =0.51), with no discrimination ability (Figure 17b). In contrast, STM classification results were generally poor, with better results only present in F3, while other fields have larger divergence in classification results under the sub-class (DP, P, and R), as shown in Figure 17c). Overall, the AutoML classification ability from UAS multispectral-VIs of CM was the best, followed by MA and STM.

5.2.4 AutoML precision-recall, prediction error, and classification report of CM recognition

Amongst the classification results of AMPs in May (Figure 17.) of four crop types that CM yielded the best ROC/AUC overall performance. Therefore, the precision-recall (PR) curves, prediction error, and classification report plots were used to gain an in-depth understanding of the classification status of CM treatments (Figure 18).

The PR curve of F4 CM shows the trade-off between a classifier's precision performance from UAS multispectral-VIs in May (Figure 18a) where a model with perfect performance is depicted at the coordinate of (1, 1). A curve that tends towards the (1, 1) coordinate represents a well-performing model, whereas a no-skill classifier is depicted as a horizontal line on the plot with a precision that is proportional to the number of positive examples in the dataset. For a balanced dataset, this value ought to be 0.5 (Saito & Rehmsmeier, 2015). The results show that the classifications of Field 1 and 2 were promising, their average PR being 0.90 and 0.85 respectively, while the results of F3 and F4 were poor (0.50 and 0.49). One can further discover from the prediction error graph (Figure 18b) in F3 and F4 that the judgment error of Cmin+ is low, and the confusions of Omin+ and Omin- were more common. The precision, recall, and F1-score results of various cultivation method sub-classes can be compared to evaluate the classification accuracy from the heatmap (Figure 18c).

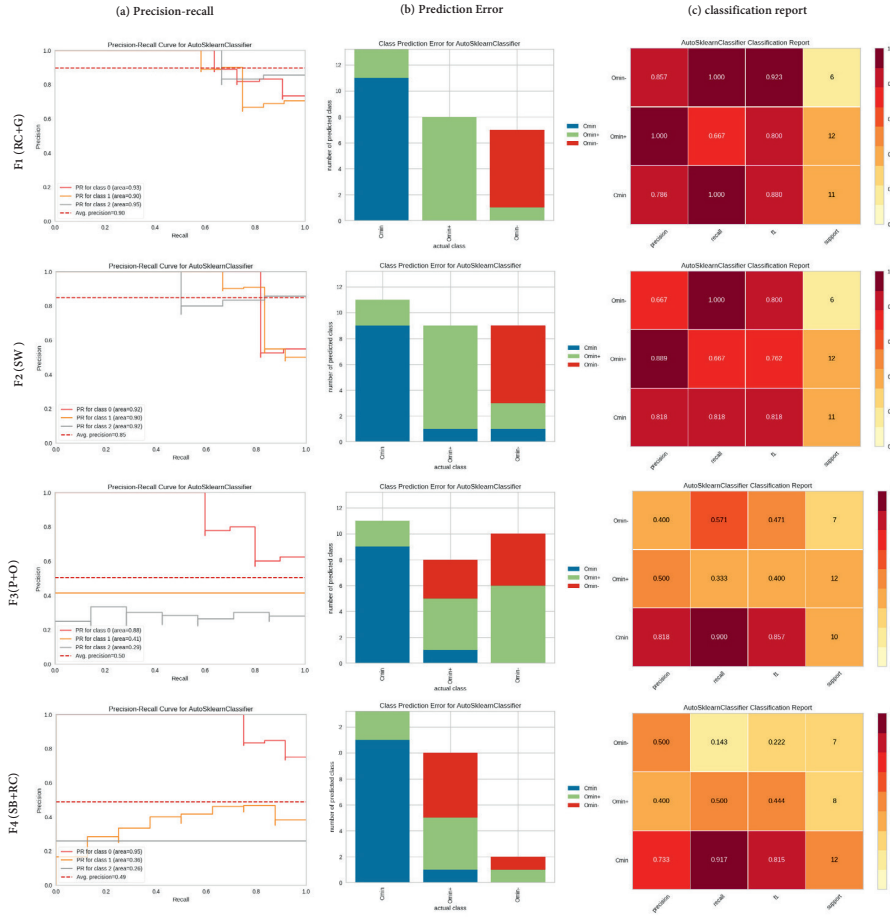

Figure 18. The evaluation of AutoML classification of AMPs from the acquisition of the UAS multispectral-VIs DN in May. (a) Precision-recall, where the class 0, 1, 2 equals to Cmin+, Omin+, and Omin- respectively (b) Prediction Error (confusion matrix), the X-axis represents the three subclass form CM result in May, and the Y-axis represents the type (with colour) and the number of correct or incorrect estimates., and (c) Classification report lists the precision, recall, and F1-score per class as a heatmap for overall comprehensive evaluation results.

5.2.5 AutoML Pipeline Visualization

An interactive AutoML visualization tool *PipelineProfiler* was conducted in this study. Figure 19 shows the CM classification results across four crop fields in May with the accuracy performance of AutoML pipelines running time set at 60 seconds, and the primitive comparison against the others and the real-time hyperparameter selection strategy. The results demonstrate that the best classifier found for Field 1 was linear

discriminant analysis (LDA) (FISHER, 1936) (Figure 19a), for Field 2 it was the Extra Trees Algorithm (Geurts et al., 2006)(Figure 19b), for Field 3 it was LDA (Figure 19c) and RF for Field 4 (Figure 19d), with each of their hyperparameters found by AutoML also being represented in the figures.

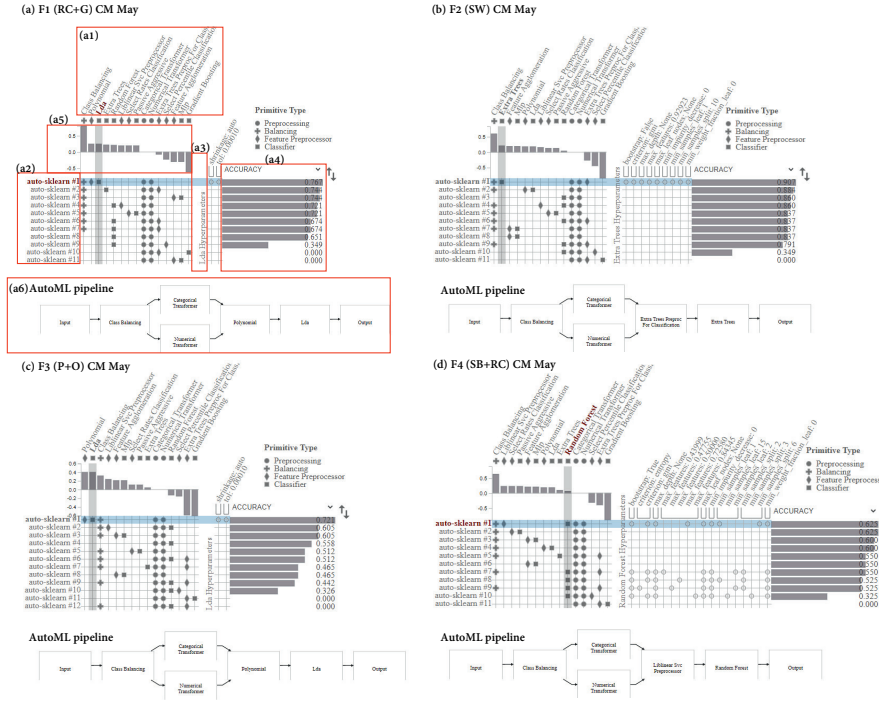

Figure 19. The interactive AutoML pipeline matrix plots with running time-limited setting 60 sec sorted by accuracy performance **(a)-(d)**, **(a)** Field 1 pipeline matrix with the Top1 classifier LDA, where **(a1)** illustrated Primitives (in columns) used by the pipelines **(a2)** (in rows, the blue line showed the best accuracy rank); **(a3)** one-hot-encoded hyperparameters (in columns) for the primitive across pipelines, **(a4)** the AutoML pipeline with the accuracy ranking; **(a5)** Primitive contribution view, showing the correlations between primitive usage and pipeline scores – in a5 displays that class balancing has the highest correlation score with accuracy; **(a6)** Step by step AutoML Pipeline flowchart. The ML box before Output represents the classifier used by this set of algorithms (in a6 LDA as the classifier) **(b)-(d)** Field 2, 3, and 4 interactive pipeline matrix sort by AutoML accuracy performance with the chosen hyperparameters (top 1 was listed).

5.2.6 Comparison of performance between AutoML and other machine learning technologies

Based on the large calculations and multiple classifier selections that were required during the initial stage of AutoML computations, the processing time setting of 60 seconds may not completely reflect the performance power of AutoML. To evaluate the effects of AutoML processing time, the times were adjusted to 1200 seconds and 60 seconds (original running time) and considered the AMPs classification accuracy with RF, SVM, and ANN algorithms (Table 8). The results demonstrate that under the permutation and combination of ML algorithms included in AutoML, classification accuracy does not perform well in 60 seconds of computing time. Furthermore, performance was the worst in F1 CM, F2 STM, and F3 CM classifications compared to RF, SVM, and ANN. However, as processing time was increased to 1200 seconds, the classification accuracy of AutoML in AMPs was shown to improve. The results also indicated that overall AutoML (1200 sec) and RF classifiers produced 5 and 3 best classification accuracy in AMPs respectively (in black bold) and did not produce the worst accuracy values (in bold red) in any instances. Regarding SVM and ANN, the classifiers performed the best in 3 and 5 cases, respectively. However, these methods consistently produced low performing classifiers compared to other AMPs.

Table 8. The AMPs classification accuracy comparison of AutoML and three other popular applied ML (RF, SYM, and ANN) algorithms in UAS multispectral-VIs.

		ML algorithms				
Field	AMPs	AutoML (1200 sec run)	AutoML (60 sec run)	RF	SVM	ANN
F1 (RC+G)	CM	0.79	0.76**	0.79	0.83	**0.86***
	MA	0.59	**0.62***	**0.62***	**0.62***	0.55**
	STM	**0.57***	0.31	0.48	0.38**	0.48
F2 (WS)	CM	0.79	0.79	0.79	**0.83***	0.72**
	MA	**0.55***	0.52	0.48	0.52	0.45**
	STM	**0.52***	0.45**	0.48	**0.45****	**0.52***
F3 (P+O)	CM	**0.55***	0.41**	**0.55***	0.48	**0.55***
	MA	0.66	0.72	**0.76***	0.62**	**0.76***
	STM	0.66	**0.69***	**0.69***	0.57**	0.59

F4	CM	0.57	**0.59***	0.56	**0.59***	0.48**
(SB+RC)	MA	**0.85***	0.78	0.67	0.78	0.63**
	STM	0.56	0.59	0.59	0.52**	**0.63***

(*) The **bold** black numerical value in the Table represents the highest accuracy classifier in the row; (**) the thin red numerical value represents the worst accuracy in the row.

5.3. The Hyperspectral Image and AutoML in Crop Yield and Biomass Estimation.

5.3.1 The Field Observation DM Data Analysis

The average actual grain yield and above ground straw mass data (fresh and dry) gathered from the SW, P+O, and SB+RC experimental regions are displayed in the violin plot (Figure 20), where the range of grain yield and straw mass data are exhibted and were assembled by fields since the treatments were interspersed within each plot. In addition, dry and fresh weight were examined separately since the accumulated rainfall of 4.1 mm (in SW and P+O fields) and 0.4 mm (in SB+RC fields) in the three days before the two harvests (on 16 August 2019, and 5 August 2019, respectively) may have contributed to increased fresh weight with additional water content.

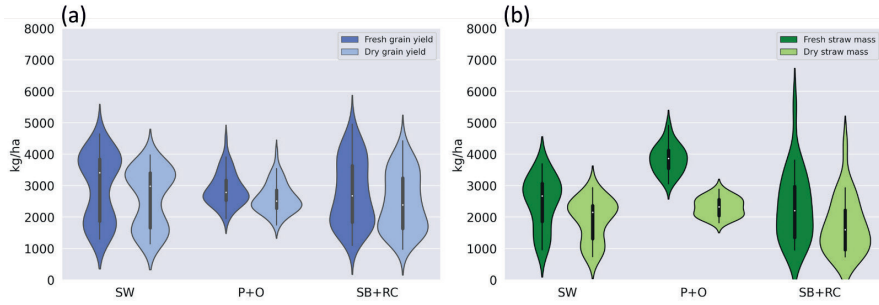

Figure 20. Violin plots of mean harvest results of fresh and dry (a) grain yield and (b) straw mass, grouped by spring wheat (SW), pea and oat mixture (P+O), and spring barley with under-sowing red clover (SB+RC) fields. White dots represent the median, while thick black bars in the centre demonstrate interquartile ranges, black lines represent the remainder of the distribution. The shape of the violins shows point density and data distribution as a whole.

5.3.2 The Hyperspectral Reflectance Signature under Various Agriculture Management Practises

Figure 21 displays a mean reflectance plot produced from hyperspectral data of SW, P+O, and SB+RC fields, with enclosed subsets categorized by (Figure 21-A) STM and (Figure 21-B) CM agricultural operations. Regarding agricultural management practices, the wavelength bands between 700-750 nm and 760-900 nm have significant identification capabilities, while the 400-700 nm region shows little differentiation between management practices. The cultivation method (figure 21-B) provides greater recognition ability (separation) in this range when compared to STM spectral information (figure 21-A). In terms of crop types, spring wheat monocropping seems to give a better ability to recognize AMPs, followed by mixed cropping systems SB+RC and P+O fields. However, since the focus of this study is on grain yield and biomass prediction, the narrowband VIs wave range were omitted based on the strong absorption bands near 760 nm were omited and thus excluded from subsequent AutoML analyses.

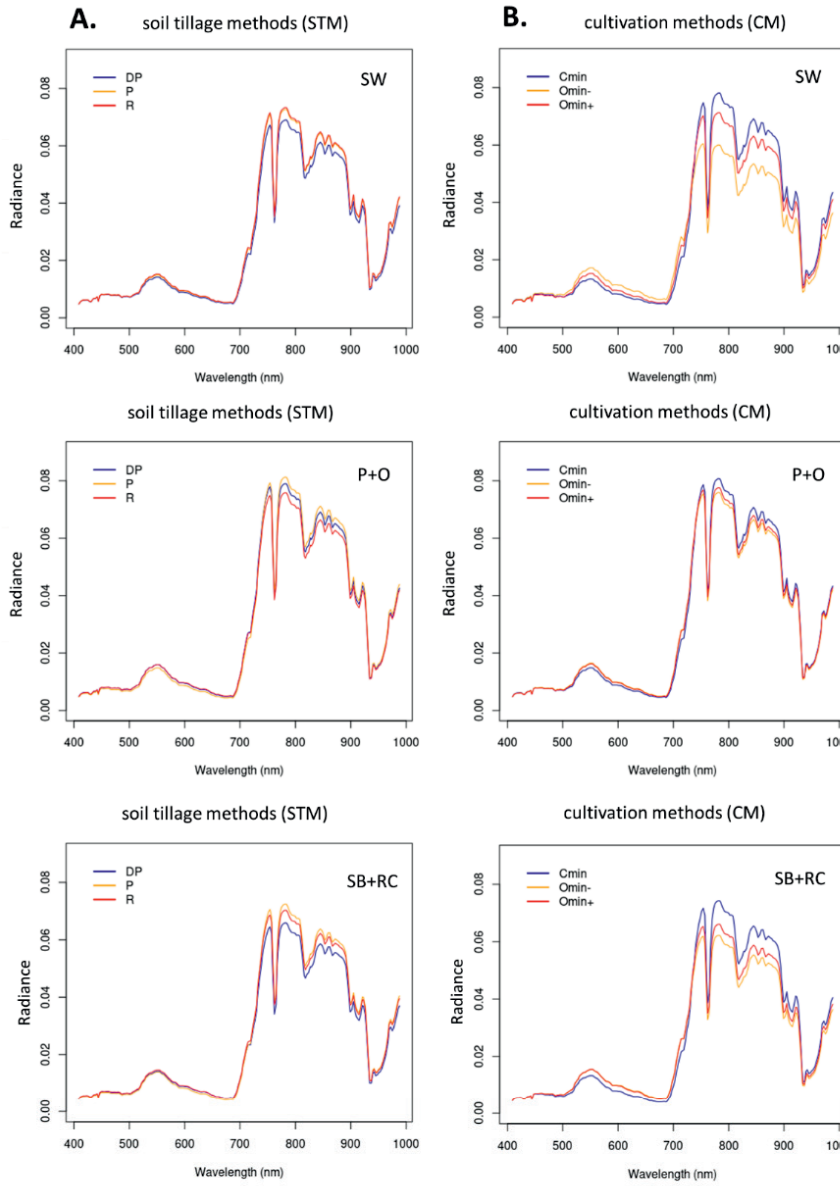

Figure 21. Mean radiance plot derived from hyperspectral data of spring wheat (SW), pea and oat mixture (P+O), and spring barley with under-sowing red clover (SB+RC) fields, grouped by (A) soil tillage method (STM) and (B) cultivation method (CM) farming operations with contained subsets. The wavelength ranges from the visible to near-infrared (VNIR, 400-1000 nm)

5.3.3 Characterization of the Correlation Coefficient with Averaged Radiance Hyperspectral Data and Field Observation

Correlation Coefficient (r) was used as exploratory analysis in this study and as a reference for subsequent modelling. Figure 22 shows the correlation coefficients (r) between each averaged hyperspectral narrow-band data with the dry mass (Figure 22-A) and fresh mass (Figure 22-B) at the plot level. The pattern of positive r values was typically obtained with reflectance between 750 – 940 nm wavelengths, whereas the strong negative correlation with reflectance was between 500 – 700 nm. Moreover, the correlation of straw mass (red line) was stronger than grain yield (blue line) at all fields in the 750-1000 nm range. By comparison, the results showed that, in the patterning of r curves, SW was closely associated with highly positive and negative r values in dry mass (Figure 22-A), while with the lower correlation nearby the oxygen absorption peak 760 nm. This tendency has been observed in the previous reflectance signature analysis as well. Among the three fields, P+O has the least correlation. Regarding the fresh mass (Figure 22-B), the correlation and spectral characteristics are comparable to the weight of the dry mass. Except for 740 - 750 nm, SW has overall the strongest correlation, followed by SB+RC, P+O.

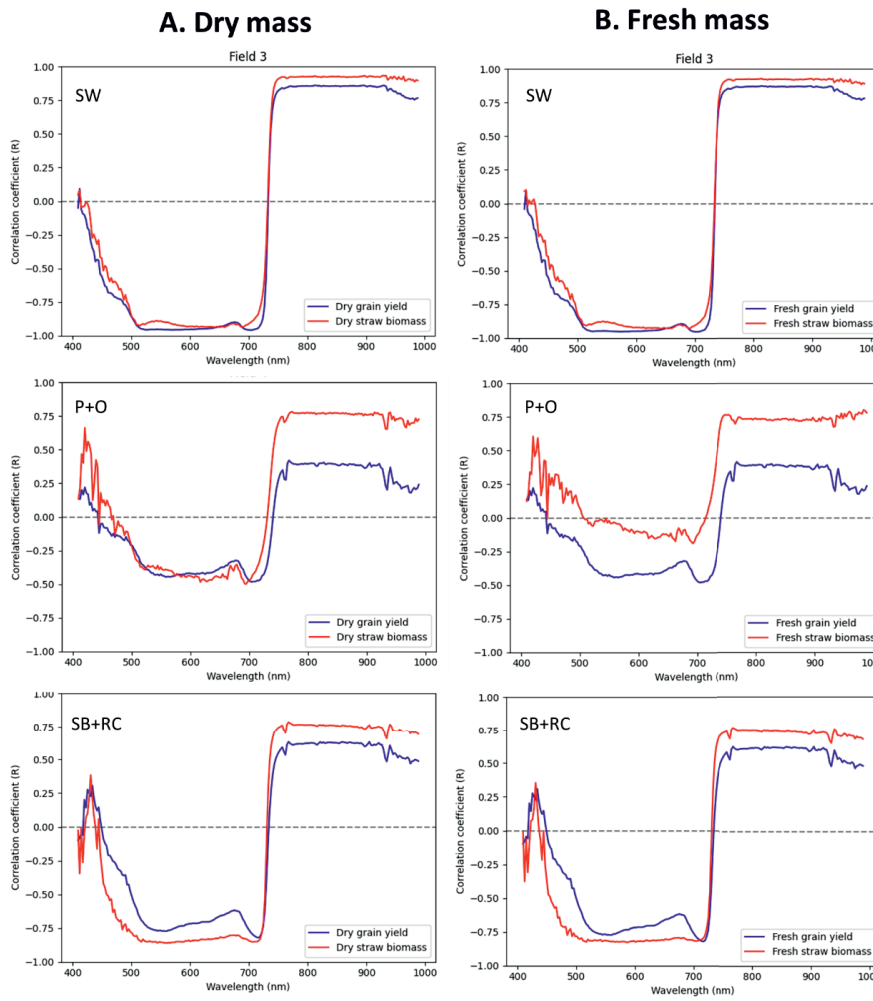

Figure 22. The Pearson Correlation Coefficient (r) between the field observation value [grain yield (A); Straw mass (B)] and averaged hyperspectral radiance at the plot level in SW, P+O, and SB+RC region.

5.3.4 The AutoML Model Prediction and Evaluation

In this study, the narrowband VIs reflectance of grain yield (n = 56) and straw mass (n = 24) based on training/testing (0.5/0.5) principles were used for AutoML modelling, respectively. The AutoML framework was used to test the appropriate combinations of data set parameters throughout the modelling process. Scatter plots representing model predictions and observed weight values (kg ha^{-1}) were compared to the

coefficient of the determination (R^2) and normalized root means square error (NRMSE) along with the 1:1 line.

Figure 23 shows the regression plots of fresh (Figure 23-A) grain yield (kg ha^{-1}) and (Figure 23-B) straw mass (kg ha^{-1}) in SW, P+O, and SB+RC fields based on narrowband VIs and AutoML methods. The results indicate that, in fresh grain yield (Figure 23-A), the AutoML model had the lowest prediction errors (NRMSE= 0.13) and the highest R^2 value (0.95) in SW field, followed by SB+RC field (NRMSE= 0.16, R^2=0.88), and P+O (NRMSE= 0.16, R^2=0.88). Even though the three models functioned well, there was a minor non-uniform bias found within the models, with an underestimation of grain yields in areas with higher output in SW and SB+RC fields. On the other hand, for fresh straw mass, the SW field remains the best performing among the other fields with (NRMSE= 0.16, R^2=0.88) followed by SB+RC field (NRMSE= 0.27, R^2=0.77) with uniform overestimation bias, and P+O (NRMSE= 0.25, R^2=0.56) (Figure 23-B). Among them, P+O prediction ability is insufficient, and the reference data collected are concentrated in the 3,000 to 5,000 (kg ha^{-1}) interval, which makes the regression model unable to be effectively extended.

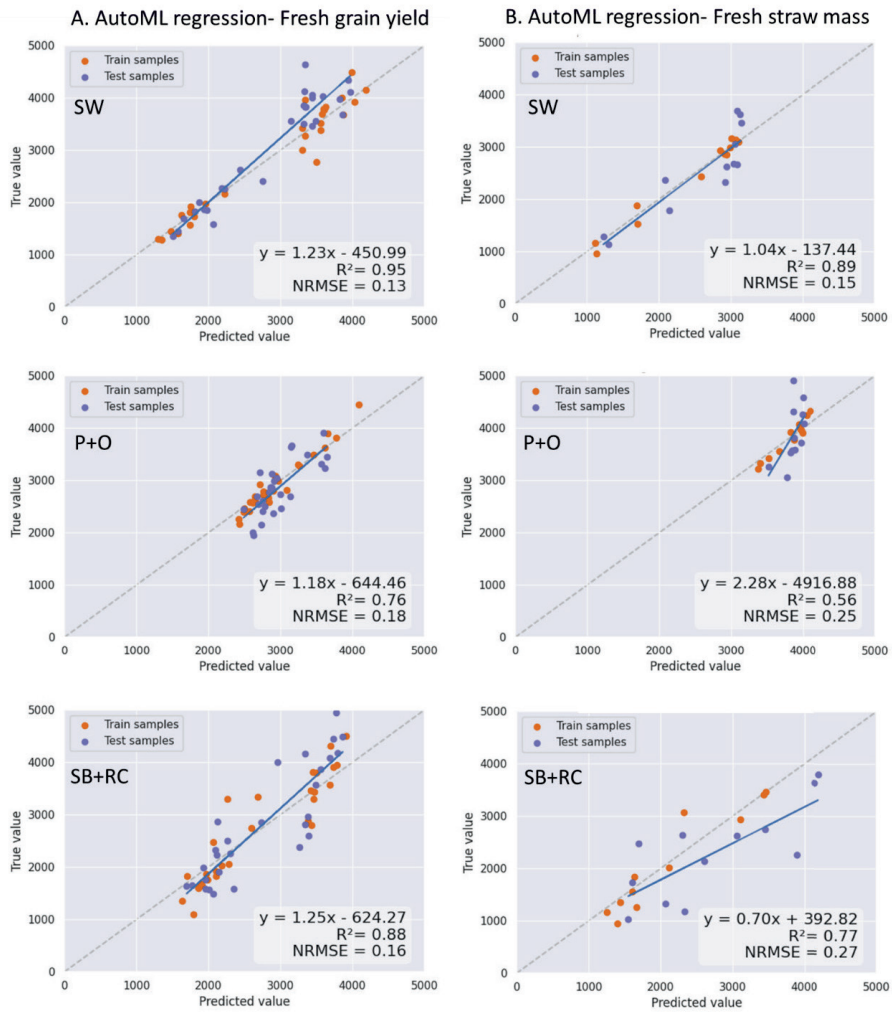

Figure 23. Regression plots of (a) fresh grain yield (kg ha^{-1}) and (b) fresh straw mass (kg ha^{-1}) in SW, P+O, and SB+RC fields based on narrowband VIs and AutoML methods. The horizontal axis in the scatter plots represents the model's projected grain yield or straw mass, while the vertical axis represents field-observed data. Where the R^2 = coefficient of determination, NRMSE = normalized root represents the squared error, while the 1:1 slope is shown by the black dotted line.

Figure 24 demonstrates the behaviour of predictive models utilizing dry (A) grain yield (kg ha^{-1}) and (B) straw mass (kg ha^{-1}) in SW, P+O, and SB+RC fields based on narrowband VIs and AutoML methods. The results specify that, in summary, SW yields the best performance for dry grain yield (NRMSE= 0.12, R^2=0.96) and straw mass (NRMSE= 0.15, R^2=0.89) among SB+RC, and P+O files (Figure 24-A). Compared to the fresh mass model, the dry performance is better in general, especially

in the dry straw model of SB+RC (NRMSE= 0.33, R^2=0.86), and P+O (NRMSE= 0.24, R^2=0.83) (Figure 24-B), although these two models have a larger degree of bias under the comparison of 1:1 slope.

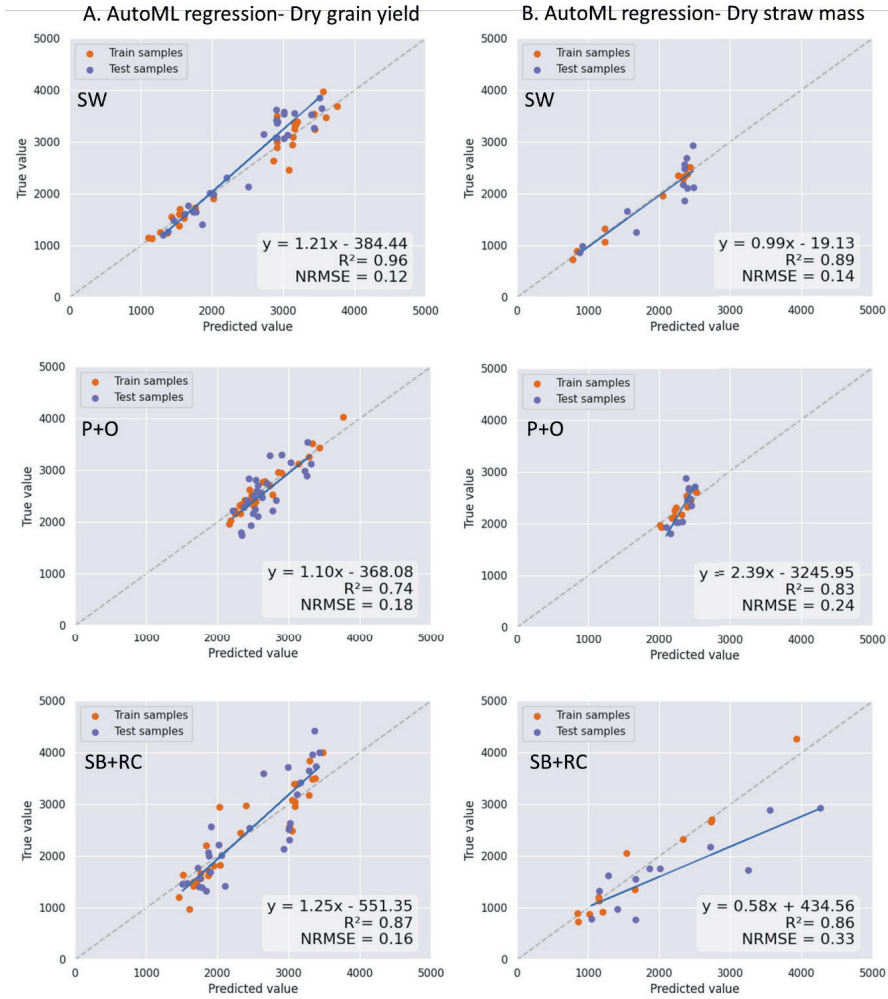

Figure 24. Regression plots of (a) dry grain yield (kg ha^{-1}) and (b) dry straw mass (kg ha^{-1}) in SW, P+O, and SB+RC fields based on narrowband VIs and AutoML methods. The horizontal axis in the scatter plots represents the model's projected grain yield or straw mass, while the vertical axis represents field-observed data. Where the R^2 = coefficient of determination, NRMSE = normalized root means squared error, and the black dotted line exemplifies the 1:1 slope.

5.3.5 The AutoML Model Pipeline Visualization

An interactive AutoML visualization tool PipelineProfiler (Ono et al., 2021) was used in this study (Figure 25). To simplify the description, the best regression modelling results across 2 crop fields (SW and SB+RC) were listed, and the evaluation performance of AutoML pipeline execution times set at 30 seconds, the primitive comparison against other regressors in the same pipeline, and real-time hyperparameter selections. The results confirmed that the best regressor found for dry grain yield was automatic relevance determination (Ard) Regression (Qi et al., 2004) for SW field (Figure 25-A), and for SB+RC field, it was the Random Forest (Breiman, 2001) (Figure 26-A), while for dry straw mass, it was Gaussian Process (Seeger, 2004) (Figure 25-B) for SW field, and Ard Regression for SB+RC field (Figure 26-B), with all hyperparameters found by AutoML also displayed in the figures.

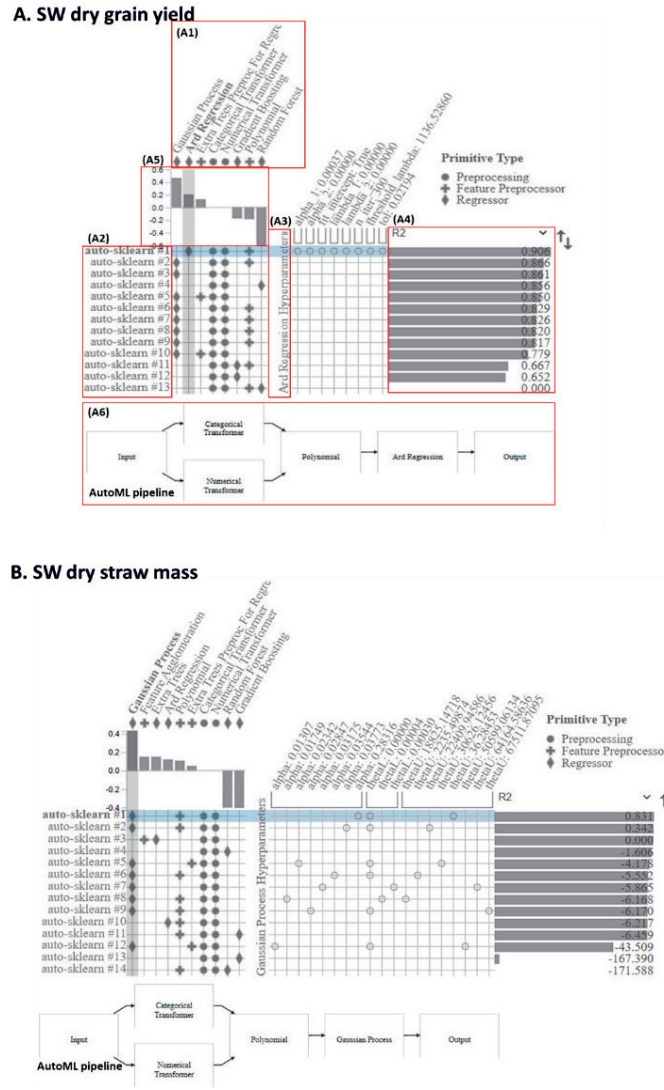

Figure 25. The interactive AutoML pipeline matrix plots with thirty-second running-time limits sorted by coefficient of determination (R^2) performance (**A**, and **B**). (**A**) Spring wheat (SW) dry grain yield pipeline matrix with the Top1 regressor, automatic relevance determination (Ard) regression, where (**A1**) illustrated Primitives (in columns) used by the pipelines (**A2**) the blue line (in rows) showed the best R^2 rank); (**A3**) one-hot-encoded hyperparameters (in columns) for the primitive across pipelines, (**A4**) R^2 performance ranking of AutoML pipelines; (**A5**) Primitive contribution view demonstrating the correlations between pipeline scores and primitive usage are displayed in A5. The Gaussian Process showed the highest correlation score regarding R^2 performance; (**A6**) Step by step AutoML Pipeline algorithm flowchart, where the box before the output represents the regressor of the model. (in A6 Ard regression as the regressor) (**B**) Spring wheat (SW) dry straw mass Field pipeline matrix with the Top1 regressor, Gaussian Process.

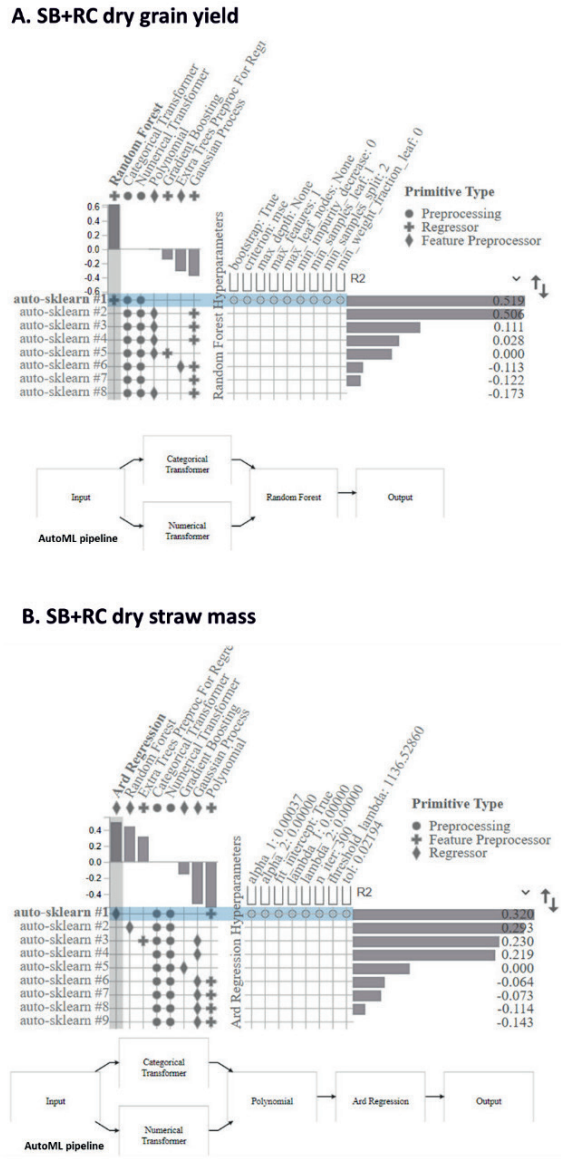

Figure 26. The interactive AutoML pipeline matrix plots with thirty-second running-time limits sorted by coefficient of determination (R^2) performance (A, and B). (A) spring barley with under-sowing red clover (SB+RC) dry grain yield pipeline matrix with the Top1 regressor, Random Forest. The rows display a blue line representing the best R^2 rank followed by its hyperparameters settings; (B) SB+RC dry straw mass pipeline matrix with the Top1 regressor, Ard regression, followed by its hyperparameters settings.

5.4. The Field Observation DM Data Analysis

Based on the AutoML models provided above (Figure 25 and Figure 26), a series of prediction maps were generated (Figure 27) for dry grain yield and straw mass for SW, P+O, and SB+RC experimental sites at the plot level. Furthermore, the SW and P+O fields' prediction capability were 60 days before the harvest date (18 Jane -16 August), whereas the SB+RC field's estimating was 49 days before harvest (18 Jane - 5 August).

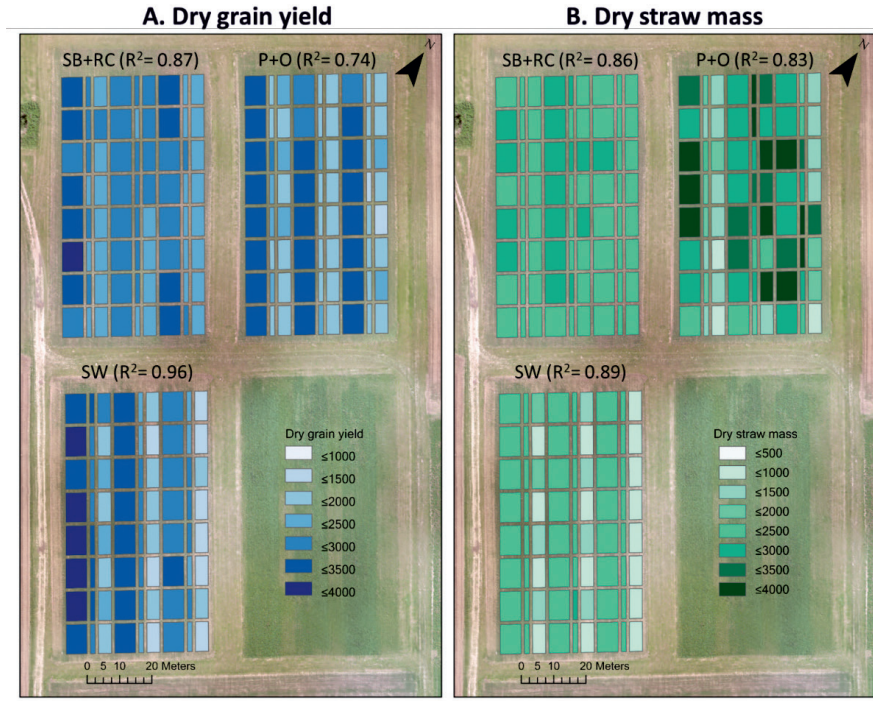

Figure 27. The spatial prediction mapping output of (a) dry grain yield (kg ha^{-1}) and (b) dry straw mass (kg ha^{-1}) in SW, P+O, and SB+RC fields based on their respective AutoML prediction models at the plot level. The performing coefficient of determination (R^2) is displayed in the previous results.

6. DISCUSSION

6.1. Applicability of this thesis

This thesis has presented a rapid, non-destructive, low-cost framework for field-based crop yield modelling as well as the recognition of the management practice. The prediction models covered three different agricultural operations (STM, CM, and MA) to represent the variable conditions in a practical farming system, which provided varied agricultural data to identify the robustness of the derived ML and AutoML models. Acquisition of UAV data conducted cover full growing periods in 2019 offered a wider range of suitable monitoring capabilities. Traditional machine learning techniques (RFR, SVM, and ANN) and SOTA AI-based, open-sourced AutoML framework for automatically exploring crop image regression and classification ability and assisting in optimizing problematic hyperparameter adjustments.

All VI information was derived from four multispectral and 216 hyperspectral bands under regular mono- and mixed cultivation. Utilizing various VIs which calculate the relative values or ratio among wavelengths can reduce the impact of radiance effects caused by individual reflectance spectra (Li et al., 2020). Besides, high-resolution multispectral imaging produces continuous and accurate indices in contrast to simple visual scores and rankings (J. Wang et al., 2019). Consequently, no additional sensors were needed, which reduced measurement errors and increased cost-efficiency. Additionally, in this thesis, the possibilities of deploying UAS multispectral-VIs systems were examined to monitor small-scale experimental fields and larger-scale farmland by employing hyperspectral imaging. The flexibility, in sensor systems, environmental capabilities, and increased flight durations could expand the application to meet a diverse range of requirements (Wachendorf et al., 2018), whilst still providing the precise yield prediction accuracy required in this study. Since the investigation was carried out under a diversity of agricultural management practices, the methods and findings can profoundly aid agronomists and farmers in designing accurate cropping systems to enhance environmental assessment.

Remote sensing related modelling research publications have significantly increased in recent years, with over a hundred articles developed since

2017. This substantial adoption of UAS and hyperspectral related approaches demonstrates its impact and the mounting interest in such research issues (Samaras et al., 2019). This framework may also be implemented in other classification and regression research, such as research employing multi-sensors [i.e., thermal, visible light, hyperspectral, radar or light detection and ranging (Lidar) sensors] across a range of contemporary agriculture classification activities (ie, weed management (David & Ballado, 2017; Torres-Sánchez et al., 2013), crop phenotyping (Chawade et al., 2019; Sankaran, Khot, Espinoza, et al., 2015; G. Yang et al., 2017; Young, 2019), disease monitoring (Vivaldini et al., 2019; X. Zhang et al., 2019); as well as research focused on ecological classification schemes, multispectral-based plant community mapping options (Villoslada et al., 2020), and coastal wetland vegetation classification results (Burnside et al., 2007).

6.2. The Impact of the Cultivated Period, Flight Times, and Farming Operations

The initial goal was to build red-clover biomass prediction models for one and two-year growing periods. Despite the results of the established models (Figure 12 and 13), the prediction accuracy of 1YC (R^2 ranges from 0.81 to 0.90, NRMSE ranges from 0.11 to 0.15) and 2YC (R^2 ranges from 0.84 to 0.89, NRMSE ranges from 0.11 to 0.15) were both adequate. Regardless, the clover-grass mix produced a heterogeneous canopy with the coverage of the two components. A previous mixed clover-grass study focused on canopy height (CH) modelling for DM yield prediction showed the models performed better when constructed independently among the two species, and cannot be easily shifted to other grassland types due to their structural properties (Grüner et al., 2019). A previous study found that although measurements performed at the ground-level were more accurate, the use of aerial systems was preferred since species identification was irrelevant when predicting the biomass of mixed-grass (Rueda-Ayala et al., 2019). Thus, these species-dependent VI events seem to be minor concerns within this study's findings. NIR-based Vis were inferred to be suitable for the estimation of DM yields in one- and two-year cultivation periods in this study.

The second objective was to compare the impact of different pre-harvest flight dates on model estimation capabilities. The choice of flight timing was crucially matched with the spectral reflectance data during various

growth periods. A previous study showed that the ideal period for forage crop assessment was one day before harvesting (Lussem et al., 2018), whereas another study suggested that the targeted silage harvesting stage was favoured (Viljanen et al., 2018b). The results from this study indicate that in clover-grass mixture fields, the estimation ability is improved when UAS imagery is collected closer to the harvest period. Moreover, the results indicate that the VIs derived from UAS images captured earlier than 38 days before harvest also have sufficient DM yield estimation capacity.

A study of grassland DM yield estimation by a UAV-RGB camera showed different nitrogen fertilizer levels with its R^2 ranging from 0.57 to 0.70 (Lussem et al., 2019). Currently, adequate fertilizer estimation remains challenging in heterogeneous plant communities such as grasslands (Lussem et al., 2018). However, legume crops could provide more positive N balance input than mineral fertilization under various tillage conditions [(Wittwer & van der Heijden, 2020). Although the input of N is not effectively quantified in this study, the DM yields of the various N input combinations could still be effectively predicted, which increases the viability of using non-destructive methods to quantify a range of, and distinct, N sources in future fertilizer management decisions.

6.3. The machine learning methods and Importance of Variable Rankings

6.3.1 The machine learning methods

Machine learning techniques are still deemed to be a novel in the realm of estimating grassland biomass (Ali et al., 2015; Bithas et al., 2019; Maimaitijiang et al., 2020; Maxwell et al., 2018). The predictive ability of three broadly adopted and reliably implemented ML methods in clover-grass DM yield was promising in this study. ANN showed better predictive accuracy eleven days prior to harvest (11DB). This result is consistent with the LOO cross-validation results. The practicality and flexibility of ANN has previously been demonstrated in studies of grassland biomass estimation(Ali et al., 2017; Xie et al., 2009), and nitrogen and phosphorus concentration modeling in mixed-species environments (Mutanga & Kumar, 2007; Mutanga & Skidmore, 2004a). Interestingly, within the study, RFR and SVR were shown to have increased predictive capability at 38DB; which is farther from the harvest period. Both RFR

and SVR were also shown to have a promising potential in clover-grass biomass prediction applications, since they are fast and require fewer training samples, when compared to the ANN (Ali et al., 2015; Z. Zhang et al., 2017). The overall accuracy of the three ML methods provided R^2 ranges from 0.81 to 0.90, and the NRMSE ranges from 0.11 to 0.15. These findings further support the asserted dominant ability of MLs as a perennial forage crop biomass estimator; demonstrated in this study for mixed-grass species.

6.3.2 Importance of Variable Rankings

Variable importance ranking is essential for predictor selection and model simplification normally. In this study, the results of ranking showed which VIs were able to capture most of the variability in vegetation characteristics from the grass fields. Different VI values at the level of leaf area indices were likely caused by the diverse canopy structures of clover (horizontal) and grass leaves (vertically orientated) (Biewer et al., 2009). A recent study confirmed that GNDVI is suitable as a biomass predictor for perennial forage crops, where $R^2 = 0.80$ for freshly-cut, and 0.66 for dry yields (Aube, 2021), as well as in the grain yield estimation in maize (Marques Ramos et al., 2020). These results resemble the RFR and ANN modelling of this study, where the GNDVI, GDVI, and MSR had the highest average contributions. The weight of SR was generally low. Similar result was also found in a study of grass DM yield prediction by Partial Least Square (PLS) and RF techniques, where the above-mentioned VIs were relatively important variables, while SR yielded the worst prediction out of twelve VIs (Grüner et al., 2020). Other previous studies have indicated that NDVI is more commonly used for pasture biomass measurements (Insua et al., 2019; Lee et al., 2015), as well as in larger-scale grassland followed by the seasonal monitoring (Q. Ma et al., 2019). However, the findings of this study indicate that NDVI may not be the most suitable VI, which was supported by the previous study (Aube, 2021). This highlights the importance of considering the saturation, sensitivity, stages of crop development, canopy structure and the type of environment when testing various vegetation indexes (Wachendorf et al., 2018). This study's findings indicate that multispectral information based on NIR and the green band may be more suitable for DM yield prediction using RFR and ANN modelling. The exception, however, is the SR indices, which has the highest contribution consistently across all periods in the SVR modelling. This distinctive finding has yet to

be found in similar crop studies in other literature. It is clear from the findings that more tests of VI should be conducted in studies to increase the collective understanding and provide improve the knowledge base.

6.4. The Impact of Algorithm Selection, Cultivated Period, and Crop types in AutoML AMP recognization

6.4.1 the Impact of the AutoML Method in UAS

The AutoML framework quickly provided usable classifiers and hyperparameter selections for unknown UAS classification tasks and parameter selection. For example, in the current study, the parameters and applicable classifiers of AMPs were unknown *a priori*. However, it provided a promising and efficient performance rating for classifiers for inclusion in modelling selection. As the results of Figure 19 show, LDA (Figure 19a and Figure 19c) and Extra Trees (Figure 19b) were chosen as the best classifiers corresponding to the VPT fields of the AMP recognition task. These ML methods have been less applied and referenced in the field of UAS (L. Ma et al., 2017a). These findings clearly illustrate that AutoML has the potential to locate alternative ML approaches that might customarily be ignored by investigators with unknown classification subjects.

In addtion, the operational efficiency of AutoML classifiers can be given a time limit and gives the researcher the flexibility to find the most suitable formula within the required time. In general, a longer time setting allows for increasingly accurate results with additional classifier combinations. Since the experiments did not involve substantially large datasets, the focus was put on time setting close to the minimum limit of AutoML calculation [60 sec of total CPU operation (this can be up to 3000 sec) and 10 seconds of a single ML algorithm computation] to highlight the flexibility and rapid performance of AutoML.

Finally, within this research, the latest released AutoML interactive visualization system *PipelineProfiler* was employed and assisted in the screening of classifiers and the reference of fine-tuning parameters when analyzing UAS data. This interaction includes adjustable time, accuracy ranking, and selection of hyperparameters in response to the requirement of customized UAS modelling. The results showed that AutoML computations within a 60-seconds-run produced between 11

and 12 pipelines (Figure 19), which might offer a beneficial foundation for providing adequate outcomes in most cases with minimal attempts and time.

6.4.2 The Impact of Algorithm Selection, Cultivated Period, and Crop types in AutoML AMP recognization

In terms of algorithm selection in the AMP classification results, different classifiers were suggested by AutoML as the best performances even within the same AMP category for different crop types (Figure 19). The conclusion is that applying AutoML in UAS-derived multispectral VI data allowed for the consideration of a variety of algorithm combinations to meet the complexity of the VPT field. The three most used ML algorithms (RF, SVM, and ANN) were compared in the UAS classification fields with AutoML algorithms (Table 8). The overall performance shows that AutoML (with 1200-second CPU duration) provided the five best (or equal best) accuracy performances (shown in bold black in Table 8). Interestingly, in all tests, the AutoML (1200) and RF methods were never found to be the worst-performing methods (shown in bold red). Moreover, when using the ANN method, despite providing five of the best classification accuracy results, this method also included five of the worst performance results. Similar outcomes were observed regarding the SVM and AutoML 60-seconds runs.

From the results, increasing the computing time was deduced to have the potential to improve the accuracy and stability of AutoML classification performance under certain AMPs conditions. However, it also highlights the potential to include AutoML methods in the computation of common classification problem-solving. Similar ranks were shown in a study that compared the results of the numerous classifiers with Auto-sklearn, where the RF classifier presented the strongest performance, and SVM showed robust performance for some datasets (Feurer, Klein, et al., 2015b). Since the Auto-Sklearn classifiers are based on Scikit-learn as a blueprint, it should theoretically capture the hyperparameters of the RF algorithm on what was selected for Table 8. Despite the strong performance of AutoML (1200 sec), there were still several results that indicated an inferior of AutoML (1200 sec) when compared to the RF classifier (i.e., Field 1 MA; Field 3 MA and STM). Moreover, in a few cases, the accuracy of AutoML (1200 sec) was even lower than the calculation result of the 60-sec set (i.e., Field 1 MA, Field 3 STM,

and Field 4 CM). It may be that the algorithm computations involve different factors other than accuracy, and the model it uses to tune the parameters actively tries to avoid overfitting. This will possibly lead to the situation where the most accurate model, on the testing or training data, will not be the one that can generalize the best on real data. Also, developers from the Auto-sklearn team have previously described that during the ensemble selection phase the methods can add numerous substandard models to the final ensemble and unregularized selection may lead to overfitting with a small number of candidate models (Feurer et al., 2018). This result shows that there is still room for improvement regarding AutoML calculation methods in the future.

In terms of cultivated period and crop type, according to the monthly performance of different crop growth stages, the PCA results indicate that the VPT with better clustering performance occurred during the flight in May; with a confidence level of 0.95 (Figure 16b). In this regard, this flight period was further used for the AMPs classification study. Conversely, in the case of more homogeneous crop types [field 3 (WS)], and despite promising as the classification result in CM, the results of MA and STM were not as effective as other crops (Figures 17 and 18). These results may suggest that even with higher heterogeneity of cultivation within the plots (i.e., F1, F3, and F4) appears to not necessarily affect the classification ability. However, concerning the Field 3 results from the PCA in May (stage of stem elongation) and July (stage of flowering), the MA clustering ability was better with a 0.95 confidence level in both months, and the accuracy was later improved from the classification analysis. The results of the study have demonstrated that, although the feature selection stage of AutoML is a black box, the potential predictive ability of the AutoML model based on PCA result can still be preliminerarily be discovered, and cost of period selection can be reduced; as was done in this study. In addition, this study has contributed evidence to the classification obstacles in the case of STM that may cause by the orientation of images taken over vegetation or soil with uniform texture and re-cursive pattern, sub-optimal flight configuration (Sona et al., 2016), or unflavored VIs selection. Some studies also suggest that the use of grey-level co-occurrence matrix (GLCM)-based texture information (David & Ballado, 2017; Kwak & Park, 2019), semantic segmentation (M. der Yang, Tseng, et al., 2020) or edge computing (M. der Yang, Boubin, et al., 2020a) can improve the accuracy of UAS-ML classification in the crop categories. This may

be an applicable technology for AMPs classification in the future. The applicability and optimization of this framework, and the visualization of feature importance, required the optimization of the AutoML programmers and UAS application feedback to improve.

Currently, multispectral indices were effectively applied in some AMP image analysis studies with the colour, texture and shape factors of the agricultural land at the satellite level. These include conservation tillage methods identification (Najafi et al., 2021) and agriculture landscapes with pixel-based or object-based classification tasks (Duro et al., 2012; H.-O. Kim & Yeom, 2012). AMPs application are indispensable for environmental monitoring and for facilitating the agricultural decision-making process, regarding the adoption practices proposed by growing conservation agriculture demand (Telles et al., 2018), and for its potential upscaling ability to accelerate land cover classification studies. Recently, combining commonly adopted management practice with UAS multispectral-VIs research has gradually gained attention and has been applied to cotton and sorghum fields (Yeom et al., 2019b). In this study, the effective application of UAS sensors to recognize multiple AMP categories has been shown. More specifically, an UAS-AutoML approach can improve the classification ability under specific crop AMPs, highlighting that, in this study site, classification performed better in CM, with overall classification performance followed by MA and STM.

6.5. The Impact of the AutoML Method in Hyperspectral Imaging

6.5.1 The Effect of Hyperspectral Signatures and the Correlation between Crop Yield and Straw Mass

The initial goal of this study was to conduct an exploratory evaluation of the hyperspectral reflectance signature and determine the ideal narrowband VIs for modelling common crop types and farming schedules in Northern Europe. To identify redundant bands and establish wavebands that could best help AutoML regression modelling, the VIs were first chosen based on prior knowledge of the literature and then filtered by the reflectance signature (Figure 21) and their Correlation Coefficients with yield and biomass (Figure 22). Although there was no general focus on a formal classification analysis in this current study,

the characteristics of hyperspectral data under different agricultural practices (i.e., STM, CM, and MA) are still worthy of attention.

Figure 21 reveals that, in general, because chlorophyll absorption is not limited to the centre wavelength but also affects adjacent bands, the reflectance values in the blue and red sections are significantly reduced, resulting in „absorption characteristics" in the spectral signature of the reflectance in all spectral results. In addition, all the reflection spectra show obvious absorption peaks at 760 nm. This spectral region is influenced by atmospheric oxygen (Riris et al., 2017) and therefore, this region was avoided while calculating VI's. Also from the results, the wavelength range 750-900 nm (NIR) has strong recognition capabilities based on the variation of reflection intensity, however, the 400-700 nm (visible bands) region was inefficient and offered little separation or discernment. The differentiation on spectra at the wavelength range of 750 – 900 nm suggests that the interior leaf structure, biochemical concentration, and water content of the target vegetation are different. A previous study pointed out that the diversity of NIR regions is usually caused by differences in internal leaf structure (A. A. Gitelson et al., 2003). While reflectance varies at the canopy level may be due to additional factors like LAI, canopy design, and backdrop soil (Darvishzadeh et al., 2008). These results will be valuable for further classification activities in agriculture management recognition.

The coefficients correlation (r) of each narrow-band with both grain yield and straw mass exhibited a similar pattern of r curves for both dry (Figure 22-A) and fresh weight (Figure 22-B) analysis, yet r in absolute values for the P+O field was observed to be less correlated than those for grain yield and straw mass, especially in the fresh weight. Since the P+O field was mixed cultivation and the source of weight is the sum of the two crops and the amount of precipitation before harvesting may indirectly bring about a lower degree of correlation. Interestingly, while the findings of these linear correlation tests all show that the straw mass has a stronger link with the spectrum, it does not depend on the empirical model's degree of fit (see Figure 23 and 24). Hence, it was discovered that grain yield (R^2) had a superior goodness-of-fit performance than straw mass in general, with lower NRMSE.

6.5.2 The Hyperspectral Narrowband VIs and AutoML modelling

Despite the opportunities afforded by hyperspectral systems to collect a multitude of spectrum data, extracting the relevant important wavelengths from a data cube can be challenging (Martínez-Usó et al., 2007). In this study, hyperspectral narrowband Vis were used as predictors for AutoML modelling. However, selecting narrowband VIs with spectrums that might be affected by atmospheric oxygen were avoided. With this in mind, the target VIs selected for analysis were extracted, calculated, and processed in the modelling stage, which reduced processing and storage demands.

Based on the empirical AutoML regression model, the estimation capacity of hyperspectral narrowband VIs was exceptional. The best coefficient of determination for mono-cultivated wheat was 0.96, for mixed peas and oats were 0.76, and for mixed legumes and spring barley was 0.88. In terms of straw mass estimation, they were 0.98, 0.83, and 0.86 respectively. The prediction ability of dry weight was discovered to be typically greater than that of fresh weight, especially in fields where mixed peas and oats, which was 27 per cent higher. This demonstrates that the crop water content has an influence on the model's estimation outputs to a certain extent.

According to a previous study, spectral measurements were taken during the Tillering II and Heading phases in wheat yielded the best results for estimating biophysical factors using narrowband VIs (Xavier et al., 2006). This is consistent with the recommended flight time. In addition, different band combinations can be effectively utilized since crop circumstances change according to factors such as management conditions, soil characteristics. Others have demonstrated that piecewise multiple regression models on narrow bands provide for greater flexibility in selecting the bands that provide the most information at a given stage of crop development (Thenkabail et al., 2000). This viewpoint has also been confirmed in this research.

6.5.3 The AutoML Method's Applicability and Impact in Hyperspectral Imaging

In this study, an AutoML framework to assist in self-regulating, instinctive regression operations, as well as enhancing challenging hyperparameter adjustments was employed. This method advances the use of hyperspectral imaging in farm-scaled environmental and crop phenotypic activity and possesses several advantages.

A general consideration is the flexibility of implementation. With the ever-increasing variability of remote sensing systems and the requirement for empirical model choices, the constraints of adjusting unidentified background parameters are being addressed. This means that many existing models that have been under-optimized in the past now have the chance to be re-modelled using artificial intelligence-based machines to relearn the performance tasks.

Another pertinent point regarding the alleviation of learning costs. Experience tells us that computer learning for remote-sensed images frequently necessitates a large number of samples and a lengthy learning period, i.e., Deep learning (Lecun et al., 2015; M. der Yang, Boubin, et al., 2020b; M. der Yang, Tseng, et al., 2020). This is incompatible with conventional agricultural experimental sampling procedures, which are limited by personnel, the complexity of the experiment design, and the number of repetitions. While, AutoML practices the Random Forest (RF) method (Breiman, 2001) for fast cross-validation, testing one-fold at a time and weeding out underperformance hyperparameter choices, for example, the combined algorithm selection and hyperparameter optimization (CASH) problems (Feurer, Klein, et al., 2015a). It boasts novel pipeline operators that increase the goodness of fit of datasets significantly. The RF approach is well-known for assessing lower sample sizes and increasing the performance of small datasets (Breiman, 2001; Luan et al., 2020). In addition, the AutoML framework quickly provided promising regressors and hyperparameter selections. In this research, each run of the regression model only took thirty seconds of learning time. This considerably improves learning efficiency, and the ability to find an appropriate formula in the time allotted and reduces the requirement for machine learning expertise (Feurer et al., 2020; Feurer & Hutter, 2018).

More specific consideration is the capacity for innovation. It is noticeable that random forest (RF), support vector machine (SVM), and artificial neural network (ANN) algorithms are among the most widely employed ML techniques in a wide range of recent remote sensing-based studies (L. Ma et al., 2017b). Their practicality and performance have been confirmed by many, but equally, there are still other similarly applicable ML methods that may have been shelved. As shown in Figures 25 and 26, the Ard regressors (Mackay, 1995; Neal, 1996) and Gaussian Processors (Seeger, 2004) were chosen as the best regressors for the grain yield and biomass tasks. These algorithms have received less attention and reference in remote sensing studies. These results indicate that AutoML can uncover alternative ML methods that would otherwise be overlooked by investigators when working with regression subjects.

6.6. The Limitations of this thesis

The location, soil types, chosen crop categories, and varieties present may be restricted in this study. A potential solution worth pursuing may be to increase the VPT sampling size and/or enhance the segmentation number of each plot; ultimately increasing the training samples for AutoML calculation. In addition, it is important to note that yield comparisons under different agricultural management approaches were not considered since the intricacy of the experimental design may have led to inadequate sampling numbers; as well as possible interaction effects. However, efforts to achieve a wide-ranging and well-considered predictor collection, through a variety of VI combinations, may lead to performance improvements. In addition, due to the limits of the current Auto-Sklearn system, not all regressors performed could be backtracked in this research to explore the individual feature importance ranking of VIs, which limits their capacity to aid in the selection of suitable VIs. However, attempts to provide a wide range of selectable VIs (19 selected VIs from AMP classification study) and continuous bands (216 narrow bands from airborne-hyperspectral yield and straw mass estimation) resulted in improved performance.

7. CONCLUSIONS

Firstly, in response to hypothesis 1, performing multispectral-UAS flights under the one- and two-year cultivated red clover-grass mixture performance trials within 38 to 11 days before the harvesting with the GSD 10 cm and combined with VI and multiple ML methods is promising in the DM yield prediction on a farm-scale with non-destructive and cost-effective way. The ML analysis results showed the best performance for ANN, followed by RFR, and SV. For VI performance, GNDVI and GDVI, and MSR performed well as predictors in ANN and RFR. However, the prediction ability of models was influenced by the farming operations. The stratified sampling based on STM had a better model performance than CM and MA. The results support the sufficiency of UAS to deal with complex experimental design development, such as tillage methods and various fertilizer inputs. However, the robustness and applicability of fertilizer quantification and the mixed legume-grass species distribution detection remain to be completely addressed in the future.

Secondly, in response to hypotheses 2 and 3, this study demonstrated a novel UAS-multispectral imaging technology and a state-of-the-art AutoML framework across multiple AMP tasks through non-destructive and cost-effective approaches. The scientific merit of this article lies in utilising artificial intelligence to replace the judgment of the human for UAS-multispectral classification analysis with its automated data pre-processing, model selection, feature engineering, and hyperparameter optimization capabilities. Furthermore, it provides innovative insights into agricultural management practices and accelerates the intellectualized progress of the in-field monitoring UAS system and establishes future crop phenotyping abilities. In our study, AutoML embodied "learning how to learn" for any given UAS subject; and it is the first study of its kind to apply an auto-learning system for AMP classification tasks in multispectral-derived VI data. In addition, we compared AutoML performance with those of three widely used ML methods. The ML comparison analysis results showed AutoML achieved the most overall classification accuracy numbers after 1200 seconds of calculation and without any of the worst-performing classifications of the given datasets. In terms of AMPs classification, the best recognized period

for data capture occurred in the crop vegetative growth stage (in May of Estonia).

Thirdly, in response to hypothesis 4, our study highlights the capability of hyperspectral analysis for yield and biomass prediction in complex design fields through the use of two significant open-sourced software systems: the R language hyperspectral processing package and Python's Auto-Sklearn machine learning technology. The performance evaluation with several types of hyperspectral vegetation indicators we employed to characterize crop production and straw mass was satisfactory. We suggest they can be further applied to other crop biophysical characteristics. The VIs we suggest, as well as automatic narrowband VI calculation, might minimize data redundancy and cleaning time, as well as the computational power hardware requirements. It is also envisaged that further agricultural cultivation practices could be classified using hyperspectral imaging in the NIR spectral region (750-900 nm) with considerable discernible changes in reflectance spectra.

In summary, this thesis focused on the integration and implementation of the multispectral and hyperspectral imaging and AutoML framework approach with various crop types under multifunctional agriculture management fields in response to crop biomass/yield estimation. Under common crops and cultivation in most Nordic countries, it will provide agricultural decision-makers with professional yield estimation and sustainable agricultural management advice; and the integration of remote sensing technologies, geoprocessing methods, and automatic systems are vital tools for increasing the knowledge of plant-environment interactions within the management of crops. The study also reveals that the anticipated yield may be advanced two months before harvest. That is, spring wheat, spring barley, and oat were approximately in the booting to heading stage, for field pea around the reproductive growth stages, whereas the red clover field was in the flowing stage (49 days before in our case). The emergence of the AutoML system has helped to increase the application and effectiveness of remote sensing-based data analysis technology. However, more research and experiments will be required in the future to advance and validate the automatic learning framework's true potential and usage.

REFERENCES

Abuleil, A. M., Taylor, G. W., & Moussa, M. (2015). An Integrated System for Mapping Red Clover Ground Cover Using Unmanned Aerial Vehicles: A Case Study in Precision Agriculture. *Proceedings -2015 12th Conference on Computer and Robot Vision, CRV 2015*. https://doi.org/10.1109/CRV.2015.43

Ali, I., Cawkwell, F., Dwyer, E., & Green, S. (2017). Modeling Managed Grassland Biomass Estimation by Using Multitemporal Remote Sensing Data-A Machine Learning Approach. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. https://doi.org/10.1109/JSTARS.2016.2561618

Ali, I., Greifeneder, F., Stamenkovic, J., Neumann, M., & Notarnicola, C. (2015). Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data. In *Remote Sensing*. https://doi.org/10.3390/rs71215841

Andrade, J. F., Rattalino Edreira, J. I., Mourtzinis, S., Conley, S. P., Ciampitti, I. A., Dunphy, J. E., Gaska, J. M., Glewen, K., Holshouser, D. L., Kandel, H. J., Kyveryga, P., Lee, C. D., Licht, M. A., Lindsey, L. E., McClure, M. A., Naeve, S., Nafziger, E. D., Orlowski, J. M., Ross, J., … Grassini, P. (2019). Assessing the influence of row spacing on soybean yield using experimental and producer survey data. *Field Crops Research*, *230*. https://doi.org/10.1016/j.fcr.2018.10.014

Annicchiarico, P., Barrett, B., Brummer, E. C., Julier, B., & Marshall, A. H. (2015). Achievements and Challenges in Improving Temperate Perennial Forage Legumes. *Critical Reviews in Plant Sciences*, *34*, 327–380. https://doi.org/10.1080/07352689.2014.898462

Anwar, M. Z., Kaleem, Z., & Jamalipour, A. (2019). Machine Learning Inspired Sound-Based Amateur Drone Detection for Public Safety Applications. *IEEE Transactions on Vehicular Technology*, *68*(3). https://doi.org/10.1109/TVT.2019.2893615

Araus, J. L., & Cairns, J. E. (2014). Field high-throughput phenotyping: The new crop breeding frontier. In *Trends in Plant Science* (Vol. 19, Issue 1). https://doi.org/10.1016/j.tplants.2013.09.008

Arturi, M. J., Aulicino, M. B., Ansín, O., Gallinger, G., & Signorio, R. (2012). Combining Ability in Mixtures of Prairie Grass and Clovers.

American Journal of Plant Sciences, *03*(10), 1355–1360. https://doi.org/10.4236/ajps.2012.310163

Ashapure, A., Jung, J., Yeom, J., Chang, A., Maeda, M., Maeda, A., & Landivar, J. (2019). A novel framework to detect conventional tillage and no-tillage cropping system effect on cotton growth and development using multi-temporal UAS data. *ISPRS Journal of Photogrammetry and Remote Sensing*. https://doi.org/10.1016/j.isprsjprs.2019.04.003

Aube, L. (2021). *Estimation of forage biomass and vegetation cover in grasslands using UAV imagery*. 1–18. https://doi.org/10.1371/journal.pone.0245784

Bajcsy, P., & Groves, P. (2004). Methodology for hyperspectral band selection. In *Photogrammetric Engineering and Remote Sensing* (Vol. 70, Issue 7). https://doi.org/10.14358/PERS.70.7.793

Ballester, C., Brinkhoff, J., Quayle, W. C., & Hornbuckle, J. (2019). Monitoring the effects ofwater stress in cotton using the green red vegetation index and red edge ratio. *Remote Sensing*, *11*(7). https://doi.org/10.3390/RS11070873

Beleites, C., & Sergo, V. (2012). hyperSpec: a package to handle hyperspectral data sets in R. *Journal of Statistical Software*.

Bender, A., & Tamm, S. (2018). Seed yield of tetraploid red clover as influenced by cover crop management. *Zemdirbyste*, *105*(2), 133–140. https://doi.org/10.13080/z-a.2018.105.017

Bengfort, B., & Bilbro, R. (2019). Yellowbrick: Visualizing the Scikit-Learn Model Selection Process. *Journal of Open Source Software*, *4*(35). https://doi.org/10.21105/joss.01075

Biewer, S., Erasmi, S., Fricke, T., & Wachendorf, M. (2009). Prediction of yield and the contribution of legumes in legume-grass mixtures using field spectrometry. *Precision Agriculture*. https://doi.org/10.1007/s11119-008-9078-9

Bithas, P. S., Michailidis, E. T., Nomikos, N., Vouyioukas, D., & Kanatas, A. G. (2019). A survey on machine-learning techniques for UAV-based communications. In *Sensors (Switzerland)* (Vol. 19, Issue 23). https://doi.org/10.3390/s19235170

Blackburn, G. A. (1998). Spectral indices for estimating photosynthetic pigment concentrations: A test using senescent

tree leaves. *International Journal of Remote Sensing*, *19*(4). https://doi.org/10.1080/014311698215919

Board, E. L. (2018). *MODERNIZATION OF THE ESTONIAN NATIONAL GNSS REFERENCE STATION NETWORK*. *44*(2), 55–62.

Boelt, B., Julier, B., Karagić, Đ., & Hampton, J. (2015). Legume Seed Production Meeting Market Requirements and Economic Impacts. *Critical Reviews in Plant Sciences*, *34*, 412–427. https://doi.org/10.1080/07352689.2014.898477

Bonesmo, H., Kaspersen, K., & Kjersti Bakken, A. (2004). Evaluating an image analysis system for mapping white clover pastures. *Acta Agriculturae Scandinavica Section B: Soil and Plant Science*, *54*(2). https://doi.org/10.1080/09064710410024462

Boyd, K., Eng, K. H., & Page, C. D. (2013). Area under the precision-recall curve: Point estimates and confidence intervals. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *8190 LNAI*(PART 3). https://doi.org/10.1007/978-3-642-40994-3_29

Breiman, L. (2001). Random forests. *Machine Learning*. https://doi.org/10.1023/A:1010933404324

Burnside, N. G., Joyce, C. B., Puurmann, E., & Scott, D. M. (2007). Use of vegetation classification and plant indicators to assess grazing abandonment in Estonian coastal wetlands. *Journal of Vegetation Science*, *18*(5). https://doi.org/10.1658/1100-9233(2007)18[645:uovcap]2.0.co;2

Carvell, C., Roy, D. B., Smart, S. M., Pywell, R. F., Preston, C. D., & Goulson, D. (2006). Declines in forage availability for bumblebees at a national scale. *Biological Conservation*. https://doi.org/10.1016/j.biocon.2006.05.008

Changchun Li, Chunyan Ma, Haojie Pei, Haikuan Feng, Jinjin Shi, Yilin Wang, Weinan Chen, Yacong Li, Xiaowei Feng, & Yonglei Shi. (2020). Estimation of Potato Biomass and Yield Based on Machine Learning from Hyperspectral Remote Sensing Data. *Journal of Agricultural Science and Technology B*, *10*(4). https://doi.org/10.17265/2161-6264/2020.04.001

Chao, Z., Liu, N., Zhang, P., Ying, T., & Song, K. (2019). Estimation methods developing with remote sensing information for energy

crop biomass: A comparative review. In *Biomass and Bioenergy*. https://doi.org/10.1016/j.biombioe.2019.02.002

Chawade, A., Armonienė, R., Berg, G., Brazauskas, G., Frostgård, G., Geleta, M., Gorash, A., Henriksson, T., Himanen, K., Ingver, A., Johansson, E., Jørgensen, L. N., Koppel, M., Koppel, R., Makela, P., Ortiz, R., Podyma, W., Roitsch, T., Ronis, A., … Weih, M. (2018). A transnational and holistic breeding approach is needed for sustainable wheat production in the Baltic Sea region. In *Physiologia Plantarum* (Vol. 164, Issue 4). https://doi.org/10.1111/ppl.12726

Chawade, A., van Ham, J., Blomquist, H., Bagge, O., Alexandersson, E., & Ortiz, R. (2019). High-throughput field-phenotyping tools for plant breeding and precision agriculture. *Agronomy*, *9*(5), 1–18. https://doi.org/10.3390/agronomy9050258

Chen, J. M. (1996). Evaluation of vegetation indices and a modified simple ratio for boreal applications. *Canadian Journal of Remote Sensing*. https://doi.org/10.1080/07038992.1996.10855178

Chen, P. F., Tremblay, N., Wang, J. H., Vigneault, P., Huang, W. J., & Li, B. G. (2010). New index for crop canopy fresh biomass estimation. *Guang Pu Xue Yu Guang Pu Fen Xi/Spectroscopy and Spectral Analysis*, *30*(2). https://doi.org/10.3964/j.issn.1000-0593(2010)02-0512-06

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, *21*(1). https://doi.org/10.1186/s12864-019-6413-7

Chlingaryan, A., Sukkarieh, S., & Whelan, B. (2018). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. In *Computers and Electronics in Agriculture* (Vol. 151). https://doi.org/10.1016/j.compag.2018.05.012

Choudhury, M. R., Das, S., Christopher, J., Apan, A., Chapman, S., Menzies, N. W., & Dang, Y. P. (2021). Improving biomass and grain yield prediction of wheat genotypes on sodic soil using integrated high-resolution multispectral, hyperspectral, 3d point cloud, and machine learning techniques. *Remote Sensing*, *13*(17). https://doi.org/10.3390/rs13173482

Colombo, R., Bellingeri, D., Fasolini, D., & Marino, C. M. (2003). Retrieval of leaf area index in different vegetation types using high

resolution satellite data. *Remote Sensing of Environment*, *86*(1). https://doi.org/10.1016/S0034-4257(03)00094-4

Costa, J. M., Marques da Silva, J., Pinheiro, C., Barón, M., Mylona, P., Centritto, M., Haworth, M., Loreto, F., Uzilday, B., Turkan, I., & Oliveira, M. M. (2019). Opportunities and Limitations of Crop Phenotyping in Southern European Countries. *Frontiers in Plant Science*, *10*(September), 1–16. https://doi.org/10.3389/fpls.2019.01125

Crews, T. E., & Peoples, M. B. (2004a). Legume versus fertilizer sources of nitrogen: Ecological tradeoffs and human needs. *Agriculture, Ecosystems and Environment*, *102*(3). https://doi.org/10.1016/j.agee.2003.09.018

Crews, T. E., & Peoples, M. B. (2004b). Legume versus fertilizer sources of nitrogen: Ecological tradeoffs and human needs. *Agriculture, Ecosystems and Environment*. https://doi.org/10.1016/j.agee.2003.09.018

Crippen, R. E. (1990). Calculating the vegetation index faster. *Remote Sensing of Environment*, *34*(1). https://doi.org/10.1016/0034-4257(90)90085-Z

Cross, M. D., Scambos, T., Pacifici, F., & Marshall, W. E. (2019). Determining Effective Meter-Scale Image Data and Spectral Vegetation Indices for Tropical Forest Tree Species Differentiation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(8). https://doi.org/10.1109/JSTARS.2019.2918487

Darvishzadeh, R., Skidmore, A., Atzberger, C., & van Wieren, S. (2008). Estimation of vegetation LAI from hyperspectral reflectance data: Effects of soil type and plant architecture. *International Journal of Applied Earth Observation and Geoinformation*, *10*(3). https://doi.org/10.1016/j.jag.2008.02.005

Datt, B. (1998). Remote sensing of chlorophyll a, chlorophyll b, chlorophyll a+b, and total carotenoid content in eucalyptus leaves. *Remote Sensing of Environment*, *66*(2). https://doi.org/10.1016/S0034-4257(98)00046-7

David, L. C. G., & Ballado, A. H. (2017). Vegetation indices and textures in object-based weed detection from UAV imagery. *Proceedings - 6th IEEE International Conference on Control System, Computing and Engineering, ICCSCE 2016*. https://doi.org/10.1109/ICCSCE.2016.7893584

DeLonge, M. S., Owen, J. J., & Silver, W. L. (2014). Greenhouse Gas Mitigation Opportunities in California Agriculture: Review of

California Rangeland Emissions and Mitigation Potential. In *NI GGMOCA R 4*.

Demmig-Adams, B., & Adams, W. W. (1992). Photoprotection and other responses of plants to high light stress. *Annual Review of Plant Physiology and Plant Molecular Biology*, *43*(1). https://doi.org/10.1146/annurev.pp.43.060192.003123

Desta, B. T., Gezahegn, A. M., & Tesema, S. E. (2021). Impacts of tillage practice on the productivity of durum wheat in Ethiopia. *Cogent Food and Agriculture*, *7*(1). https://doi.org/10.1080/23311932.2020.1869382

Dong, T., Meng, J., Shang, J., Liu, J., & Wu, B. (2015). Evaluation of Chlorophyll-Related Vegetation Indices Using Simulated Sentinel-2 Data for Estimation of Crop Fraction of Absorbed Photosynthetically Active Radiation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *8*(8). https://doi.org/10.1109/JSTARS.2015.2400134

Doyle, C. J., & Topp, C. F. E. (2004). The economic opportunities for increasing the use of forage legumes in north European livestock systems under both conventional and organic management. *Renewable Agriculture and Food Systems*, *19*(01), 15–22. https://doi.org/10.1079/rafs200355

Drinkwater, L. E., & Snapp, S. S. (2007). Nutrients in Agroecosystems: Rethinking the Management Paradigm. In *Advances in Agronomy* (Vol. 92). https://doi.org/10.1016/S0065-2113(04)92003-2

Duro, D. C., Franklin, S. E., & Dubé, M. G. (2012). A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sensing of Environment*, *118*, 259–272. https://doi.org/10.1016/j.rse.2011.11.020

ESRI. (2016). ArcGIS PRO : Essential Workflows. *Esri*.

Fanigliulo, R., Antonucci, F., Figorilli, S., Pochi, D., Pallottino, F., Fornaciari, L., Grilli, R., & Costa, C. (2020). Light drone-based application to assess soil tillage quality parameters. *Sensors (Switzerland)*. https://doi.org/10.3390/s20030728

FAO. (2006). World Reference Base for Soil Resources. World Soil Resources Report 103 . Rome. In *Food and Agriculture Organization of the United Nations (2006), pp. 132*.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8). https://doi.org/10.1016/j.patrec.2005.10.010

Feng, H., Tao, H., Zhao, C., Li, Z., & Yang, G. (n.d.). *Comparison of UAV RGB imagery and hyperspectral remote-sensing data for monitoring winter-wheat growth*.

Feng, J., Jiao, L., Liu, F., Sun, T., & Zhang, X. (2016). Unsupervised feature selection based on maximum information and minimum redundancy for hyperspectral images. *Pattern Recognition*, *51*. https://doi.org/10.1016/j.patcog.2015.08.018

Fernández-Manso, A., Fernández-Manso, O., & Quintano, C. (2016). SENTINEL-2A red-edge spectral indices suitability for discriminating burn severity. *International Journal of Applied Earth Observation and Geoinformation*, *50*. https://doi.org/10.1016/j.jag.2016.03.005

Feurer, M., Eggensperger, K., Falkner, S., Lindauer, M., & Hutter, F. (2018). Practical Automated Machine Learning for the AutoML Challenge 2018. *International Workshop on Automatic Machine Learning at ICML*, *Section 2*.

Feurer, M., Eggensperger, K., Falkner, S., Lindauer, M., & Hutter, F. (2020). Auto-Sklearn 2.0: The Next Generation. In *arXiv*.

Feurer, M., & Hutter, F. (2018). Towards Further Automation in AutoML. *ICML 2018 AutoML Workshop*.

Feurer, M., Klein, A., Eggensperger, K., Springenberg, J. T., Blum, M., & Hutter, F. (2015a). Efficient and robust automated machine learning. *Advances in Neural Information Processing Systems*, *2015-January*.

Feurer, M., Klein, A., Eggensperger, K., Springenberg, J. T., Blum, M., & Hutter, F. (2015b). Efficient and robust automated machine learning. *Advances in Neural Information Processing Systems*.

Feurer, M., Springenberg, J. T., & Hutter, F. (2015). Initializing Bayesian hyperparameter optimization via meta-learning. *Proceedings of the National Conference on Artificial Intelligence*, *2*.

FISHER, R. A. (1936). THE USE OF MULTIPLE MEASUREMENTS IN TAXONOMIC PROBLEMS. *Annals of Eugenics*, *7*(2). https://doi.org/10.1111/j.1469-1809.1936.tb02137.x

Franceschi, L., Frasconi, P., Salzo, S., Grazzi, R., & Pontil, M. (2018). Bilevel programming for hyperparameter optimization and meta-learning. In *arXiv*.

Frazier, A. E. (2015). Landscape heterogeneity and scale considerations for super-resolution mapping. *International Journal of Remote Sensing*, *36*(9). https://doi.org/10.1080/2150704X.2015.1040130

Gardner, J. B., & Drinkwater, L. E. (2009). The fate of nitrogen in grain cropping systems: A meta-analysis of 15N field experiments. *Ecological Applications*. https://doi.org/10.1890/08-1122.1

Ge, Y., Chen, Y., Stein, A., Li, S., & Hu, J. (2016). Enhanced Subpixel Mapping with Spatial Distribution Patterns of Geographical Objects. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(4). https://doi.org/10.1109/TGRS.2015.2499790

Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, *63*(1). https://doi.org/10.1007/s10994-006-6226-1

Gianelle, D., & Vescovo, L. (2007). Determination of green herbage ratio in grasslands using spectral reflectance. Methods and ground measurements. *International Journal of Remote Sensing*, *28*(5). https://doi.org/10.1080/01431160500196398

Gianelle, D., Vescovo, L., Marcolla, B., Manca, G., & Cescatti, A. (2009). Ecosystem carbon fluxes and canopy spectral reflectance of a mountain meadow. *International Journal of Remote Sensing*, *30*(2). https://doi.org/10.1080/01431160802314855

Gitelson, A. A. (2004). Wide Dynamic Range Vegetation Index for Remote Quantification of Biophysical Characteristics of Vegetation. *Journal of Plant Physiology*, *161*(2). https://doi.org/10.1078/0176-1617-01176

Gitelson, A. A., Gritz, Y., & Merzlyak, M. N. (2003). Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology*, *160*(3). https://doi.org/10.1078/0176-1617-00887

Gitelson, A. A., Kaufman, Y. J., & Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment*. https://doi.org/10.1016/S0034-4257(96)00072-7

Gitelson, A. A., Kaufman, Y. J., Stark, R., & Rundquist, D. (2002). Novel algorithms for remote estimation of vegetation fraction. *Remote Sensing of Environment*. https://doi.org/10.1016/S0034-4257(01)00289-9

Gitelson, A., & Merzlyak, M. N. (1994a). Quantitative estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves. *Journal of Photochemistry and Photobiology, B: Biology*, *22*(3). https://doi.org/10.1016/1011-1344(93)06963-4

Gitelson, A., & Merzlyak, M. N. (1994b). Spectral Reflectance Changes Associated with Autumn Senescence of Aesculus hippocastanum L. and Acer platanoides L. Leaves. Spectral Features and Relation to Chlorophyll Estimation. *Journal of Plant Physiology*. https://doi.org/10.1016/S0176-1617(11)81633-0

Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., Pretty, J., Robinson, S., Thomas, S. M., & Toulmin, C. (2010). Food security: The challenge of feeding 9 billion people. In *Science*. https://doi.org/10.1126/science.1185383

Grüner, E., Astor, T., & Wachendorf, M. (2019). Biomass prediction of heterogeneous temperate grasslands using an SFM approach based on UAV imaging. *Agronomy*, *9*(2). https://doi.org/10.3390/agronomy9020054

Grüner, E., Wachendorf, M., & Astor, T. (2020). The potential of UAV-borne spectral and textural information for predicting aboveground biomass and N fixation in legume-grass mixtures. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0234703

Gunčar, G., Kukar, M., Notar, M., Brvar, M., Černelč, P., Notar, M., & Notar, M. (2018). An application of machine learning to haematological diagnosis. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-017-18564-8

Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., & Strachan, I. B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment*, *90*(3). https://doi.org/10.1016/j.rse.2003.12.013

Haboudane, D., Miller, J. R., Tremblay, N., Zarco-Tejada, P. J., & Dextraze, L. (2002). Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing of Environment*, *81*(2–3). https://doi.org/10.1016/S0034-4257(02)00018-4

Hanson, J., & Ellis, R. H. (2020). Progress and challenges in ex situ conservation of forage germplasm: Grasses, herbaceous legumes

and fodder trees. In *Plants* (Vol. 9, Issue 4). https://doi.org/10.3390/plants9040446

Hassan, M. A., Yang, M., Rasheed, A., Yang, G., Reynolds, M., Xia, X., Xiao, Y., & He, Z. (2019). A rapid monitoring of NDVI across the wheat growth cycle for grain yield prediction using a multi-spectral UAV platform. *Plant Science*. https://doi.org/10.1016/j.plantsci.2018.10.022

He, Q. (1999). *Neural Network and Its Application in IR*.

He, X., Zhao, K., & Chu, X. (2021). AutoML: A survey of the state-of-the-art. *Knowledge-Based Systems*, *212*. https://doi.org/10.1016/j.knosys.2020.106622

Heiskanen, J., Rautiainen, M., Stenberg, P., Mõttus, M., & Vesanto, V. H. (2013). Sensitivity of narrowband vegetation indices to boreal forest LAI, reflectance seasonality and species composition. *ISPRS Journal of Photogrammetry and Remote Sensing*, *78*. https://doi.org/10.1016/j.isprsjprs.2013.01.001

Henebry, G., Viña, A., & Gitelson, A. (2004). The wide dynamic range vegetation index and its potential utility for gap analysis. *Papers in Natural Resources*.

Hernández-Clemente, R., Navarro-Cerrillo, R. M., & Zarco-Tejada, P. J. (2012). Carotenoid content estimation in a heterogeneous conifer forest using narrow-band indices and PROSPECT+DART simulations. *Remote Sensing of Environment*, *127*. https://doi.org/10.1016/j.rse.2012.09.014

Herwitz, S. R., Johnson, L. F., Dunagan, S. E., Higgins, R. G., Sullivan, D. v., Zheng, J., Lobitz, B. M., Leung, J. G., Gallmeyer, B. A., Aoyagi, M., Slye, R. E., & Brass, J. A. (2004). Imaging from an unmanned aerial vehicle: Agricultural surveillance and decision support. *Computers and Electronics in Agriculture*, *44*(1). https://doi.org/10.1016/j.compag.2004.02.006

Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, *25*(3). https://doi.org/10.1016/0034-4257(88)90106-X

Hunt, E. R., Hively, W. D., Daughtry, C. S. T., Mccarty, G. W., Fujikawa, S. J., Ng, T. L., Tranchitella, M., Linden, D. S., & Yoel, D. W. (2008). Remote Sensing of Crop Leaf Area Index Using Unmanned Airborne Vehicles. *Proceedings of the Pecora 17 Symposium*.

Hutter, F., Hoos, H. H., & Leyton-Brown, K. (2011). Sequential model-based optimization for general algorithm configuration. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6683 LNCS*. https://doi.org/10.1007/978-3-642-25566-3_40

Insua, J. R., Utsumi, S. A., & Basso, B. (2019). Estimation of spatial and temporal variability of pasture growth and digestibility in grazing rotations coupling unmanned aerial vehicle (UAV) with crop simulation models. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0212773

Jordan, C. F. (1969). Derivation of Leaf-Area Index from Quality of Light on the Forest Floor. *Ecology*. https://doi.org/10.2307/1936256

Kambhatla, N., & Leen, T. K. (1997). Dimension Reduction by Local Principal Component Analysis. *Neural Computation*, *9*(7). https://doi.org/10.1162/neco.1997.9.7.1493

Karlen, D. L., Kovar, J. L., Cambardella, C. A., & Colvin, T. S. (2013). Thirty-year tillage effects on crop yield and soil fertility indicators. *Soil and Tillage Research*, *130*, 24–41. https://doi.org/10.1016/j.still.2013.02.003

Karthikeyan, L., Chawla, I., & Mishra, A. K. (2020). A review of remote sensing applications in agriculture for food security: Crop growth and yield, irrigation, and crop losses. *Journal of Hydrology*, *586*. https://doi.org/10.1016/j.jhydrol.2020.124905

Kearns, M., & Ron, D. (1999). Algorithmic stability and sanity-check bounds for leave-one-out cross-validation. *Neural Computation*. https://doi.org/10.1162/089976699300016304

Kelcey, J., & Lucieer, A. (2012). Sensor correction of a 6-band multispectral imaging sensor for UAV remote sensing. *Remote Sensing*, *4*(5). https://doi.org/10.3390/rs4051462

Kim, H.-O., & Yeom, J.-M. (2012). A Study on Object-Based Image Analysis Methods for Land Cover Classification in Agricultural Areas. *Journal of the Korean Association of Geographic Information Studies*, *15*(4). https://doi.org/10.11108/kagis.2012.15.4.026

Kim, M. S., Daughtry, C. S. T., Chappelle, E. W., McMurtrey, J. E., & Walthall, C. L. (1994). The use of high spectral resolution bands for estimating absorbed photosynthetically active radiation (A par).

Proceedings of the 6th Symposium on Physical Measurements and Signatures in Remote Sensing.

Komer, B., Bergstra, J., & Eliasmith, C. (2014). Hyperopt-Sklearn: Automatic Hyperparameter Configuration for Scikit-Learn. *Proceedings of the 13th Python in Science Conference*. https://doi.org/10.25080/majora-14bd3278-006

Kross, A., McNairn, H., Lapen, D., Sunohara, M., & Champagne, C. (2015). Assessment of RapidEye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops. *International Journal of Applied Earth Observation and Geoinformation*, *34*(1). https://doi.org/10.1016/j.jag.2014.08.002

Kumar, A., McCann, R., Naughton, J., & Patel, J. (2016). Model Selection Management Systems: The Next Frontier of Advanced Analytics. *ACM SIGMOD Record*, *44*(4). https://doi.org/10.1145/2935694.2935698

Kwak, G. H., & Park, N. W. (2019). Impact of texture information on crop classification with machine learning and UAV images. *Applied Sciences (Switzerland)*, *9*(4). https://doi.org/10.3390/app9040643

Laidig, F., Piepho, H. P., Drobek, T., & Meyer, U. (2014). Genetic and non-genetic long-term trends of 12 different crops in German official variety performance trials and on-farm yield trends. *Theoretical and Applied Genetics*, *127*(12). https://doi.org/10.1007/s00122-014-2402-z

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. Journal of Statistical Software. *J. of Statistical Software*, *25*(1).

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. In *Nature* (Vol. 521, Issue 7553, pp. 436–444). Nature Publishing Group. https://doi.org/10.1038/nature14539

Lee, H., Lee, H.-J., Jung, J.-S., & Ko, H.-J. (2015). Mapping Herbage Biomass on a Hill Pasture using a Digital Camera with an Unmanned Aerial Vehicle System. *Journal of The Korean Society of Grassland and Forage Science*. https://doi.org/10.5333/kgfs.2015.35.3.225

Li, B., Xu, X., Zhang, L., Han, J., Bian, C., Li, G., Liu, J., & Jin, L. (2020). Above-ground biomass estimation and yield prediction in potato by using UAV-based RGB and hyperspectral imaging. *ISPRS Journal of Photogrammetry and Remote Sensing*, *162*(February), 161–172. https://doi.org/10.1016/j.isprsjprs.2020.02.013

Liu, T., Abd-Elrahman, A., Morton, J., & Wilhelm, V. L. (2018). Comparing fully convolutional networks, random forest, support vector machine, and patch-based deep convolutional neural networks for object-based wetland mapping using images from small unmanned aircraft system. *GIScience and Remote Sensing*. https://doi.org/10.1080/15481603.2018.1426091

Loide, V. (2019). The results of an NPK-fertilisation trial of long-term crop rotation on carbonate-rich soil in Estonia. *Acta Agriculturae Scandinavica Section B: Soil and Plant Science*, *69*(7), 596–605. https://doi.org/10.1080/09064710.2019.1628292

Lollato, R. P., Roozeboom, K., Lingenfelser, J. F., da Silva, C. L., & Sassenrath, G. (2020). Soft winter wheat outyields hard winter wheat in a subhumid environment: Weather drivers, yield plasticity, and rates of yield gain. *Crop Science*, *60*(3). https://doi.org/10.1002/csc2.20139

Luan, J., Zhang, C., Xu, B., Xue, Y., & Ren, Y. (2020). The predictive performances of random forest models with limited sample size and different species traits. *Fisheries Research*. https://doi.org/10.1016/j.fishres.2020.105534

Lüscher, A., Mueller-Harvey, I., Soussana, J. F., Rees, R. M., & Peyraud, J. L. (2014). Potential of legume-based grassland-livestock systems in Europe: A review. *Grass and Forage Science*, *69*(2), 206–228. https://doi.org/10.1111/gfs.12124

Lussem, U., Bolten, A., Gnyp, M. L., Jasper, J., & Bareth, G. (2018). Evaluation of RGB-based vegetation indices from UAV imagery to estimate forage yield in Grassland. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *42*(3), 1215–1219. https://doi.org/10.5194/isprs-archives-XLII-3-1215-2018

Lussem, U., Bolten, A., Menne, J., Gnyp, M. L., & Bareth, G. (2019). Ultra-high spatial resolution uav-based imagery to predict biomass in temperate grasslands. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*. https://doi.org/10.5194/isprs-archives-XLII-2-W13-443-2019

Ma, L., Li, M., Ma, X., Cheng, L., Du, P., & Liu, Y. (2017a). A review of supervised object-based land-cover image classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *130*, 277–293. https://doi.org/10.1016/j.isprsjprs.2017.06.001

Ma, L., Li, M., Ma, X., Cheng, L., Du, P., & Liu, Y. (2017b). A review of supervised object-based land-cover image classification. In *ISPRS Journal of Photogrammetry and Remote Sensing*. https://doi.org/10.1016/j.isprsjprs.2017.06.001

Ma, Q., Chai, L., Hou, F., Chang, S., Ma, Y., Tsunekawa, A., & Cheng, Y. (2019). Quantifying grazing intensity using remote sensing in alpine meadows on Qinghai-Tibetan Plateau. *Sustainability (Switzerland)*. https://doi.org/10.3390/su11020417

Mackay, D. J. C. (1995). Probable networks and plausible predictions - a review of practical bayesian methods for supervised neural networks. In *Network: Computation in Neural Systems* (Vol. 6, Issue 3). https://doi.org/10.1088/0954-898X_6_3_011

Maimaitijiang, M., Sagan, V., Sidike, P., Daloye, A. M., Erkbol, H., & Fritschi, F. B. (2020). Crop monitoring using satellite/UAV data fusion and machine learning. *Remote Sensing*, *12*(9). https://doi.org/10.3390/RS12091357

Mandal, A., Majumder, A., Dhaliwal, S. S., Toor, A. S., Mani, P. K., Naresh, R. K., Gupta, R. K., & Mitran, T. (2020). Impact of agricultural management practices on soil carbon sequestration and its monitoring through simulation models and remote sensing techniques: A review. *Critical Reviews in Environmental Science and Technology*. https://doi.org/10.1080/10643389.2020.1811590

Marques Ramos, A. P., Prado Osco, L., Elis Garcia Furuya, D., Nunes Gonçalves, W., Cordeiro Santana, D., Pereira Ribeiro Teodoro, L., Antonio da Silva Junior, C., Fernando Capristo-Silva, G., Li, J., Henrique Rojo Baio, F., Marcato Junior, J., Eduardo Teodoro, P., & Pistori, H. (2020). A random forest ranking approach to predict yield in maize with uav-based vegetation spectral indices. *Computers and Electronics in Agriculture*. https://doi.org/10.1016/j.compag.2020.105791

Martínez-Usó, A., Pla, F., Sotoca, J. M., & García-Sevilla, P. (2007). Clustering-based hyperspectral band selection using information measures. *IEEE Transactions on Geoscience and Remote Sensing*, *45*(12). https://doi.org/10.1109/TGRS.2007.904951

Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. In *International Journal of Remote Sensing* (Vol. 39, Issue 9). https://doi.org/10.1080/01431161.2018.1433343

Mendoza, H., Klein, A., Feurer, M., Springenberg, J. T., & Hutter, F. (2016). Towards automatically-tuned neural networks. *Workshop on Automatic Machine Learning*.

Merton, R., & Huntington, J. (1999). Early Simulation Results of the Aries-1 Satellite Sensor for Multi-Temporal Vegetation Research Derived From Aviris. *Proceedings of the Eighth Annual JPL …*, *January*, 1–10. http://www.eoc.csiro.au/hswww/jpl_99.htm

Ministry of Rural Affairs. (2021). *Crop farming*. https://www.stat.ee/en/find-statistics/methodology-and-quality/esms-metadata/21214#2-Metadata-update-1

Mokroš, M., Surov, P., & Grzn, A. (2019). *UAV RTK / PPK Method — An Optimal Solution for Mapping Inaccessible Forested Areas ?* https://doi.org/10.3390/rs11060721

Monteiro, P. F. C., Filho, R. A., Xavier, A. C., & Monteiro, R. O. C. (2012). Assessing biophysical variable parameters of bean crop with hyperspectral measurements. *Scientia Agricola*, *69*(2). https://doi.org/10.1590/S0103-90162012000200001

Mozgeris, G., Jonikavičius, D., Jovarauskas, D., Zinkevičius, R., Petkevičius, S., & Steponavičius, D. (2018). Imaging from manned ultra-light and unmanned aerial vehicles for estimating properties of spring wheat. *Precision Agriculture*. https://doi.org/10.1007/s11119-018-9562-9

Mulla, D. J. (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering*, *114*, 358–371. https://doi.org/10.1016/j.biosystemseng.2012.08.009

Munaro, L. B., Hefley, T. J., DeWolf, E., Haley, S., Fritz, A. K., Zhang, G., Haag, L. A., Schlegel, A. J., Edwards, J. T., Marburger, D., Alderman, P., Jones-Diamond, S. M., Johnson, J., Lingenfelser, J. E., Unêda-Trevisoli, S. H., & Lollato, R. P. (2020). Exploring long-term variety performance trials to improve environment-specific genotype × management recommendations: A case-study for winter wheat. *Field Crops Research*. https://doi.org/10.1016/j.fcr.2020.107848

Muñoz-Huerta, R. F., Guevara-Gonzalez, R. G., Contreras-Medina, L. M., Torres-Pacheco, I., Prado-Olivarez, J., & Ocampo-Velazquez, R. v. (2013). A review of methods for sensing the nitrogen status

in plants: Advantages, disadvantages and recent advances. In *Sensors (Switzerland)* (Vol. 13, Issue 8). https://doi.org/10.3390/s130810823

Mutanga, O., & Kumar, L. (2007). Estimating and mapping grass phosphorus concentration in an African savanna using hyperspectral image data. *International Journal of Remote Sensing*. https://doi.org/10.1080/01431160701253253

Mutanga, O., & Skidmore, A. K. (2004a). Integrating imaging spectroscopy and neural networks to map grass quality in the Kruger National Park, South Africa. *Remote Sensing of Environment*. https://doi.org/10.1016/j.rse.2003.12.004

Mutanga, O., & Skidmore, A. K. (2004b). Narrow band vegetation indices overcome the saturation problem in biomass estimation. *International Journal of Remote Sensing*, *25*(19). https://doi.org/10.1080/01431160310001654923

Najafi, P., Feizizadeh, B., & Navid, H. (2021). A comparative approach of fuzzy object based image analysis and machine learning techniques which are applied to crop residue cover mapping by using sentinel-2 satellite and uav imagery. *Remote Sensing*, *13*(5). https://doi.org/10.3390/rs13050937

Näsi, R., Viljanen, N., Kaivosoja, J., Alhonoja, K., Hakala, T., Markelin, L., & Honkavaara, E. (2018). Estimating biomass and nitrogen amount of barley and grass using UAV and aircraft based spectral and photogrammetric 3D features. *Remote Sensing*, *10*(7). https://doi.org/10.3390/rs10071082

Nawar, S., Corstanje, R., Halcro, G., Mulla, D., & Mouazen, A. M. (2017). Delineation of Soil Management Zones for Variable-Rate Fertilization: A Review. In *Advances in Agronomy* (Vol. 143). https://doi.org/10.1016/bs.agron.2017.01.003

Neal, R. (1996). Bayesian Learning for Neural Networks. *LECTURE NOTES IN STATISTICS -NEW YORK- SPRINGER VERLAG-*, *1*(118).

Oliveira, R. A., Näsi, R., Niemeläinen, O., Nyholm, L., Alhonoja, K., Kaivosoja, J., Jauhiainen, L., Viljanen, N., Nezami, S., Markelin, L., Hakala, T., & Honkavaara, E. (2020). Machine learning estimators for the quantity and quality of grass swards used for silage production using drone-based imaging spectrometry and photogrammetry.

Remote Sensing of Environment, *246*. https://doi.org/10.1016/j.rse.2020.111830

Olson, R. S., Urbanowicz, R. J., Andrews, P. C., Lavender, N. A., Kidd, L. C., & Moore, J. H. (2016). Automating biomedical data science through tree-based pipeline optimization. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9597*. https://doi.org/10.1007/978-3-319-31204-0_9

Ono, J. P., Castelo, S., Lopez, R., Bertini, E., Freire, J., & Silva, C. (2021). PipelineProfiler: A Visual Analytics Tool for the Exploration of AutoML Pipelines. *IEEE Transactions on Visualization and Computer Graphics*, *27*(2). https://doi.org/10.1109/TVCG.2020.3030361

Osco, L. P., Ramos, A. P. M., Pereira, D. R., Moriya, érika A. S., Imai, N. N., Matsubara, E. T., Estrabis, N., de Souza, M., Junior, J. M., Gonçalves, W. N., Li, J., Liesenberg, V., & Creste, J. E. (2019). Predicting canopy nitrogen content in citrus-trees using random forest algorithm associated to spectral vegetation indices from UAV-imagery. *Remote Sensing*. https://doi.org/10.3390/rs11242925

Pavón-Pulido, N., López-Riquelme, J. A., Torres, R., Morais, R., & Pastor, J. A. (2017). New trends in precision agriculture: a novel cloud-based system for enabling data storage and agricultural task planning and automation. *Precision Agriculture*, *18*(6). https://doi.org/10.1007/s11119-017-9532-7

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*.

Pencina, M. J., D'Agostino, R. B., D'Agostino, R. B., & Vasan, R. S. (2008). Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. *Statistics in Medicine*, *27*(2). https://doi.org/10.1002/sim.2929

Pieruschka, R., & Schurr, U. (2019). Plant Phenotyping: Past, Present, and Future. *Plant Phenomics*, *2019*, 1–6. https://doi.org/10.34133/2019/7507131

Pilgrim, M. (2009). Dive into python 3. In *Dive Into Python 3*. https://doi.org/10.1007/978-1-4302-2416-7

Poley, L. G., & McDermid, G. J. (2020). A systematic review of the factors influencing the estimation of vegetation aboveground biomass using unmanned aerial systems. In *Remote Sensing*. https://doi.org/10.3390/rs12071052

Poncet, A. M., Knappenberger, T., Brodbeck, C., Fogle, M., Shaw, J. N., & Ortiz, B. v. (2019). Multispectral UAS data accuracy for different radiometric calibration methods. *Remote Sensing*. https://doi.org/10.3390/rs11161917

Qi, Y., Minka, T. P., Picard, R. W., & Ghahramani, Z. (2004). Predictive automatic relevance determination by expectation propagation. *Proceedings, Twenty-First International Conference on Machine Learning, ICML 2004*. https://doi.org/10.1145/1015330.1015418

R Core Team (2020). (2020). R: A language and environment for statistical computing. In *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.*

Raeva, P. L., Šedina, J., & Dlesk, A. (2019). Monitoring of crop fields using multispectral and thermal imagery from UAV. *European Journal of Remote Sensing*, *52*(sup1). https://doi.org/10.1080/22797254.2018.1527661

Remeseiro, B., & Bolon-Canedo, V. (2019). A review of feature selection methods in medical applications. In *Computers in Biology and Medicine* (Vol. 112). https://doi.org/10.1016/j.compbiomed.2019.103375

Riris, H., Rodriguez, M., Mao, J., Allan, G., & Abshire, J. (2017). Airborne demonstration of atmospheric oxygen optical depth measurements with an integrated path differential absorption lidar. *Optics Express*, *25*(23). https://doi.org/10.1364/oe.25.029307

Rondeaux, G., Steven, M., & Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment*, *55*(2). https://doi.org/10.1016/0034-4257(95)00186-7

Roujean, J. L., & Breon, F. M. (1995). Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sensing of Environment*, *51*(3). https://doi.org/10.1016/0034-4257(94)00114-3

Rouse, J. W., Hass, R. H., Schell, J. A., Deering, D. W., & Harlan, J. C. (1974). Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. *Final Report, RSC 1978-4, Texas A & M University, College Station, Texas.*

Rueda-Ayala, V. P., Peña, J. M., Höglind, M., Bengochea-Guevara, J. M., & Andújar, D. (2019). Comparing UAV-based technologies and RGB-D reconstruction methods for plant height and biomass monitoring on grass ley. *Sensors (Switzerland)*. https://doi.org/10.3390/s19030535

Sahoo, R. N., Ray, S. S., & Manjunath, K. R. (2015). Hyperspectral remote sensing of agriculture. *Current Science*, *108*(5). https://doi.org/10.18520/cs/v108/i5/848-859

Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS ONE*, *10*(3). https://doi.org/10.1371/journal.pone.0118432

Samaras, S., Diamantidou, E., Ataloglou, D., Sakellariou, N., Vafeiadis, A., Magoulianitis, V., Lalas, A., Dimou, A., Zarpalas, D., Votis, K., Daras, P., & Tzovaras, D. (2019). Deep learning on multi sensor data for counter Uav applications—a systematic review. In *Sensors (Switzerland)* (Vol. 19, Issue 22). https://doi.org/10.3390/s19224837

Sankaran, S., Khot, L. R., & Carter, A. H. (2015). Field-based crop phenotyping: Multispectral aerial imaging for evaluation of winter wheat emergence and spring stand. *Computers and Electronics in Agriculture*, *118*. https://doi.org/10.1016/j.compag.2015.09.001

Sankaran, S., Khot, L. R., Espinoza, C. Z., Jarolmasjed, S., Sathuvalli, V. R., Vandemark, G. J., Miklas, P. N., Carter, A. H., Pumphrey, M. O., Knowles, R. R. N., & Pavek, M. J. (2015). Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review. In *European Journal of Agronomy* (Vol. 70). https://doi.org/10.1016/j.eja.2015.07.004

Schmidt, K. S., & Skidmore, A. K. (2003). Spectral discrimination of vegetation types in a coastal wetland. *Remote Sensing of Environment*, *85*(1). https://doi.org/10.1016/S0034-4257(02)00196-7

Seeger, M. (2004). Gaussian processes for machine learning. In *International journal of neural systems* (Vol. 14, Issue 2). https://doi.org/10.1142/S0129065704001899

Seitz, S., Goebes, P., Puerta, V. L., Pereira, E. I. P., Wittwer, R., Six, J., van der Heijden, M. G. A., & Scholten, T. (2019). Conservation tillage and organic farming reduce soil erosion. *Agronomy for Sustainable Development*, *39*(1). https://doi.org/10.1007/s13593-018-0545-z

senseFly. (2016). *eBee: senseFly SA*. Website.

Serpico, S. B., D'Inca, M., Melgani, F., & Moser, G. (2003). <title>Comparison of feature reduction techniques for classification of hyperspectral remote sensing data</title>. *Image and Signal Processing for Remote Sensing VIII*, *4885*. https://doi.org/10.1117/12.463524

Serrano, L., Filella, I., & Peñuelas, J. (2000). Remote sensing of biomass and yield of winter wheat under different nitrogen supplies. *Crop Science*. https://doi.org/10.2135/cropsci2000.403723x

Shin, J., Jin Kim, H., Park, S., & Kim, Y. (2010). Model predictive flight control using adaptive support vector regression. *Neurocomputing*. https://doi.org/10.1016/j.neucom.2009.10.002

Sims, D. A., & Gamon, J. A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment*, *81*(2–3). https://doi.org/10.1016/S0034-4257(02)00010-X

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing and Management*, *45*(4). https://doi.org/10.1016/j.ipm.2009.03.002

Sona, G., Passoni, D., Pinto, L., Pagliari, D., Masseroni, D., Ortuani, B., & Facchi, A. (2016). UAV multispectral survey to map soil and crop for precision farming applications. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *2016-January*. https://doi.org/10.5194/isprsarchives-XLI-B1-1023-2016

Sripada, R. P., Heiniger, R. W., White, J. G., & Meijer, A. D. (2006). Aerial color infrared photography for determining early in-season nitrogen requirements in corn. *Agronomy Journal*. https://doi.org/10.2134/agronj2005.0200

Stagakis, S., Markos, N., Sykioti, O., & Kyparissis, A. (2010). Monitoring canopy biophysical and biochemical parameters in ecosystem scale using satellite hyperspectral imagery: An application on a Phlomis fruticosa Mediterranean ecosystem using multiangular CHRIS/PROBA observations. *Remote Sensing of Environment*, *114*(5). https://doi.org/10.1016/j.rse.2009.12.006

Strong, C. J., Burnside, N. G., & Llewellyn, D. (2017). The potential of small-Unmanned Aircraft Systems for the rapid detection of threatened unimproved grassland communities using an Enhanced

Normalized Difference Vegetation Index. *PLoS ONE*, *12*(10), 1–16. https://doi.org/10.1371/journal.pone.0186193

Suykens, J. A. K., & Vandewalle, J. (1999). Least squares support vector machine classifiers. *Neural Processing Letters*, *9*(3). https://doi.org/10.1023/A:1018628609742

Telles, T. S., Reydon, B. P., & Maia, A. G. (2018). Effects of no-tillage on agricultural land values in Brazil. *Land Use Policy*, *76*. https://doi.org/10.1016/j.landusepol.2018.04.053

The Mathworks Inc. (2019). MATLAB (R2019a), version 9.6.0.1072779. In *The MathWorks Inc.*

Thenkabail, P. S., Smith, R. B., & de Pauw, E. (2000). Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sensing of Environment*, *71*(2). https://doi.org/10.1016/S0034-4257(99)00067-X

Thilakarathna, M. S., McElroy, M. S., Chapagain, T., Papadopoulos, Y. A., & Raizada, M. N. (2016). Belowground nitrogen transfer from legumes to non-legumes under managed herbaceous cropping systems. A review. *Agronomy for Sustainable Development*, *36*(4). https://doi.org/10.1007/s13593-016-0396-4

Thornton, C., Hutter, F., Hoos, H. H., & Leyton-Brown, K. (2013). Auto-WEKA: Combined selection and hyperparameter optimization of classification algorithms. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, *Part F128815*. https://doi.org/10.1145/2487575.2487629

Thorp, K. R., Wang, G., Bronson, K. F., Badaruddin, M., & Mon, J. (2017). Hyperspectral data mining to identify relevant canopy spectral features for estimating durum wheat growth, nitrogen status, and grain yield. *Computers and Electronics in Agriculture*, *136*. https://doi.org/10.1016/j.compag.2017.02.024

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

Torres-Sánchez, J., López-Granados, F., de Castro, A. I., & Peña-Barragán, J. M. (2013). Configuration and Specifications of an Unmanned

Aerial Vehicle (UAV) for Early Site Specific Weed Management. *PLoS ONE*, *8*(3). https://doi.org/10.1371/journal.pone.0058210

Tripicchio, P., Satler, M., Dabisias, G., Ruffaldi, E., & Avizzano, C. A. (2015). Towards Smart Farming and Sustainable Agriculture with Drones. *Proceedings - 2015 International Conference on Intelligent Environments, IE 2015*. https://doi.org/10.1109/IE.2015.29

Triplett, G. B., & Dick, W. A. (2008). No-tillage crop production: A revolution in agriculture! *Agronomy Journal*, *100*(3 SUPPL.), 153–165. https://doi.org/10.2134/agronj2007.0005c

Tsouros, D. C., Bibi, S., & Sarigiannidis, P. G. (2019). A review on UAV-based applications for precision agriculture. In *Information (Switzerland)* (Vol. 10, Issue 11). https://doi.org/10.3390/info10110349

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, *8*(2). https://doi.org/10.1016/0034-4257(79)90013-0

van Ittersum, M. K., Cassman, K. G., Grassini, P., Wolf, J., Tittonell, P., & Hochman, Z. (2013). Yield gap analysis with local to global relevance-A review. *Field Crops Research*. https://doi.org/10.1016/j.fcr.2012.09.009

Vasudevan, A., Kumar, D. A., & Bhuvaneswari, N. S. (2016). Precision farming using unmanned aerial and ground vehicles. *Proceedings - 2016 IEEE International Conference on Technological Innovations in ICT for Agriculture and Rural Development, TIAR 2016*. https://doi.org/10.1109/TIAR.2016.7801229

Vescovo, L., Wohlfahrt, G., Balzarolo, M., Pilloni, S., Sottocornola, M., Rodeghiero, M., & Gianelle, D. (2012). New spectral vegetation indices based on the near-infrared shoulder wavelengths for remote detection of grassland phytomass. *International Journal of Remote Sensing*, *33*(7). https://doi.org/10.1080/01431161.2011.607195

Viljanen, N., Honkavaara, E., Näsi, R., Hakala, T., Niemeläinen, O., & Kaivosoja, J. (2018a). A novel machine learning method for estimating biomass of grass swards using a photogrammetric canopy height model, images and vegetation indices captured by a drone. *Agriculture (Switzerland)*, *8*(5). https://doi.org/10.3390/agriculture8050070

Viljanen, N., Honkavaara, E., Näsi, R., Hakala, T., Niemeläinen, O., & Kaivosoja, J. (2018b). A novel machine learning method for estimating biomass of grass swards using a photogrammetric canopy

height model, images and vegetation indices captured by a drone. *Agriculture (Switzerland)*. https://doi.org/10.3390/agriculture8050070

Villoslada, M., Bergamo, T. F., Ward, R. D., Burnside, N. G., Joyce, C. B., Bunce, R. G. H., & Sepp, K. (2020). Fine scale plant community assessment in coastal meadows using UAV based multispectral data. *Ecological Indicators*. https://doi.org/10.1016/j.ecolind.2019.105979

Vivaldini, K. C. T., Martinelli, T. H., Guizilini, V. C., Souza, J. R., Oliveira, M. D., Ramos, F. T., & Wolf, D. F. (2019). UAV route planning for active disease classification. *Autonomous Robots*, *43*(5). https://doi.org/10.1007/s10514-018-9790-x

Vleugels, T., Amdahl, H., Roldán-Ruiz, I., & Cnops, G. (2019). Factors underlying seed yield in red clover: Review of current knowledge and perspectives. *Agronomy*, *9*(12), 1–15. https://doi.org/10.3390/agronomy9120829

Wachendorf, M., Fricke, T., & Möckel, T. (2018). Remote sensing as a tool to assess botanical composition, structure, quantity and quality of temperate grasslands. *Grass and Forage Science*, *73*(1), 1–14. https://doi.org/10.1111/gfs.12312

Walsh, O. S., Shafian, S., Marshall, J. M., Jackson, C., McClintick-Chess, J. R., Blanscet, S. M., Swoboda, K., Thompson, C., Belmont, K. M., & Walsh, W. L. (2018). Assessment of UAV Based Vegetation Indices for Nitrogen Concentration Estimation in Spring Wheat. *Advances in Remote Sensing*. https://doi.org/10.4236/ars.2018.72006

Wang, J., Badenhorst, P., Phelan, A., Pembleton, L., Shi, F., Cogan, N., Spangenberg, G., & Smith, K. (2019). Using Sensors and Unmanned Aircraft Systems for High-Throughput Phenotyping of Biomass in Perennial Ryegrass Breeding Trials. *Frontiers in Plant Science*. https://doi.org/10.3389/fpls.2019.01381

Wang, Q., Ming, Y., Jin, Z., Shen, Q., Liu, D., Smith, M. J., Veeramachaneni, K., & Qu, H. (2019). ATMSeer: Increasing Transparency and Controllability in Automated Machine Learning. In *arXiv*.

Wen, W., Timmermans, J., Chen, Q., & van Bodegom, P. M. (2021). A review of remote sensing challenges for food security with respect to salinity and drought threats. In *Remote Sensing* (Vol. 13, Issue 1). https://doi.org/10.3390/rs13010006

Wittwer, R. A., & van der Heijden, M. G. A. (2020). Cover crops as a tool to reduce reliance on intensive tillage and nitrogen fertilization

in conventional arable cropping systems. *Field Crops Research*. https://doi.org/10.1016/j.fcr.2020.107736

Wu, C., Niu, Z., Tang, Q., & Huang, W. (2008). Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Agricultural and Forest Meteorology*, *148*(8–9). https://doi.org/10.1016/j.agrformet.2008.03.005

Wu, W. (2014). The Generalized Difference Vegetation Index (GDVI) for dryland characterization. *Remote Sensing*, *6*(2). https://doi.org/10.3390/rs6021211

Xavier, A. C., Theodor Rudorff, B. F., Moreira, M. A., Alvarenga, B. S., de Freitas, J. G., & Salomon, M. V. (2006). Hyperspectral field reflectance measurements to estimate wheat grain yield and plant height. *Scientia Agricola*, *63*(2), 130–138. https://doi.org/10.1590/s0103-90162006000200004

Xiang, H., & Tian, L. (2011). Development of a low-cost agricultural remote sensing system based on an autonomous unmanned aerial vehicle (UAV). *Biosystems Engineering*, *108*(2). https://doi.org/10.1016/j.biosystemseng.2010.11.010

Xie, Y., Sha, Z., Yu, M., Bai, Y., & Zhang, L. (2009). A comparison of two models with Landsat data for estimating above ground grassland biomass in Inner Mongolia, China. *Ecological Modelling*. https://doi.org/10.1016/j.ecolmodel.2009.04.025

Xue, J., & Su, B. (2017). Significant remote sensing vegetation indices: A review of developments and applications. In *Journal of Sensors* (Vol. 2017). https://doi.org/10.1155/2017/1353691

Yang, M. der, Boubin, J. G., Tsai, H. P., Tseng, H. H., Hsu, Y. C., & Stewart, C. C. (2020a). Adaptive autonomous UAV scouting for rice lodging assessment using edge computing with deep learning EDANet. *Computers and Electronics in Agriculture*, *179*. https://doi.org/10.1016/j.compag.2020.105817

Yang, M. der, Boubin, J. G., Tsai, H. P., Tseng, H. H., Hsu, Y. C., & Stewart, C. C. (2020b). Adaptive autonomous UAV scouting for rice lodging assessment using edge computing with deep learning EDANet. *Computers and Electronics in Agriculture*, *179*. https://doi.org/10.1016/j.compag.2020.105817

Yang, M. der, Huang, K. H., & Tsai, H. P. (2020). Integrating MNF and HHT transformations into artificial neural networks for hyperspectral

image classification. *Remote Sensing*, *12*(14). https://doi.org/10.3390/rs12142327

Yang, M. Der, Huang, K. S., Kuo, Y. H., Tsai, H. P., & Lin, L. M. (2017). Spatial and spectral hybrid image classification for rice lodging assessment through UAV imagery. *Remote Sensing*. https://doi.org/10.3390/rs9060583

Yang, M. der, Tseng, H. H., Hsu, Y. C., & Tsai, H. P. (2020). Semantic segmentation using deep learning with vegetation indices for rice lodging identification in multi-date UAV visible images. *Remote Sensing*. https://doi.org/10.3390/rs12040633

Yang, G., Liu, J., Zhao, C., Li, Z., Huang, Y., Yu, H., Xu, B., Yang, X., Zhu, D., Zhang, X., Zhang, R., Feng, H., Zhao, X., Li, Z., Li, H., & Yang, H. (2017). Unmanned aerial vehicle remote sensing for field-based crop phenotyping: Current status and perspectives. In *Frontiers in Plant Science* (Vol. 8). https://doi.org/10.3389/fpls.2017.01111

Yang, X. M., Drury, C. F., Reynolds, W. D., & Reeb, M. D. (2019). Legume cover crops provide nitrogen to corn during a three-year transition to organic cropping. *Agronomy Journal*. https://doi.org/10.2134/agronj2018.10.0652

Yang, Y., Reilly, E. C., Jungers, J. M., Chen, J., & Smith, T. M. (2019). Climate Benefits of Increasing Plant Diversity in Perennial Bioenergy Crops. *One Earth*. https://doi.org/10.1016/j.oneear.2019.11.011

Yao, Q., Wang, M., Chen, Y., Dai, W., Hu, Y. Q., Li, Y. F., Tu, W. W., Yang, Q., & Yu, Y. (2018). Taking the human out of learning applications: A survey on automated machine learning. In *arXiv*.

Yeom, J., Jung, J., Chang, A., Ashapure, A., Maeda, M., Maeda, A., & Landivar, J. (2019a). Comparison of vegetation indices derived from UAV data for differentiation of tillage effects in agriculture. *Remote Sensing*, *11*(13). https://doi.org/10.3390/rs11131548

Yeom, J., Jung, J., Chang, A., Ashapure, A., Maeda, M., Maeda, A., & Landivar, J. (2019b). Comparison of vegetation indices derived from UAV data for differentiation of tillage effects in agriculture. *Remote Sensing*. https://doi.org/10.3390/rs11131548

Young, S. N. (2019). Framework for evaluating field-based, high-throughput phenotyping systems: A meta-analysis. *Sensors (Switzerland)*, *19*(16). https://doi.org/10.3390/s19163582

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 Sensing*. https://doi.org/10.3390/rs9070708

Zarco-Tejada, P. J. (2000). Hyperspectral Remote Sensing of Closed Forest Canopies : Estimation of Chlorophyll Fluorescence and Pigment Content. In *Space Technology*.

Zarco-Tejada, P. J., Miller, J. R., Noland, T. L., Mohammed, G. H., & Sampson, P. H. (2001). Scaling-up and model inversion methods with narrowband optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(7). https://doi.org/10.1109/36.934080

Zarco-Tejada, P. J., Ustin, S. L., & Whiting, M. L. (2005). Temporal and spatial relationships between within-field yield variability in cotton and high-spatial hyperspectral remote sensing imagery. *Agronomy Journal*, *97*(3). https://doi.org/10.2134/agronj2003.0257

Zarza, R., Rebuffo, M., La Manna, A., & Balzarini, M. (2020). Red clover (Trifolium pratense L.) seedling density in mixed pastures as predictor of annual yield. *Field Crops Research*, *256*(January). https://doi.org/10.1016/j.fcr.2020.107925

Zhang, C., & Kovacs, J. M. (2012a). The application of small unmanned aerial systems for precision agriculture: A review. In *Precision Agriculture* (Vol. 13, Issue 6). https://doi.org/10.1007/s11119-012-9274-5

Zhang, C., & Kovacs, J. M. (2012b). The application of small unmanned aerial systems for precision agriculture: A review. *Precision Agriculture*, *13*(6), 693–712. https://doi.org/10.1007/s11119-012-9274-5

Zhang, J., Huang, W., & Zhou, Q. (2014). Reflectance variation within the in-chlorophyll centre waveband for robust retrieval of leaf chlorophyll content. *PLoS ONE*, *9*(11). https://doi.org/10.1371/journal.pone.0110812

Zhang, W., Zhang, Z., Chao, H. C., & Guizani, M. (2019). Toward Intelligent Network Optimization in Wireless Networking: An Auto-Learning Framework. *IEEE Wireless Communications*, *26*(3). https://doi.org/10.1109/MWC.2019.1800350

Zhang, X., Han, L., Dong, Y., Shi, Y., Huang, W., Han, L., González-Moreno, P., Ma, H., Ye, H., & Sobeih, T. (2019). A deep learning-based approach for automated yellow rust disease detection from

high-resolution hyperspectral UAV images. *Remote Sensing*. https://doi.org/10.3390/rs11131554

Zhang, Z., Masjedi, A., Zhao, J., & Crawford, M. M. (2017). Prediction of sorghum biomass based on image based features derived from time series of UAV images. *International Geoscience and Remote Sensing Symposium (IGARSS)*. https://doi.org/10.1109/IGARSS.2017.8128413

Zheng, H., Cheng, T., Zhou, M., Li, D., Yao, X., Tian, Y., Cao, W., & Zhu, Y. (2019). Improved estimation of rice aboveground biomass combining textural and spectral analysis of UAV imagery. *Precision Agriculture*. https://doi.org/10.1007/s11119-018-9600-7

Zhu-Barker, X., & Steenwerth, K. L. (2018a). Chapter Six - Nitrous Oxide Production From Soils in the Future: Processes, Controls, and Responses to Climate Change. In *Climate Change Impacts on Soil Processes and Ecosystem Properties*.

Zhu-Barker, X., & Steenwerth, K. L. (2018b). *Nitrous Oxide Production From Soils in the Future* (Vol. 35, pp. 131–183). Elsevier. https://doi.org/10.1016/b978-0-444-63865-6.00006-5

Zikeli, S., Gruber, S., Teufel, C. F., Hartung, K., & Claupein, W. (2013). Effects of reduced tillage on crop yield, plant available nutrients and soil organic matter in a 12-year long-term trial under organic management. *Sustainability (Switzerland)*, *5*(9), 3876–3894. https://doi.org/10.3390/su5093876

SUMMARY IN ESTONIAN

Title. Unmanned aircraft systems and image analysis in yield estimation and agricultural management

Estonian title: Mehitamata õhusõiduki rakendamine põllukultuuride saagikuse ja maa harimisviiside tuvastamisel

SISSEJUHATUS

Doktoritöö eesmärk oli uurida, kuidas masinõppe (MÕ) tehnoloogiad võimaldavad edusamme täppispõllumajanduse valdkonna pildianalüüsis. Multimodaalsed arvutustehnoloogiad laiendavad masinõppe kasutamist põllumajanduses andmete kogumisel ja valimisel (Nawar *et al.*, 2017). Selline täpsemal informatsioonil põhinev tehnoloogia võimaldab keerukate viljelussüsteemide puhul teha otsuseid inimese vähema sekkumisega, ja loob skaleeritava raamistiku täppispõllumajanduse jaoks (Chlingaryan *et al.*, 2018). Põllukultuuride katsete korral on komplekssete masinõppemudelite kasutamine keerukas, sest alad on piiratud ning valimi suurus ei ole piisav; vaja on testandmebaase, kindlaid aja- ja ruumitingimusi ning keskkonnategureid. See komplitseerib parameetrite valikut ning muudab ebapraktiliseks ühe empiirilise mudeli kasutamise terves piirkonnas. Siinse uurimuse algetapis rakendati suhteliselt traditsioonilist masinõppemeetodit, et lahendada saagikuse ja biomassi prognoosimise regressiooniprobleem (otsustusmetsa regression, tugivektori regressioon ja tehisnärvivõrk) punase ristiku prognoositava kuivaine saagikuse suhtes. Saadi sobivaid tulemusi, kuid hüperparameetrite valimine, pikk algoritmide valimisprotsess, andmete puhastamine ja kollineaarsusprobleemid takistasid masinõpet oluliselt.

Automatiseeritud masinõppe (AMÕ) uusimate suundumustena rakendatakse tehisintellekti, et lahendada põhiprobleemid automatiseeritud algoritmi valiku ja rakendatava *pipeline*-mudeli hüperparameetrite optimeerimise abil. Seni napib teadmisi MÕ tehnoloogia integreerimiseks mehitamata õhusõidukite ning hüperspektripõhiste pildiandmete kategoriseerimise ja regressioonirakendustega. Väitekirjas uuriti nüüdisaegset ja avatud lähtekoodiga AMÕ tehnoloogiat Auto-sklearn, mis on ühe enimkasutatava masinõppesüsteemi Scikit-learn edasiarendus. Süsteemiga liideti kaks unikaalset AMÕ visualiseerimisrakendust, et

uurida mehitamata õhusõidukiga kogutud andmete multispektraalsete taimkatteindeksite ja hüperspektraalsete kitsaribaandmete taimkatteindeksite tuvastamist ja rakendamist põllumajanduses. Neid võtteid kasutatakse mullaharimisel, kultiveerimisel ja sõnnikuga väetamisel nelja kultuuriga põldudel (punase ristiku rohusegu, suvinisu, herne-kaera segu, suvioder). Neid ei ole põhjalikult hinnatud, samuti ei hõlma need omadusi, mida kasutatatakse põllumajanduses kaugseire rakendustes.

Uurimus käsitleb biomassi ja saagikuse seni uurimata analüüsivõimalusi oluliste põllukultuuride ja viljelusmeetodite näitel. Hinnatakse ka kaugseirelahenduste potentsiaali põllupõhiste ja multifunktsionaalsete platvormide kasutamisel täppispõllumajanduses. Uurimus tutvustab kiiret, keskkonna suhtes kahjutut ja mõõduka hinnaga tehnoloogiat põllupõhise biomassi ja teraviljasaagi modelleerimiseks, et leida sobiv viljelusviis. Töö tulemused võimaldavad põllumajandustootjatel ja agronoomidel tõhusamalt valida põllundustehnoloogiaid ning arvestada täpsemalt keskkonnatingimustega.

HÜPOTEESID JA UURIMUSE EESMÄRK

Uurimuse eesmärk on luua uudne kaugseiresüsteem nutika täppispõllumajanduse arendamiseks ning keskkonnatingimuste paremaks jälgimiseks. Selleks kasutatakse multispektraalse kaameraga varustatud mehitamata õhusõidukit, hüperspektraalsete andmete kogumist ning AMÕ-d, et uurida Eestis levinud põllukultuuride saagi prognoosimise ja täppisviljeluse võimalusi.

Uurimuse aluseks olid järgmised eesmärgid ja hüpoteesid.

1. Uurida seoseid põllukultuuri vegetatsaiooniperioodi, asukoha ja viljelustehnika vahel ning multispektraalse kaameraga varustatud mehitamata õhusõiduki kasutamist söödakultuuride toodangu ja biomassi hindamiseks.

Hüpotees: Multispektraalse kaameraga varustatud mehitamata õhusõidukiga kogutud andmed on rakendatavad põllukultuuride saagi kiireks ja täpseks prognoosimiseks eri aastaaegadel, perioodidel ja erinevate viljelusviiside korral.

2. Uurida AMÕ kasutamist saagi pildiandmete analüüsimisel, et töötada välja tõhusaid regressiooni- ja klassifitseerimismudelid.

Hüpotees: Klassikalisi masinõppemeetodeid ühendav AMÕ võib põllumajanduses tõhusalt ja kiiresti lahendada pildiandmete regressiooni- ja klassifitseerimisprobleeme.

3. Uurida mehitamata õhusõidukite sobivust põllumajanduses kultiveerimisvõtete valimisel ja loomisel.

Hüpotees: Mehitamata õhusõidukitega saadud pildiandmeid saab kasutada mitmeti: Eestis levinud põllukultuuride kasvatamisel mullaharimis- ja kultiveerimismeetodite ning sõnniku kasutamise tuvastamisel.

4. Mõista hüperspektraalse pildistamise abil saagi prognoosimise võimalusi ja muude analüüsimeetoditega ühitamisest tulenevaid eeliseid taimekasvatuses ja seires.

Hüpotees: Automatiseeritud süsteemi ja avatud lähtekoodiga süsteemide (R ja Python) ühendamisega on lahendatav hüperspektraalsete andmete ülikülluse probleem ning lüheneb andmetöötluse aeg.

MEETOD

Uurimuse eesmärk oli luua kiire, keskkonna suhtes kahjutu ja mõõduka hinnaga süsteem biomassi ja saagikuse modelleerimiseks erinevate põllukultuuride ja viljelusmeetodite jaoks, ning analüüsida täppispõllumajanduse vajadustele vastavaid põllupõhise kaugseire multifunktsionaalseid lahendusi.

Uurimuse käigus kasutati multispektraalsete anduritega varustatud mehitamata õhusõidukit ja lennukilt kogutud hüperspektraalseid andmeid, et hinnata saagikust ja kategoriseerida põllud pildiandmetöötluse, analüüsi ja masinõppe arvutuste abil.

TULEMUSED JA ARUTELU

Multispektraalsete anduritega varustatud mehitamata õhusõiduki lennud põldude kohal 38–11 päeva enne saagi koristamist, kus kasvatati ühe- ja kaheaastast punase ristiku rohusegu, eraldusvõimega (GSD)

10 cm ning kombineerituna taimkatteindeksiga ja mitme masinõppe meetodiga on ühe põllumajandusettevõtte mastaabis kuivaine saagikuse prognoosimiseks paljulubav, keskkonda säästev ja kulutõhus. Masinõppe analüüsitulemused olid kõige paremad tehisnärvivõrgu korral, järgnesid otsustusmetsa ja tugivektori regressioon. Taimkatteindeksi määramisel sobisid tehisnärvivõrgu ja otsustusmetsa sisenditena hästi GNDVI ja GDVI ja MSR. Samal ajal toimuv põlluharimine mõjutas mudelite prognoosivõimet. Mullaharimisel põhinev stratifitseeritud valim toimis paremini kui kultiveerimisel ja sõnnikuga väetamisel põhinevad mudelid. Tulemused kinnitavad, et mehitamata õhusõiduk sobib keerukate protsesside väljatöötamiseks, nagu seda on mullaharimismeetodid ja väetiste lisamine. Kuid väetiste koguste ning liblikõieliste ja heintaimede leviku tuvastamiskindlust tuleb edaspidi täpsemalt käsitleda.

Teiseks esitleti uurimuse käigus uudset mehitamata õhusõiduki tehnoloogiat ning modernset AMÕ süsteemi, millega saab keskkonna suhtes kahjutult ja kulutõhusalt lahendada põllumajanduse ülesandeid. Uurimuse teaduslik väärtus seisneb tehisintellekti kasutamises, nii et see asendab inimese hinnangu mehitamata õhusõiduki abil tehtud analüüsiga koos andmete automatiseeritud eeltöötluse, mudelivaliku, funktsioonide projekteerimise ja hüperparameetrite optimeerimisega. Uurimus edastab põllumajanduse jaoks uusi teadmisi, kiirendab mehitamata õhusõidukite täiustamist põldude seireks ning aitab kaasa põllukultuuride fenotüüpimisele. Uurimuses võimaldas AMÕ süsteem mehitamata õhusõidukilt kogutud andmete abil mistahes objekti tundmõppimist. See on esimene omataoline uurimus, milles rakendatakse AMÕ süsteemi põllumajanduses klassifitseerimisülesannetes multispektraalsetel andmetel taimkatteindeksi andmete korral. Lisaks võrreldi AMÕ süsteemi tööd kolme laialdaselt kasutatava MÕ meetodiga. "MÕ võrdlusanalüüs näitas, et AMÕ saavutas pärast 1200 sekundit kestnud arvutamist kõige paremad klassifikatsiooni täpsuse numbrid, kusjuures ja sealjuures ei saadud ühtegi antud andekogu kohta käivat mitte-sobivamat klassifikatsiooni. Põllumajanduse klassifikatsiooni järgi oli põllukultuuride vegetatiivne kasvufaas andmehõiveks parim periood (Eesti puhul maikuus).

Kolmandaks tõstab uurimus esile hüperspektraalanalüüsi võimalusi saagikuse ja biomassi prognoosimiseks keeruliste juhtude korral, kasutades kahte olulist avatud lähtekoodiga tarkvarasüsteemi: R-keele hüperspektraalse töötlemise paketti ja Pythoni Auto-Sklearni

masinõppetehnoloogiat. Eri tüüpi hüperspektraalsete, saagikust ja põhumassi iseloomustavate taimkattenäitajatega analüüsimise tulemuslikkust sai hinnata rahuldavaks. Soovitame neid näitajaid kasutada ka põllukultuuride muude biofüüsikaliste omaduste analüüsimiseks. Uurimuses esitatud taimkatteindeksid ja automaatne kitsariba taimkatteindeksiarvutus võivad minimeerida andmete liiasust ja puhastusaega, samuti nõudeid riistvara arvutusvõimsusele. Viljelust saab põllumajanduses edaspidi tõenäoliselt klassifitseerida hüperspektraalse pildistamise abil NIR-spektripiirkonnas (750–900 nm), millel on märkimisväärsed muutused peegeldusspektris.

Kokkuvõttes keskendus uurimus multispektraalsele ja hüperspektraalsele pildistamisele ning AMÕ integreerimisele ja rakendamisele eri põllukultuuride korral multifunktsionaalses põllumajanduses põllukultuuride biomassi ja saagikuse hindamiseks. Põhjamaades pakub see metoodika põllumajanduses juhtivatele otsustajatele saagikuse professionaalse hindamise võimalust ja lahendusi säästvaks põllumajanduseks. Kaugseiretehnoloogiate, geotöötlusmeetodite ja automaatsüsteemide integreerimine on tõhusad vahendid, et avardada põllukultuuride kasvatuses teadmisi taimede ja keskkonna seostest. Uurimus näitas ka, et eeldatavat saagikust võib prognoosida kaks kuud enne saagikoristust. See tähendab, et suvinisu, suvioder ja kaer olid valdavalt kasvu algfaasis, põldhernes reproduktiivse kasvu järgus ning punase ristiku põld kasvufaasis (siinses uurimuses 49 päeva varem). AMÕ süsteem toetab kaugseirel põhineva andmeanalüüsi tehnoloogia tõhustamist ja kasutuselevõttu. AMÕ potentsiaali ja kasutamise hoogustamiseks tuleb edaspidi teha rohkem uuringuid ja katseid.

ACKNOWLEDGEMENTS

This thesis was completed under the attentive supervision of Professor Kalev Sepp. I've been profoundly influenced and impressed by his serious scientific perspective, rigorous academic spirit, and work style of excellence, however, maintaining his unique humorous attitude, has deeply infected and inspired me. From the selection of the topic through the execution of the assignment, Kalev has provided me with unwavering support. I would like to express my sincere thanks and high respect to Kalev.

I am also grateful to my supervisor, Professor Niall Burnside, for his meticulous guidance on my research, for directing me toward my research work, and for responding my inquiries. Niall's rigorous and realistic work style, keen thinking, and profound knowledge have left a deep impression on me, and have left a lasting influence and provided me with goals and models for my future endeavours.

I would also like to thank my supervisor Dr Ants Vain, who is a dedicated, conscientious, and enthusiastic advisor. During the process of writing all of my published papers and PhD dissertations, he provided me with meticulous direction and extensive edits.

I would like to thank Professor Tiit Kuster contributed to the evaluation of my thesis and acting as my opponent for the preliminary defence. He has clear thoughts and is adept at detecting and summarizing, enabling me to learn and improve via the exchange and acceptance of critical scientific perspectives.

In addition, I would like to thank my dear colleagues, Dr Miguel Villoslada, who enlightened my knowledge of UAS, Raul Sampaio de Lima, who provided me with guidance and advice on analysis techniques, and Kaupo, who assisted me in compiling a thesis summary in Estonian, as well as my laboratory colleagues, Janar, Thaisa, Diana, Henri, Siiri, Ricardo, Mariana, and Volha, these outstanding researchers as well as friends I made here made my four years of study and research life more complete.

From the joy and anxiety when I first entered EMÜ, my doctoral study life has come to an end. During these four years of studies and life, I'd want to express my gratitude thank to my parents and family, both in South Africa and Taiwan, especially my respected mother, all of whom have forgiven me for choosing this challenging path; your love is my greatest wealth.

And my heartfelt thanks go to my dearest friend and soulmate, my spouse Alain Kermis Li, has provided me with the most compassionate and steadfast support, unconditional love, and accompanied me through my beautiful research voyage.

Finally, I'd want to thank all of the professionals, academics, and opponent who took your precious time to evaluate this thesis.

I

ORIGINAL PUBLICATIONS

Li, K. Y., Burnside, N. G., de Lima, R. S., Villoslada, M., Sepp, K., Yang, M. D., Raet, J. Vain, A., Selge, A. & Sepp, K. (2021). The Application of an Unmanned Aerial System and Machine Learning Techniques for Red Clover-Grass Mixture Yield Estimation under Variety Performance Trials. Remote Sensing, 13(10), 1994.

remote sensing

MDPI

Article

The Application of an Unmanned Aerial System and Machine Learning Techniques for Red Clover-Grass Mixture Yield Estimation under Variety Performance Trials

Kai-Yun Li [1,*], Niall G. Burnside [2], Raul Sampaio de Lima [1], Miguel Villoslada Peciña [1], Karli Sepp [3], Ming-Der Yang [4], Janar Raet [1], Ants Vain [1], Are Selge [1] and Kalev Sepp [1]

1 Institute of Agriculture and Environmental Sciences, Estonian University of Life Sciences, Kreutzwaldi 5, EE-51006 Tartu, Estonia; raul.sampaio@student.emu.ee (R.S.d.L.); Miguel.Pecina@emu.ee (M.V.P.); janar.raet@emu.ee (J.R.); a.vain@ts.ee (A.V.); are.selge@emu.ee (A.S.); kalev.sepp@emu.ee (K.S.)
2 School of Environment & Technology, University of Brighton, Lewes Road, Brighton BN2 4JG, UK; N.G.Burnside@brighton.ac.uk
3 Agricultural Research Center, 4/6 Teaduse St., 75501 Saku, Estonia; karli.sepp@pmk.agri.ee
4 Department of Civil Engineering, and Innovation and Development Center of Sustainable Agriculture, National Chung Hsing University, Taichung 402, Taiwan; mdyang@dragon.nchu.edu.tw
* Correspondence: kai-yun.li@student.emu.ee

check for updates

Citation: Li, K.-Y.; Burnside, N.G.; Sampaio de Lima, R.; Villoslada Peciña, M.; Sepp, K.; Yang, M.-D.; Raet, J.; Vain, A.; Selge, A.; Sepp, K. The Application of an Unmanned Aerial System and Machine Learning Techniques for Red Clover-Grass Mixture Yield Estimation under Variety Performance Trials. *Remote Sens.* **2021**, *13*, 1994. https://doi.org/10.3390/rs13101994

Academic Editor: Magaly Koch

Received: 28 March 2021
Accepted: 18 May 2021
Published: 19 May 2021

Abstract: A significant trend has developed with the recent growing interest in the estimation of aboveground biomass of vegetation in legume-supported systems in perennial or semi-natural grasslands to meet the demands of sustainable and precise agriculture. Unmanned aerial systems (UAS) are a powerful tool when it comes to supporting farm-scale phenotyping trials. In this study, we explored the variation of the red clover-grass mixture dry matter (DM) yields between temporal periods (one- and two-year cultivated), farming operations [soil tillage methods (STM), cultivation methods (CM), manure application (MA)] using three machine learning (ML) techniques [random forest regression (RFR), support vector regression (SVR), and artificial neural network (ANN)] and six multispectral vegetation indices (VIs) to predict DM yields. The ML evaluation results showed the best performance for ANN in the 11-day before harvest category (R^2 = 0.90, NRMSE = 0.12), followed by RFR (R^2 = 0.90 NRMSE = 0.15), and SVR (R^2 = 0.86, NRMSE = 0.16), which was furthermore supported by the leave-one-out cross-validation pre-analysis. In terms of VI performance, green normalized difference vegetation index (GNDVI), green difference vegetation index (GDVI), as well as modified simple ratio (MSR) performed better as predictors in ANN and RFR. However, the prediction ability of models was being influenced by farming operations. The stratified sampling, based on STM, had a better model performance than CM and MA. It is proposed that drone data collection was suggested to be optimum in this study, closer to the harvest date, but not later than the ageing stage.

Keywords: unmanned aerial system; red clover; random forest; support vector regression; artificial neural network; tillage; fertilizing; manure; forage legume; yield estimation

1. Introduction

Red clover (*Trifolium pratense* L.) is the principal perennial forage crop legume species in most countries of northern Europe, including Estonia [1,2]. Legumes have the ability to increase the productivity of grass pastures by fixing atmospheric nitrogen into the soil; via the symbiotic rhizobia in their root nodules [3]. This fixation of atmospheric nitrogen makes red clover an ideal rotational crop; particularly in organic agricultural systems where no synthetic nitrogen fertilizers are used [4]. A range of studies have observed that the establishment of red clover was more successful when sown in mixtures with grass species rather than in pure, monocultural stands [5,6]. In Estonia, the application of red clover in trials results in a mixed-species approach with other grass species blended

to increase its commercial application value, where the complexity of estimation might be higher than monocropping systems. Legume-based systems, and in particular red clover practices, are economically attractive to dairy farmers in northern Europe and are essential for ensuring that organic systems can compete in terms of profitability with more conventional or artificially improved systems [7]. The success of red clover in pasture farming systems makes it a vital economical crop regardless of whether it is integrated into conventional or organic farming operations. Besides their positive effects on farming productivity, legumes can also play a part in lowering greenhouse gas emissions by reducing the use of inorganic nitrogen fertilizers and replacing them with symbiotically fixed nitrogen, and the use of perennial grass species, a common practice, to reduce carbon loss in cultivated soil [8]. This approach improves the sustainability of the agricultural ecosystem compared to monocropping systems, as well as contributes to the conservation value for threatened bumblebee species [9]. In Estonia, the cultivation of clover-grass mixtures had played significant agronomic purposes in co-cultivation and increased the feed value of the mixture, sequestering nitrogen and thereby reducing the amount of fertilizer [10], and achieving C-balance and carbon sequestration [11] in crop rotation through perennial grassland.

However, it is important to note that with the present trend towards global trade, the increased importation of grain legumes into Europe has led to lower local production in many countries [12]. This trend has raised questions regarding the sustainability and security of protein supplies [13]. Perhaps in opposition to this, the main direction of red clover cultivation goals has focused on forage yields and persistence [14], which may directly improve business competitiveness and increase agricultural versatility. This also highlights the importance of the estimation and quantification of high-yield clover and grass mixtures, and in particular extending from the laboratory to field-based performance trials and studies. Traditional on-site destructive silage and forage biomass sampling and measurements provide accurate reference data for building and assessing yield models. However, it is time-consuming, highly labor-intensive, and limited by large-scale spatial quantity parameter collection [15]. Also, trials and phenotyping techniques have become increasingly criticized in recent years. One of the objectives of this study is to explore and expand the monitoring range from controlled environments, such as laboratories and greenhouses, to bare soil-based situations [16,17]. Variety performance trials (VPT), as an alternative, is a randomized controlled field-based experimental design to improve environment $\times$ management scenario recommendations for variety comparison and breeding selection [18,19]. A typical approach of VPT is recognizing multiple crop phenotypes and their response to management practices. With the current raised awareness of environmental protection and the concept of sustainable agriculture, the adaptability of eco-friendly cultivation techniques, such as reduced tillage and the application of various minerals and organic fertilizers are emerging [20]. However, the description of these managements in current VPT systems is indispensable for vigorous simulation and rigorously assessed for their ability to reproduce measured crop yields that established near-optimal management practices [21].

These limitations and challenges associated with silage and forage legumes have, to some degree, promoted and encouraged the development of remote sensing (RS) solutions for legume biomass estimation tasks [22]. The use of RS offers cost-efficient, non-destructive, and spatially extensive approaches for crop monitoring and trait decision-making. Equally, RS approaches can support yield prediction and help to determine and assess a multifaceted range of plant traits when combined with modeling of phenotypic features [23]. Its application in agriculture is also a vital tool in further understanding plant-environment interactions within the management of crops [24]. Therefore, it is necessary to broaden the horizon of rapid and accurate RS approaches in the analyses of red clover-grass mixture trials in response to multiple farming techniques and operations.

More recently, Unmanned Aerial Systems (UAS) offer significant potential to support and develop current and future red clover trial studies. It has arisen as a cost-efficient remote-sensing platform for capturing high-resolution imagery [25]. The recent expansion

of sensors and cameras that can be integrated with, and mounted on, these systems offers the capability to distinguish and compare plant changes both spatially and temporally for the detection and differentiation of local agricultural practices [26]. For example, the reflectance of vegetation information captured by UAS-borne sensors is verified by biological and morphological features of the surface of tissues or leaves [27]. Depending upon the sensor types (e.g., visible spectrum or multispectral), the vegetation light spectra that can be captured can range from the ultraviolet region (UV), the visible region (RGB), through to near-infrared region (NIR). Furthermore, when captured across a series of these ranges, this information can be used to calculate several vegetation indices (VIs). VIs are widely utilized in the assessment of crop characteristics with the ability to reduce soil or environmental noise and enhance their sensitivity for a target characteristics [15]. One of the most applied multispectral VI is the Normalized Difference Vegetation Index (NDVI) with its ratio between the red and near-infrared bands [28]. However, NDVI is not only sensitive to soil and atmospheric effects, but also certain spectrum ranges were found to have an asymptotic relation, as applicability is limited for higher biomass levels [29,30]. Also, the ability of the reflectance sensor in biomass prediction could be limited by aging crop materials and diverse canopy structures caused by mixed species [15]. Therefore, an alternative for increasing the accuracy of various crop modeling tasks is by increasing the varieties and combinations of adjusted and optimized VIs [31].

Several UAS-based forage clover studies have been conducted in recent years. UAS-RGB-based vegetation indexes and linear regression models were utilized in estimating the red clover dry matter (DM) yield with the best performance R^2 values 0.62 [32]. UAS-RGB-based point cloud data generated into photogrammetric canopy height models (CHM) can also be utilized in forage legumes DM prediction; clover-grass canopies showed better performance than lucerne-grass mixtures for DM prediction [33]. Concerning another study combining CHM, RGB, and VIs with machine learning (ML) techniques for grass swards silage prediction, the Pearson correlation coefficients reached 0.98 [34]. Equally, clover-related phenotypic research has also received much attention in recent years: it included clover-grass pasture coverage and spatial dynamics monitoring [35,36], and quality parameters, such as the digestibility of organic matter, water-soluble carbohydrates, the nitrogen concentration, and uptake [37]. These studies largely used UAS-derived images combined with state-of-the-art ML technologies for qualitative or quantitative analysis. However, red clover silage and forage biomass studies require the establishment and characterization of multiple management options, such as fertilizers, tillage methods, and farming systems, in order to meet the actual planting condition and provide valuable feedback to local farmers and policymakers. There is little scrutiny on the application of DM modeling in the response to the above-mentioned factors.

A gap, therefore, currently exists in the knowledge base for legume biomass analysis and the further understanding of the potential for remotely-sensed solutions to field-based and multifunctional platforms for the demands of precision agriculture. To address this knowledge gap, this study presents a rapid, non-destructive, low-cost framework for field-based red-clover DM yield modeling. The outputs could potentially assist agronomists and farmers in developing precise farming systems and increase the effective monitoring of environmental conditions. More specifically, the objectives of this study are the comparison of two temporal pre-harvest (11 days and 38 days) DM prediction capabilities under one- and two-year clover-grass cultivation fields with three different treatments; and the comparison of the performance of three machine learning algorithms and their corresponding variable importance rankings in estimating clover-grass mixture DM. To address these aims, UAS-multispectral data and six VIs were extracted and used for training and evaluation models.

2. Materials and Methods

2.1. Study Area and Experiment Layout

This study was undertaken at the Agricultural Research Centre (ARC) in Kuusiku (58°58′52.7″N 24°42′59.1″E), Estonia (Figure 1a). The ARC was established in 1924 by an official institution under the governance of the Ministry of Agriculture and consists of consolidated laboratories and field testing centers. The experimental area covers 226 hectares, of which the variety performance trial area we selected to consist of two soil types: Calcaric Cambisol and Calcari-Leptic Regosol [38].

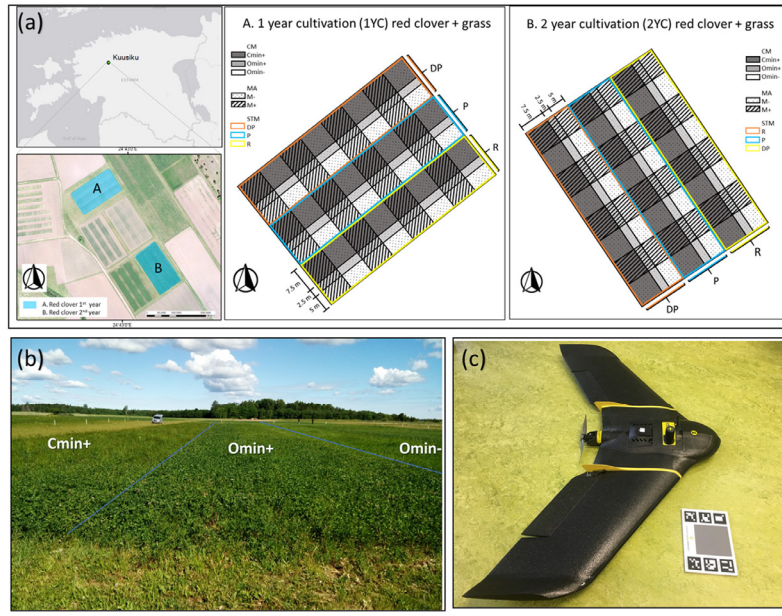

Figure 1. (**a**) The Agricultural Research Centre (ARC) is situated in Kuusiku, Estonia. A. The experiment layout containing three treatments: 1. soil tillage method (STM), 2. cultivation method (CM), and manure application (MA) for one-year cultivation (1YC) in Field A with a total of 72 observation plots, and two-year cultivation (2YC) in Field B equally with a total of 72 plots. For caption descriptions, see Table 1. (**b**) A visual demonstration of the different CM treatments within the 2YC DP area. (**c**) The eBee Plus device was used to capture the multispectral imaging data and the Airinov target was used for radiometric calibration in this study.

In this study, red clover was used in grass-mixture for the representation in practical farming purposes. More specifically, approximately 75% of the mixture consisted of red clover (*Trifolium pratense* L.) tetraploids variety 'Varte', an early variety normally cropped in Estonia, and the remaining 25% of the mix comprised of meadow fescue (*Festuca pratensis*)—variety 'Jõgeva 47'. To assess the potential of UAS-based DM prediction capacity in the agricultural application of the clover-grass mixture field, the experiment was designed with three principal experimental factors (see Table 1). More specifically, these included: (1) soil tillage methods (STM), considering reduced tillage (R) (8–10 cm), ploughing (P) at a depth traditionally used in conventional tillage (18–20 cm), and disking (DP) (8–10 cm)

as treatments; (2) cultivation methods (CM) (as shown in Figure 1b), considering conventional farming with mineral fertilizer application (CMin+), organic farming with mineral fertilizer application (OMin+), and organic farming without mineral fertilizer (OMin-); and (3) manure applications (MA). Considering the convenience of ploughing and fertilizer applications, we present the design of the field and location as shown in Figure 1a, and the treatment details in Table 1. The mixture's fresh aboveground biomass was cut twice in two fields (1YC and 2YC). Each field contains 72 plots, which means a total of 144 plots were sampled; the first cut took place on 10/06/2019, and the second took place on 16/08/2019. The fresh biomass was weighed by plot and dried to verify its DM yield measured in kilograms per hectare.

Table 1. The farming operation and treatment of the red clover experiment fields.

Farming Operation	Treatment	Description
Soil tillage methods (STM)	Reduced tillage (R)	R (8–10 cm)
	Ploughing (p)	P (18–20 cm)
	Disking and ploughing (DP)	D (8–10 cm) & P (18–20 cm)
Cultivation methods (CM)	Conventional framing with fertilizer (Cmin+)	NPK 5-10-25 [1]
	Organic farming with mineral fertilizer (Omin+)	Patentkali [2]
	Organic farming without mineral fertilizer (Omin−)	N/A
Manure application (MA)	With manure application (M+)	M (30,000 kg ha^{-1}) [3]
	Without manure application (M−)	N/A

[1] NPK 5-10-25 (chemical fertilizer) 291 kg ha^{-1} (N-14 kg ha^{-1}, P-13 kg ha^{-1}, and K-60 kg ha^{-1}); [2] Patentkali (mineral fertilizer) 240 kg ha^{-1} (K-60 kg ha^{-1}, S- 41 kg ha^{-1}, M-14 kg ha^{-1}); [3] Manure 30,000 kg ha-1 (N-234 kg ha^{-1}, P-20 kg ha^{-1}, and K-216 kg ha^{-1}).

2.2. Image Acquisition

Figure 2 presents a workflow of the methodology used to combine the UAS-based image collection, processing, biomass sampling, as well as modeling and evaluation within ML algorithms. To capture data for both image processing and biomass evaluations, the UAS imaging was conducted twice [i.e., 11 days before 1st cut (11DB) and 38 days before 2nd cut (38DB) harvesting] in the summer of 2019. Due to the needs of the other experimental areas, the data of 80 hectares were collected, of which 2.4 hectares were used in this study. An eBee Plus device (Figure 1c), with onboard GNSS post-processed kinematic (PPK) capabilities, was deployed and equipped with a Parrot Sequoia multispectral sensor. The Parrot Sequoia sensor captured imagery across four spectral bands: near-infrared (770–810 nm); red-edge (730–740 nm); red (640–680 nm); and green (530–570 nm). The flight lines overlap was set with a frontal image overlap of 80% and lateral image overlap of 75%. All the operations took place between 10 a.m. to 2 p.m. to ensure consistency with the sun's angle, and to reduce lateral shading within the experimental fields. The images were captured from a height of 120 meters, and the resulting images had ground sampling distance (GSD) of 10 cm per pixel. Prior to each flight mission, an Airinov radiometric calibration target and one-point calibration method [39] was used to facilitate post-flight radiometric correction of the multispectral imagery. Table 2 provides a summary of environmental conditions for the two flights conducted prior to clover-grass biomass harvesting.

Table 2. Specification of the weather condition and the corresponding time before harvest.

Date of Flight	Weather	Wind Speed (km/h)	Wind Direction	Temperature (min-max°C)	Humidity	Operation
30 May 2019	Sunny	11	S	15–16	35%	11 days before 1st cut (11DB)
1 July 2019	Overcast	12	WSW	19–20	64%	38 days before 2nd cut (38DB)

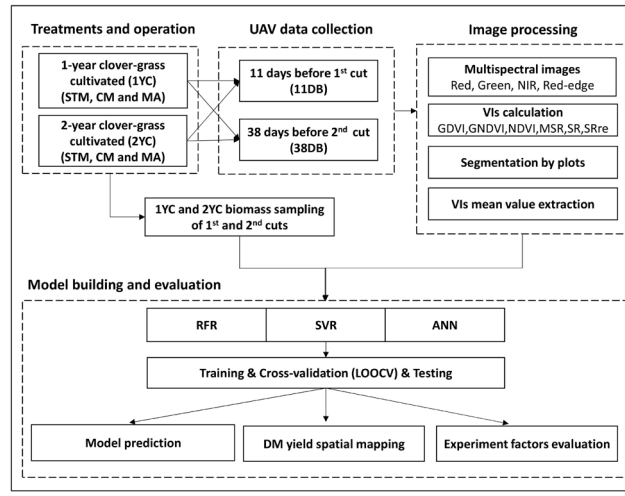

Figure 2. The methodology flowchart of red clover-grass mixture DM modeling. The image processing rectangular dotted box contains all predictors extracted from the UAV images, and the modeling building, and evaluation rectangular dotted box contains all modeling methods and analysis procedure.

2.3. Image Processing and Analysis

The UAS data was post-processed in SenseFly eMotion 3 [40] using receiver independent exchange (RINEX) format data provided by the GNSS CORS (Continuously Operating Reference Station) of Estonia [41] for post-processing kinematics (PPK) corrections. This post-process provided an increase in the geotagging accuracy [42] of the UAS images from 5 m error to under 0.06 m, where the method and accuracy obtained is similar to [43]; and thus less than the one-pixel size in our study. Pix4D v.4.3.31® (Pix4D SA, 1015 Lausanne, Switzerland) software was utilized to process and radiometrically correct (default in Pix4D) the imagery and generate the multispectral orthomosaics. These images were subsequently clipped to represent only the extent of the experimental area.

2.4. Vegetation Indices Calculation and Extraction

In this study, six VIs were calculated using R version 4.0.2 [44] (Table 3). The normalized difference vegetation index (NDVI) utilizes the reflectance (ρ) in the NIR and Red wavelengths, and the outputs range from −1.0 to 1.0. This index was selected for this study as it has a sensitive response to track physiological dynamics and biomass [45]. However, NDVI reaches saturation when leaf area index (LAI) values are about 2.5–3 or in dense crop canopies [46,47]. The green normalized difference vegetation index (GNDVI) was also calculated and outputs values range from 0 to 1. Previous studies have shown GNDVI to be linearly correlated with LAI and biomass, with the ability to reduce the effects of soil reflectance and estimate nitrogen conditions [48]. Similarly, the Simple Ratio (SR), a normalization of ρ NIR against ρ Red, was calculated as it has been previously shown that this index can better indicate the strength of canopy photosynthetic material and yield prediction better than NDVI under different nitrogen supplies [49]. The Red-Edge Simple Ratio (SRre) formula was calculated by replacing the ρ Red band with the ρ Red-edge. Its inclusion in the assessment was due to previous studies indicating a higher correlation with

plant nitrogen concentration compared to ρ Red based VIs. This can lessen the soil background influence on crop reflectance [50]. Finally, the Modified Simple Ratio (MSR), as a potentially improved version of the Renormalized Difference Vegetation Index (RDVI), was calculated to linearize the relationship between biophysical parameters [51] and enhance the sensitivity of vegetation occurrences, which can be observed in other VIs.

Table 3. Descriptions and formulas of NIR-related VIs used in this study.

Vegetation Index	Description	Equation	Reference
NDVI	Normalized Difference Vegetation Index	$(\rho\ NIR - \rho\ R\ ^{1})/(\rho\ NIR + \rho\ R)$	[28]
GNDVI	Green Normalized Difference Vegetation Index	$(\rho\ NIR - \rho\ G\ ^{2})/(\rho\ NIR + \rho\ G)$	[52]
GDVI	Green Difference Vegetation Index	$\rho\ NIR\ ^{3} - \rho\ G$	[53]
SR	Simple Ratio	$\rho\ NIR/\rho\ R$	[54]
SRre	Red-edge simple ratio	$\rho\ NIR/\rho\ REG\ 4$	[55]
MSR	Modified simple ratio	$((\rho\ NIR - \rho\ R) - 1)/(((\rho\ NIR + \rho\ R) * (0.5)) + 1)$	[51]

[1] ρ R refers to red band, [2] ρ G refers to green band, [3] ρ NIR refers to near-infrared, and [4] ρ REG refers to the red edge.

The two experimental fields [i.e., 1YC (n = 72) and 2YC (n = 72)], with a total of 144 plots were digitized in ArcGIS Pro 2.6.3 [56]. The average VIs within each plot were extracted and calculated as the VIs of each plot at the experiment site. To avoid potential edge effects in the fertilizer treatment, a one-meter buffer zone was extended inwards from each plot boundary, and data sampled within this target region (Figure 3). These extracted values were further used in this study when building ML algorithms for clover-grass mixture DM yield estimation and evaluation.

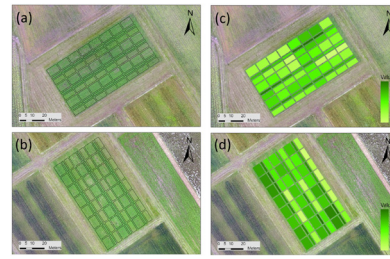

Figure 3. A demonstration of VIs (GDVI as an example) zonal statistics in 1YC and 2YC fields. (**a**) RGB image with 1-meter buffer zone plot polygons from 1YC11DB, (**b**) RGB image with 1-meter buffer zone plot polygons from 2YC11DB, (**c**) GDVI zonal statistic with a region of interest (ROI) in 1YC11DB, and (**d**) GDVI zonal statistics with ROI from 2YC11DB.

2.5. Machine Learning Techniques

Parametric regression models may lead to multicollinearity between covariates and overfitting, which renders them impractical when dealing with highly dimensional remotely sensed data [57]. Conversely, machine learning algorithms can handle high volumes of predictor variables that are interrelated and have a non-linear relationship with response variables [58]. A recent remote sensing-based ML study collected data from 220 related articles and found that random forest (RF), support vector machine (SVM), and artificial neural network (ANN) algorithms were amongst the most used ML techniques [59]. Therefore, these derived ML regressions [random forest regression (RFR), support vector regression (SVR), and artificial neural network (ANN)] were chosen for modeling DM in this study. These algorithms were programmed in Python [60] (version 3.8). The VI values presented

in Table 3 were used as continuous predictor variables of the DM regression models, which were divided into training sites and prediction sites, and the parameters of each algorithm were adapted to ensure the performance as effectively as possible for our training and testing dataset.

First, to reduce the potential over-fitting problem of the model, a leave-one-out cross-validation (LOOCV) procedure [61] was conducted to validate the three ML techniques. The LOOCV procedure involves creating a model by separating one sample for testing and the rest (n = 36) for validation in every iteration. (Figure 4a). Second, all training sites were used to model and predict the three ML methods (Figure 4b). As the training and testing dataset comprised two repetitions of results from each treatment, a comprehensive range of crop conditions were covered by the modeling. The variable importance of the VIs was calculated for each ML technique differently and was listed per each model's VIs importance scores, and the suitable models for different periods for DM yield spatial mapping were demonstrated. Finally, experimental treatments (i.e., STM, CM, and MA) were used to explore the relationship between different experiment factors and models (Figure 4d). The testing sites were sampled following a stratified approach based on the three different farming operations, 3 strata with 12 samples in each one in STM (DP, P, and R) and CM (Cmin, Omin+, and Omin−) groups, and 2 strata with 18 samples in MA (M+, and M−). The ML methods, selected parameters, model evaluation techniques, and variable importance calculations are described below.

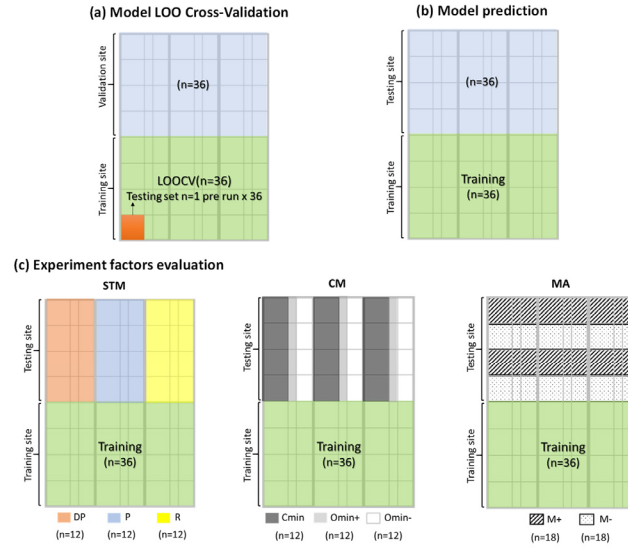

Figure 4. Examples of different sampling methods and regions in field 2YC, the training site (1/2, n = 36) contains two repeated trial plots. (**a**) The evaluation of RFR, SVR, and ANN using the cross-validation method LOOCV in the training site (1/2, n = 36) and validation site (1/2, n = 36). (**b**) Model construction including training site (1/2, n = 36) and testing site (1/2, n = 36). (**c**) Evaluation of the model efficiency across three different treatments: STM (n = 12 for each subset), CM (n = 12 for each subset), and MA (n = 18 for each subset).

2.5.1. Random Forest Regression

An adaptation of the Random Forest (RF) algorithm [62] was conducted for DM regression models (i.e., RFR). The RFR algorithm fits an ensemble of decision tree models to a set of data. The regression tree algorithm creates individual decision trees automatically based on randomly chosen samples and subsets of the training data. For random forest construction, the best split is selected among a random subset of the predictors at each node. Calculations were conducted with 100 trees, the minimum number of samples required to split an internal node was set to 2, and the minimum number of samples required to be at a leaf node was set to 1. Tests were run to confirm regression accuracy by using different amounts of trees ranging from 100 to 500, and it was noted that accuracy did not vary substantially with this parameter. Similar results have also been observed in other RF studies [63]. In terms of variable importance, the feature importance values were extracted using the *feature_importances* object located in the *sklearn.ensemble*. *RandomForestRegressor* class. The algorithm calculates these percentage values based on how every feature decreases the impurity of the split (mean decrease impurity) in each decision tree. The average across all trees in the forest represents the feature importance.

2.5.2. Support Vector Regression

Support vector regression (SVR), which is a Kernel-based machine learning method, was used for its low dimensional and quadratic programming (QP) problem converted ability with usually only a scarce training data set needed [64]. For this study, a linear kernel was used. Three extra parameters were set for the algorithm. The first included the regularization parameter (*C*, cost) set at 500. This parameter controls the trade-off between achieving a low error on the training data and minimizing the norm of the weights. The second parameter, gamma, was set at 0.5. It defines how far the influence of a single training example reaches. The third parameter, epsilon, gives a margin of tolerance and was set at 0.01. In terms of variable importance, the coefficients of all six predictors estimated by the inner sklearn algorithm were extracted from the created SVR model using the *coef_ value* located in the *sklearn.svm.SVC* class, and then rescaled to be in terms of percentage.

2.5.3. Artificial Neural Network Regression

The gradient-based artificial neural network (ANN), which is also called multi-layer perceptron, is a supervised algorithm that can learn nonparametric and nonlinear features that simulate human brain neural network spreading between layers and receivers and information processing [65] for classification or regression tasks. Execution of the ANN algorithm required fine-tuning of certain parameters. In this study, *lbfgs*, which stands for *Limited-memory Broyden–Fletcher–Goldfarb–Shanno*, was used as the solver since it was most optimal in saving memory. The *MLPRegressor* algorithm was executed using one layer with fifteen hidden units, with the regularization parameter (*alpha*) set at 0.00005. The maximum number of iterations allowed for this algorithm was set to 100,000. In terms of variable importance, the weights of all six predictors assigned by the inner *MLPRegressor algorithm* were extracted from the created model using the *coefs_* object located in the *sklearn.linear_module.Perceptron* class, and then rescaled to be in terms of percentage. The SVR and ANN importance scores were similarly extracted but were rescaled to also be in terms of percentage.

2.5.4. Model Evaluation

The evaluation was performed for LOO cross-validation, model prediction and experimental factor assessment (Figure 4). For the evaluation of each model, the accuracy evaluation method described by [66] was used. The models' accuracies were measured

by the coefficient of determination (R^2) (Equation (1)) and normalized root means square error (NRMSE) (Equation (2)). The equations used are as follows:

$$R^2 = 1 - \frac{\sum(\hat{y}_i - y_i)^2}{\sum(y_i - \overline{y})^2} \tag{1}$$

$$NRMSE = \frac{\sqrt{\sum((\hat{y}_i - y_i)^2)/n}}{\Delta y} \tag{2}$$

where: y_i represents the ith observation value of the training dataset; $\overline{y}$ represents mean value of the training dataset, $\hat{y}_i$ is the model predictions, n is the number of observations, and Δy represents the difference between the maximum and minimum values of the training dataset.

3. Results

3.1. The Field Observation DM Data Analysis

The violin plot (Figure 5) displays the average true ground DM data collected from 144 plots taken during two harvest periods (1st cut and 2nd cut) of two experimental areas. Since the treatments were mixed with each plot, we only displayed the range of DM yield data, and grouped them by three farming operations, namely STM, CM, and MA. The potential interaction effects between treatments were not addressed during the analyses. The results of 1YC (Figure 5a) for STM showed that treatment P had a relatively lower DM than the R and DP treatments with a mean DM of 5195 kg ha-1. However, the effect of STM was noticeable in the two-year cultivation (2YC), in which the DM yield of R and P decreased in the 2YC field (Figure 5b). In terms of CM, DM from organically cultivated crops without mineral fertilizer (Omin-) can be expected to show lower yields in both fields. The DM yield of 2YC was slightly less than that of 1YC.

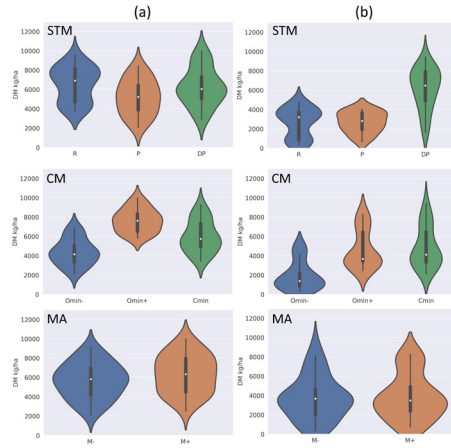

Figure 5. Violin plots of mean harvest results of clover-grass mixture dry matter (DM) yield in the (**a**) 1YC field and the (**b**) 2YC field, grouped by STM, CM, and MA farming operations. Biomass data were obtained during two separate harvests (1st cut and 2nd cut). The white dot indicates the median, while the thick black bars in the center show the interquartile range, the black line represents the rest of the distribution, and violins show point density and data distribution.

3.2. The Red Clover-Grass Mixture DM Modeling and LOOCV

For better evaluation evaluating of the model performance, LOOCV was firstly conducted in this research to assess three machine learning methods' (RFR, SVR, and ANN) abilities to predict DM yield using six VIs. (Figure 4a) The distributions of R^2 and NRMSE values under thirty-six LOOCV iterations of the three models are shown in box plots (Figure 6). The results showed that in terms of flight dates, 11DB generally performed better than 38DB. In terms of model performance, ANN had the best performance in 11DB, with the highest R^2 values (1YC = 0.84, 2YC = 0.85), and the lowest median NRMSE with a stable distribution of outliers. RFR's accuracy was slightly smaller than for ANN. In terms of regional prediction, 2YC was on average better than 1YC. Overall, the results of LOOCV showed that the three models have moderate to high accuracies in two locations and different flight dates, and the best R^2 values were observed in 2YC11D (0.80 to 0.85), while the worst were observed in 1YC38DB (0.64 to 0.70).

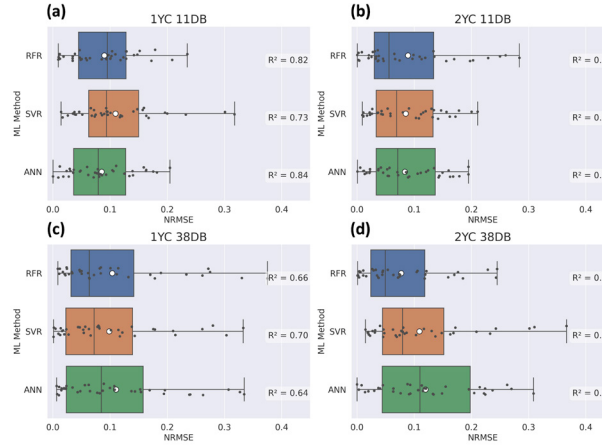

Figure 6. Comparison of the NRMSE and R^2 values resulting from 3 different ML methods (RFR, SVR, and ANN) of LOOCV in (**a**) 1YC11DB (**b**) 1YC38Db (**c**) 2YC11DB (**d**) 2YC38DB. Each model performed 36 times LOOCV calculations. The R^2 for LOOCV was calculated using the average variance between the actual and prediction value for every iteration of the cross-validation. The black dots showed the NRMSE results of each cross-validation, and the white dots represent its average value. The median line in the box shows the middle value, and the interquartile range of the box (shown in blue, red, and green) represent the 25th to the 75th percentile.

3.3. The Red Clover-Grass Mixture Model Prediction and Variable Importance

After the cross-validation, the training dataset (n = 36) was used for the calculation of models separately (Figure 4b). During the modeling phase, the appropriate combinations of the parameters of the data set were tested. The scatter plots with model predictions and observed DM values were compared to the 1:1 line, and their corresponding variable importance values are shown in Figures 7 and 8.

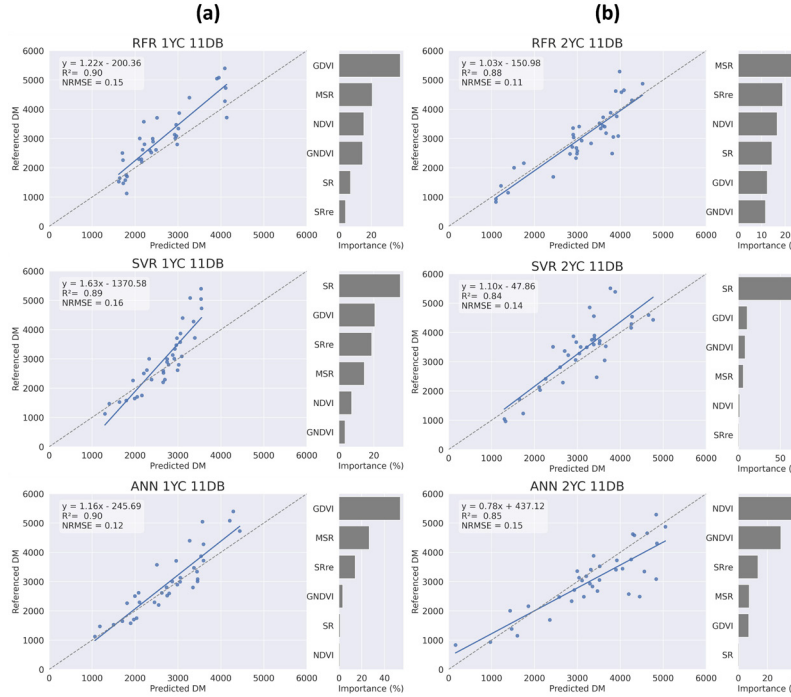

Figure 7. Regression plots of 1YC and 2YC fields based on RFR, SVR, and ANN methods in 11DB flight. The plots correspond to (**a**) 1YC11DB and (**b**) 2YC11DB. The horizontal bar plots on the right side of each graph shows the variable importance estimation based on the models. The horizontal axis in the scatter plots describes the predicted DM yield acquired from the model, and the vertical axis stands for the field-observed DM yield. The R^2 = coefficient of determination, NRMSE = normalized root means squared error, and the black dotted line exemplifies the 1:1 slope.

Figure 7 shows three red clover-grass mixture DM models across the 1YC and 2YC fields where the images were captured eleven days before harvesting (11DB). The results indicate that, in 1YC11DB (Figure 7a), the ANN model had the lowest prediction errors (NRMSE = 0.12) and the highest R^2 value (0.90). RFR had a similar performance, but with higher NRMSE. Although the three models performed well, a slight uniform underestimation of DM yield appeared in both the RFR and ANN models. On the other hand, a non-uniform bias appeared in the SVR model, which overestimated small DM values and underestimated large DM values. According to the ranking of variable importance, GDVI and MSR provided higher contributions to the RFR and ANN models. Concerning the SVM model, larger contributions were found for SR and GDVI. The results of 1YC38DB (Figure 7b) show that the three models performed relatively well (R^2 from 0.84 to 0.88). The slope of RFR was closest to the 1:1 line, with the smallest NRMSE (0.11) and highest R^2 value (0.88); ANN and SVR had similar predictive capabilities. With regard to variable importance ranking, SVR showed similar results compared to 11DB with the highest con-

tribution of SR, while in the results of RFR and ANN, the contribution of VIs did not show obvious similarity with the highest ranking in MSR and NDVI, respectively.

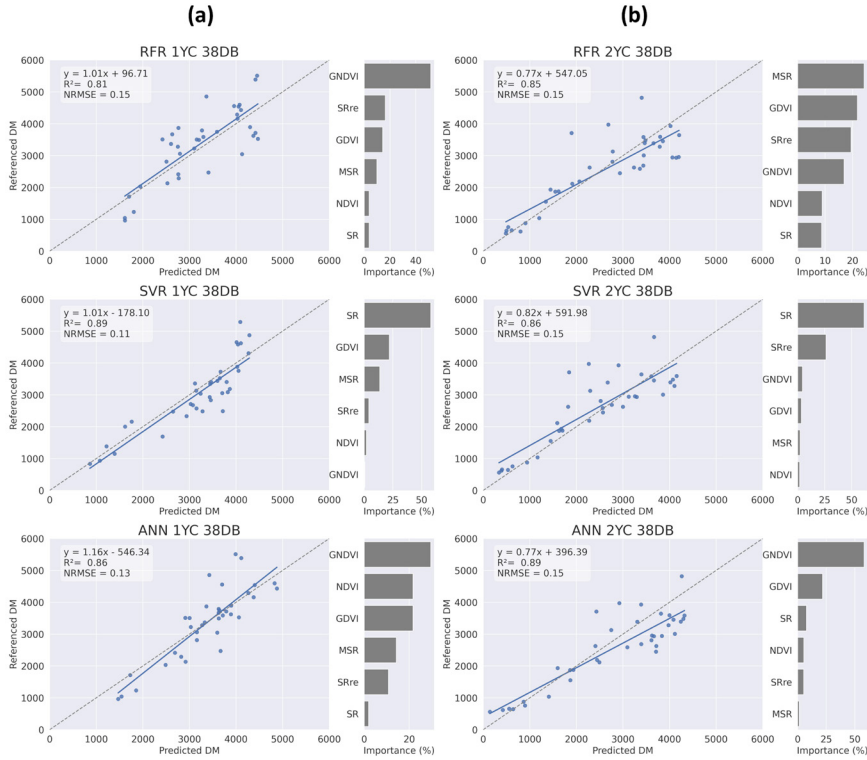

Figure 8. Regression plots of 1YC and 2YC fields based on RFR, SVR, and ANN methods in 38DB flight. The plots correspond to (**a**) 1YC38DB and (**b**) 2YC38DB. The horizontal bar plots on the right side of each graph shows the variable importance estimation based on the models. The horizontal axis in the scatter plots describes the predicted DM yield acquired from the model, and the vertical axis stands for the field observed DM yield. The R^2 = coefficient of determination, NRMSE = normalized root means squared error, and the black dotted line exemplifies the 1:1 slope.

Figure 8 shows the behavior of predictive models using the thirty-eight days before harvest (38DB) datasets. The 1YC38DB results (Figure 8a) showed that SVR had the highest R^2 (0.89) and the smallest NRMSE (0.11), where the slope was close to the 1:1 line. In contrast, ANN and RFR relatively had weaker performances. However, the overall performance of the models for the 2YC dataset was slightly inferior to the result of 1YC, showing a higher bias of the slopes. In terms of 2YC38DB results (Figure 8b), ANN had the best performance among the three algorithms (R^2 = 0.89, NRMSE = 0.15). In the ranking of predictor variables in 1YC and 2YC, the GNDVI, MSR, NDVI, and SRre played crucial roles in both RFR and ANN models. Besides, the GNDVI was the most important variable

in both 1YC and the model of ANN in the 2YC field. In contrast, SR stably ranked as the most important in the models of SVR, regardless of flight dates or regions.

Based on the evaluation of models (Figures 7 and 8) and their suitability for different periods (Figure 6), we generated prediction maps (Figure 9) of DM yield for both experimental sites.

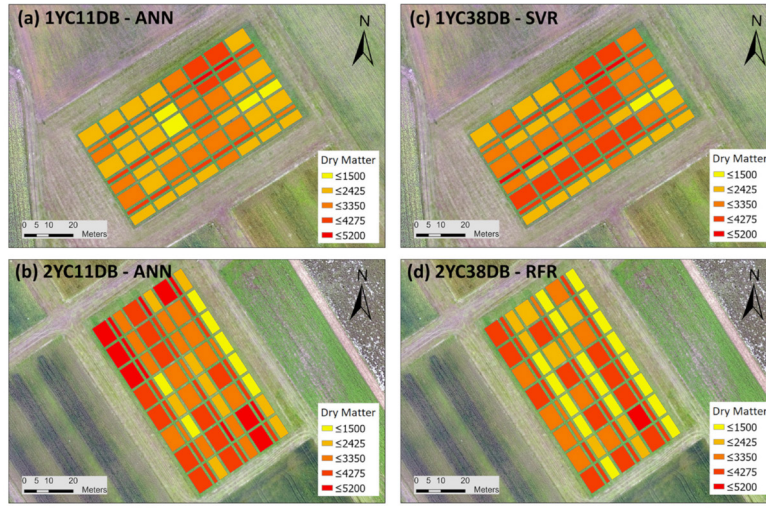

Figure 9. The spatial testing (prediction) site mapping output of DM yields (kg ha^{-1}) in 1YC and 2YC fields based on 11DB and 38DB flights by machine learning techniques at the plot level: (**a**) 1YC11DB, (**b**) 2YC11DB, (**c**) 1YC38DB, and (**d**) 2YC38DB. The best performing algorithm was chosen for each of the four categories, as shown in the previous results.

3.4. The Response of DM and VIs to Different Soil Tillage Methods (STM), Cultivation Method (CM), and Manure (MA) Treatments

To understand the detection capability of the relationship between VIs and DM yields under the influence of different soil tillage methods (STM), cultivation method (CM), and manure treatment (MA), three models were trained from previous 11DB datasets (Figure 4d). Testing sites were stratified sampling based on the three different farming operations to evaluate the predictive ability and sensitivity of detecting DM yields. The NRMSE and the R^2 values were performed for the goodness-of-fit measurement in 1YC (Table 4) and 2YC(Table 5) fields. Table 4 shows that, based on STM, the three models performed well in the DP, P, and R treatments, with the overall R^2 value ranging from 0.85 to 0.94, and NRMSE ranging from 0.19 to 0.37. In contrast, in CM, the performance of the three models showed lower accuracy in Cmin and Omin+. SVR had the worst performance with R^2 (0.62 and 0.74, respectively) and NRMSE (0.48 and 0.26). While in MA, the prediction ability of the three models was satisfactory. As Table 5 shows, the predictive ability of STM remained steady in the 2YC field. The R^2 ranges from 0.71 to 0.96, and NRMSE ranges from 0.08 to 0.33 in three models. In CM, Omin+ showed the worst R^2 and NRMSE values among other treatments with RFR (0.70 and 0.84, respectively) and, SVR (0.81 and 0.25) and ANN (0.49 and 0.35). In terms of the prediction ability of MA treatments, M- performed better than M+ with the R^2 values higher than 0.95, and the NRMSE values were less than or equal to 0.11 in all three models.

Table 4. The evaluation of the models between observed and predicted DM yields under different experimental factors in 1YC11DB, the corresponding models for use was equivalent to Figure 7a. n = number of testing samples; R^2 = coefficient of determination; NRMSE = normalized root mean square error.

1YC11DB			RFR		SVR		ANN	
Treatments		***n***	**R^2**	**NRMSE**	**R^2**	**NRMSE**	**R^2**	**NRMSE**
STM	DP	12	0.85	0.26	0.81	0.34	0.92	0.29
	P	12	0.94	0.29	0.90	0.31	0.92	0.37
	R	12	0.90	0.20	0.85	0.23	0.90	0.19
CM	Cmin	12	0.75	0.31	0.62	0.48	0.80	0.31
	Omin+	12	0.69	0.31	0.74	0.26	0.60	0.47
	Omin−	12	0.89	0.18	0.85	0.27	0.82	0.20
MA	M+	18	0.88	0.19	0.82	0.21	0.86	0.23
	M−	18	0.91	0.21	0.84	0.22	0.88	0.24

Table 5. The evaluation of the models between observed and predicted DM yields under different experimental factors in 2YC11DB, the corresponding models for use was equivalent to Figure 8a. n = number of testing samples; R^2 = coefficient of determination; NRMSE = normalized root mean square error.

2YC11DB			RFR		SVR		ANN	
Treatments		***n***	**R^2**	**NRMSE**	**R^2**	**NRMSE**	**R^2**	**NRMSE**
STM	DP	12	0.93	0.14	0.92	0.13	0.96	0.08
	P	12	0.85	0.20	0.89	0.20	0.71	0.33
	R	12	0.91	0.16	0.94	0.16	0.87	0.24
CM	Cmin	12	0.92	0.19	0.96	0.19	0.76	0.25
	Omin+	12	0.70	0.24	0.81	0.25	0.49	0.35
	Omin−	12	0.90	0.14	0.93	0.12	0.91	0.15
MA	M+	18	0.79	0.17	0.81	0.16	0.69	0.24
	M−	18	0.95	0.10	0.95	0.11	0.96	0.11

4. Discussion

This study has presented a rapid, non-destructive, low-cost framework for field-based red-clover DM yield modeling. The outputs have the potential to markedly assist agronomists and farmers in developing precise farming systems and increase the effective monitoring of environmental conditions.

4.1. Applicability of the Method

The prediction models covered three different agricultural operations (STM, CM, and MA) to represent the variable conditions in a practical farming system, which provided varied dry mass data to identify the robustness of the derived ML models. Acquisition of data conducted during two different periods offered a wider range of suitable monitoring capabilities. Three machine learning techniques (RFR, SVM, and ANN) were conducted to explore the DM yield prediction ability in legume pasture fields. All VI information was derived from four multispectral bands. Utilizing various Vis, which calculate the relative values or ratio among wavelengths can reduce the impact of radiance effects caused by individual reflectance spectra [67]. Besides, high-resolution multispectral imaging produces continuous and accurate indices in contrast to simple visual scores and rankings [68]. Consequently, no additional sensors were needed, which reduced measurement errors and increased cost-efficiency. It is important to consider that the eBee platform used in this study is potentially less practical for small farm area investigation (when compared to multi-rotor drones) [69]. The flexibility, however, in sensor systems, environmental capabilities, and increased flight durations could expand application to meet a diverse

range of requirements [15] (e.g., larger-scale farmland or coverage area with shorter tasks duration); whilst still providing precise yield prediction accuracy required in our study.

4.2. The Impact of the Cultivated Period, Flight Times, and Farming Operations

The first specific objective was the development of red-clover biomass prediction models for one- and two-year cultivation periods, following common farming schedules in Northern Europe. Despite the results of the established models (Figures 7 and 8), the prediction accuracy of 1YC (R^2 ranges from 0.81 to 0.90, NRMSE ranges from 0.11 to 0.15) and 2YC (R^2 ranges from 0.84 to 0.89, NRMSE ranges from 0.11 to 0.15) were both adequate. Nevertheless, the combination of clover-grass resulted in a heterogeneous canopy with the coverage of the two components. A previous mixed clover-grass study focused on canopy height (CH) modelling for DM yield prediction and showed the models performed better when established separately among the two species, and cannot be easily shifted to other grassland types owing to their structural characteristics [33]. Another study indicated that in terms of legume cover crops, the performance of NIR-based VIs did not perform much better than CH at the end of its crop cycle in terms of DM yield estimation [70]. This limitation of applying VIs in the mature growth stage of legumes does not necessarily impair the detection capacity, as forage biomass is usually harvested during the vegetative growth cycle [71]. A previous study found that although measurements performed at the ground-level were more accurate, the use of aerial systems was preferred since species identification was irrelevant when predicting the biomass of mixed-grass [72]. Therefore, these species-dependent VI phenomena seem to be a minor concern within the results presented in our study. We can infer that NIR-based VIs are suitable for the estimation of DM yields in one- and two-year cultivation periods in this study.

The second objective was to compare the impact of different pre-harvest flight dates on model estimation capabilities. The choice of flight timing was crucially matched with the spectral reflectance data during various growth periods. A previous study showed that the ideal period for forage crop assessment was one day before harvesting [32], whereas another study suggested that the targeted silage harvesting stage was favored [34]. Interestingly, studies focusing on other crop species have identified optimum monitoring periods in the case of maize yield prediction. The choice of performing a flight 100 days before harvest times had the best accuracy [73], whereas, for rice grain, it was optimal around the booting stage of growth [74]. In the case of our study on clover-grass mixes, the LOOCV and ML analysis demonstrate that the NRMSE average values from three models of 11DB were typically lower than those of 38DB (Figure 6). However, when compared to the bias of the slope of 11DB and 38DB from 1YC with the 1:1 line, 38DB provided a better model fit whilst at the same time being less prone to the under- or over-estimation of DM yield (Figures 7a and 8a). Although the 11DB flight was in the red clover flowering stage, it seemed insignificantly influenced by the multispectral reflectance values since the flower size was small and normally less than 0.25 pixel in our study. Thus, the results indicate that in clover-grass mixture fields, the estimation ability is improved when UAS imagery is collected closer to the harvest period, but not later than the yellowing stage. Moreover, the results also indicate that the VIs derived from UAS images captured earlier than 38 days before harvest also have sufficient DM-yield estimation capacity and provide the potential to estimate DM earlier, while the accuracy might be partially lost.

The third objective was to explore, in detail, selected ML sensitivity regarding different farming operations. Relative research on the relationship between STM, CM, MA, and forage crop DM yields are still scarce in remote sensing studies. We redivided the testing site and stratified sampling the plots based on the differences in farming operations (Figure 4d), and the evaluation results of prediction ability based on STM was generally better in both 1YC and 2YC, followed by MA and CM treatments. Recently, remotely-sensed soil tillage quality evaluation on the bare soil level has contributed to compare several tillage techniques to increase the quality of farming work and energy conservation [75]. However, there are still few applications of the assessment of tillage methods and yield performance

during the growth period of forage crops. The results of this study, which considered crop growth period when measuring DM yield, will assist farmers in making comparisons between several tillage methods for covering their basic agricultural needs.

On the other hand, a study confirmed that a biomass coverage index of 43.8% represents the best performance of disking operations [76]. Similarly in our study, a higher DM of 2YC red clover in DP treatment was observed (Figure 5b) and could be accurately predicted. However, the general predictive ability of CM was low, especially in the Omin+ treatment group, which may be attributed to the larger DM and spectral reflectance divergence of this area. In terms of tillage modeling results, previous studies have shown the non-tillage (NT) fields had more vigorous and abundant crops than conventional tillage (CT) methods, and the NIR-based VIs showed better discrimination performance than RGB-based VIs [26,77]. In our research, we have also found that the reduced tillage (R) areas yielded better accuracies, in comparative terms. A study of grassland DM yield estimation by a UAV-RGB camera showed different nitrogen fertilizer levels with its R^2 ranging from 0.57 to 0.70 [78]. Currently, adequate fertilizer estimation remains challenging in heterogeneous plant communities such as grasslands [32]. However, legume crops could provide more positive N balance input than mineral fertilization under various tillage conditions [79]. Although the input of N is not effectively quantified in this study, the DM yields of the various N input combinations could still be effectively predicted, which increases the viability of using non-destructive methods to quantify a range of, and distinct, N sources in future fertilizer management decisions.

4.3. The Machine Learning Methods

Machine learning techniques are still deemed to be novel in the realm of estimating grassland biomass [80]. The predictive ability of three broadly adopted and reliably implemented ML methods in clover-grass DM yield was promising in this study. ANN showed better predictive accuracy eleven days prior to harvest (11DB). This result is consistent with the LOO cross-validation results. The practicality and flexibility of ANN has previously been demonstrated in studies of grassland biomass estimation [81,82], and nitrogen and phosphorus concentration modeling in mixed-species environments [83,84]. Interestingly, within our study, RFR and SVR were shown to have increased predictive capability at 38DB; which is farther from the harvest period. Both RFR and SVR were also shown to have a promising potential in clover-grass biomass prediction applications, since they are fast and require fewer training samples, when compared to the ANN [80,85]. The overall accuracy of the three ML methods provided R^2 ranges from 0.81 to 0.90, and the NRMSE ranges from 0.11 to 0.15. These findings further support the asserted dominant ability of MLs as a perennial forage crop biomass estimator; demonstrated in this study for mixed-grass species.

4.4. Importance of Variable Rankings

Variable importance ranking is essential for predictor selection and model simplification normally. In our study, the results of ranking showed which VIs were able to capture most of the variability in vegetation characteristics from the grass fields. Different VI values at the level of leaf area indices were likely caused by the diverse canopy structures of clover (horizontal) and grass leaves (vertically orientated) [71]. A recent study confirmed that GNDVI is suitable as a biomass predictor for perennial forage crops, where $R^2 = 0.80$ for freshly-cut, and 0.66 for dry yields [86], as well as in the grain yield estimation in maize [87]. These results resemble the RFR and ANN modelling of this study, where the GNDVI, GDVI, and MSR had the highest average contributions. The weight of SR was generally low. A similar result was also found in a study of grass DM yield prediction by Partial Least Square (PLS) and RF techniques, where the above-mentioned VIs were relatively important variables, while SR yielded the worst prediction out of twelve VIs [22].

Other previous studies have indicated that NDVI is more commonly used for pasture biomass measurements [88,89], as well as in larger-scale grassland followed by seasonal

monitoring [90]. However, the findings of this study indicate that NDVI may not be the most suitable VI, which was supported by a previous study [86]. This highlights the importance of considering the saturation, sensitivity, stages of crop development, canopy structure, and the type of environment when testing various vegetation indexes [15]. Our findings indicate that multispectral information based on NIR and the green band may be more suitable for DM yield prediction using RFR and ANN modelling. The exception, however, is the SR indices, which have the highest contribution consistently across all periods in the SVR modelling. This distinctive finding has yet to be found in similar crop studies in other literature. It is clear from our findings that more tests of VI should be conducted in studies to increase the collective understanding and improve the knowledge base.

4.5. The Limitations in This Study

In this study, the results may be limited by the red-clover and grass varieties present, and the study area investigated. This, however, can easily be addressed by including a wider range of species, and study regions, in future investigations. Further limitations may exist due to the inherent complexity and repeatability of field trial design, the small sample size, and the potential interaction effects between treatments; none of which are fully addressed through this study. Although RFR is known to be a suitable methodology for measuring smaller sample sizes [91], basic tree learners within the RFR algorithm benefited the performance of small datasets [62], whereas a higher number of sample points will typically lead to more accurate predictions. However, to address the potential for model over-fitting, we also implemented LOOCV in the training site. Here the results showed that the stability and divergence of variance were in an acceptable range. Nevertheless, increasing the number of samples will still be worth pursuing in future investigations, despite it creating an increased burden for onsite sampling and measurement.

5. Conclusions

Agriculture is experiencing a reimagined technological revolution supported by remote sensing technologies. In our study, UAS offers a significant sensor-based platform that supports VPT to develop current and future smart farming trends to achieve the assessment of eco-efficiency agriculture management practices and above-ground biomass estimation. It offers scientists and practitioners the capability to distinguish and compare trial changes, both spatially and temporally, for the monitoring and optimization of local agricultural practices. Our study has highlighted the potential for innovative machine learning methods to compensate for reduced sample sizes, reducing human efforts, and maximizing the utilization of available resources when implementing and simulating the actual activities in perennial clover-grass mixture trials.

We performed multispectral-UAS flights, under the one- and two-year cultivated red clover-grass mixture performance trials, within 38 to 11 days before the harvesting, with the GSD 10 cm and combined the resultant VI's within multiple ML methods. Our findings present a robust DM yield prediction method, which can operate at the farm-scale; which is both non-destructive and cost-effective. The ML analysis results showed the best performance for ANN in the 11DB (R^2 = 0.90, NRMSE = 0.12), followed by RFR (R^2 = 0.90 NRMSE = 0.15), and SVM (R^2 = 0.86, NRMSE = 0.16). For VI performance, GNDVI and GDVI, and MSR performed well as predictors in ANN and RFR. While the prediction ability of models was being influenced by the farming operations, the stratified sampling based on STM provided a better model performance than CM and MA. The results indicate the potential of UAS to deal with complex experimental design development, such as tillage methods and varying fertilizer inputs.

However, the robustness and applicability of fertilizer quantification, and the mixed legume-grass species distribution detection, still remain to be addressed. Accurate and real-time phenotypic information of crops under diverse agri-environment schemes and their morphological and physiological states remains a further obstacle to be well-quantified by

UAS, as well as providing theoretical and technical support for sustainable agricultural development in response to forage crop biomass/yield estimation in the future. The proposed methods in this study could also be improved further through the implementation of other practical techniques; such as texture analysis, which could offer the potential to measure spatial heterogeneity and improve accuracy [92]. In addition, further improvement could be developed by employing other ML network systems, like deep learning methods, to lower the misinterpretation rate [93]. However, this study does demonstrate the effectiveness, and potential of, short-term forage crop management monitoring by UAV to aid decision making, and presents a foundation on which to develop new possibilities for larger-scale UAV field-based phenotyping platforms to accelerate the crop breeding process.

Author Contributions: All authors contributed significantly to the work presented in this paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the European Regional Development Fund within the Estonian National Programme for Addressing Socio-Economic Challenges through R&D (RITA): L180283PKKK and the Doctoral School of Earth Sciences and Ecology, financed by the European Union, European Regional Development Fund (Estonian University of Life Sciences ASTRA project "Value-chain based bio-economy").

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the data size.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Annicchiarico, P.; Barrett, B.; Brummer, E.C.; Julier, B.; Marshall, A.H. Achievements and Challenges in Improving Temperate Perennial Forage Legumes. *Crit. Rev. Plant Sci.* **2015**, *34*, 327–380. [CrossRef]
2. Bender, A.; Tamm, S. Seed yield of tetraploid red clover as influenced by cover crop management. *Zemdirb. Agric.* **2018**, *105*, 133–140. [CrossRef]
3. Thilakarathna, M.S.; McElroy, M.S.; Chapagain, T.; Papadopoulos, Y.A.; Raizada, M.N. Belowground nitrogen transfer from legumes to non-legumes under managed herbaceous cropping systems. A review. *Agron. Sustain. Dev.* **2016**, *36*, 58. [CrossRef]
4. Vleugels, T.; Amdahl, H.; Roldán-Ruiz, I.; Cnops, G. Factors Underlying Seed Yield in Red Clover: Review of Current Knowledge and Perspectives. *Agronomy* **2019**, *9*, 829. [CrossRef]
5. Arturi, M.J.; Aulicino, M.B.; Ansín, O.; Gallinger, G.; Signorio, R. Combining Ability in Mixtures of Prairie Grass and Clovers. *Am. J. Plant Sci.* **2012**, *3*, 1355–1360. [CrossRef]
6. Zarza, R.; Rebuffo, M.; La Manna, A.; Balzarini, M. Red clover (Trifolium pratense L.) seedling density in mixed pastures as predictor of annual yield. *Field Crop. Res.* **2020**, *256*, 107925. [CrossRef]
7. Doyle, C.; Topp, C. The economic opportunities for increasing the use of forage legumes in north European livestock systems under both conventional and organic management. *Renew. Agric. Food Syst.* **2004**, *19*, 15–22. [CrossRef]
8. Hanson, J.; Ellis, R.H. Progress and Challenges in Ex Situ Conservation of Forage Germplasm: Grasses, Herbaceous Legumes and Fodder Trees. *Plants* **2020**, *9*, 446. [CrossRef]
9. Carvell, C.; Roy, D.B.; Smart, S.M.; Pywell, R.F.; Preston, C.D.; Goulson, D. Declines in forage availability for bumblebees at a national scale. *Biol. Conserv.* **2006**, *132*, 481–489. [CrossRef]
10. Yang, X.; Drury, C.; Reynolds, W.; Reeb, M. Legume Cover Crops Provide Nitrogen to Corn During a Three-Year Transition to Organic Cropping. *Agron. J.* **2019**, *111*, 3253–3264. [CrossRef]
11. Yang, Y.; Reilly, E.C.; Jungers, J.M.; Chen, J.; Smith, T.M. Climate Benefits of Increasing Plant Diversity in Perennial Bioenergy Crops. *One Earth* **2019**, *1*, 434–445. [CrossRef]
12. Godfray, H.C.J.; Beddington, J.R.; Crute, I.R.; Haddad, L.; Lawrence, D.; Muir, J.F.; Pretty, J.; Robinson, S.; Thomas, S.M.; Toulmin, C. Food Security: The Challenge of Feeding 9 Billion People. *Science* **2010**, *327*, 812–818. [CrossRef] [PubMed]
13. Lüscher, A.; Mueller-Harvey, I.; Soussana, J.-F.; Rees, R.M.; Peyraud, J.-L. Potential of legume-based grassland–livestock systems in Europe: A review. *Grass Forage Sci.* **2014**, *69*, 206–228. [CrossRef] [PubMed]
14. Boelt, B.; Julier, B.; Karagić, Đ.; Hampton, J. Legume Seed Production Meeting Market Requirements and Economic Impacts. *Crit. Rev. Plant Sci.* **2014**, *34*, 412–427. [CrossRef]
15. Wachendorf, M.; Fricke, T.; Möckel, T. Remote sensing as a tool to assess botanical composition, structure, quantity and quality of temperate grasslands. *Grass Forage Sci.* **2018**, *73*, 1–14. [CrossRef]
16. Araus, J.L.; Cairns, J.E. Field high-throughput phenotyping: The new crop breeding frontier. *Trends Plant Sci.* **2014**, *19*, 52–61. [CrossRef]

17. Yang, G.; Liu, J.; Zhao, C.; Li, Z.; Huang, Y.; Yu, H.; Xu, B.; Yang, X.; Zhu, D.; Zhang, X.; et al. Unmanned Aerial Vehicle Remote Sensing for Field-Based Crop Phenotyping: Current Status and Perspectives. *Front. Plant Sci.* **2017**, *8*, 1111. [CrossRef]
18. Laidig, F.; Piepho, H.-P.; Drobek, T.; Meyer, U. Genetic and non-genetic long-term trends of 12 different crops in German official variety performance trials and on-farm yield trends. *Theor. Appl. Genet.* **2014**, *127*, 2599–2617. [CrossRef]
19. Lollato, R.P.; Roozeboom, K.; Lingenfelser, J.F.; Da Silva, C.L.; Sassenrath, G. Soft winter wheat outyields hard winter wheat in a subhumid environment: Weather drivers, yield plasticity, and rates of yield gain. *Crop. Sci.* **2020**, *60*, 1617–1633. [CrossRef]
20. Zhu-Barker, X.; Steenwerth, K.L. Nitrous Oxide Production from Soils in the Future. *Chem. Bioavailab. Terr. Environ.* **2018**, *35*, 131–183. [CrossRef]
21. Van Ittersum, M.K.; Cassman, K.G.; Grassini, P.; Wolf, J.; Tittonell, P.; Hochman, Z. Yield gap analysis with local to global relevance—A review. *Field Crop. Res.* **2013**, *143*, 4–17. [CrossRef]
22. Grüner, E.; Wachendorf, M.; Astor, T. The potential of UAV-borne spectral and textural information for predicting aboveground biomass and N fixation in legume-grass mixtures. *PLoS ONE* **2020**, *15*, e0234703. [CrossRef] [PubMed]
23. Costa, J.M.; Da Silva, J.M.; Pinheiro, C.; Barón, M.; Mylona, P.; Centritto, M.; Haworth, M.; Loreto, F.; Uzilday, B.; Turkan, I.; et al. Opportunities and Limitations of Crop Phenotyping in Southern European Countries. *Front. Plant Sci.* **2019**, *10*, 1125. [CrossRef] [PubMed]
24. Pieruschka, R.; Schurr, U. Plant Phenotyping: Past, Present, and Future. *Plant Phenomics* **2019**, *2019*, 1–6. [CrossRef]
25. Mozgeris, G.; Jonikavičius, D.; Jovarauskas, D.; Zinkevičius, R.; Petkevičius, S.; Steponavičius, D. Imaging from manned ultra-light and unmanned aerial vehicles for estimating properties of spring wheat. *Precis. Agric.* **2018**, *19*, 876–894. [CrossRef]
26. Yeom, J.; Jung, J.; Chang, A.; Ashapure, A.; Maeda, M.; Maeda, A.; Landivar, J. Comparison of Vegetation Indices Derived from UAV Data for Differentiation of Tillage Effects in Agriculture. *Remote Sens.* **2019**, *11*, 1548. [CrossRef]
27. Zhang, C.; Kovacs, J.M. The application of small unmanned aerial systems for precision agriculture: A review. *Precis. Agric.* **2012**, *13*, 693–712. [CrossRef]
28. Rouse, J.W.; Hass, R.H.; Schell, J.A.; Deering, D.W.; Harlan, J.C. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*; Final Report, RSC 1978-4; Texas A&M University: College Station, TX, USA, 1974.
29. Xue, J.; Su, B. Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. *J. Sens.* **2017**, *2017*, 1–17. [CrossRef]
30. Chao, Z.; Liu, N.; Zhang, P.; Ying, T.; Song, K. Estimation methods developing with remote sensing information for energy crop biomass: A comparative review. *Biomass- Bioenergy* **2019**, *122*, 414–425. [CrossRef]
31. Osco, L.P.; Ramos, A.P.M.; Pereira, D.R.; Moriya Érika, A.S.; Imai, N.N.; Matsubara, E.T.; Estrabis, N.; De Souza, M.; Junior, J.M.; Gonçalves, W.N.; et al. Predicting Canopy Nitrogen Content in Citrus-Trees Using Random Forest Algorithm Associated to Spectral Vegetation Indices from UAV-Imagery. *Remote. Sens.* **2019**, *11*, 2925. [CrossRef]
32. Lussem, U.; Bolten, A.; Gnyp, M.L.; Jasper, J.; Bareth, G. EVALUATION OF RGB-BASED VEGETATION INDICES FROM UAV IMAGERY TO ESTIMATE FORAGE YIELD IN GRASSLAND. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *XLII-3*, 1215–1219. [CrossRef]
33. Grüner, E.; Astor, T.; Wachendorf, M. Biomass Prediction of Heterogeneous Temperate Grasslands Using an SfM Approach Based on UAV Imaging. *Agronomy* **2019**, *9*, 54. [CrossRef]
34. Viljanen, N.; Honkavaara, E.; Näsi, R.; Hakala, T.; Niemeläinen, O.; Kaivosoja, J. A Novel Machine Learning Method for Estimating Biomass of Grass Swards Using a Photogrammetric Canopy Height Model, Images and Vegetation Indices Captured by a Drone. *Agriculture* **2018**, *8*, 70. [CrossRef]
35. Kaspersen, K.; Bakken, A.K. Evaluating an image analysis system for mapping white clover pastures. *Acta Agric. Scand. Sect. B Plant Soil Sci.* **2004**, *54*, 76–82. [CrossRef]
36. Abuleil, A.M.; Taylor, G.W.; Moussa, M. An Integrated System for Mapping Red Clover Ground Cover Using Unmanned Aerial Vehicles: A Case Study in Precision Agriculture. In Proceedings of the 2015 12th Conference on Computer and Robot Vision, Halifax, NS, Canada, 3–5 June 2015; pp. 277–284.
37. Oliveira, R.A.; Näsi, R.; Niemeläinen, O.; Nyholm, L.; Alhonoja, K.; Kaivosoja, J.; Jauhiainen, L.; Viljanen, N.; Nezami, S.; Markelin, L.; et al. Machine learning estimators for the quantity and quality of grass swards used for silage production using drone-based imaging spectrometry and photogrammetry. *Remote Sens. Environ.* **2020**, *246*, 111830. [CrossRef]
38. FAO. World Reference Base for Soil Resources. In *World Soil Resources Report 103*; FAO: Rome, Italy, 2006.
39. Poncet, A.M.; Knappenberger, T.; Brodbeck, C.; Fogle, J.M.; Shaw, J.N.; Ortiz, B.V. Multispectral UAS Data Accuracy for Different Radiometric Calibration Methods. *Remote Sens.* **2019**, *11*, 1917. [CrossRef]
40. senseFly. eBee: senseFly SA. Available online: https://www.sensefly.com/drones/ (accessed on 12 October 2019).
41. Metsar, J.; Kollo, K.; Ellmann, A. Modernization of the estonian national GNSS reference station network. *Geod. Cartogr.* **2018**, *44*, 55–62. [CrossRef]
42. Tomaštík, J.; Mokroš, M.; Surový, P.; Grznárová, A.; Merganič, J. UAV RTK/PPK Method—An Optimal Solution for Mapping Inaccessible Forested Areas? *Remote Sens.* **2019**, *11*, 721. [CrossRef]
43. Villoslada, M.; Bergamo, T.; Ward, R.; Burnside, N.; Joyce, C.; Bunce, R.; Sepp, K. Fine scale plant community assessment in coastal meadows using UAV based multispectral data. *Ecol. Indic.* **2020**, *111*, 105979. [CrossRef]
44. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020.

45. Hassan, M.A.; Yang, M.; Rasheed, A.; Yang, G.; Reynolds, M.; Xia, X.; Xiao, Y.; He, Z. A rapid monitoring of NDVI across the wheat growth cycle for grain yield prediction using a multi-spectral UAV platform. *Plant Sci.* **2019**, *282*, 95–103. [CrossRef]
46. Gitelson, A.A.; Kaufman, Y.J.; Stark, R.; Rundquist, D. Novel algorithms for remote estimation of vegetation fraction. *Remote Sens. Environ.* **2002**, *80*, 76–87. [CrossRef]
47. Tilly, N.; Hoffmeister, D.; Cao, Q.; Huang, S.; Lenz-Wiedemann, V.; Miao, Y.; Bareth, G. Multitemporal crop surface models: Accurate plant height measurement and biomass estimation with terrestrial laser scanning in paddy rice. *J. Appl. Remote Sens.* **2014**, *8*, 083671. [CrossRef]
48. Hunt, E.R.; Hively, W.D.; Daughtry, C.S.T.; Mccarty, G.W.; Fujikawa, S.J.; Ng, T.L.; Tranchitella, M.; Linden, D.S.; Yoel, D.W. Remote Sensing of Crop Leaf Area Index Using Unmanned Airborne Vehicles. In Proceedings of the Pecora 17 Symposium, Denver, CO, USA, 18 November 2008.
49. Serrano, L.; Filella, I.; Penuelas, J. Remote Sensing of Biomass and Yield of Winter Wheat under Different Nitrogen Supplies. *Crop. Sci.* **2000**, *40*, 723–731. [CrossRef]
50. Walsh, O.S.; Shafian, S.; Marshall, J.M.; Jackson, C.; McClintick-Chess, J.R.; Blanscet, S.M.; Swoboda, K.; Thompson, C.; Belmont, K.M.; Walsh, W.L. Assessment of UAV Based Vegetation Indices for Nitrogen Concentration Estimation in Spring Wheat. *Adv. Remote Sens.* **2018**, *7*, 71–90. [CrossRef]
51. Chen, J.M. Evaluation of Vegetation Indices and a Modified Simple Ratio for Boreal Applications. *Can. J. Remote Sens.* **1996**, *22*, 229–242. [CrossRef]
52. Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* **1996**, *58*, 289–298. [CrossRef]
53. Sripada, R.P.; Heiniger, R.W.; White, J.G.; Meijer, A.D. Aerial Color Infrared Photography for Determining Early In-Season Nitrogen Requirements in Corn. *Agron. J.* **2006**, *98*, 968–977. [CrossRef]
54. Jordan, C.F. Derivation of Leaf-Area Index from Quality of Light on the Forest Floor. *Ecology* **1969**, *50*, 663–666. [CrossRef]
55. Gitelson, A.; Merzlyak, M.N. Spectral Reflectance Changes Associated with Autumn Senescence of Aesculus hippocastanum L. and Acer platanoides L. Leaves. Spectral Features and Relation to Chlorophyll Estimation. *J. Plant Physiol.* **1994**, *143*, 286–292. [CrossRef]
56. ESRI. *ArcGIS PRO: Essential Workflows*; ESRI: Redlands, CA, USA, 2016; Available online: https://community.esri.com/t5/esri-training-documents/arcgis-pro-essential-workflows-course-resources/ta-p/914710 (accessed on 1 June 2020).
57. Zheng, H.; Cheng, T.; Zhou, M.; Li, D.; Yao, X.; Tian, Y.; Cao, W.; Zhu, Y. Improved estimation of rice aboveground biomass combining textural and spectral analysis of UAV imagery. *Precis. Agric.* **2018**, *20*, 611–629. [CrossRef]
58. Poley, L.G.; McDermid, G.J. A Systematic Review of the Factors Influencing the Estimation of Vegetation Aboveground Biomass Using Unmanned Aerial Systems. *Remote Sens.* **2020**, *12*, 1052. [CrossRef]
59. Ma, L.; Li, M.; Ma, X.; Cheng, L.; Du, P.; Liu, Y. A review of supervised object-based land-cover image classification. *ISPRS J. Photogramm. Remote Sens.* **2017**, *130*, 277–293. [CrossRef]
60. Pilgrim, M. *Dive into Python 3*; Apress: New York, NY, USA, 2009.
61. Kearns, M.; Ron, D. Algorithmic Stability and Sanity-Check Bounds for Leave-One-Out Cross-Validation. *Neural Comput.* **1999**, *11*, 1427–1453. [CrossRef] [PubMed]
62. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
63. Liu, T.; Abd-Elrahman, A.; Morton, J.; Wilhelm, V.L. Comparing fully convolutional networks, random forest, support vector machine, and patch-based deep convolutional neural networks for object-based wetland mapping using images from small unmanned aircraft system. *GIScience Remote Sens.* **2018**, *55*, 243–264. [CrossRef]
64. Shin, J.; Kim, H.J.; Park, S.; Kim, Y. Model predictive flight control using adaptive support vector regression. *Neurocomputing* **2010**, *73*, 1031–1037. [CrossRef]
65. He, Q. *Neural Network and Its Application in IR*; Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign Spring: Champaign, IL, USA, 1999.
66. Yue, J.; Yang, G.; Li, C.; Li, Z.; Wang, Y.; Feng, H.; Xu, B. Estimation of Winter Wheat Above-Ground Biomass Using Unmanned Aerial Vehicle-Based Snapshot Hyperspectral Sensor and Crop Height Improved Models. *Remote Sens.* **2017**, *9*, 708. [CrossRef]
67. Li, B.; Xu, X.; Zhang, L.; Han, J.; Bian, C.; Li, G.; Liu, J.; Jin, L. Above-ground biomass estimation and yield prediction in potato by using UAV-based RGB and hyperspectral imaging. *ISPRS J. Photogramm. Remote Sens.* **2020**, *162*, 161–172. [CrossRef]
68. Wang, J.; Badenhorst, P.; Phelan, A.; Pembleton, L.; Shi, F.; Cogan, N.; Spangenberg, G.; Smith, K. Using Sensors and Unmanned Aircraft Systems for High-Throughput Phenotyping of Biomass in Perennial Ryegrass Breeding Trials. *Front. Plant Sci.* **2019**, *10*, 1381. [CrossRef] [PubMed]
69. Chaudhry, M.H.; Ahmad, A.; Gulzar, Q. A comparative study of modern UAV platform for topographic mapping. In Proceedings of the 10th IGRSM International Conference and Exhibition on Geospatial & Remote Sensing, Kuala Lumpur, Malaysia, 20–21 October 2020; Volume 540. [CrossRef]
70. Roth, L.; Streit, B. Predicting cover crop biomass by lightweight UAS-based RGB and NIR photography: An applied photogrammetric approach. *Precis. Agric.* **2017**, *19*, 93–114. [CrossRef]
71. Biewer, S.; Erasmi, S.; Fricke, T.; Wachendorf, M. Prediction of yield and the contribution of legumes in legume-grass mixtures using field spectrometry. *Precis. Agric.* **2009**, *10*, 128–144. [CrossRef]

72. Rueda-Ayala, V.P.; Peña, J.M.; Höglind, M.; Bengochea-Guevara, J.M.; Andújar, D. Comparing UAV-Based Technologies and RGB-D Reconstruction Methods for Plant Height and Biomass Monitoring on Grass Ley. *Sensors* **2019**, *19*, 535. [CrossRef]
73. Michez, A.; Bauwens, S.; Brostaux, Y.; Hiel, M.-P.; Garré, S.; Lejeune, P.; Dumont, B. How Far Can Consumer-Grade UAV RGB Imagery Describe Crop Production? A 3D and Multitemporal Modeling Approach Applied to Zea mays. *Remote Sens.* **2018**, *10*, 1798. [CrossRef]
74. Zhou, X.; Zheng, H.; Xu, X.; He, J.; Ge, X.; Yao, X.; Cheng, T.; Zhu, Y.; Cao, W.; Tian, Y. Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery. *ISPRS J. Photogramm. Remote Sens.* **2017**, *130*, 246–255. [CrossRef]
75. Fanigliulo, R.; Antonucci, F.; Figorilli, S.; Pochi, D.; Pallottino, F.; Fornaciari, L.; Grilli, R.; Costa, C. Light Drone-Based Application to Assess Soil Tillage Quality Parameters. *Sensors* **2020**, *20*, 728. [CrossRef] [PubMed]
76. Fanigliulo, R.; Biocca, M.; Pochi, D. Effects of six primary tillage implements on energy inputs and residue cover in Central Italy. *J. Agric. Eng.* **2016**, *47*, 177–180. [CrossRef]
77. Ashapure, A.; Jung, J.; Yeom, J.; Chang, A.; Maeda, M.; Maeda, A.; Landivar, J. A novel framework to detect conventional tillage and no-tillage cropping system effect on cotton growth and development using multi-temporal UAS data. *ISPRS J. Photogramm. Remote Sens.* **2019**, *152*, 49–64. [CrossRef]
78. Lussem, U.; Bolten, A.; Menne, J.; Gnyp, M.L.; Bareth, G. Ultra-high spatial resolution uav-based imagery to predict biomass in temperate grasslands. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *XLII-2/W13*, 443–447. [CrossRef]
79. Wittwer, R.A.; van der Heijden, M.G. Cover crops as a tool to reduce reliance on intensive tillage and nitrogen fertilization in conventional arable cropping systems. *Field Crop. Res.* **2020**, *249*, 107736. [CrossRef]
80. Ali, I.; Greifeneder, F.; Stamenkovic, J.; Neumann, M.; Notarnicola, C. Review of Machine Learning Approaches for Biomass and Soil Moisture Retrievals from Remote Sensing Data. *Remote Sens.* **2015**, *7*, 16398–16421. [CrossRef]
81. Xie, Y.; Sha, Z.; Yu, M.; Bai, Y.; Zhang, L. A comparison of two models with Landsat data for estimating above ground grassland biomass in Inner Mongolia, China. *Ecol. Model.* **2009**, *220*, 1810–1818. [CrossRef]
82. Ali, I.; Cawkwell, F.; Dwyer, E.; Green, S. Modeling Managed Grassland Biomass Estimation by Using Multitemporal Remote Sensing Data—A Machine Learning Approach. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *10*, 3254–3264. [CrossRef]
83. Mutanga, O.; Skidmore, A. Integrating imaging spectroscopy and neural networks to map grass quality in the Kruger National Park, South Africa. *Remote Sens. Environ.* **2004**, *90*, 104–115. [CrossRef]
84. Mutanga, O.; Kumar, L. Estimating and mapping grass phosphorus concentration in an African savanna using hyperspectral image data. *Int. J. Remote Sens.* **2007**, *28*, 4897–4911. [CrossRef]
85. Zhang, Z.; Masjedi, A.; Zhao, J.; Crawford, M.M. Prediction of sorghum biomass based on image based features derived from time series of UAV images. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 6154–6157. [CrossRef]
86. Théau, J.; Lauzier-Hudon, Étienne; Aubé, L.; Devillers, N. Estimation of forage biomass and vegetation cover in grasslands using UAV imagery. *PLoS ONE* **2021**, *16*, e0245784. [CrossRef]
87. Ramos, A.P.M.; Osco, L.P.; Furuya, D.E.G.; Gonçalves, W.N.; Santana, D.C.; Teodoro, L.P.R.; Junior, C.A.D.S.; Capristo-Silva, G.F.; Li, J.; Baio, F.H.R.; et al. A random forest ranking approach to predict yield in maize with uav-based vegetation spectral indices. *Comput. Electron. Agric.* **2020**, *178*, 105791. [CrossRef]
88. Insua, J.R.; Utsumi, S.A.; Basso, B. Estimation of spatial and temporal variability of pasture growth and digestibility in grazing rotations coupling unmanned aerial vehicle (UAV) with crop simulation models. *PLoS ONE* **2019**, *14*, e0212773. [CrossRef]
89. Lee, H.; Lee, H.-J.; Jung, J.-S.; Ko, H.-J. Mapping Herbage Biomass on a Hill Pasture using a Digital Camera with an Unmanned Aerial Vehicle System. *J. Korean Soc. Grassl. Forage Sci.* **2015**, *35*, 225–231. [CrossRef]
90. Ma, Q.; Chai, L.; Hou, F.; Chang, S.; Ma, Y.; Tsunekawa, A.; Cheng, Y. Quantifying Grazing Intensity Using Remote Sensing in Alpine Meadows on Qinghai-Tibetan Plateau. *Sustainability* **2019**, *11*, 417. [CrossRef]
91. Luan, J.; Zhang, C.; Xu, B.; Xue, Y.; Ren, Y. The predictive performances of random forest models with limited sample size and different species traits. *Fish. Res.* **2020**, *227*, 105534. [CrossRef]
92. Yang, M.-D.; Huang, K.-S.; Kuo, Y.-H.; Tsai, H.P.; Lin, L.-M. Spatial and Spectral Hybrid Image Classification for Rice Lodging Assessment through UAV Imagery. *Remote Sens.* **2017**, *9*, 583. [CrossRef]
93. Yang, M.-D.; Tseng, H.-H.; Hsu, Y.-C.; Tsai, H.P. Semantic Segmentation Using Deep Learning with Vegetation Indices for Rice Lodging Identification in Multi-date UAV Visible Images. *Remote Sens.* **2020**, *12*, 633. [CrossRef]

II

Li, K. Y., Burnside, N. G., de Lima, R. S., Villoslada, M., Sepp, K., Yang, M. D., Raet, J. Vain, A., Selge, A. & Sepp, K. (2021). The Application of an Unmanned Aerial System and Machine Learning Techniques for Red Clover-Grass Mixture Yield Estimation under Variety Performance Trials. Remote Sensing, 13(10), 1994.

remote sensing

MDPI

Article

An Automated Machine Learning Framework in Unmanned Aircraft Systems: New Insights into Agricultural Management Practices Recognition Approaches

Kai-Yun Li [1,*], Niall G. Burnside [2], Raul Sampaio de Lima [1], Miguel Villoslada Peciña [1], Karli Sepp [3], Victor Henrique Cabral Pinheiro [4], Bruno Rucy Carneiro Alves de Lima [4], Ming-Der Yang [5], Ants Vain [1] and Kalev Sepp [1]

1 Institute of Agriculture and Environmental Sciences, Estonian University of Life Sciences, Kreutzwaldi 5, 51006 Tartu, Estonia; raul.sampaio@student.emu.ee (R.S.d.L.); Miguel.Pecina@emu.ee (M.V.P.); a.vain@ts.ee (A.V.); kalev.sepp@emu.ee (K.S.)
2 School of Environment & Technology, University of Brighton, Lewes Road, Brighton BN2 4JG, UK; N.G.Burnside@brighton.ac.uk
3 Agricultural Research Center, 4/6 Teaduse St., 75501 Saku, Estonia; karli.sepp@pmk.agri.ee
4 Institute of Computer Science, Faculty of Science and Technology, University of Tartu, 50090 Tartu, Estonia; victor.pinheiro@ut.ee (V.H.C.P.); bruno.rucy.carneiro.alves.de.lima@ut.ee (B.R.C.A.d.L.)
5 Department of Civil Engineering, Innovation and Development Center of Sustainable Agriculture, National Chung Hsing University, Taichung 402, Taiwan; mdyang@dragon.nchu.edu.tw
* Correspondence: kai-yun.li@student.emu.ee

check for updates

Citation: Li, K.-Y.; Burnside, N.G.; de Lima, R.S.; Peciña, M.V.; Sepp, K.; Cabral Pinheiro, V.H.; de Lima, B.R.C.A.; Yang, M.-D.; Vain, A.; Sepp, K. An Automated Machine Learning Framework in Unmanned Aircraft Systems: New Insights into Agricultural Management Practices Recognition Approaches. *Remote Sens.* **2021**, *13*, 3190. https://doi.org/10.3390/rs13163190

Academic Editor: Emanuel Peres

Received: 12 July 2021
Accepted: 9 August 2021
Published: 12 August 2021

Abstract: The recent trend of automated machine learning (AutoML) has been driving further significant technological innovation in the application of artificial intelligence from its automated algorithm selection and hyperparameter optimization of the deployable pipeline model for unravveling substance problems. However, a current knowledge gap lies in the integration of AutoML technology and unmanned aircraft systems (UAS) within image-based data classification tasks. Therefore, we employed a state-of-the-art (SOTA) and completely open-source AutoML framework, Auto-sklearn, which was constructed based on one of the most widely used ML systems: Scikit-learn. It was combined with two novel AutoML visualization tools to focus particularly on the recognition and adoption of UAS-derived multispectral vegetation indices (VI) data across a diverse range of agricultural management practices (AMP). These include soil tillage methods (STM), cultivation methods (CM), and manure application (MA), and are under the four-crop combination fields (i.e., red clover-grass mixture, spring wheat, pea-oat mixture, and spring barley). Furthermore, they have currently not been efficiently examined and accessible parameters in UAS applications are absent for them. We conducted the comparison of AutoML performance using three other common machine learning classifiers, namely Random Forest (RF), support vector machine (SVM), and artificial neural network (ANN). The results showed AutoML achieved the highest overall classification accuracy numbers after 1200 s of calculation. RF yielded the second-best classification accuracy, and SVM and ANN were revealed to be less capable among some of the given datasets. Regarding the classification of AMPs, the best recognized period for data capture occurred in the crop vegetative growth stage (in May). The results demonstrated that CM yielded the best performance in terms of classification, followed by MA and STM. Our framework presents new insights into plant–environment interactions with capable classification capabilities. It further illustrated the automatic system would become an important tool in furthering the understanding for future sustainable smart farming and field-based crop phenotyping research across a diverse range of agricultural environmental assessment and management applications.

Keywords: unmanned aircraft system; automated machine learning; agricultural management practices; image classification; precision agriculture; variety performance trials; crop breeding; crop phenotyping; agriculture decision-making

1. Introduction

Unmanned Aerial Systems (UAS) are considered one of the most significant technologies for the further development of precision agriculture (PA) [1] and sustainable smart farming [2]. UAS are frequently employed for the surveillance of cultivated lands, providing effective solutions for accurate decision support, increasing farming efficiency, enhancing profitability, reducing environmental impacts, and driving further technological innovation [1,3,4]. UAS equipped with various novel sensor types can be exploited to improve agreement and synergy between imagery and field reference data. In addition, these systems can also identify the regional monitoring requirements, such as disease detection, growth observation, yield estimation, and weed management [5,6]. In PA, vegetation indices (VI) are one of the most widely used outputs from UAS imagery applications and assist in the delivery of dependable spatial and temporal information across multiple agricultural activities. VIs typically constitute mathematical combinations of individual or groups of bands from the electromagnetic spectrum and are intended to minimize the effect of external confounding factors while enhancing the detectability of vegetation characteristics [5,7]. Currently, UAS-based remote sensing techniques offer a notable contribution in field-based crop phenotyping investigations [8]. Immediate and accurate acquisition of crop phenotypic information in various agri-environments supports the exploration of genetic–environmental interactions from critical production traits to determine the inheritance information and expression patterns to increase crop yields and tolerance to abiotic/biotic stresses [9,10]. However, it is crucial to take into account that field conditions are notoriously diverse compared to experimental environments, such as greenhouses or laboratories. Moreover, the outputs and findings collected from controlled environments can be difficult to extrapolate onto field settings and can impair the interpretation and application of research schemes [10].

Therefore, a common approach when identifying multiple crop management procedures and their interaction with the environment involves a well-conducted randomized experimental design, in which different agricultural management practices (AMP) are imposed on crops [11]. Variety performance trials (VPT) are a valuable method to address this issue. VPTs are regularly implemented in AMP research activities to improve the understanding of diverse systems and develop environmental management recommendations for variety selection [12,13]. Concerning the AMPs trial criteria chosen and the recent growth in environmental protection awareness under the concepts of sustainable agriculture, the flexibility of environmentally friendly cultivation methods, such as reduced tillage and the application of various minerals and organic fertilizers, are being developed [14]. For example, tillage reduction is an essential characteristic of agricultural management that changes the soil either physically, chemically, mechanically, or biologically to create the appropriate conditions for seedling sprouting and healthy plant growth [15,16], whereas organic additions, such as manure or organic fertilizers, are widely used methods to enhance soil fertility [17]. Studying VPT datasets, however, provides unique analysis problems due to the structure, nature, and husbandry variations of each trial. The evaluation of differences in management practices could potentially be confounded due to their nested structure (e.g., as opposed to controlled replicated treatments) [18]. These AMPs have been increasingly proposed as an ecological method involving nutrient management, increased water holding capacity, and recoupled C and N cycling in agricultural ecosystems to improve sustainability [19,20]. Although the specification of weather, soil, and management practices in current cropping systems are vital for robust model simulation and evaluation, these data are usually inaccessible for most cropping systems with adequate geospatial detail and lack of ability to replicate measured yields of field crops that received the best possible AMPs across a broad range of environments [21]. Recently, the application of UAS combined with popular machine learning (ML) systems drives a significant contribution to VPT crop biomass estimation. These results deepen the possibility of applying machine learning technology to diverse and complex AMP farmland classification applications.

Incorporating multisensory computing science approaches provides a wide range of valuable information for the expansion of precision farming practices [22]. ML techniques may not provide a universal solution in precision farming; however, these approaches enable better determination in verisimilitude scenarios with minimum human intervention. They provide not only a powerful and flexible framework for decision-making but also facilitate the integration of expert knowledge into the PA system [23]. Complexities become a drawback in VPTs since desired models need to contain training and testing databases and are often restricted by the number of pure line seeds, various AMPs (fertility test, tillage category, disease resistance, etc.), and confined areas with small sampling sizes to compensate for the labor-intensive fieldwork. Likewise, environmental factor interventions enhance obstacles in parameter selection in ML systems owing to the differences in location, climate, and soil properties [24]. Occasionally, even the same crop genotypes may not express similar spectral characteristics in UAS, which renders the models invalid. If the reference parameters exist to formulate relationship functions, the genuine implementation results are frequently unsatisfactory owing to mismatches between concepts and realities.

As an alternative, the innovative concept of automated machine learning (AutoML) has arisen to reduce these data-driven costs while becoming a significant topic as the exponential growth of computing power continues [25]. AutoML is defined as a combination of algorithm selection and hyperparameter optimization, which aims to recognize the mixture of algorithm components with the best (cross-validated) performance by covering from raw datasets to the deployable pipeline ML model to unravel substance problems [26]. AutoML is built to decrease the time demands of data scientists and save time by empowering specialists to build ML applications automatically without requiring widespread knowledge of ML [27] and entails the automated construction of an ML pipeline based on limited computational constraints [28]. Recent advancements in AutoML systems, such as Auto-WEKA [29] and Auto-sklearn, [30] are recommended as an artificial intelligence-based solution for the expanding challenge of ML applications by combining a highly parametric ML framework with a Bayesian optimization method for a given dataset, significantly streamlining these steps for non-experts [30]. The standard procedure of ML modeling involves data pre-processing, feature engineering, feature extraction, feature selection, algorithm selection, and hyperparameter optimization to increase the model's predictive performance [31].

Although AutoML has promoted great achievements in computer science and recently UAS applications, for example, the approximation of root-zone soil moisture [32] by AutoML interface H2O AutoML [33] and RGB-based crop phenotyping [34] by neural architecture search system AutoKeras [35], it has not been widely applied in multispectral image analysis. A current gap persists in the knowledge base for multispectral-based AMP analysis and agriculture land use studies in addition to the further understanding the potential for remotely sensed solutions to field-based and multifunctional platforms for the demands of plant phenotyping and smart farming management. To solve this knowledge gap, this study employed a state-of-the-art (SOTA) and completely open-source AutoML system, Auto-sklearn [30], which is constructed based on one of the most widely used ML systems, Scikit-learn, in the scientific Python community [36], combined with two novel AutoML visualization tools to explore UAS-derived multispectral vegetation indices (VI) as an example for handling the AMPs classification tasks.

More precisely, the aims of this study were to (1) build an AutoML framework for UAS classification tasks; (2) explore the applicability of UAS sensors to recognize multiple AMP categories, namely soil tillage methods (STM), cultivation methods (CM), and manure application (MA), which have currently not been efficiently examined and are absent of accessible parameters in both UAS and ML fields; and (3) compare AutoML's ability using different ML classifiers to identify image-based AMPs for diverse crop categories and its appropriate growth stages. To our knowledge, this paper is the first study to use an AutoML system with UAS -derived multispectral VIs, for the agricultural classification task. Moreover, this paper is the first to provide a novel AutoML framework, across multiple

AMP activities, and present new insights into UAS and ML optimization methods for future PA and crop phenotyping research.

2. Materials and Methods

2.1. Study Area and Experiment Layout

This study commenced at the Agricultural Research Centre (ARC) in Kuusiku (58°58′52.7″N 24°42′59.1″E), Estonia (Figure 1a). The experimental area used in this study covered 226 hectares, of which the 2.87-hectares variety performance trial (VPT) area consists of two soil types: Calcaric Cambisol and Calcaric-Leptic Regosol [37]. The experimental layout consisted of four types of common crop and their regular combinations in Estonia, i.e., Field 1: red clover 75% (*Trifolium pratense* L.) with grass 25% (*Festuca pratensis*) (RC + G). Field 2: spring wheat (SW), Field 3: pea and oat mixture (P + O), and Field 4: spring barley with under-sowing red clover (SB + RC) in 2019 (Figure 1b). This experimental design was developed to facilitate the understanding of the physiological conditions and yield performance capabilities of the chosen varieties and their combinations under three types of AMPs. To assess the UAS-based AMP detection capacity, the experiment was put together with three principal experimental factors (Figure 1c), which included: (1) soil tillage methods (STM), considering reduced tillage (R) (8–10 cm), ploughing (P) at a depth traditionally used in conventional tillage (18–20 cm), and disking (DP) (8–10 cm) as treatments; (2) cultivation methods (CM), considering conventional farming with mineral fertilizer application (Cmin+), organic farming with mineral fertilizer application (Omin+), and organic farming without mineral fertilizer (Omin−); and (3) manure applications (MA). Each field comprised 72 plots, which amounted to a total of 288 plots sampled within our study area.

2.2. UAS Image Acquisition

Figure 2 shows the workflow utilized to combine the UAS-based image collection, processing, sampling, and AutoML framework modified from [30]. A fixed-wing UAS eBee Plus (Sensefly Inc., Cheseaux–Lausane, Switzerland) equipped with GNSS post-processed kinematic (PPK) capabilities was deployed with a Parrot Sequoia multispectral sensor (version 1.2.1, Parrot, Paris, France). This UAS platform and sensor were used for image acquisition and captured imagery across four spectral bands: green (530–570 nm), red (640–680 nm), red-edge (730–740 nm), and near-infrared (770–810 nm). To facilitate seasonal image processing and AMP recognition, UAS images were captured over three timeslots in 2019 at the Kuusiku Research Center: 23 April (temperature: 16 °C, wind speed: 11 km h^{-1} S, sunny), 30 May (temperature: 19 °C, wind speed: 12 km h^{-1} WSW, overcast), and 10 July (temperature 20 °C, wind speed: 3.6 km h^{-1} NW, sun with minor cloud cover). The weather conditions in the 6 days prior to the image acquisition are displayed in Supplementary Figure S1. The originally designed flight time was 37 min and 30 s per task over an area of 65.8 hectares (with areas of interest 2.87 hectares in this study). However, depending on the weather conditions and wind speed of the day, the eBee flight time might have been slightly different from the number of battery replacements (the endurance of one battery was approximately 20–30 min). This data capture protocol was designed to represent the reflectance spectrum characteristics of crops during different growth stages. Flight-line overlap was set using a frontal image overlap of 80% and a lateral overlap of 75% with a target altitude of 120 m above ground level (AGL), resulting in a ground sampling distance (GSD) of 10 cm per pixel. All image data capture procedures were undertaken between the hours of 10 a.m. to 2 p.m. to guarantee the consistency of photo collection quality and to minimize lateral shading of crops within the VPT fields. An Airinov radiometric calibration target (Airinov, Paris, France) and a one-point calibration method [38] were used to enable post-flight radiometric correction of the multispectral imagery before each flight to remove dark current and lens vignetting effects while post-processing the image [39].

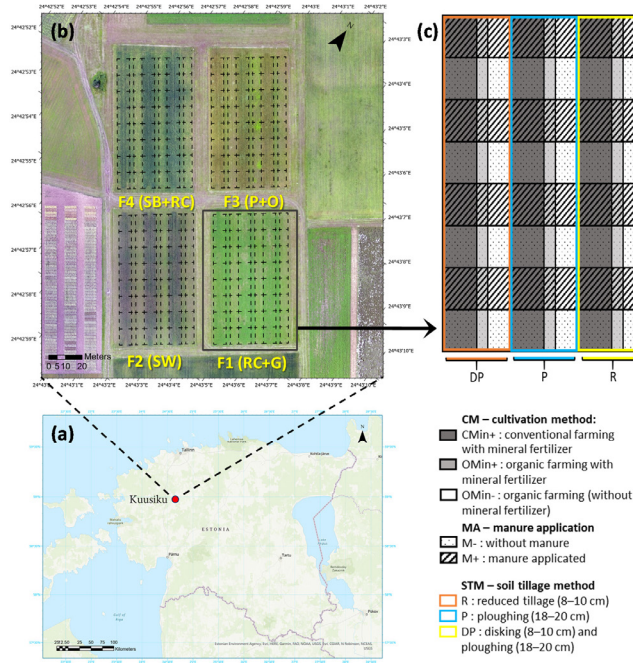

Figure 1. (**a**) The study area located at the Kuusiku agriculture center, Estonia. (**b**) The RGB orthomosaic image from 30 May of the experimental layout fields with four crop types, i.e., (F1. (RC + G), F2. (SW), Field 3. (PO), and Field 4. (SB + RC)) (**c**) The VPT with three agricultural management practices (AMP): cultivation method (CM) with three levels (Cmin+, NPK 5-10-25 291 kg ha^{-1} (N-14 kg ha^{-1}, P-13 kg ha^{-1}, and K-60 kg ha^{-1}); Cmin−, mineral fertilizer 240 kg ha^{-1} (K-60 kg ha^{-1}, S-41 kg ha^{-1}, M-14 kg ha^{-1}); and Omin−), manure application (MA) with two levels (M+, manure 30,000 kg ha^{-1} (N-234 kg ha^{-1}, P-20 kg ha^{-1}, and K-216 kg ha^{-1}), and M−), and soil tillage method (STM) with three levels (DP, P, and R) were conducted in this study.

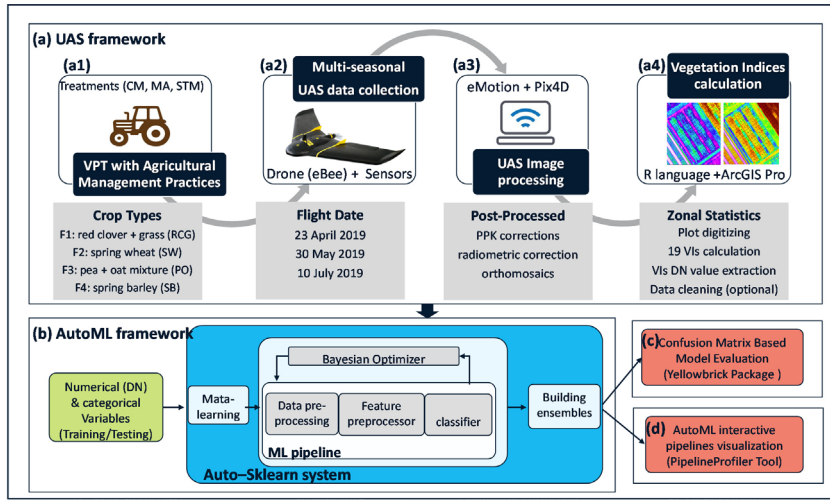

Figure 2. The flowchart of the UAS and AutoML framework in this study. (**a**) The UAS framework, where (**a1**) Three types of AMPs were processed for four crop categories. (**a2**) The eBee plus with Parrot Sequoia multispectral sensor with the time series flight (April, May, and July) to collect spectral information from different crop periods. (**a3**) UAS image post-processed in SenseFly eMotion with PPK corrections and orthomosaics in Pix4D. (**a4**) 19 VIs calculation, segmentation and corresponding plot digital number (DN) extraction for AutoML modeling. (**b**) The Auto-sklearn framework constructed ML pipelines automatically, which were proposed by the Bayesian optimization method with warm-started meta-learning and joint with post hoc ensemble building approach to achieve robust performance (adapted from [30,40]). (**c**) Yellowbrick visualization package was conducted for AutoML model evaluation. (**d**) PipelineProfiler was conducted for AutoML interactive pipelines visualization tool allows the examination of the solution space of end-to-end ML pipelines.

2.3. UAS Image Processing

For pre-processing UAS images, we used SenseFly eMotion 3, applying differential correction data (RINEX) provided by the GNSS CORS (Continuously Operating Reference Station) of Estonia for post-processing kinematics (PPK) corrections [41]. PPK was reported to increase the higher horizontal and vertical geotagging accuracy when compared to ground control points (GCP) [42,43]. In our study, the UAS image corrections were decreased from 5 m error to under 0.06 m (less than one-pixel size). Pix4D v.4.3.31® (Pix4D SA, 1015 Lausanne, Switzerland) software was utilized to process and radiometrically correct (calibrated according to the variances between the measured value and target actual reflectance [38]) the imagery, as well as to generate the multispectral orthomosaics. These images were subsequently clipped with a one-meter inward buffer zone from each plot to represent only the extent of the area of the VPTs.

2.4. Vegetation Indices Calculation

In this study, nineteen VIs were chosen and calculated to address the issues of heterogeneous crop classes, soil types, and the current absence of valuable UAS referenced parameters in AMPs (see Table 1). More specifically, Datt4, SRre, NDVIre were selected due to their positive correlation with chlorophyll content [44–46]; MTVI, MSR, MSRre, RVIS, WDRVI [47–51] are known to be sensitive to variations in leaf area index (LAI); GDVI was used for better lower vegetal land cover estimates and characterization [52]; GIPVI

was calculated for its potential in grassland communities detection [53]; GNDVI, NDVI, RTVIcore were utilized due to their high performance in crop above-ground biomass (AGB) estimation [54,55]; GDI, GRDI, and RDVI were included due to their ability to compensate for NDVI saturation problems, and the potential effects of soil and sun viewing geometry [56,57]; GRVI was applied for its sensitivity to soil moisture [58], SR for strongly correlated with comprehensive growth index (CGI) [59] and REGVI was included for its sensitivity to deviations in senescence and vegetation stress [60].

Table 1. Descriptions and formulas of multispectral UAS derived VIs used in this study. The ρ R refers to the reflectance of the red band, ρ G refers to the reflectance of the green band, ρ REG refers to the reflectance of the red edge, and ρ NIR refers to the reflectance of the near-infrared.

Vegetation Index	Equation	Reference
Datt4	ρ R/(ρ G * ρ REG)	[61]
Green Infrared Percentage Vegetation Index (GIPVI)	ρ NIR/(ρ NIR + ρ G)	[62]
Green Normalized Difference Vegetation Index (GNDVI)	(ρ NIR − ρ G)/(NIR + ρ G)	[63]
Green Difference Vegetation Index (GDVI)	ρ NIR − ρ G	[64]
Green Ration Vegetation Index (GRVI)	ρ NIR/ρ G	[64]
Green Difference Index (GDI)	ρ NIR − ρ R + ρ G	[65]
Green Red Difference Index (GRDI)	(ρ G − ρ R)/(ρ G + ρ R)	[65]
Normalized Difference Vegetation Index (NDVI)	(ρ NIR − ρ R)/(ρ NIR + ρ R)	[66]
Red-Edge Normalized Difference Vegetation Index (NDVIre)	(ρ NIR − ρ REG)/(ρ NIR + ρ REG)	[46]
Red-Edge Simple Ratio (SRre)	ρ NIR/ρ REG	[46]
Renormalized Difference Vegetation Index (RDVI)	((ρ NIR − ρ R)/((ρ NIR + ρ R) ** (0.5)))	[67]
Red-Edge Modified Simple Ratio (MSRre)	((ρ NIR − ρ REG) − 1)/(((ρ NIR + ρ REG) ** (0.5)) + 1)	[49]
Red-Edge Triangular Vegetation Index (RTVIcore)	(100 * (ρ NIR − ρ REG)) − (10 * (ρ NIR − ρ G))	[55]
Red-Edge Vegetation Stress Index (RVSI)	((ρ R + ρ NIR)/2) − ρ REG	[50]
Red-Edge Greenness Vegetation Index (REGVI)	(ρ REG − ρ G)/(ρ REG + ρ G)	[68]
Simple Ratio (SR)	ρ NIR/ρ R	[69]
Modified Simple Ratio (MSR)	((ρ NIR − ρ R) − 1)/(((NIR + ρ R) ** (0.5)) + 1)	[48]
Modified Triangular Vegetation Index (MTVI)	1.2 * ((1.2 * (ρ NIR − ρ G)) − (2.5 * (ρ R − ρ G)))	[47]
Wide Dynamic Range Vegetation Index (WDRVI)	(((0.2 * ρ NIR) − ρ R)/((0.2 * ρ NIR) + ρ R))	[70]

2.5. Principal Component Analysis and VI Extraction

In this study, principal component analysis (PCA) was used to decrease the dimensionality of data through the calculation of a series of new variables, or principal components, through linear combinations of the original parameters [71]. PCA was employed as an exploratory data analysis (EDA) technique to describe the relationship between three different agricultural management types (CM, MA, and STM) and multispectral UAS-VIs. The PCA was used for testing whether or not it could improve the classification efficiency of AMPs. PCA was conducted using R version 4.0.2 [72] and the *FactoMineR* package [73]. For extraction of the digital number (DN) values from each VIs of four experimental fields (72 plots in each field), a total of 288 plots were digitized in ArcGIS Pro 2.6.3 [74]. As stated previously, a one-meter buffer zone was extended inwards from each plot boundary to address potential edge effects from agricultural management, and the average VIs were isolated and calculated. These extracted values were further used in this study when building ML algorithms and for AutoML assessment and evaluation.

2.6. AutoML Modeling with Auto-Sklearn

Auto-sklearn [30], a robust and efficient AutoML system first introduced in 2015 and upgraded in 2020 [75], was utilized in this study. Auto-sklearn is developed on the Python Scikit-learn machine learning package. It uses 15 classifiers, 14 feature pre-processing methods, and four data pre-processing methods, giving rise to a structured hypothesis space with 110 hyperparameters [76]. It improves on existing AutoML methods by automatically considering the previous performance on similar datasets, and by constructing ensembles

from the models evaluated during the optimization process. At its core, this method combines the highly parametric ML framework with automatically constructed ML pipelines suggested by the Bayesian optimization method sequential model-based algorithm configuration (SMAC) [77]. SMAC can automatically construct ML pipelines that include feature selection (i.e., removing insignificant features), transformation (i.e., dimensionality reduction), classifier selection comprising support vector machines (SVM) [78], Random Forest (RF) [79], and other algorithms, hyperparameter optimization, etc. Subsequently, it then utilizes a Random Forest technique for swift cross-validation by evaluating one-fold at a time, while at the same time discarding poor-performing hyperparameter settings during early stages. It achieves competitive classification accuracy, in addition to novel pipeline operators that significantly increase classification accuracy on the datasets [80]. During the feature selection stage, any highly correlated VIs were removed to eradicate the influence of collinearity. This step was omitted here since Auto-sklearn deals with the low dimensional optimization problems [81].

In this study, all calculations were done in the open-source operating system LINUX with Intel Core i5-1035G1 CPU (1.00 GHz) and 16 GB RAM. For the AutoML framework, the steps described in [30] were followed, with some modifications for this study (Figure 2b). First, the system used a supplementary approach of extensively applied meta-learning methods to train machine learning models over statistical attributes of datasets and estimated the parameter of models that yielded the best precision [82]. Second, the system automatically built ensembles of the models considered by Bayesian optimization. Third, the system constructed a highly parameterized ML framework from high-performing classifiers and pre-processors implemented within the ML framework. Finally, the system performed broad empirical analysis using a diverse collection of datasets to demonstrate the resulting Auto-sklearn system outperformed preceding AutoML methods. The major AutoML parameter settings of this study are described in Table 2. Due to computational resource constraints and to test the efficiency of AutoML, we first limited the CPU time for each run to 60 s and the running time for evaluating a single model to 10 s as an example of rapid model selection. Subsequently, we then used a total of 1200 s with a 10-s single model computing time as a representative of the better processing of AutoML models. The data were analyzed separately according to the four crop fields (F1–F4), with each field containing 72 plots (n = 72) with a split in the training site and validation site (0.6/0.4) for classification modeling.

Table 2. The AutoML main parameters and descriptions used in this study.

Parameter Name	Range Value	Description
time_left_for_this_task	60–1200 s	The time limit for the search of appropriate models.
per_run_time_limit	10 s	The time limit for a single call to the machine learning model.
ensemble_size	50 (default)	The number of models added to the ensemble built by *Ensemble selection from libraries of models*.
ensemble_nbest	50 (default)	The number of best models for building an ensemble model.
resampling_strategy	CV; folds = 3	(CV = cross-validation); to handle overfitting
seed	47	Used to seed SMAC.
training/testing split	(0.6; 0.4)	Data partitioning way

The other parameters that are not listed on the table were run in default mode.

A recent review study of supervized ML methods applied in land-cover image classification disclosed that Random Forest (RF), support vector machine (SVM), and artificial neural network (ANN) classifiers were among the most commonly used ML techniques from 220 related articles [83]. Therefore, in this study, these popular ML classifiers were selected for comparison against the accuracy performance of AutoML (with 60-s run, and 1200-s run of Auto-sklearn). These algorithms were programmed in Python by the robust ML library Scikit-learn (0.24.2) [76] with the perimeter setting as following: *sklearn.ensemble*.RandomForestClassifier (100 trees; *min_samples_split* (2); *leaf_node* (1)); *sklearn.svm*.SVC (cost (C = 500); gamma (0.5); epsilon (0,01)), and *sklearn.neural_network*.

MLPClassifier (alpha (0.00005); the maximum number of iterations (100,000)) The parameters not mentioned were computed as default settings from Scikit-learn, and for the accuracy, calculation referring to Table 3.

Table 3. The confusion matrix-based accuracy evaluation equations used throughout this study.

Indices	Equations
Recall	TP/(TP + FN)
Precision	TP/(TP + FP)
Specificity	TN/(TN + FP)
Accuracy	TP/(TP + TN + FP + FN)
F1-score	2 * Precision * Recall/(Precision + Recall)
False Positive Rate (FPR)	1 − Specificity = FP/(FP + TN)
True Positive Rate (TPR)	Sensitivity = TP/(TP + FN)

2.7. AutoML Model Evaluation and Visualization

For the visualization and evaluation of the Auto-sklearn model, the workflow included, in general, multiple iterations through feature engineering, algorithm selection, and hyperparameter tuning [84]. In this study, an open-source visual steering tool Yellowbrick visualization package (essentially a wrapper for the Sklearn documentation) was conducted for AutoML evaluation [85]. Yellowbrick contributes to assessing the stability and predictive values of ML models and delivers visualizations for our AutoML classification models. The accuracy evaluation based on the confusion matrix system of the AutoML classification parameters were defined as follows: true positive (TP), false positive (FP), true negative (TN), and false negative (FN), which have been well described in [86]. The equations used in this study are described in Table 3. The derived receiver operating characteristic curve (ROC) graph with the *x*-axis showing FPR and the *y*-axis showing TPR was used in this study to show the relationship among specificity and sensitivity for each possible cut-off [87] and the area under the curve (AUC) ranges from 0 to 1 to visualize the trade-off between the classifier's sensitivity and specificity [87,88]. Macro- and micro-averaging ROC were calculated to evaluate overall classifier performance in multi-class problems. In this approach, the ROC curve was calculated anew, based upon the true positive and false positive rates for all dataset (by weighting curves by the relative frequencies of the dataset and then averaging them) [89,90]. In addition, the precision–recall curve (PR) was calculated for different probability thresholds. PR curves were conducted in cases where there was an imbalance in the observations between the classes [91] as another classification evaluation standard to assist with the ROC curve. The prediction errors (confusion matrix) and classification report that displays precision, recall, and F1-score [92] (Table 3) per class as a heatmap in our study.

Alternatively, even though the AutoML framework facilitates the construction of models, given their black-box nature, the complication of the underlying algorithms and the large number of pipelines they derive leads to the reduced trust of AutoML pipelines systems [93]. Therefore, in our study, PipelineProfiler [94] was conducted for AutoML pipelines visualization. PipelineProfiler is a SOTA in visual analytics for AutoML interactive visualization tool that allows the examination of the solution space of end-to-end ML pipelines. It offers a recovering understanding of how the AutoML algorithms are generated and the perceptions of how they can be optimized. As the outcome of the interactive AutoML pipeline matrix plots, where illustrated Pipeline flowchart, primitives used by the pipelines; one-hot-encoded hyperparameters for the primitive across pipelines; the accuracy ranking; primitive contribution view; and the class balancing of correlation score with accuracy. These calculations and expressions are clearly detail described in the [94] article.

3. Results

3.1. The AMPs Observation in VPTs and VIs Calculation

Figure 3 displays the observation of onsite crop VPTs (i.e., Field 1 (F1) (Figure 3a) and Field 2 (F2) (Figure 3b) with CM treatments) and one of the VIs (NDVI; Figure 3c) captured on 10 July from F1 and F2. It can be observed from the onsite AMPs treatment photographs of F1 and F2 in July that it was not readily distinguishable. In addition, it can be seen from the NDVI image that the heterogeneity within the plot may be caused by edge effects or uneven fertilization. For this reason, we used the plot average value considering the pixels inward boundary clipping to decrease the noise.

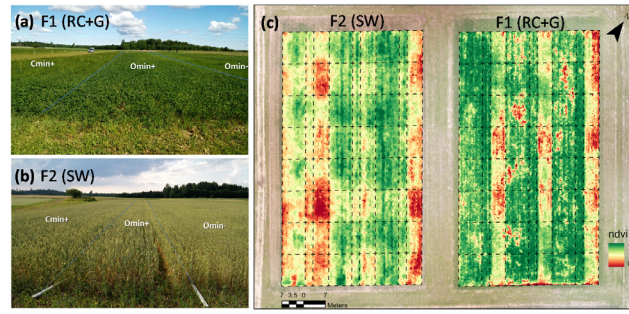

Figure 3. Interpretation diagrams representing onsite crop VPTs and the calculation of VIs per the image captured on 10 July (**a**) Field 1: red clove + grass (RC + G) with CM treatment. (**b**) Field 2: Spring wheat (SW) field with CM treatment. (**c**) Normalized Difference Vegetation Index (NDVI) image captured of F1 (RC + G) and F2 (SW) VPT.

3.2. Monthly PCA Analysis in Various Crop Growth Periods

PCA was conducted as the first step of data exploration in this study to gain an understanding of the relationship between VIs and different AMP categories during the three flight periods (April, May, and July) with their corresponding growing stages (Figure 4). The results showed that on 30 May and 10 July, the PC1 and PC2 captured most of the variation from the F1 to F4 fields with 98.3%, 98.7%, 97.3%, and 97.6%, respectively, on 30 May (Figure 4b), and with 98.7%, 94.0%, 95.4%, and 95.4%, respectively, on 10 July (Figure 4c); followed by 23 April (Figure 4a). In addition, during the three flight periods, the PCA results in May and July provide better separation of the three AMP categories throughout the four crop cultivation areas based on the colored concentration ellipses where the sizes were determined by a 0.95-probability level. In terms of the AMP category, the subclasses of CM (Cmin+ Cmin+ and the other two categories) and MA (M+ and M−) seemed easier for non-overlapping AMP clustering, followed by STM. In terms of crop types, F1 (SW) were better clustered in April, while F2 (SW), F3 (P + O), and F4 (SB + RC) were better clustered in May or July. Given the better clustering performance in May, follow-up AutoML analysis was conducted on the UAS-VIs data of this month. In general, feature selection (finding the most relevant spectral bands) and extraction (reduced set of new significant variables) are commonly used to solve the collinearity and overfitting problems in the dimensionality reduction process [95]. However, after our test results, using PCA, 95% feature extraction in our preliminary experiments could not significantly improve the classification efficiency. Therefore, these PCA results were simply used as a reference basis for AutoML classification.

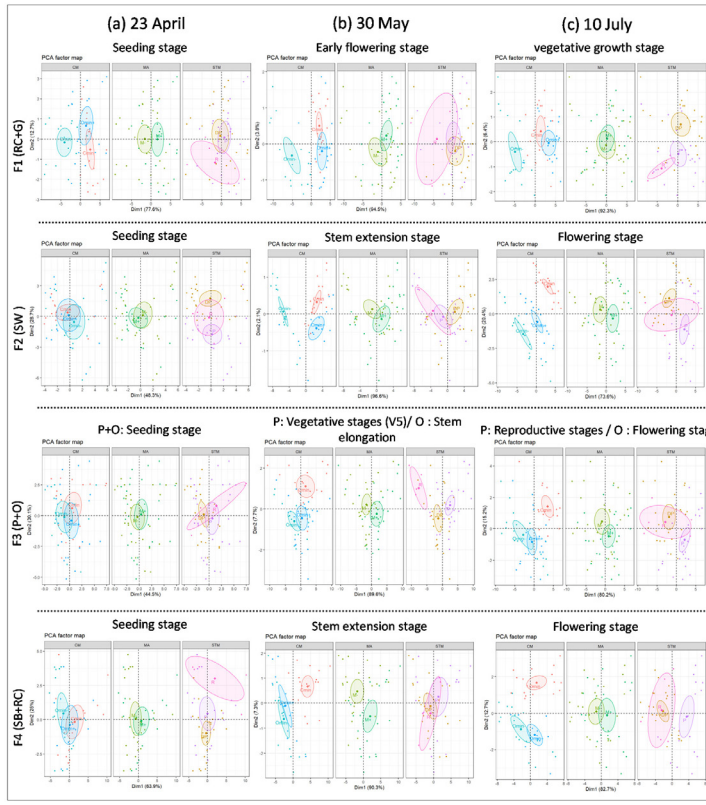

Figure 4. PCA biplot of 19 VI variables (n = 72) of each crop field on (**a**) 23 April, (**b**) 30 May, and (**c**) 10 July. Each biplot shows the PCA individuals (three AMPs) (i.e., CM (Cmin, Omin+, Omin−), MA (M+, M−), STM (DP, P, R)) of the first (*x*-axis: PC1 score) and second (*y*-axis: PC2 score) principal components (the variation explained by the dimensions are shown on the axes); four crop categories (F1–F4) and its corresponding growing stage from top to bottom. Colored concentration ellipses (size determined by a 0.95-probability level) show the observations grouped by marked AMP sub-classes.

3.3. *AutoML ROC and AUC Evaluation of AMP Recognition in May*

The different subclasses and average results of ROC/AUC were calculated for evaluation of the AutoML performance for the AMP classification ability in UAS-VIs that were captured in May (Figure 5), where AUC values were categorized in this study as AUC = 0.5: no discrimination; $0.7 \leq$ AUC ≤ 0.8 (acceptable discrimination); $0.8 \leq$ AUC ≤ 0.9 (excellent discrimination); $0.9 \leq$ AUC ≤ 1.0 (outstanding discrimination) [87].

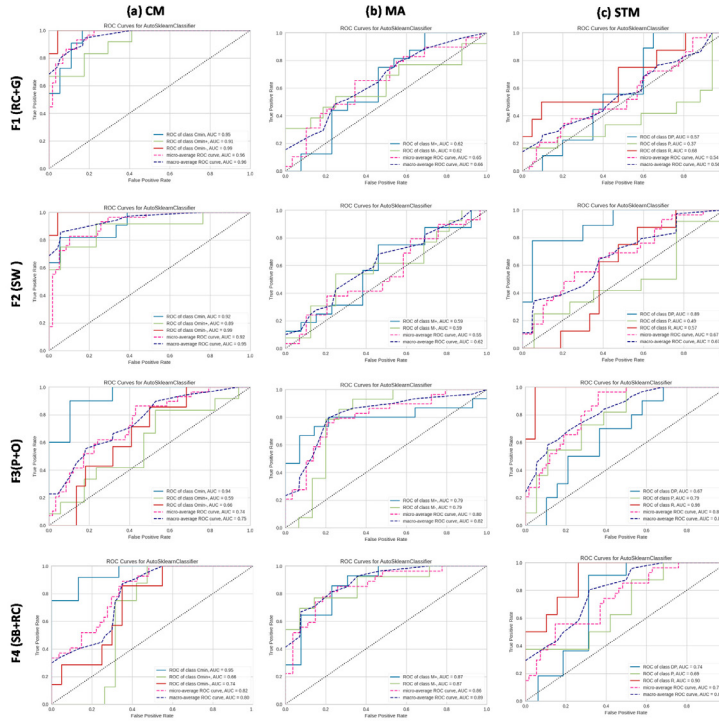

Figure 5. ROC curves and AUC of the AutoML classification corresponding to the subclasses within the AMPs for the acquisition of the UAS-VIs DN in May. From left to right, the ROC curves computed on (**a**) CM (Cmin+ (blue lines), Omin+ (green lines), Omin− (red lines)); (**b**) MA [M+ (blue lines), M− (green lines)]; (**c**) STM (DP (blue lines), P (green lines), R (red lines)); and their micro (pink dotted line) and macro (dark blue dotted line) average performance. Four crop categories (F1–F4) from top to bottom.

The AutoML results showed that the micro-average ROC of CM's classification resulted in F1 (RC + G) and F2 (SW) being higher (AUC = 0.95, and 0.92, respectively). Especially in the subclass Omin−, the AUC both reached 0.99 for the micro-average ROC, followed by F4, and F3 (P + O), with 0.86 and 0.75, respectively) (Figure 5a). On the contrary, MA classification results showed that the micro-average AUC in F3 and F4 were higher (AUC = 0.83, and 0.89, respectively), followed by F1 (AUC = 0.71). F2 performance for MA was the worst (AUC = 0.51), with no discrimination ability (Figure 5b). In contrast, STM classification results were generally poor, with better results only present in F3, while other fields have larger divergence in classification results under the sub-class (DP, P, and R), as shown in Figure 5c). Overall, the AutoML classification ability from UAS-VIs of CM was the best, followed by MA and STM.

3.4. AutoML Precision–Recall, Prediction Error, and Classification Report of CM Recognition

Among the classification results of AMPs in May (Figure 5) of four crop types, CM yielded the best ROC/AUC overall performance. Therefore, we used the precision–recall (PR) curves, prediction error, and classification report plots to gain an in-depth understanding of the classification status of CM treatments (Figure 6).

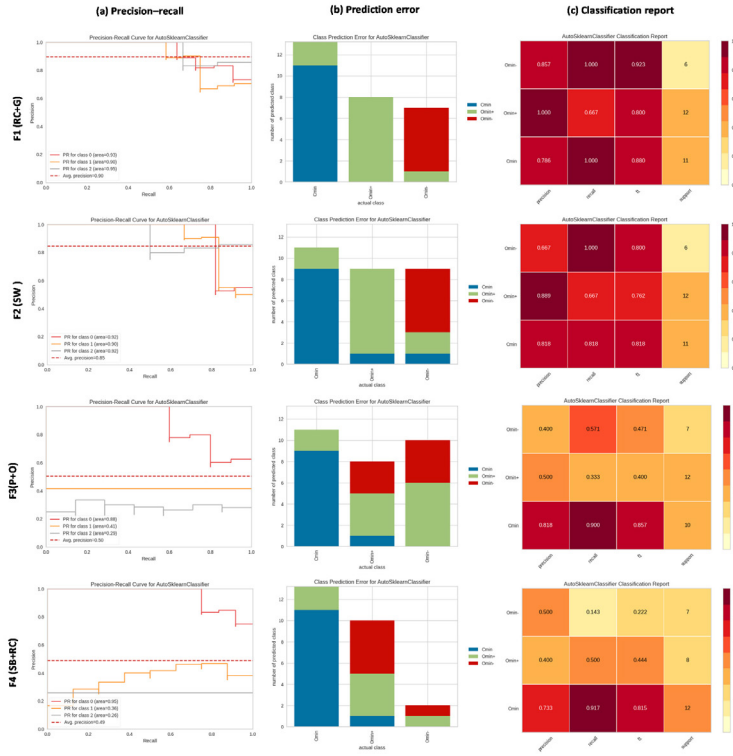

Figure 6. The evaluation of AutoML classification of AMPs from the acquisition of the UAS-VIs DN in May. (**a**) Precision–recall, where the class 0, 1, 2 equals to Cmin+, Omin+, and Omin−, respectively (**b**) Prediction error (confusion matrix), the *X*-axis represents the three subclass form CM result in May, and the *Y*-axis represents the type (with color), and the number of correct or incorrect estimates., and (**c**) Classification report lists the precision, recall, and F1-score per class as a heatmap for overall comprehensive evaluation results. The calculation methods used in this figure are shown in Table 3.

The PR curve of F4 CM shows the trade-off between a classifier's precision performance from UAS VIs in May (Figure 6a), where a model with perfect performance is depicted at the coordinate of (1,1). A curve that tends towards the (1, 1) coordinate represents a well-performing model, whereas a no-skill classifier is depicted as a horizontal line on the plot with a precision that is proportional to the number of positive examples in the dataset. For a balanced dataset, this value ought to be 0.5 [96]. The results showed

that the classifications of Fields 1 and 2 were promising, their average PR being 0.90 and 0.85, respectively, while the results of F3 and F4 were poor (0.50 and 0.49). We can further discover from the prediction error graph (Figure 6b) in F3 and F4 that the judgment error of Cmin+ was low, and the confusions of Omin+ and Omin− were more common. We can also compare the precision, recall, and F1-score results of various cultivation method sub-classes to evaluate the classification accuracy from the heatmap (Figure 6c).

3.5. AutoML Pipeline Visualization

An interactive AutoML visualization tool *PipelineProfiler* was used in this study. Figure 6 shows the CM classification results across four crop fields in May with the accuracy performance of AutoML pipelines running time set at 60 s, and the primitive comparison against the others, and the real-time hyperparameter selection strategy (Figure 7). The results demonstrated that the best classifier found for Field 1 was linear discriminant analysis (LDA) [97] (Figure 7a), for Field 2, it was the Extra Trees Algorithm [98] (Figure 7b), for Field 3, it was LDA (Figure 7c), and Random Forest (RF) for Field 4 (Figure 7d), with each of their hyperparameters found by AutoML also being represented in the figures.

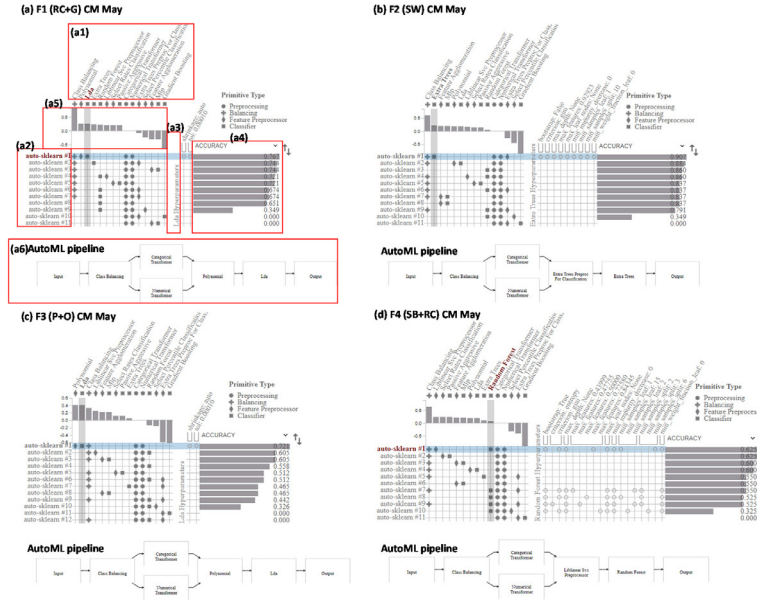

Figure 7. The interactive AutoML pipeline matrix plots with running time-limited setting 60 s sorted by accuracy performance (**a**–**d**), (**a**) Field 1 pipeline matrix with the Top1 classifier LDA, where (**a1**) illustrated Primitives (in columns) used by the pipelines (**a2**) (in rows, the blue line showed the best accuracy rank); (**a3**) one-hot-encoded hyperparameters (in columns) for the primitive across pipelines, (**a4**) the AutoML pipeline with the accuracy ranking; (**a5**) Primitive contribution view, showing the correlations between primitive usage and pipeline scores—in a5 displays that class balancing has the highest correlation score with accuracy; (**a6**) Step by step AutoML Pipeline flowchart. The ML box before Output represents the classifier used by this set of algorithms (in a6 LDA as the classifier) (**b**–**d**) Fields 2, 3, and 4 interactive pipeline matrix sort by AutoML accuracy performance with the chosen hyperparameters (top 1 was listed).

3.6. Comparison of Performance between AutoML and Other Machine Learning Technologies

Based on the large calculations and multiple classifier selections that were required during the initial stage of AutoML computations, the processing time setting of 60 s may not completely reflect the performance power of AutoML. To evaluate the effects of AutoML processing time, we adjusted the times to 1200 s and 60 s (original running time) and considered the AMPs' classification accuracy with RF, SVM, and ANN algorithms (Table 4). The results demonstrated that under the permutation and combination of ML algorithms included in AutoML, classification accuracy did not perform well in 60 s of computing time. Furthermore, performance was the worst in F1 CM, F2 STM, and F3 CM classifications compared to RF, SVM, and ANN. However, as processing time was increased to 1200 s, the classification accuracy of AutoML in AMPs was shown to improve. The results also indicated that overall AutoML (1200 s) and RF classifiers produced 5 and 3 best classification accuracy in AMPs, respectively (in black bold) and did not produce the worst accuracy values (in bold red) in any instances. Regarding SVM and ANN, the classifiers performed the best in 3 and 5 cases, respectively. However, these methods consistently produced low-performing classifiers compared to other AMPs.

Table 4. The AMPs classification accuracy comparison of AutoML and three other popular applied ML (RF, SYM, and ANN) algorithms in UAS.

Field	AMPs	ML Algorithms				
		AutoML (1200 s Run)	AutoML (60 s Run)	RF	SVM	ANN
F1 (RC + G)	CM	0.79	**0.76** **	0.79	0.83	**0.86** *
	MA	0.59	**0.62** *	**0.62** *	**0.62** *	**0.55** **
	STM	**0.57** *	0.31	0.48	**0.38** **	0.48
F2 (WS)	CM	0.79	0.79	0.79	**0.83** *	**0.72** **
	MA	**0.55** *	0.52	0.48	0.52	**0.45** **
	STM	**0.52** *	**0.45** **	0.48	**0.45** **	**0.52** *
F3 (P + O)	CM	**0.55** *	**0.41** **	**0.55** *	0.48	**0.55** *
	MA	0.66	0.72	**0.76** *	**0.62** **	**0.76** *
	STM	0.66	**0.69** *	**0.69** *	**0.57** **	0.59
F4 (SB + RC)	CM	0.57	**0.59** *	0.56	**0.59** *	**0.48** **
	MA	**0.85** *	0.78	0.67	0.78	**0.63** **
	STM	0.56	0.59	0.59	**0.52** **	**0.63** *

(*) The bold black numerical value in the Table represents the highest accuracy classifier in the row; (**) the thin red numerical value represents the worst accuracy in the row.

4. Discussion

This paper is the first study to use an auto-learning system, with UAS multispectral-derived VIs, for agricultural classification purposes. The study provides a novel AutoML framework across multiple AMP activities and presents a UAS and ML methodology optimized for future PA and crop phenotyping research.

4.1. Applicability and the Impact of the AutoML Method in UAS

In this study, we employed a SOTA, open-sourced AutoML framework for automatic, rapid multispectral image classification strategies and assistance in optimizing problematic hyperparameter adjustments. This technology brings several benefits and enhances the application of UAS for environmental and ecological research classification tasks.

First, UAS related classification research publications have significantly increased within recent years, with over a hundred articles developed since 2017. This substantial adoption of UAS related classification approaches demonstrates its impact and the mounting interest in such research issues [99]. Our UAS-AutoML framework may also be implemented in other UAS classification research, such as research employing multi-sensors (i.e., thermal, visible light, hyperspectral, radar or light detection and ranging

(Lidar) sensors) across a range of contemporary agriculture classification activities (i.e., weed management [100,101], crop phenotyping [9,102–104], disease monitoring [105,106], etc.), as well as research focused on ecological classification schemes, multispectral-based plant community mapping options [107], and coastal wetland vegetation classification results [108].

Second, the AutoML framework quickly provided usable classifiers and hyperparameter selections for unknown UAS classification tasks and parameter selection. For example, in the current study, the parameters and applicable classifiers of AMPs were unknown a priori. However, it provided a promising and efficient performance rating for classifiers for inclusion in modeling selection. As the results of Figure 6 show, LDA (Figure 7a,c) and Extra Trees (Figure 7b) were chosen as the best classifiers corresponding to the VPT fields of the AMP recognition task. These ML methods have been less applied and referenced in the field of UAS [83]. These findings clearly illustrate that AutoML has the potential to locate alternative ML approaches that might customarily be ignored by investigators with unknown classification subjects.

Third, the operational efficiency of AutoML classifiers can be given a time limit and gives the researcher the flexibility to find the most suitable formula within the required time. In general, a longer time setting allows for increasingly accurate results with additional classifier combinations. Since our experiments did not involve substantially large datasets, the focus was put on time setting close to the minimum limit of AutoML calculation (60 s of total CPU operation (this can be up to 3000 s) and 10 s of a single ML algorithm computation) to highlight the flexibility and rapid performance of AutoML.

Finally, within our research, the latest released AutoML interactive visualization system *PipelineProfiler* was employed and assisted in the screening of classifiers and the reference of fine-tuning parameters when analyzing UAS data. This interaction included adjustable time, accuracy ranking, and selection of hyperparameters in response to the requirement of customized UAS modeling. Our results showed that AutoML computations within 60-s-run produced between 11 and 12 pipelines (Figure 7), which might offer a beneficial foundation for providing adequate outcomes in most cases with minimal attempts and time.

4.2. The Impact of Algorithm Selection, Cultivated Period, and Crop Types in AutoML AMP Recognization

In terms of algorithm selection in our AMP classification results, different classifiers were suggested by AutoML as the best performances even within the same AMP category for different crop types (Figure 7). We can conclude that applying AutoML in UAS-derived multispectral VI data allowed for the consideration of a variety of algorithm combinations to meet the complexity of the VPT field. We also compared the three most used ML algorithms (RF, SVM, and ANN) in the UAS classification fields with AutoML algorithms (Table 4). The overall performance showed that AutoML (with 1200-s CPU duration) provided the five best (or equal best) accuracy performances (shown in bold black in Table 4). Interestingly, in all tests, the AutoML (1200) and RF methods were never found to be the worst-performing methods (shown in bold red). Moreover, when using the ANN method, despite providing five of the best classification accuracy results, this method also included five of the worst performance results. Similar outcomes were observed regarding the SVM and AutoML 60-s runs.

From our results, we can deduce that increasing the computing time has the potential to improve the accuracy and stability of AutoML classification performance under certain AMPs conditions. However, it also highlights the potential to include AutoML methods in the computation of common classification problem-solving. Similar ranks were shown in a study that compared the results of the numerous classifiers with Auto-sklearn, where the RF classifier presented the strongest performance, and SVM showed robust performance for some datasets [30]. Since the Auto-sklearn classifiers are based on Scikit-learn as a blueprint, it should theoretically capture the hyperparameters of the RF algorithm on what was selected for Table 4. Despite the strong performance of AutoML (1200 s), there were

still several results that indicated the inferiority of AutoML (1200 s) when compared to the RF classifier (i.e., Field 1 MA; Field 3 MA, and STM). Moreover, in a few cases, the accuracy of AutoML (1200 s) was even lower than the calculation result of the 60-s set (i.e., Field 1 MA, Field 3 STM, and Field 4 CM). It may be that the algorithm computations involve different factors other than accuracy, and the model it uses to tune the parameters actively tries to avoid overfitting. This will possibly lead to the situation where the most accurate model, on the testing or training data, will not be the one that can generalize the best on real data. In addition, developers from the Auto-sklearn team have previously described that during the ensemble selection phase, the methods can add numerous substandard models to the final ensemble, and unregularized selection may lead to overfitting with a small number of candidate models [40]. This result shows that there is still room for improvement regarding AutoML calculation methods in the future.

In terms of cultivated period and crop type, according to the monthly performance of different crop growth stages, the PCA results indicated that the VPT with better clustering performance occurred during the flight in May, with a confidence level of 0.95 (Figure 4b). In this regard, this flight period was further used for our AMPs' classification study. Conversely, in the case of more homogeneous crop types (Field 3 (WS)), and despite the promising classification result in CM, the results of MA and STM were not as effective as other crops (Figures 5 and 6). These results may suggest that even with higher heterogeneity of cultivation within the plots (i.e., F1, F3, and F4), it appears not necessarily to affect the classification ability. However, concerning the Field 3 results from the PCA in May (stage of stem elongation) and July (stage of flowering), the MA clustering ability was better with a 0.95 confidence level in both months, and the accuracy was later improved from the classification analysis. The results of our study have demonstrated that, although the feature selection stage of AutoML is a black box, we can still preliminarily determine the potential predictive ability of the AutoML model based on PCA result and reduce the cost of period selection as we did in this study. In addition, this study has contributed evidence to the classification obstacles in the case of STM that may be caused by the orientation of images taken over vegetation or soil with uniform texture and re-cursive pattern, sub-optimal flight configuration [109], or unflavored VIs selection. Some studies also suggest that the use of grey-level co-occurrence matrix (GLCM)-based texture information [100,110], semantic segmentation [111], or edge computing [112] can improve the accuracy of UAS-ML classification in the crop categories. This may be an applicable technology for AMPs classification in the future. The applicability and optimization of this framework, and the visualization of feature importance, required the optimization of the AutoML programmers and UAS application feedback to improve.

Currently, multispectral indices have been effectively applied in some AMP image analysis studies with the color, texture, and shape factors of the agricultural land at the satellite level. These include conservation tillage methods identification [113] and agriculture landscapes with pixel-based or object-based classification tasks [114,115]. AMPs application are indispensable for environmental monitoring and for facilitating the agricultural decision-making process, regarding the adoption practices proposed by growing conservation agriculture demand [116], and for its potential upscaling ability to accelerate land cover classification studies. Recently, combining commonly adopted management practice with UAS multispectral-VIs research has gradually gained attention and has been applied to cotton and sorghum fields [117]. In our study, the effective application of UAS sensors to recognize multiple AMP categories has been shown. More specifically, an UAS-AutoML approach can improve the classification ability under specific crop AMPs, highlighting that, in this study site, classification performed better in CM, with overall classification performance followed by MA and STM.

4.3. The Limitations of Our Method

In this study, not all classifiers computed within the Auto-sklearn system were able to be backtracked and reviewed to investigate the individual feature importance rankings

of VIs, which has limitations in terms of their ability to assist in the selection of suitable VIs for AMPs classification tasks. However, our efforts to achieve a wide-ranging and well-considered predictor collection through a variety of VI combinations may lead to performance improvements. This study may also be limited by the location, crop categories selected, and varieties present at the study site. However, these issues can be simply addressed by including a wider range of VPT at multiple study sites and across a greater diversity of crop types in future investigations. Due to the characteristic complexity and repeatability of VPT, we need to recognize that the small sample size, and the potential interaction effects between trials, were not fully addressed. A potential solution worth pursuing may be to increase the VPT sampling size and/or enhance the segmentation number of each plot, ultimately increasing the training samples for AutoML calculation. Currently, the applicability of the AutoML framework will still require more UAS-based tests in the future to demonstrate its true potential and effectiveness.

5. Conclusions

First, our study demonstrated a novel UAS technology and a state-of-the-art AutoML framework across multiple AMP tasks through non-destructive and cost-effective approaches. The scientific merit of this article lay in utilizing artificial intelligence to replace the judgment of the human for UAS classification analysis with its automated data pre-processing, model selection, feature engineering, and hyperparameter optimization capabilities. Furthermore, it provided innovative insights into agricultural management practices and accelerated the intellectualized progress of the in-field monitoring UAS system and established future crop phenotyping abilities. In our study, AutoML embodied "learning how to learn" for any given UAS subject; and it is the first study of its kind to apply an auto-learning system for AMP classification tasks in multispectral-derived VI data.

Second, in this study, we employed an AutoML workflow combined with two innovative visualization tools. We performed three multispectral-UAS flights at the farm-scale, under the four crop types (RC + G, SW, PO, and SB + RC) of VPT within three AMPs (CM, MA, and STM) treatments. In addition, we compared AutoML performance with those of three widely used ML methods. The ML comparison analysis results showed AutoML achieved the most overall classification accuracy numbers after 1200 s of calculation and without any of the worst-performing classifications of the given datasets. In terms of AMPs classification, the best recognized period for data capture occurred in the crop vegetative growth stage (in May of Estonia). The result demonstrated that CM yielded the best performance in terms of treatment, followed by MA, and STM; the last was shown to be the worst-performing treatment. These conclusions may be attributed to the low heterogeneity of the spectral reflectance value in the corresponding AMP treatment.

Third, the flexibility of fixed-wing imaging technology provides longer flight durations and thus allows for larger applications, such as commercial farmland, grasslands, forests, etc. Furthermore, the multispectral dataset produces various precise VIs without the need for any supplementary sensors, which reduces measurement errors and significantly reduces costs. In addition, given the AutoML's open-sourced platform and the powerful capabilities of automation, the complexities surrounding parameter selection in machine learning are greatly reduced, while it also has the potential to select long-ignored but highly efficient ML algorithms. Regarding the choice of AutoML systems or interfaces, although many of them have been developed successively (i.e., Auto-sklearn, H2O AutoML, AutoKeras), it is necessary to identify whether their subsequent updates and revisions keep up to date with current times.

Lastly, this UAS-AutoML solution has the potential to be implemented across a variety of other UAS classification research, such as contemporary agricultural classification methods, multispectral-based plant community mapping, ecological or wetland plant community recognition. Other remote sensing classification methods that lack algorithm

and hyperparameter backgrounds may also be considered and benefited from our findings and insights.

In summary, our study, the UAS application particularly focused on the adoption and application of AutoML method across a diverse range of agricultural environmental assessment and management applications. Our approach demonstrated that UAS based on our AutoML framework, can recognize multiple agricultural management practices under certain conditions and that the integration of UAS technologies, geoprocessing methods, and automatic systems are vital tools for increasing the knowledge of plant–environment interactions within the management of crops. The framework also considerably contributes towards the simplified advancement of image-driven analytical pipelines for current VPT systems used in most countries. At the end of preparing this study, the Google Cloud AutoML also came out in 2019 for image-recognition use cases [118], showing that automatic learning will drive a non-negligible impact in the UAS field and provide new insight into the potential for remotely sensed solutions to field-based and multifunctional platforms for the demands of precision agriculture in the future.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/rs13163190/s1, Figure S1: Daily climograph of the study area (Kuusiku) during the flying period, including the previous 6 days (a. 17–23 April, b. 24–30 May, and c. 4–10 July) in 2019. Blue bars and the red line represent the daily average of rainfall and temperature, respectively.

Author Contributions: All authors contributed significantly to the work presented in this paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the European Regional Development Fund within the Estonian National Programme for Addressing Socio-Economic Challenges through R&D (RITA): L180283PKKK and the Doctoral School of Earth Sciences and Ecology, financed by the European Union, European Regional Development Fund (Estonian University of Life Sciences ASTRA project "Value-chain based bio-economy").

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the data size.

Acknowledgments: We give great thanks to the community that develop and maintain the scikit-learn framework, as well as the Auto-sklearn interface developed at the University of Freiberg. In addition, this work was supported by Kuusiku Variety Testing Centre, Estonia, with the tests carried out by order of Agricultural Board. Information about variety listings are currently available in www.pma.agri.ee.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mulla, D.J. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosyst. Eng.* **2013**, *114*, 358–371. [CrossRef]
2. Tripicchio, P.; Satler, M.; Dabisias, G.; Ruffaldi, E.; Avizzano, C.A. Towards Smart Farming and Sustainable Agriculture with Drones. In Proceedings of the 2015 International Conference on Intelligent Environments, IEEE, Rabat, Morocco, 15–17 July 2015; pp. 140–143. [CrossRef]
3. Herwitz, S.R.; Johnson, L.F.; Dunagan, S.E.; Higgins, R.G.; Sullivan, D.V.; Zheng, J.; Lobitz, B.M.; Leung, J.G.; Gallmeyer, B.A.; Aoyagi, M.; et al. Imaging from an unmanned aerial vehicle: Agricultural surveillance and decision support. *Comput. Electron. Agric.* **2004**, *44*, 49–61. [CrossRef]
4. Zhang, C.; Kovacs, J.M. The application of small unmanned aerial systems for precision agriculture: A review. *Precis. Agric.* **2012**, *13*, 693–712. [CrossRef]
5. Tsouros, D.C.; Bibi, S.; Sarigiannidis, P.G. A Review on UAV-Based Applications for Precision Agriculture. *Information* **2019**, *10*, 349. [CrossRef]
6. Xiang, H.; Tian, L. Development of a low-cost agricultural remote sensing system based on an autonomous unmanned aerial vehicle (UAV). *Biosyst. Eng.* **2011**, *108*, 174–190. [CrossRef]

7. Raeva, P.L.; Šedina, J.; Dlesk, A. Monitoring of crop fields using multispectral and thermal imagery from UAV. *Eur. J. Remote Sens.* **2018**, *52*, 192–201. [CrossRef]
8. Sankaran, S.; Khot, L.R.; Carter, A.H. Field-based crop phenotyping: Multispectral aerial imaging for evaluation of winter wheat emergence and spring stand. *Comput. Electron. Agric.* **2015**, *118*, 372–379. [CrossRef]
9. Yang, G.; Liu, J.; Zhao, C.; Li, Z.; Huang, Y.; Yu, H.; Xu, B.; Yang, X.; Zhu, D.; Zhang, X.; et al. Unmanned Aerial Vehicle Remote Sensing for Field-Based Crop Phenotyping: Current Status and Perspectives. *Front. Plant Sci.* **2017**, *8*, 1111. [CrossRef]
10. Araus, J.L.; Cairns, J. Field high-throughput phenotyping: The new crop breeding frontier. *Trends Plant Sci.* **2014**, *19*, 52–61. [CrossRef]
11. Andrade, J.; Edreira, J.I.R.; Mourtzinis, S.; Conley, S.; Ciampitti, I.; Dunphy, J.E.; Gaska, J.M.; Glewen, K.; Holshouser, D.L.; Kandel, H.J.; et al. Assessing the influence of row spacing on soybean yield using experimental and producer survey data. *Field Crop. Res.* **2018**, *230*, 98–106. [CrossRef]
12. Laidig, F.; Piepho, H.-P.; Drobek, T.; Meyer, U. Genetic and non-genetic long-term trends of 12 different crops in German official variety performance trials and on-farm yield trends. *Theor. Appl. Genet.* **2014**, *127*, 2599–2617. [CrossRef]
13. Lollato, R.P.; Roozeboom, K.; Lingenfelser, J.F.; Da Silva, C.L.; Sassenrath, G. Soft winter wheat outyields hard winter wheat in a subhumid environment: Weather drivers, yield plasticity, and rates of yield gain. *Crop. Sci.* **2020**, *60*, 1617–1633. [CrossRef]
14. Zhu-Barker, X.; Steenwerth, K.L. Nitrous Oxide Production from Soils in the Future: Processes, Controls, and Responses to Climate Change, Chapter Six. In *Climate Change Impacts on Soil Processes and Ecosystem Properties*; Elsevier: Amsterdam, The Netherlands, 2018; pp. 131–183. [CrossRef]
15. De Longe, M.S.; Owen, J.J.; Silver, W.L. *Greenhouse Gas Mitigation Opportunities in California Agriculture: Review of California Rangeland Emissions and Mitigation Potential*; Nicholas Institute for Environ, Policy Solutions Report; Duke University: Durham, NC, USA, 2014.
16. Zhu-Barker, X.; Steenwerth, K.L. Nitrous Oxide Production from Soils in the Future. In *Developments in Soil Science*; Elsevier: Amsterdam, The Netherlands, 2018; Volume 35, pp. 131–183. [CrossRef]
17. Crews, T.; Peoples, M. Legume versus fertilizer sources of nitrogen: Ecological tradeoffs and human needs. *Agric. Ecosyst. Environ.* **2004**, *102*, 279–297. [CrossRef]
18. Munaro, L.B.; Hefley, T.J.; DeWolf, E.; Haley, S.; Fritz, A.K.; Zhang, G.; Haag, L.A.; Schlegel, A.J.; Edwards, J.T.; Marburger, D.; et al. Exploring long-term variety performance trials to improve environment-specific genotype × management recommendations: A case-study for winter wheat. *Field Crop. Res.* **2020**, *255*, 107848. [CrossRef]
19. Gardner, J.B.; Drinkwater, L.E. The fate of nitrogen in grain cropping systems: A meta-analysis of 15N field experiments. *Ecol. Appl.* **2009**, *19*, 2167–2184. [CrossRef]
20. Drinkwater, L.; Snapp, S. Nutrients in Agroecosystems: Rethinking the Management Paradigm. *Adv. Agron.* **2007**, *92*, 163–186. [CrossRef]
21. van Ittersum, M.K.; Cassman, K.G.; Grassini, P.; Wolf, J.; Tittonell, P.; Hochman, Z. Yield gap analysis with local to global relevance—A review. *Field Crop. Res.* **2013**, *143*, 4–17. [CrossRef]
22. Nawar, S.; Corstanje, R.; Halcro, G.; Mulla, D.; Mouazen, A.M. Delineation of Soil Management Zones for Variable-Rate Fertilization. *Adv. Agron.* **2017**, *143*, 175–245. [CrossRef]
23. Chlingaryan, A.; Sukkarieh, S.; Whelan, B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Comput. Electron. Agric.* **2018**, *151*, 61–69. [CrossRef]
24. Zhang, W.; Zhang, Z.; Chao, H.-C.; Guizani, M. Toward Intelligent Network Optimization in Wireless Networking: An Auto-Learning Framework. *IEEE Wirel. Commun.* **2019**, *26*, 76–82. [CrossRef]
25. He, X.; Zhao, K.; Chu, X. AutoML: A survey of the state-of-the-art. *Knowl. Based Syst.* **2020**, *212*, 106622. [CrossRef]
26. Mendoza, H.; Klein, A.; Feurer, M.; Springenberg, J.T.; Hutter, F. Towards Automatically-Tuned Neural Networks. In Proceedings of the Workshop on Automatic Machine Learning, New York, NY, USA, 24 June 2016.
27. Zöller, M.-A.; Huber, M.F. Benchmark and Survey of Automated Machine Learning Frameworks. *J. Artif. Intell. Res.* **2021**, *70*, 409–472. [CrossRef]
28. Yao, Q.; Wang, M.; Chen, Y.; Dai, W.; Hu, Y.Q.; Li, Y.F.; Tu, W.W.; Yang, Q.; Yu, Y. Taking the Human out of Learning Applications: A Survey on Automated Machine Learning. *arXiv* **2018**, arXiv:1810.13306.
29. Thornton, C.; Hutter, F.; Hoos, H.H.; Leyton-Brown, K. Auto-WEKA: Combined Selection and Hyperparame-ter Optimization of Classification Algorithms. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Chicago, IL, USA, 11–14 August 2013; pp. 847–855.
30. Feurer, M.; Klein, A.; Eggensperger, K.; Springenberg, J.T.; Blum, M.; Hutter, F. Auto-sklearn: Efficient and Robust Automated Machine Learning. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 8–14 December 2019; pp. 113–134. [CrossRef]
31. Remeseiro, B.; Bolon-Canedo, V. A review of feature selection methods in medical applications. *Comput. Biol. Med.* **2019**, *112*, 103375. [CrossRef]
32. Babaeian, E.; Paheding, S.; Siddique, N.; Devabhaktuni, V.K.; Tuller, M. Estimation of root zone soil moisture from ground and remotely sensed soil information with multisensor data fusion and automated machine learning. *Remote Sens. Environ.* **2021**, *260*, 112434. [CrossRef]

33. Ledell, E.; Poirier, S. H_2O AutoML: Scalable Automatic Machine Learning. In Proceedings of the 7th ICML Workshop on Automated Machine Learning, Vienna, Austria, 17 July 2020; Available online: https://www.automl.org/wp-content/uploads/2020/07/AutoML_2020_paper_61.pdf?fbclid=IwAR2QaAJWDbgi1jIfnhK83x2g3hV6APfvTZoeUblcf4q44wxqT1z5oRTiEVo (accessed on 20 October 2020).
34. Koh, J.C.O.; Spangenberg, G.; Kant, S. Automated Machine Learning for High-Throughput Image-Based Plant Phenotyping. *Remote Sens.* **2021**, *13*, 858. [CrossRef]
35. Jin, H.; Song, Q.; Hu, X. Auto-Keras: An Efficient Neural Architecture Search System. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Anchorage, AL, USA, 4–8 August 2019. [CrossRef]
36. Komer, B.; Bergstra, J.; Eliasmith, C. Hyperopt-Sklearn: Automatic Hyperparameter Configuration for Scikit-Learn. In Proceedings of the 13th Python in Science Conference, Austin, TX, USA, 6–12 July 2014; pp. 34–40. [CrossRef]
37. FAO. *World Reference Base for Soil Resources*; World Soil Resources Report 103; FAO: Rome, Italy, 2006; ISBN 9251055114.
38. Poncet, A.M.; Knappenberger, T.; Brodbeck, C.; Fogle, J.M.; Shaw, J.N.; Ortiz, B.V. Multispectral UAS Data Accuracy for Different Radiometric Calibration Methods. *Remote Sens.* **2019**, *11*, 1917. [CrossRef]
39. Kelcey, J.; Lucieer, A. Sensor Correction of a 6-Band Multispectral Imaging Sensor for UAV Remote Sensing. *Remote Sens.* **2012**, *4*, 1462–1493. [CrossRef]
40. Feurer, M.; Eggensperger, K.; Falkner, S.; Lindauer, M.; Hutter, F. Practical Automated Machine Learning for the AutoML Challenge 2018. In Proceedings of the International Workshop on Automatic Machine Learning at ICML, Stockholm, Sweden, 10–15 July 2018.
41. Metsar, J.; Kollo, K.; Ellmann, A. Modernization of the Estonian National GNSS Reference Station Network. *Geod. Cartogr.* **2018**, *44*, 55–62. [CrossRef]
42. de Lima, R.; Lang, M.; Burnside, N.; Peciña, M.; Arumäe, T.; Laarmann, D.; Ward, R.; Vain, A.; Sepp, K. An Evaluation of the Effects of UAS Flight Parameters on Digital Aerial Photogrammetry Processing and Dense-Cloud Production Quality in a Scots Pine Forest. *Remote Sens.* **2021**, *13*, 1121. [CrossRef]
43. Tomaštík, J.; Mokroš, M.; Surový, P.; Grznárová, A.; Merganič, J. UAV RTK/PPK Method—An Optimal Solution for Mapping Inaccessible Forested Areas? *Remote Sens.* **2019**, *11*, 721. [CrossRef]
44. Zhang, J.; Huang, W.; Zhou, Q. Reflectance Variation within the In-Chlorophyll Centre Waveband for Robust Retrieval of Leaf Chlorophyll Content. *PLoS ONE* **2014**, *9*, e110812. [CrossRef]
45. Dong, T.; Meng, J.; Shang, J.; Liu, J.; Wu, B. Evaluation of Chlorophyll-Related Vegetation Indices Using Simulated Sentinel-2 Data for Estimation of Crop Fraction of Absorbed Photosynthetically Active Radiation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 4049–4059. [CrossRef]
46. Gitelson, A.; Merzlyak, M.N. Spectral Reflectance Changes Associated with Autumn Senescence of Aesculus hippocastanum L. and Acer platanoides L. Leaves. Spectral Features and Relation to Chlorophyll Estimation. *J. Plant Physiol.* **1994**, *143*, 286–292. [CrossRef]
47. Haboudane, D.; Miller, J.R.; Pattey, E.; Zarco-Tejada, P.J.; Strachan, I.B. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sens. Environ.* **2004**, *90*, 337–352. [CrossRef]
48. Chen, J.M. Evaluation of Vegetation Indices and a Modified Simple Ratio for Boreal Applications. *Can. J. Remote Sens.* **1996**, *22*, 229–242. [CrossRef]
49. Wu, C.; Niu, Z.; Tang, Q.; Huang, W. Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Agric. For. Meteorol.* **2008**, *148*, 1230–1241. [CrossRef]
50. Merton, R.; Huntington, J. Early Simulation Results of the Aries-1 Satellite Sensor for Multi-Temporal Vege-tation Research Derived from Aviris. In Proceedings of the Eighth Annual JPL, Orlando, FL, USA, 7–14 May 1999; Available online: http://www.eoc.csiro.au/hswww/jpl_99.htm (accessed on 22 October 2020).
51. Henebry, G.; Viña, A.; Gitelson, A. The Wide Dynamic Range Vegetation Index and Its Potential Utility for Gap Analysis. Papers in Natural Resources 2004. Available online: https://digitalcommons.unl.edu/natrespapers/262 (accessed on 22 October 2020).
52. Wu, W. The Generalized Difference Vegetation Index (GDVI) for Dryland Characterization. *Remote Sens.* **2014**, *6*, 1211–1233. [CrossRef]
53. Strong, C.J.; Burnside, N.G.; Llewellyn, D. The potential of small-Unmanned Aircraft Systems for the rapid detection of threatened unimproved grassland communities using an Enhanced Normalized Difference Vegetation Index. *PLoS ONE* **2017**, *12*, e0186193. [CrossRef]
54. Kross, A.; McNairn, H.; Lapen, D.; Sunohara, M.; Champagne, C. Assessment of RapidEye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *34*, 235–248. [CrossRef]
55. Chen, P.F.; Tremblay, N.; Wang, J.H.; Vigneault, P.; Huang, W.J.; Li, B.G. New Index for Crop Canopy Fresh Biomass Estimation. *Spectrosc. Spectr. Anal.* **2010**, *30*, 512–517. [CrossRef]
56. Mutanga, O.; Skidmore, A. Narrow band vegetation indices overcome the saturation problem in biomass estimation. *Int. J. Remote Sens.* **2004**, *25*, 3999–4014. [CrossRef]
57. Vasudevan, A.; Kumar, D.A.; Bhuvaneswari, N.S. Precision farming using unmanned aerial and ground vehicles. In Proceedings of the 2016 IEEE International Conference on Technological Innovations in ICT for Agriculture and Rural Development, TIAR, Chennai, India, 15–16 July 2016; pp. 146–150. [CrossRef]

58. Ballester, C.; Brinkhoff, J.; Quayle, W.C.; Hornbuckle, J. Monitoring the Effects of Water Stress in Cotton using the Green Red Vegetation Index and Red Edge Ratio. *Remote Sens.* **2019**, *11*, 873. [CrossRef]
59. Feng, H.; Tao, H.; Zhao, C.; Li, Z.; Yang, G. Comparison of UAV RGB Imagery and Hyperspectral Remote-Sensing Data for Monitoring Winter-Wheat Growth. *Res. Sq.* **2021**. [CrossRef]
60. Cross, M.D.; Scambos, T.; Pacifici, F.; Marshall, W.E. Determining Effective Meter-Scale Image Data and Spectral Vegetation Indices for Tropical Forest Tree Species Differentiation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 2934–2943. [CrossRef]
61. Datt, B. Remote Sensing of Chlorophyll a, Chlorophyll b, Chlorophyll a+b, and Total Carotenoid Content in Eucalyptus Leaves. *Remote Sens. Environ.* **1998**, *66*, 111–121. [CrossRef]
62. Crippen, R. Calculating the vegetation index faster. *Remote Sens. Environ.* **1990**, *34*, 71–73. [CrossRef]
63. Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* **1996**, *58*, 289–298. [CrossRef]
64. Sripada, R.P.; Heiniger, R.W.; White, J.; Meijer, A.D. Aerial Color Infrared Photography for Determining Early In-Season Nitrogen Requirements in Corn. *Agron. J.* **2006**, *98*, 968–977. [CrossRef]
65. Gianelle, D.; Vescovo, L. Determination of green herbage ratio in grasslands using spectral reflectance. Methods and ground measurements. *Int. J. Remote Sens.* **2007**, *28*, 931–942. [CrossRef]
66. Rouse, J.W.; Hass, R.H.; Schell, J.A.; Deering, D.W.; Harlan, J.C. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*; Final Report, RSC 1978-4; Texas A&M University: College Station, TX, USA, 1974.
67. Roujean, J.-L.; Breon, F.-M. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sens. Environ.* **1995**, *51*, 375–384. [CrossRef]
68. Sims, D.A.; Gamon, J.A. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sens. Environ.* **2002**, *81*, 337–354. [CrossRef]
69. Jordan, C.F. Derivation of Leaf-Area Index from Quality of Light on the Forest Floor. *Ecology* **1969**, *50*, 663–666. [CrossRef]
70. Gitelson, A.A. Wide Dynamic Range Vegetation Index for Remote Quantification of Biophysical Characteristics of Vegetation. *J. Plant Physiol.* **2004**, *161*, 165–173. [CrossRef]
71. Kambhatla, N.; Leen, T.K. Dimension Reduction by Local Principal Component Analysis. *Neural Comput.* **1997**, *9*, 1493–1516. [CrossRef]
72. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020.
73. Lê, S.; Josse, J.; Husson, F. FactoMineR: AnRPackage for Multivariate Analysis. *J. Stat. Softw.* **2008**, *25*, 1–18. [CrossRef]
74. ESRI. *ArcGIS PRO: Essential Workflows*; ESRI: Redlands, CA, USA, 2016; Available online: https://community.esri.com/t5/esritraining-documents/arcgis-pro-essential-workflows-course-resources/ta-p/914710 (accessed on 1 May 2021).
75. Feurer, M.; Eggensperger, K.; Falkner, S.; Lindauer, M.; Hutter, F. Auto-Sklearn 2.0: The Next Generation. *arXiv* **2020**, arXiv:2007.04074.
76. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
77. Hutter, F.; Hoos, H.H.; Leyton-Brown, K. Sequential Model-Based Optimization for General Algorithm Configuration. In Proceedings of the Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Rome, Italy, 17–21 January 2011; pp. 507–523. [CrossRef]
78. Suykens, J.; Vandewalle, J. Least Squares Support Vector Machine Classifiers. *Neural Process. Lett.* **1999**, *9*, 293–300. [CrossRef]
79. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
80. Olson, R.S.; Urbanowicz, R.J.; Andrews, P.C.; Lavender, N.A.; Kidd, L.C.; Moore, J.H. Automating Biomedical Data Science Through Tree-Based Pipeline Optimization. In Proceedings of the Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Porto, Portugal, 30 March–1 April 2016; pp. 123–137. [CrossRef]
81. Feurer, M.; Springenberg, J.T.; Hutter, F. Initializing Bayesian Hyperparameter Optimization via Me-ta-Learning. In Proceedings of the Proceedings of the National Conference on Artificial Intelligence, Austin, TX, USA, 25–30 January 2015; Volume 2. Available online: https://ojs.aaai.org/index.php/AAAI/article/view/9354 (accessed on 20 October 2020).
82. Franceschi, L.; Frasconi, P.; Salzo, S.; Grazzi, R.; Pontil, M. Bilevel Programming for Hyperparameter Opti-mization and Meta-Learning. *arXiv* **2018**, arXiv:1806.04910.
83. Ma, L.; Li, M.; Ma, X.; Cheng, L.; Du, P.; Liu, Y. A review of supervised object-based land-cover image classification. *ISPRS J. Photogramm. Remote Sens.* **2017**, *130*, 277–293. [CrossRef]
84. Kumar, A.; McCann, R.; Naughton, J.; Patel, J.M. Model Selection Management Systems: The Next Frontier of Advanced Analytics. *ACM SIGMOD Rec.* **2016**, *44*, 17–22. [CrossRef]
85. Bengfort, B.; Bilbro, R. Yellowbrick: Visualizing the Scikit-Learn Model Selection Process. *J. Open Source Softw.* **2019**, *4*, 1075. [CrossRef]
86. Anwar, M.Z.; Kaleem, Z.; Jamalipour, A. Machine Learning Inspired Sound-Based Amateur Drone Detection for Public Safety Applications. *IEEE Trans. Veh. Technol.* **2019**, *68*, 2526–2534. [CrossRef]
87. Fawcett, T. An introduction to ROC analysis. *Pattern Recognit. Lett.* **2005**, *27*, 861–874. [CrossRef]

88. Pencina, M.J.; Agostino, R.B.D.; Vasan, R.S. Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. *Stat. Med.* **2007**, *27*, 157–172. [CrossRef] [PubMed]
89. Gunčar, G.; Kukar, M.; Notar, M.; Brvar, M.; Černelč, P.; Notar, M.; Notar, M. An application of machine learning to haematological diagnosis. *Sci. Rep.* **2018**, *8*, 1–12. [CrossRef]
90. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* **2009**, *45*, 427–437. [CrossRef]
91. Boyd, K.; Eng, K.H.; Page, C.D. Area under the Precision-Recall Curve: Point Estimates and Confidence In-tervals. In Proceedings of the Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Prague, Czech Republic, 23–27 September 2013; Volume 8190 LNAI.
92. Chicco, D.; Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genom.* **2020**, *21*, 1–13. [CrossRef] [PubMed]
93. Wang, Q.; Ming, Y.; Jin, Z.; Shen, Q.; Liu, D.; Smith, M.J.; Veeramachaneni, K.; Qu, H. ATMSeer: Increasing Transparency and Controllability in Automated Machine Learning. *arXiv* **2019**, arXiv:1902.05009.
94. Ono, J.P.; Castelo, S.; Lopez, R.; Bertini, E.; Freire, J.; Silva, C. PipelineProfiler: A Visual Analytics Tool for the Exploration of AutoML Pipelines. *IEEE Trans. Vis. Comput. Graph.* **2021**, *27*, 390–400. [CrossRef] [PubMed]
95. Serpico, S.B.; D'Inca, M.; Melgani, F.; Moser, G. Comparison of Feature Reduction Techniques for Classification of Hyperspectral Remote Sensing Data. In Proceedings of the Image and Signal Processing for Remote Sensing VIII, Crete, Greece, 13 March 2003; Volume 4885.
96. Saito, T.; Rehmsmeier, M. The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. *PLoS ONE* **2015**, *10*, e0118432. [CrossRef] [PubMed]
97. Fisher, R.A. The use of multiple measurements in taxonomic problems. *Ann. Eugen.* **1936**, *7*, 179–188. [CrossRef]
98. Geurts, P.; Ernst, D.; Wehenkel, L. Extremely randomized trees. *Mach. Learn.* **2006**, *63*, 3–42. [CrossRef]
99. Samaras, S.; Diamantidou, E.; Ataloglou, D.; Sakellariou, N.; Vafeiadis, A.; Magoulianitis, V.; Lalas, A.; Dimou, A.; Zarpalas, D.; Votis, K.; et al. Deep Learning on Multi Sensor Data for Counter UAV Applications—A Systematic Review. *Sensors* **2019**, *19*, 4837. [CrossRef]
100. David, L.C.; Ballado, A.J. Vegetation indices and textures in object-based weed detection from UAV imagery. In Proceedings of the 6th IEEE International Conference on Control System, Computing and Engineering, ICCSCE, Penang, Malaysia, 25–27 November 2016. [CrossRef]
101. Torres-Sánchez, J.; Lopez-Granados, F.; De Castro, A.I.; Peña-Barragan, J.M. Configuration and Specifications of an Unmanned Aerial Vehicle (UAV) for Early Site Specific Weed Management. *PLoS ONE* **2013**, *8*, e58210. [CrossRef]
102. Sankaran, S.; Khot, L.R.; Espinoza, C.Z.; Jarolmasjed, S.; Sathuvalli, V.R.; VanDeMark, G.J.; Miklas, P.N.; Carter, A.; Pumphrey, M.; Knowles, N.R.; et al. Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review. *Eur. J. Agron.* **2015**, *70*, 112–123. [CrossRef]
103. Chawade, A.; Van Ham, J.; Blomquist, H.; Bagge, O.; Alexandersson, E.; Ortiz, R. High-Throughput Field-Phenotyping Tools for Plant Breeding and Precision Agriculture. *Agronomy* **2019**, *9*, 258. [CrossRef]
104. Young, S.N. A Framework for Evaluating Field-Based, High-Throughput Phenotyping Systems: A Meta-Analysis. *Sensors* **2019**, *19*, 3582. [CrossRef]
105. Vivaldini, K.C.T.; Martinelli, T.H.; Guizilini, V.C.; Souza, J.; Oliveira, M.D.; Ramos, F.T.; Wolf, D.F. UAV route planning for active disease classification. *Auton. Robot.* **2018**, *43*, 1137–1153. [CrossRef]
106. Zhang, X.; Han, L.; Dong, Y.; Shi, Y.; Huang, W.; Han, L.; González-Moreno, P.; Ma, H.; Ye, H.; Sobeih, T. A Deep Learning-Based Approach for Automated Yellow Rust Disease Detection from High-Resolution Hyperspectral UAV Images. *Remote Sens.* **2019**, *11*, 1554. [CrossRef]
107. Villoslada, M.; Bergamo, T.; Ward, R.; Burnside, N.; Joyce, C.; Bunce, R.; Sepp, K. Fine scale plant community assessment in coastal meadows using UAV based multispectral data. *Ecol. Indic.* **2020**, *111*, 105979. [CrossRef]
108. Burnside, N.G.; Joyce, C.B.; Puurmann, E.; Scott, D.M. Use of Vegetation Classification and Plant Indicators to Assess Grazing Abandonment in Estonian Coastal Wetlands. *J. Veg. Sci.* **2007**, *18*, 645–654. [CrossRef]
109. Sona, G.; Passoni, D.; Pinto, L.; Pagliari, D.; Masseroni, D.; Ortuani, B.; Facchi, A. UAV multispectral survey to map soil and crop for precision farming applications. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *XLI-B1*, 1023–1029. [CrossRef]
110. Kwak, G.-H.; Park, N.-W. Impact of Texture Information on Crop Classification with Machine Learning and UAV Images. *Appl. Sci.* **2019**, *9*, 643. [CrossRef]
111. Yang, M.-D.; Tseng, H.-H.; Hsu, Y.-C.; Tsai, H.P. Semantic Segmentation Using Deep Learning with Vegetation Indices for Rice Lodging Identification in Multi-date UAV Visible Images. *Remote Sens.* **2020**, *12*, 633. [CrossRef]
112. Yang, M.-D.; Boubin, J.G.; Tsai, H.P.; Tseng, H.-H.; Hsu, Y.-C.; Stewart, C.C. Adaptive autonomous UAV scouting for rice lodging assessment using edge computing with deep learning EDANet. *Comput. Electron. Agric.* **2020**, *179*, 105817. [CrossRef]
113. Najafi, P.; Feizizadeh, B.; Navid, H. A Comparative Approach of Fuzzy Object Based Image Analysis and Machine Learning Techniques Which Are Applied to Crop Residue Cover Mapping by Using Sentinel-2 Satellite and UAV Imagery. *Remote Sens.* **2021**, *13*, 937. [CrossRef]

114. Duro, D.; Franklin, S.; Dubé, M.G. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sens. Environ.* **2012**, *118*, 259–272. [CrossRef]
115. Kim, H.-O.; Yeom, J.-M. A Study on Object-Based Image Analysis Methods for Land Cover Classification in Agricultural Areas. *J. Korean Assoc. Geogr. Inf. Stud.* **2012**, *15*, 26–41. [CrossRef]
116. Telles, T.S.; Reydon, B.P.; Maia, A.G. Effects of no-tillage on agricultural land values in Brazil. *Land Use Policy* **2018**, *76*, 124–129. [CrossRef]
117. Yeom, J.; Jung, J.; Chang, A.; Ashapure, A.; Maeda, M.; Maeda, A.; Landivar, J. Comparison of Vegetation Indices Derived from UAV Data for Differentiation of Tillage Effects in Agriculture. *Remote Sens.* **2019**, *11*, 1548. [CrossRef]
118. Bisong, E. Google AutoML: Cloud Vision. In *Building Machine Learning and Deep Learning Models on Google Cloud Platform*; Springer Nature: Cham, Switzerland, 2019; pp. 581–598. [CrossRef]

III

Li, K. Y., Sampaio de Lima, R., Burnside, N. G., Vahtmäe, E., Kutser, T., Sepp, K., Cabral Pinheiro, V.H., Yang, M.-D., Vain, A. & Sepp, K. (2022). Toward Automated Machine Learning-Based Hyperspectral Image Analysis in Crop Yield and Biomass Estimation. Remote Sensing, 14(5), 1114.

remote sensing

MDPI

Article

Toward Automated Machine Learning-Based Hyperspectral Image Analysis in Crop Yield and Biomass Estimation

Kai-Yun Li [1,*], Raul Sampaio de Lima [1], Niall G. Burnside [2], Ele Vahtmäe [3], Tiit Kutser [3], Karli Sepp [4], Victor Henrique Cabral Pinheiro [5], Ming-Der Yang [6], Ants Vain [1] and Kalev Sepp [1]

1 Institute of Agriculture and Environmental Sciences, Estonian University of Life Sciences, Kreutzwaldi 5, 51006 Tartu, Estonia; raul.sampaio@student.emu.ee (R.S.d.L.); ants.vain1@gmail.com (A.V.); kalev.sepp@emu.ee (K.S.)
2 Centre for Earth Observation, School of Applied Sciences, University of Brighton, Lewes Road, Brighton BN2 4GJ, UK; n.g.burnside@brighton.ac.uk
3 Estonian Marine Institute, University of Tartu, Mäealuse 14, 12618 Tallinn, Estonia; ele.vahtmae@ut.ee (E.V.); tiit.kutser@ut.ee (T.K.)
4 Agricultural Research Center, 4/6 Teaduse St., 75501 Saku, Estonia; karli.sepp@pmk.agri.ee
5 Institution of Computer Science, Faculty of Science and Technology, University of Tartu, 50090 Tartu, Estonia; victor.pinheiro@ut.ee
6 Department of Civil Engineering, and Innovation and Development Center of Sustainable Agriculture, National Chung Hsing University, Taichung 402, Taiwan; mdyang@dragon.nchu.edu.tw
* Correspondence: kai-yun.li@student.emu.ee

check for updates

Citation: Li, K.-Y.; Sampaio de Lima, R.; Burnside, N.G.; Vahtmäe, E.; Kutser, T.; Sepp, K.; Cabral Pinheiro, V.H.; Yang, M.-D.; Vain, A.; Sepp, K. Toward Automated Machine Learning-Based Hyperspectral Image Analysis in Crop Yield and Biomass Estimation. *Remote Sens.* **2022**, *14*, 1114. https://doi.org/10.3390/rs14051114

Academic Editors: Giovanni Avola and Alessandro Matese

Received: 27 January 2022
Accepted: 22 February 2022
Published: 24 February 2022

Abstract: The incorporation of autonomous computation and artificial intelligence (AI) technologies into smart agriculture concepts is becoming an expected scientific procedure. The airborne hyperspectral system with its vast area coverage, high spectral resolution, and varied narrow-band selection is an excellent tool for crop physiological characteristics and yield prediction. However, the extensive and redundant three-dimensional (3D) cube data processing and computation have made the popularization of this tool a challenging task. This research integrated two important open-sourced systems (R and Python) combined with automated hyperspectral narrowband vegetation index calculation and the state-of-the-art AI-based automated machine learning (AutoML) technology to estimate yield and biomass, based on three crop categories (spring wheat, pea and oat mixture, and spring barley with red clover) with multifunctional cultivation practices in northern Europe and Estonia. Our study showed the estimated capacity of the empirical AutoML regression model was significant. The best coefficient of determination (R^2) and normalized root mean square error (NRMSE) for single variety planting wheat were 0.96 and 0.12 respectively; for mixed peas and oats, they were 0.76 and 0.18 in the booting to heading stage, while for mixed legumes and spring barley, they were 0.88 and 0.16 in the reproductive growth stages. In terms of straw mass estimation, R^2 was 0.96, 0.83, and 0.86, and NRMSE was 0.12, 0.24, and 0.33 respectively. This research contributes to, and confirms, the use of the AutoML framework in hyperspectral image analysis to increase implementation flexibility and reduce learning costs under a variety of agricultural resource conditions. It delivers expert yield and straw mass valuation two months in advance before harvest time for decision-makers. This study also highlights that the hyperspectral system provides economic and environmental benefits and will play a critical role in the construction of sustainable and intelligent agriculture techniques in the upcoming years.

Keywords: hyperspectral; automated machine learning; vegetation index; yield estimates; biomass estimation; precision agriculture; narrowband; spring wheat; spring barley; pea and oat

1. Introduction

Fresh trends in precision agriculture (PA) and the development of automated systems for agricultural resource management have been widely explored and deployed in

recent years [1]. The emergence of these techniques seeks to increase crop growth and production, maximize profitability through empirical models and data assimilation, and make a substantial contribution to food security [2,3], agricultural disasters risk management [4], and, more importantly, address concerns relating to climate change mitigation [5]. Image-based remote sensing (RS) technologies are regarded as a vital instrument in this context for providing valuable information that is currently unavailable or inaccurate for achieving sustainable and efficient farming operations [6]. The use of RS technologies provides timely, non-destructive, spatial estimates for measuring and tracking specific vegetation attributes [7], as well as continuing to improve crop yield production and quality, thereby assisting in future food security and reducing the negative impacts of agricultural practices [5,8,9]. Moreover, agriculture management practices based on the concept of sustainable cropping ideas (such as reduced tillage intensity [10–15], fertilizer input [16], and organic farming [17,18]) combined with mixed cropping systems, particularly legume-based, can effectively diminish greenhouse gas emissions by reducing the use of inorganic nitrogen fertilizers and replacing them with symbiotically fixed nitrogen [19], as well as carbon loss [5,20,21] and soil erosion [22] in cultivated soil. Furthermore, they can contribute to productivity and economic appeal to Northern European farmers, which is crucial for ensuring that these ecologically friendly systems can compete in terms of profitability with more traditional or artificially generated systems [23]. Variety performance trials (VPTs) with a well-designed randomized design, for example, are an excellent technique to assess a variety of management procedures and their interactions with the agri-environment [24–26]. However, owing to the variability in the structure, character, and husbandry of each experiment, investigations of VPT datasets can provide diverse outcomes [27].

Despite weather conditions, soil, and management in current trials with rigorous model simulation, the challenge of sampling and model development is exacerbated by landscape heterogeneity [28] and varied spatial distribution patterns of geographical items [29]. To face these challenges, RS technology provides the opportunity to measure biophysical indicators in research sites. In addition to detecting and quantifying their geographical variability, it can potentially play a pivotal role in the provision of time-specific information for decision supporting systems [1,6] and improve operations by making them more cost-effective and time-efficient.

Currently, a primary objective of agronomic remote sensing is to identify those bands of light-spectrum which are most sensitive to canopy reflectance, and the derived parameters that distinguish vegetation features, identify growth status, and quantify the relationships which exist between spectral properties and agronomic parameters [30]. Vegetation indices (VIs) are one of the most extensively utilized precision farming tools for supplying reliable spatial and temporal information on vegetation cover across a variety of agricultural operations. In visible/near-infrared imagery, vegetation has a distinct spectral signature that permits it to be distinguished from other forms of land cover [31]. VIs utilize a mathematical combination from at least two spectral bands of the electromagnetic spectrum, intending to reduce confusing factors (i.e., soil disturbance and other environmental noises) while increasing the importance of plant features [32,33]. As an example, a traditional agricultural yield estimation methodology, such as the Normalized Difference Vegetation Index (NDVI), calculates the difference between the red and near-infrared bands from multispectral sensors and provides a measure of chlorophyll pigmentation. Furthermore, a variety of new indicators were developed in the early years to correct for soil backgrounds and the effects of climatic environments [34–37]. Multi-spectral, broadband-based remote sensing has had longstanding success in established correlations between conventional indices with yield and crop status. However, due to saturation in dense vegetation at larger leaf area index (LAI) values, multilayered canopies, and various farming systems, the calculated indices can occasionally produce inaccurate measurements and pose limits for quantitative estimation of biochemical properties owing to lower spectral resolution [7,38–40].

As an alternative technology, a high-spectral-resolution imaging system (i.e., hyperspectral imaging) creates the opportunity to enable increasingly sophisticated agricultural applications. The necessity for research in identifying optimum wavebands to predict crop biophysical characteristics is vital as hyperspectral remote sensing data becomes ever more available and significant [41,42]. With the use of narrow spectral channels of less than 10 nm, hyperspectral remote sensing data has the potential to identify more nuanced differences in vegetation than multispectral data [43]. It has been suggested that hyperspectral data analysis may present a format to provide a deeper understanding of the mechanisms governing spectral reflectance from field scales and canopy levels [44,45]. These reduced-range channels allow for the detection of detailed plant and crop characteristics that would typically be obscured by broader-band multispectral channels. Innovative approaches for analyzing spectral reflectance data are being established as a result of advances within hyperspectral remote sensing technology [41,46]. Whilst hyperspectral sensors provide a more detailed depiction of plant canopy reflectance than more traditional multispectral sensors, they come with concerns regarding data redundancy and spectral autocorrelation [31,47,48]. In an attempt to redress and resolve these challenges, the reduction of data dimensionality is proposed, which can often be achieved via feature extraction, i.e., translating the spectra to a lower-dimensional representation, or selecting only a subset of essential bands or spectral characteristics for analysis. [49]. One proposed technique to investigate imaging spectroscopy via spectral characteristics is to use application-specific optimal bands' combination, i.e., narrowband VIs. These narrowband VIs have significantly improved crop characteristics and deliver substantially advanced variability information with a superior dynamic range and considerable improvements over broad bands [7]. There is mounting evidence that narrowband VIs can improve biomass estimations for many land-cover types [50]. Recently, a study regarding wheat grain yields also revealed that when compared with broadband VIs, hyperspectral indices provided greater estimation ability of grain production and biophysical factors [42]. As a result of the emergence of hyperspectral systems, there now exists the possibility to both refine previous spectral indices and build novel approaches that make use of the increased spectral resolution of hyperspectral data. Alternatively, the analysis might suggest that narrow-band, continuous reflectance data from a hyperspectral sensor are preferred and potentially more accurate for certain remote sensing applications [31].

Hyperspectral data, when paired with popular machine learning (ML) algorithms, have made a substantial contribution to crop biomass and yield estimation [51–53]. These multimodal computing technologies broaden the application of ML to a wider range of beneficial data collection and selection for the progression of agriculture practices [54] These approaches will contribute to improved decision-making within complex systems, with minimal human interaction, and provide a scalable framework for integrating expert knowledge of the PA system [55]. Complexity can be seen as a disadvantage in crop trials since the ML modelling includes training/testing databases, limited areas with insignificant sampling sizes, time and space-specificity, and environmental factor interventions, which raise problems in parameter selection and make use of a single empirical model for an entire region impractical [56,57]. Instead, the robust artificial intelligence-based notion of automated machine learning (AutoML) has emerged to minimize such data-driven expenses and enables experts to build self-regulating machine learning applications [58,59]. AutoML is characterized as a combination of selecting an algorithm and hyperparameter optimization based on the Bayesian optimization method that seeks to identify the optimum (cross-validated) combination of algorithm components by encompassing data from raw datasets to a deployable pipeline ML model, which greatly simplifies these stages for people with limited expertise [60–62]. For improving the model's prediction performance, the common technique for ML modelling includes data pre-processing, feature and algorithm selection, extraction, and engineering, as well as hyperparameter optimization [63].

However, although AutoML has made significant contributions to computer science and, more recently, remote sensing applications, such as soil moisture monitoring and plant

phenotyping [64,65], it has yet to be broadly adopted in the disciplines of hyperspectral imaging and PA systems. This study used an open-source, cutting-edge Auto-Sklearn algorithm to close the knowledge gap [62]. It is based on the widely-used ML system Scikit-learn platform in Python [66] In addition, the hyperspectral data analysis (*hsdar*) package [67] was utilized in software R [68] to address crop yield and biomass regression tasks. To be more specific, our goals were to use a novel AutoML system to (1) construct an AutoML framework for hyperspectral imaging regression tasks, and (2) explore the applicability of the AutoML models to estimate spring wheat, spring barley, pea and oat mixture grain yields and straw mass in regular mono- or mixed cropping systems in Northern Europe and Estonia. In this study, we presented a comprehensive AutoML infra-structure for a wider range of crop management practices tasks, as well as innovative AutoML- hyperspectral fusion methodologies for future PA and crop phenotyping research.

2. Materials and Methods

2.1. Research Site and Experiment Layout

This experiment was conducted in the Agricultural Research Centre (ARC) in Kuusiku (58°58′52.7″N 24°42′59.1″E), Estonia (Figure 1a), which is the division of the Estonian Ministry of Agriculture. Over 2.1 hectares of the variety performance trial (VPT) area were involved in this study, and the area consisted of two soil types: Calcaric Cambisol and Calcaric-Leptic Regosol [69]. The ARC experimental area had a temperate climate with an average annual temperature of 5.3 °C, where the average daytime temperature was 9.5 °C, and 0.8 °C as night temperature. The annual precipitation was 75 cm. The daily climograph of the study area (see Figure A1) shows precipitation and temperature fluctuations for the crop growing period from April to August 2019. The experimental fields consisted of three commonly cultivated crop categories and their regular cropping combinations in Estonia (Figure 1b), i.e., Field 1: spring wheat (SW) (Figure 1c), as representative of the uniform variety planting field; Field 2: pea and oat mixture (P + O); and Field 3: spring barley with under-sowing red clover (SB + RC) (Figure 1b) as representative of the mixed planting fields. All three fields are part of common crop rotation with a spatial and temporal arrangement.

The experimental strategy was established to aid in the recognition of physiological parameters and comparison of yield abilities of the selected varieties and their combinations under three forms of agriculture management practices (AMP): (1) soil tillage methods (STM); (2) cultivation methods (CM); and (3) manure applications (MA), as well as to demonstrate appropriate farming methods to local farmers. Figure 2 shows the AMPs and their specific arrangement in SW, P + O, and SB + RC fields. Every field comprised 72 plots, with a total of 216 plots. Based on considerations of budget limitations, labor shortages, excessive scope, and repetitiveness, the sampling of grain yield was taken from 56 out of 72 plots (n = 56), and straw biomass was sampled from 24 out of 72 plots (n = 24) specific from the disking and ploughing (DP) area (Figure 2). The harvesting took place on 5 August 2019 in field SB + RC and on 16 August 2019 in fields SW and P + O. The fresh grain and biomass were weighed by plot and dried to verify the dry grain yield and fresh straw mass measured in kilograms per hectare. However, regarding the mixture P + O field, the total weight of the two crops was calculated, while in the SB + RC field only the SB grain yield and straw mass.

2.2. Hyperspectral Image Data Collection

Airborne measurements were carried out in Kuusiku Agricultural Research Centre on 18 June 2019 using hyperspectral imager HySpex (Norsk Elektro Optikk AS (NEO), Oslo, Norway) owned by Estonian Marine Institute and operated by the Estonian Land Board. HySpex was flown at an altitude of 900 m which resulted in a spatial resolution of 40 cm (Figure 1a). The spectral resolution of HySpex is approximately 2.69 nm (216 spectral bands ranging from visible to near-infrared with centers between 409 nm and 989 nm). The day was sunny with a wind speed of 2.6 m/s, average air temperature of 10 °C. Regarding the growth stages of the main crops on the flight date, spring wheat, spring barley, and oat

were approximately in the booting to heading stage. The mixed crops, i.e., field pea and red clover were in the reproductive growth stages and the flowing stage, respectively.

Raw HySpex image data were converted into units of spectral radiance (W m^{-2} nm^{-1} sr^{-1}) using Rad software developed by the NEO. PARGE (Parametric Geocoding, ReSe Applications Schäpfler, University of Zurich) geo-coding software was used for geo correction of the flight lines utilizing accurate altitude and location measurements provided by the GPS/INS unit. The captured Hyspex flight line used in this study is shown in Figure 2. Atmospheric influence at such a low altitude was considered minimal and therefore atmospheric correction was not applied to the imagery.

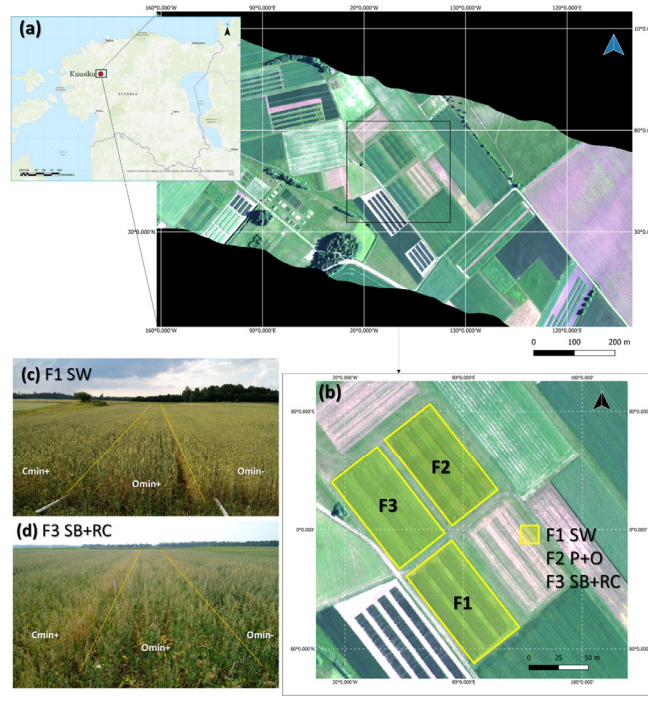

Figure 1. Airborne push-broom hyperspectral image in the Agricultural Research Centre (ARC), Kuusiku, Estonia. (**a**) Hyperspectral image with the band combination: band 83 (630 nm), band 47 (532 nm), and band 22 (465 nm) light in. (**b**) The experiment fields of this study, where Field 1 (F1): spring wheat (SW), Field 2 (F2): pea and oat mixture (P + O), and Field 3 (F3): spring barley with under-sowing red clover (SB + RC). The interpretation diagrams represent on-site (**c**) single variety planting SW, and (**d**) mixed planting SB + RC.

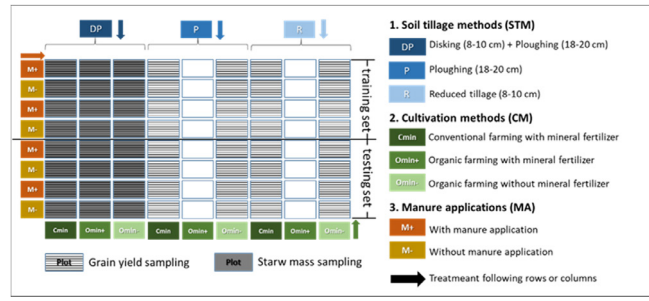

Figure 2. The structure of agriculture management practices (AMPs) and the sampling method of grain yield and straw mass in the SW, P + O, and SB + RC fields. The AMPs contain three treatments: 1. soil tillage method (STM), 2. cultivation method (CM), and manure application (MA), where the grain yield (n = 56) (black striped rectangle box) and straw mass (n = 24) (grey rectangle box). To guarantee that the training area contained all combinations of AMPs, each field was split into training and testing areas equally from the center. The special arrangements of AMP categories and the sampling method were the same in the three fields.

2.3. Hyperspectral Image Processing

Most hyperspectral processing techniques now employ commercial software such as Erdas Imagine, ENVI, or the MATLAB hyperspectral toolbox [70]. These technologies are often expensive and can have limited statistical analysis capabilities. Therefore, we employed a new package that was built on the open-source software R in 2019. The hyperspectral data analysis (*Hsdar*) package incorporates several important hyperspectral capabilities from the *HyperSpec* package [71], with an emphasis on the analysis of large data sets collected in the field for vegetation remote sensing. It is available at https://CRAN.R-project.org/package=hsdar (accessed on 20 July 2021) on the Comprehensive R Archive Network (CRAN).

In our study, hyperspectral data were reconstructed into a class named 'Speclib' to offer a framework for handling huge sets in R. This allows the user to store three-dimensional (3D) cube data together with extra adding information into a matrix. This matrix, together with the wavelength information can then be utilized in the *Hsdar* software and used to manage subsequent calculations. A Savitz-ky-Golay filter (method "*sgolay*") with a length of 15 nm was used in the initial preprocessing stage to reduce noise from the spectra. By fitting a polynomial function to the reflectance data, the filter minimizes noise and removes minor discrepancies between adjacent bands. These noise-reduced hyperspectral data were calculated zonal statistics and converted to a (216 (wavelength bands) multiplied by 216 (plot Shapefile)) table. This table was then subsequently used for preliminary correlation analysis between grain yield and straw mass with the mean wavelength reflectance value into plot level (Figure 3A). The correlation analysis results of each narrowband band can be utilized as a consideration in the following selection of narrowband vegetation indexes.

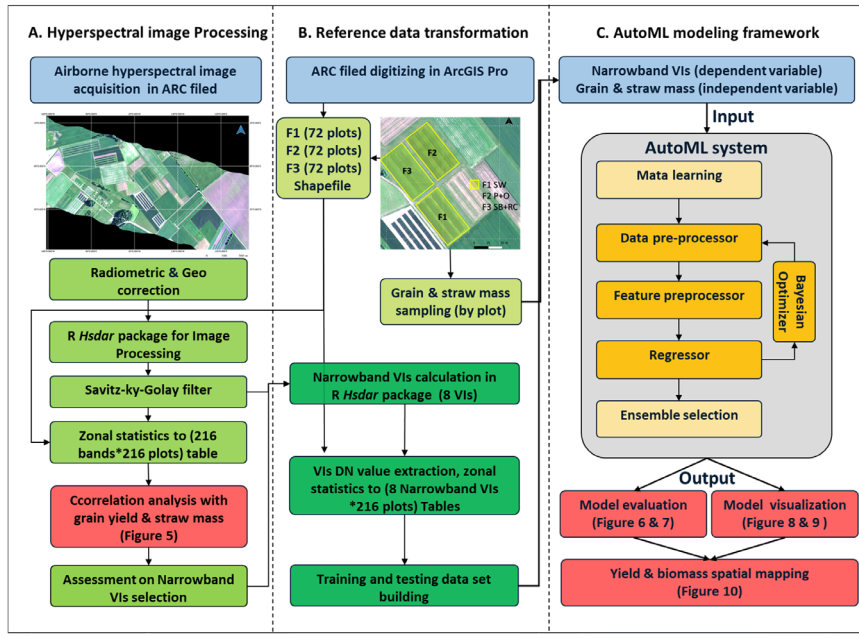

Figure 3. The flowchart of the hyperspectral image processing and AutoML framework was utilized in this study. (**A**) The hyperspectral image processing framework where hyperspectral imager HySpex was conducted and R *Hsdar* package was employed in the processing steps. (**B**) Field reference data transformation, ARC field were digitized based on each field and following AMP treatments. The grain yield and straw mass data were collected according to plots. Eight narrowband VIs were selected and calculated and segmented into corresponding plot digital numbers (DN) for AutoML modelling. (**C**) To achieve robust performance, the Auto-sklearn framework automatically built ML pipelines that were provided by the Bayesian optimization method with warm-started meta-learning and combined with a post hoc ensemble building strategy (Adapted with permission from ref. [62] 2019 Springer).

2.4. Narrowband Vegetation Index

Optical indices for chlorophyll estimation studies have focused on analyzing reflectance in specific narrow bands, ratios, combinations, and the properties of derivative spectra to minimize extraneous factor changes and increase sensitivity to chlorophyll content [6]. In this study, we targeted VIs that were sensitive to canopy structure, biochemistry, and physiology, and those that might potentially indicate variance in grain yields and biomass in our study. Pigments (i.e., chlorophyll a, chlorophyll b, and carotenoids) exhibit varied spectral behavior from an optical standpoint, with specific absorption properties at different wavelengths [72]. Therefore, we employed pre-defined indices in the *Hsdar* R package to automatically fit provided wave-length positions and compute corresponding VIs to reduce the intricacy of computation and boost the repeatability of this research (Table 1).

During our study, the Normalized Difference Vegetation Index (NDVI) was adopted based on it is sensitivity to green leaf area or green leaf biomass, and it can be used to monitor photosynthetically active vegetation biomass distribution using linear combinations of red and infrared radiances [73]. However, it is crucial to note that NDVI has a saturation effect at richer vegetation covers [74]. To solve the probable saturation problem, NDVI2 was applied with its ability to adequately determine chlorophyll in the presence of a high-pigment concentration background [75]. The renormalized difference vegetation index (RDVI) narrow band was employed in this study due to its capacity in identifying mixture phytomass in grassland [76]. The prospect for using the Transformed Chlorophyll Absorption in Reflectance Index (TCARI) in an operational remote sensing situation in the context of precision agriculture was investigated. The R_{700}/R_{670} ratio was chosen to reduce the combined impacts of underlying soil reflectance and non-photosynthetic materials. The changes in reflectance characteristics of background materials (soil and non-photosynthetic components) and the R_{700}/R_{550} ratio are strongly connected to differences in background materials [6,77]. Soil-Adjusted Vegetation Index (SAVI) was conducted to reduce soil-induced fluctuations in vegetations using a transformation approach to decrease soil brightness impacts by counting red and near-infrared wavelengths from spectral data [78]. Optimized Soil-Adjusted Vegetation Index (OSAVI) with two types of reflectance combinations (OSAVI and OSAVI2) was selected for its simplicity of use in the context of deployable observations on agricultural landscapes, as its estimation requires no knowledge of soil optical properties, and it also provided the best results for most crops [79], as well as the distinction of tillage effects in an economically RGB UAV application [80]. In addition, the choice of Simple Ratio (SR) narrow-band indices (R_{515}/R_{550}), different from chlorophyll pigment content detection, was based on its feasibility to predict carotene content on hyperspectral imagery in heterogeneous canopies [81]. Carotenoid concentrations reveal important information about plant physiological state [82], and offering a heterogeneous VI source may improve model predictability and minimize collinearity.

Table 1. Descriptions and formulae of narrowband VIs were utilized in this study. Narrowband VIs were calculated, which were closest to the wavelengths given in the original *Hsdar* R package references.

Vegetation Index	Description	Equation	Reference
NDVI	Normalized Difference Vegetation Index	$(R_{800} - R_{680})/(R_{800} + R_{680})$	[73]
NDVI2	Normalized Difference Vegetation Index 2	$(R_{750} - R_{705})/(R_{750} + R_{705})$	[75]
OSAVI	Optimized Soil Adjusted Vegetation Index	$(1 + 0.16) \times (R_{800} - R_{670})/(R_{800} + R_{670} + 0.16)$	[79]
OSAVI2	Optimized Soil Adjusted Vegetation Index 2	$(1 + 0.16) \times (R_{750} - R_{705})/(R_{750} + R_{705} + 0.16)$	[83]
RDVI	Renormalized Difference Vegetation Index	$(R_{800} - R_{670})/\sqrt{(R_{800} + R_{670})}$	[84]
SR	Simple Ratio	R_{515}/R_{550}	[81]
SAVI	Soil-Adjusted Vegetation Index	$(1 + L\ ^{1}) \times (R_{800} - R_{670})/(R_{800} + R_{670} + L)$	[78]
TCARI	Transformed Chlorophyll Absorption Reflectance Index	$((R_{700} - R_{670}) - 0.2 \times (R_{700} - R_{550}) \times (R_{700}/R_{670})$	[6]

[1] L, a soil brightness adjustment factor (L) established as 0.5 to suit the majority of land cover types for the SAVI index.

These narrowband VIs were computed and saved in TIFF file format (https://www.adobe.io/open/standards/TIFF.html accessed on 15 July 2021), which were then utilized to extract spatial information in the SW, P + O, and SB + RC experimental fields. For extraction, a total of 216 plots were digitized in ArcGIS Pro [85]. Average VIs across every plot were extracted and determined at each plot at the research location, while one-meter buffer zones were calculated inwards from each plot boundary to eliminate unexpected boundary effects. Considering the potential variances in the treatment of each AMP, we divided the field from the center of the area into training and testing areas equally, ensuring that the training area contained all combinations of AMPs (Figure 2). These collected parameters were then utilized in this study to create AutoML algorithms for estimating and evaluating grain production and straw mass.

2.5. AutoML Regression with Auto-Sklearn

This study employed the robust and frequently updated AutoML system, Auto-sklearn, based on the scikit-learn ML library in Python [86]. It employs 15 classifiers, 14 feature processing, and four data pre-processing methods, yielding a 110-hyperparameter structured hypothesis space [62,87]. It offers an advancement on existing AutoML approaches by incorporating prior performance on comparable datasets and generates ensembles from the models that were examined throughout the optimization procedure (Figure 3C). This technique involves the largely configurable ML prototype with the automatically generated ML pipelines, i.e., feature selection (deleting trivial features), transformation (reducing dimensionality), and hyperparameter optimization based on Bayesian optimization strategy sequential model-based algorithm configuration (SMAC) [88]. Following that, a Random Forest [89] approach was utilized for fast cross-validation, assessing one-fold at a time and eliminating poor-performing hyperparameter configurations during the initial phases. The Random Forest approach delivers a superior accuracy rate, as well as alternative pipeline operators that boost regression performance within the datasets [62,90].

All computations in this study were performed on an Intel Core i5-1035G1 CPU (1.00 GHz) with 16 GB RAM utilizing the LINUX open-source operating system. The processes outlined in [62] were executed for the AutoML framework. To begin with, the system employs a supplemental technique based on widely used meta-learning procedures to train machine learning models over the statistical features of datasets and evaluates the model parameters that produce the greatest performance [91]. Second, the system creates ensembles of the models that Bayesian optimization examined, using high-performing regressors and pre-processors employed within the ML framework. Finally, the program works a wide range of empirical examinations on a diverse set of data to determine whether the AutoML regression offers better outcomes than previous regressions. However, any strongly correlated VIs should be eliminated during the feature selection step to avoid the effects of collinearity. Since Auto-sklearn works with low-dimensional optimization issues [92], this step was bypassed in this stage. Table 2 lists the principal AutoML regression parameters employed in this study. To perform tests, as a demonstration of the practicability and efficiency of AutoML model selection, CPU timing for each task was restricted to 30 s, and the runtime for assessing a single model to 10 s. The analyses were performed separately for each of the crop fields, with grain yield consisting of 56 plots (n = 56) and straw mass divided in the training and test sites (0.5/0.5) for regression modelling (Figures 6 and 7).

Table 2. The AutoML regression parameters and descriptions that were employed in this study.

Parameter Name	Range Value	Description
time_left_for_this_task	30 s	The time restriction for seeking suitable models.
per_run_time_limit	10 s	The maximum amount of time a single call to the ML model could perform.
ensemble_size	50 (default)	Several models were added to the ensemble from *Ensemble libraries*.
ensemble_nbest	50 (default)	The amount of best models for building an ensemble model.
resampling_strategy	CV; folds = 3	(CV = cross-validation); to deal with overfitting
seed	47	Used to seed SMAC.
training/testing split	(0.5; 0.5)	Data partitioning way

Note: Other options and parameters that aren't shown in the table were set to default.

2.6. Model Evaluation

The assessment was carried out for the prediction of AutoML models (Figures 6 and 7). Performance evaluation approaches proposed by [19,93] were utilized to evaluate each model. The coefficient of determination (R^2) (Equation (1) and normalized root means square error (NRMSE) (Equation (2)) were used to evaluate the models' accuracy. The following are the equations that were used:

$$R^2 = 1 - \frac{\sum(\hat{y}_i - y_i)^2}{\sum(y_i - \overline{y})^2} \quad (1)$$

$$NRMSE = \frac{\sqrt{\sum((\hat{y}_i - y_i)^2)/n}}{\Delta y} \quad (2)$$

where: y_i is the training dataset's ith observation value represents the observation value; $\overline{y}$ denotes the training dataset's mean value; $\hat{y}_i$ denotes the model predictions, n denotes the number of observations; and Δy represents the difference between the training dataset's lowest and highest values.

3. Results

3.1. The Field Observation DM Data Analysis

The average actual grain yield and above ground straw mass data (fresh and dry) gathered from the SW, P + O, and SB + RC experimental regions are displayed in the violin plot (Figure 4), where we exhibited the range of grain yield and straw mass data and assembled them by fields since the treatments were interspersed within each plot. In addition, we opted to examine at dry and fresh weight separately since the accumulated rainfall of 4.1 mm (in SW and P + O fields) and 0.4 mm (in SB + RC fields) in the three days before the two harvests (on 16 August 2019, and 5 August 2019, respectively) may have contributed to increased fresh weight with additional water content.

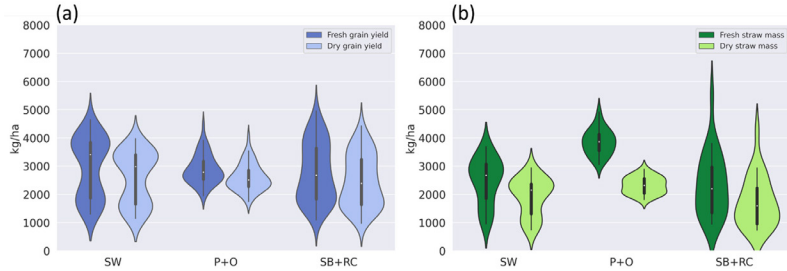

Figure 4. Violin plots of mean harvest results of fresh and dry (**a**) grain yield and (**b**) straw mass, grouped by spring wheat (SW), pea and oat mixture (P + O), and spring barley with under-sowing red clover (SB + RC) fields. White dots represent the median, while thick black bars in the center demonstrate interquartile ranges, and black lines represent the remainder of the distribution. The shape of the violins shows point density and data distribution as a whole.

3.2. The Hyperspectral Reflectance Signature under Various Agriculture Management Practises

Figure 4 displays a mean reflectance plot produced from hyperspectral data of SW, P + O, and SB + RC fields, with enclosed subsets categorized by (Figure 5A) soil tillage method (STM) and (Figure 5B) cultivation method (CM) agricultural operations. Regarding agricultural management practices, the wavelength bands between 700–750 nm and 760–900 nm had significant identification capabilities, while the 400–700 nm region showed little differentiation between management practices. The cultivation method (Figure 4B) provides greater recognition ability (separation) in this range when compared with STM spectral information (Figure 5A). In terms of crop types, spring wheat monocropping seemed to give a better ability to recognize AMPs, followed by mixed cropping systems SB + RC and P + O fields. However, since the focus of this study was on grain yield and biomass prediction, we omitted the narrowband VIs wave range based on the strong absorption bands near 760 nm and excluded them from subsequent AutoML analyses.

3.3. Characterization of the Correlation Coefficient with Averaged Radiance Hyperspectral Data and Field Observation

Correlation Coefficient (r) was used as exploratory analysis in our study and as a reference for subsequent modelling. Figure 5 shows the correlation coefficients (r) between each averaged hyperspectral narrow-band data with the dry mass (Figure 6A) and fresh mass (Figure 6B) at the plot level. The pattern of positive r values was typically obtained with reflectance between 750–940 nm wavelengths, whereas the strong negative correlation with reflectance was between 500–700 nm. Moreover, we also observed that the correlation of straw mass (red line) was stronger than grain yield (blue line) at all fields in the 750–1000 nm range. By comparison, the results showed that, in the patterning of r curves, SW was closely associated with highly positive and negative r values in dry mass (Figure 6A), while with the lower correlation nearby the oxygen absorption peak was 760 nm. This tendency was observed in our previous reflectance signature analysis as well. Among the three fields, P + O had the least correlation. Regarding the fresh mass (Figure 5B), the correlation and spectral characteristics were comparable to the weight of the dry mass. Except for 740–750 nm, SW had overall the strongest correlation, followed by SB + RC and P + O.

3.4. The AutoML Model Prediction and Evaluation

In this study, the narrowband VIs reflectance of grain yield (n = 56) and straw mass (n = 24) based on training/testing (0.5/0.5) principles were used for AutoML modelling, respectively. The AutoML framework was used to test the appropriate combinations of data set parameters throughout the modelling process. Scatter plots representing model predictions and observed weight values (kg ha^{-1}) were compared to the coefficient of the determination (R^2) and normalized root means square error (NRMSE) along with the 1:1 line.

Figure 7 shows the regression plots of fresh (Figure 7A) grain yield (kg ha^{-1}) and (Figure 7B) straw mass (kg ha^{-1}) in SW, P + O, and SB + RC fields based on narrowband VIs and AutoML methods. The results indicated that, in fresh grain yield (Figure 7A), the AutoML model had the lowest prediction errors (NRMSE = 0.13) and the highest R^2 value (0.95) in SW field, followed by SB + RC field (NRMSE = 0.16, R^2 = 0.88) and P + O (NRMSE = 0.16, R^2 = 0.88). Even though the three models functioned well, there was a minor non-uniform bias found within the models, with an underestimation of grain yields in areas with higher output in SW and SB + RC fields. On the other hand, for fresh straw mass, the SW field remains the best performing among the other fields with (NRMSE = 0.16, R^2 = 0.88) followed by the SB + RC field (NRMSE = 0.27, R^2 = 0.77) with uniform overestimation bias, and P + O (NRMSE = 0.25, R^2 = 0.56) (Figure 7B). Among them, P + O's prediction ability was insufficient, and the reference data collected were concentrated in the 3000 to 5000 (kg ha^{-1}) interval, which makes the regression model unable to be effectively extended.

Figure 8 demonstrates the behavior of predictive models utilizing dry (A) grain yield (kg ha^{-1}) and (B) straw mass (kg ha^{-1}) in SW, P + O, and SB + RC fields based on narrowband VIs and AutoML methods. The results specified that, in summary, SW yielded the best performance for dry grain yield (NRMSE = 0.12, R^2 = 0.96) and straw mass (NRMSE = 0.15, R^2 = 0.89) among SB + RC, and P + O files (Figure 8A). Compared with the fresh mass model, the dry performance was better in general, especially in the dry straw model of SB + RC (NRMSE = 0.33, R^2 = 0.86) and P + O (NRMSE = 0.24, R^2 = 0.83) (Figure 8B), although these two models had a larger degree of bias under the comparison of 1:1 slope.

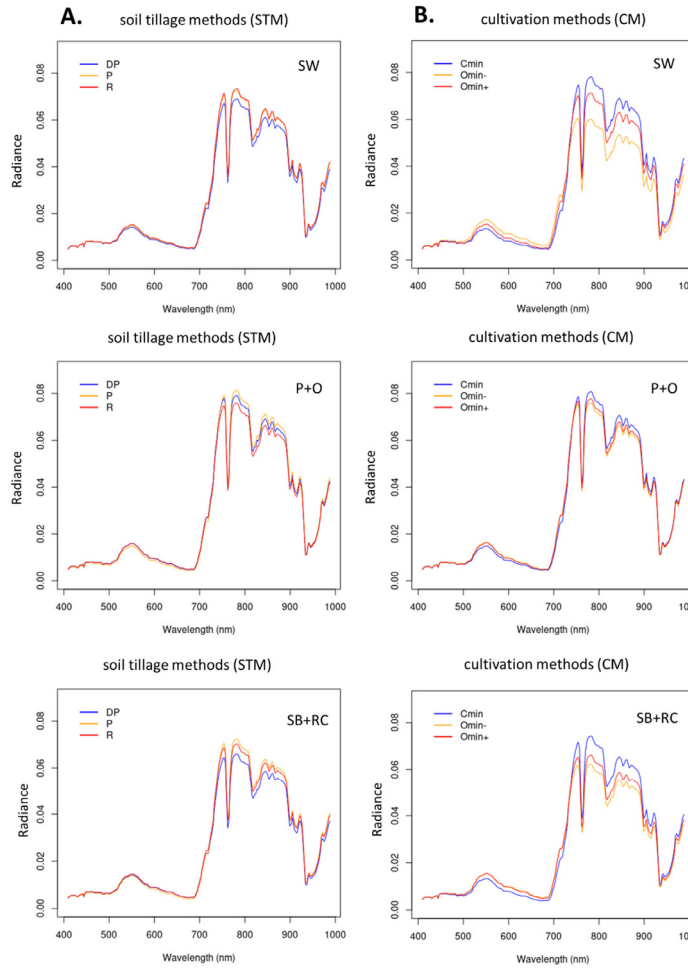

Figure 5. Mean radiance plot derived from hyperspectral data of spring wheat (SW), pea and oat mixture (P + O), and spring barley with under-sowing red clover (SB + RC) fields, grouped by (**A**) soil tillage method (STM) and (**B**) cultivation method (CM) farming operations with contained subsets. The wavelength ranges from the visible to near-infrared (VNIR, 400–1000 nm).

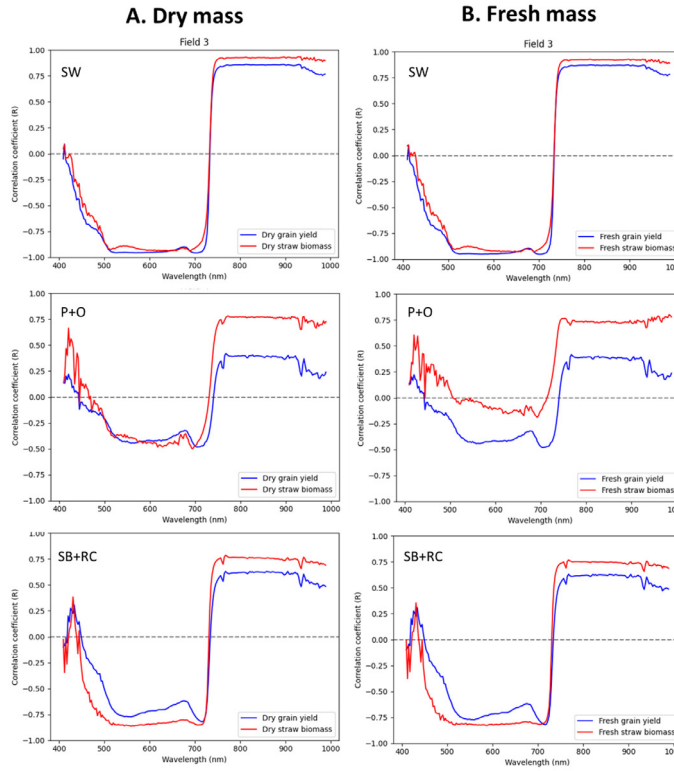

Figure 6. The Pearson Correlation Coefficient (r) between the field observation value (grain yield (**A**); Straw mass (**B**)) and averaged hyperspectral radiance at the plot level in SW, P + O, and SB + RC region.

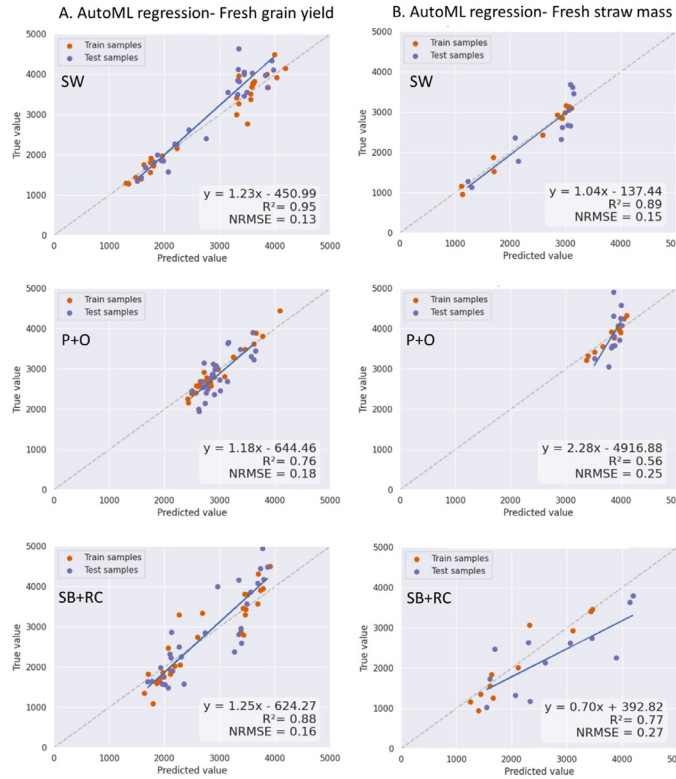

Figure 7. Regression plots of (**A**) fresh grain yield (kg ha^{-1}) and (**B**) fresh straw mass (kg ha^{-1}) in SW, P + O, and SB + RC fields based on narrowband VIs and AutoML methods. The horizontal axis in the scatter plots represents the model's projected grain yield or straw mass, while the vertical axis represents field-observed data. Where the R^2 = coefficient of determination, NRMSE = normalized root represents the squared error, while the 1:1 slope is shown by the black dotted line.

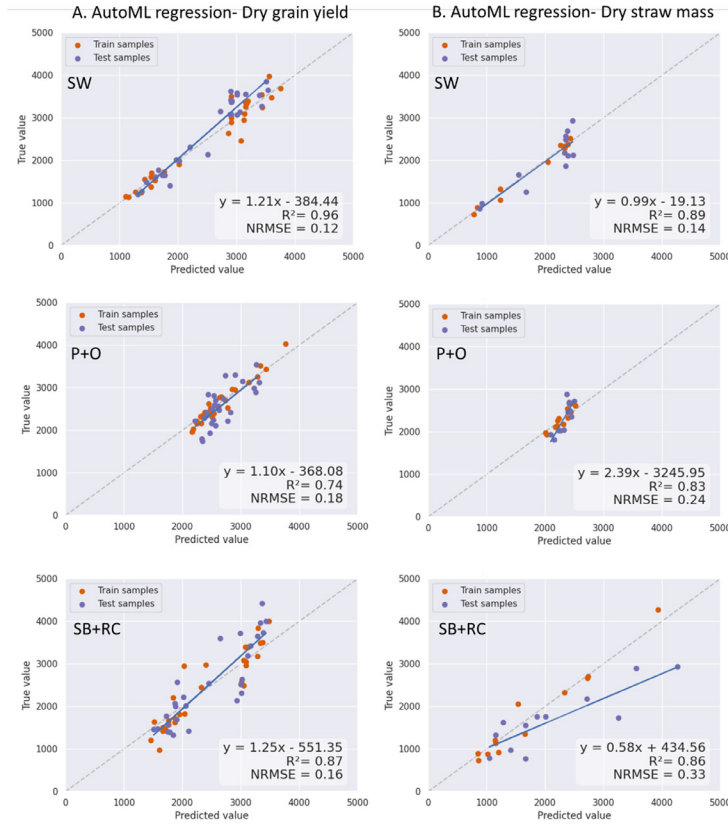

Figure 8. Regression plots of (**A**) dry grain yield (kg ha^{-1}) and (**B**) dry straw mass (kg ha^{-1}) in SW, P + O, and SB + RC fields based on narrowband VIs and AutoML methods. The horizontal axis in the scatter plots represents the model's projected grain yield or straw mass, while the vertical axis represents field-observed data. Where R^2 = coefficient of determination, NRMSE = normalized root means squared error, and the black dotted line exemplifies the 1:1 slope.

3.5. The AutoML Model Pipeline Visualization

An interactive AutoML visualization tool PipelineProfiler [94] was used in this study (Figure 9). To simplify the description, we only list the best regression modelling results across two crop fields (SW and SB + RC) with the evaluation performance of AutoML pipeline execution times set at 30 s, the primitive comparison against other regressors in the same pipeline, and real-time hyperparameter selections. The results confirmed that the best regressor found for dry grain yield was automatic relevance determination (Ard) Regression [95] for the SW field (Figure 9A), and for the SB + RC field, it was the Random Forest [89] (Figure 10A), while for dry straw mass, it was Gaussian Process [96]

(Figure 9B) for the SW field, and Ard Regression for the SB + RC field (Figure 10B), with all hyperparameters found by AutoML also displayed in the figures.

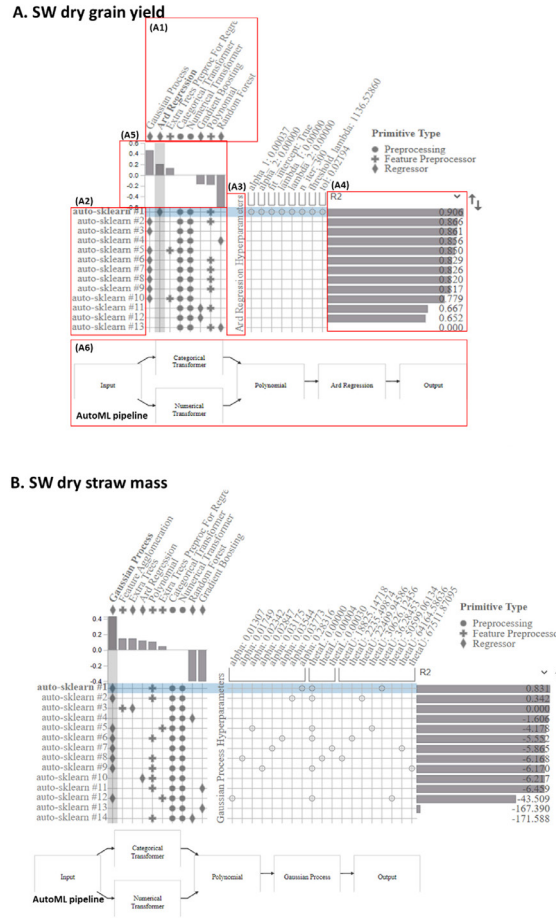

Figure 9. The interactive AutoML pipeline matrix plots with thirty-second running-time limits sorted by coefficient of determination (R^2) performance (**A**,**B**). (**A**) Spring wheat (SW) dry grain yield pipeline matrix with the Top1 regressor, automatic relevance determination (Ard) regression, where (**A1**) illustrated Primitives (in columns) used by the pipelines (**A2**) the blue line (in rows) showed the best R^2 rank); (**A3**) one-hot-encoded hyperparameters (in columns) for the primitive across pipelines, (**A4**) R^2 performance ranking of AutoML pipelines; (**A5**) Primitive contribution view demonstrating the correlations between pipeline scores and primitive usage are displayed in A5. The Gaussian

Process showed the highest correlation score regarding R^2 performance; (**A6**) step-by-step AutoML Pipeline algorithm flowchart, where the box before the output represents the regressor of the model (in A6 Ard regression as the regressor). (**B**) Spring wheat (SW) dry straw mass field pipeline matrix with the Top1 regressor, Gaussian Process.

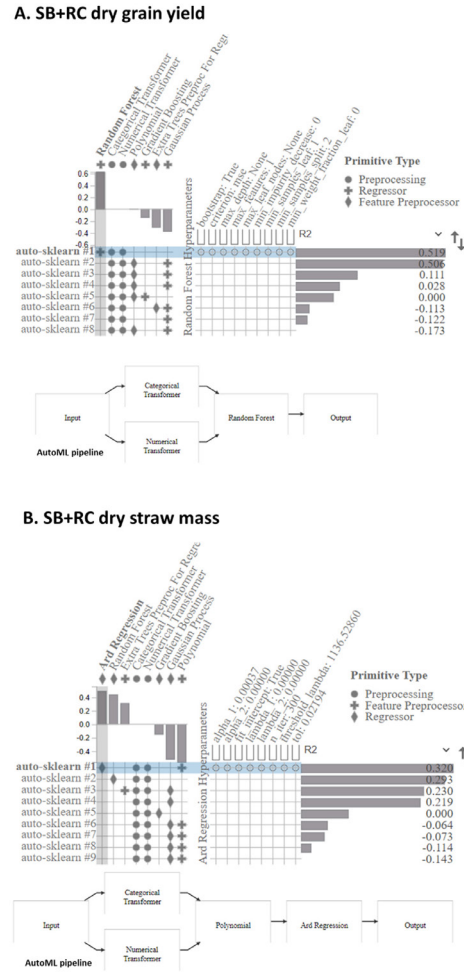

Figure 10. The interactive AutoML pipeline matrix plots with thirty-second running-time limits sorted by coefficient of determination (R^2) performance (**A**,**B**). (**A**) spring barley with under-sowing

red clover (SB + RC) dry grain yield pipeline matrix with the Top1 regressor, Random Forest. The rows display a blue line representing the best R^2 rank followed by its hyperparameters settings; (**B**) SB + RC dry straw mass pipeline matrix with the Top1 regressor, Ard regression, followed by its hyperparameters settings.

3.6. The Field Observation DM Data Analysis

Based on the AutoML models provided above (Figures 7 and 8), a series of prediction maps were generated (Figure 11) for dry grain yield and straw mass for SW, P + O, and SB + RC experimental sites at the plot level. Furthermore, the SW and P + O fields' prediction capabilities were 60 days before the harvest date (18 June–16 August), whereas the SB + RC field's estimating was 49 days before harvest (18 June–5 August).

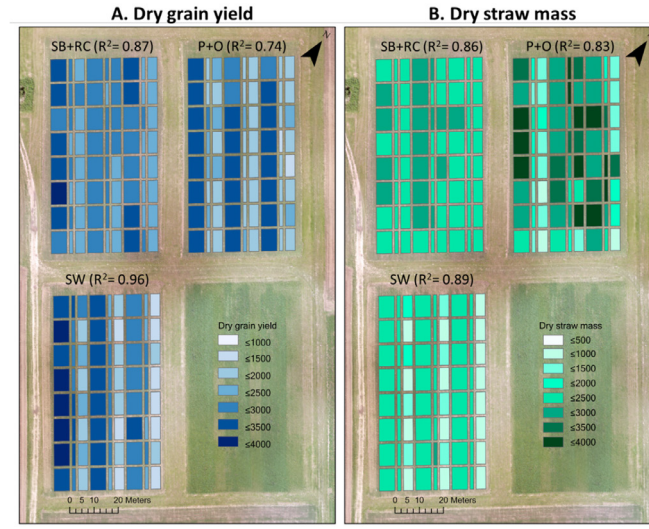

Figure 11. The spatial prediction mapping output of (**A**) dry grain yield (kg ha^{-1}) and (**B**) dry straw mass (kg ha^{-1}) in SW, P + O, and SB + RC fields based on their respective AutoML prediction models at the plot level. The performing coefficient of determination (R^2) is displayed in the previous results.

4. Discussion

This research demonstrated an automatic, open-sourced, rapid, and non-destructive framework by using hyperspectral narrow-band vegetation indexes under regular mono- and mixed cultivation for crop grain yield and straw mass modelling. Since the investigation was carried out under a diversity of agricultural management practices, the methods and findings can profoundly aid agronomists and farmers in designing accurate cropping systems to enhance environmental assessment.

4.1. The Effect of Hyperspectral Signatures and the Correlation between Crop Yield and Straw Mass

The initial goal of this study was to conduct an exploratory evaluation of the hyperspectral reflectance signature and determine the ideal narrowband VIs for modelling common crop types and farming schedules in Northern Europe. To identify redundant

bands and establish wavebands that could best help AutoML regression modelling, the VIs were first chosen based on prior knowledge of the literature and then filtered by the reflectance signature (Figure 5) and their Correlation Coefficients with yield and biomass (Figure 6). Although there was no general focus on a formal classification analysis in our current study, the characteristics of hyperspectral data under different agricultural practices (i.e., STM, CM, and MA) are still worthy of attention.

Figure 5 reveals that, in general, because chlorophyll absorption is not limited to the center wavelength but also affects adjacent bands, we can see that reflectance values in the blue and red sections are significantly reduced, resulting in "absorption characteristics" in the spectral signature of the reflectance in all spectral results. In addition, all the reflection spectra showed obvious absorption peaks at 760 nm. This spectral region is influenced by atmospheric oxygen [97,98] and, therefore, this region was avoided while calculating VI's. Additionally, from the results, the wavelength range 750–900 nm (NIR) had strong recognition capabilities based on the variation of reflection intensity; however, the 400–700 nm (visible bands) region was inefficient and offered little separation or discernment. The differentiation on spectra at the wavelength range of 750–900 nm suggested that the interior leaf structure, biochemical concentration, and water content of the target vegetation are different. A previous study pointed out that the diversity of NIR regions is usually caused by differences in internal leaf structure [99]. However, reflectance variation at the canopy level may be due to additional factors like LAI, canopy design, and backdrop soil [100]. These results will be valuable for further classification activities in agriculture management recognition.

The coefficients correlation (r) of each narrow-band with both grain yield and straw mass exhibited a similar pattern of r curves for both dry (Figure 6A) and fresh weight (Figure 6B) analysis, yet r in absolute values for the P + O field was observed to be less correlated than those for grain yield and straw mass, especially in the fresh weight. This is because the P + O field was mixed cultivation and the source of weight is the sum of the two crops and the amount of precipitation before harvesting may indirectly bring about a lower degree of correlation. Interestingly, while the findings of these linear correlation tests all showed that the straw mass has a stronger link with the spectrum, it does not depend on the empirical model's degree of fit (see Figures 7 and 8). Hence, we discovered that grain yield (R^2) had a superior goodness-of-fit performance than straw mass in general, with lower NRMSE.

4.2. The Hyperspectral Narrowband VIs and AutoML Modelling

Despite the opportunities afforded by hyperspectral systems to collect a multitude of spectrum data, extracting the relevant important wavelengths from a data cube can be challenging [101]. In our study, we used hyperspectral narrowband VIs as predictors for AutoML modelling. However, we avoided selecting narrowband VIs with spectrums that might be affected by atmospheric oxygen. With this in mind, the target VIs selected for analysis were extracted, calculated, and processed in the modelling stage, which reduced processing and storage demands.

Based on the empirical AutoML regression model, the estimation capacity of hyperspectral narrowband VIs was exceptional. The best coefficient of determination for mono-cultivated wheat was 0.96, for mixed peas and oats was 0.76, and for mixed legumes and spring barley was 0.88. In terms of straw mass estimation, they were 0.98, 0.83, and 0.86 respectively. We determined that the prediction ability of dry weight was typically greater than that of fresh weight, especially in fields where mixed peas and oats, which was 27 per cent higher. This demonstrated that the crop water content has an influence on the model's estimation outputs to a certain extent.

According to a previous study, spectral measurements were taken during the Tillering II and Heading phases in wheat yielded the best results for estimating biophysical factors using narrowband VIs [42]. This is consistent with our recommended flight time. In addition, different band combinations can be effectively utilized since crop circumstances

change according to factors such as management conditions and soil characteristics. Others have demonstrated that piecewise multiple regression models on narrow bands provide for greater flexibility in selecting the bands that provide the most information at a given stage of crop development [102]. This viewpoint has also been confirmed in our research.

4.3. The AutoML Method's Applicability and Impact in Hyperspectral Imaging

In this study, we employed an AutoML framework to assist in self-regulating, instinctive regression operations, as well as enhancing challenging hyperparameter adjustments. This method advances the use of hyperspectral imaging in farm-scaled environmental and crop phenotypic activity and possesses several advantages.

Firstly, the flexibility of implementation. With the ever-increasing variability of remote sensing systems and the requirement for empirical model choices, the constraints of adjusting unidentified background parameters are being addressed. This means that many existing models that have been under-optimized in the past now have the chance to be re-modelled using artificial intelligence-based machines to relearn the performance tasks.

Secondly, the alleviation of learning costs. Experience tells us that computer learning for remote-sensed images frequently necessitates a large number of samples and a lengthy learning period, i.e., deep learning [103–105]. This is incompatible with conventional agricultural experimental sampling procedures, which are limited by personnel, the complexity of the experiment design, and the number of repetitions. While, AutoML practices the Random Forest (RF) method [89] for fast cross-validation, testing one-fold at a time and weeding out underperformance hyperparameter choices, for example, the combined algorithm selection and hyperparameter optimization (CASH) problems [62]. It boasts novel pipeline operators that increase the goodness of fit of datasets significantly. The RF approach is well-known for assessing lower sample sizes and increasing the performance of small datasets. [89,106]. In addition, the AutoML framework quickly provided promising regressors and hyperparameter selections. In our research, each run of the regression model only took thirty seconds of learning time. This considerably improves learning efficiency, the ability to find an appropriate formula in the time allotted, and reduces the requirement for machine learning expertise [87,107].

Thirdly, the capacity of innovation. It is noticeable that random forest (RF), support vector machine (SVM), and artificial neural network (ANN) algorithms are among the most widely employed ML techniques in a wide range of recent remote sensing-based studies [108]. Their practicality and performance have been confirmed by many, but equally, there are still other similarly applicable ML methods that may have been shelved. As shown in Figures 9 and 10, the Ard regressors [109,110] and Gaussian Processors [96] were chosen as the best regressors for the grain yield and biomass tasks. These algorithms have received less attention and reference in remote sensing studies. These results indicated that AutoML can uncover alternative ML methods that would otherwise be overlooked by investigators when working with regression subjects.

4.4. The Limitations in This Study

The location, soil types, chosen crop categories, and varieties present may be restricted in this study. In addition, it is important to note that we did not address yield comparisons under different agricultural management approaches since the intricacy of the experimental design may have led to inadequate sampling numbers, as well as possible interaction effects. However, we have presented a framework that can be applied to numerous test regions and the necessity to moderately reduce the number of samples by using AutoML. In addition, due to the limits of the current Auto-Sklearn system, not all regressors performed could be backtracked in our research to explore the individual feature importance ranking, which limits their capacity to aid in the selection of suitable VIs. However, our attempts to provide a wide range of continuous and selectable narrow-band spectral information (over 216 spectral bands) resulted in improved performance.

5. Conclusions

Our study highlights the capability of hyperspectral analysis for yield and biomass prediction in complex design fields through the use of two significant open-sourced software systems: the R language hyperspectral processing package and Python's Auto-Sklearn machine learning technology. The performance evaluation with several types of hyperspectral vegetation indicators we employed to characterize crop production and straw mass was satisfactory. We suggest they can be further applied to other crop biophysical characteristics. The VIs we suggest, as well as automatic narrowband VI calculation, might minimize data redundancy and cleaning time, as well as the computational power hardware requirements. It is also envisaged that further agricultural cultivation practices could be classified using hyperspectral imaging in the NIR spectral region (750–900 nm) with considerable discernible changes in reflectance spectra.

However, the aerial hyperspectral platform utilized in this study may be less cost-effective than fixed-wing or rotary-wing drone systems, which may be more viable for farm-scale exploration. Comprehensive and contemporaneous phenotypic information of products under various agri-environment schemes, as well as their field-based biochemical conditions, reminds us of further challenges which likely exist for remote sensing technology to overcome. Nevertheless, hyperspectral imaging combined with complementary modelling precision, the abundance of spectrum selection flexibility, and extensive flight coverage still have an important role at this stage.

In conclusion, our research focused on the integration and implementation of the hyperspectral imaging and AutoML framework approach with various crop types under multifunctional agriculture management fields in response to crop biomass/yield estimation. Under common crops and cultivation in most Nordic countries, it will provide agricultural decision-makers with professional yield estimation and sustainable agricultural management advice. The study also revealed that the anticipated yield may be advanced two months before harvest. That is, spring wheat, spring barley, and oat were approximately in the booting to heading stage, field pea was around the reproductive growth stages, and the red clover field was in the flowering stage (49 days before in our case). The emergence of the AutoML system has helped to increase the application and effectiveness of remote sensing-based data analysis technology. However, more research and experiments will be required in the future to advance and validate the automatic learning framework's true potential and usage.

Author Contributions: Conceptualization, K.-Y.L.; methodology, K.-Y.L. and R.S.d.L.; software, K.-Y.L., R.S.d.L., E.V. and V.H.C.P.; validation, N.G.B.; formal analysis, K.-Y.L. and V.H.C.P.; investigation, R.S.d.L., E.V. and K.S. (Karli Sepp); resources, T.K. and K.S. (Kalev Sepp); data curation, K.-Y.L., R.S.d.L., E.V., T.K. and K.S. (Karli Sepp); writing—original draft preparation, K.-Y.L.; writing—review and editing, N.G.B., T.K., K.S. (Karli Sepp), V.H.C.P., M.-D.Y., A.V. and K.S. (Kalev Sepp); supervision, N.G.B., T.K. and K.S. (Kalev Sepp); funding acquisition, K.S. (Kalev Sepp). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the European Regional Development Fund within the Estonian National Programme for Addressing Socio-Economic Challenges through R&D (RITA): L180283PKKK, Estonian Research Council grant PUT PRG302, Estonian IT Academy financed by European Social Fund, and the Doctoral School of Earth Sciences and Ecology, financed by the European Union, European Regional Development Fund (Estonian University of Life Sciences ASTRA project "Value-chain based bio-economy").

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the data size.

Acknowledgments: We give respect and gratitude to the scikit-learn framework's developers and maintainers, as well as the Auto-sklearn interface developed at the University of Freiberg. This work was also supported by the Kuusiku Variety Testing Centre of Agricultural Research Centre in Estonia, and the Estonian IT Academy (English brand name StudyITin.ee) which has been financed by the European Social Fund.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

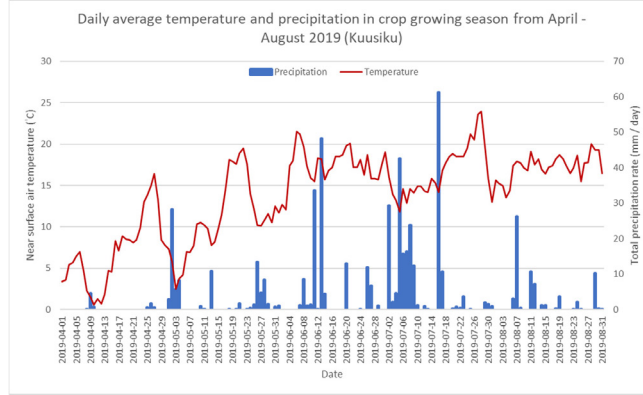

Figure A1. Daily climograph of the study area (Kuusiku) during the crop growing period from April to August in 2019. The blue bars and the red line represent the daily average of rainfall and temperature, respectively.

References

1. Pavón-Pulido, N.; López-Riquelme, J.A.; Torres, R.; Morais, R.; Pastor, J.A. New Trends in Precision Agriculture: A Novel Cloud-Based System for Enabling Data Storage and Agricultural Task Planning and Automation. *Precis. Agric.* **2017**, *18*, 1038–1068. [CrossRef]
2. Karthikeyan, L.; Chawla, I.; Mishra, A.K. A Review of Remote Sensing Applications in Agriculture for Food Security: Crop Growth and Yield, Irrigation, and Crop Losses. *J. Hydrol.* **2020**, *586*, 124905. [CrossRef]
3. Wen, W.; Timmermans, J.; Chen, Q.; van Bodegom, P.M. A Review of Remote Sensing Challenges for Food Security with Respect to Salinity and Drought Threats. *Remote Sens.* **2021**, *13*, 6. [CrossRef]
4. Yang, M.D.; Huang, K.S.; Kuo, Y.H.; Tsai, H.P.; Lin, L.M. Spatial and Spectral Hybrid Image Classification for Rice Lodging Assessment through UAV Imagery. *Remote Sens.* **2017**, *9*, 583. [CrossRef]
5. Mandal, A.; Majumder, A.; Dhaliwal, S.S.; Toor, A.S.; Mani, P.K.; Naresh, R.K.; Gupta, R.K.; Mitran, T. Impact of Agricultural Management Practices on Soil Carbon Sequestration and Its Monitoring through Simulation Models and Remote Sensing Techniques: A Review. *Crit. Rev. Environ. Sci. Technol.* **2020**, *52*, 1–49. [CrossRef]
6. Haboudane, D.; Miller, J.R.; Tremblay, N.; Zarco-Tejada, P.J.; Dextraze, L. Integrated Narrow-Band Vegetation Indices for Prediction of Crop Chlorophyll Content for Application to Precision Agriculture. *Remote Sens. Environ.* **2002**, *81*, 416–426. [CrossRef]
7. Sahoo, R.N.; Ray, S.S.; Manjunath, K.R. Hyperspectral Remote Sensing of Agriculture. *Curr. Sci.* **2015**, *108*, 848–859. [CrossRef]
8. Windfuhr, M.; Jonsén, J. *Food Sovereignty towards Democracy in Localized Food Systems*; FAO: Rome, Italy, 2005; Volume 339.
9. Gómez, D.; Salvador, P.; Sanz, J.; Casanova, J.L. Potato Yield Prediction Using Machine Learning Techniques and Sentinel 2 Data. *Remote Sens.* **2019**, *11*, 1745. [CrossRef]
10. Triplett, G.B.; Dick, W.A. No-Tillage Crop Production: A Revolution in Agriculture! *Agron. J.* **2008**, *100*, 153–165. [CrossRef]
11. Karlen, D.L.; Kovar, J.L.; Cambardella, C.A.; Colvin, T.S. Thirty-Year Tillage Effects on Crop Yield and Soil Fertility Indicators. *Soil Tillage Res.* **2013**, *130*, 24–41. [CrossRef]
12. Ashapure, A.; Jung, J.; Yeom, J.; Chang, A.; Maeda, M.; Maeda, A.; Landivar, J. A Novel Framework to Detect Conventional Tillage and No-Tillage Cropping System Effect on Cotton Growth and Development Using Multi-Temporal UAS Data. *ISPRS J. Photogramm. Remote Sens.* **2019**, *152*, 49–64. [CrossRef]
13. Telles, T.S.; Reydon, B.P.; Maia, A.G. Effects of No-Tillage on Agricultural Land Values in Brazil. *Land Use Policy* **2018**, *76*, 124–129. [CrossRef]

14. Fanigliulo, R.; Antonucci, F.; Figorilli, S.; Pochi, D.; Pallottino, F.; Fornaciari, L.; Grilli, R.; Costa, C. Light Drone-Based Application to Assess Soil Tillage Quality Parameters. *Sensors* **2020**, *20*, 728. [CrossRef] [PubMed]
15. Desta, B.T.; Gezahegn, A.M.; Tesema, S.E. Impacts of Tillage Practice on the Productivity of Durum Wheat in Ethiopia. *Cogent Food Agric.* **2021**, *7*, 1869382. [CrossRef]
16. Crews, T.E.; Peoples, M.B. Legume versus Fertilizer Sources of Nitrogen: Ecological Tradeoffs and Human Needs. *Agric. Ecosyst. Environ.* **2004**, *102*, 279–297. [CrossRef]
17. Zikeli, S.; Gruber, S.; Teufel, C.F.; Hartung, K.; Claupein, W. Effects of Reduced Tillage on Crop Yield, Plant Available Nutrients and Soil Organic Matter in a 12-Year Long-Term Trial under Organic Management. *Sustainability* **2013**, *5*, 3876–3894. [CrossRef]
18. Yang, X.M.; Drury, C.F.; Reynolds, W.D.; Reeb, M.D. Legume Cover Crops Provide Nitrogen to Corn during a Three-Year Transition to Organic Cropping. *Agron. J.* **2019**, *111*, 3253–3264. [CrossRef]
19. Li, K.Y.; Burnside, N.G.; de Lima, R.S.; Peciña, M.V.; Sepp, K.; der Yang, M.; Raet, J.; Vain, A.; Selge, A.; Sepp, K. The Application of an Unmanned Aerial System and Machine Learning Techniques for Red Clover-Grass Mixture Yield Esti-Mation under Variety Performance Trials. *Remote Sens.* **2021**, *13*, 1994. [CrossRef]
20. Loide, V. The Results of an NPK-Fertilisation Trial of Long-Term Crop Rotation on Carbonate-Rich Soil in Estonia. *Acta Agric. Scand. Sect. B Soil Plant Sci.* **2019**, *69*, 596–605. [CrossRef]
21. Gianelle, D.; Vescovo, L.; Marcolla, B.; Manca, G.; Cescatti, A. Ecosystem Carbon Fluxes and Canopy Spectral Reflectance of a Mountain Meadow. *Int. J. Remote Sens.* **2009**, *30*, 435–449. [CrossRef]
22. Seitz, S.; Goebes, P.; Puerta, V.L.; Pereira, E.I.P.; Wittwer, R.; Six, J.; van der Heijden, M.G.A.; Scholten, T. Conservation Tillage and Organic Farming Reduce Soil Erosion. *Agron. Sustain. Dev.* **2019**, *39*, 4. [CrossRef]
23. Doyle, C.J.; Topp, C.F.E. The Economic Opportunities for Increasing the Use of Forage Legumes in North European Livestock Systems under Both Conventional and Organic Management. *Renew. Agric. Food Syst.* **2004**, *19*, 15–22. [CrossRef]
24. Laidig, F.; Piepho, H.P.; Drobek, T.; Meyer, U. Genetic and Non-Genetic Long-Term Trends of 12 Different Crops in German Official Variety Performance Trials and on-Farm Yield Trends. *Theor. Appl. Genet.* **2014**, *127*, 2599–2617. [CrossRef] [PubMed]
25. Lollato, R.P.; Roozeboom, K.; Lingenfelser, J.F.; da Silva, C.L.; Sassenrath, G. Soft Winter Wheat Outyields Hard Winter Wheat in a Subhumid Environment: Weather Drivers, Yield Plasticity, and Rates of Yield Gain. *Crop Sci.* **2020**, *60*, 1617–1633. [CrossRef]
26. Andrade, J.F.; Rattalino Edreira, J.I.; Mourtzinis, S.; Conley, S.P.; Ciampitti, I.A.; Dunphy, J.E.; Gaska, J.M.; Glewen, K.; Holshouser, D.L.; Kandel, H.J.; et al. Assessing the Influence of Row Spacing on Soybean Yield Using Experimental and Producer Survey Data. *Field Crops Res.* **2019**, *230*, 98–106. [CrossRef]
27. Munaro, L.B.; Hefley, T.J.; DeWolf, E.; Haley, S.; Fritz, A.K.; Zhang, G.; Haag, L.A.; Schlegel, A.J.; Edwards, J.T.; Marburger, D.; et al. Exploring Long-Term Variety Performance Trials to Improve Environment-Specific Genotype × Management Recommendations: A Case-Study for Winter Wheat. *Field Crops Res.* **2020**, *255*, 107848. [CrossRef]
28. Frazier, A.E. Landscape Heterogeneity and Scale Considerations for Super-Resolution Mapping. *Int. J. Remote Sens.* **2015**, *36*, 2395–2408. [CrossRef]
29. Ge, Y.; Chen, Y.; Stein, A.; Li, S.; Hu, J. Enhanced Subpixel Mapping with Spatial Distribution Patterns of Geographical Objects. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 2356–2370. [CrossRef]
30. Feng, W.; Yao, X.; Zhu, Y.; Tian, Y.C.; Cao, W.X. Monitoring Leaf Nitrogen Status with Hyperspectral Reflectance in Wheat. *Eur. J. Agron.* **2008**, *28*, 394–404. [CrossRef]
31. Thorp, K.R.; Wang, G.; Bronson, K.F.; Badaruddin, M.; Mon, J. Hyperspectral Data Mining to Identify Relevant Canopy Spectral Features for Estimating Durum Wheat Growth, Nitrogen Status, and Grain Yield. *Comput. Electron. Agric.* **2017**, *136*, 1–12. [CrossRef]
32. Tsouros, D.C.; Bibi, S.; Sarigiannidis, P.G. A Review on UAV-Based Applications for Precision Agriculture. *Information* **2019**, *10*, 349. [CrossRef]
33. Raeva, P.L.; Šedina, J.; Dlesk, A. Monitoring of Crop Fields Using Multispectral and Thermal Imagery from UAV. *Eur. J. Remote Sens.* **2019**, *52*, 192–201. [CrossRef]
34. Baret, F.; Guyot, G.; Major, D.J. TSAVI: A Vegetation Index Which Minimizes Soil Brightness Effects on LAI and APAR Estimation. In Proceedings of the Digest—International Geoscience and Remote Sensing Symposium (IGARSS), Vancouver, BC, Canada, 10–14 July 1989; Volume 3.
35. Qi, J.; Chehbouni, A.; Huete, A.R.; Kerr, Y.H.; Sorooshian, S. A Modified Soil Adjusted Vegetation Index. *Remote Sens. Environ.* **1994**, *48*, 119–126. [CrossRef]
36. Rhyma, P.P.; Norizah, K.; Hamdan, O.; Faridah-Hanum, I.; Zulfa, A.W. Integration of Normalised Different Vegetation Index and Soil-Adjusted Vegetation Index for Mangrove Vegetation Delineation. *Remote Sens. Appl. Soc. Environ.* **2020**, *17*, 100280. [CrossRef]
37. Zhou, G. Determination of Green Aboveground Biomass in Desert Steppe Using Litter-Soil-Adjusted Vegetation Index. *Eur. J. Remote Sens.* **2014**, *47*, 611–625. [CrossRef]
38. Mutanga, O.; Skidmore, A.K. Narrow Band Vegetation Indices Overcome the Saturation Problem in Biomass Estimation. *Int. J. Remote Sens.* **2004**, *25*, 3999–4014. [CrossRef]
39. Zarco-Tejada, P.J.; Ustin, S.L.; Whiting, M.L. Temporal and Spatial Relationships between Within-Field Yield Variability in Cotton and High-Spatial Hyperspectral Remote Sensing Imagery. *Agron. J.* **2005**, *97*, 641–653. [CrossRef]

40. Haboudane, D.; Miller, J.R.; Pattey, E.; Zarco-Tejada, P.J.; Strachan, I.B. Hyperspectral Vegetation Indices and Novel Algorithms for Predicting Green LAI of Crop Canopies: Modeling and Validation in the Context of Precision Agriculture. *Remote Sens. Environ.* **2004**, *90*, 337–352. [CrossRef]
41. Monteiro, P.F.C.; Filho, R.A.; Xavier, A.C.; Monteiro, R.O.C. Assessing Biophysical Variable Parameters of Bean Crop with Hyperspectral Measurements. *Sci. Agric.* **2012**, *69*, 87–94. [CrossRef]
42. Xavier, A.C.; Theodor Rudorff, B.F.; Moreira, M.A.; Alvarenga, B.S.; de Freitas, J.G.; Salomon, M.V. Hyperspectral Field Reflectance Measurements to Estimate Wheat Grain Yield and Plant Height. *Sci. Agric.* **2006**, *63*, 130–138. [CrossRef]
43. Stagakis, S.; Markos, N.; Sykioti, O.; Kyparissis, A. Monitoring Canopy Biophysical and Biochemical Parameters in Ecosystem Scale Using Satellite Hyperspectral Imagery: An Application on a Phlomis Fruticosa Mediterranean Ecosystem Using Multiangular CHRIS/PROBA Observations. *Remote Sens. Environ.* **2010**, *114*, 977–994. [CrossRef]
44. Zarco-Tejada, P.J.; Miller, J.R.; Noland, T.L.; Mohammed, G.H.; Sampson, P.H. Scaling-up and Model Inversion Methods with Narrowband Optical Indices for Chlorophyll Content Estimation in Closed Forest Canopies with Hyperspectral Data. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 1491–1507. [CrossRef]
45. Zarco-Tejada, P.J. *Hyperspectral Remote Sensing of Closed Forest Canopies: Estimation of Chlorophyll Fluorescence and Pigment Content*; York University: Toronto, ON, Canada, 2000.
46. Schmidt, K.S.; Skidmore, A.K. Spectral Discrimination of Vegetation Types in a Coastal Wetland. *Remote Sens. Environ.* **2003**, *85*, 92–108. [CrossRef]
47. Feng, J.; Jiao, L.; Liu, F.; Sun, T.; Zhang, X. Unsupervised Feature Selection Based on Maximum Information and Minimum Redundancy for Hyperspectral Images. *Pattern Recognit.* **2016**, *51*, 295–309. [CrossRef]
48. Yang, M.D.; Huang, K.H.; Tsai, H.P. Integrating MNF and HHT Transformations into Artificial Neural Networks for Hyperspectral Image Classification. *Remote Sens.* **2020**, *12*, 2327. [CrossRef]
49. Bajcsy, P.; Groves, P. Methodology for Hyperspectral Band Selection. *Photogramm. Eng. Remote Sens.* **2004**, *70*, 793–802. [CrossRef]
50. Heiskanen, J.; Rautiainen, M.; Stenberg, P.; Mõttus, M.; Vesanto, V.H. Sensitivity of Narrowband Vegetation Indices to Boreal Forest LAI, Reflectance Seasonality and Species Composition. *ISPRS J. Photogramm. Remote Sens.* **2013**, *78*, 1–14. [CrossRef]
51. Näsi, R.; Viljanen, N.; Kaivosoja, J.; Alhonoja, K.; Hakala, T.; Markelin, L.; Honkavaara, E. Estimating Biomass and Nitrogen Amount of Barley and Grass Using UAV and Aircraft Based Spectral and Photogrammetric 3D Features. *Remote Sens.* **2018**, *10*, 1082. [CrossRef]
52. Li, C.; Ma, C.; Pei, H.; Feng, H.; Shi, J.; Wang, Y.; Chen, W.; Li, Y.; Feng, X.; Shi, Y. Estimation of Potato Biomass and Yield Based on Machine Learning from Hyperspectral Remote Sensing Data. *J. Agric. Sci. Technol. B* **2020**, *10*, 195–213. [CrossRef]
53. Choudhury, M.R.; Das, S.; Christopher, J.; Apan, A.; Chapman, S.; Menzies, N.W.; Dang, Y.P. Improving Biomass and Grain Yield Prediction of Wheat Genotypes on Sodic Soil Using Integrated High-Resolution Multispectral, Hyperspectral, 3d Point Cloud, and Machine Learning Techniques. *Remote Sens.* **2021**, *13*, 3482. [CrossRef]
54. Nawar, S.; Corstanje, R.; Halcro, G.; Mulla, D.; Mouazen, A.M. Delineation of Soil Management Zones for Variable-Rate Fertilization: A Review. *Adv. Agron.* **2017**, *143*, 175–245.
55. Chlingaryan, A.; Sukkarieh, S.; Whelan, B. Machine Learning Approaches for Crop Yield Prediction and Nitrogen Status Estimation in Precision Agriculture: A Review. *Comput. Electron. Agric.* **2018**, *151*, 61–69.
56. Zhang, W.; Zhang, Z.; Chao, H.C.; Guizani, M. Toward Intelligent Network Optimization in Wireless Networking: An Auto-Learning Framework. *IEEE Wirel. Commun.* **2019**, *26*, 76–82. [CrossRef]
57. Colombo, R.; Bellingeri, D.; Fasolini, D.; Marino, C.M. Retrieval of Leaf Area Index in Different Vegetation Types Using High Resolution Satellite Data. *Remote Sens. Environ.* **2003**, *86*, 120–131. [CrossRef]
58. He, X.; Zhao, K.; Chu, X. AutoML: A Survey of the State-of-the-Art. *Knowl.-Based Syst.* **2021**, *212*, 106622. [CrossRef]
59. Mendoza, H.; Klein, A.; Feurer, M.; Springenberg, J.T.; Hutter, F. Towards Automatically-Tuned Neural Networks. In Proceedings of the Workshop on Automatic Machine Learning, New York, NY, USA, 24 June 2016.
60. Yao, Q.; Wang, M.; Chen, Y.; Dai, W.; Hu, Y.Q.; Li, Y.F.; Tu, W.W.; Yang, Q.; Yu, Y. Taking the Human out of Learning Applications: A Survey on Automated Machine Learning. *arXiv* **2018**, arXiv:1810.13306.
61. Thornton, C.; Hutter, F.; Hoos, H.H.; Leyton-Brown, K. Auto-WEKA: Combined Selection and Hyperparameter Optimization of Classification Algorithms. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Chicago, IL, USA, 11–14 August 2013; pp. 847–855. [CrossRef]
62. Feurer, M.; Klein, A.; Eggensperger, K.; Springenberg, J.T.; Blum, M.; Hutter, F. Efficient and Robust Automated Machine Learning. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 7–12 December 2015; Volume 2015.
63. Remeseiro, B.; Bolon-Canedo, V. A Review of Feature Selection Methods in Medical Applications. *Comput. Biol. Med.* **2019**, *112*, 103375. [CrossRef]
64. Babaeian, E.; Paheding, S.; Siddique, N.; Devabhaktuni, V.K.; Tuller, M. Estimation of Root Zone Soil Moisture from Ground and Remotely Sensed Soil Information with Multisensor Data Fusion and Automated Machine Learning. *Remote Sens. Environ.* **2021**, *260*, 112434. [CrossRef]
65. Koh, J.C.O.; Spangenberg, G.; Kant, S. Automated Machine Learning for High-throughput Image-based Plant Phenotyping. *Remote Sens.* **2021**, *13*, 858. [CrossRef]

66. Komer, B.; Bergstra, J.; Eliasmith, C. Hyperopt-Sklearn: Automatic Hyperparameter Configuration for Scikit-Learn. In Proceedings of the 13th Python in Science Conference, Austin, TX, USA, 6–12 July 2014.
67. Lehnert, L.W.; Meyer, H.; Obermeier, W.A.; Silva, B.; Regeling, B.; Thies, B.; Bendix, J. Hyperspectral Data Analysis in R: The Hsdar Package. *J. Stat. Softw.* **2019**, *89*, 1–23. [CrossRef]
68. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020.
69. FAO. *World Reference Base for Soil Resources*; World Soil Resources Report 103; FAO: Rome, Italy, 2006; ISBN 9251055114.
70. The Mathworks Inc. *MATLAB (R2019a)*, Version 9.6.0.1072779; The MathWorks Inc.: Natick, MA, USA, 2019.
71. Beleites, C. Unstable Laser Emission Vignette for the Data Set Laser of the R Package HyperSpec. *Spectrosc. Laser* **2021**, *75*, 1–6.
72. Blackburn, G.A. Spectral Indices for Estimating Photosynthetic Pigment Concentrations: A Test Using Senescent Tree Leaves. *Int. J. Remote Sens.* **1998**, *19*, 657–675. [CrossRef]
73. Tucker, C.J. Red and Photographic Infrared Linear Combinations for Monitoring Vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [CrossRef]
74. Fernández-Manso, A.; Fernández-Manso, O.; Quintano, C. SENTINEL-2A Red-Edge Spectral Indices Suitability for Discriminating Burn Severity. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *50*, 170–175. [CrossRef]
75. Gitelson, A.; Merzlyak, M.N. Quantitative Estimation of Chlorophyll-a Using Reflectance Spectra: Experiments with Autumn Chestnut and Maple Leaves. *J. Photochem. Photobiol. B Biol.* **1994**, *22*, 247–252. [CrossRef]
76. Vescovo, L.; Wohlfahrt, G.; Balzarolo, M.; Pilloni, S.; Sottocornola, M.; Rodeghiero, M.; Gianelle, D. New Spectral Vegetation Indices Based on the Near-Infrared Shoulder Wavelengths for Remote Detection of Grassland Phytomass. *Int. J. Remote Sens.* **2012**, *33*, 2178–2195. [CrossRef]
77. Kim, M.S.; Daughtry, C.S.T.; Chappelle, E.W.; McMurtrey, J.E.; Walthall, C.L. The Use of High Spectral Resolution Bands for Estimating Absorbed Photosynthetically Active Radiation (A Par). In Proceedings of the 6th Symposium on Physical Measurements and Signatures in Remote Sensing, Val d'Isère, France, 17–21 January 1994.
78. Huete, A.R. A Soil-Adjusted Vegetation Index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [CrossRef]
79. Rondeaux, G.; Steven, M.; Baret, F. Optimization of Soil-Adjusted Vegetation Indices. *Remote Sens. Environ.* **1996**, *55*, 95–107. [CrossRef]
80. Yeom, J.; Jung, J.; Chang, A.; Ashapure, A.; Maeda, M.; Maeda, A.; Landivar, J. Comparison of Vegetation Indices Derived from UAV Data for Differentiation of Tillage Effects in Agriculture. *Remote Sens.* **2019**, *11*, 1548. [CrossRef]
81. Hernández-Clemente, R.; Navarro-Cerrillo, R.M.; Zarco-Tejada, P.J. Carotenoid Content Estimation in a Heterogeneous Conifer Forest Using Narrow-Band Indices and PROSPECT + DART Simulations. *Remote Sens. Environ.* **2012**, *127*, 32–46. [CrossRef]
82. Demmig-Adams, B.; Adams, W.W. Photoprotection and Other Responses of Plants to High Light Stress. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1992**, *43*, 599–626. [CrossRef]
83. Wu, C.; Niu, Z.; Tang, Q.; Huang, W. Estimating Chlorophyll Content from Hyperspectral Vegetation Indices: Modeling and Validation. *Agric. For. Meteorol.* **2008**, *148*, 1230–1241. [CrossRef]
84. Roujean, J.L.; Breon, F.M. Estimating PAR Absorbed by Vegetation from Bidirectional Reflectance Measurements. *Remote Sens. Environ.* **1995**, *51*, 375–384. [CrossRef]
85. ESRI. *ArcGIS PRO: Essential Workflows*; ESRI: Redlands, CA, USA, 2016.
86. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
87. Feurer, M.; Eggensperger, K.; Falkner, S.; Lindauer, M.; Hutter, F. Auto-Sklearn 2.0: The Next Generation. *arXiv* **2020**, arXiv:2007.0407424.
88. Hutter, F.; Hoos, H.H.; Leyton-Brown, K. Sequential Model-Based Optimization for General Algorithm Configuration. In *Proceedings of the Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2011; Volume 6683.
89. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *24*, 123–140. [CrossRef]
90. Olson, R.S.; Urbanowicz, R.J.; Andrews, P.C.; Lavender, N.A.; Kidd, L.C.; Moore, J.H. Automating Biomedical Data Science through Tree-Based Pipeline Optimization. In *Proceedings of the Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2016; Volume 9597.
91. Franceschi, L.; Frasconi, P.; Salzo, S.; Grazzi, R.; Pontil, M. Bilevel Programming for Hyperparameter Optimization and Meta-Learning. In *International Conference on Machine Learning*; PMLS; 2018; Volume 80, pp. 1568–1577. Available online: http://proceedings.mlr.press/v80/franceschi18a.html?ref=https://githubhelp.com (accessed on 26 January 2022).
92. Feurer, M.; Springenberg, J.T.; Hutter, F. Initializing Bayesian Hyperparameter Optimization via Meta-Learning. In Proceedings of the National Conference on Artificial Intelligence, Austin, TX, USA, 25–30 January 2015; Volume 2.
93. Yue, J.; Yang, G.; Li, C.; Li, Z.; Wang, Y.; Feng, H.; Xu, B. Estimation of Winter Wheat Above-Ground Biomass Using Unmanned Aerial Vehicle-Based Snapshot Hyperspectral Sensor and Crop Height Improved Models. *Remote Sens.* **2017**, *9*, 708. [CrossRef]
94. Ono, J.P.; Castelo, S.; Lopez, R.; Bertini, E.; Freire, J.; Silva, C. PipelineProfiler: A Visual Analytics Tool for the Exploration of AutoML Pipelines. *IEEE Trans. Vis. Comput. Graph.* **2021**, *27*, 390–400. [CrossRef]
95. Qi, Y.; Minka, T.P.; Picard, R.W.; Ghahramani, Z. Predictive Automatic Relevance Determination by Expectation Propagation. In Proceedings of the Twenty-First International Conference on Machine Learning, ICML 2004, Banff, AB, Canada, 4–8 July 2004.
96. Seeger, M. Gaussian Processes for Machine Learning. *Int. J. Neural Syst.* **2004**, *14*, 69–106. [CrossRef]

97. Van Diedenhoven, B.; Hasekamp, O.P.; Aben, I. Surface Pressure Retrieval from SCIAMACHY Measurements in the O2A Band: Validation of the Measurements and Sensitivity on Aerosols. *Atmos. Chem. Phys. Discuss.* **2005**, *5*, 2109–2120. [CrossRef]
98. Riris, H.; Rodriguez, M.; Mao, J.; Allan, G.; Abshire, J. Airborne Demonstration of Atmospheric Oxygen Optical Depth Measurements with an Integrated Path Differential Absorption Lidar. *Opt. Express* **2017**, *25*, 29307–29327. [CrossRef]
99. Gitelson, A.A.; Gritz, Y.; Merzlyak, M.N. Relationships between Leaf Chlorophyll Content and Spectral Reflectance and Algorithms for Non-Destructive Chlorophyll Assessment in Higher Plant Leaves. *J. Plant Physiol.* **2003**, *160*, 271–282. [CrossRef] [PubMed]
100. Darvishzadeh, R.; Skidmore, A.; Atzberger, C.; van Wieren, S. Estimation of Vegetation LAI from Hyperspectral Reflectance Data: Effects of Soil Type and Plant Architecture. *Int. J. Appl. Earth Obs. Geoinf.* **2008**, *10*, 358–373. [CrossRef]
101. Martínez-Usó, A.; Pla, F.; Sotoca, J.M.; García-Sevilla, P. Clustering-Based Hyperspectral Band Selection Using Information Measures. *Proc. IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 4158–4171.
102. Thenkabail, P.S.; Smith, R.B.; de Pauw, E. Hyperspectral Vegetation Indices and Their Relationships with Agricultural Crop Characteristics. *Remote Sens. Environ.* **2000**, *71*, 158–182. [CrossRef]
103. Lecun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
104. Yang, M.D.; Tseng, H.H.; Hsu, Y.C.; Tsai, H.P. Semantic Segmentation Using Deep Learning with Vegetation Indices for Rice Lodging Identification in Multi-Date UAV Visible Images. *Remote Sens.* **2020**, *12*, 633. [CrossRef]
105. Yang, M.D.; Boubin, J.G.; Tsai, H.P.; Tseng, H.H.; Hsu, Y.C.; Stewart, C.C. Adaptive Autonomous UAV Scouting for Rice Lodging Assessment Using Edge Computing with Deep Learning EDANet. *Comput. Electron. Agric.* **2020**, *179*, 683–696. [CrossRef]
106. Luan, J.; Zhang, C.; Xu, B.; Xue, Y.; Ren, Y. The Predictive Performances of Random Forest Models with Limited Sample Size and Different Species Traits. *Fish. Res.* **2020**, *227*, 105534. [CrossRef]
107. Feurer, M.; Hutter, F. Towards Further Automation in AutoML. In Proceedings of the ICML 2018 AutoML Workshop, Bangkok, Thailand, 25–28 October 2018.
108. Ma, L.; Li, M.; Ma, X.; Cheng, L.; Du, P.; Liu, Y. A Review of Supervised Object-Based Land-Cover Image Classification. *ISPRS J. Photogramm. Remote Sens.* **2017**, *130*, 277–293. [CrossRef]
109. Neal, R. Bayesian Learning for Neural Networks. In *Lecture Notes in Statistics*; Springer: New York, NY, USA, 1996; Volume 1.
110. Mackay, D.J.C. Probable Networks and Plausible Predictions—A Review of Practical Bayesian Methods for Supervised Neural Networks. *Netw. Comput. Neural Syst.* **1995**, *6*, 469. [CrossRef]

CURRICULUM VITAE

Name: Kai-Yun
Surnames: Li
Date of Birth: 24.04.1988
E-mail: kai-yun.li@student.emu.ee

Education

2018-2022	PhD studies in Environmental Sciences and Applied Biology
2010-2012	M.Sc. in Agriculture and Natural Resources. National Chung Hsing University, Taiwan.
2006-2010	B.Sc. in Agronomy. National Chiayi University, Taiwan.

Additional training

2019	Advanced Training School on Remote Sensing/ Earth Observation. Võru, Estonia.
2021	High throughput field phenotyping training, Norwegian University of Life Sciences, Department of Plant Sciences.

Languages

Chinese (Mother tongue) English (Advanced)

Professional employment:

2020 – Present	Estonian University of Life Sciences. Institute of Agricultural and Environmental Sciences. Chair of Environmental Protection and Landscape Management.
2015 – 2018	Agriculture Department Manager, Shoufeng Township Office, Hualien, Taiwan.
2014 – 2015	Associate Technical Agriculture Specialist, Shoufeng Township Office, Hualien, Taiwan.

2012 – 2013 Project Assistant,
International Cooperation and Development Fund
Taiwan Technical Mission, Solomon Islands.

Projects

1.01.2022 -31.12.2022 EAG204. "Innovative drone-based remote sensing tools for agricultural management and nature conservation".

LIST OF PUBLICATIONS

Scholarly articles indexed by Web of Science Science Citation Index Expanded, Social Sciences Citation Index, Arts & Humanities Citation Index and/or indexed by Scopus

I **Li, K. Y.**, Burnside, N. G., de Lima, R. S., Villoslada, M., Sepp, K., Yang, M. D., Raet, J. Vain, A., Selge, A. & Sepp, K. (2021). The Application of an Unmanned Aerial System and Machine Learning Techniques for Red Clover-Grass Mixture Yield Estimation under Variety Performance Trials. Remote Sensing, 13(10), 1994.

II **Li, K. Y.**, Burnside, N. G., de Lima, R. S., Villoslada, M., Sepp, K., Cabral Pinheiro, V. H., de Lima, B.R.C.A., Yang, M. D., Vain, A., & Sepp, K. (2021). An Automated Machine Learning Framework in Unmanned Aircraft Systems: New Insights into Agricultural Management Practices Recognition Approaches. Remote Sensing, 13(16), 3190.

III **Li, K. Y.**, de Lima, R. S., Burnside, N. G., Vahtmäe, E., Kutser, T., Sepp, K., Cabral Pinheiro, V.H., Yang, M. D., Vain, A. & Sepp, K. (2022). Toward Automated Machine Learning-Based Hyperspectral Image Analysis in Crop Yield and Biomass Estimation. Remote Sensing, 14(5), 1114.

IV de Lima, R. S.; **Li, K. Y.**; Vain, A.; Lang, M.; Bergamo, T.F.; Kokamägi, K.; Burnside, N.G.; Ward, R.D.; Sepp, K. The Potential of Optical UAS Data for Predicting Surface Soil Moisture in a Peatland across Time and Sites. Remote Sens. 2022, 14, 2334.

VIIS VIIMAST KAITSMIST

JORDI ESCUER GATIUS
MITIGATION OF NITROUS OXIDE EMISSIONS FROM ARABLE SOILS
PÕLLUMULLAST ERALDUVA DILÄMMASTIKOKSIIDI VÄHENDAMINE
Kaasprofessor Merrit Shanskiy, kaasprofessor Kaido Soosaar, professor Alar Astover
30. juuni 2022

MARTA MARIA ALOS ORTI
URBAN ECOLOGY: NOVEL ECOSYSTEMS, NOVEL CHALLENGES
LINNAÖKOLOOGIA VÄLJAKUTSED: UUED ÖKOSÜSTEEMID, UUED LAHENDUSED
Professor Lauri Laanisto
29. august 2022

KÄTLIN PITMAN
BIOSENSOR ARRAY FOR BOD MEASUREMENTS
IN DIFFERENT TYPES OF WASTEWATER
BIOSENSOR-RIVI ERINEVATE REOVETE BIOKEEMILISE HAPNIKUTARBE UURIMISEKS
Kaasprofessor Merlin Raud, kaasprofessor Jaak Nerut, professor Timo Kikas
29. august 2022

GRETE TÕNISALU
SMALL MAMMALS, THE LESSER SPOTTED EAGLE, AND ECOTONES: A CASE STUDY ON PREDATOR-PREY-HABITAT RELATIONSHIPS IN AGRICULTURAL LANDSCAPE
PISIIMETAJAD, VÄIKE-KONNAKOTKAS JA SERVAALAD: SAAKLOOMA, KISKJA JA ELUPAIGA SEOSTE UURING PÕLLUMAJANDUSMAASTIKUS
Vanemteadur Ülo Väli
5. september 2022

FRANCESCA CARNOVALE
WELFARE OF LIVESTOCK SHEEP TRANSPORT IN HOT AND COLD CLIMATES
LAMMASTE HEAOLU NENDE TRANSPORDIL KUUMAS JA KÜLMAS KLIIMAS
Professor David Arney, vanemlektor Andres Aland, professor Fabio Napolitano
3. oktoober 2022

ISSN 2382-7076
ISBN 978-9916-669-63-1 (trükis)
ISBN 978-9916-669-64-8 (pdf)