Student thesis series INES nr 667

# Estimating dry matter content in forage grasslands using Sentinel-2 satellite data

**Fan Wang**

2023
Department of
Physical Geography and Ecosystem Science
Lund University
Sölvegatan 12
S-223 62 Lund
Sweden

Fan Wang (2023).

**Estimating dry matter content in forage grasslands using Sentinel-2 satellite data**

Master's degree thesis, 30 credits in Subject of Degree

Department of Physical Geography and Ecosystem Science, Lund University

Level: Master of Science (MSc)

Course Duration: January 2023 –June 2023

# Estimating dry matter content in forage grasslands using Sentinel-2 satellite data

---

Fan Wang

Master thesis, 30 credits, in *Geomatics*

Supervisors:

Zhanzhang Cai

Zheng Duan

Dep. of Physical Geography and Ecosystem Science, Lund University

Exam committee:

Hongxiao Jin

Babak Mohammadi

Dep. of Physical Geography and Ecosystem Science, Lund University

# Acknowledgements

I would like to thank ZhanZhang Cai for great supervision and valuable guidance. I am grateful for his willingness to engage in open discussions and exchange ideas. I also want to thank Zheng Duan for providing insightful advice in defining the topic of this thesis. Special thanks to Dr. Min Wang from Vultus, Sweden and Alexandre Teixeira from TKS Agri, Norway for allowing me to use the field data, without which the study would not have been conducted. I would also like to express my gratitude to my beloved family and friends for unwavering support and companionship.

# Abstract

Over the past two decades, the integration of sensor technology, automation, and data analysis has seen increasing use in agricultural applications. Dry matter content as percentage (DMCaP), defined as the ratio of dry mass to fresh mass, is an essential quality parameter for forage. In this thesis, we utilized in-situ measurements of DMCaP obtained from forage grasslands to evaluate the capabilities of Sentinel-2 data, and combined with environmental variables in estimating community-level DMCaP in forage grasslands in Southern Norway. Two types of models were developed: an interpretable model and a Random Forest Regression (RFR) model. The RFR model demonstrated superior performance compared to the interpretable model, exhibiting a high potential for accurately estimating DMCaP (RMSE = 3.88%). Adding environmental variables further improved the accuracy (RMSE = 2.90%). For the interpretable model, little difference was observed in the use of different vegetation indices. Despite the inherent uncertainties in the data, the RFR model proved to be a viable option for estimating DMCaP in local-scale forage grasslands. This study underscores the potential of satellite remote sensing for estimating the DMCaP quality factor in agricultural applications, providing valuable insights for the forage-harvesting process and pasture management. It addresses the gap in estimating vegetation dry matter content between the agricultural and academic communities, as well as the limitations observed in previous studies when matching satellite and in-situ data pairs.

***Keywords:*** *dry matter content, machine learning, remote sensing, forage grasslands, vegetation index*

# Abbreviations

| | |
|---|---|
| **ANN** | Artificial Neural Network |
| **ASL** | Above Sea Level |
| **BOA** | Bottom-of-Atmosphere |
| **CV** | Cross-Validation |
| **CVI** | Chlorophyll Vegetation Index |
| **DMC** | Dry Matter Content |
| **DMCaP** | Dry Matter Content as Percentage |
| **DT** | Decision Tree |
| **EM** | Electromagnetic |
| **EVI** | Enhanced Vegetation Index |
| **EWT** | Equivalent Water Thickness |
| **GEE** | Google Earth Engine |
| **GNDVI** | Green Normalized Difference Vegetation Index |
| **LAI** | Leaf Area Index |
| **LMA** | Leaf Mass per Area |
| **MAE** | Mean Absolute Error |
| **ML** | Machine Learning |
| **MSI** | Multi-Spectral Instrument |
| **NDMI** | Normalized Dry Matter Index |
| **NDRE** | Normalized Difference Red-Edge |
| **NDVI** | Normalized Difference Vegetation Index |
| **NDWI** | Normalized Difference Water Index |
| **NIR** | Near-Infrared |
| **$R^2$** | Coefficient of determination |
| **RFR** | Random Forest Regression |
| **RMSE** | Root Mean Square Error |
| **RS** | Remote Sensing |
| **RT** | Radiative Transfer |
| **S2** | Sentinel-2 |
| **SCL** | Scene Classification Layer |
| **RVI** | Ratio Vegetation Index |
| **SVR** | Support Vector Regression |
| **SWIR** | Shortwave-infrared |

| | |
|---|---|
| **TOA** | Top-of-Atmosphere |
| **VIs** | Vegetation Indices |
| **VIS** | Visible |
| **WC** | Water Content |

# List of Figures

# List of Tables

# Table of Contents

# 1. Introduction

The continuing growth of the world population requires effective agricultural practices to provide sufficient food. Animal-based food (meat, dairy, eggs) currently plays a crucial role in global food supply, with the European region having the highest level of consumption of animal-based foods, and the demand for animal-sourced food is still predicted to increase from a global perspective (Henchion et al., 2021). Technological innovation has greatly influenced the society over the past two decades, which has been labelled the fourth industrial revolution (Shepherd et al., 2018). These technological advancements have also had a large impact on food systems, and livestock systems in particular, is on the cusp of a 'revolution' with the use of sensor technology, automation, and big data along with data analysis (Barrett & Rose, 2022; Newton et al., 2020).

Forage grasslands, also known as pastures, are for forage production and harvest by grazing, cutting or both (Allen et al., 2011). They have been estimated to represent 26% of the global land area and 70% of the agricultural area (Conant, 2010). From the perspective of agricultural industry, they provide the most cost-effective source of herbage for the livestock sector. Forage crop could be fed directly to livestock through grazing or in a processed form, such as partial drying (as hay) or pre-digestion (as silage) (Capstaff & Miller, 2018).

The dry matter content (DMC) of forage, which refers to the solid substances that remain after water has been removed. The composition of DMC includes water-soluble carbohydrates, proteins, lipids, fibers and other chemical and structural compounds. The dry matter content as percentage (DMCaP), which is defined as the ratio of forage dry mass to fresh mass, is an essential quality parameter for getting the maximum yield of nutrients and producing palatable forage. DMCaP could indicate the nutritional and energy level of forage, and thus, provide feeding guidance to animal keepers (Capstaff & Miller, 2018).

DMCaP is especially important when engaging in silage production practices. The silage-making process involves the following phases: the initial aerobic condition in a silo immediately after harvesting, the fermentation phase after compacting and sealing to exclude air, the stable storage phase and finally the animal-feeding phase (Wilkinson

& Davies, 2013). The proper DMCaP value at harvesting ensures good fermentation and optimum intakes while minimizing the risk of aerobic instability (*Grassland guide: Grass nutritional value*, n.d.).The ideal range of DMCaP for making silage is between 20% and 40% depending on the incoming type of silage, for example, the optimal range for grass is 25% to 30%. The silage would have a lower nutrition level and may spoil with very low DMCaP, whereas too high DMC would make it difficult to ferment (*Getting Silage Dry Matter Right*, n.d.).

A conventional and the most precise method to determine DMCaP is the gravimetric method through destructive cutting (Hopkins, 2000). Portable field near-infrared spectroradiometers have also been used to estimate DMCaP (Cevoli et al., 2022). However, either of these two methods is small-plot based, and could be time-consuming and labor-intensive. Remote sensing provides spectral, temporal and spatial information, enabling the estimation of DMCaP of extensive areas simultaneously. This approach minimizes the need for field measurements and laboratory procedures, ultimately easing efforts of monitoring DMCaP for agricultural practices.

Both DMC and Water Content (WC) are involved in estimating DMCaP. Past studies have mostly focused on estimating leaf-level traits of DMC and WC using RS (e.g., Wang et al. (2011), Yilmaz et al. (2008), and Shah et al. (2019)). Few studies have explored community-level DMCaP. Li et al. (2018) developed a linear model using field-measured community-level plant DMCaP and Landsat-8 vegetation indices (VIs) to estimate DMCaP of Alphine Grasslands for the whole Qinghai-Tibetan Plateau, showing the highest accuracy ($R^2$ =0.53, rRMSE = 0.144) of using enhanced vegetation index (EVI) in comparison with other indices. Bretas et al. (2021) modelled the relationship between community-level DMCaP and Landsat-8/ Sentinel-2 VIs combined with meteorological data to predict DMCaP in four pastures in Brazil, where Random Forest (RF) algorithm had the best performance ($R^2$ = 0.85, rRMSE = 0.079) among all models. These models were built based on simple linear regression and/or machine learning algorithm. Both studies linked field and satellite measurements by extracting the closet pixel values with respect to the individual field sampling locations and dates, which could result in significant errors with respect to timeliness. To date, very little is known about the potential of utilizing satellite remote sensing to monitor community-level DMCaP in pastures for agricultural purposes in Northern Europe.

The overall aim of this study is to explore the capabilities of primarily Sentinel-2 satellite data, and combined with environmental variables, in estimating DMCaP in local-scale forage grasslands in Southern Norway. Mapping DMCaP would provide valuable insights in forage-harvesting process and pasture management. Three objectives of this study are as follows:

1) Development of models for estimating community-level DMCaP by utilizing primarily Sentinel-2 data. The aim is to establish a robust framework that bridges the temporal gaps and overcome the limitation of temporal inconsistency between field measurements and satellite data.

2) Evaluation of the performance of the interpretable model and the Random Forest Regression (RFR) model in estimating forage grasslands DMCaP. The evaluation will be conducted using a regional-scale dataset of community-level DMCaP samples.

3) Upscaling of plot-based measurements to the entire field using the model of the best performance, and analysis of spatial and temporal variation of DMCaP.

# 2. Background

## 2.1 Satellite Remote Sensing

Remote sensing (RS) refers to techniques that observe objects from a distance without direct contact (Lillisand et al., 2008). It started from afar when creating maps with the use of balloons, and nowadays, the term RS refers to data acquired from satellites and airborne platforms carrying optical and radar sensors, as well as airborne surveying and photogrammetry (Campbell & Wynne, 2011). To distinguish between different platforms, depending on how far the platform is from the ground, RS could be broadly categorized to satellite-based, airborne-based and ground-based. Satellite RS refers to the satellite equipped with sensors collecting data about the Earth's surface and atmosphere.

Depending on the type of the sensor, RS technologies can be broadly categorized into two types: passive RS and active RS. Passive RS relies on detecting natural radiation emitted or reflected by an external source, such as the sun, not by the sensor. This can include visible (VIS) and near-infrared (NIR) radiation from the sun, as well as thermal-infrared (TIR) radiation emitted by the Earth's surface. Passive sensors, such as photographic cameras and radiometers, are also referred to as optical sensors. Examples of passive RS applications include land cover mapping (Kroupi et al., 2019), vegetation monitoring (Li et al., 2018), grassland management (Xu et al., 2019), and atmospheric composition analysis (Clarmann et al., 2002). Active RS, on the other hand, captures the reflected or backscattered signal emitted by the sensor. Typical active sensors include radar, lidar and sonar. Examples of active RS are topography mapping (Lo Re et al., 2018), lake volume monitoring (Cretaux et al., 2016), vegetation assessment (Zhang & Shao, 2021), and ocean properties retrieval (Behrenfeld et al., 2022). Both optical and active RS have limitations, and the former is often influenced by cloud cover, while the latter can penetrate through clouds but with higher noise level.

### 2.1.1 Optical and active sensing in monitoring vegetation

Optical and active RS have been both employed to monitor vegetation. The utilization of RS data for vegetation monitoring relies on the sensitivity of electromagnetic (EM) radiation within their respective spectral range to specific biochemical and/or biophysical parameters of plants, such as chlorophyll content, water content, and leaf

area. Different spectral ranges (e.g., optical or radar) can detect different parameters. Optical RS is based on the reflection and refraction characteristics of the canopy surface, and soil as well, if the vegetation cover is below 100%, providing information on vegetative biophysical or biochemical processes. Active radiation is measured through the reflected or backscattered signals, which are influenced by the geometrical structure and the dielectric characteristics of the canopy. Depend on the frequency, it can penetrate through the canopy down into the soil (Gerstl, 1990).

In agricultural vegetation monitoring, optical data is more frequently used compared to radar data. This is because optical data, particularly vegetation indices (VIs), are easily accessible and have been proven to be suitable for monitoring agricultural vegetation (Liu et al., 2019; Weiss et al., 2020). However, the limitation of dependence on cloudless skies and solar illumination for optical RS can result in fragmented time series. In contrast, active RS is able to overcome this limitation, providing high temporal resolution regardless of the weather and time of the day. Disadvantages of radar data include its difficulty in comprehension due to the ill-posed nature of radar signal interpretation and its higher noise interference. Some radar vegetation indices derived from synthetic aperture radar data, have been developed and used in crop monitoring (Kim et al., 2012; Kumar et al., 2013).

Efforts have also been made to integrate optical and radar data (Orynbaikyzy et al., 2019). In one approach for example, Sentinel-1 (S1) and Sentinel-2 (S2) data were handled separately in the first step and then the two datasets were used jointly in regression analyses to estimate biomass (Forkuor et al., 2020), leaf area index (LAI) (Wang et al., 2019), yield (Mateo-Sanchis et al., 2019), or soil moisture (Amazirh et al., 2018).

### 2.1.2 Popular optical Satellite Remote Sensing systems

Currently, there are over one thousand functioning satellites orbiting around the Earth, many of which are for RS purposes. Satellites are equipped with one or more sensors or instruments depending on their intended use. These sensors collect different types of data about the surface of the Earth, including land, water, and atmosphere. The satellites and their sensors continuously capture images of the Earth's surface at varying spatial, spectral, radiometric, and temporal resolutions. Spatial resolution is defined by the size of the pixel that represents a ground area. If the sensor has a small footprint, it tends to

have high spatial resolution, whereas a large footprint leads to low spatial resolution. Radiometric resolution refers to the ability of a sensor to measure and distinguish between different levels of EM energy reflected or emitted from the Earth's surface, which determines the level of detail that can be captured in an image. Temporal resolution refers to the time interval between two successive observations of the same area on the ground, which is considered to be associated with the satellite platform rather than the sensor. A high temporal resolution means that the sensor can capture images of the same location at short time intervals, resulting in more frequent updates of the area. Conversely, a low temporal resolution means that the sensor takes longer to revisit the same location, leading to fewer updates over time. Spectral resolution of a sensor is indicated the number of bands captured within a specific range of the EM spectrum (Nowatzki et al., 2004).

It is critical to have appropriate spatial and temporal resolutions with agricultural applications. This may vary relying on multiple factors including the goals of management, the size the field, and the capacity of farm machinery to adjust inputs such as fertilizer, irrigation, and pesticides. The use of RS in agriculture has a long history, predating the introduction of the term "remote sensing" in 1958 (Nellis et al., 2009). Back then, these methods typically required extensive fieldwork and laboratory analysis, as low-altitude photography and ground crews were usually involved (Still & Shih, 1985). The introduction of satellite RS in later years enabled a more efficient and effective way for agricultural purposes at regional, national, and global scales.

The era of satellite RS for agricultural applications began when the National Aeronautics and Space Administration (NASA) launched Landsat 1 (previously called the Earth Resources Technology Satellite, or ERTS) in 1972. Subsequently, a sequence of Landsat satellites (Landsat 2-8) was launched to provide high-quality imagery to researchers, policymakers, and land managers in managing natural resources. In 1984, Landsat 5 Thematic Mapper (TM) with visible and near-infrared bands was launched, providing data with higher spectrum and spatial resolution (30m). France and India launched their own satellites, the SPOT 1 and IRS-1A in 1986 and 1988, respectively. More recently, there have been more satellites RS systems for agricultural purposes. Table 2.1 summarizes sensed EM spectrum (bands) and spatiotemporal resolutions from popular optical satellite RS systems exploited by relevant research regarding vegetation.

*Table 2.1. A summary of popular optical satellite remote sensing systems in vegetation monitoring. (MS: multi-spectral, TIR: thermal infrared)*

| Satellite (Years Active) | Bands (Spatial resolution) | Temporal resolution | Applications |
|---|---|---|---|
| **Landsat 8 (2013-) Landsat 9 (2021-)** | MS (30m-VIS, NIR, SWIR) and TIR (100m) | 16 days | Plant traits estimation (Li et al., 2018); Vegetation mapping (Sharma et al., 2017); Biomass and cover estimates (Jansen et al., 2018). |
| **MODIS (1999-)** | MS (250m, 500m) and TIR (1km) | 1-2 days | Vegetation dry matter (Khanna et al., 2007); Leaf dry matter content estimation (Adab et al., 2016); Vegetation water content (Zarco-Tejada et al., 2003). |
| **Sentinel-2 (2015-)** | MS (VIS-10m, NIR-10m, 20m, SWIR-20m, 60m) | 2-5 days | Vegetation mapping (Sharma et al., 2017); Plant traits (Li et al., 2018); Crop yield and type(Elders et al., 2022); Leaf chlorophyll content (Darvishzadeh et al., 2019); Canopy traits (Gara et al., 2019). |
| **SPOT 6 (2012-) SPOT 7 (2014-)** | MS (6m-VIS, NIR) | 1 day | Disease (Yuan et al., 2016); Vegetation Mapping (Hubert-Moy et al., 2020). |
| **AVHRR** | MS (1.1km-VIS, NIR) and TIR (4km) | 1 day | Vegetation mapping (Ju & Masek, 2016); Phenology (Shen et al., 2014). |

## 2.2 Dry matter content

DMC can be expressed in multiple ways depending on the purposes and different scales. DMCaP is such one example for agricultural purposes, and this term mentioned in this study all refers to the measurements at community level. DMC is often used as a term to measure biomass yield with the unit of kg DM/ha when quantifying herbage production (Oliveira et al., 2020). In terms of traits at leaf-level for vegetation monitoring, it is known as leaf dry matter content (LDMC) or the leaf mass per area (LMA), defined as the ratio of leaf dry mass to leaf area (unit, $g/cm^2$) (Jacquemoud et al., 1996). LDMC and together with leaf area index (LAI) could also provide an estimation of canopy biomass (Baret & Fourty, 1997). The ratio between aboveground plant dry biomass and fresh biomass is referred to as plant dry matter content (PDMC) when it comes to vegetation that does not have prominent aboveground stems and those stems are green as well such as grass (Li et al., 2018).

To date, most studies have focused on estimating DMC at leaf-level, or LMA. However,

dry matter and water content are both involved for estimating DMCaP at plant community-level. Vegetation water content is typically expressed using two metrics: (i) equivalent water thickness (EWT), defined as the ratio of water mass to leaf area (unit, $g/cm^2$) (Danson et al., 1992) and (ii) fuel moisture content (FMC), defined as the ratio of water mass to either the fresh weight (Mbow, 1999) or the dry weight (Burgan, 1996; Chuvieco et al., 1999). The latter term is more commonly used in forest fire research (Sow et al., 2013), and plus with its ambiguous definition, EWT is chosen to represent vegetation water content in this study. Therefore, vegetation dry matter and water content at leaf-level are expressed as LMA and EWT respectively. In this section, relevant studies regarding LMA and EWT estimation will be reviewed to provide a foundation for this study, especially for building the interpretable model.

### 2.2.1 Field and laboratory measurements

DMCaP can be obtained by field measurements and RS techniques. The destructive small-plot cutting method is the conventional and the most precise approach for determining DMCaP, which includes basic operations of cutting and weighing a fresh vegetation sample, followed by drying the sample in the oven and calculate the DMCaP as the ratio of the weight of dried samples and fresh samples (Hopkins, 2000). Portable field spectroradiometers have also been used to measure canopy reflectance to estimate vegetation biophysical and biochemical characteristics such as leaf chlorophyll concentration and lead nitrogen concentration (Hansen & Schjoerring, 2003). However, small-plot methods, either destructive cutting or portable spectroradiometer methods, are time-consuming and labor-intensive. Especially for large areas, it is not practical to reach full spatial coverage of DMCaP. Alternatively, remote sensing techniques have shown the potential to provide both high temporal and spatial information on DMCaP with the advantages of being nondestructive and covering large areas (Cheng et al., 2014; Conejo et al., 2015).

### 2.2.2 Satellite imagery estimation

The estimation of LMA and EWT using remote sensing techniques can be broadly divided into two categories: physical-based methods and data-driven methods. The physical-based methods use radiative transfer (RT) models to simulate the interaction between spectral information and vegetation canopy, allowing for forward modelling of leaf optical properties as well as backward inverse estimation of biophysical and

biochemical traits from leaf optical properties (Feret et al., 2021). Numerous radiative transfer models have been employed to model leaf optical properties, such as the PROSPECT (Feret et al., 2017; Jacquemoud & Baret, 1990), LIBERTY (Dawson et al., 1998; Wang & Ju, 2017) , and SLOP models (Maier et al., 1999). PROSPECT model is the most widely used among them owing to its simplicity and satisfactory performances. Biophysical traits such as LMA and EWT could be estimated using model inversion methods of either look-up-table method (Ali et al., 2016) or iterative optimization method (Li & Wang, 2011). These methods were proposed under the assumption that the leaf-specific absorption spectra of all vegetation species remain constant and unable to accommodate the spectral variability of leaf biochemical constituents (Koirala et al., 2020). Hence, the inversion estimation is challenging due to the significant number of unknowns relative to the independent information embedded in the spectral signature. Moreover, the inversion algorithms can be computational-demanding.

The data-driven methods are built relying on the statistical relationship between the optical characteristics of leaves (mostly at canopy-level) and their biochemical constituents. They are widely used due to their simplicity and computational efficiency, which make them a preferred method for fast estimation in agricultural practices. To establish statistical models, various regression techniques are employed to calibrate the training datasets. These regression techniques include simple linear regression (Li et al., 2018; Romero et al., 2012), partial least square regression (PLSR) (Hansen & Schjoerring, 2003), as well as complex machine learning (ML) models such as support vector regression (Feret et al., 2019; Yao et al., 2015), artificial neural networks (Jin & Liu, 1997) , and Random Forest Regression (Shah et al., 2019). Optical characteristics of leaves comprise the leaf reflectance, its derivates and combinations such as vegetation indices (VIs) (Feret et al., 2011; le Maire et al., 2008), and red-edge positions (Frampton et al., 2013).

The VIs are mathematical combinations of leaf reflectance, which are indicators of the presence, health, and productivity of vegetation, used as substitutes for time-series analysis (Hmimina et al., 2013) and vegetation states monitoring (Hill, 2013). They could reduce the volume of data and noise, making it easier to estimate biophysical and biochemical traits of vegetation.

*Table 2.2. Remotely sensed indices related to vegetation dry matter and water content extracted from literature. NDMI (Normalized dry matter index); NDWI (Normalized Difference Water Index); NDII (Normalized Difference Infrared Index); RMSI (Reciprocal of moisture stress index), SRWI (Simple ratio water index), LWVI1(Leaf Water Vegetation Index 1), LWVI2(Leaf Water Vegetation Index 2)*

| Type | Abbreviation | Formula | Reference |
|---|---|---|---|
| **LMA** | NDMI | $\frac{\rho 1649 - \rho 1722}{\rho 1649 + \rho 1722}$ | Wang et al. (2011) |
| | NDLMA | $\frac{\rho 1368 - \rho 1722}{\rho 1368 + \rho 1722}$ | le Maire et al. (2008) |
| | ND | $\frac{\rho 2295 - \rho 1550}{\rho 2295 + \rho 1550}$ | le Maire et al. (2008) |
| | RI | $\frac{\rho 1368}{\rho 1722}$ | Feret et al. (2011) |
| **EWT** | NDWI | $\frac{\rho 860 - \rho 1240}{\rho 860 + \rho 1240}$ | Gao (1995) |
| | NDII | $\frac{\rho 820 - \rho 1650}{\rho 820 + \rho 1650}$ | Klemas and Smart (1983) |
| | RMSI | $\frac{\rho 860}{\rho 150}$ | Hunt and Rock (1989) |
| | SRWI | $\frac{\rho 860}{\rho 1240}$ | Zarco-Tejada et al. (2003) |
| | LWVI1 | $\frac{\rho 1094 - \rho 983}{\rho 1094 + \rho 983}$ | Galvao et al. (2005) |
| | LWVI2 | $\frac{\rho 1094 - \rho 1205}{\rho 1094 + \rho 1205}$ | Galvao et al. (2005) |

Indices for estimation of vegetation LMA and EWT are combinations of reflectance at two or more bands in the spectral region of 900-2400 nm (Liu et al., 2017), where strong/ weak absorption of water and/or dry matter occurred. For EWT estimation, the near-infrared (NIR) and shortwave-infrared (SWIR) spectral regions were selected to build indices (Ceccato et al., 2002; Colombo et al., 2008; Yilmaz et al., 2008). It has been reported to be difficult to estimate LMA because of predominate absorption of water content in the spectral region of 1300-2400 nm (Li & Wang, 2011; Qiu et al., 2018). To suppress the impact of water, studies typically utilize the absorption of the C-H bond stretch at approximately 1700 nm (Wang et al., 2011). This explains why dry matter related indices usually involve spectral response at around 1700 nm (Feret et al., 2011; le Maire et al., 2008; Wang et al., 2011).

The indices have multiple types, ranging from simple ratios of different bands, normalized differences of two bands to complex combinations. Typically, simple ratios

and normalized differences are easier to comprehend and employ and thus are commonly utilized for estimating water and dry matter content. Table 2.2 summarizes commonly used indices exploited by relevant research. One typical dry matter index is normalized dry matter index (NDMI). It was developed by examining the spectral response of fresh leaves across a wide range of species (Wang et al., 2011). An alternative is to use vegetation indices for dry matter content (Bretas et al., 2021; Li et al., 2018).

## 2.3 Machine Learning and Random Forest

The growth of urban big data, advancements in computing power, and the availability of open-source software tools have contributed to the rapid development of Machine Learning (ML) techniques in recent years, making it an increasingly popular tool in a wide range of fields including vegetation monitoring. In recent years, ML techniques have been extensively employed in the remote sensing of leaf biochemical constituents, owing to their ability to address non-linear tasks.

ML is a subfield of artificial intelligence (AI) that involves the use of statistical models and algorithms to enable computer systems to improve their performance on a specific task by learning from data without being explicitly programmed. The aim of ML is to learn from data and make predictions or decisions based on patterns in the data they analyze, rather than relying on prior knowledge. There are a variety of ML algorithms, and depend on the specific task, they work differently and can be categorized to classifiers and regressors. The retrieval of vegetation biochemical constituents is categorized as a regression task, and for this purpose, regressors would be utilized. As aforementioned, very few studies have explored remote sensing of community DMCaP, and correspondingly, the use of machine learning techniques in this area is also scarce. Therefore, ML regression techniques used in relevant area (e.g., retrieval of leaf biochemical contents) would also be reviewed. Owing to resource and time limitations, not all ML algorithms would be covered in this thesis. Several popular algorithms will be identified and one ML algorithm would be chosen to build the model in this study.

Some renowned ML regression algorithms employed in remote sensing of vegetation are Artificial Neural Network (ANN), Support Vector Regression (SVR), Random Forest Regression (RFR).

Jin and Liu (1997) have demonstrated the capability of ANN to retrieve wheat dry matter fraction as well as other biomass parameters from high-dimensional (multi-day, dual-polarization, active and passive) RS data. A study by Feret et al. (2019) showed that SVR could provide accurate estimates of both LMA and EWT, but had weaker generalization ability than RT models. Yao et al. (2015) applied both ANN and SVR models to monitor wheat leaf nitrogen content, showing that SVR outperformed ANN in terms of accuracy. Shah et al. (2019) used RFR to do a stepwise reduction of input predictors among 2102 bands across the 400–2500 nm spectral range and 45 VIs, achieving significantly reduced errors than standard linear regression in the retrieval of leaf chlorophyll content in wheat. Bretas et al. (2021) also used RFR to predict DMCaP and aboveground dry matter in pastures in Brazil, suggesting that RFR has higher potential than linear regression.

A distinctive feature of RFR is its non-parametric characteristic, while ANN and SVR are both parametric algorithms. In training a ML algorithm, this feature is extremely important as the data distribution needs to be considered in parametric algorithms. They usually require scaled, and sometimes normalized data, while non-parametric algorithms such as RFR, do not require prior data preparation. Another important feature of RFR is that it usually requires fewer training samples than other algorithms (Ali et al., 2015). Other advantages of RFR are its ability to attain high accuracies, fast training even if on large datasets, and the unbiased out-of-bag accuracy (Douglass, 2020; Shah et al., 2019). Due to these characteristics, RFR is selected in this study to build the ML model using a relatively small dataset, which also enables fast estimation of DMCaP for agricultural purposes.

RFR is an ensemble learning method that combines predictions from multiple decision trees (DTs). The term DT is used for algorithm with a tree-like structure, where the child nodes are evaluated based on specific conditions, and the final result is presented in one of the leaf nodes. As shown in Fig 2.1, a Random Forest Regression is made up of N DTs, and subsets from the dataset would be assigned to DTs. Each DT is built using a single subset only, and the tree-building process continues until there are equal to or fewer than n samples in each node (typically, n is set to 5 for regression tasks). The final result for the regression task is obtained by averaging the predictions of each DT.

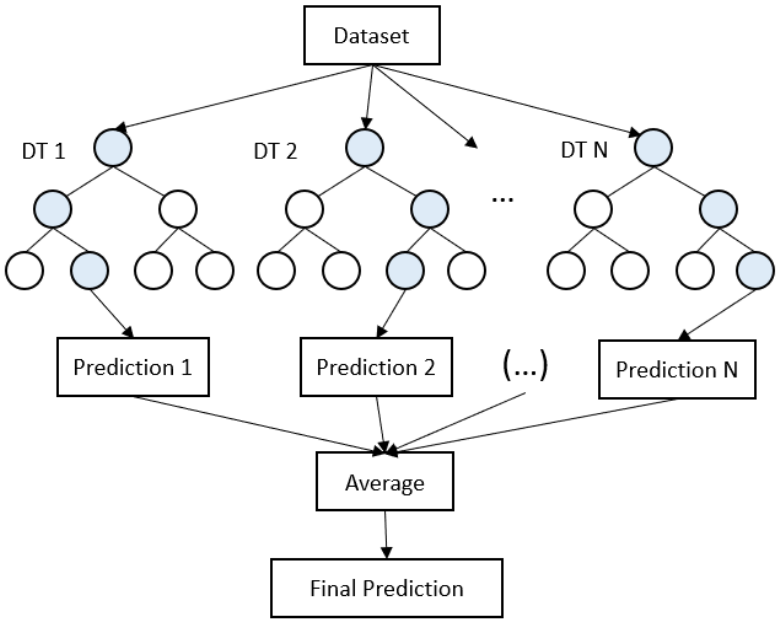


*Fig 2.1. Schematic of Random Forest Regression. All nodes at the have no child are leaf nodes, and others are split nodes.*

# 3. Data and methodology

The general workflow from data preprocessing until model evaluation is illustrated in Fig 3.1. It should be noted that the interpretable model was built using Sentinel-2 (S2) data alone, whereas the RFR model was built using S2 data, and a combination of S2 data and environmental variables. The interpretable model was built using data from all fields and individual fields, respectively, to explore its validity, as presented in Section 4.2.

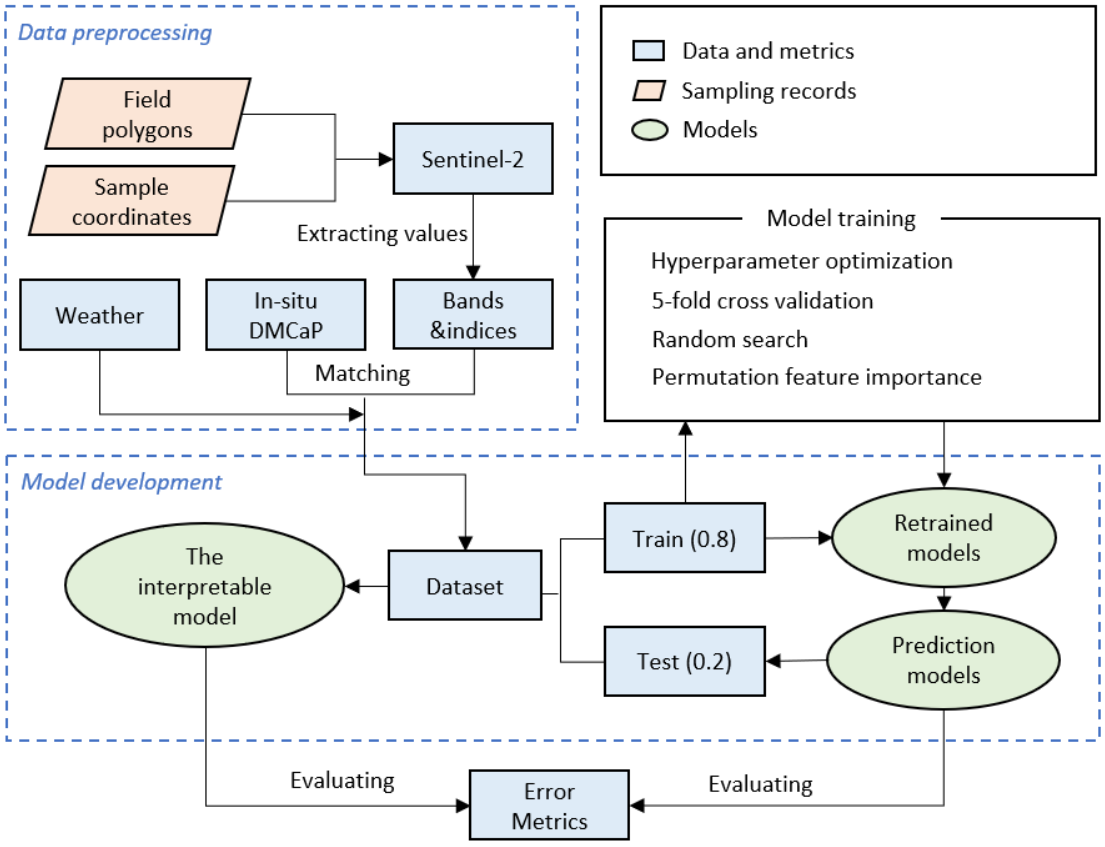


*Fig 3.1. General workflow. For data preprocessing, field polygons are used to clip Sentinel-2 (S2) satellite imageries and then sampling plots with unique geographic coordinates are used to extract S2 measurements on these imageries. Spatio-temporal matching is performed between S2 measurements and in-situ dry matter content as percentage (DMCaP) measurements. The dataset contains S2, in-situ measurements, and a homogeneous weather history for all four fields. Model development includes building statistical relationships for the interpretable model and training the Random Forest Regression (RFR) models. After model development, error metrics were computed to determine the optimal model to map DMCaP.*

## 3.1 Experimental area

The study is conducted with the help of Vultus AB, Sweden and TKS Agri, Norway. Field measurements of DMCaP were owned by TKS Agri, Norway. The experiment was carried out in four trial fields, of TKS Agri with a total area of 19.34 ha, all located near Stavanger city in the south-west of Rogaland County, Norway (Figure 3.2). These trial fields are named as Field 1-4 respectively, and are all in irregular shapes. Each field has similar soil quality and slightly different climatic conditions, and the latter will be described in more detail in the following contents.

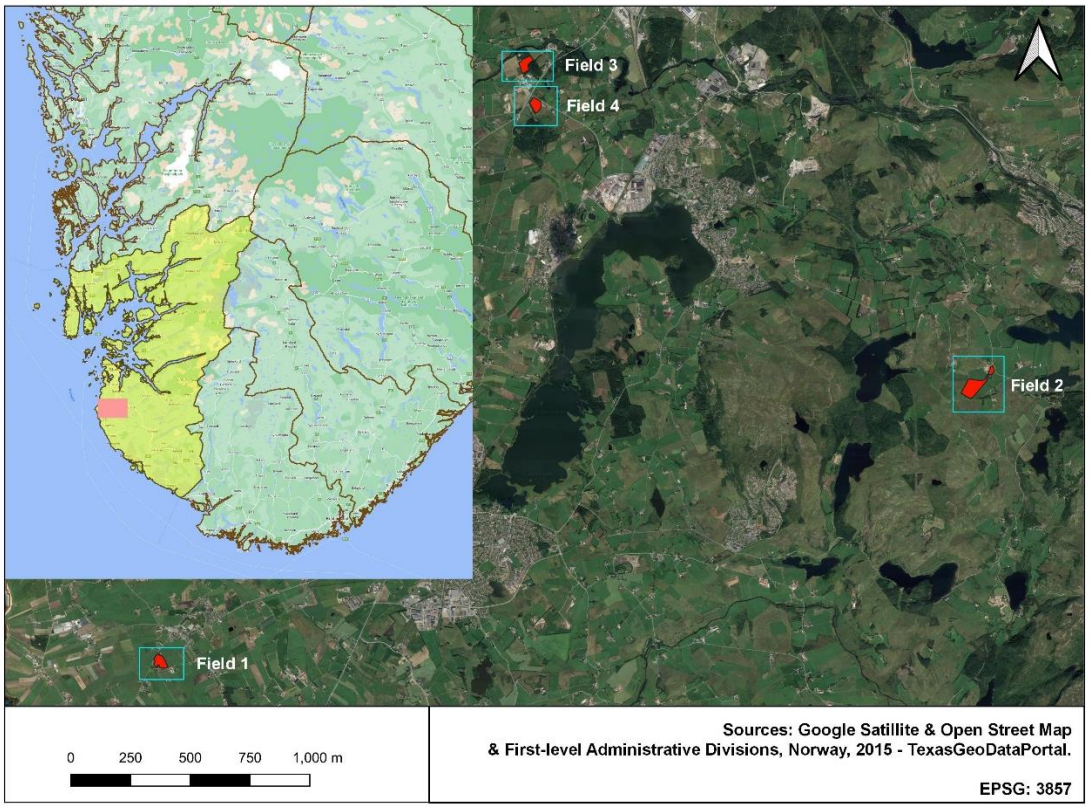


*Fig 3.2. Map of experimental area, with brown lines in the overview map for first-level administrative divisions in southern Norway, area rendered in yellow for Rogaland County, rectangle rendered in pink for an overview of experimental area, and polygons rendered in red for 4 trial fields.*

Field 1 (5.564700 °E, 58.713416 °N) is situated in the south-west of Stavanger city with an area of 3.63 ha. The terrain is overall flat, and the field has good sunlight conditions. This field tends to become very wet following heavy rain and often exposed to harsh weather conditions (wind and fog). The field has an average height of around

29m (+/- 1 m) above sea level (ASL), with oceanic climate due to its proximity to the sea. Raining season is from November to April, while the cattle-grazing season runs from mid-June to mid-September, with the growing period spanning from May to the mid-to-end of September. The field receives a long-term annual rainfall of 855.2mm, dropping to around 660 mm in dry years and rising to 1100 mm in rainy years. The grass grown in this field primarily consists of perennials (80-90%), including Ryegrass *Lolium perenne*, Blue grass *Poa pratensis*, White clover *Trifolium pratense*, Timothy *Phleum pratense*, and traces of Dandelions Taraxacum sp.

Field 2 (5.799364 °E, 58.754540 °N) is situated south of Stavanger city with an area of 8.89 ha. This field lies on a slope of a hill with average altitude of 156m ASL, tilted towards the southwest. Some areas of the field are often covered with water about 5cm in depth, despite having three drainage wells installed. The field has good sunlight condition but are prone to weather due to lack of hedgerows. The climate is oceanic, with rainfall typically occurring from November to April. On average, the field receives 1248.2mm of rainfall per year, dropping to 760mm in dry years and rising to 1100mm in rainy years. The growing period is from May to September and the cattle-grazing season is from mid-June to mid-September. The grass composition is mainly perennials (80-90%), including Ryegrass *Lolium perenne*, White Clover *Trifolium pratense*, Timothy *Phleum pratense*, and traces of Dandelions *Taraxacum sp*.

Field 3 (5.669885 °E, 58.754540 °N) is in the south of Stavanger, with an area of 3.67 ha. It is situated on a slope tilting towards the north-northwest, surrounded by forests from the two sides: north and south. The field is relatively dry and sunlight exposition partially is non-optimal. The altitude ranges from 9 to 21m ASL. The climate is mostly oceanic, with rainfall typically occurring from October to March. The grass growing period extends is from May until the middle or end of September. The grass composition is predominately perennials (80-90%), including Ryegrass *Lolium perenne*, Blue grass *Poa pratensis*, White Clover *Trifolium pratense*, Timothy *Phleum pratense*, and traces of Dandelions *Taraxacum sp*.

Field 4 (5.673124 °E, 58.796697 °N) is situated in the south of Stavanger, with an area of 3.15 ha, near Field 3. It is almost flat and levelled, with an evenly distributed growth of perennials and excellent sun exposure. The field has an average height of 11m ASL. The climate is oceanic, with rainfall primarily occurring from October to March. The

growing period is from May until the middle or end of September. The grass composition is mainly perennials (80-90%), including Ryegrass *Lolium perenne*, Blue grass *Poa pratensis*, White Clover *Trifolium pratense*, Timothy *Phleum pratense*, and traces of Dandelions *Taraxacum sp.*

Table 3.1 summarizes the conditions in all fields described above. Species distribution is relevant for determining community-level DMCaP; however, a thorough investigation of this has not been conducted. As the dominant vegetation types are the same in the four trial fields, we hypothesized that there is no significant difference in species distribution between the fields. The sampling method of DMCaP values will be described in the section 3.2.

*Table 3.1. A summary of traits of trial fields. (1. Ryegrass Lolium perenne, 2.Blue grass Poa pratensis, 3.White clover Trifolium pratense, 4.Timothy Phleum pratense)*

| Name | Area | Characteristics | Dominant vegetation types |
|---|---|---|---|
| **Field 1** | 3.63 ha | Overall flat terrain;<br>good sunlight condition;<br>annual rainfall of 855.2mm. | 1,2,3,4 |
| **Field 2** | 8.89 ha | Slope terrain;<br>extremely moist in some areas;<br>good sunlight condition;<br>annual rainfall of 1248.2mm. | 1,2,3,4 |
| **Field 3** | 3.67 ha | Slope terrain;<br>forests surrounded;<br>non-optimal sunlight condition. | 1,2,3,4 |
| **Field 4** | 3.15 ha | Overall flat terrain;<br>good sunlight condition. | 1,2,3,4 |

## 3.2 Field data

Field measurements of DMCaP values were collected in 56 sampling plots across all 4 trial fields, spanning from May to August 2022, with time intervals of 1-2 weeks. The sampling plots (10m*10m) were selected using stratified random sampling. Each plot with unique geographical coordinates was marked with poles in the center. DMCaP values were obtained using gravimetric method, which included cutting and weighing fresh grass samples, followed by drying the samples in the oven and calculate the ratio

of the weight of dried samples and fresh samples. Grass samples were cut off at around 5cm above the ground, which is the same height as mechanical harvesting, and dried in the oven in 100 °C for 2 hours. Based on the fact that all samples were collected through destructive cutting, each grass sample consisted of subsamples collected in a specific sampling plot instead of collecting in the exact same location each time.

The sampling plots could be different for different collecting dates, and all plots used are shown in Fig 3.3. The number of sampling plots and samples of each plot could be different for each field, details are described in Table 3.2. The dates that samples being collected could be different as well (see Table B.1).


*Fig 3.3. Spatial distributions of sampling plots, with Field 1 in top-left, Field 2 in top-right, Field 3 in bottom-left, and Field 4 in bottom-right.*

The samples included 200 DMCaP values in total and 7 values were excluded due to

long temporal gaps between field and RS measurements. The samples had a minimum value of 15.8% and a maximum value of 36.9%, and with most of them between 20%-30% (Fig 3.4).

A homogeneous weather history was retrieved from the nearest weather station and was used to represent the whole area. It contains mean air temperature, precipitation, radiation, soil temperature at 10 cm and 20 cm depth, relative air humidity, and wind speed.

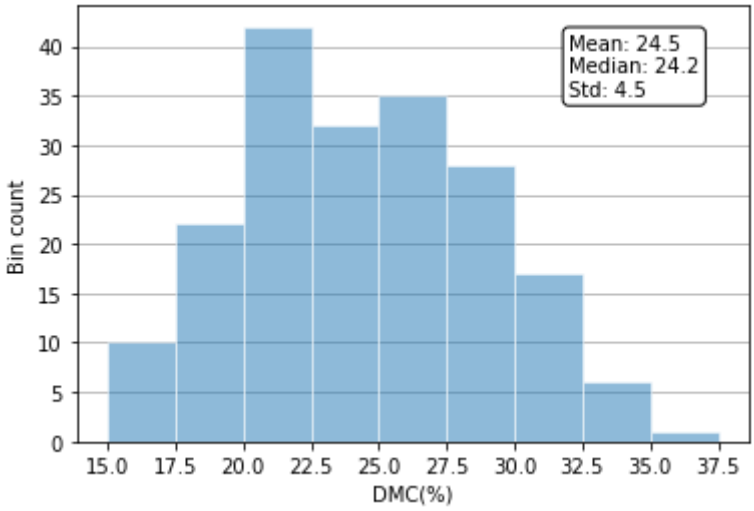


*Fig 3.4. Frequency distribution of in-situ measured dry matter content as percentage (N=193).*

*Table 3.2. Number of sampling plots of each field and number of samples of each plot, with comments showing the number of missing samples due to unexpected situation (No. = number).*

| Name | No. of plots | No. of samples per plot | Total samples collected | Fresh weight per sample/g | Comments |
|---|---|---|---|---|---|
| **Field 1** | 7 | 7 | 49 | 400 | |
| **Field 2** | 8 | 8 | 64 | 400 | |
| **Field 3** | 9 | 5 | 43 | 400 | 2 missing |
| **Field 4** | 9 | 5 | 44 | 400 | 1 missing |

## 3.3 Remote sensing data

Popular satellite RS systems applied in vegetation monitoring are described in section 2.1.2. The following content will explain why Sentinel-2 (S2) has been chosen to conduct this study. As the study material covers 4 trial fields at a regional scale, high

spatial resolution data is required. Moreover, high temporal resolution is necessary to effective estimation of DMCaP. In addition, the VIS, NIR, and SWIR bands are necessary for estimating dry matter content and water content, both of which are involved in the estimation of DMCaP. Due to these requirements, the optimal candidate is S2 of Copernicus program.

S2 is a wide-swath satellite RS system equipped with MultiSpectral Instrument (MSI), and has high resolution. Its purpose is to typically supports land monitoring studies including vegetation, soil, and water coverage monitoring, as well as inland waterways and coastal area observation. This mission comprises a twin polar-orbiting satellites (A&B) at a mean altitude of 786 km, phased at 180° to each other, decreasing the 10-day revisit time of one satellite to 5 days of the twin-satellite system at the equator, and 2-3 days in mid-latitude areas (SUHET, 2015). This system captures 13 wavelength bands of the EM spectrum, including 10 VIS-NIR bands and 3 SWIR bands. As described in 2.2.2, these MSI bands have been widely used for LMA and EWT estimation.

When satellite sensors capturing images, the atmosphere affects EM radiation from the sun before it reaches to the surface. This means that the radiation reflected back to the sensor is also affected and weakened, making atmospheric correction necessary. ESA has been distributing several product types of S2, which includes the S2 Level-2A (L2A) bottom-of-atmosphere (BOA) reflectance product. S2 L2A products, which were used in this study, are processed from Level-1C top-of atmosphere (TOA) products, by applying atmospheric, terrain, and cirrus corrections.

L2A products, besides BOA reflectance of 13 wavelength bands, also include scene classification layer (SCL) with eleven classes together with quality indicators (QA) for cloud and snow probabilities, aerosol optical thickness (AOT) and water vapor (WV). SCL provides 12 different classes including clouds, vegetation, and bare soils/desert. SCL was used in this study to distinguish between cloudy pixels and clear pixels.

Google Earth Engine (GEE) is a cloud-based platform developed and managed by Google Inc., which stores massive amounts of satellite imagery data. It provides a valuable tool for RS scientists, as well as for other users, to access data on a planetary scale. GEE also offers a high-performance computation service that allows users to process and analyze data within the same environment where the data is stored. The

data stored in GEE is gathered by Google Inc. directly from data providers such as NASA and ESA. In addition to accessing the public data catalog, users can also upload and merge their own data with the existing data to solve customized and more advanced problems.

S2 L2A reflectance of corresponding ground points are retrieved in GEE. In GEE, when matching ground-measured DMCaP values with their corresponding RS measurements, pixel values of unique geographic coordinates can be extracted without downloading the whole image. This was done using JavaScript in a web graphic user interface. To minimize the impact of cloud, the imagery was masked using SCL. Pixels classified to classes excluding vegetation and bare soil/ desert would be masked out and not retrieved. This brings to another detail that, on one sing date, several ground sampling plots would have two corresponding S2 images if they are in the overlap areas of two tiles. This means one DMCaP value could have two corresponding RS values. In this case, Field 1 is in the overlap area of tile 32VLL and tile 32VKL. To ensure the consistency of atmospheric correction, only tile 32VLL were used to retrieve S2 data and subsequent work as all four fields can be covered in tile 32VLL. When extracting pixel values from satellite images, the geographic coordinates of sampling plots were usually not located in the center of one pixel, a 5m-radius circle buffer to match the spatial resolution of 10m was created centered in each plot and the median of all 4 neighboring pixels was retrieved (Fig 3.5).

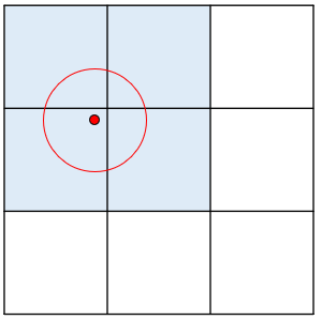

*Fig 3.5. Diagram of extracting remote sensing measurements based on location of sampling plots, with the red dots indicating the center of the sampling plot, the red circle for buffer created, and blue for neighboring pixels.*

A total of 75 S2 images from May to August 2022, which corresponded to the period of sample collecting, were retrieved. In these 75 images, some of the images may contain a large number of cloudy pixels, which were masked out using SCL. As a result,

pixel values from 35 images were retrieved for all sampling plots, which means one sampling plot would have a time series of RS measurements containing a maximum of 35 values, if this sampling plot was could-free in all 35 images. Fig A.1 shows the temporal gaps of these 35 images.

Besides the VIS, NIR and SWIR bands, indices were also calculated and used for build models. Two types of indices were calculated: vegetation and water indices, which represent dry matter and water content respectively. Table 2.2 describes indices that are sensitive to dry mater and water content respectively. However, a few of them might be directly calculated from S2 data because of its limitation of the spectral resolution. There is a wide range of VIs proposed for different purposes and validated using different datasets at varying levels. Each of these indices has its own strengths and weaknesses, and some are more optimal in certain applications than others. Table 3.3 summarizes all indices used in this study. Six VIs were used as substitutes for detecting dry matter content and NDWI (Normalized Difference Water index) was used for detecting water content. These VIs ranged from ratios, normalized differences, to more sophisticated forms. The rationale for including multiple vegetation indices is to compare and determine the optimal index for the specific case.

VIs were initially developed to quantify vegetation greenness by contrasting the high reflectance in the NIR region of the EM spectrum with the strong absorption by chlorophyll in the red region. Jordan (1969) proposed using the ratio of NIR to red band to estimate canopy chlorophyll content and leaf area index (LAI), which became known as the Simple Ratio (SR) or Ratio Vegetation Index (RVI). The RVI was further developed into the NDVI by Rouse et al. (1974), and it has since become the most commonly used vegetation index for various applications. There are other studies challenged the approach of using red band and used green band instead such as green normalized difference vegetation index (GNDVI) proposed by (Gitelson et al., 1996). Some have argued that it is at least five times more responsive to levels of chlorophyll-a compared to the NDVI and especially useful in distinguishing in vegetation that is under stress and senescence. It is known that NDVI can be sensitive to external factors, such as soils and atmospheric conditions, which can lead to potentially misleading results in some cases. Other indices that address these limitations were proposed, such as the soil adjusted vegetation index (SAVI) (Huete, 1988), atmospherically resistant vegetation index (ARVI) (Kaufman & Tanre, 1992). Huete et al. (2002) presented a

combination of these two as the enhanced vegetation index (EVI). Normalized difference red-edge (NDRE) utilizes the developments in spectral capabilities to obtain a more detailed characterization of the red-edge region, which is a significant spectral response of vegetation found between the red absorption maximum and high reflectance in the NIR region. Chlorophyll vegetation index (CVI) is also an index used for biomass monitoring.

NDWI is often using NIR-SWIR combination, which is used synonymously with NDMI (Normalized Difference Moisture Index). It seems that NDMI is consistently described using NIR-SWIR combination while NDWI is also described using GREEN-NIR combination. The indices with these two combinations function very differently, with NIR-SWIR highlighting variances in water content of leaves and GREEN-NIR highlighting variances in water level of water bodies (McFeeters, 1996). In this study, NIR-SWIR combination is referred as NDWI in order to separate from NDMI (Normalizes Dry Matter Index) mentioned in section 2.2.2.

*Table 3.3. A list of Sentinel-2 derived indices used for modelling.*

| Type | Name | Formula | S2 Formula |
|---|---|---|---|
| **Vegetation index** | RVI | $\frac{NIR}{Red}$ | $\frac{B8}{B4}$ |
| | NDVI | $\frac{NIR - Red}{NIR + Red}$ | $\frac{B8 - B4}{B8 + B4}$ |
| | GNDVI | $\frac{NIR - Green}{NIR + Green}$ | $\frac{B8 - B3}{B8 + B3}$ |
| | EVI | $2.5 * \frac{NIR - Red}{NIR + Red + 1}$ | $2.5 * \frac{B8 - B4}{B8 + B4 + 1}$ |
| | CVI | $\frac{NIR * Red}{Green * Green}$ | $\frac{B8 * B4}{B3 * B3}$ |
| | NDRE | $\frac{NIR - Red\_Edge}{NIR + Red\_Edge}$ | $\frac{B8 - B5}{B8 + B5}$ |
| **Water index** | NDWI | $\frac{NIR - SWIR}{NIR + SWIR}$ | $\frac{B8 - B11}{B8 + B11}$ |

## 3.4 Linking field and satellite measurements

Section 3.3 describes satellite RS measurements that were obtained, as well as the methodology used to retrieve these measurements. They are a collection of plot-based pixel values of all plots and all available sensing dates between May to August 2022. RS measurements are multi-dimensional, encompassing bands, sampling plots, fields,

and dates. In order to get them organized and match them to ground measurements of DMCaP values, each sampling plot was given a unique identifier which contained the information of field and sampling number. In this way, dispersed times series of RS measurements were retrieved for each plot. Despite the use of SCL mask (classes of soil/ desert and vegetation) in extracting RS measurements, the data is still noisy. Thus, a smoothing gap-filling, instead of interpolation alone, was applied to RS measurements of each plot to get smoothed and gap-filled time series (daily), which were then linked to field measurements based on dates. To summarize, RS measurements were smoothed and interpolated values while field measurements retrained their original values.

Smoothing spline was used to smooth and gap-fill the dispersed time series. It is a mathematical function that fits local piecewise third-order polynomials to small number of data points and then integrate them together so that they join smoothly(De Boor, 1978). The resulting smooth function is then used to estimate values between the given data points. Cubic spline smoothing could achieve a good balance between smoothing and retaining important features within the data, and thus selected to perform on the RS measurements. This process was done using the cubic smoothing spline function of Curve Fitting Toolbox in MATLAB. The smoothing parameter, s, ranges from 0 to 1, with 0 representing the least-squares straight line fit to the data and 1 representing cubic spline interpolation (*Cubic smoothing spline - csaps*, n.d.). Fig 3.6 shows time series of the original RS measurements and smoothed ones of one sampling plot in Field 1. A value between 0 and 1 represents a trade-off between the two extremes. As shown in Fig 3.6, a smoothing parameter between the range of 0.01 to 0.09 is sensible and the intermediate value, 0.05, was used to match RS and field measurements.

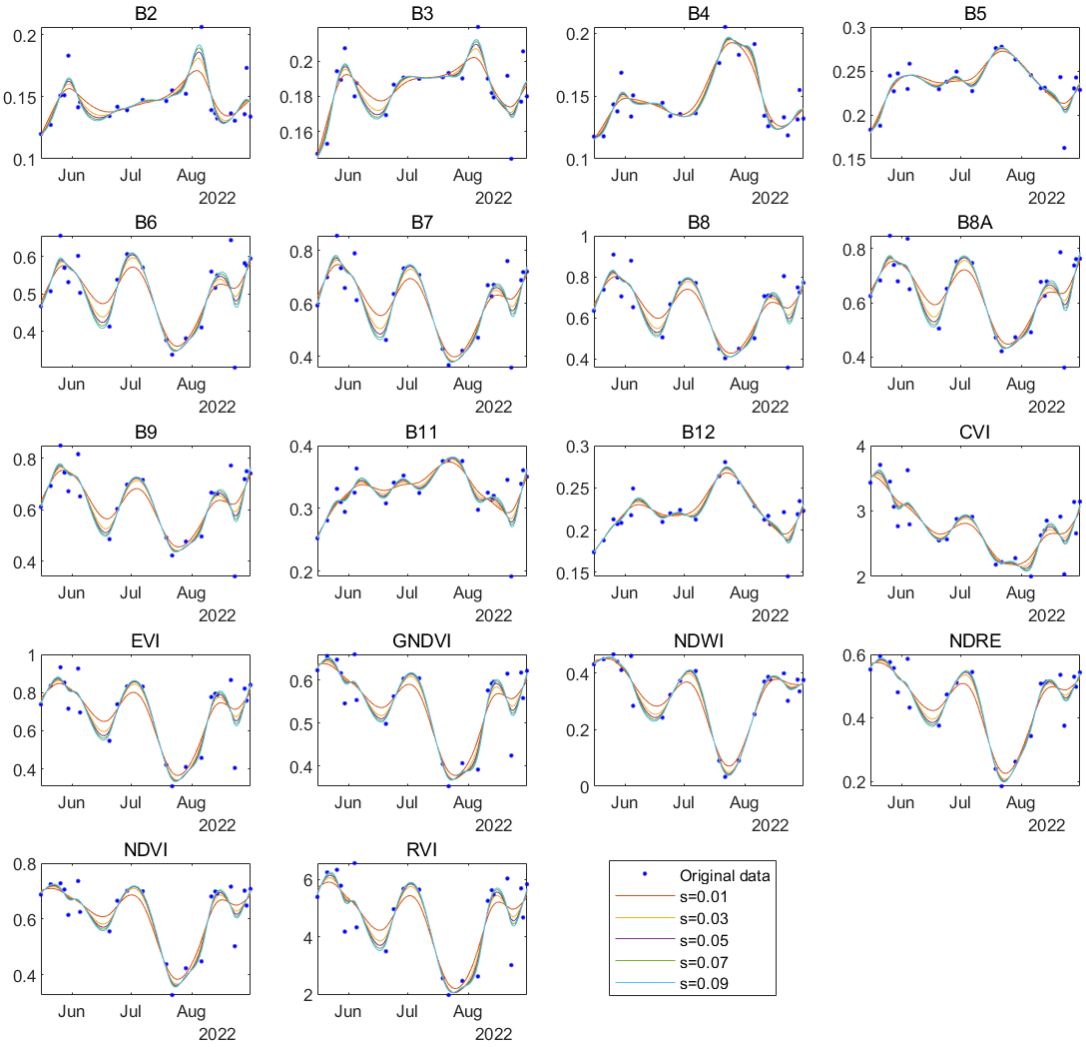


*Fig 3.6. Time series of RS measurements before and after smoothing in one sampling plot, with s indicating the smoothing parameter.*

## 3.5 Modelling

Based on literature review, two types of models, an interpretable model and a ML RFR model described in Section 2.3 were built. Both models were implemented using Python. The interpretable model was built using Python SciPy package and the RFR model was built using Scikit-learn package.

### 3.5.1 The interpretable model

DMCaP (DMC%) was calculated using the ratio of dry weight and fresh weight. It applies to at both leaf-level (Eq. 1) and canopy-level (Eq. 2). As aforementioned in Section 2.2, LMA is a metric used for determining leaf dry matter content and EWT for

leaf water content.

$$DMC\% = \frac{\mathrm{W_m}}{\mathrm{W_m} + \mathrm{W_w}} = \frac{\frac{\mathrm{W_m}}{A}}{\frac{\mathrm{W_m}}{A} + \frac{\mathrm{W_A}}{A}} = \frac{\mathrm{C_m}}{\mathrm{C_m} + \mathrm{C}_w} \quad (1)$$

where $\mathrm{W_m}$ is leaf dry matter weight (g), $\mathrm{W_w}$ is the leaf water weight (g), A is the leaf area ($cm^2$), $\mathrm{C_m}$ is the LMA ($g/cm^2$), and $\mathrm{C}_w$ is the EWT ($g/cm^2$).

$$DMC\% = \frac{\mathrm{C_m} * LAI}{\mathrm{C_m} * LAI + \mathrm{C}_w * LAI} \quad (2)$$

where LAI is the leaf area index ($cm^2/cm^2$), and canopy dry matter content and water content are LMA and EWT multiplied LAI respectively.

Data-driven methods for estimating dry matter (or dry biomass) and/or water content either at leaf- or canopy-level, rely on building a linear (Boutton & Tieszen, 1983; Bretas et al., 2021; Li & Guo, 2018; Sow et al., 2013) or an exponential relationship (Romero et al., 2012), or both (Yang et al., 2020) between biochemical continents and the indices. The linear function is the simplest and most widely used and it was adopted in this study. The difference between leaf dry mass and aboveground dry biomass is that the latter refers to the stems as well. These two terms will not be strictly distinguished as DMCaP referred to the measurements of a fraction of above-ground parts (cut off at around 5cm above the ground), and grass does not have prominent aboveground stems and those stems are green as well.

The interpretable model (Eq. 3) was built based on the assumptions that there is a linear relationship between: (i) dry biomass and VIs and, (ii) water content and water index.

$$\begin{aligned} DMC\% &= \frac{\mathrm{scalar_a} * \mathrm{VI} + \mathrm{offset_a}}{\mathrm{scalar_a} * \mathrm{VI} + \mathrm{scalar_b} * \mathrm{WI} + \ \mathrm{offset}_c} \\ &= \frac{(\mathrm{scalar_a} * \mathrm{VI} + \mathrm{offset_a})/\mathrm{scalar_a}}{(\mathrm{scalar_a} * \mathrm{VI} + \mathrm{scalar_b} * \mathrm{WI} + \ \mathrm{offset}_c)/\mathrm{scalar_a}} \\ &= \frac{\mathrm{VI} + \frac{\mathrm{offset_a}}{\mathrm{scalar_a}}}{\mathrm{VI} + \frac{\mathrm{scalar_b}}{\mathrm{scalar_a}} * \mathrm{WI} + \frac{\mathrm{offset}_c}{\mathrm{scalar_a}}} \\ &= \frac{\mathrm{VI} + \beta_1}{\mathrm{VI} + \beta_2 * \mathrm{WI} + \ \beta_3} \end{aligned} \quad (3)$$

where, $\mathrm{scalar}$ is the slope, offset is the intercept, VI is the vegetation index, WI is the water index, and $\mathrm{offset}_c$ in the denominator is the sum of offset for VI and WI. By dividing both numerator and denominator by $\mathrm{scalar}_a$, we get a simplified formula (Eq. 3) with three coefficients ($\beta_1$, $\beta_2$, and $\beta_3$) to estimate in the parameterization process. This reduces the complexity of the model and can make it more robust. VI and WI applied were described in table 3.3. In this non-linear fractional equation, the numerator represents dry biomass, and the numerator represents fresh biomass.

### 3.5.2 Random Forest Regression

The RFR algorithm was implemented in this study to estimate DMCaP values. An 8:2 ratio was used to split the dataset into training and testing set. The training set was used to train and validate the model, and the testing set was used to evaluate its performance. RS measurements, including band 2-4 (VIS), band 5-7(Red-edge bands), band 8-9 (NIR), band11-12(SWIR), and dry matter index and water index listed table 3.3 were fed in to RFR model as predictors, with DMCaP as the response variable. A second RFR model incorporated both RS measurements and environmental predictors. The input predictors were then ranked according to permutation feature importance, and a step-wise reduction in the number of input predictors was adopted**.**

#### (1) Cross-Validation

Cross-validation (CV) is a technique used to evaluate the ability of a model to predict the 'unseen' data, which is not used to train the model. It is useful when we have limited amount of data, and it can detect the problem of over-fitting as well. The measurements used in this study are limited and are unevenly spatiotemporal distributed within fields and between fields (Fig A.2, and Table 3.2). Therefore, a 5-fold cross-validation was performed on the training set. Fig 3.7 shows the general work flow of a 5-fold CV. The training dataset is split into five equal parts, and each is called a fold. Then the model is trained using four folds and validated using the trained model on the leaf-out fold. This process will be iterated five times until each fold has been a validation fold once. The model performance is usually summarized by averaging the error in each iteration, and the commonly used metric is root mean squared error (RMSE).

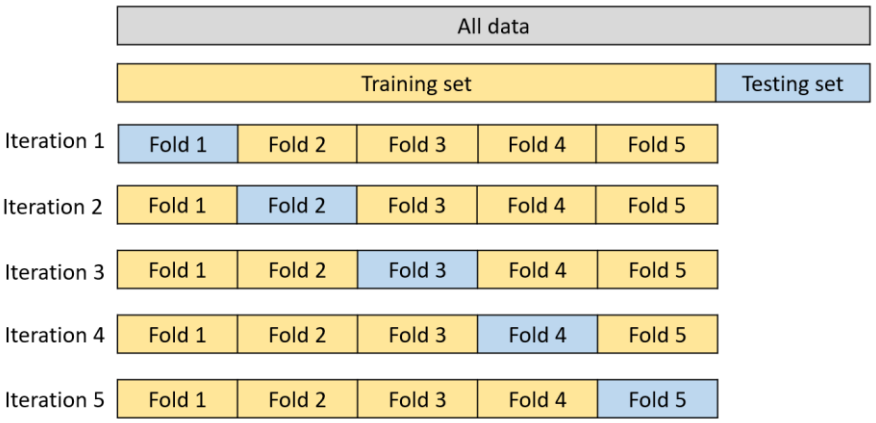


*Fig 3.7. Schematic diagram of 5-fold Cross-Validation (Modified from Pedregosa et al. (2011)).*

**(2) Hyperparameter tuning**

In ML models, there are two types of parameters, model parameters and hyperparameters. The former type is parameters that being initialized and adjusted during the training phase. The hyperparameters are user-defined parameters to structure a model before training the model, which cannot be learned directly from the data, and they will remain constant during the learning process. RFR algorithm is a non-parametric method, therefore, a more detailed description of model parameters will be omitted here.

The optimal hyperparameters are criteria-specific, and the process of exploring the range of combinations and finding the optimal hyperparameter configuration is called hyperparameter tuning (Bergstra et al., 2011). Two of the most popular methods of optimizing hyperparameters are grid search and random search proposed by Bergstra and Bengio (2012). The success of these techniques is mainly attributed to their simplicity, ease of implementation, and better performance over purely manual optimization. When performing grid search, the model is trained on each combination of hyperparameter values in a set of possible hyperparameters placed in a matrix-like structure, to find the optimal configuration. Random search is similar to grid search, but instead of conducting an exhaustive search on all possible combinations, it randomly selects them based on a set of hyperparameters. The number of total runs can be specified by the user. Compared to grid search, random search can find the optimal configuration by effectively searching a broad range of values (Bergstra & Bengio, 2012). Hence, random search was used in this study to perform 100 combinations on a 5-fold CV, leading to a total of 500 fits. The optimal configuration found by random

each is listed in the appendix (Table B.2).

## 3.6 Model evaluation

Recommendations for commonly used metrics of evaluating the performance of a model can be found in a study by Seegers et al. (2018). In this study, a direct and robust combination of the root mean square error (RMSE), the mean absolute error (MAE), bias, and the coefficient of determination ($R^2$) was used. RMSE (Eq. 4) is the square root of mean squared difference between the true and predicted values, and it indicates how far predictions fall from measured true values using Euclidean distance. The MAE (Eq. 5) is the mean absolute error between the true and predicted values. MAE varies with the average error, and RMSE differs from MAE by considering the distribution of error magnitudes and the size of the sample. In other words, RMSE addresses its sensitivity to data distributions and outliers. Bias (Eq. 6) quantifies the average difference between the true and predicted values, and it is an indicator of the direction of systematic error as either overestimating or underestimating the predicted values. $R^2$ (Eq. 7) explains the how much of the variance of the dependent variable is explained by the independent variables. It could be an indicator of goodness of the fit.

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

$$MAE = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{n} \tag{5}$$

$$Bias = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)}{n} \tag{6}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{7}$$

where, $n$ is the number of observations/samples, the subscript $i$ denotes individual data points, $y$ is the true values, and $\hat{y}$ is the predicted values.

# 4. Results

## 4.1 The interpretable model

Table 4.1 provides an overview of the metrics used to evaluate the performance of the interpretable model. This model was built using data of all fields and data of individual fields respectively. The evaluating metrics for the model using individual fields as the input can be found in the Appendix (Table B.3).

*Table 4.1. Performance of the interpretable model using different combinations of water and vegetation indices. NDWI was the only used vegetation water index, and vegetation indices used were those listed in Table 3.3. RMSE, MAE and Bias can be interpreted as DMC (%). All fields is referring that all samples from Field 1-4 were put in the model (running once) and fields average is referring that samples from Field 1-4 were put in the model respectively (running four times) and the average of them was calculated. |Bias| is calculated using the absolute values of individual bias of fields, to avoid cancelling out of positive and negative values.*

| | **All fields** | | | | **Fields average** | | | |
|---|---|---|---|---|---|---|---|---|
| | **RMSE** | **MAE** | **Bias** | **R2** | **RMSE** | **MAE** | **\|Bias\|** | **R2** |
| **RVI** | 4.42 | 3.70 | 0.0000 | 0.03 | 3.92 | 3.24 | 0.0028 | 0.11 |
| **NDVI** | 4.38 | 3.62 | -0.0007 | 0.04 | 3.87 | 3.19 | 0.0032 | 0.13 |
| **GNDVI** | 4.43 | 3.71 | 0.0004 | 0.02 | 3.89 | 3.18 | 0.0044 | 0.13 |
| **EVI** | 4.43 | 3.71 | 0.0000 | 0.02 | 3.96 | 3.33 | 0.0032 | 0.10 |
| **CVI** | 4.48 | 3.74 | -0.0000 | 0.00 | 3.97 | 3.32 | 0.0009 | 0.10 |
| **NDRE** | 4.42 | 3.70 | 0.0004 | 0.03 | 3.80 | 3.09 | 0.0037 | 0.16 |

In general, the performance of the interpretable model built using either all data or data of individual fields exhibited high uncertainties. Using data of individual fields to build the model led to minor improvements. There was no significant difference in model performance using different water index and vegetation index combinations. $R^2$ can be understood as the reverse situation of standard deviation of the residuals, and it approaches 1 standard deviation of the residuals decreases. It equals to 1 when the model predicts values absolutely accurate without any error, which is rarely observed in real cases. A more common situation is a value between 0 and 1, a value greater than zero suggests a greater predicting ability of the model than simply using the mean of true values. As shown in Table 4.1, for using all data, $R^2$ score is slightly larger than zero, indicating a close-to-mean predicting ability of the model. Bias is close to zero,

and RSME and MAE are around 4.4% and 3.7% respectively. Models built on data from individual fields respectively had minor improvements on RMSE, MAE and $R^2$. We could infer that the interpretable model is not robust, and it exhibits improved performance when using highly local-scale data.

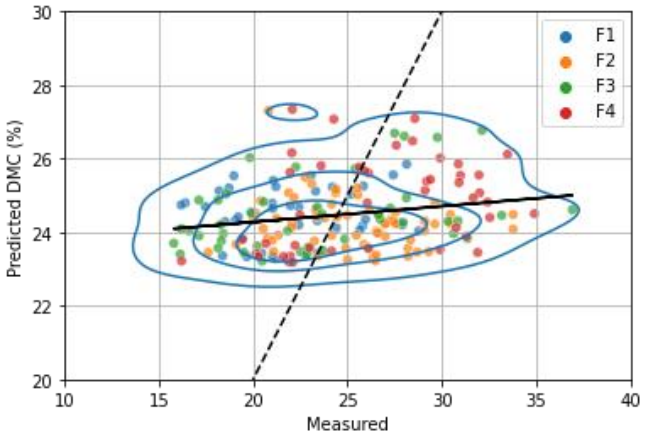


*Fig 4.1. Scatter plot showing the performance of the interpretable model built on all data using NDVI and NDWI combination. The x axis shows the measured DMCaP and the y axis the predicted DMCaP. Contour lines are included to show the data bivariate distributions. The dashed black line (equality line, y=x) and solid black line (fitted slope) are also included to a better illustration of $R^2$. The legend F1-4 represents data of Field 1-4.*

The performance of the interpretable model built on either all data (Fig 4.1) and data of individual fields (Fig 4.2) using NDVI and NDWI combination are further illustrated in scatter plots. The fitted slopes are generally closer to the equality line when the model is built on data of individual fields, where $R^2$ rose from nearly zero to above 0.10. The data points together with contour lines, compared to the equality line indicate poor performance with either overestimation or underestimation of DMCaP, with few points centered around the equality line. Also, an overall narrower range of predicted DMCaP values was observed (Fig 4.1 and Fig 4.2). As shown in Fig 4.1, measured DMCaP values range from 15% to 37% while predicted values range from 22% to 28%. Fig 4.3 shows the residuals of the interpretable model built on all data using NDVI and NDWI combination. The residuals are randomly scattered around the zero with no clear pattern and have fairly constant variance. However, some residuals are quite large, suggesting significant prediction errors.

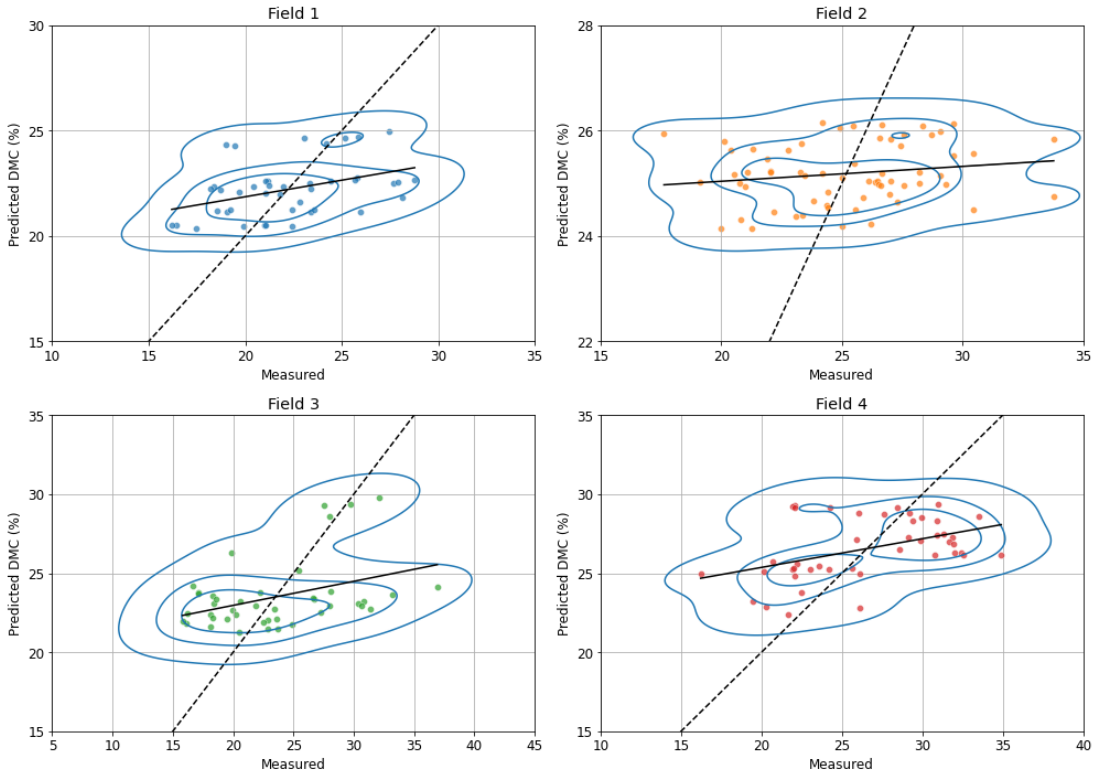


*Fig 4.2. Scatter plots showing the performance of the interpretable model built on individual fields using NDVI and NDWI combination.*

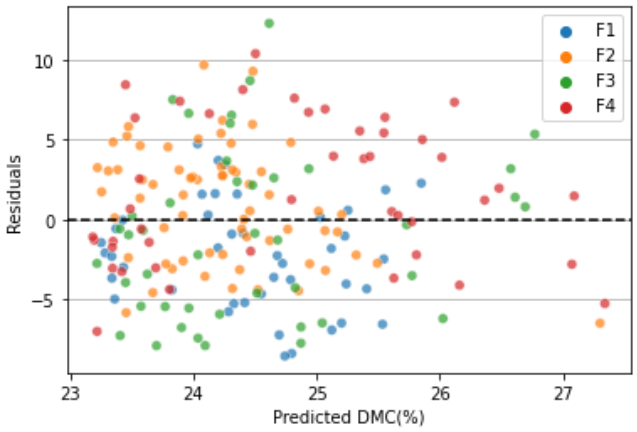


*Fig 4.3 Residua plot of the interpretable model built on all data using NDVI and NDWI combination. The legend F1-4 represents data of Field 1-4.*

## 4.2 The RFR model

Table 4.2 provides an overview of the performance of the RFR model evaluated on the testing set. The model with both RS measurements and environmental variables as predictors has lower RMSE, MAE, and bias, as well as a larger $\mathbf{R^2}$ compared to the model with only RS measurements as predictors. The CV-score (RMSE) based on the

training set is 3.09% for the former and 4.14% for the latter. As aforementioned in Section 3.5.2, a step-wise reduction in the number of input predictors was used. The number of input predictors was reduced to 12 eventually by observing permutation feature importance. The RFR model has slightly lower RMSEs and MAEs but higher biases compared to the interpretable model when using the entire dataset. Overall, the RFR model outperform the interpretable model, and the one with environmental predictors performs the best among the three. Notably, the $R^2$ values for all three models are relatively low to moderate. Given the fact of low volume and high complexity of the data, it would be premature to conclude that the models are ineffective. This will be further discussed in the next chapter.

*Table 4.2. Performance of the Random Forest Regression (RFR) model. RMSE, MAE and Bias can be interpreted as DMC (%).*

| | **Sentinel-2 Data** | | | | **+ Environmental Variables** | | | |
|---|---|---|---|---|---|---|---|---|
| | **RMSE** | **MAE** | **Bias** | **$R^2$** | **RMSE** | **MAE** | **Bias** | **$R^2$** |
| **RFR** | 3.88 | 3.14 | -1.09 | 0.06 | 2.90 | 2.43 | -0.87 | 0.47 |

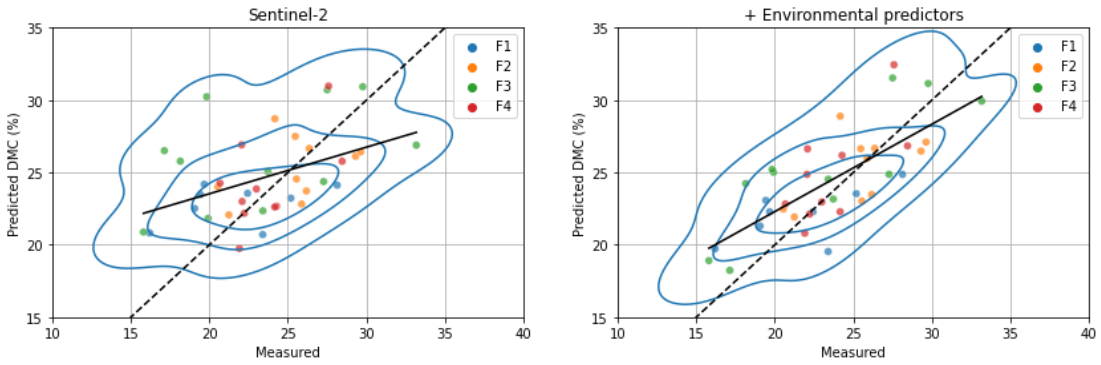


*Fig 4.4. Scatter plot showing the performance of the Random Forest Regression (RFR) model on the testing set. Contour lines are included to show the data bivariate distributions. The dashed black line (equality line, y=x) and solid black line (fitted slope) are also included to better illustration of $R^2$. The legend F1-4 represents data of Field 1-4.*

As shown in Fig 4.4, despite the limited number of testing data points, the RFR model demonstrates an acceptable performance. The majority of the data points are centered around the equality line (small residuals) and the fitted slopes are close to the equality line (good explained variance).

The distribution of the residuals further illustrates better performance of the RFR model

with environmental predictors (Fig 4.5). The interquartile range and whiskers are smaller with the RFR model trained using both Sentinel-2 data and environmental variables as predictors.

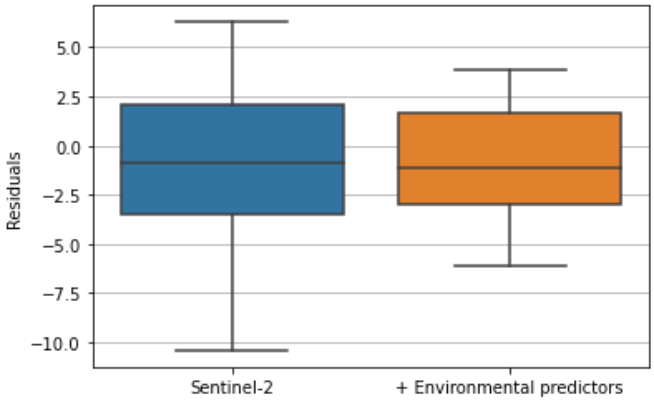


*Fig 4.5. Box plot of residuals of the Random Forest Regression (RFR) model using only Sentinel-2 data and using Sentinel-2 data and environmental variables as predictors.*

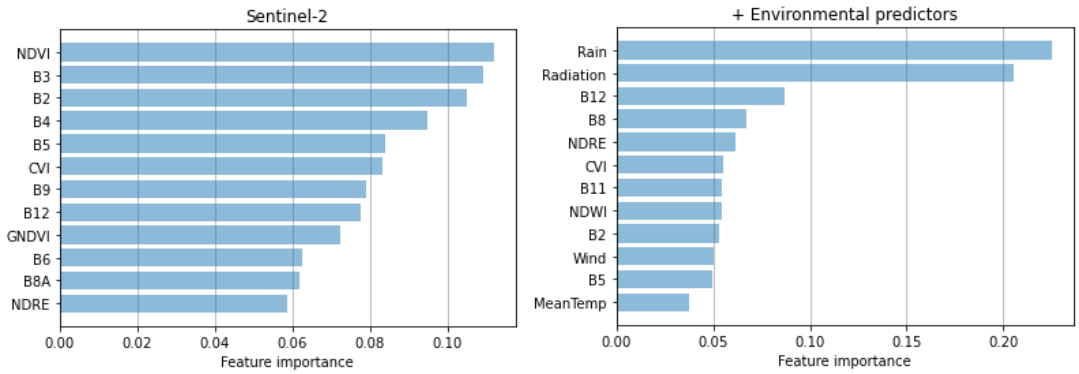


*Fig 4.6. Feature importance of the Random Forest Regression (RFR) model using only Sentinel-2 data and using Sentinel-2 data and environmental variables as predictors. Feature importance is quantified by the increase or decrease in mean squared error (MSE) when we permute a feature. Be aware of that the scores do not indicate the inherent predictive value of an individual feature, the but rather highlight the importance of the feature for a particular model.*

Fig 4.6 shows the permutation feature importance of the RFR model. For the model trained using only Sentinel-2 data, including NDVI and the VIS bands (B2-B4) leads to the highest MSE reduction, suggesting NDVI and the VIS bands are the most important predictors for the model. For the model trained using both Sentinel-2 data and environmental variables, including precipitation and radiation leads to the highest MSE reduction, followed by SWIR (B12) and NIR (B8) bands. This indicates that the RFR model tends to rely more on environmental predictors compared to Sentinel-2 data.

## 4.3 Visual assessments

Based on model evaluation, it has been identified that the RFR model using Sentinel-2 data and environmental variables as predictors outperforms both the RFR model using only Sentinel-2 data as predictors and the interpretable model. Hence, is selected as the best model for estimating field-scale DMCaP. As shown in Fig 4.7, Field 2, which has the largest area and the most sampling plots among all four fields, has been selected for demonstration. The model was applied to map the DMCaP dynamics from May to August, 2022. Sentinel-2 images were downloaded from GEE, and any images with cloud cover were excluded. A total of six images were selected to showcase estimating DMCaP at the field-scale.

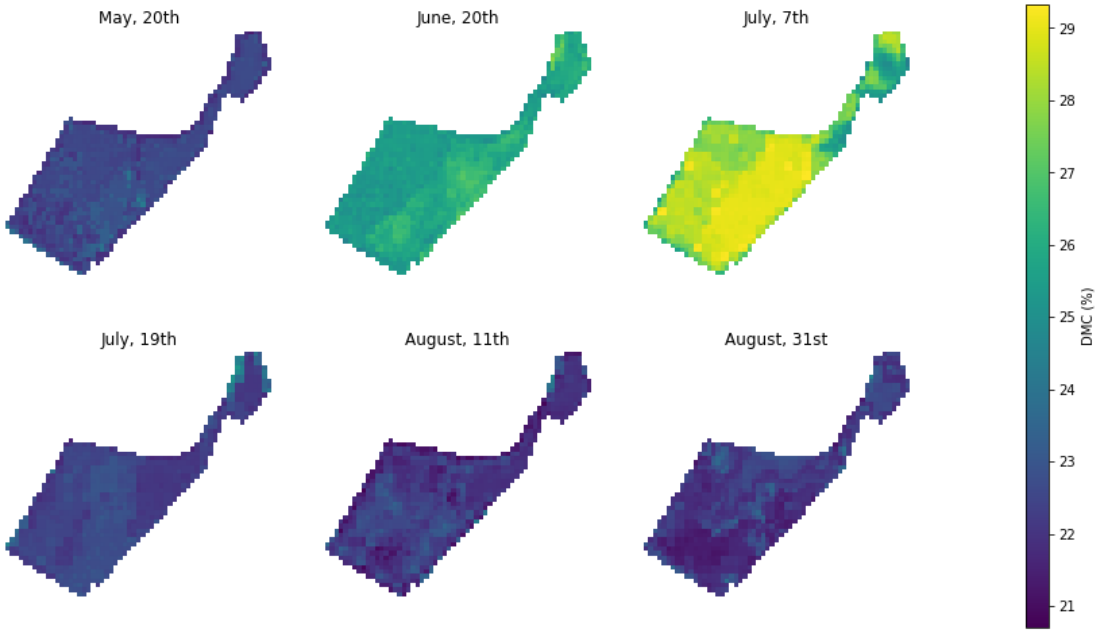


*Fig 4.7. Estimation of dry matter content as percentage (DMCaP) in Field 2 on six dates in 2022 utilizing the Random Forest Regression (RFR) model using both Sentinel-2 data and environmental variables as predictors.*

The estimated DMCaP reveals spatial and temporal variances. On June 20$^{th}$ and July 7$^{th}$, higher DMCaP values are observed throughout the field compared to other dates, particularly in the south-eastern part. No distinct spatial pattern or obvious anomalies are observed.

# 5. Discussion

## 5.1 Model performance evaluation using $R^2$

The evaluating metrics of the interpretable model built on individual fields and the RFR model suggests acceptable prediction errors (RMSE and MAE) and systematic errors (Bias), but poor to moderate explained variance (small $R^2$) of the DMCaP by Sentinel-2 data and environmental variables. The interpretation of $R^2$ would be thoroughly discussed in the following paragraphs.

$R^2$ represents the proportion of the variance of the dependent variable that is explained by the independent variables in the model. The best possible score is 1.0, and it is not necessarily the square of correlation efficient (R) and it can be negative if the model is arbitrary worse. A model that consistently predicts the mean of the dependent variable without considering the input features would get a score of zero. $R^2$ can be simply understood as using the mean value as the error benchmark to determine whether the prediction is better or worse than the benchmark. However, a high $R^2$ does not necessarily indicate a good model. One example is when doing multiple regression, adding irrelevant predictor variables to the model will always increases $R^2$ regardless of their contribution to the model, where larger $R^2$ is indicating overfitting (Figueiredo Filho et al., 2011). Likewise, a low $R^2$ does not imply the model is bad as it is a reflect of both the model and the data. $R^2$ will be poor even in a perfect model if there is large variance in the noise term.

It has been long and widely acknowledged in mathematical literature that $R^2$ is an inadequate metric for evaluating the goodness of fit for nonlinear models. Nevertheless, it is still been frequently used in agricultural and environmental remote sensing fields for analyzing and interpreting nonlinear models (Bretas et al., 2021; Jing et al., 2009; Lv et al., 2014). Spiess and Neumeyer (2010) have conducted thousands of simulations in their research, demonstrating how using $R^2$ to assess the performance of nonlinear models leads to incorrect conclusions.

To wrap up, a high $R^2$ score does not necessarily suggesting a good model and a low $R^2$ score does not necessarily suggesting a good model. $R^2$ is being good for summarizing the strength of a relationship as a general statistic, but not a decisive statistic. This means that, with all conditions being equal, a model that explained 80% of the variance

is more likely (not absolutely) to be better than one that explains 20% of the variance. Solely relying on $R^2$ score to evaluate the model can lead to incorrect conclusions as there are other factors (e.g., noise in the data) influencing the score and causality varies case by case. Case-specific context needs to be taken in consideration and, other evaluating metrics are usually involved to assess a model.

In this study, regarding to the interpretable model, $R^2$ scores of the model built on 4 individual fields are higher than that of the model built on data of all fields. With the model structure and input variables being equal, we can infer that the samples from individual fields lead to different $R^2$ scores. The interpretable model exhibits better performance on individual fields, suggesting that it is more suitable with highly local-scaled data. Together with low $R^2$ score of the RFR model using only Sentinel-2 data as predictors, we argue that, part of the reason on low $R^2$ values lies in massive noise and uncertainties in the dataset so a high $R^2$ is hard to achieve. Further discussion regarding to the data and models themselves will be presented in Section 5.2 and Section 5.3. If we do not see $R^2$ as a decisive metric, the RFR model retains its relevance and utility despite low $R^2$ scores, as the RMSE, MAE, and bias values are within an acceptable range for agricultural applications.

## 5.2 The estimation models

For the interpretable model, we conducted several additional tests using different input samples. The general observation is that the model is not robust and is very sensitive to the input data. The coefficients and constants can vary drastically depending on different input samples. Given the nonlinear characteristic of the interpretable model, the final estimation of coefficients and constants is also influenced by the initial guesses and the number of iterations, which is out of scope of this study and will not be further discussed. The dependency on these factors of the interpretable model necessitates artificial verification of coefficients and constants based on empirical experience from previous fittings, which introduces subjectivity.

Moving on to the formula of the interpretable model, DMCaP is calculated by using ratio of dry weight and fresh weight. The main concept behind this model is to quantify the dry mass and water mass using vegetation and water index respectively. The model is built on the assumption that there is a linear correlation between two key elements: (i) dry biomass and VIs, and (ii) water content and water index. Although there are

studies suggesting a relationship between dry mass and VIs (Boutton & Tieszen, 1983; Bretas et al., 2021; Kong et al., 2019), it is important to note that this relationship is not always linear. Typically, these studies involve preliminary data exploration before starting to develop statistical models. The relationship between vegetation biochemical constituents and indices depends on various factors such as vegetation species, experimental areas, and growth stages. In this study, the ideal approach is to estimate dry mass and water or fresh mass separately, and then calculating DMCaP. However, due to unavailability of the qualitative data on dry mass and water mass of the vegetation, we relied on VIs and the water index to estimate a quality value (DMCaP) based on certain assumptions. Uncertainty is introduced in both the numerator and denominator of the estimation. It becomes challenging to determine whether the assumptions are incorrect, or if the presence of significant uncertainties and noise is hindering satisfactory model performance.

Studies (Feret et al., 2021; Romero et al., 2012) using radiative transfer models have revealed that the optical properties of vegetation are influenced not only by biochemical constituents but also by other parameters, such as the vertical structure and density of vegetation. This introduces an additional layer of uncertainty to the model.

For the RFR model, we decreased the number of predictors in the training process by observing permutation feature importance. It is important to note that the method of permutation feature importance does have certain limitations. It tends to prioritize numerical features and categorical features with a high number of unique values (high cardinality). When dealing with correlated features, it may choose one feature over the other and disregard the significance of the second feature, potentially resulting in misleading conclusions (Płoński, 2020).

By looking at the permutation feature importance, one of the commonly raised questions is: What are the important predictors that the RFR model uses to make the predictions? From a machine learning perspective, the answer is: They are actually a combination of many predictors. As shown in Fig 4.6, NDVI and precipitation are the most important predictors for the two RFR models respectively. However, we should not try to associate the inherent predictive value of NDVI or precipitation to DMCaP group differences, even if such association exists. There is a clear difference between association and prediction studies: the former aims to gain a better understanding of a

phenomenon while the latter aims to make accurate predictions.

Another thought regarding cross-validation (CV) in the training process of the RFR model is that, for time series data, the recommended approach is to use walk-forward validation. This entails ensuring that the validation data always comes after the training data in terms of time. In this study, walk-forward validation is not implemented as it is challenging due to uneven time intervals (see Appendix A.2) and the dataset consisting of samples from four different trial fields.

It is important to note that the models were developed using localized samples from non-natural forage grasslands. As DMCaP characterizes community-level traits associated with species distribution and anthropogenic activities, the generalization ability to larger scales and diverse land types may be limited.

## 5.3 The quality of input data

The dataset is a combination of RS optical data, in-situ measurements of DMCaP, and weather history. In the following discussion, we will address three aspects: the quality of in-situ data, the quality of RS data, and the matching of RS and in-situ data pairs.

Regarding the in-situ data, it includes the geographic coordinates of each sampling plot and the corresponding DMCaP value. There are GPS positioning errors when recording the coordinates. Each sample consists of manually harvested subsamples within the plot, without strict control over the total weight of each sample. The selection of subsamples relies on visual judgment, which introduces subjectivity. Considering the rather even spatial distribution of samples within each field and the fact that DMCaP is a quality value rather than a quantity value, we argue that the overall quality of the in-situ data is sufficient for agricultural purposes.

As for the RS data, L2A BOA reflectance was used to minimize the atmospheric interference, and SCL was employed to mask out cloudy pixels. Although errors may exist in both BOA reflectance and SCL, we still consider the L2A products to have overall good quality.

When it comes to matching RS and in-situ data pairs, larger uncertainties and/or errors are introduced. The geographical location of the trial fields experiences a high frequency of cloudy weather, resulting in a limited number of available S2 images for

analysis. There are significant temporal gaps exceeding 5 days between successive S2 images used for extracting satellite measurements (see Fig A.1). Noise is present in the RS data, and the occurrence of two harvesting events adds another level of complexity to the time series. The spatial resolution of RS data is at most 10 m and may not well represent the optical properties of according field sampling plots. The introduction of artificially selected smoothing parameters further contributes to the uncertainty. These factors combined make it challenging to accurately match RS and in-situ data pairs.

Overall, errors and uncertainties were introduced from the RS data, in-situ data and matching them as pairs. The dataset is complex as it contains information about time series, spatial variation within and between each field. Consequently, the overall quality of the dataset is average. The average quality, combined with small data volume, contribute to errors and uncertainties of both the interpretable model and RFR model.

If the objective is to improve the quality of the dataset, it would be more beneficial to utilize ground-based aerial platforms such as unmanned aerial vehicles (UAVs). UAVs typically provide higher spatial resolution (<5 m) images compared to satellites, and the temporal resolution can be easily adjusted. Customized sensors could also be deployed on the UAV platform, allowing for more optimal indices related to vegetation dry matter and water content. However, this approach can be expensive and time-consuming. There is always a trade-off between accuracy and cost. For agricultural purposes, prioritizing a fast and low-cost method is important, even if it means sacrificing a certain level of accuracy.

## 5.4 Experience and issues with Google Earth Engine and Python

The retrieval of all RS data was conducted using GEE and overall, the experience was satisfying. GEE helped ease concerns regarding storage space that is required to store multi-temporal images. The built-in functions and provided demos made it convenient to retrieve customized data.

During the data retrieval phase, ‘computation timed out’ and ‘memory capacity exceeded’ errors were frequently encountered. These issues were resolved by either waiting for a longer period or reducing the required time period of satellite images. Since this study only required local-scale data and the total number of images used for analysis was not large, no efforts were made to optimize the code to maximize

computing power and memory usage.

A bug was hit when trying to aggregate and export the data within each field. The use of the Scene Classification Layer (SCL) resulted in cloudy pixels/regions being masked out, leading to null values. When transforming a list to a feature collection ready for export, all properties disappeared. This occurred because the system treats a property with a null value and the absence of a property as indistinguishable. As a result, when creating properties (fields) based on the first element in the list, if the first element has properties with null values, no values for those properties will be included, even if other elements in the list have values for those properties. To keep the properties for other elements, we must manually set null values of the first element in the list to some other value.

It is worth noting that there is no distinction between null values and no data in GEE, which can be advantageous during the debugging process by saving considerable effort. Python Pandas library was utilized to organize and manipulate data. This library is recommended for data analysis, especially for time series data.

## 5.5 Limitations of this study and recommendations for future studies

The high frequency of cloudy weather in the trial fields limits the temporal resolution of the S2 optical data. The temporal gaps between two successive images used for extracting pixel values contribute to the uncertainties in the dataset. A potential future development of this study is to increase temporal resolution by incorporating data from active sensing sources, such as Sentinel-1. The utilization of multi-source RS data has the potential to significantly reduce uncertainties when matching RS and in-situ data pairs (Li et al., 2012; Xie et al., 2012).

Another limitation is the unaddressed significant noise in the time series during spline smoothing interpolation. Additional efforts could be made to reduce noise and improve the quality of the time series data. Furthermore, the spline smoothing approach used was point-based and did not consider the spatial pattern of the region. For constructing time series, professional tools/software such as TIMESAT could be used to better capture the spatial dynamics of the area.

Bretas et al. (2021) utilized a Random Forest (RF) model to predict DMCaP in pastures by integrating meteorological data and satellite imagery. Although weather data was

included in the study, it was homogenous for all fields. It would be beneficial for future studies to incorporate higher spatial-resolution meteorological data and meteorological data from diverse sources to investigate if it enhances the estimation of DMCaP in forage grasslands.

# 6. Conclusions

Monitoring vegetation dry matter and water content at leaf-level has been long conducted, but few studies have investigated the capabilities of satellite remote sensing in estimating dry matter content as percentage (DMCaP) for agricultural applications. In this study, we utilized in-situ measurements of DMCaP obtained from forage grasslands to evaluate the capabilities of Sentinel-2 data, and combined with environmental variables in estimating community-level DMCaP in forage grasslands in Southern Norway. Two types of models were developed, the interpretable model and the Random Forest Regression (RFR) model. The evaluation of the model performance showed that the $R^2$ values were relatively low to moderate for both types of models, while the RSMEs and MAEs were generally within an acceptable range. The RFR model outperformed the interpretable model, achieving a RMSE of 3.88%. Adding environmental variables to the RFR model improved the accuracy (RMSE = 2.90%). The interpretable model performed better when applied to individual fields, indicating its suitability for highly local-scaled data. There was no significant difference in performance when using different combinations of water index and vegetation index.

Given the complexity of the dataset and the considerable uncertainties involved, it is challenging to determine whether the lower performance of the interpretable model is due to incorrect assumptions or the inherent uncertainties within the dataset. The interpretable model, with its lower performance and dependency on multiple factors, is deemed less favorable. The relatively low to moderate $R^2$ values of both types of models are more likely attributed to the characteristics of the dataset.

Although we have concluded that the RFR model is a more favorable choice compared to the interpretable model, it is important to note that it was built using local-scale samples from non-natural forage grasslands. The DMCaP, which represents community-level traits that associates to species distribution and anthropogenic activities, may result in limited generalization capability of the model in larger scales and different land use types.

To reduce uncertainties in matching RS and in-situ data pairs, the use of multi-source RS data is recommended. Professional tools or software can be beneficial in reducing noise and considering spatial patterns when constructing time series. Future studies

could also consider to incorporate higher spatial-resolution meteorological data and meteorological data from diverse sources to investigate if it enhances the estimation of DMCaP in forage grasslands.

In conclusion, primarily Sentinel-2 data, and combined with environmental variables were utilized to develop models for estimating community-level DMCaP values in forage grasslands for agricultural purposes in southern Norway. This study addresses the gap in estimating vegetation dry matter content between the agricultural and academic communities, as well as the limitations observed in previous studies when attempting to match satellite and in-situ data pairs. It highlights the potential of satellite RS in estimating the DMCaP quality factor for agricultural applications, providing valuable insights for the forage-harvesting process and pasture management.

# References


Adab, H., Kanniah, K. D., & Beringer, J. (2016). Estimating and Up-Scaling Fuel Moisture and Leaf Dry Matter Content of a Temperate Humid Forest Using Multi Resolution Remote Sensing Data [Article]. *Remote Sensing*, *8*(11), Article 961. https://doi.org/10.3390/rs8110961

Ali, A. M., Darvishzadeh, R., Skidmore, A. K., van Duren, I., Heiden, U., & Heurich, M. (2016). Estimating leaf functional traits by inversion of PROSPECT: Assessing leaf dry matter content and specific leaf area in mixed mountainous forest [Article]. *International Journal of Applied Earth Observation and Geoinformation*, *45*, 66-76. https://doi.org/10.1016/j.jag.2015.11.004

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 [Review]. *Remote Sensing*, *7*(12), 16398-16421. https://doi.org/10.3390/rs71215841

Allen, V. G., Batello, C., Berretta, E. J., Hodgson, J., Kothmann, M., Li, X., McLvor, J., Milne, J., Morris, C., Peeters, A., Sanderson, M., & The Forage Grazing Terminology, C. (2011). An international terminology for grazing lands and grazing animals [Article]. *Grass and Forage Science*, *66*(1), 2-28. https://doi.org/10.1111/j.1365-2494.2010.00780.x

Amazirh, A., Merlin, O., Er-Raki, S., Gao, Q., Rivalland, V., Malbeteau, Y., Khabba, S., & Jose Escorihuela, M. (2018). Retrieving surface soil moisture at high spatio-temporal resolution from a synergy between Sentinel-1 radar and Landsat thermal data: A study case over bare soil [Review]. *Remote Sensing of Environment*, *211*, 321-337. https://doi.org/10.1016/j.rse.2018.04.013

Baret, F., & Fourty, T. (1997). Estimation of leaf water content and specific leaf weight from reflectance and transmittance measurements [Article]. *Agronomie*, *17*(9-10), 455-464. https://doi.org/10.1051/agro:19970903

Barrett, H., & Rose, D. C. (2022). Perceptions of the Fourth Agricultural Revolution: What's In, What's Out, and What Consequences are Anticipated? *Sociologia Ruralis*, *62*(2), 162-189. https://doi.org/10.1111/soru.12324

Behrenfeld, M. J., Hu, Y., Bisson, K. M., Lu, X., & Westberry, T. K. (2022). Retrieval of ocean optical and plankton properties with the satellite Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) sensor: Background, data processing, and validation status [Review]. *Remote Sensing of Environment*, *281*, Article 113235. https://doi.org/10.1016/j.rse.2022.113235

Bergstra, J., Bardenet, R., Bengio, Y., & Kégl, B. (2011). Algorithms for hyper-parameter optimization. *Advances in neural information processing systems*, *24*.

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of machine learning research*, *13*(2).

Boutton, T. W., & Tieszen, L. L. (1983). Estimation of plant biomass by spectral reflectance in an East African grassland. *Rangeland Ecology & Management/Journal of Range Management Archives*, *36*(2), 213-216.

Bretas, I. L., Valente, D. S. M., Silva, F. F., Chizzotti, M. L., Paulino, M. F., D'Aurea, A. P., Paciullo, D. S. C., Pedreira, B. C., & Chizzotti, F. H. M. (2021). Prediction of aboveground biomass and dry-matter content in brachiaria pastures by combining meteorological data and satellite imagery [Article]. *Grass and Forage Science*, *76*(3), 340-352. https://doi.org/10.1111/gfs.12517

Burgan, R. E. (1996). Use of remotely sensed data for fire danger estimation. *EARSel*

*Advances in remote sensing*, *4*, 1-8.
Campbell, J. B., & Wynne, R. H. (2011). *Introduction to remote sensing*. Guilford Press.
Capstaff, N. M., & Miller, A. J. (2018). Improving the Yield and Nutritional Quality of Forage Crops. *Frontiers in Plant Science*, *9*, Article 535. https://doi.org/10.3389/fpls.2018.00535
Ceccato, P., Flasse, S., & Gregoire, J. M. (2002). Designing a spectral index to estimate vegetation water content from remote sensing data - Part 2. Validation and applications [Article]. *Remote Sensing of Environment*, *82*(2-3), 198-207, Article Pii s0034-4257(02)00036-6. https://doi.org/10.1016/s0034-4257(02)00036-6
Cevoli, C., Di Cecilia, L., Ferrari, L., Fabbri, A., & Molari, G. (2022). Evaluation of cut alfalfa moisture content and operative conditions by hyperspectral imaging combined with chemometric tools: In-field application. *Biosystems Engineering*, *222*, 132-141. https://doi.org/https://doi.org/10.1016/j.biosystemseng.2022.08.004
Cheng, T., Rivard, B., Sanchez-Azofeifa, A. G., Feret, J.-B., Jacquemoud, S., & Ustin, S. L. (2014). Deriving leaf mass per area (LMA) from foliar reflectance across a variety of plant species using continuous wavelet analysis [Article]. *Isprs Journal of Photogrammetry and Remote Sensing*, *87*, 28-38. https://doi.org/10.1016/j.isprsjprs.2013.10.009
Chuvieco, E., Deshayes, M., Stach, N., Cocero, D., & Riaño, D. (1999). Short-term fire risk: foliage moisture content estimation from satellite data. *Remote Sensing of Large Wildfires: in the European Mediterranean Basin*, 17-38.
Clarmann, T. V., Chineke, T. C., Fischer, H., Funke, B., Garcia-Comas, M., Gil-Lopez, S., Glatthor, N., Grabowski, U., Hopfner, M., Kellmann, S., Kiefer, M., Linden, A., Lopez-Puertas, M., Lopez-Valverde, M. A., Tsidu, G. M., Milz, M., Steck, T., & Stiller, G. P. (2002, 2003 Sep 24-27). Remote sensing of the middle atmosphere with MIPAS.*Proceedings of SPIE* [Remote sensing of clouds and the atmosphere vii]. Conference on Remote Sensing of Clouds and the Atmosphere VII, Agia Pelagia, Greece.
Colombo, R., Merom, M., Marchesi, A., Busetto, L., Rossini, M., Giardino, C., & Panigada, C. (2008). Estimation of leaf and canopy water content in poplar plantations by means of hyperspectral indices and inverse modeling [Article]. *Remote Sensing of Environment*, *112*(4), 1820-1834. https://doi.org/10.1016/j.rse.2007.09.005
Conant, R. T. (2010). Challenges and opportunities for carbon sequestration in grassland systems: a technical report on grassland management and climate change mitigation.
Conejo, E., Frangi, J.-P., & De Rosny, G. (2015). Neural network implementation for a reversal procedure for water and dry matter estimation on plant leaves using selected LED wavelengths [Article]. *Applied Optics*, *54*(17), 5453-5460. https://doi.org/10.1364/ao.54.005453
Cretaux, J. F., Abarca-del-Rio, R., Berge-Nguyen, M., Arsen, A., Drolon, V., Clos, G., & Maisongrande, P. (2016). Lake Volume Monitoring from Space [Review]. *Surveys in Geophysics*, *37*(2), 269-305. https://doi.org/10.1007/s10712-016-9362-6
*Cubic smoothing spline - csaps*. (n.d.). The MathWorks Inc. Retrieved April 30, 2023 from https://ww2.mathworks.cn/help/curvefit/csaps.html
Danson, F. M., Steven, M. D., Malthus, T. J., & Clark, J. A. (1992). HIGH-SPECTRAL RESOLUTION DATA FOR DETERMINING LEAF WATER-CONTENT

[Article]. *International Journal of Remote Sensing*, *13*(3), 461-470. https://doi.org/10.1080/01431169208904049

Darvishzadeh, R., Skidmore, A., Abdullah, H., Cherenet, E., Ali, A., Wang, T., Nieuwenhuis, W., Heurich, M., Vrieling, A., O'Connor, B., & Paganini, M. (2019). Mapping leaf chlorophyll content from Sentinel-2 and RapidEye data in spruce stands using the invertible forest reflectance model [Article]. *International Journal of Applied Earth Observation and Geoinformation*, *79*, 58-70. https://doi.org/10.1016/j.jag.2019.03.003

Dawson, T. P., Curran, P. J., & Plummer, S. E. (1998). LIBERTY - Modeling the effects of leaf biochemical concentration on reflectance spectra [Article]. *Remote Sensing of Environment*, *65*(1), 50-60. https://doi.org/10.1016/s0034-4257(98)00007-8

De Boor, C. (1978). A practical guide to splines (Vol. 27, p. 325). *New York： springer-Verlag*.

Douglass, M. J. J. (2020). Hands-on Machine Learning with Scikit-Learn, Keras, and Tensorflow, 2nd edition [Book Review]. *Physical and Engineering Sciences in Medicine*, *43*(3), 1135-1136. https://doi.org/10.1007/s13246-020-00913-z

Elders, A., Carroll, M. L., Neigh, C. S. R., D'Agostino, A. L., Ksoll, C., Wooten, M. R., & Brown, M. E. (2022). Estimating crop type and yield of small holder fields in Burkina Faso using multi-day Sentinel-2 [Article]. *Remote Sensing Applications-Society and Environment*, *27*, Article 100820. https://doi.org/10.1016/j.rsase.2022.100820

Feret, J.-B., Berger, K., de Boissieu, F., & Malenovsky, Z. (2021). PROSPECT-PRO for estimating content of nitrogen-containing leaf proteins and other carbon-based constituents [Article]. *Remote Sensing of Environment*, *252*, Article 112173. https://doi.org/10.1016/j.rse.2020.112173

Feret, J.-B., Francois, C., Gitelson, A., Asner, G. P., Barry, K. M., Panigada, C., Richardson, A. D., & Jacquemoud, S. (2011). Optimizing spectral indices and chemometric analysis of leaf chemical properties using radiative transfer modeling [Article]. *Remote Sensing of Environment*, *115*(10), 2742-2750. https://doi.org/10.1016/j.rse.2011.06.016

Feret, J. B., Gitelson, A. A., Noble, S. D., & Jacquemoud, S. (2017). PROSPECT-D: Towards modeling leaf optical properties through a complete lifecycle [Article]. *Remote Sensing of Environment*, *193*, 204-215. https://doi.org/10.1016/j.rse.2017.03.004

Feret, J. B., le Maire, G., Jay, S., Berveiller, D., Bendoula, R., Hmimina, G., Cheraiet, A., Oliveira, J. C., Ponzoni, F. J., Solanki, T., de Boissieu, F., Chave, J., Nouvellon, Y., Porcar-Castell, A., Proisy, C., Soudani, K., Gastellu-Etchegorry, J. P., & Lefevre-Fonollosa, M. J. (2019). Estimating leaf mass per area and equivalent water thickness based on leaf optical properties: Potential and limitations of physical modeling and machine learning [Article]. *Remote Sensing of Environment*, *231*, Article 110959. https://doi.org/10.1016/j.rse.2018.11.002

Figueiredo Filho, D. B., Júnior, J. A. S., & Rocha, E. C. (2011). What is R2 all about? *Leviathan (São Paulo)*(3), 60-68. https://doi.org/https://doi.org/10.11606/issn.2237-4485.lev.2011.132282

Forkuor, G., Zoungrana, J.-B. B., Dimobe, K., Ouattara, B., Vadrevu, K. P., & Tondoh, J. E. (2020). Above-ground biomass mapping in West African dryland forest using Sentinel-1 and 2 datasets - A case study [Article]. *Remote Sensing of Environment*, *236*, Article 111496. https://doi.org/10.1016/j.rse.2019.111496

Frampton, W. J., Dash, J., Watmough, G., & Milton, E. J. (2013). Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation [Article]. *Isprs Journal of Photogrammetry and Remote Sensing*, *82*, 83-92. https://doi.org/10.1016/j.isprsjprs.2013.04.007

Galvao, L. S., Formaggio, A. R., & Tisot, D. A. (2005). Discrimination of sugarcane varieties in southeastern brazil with EO-1 hyperion data [Article]. *Remote Sensing of Environment*, *94*(4), 523-534. https://doi.org/10.1016/j.rse.2004.11.012

Gao, B. C. (1995). A normalized difference water index for remote sensing of vegetation liquid water from space.*Proceedings of the Society of Photo-Optical Instrumentation Engineers (Spie)* [Imaging spectrometry]. 1995 Imaging Spectrometry Conference, Orlando, Fl.

Gara, T. W., Darvishzadeh, R., Skidmore, A. K., Wang, T., & Heurich, M. (2019). Accurate modelling of canopy traits from seasonal Sentinel-2 imagery based on the vertical distribution of leaf traits [Article]. *Isprs Journal of Photogrammetry and Remote Sensing*, *157*, 108-123. https://doi.org/10.1016/j.isprsjprs.2019.09.005

Gerstl, S. A. W. (1990). PHYSICS CONCEPTS OF OPTICAL AND RADAR REFLECTANCE SIGNATURES - A SUMMARY REVIEW [Article; Proceedings Paper]. *International Journal of Remote Sensing*, *11*(7), 1109-1117. https://doi.org/10.1080/01431169008955083

*Getting Silage Dry Matter Right*. (n.d.). Retrieved Jan 1, 2023 from https://www.silageconsultant.co.uk/blog/getting-silage-dry-matter-percentage-right

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*, *58*(3), 289-298. https://doi.org/https://doi.org/10.1016/S0034-4257(96)00072-7

*Grassland guide: Grass nutritional value*. (n.d.). Retrieved March 4, 2023 from https://germinal.co.nz/knowledge-hub/grass-nutritional-value/

Hansen, P. M., & Schjoerring, J. K. (2003). Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression [Article]. *Remote Sensing of Environment*, *86*(4), 542-553. https://doi.org/10.1016/s0034-4257(03)00131-7

Henchion, M., Moloney, A. P., Hyland, J., Zimmermann, J., & McCarthy, S. (2021). Review: Trends for meat, milk and egg consumption for the next decades and the role played by livestock systems in the global production of proteins [Review]. *Animal*, *15*, Article 100287. https://doi.org/10.1016/j.animal.2021.100287

Hill, M. J. (2013). Vegetation index suites as indicators of vegetation state in grassland and savanna: An analysis with simulated SENTINEL 2 data for a North American transect [Article]. *Remote Sensing of Environment*, *137*, 94-111. https://doi.org/10.1016/j.rse.2013.06.004

Hmimina, G., Dufrene, E., Pontailler, J. Y., Delpierre, N., Aubinet, M., Caquet, B., de Grandcourt, A., Burban, B., Flechard, C., Granier, A., Gross, P., Heinesch, B., Longdoz, B., Moureaux, C., Ourcival, J. M., Rambal, S., Saint Andre, L., & Soudani, K. (2013). Evaluation of the potential of MODIS satellite data to predict vegetation phenology in different biomes: An investigation using ground-based NDVI measurements [Article]. *Remote Sensing of Environment*,

*132*, 145-158. https://doi.org/10.1016/j.rse.2013.01.010
Hopkins, A. (2000). *Grass: Its Production and Utilization*. Wiley. https://books.google.se/books?id=MHKuQgAACAAJ
Hubert-Moy, L., Fabre, E., & Rapinel, S. (2020). Contribution of SPOT-7 multi-temporal imagery for mapping wetland vegetation [Article]. *European Journal of Remote Sensing*, *53*(1), 201-210. https://doi.org/10.1080/22797254.2020.1795727
Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*(1), 195-213. https://doi.org/https://doi.org/10.1016/S0034-4257(02)00096-2
Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, *25*(3), 295-309. https://doi.org/https://doi.org/10.1016/0034-4257(88)90106-X
Hunt, E. R., & Rock, B. N. (1989). Detection of changes in leaf water-content using near-infrared and middle infrared reflectances. [Article]. *Remote Sensing of Environment*, *30*(1), 43-54. https://doi.org/10.1016/0034-4257(89)90046-1
Jacquemoud, S., & Baret, F. (1990). PROSPECT - A MODEL OF LEAF OPTICAL-PROPERTIES SPECTRA [Article]. *Remote Sensing of Environment*, *34*(2), 75-91. https://doi.org/10.1016/0034-4257(90)90100-z
Jacquemoud, S., Ustin, S. L., Verdebout, J., Schmuck, G., Andreoli, G., & Hosgood, B. (1996). Estimating leaf biochemistry using the PROSPECT leaf optical properties model [Article]. *Remote Sensing of Environment*, *56*(3), 194-202. https://doi.org/10.1016/0034-4257(95)00238-3
Jansen, V. S., Kolden, C. A., & Schmalz, H. J. (2018). The Development of Near Real-Time Biomass and Cover Estimates for Adaptive Rangeland Management Using Landsat 7 and Landsat 8 Surface Reflectance Products [Article]. *Remote Sensing*, *10*(7), Article 1057. https://doi.org/10.3390/rs10071057
Jin, Y. Q., & Liu, C. (1997). Biomass retrieval from high-dimensional active/passive remote sensing data by using artificial neural networks [Article]. *International Journal of Remote Sensing*, *18*(4), 971-979. https://doi.org/10.1080/014311697218863
Jing, X., Huang Wen, J., Wang Ji, H., Wang Jin, D., & Wang, K.-r. (2009). Hyperspectral Inversion Models on Verticillium Wilt Severity of Cotton Leaf [Article]. *Spectroscopy and Spectral Analysis*, *29*(12), 3348-3352. https://doi.org/10.3964/j.issn.1000-0593(2009)12-3348-05
Jordan, C. F. (1969). Derivation of Leaf-Area Index from Quality of Light on the Forest Floor. *Ecology*, *50*(4), 663-666. https://doi.org/https://doi.org/10.2307/1936256
Ju, J., & Masek, J. G. (2016). The vegetation greenness trend in Canada and US Alaska from 1984-2012 Landsat data [Article]. *Remote Sensing of Environment*, *176*, 1-16. https://doi.org/10.1016/j.rse.2016.01.001
Kaufman, Y. J., & Tanre, D. (1992). ATMOSPHERICALLY RESISTANT VEGETATION INDEX (ARVI) FOR EOS-MODIS [Article]. *Ieee Transactions on Geoscience and Remote Sensing*, *30*(2), 261-270. https://doi.org/10.1109/36.134076
Khanna, S., Palacios-Orueta, A., Whiting, M. L., Ustin, S. L., Riano, D., & Litago, J. (2007). Development of angle indexes for soil moisture estimation, dry matter detection and land-cover discrimination [Article]. *Remote Sensing of Environment*, *109*(2), 154-165. https://doi.org/10.1016/j.rse.2006.12.018
Kim, Y., Jackson, T., Bindlish, R., Lee, H., & Hong, S. (2012). Radar Vegetation Index

for Estimating the Vegetation Water Content of Rice and Soybean [Article]. *Ieee Geoscience and Remote Sensing Letters*, *9*(4), 564-568. https://doi.org/10.1109/lgrs.2011.2174772

Klemas, V., & Smart, R. (1983). The influence of soil salinity, growth form, and leaf moisture on-the spectral radiance of. *Photogramm. Eng. Remote Sens*, *49*, 77-83.

Koirala, B., Zahiri, Z., & Scheunders, P. (2020). A Machine Learning Framework for Estimating Leaf Biochemical Parameters From Its Spectral Reflectance and Transmission Measurements [Article]. *Ieee Transactions on Geoscience and Remote Sensing*, *58*(10), 7393-7405. https://doi.org/10.1109/tgrs.2020.2982263

Kong, B., Yu, H., Du, R., & Wang, Q. (2019). Quantitative Estimation of Biomass of Alpine Grasslands Using Hyperspectral Remote Sensing [Article]. *Rangeland Ecology & Management*, *72*(2), 336-346. https://doi.org/10.1016/j.rama.2018.10.005

Kroupi, E., Kesa, M., Navarro-Sanchez, V. D., Saeed, S., Pelloquin, C., Alhaddad, B., Moreno, L., Soria-Frisch, A., & Ruffini, G. (2019). Deep convolutional neural networks for land-cover classification with Sentinel-2 images. *Journal of Applied Remote Sensing*, *13*(2), Article 024525. https://doi.org/10.1117/1.Jrs.13.024525

Kumar, D., Rao, S., & Sharma, J. (2013). Radar Vegetation Index as an alternative to NDVI for monitoring of soyabean and cotton. Proceedings of the XXXIII INCA International Congress (Indian Cartographer), Jodhpur, India,

le Maire, G., Francois, C., Soudani, K., Berveiller, D., Pontailler, J.-Y., Breda, N., Genet, H., Davi, H., & Dufrene, E. (2008). Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass [Article]. *Remote Sensing of Environment*, *112*(10), 3846-3864. https://doi.org/10.1016/j.rse.2008.06.005

Li, A., Jiang, J., Bian, J., & Deng, W. (2012). Combining the matter element model with the associated function of probability transformation for multi-source remote sensing data classification in mountainous regions [Article]. *Isprs Journal of Photogrammetry and Remote Sensing*, *67*, 80-92. https://doi.org/10.1016/j.isprsjprs.2011.10.008

Li, C., Wulf, H., Schmid, B., He, J.-S., & Schaepman, M. E. (2018). Estimating Plant Traits of Alpine Grasslands on the Qinghai-Tibetan Plateau Using Remote Sensing [Article]. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(7), 2263-2275. https://doi.org/10.1109/jstars.2018.2824901

Li, P., & Wang, Q. (2011). Retrieval of Leaf Biochemical Parameters Using PROSPECT Inversion: A New Approach for Alleviating Ill-Posed Problems [Article]. *Ieee Transactions on Geoscience and Remote Sensing*, *49*(7), 2499-2506. https://doi.org/10.1109/tgrs.2011.2109390

Li, Z., & Guo, X. (2018). Non-photosynthetic vegetation biomass estimation in semiarid Canadian mixed grasslands using ground hyperspectral data, Landsat 8 OLI, and Sentinel-2 images. *International Journal of Remote Sensing*, *39*(20), 6893-6913.

Lillisand, T., Kiefer, R., & Chipman, J. (2008). Remote Sensing and image interpretation, Hoboken. In: NJ: Wiley & Sons.

Liu, C.-a., Chen, Z.-x., Shao, Y., Chen, J.-s., Hasi, T., & Pan, H.-z. (2019). Research advances of SAR remote sensing for agriculture applications: A review

[Review]. *Journal of Integrative Agriculture*, *18*(3), 506-525. https://doi.org/10.1016/s2095-3119(18)62016-7

Liu, L., Song, B., Zhang, S., & Liu, X. (2017). A Novel Principal Component Analysis Method for the Reconstruction of Leaf Reflectance Spectra and Retrieval of Leaf Biochemical Contents [Article]. *Remote Sensing*, *9*(11), Article 1113. https://doi.org/10.3390/rs9111113

Lo Re, G., Fuller, I. C., Sofia, G., & Tarolli, P. (2018). High-resolution mapping of Manawatu palaeochannels. *New Zealand Geographer*, *74*(2), 77-91. https://doi.org/10.1111/nzg.12186

Lv, X., Wang, J., Sun, Q., Yao, G., & Gao, F. (2014). Research on the hyperspectral remote sensing estimation models for the fresh yield of alfalfa grassland [article]. *Acta Prataculturae Sinica*, *23*(3), 84-91. <Go to ISI>://CABI:20143192874

Maier, S. W., Ludeker, W., & Gunther, K. P. (1999). SLOP: A revised version of the stochastic model for leaf optical properties [Article; Proceedings Paper]. *Remote Sensing of Environment*, *68*(3), 273-280. https://doi.org/10.1016/s0034-4257(98)00118-7

Mateo-Sanchis, A., Piles, M., Munoz-Mari, J., Adsuara, J. E., Perez-Suay, A., & Camps-Valls, G. (2019). Synergistic integration of optical and microwave satellite data for crop yield estimation [Article]. *Remote Sensing of Environment*, *234*, Article 111460. https://doi.org/10.1016/j.rse.2019.111460

Mbow, C. (1999). Proposition of a method for early fires planning using ground and satellite (NDVI/NOAA-AVHRR) data from Niokolo Koba National Park (Southeast Senegal). Proceedings for the Second International Symposium on Operationalization of Remote Sensing. Poster presentation. ITC, Enschede, The Netherlands,

McFeeters, S. K. (1996). The use of the normalized difference water index (NDWI) in the delineation of open water features [Article]. *International Journal of Remote Sensing*, *17*(7), 1425-1432. https://doi.org/10.1080/01431169608948714

Nellis, M. D., Price, K. P., & Rundquist, D. (2009). Remote sensing of cropland agriculture. *The SAGE handbook of remote sensing*, *1*, 368-380.

Newton, J. E., Nettle, R., & Pryce, J. E. (2020). Farming smarter with big data: Insights from the case of Australia's national dairy herd milk recording scheme [Article]. *Agricultural Systems*, *181*, Article 102811. https://doi.org/10.1016/j.agsy.2020.102811

Nowatzki, J., Andres, R., & Kyllo, K. (2004). Agricultural remote sensing basics.

Oliveira, R. A., Nasi, R., Niemelainen, 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 [Article]. *Remote Sensing of Environment*, *246*, Article 111830. https://doi.org/10.1016/j.rse.2020.111830

Orynbaikyzy, A., Gessner, U., & Conrad, C. (2019). Crop type classification using a combination of optical and radar remote sensing data: a review [Review]. *International Journal of Remote Sensing*, *40*(17), 6553-6595. https://doi.org/10.1080/01431161.2019.1569791

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, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of machine learning research*, *12*, 2825-2830.

Płoński, P. (2020). *Random Forest Feature Importance Computed in 3 Ways with Python*. Retrieved May 11th, 2023 from https://mljar.com/blog/feature-importance-in-random-forest/

Qiu, F., Chen, J. M., Ju, W., Wang, J., Zhang, Q., & Fang, M. (2018). Improving the PROSPECT Model to Consider Anisotropic Scattering of Leaf Internal Materials and Its Use for Retrieving Leaf Biomass in Fresh Leaves [Article]. *Ieee Transactions on Geoscience and Remote Sensing*, *56*(6), 3119-3136. https://doi.org/10.1109/tgrs.2018.2791930

Romero, A., Aguado, I., & Yebra, M. (2012). Estimation of dry matter content in leaves using normalized indexes and PROSPECT model inversion [Article]. *International Journal of Remote Sensing*, *33*(2), 396-414. https://doi.org/10.1080/01431161.2010.532819

Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *NASA Spec. Publ*, *351*(1), 309.

Seegers, B. N., Stumpf, R. P., Schaeffer, B. A., Loftin, K. A., & Werdell, P. J. (2018). Performance metrics for the assessment of satellite data products: an ocean color case study. *Optics Express*, *26*(6), 7404-7422. https://doi.org/10.1364/oe.26.007404

Shah, S. H., Angel, Y., Houborg, R., Ali, S., & McCabe, M. F. (2019). A Random Forest Machine Learning Approach for the Retrieval of Leaf Chlorophyll Content in Wheat [Article]. *Remote Sensing*, *11*(8), Article 920. https://doi.org/10.3390/rs11080920

Sharma, R. C., Hara, K., & Tateishi, R. (2017). High-Resolution Vegetation Mapping in Japan by Combining Sentinel-2 and Landsat 8 Based Multi-Temporal Datasets through Machine Learning and Cross-Validation Approach. *Land*, *6*(3), Article 50. https://doi.org/10.3390/land6030050

Shen, M., Zhang, G., Cong, N., Wang, S., Kong, W., & Piao, S. (2014). Increasing altitudinal gradient of spring vegetation phenology during the last decade on the Qinghai-Tibetan Plateau [Article]. *Agricultural and Forest Meteorology*, *189*, 71-80. https://doi.org/10.1016/j.agrformet.2014.01.003

Shepherd, M., Turner, J. A., Small, B., & Wheeler, D. (2018). Priorities for science to overcome hurdles thwarting the full promise of the 'digital agriculture' revolution. *Journal of the Science of Food and Agriculture*, *100*(14), 5083-5092. https://doi.org/https://doi.org/10.1002/jsfa.9346

Sow, M., Mbow, C., Hély, C., Fensholt, R., & Sambou, B. (2013). Estimation of herbaceous fuel moisture content using vegetation indices and land surface temperature from MODIS data. *Remote Sensing*, *5*(6), 2617-2638.

Spiess, A.-N., & Neumeyer, N. (2010). An evaluation of R-2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: a Monte Carlo approach [Article]. *BMC Pharmacology*, *10*, 6-Article No.: 6. https://doi.org/10.1186/1471-2210-10-6

Still, D. A., & Shih, S. F. (1985). USING LANDSAT DATA TO CLASSIFY LAND-USE FOR ASSESSING THE BASINWIDE RUNOFF INDEX [Article]. *Water Resources Bulletin*, *21*(6), 931-940. <Go to ISI>://WOS:A1985AYF9100004

SUHET. (2015). *Sentinel-2 User Handbook*. https://sentinels.copernicus.eu/documents/247904/685211/Sentinel-2_User_Handbook

Wang, B. J., & Ju, W. (2017). Limitations and Improvements of the Leaf Optical Properties Model Leaf Incorporating Biochemistry Exhibiting Reflectance and Transmittance Yields (LIBERTY) [Article]. *Remote Sensing*, *9*(5), Article 431.

https://doi.org/10.3390/rs9050431
Wang, J., Xiao, X., Bajgain, R., Starks, P., Steiner, J., Doughty, R. B., & Chang, Q. (2019). Estimating leaf area index and aboveground biomass of grazing pastures using Sentinel-1, Sentinel-2 and Landsat images [Article]. *Isprs Journal of Photogrammetry and Remote Sensing*, *154*, 189-201. https://doi.org/10.1016/j.isprsjprs.2019.06.007
Wang, L., Qu, J. J., Hao, X., & Hunt, E. R., Jr. (2011). Estimating dry matter content from spectral reflectance for green leaves of different species [Article]. *International Journal of Remote Sensing*, *32*(22), 7097-7109. https://doi.org/10.1080/01431161.2010.494641
Weiss, M., Jacob, F., & Duveiller, G. (2020). Remote sensing for agricultural applications: A meta-review [Review]. *Remote Sensing of Environment*, *236*, Article 111402. https://doi.org/10.1016/j.rse.2019.111402
Wilkinson, J. M., & Davies, D. R. (2013). The aerobic stability of silage: key findings and recent developments [Review]. *Grass and Forage Science*, *68*(1), 1-19. https://doi.org/10.1111/j.1365-2494.2012.00891.x
Xie, Y., Yi, S., & Tang, Z. (2012). Uncertainty Multi-source Information Fusion for Intelligent Flood Risk Analysis Based on Random Set Theory [Article]. *International Journal of Computational Intelligence Systems*, *5*(5), 975-984. https://doi.org/10.1080/18756891.2012.733237
Xu, D., Chen, B., Yan, R., Yan, Y., Sun, X., Xu, L., & Xin, X. (2019). Quantitative monitoring of grazing intensity in the temperate meadow steppe based on remote sensing data [Article]. *International Journal of Remote Sensing*, *40*(5-6), 2227-2242. https://doi.org/10.1080/01431161.2018.1500733
Yang, B., Lin, H., & He, Y. (2020). Data-Driven Methods for the Estimation of Leaf Water and Dry Matter Content: Performances, Potential and Limitations. *Sensors*, *20*(18), 5394.
Yao, X., Huang, Y., Shang, G., Zhou, C., Cheng, T., Tian, Y., Cao, W., & Zhu, Y. (2015). Evaluation of Six Algorithms to Monitor Wheat Leaf Nitrogen Concentration [Article]. *Remote Sensing*, *7*(11), 14939-14966. https://doi.org/10.3390/rs71114939
Yilmaz, M. T., Hunt, E. R., Jr., & Jackson, T. J. (2008). Remote sensing of vegetation water content from equivalent water thickness using satellite imagery [Article]. *Remote Sensing of Environment*, *112*(5), 2514-2522. https://doi.org/10.1016/j.rse.2007.11.014
Yuan, L., Pu, R., Zhang, J., Wang, J., & Yang, H. (2016). Using high spatial resolution satellite imagery for mapping powdery mildew at a regional scale [Article]. *Precision Agriculture*, *17*(3), 332-348. https://doi.org/10.1007/s11119-015-9421-x
Zarco-Tejada, P. J., Rueda, C. A., & Ustin, S. L. (2003). Water content estimation in vegetation with MODIS reflectance data and model inversion methods [Article]. *Remote Sensing of Environment*, *85*(1), 109-124. https://doi.org/10.1016/s0034-4257(02)00197-9
Zhang, Y., & Shao, Z. (2021). Assessing of Urban Vegetation Biomass in Combination with LiDAR and High-resolution Remote Sensing Images [Article]. *International Journal of Remote Sensing*, *42*(3), 964-985. https://doi.org/10.1080/01431161.2020.1820618

# Appendices

The appendices include two figures (Appendix A) and three tables (Appendix B). The first figure (Fig A.1) shows the frequency histogram of temporal gaps between two successive Sentinel-2 images used for extracting remote sensing measurements. The second figure (Fig A.2) presents the sampling frequency of in-situ measurements of dry matter content as percentage. The first table (Table B.1) shows the dates of collecting in-situ measurements for four fields. The second table (Table B.2) lists the tuned hyperparameters using random search of the Random Forest Regression model. The third table (Table B.3) summarizes the performance metrics of the interpretable model using data of four fields respectively.

## Appendix A. Additional figures

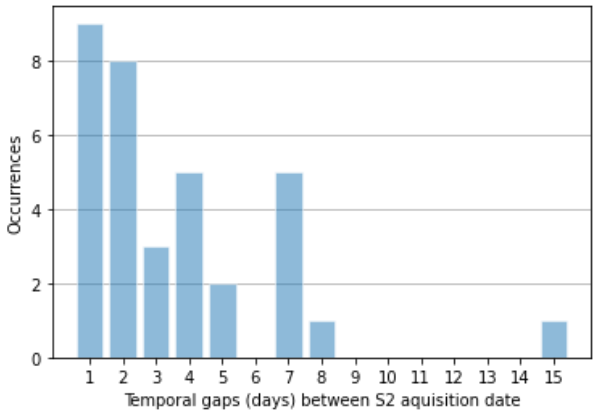


*Fig A.1. Frequency histogram of temporal gaps between two successive Sentinel-2 images used for extracting remote sensing measurements.*

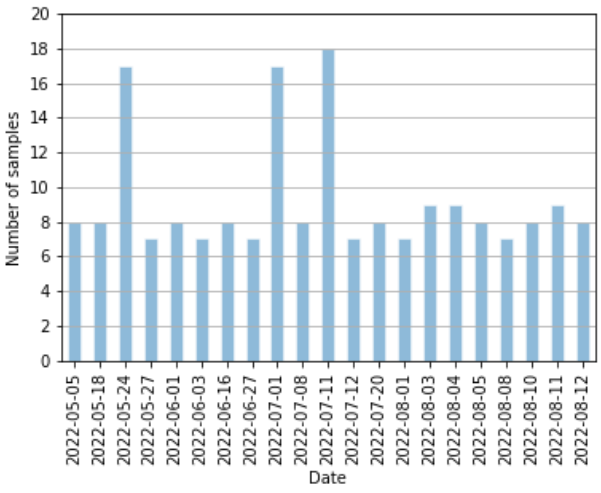


*Fig A.2. Sampling frequency of in-situ measurements of dry matter content as percentage.*

# Appendix B. Additional tables

*Table B.1. Dates collecting in-situ measurements for all fields, all dates referred to in the year of 2022.*

| Field name | Dates |
|---|---|
| **Field 1** | May 12th, May 27th, June 3rd, June 27th, July 12th, August 1st, August 8th |
| **Field 2** | May 5th, May 18th, June 1st, June 16th, July 8th, July 20th, August 5th, August 12th |
| **Field 3** | May 24th, July 1st, July 11th, August 3rd, August 10th |
| **Field 4** | May 24th, July 1st, July 11th, August 4rd, August 11th |

*Table B.2. Tuned hyperparameters using random search of the Random Forest Regression (RFR) model using only Sentinel-2 data and using Sentinel-2 data and environmental variables as predictors. Default settings provided by the Python Scikit-learn package were applied for hyperparameters not listed in the table.*

| Hyperparameter | RFR (Sentinel-2) | RFR (+ Environmental variables) |
|---|---|---|
| **Number of trees** | 300 | 300 |
| **Minimum samples split** | 2 | 2 |
| **Minimum samples leaf** | 1 | 2 |
| **Maximum features** | auto | sqrt |
| **Maximum depth** | 10 | 7 |
| **Bootstrap** | True | False |

*Table B.3. Performance metrics of the interpretable model using data of Field 1-4 respectively (running four times). F1-4 represents Field 1-4.*

| | **RMSE (F1-4)** | **MAE (F1-4)** | **Bias (F1-4)** | **R2 (F1-4)** |
|---|---|---|---|---|
| **RVI** | 3.01, 3.35, 5.25, 4.08 | 2.47, 2.71, 4.39, 3.38 | -0.0020, -0.0000, 0.0010, -0.0103 | 0.16, 0.03, 0.05,0.22 |
| **NDVI** | 3.01, 3.35, 4.96, 4.16 | 2.47, 2.71, 4.07, 3.51 | -0.0020, 0.0001, -0.0021, -0.0085 | 0.16, 0.03, 0.15, 0.19 |
| **GNDVI** | 3.01, 3.35, 5.11, 4.09 | 2.47, 2.71, 4.20, 3.36 | -0.0021, -0.0001, 0.0034 -0.0076 | 0.16, 0.03, 0.10, 0.22 |
| **EVI** | 3.01, 3.33, 5.29, 4.21 | 2.47, 2.74, 4.54, 3.56 | -0.0021, 0.0006, -0.0010, -0.0089 | 0.16, 0.04, 0.04, 0.17 |
| **CVI** | 3.01, 3.35, 5.29, 4.22 | 2.47, 2.71, 4.54, 3.57 | -0.0021, -0.0001, -0.0010, -0.0003 | 0.16, 0.03, 0.04, 0.17 |
| **NDRE** | 2.96, 3.35, 5.08, 3.82 | 2.41, 2.71, 4.16, 3.08 | -0.0037, -0.001, 0.0034, -0.0067 | 0.19, 0.03, 0.11, 0.32 |