Using Sentinel-2 Satellite Images to Estimate Traits of Forage Grasslands

Niklas Zeiner

Degree project/Independent project • 30 hp
Swedish University of Agricultural Sciences, SLU
Faculty of Natural Resources and Agricultural Sciences/Department of Agricultural Research for Northern Sweden
Agriculture Programme – Soil and Plant Sciences
Umeå 2020

Using Sentinel-2 Satellite Images to Estimate Traits of Forage Grasslands

Niklas Zeiner

Supervisor: **Julien Morel, Swedish University of Agricultural Sciences, Department of Agricultural Research in Northern Sweden**

Examiner: Bo Stenberg, Swedish University of Agricultural Sciences, Department of Soil and Environment

Credits: 30 hp

Level: Second Cycle, A2E

Course title: **Independent Project in Biology**

Course code: **EX0898**

Programme/education: **Agriculture Programme – Soil and Plant Sciences**

Course coordinating dept: **Department of Agricultural Research in Northern Sweden**

Place of publication: **Umeå**

Year of publication: **2020**

Cover picture: Niklas Zeiner

Keywords: **Remote sensing, Forage grasslands, Sentinel-2, Dry matter yield, Canopy average height, Total leaf chlorophyll content, Partial least squares, Support vector machines**

Swedish University of Agricultural Sciences

Faculty of Natural Resources and Agricultural Sciences

Department of Agricultural Research for Northern Sweden

Publishing and archiving

Approved students' theses at SLU are published electronically. As a student, you have the copyright to your own work and need to approve the electronic publishing. If you check the box for **YES**, the full text (pdf file) and metadata will be visible and searchable online. If you check the box for **NO**, only the metadata and the abstract will be visible and searchable online. Nevertheless, when the document is uploaded it will still be archived as a digital file.

If you are more than one author you all need to agree on a decision. Read about SLU's publishing agreement here: https://www.slu.se/en/subweb/library/publish-and-analyse/register-and-publish/agreement-for-publishing/.

☒ YES, I/we hereby give permission to publish the present thesis in accordance with the SLU agreement regarding the transfer of the right to publish a work.

☐ NO, I/we do not give permission to publish the present work. The work will still be archived and its metadata and abstract will be visible and searchable.

Abstract

In this project, regression models based on data from field measurements and spectral information extracted from satellite imagery were used to estimate traits of forage grasslands; dry matter yield, canopy average height and total leaf chlorophyll. Four fields at SLUs Röbäcksdalen field station were sampled on 22 occasions and a total of 198 samples, including measurement of the highest plant, canopy height, leaf chlorophyll content, canopy spectral reflectance and biomass were collected. Two regression methods, partial least squares (PLS) and support vector machines (SVM), were used to build regression models using different subsets of the available spectral information. Model calibration was performed with 2/3 of the dataset and model validation was performed with the remaining 1/3 of the dataset. It was shown that the models built with SVM outperformed the models built with PLS, during both calibration and validation as well as for all different traits and subsets of spectral information. Field measurement and regression model results were discussed and limitations, their significance and possible improvements were considered. It was concluded that using spectral information from satellite images is a promising approach for estimation of traits in the field and could be used to build tools as a tool to support farmers' decision making.

Keywords: Remote sensing, Forage grasslands, Sentinel-2, Dry matter yield, Canopy average height, Total leaf chlorophyll content, Partial least squares, Support vector machines

Populärvetenskaplig sammanfattning

Vall är en viktig del av svenskt lantbruk och produceras på ungefär 45% av den tillgängliga åkermarken. Vall används främst för produktion av djurfoder men även som en del av växtföljder för att minska trycket från skadegörare eller ogräs och för att främja den biologiska mångfalden. På grund av förhållandevis låga vinstmarginaler vid försäljning av djurfoder investerar lantbruksföretag inte mycket i optimering av vallproduktion. En mindre kostsam investering kan vara precisionsodling, som är ett sätt att optimera växtodling genom att övervaka jordbruksgrödor automatiskt och genomföra insatser såsom applicering av gödsel och växtskyddsmedel eller själva skörden vid rätt tidpunkt och på rätt plats. Inom lantbruk tillämpas övervakningen, också kallad fjärranalys, ofta med sensorer som mäter hur synligt och osynligt ljus av olika våglängder reflekteras av en yta. I det här fallet har denna information använts i form av satellitbilder från Sentinel-2 satelliterna som är tillgängliga varannan dag. Satellitbilderna kan sedan användas för att bygga modeller genom att koppla ihop resultat av mätningar i fält med informationen från satellitbilden.

Syftet med detta projekt var att utföra mätningar i fält och bygga modeller baserat på dessa mätningar samt informationen från satellitbilderna. Hypotesen var att bilderna från Sentinel-2 satelliterna gör det möjligt att med hjälp av modellerna uppskatta mätvärden i fält.

Inom projektet genomfördes mätningar av biomassa, planthöjd, klorofyllinnehåll och reflektans på 4 olika fält med vall vid SLUs Röbäcksdalen forskningsstation i Umeå. Två regressionsmetoder, partial least squares (PLS) och support vector machines (SVM), användes för att bygga modellerna. Kalibrering och anpassning av modellerna utfördes med 2/3 av satellitbilderna och fältmätningarna, kontroll och validering av modellerna utfördes med resterande 1/3 av satellitbilderna och fältmätningarna.

Resultaten visade att SVM modellerna fungerade bättre än PLS modellerna vid uppskattning av mätvärden i fält. En utvärdering av modellerna på andra fält på annan ort har inte genomförts och det är okänt hur modellerna fungerar under olika förhållanden. En intressant aspekt som upptäcktes var att modeller byggda med satellitbilderna som motsvarar synligt ljus gav förhållandevis bra resultat. Det innebär att kamerabilder, till exempel från en drönare, skulle kunna användas för att bygga modeller vilket öppnar för mätningar vid behov. Metoden är lovande och skulle som ett verktyg kunna användas av lantbrukare för övervakning av jordbruksgrödor och hjälpa dem att fatta beslut kring insatser.

Table of contents

List of tables

List of figures

Abbreviations

CSR	Canopy Spectral Reflectance
DMY	Dry Matter Yield
ESA	European Space Agency
NIR	Near-infrared
NDRE	Normalized Difference Red Edge
NDVI	Normalized Difference Vegetation Index
PLS	Partial Least Squares
REDGE	Red Edge
RMSE	Root Mean Square Error
SLU	Swedish University of Agricultural Sciences
SVM	Support Vector Machine

1. Introduction

1.1. Forage Grasslands

Forage is an important part of agriculture in Sweden and is estimated to represent around 45% of the total arable land use in 2020 (Olsson 2020). Apart from fodder crops like maize and certain cereals, forage grasslands are usually composed of a mix of grasses and legumes. The crops can be fed to livestock directly through grazing or be dried or pre-processed before feeding (Capstaff & Miller 2018). Efforts to optimize agricultural systems have often been concentrated on human food crops due to higher profit margins. Nevertheless, forage grasslands are gaining importance, especially in crop rotations. In organic farming systems they can be useful in avoiding or limiting pest prevalence (Bengtsson *et al.* 2019) or providing pre-crop effects (Eriksen *et al.* 2006). Increased biodiversity and land conservation is often a welcome side-effect of having forage grasslands in a crop rotation or nearby (Zhou *et al.* 2019b) but can also be the main purpose.

1.2. Precision Agriculture

One promising and possibly inexpensive method to optimize agricultural systems is precision agriculture which integrates monitoring of livestock, crops or both and the application of location- and time-based inputs (Pierce & Nowak 1999). In crop production, pesticide and fertilizer use can for example be limited to where and when it is needed which saves time and money and reduces the environmental impact (Stafford 2000). Another possible application of precision agriculture is harvest management depending on the growth stage and quality of the monitored crop including optimized harvesting strategies from a logistics point of view (Yagoubi *et al.* 2018).

1.3. Remote Sensing

Remote sensing is a useful tool for monitoring vegetation in general and crops in particular. In remote sensing, information about objects is collected from a distance. For agricultural purposes, this is often accomplished with sensors that measure reflectance of the crops in different parts of the light spectrum. Visible light range, for instance, spans from 400 to 700 nm, and carries information on the pigment content of the canopy. Near-infrared light, ranging from 700 to 1500 nm, contains information about the complexity of the canopy (e.g. its structure and related biomass accumulation). The shortwave infrared range, from 1500 to 2500 nm, contains information about the water and protein content. Such sensors can be mounted on satellites, planes or drones, the choice of the coupled carrier and sensor depending on the application and the requirements. A satellite can collect data from large areas in a short period of time, but it is limited in spatial resolution due to its distance, cannot be used for on-demand applications due to predefined orbital path and the line-of-sight from orbit can be obscured by clouds. Similar measurements from a drone are time-consuming and the sensor specifications are limited by the payload of the drone. They can on the other hand provide a higher spatial resolution, measurements can be taken on demand and are less impacted by clouds. Spectral information can then be analysed and linked to traits that are directly involved in radiative transfer mechanisms (e.g. leaf area index, surface temperature) or are a product of the combination of the aforementioned traits (e.g. biomass, nitrogen content) (Weiss *et al.* 2020).

1.4. Vegetation Indices

One of the most common indices used to monitor vegetation and build regression models for biomass accumulation is the Normalized Difference Vegetation Index (NDVI) which was first described in the 70s by Rouse *et al.* (Rouse J.W. *et al.* 1974). It is based on the relationship between the absorption and reflection of light in the red and near-infrared (NIR) part of the light spectrum by plants (Boschetti *et al.* 2007).

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

Denser vegetation with more biomass absorbs more red light resulting in an NDVI closer to 1 while areas with less dense vegetation, absorbing less red light, are closer to 0 or negative for highly reflective surfaces like water, snow or clouds resulting in one variable to describe the vegetation.

Another well-known vegetation index is the Normalized Difference Red Edge (NDRE) Index. It is similar to the NDVI but is based on vegetation's absorption and reflection of light in the red and red edge (REDGE) part of the light spectrum and is also used to build regression models for chlorophyll content (Boiarskii & Hasegawa 2019).

$$NDRE = \frac{NIR - Red\ Edge}{NIR + Red\ Edge}$$

While the red light range is absorbed by the canopy's top layer, using information from the REDGE part of the spectrum, which is not absorbed as strongly by the top layer, can provide more information from layers below the canopy's topmost layer. This is especially useful in dense vegetation or in the vegetation's later growth stages where the NDVI is rather insensitive and can easily saturate while the NDRE still provides useful information (Eitel *et al.* 2011).

1.5. Models

By linking spectral data to traits, or properties connected to traits, models can be created which can then be used to estimate these traits and track their changes through both time and space. A model is essentially a simplified reproduction of the real world and is limited to only portraying certain aspects of interest of reality (Smith & Smith 2007). Models can for example be based on regression methods defined by mathematical equations that transform input data into output data with as little errors as possible. Usually, models are first calibrated, which means that they are optimized and trained to transform a set of experimental input data into a set of known output data. Then these models are validated in order to determine how they perform in predicting output data from input data that has not been used during calibration. In other words, an evaluation of how well the defined mathematical equations approximate the chosen aspect of the real world in terms of errors and uncertainties is performed (Trucano *et al.* 2006). The process of calibration and validation can be repeated with different datasets, e.g. from different locations, until the model performs well enough for the chosen application. This model can then provide farmers with useful information for monitoring their crops and planning inputs accordingly.

1.6. Objective and Hypothesis

The objective of this project was to estimate biomass and chlorophyll content of forage grasslands by building regression models based on field measurements and spectral information from satellite imagery.

The hypothesis was that the spectral and spatial characteristics of the Sentinel-2 constellation can provide useful information to estimate the aforementioned traits.

2. Methods and Materials

The workflow of this project was divided into (i) data collection in the field during the growing season, (ii) data analysis and (iii) regression model building.

2.1. Field Sampling

Field data were collected on four different forage grassland fields (F1 - F4) at the SLU Röbäcksdalen field station (SLU, Umeå, Figure 1).

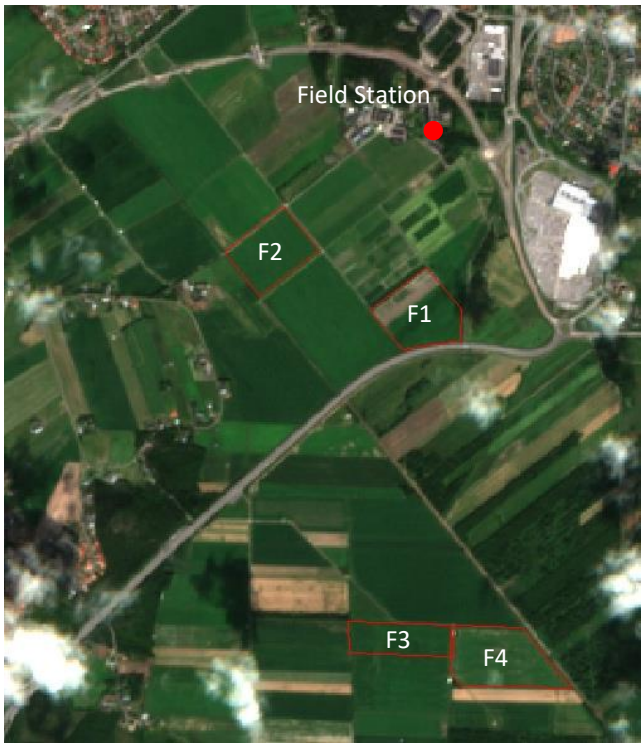

Figure 1. Overview of sampled fields F1 – F4 around Röbäcksdalen field station.

The fields contained a mixture of timothy-grass and red clover and were included in a 3-harvests cycle over the growing season. Table 1 summarizes the number of sampling occasions for each field.

Table 1. Overview of number of sampling occasion per cycle and field

Cycle	Field 1	Field 2	Field 3	Field 4
1	3	2	1	1
2	2	3	1	0
3	4	2	1	2
Total	9	7	3	3

For each of the sampling occasions, one field was selected and three 10 x 10 m areas, called plots, with even plant growth were chosen. Within each of these plots, three subplots of 50 x 50 cm were chosen in order to account for potential spatial heterogeneity. Sample data was collected and aggregated on the subplot level.

GPS positions were recorded using a Trimble Geo 7X with a centimetre precision (on one occasion an Apple iPhone SE was used instead). Height of the tallest plant, average height of the canopy, total leaf chlorophyll content (C_{ab}) and reflectance quantities of the canopy were measured. Biomass samples were collected using a 50 x 50 cm sampling quadrat. In order to reproduce farming conditions, 8 cm of stubble were left uncut. In addition, field conditions (e.g. bare ground spots) and environmental conditions that could possibly influence measurements (e.g. intermittent overcast or wet grass) were recorded.

Figure 2 shows an overview of the different field sampling steps.

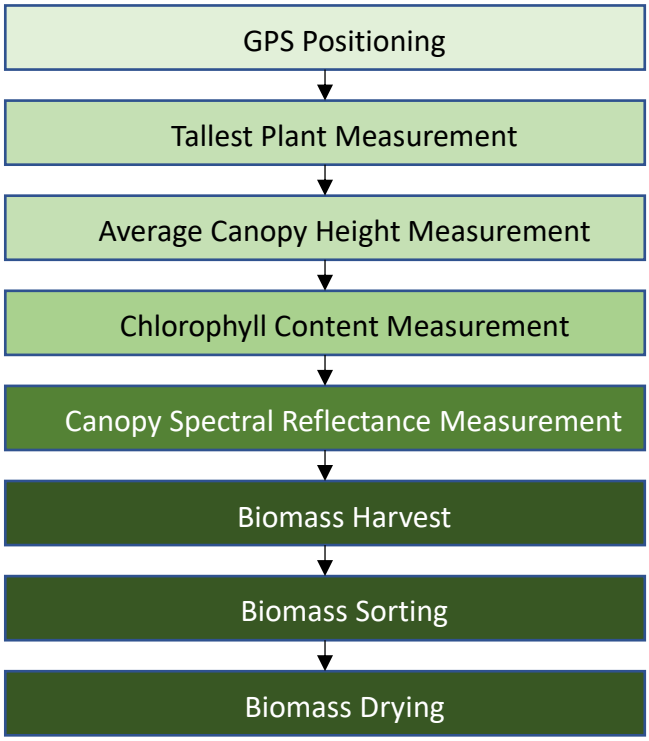

Figure 2. Overview of field sampling workflow

2.1.1. Plant Height Measurement and Biomass Harvest

In each subplot, the height of the tallest grass plant and the average height of the canopy were measured using a ruler. Harvested samples were stored in a cooling box. In the laboratory, the samples were measured for their fresh weight and hand separated into grass, legumes and weed groups for botanical composition determination. Each group was then separately oven-dried for 48 hours at 60° Celsius. Legume and weed samples below 1g were oven-dried together with the grass. Finally, dry-weights were recorded. All data were stored in a Google spreadsheet for further analysis.

2.1.2. Total Leaf Chlorophyll Content Measurement

In each subplot the average and standard deviation total leaf chlorophyll content of 9 randomly selected grass and legume leaves (when available) were measured using a Dualex DX18093 (Figure 3), which estimates chlorophyll content from measuring leaf transmittance in the red region of the light spectrum (Cerovic *et al.* 2012).

Figure 3. Dualex DX18093 leaf chlorophyll meter.

2.1.3. Canopy Spectral Reflectance Measurement

The spectra for canopy spectral reflectance (CSR) as defined by Schaepman-Strub *et al.* (Schaepman-Strub *et al.* 2006) were measured with a hand-held Yara-N sensor (Figure 4). The Yara N-sensor is a field spectrometer with a field of view of 25°, measuring the canopy-reflected light in 60 discrete bands from 400 nm to 1000 nm (visible to near-infrared regions), with a spectral sampling of 10 nm. It consists of an incoming light sensor (upward facing) and a reflectivity sensor (downward

facing). The measurements for each subplot were performed from 4 directions to account for variations in the sun-sensor geometry with a zenith viewing angle of 45°. Acquired spectra were converted to reflectance using a dedicated software (Yara FieldSpec, Yara), and the subplot-averaged reflectance spectrum further calculated by averaging the four subplots.

Figure 4. Hand-held Yara-N sensor (field spectrometer)

2.2. Satellite Imagery

Level 2A (bottom of atmosphere reflectance) Sentinel-2 satellite images were used for the regression models and downloaded from the Copernicus Open Access Hub (ESA 2020). The Sentinel-2 constellation consists of 2 satellites that provide open-access images with a high spatial resolution (ranging from 10 to 60 m) and a high frequency of revisit (approximately every second day at Sweden's latitudes). Multispectral images obtained consist of 13 bands ranging from the visible to the shortwave infrared regions. These images have been pre-processed by the European Space Agency (ESA) to account for geometric distortion and radiometric calibration. An automatically created 20 m resolution scene classification map is

also available for masking clouds and shadows. For this project 11 bands and the scene classification map were selected (Table 2).

Table 2. Details of the Sentinel-2 spectral band chosen for this project (adapted from ESA 2020*)*

Band	Name	Central Wavelength	Resolution
B02	Blue	490 nm	10 m
B03	Green	560 nm	10 m
B04	Red	665 nm	10 m
B05	Red Edge 1	705 nm	20 m
B06	Red Edge 2	740 nm	20 m
B07	Red Edge 3	783 nm	20 m
B08	Near-infrared (NIR)	842 nm	10 m
B8A	Red Edge 4	865 nm	20 m
B09	Water Vapor	940 nm	60 m
B11	Shortwave Infrared 1 (SWIR 1)	1610 nm	20 m
B12	Shortwave Infrared 2 (SWIR 2)	2190 nm	20 m
SCL	Scene Classification Map	N/A	20 m

Satellite images were cropped to the area of the four sampling fields for faster processing and the bands in 20 m and 60 m resolution were resampled to 10 m resolution. Scene classification maps were used to mask pixels with cloud cover. The reflectance values from pixel coordinates corresponding to the respective subplot coordinates were extracted from the satellite images. Then a check on which satellite image dates coincide with the sampling dates was performed. For the sampling dates where no satellite image was available, the reflectance was interpolated by calculating the average reflectance between satellite images before and after the sampling date. The average was weighed according to the time differences between the two satellite image dates and the sampling date. An overview of the satellite image processing workflow is shown in Figure 5.

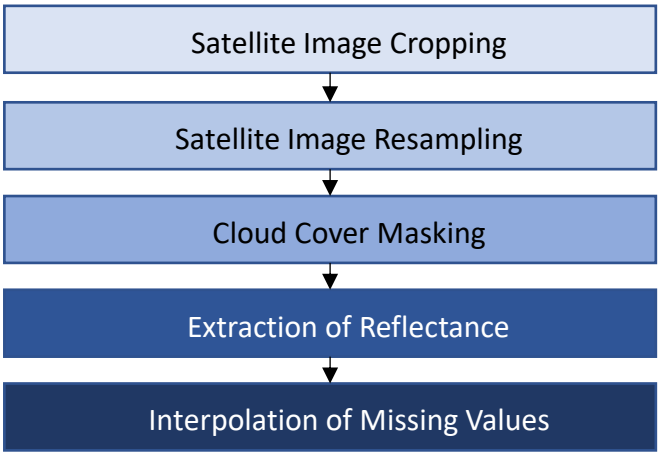

Figure 5. Overview of satellite image processing workflow.

All data processing was performed in R (R Core Team 2013) under macOS Catalina except for image downloading, resampling and cropping which was performed in R under Windows 10. The source code for this project is available on request. Satellite image processing and coordinate conversion from the SWEREF99 (Trimble Geo 7x) and ESRI:4023 (Apple iPhone SE) formats into the EPSG:32634 format of the satellite images was performed with the *raster* R-package (Hijmans 2020) using standardized format definitions from the EPSG website (MapTiler Team 2020). For date and time conversion operations the *lubridate* R-package was used (Grolemund & Wickham 2011). In addition to the built-in plot function, the *ggplot2* R-package (Wickham 2016) was used for graphs and diagrams. Field limit shapefiles were created with QGIS (QGIS Development Team 2009). The sampling data from the Google spreadsheet were imported using the *googledrive* R-package (D'Agostino McGowan & Bryan 2020).

2.3. Regression Models

The dataset used for building the regression models contained 198 samples. The predictor variables consisted of the 11 different bands from the satellite images and the NDVI. The NDVI was calculated from the red and NIR bands of the satellite images. The initial approach to use a simple linear regression with the NDVI as a baseline to evaluate the performance of other regression methods was dropped as this approach showed a very weak correlation at best. The NDRE showed a similarly weak correlation which is why the NDVI was used as a one-variable predictor set for the other regression models. In addition to using all spectral information, three different combinations were tested for building regression models by (i) adding the NDVI to all bands to investigate how adding a set of predictors that is dependent on existing predictors affects the models, (ii) building models with the NDVI only to test whether data from commercially available instruments that estimate the NDVI (e.g. Trimble GreenSeeker) could be used for this purpose and (iii) using the RGB bands to explore whether comparatively inexpensive equipment like an RGB camera would achieve similar results (Table 3).

Table 3. Spectral bands/NDVI combinations chosen as predictors.

Combination	Bands
All bands	B02, B03, B04, B05, B06, B07, B08, B8A, B09, B11, B12
All bands + NDVI	B02, B03, B04, B05, B06, B07, B08, B8A, B09, B11, B12, NDVI
NDVI	NDVI
RGB	B02, B03, B04

The NDVI was also calculated from the canopy spectral reflectance data collected with the Yara-N sensor but the small advantage in performance of the regression models using all bands and the Yara-N NDVI and using only the Yara-N NDVI did not justify further investigation at that point.

The response variables dry matter yield, canopy average height and total leaf chlorophyll content (primary grass) were selected for analysis due to their practical use as indicators for amount and quality of biomass in the field and as a check to verify that the correlated dry matter yield and canopy average height showed similar trends.

The regression models were built using partial least square (PLS) regression and support vector machines (SVM). They were calibrated with a randomly selected subset of the dataset containing 2/3 of the samples. Then the models were validated, using the remaining 1/3 of the samples to predict the response variables. Both calibration and validation subset were the same for every regression method and response variable. Finally, the results of both the calibration and the validation (prediction performance) was compared. For the SVM models, a check on whether prediction errors coincided with instances where the reflectance information from the satellite images had been interpolated was performed as well. An overview of the regression model building workflow is shown in Figure 6.

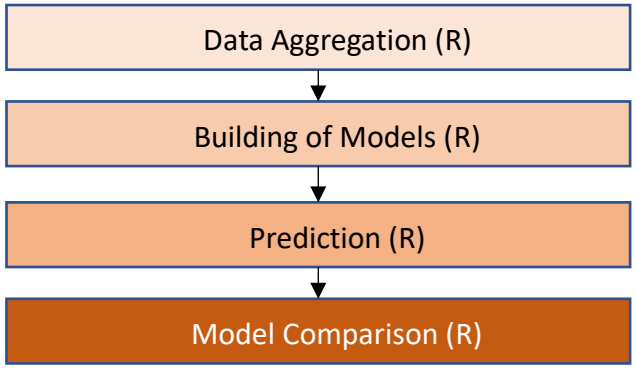

Figure 6. Overview of regression model building workflow.

2.3.1. Partial Least Squares

PLS regression is an analysis method for linear relationships between two datasets, the predictor variables and response variables. This method is especially useful when dealing with large sets of correlated predictor variables (Wold *et al.* 2001).

The PLS regression models for this project were created with the *pls* R-package (Mevik *et al.* 2020), using a leave-one-out cross validation for model calibration. A leave-one-out cross validation aims to use all the available samples of a *n*-size dataset to assess the performance of a model, using *n*-1 samples for calibration and

the remaining samples for validation and repeating the process until every sample has been used for both validation and calibration.

2.3.2. Support Vector Machines

SVM is an algorithm for classification and regression purposes developed by Vapnik (Vapnik 1982). The principle of SVM regression is based on mapping predictor and response variables into a higher dimensional space in order to find a relationship between them while maintaining low error rates in the original dimension.

For this project the SVM regression models were created with the *liquidSVM* R-package (Steinwart & Thomann 2017a). For building the model, a 10-fold cross validation was used for model calibration. A leave-one-out cross validation was tested for calibration as well but the models built with a 10-fold cross validation performed slightly better. Using a 10-fold cross validation, the dataset is split into 10 sub-datasets that are iteratively used for calibration and validation.

3. Results

3.1. Field Sampling

Due to the low sampling frequency, development trends could not be assessed properly in fields 3 and 4. These results have therefore been omitted here

3.1.1. Dry Matter Yield

Dry matter yield accumulation showed various rates of increase, depending on the field and growth/regrowth cycle both at the field level and plot levels (Figure 7).

The average and maximum dry matter yield did not reach the same quantities in the second cycle as in the first cycle and were even lower in the third cycle. The extent of spreading in sampling values decreased towards the end of the growing season.

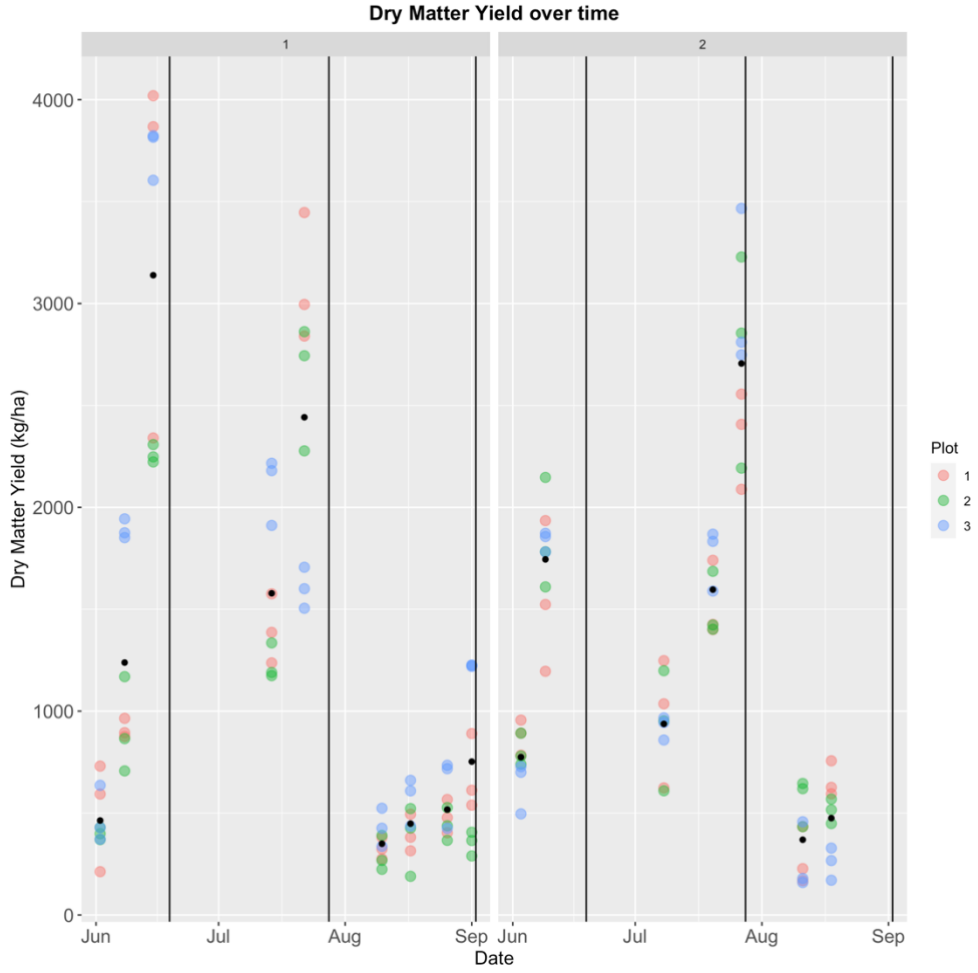

Figure 7. Dry Matter Yield (kg/ha) over time in forage grassland fields 1 (left) and 2 (right) including means at field level (black dots) and harvest dates (black vertical lines).

3.1.2. Canopy Average Height

The development of the canopy average height followed a similar pattern as the development of dry matter yield. An increase in the mean canopy average height over the first cycle was followed by a slower increase and reduced height in subsequent cycles. The spreading of measurement results decreased over the growing season. The development of the average height of the canopy for fields 1 and 2 are shown in Figure 8.

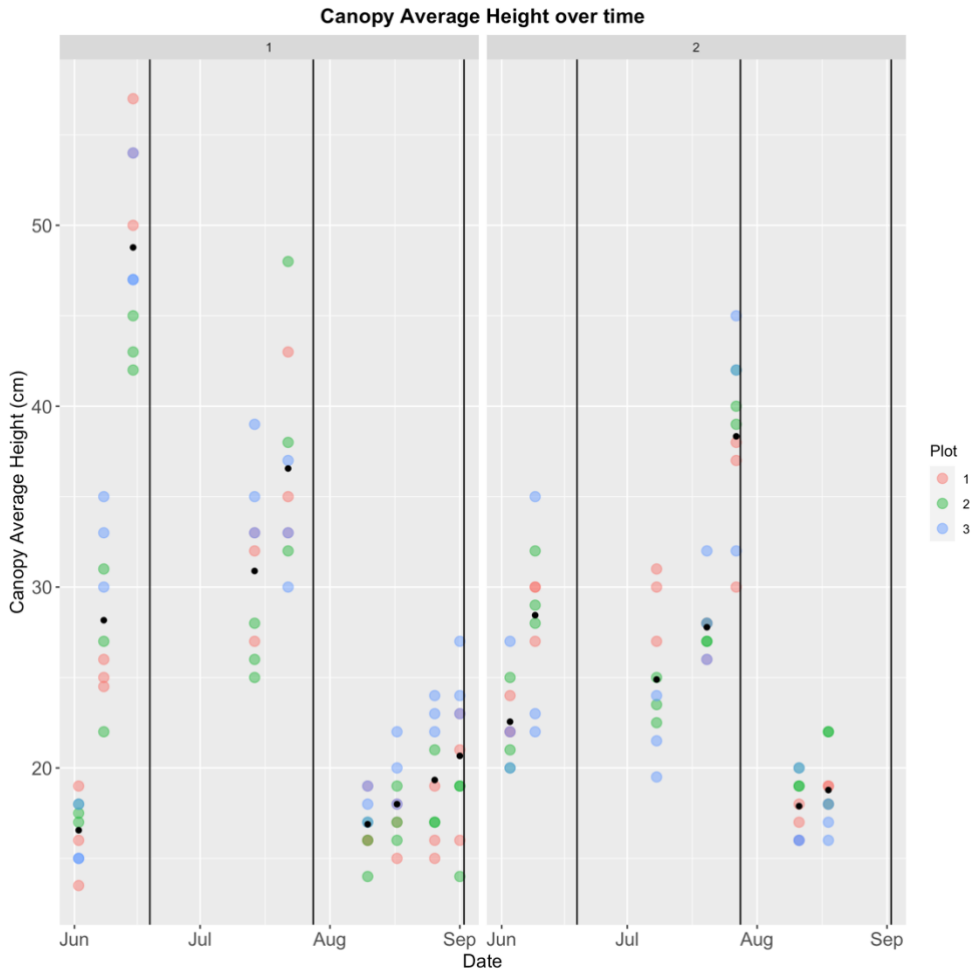

Figure 8. Canopy Average Height (cm) over time in forage grassland fields 1 (left) and 2 (right) including means at field level (black dots) and harvest dates (black vertical lines).

3.1.3. Total Leaf Chlorophyll Content

The total leaf chlorophyll content showed high standard deviation for individual sampling occasions, especially during the first and third cycle in field 1 and in the first cycle in field 2. The mean of the total leaf chlorophyll content decreased during the first cycle in both field 1 and field 2 and increased during the second cycle. During the third cycle an initial decrease was followed by an increase in the mean in field 1 while it increased in field 2. This is illustrated Figure 9.

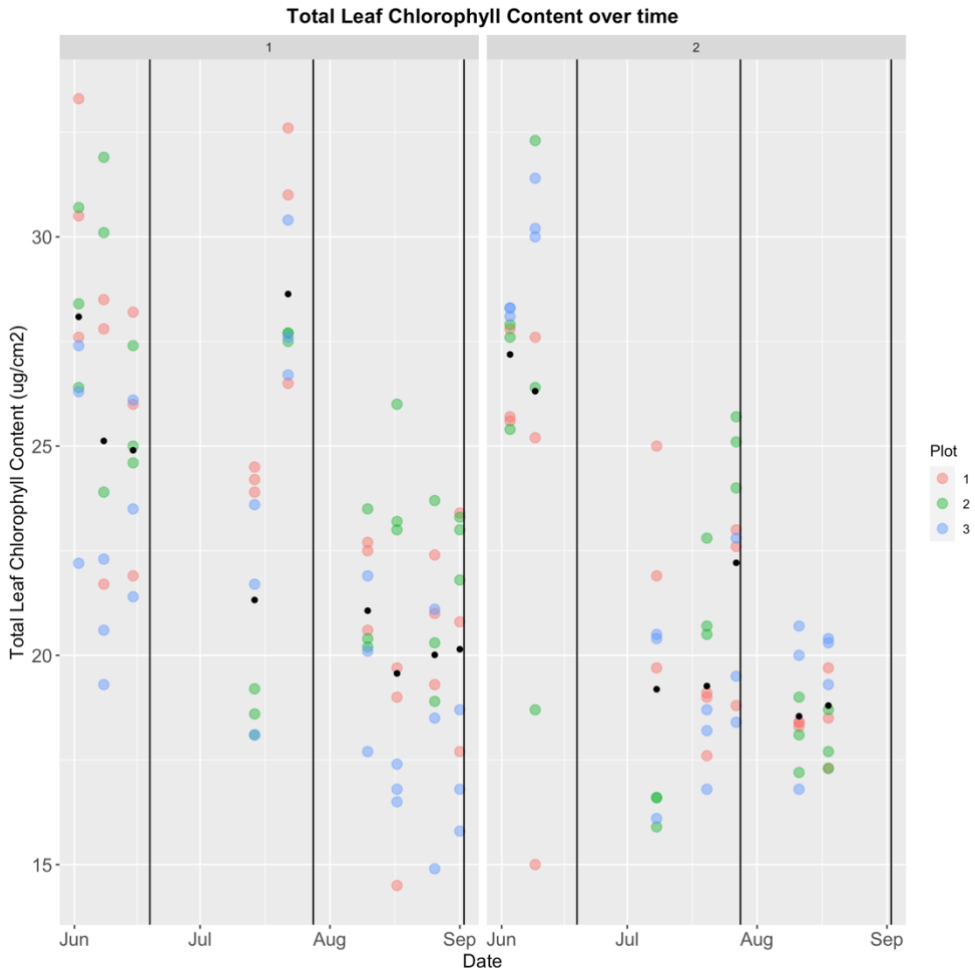

Figure 9. Total Leaf Chlorophyll Content (µg/cm^2) over time in forage grassland fields 1 (left) and 2 (right) including means at field level (black dots) and harvest dates (black vertical lines).

3.1.4. Canopy Spectral Reflectance

The dynamic of CSR is illustrated by measurements taken in field 1. The CSR signatures showed a peak in the green part of the visible spectrum (around 550 nm) and a sharp increase in the red edge part of the spectrum (680 – 740 nm). This increase flattened out in a plateau in the NIR part (740 – 1000 nm) and concluded in a decrease in the NIR part (around 970 nm). The green peak decreased in intensity during the first and second cycle and increased in the third cycle. The reflectance in the NIR part increased over the length of the cycles except for one plot in the second cycle and one sampling occasion in the third cycle. The CSR signatures for field 1 are illustrated in Figure 10 (first cycle), Figure 11 (second cycle), Figure 12 (third cycle, first part) and Figure 13 (third cycle, second part).

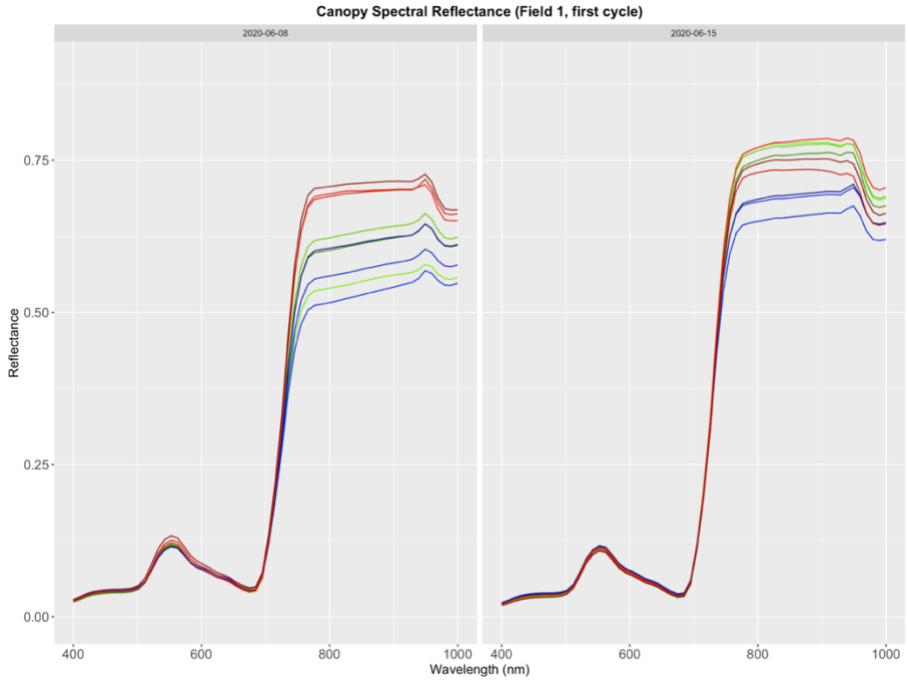

Figure 10. CSR profile, first cycle field 1 (colour groups represent different plots, colour shades represent different subplots)

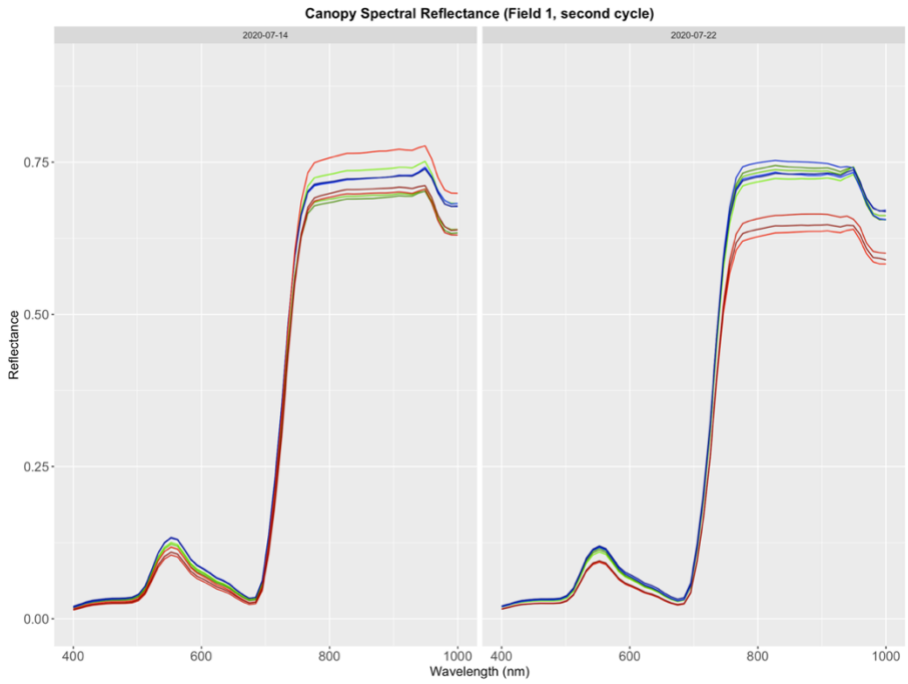

Figure 11. CSR profile, second cycle field 1 (colour groups represent different plots, colour shades represent different subplots)

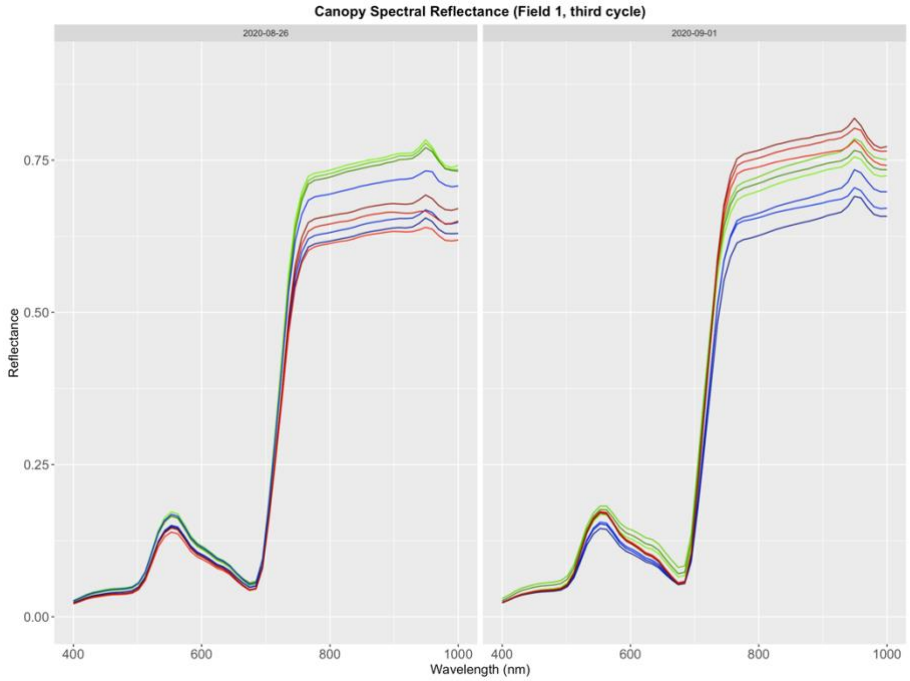

Figure 12. CSR profile, third cycle (first part) field 1 (colour groups represent different plots, colour shades represent different subplots)

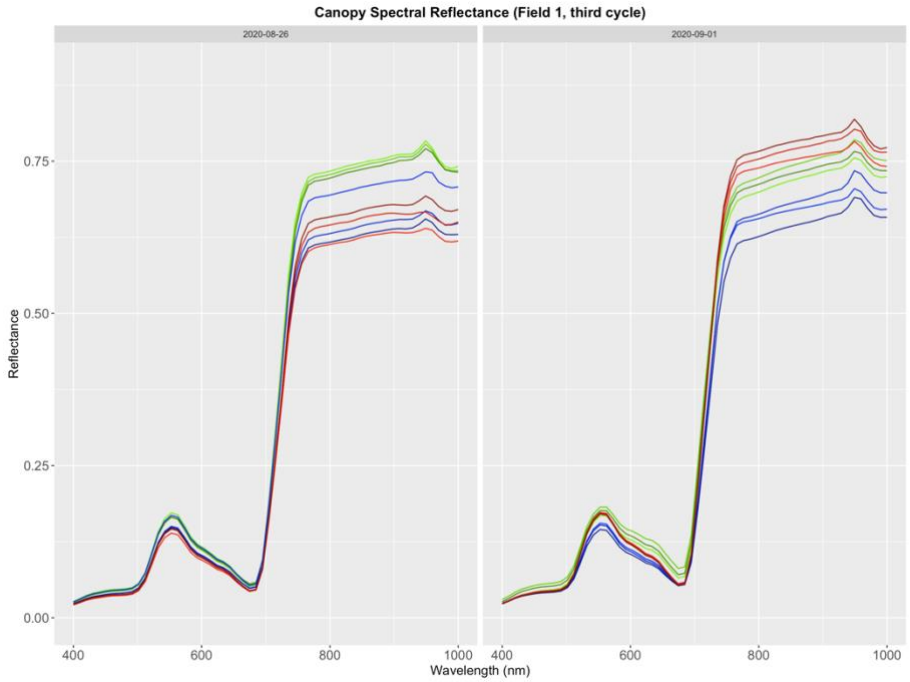

Figure 13. CSR profile, third cycle (second part) field 1 (colour groups represent different plots, colour shades represent different subplots)

3.2. Regression Models

3.2.1. Dry Matter Yield

The validation of the PLS model for dry matter yield showed a relatively poor prediction performance of the model, with high deviations between measured and predicted values and slopes and intercepts far from 1 and 0, respectively. The accuracy of the regression generally decreased when using less spectral information. However, too much spectral information leads to overestimation (Figure 14).

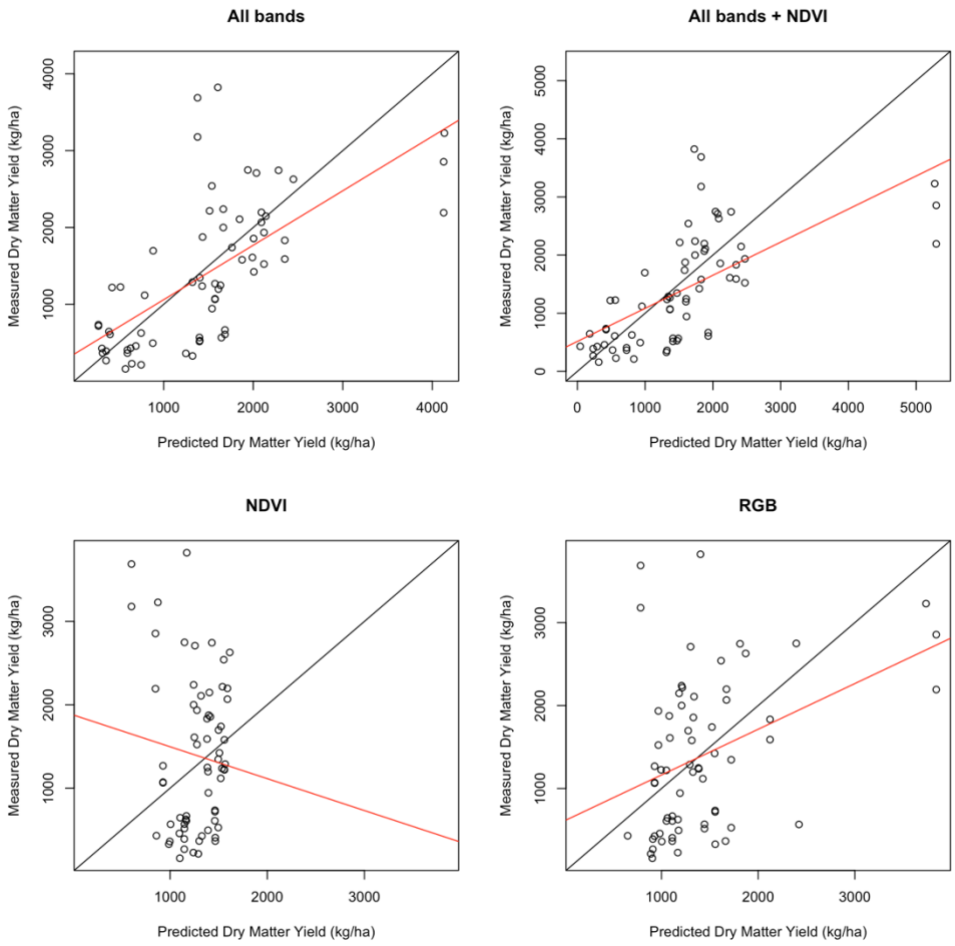

Figure 14. PLS Model Evaluation Results - Measured Dry Matter Yield (y-axes) vs. Predicted Dry Matter Yield (x-axes) in forage grasslands using different band combinations including 1:1 lines (black) and regression lines (red).

The SVM model validation for dry matter yield showed good performances. Slopes and intercepts of the different models were close to 1 and 0, respectively, but variance increases with increasing dry matter yield. Reducing the number of

predictors by limiting the number of bands generally increased the errors (Figure 15).

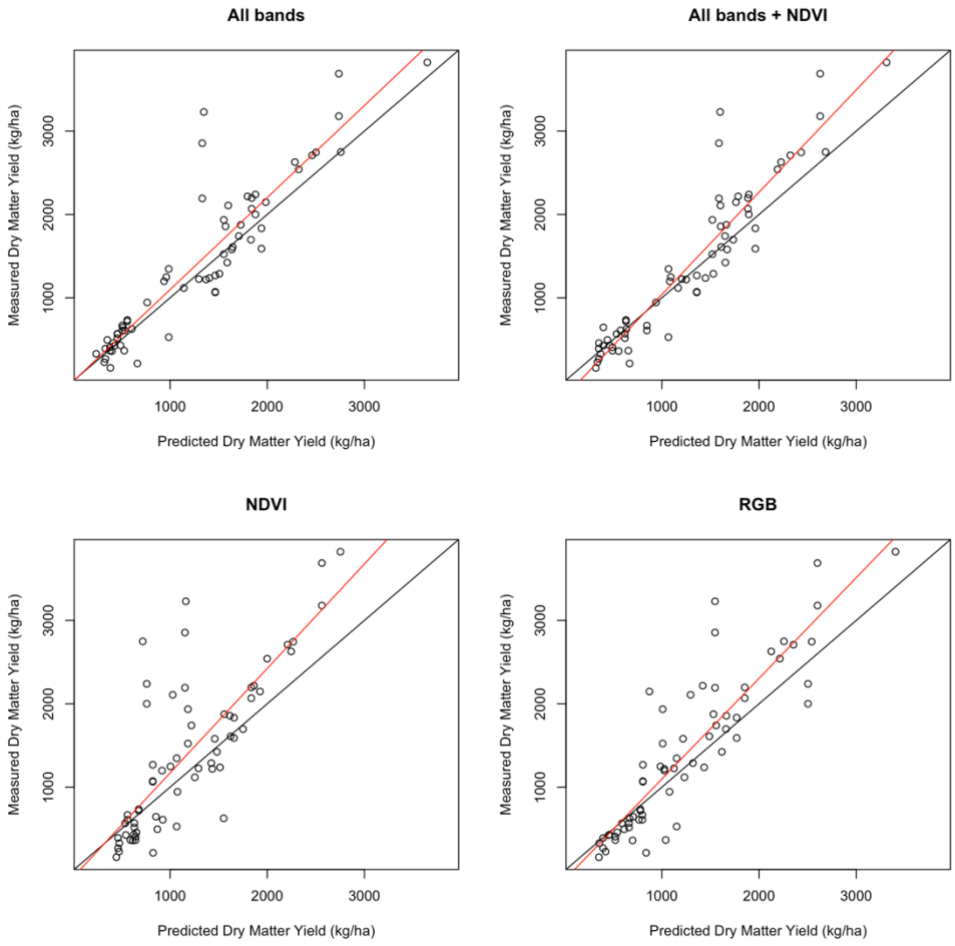

Figure 15. SVM Model Evaluation Results - Measured Dry Matter Yield (y-axes) vs. Predicted Dry Matter Yield (x-axes) in forage grasslands using different band combinations including 1:1 lines (black) and regression lines (red).

A direct comparison between R^2 and RMSE values for the PLS and SVM model calibration and validation for dry matter yield underlined the performance difference. The relative performance drop from calibration results to validation results was more pronounced for SVM, especially for a limited number of predictors. However, the poorest performing SVM model still achieved higher R^2 and lower RMSE scores during validation than the best performing PLS model during calibration (Table 4).

Table 4. R^2 and RMSE for Dry Matter Yield PLS and SVM models for forage grasslands showing calibration and validation statistics.

	Calibration R^2		Validation R^2		Calibration RMSE		Validation RMSE	
	PLS	SVM	PLS	SVM	PLS	SVM	PLS	SVM
All	0.47	0.98	0.42	0.84	654	134	751	403
All + NDVI	0.51	0.97	0.41	0.88	630	191	854	379
NDVI	0.06	0.96	0.01	0.63	874	278	989	623
RGB	0.18	0.97	0.14	0.79	817	207	904	470

3.2.2. Canopy Average Height

Similarly to dry matter yield, PLS-based regression models for estimating the canopy average height showed poor performance and high dispersion of the sample points around the 1:1 line. Scattering increased when limiting the number of bands used but overestimation tendencies decreased (Figure 16).

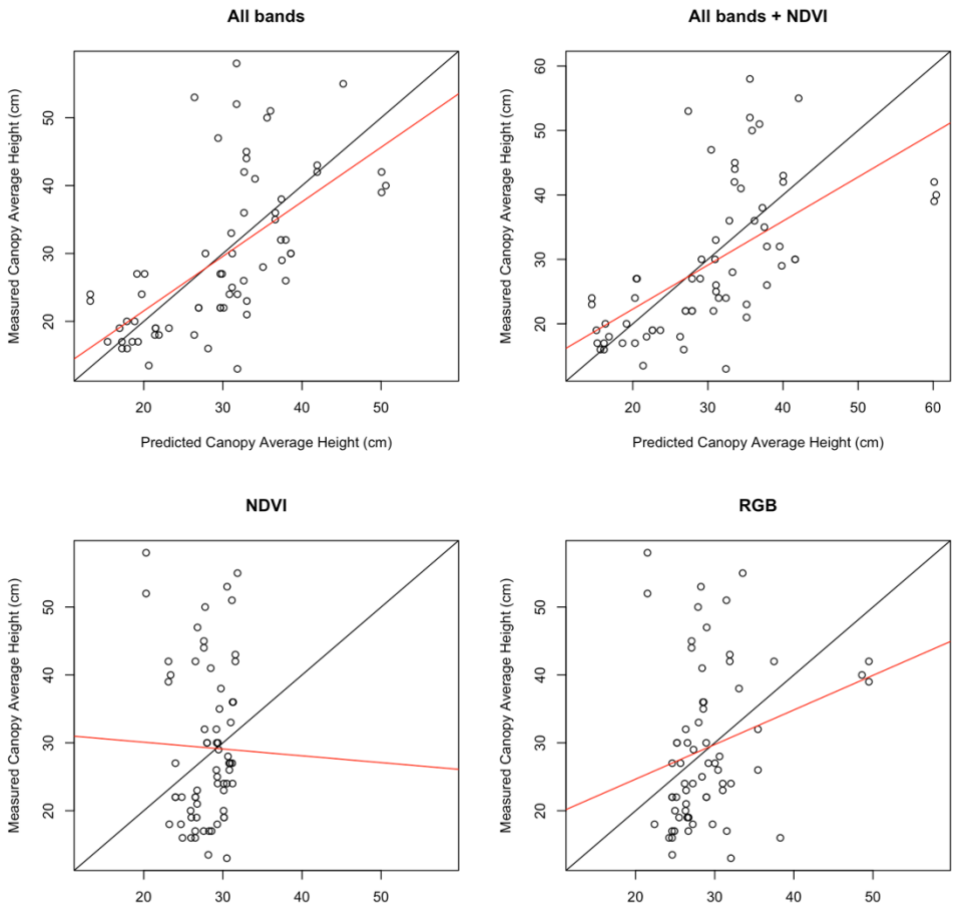

Figure 16. PLS Model Evaluation Results - Measured Canopy Average Height (y-axes) vs. Predicted Canopy Average Height (x-axes) in forage grasslands using different band combinations including 1:1 lines (black) and regression lines (red).

The SVM model for canopy average height showed lower error rates and scattering than the PLS model. Reducing the number of predictors by limiting the number of bands generally increased the errors (Figure 17).

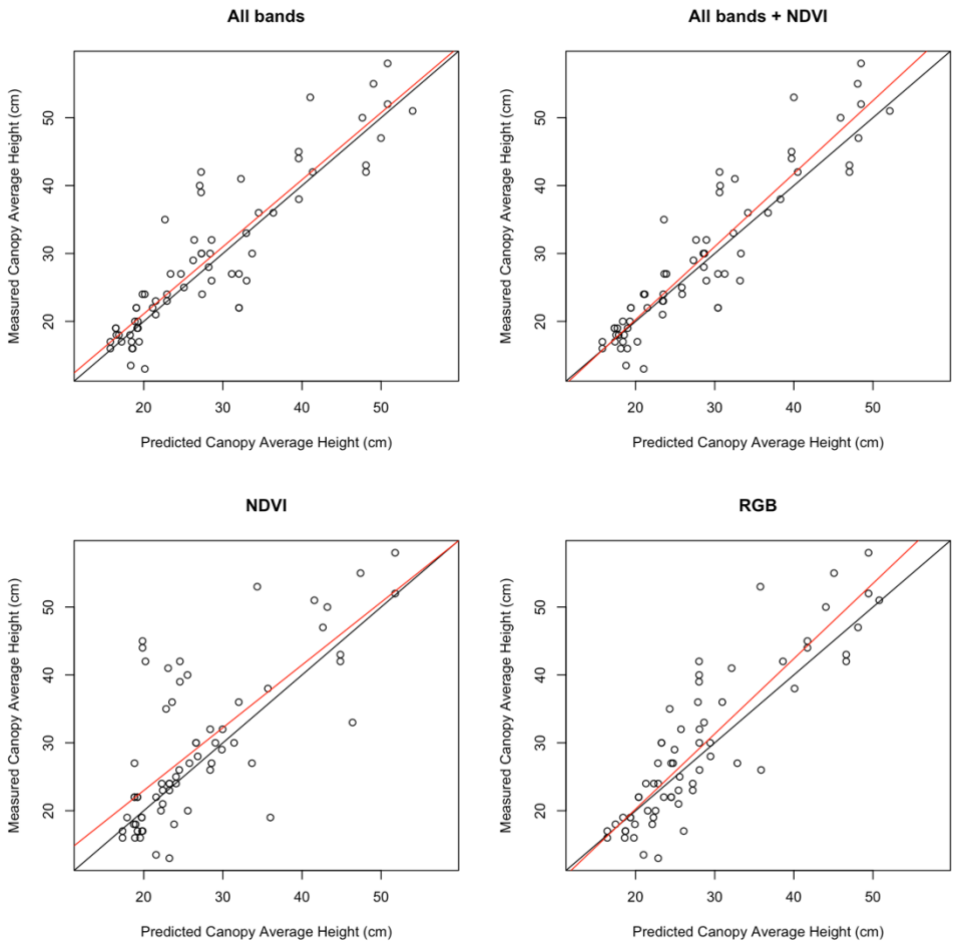

Figure 17. SVM Model Evaluation Results - Measured Canopy Average Height (y-axes) vs. Predicted Canopy Average Height (x-axes) in forage grasslands using different band combinations including 1:1 lines (black) and regression lines (red).

A comparison between R^2 and RMSE for the PLS and SVM model calibration and validation for canopy average height showed the differences in performance. The relative performance drop from calibration results to validation results was more pronounced for SVM, especially with a limited number of predictors. Using all bands for example, PLS reached the same R^2 for validation and calibration while SVM showed a 15% drop from calibration to validation. However, SVM performed better during validation than PLS during calibration in all scenarios (Table 5).

Table 5. R^2 and RMSE for Canopy Average Height PLS and SVM models for forage grasslands showing calibration and validation statistics.

	Calibration R^2		Validation R^2		Calibration RMSE		Validation RMSE	
	PLS	SVM	PLS	SVM	PLS	SVM	PLS	SVM
All	0.38	0.95	0.38	0.81	8.7	3.2	9.2	5.1
All + NDVI	0.40	0.96	0.37	0.85	8.6	2.3	9.7	4.6
NDVI	0.05	0.95	0.00	0.52	10.9	3.3	12	8.4
RGB	0.09	0.92	0.06	0.78	10.6	3.6	11.5	5.6

3.2.3. Total Leaf Chlorophyll Content

PLS-estimation for total leaf chlorophyll content showed a similar pattern as for dry matter yield and average canopy height, as the model poorly managed to capture the variability in the dataset. However, for total leaf chlorophyll content, PLS appeared to be less impacted by reducing the number of predictors (Figure 18).

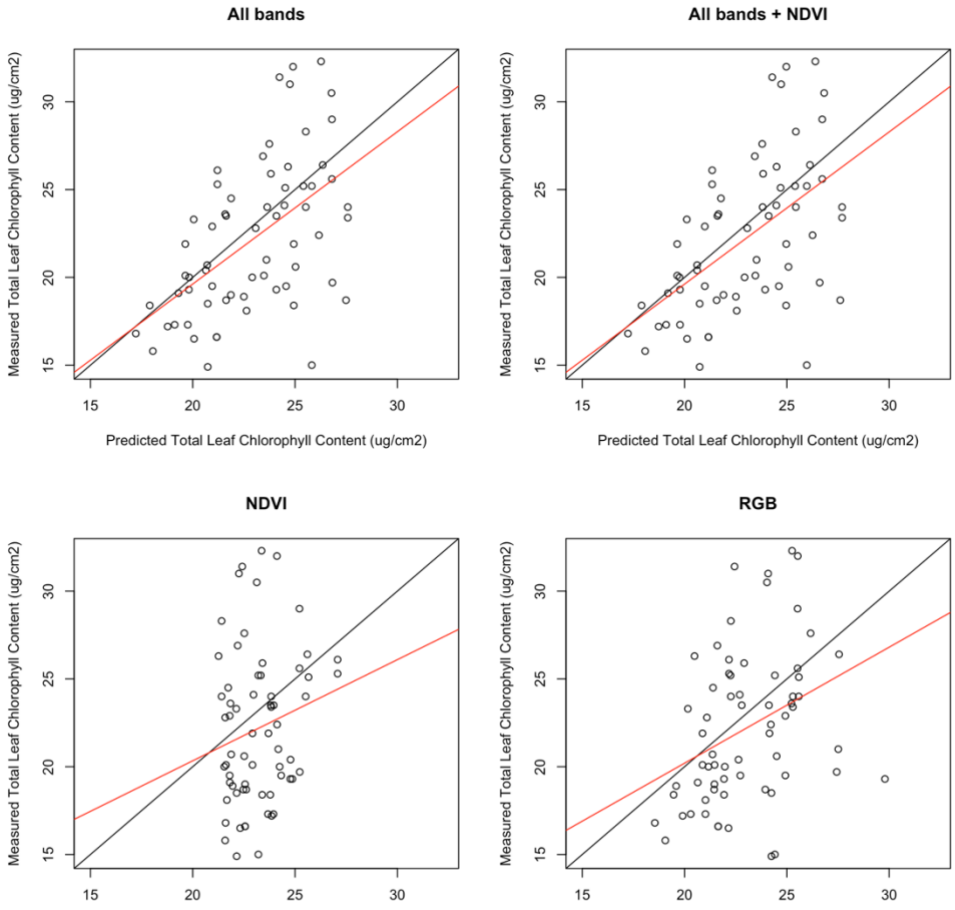

Figure 18. PLS Model Evaluation Results - Measured Total Leaf Chlorophyll Content (y-axes) vs. Predicted Total Leaf Chlorophyll Content (x-axes) in forage grasslands using different band combinations including 1:1 lines (black) and regression lines (red).

The SVM model for total leaf chlorophyll content showed lower error rates and scattering than the PLS model. Reducing the number of predictors by limiting the number of bands had a higher impact on errors than for dry matter yield and canopy average height (Figure 19).

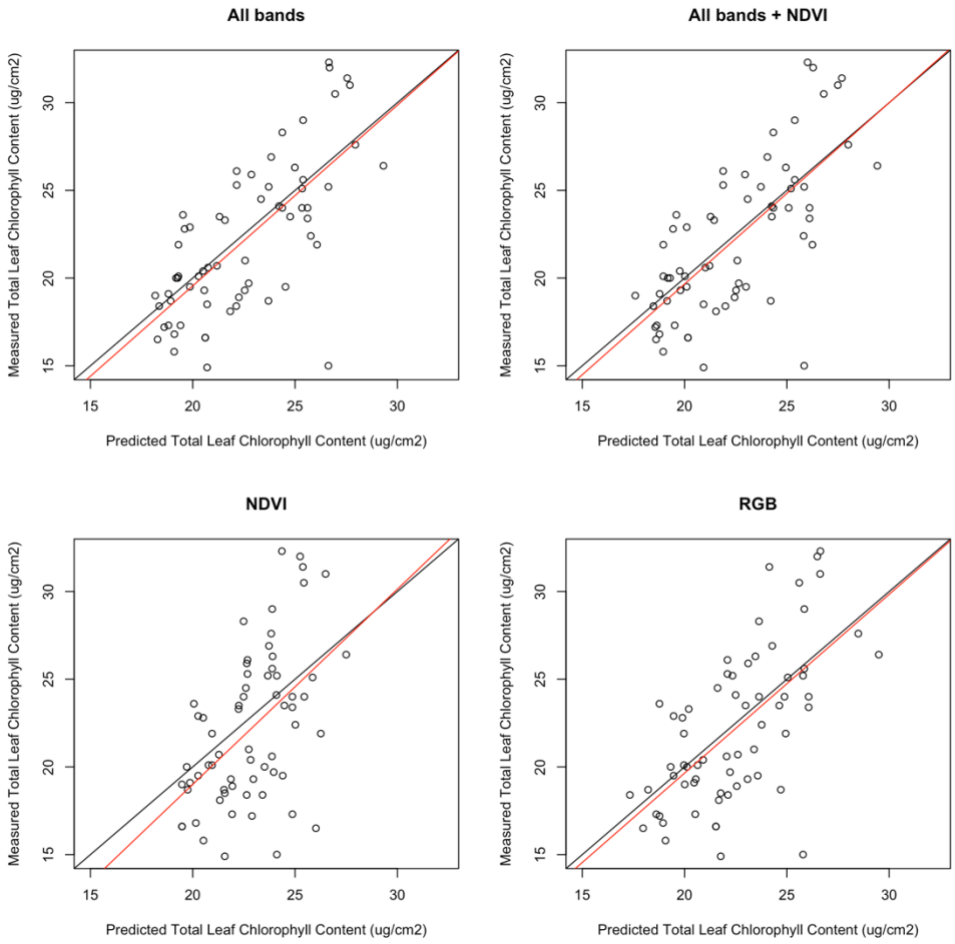

Figure 19. SVM Model Evaluation Results - Measured Total Leaf Chlorophyll Content (y-axes) vs. Predicted Total Leaf Chlorophyl Content (x-axes) in forage grasslands using different band combinations including 1:1 lines (black) and regression lines (red).

A comparison between R^2 and RMSE for the PLS and SVM model calibration and validation for total leaf chlorophyll content showed the differences in performance. The relative performance drop from calibration results to validation was almost the same for PLS and SVM when looking at R^2. Looking at RMSE, the performance drop was more pronounced for SVM. However, SVM performed almost equally well during validation than PLS during calibration and sometimes better than PLS during calibration (Table 6).

Table 6. R^2 and RMSE for Total Leaf Chlorophyll Content PLS and SVM models for forage grasslands showing calibration and validation statistics.

	Calibration R^2		Validation R^2		Calibration RMSE		Validation RMSE	
	PLS	SVM	PLS	SVM	PLS	SVM	PLS	SVM
All	0.5	0.84	0.29	0.5	2.88	1.73	3.76	3.11
All + NDVI	0.5	0.85	0.29	0.5	2.87	1.69	3.75	3.09
NDVI	0.09	0.74	0.03	0.26	3.87	2.51	4.43	3.81
RGB	0.33	0.68	0.12	0.40	3.32	2.33	4.22	3.37

4. Discussion

In comparison with previous years there was a lack of legumes in the forage grassland fields in Röbäcksdalen in 2020, which is due to the harsh winter of 2019/2020.

In general, the growth of the primary grass was not very homogenous, especially in the beginning of the growth season. One possible reason for this is the uneven topographic of the fields that could have caused water to collect in sinks in autumn. During spring, the ice in sinks takes longer to thaw and could have delayed plant growth in these locations.

While weed growth in the fields was negligible during the first growth/regrowth cycle, considerable amounts of weed were observed in almost all sampled fields after the first harvest. The amount of legumes tends to increase after the first cut and the very low amounts of legumes in the second and third cycle could have facilitated weed emergence and establishment due to the lack of competition by legumes.

The resulting heterogeneity in the fields had an impact on the initial sampling plan. Instead of sampling every field once a week most of the samples were collected in fields 1 and 2. For the same reason canopy average height was chosen as a trait instead of the height of the tallest plant. While measurement of the tallest plant is less subject to operator bias, it does not accurately represent the height of the vegetation in the field in the case of a very heterogenous field.

4.1. Field sampling

4.1.1. Canopy Average Height and Biomass

The aforementioned heterogeneity can explain the variability in the dry matter yield and canopy average height measurement results at both field level and plot level. After the initial variations in soil conditions caused by water in sinks, the impact of the topographical differences in the field decreases which could explain the lower degree of spreading in cycles 2 and 3. The lower level of dry matter yield and canopy average height in cycle 2 and 3 can be explained by the slow rate of

regrowth in northern timothy varieties (Nissinen *et al.* 2010), the competition with weeds, the decrease in day-length and possibly solar radiation and temperature.

4.1.2. Total Leaf Chlorophyll Content

Differences in the stage of growth can partly explain the spreading of the total leaf chlorophyll content measurement results in the first cycle in fields 1 and 2 as the chlorophyll content and leaf greenness change during different developmental stages (Radkowski 2013).

In the second cycle in both fields and the third cycle in field 2 the results are consistent with expectations that the chlorophyll content in the leaves increases towards the end of the growth cycle.

The reason for the excessive variability in the third cycle in field 1 was the prevalence of a disease in many locations. Many leaves had yellowish or brownish to black spots which suggests chloroses and consequently low total leaf chlorophyll content values as these parts of the leaves have lost a portion of their chlorophyll.

4.1.3. Canopy Spectral Reflectance

The CSR results are consistent with our expectations, as they show a peak in the visible spectrum concentrated around green, the steep incline in the red-edge part of the spectrum followed by a plateau in the NIR part. The decrease around 970 nm is due to the water content in the leaves (Peñuelas *et al.* 1993).

The lower reflectance in the NIR in one plot in the first cycle and one sampling day in the third cycle are probably due to the changing overcast conditions that were noted on these two days.

4.2. Regression Models

The SVM models outperform the PLS models in predicting dry matter yield, canopy average height and total leaf chlorophyll content as the graphs and R^2- and RMSE-values for calibration and validation show. This was expected as PLS is aimed to model linear relationships between predictor and response variables and the linear relationship between spectral information and traits analysed in this project is very weak. In projects estimating DMY (Zhou *et al.* 2019a) and wheat leaf nitrogen concentration (Yao *et al.* 2015) from spectral information, the robustness of SVM and its ability to deal with large sets of predictors has already been shown. While SVM in the majority of cases performed better during validation than PLS during calibration, the relative performance drop from calibration to validation in terms of R^2 and RMSE was generally higher for SVM than for PLS. With a limited set of predictors, this relative performance drop increased which

suggests that in practical applications one would in this case use SVM with all bands and the NDVI.

In general, the more bands SVM used to build the models, the closer the predictions were to the measured values. However, using only the RGB bands produced better result than expected. The reason for building models with spectral information from RGB bands was the question whether RGB drone images could be used as a complement or substitute for satellite images in areas with low revisit frequencies, frequent cloud cover or for on-demand applications. Common RGB cameras have a certain wavelength overlap between the RGB channels while sensors used on satellites are recording independent wavebands (Brown *et al.* 2016). This suggests that models built with spectral information from RGB images will not perform as well as models built with full Sentinel-2 spectral information, but more research is required to assess whether these models are accurate enough for practical applications.

One aspect that has to be considered concerning the method for calibration and validation of the models in this project is the dependency between the calibration and validation set. The choice which samples were used for calibration and which for validation was completely random in this project. This means that it is very likely that samples that have been collected on the same day and/or at the same location ended up in both the calibration and validation dataset. The prediction results shown here are therefore possibly better than they would be with completely independent calibration and validation datasets. Creating different random selections for calibration and validation dataset has been shown to influence the results both positively and negatively, depending on the samples chosen for each set. However, both regression methods seemed to be influenced in the same direction and to a similar extent.

Another aspect that has to be considered is that solving regression problems with SVM involves random elements and that the process is dependent on the computer architecture (Steinwart & Thomann 2017b). The solution and therefore the calibrated model will be near-optimal and results presented in this report cannot be reproduced exactly. A check with the sampling dataset and spectral information used in this report showed that different models built consecutively were slightly different and produced different predictions. The RMSE and R^2 of these models were however almost identical.

One of the issues when building regression is the risk of over-fitting. When this happens, the created model will perform well with the dataset used to create it but it will be too specific and perform worse with an unknown dataset of the same type (Mucherino *et al.* 2009). Due to the unusual heterogeneity of the fields and a lack of comparable sample data from other fields, the models in this project were only tested on data from Röbäcksdalen. For further applications, the ability of these models to generalize has to be evaluated.

Interpolation of reflectance for days with no available satellite image introduces an uncertainty which was quantified as the number of days between the images used for interpolation. Tests showed that a higher uncertainty did not consistently coincide with a higher prediction error. Some of the outliers have however been identified as samples where reflectance was interpolated over a harvest but have not been investigated further. Removing samples with high uncertainty regarding the spectral information before building the models could improve the prediction results but has not been thoroughly tested in this project.

5. Conclusion

In this project, spectral information from satellite images has successfully been used to build regression models for estimation of dry matter yield, average canopy height and total leaf chlorophyll content of forage grassland fields. Two regression methods were applied, namely partial least squares and support vector machines, and it has been shown that the latter outperformed the former both during calibration and validation.

While the support vector machines models built in this project performed well on the dataset built during this project, the question of how well these models can be generalized to other fields and locations is subject for further research.

Estimating field traits using spectral information from satellite images is a promising approach for building tools that help farmers monitor their fields and help in decision making regarding required inputs. For applications where satellite images are not available, using RGB drone images as an inexpensive alternative to using multispectral drone images might be sufficient but has to be tested.

References

Bengtsson, J., Bullock, J.M., Egoh, B., Everson, C., Everson, T., O'Connor, T., O'Farrell, P.J., Smith, H.G. & Lindborg, R. (2019). Grasslands—more important for ecosystem services than you might think. *Ecosphere*, vol. 10 (2), p. e02582 John Wiley & Sons, Ltd. DOI: https://doi.org/https://doi.org/10.1002/ecs2.2582

Boiarskii, B. & Hasegawa, H. (2019). Comparison of NDVI and NDRE Indices to Detect Differences in Vegetation and Chlorophyll Content. *International Conference on Applied Science, Technology and Engineering J. Mech. Cont.& Math. Sci.*, (4), pp. 20–29

Boschetti, M., Bocchi, S. & Brivio, P.A. (2007). Assessment of pasture production in the Italian Alps using spectrometric and remote sensing information. *Agriculture, Ecosystems & Environment*, vol. 118 (1), pp. 267–272. DOI: https://doi.org/https://doi.org/10.1016/j.agee.2006.05.024

Brown, T., Hultine, K., Steltzer, H., Denny, E., Denslow, M., Granados, J., Henderson, S., Moore, D., Nagai, S., Sanclements, M., Sanchez-Azofeifa, G.A., Sonnentag, O., Tazik, D. & Richardson, A. (2016). Using phenocams to monitor our changing Earth: toward a global phenocam network. *Frontiers in Ecology and the Environment*, vol. 14, pp. 84–93

Capstaff, N.M. & Miller, A.J. (2018). Improving the Yield and Nutritional Quality of Forage Crops. *Frontiers in plant science*, vol. 9, p. 535 Frontiers Media S.A. DOI: https://doi.org/10.3389/fpls.2018.00535

Cerovic, Z.G., Masdoumier, G., Ghozlen, N. Ben & Latouche, G. (2012). A new optical leaf-clip meter for simultaneous non-destructive assessment of leaf chlorophyll and epidermal flavonoids. *Physiologia Plantarum*, vol. 146 (3), pp. 251–260. DOI: https://doi.org/10.1111/j.1399-3054.2012.01639.x

D'Agostino McGowan, L. & Bryan, J. (2020). googledrive: An Interface to Google Drive. Available at: https://cran.r-project.org/package=googledrive

Eitel, J.U.H., Vierling, L.A., Litvak, M.E., Long, D.S., Schulthess, U., Ager, A.A., Krofcheck, D.J. & Stoscheck, L. (2011). Broadband, red-edge information from satellites improves early stress detection in a New Mexico conifer woodland. *Remote sensing of environment.*, vol. 115 (12), p. 3640—3646. DOI: https://doi.org/10.1016/j.rse.2011.09.002

Eriksen, J., Pedersen, L. & Jørgensen, J.R. (2006). Nitrate leaching and bread-making quality of spring wheat following cultivation of different grasslands. *Agriculture, Ecosystems & Environment*, vol. 116 (3), pp. 165–175. DOI: https://doi.org/https://doi.org/10.1016/j.agee.2006.02.004

ESA (2020). *Copernicus Open Access Hub*. Available at: https://scihub.copernicus.eu [2020-09-30]

Grolemund, G. & Wickham, H. (2011). Dates and Times Made Easy with {lubridate}. *Journal of Statistical Software*, vol. 40 (3), pp. 1–25. Available at: http://www.jstatsoft.org/v40/i03/

Hijmans, R.J. (2020). raster: Geographic Data Analysis and Modeling. Available at: https://cran.r-project.org/package=raster

MapTiler Team (2020). *EPSG.io: Coordinate Systems Worldwide.* Available at: https://epsg.io [2020-09-30]

Mevik, B.-H., Wehrens, R. & Liland, K.H. (2020). pls: Partial Least Squares and Principal Component Regression. Available at: https://cran.r-project.org/package=pls

Mucherino, A., Papajorgji, P. & Pardalos, P.M. (2009). A survey of data mining techniques applied to agriculture. *Operational Research*, vol. 9 (2), pp. 121–140. DOI: https://doi.org/10.1007/s12351-009-0054-6

Nissinen, O., Kallainen, P. & Jauhiainen, L. (2010). Development of yield and nutritive value of timothy in primary growth and regrowth in northern growing conditions. *Agricultural and Food Science*, vol. 19 (3 SE-Articles). DOI: https://doi.org/10.2137/145960610792912602

Olsson, Y. (2020). *Use of agricultural land in 2020, preliminary statistics* Available at: https://www.scb.se/contentassets/2e011f0876324b1a918c1e70b5ef088a/jo0104_2020a01_sm_jo10sm2001.pdf

Peñuelas, J., Filella, I., Biel, C., Serrano, L. & Savé, R. (1993). The reflectance at the 950–970 nm region as an indicator of plant water status. *International Journal of Remote Sensing*, vol. 14 (10), pp. 1887–1905 Taylor & Francis. DOI: https://doi.org/10.1080/01431169308954010

Pierce, F.J. & Nowak, P. (1999). Aspects of Precision Agriculture. In: Sparks, D.L.B.T.-A. in A. (ed.) Academic Press, pp. 1–85.

QGIS Development Team (2009). QGIS Geographic Information System. Available at: http://qgis.osgeo.org

R Core Team (2013). R: A Language and Environment for Statistical Computing. Vienna, Austria. Available at: http://www.r-project.org/

Radkowski, A. (2013). Leaf Greenness (SPAD) Index in Timothy-Grass Seed - Plantation at Different Doses of Titanium Foliar Fertilization. *Ecological Chemistry and Engineering. A*, vol. 20

Rouse J.W., J., Haas, R.H., Schell, J.A. & Deering, D.W. (1974). Monitoring Vegetation Systems in the Great Plains with Erts. *NASA Special Publication*. p. 309.

Schaepman-Strub, G., Schaepman, M.E., Painter, T.H., Dangel, S. & Martonchik, J. V (2006). Reflectance quantities in optical remote sensing—definitions and case studies. *Remote Sensing of Environment*, vol. 103 (1), pp. 27–42. DOI: https://doi.org/https://doi.org/10.1016/j.rse.2006.03.002

Smith, J.U. (viaf)61665621 & Smith, P. (2007). *Introduction to environmental modelling*. Oxford ; New York : Oxford University Press. Available at: http://lib.ugent.be/catalog/rug01:001996439

Stafford, J. V (2000). Implementing Precision Agriculture in the 21st Century. *Journal of Agricultural Engineering Research*, vol. 76 (3), pp. 267–275. DOI: https://doi.org/https://doi.org/10.1006/jaer.2000.0577

Steinwart, I. & Thomann, P. (2017a). {liquidSVM}: A Fast and Versatile SVM package. *ArXiv e-prints 1702.06899*,. Available at: http://www.isa.uni-stuttgart.de/software

Steinwart, I. & Thomann, P. (2017b). *liquidSVM Vignette*. Available at: https://cran.r-project.org/web/packages/liquidSVM/vignettes/demo.html [2020-09-30]

Trucano, T.G., Swiler, L.P., Igusa, T., Oberkampf, W.L. & Pilch, M. (2006). Calibration, validation, and sensitivity analysis: What's what. *Reliability Engineering & System Safety*, vol. 91 (10), pp. 1331–1357. DOI: https://doi.org/https://doi.org/10.1016/j.ress.2005.11.031

Vapnik, V. (1982). *Estimation of Dependences Based on Empirical Data: Springer Series in Statistics (Springer Series in Statistics)*. Berlin, Heidelberg: Springer-Verlag.

Weiss, M., Jacob, F. & Duveiller, G. (2020). Remote sensing for agricultural applications: A meta-review. *Remote Sensing of Environment*, vol. 236, p.

111402. DOI: https://doi.org/https://doi.org/10.1016/j.rse.2019.111402
Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. Available at: https://ggplot2.tidyverse.org
Wold, S., Sjöström, M. & Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, vol. 58 (2), pp. 109–130. DOI: https://doi.org/https://doi.org/10.1016/S0169-7439(01)00155-1
Yagoubi, R.S., Naud, O., Dejean, K.G. & Crestani, D. (2018). New Approach for Differential Harvest Problem: The model checking way. *IFAC-PapersOnLine*, vol. 51 (7), pp. 57–63. DOI: https://doi.org/https://doi.org/10.1016/j.ifacol.2018.06.279
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. *Remote Sensing*, vol. 7 (11), pp. 14939–14966 Multidisciplinary Digital Publishing Institute.
Zhou, Z., Morel, J., Parsons, D., Kucheryavskiy, S. V & Gustavsson, A.-M. (2019a). Estimation of yield and quality of legume and grass mixtures using partial least squares and support vector machine analysis of spectral data. *Computers and Electronics in Agriculture*, vol. 162, pp. 246–253. DOI: https://doi.org/https://doi.org/10.1016/j.compag.2019.03.038
Zhou, Z., Palmborg, C., Ericson, L., Dryler, K., Lindgren, K., Bergkvist, G. & Parsons, D. (2019b). A 60-years old field experiment demonstrates the benefit of leys in the crop rotation. *Acta Agriculturae Scandinavica, Section B — Soil & Plant Science*, vol. 69 (1), pp. 36–42 Taylor & Francis. DOI: https://doi.org/10.1080/09064710.2018.1492010

Acknowledgements

A big thank you to my supervisor Julien Morel and assistant supervisor David Parsons for introducing me to the possibilities remote sensing offers and for putting in a lot of effort to guide me through this project. I also want to thank Sanna Bergqvist and Johan S. for helping me out with the field work. And I could not have done this project without Harriet and Kjell Öquist who watched my dog when I needed to concentrate on sampling in the field. Last but not least, I want to thank everyone at Skogis and the Röbäcksdalen field station for making me feel welcome and helping me with all the practical administrative issues.