Spring Wheat Yield prediction based on UAV Imagery

Data Science Report, NMBU, Spring 2021

Muhammad Fahad Ijaz
muhammad.fahad.ijaz@nmbu.no

ABSTRACT

This report is a part of the vPheno (virtual phenotyping) project at NMBU, started 2017. The aim of the project is to employ the image analysis and machine learning techniques to reduce the time to develop more robust cultivars, which is anywhere around 7 to 15 years by current practices. Using the data that has been collected as part of this project, I tried to develop machine learning models to predict the spring wheat grain yield, using Unmanned Aerial Vehicles (UAVs) images. The extracted features from the images were employed as input to the machine learning algorithms for yield prediction. The data from different dates was integrated using Simpson's integration. The best model parameters were estimated using a brute force search. Random Forest, Gradient Boosting and CatBoost regression models showed a good capability for grain yield prediction. Suggestion are included to improve the data collection methods and, consequently, the results for future studies.

1 INTRODUCTION

Wheat is the most important food source all around the world [Igrejas and Branlard 2020]. The global demand for wheat is projected to more than double by the year 2050 [Tilman et al. 2011]. This stresses the need for the increase in food crops' yield. Compared to many other countries, cultivable land is limited in Norway and it costs a lot to cultivate barren land [Ministry of Agriculture and Food 2014], the better approach would be to focus on increasing the yield from the existing cultivable land. Moreover, cultivating more land would naturally cost more. So, the reason and motivation behind plant phenotyping is to increase the yield from the same fields.

In current state of knowledge it takes something from the realm of 7-15 years to develop a new variety from scratch. The current breeding techniques are mass selection, pure-line breeding, pedigree breeding, single seed descend (SSD), double haploids (DH) to name but a few. These techniques can be used to improve certain characteristics of the plants. Some focus on the traits like nutrition contents, while other may be oriented towards improving the yield, crop maturity time, to name a few. This is a very time-consuming process and involves a substantial cost as well.

Remote sensing systems can be deployed to get information in real time and the data collected can be used for yield prediction. One such technique is to deploy Unmanned Aerial Vehicles (UAVs) to gather data from the field. That data is used to predict different characteristics of the plants. Those traits include grain yield as well. Remote sensing does not reduce the time. It is in the experimental phase with advances happening often. Remote sensing mostly aims to improve selection accuracy and the resource pool available. Selection accuracy is improved with more precise phenotyping. It can also increase resource pool available. With bigger resource pool and higher accuracy, annual genetic gains are increased and therefore, in theory, the number of cycles to get a new variety will be fewer. In general, adding UAV multispectral phenotypes to genomic prediction models usually enhances their prediction ability. My goal is to develop a model which can predict the crop yield with greater accuracy.

Remote sensing enables a non-intrusive, high-throughput monitoring of plant physiological characteristics[Tattaris et al. 2016]. Remotely sensed spectral measurements are a result of the interaction between incoming radiation and target objects. It creates a characteristic signature of reflected light. Spectral indices are a function of the light absorption properties of the plant and are calculated from these signatures [Tattaris et al. 2016].

This report contributes to the research project vPheno (virtual phenomics) at the Faculty of Bioscience at NMBU. The project, started in 2017, aims to help plant breeders to reduce the time required to select better, high yield varieties of crops and speedup the process using image analysis techniques. Earlier results related to this work have been reported by Burud et al. [2017], Grindbakken [2018], Lied [2019] and [Shafiee et al. 2021]. This report explores a different way of data integration, using Simpson's integration to aggregate the data collected at different dates.

The objective of this report is to generate a machine learning model that can predict wheat grain yield, based on the image data collected from the field at different times during the season.

The first section deals with the theory of the process and spectral indices. It follows the process of data collection and pre-processing, feature extraction from images and the explanation of machine learning methods. In the results section the prediction results are outlined which are discussed in the discussion section where a comparison is made to one of the previous studies by Lied [2019].

2 THEORY

2.1 Plant Phenotyping

Plant phenotyping is the study of how plants traits, termed as phenome, develop from their interaction with the environment and how their genome can be related to those traits [Minervini et al. 2014]. These traits include, among others, the plant yield, maturity time, plant height, etc. The importance of phenotyping has increased significantly with the increased demand for improving the yield. Plant phenotyping investigates how the physical features are affected by the genome of the plant. Efficient and sustainable agriculture practices stress the need and importance of such methods [Minervini and Tsaftaris 2013].

Image-based phenotyping methods are fast and nondestructive. They can predict the results and save time to make decision about the choice of the variety of wheat to be chosen for the next phase.

2.2 Spectral Indices

Spectral indices are derived values from one or more reflection bands. They are used to indicate the presence of the required features in an image and are derived by combining the spectral reflectance from multiple wavelengths of light reflected from a surface [L3Harris Geospatial 2021a]. There are several different spectral indices. The ones related to plants are called vegetation indices [Huete et al. 2002]. Vegetation indices are designed to highlight specific properties of vegetation [L3Harris Geospatial 2021b].For this study, we are focusing on the ones that can be derived from the spectral bands data we have gathered for the project. These indices have shown good results in predicting the grain yield [Zhang and Liu 2014]. The five bands, collected by the by the Micasense RedEdge-M multi-spectral camera, are listed in the Table 1.

Table 1: Table with details of bands data collected by the Micasense RedEdge-M [Lied 2019; MicaSense 2019]

Band Name	Center wavelength (nm)	Bandwidth (nm)
Blue	475	20
Green	560	20
Red	668	10
RedEdge	717	10
NIR	840	40

The relevant vegetation indices and the formula to calculate them are as follows.

2.2.1 Normalized difference vegetation index - NDVI. NDVI is used for analyzing vegetation. This was calculated by measuring the difference in reflection between near–infrared (NIR) and red light (RED). Chlorophyll is an indication of healthy vegetation. It absorbs more red light while reflects NIR and green light. NDVI is calculated using the equation 1 [Rouse et al. 1974].

$$NDVI = \frac{NIR - RED}{NIR + RED} \tag{1}$$

NDVI Values range from negative one (-1) to positive one (+1). Negative values indicate water bodies and are not related to vegetation. Values approaching zero indicate no green vegetation whereas the values close to one indicate the presence of green vegetation [The Earth Observatory 2000]. NDVI is an indicator of plant health. A value between 0 and 0.33 a stressed plant, 0.33 to 0.66 is moderately healthy, and 0.66 to 1 is very healthy [Taipale 2020]. These numbers vary depending on the plant type and other conditions. But they can be used as a rule of thumb.

NDVI suffers from saturation problem in well vegetated areas. This is mainly because of the red band whose energy is strongly absorbed by pigments. If there is a certain amount of pigment in the leaf, the reflectance in the red band will will almost remain unchanged even if more pigment is present[Wang et al. 2003].

2.2.2 MERIS terrestrial chlorophyll index - MTCI. This index is a measure of the chlorophyll content. Spectral data indicates that important information about the leaf area index can be found in the RedEdge band. Sharp changes in the reflectance are observed between wavelengths 690 and 750 nm. So a positive change in this region is indicative of chlorophyll content in the leaf[Delegido et al. 2013]. MTCI integrates the red-edge band (REG) into the equation for that reason, in addition to Red and NIR, which are included in NDVI as well. MTCI is calculated by the equation 2 [Dash and Curran 2004]:

$$MTCI = \frac{NIR - RedEdge}{RedEdge - Red} \tag{2}$$

In comparison with NDVI, MTCI is better at yield prediction [Zhang and Liu 2014].

2.2.3 Enhanced vegetation index - EVI. EVI is calculated like NDVI. However, it corrects for some distortions in the reflected light caused by particles in the air and the ground cover below the vegetation. The EVI value does not become saturated as quickly as the NDVI when viewing areas with a high amount of chlorophyll [Lein 2012]. EVI is calculated using the equation 3 [Agapiou et al. 2012; Huete et al. 1994; Lein 2012].

$$EVI = 2.5 \times \frac{NIR - RED}{NIR + (6 \times RED) - (7.5 \times BLUE) + 1} \tag{3}$$

3 METHOD

The data is collected from one field where different varieties of spring wheat has been planted in several sub-plots. Data collection is done by using a UAV equipped with an RGB Camera, in addition to a Micasense RedEdge-M multi-spectral camera, a GPS Module and an ambient light sensor. The UAV is flown over the field on a predefined grid pattern path and it captures the image data of the field. The image data is then stitched to get an overall image of the field. The data collected by the system consists of blue, green, red, red edge and near-infrared bands. The median band values for each band are taken for each plot and are used to calculate the vegetation indices, i.e. Normalized Difference Vegetation Index (NDVI), MERIS Terrestrial Chlorophyll Index (MTCI), and Enhanced Vegetation Index (EVI).

3.1 Test Site

The test site is located at Vollebekk Research Farm, close to NMBU in Ås, in south-eastern Norway. The coordinates of the site are 59°39'23.6"N 10°45'13.8"E. One spring wheat field with 96 sub-plots is included in this study. The cultivars were sown on 20th April, 2020. Twenty-four different candidate cultivars are cultivated in the sub-plots with each cultivar planted in four sub-plots. The placement of the cultivars is random with randomization being done in an alpha-lattice design [Kumar et al. 2020].

Each subplot is separated from the neighboring subplots by a border row. This is done to reduce the border effect [May and Morrison 1986; Wang et al. 2017]. As a standard agronomic practice, herbicides and fungicides are used on the crops to ensure better yield. In addition to the image data, manual measurements are also done to determine the number of days to maturity and number of days to heading of the plant. Heading is a growth stage of the wheat plant when the wheat head/fruit has fully emerged[White and Edwards 2008]. Maturity comes later and is defined as the stage where the complete plant material has turned golden and is ready for harvest[Hanft and Wych 1982]. Plant height is also measured

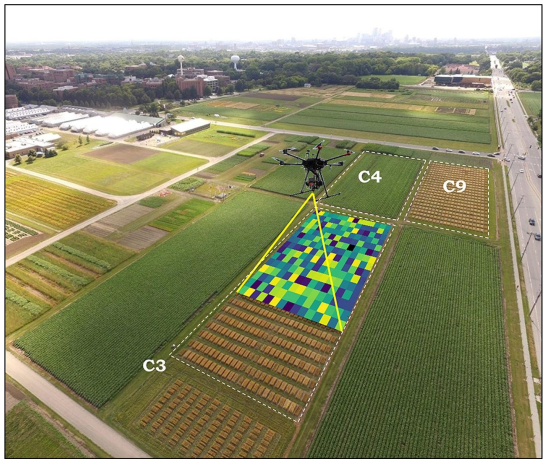

Figure 1: A illustration of the data collection process[Moghimi et al. 2020]. This illustration is to explain the data collection process and is not of the field and UAV used for this study.

once before harvest, when the plants are at their maximum growth height. Grain yield is measured in g/m^2.

3.2 Image Data Collection

The image data is collected by a UAV which had been mounted with necessary equipment to gather the images and multi-spectral data. The UAV used is DJI Matrice 100 [DJI 2019]. It comes with a GPS Module. Zenmuse X3 camera module from DJI is installed on the UAV which is used to gather the RGB image data. The RGB Camera has a 1/2.3" CMOS sensor with 12.4 effective megapixels [DJI 2017]. In addition to that, a RedEdge Downwelling Light Sensor (DLS) [MicaSense 2017] from Micasense and a RedEdge-M multi-spectral by Micasense [MicaSense 2019] are used as well [Lied 2019]. The image resolution of the RGB camera is 3000×4000 pixels.

The lighting conditions vary from season to season, depending on weather. This affects the consistency of the collected data. This effect needs to be compensated to acquire reliable data in the long run. The ambient light sensors are used to compensate for this variation [MicaSense 2021].

3.3 Image Processing

Pix4D and QGIS are the software's used to process the image data and convert it into tabular form to be used for machine learning. After the images are collected using the equipment described, the next step is to stitch them together to form one big reflectance map of the field. This is done in Pix4D. This is detailed in Lied [2019] and Grindbakken [2018].

The QGIS application is used to extract the required data from the reflectance map generated from the data collected from the cameras on the UAV. Geographical Information Systems are commonly referred to as GIS. QGIS is an open-source GIS application. Geo-processing, sampling and vector analysis are some of the applications of QGIS. For processing the reflectance map, a custom grid is generated, using QGIS, matching the setup of subplots in the field [Lied 2019]. It extracts the required values of the bands for each subplot from the grid. It was used to plot and index the median value of pixels in each sub-plot. The data is exported to a csv file.

3.4 Detail of available data

Table 2 is representing the dataset for the current study. The data collected in dates between heading and maturity is of most importance since it is more relevant and practical to use that data for yield predictions [Shafiee et al. 2021].

Table 2: Remote sensing data Robot Field

Date	No. of Sub-plots	Comments
18/06/2020	96	Complete
23/06/2020	96	Complete
24/06/2020	96	Complete
25/06/2020	96	Complete
29/06/2020	96	Complete
01/07/2020	96	Complete
07/07/2020	96	Complete
13/07/2020	96	Complete
16/07/2020	96	NIR data missing
20/07/2020	96	Complete
22/07/2020	96	Complete
27/07/2020	96	Complete
30/07/2020	96	Complete
04/08/2020	96	After maturity
07/08/2020	96	After maturity
12/08/2020	96	After maturity
14/08/2020	96	After maturity

NIR data was missing in one of the dates which was dropped. Twelve remaining datasets marked 'Complete' in Table 2 were used for further processing.

3.5 Data Features

The features in the available datasets are listed in Table 3. Eight features have been used as input features while grain yield is the target variable. There are other features also available from the data. For example, maturity, plant height, protein content, etc. but these can only be determined at a later stage of growth. Since, in a real-world scenario, this data will not be available to the model while making predictions for new data, these were not included in model training. Heading is a feature that is available before harvest. It was not used since it is practically the same for all sub-plots in this field. It can be useful to include this when using the data from multiple fields. Since in this case, we have the date from only one field, this was not used for model training and evaluation.

Table 3: Details of Features from the dataset

Name	Explanation	Selection	Type
Input Features			
Blue	Blue band	Median	Band Data
Green	Green Band	Median	Band Data
Red	Red Band	Median	Band Data
Red Edge	Red-NIR transition zone	Median	Band Data
NIR	Near Infrared	Median	Band Data
NDVI	Normalized Difference Vegetation Index	Median	Vegetation index
MTCI	The MERIS terrestrial chlorophyll index	Median	Vegetation index
EVI	Enhanced vegetation index	Median	Vegetation index
Target Variable			
Grain Yield	Grain Yield in g/m^2	-	Target Variable
Features applicable post-harvest (except heading)			
Heading	Days to heading from sowing	-	Days
Maturity	Days to maturity from sowing	-	Days
Plant height	Plant height at maturity in cm	-	cm
Protein content	Protein content in %	-	%
Spikes	Number of Spikes per m^2	-	Number

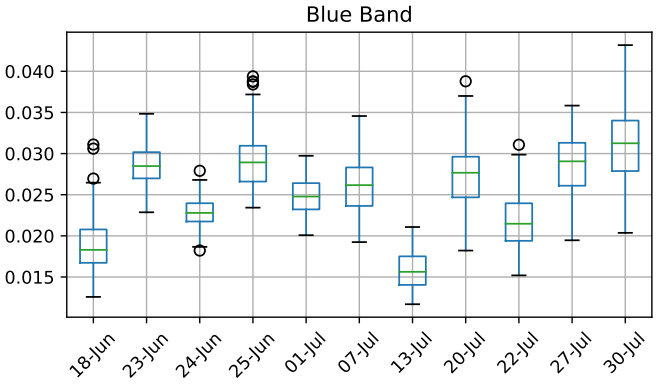

Figure 2: Boxplots of Blue Band data from all datasets

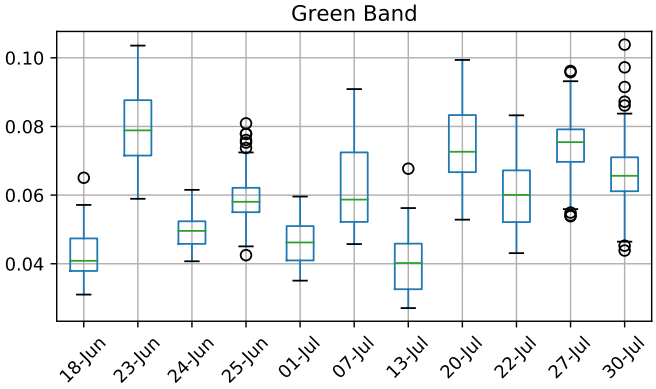

Figure 3: Boxplots of Green Band data from all datasets

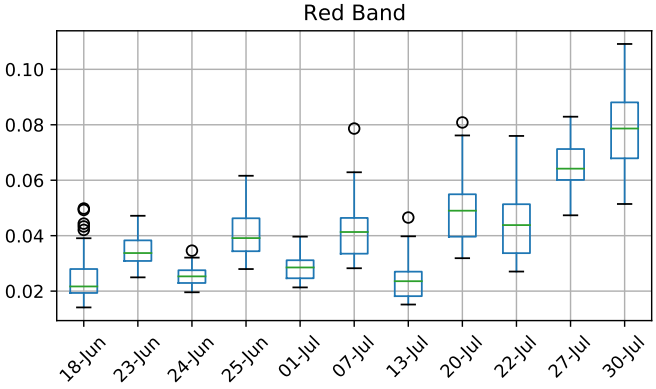

Figure 4: Boxplots of Red Band Data from all datasets

3.6 Data pre-processing

Each sub-plot has a single median value for each band for which the data has been collected in the images. This is generated from the pixels of each sub-plot. The reason for using the median values for all bands is because median values are not affected much by outliers' presences in the data. Median value can compensate for the presence of soil pixels in the image, whose band values would be on a different than of the pixels for the vegetation.

3.6.1 Irregularities in the data. All the bands data from each date was plotted as box plots. Examining the plots, it is clear that the data from 29-Jun does not follow the trends in all the bands. There might be some problems with data collection process. The data from this date has been dropped from further investigation. Remaining eleven datasets have been used further.

3.6.2 Data Compilation Strategy. The previous studies on similar data at NMBU [Burud et al. 2017], [Grindbakken 2018] and [Lied 2019], [Shafiee et al. 2021] have taken a different approach to data compilation for machine learning. They compiled the data sets by adding the data collected on different dates as new features in the dataset. This resulted in an enormous increase in the number of features in the dataset, but the number of samples remained the same. The reasoning for this strategy was that since for one plot, although the data is being collected several times in one season, but at the end of the season, one sub-plot has only one grain yield value. So, it is better to add new features to the same samples than to add

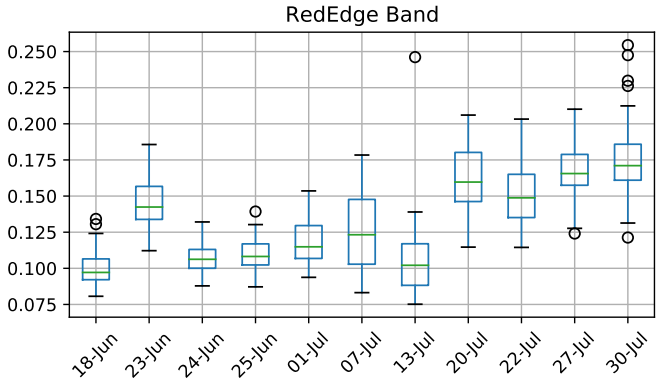

Figure 5: Boxplots of RedEdge Band data from all datasets

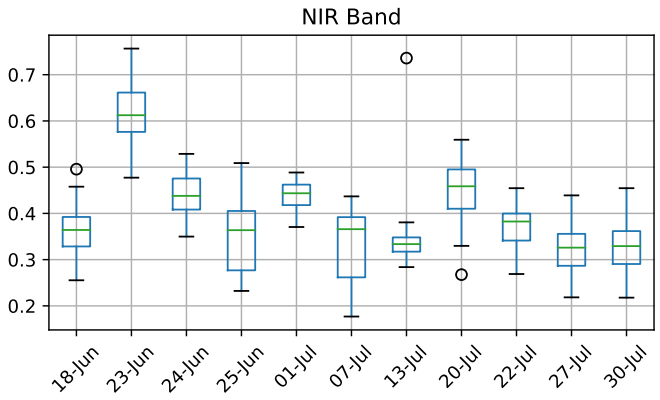

Figure 6: Boxplots of NIR Band data from all datasets

them as new samples as all the feature of one sub-plot correspond to the same grain yield.

I took a different approach by integrating the data for each sub-plot by using Simpson's Integration[Weisstein 2005]. It uses the number of days from sowing for each date and integrates all the values for one band into a single value by calculating the area under the band value and number of days curve. It integrated the 11 values for each band for one plot into one integrated value. So, for the 96 sub-plots, there are eight input features which are the integration of the same features of the data from ten different dates.

In case of multiple fields, the data collected is not usually equally spaced in time. There are several practical reason for not being able to collect data as planned. Weather plays an important role. Data is usually not collected on a cloudy day or during bad weather or rain. The real world data ends up not equally spaced thus not comparable from one season to the other. By integrating the data, we can get a comparable approximation of the data if there are same number of days between the first and last date of data collection.

Figure 7 shows the correlation between all the input features and with the target variable,i.e. grain yield.

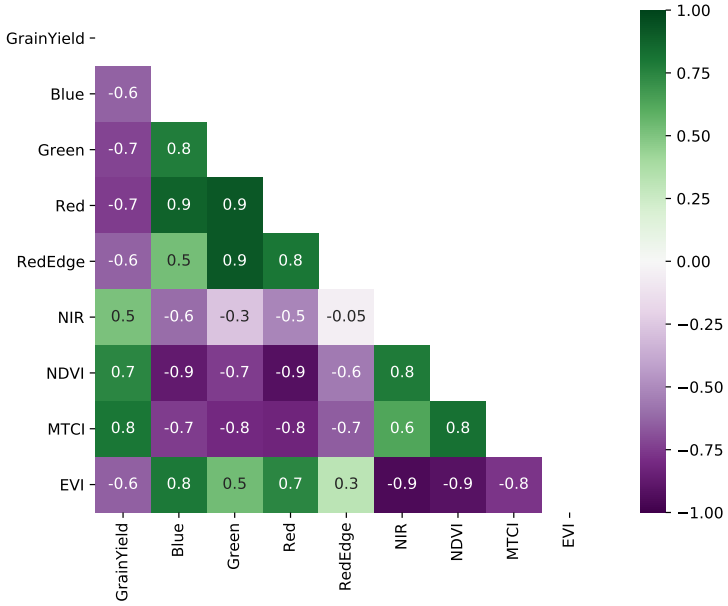

Figure 7: Pearson Correlation Heat map indicating the correlation between the input and target variables

3.7 Machine Learning

The machine learning analysis has been programmed in Python programming language [Van Rossum and Drake 2009]. The machine learning algorithms are from Scikit-Learn (0.23.2) [Pedregosa et al. 2011] and CatBoost (version 0.24.3) [Dorogush et al. 2018] libraries. The analysis and predictions have been executed in a Jupyter Notebook [Kluyver et al. 2016] with Python version 3.8.5.

Table 4: Versions of files used for this report; Gitlab repository https://gitlab.com/fahadijaz/ds-report-spring-2021

File	Git hash
1. Data Cleaning.ipynb	4d06c51a
2. ML-Simps Integration.ipynb	b50c6a21

3.7.1 Git hashes.

3.7.2 Sampling and data partitioning. The dataset divided into trai (70%) and test (30%) sets. The test_train_split function from Scikit-Learn is used for making the split. The data is shuffled during the split. The default values of the hyperparameters are used.

Random sampling was used as a means to validate the model results. It splits the data into random splits and those splits are used to train and test the model. This was done using random seeds in the range from 0 to 9. The average of the ten results is used as measure of the prediction accuracy of the model.

3.7.3 Standardization. Standardization has been used for the Lasso model as it resulted in an improvement in the prediction accuracy. It has been done using the StandardScaler from Scikit Learn. StandardScaler scales the features to unit variance. This improves the

predictions results for most of the machine learning estimators [Raschka 2015, Ch. 2]. Default values have been used for the hyperparameters for StandardScaler.

3.7.4 Machine Learning Models. This problem requires the machine learning models to predict a continuous target. This makes it a regression problem. So, regression estimators have been used for that purpose. From Scikit learn, Lasso, Random Forest Regressor and Gradient Boosting Regressor are used. In addition, CatBoostRegressor is also used.

Random forest is an ensemble model. It combined multiple decision trees and aggregate their results using a statistical technique called bagging, to formulate the results. This technique results in a robust model that is less prone to over fitting [?]. Random Forest comes up with a robust, accurate model that can handle large varieties of input data with binary, categorical, continuous features.

Gradient Boosting Machines (GBMs) is explained by two concepts: Decision trees and boosting. Decision trees are machine learning models which use specific connections between variables to classify the data points correctly. Boosting is a technique that creates several decision trees based on the results of the previous trees. For every tree made, a greater weight is put on the changing/altering of the weights that are hard to classify. Gradient Boosting Machines uses a cost function to determine the properties of each subsequent tree [Singh 2018].

CatBoost uses a Gradient Boosting algorithm for prediction of the classifier. It uses a framework for gradient boosting based on decision trees[CatBoost 2019].

3.7.5 Model Evaluation. For model evaluation, Coefficient of Determination (R^2), Mean Absolute Error and Root Mean Squared Error(RMSE) and were used.

The coefficient of determination is a measure of regression score where the best possible score is 1.0. The values lower than that indicate worse prediction results.

RMSE determines the regression loss. It is calculated adding the squared difference between the prediction and the target variable and dividing the sum by the number of samples. The square root of the resulting value is the RMSE.

MAE is a measure of the average magnitude of error in the predictions. It is calculated adding the absolute difference between the prediction and the target variable and dividing the sum by the number of samples.

3.7.6 Hyperparameters Tuning. The hyperparameters of the models were tuned one by one using Brute force search method. One hyperparameter was tuned based on the *r2_score* metric from sklearn. The search was carried out step by step, tuning one parameter at a time. The best hyperparameters of each of models are given in Table 5, Table 6, Table 7 and Table 8.

Table 5: Hyperparameters of Lasso Model

Hyperparameter	Value
`alpha`	`4.5`

In terms of hyperparameter tuning, most of the hyperparameter performed the best with their default values. Tuning them did not result in any improvement in their accuracy. Those of them which resulted in better predictions have been used and reported in this report.

Table 6: Hyperparameters of RandomForestRegressor Model

Hyperparameter	Value
`max_depth`	`250`
`min_samples_split`	`14`
`min_samples_leaf`	`3`
`random_state`	`1`

Table 7: Hyperparameters of GradientBoostingRegressor Model

Hyperparameter	Value
`subsample`	`0.8`
`learning_rate`	`0.4`
`random_state`	`500`
`random_state`	`1`

Table 8: Hyperparameters of CatBoostRegressor Model

Hyperparameter	Value
`depth`	`8`

4 RESULTS

This report investigates the potential of using Simpson's integration technique for aggregating the data from several dates to accurately predict the grain yield of spring wheat. In addition, this report also evaluates the ability of the vegetation indices to make yield prediction.

4.1 Evaluation Results

The results from the estimators are listed in 9. Standardized data, as explained in section 3.7.3, was only used for Lasso Model. Tree based models used are not affected by any monotonic transformation of the features. So, non-transformed data was used for the remaining models.

Random Forest Regressor performed the best in terms of the three matrices in the Table 9. The scores of Gradient Boosting Regressor and CatBoost Regressor are close, with CatBoost Regressor trailing just behind the Gradient Boosting Regressor. Lasso performed the worst.

Table 9: Prediction results

Model	**Standardization**	R^2	**MAE**	**RMSE**
Lasso	Yes	0.617	46.29	58.11
RandomForest	No	0.736	38.21	48.21
GradientBoosting	No	0.714	38.94	49.96
CatBoost	No	0.713	39.32	50.28

4.2 Feature Importance

The feature importance from the regression models used is plotted in Fig. 8, Fig. 9 and Fig. 10. Lasso model from Scikit learn does not have a method to return the feature importance of the input features.

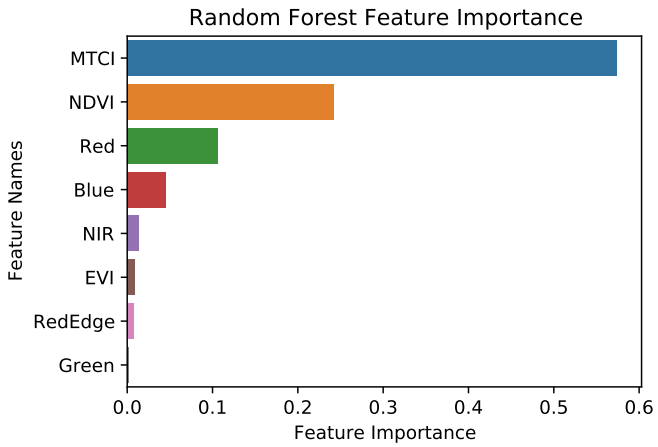

Figure 8: Feature Importance of Simpson's integrated vegetation indices calculated by the Random Forest model

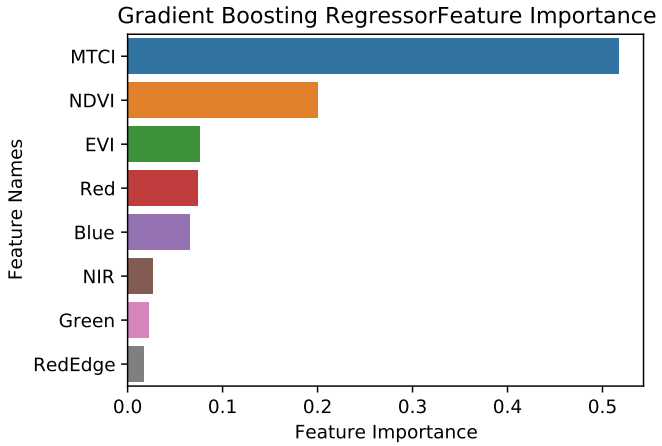

Figure 9: Feature Importance of Simpson's integrated vegetation indices calculated by the Random Forest model

5 DISCUSSION

5.1 Assessment of model performance

From the results, it is clear that Random Forest Regressor performed the best but the difference in the results is not huge. We can say that the Random Forest, Gradient Boosting and CatBoost performed well for this data set. It also indicates that it is better to use accumulative data using Simpson's integration. The results could be improved with the inclusion of more data. More data can help in improving

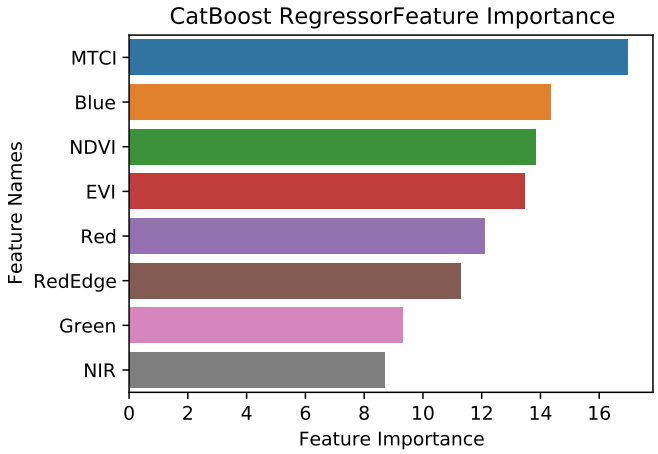

Figure 10: Feature Importance of Simpson's integrated vegetation indices calculated by the Random Forest model

the results, in addition to defining and collecting more features related to the crop, and the conditions surrounding it.

5.2 Feature Importance

Among the three best performing models, MTCI is a feature that has contributed the most towards the model accuracy. NDVI also is among the top three important features. EVI has a varying degree of importance in predicting the results. It is important to note that the vegetation indices MTCI and NDVI, which are derived from the bands data, perform better than the original bands themselves.

5.3 Comparison with results from Lied [2019]

The results from a previous study on the similar data [Lied 2019] , had varying degree of results accuracy with better results on one data set while worse results than in Table 10 for two other datasets.

Table 10: Table 4.5 from [Lied 2019], showing the results without maturity as a feature

	Field A-18	Field B-18	Field C-18
R^2	0.915	0.6	0.7
MAE	27.2	46.6	18.9

There are few inconsistencies in the method used there. Cross validation was not used and the reported results are the results for one random split of data. That can a lucky good model for A-18 field while the splits for B-18 and C-18 did not result in a good prediction accuracy.

Moreover, the data aggregation strategy, described in 3.6.2, has some limitations. Since each field different number of features, according to the previous data aggregation strategy, the data could not be used together to train one model. Instead, separate models had to be trained and they could only work for the field and dataset they were trained on since the data from other fields would have more or less features than the others. So, even if the prediction

results are good for one model, they are not necessarily usable on the data from other fields.

I have tried to overcome those limitation and have tested the potential of Simpson's integration technique to make a robust and universal model that can be used for other fields, regardless of the dates when the data was collected.

5.4 Suggestions & Conclusion

Comparing the results to Lied [2019], the prediction score in this report are promising and indicate that Simpson's integration technique is promising and can be used to train a generalised model. The prediction score is expected to improve with the availability of more data for training.

There is a need to collect more data and especially the variables that affect the grain yield. Weather and soil play an important role in crop growth. The amount of rain during the season is an important feature. The number of days of complete sunshine, or even number of hours of sunshine could be measured. Fertilizers and pesticides also impact the plant.

Moreover, the data collection strategies can be improved. Better equipment, high resolution cameras and an on-ground robot/platform on wheels instead of a UAV can collect better pictures of the field. The pictures from drone can be blurry since it is moving while it is taking pictures. Or it can also be tried to program the path of the drone such that it stops to take the picture at a point before moving a few meters and take the next picture. These techniques can improve the quality and reliability of the image data captured for this purpose.

This approach will not accelerate breeding research in short term. However, the additional dimensions we can capture will widen the spectrum of plant research, including breeding research. With time and understanding what do the vegetation indices point to, they may become new breeding goals. Multi spectral imaging can play a big role in beating the challenges put forth by the changing climate in future by capturing additional aspects of plant physiology. The fact that high-throughput phenotyping allows us to screen large collections in a fraction of normal time will also contribute to plant research advances.

Summing up, image-based plant phenotyping techniques have a great potential for yield prediction. The predictions can be improved further to reach a realistically usable level by incorporating more variables and collecting more data over a longer period of time.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to Dr. Sahameh Shafiee , Postdoctoral Fellow at Faculty of Biosciences, NMBU, for providing me with the background knowledge, relevant data, and research material to form the basis of my study. I am also grateful to Professor Ingunn Burud, at the Faculty of Science and Technology, NMBU, for allowing me to work on this project and supervising it. I am grateful to Professor Hans Ekkehard Plesser for his multiple reviews and valuable remarks. I would also like to thank my wife who supported in writing this report

REFERENCES

Athos Agapiou, Diofantos G. Hadjimitsis, and Dimitrios D. Alexakis. 2012. Evaluation of Broadband and Narrowband Vegetation Indices for the Identification of Archaeological Crop Marks. *Remote Sensing* 4, 12 (2012), 3892–3919. https://doi.org/10.3390/rs4123892

Ingunn Burud, Gunnar Lange, Morten Lillemo, Eivind Bleken, Lars Grimstad, and Pål Johan From. 2017. Exploring Robots and UAVs as Phenotyping Tools in Plant Breeding. *IFAC-PapersOnLine* 50, 1 (2017), 11479–11484. https://doi.org/10.1016/j.ifacol.2017.08.1591 20th IFAC World Congress.

CatBoost. 2019. CatBoost - open-source gradient boosting library. https://catboost.ai/ (Accessed on 05/12/2021).

Jadunandan Dash and PJ Curran. 2004. The MERIS terrestrial chlorophyll index. (2004).

J. Delegido, J. Verrelst, C.M. Meza, J.P. Rivera, L. Alonso, and J. Moreno. 2013. A red-edge spectral index for remote sensing estimation of green LAI over agroecosystems. *European Journal of Agronomy* 46 (2013), 42–52. https://doi.org/10.1016/j.eja.2012.12.001

DJI. 2017. Zenmuse X3 - Product Information - DJI. Retrieved 2021-05-13 from https://www.dji.com/no/zenmuse-x3/info

DJI. 2019. Matrice 100 - DJI. Retrieved 2021-05-13 from https://www.dji.com/no/matrice100

Anna Dorogush, Vasily Ershov, and Andrey Gulin. 2018. CatBoost: gradient boosting with categorical features support. (10 2018). arXiv:https://arxiv.org/abs/1810.11363

Ole Kristian Grindbakken. 2018. *Phenotyping studies of spring wheat by multispectral image analysis.* Master's thesis. Ås, Norway. http://hdl.handle.net/11250/2561258

J. M. Hanft and R. D. Wych. 1982. Visual Indicators of Physiological Maturity of Hard Red Spring Wheat 1. *Crop Science* 22, 3 (05 1982), 584–588. https://doi.org/10.2135/cropsci1982.0011183x00220003036x

A. Huete, K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira. 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment* 83, 1-2 (11 2002), 195–213. https://doi.org/10.1016/S0034-4257(02)00096-2

A Huete, C Justice, and H Liu. 1994. Development of vegetation and soil indices for MODIS-EOS. *Remote Sensing of Environment* 49, 3 (1994), 224–234. https://doi.org/10.1016/0034-4257(94)90018-3

Gilberto Igrejas and Gérard Branlard. 2020. *The Importance of Wheat.* Springer International Publishing, Cham, Switzerland, 1–7. https://doi.org/10.1007/978-3-030-34163-3_1

Thomas Kluyver, Benjamin Ragan-Kelley, Fernando Pérez, Brian Granger, Matthias Bussonnier, Jonathan Frederic, Kyle Kelley, Jessica Hamrick, Jason Grout, Sylvain Corlay, Paul Ivanov, Damián Avila, Safia Abdalla, and Carol Willing. 2016. Jupyter Notebooks – a publishing format for reproducible computational workflows. In *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, F. Loizides and B. Schmidt (Eds.). IOS Press, 87 – 90.

A. Kumar, B. Bharti, J. Kumar, D. Bhatia, G. P. Singh, J. P. Jaiswal, and R. Prasad. 2020. Improving the efficiency of wheat breeding experiments using alpha lattice design over randomised complete block design. *Cereal Research Communications* 48, 1 (01 03 2020), 95–101. https://doi.org/10.1007/s42976-020-00014-3

L3Harris Geospatial. 2021a. Spectral Indices. Retrieved 2021-04-15 from https://www.l3harrisgeospatial.com/docs/spectralindices.html

L3Harris Geospatial. 2021b. Vegetation Indices. Retrieved 2021-04-15 from https://www.l3harrisgeospatial.com/docs/vegetationindices.html

James K. Lein. 2012. *Environmental Sensing: Analytical Techniques for Earth Observation, first edition.* Springer New York. https://doi.org/10.1007/978-1-4614-0143-8

Lars Martin Bøe Lied. 2019. *Multispectral image analysis of spring wheat using UAV and machine learning.* Master's thesis. Ås, Norway. http://hdl.handle.net/11250/2625385

K. May and R. Morrison. 1986. EFFECT OF DIFFERENT PLOT BORDERS ON GRAIN YIELDS IN BARLEY AND WHEAT. *Canadian Journal of Plant Science* 66 (1986), 45–51.

MicaSense. 2017. Light Sensors: the Basics (DLS and Sunshine Sensor) – MicaSense Knowledge Base. Retrieved 2021-03-12 from https://support.micasense.com/hc/en-us/articles/115002782008-Light-Sensors-the-Basics-DLS-and-Sunshine-Sensor-

MicaSense. 2019. RedEdge-M User Manual (PDF) – MicaSense Knowledge Base. Retrieved 2021-03-12 from https://support.micasense.com/hc/en-us/articles/115003537673-RedEdge-M-User-Manual-PDF-

MicaSense. 2021. Best practices: Collecting Data with MicaSense Sensors – MicaSense Knowledge Base. Retrieved 2021-03-12 from https://support.micasense.com/hc/en-us/articles/224893167

Massimo Minervini, Mohammed M. Abdelsamea, and Sotirios A. Tsaftaris. 2014. Image-based plant phenotyping with incremental learning and active contours. *Ecological Informatics* 23 (2014), 35–48. https://doi.org/10.1016/j.ecoinf.2013.07.004 Special Issue on Multimedia in Ecology and Environment.

Massimo Minervini and Sotirios A. Tsaftaris. 2013. Application-aware image compression for low cost and distributed plant phenotyping. In *2013 18th International Conference on Digital Signal Processing (DSP)*. 1–6. https://doi.org/10.1109/ICDSP.2013.6622670

Ministry of Agriculture and Food. 2014. Soil conservation - regjeringen.no. Retrieved 2021-05-05 from https://www.regjeringen.no/en/topics/food-fisheries-and-agriculture/landbrukseiendommer/innsikt/jordvern/soil-conservation/id2009556/

Ali Moghimi, Ce Yang, and James A. Anderson. 2020. Aerial hyperspectral imagery and deep neural networks for high-throughput yield phenotyping in wheat. *Computers*

and Electronics in Agriculture 172 (2020), 105299. https://doi.org/10.1016/j.compag.2020.105299

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12 (2011), 2825–2830.

Sebastian Raschka. 2015. *Python Machine Learning.* Packt Publishing.

Jhon W Rouse, R Hect Haas, John A Schell, Donald W Deering, and James C Harlan. 1974. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. *NASA/GSFC Type III Final Report, Greenbelt, Md* 371 (1974).

Sahameh Shafiee, Lars Martin Lied, Ingunn Burud, Jon Arne Dieseth, Muath Alsheikh, and Morten Lillemo. 2021. Sequential forward selection and support vector regression in comparison to LASSO regression for spring wheat yield prediction based on UAV imagery. *Computers and Electronics in Agriculture* 183 (2021), 106036. https://doi.org/10.1016/j.compag.2021.106036

Harshdeep Singh. 2018. Understanding Gradient Boosting Machines | by Harshdeep Singh | Towards Data Science. Retrieved 2021-05-12 from https://towardsdatascience.com/understanding-gradient-boosting-machines-9be756fe76ab

Eric Taipale. 2020. NDVI Plant Health and Your Farm | Sentera. Retrieved 2021-05-07 from https://sentera.com/understanding-ndvi-plant-health/

Maria Tattaris, Matthew P. Reynolds, and Scott C. Chapman. 2016. A Direct Comparison of Remote Sensing Approaches for High-Throughput Phenotyping in Plant Breeding. *Frontiers in Plant Science* 7 (8 2016). https://doi.org/10.3389/fpls.2016.01131

The Earth Observatory. 2000. Measuring Vegetation (NDVI & EVI). Retrieved 2021-05-07 from https://earthobservatory.nasa.gov/features/MeasuringVegetation/measuring_vegetation_2.php

David Tilman, Christian Balzer, Jason Hill, and Belinda L. Befort. 2011. Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences* 108, 50 (2011), 20260–20264. https://doi.org/10.1073/pnas.1116437108 arXiv:https://www.pnas.org/content/108/50/20260.full.pdf

Guido Van Rossum and Fred L. Drake. 2009. *Python 3 Reference Manual.* CreateSpace, Scotts Valley, CA.

Zikui Wang, Xining Zhao, Pute Wu, Ying Gao, Qian Yang, and Yuying Shen. 2017. Border row effects on light interception in wheat/maize strip intercropping systems. *Field Crops Research* 214 (2017), 1–13. https://doi.org/10.1016/j.fcr.2017.08.017

Z.-X Wang, C. Liu, and Alfredo Huete. 2003. From AVHRR-NDVI to MODIS-EVI: Advances in vegetation index research. *Acta Ecologica Sinica* 23 (01 2003), 979–987.

Eric W. Weisstein. 2005. Simpson's Rule – from Wolfram MathWorld. Retrieved 2021-05-12 from https://mathworld.wolfram.com/SimpsonsRule.html

J. White and J. Edwards. 2008. *Wheat growth and development.* NSW Department of Primary Industries, Orange, Australia.

Su Zhang and Liangyun Liu. 2014. The potential of the MERIS Terrestrial Chlorophyll Index for crop yield prediction. *Remote Sensing Letters* 5, 8 (2014), 733–742. https://doi.org/10.1080/2150704X.2014.963734