MAPPING NITROGEN STATUS IN RICE CROPS USING UNMANNED AERIAL VEHICLE (UAV) DATA, MULTIVARIATE METHODS AND MACHINE LEARNING ALGORITHMS

FANSHU MA
June, 2020

SUPERVISORS:

Dr. R. Darvishzadeh
Prof. Dr. A.D. Nelson

MAPPING NITROGEN STATUS IN RICE CROPS USING UNMANNED AERIAL VEHICLE (UAV) DATA, MULTIVARIATE METHODS AND MACHINE LEARNING ALGORITHMS
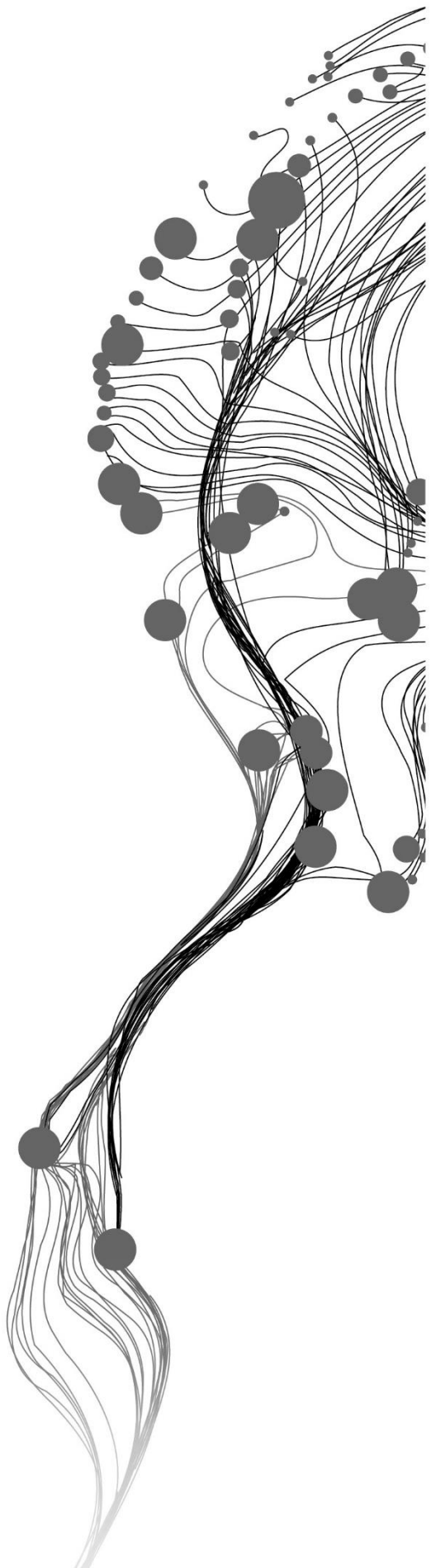
FANSHU MA

Enschede, The Netherlands, June, 2020

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Natural Resources Management

SUPERVISORS:
Dr. R. Darvishzadeh
Prof. Dr. A.D. Nelson

THESIS ASSESSMENT BOARD:
Dr. T. Wang (Chair)
Dr. Mirco Boschetti (External Examiner, Italian National Research Council)

ABSTRACT

Rice as one of the most important food crops feeds more people than any other crops in the world. Nitrogen (N) is an essential element during the rice-growing stages which affect rice yield and production. Although raised N application is used to increase the yield, in order to meet the demand for food, excess of its application would cause a series of environmental problems and even would decrease yield. Therefore, estimating nitrogen in rice is important to precision N application, environmental pollution reduction, and global carbon and N cycle. This study aims to use time-series multispectral Unmanned aerial vehicle (UAV) data and field observations of nitrogen together with multivariate methods and machine learning algorithms for estimating and mapping of nitrogen in different rice-growing seasons. The study area is in IRRI (International Rice Research Institution) experimental fields in Los Baños, the Philippines. Rice N was measured in 2016 early wet season (EWS) and 2017 dry season (DS) destructively (referred to as tissue analysis) from different parts of crops (stem, grain and whole plant parts) at the end of the growing season. SPAD and leaf colour chart readings (LCC) (as nitrogen proxies) were obtained nine times during the growing seasons whereas, other relevant measurements, such as leaf area index (LAI) was measured four times during the growing seasons. Further, SPAD and LAI values were used to calculate the canopy chlorophyll content (CCC). The relationships between SPAD, LCC and tissue analysis of the whole plant parts (referred to as plant nitrogen accumulation) (PNA) were firstly explored to understand the relationship between nitrogen measurements obtained destructively and those nitrogen proxies obtained non-destructively. In order to choose the best vegetation index (VI) for N estimation, the correlation coefficients between VIs and field measurements (PNA, SPAD and CCC) were examined. PNA was then used for further analysis, and the VI, which had the highest correlation coefficient with PNA was used in simple linear and stepwise regression models for PNA estimation. The partial-least square regression (PLSR), support vector machine (SVR) and random forest (RF) were then compared for PNA estimation using R^2, RMSE and NRMSE between measured and estimated PNA. Finally, the most accurate algorithm was used for mapping rice PNA. Results are as follows, 1) Strong correlations were observed among the PNA, SPAD and LCC in the rice panicle initiation and heading stages; 2) The GNDVI derived from the multispectral UAV images was the best performing VI for PNA estimation in both seasons; 3) comparison between simple linear and stepwise regressions revealed that using simple linear regression models (SR) and GNDVI from rice panicle initiation and heading stages are sufficient for PNA estimation; 4) among the machine learning algorithms, the RF was the most accurate machine learning algorithm for PNA estimation in 2016EWS (R^2=0.9, RMSE=8.37, NRMSE=10.9%) and 2017DS (R^2=0.93, RMSE=9.93, NRMSE=8.1%); 5) PNA estimation maps were generated for the whole study site using the RF model in both seasons. Further investigation for more accurate N status based on rice hills level and different input for machine learning algorithms could be examined in future studies.

Keywords: machine learning, regression, vegetation index, UAV, multispectral imagery, plant nitrogen accumulation, rice

ACKNOWLEDGEMENTS

<div align="right">
Fanshu Ma
Enschede, June 13th 2020
</div>

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

2016EWS	early wet season in 2016
2017DS	dry season in 2017
CCC	canopy chlorophyll content
CI	Chlorophyll Index
CM	chlorophyll meter
DAT	days after transplanting
FAO	Food and Agriculture Organization
GNDVI	Green Normalized Difference Vegetation Index
GRVI	Green Ratio Vegetation Index
IRRI	International Rice Research Institution
LAI	leaf area index
LCC	leaf colour chart
LTCCE	Long-Term Continuous Cropping Experiment
ML	machine learning
N	Nitrogen
NDRE	Normalized Difference Red-edge Index
NDVI	Normalized Difference Vegetation Index
NIR	near-infrared
NRMSE	Normalized Root Mean Square
PLSR	partial least square regression
PNA	Plant Nitrogen Accumulation
r	Pearson's Correlation Coefficient
R^2	Coefficient of determination
RE	red-edge
RF	random forest
RMSE	Root Mean Square Error
RTVI	Red-edge Triangular Vegetation Index
RVI	Ratio Vegetation Index
SPAD	Soil-plant Analyses Development
SVR	support vector regression
UAV	Unmanned Aerial Vehicle
VI	vegetation index

1. INTRODUCTION

1.1. Background

Population growth and increasing consumption of resources are increasing the global demand for food, and this trend is expected to continue (Conforti, 2011). The growing competition for water, land, and energy resources will impact current food systems (Godfray et al., 2010). Also, climate change will largely have negative effects on crop yield because of extreme abiotic factors like high and low temperatures, excess rainfall and droughts (Dabi & Khanna, 2018). Due to these reasons, food security continues to attract global attention. The World Food Summit (1996) reinforced the definition of food security: Food security exists when all people, at all times have physical and economic access to sufficient, safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life. The Food and Agriculture Organization (FAO) is a specialized agency of the United Nations that leads international efforts to defeat hunger. A recent FAO reports documents that world hunger appears to be on the rise again after a prolonged decline. According to this report, the estimated number of undernourished people has increased from 777 million in 2015 to 815 million in 2016 (FAO, 2017). Besides the demand for increased crop production, to achieve food security, people need to have access to healthy and nutritious food from different sources, so food quality is also considered a crucial aspect for the agricultural industry. Therefore, it is not hard to see that food security plays a significant role not only in crop production but also on human health.

Rice is a staple food for the world's population. It provides 21% of the global calorific needs and 15% of protein (IRRI, n.d.). In the meantime, rice also supplies minerals, vitamins, and fibre. Apart from its nutritional value, it is also a source of income for Asian rice farmers (Barker et al.,1985), and its financial benefits cannot be underestimated. Culturally, rice is the most important food grain in much of Asia. Rice continues to be the primary food (Yoshida, 1981), and it is the most widely cultivated crop in Asia. Therefore, maintaining and increasing rice production is an objective in many Asian countries. The Philippines, as a major food producer and one of the top rice importers in the world, also faces insufficient rice production problems (Manglapus, 1974). The Philippine Statistics Authority has published a report about the rice situation and outlook from 2016 to 2018 in rice production. The result shows that the country's palay (rice) production from October to December 2018, at 7.16 million metrics tons production in 2017 or by 2.2 percent. Harvest area contracted by 16 thousand hectares from the previous year's level of 1,864 thousand hectares. Yield per hectare dropped from 3.93 metric tons in 2017 to 3.87 metric tons in 2018 (Philippine Statistics Authority, 2019).

Besides environmental disturbances, nutrient management is an important part of rice crops management during the growing season. Nitrogen is the most important mineral nutrient that plants can take up from the soil in different growing stages (Zahir, 2014) as it contributes to plant biomass and yield production as well as protein. Nitrogen deficiency is a major limiting factor in the productivity of major crops (Glass, 2003). Because of the lack of nutrients in soils, especially in intensively farmed areas, nitrogen (N) fertilizer application is an effective and direct method to increase nutrient availability for the plant and thus increase biomass production and eventually yield.

Proper application of N fertilizer in the appropriate quantity and time is vital to increase crop growth and grain yields. Although N fertilizer contributes substantially to yield enhancement, its excessive use will

cause serious problems and a negative impact on both the environment and human health (Ahmed et al., 2017). For example, the excessive application of N fertilizer will increase the risk of environmental pollution, causing the leakage of N into the water and atmosphere, and it will result in water eutrophication, increased nitrate content in the subsurface water and greenhouse gas emission (Ju et al., 2009). Therefore, knowledge of crop N status is an important aspect of crop management. In the field, crop nitrogen can be measured through destructive sampling and tissue analysis of dried leaves, e.g., using Kjeldahl Digestion and Dumas Combustion (Muñoz-Huerta et al., 2013). Moreover, various diagnosis instruments have been developed for non-destructive measurements in the field, e.g., chlorophyll meter (CM) such as SPAD, lead colour chart (LCC) and Green Seeker to represent nitrogen proxies. In this regard, Bijay et al. (2002) applied the SPAD chlorophyll meter for N management on rice and wheat, the result shows that plant need-based N management through chlorophyll reduces the N requirements on rice with no yield loss, which illustrated the potential use of the SPAD chlorophyll meter on N treatment.

Field measurement of nitrogen, through destructive or non-destructive methods, is time-consuming and costly and can only be performed for a small scale at limited growing stages. Remote sensing can be used as an effective method to monitor crop parameters such as nitrogen during the entire growing season for large areas. The next section reviews existing remote sensing-based approaches for estimating crop N status, which will then lead to the problem statement and research gap for this research.

1.2. Literature Review

Different remote sensing data such as hyperspectral and multispectral images can be used to estimate nitrogen status during the rice-growing season. Hyperspectral data, providing large numbers of spectral bands, has shown great potential to estimate crop parameters such as crop nitrogen status (Tan et al., 2018). For multispectral data, a few relevant bands can be used to calculate different vegetation indices, which correlate well with N indicators (Brinkhoff et al., 2019). Although satellite remote sensing provides a high possibility for large-scale crop growth monitoring and precision management, the quality of remote sensing images from passive sensors is affected by unfavourable weather conditions such as the presence of fog and clouds. With the rise of drone technology in recent years, Unmanned Aerial Vehicle (UAV)-based remote sensing has gradually become a promising approach to overcome these problems. The high spatial resolution, relatively low operational costs and the near real-time image acquisition can overcome the limitations of ground sensing and optical satellite remote sensing. On the other hand, the acquisition of UAV imagery is limited by local weather conditions such as wind, heavy rain and changing light conditions during a flight.

To estimate crop N status using remote sensing data, various methods can be used. Vegetation indices are the most commonly used methods for N estimation. Cao et al. (2013) used the crop circle multispectral active canopy sensor to identify vegetation indices by using green, red edge, and near-infrared (NIR) bands at key growth stages to estimate N nutrition indices (NNI). Their results according to R^2 of the regression model indicated that four red edge-based indices, the red edge soil adjusted vegetation index (RESAVI), modified RESAVI (MRESAVI), red edge difference vegetation index (REDVI), and red edge re-normalized difference vegetation index (RERDVI), performed well for estimating NNI across growing stages. Zhang et al. (2006) used multispectral data to predict Nitrogen status at the canopy scale based on vegetation indices. The result indicated that canopy reflectance measurements converted to ratio vegetation index (RVI) and normalized difference vegetation index (NDVI) provided a better prediction of rice N status. The developed regression models using RVI and NDVI to predict N status proved a high with R^2 ranging from 0.82 to 0.94. In addition to rice, this method is also used on other crops. A good correlation between leaf area index (LAI) and the green normalized difference vegetation index (GNDVI)

was tested by the UAV-camera system over two variably fertilized fields on winter wheat (Zheng et al., 2018).

Although traditional vegetation indices have been used and compared for nitrogen estimation at different rice-growing stages, they have not been used for nitrogen estimation under different seasonal conditions. Since remote sensing-based approaches generally require the processing of huge amounts of data from different platforms, great attention is currently devoted to machine learning (ML) methods. Machine learning-based methods can process a large number of inputs, and also handle the non-linear problems using datasets from multiple sources (Chlingaryan et al., 2018). Based on the advantages of machine learning, it can be widely used in crop yield prediction and nitrogen state estimation (Shibayama et al., 2012).

Shao et al. (2012) compared three different methods including partial least square regression (PLSR), and least squares support vector (LS-SVM) machine for N status estimation using canopy spectral reflectance measured at visible and near-infrared regions through spectroscopy and nitrogen measurements using SPAD meter readings in rice fields. Their comparative analysis showed that the LS-SVM was superior in predicting SPAD values on rice. Apart from PLSR and SVM methods, the Random Forest (RF) algorithm also has shown the potential to accurately predict leaf N concentration using hyperspectral data (Farifteh et al., 2007). In another study, Kim et al. (2016) applied satellite remote sensing data to four ML techniques, SVM, Random Forest (RF), Extremely Randomized Trees (ERT), and Deep Learning (DL), to estimate corn yield in Iowa State. Comparisons of the validation statistics showed that DL provided more stable results by overcoming the overfitting problem of generic machine learning approaches.

The review of the above literature further reveals that no study has compared the PLSR, SVM, and RF for nitrogen estimation of rice crops in different seasons. Therefore, these multivariate method and machine learning algorithms were examined for the first time for the rice N estimation in this research.

1.3.   Research aim and objectives

The aim of the study is to accurately estimate and map the rice nitrogen (N) status during different growing seasons (2016 early wet season and 2017 dry season) in IRRI experimental fields using UAV data and field observations. Based on the aim and the available datasets of this study, several research objectives are proposed:

a.   To evaluate the relationships between nitrogen measurements obtained destructively in rice crops (from tissue analysis), and nitrogen proxies obtained un-destructively using SPAD readings, and colour chart during the dry and wet seasons.

b.   To compare the performance of common vegetation indices (such as red-edge index, green index) for estimation of plant nitrogen accumulation on rice crops using UAV data for both dry and wet seasons.

c.   To evaluate the performance of different multivariate methods such as stepwise and partial least square regression, support vector regression and random forest for estimation of plant nitrogen accumulation on rice crops using UAV data at dry and wet season.

d.   To map the plant nitrogen accumulation of rice crops in both dry and wet seasons using UAV data and the most accurate method (based on R^2, RMSE and NRMSE of the methods) studied.

1.4.   Research questions

a.   What is the relationship between different rice nitrogen measurements obtained using tissue analysis (destructive) and nitrogen proxies measured using non-destructive methods?

b.   What is the optimum vegetation index (in terms of R^2) calculated from UAV data for rice plant nitrogen accumulation estimation in dry and wet seasons?

c.   What is the most accurate machine learning method (in terms of R^2, RMSE and NRMSE) to estimate plant nitrogen accumulation in the dry season and wet season?

2.   STUDY AREA AND DATA

This chapter consists of two main sections. Section 2.1 describes the study site, and section 2.2 explains the field and UAV data acquisition which were used for the analysis in this study.

2.1.   Study Area

The study area is in the IRRI (International Rice Research Institution) Zeigler Experiment Station (also known as the IRRI farm) in Los Baños, the Philippines. This area is 21m above sea level with a tropical climate, which is suitable for rice growth all year round. IRRI established a Long-Term Continuous Cropping Experiment (LTCCE) in 1962, which is the longest-running rice field trial in the world. The LTCCE covers an area of one hectare and is located at 14° 10' 5.6424" N, 121° 15' 21.1788" E (Figure 1). It aims to determine the impact of growing irrigated rice continuously, season after season, and year after year, on crop productivity and soil health. To apply different rice experiments on the field, the 100m x 100m LTCCE site is divided into 108 main plots of 8m x 8m and 72 subplots of 4m x 8m. The layout of LTCCE may change every season; therefore, field data acquisition (including UAV and field biophysical sampling) need to be collected for different wet and dry seasons. In 2016, rice was planted on April 20th and harvested in the middle of August. Then, rice was planted again on December 21st and harvested at the middle of April in 2017. Due to the precipitation is mainly concentrated from April to October in Los Baños (Figure 2), the rice season in 2016 was defined as the early wet season (EWS) and the season in 2017 was defined as dry season (DS).
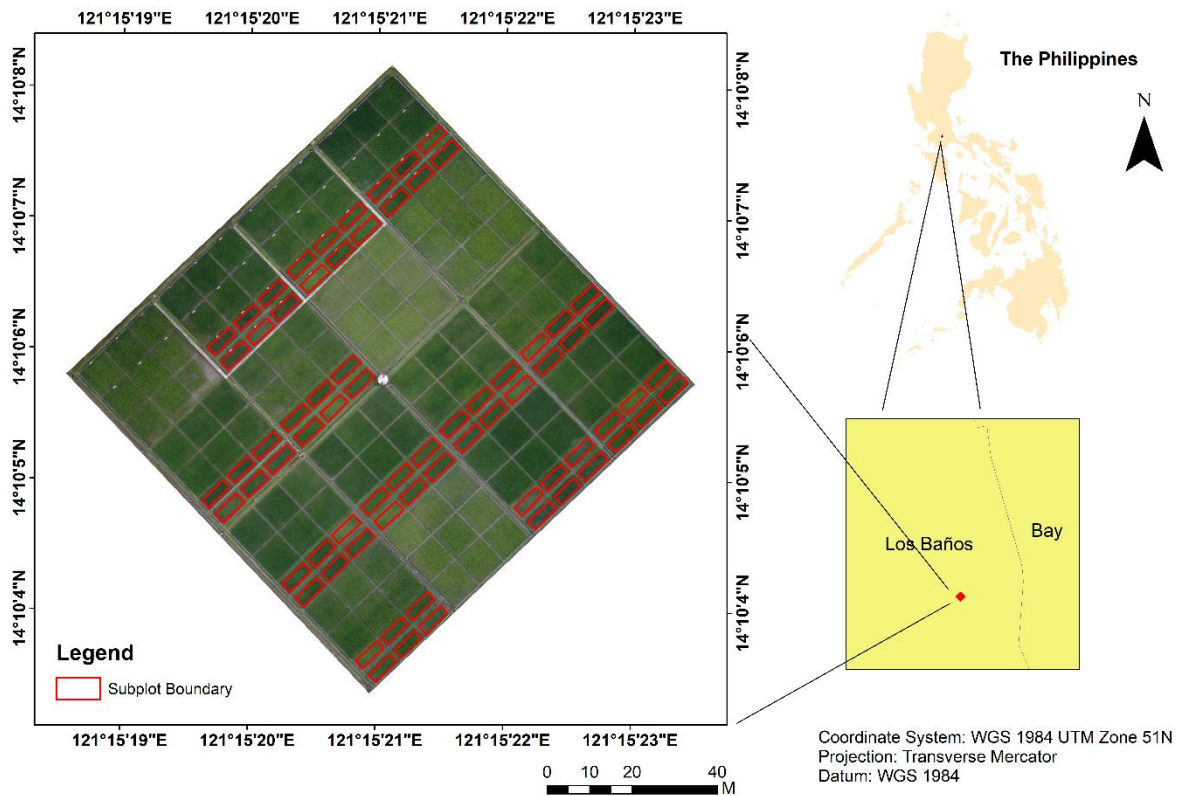
Figure 1: The LTCCE in the IRRI (International Rice Research Institution) Zeigler Experiment Station, Los Baños, the Philippines and the location of subplots (red polygons) used as samples in this study. The background RGB image was obtained by UAV on June 1st, 2016.
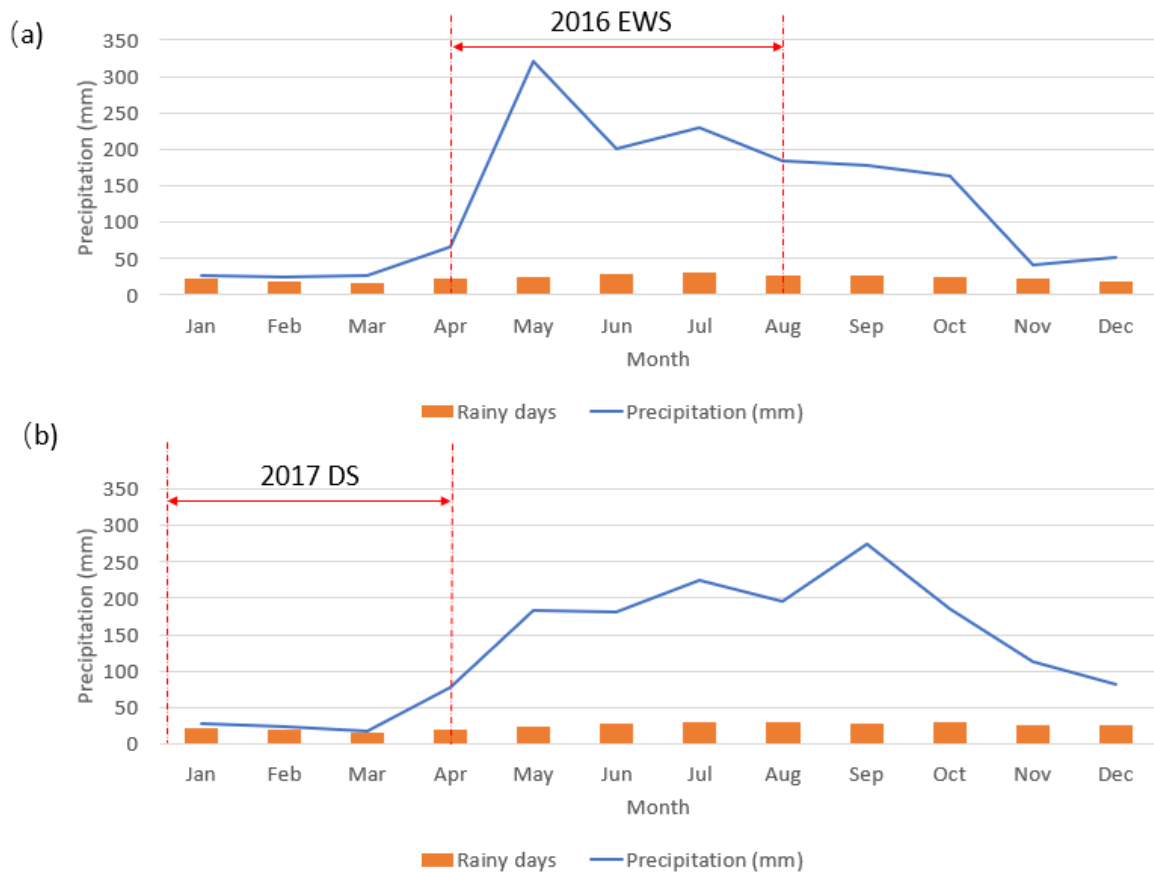
Figure 2: The average monthly precipitation (mm) in Los Baños, the Philippines in 2016 (a) and 2017 (b) ("World Weather Online," n.d.).

2.2.   Data

2.2.1.   Field data

The field data was provided as a Microsoft Excel file and contained different components. It included the details of field measurement methods, nitrogen (N) treatment dates, metadata for each variable, the 2016EWS and 2017DS field layout, variety treatment, etc. All the field data information was provided by Dr Roland Buresh (IRRI).

Three rice varieties: IRRI 146 (V4), IR111690H (V7), and IR²-10-L1-Y1-L2 (V8) were planted in the subplots with eight different N applications. The layouts for the field experiments in 2016EWS and 2017DS are shown in Figure 3, with the subplots shown in white. In both seasons, the experiments used the same rice varieties, but different N applications. Table 1 shows the different N applications in 2016EWS and 2017DS. It can be observed that N was applied for several times in different rice growth stages. The specific N application dates were managed by the crop calendar, and the operation dates are listed in Table 2. In this study, all the field operations were taken after the transplanting dates, and therefore are referred to as days after transplanting (DAT). For the whole rice-growing season, the sowing, transplanting and harvest time was clearly recorded in the field data. However, the rice panicle initiation and heading stages dates need to be estimated according to the duration of the rice crops in the field. The rice heading started about 30 days before the physiological maturity. As for the panicle initiation stage, it

was counted about 60 days before the physiological maturity. Then, Table 3 shows the date for different rice-growing stages in 2016EWS and 2017DS.

In the LTCCE, the destructive and non-destructive sampling methods were applied for field measurements. These include N tissue analysis, SPAD-502 chlorophyll meter readings, leaf colour chart (LCC) readings and leaf area index (LAI) measurement. These parameters were measured on different dates during the rice-growing season (Table 4). The LAI and tissue analysis measurements were performed using destructive sampling. LAI was measured four times during the rice-growing season and only on green leaves using a LI-3000C Portable Leaf Area Meter. As for the tissue analysis, leaves were collected from the field and used established equations to retrieve plant N accumulation (PNA) from straw and grain dry mass. All samples were oven-dried for 72 hours at 80°C and weighted to determine dry weight, and the straw and grain N concentration were measured using micro-Kjeldahl digestion method (Lang, 1958). Then, using their respective concentration times, their oven-dry weight obtained the straw and grain N accumulation. Finally, the PNA was calculated as the sum of the straw and grain N and expressed in kg of N per hectare. The tissue analysis was performed only once at the physical maturity stage. PNA represents the total amount of N accumulated in the aboveground plant biomass at harvest time, and it is more representative and comprehensive to describe the N status of rice crops, it was used for further analysis in this study.

The SPAD, and leaf colour chart (LCC) measurements were taken from the leaves of rice crops nine times during the growing season. The SPAD chlorophyll meter is a commonly used instrument for non-destructive chlorophyll measurements. SPAD readings have shown to have a strong correlation with leaf chlorophyll concentration (Uddling et al., 2007). Chlorophyll concentration in leaves and canopies can be an indicator of photosynthetic capacity, developmental stage, plant productivity, and N concentration (Ustin, 1999). Therefore, SPAD readings can be used as a proxy for N status. In this study, SPAD readings were used as the proxy for leaf chlorophyll content. The canopy chlorophyll content was then calculated by SPAD value times the LAI value. Because the LAI was only measured four times during the season in 2016EWS, the CCC values were obtained for the same dates as the LAI measurements. It is calculated using LAI times SPAD value. However, in the 2017DS, the CCC values were only calculated for three different dates after matching the SPAD and LAI values with the same measurement dates.

The leaf colour chart (LCC) is an inexpensive alternative to the chlorophyll meter (Byju wt al., 2009). The LCC measurements were taken using a four-panel colour chart made by IRRI. After comparing the greenness of rice leaf with the critical colour shade, the leaf colour threshold could be decided, which indicated its N content (Peng et al., 1996).

After acquiring the field data, checking on data distribution and variation is an important step before any data analysis. An outlier could indicate that there is an error on the observation during the measurement, and this might cause a misleading result (Bansal et al., 2016). Only the field measurements correlated with rice N status were used in this study. Exploratory data analysis was performed for SPAD, LCC, and LAI time-series data per rice variety using line charts, box plots and histogram distributions. The outliers were mainly detected by the time-series data variation of line chart (Figure A3) and boxplots (Figure A4) and therefore, were excluded from further analysis. After confirming with Dr Roland Buresh from IRRI, one SPAD measurement from V8 on DAT43 in 2016EWS, two LAI measurements from V4 in 2017DS on DAT 43 and DAT 68 were detected as outliers and removed from further analysis.
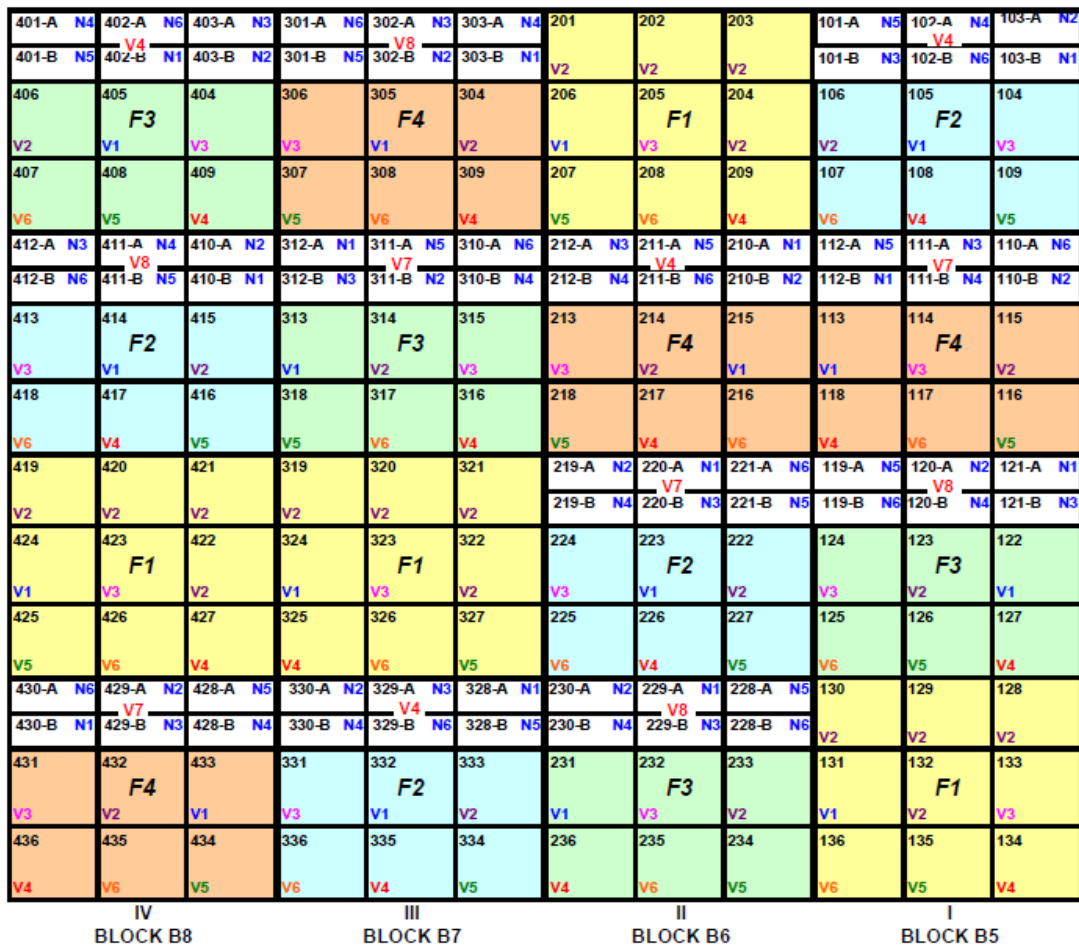
Figure 3: LTCCE field layout for 2016EWS and 2017DS in IRRI Block B5-8. The subplots are white.

Table 1: N treatment applications in the subplots in 2016EWS and 2017DS in the LTCCE in the IRRI Zeigler Experiment Station, Los Baños, the Philippines.

Season	N treatment	N treatment description	Fertilizer N applied (kg N/ha)				
			N rate	Basal	Tillering	Panicle initiation	Booting
2016EWS	N1	No added N	0	0	0	0	0
	N2	Low target yield	90	30	30	30	0
	N3	High target yield	135	45	45	45	0
	N4	Intermediate target yield, 3 splits	115	30	43	43	0
	N5	Intermediate target yield, 3 splits	115	30	30	55	0
	N6	Intermediate target yield, 4 splits	115	30	30	40	15
2017DS	N1	No added N	0	0	0	0	0
	N2	Low target yield	130	38	38	38	16
	N3	High target yield	195	57	57	57	24
	N4	Intermediate target yield, 3 splits	165	45	60	60	0
	N5	Intermediate target yield, 3 splits	165	45	50	70	0
	N6	Intermediate target yield, 4 splits	165	45	50	50	20

Table 2: N application dates in different rice-growing stages in 2016EWS and 2017DS on LTCCE in IRRI Experimental Station in Los Baños, the Philippines.

	Sequence	Calendar date	Days after transplanting (DAT)
2016EWS	Basal	2016-05-11	7
	Tillering	2016-05-27	23
	Panicle initiation	2016-06-16	43
	Booting	2016-07-06	63
2017DS	Basal	2017-01-11	7
	Tillering	2017-01-28	24
	Panicle initiation	2017-02-17	44
	Booting	2017-03-09	64

Table 3: Rice growth stages in 2016EWS and 2017DS in the LTCCE in the IRRI Zeigler Experiment Station, Los Baños, the Philippines.

Rice growth stage	2016EWS	2017DS
	Date	Date
Seed sowing	2016-04-20	2016-12-21
Transplanting	2016-05-04	2017-01-04
Panicle initiation	2016-06-05	2017-02-12
Heading	2016-07-05	2017-03-14
Harvest	2016-08-04	2017-04-13

Table 4: Dates of different field measurements used in this study taken in the LTCCE in the IRRI Zeigler Experiment Station, Los Baños, the Philippines

Field Measurement	Sequence	2016EWS		2017DS	
		Calendar date	Days after transplanting (DAT)	Calendar date	Days after transplanting (DAT)
LAI	1	2016-05-26	22	2017-01-27	23
	2	2016-06-15	42	2017-02-16	43
	3	2016-06-28	55	2017-03-02	57
	4	2016-07-05	62	2017-03-13	68
SPAD & LCC	1	2016-05-26	22	2017-01-27	23
	2	2016-06-02	29	2017-02-02	29
	3	2016-06-09	36	2017-02-09	36
	4	2016-06-15	42	2017-02-16	43
	5	2016-06-22	49	2017-02-23	50
	6	2016-06-29	56	2017-03-02	57
	7	2016-07-06	63	2017-03-09	64
	8	2016-07-14	71	2017-03-16	71
	9	2016-07-21	78	2017-03-23	78

2.2.2. UAV data

The UAV data was collected with a fixed-wing UAV-Sensefly eBee during two growing seasons. Different sensors were used: RGB; 4-band multispectral; and thermal images. The RGB image resolution is 2-3 cm/px, the multispectral image has 6-8 cm/px and the thermal images have 10-15cm/px. After exploring the UAV images with different sensors, the multispectral images with green (550nm), red (660nm), red-edge (735nm), and near-infrared (190nm) bands were used for data analysis (Figure 4). The images were captured weekly during the whole rice-growing season in 2016EWS and 2017DS. In general, the UAV was flown 12 times during the season for multispectral image acquisition.

Table 5 shows the basic information about the UAV images used in this study. All images were corrected for radiometric and geometric errors by IRRI using Pix4D software. The mosaics were spatially corrected using ground control points, and Pix4D automatically discarded images if they were too blurry. The UAV data was already mosaiced by Mr Steve Klassen from IRRI, and it could be directly used for reflectance extraction and further analysis.
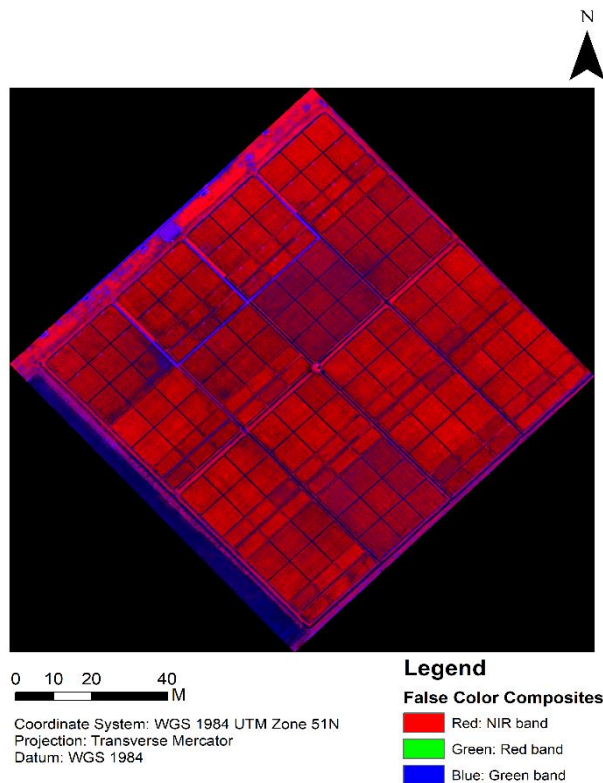
Figure 4: False colour composites using the near-infrared band as red, the green band as green, and the red band as blue. Multispectral images were taken on June 1st, 2016 in the LTCCE in the IRRI Zeigler Experiment Station, Los Baños, the Philippines.

Table 5: Multispectral UAV images information obtained from IRRI and used in the thesis.

	Image folder	Date of acquisition	Reflectance resolution (cm)
2016EWS	2016-05-23_mslte	2016-05-23	7.6
	2016-05-25_mslte	2016-05-25	7.4
	2016-06-01_mslte	2016-06-01	7.4
	2016-06-08_mslte	2016-06-08	7.6
	2016-06-14_mslte	2016-06-14	7.1
	2016-06-21_mslte	2016-06-21	7.3
	2016-06-28_mslte	2016-06-28	7.3
	2016-07-05_mslte	2016-07-05	7.7
	2016-07-06_mslte	2016-07-06	7.4
	2016-07-22_mslte	2016-07-22	7.2
	2016-07-27_mslte	2016-07-27	7.1
	2016-08-03_mslte	2016-08-03	7.4
2017 DS	2017-01-13_mslowland	2017-01-13	7.28
	2017-01-19_mslowland	2017-01-19	7.05
	2017-02-08_mslln	2017-02-08	7.49
	2017-02-16_mslowland	2017-02-16	7.34
	2017-02-23_mslowland	2017-02-23	7.37
	2017-03-02_mslowland	2017-03-02	7.19
	2017-03-08_mslowland	2017-03-08	7.6
	2017-03-15_mslowland	2017-03-15	7.6
	2017-03-22_mslte	2017-03-22	6.2
	2017-03-29_mslowland	2017-03-29	6.78
	2017-04-05_mslowland	2017-04-05	6.98
	2017-04-11_mslowland	2017-04-11	7.1

3. METHODS

This chapter explains the methods used in this study to estimate PNA of rice crops at harvest time. Figure 5 presents an overview of the methodology used in this thesis. This chapter consists of seven sections. Section 3.1 describes the pre-processing procedures of the reflectance extraction from the time-series multispectral UAV imagery. Section 3.2 aims to examine the relationship between three field measurements (PNA, SPAD and LCC) used to describe the N in the rice fields. Section 3.3 describes the VIs selection and their correlation with agronomic parameters (SPAD, CCC and PNA). In section 3.4, linear regression models, including simple linear regression (SR) and stepwise multiple linear regression (SMLR) were built using the best performance VI to predict PNA. Section 3.5 explains the use of multivariate method (PLSR) and machine learning algorithms (SVR and RF) for PNA estimation, including the details of model selection and hyperparameter tuning. Section 3.6 describes the validation and accuracy assessment of multivariate methods and machine learning algorithms. Section 3.7 explains the procedure of mapping the PNA for the 2016EWS and 2017DS using the method with the highest accuracy.
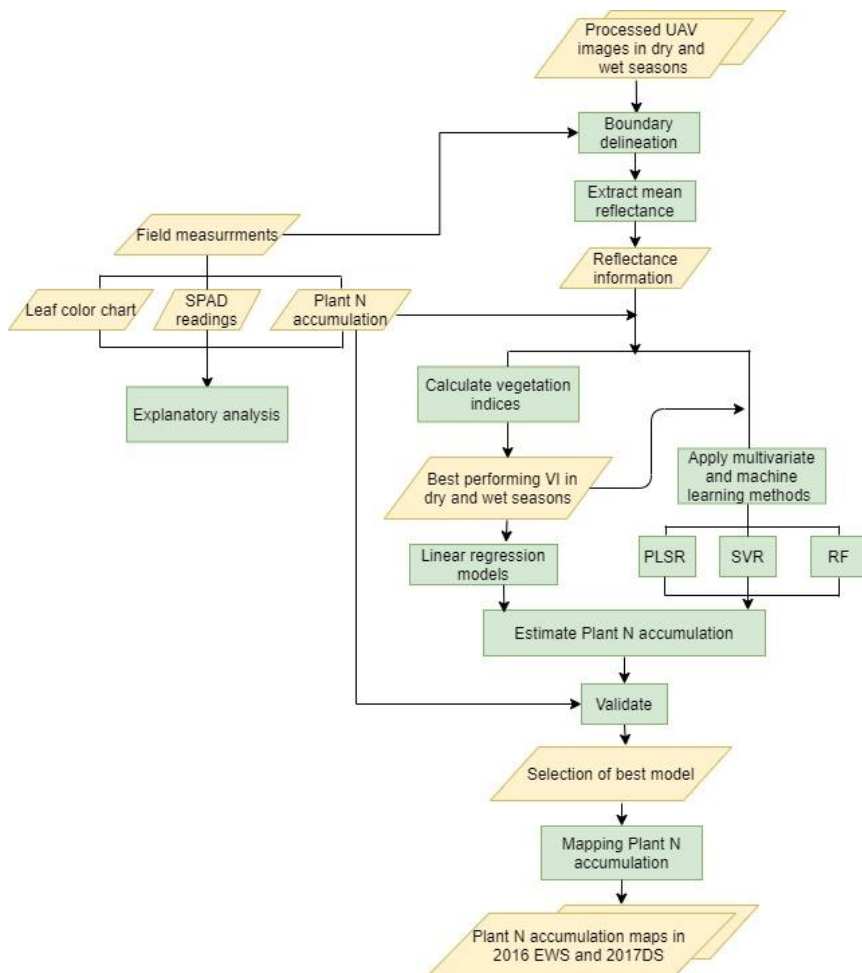
Figure 5: Methodology flowchart of the study.

3.1.   Data processing

Multispectral UAV images (see 2.2) were available for the 100m × 100m experimental field site. In this study, samples belong to the subplots where different N treatments were applied on different rice varieties and where field measurements have been taken throughout (SPAD, LAI, LCC) and at the end of the season (PNA) and represented the average value of each subplot. Therefore, the mean reflectance data of each subplot was extracted from all multispectral images. To be able to extract the reflectance at the subplot level, the subplot boundaries were digitized in ArcGIS 10.7. To avoid the shadow influence at the edge of subplots, inner subplots boundaries were manually digitized. The inner subplots boundaries were 0.5m away from the main subplot boundaries (Figure 6). After generating the inner subplots boundaries, the mean reflectance was extracted using the 'extract' function in the raster package in R (Hijmanset al., 2020). This function could directly extract the reflectance values of pixels within the subplot polygons and calculate statistics. In this process, the mean and standard deviation are the main outcomes of the extraction process. The mean reflectance values of the spectral bands were then used for the VI calculations and models input, and the standard deviations were used to check the variation of the reflectance values and to detect the possible outliers per subplot.

Legend

☐ Subplot boundary

0 5 10 20 30 40
▬▬▬▬▬ M
Coordinate System: WGS 1984 UTM Zone 51N
Projection: Transverse Mercator
Datum: WGS 1984
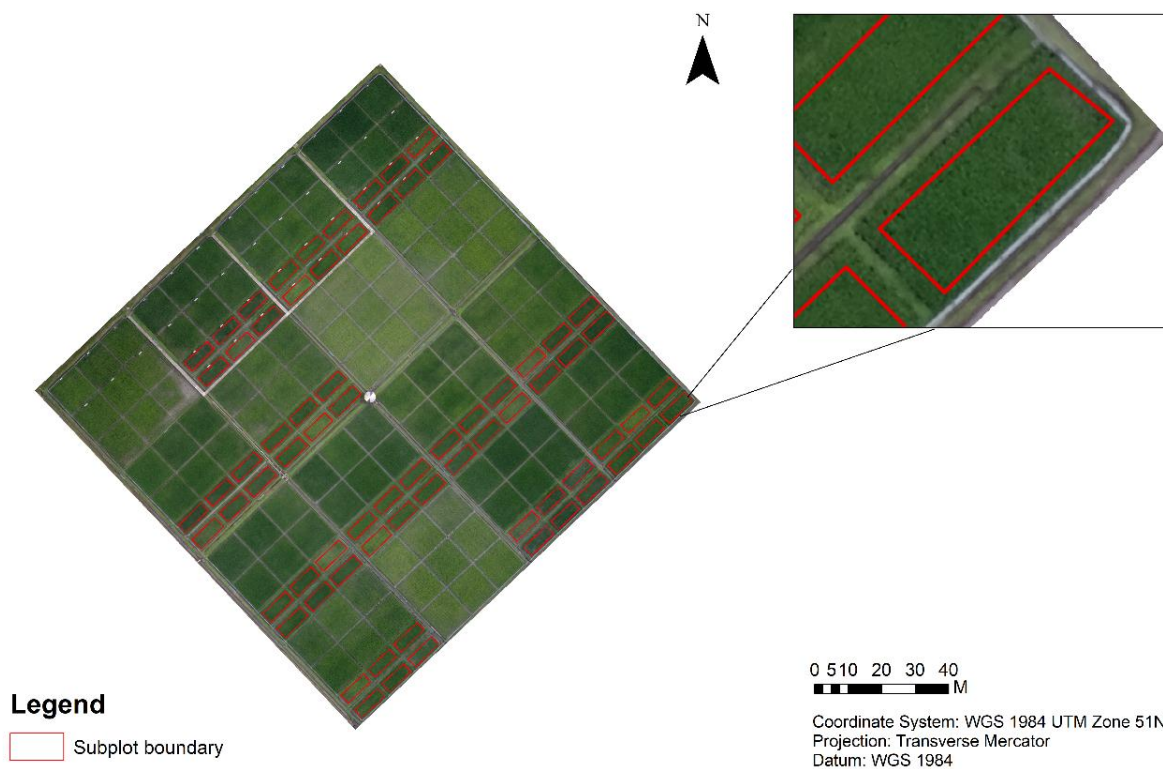
Figure 6: Digitized inner subplot boundaries with 0.5m distance away from the subplot boundary (in red). The background colour composite image (RGB) shows the LTCCE in the IRRI Experimental Station in Los Baños, the Philippines on June 1st, 2016.

3.2. Relationships between Agronomic Parameters

To understand the relationships between field measurements used as N proxies, SPAD and LCC and nitrogen measurements obtained destructively, the correlation among them were explored. The Pearson correlation coefficients (r) were therefore calculated between SPAD, LCC and PNA. Pearson correlation coefficient is a statistical method to judge the linear correlation between two variables. It ranges from -1 to+1. The higher absolute "r" value refers to two variables have a strong linear relationship with each other.

3.3. Relationships between VIs and Agronomic Parameters

VI is a spectral transformation by combining information from two or more spectral bands for enhancing the detection of vegetation properties and comparing the terrestrial photosynthetic activity and canopy structural variations (Huete et al., 2002). The VIs selection was based on the reflectance variation (Figure A1 and Figure A2) and different VIs characteristics. Seven VIs are selected in this research, and they are respectively NDVI, GNDVI, RVI, GRVI, CI, NDRE and RTVI (Table 6). All the VIs were calculated in R using the mean reflectance from each subplot.

The agronomic parameters of SPAD, CCC and PNA were correlated with selected VIs in different dates during the whole growing season, then the date with the highest correlation coefficient was used for PNA estimation. Although LCC is also related to N measurements, the measuring result is largely dependent on the subjective judgement of the observer, and it is less accurate compared to SPAD measurements (Duyet al., 2019). Therefore, it's relation with VIs was not considered in this section. In this study, the SPAD and CCC were time-series data (measured nine times per season), therefore, they could be correlated with VIs from the same date. The PNA was only measured at the harvest time, so the correlations were built between VIs in different growing stages with the harvest PNA.

Table 6: Vegetation selected in this study.

Vegetation Index	Formula	Reference
Normalized Difference Vegetation Index (NDVI)	(NIR - R) / (NIR + R)	(Tucker, 1979)
Green Normalized Difference Vegetation Index (GNDVI)	(NIR - G) / (NIR + G)	(Gitelson et al., 1996)
Ratio Vegetation Index (RVI)	NIR / R	(Tucker, 1979)
Green Ratio Vegetation Index (GRVI)	NIR / G	(Blackmer et al., 1996)
Chlorophyll Index (CI)	NIR / G - 1	(Gitelson, 2005)
Normalized Difference Red Edge Index (NDRE)	(NIR - RE) / (NIR + RE)	(Carneiro et al., 2019)
Red-edge Triangular Vegetation Index (RTVI)	100*(NIR - RE) - 10*(NIR -G)	(Chen et al., 2010)

3.4. Linear regression models between best performing VI and PNA

According to the correlation coefficient between the VIs and agronomic parameters, the best performing VI in specific rice-growing date was identified in both the 2016EWS and 2017DS. The crucial dates were selected according to the higher correlation coefficient between best performing VI and agronomic parameter during the rice-growing season. Then, linear regression models were built using the best performing VI in crucial dates for PNA estimation at harvest time.

In addition to estimating PNA using VI from a single date, applying the best performing VI in the whole growing season is also worth trying. To understand if all the VIs for the whole rice-growing season are

needed or just the crucial date is already sufficient for PNA estimation, the performance of stepwise multiple linear regression models was investigated.

Stepwise multiple regression models between the best performing VI for the whole growing season and the PNA were built to achieve the goal. Stepwise regression is a method of fitting regression models in which the choice of predictive variables is carried out by an automatic procedure (Halinski et al., 1970). In each step, the explanatory variables) are added or removed based on a prespecified criterion normally by a sequence of F-tests or t-tests, but other metrics are also possible, such as the Akaike information criterion (AIC), which is an estimator of out-off-sample prediction error. AIC can also provide information on model quality (Aho et al., 2014). In this study, the stepwise multiple regression models were built in R with the time-series VIs data as the explanatory variables, and the PNA of the rice crops is the dependent variable. With the backward direction setting, the model will remove the best performing VIs in non-significant date in each step until it got the smallest AIC value, then the remaining best performing VI with different dates with the most explanatory power are used for estimating the PNA. The PNA was estimated using stepwise multiple linear regression models for 2016EWS and 2017DS.

3.5.   Multivariate method and machine learning algorithms

Machine learning is a study of computer algorithms, and it provides systems with the ability of automatically learning and improving from experience (Mitchell et al., 1997). Therefore, machine learning is deemed as a subset of artificial intelligence. The main task of machine learning algorithms is to make predictions which are closely related to computation statistics. In addition to the strong learning ability, machine learning can also handle massive, multi-collinearity, non-linearity and noisy data.

In this study, the input data was the same for all models. It contained the mean reflectance for all individual bands and the best performing VI from the previous analysis for the whole rice-growing season as the explanatory variables, as well as PNA at harvest time as the independent variable. Once the variable data were prepared, multivariate method and machine learning methods were applied to estimate PNA in this study. Figure 7 shows the procedure of using multivariate method and machine learning algorithms for parameter estimation.
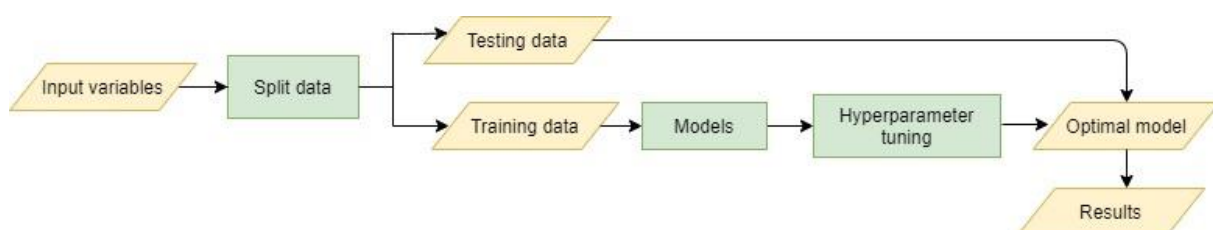
Figure 7: Multivariate and machine learning methods procedure overview

3.5.1.   Model Selection

In this study, the PNA value was the predicting variable with continuous numeric values, therefore, several linear and non-linear models were involved for PNA estimation: the partial least square regression (PLSR), the support vector regression (SVR) and random forest (RF). The basic theories of these models were explained below:

(1) Partial Least Square Regression

Partial least square regression (PLSR) is a popular linear model, the main idea of this technique is to reduce the predictors to a smaller set of uncorrelated components, and then to perform least squares regression on these components. It has been used in different field of studies, including remote sensing (Geladi et al., 1986). PLSR is an alternative method of multiple linear regression models. However, PLSR is more robust than traditional regression models (Almergren et al., 2009). PLSR finds the multidimensional direction in the independent space, which explains the maximum variance direction in the dependent space PNA in an iterative manner. In general, the PLSR can fit multiple response variables in a single model. In this study, the main parameter used in this algorithm was the ncomp, which refers to the component number of the models (Table 7). After applying the input data into the algorithm, the package could directly train and select the optimal PLSR model (determined by RMSE) with the best ncomp as the result. In the meantime, the importance of input features could also be known as an intermediate result in the study.

(2) Support Vector Regression

Support vector regression (SVR) is a supervised learning model used for regression analysis (Drucker et al., 1997). SVR gives flexibility to define how much error is acceptable in the model and will find an appropriate line (or hyperplane) in higher dimensions to fit the data (Awad et al., 2015). SVR could construct both linear and nonlinear regressions applying different kernels. One of the main benefits of SVR is that the computational complexity does not depend on the dimensionality of the input space. In addition, SVR offers excellent generalization capability and high prediction accuracy (Wu et al., 2008). Table 7 listed the parameters used in the SVR model in this study. e1071 package helped to train and determine the optimal SVR algorithm (based on RMSE) after applying the input variables in the algorithm for PNA estimation. The parameters of the optimal SVR model could be checked in the model result.

(3) Random Forest

Random forest (RF) is an ensemble learning algorithm for classification and regression, which is a bagging technique by constructing masses of decision trees while training and generating a mean prediction (regression) of individual trees (Ho, 1995; Ho, 1998). In the standard decision tress, each node was split according to the best split among all variables. However, each node in RF is split using the best among a subset of randomly selected predictors at that node (Liawwt al., 2002), which highly improves the model's generalization performance and its training accuracy. In addition, RF can handle thousands of input variables and list variables importance to the model as the analysis result. Based on the above advantages of RF, it is considered as a powerful machine learning algorithm in this study. Some important parameters were listed in Table 7 for training RF models in the study. After applying the input variables in RF algorithm, the optimal model could be selected and determined (based on RMSE) by grid search method by using the Caret package. Then, parameter values of the optimal RF model and the importance of input features could be known by checking the summary of the model.

3.5.2.   Hyperparameter tuning

Hyperparameter optimization is used to choose various parameters to get the optimal model to solve the prediction task. In different models and different statistical packages, the tuning parameters are different. Table 7 shows the parameters and packages used in R for different models. In this study, different

packages were used to select the best value or methods for significant parameters in multivariate and machine learning models.

Table 7: Multivariate methods and machine learning algorithms with significant parameters and packages used in R.

	Tuning parameters	Parameter explanations	Packages
PLSR	"ncomp"	The number of components to include in the model	Caret (Kuhn et al., 2020)
SVR	"epsilon"	Epsilon in the insensitive-loss function	e1071 (Meyer et al., 2019)
	"kernel"	The kernel used in training and predicting. Different kernel type could be selected in the model.	
	"cost"	Cost of constraints violation	
RF	"mtry"	Number of variables randomly sampled as candidates at each split. Note that the default values are different for classification (sqrt(p) where p is number of variables in x) and regression (p/3)	Caret (Kuhn et al., 2020)
	"ntree"	Number of trees to grow. This should not be set to too small a number, to ensure that every input row gets predicted at least a few times	

3.6.   Validation and Accuracy assessment

In order to have the same sample size for each rice variety, stratified random sampling was used to separate the whole dataset into calibration and validation in linear regression models. As for the splitting method, 70% of the data was used for training and 30% was used for the validation. Therefore, the relationship between best performing VI in crucial dates and PNA were studied using the calibration dataset, then the validation data was used for comparing the estimated PNA with the actual PNA.

As for the PLSR and machine learning models, the input data was split into training (70%) and testing (30%) datasets by stratified random sampling based on rice varieties. To make the model accurately work for prediction, it always needs to validate the stability of the models. It is significant to assure that the model has got most of the patterns from the data correctly without too much noise or bias (Cawley et al., 2010). Cross-validation is a statistical technique for testing the performance models, and it also helps with avoiding overfitting. As there are 51 training samples in this study, the 10-fold cross-validation was applied for PLSR, SVR and RF training models. In the 10-fold cross-validation, the training data is divided into 10 subsets, and the holdout method was repeated for ten times. At each time, one of the ten subsets was used as the validation set, and other subsets were put as the training set. The error estimation is calculated by the average of ten trials error for the final model effectiveness. Using 10-fold cross-validation could significantly reduce the bias for the training data fitting and the variance of validation. In conclusion, the training dataset was used for training the model for PNA estimation, and the testing data was used for comparing the estimated PNA with measured PNA.

The coefficient of determination (R^2), root mean square error (RMSE), and normalize root mean square error (NRMSE) were calculated for evaluating model performance in this study. R^2 represents the proportion of the variance for the dependent variable explained by an independent variable or variables in

regression models. The R^2 value normally ranges from 0-1. The higher R^2 indicate a higher correlation between observations and estimations. RMSE mainly reflects the difference between the observation and prediction value by models. It also explains as the standard deviation of the residuals (prediction errors) that shows how the concentration of data is around the best fitting line. The RMSE is commonly used in regression analysis. Both R^2, RMSE were calculated by the Caret package in R. NRMSE is calculated for comparison between different scale models or parameters. The formula is shown below:

$$NRMSE = \frac{RMSE}{Omax - Omin}$$

The Omax represents the maximum value of the observation data, and the Omin is the minimum of the observation data.

3.7. Mapping N status

PNA estimation maps were generated based on the most accurate method (based on the lowest RMSE) using individual band reflectance with the best performing VI time-series data as input. The predicted and actual PNA maps in 2016EWS and 2017DS were generated in ArcGIS 10.7. The PNA values were classified into four categories (≤70kg/ha, 70-100kg/ha,100-150kg/ha,>130 kg/ha) to be easily compared and explained. The difference between the estimated and measured PNA were also shown in the form of maps for 2016EWS and 2017DS.

4.   RESULTS

The following section covers the main findings of this research, including those obtained from the analysis of field measurements and UAV data. This chapter consists of five sections. Section 4.1 illustrates the relationship between different N related agronomic parameters (PNA, SPAD and LCC). Section 4.2 presents the correlation coefficient tables between VIs and agronomic parameters (SPAD, CCC and PNA) in order to select the best performing VI. Section 4.3 shows the results of applying the best performing VI for PNA estimation using regression models. The result of using multivariate method and machine learning algorithms for PNA estimation are shown in section 4.4. The last section presents the maps of actual and estimated PNA using the most accurate model for 2016EWS and 2017DS.

4.1.   Relationship among field agronomic parameters

Figure 8 shows the variation of the correlation coefficient between different agronomic parameters in wet (a) and dry (b) season. As can be observed from this figure, similar trends exist between SPAD and PNA, SPAD and LCC in the wet and dry season. The correlation starts increasing in the third week after transplanting and becomes stable during the panicle initiation stage (r>0.7). At the end of the season, the correlation coefficients decrease but remain at a high value. Based on the different combinations, the correlation between SPAD and LCC shows the best performance for both seasons. In conclusion, these three field measurements are well correlated in the panicle initiation and heading stage.
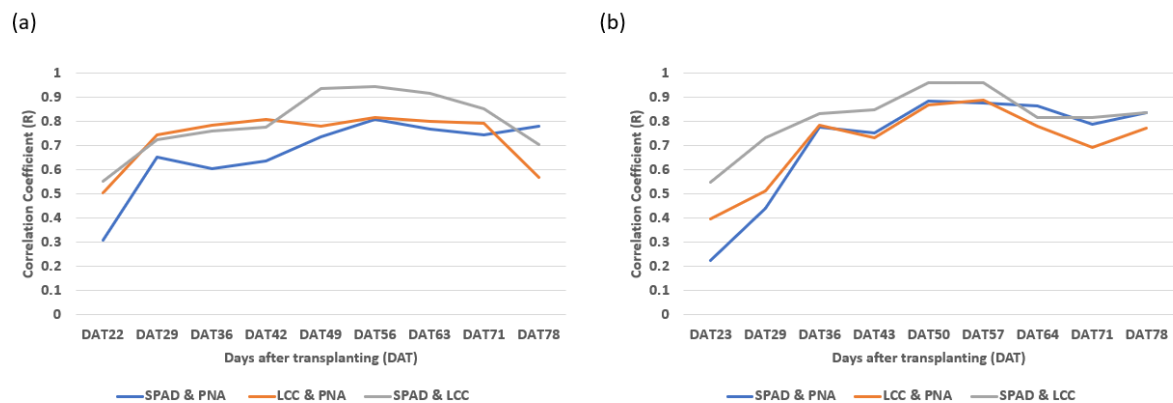
Figure 8: Correlation coefficient among SPAD, LCC, and PNA measured in rice fields during 2016EWS (a) and 2017DS (b) in LTCCE field in IRRI Zeigler Experiment Station, Los Baños, the Philippines.

4.2.   Relationship between vegetation indices and agronomic parameters

Table 8 and Table 9 list the correlation coefficient of VIs with SPAD and CCC values in critical days in 2016EWS and 2017DS. The correlation coefficient between VIs and agronomic parameters was calculated for the whole rice-growing seasons, then field measurement days with relatively high r-value were deemed as crucial days in the rice growing season. They are DAT 56 and DAT 63 in 2016EWS, and DAT 57 and DAT 68 in 2017DS. The results indicated that the NDVI, GNDVI, RVI, GRVI, CI, NDRE and RTVI had positive relationships with SPAD value and CCC value. In general, the correlation coefficients are higher than 0.8, which means that these selected VIs are highly correlated with the chlorophyll content in

both leaf and canopy levels, and the correlation with leaf level is better than the canopy level. Figure 9 shows the significance test between VIs and SPAD on DAT 56 in 2016EWS, the p value shows that the calculated correlation is significance with 72 samples in the study.

Table 10 shows the correlation coefficients between VIs in two critical days and PNA measured in harvest day in 2016EWS and 2017DS. The correlation between VIs and PNA shows a high correlation in these two crucial days. Although the PNA was only measured once at the end of the season, it has a positive correlation with the selected two dates. The correlation coefficient is high, such as GNDVI, CI, NDRE, RTVI, etc., that the r-value could reach to 0.8. Therefore, the VIs calculated in critical rice-growing days could be used for PNA estimation in different seasons.

Table 8: Correlation Coefficients (r) between VIs and SPAD meter readings and CCC for each variety (n=24) and all varieties (n=72) in DAT 56 and DAT 63 in 2016EWS.

| | SPAD | | | | | | | | CCC | | | | | | | |
| | DAT 56 | | | | DAT 63 | | | | DAT 56 | | | | DAT 63 | | | |
	All	V4	V7	V8	All	V4	V7	V8	All	V4	V7	V8	All	V4	V7	V8
NDVI	0.94	0.96	0.96	0.93	0.89	0.95	0.96	0.96	0.87	0.81	0.88	0.92	0.84	0.93	0.90	0.90
GNDVI	0.93	0.95	0.97	0.92	0.93	0.94	0.97	0.96	0.88	0.84	0.90	0.93	0.87	0.92	0.91	0.87
RVI	0.88	0.91	0.93	0.86	0.79	0.87	0.94	0.90	0.89	0.86	0.88	0.94	0.78	0.91	0.90	0.90
GRVI	0.88	0.96	0.94	0.87	0.89	0.90	0.96	0.93	0.89	0.85	0.91	0.94	0.85	0.90	0.92	0.86
CI	0.88	0.90	0.94	0.87	0.89	0.88	0.96	0.93	0.89	0.85	0.91	0.94	0.85	0.90	0.92	0.86
NDRE	0.90	0.94	0.91	0.88	0.91	0.91	0.92	0.92	0.88	0.82	0.89	0.93	0.89	0.93	0.89	0.92
RTVI	0.83	0.90	0.86	0.75	0.85	0.89	0.90	0.88	0.83	0.81	0.85	0.83	0.87	0.93	0.88	0.93


```
(a)        Pearson's product-moment correlation

data:  NDVI_2016$`2016_06_28` and d$SPAD56.DAT
t = 23.298, df = 70, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.9073106 0.9628797
sample estimates:
      cor
0.9411532
```

```
(b)        Pearson's product-moment correlation

data:  GNDVI_2016$`2016_06_28` and d$SPAD56.DAT
t = 21.675, df = 70, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.8945826 0.9576128
sample estimates:
      cor
0.9329089
```

```
(c)        Pearson's product-moment correlation

data:  RVI_2016$`2016_06_28` and d$SPAD56.DAT
t = 15.598, df = 70, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.8161955 0.9242195
sample estimates:
      cor
0.8812343
```

```
(d)        Pearson's product-moment correlation

data:  GRVI_2016$`2016_06_28` and d$SPAD56.DAT
t = 15.681, df = 70, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.8177404 0.9248940
sample estimates:
      cor
0.882268
```

```
(e)        Pearson's product-moment correlation

data:  CI_2016$`2016_06_28` and d$SPAD56.DAT
t = 15.681, df = 70, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.8177404 0.9248940
sample estimates:
      cor
0.882268
```

```
(f)        Pearson's product-moment correlation

data:  NDRE_2016$`2016_06_28` and d$SPAD56.DAT
t = 16.93, df = 70, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.8391063 0.9341535
sample estimates:
      cor
0.8965001
```
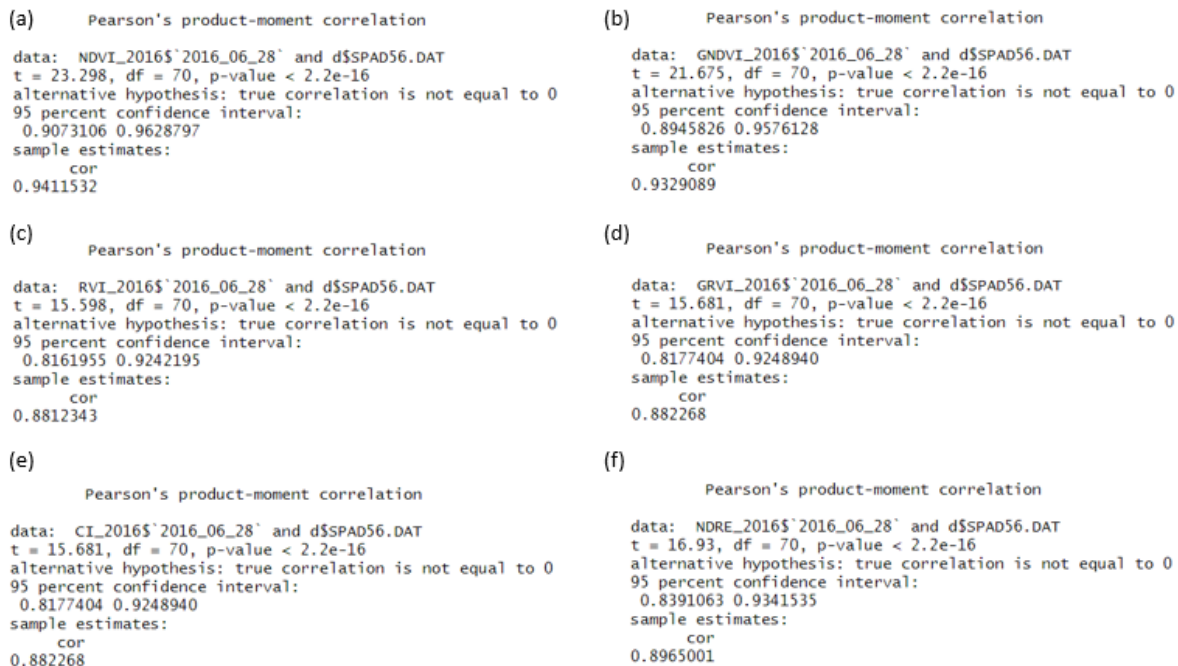
Figure 9: Correlation significance tests between VIs and SPAD measurements for all rice varieties on DAT 56 in 2016EWS (n=72).

Table 9: Correlation Coefficients (r) between VIs and SPAD meter readings and CCC for each variety (n=24) and all varieties (n=72) in DAT 57 and DAT 68 in 2017DS.

| | SPAD | | | | | | | | CCC | | | | | | | |
| | DAT 57 | | | DAT 68 | | | | DAT 57 | | | | DAT 68 | | | |
	All	V4	V7	V8	All	V4	V7	V8	All	V4	V7	V8	All	V4	V7	V8
NDVI	0.96	0.98	0.97	0.97	0.85	0.94	0.91	0.95	0.83	0.90	0.86	0.76	0.76	0.63	0.85	0.89
GNDVI	0.97	0.99	0.98	0.97	0.88	0.95	0.92	0.95	0.84	0.91	0.88	0.78	0.78	0.64	0.87	0.90
RVI	0.92	0.97	0.97	0.94	0.82	0.95	0.87	0.93	0.88	0.92	0.89	0.82	0.74	0.61	0.85	0.90
GRVI	0.94	0.97	0.97	0.94	0.89	0.95	0.90	0.94	0.86	0.92	0.89	0.82	0.78	0.65	0.87	0.91
CI	0.94	0.97	0.97	0.94	0.89	0.95	0.90	0.94	0.86	0.92	0.89	0.82	0.78	0.65	0.87	0.91
NDRE	0.95	0.96	0.97	0.94	0.87	0.93	0.93	0.94	0.85	0.91	0.91	0.81	0.79	0.65	0.89	0.91
RTVI	0.94	0.96	0.97	0.94	0.83	0.92	0.94	0.94	0.88	0.92	0.91	0.82	0.78	0.64	0.89	0.92

Table 10: Correlation Coefficients (r) between VIs and PNA for each variety (n=24) and all varieties (n=72) in DAT 56 and DAT 63 in 2016EWS, DAT 57 and DAT 68 in 2017DS.

| | PNA in 2016EWS | | | | | | | | PNA in 2017DS | | | | | | | |
| | DAT 56 | | | DAT 63 | | | | DAT 57 | | | | DAT 68 | | | |
	All	V4	V7	V8	All	V4	V7	V8	All	V4	V7	V8	All	V4	V7	V8
NDVI	0.86	0.85	0.84	0.90	0.82	0.84	0.79	0.90	0.90	0.89	0.93	0.90	0.90	0.89	0.92	0.90
GNDVI	0.87	0.87	0.85	0.92	0.83	0.87	0.78	0.90	0.91	0.90	0.95	0.92	0.92	0.91	0.93	0.92
RVI	0.90	0.88	0.91	0.92	0.79	0.82	0.80	0.91	0.91	0.91	0.96	0.95	0.91	0.91	0.90	0.93
GRVI	0.89	0.87	0.88	0.94	0.82	0.86	0.77	0.91	0.87	0.88	0.94	0.91	0.84	0.85	0.89	0.89
CI	0.89	0.87	0.88	0.94	0.82	0.86	0.77	0.91	0.92	0.92	0.96	0.95	0.93	0.93	0.93	0.93
NDRE	0.86	0.85	0.85	0.91	0.83	0.85	0.80	0.90	0.93	0.91	0.96	0.93	0.94	0.94	0.94	0.94
RTVI	0.84	0.81	0.86	0.85	0.82	0.81	0.82	0.86	0.93	0.92	0.96	0.94	0.94	0.95	0.95	0.95

4.3.   Traditional Regression models for Nitrogen status estimation using VIs

Linear regression and multiple linear regression models are firstly used for PNA estimation in this study. Figure 10 shows the calibration result of linear regression models using GNDVI in crucial days for PNA estimation in 2016EWS and 2017DS. The p-value of all regression models, which are less than 0.05, illustrate that the models are significant for PNA estimation. In addition, the significant test for each input variable was shown in the model summary. According to the p-value of the input variable, they are statistically significant coefficients in the regression models. Figure 11 and Figure 12 demonstrate the calibration and validation plots of the linear regression model using GNDVI value in crucial days to predict PNA in 2016EWS and 2017DS. The calibration model generates the equation between GNDVI and PNA, and then the validation data set was applied in the model and plotted the estimated PNA and measured PNA in different days of rice growing season. In 2016EWS, the result shows that the prediction in DAT 56 (Figure 11c) is better than in DAT 63 (Figure 11d) according to the R^2, RMSE, and NRMSE. In 2017DS, the relationship between GNDVI and PNA in DAT 68 (Figure 12c) is stronger than in DAT 57 (Figure 12d).

(a)
```
MODEL INFO:
Observations: 51
Dependent Variable: Plant.N..kg.ha.
Type: OLS linear regression

MODEL FIT:
F(1,49) = 138.14, p = 0.00
R² = 0.74
Adj. R² = 0.73

Standard errors: OLS
--------------------------------------------------------
                      Est.    S.E.   t val.      p
--------------------------------------------------------
(Intercept)        -138.11   20.25   -6.82    0.00
GNDVI2016_06_28     309.33   26.32   11.75    0.00
--------------------------------------------------------
```

(b)
```
MODEL INFO:
Observations: 51
Dependent Variable: Plant.N..kg.ha.
Type: OLS linear regression

MODEL FIT:
F(1,49) = 91.08, p = 0.00
R² = 0.65
Adj. R² = 0.64

Standard errors: OLS
--------------------------------------------------------
                      Est.    S.E.   t val.      p
--------------------------------------------------------
(Intercept)        -156.24   26.82   -5.82    0.00
GNDVI2016_07_06     339.58   35.58    9.54    0.00
--------------------------------------------------------
```

(c)
```
MODEL INFO:
Observations: 51
Dependent Variable: Plant.N..kg.ha.
Type: OLS linear regression

MODEL FIT:
F(1,49) = 234.33, p = 0.00
R² = 0.83
Adj. R² = 0.82

Standard errors: OLS
--------------------------------------------------------
                      Est.    S.E.   t val.      p
--------------------------------------------------------
(Intercept)        -179.71   19.11   -9.40    0.00
GNDVI2017_03_02     368.42   24.07   15.31    0.00
--------------------------------------------------------
```

(d)
```
MODEL INFO:
Observations: 51
Dependent Variable: Plant.N..kg.ha.
Type: OLS linear regression

MODEL FIT:
F(1,49) = 240.13, p = 0.00
R² = 0.83
Adj. R² = 0.83

Standard errors: OLS
--------------------------------------------------------
                      Est.    S.E.   t val.      p
--------------------------------------------------------
(Intercept)        -233.12   22.31  -10.45    0.00
GNDVI2017_03_15     447.08   28.85   15.50    0.00
--------------------------------------------------------
```
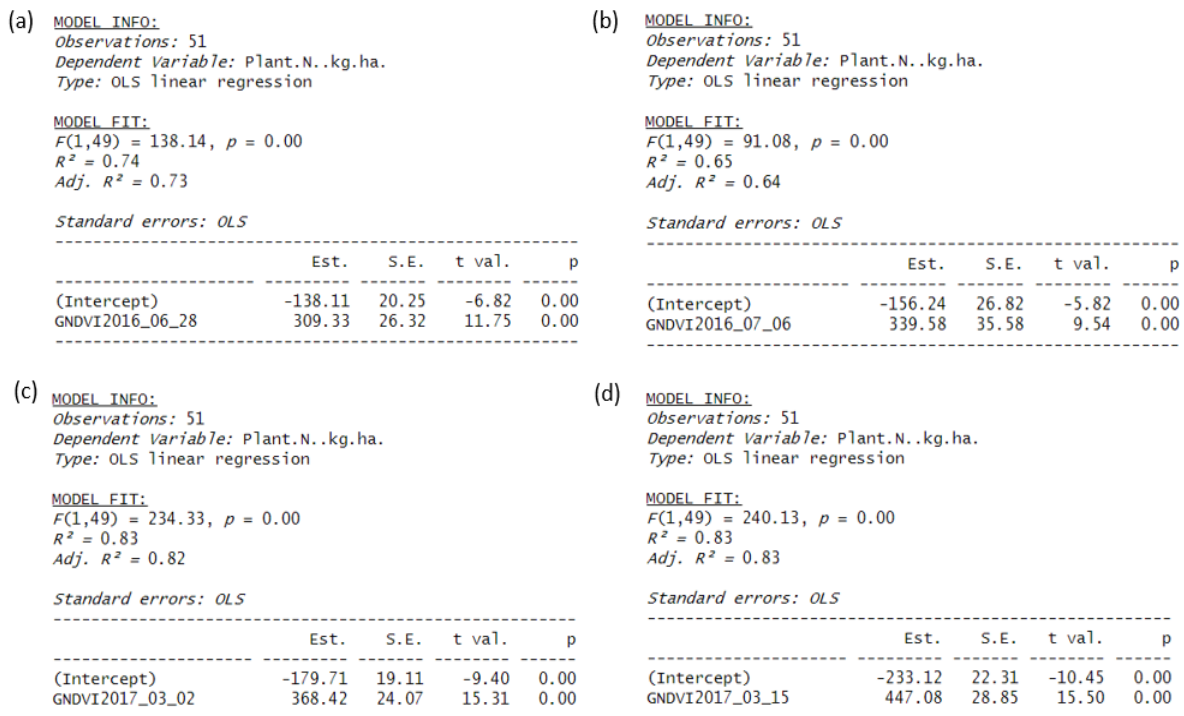
Figure 10: Linear regression model performance using single days for PNA estimation in 2016EWS (a, b)(n=51) and 2017DS (c, d) (n=51).
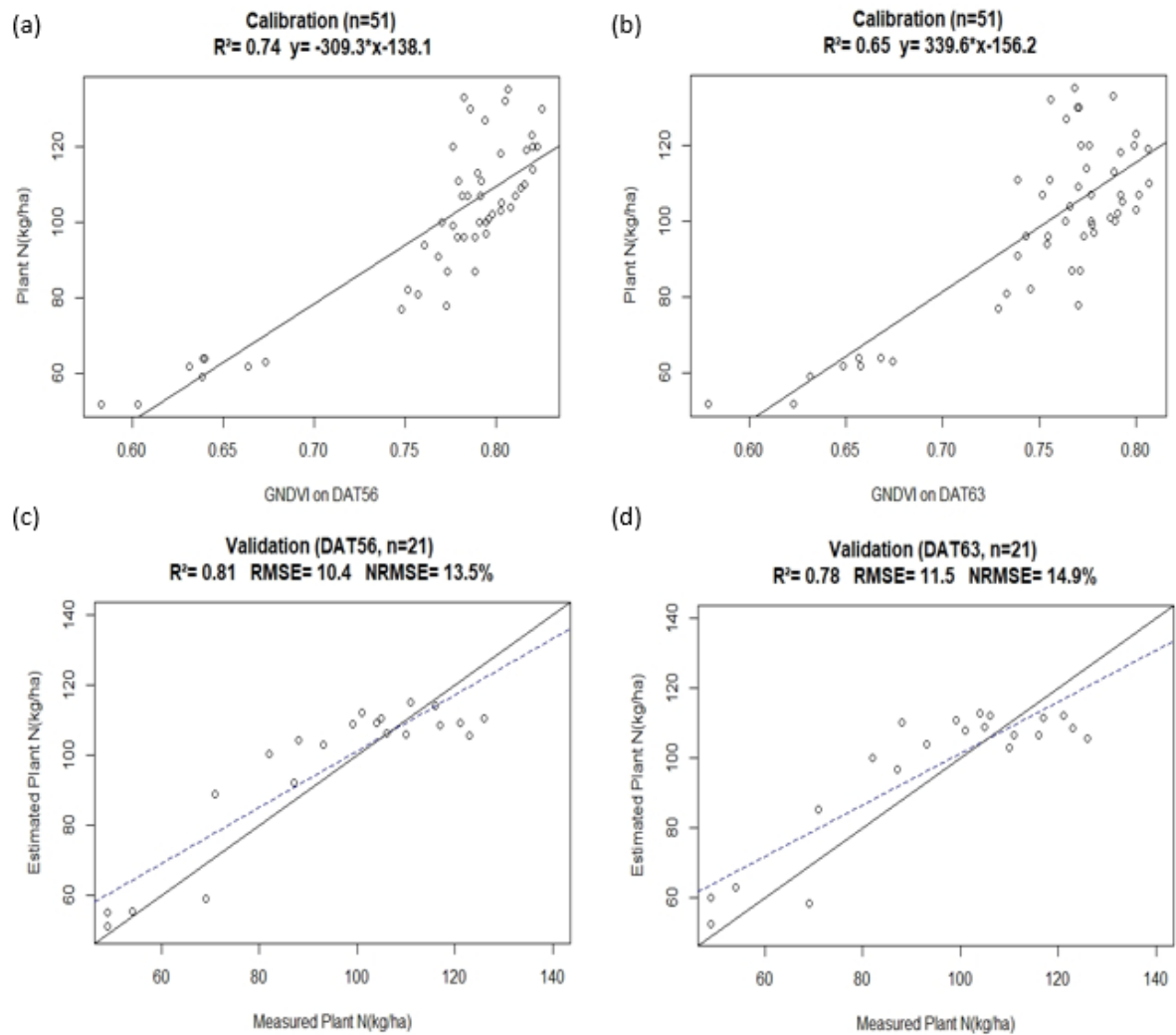
Figure 11: Relationships between measured and estimated PNA using GNDVI-based linear regression models. Calibration (n=51) and validation (n=21) plots for N status prediction in DAT 56 (a)(c) and DAT 63 (b)(d) in 2016EWS.
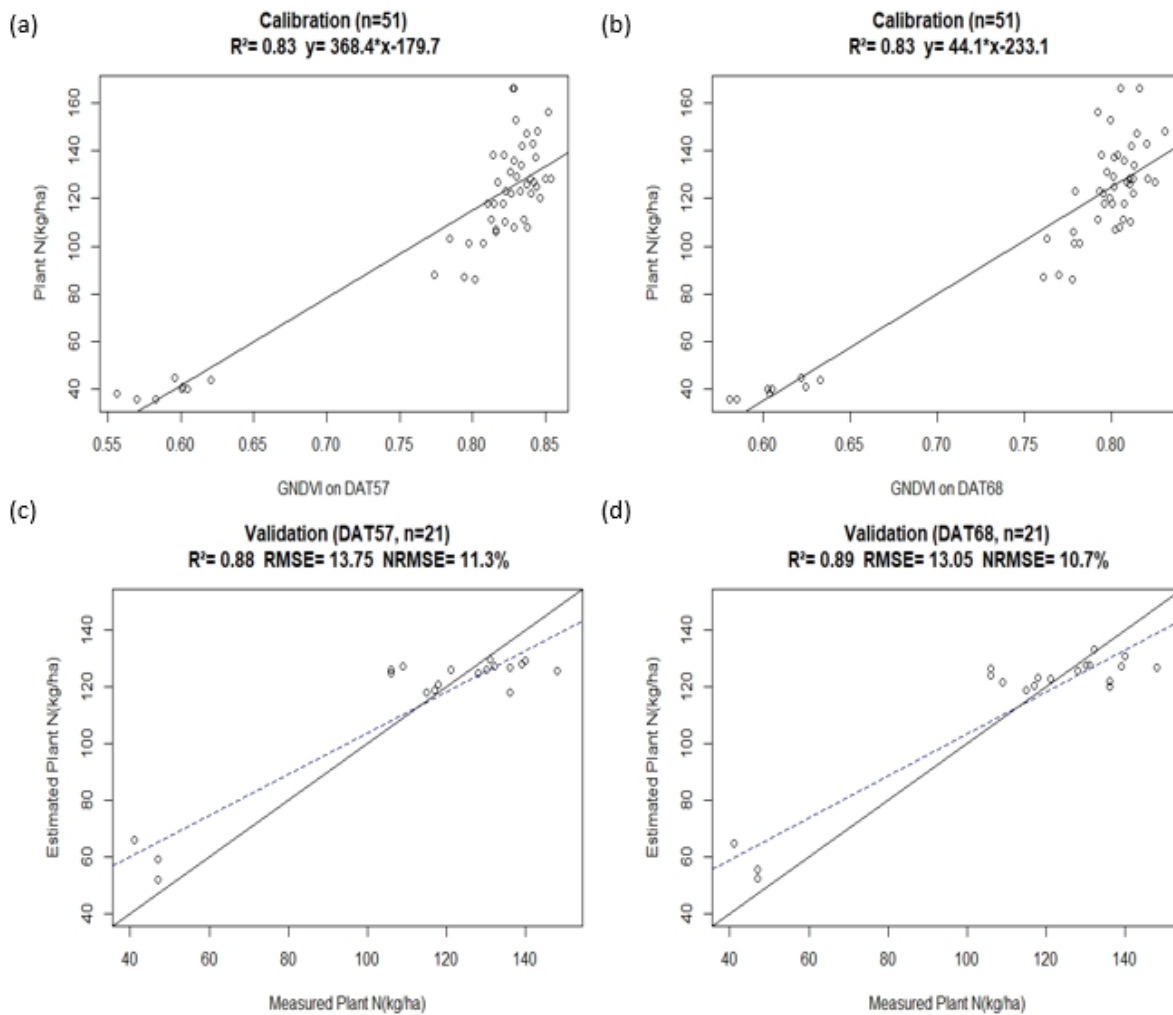
Figure 12: Relationships between measured and estimated PNA using GNDVI-based linear regression models. Calibration (n=51) and validation (n=21) plots for PNA estimation in DAT 57 (a)(c) and DAT 68 (b)(d) in 2017DS.

In addition to using simple linear regression on critical days for N status estimation, stepwise multiple linear regression models between GNDVI time-series data and PNA were also applied in 2016EWS (Figure 13a) and 2017DS (Figure 13b). The p value and F-test for each model describes that the model performs well for PNA estimation. It also shows that the GNDVI in different dates as input variables are significant for the model according to the p-value which are closed to zero. Figure 14 shows that the measured PNA and estimated PNA are highly correlated ($R^2 > 0.75$) in both seasons.

Comparing the simple linear regression models with stepwise multiple linear regression models, they have similar R^2, RMSE, NRMSE result which means that using GNDVI on a single specific day is already sufficient for PNA estimation in rice-growing season.

(a)

```
MODEL INFO:
Observations: 51
Dependent Variable: Plant.N..kg.ha.
Type: OLS linear regression

MODEL FIT:
F(4,46) = 49.83, p = 0.00
R² = 0.81
Adj. R² = 0.80

Standard errors: OLS
------------------------------------------------
                    Est.    S.E.   t val.    p
------------------------------------------------
(Intercept)      -227.41   41.92   -5.43   0.00
GNDVI2016_06_08   291.74  119.43    2.44   0.02
GNDVI2016_06_28   442.38  108.29    4.09   0.00
GNDVI2016_07_05  -491.35  139.08   -3.53   0.00
GNDVI2016_08_03   240.09   59.62    4.03   0.00
------------------------------------------------
```

(b)

```
MODEL INFO:
Observations: 51
Dependent Variable: Plant.N..kg.ha.
Type: OLS linear regression

MODEL FIT:
F(4,46) = 115.51, p = 0.00
R² = 0.91
Adj. R² = 0.90

Standard errors: OLS
------------------------------------------------
                    Est.    S.E.   t val.    p
------------------------------------------------
(Intercept)      -251.88   22.70  -11.09   0.00
GNDVI2017_01_13   -91.63   43.51   -2.11   0.04
GNDVI2017_02_16   480.02   93.19    5.15   0.00
GNDVI2017_02_23  -138.25   75.30   -1.84   0.07
GNDVI2017_04_11   252.63   45.50    5.55   0.00
------------------------------------------------
```

Figure 13: Stepwise multiple linear regression models performance using different days during the rice growth season for PNA estimation in 2016EWS(a) (n=51) and 2017DS (b) (n=51).
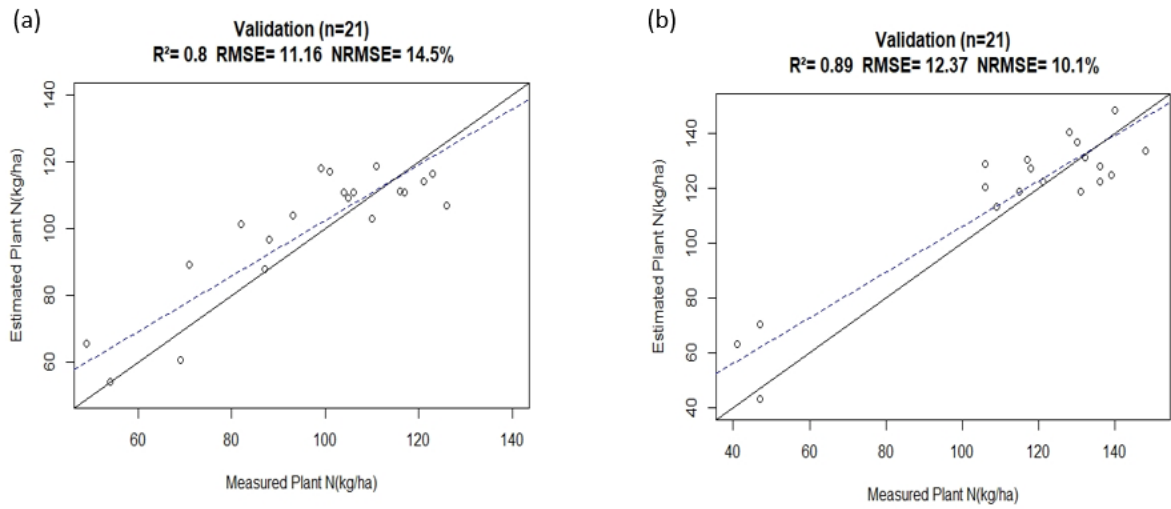

Figure 14: Relationships between measured and estimated PNA using stepwise multiple linear regression models based on UAV-derived GNDVI time-series data for estimating PNA at harvest time in 2016EWS (a) and 2017DS (b).

4.4. Multivariate methods and machine learning algorithms comparison for Nitrogen status estimation

Different methods were applied in the study for PNA estimation, so the model quality and accuracy results were judged in this section for the best model selection.

Tuning parameter is always the most important step for training model. Figure 15 shows the PLSR model parameter selection, the optimal model which the ncomp equals five with the smallest RMSE was selected in both seasons. As for the SVR model tuning, Figure 16 presents important parameters for the best SVR model after grid search. It could be noticed that the kernel type of the selected SVR model is radial which is non-linear method. The RF algorithm parameter tuning result (Figure 17) shows that when the mtry value equals to 47 and 11 in the model, the RMSE value is the lowest.
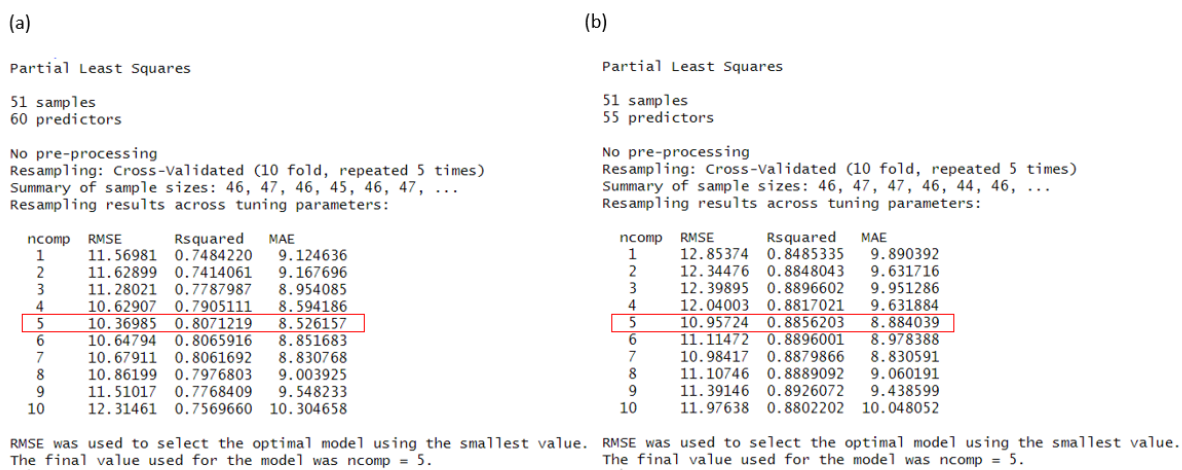

Figure 15: PLSR algorithm for parameter tuning using 10-fold cross-validation and grid search in 2016EWS(a) and 2017DS(b).
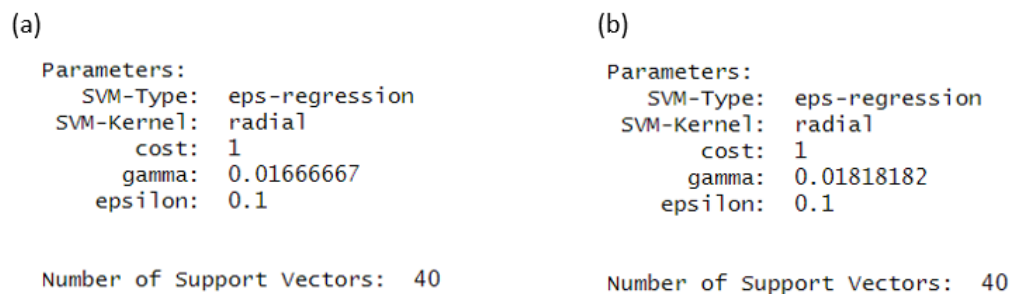

Figure 16: SVR algorithm for parameter tuning using 10-fold cross validation and grid search in 2016EWS (a) and 2017DS(b).

(a)

```
Random Forest

51 samples
60 predictors

No pre-processing
Resampling: Cross-Validated (10 fold, repeated 3 times)
Summary of sample sizes: 46, 47, 47, 44, 45, 46, ...
Resampling results across tuning parameters:

  mtry  RMSE      Rsquared   MAE
   5    11.31414  0.7087852  9.250832
   9    11.34956  0.7073095  9.275514
  13    11.26477  0.7054822  9.179120
  16    11.10584  0.7098492  9.007282
  21    11.06504  0.7143494  9.008061
  22    11.15882  0.7097951  9.113258
  23    11.20696  0.7012546  9.143399
  24    11.14436  0.7108706  9.043169
  25    11.13692  0.7058600  9.066260
  27    11.12923  0.7072945  9.067592
  28    11.12137  0.7069941  9.078751
  30    11.13094  0.7104916  9.090981
  47    10.98137  0.7077780  8.874854
  59    11.00821  0.7132441  8.871177

RMSE was used to select the optimal model using the smallest value.
The final value used for the model was mtry = 47.
```

(b)

```
Random Forest

51 samples
55 predictors

No pre-processing
Resampling: Cross-Validated (10 fold, repeated 5 times)
Summary of sample sizes: 47, 45, 47, 46, 46, 46, ...
Resampling results across tuning parameters:

  mtry  RMSE      Rsquared   MAE
  10    12.09211  0.8491787  9.365617
  11    12.08119  0.8493948  9.329681
  12    12.08835  0.8526707  9.411106
  23    12.20084  0.8467613  9.439572
  24    12.19085  0.8436189  9.437445
  30    12.36511  0.8380771  9.555155
  32    12.27272  0.8410930  9.451082
  33    12.37266  0.8358201  9.558550
  34    12.33525  0.8398704  9.552994
  39    12.35806  0.8393946  9.553282
  42    12.45520  0.8338230  9.608905
  43    12.43199  0.8327032  9.610413
  47    12.48795  0.8306883  9.618979
  53    12.51515  0.8308135  9.626880

RMSE was used to select the optimal model using the smallest value.
The final value used for the model was mtry = 11.
```
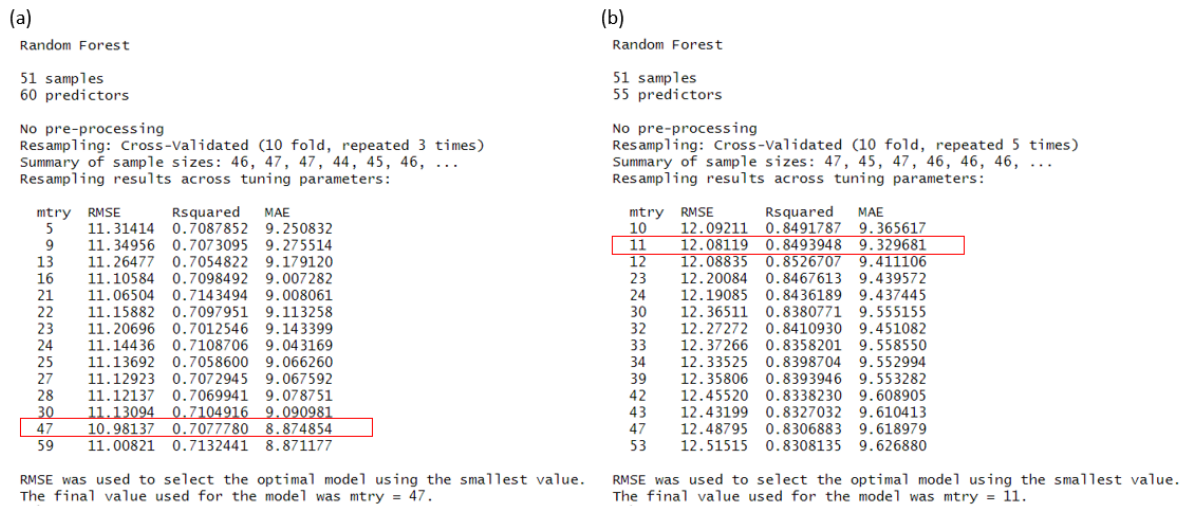
Figure 17: RF algorithm for parameter tuning using 10-fold cross-validation and grid search in 2016EWS (a) and 2017DS(b).

With the optimal model selection for different methods, the model performance was assessed in the following section. The importance of input variables was firstly ranked for PLSR (Figure 18) and SVR (Figure 19) models. Comparing the important variables at the top of the list in same season, there is similarity using PLSR and RF models for PNA estimation. In 2016EWS, the GNDVI on June 28th and the NIR band reflectance on July 22nd account for large promotion in both PLSR and RF models. As for the calibration model in 2017DS, the red and green band reflectance on March 2nd are highly important for both algorithms. Moreover, there is also difference between the two models, such as the number of highly important variables in PLSR model are more than those in RF models.
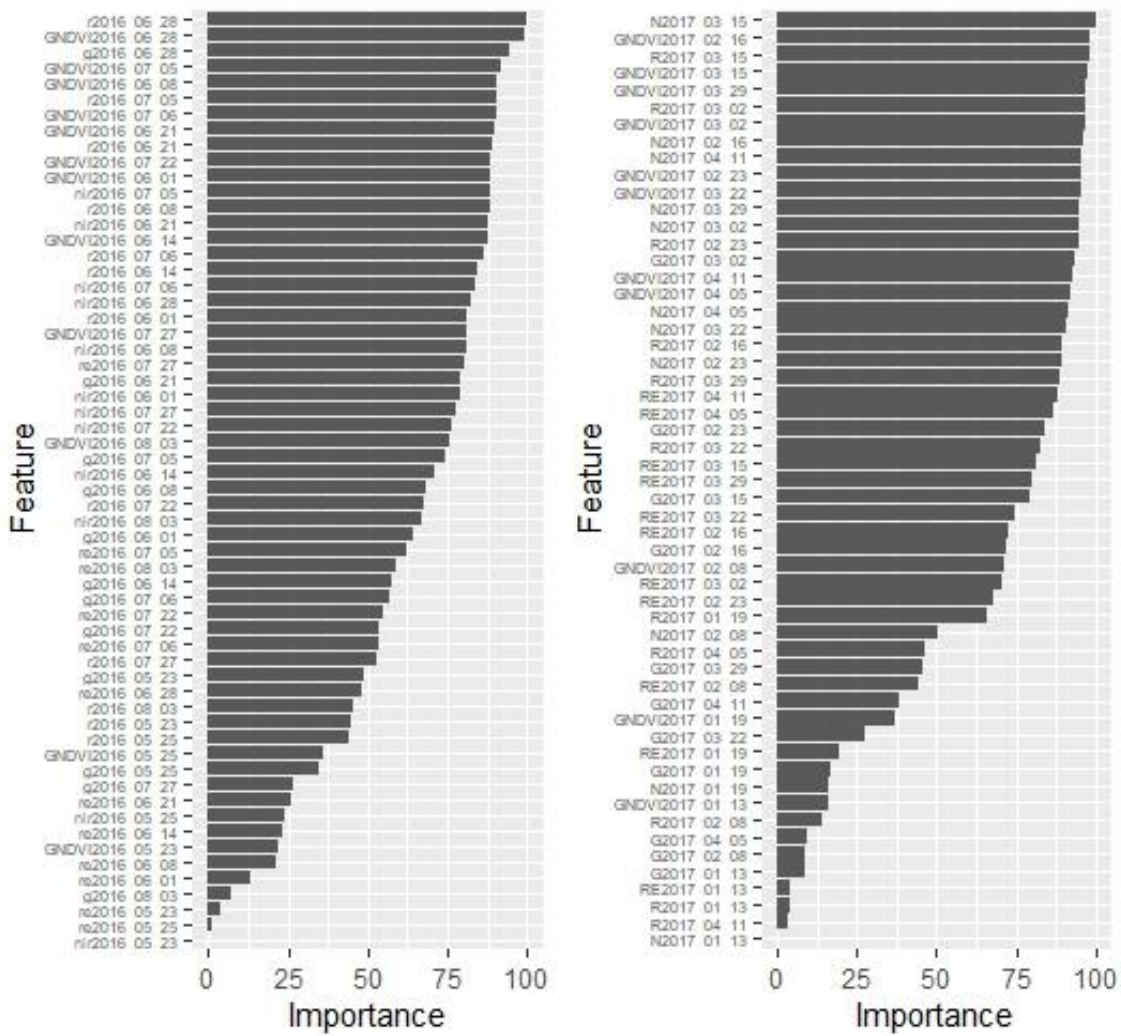
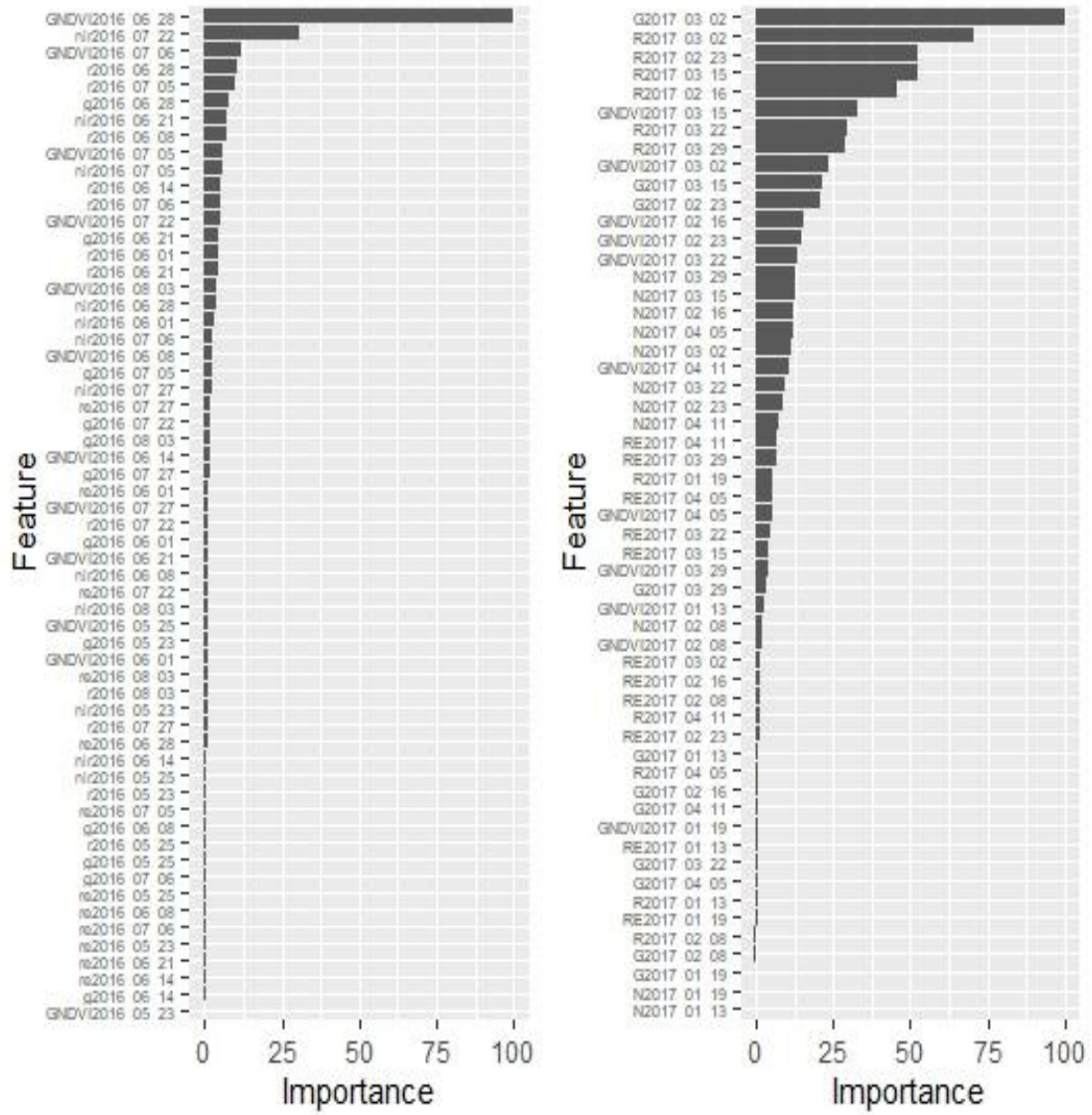Figure 18: Feature importance of PLSR model in 2016EWS(a) and 2017DS(b).

Figure 19: Feature importance of RF model in 2016EWS(a) and 2017DS(b).

The relationship between estimated and measured PNA using PLSR, SVR and RF are shown in Figure 20. Models performance were assessed by R², RMSE and NRMSE in this session. In 2016EWS (Figure 12a-c), the best performing model is RF in Figure 3c (R²= 0.9, RMSE=8.37, NRMSE=10.9%), which shows a higher correlation and lower error in between estimated and measured PNA. Then, the SVR (Figure 12b) also explains a strong relationship between prediction and measurement. In 2017DS (Figure 12d-f), the RF is still the most representative algorithm for N status prediction (R²= 0.93, RMSE=9.93, NRMSE=8.1%). PLSR model with more than 0.9 correlation between the measured and estimated value (R²= 0.91, RMSE=10.65, NRMSE=8.7%) also performs well.
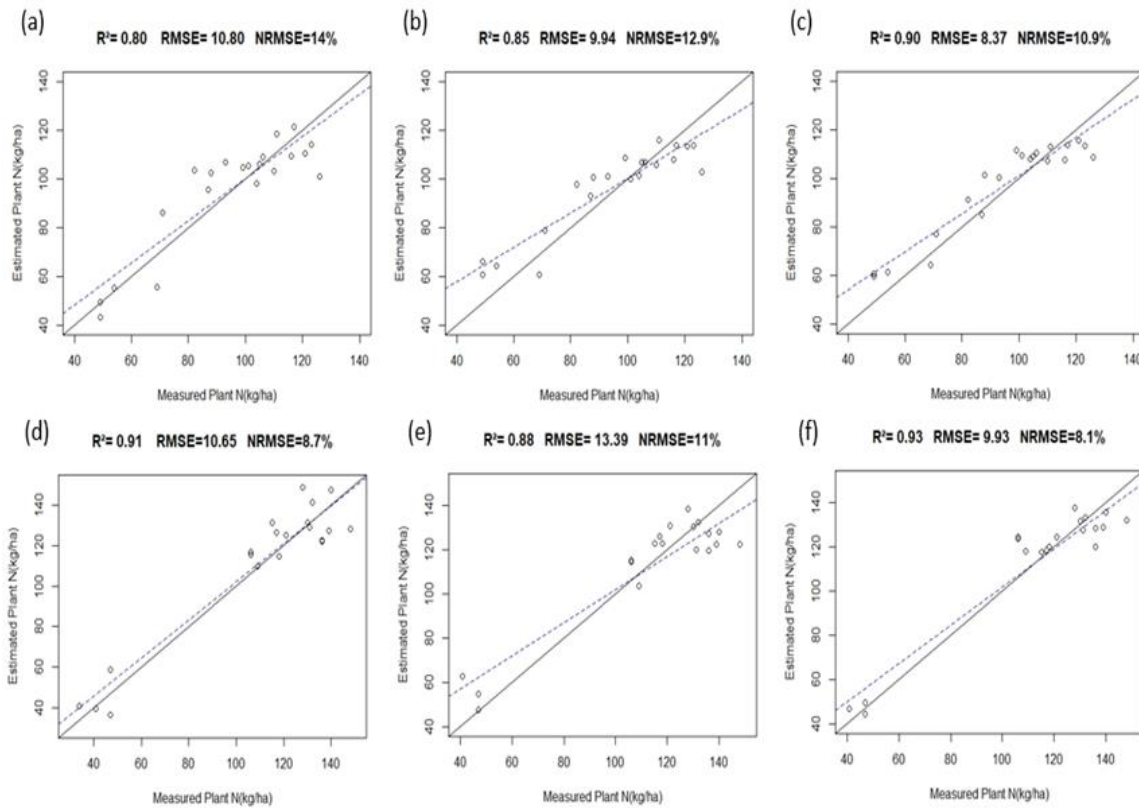

Figure 20: Relationships between measured PNA and estimated PNA using PLSR, SVM and RF in 2016EWS (a-c) and 2017DS (d-f) based on testing dataset (n=21).

Table 11 and Table 12 show all methods for PNA estimation performance in both seasons. Comparing with linear regression models with a single vegetation index, the results have improved using machine learning algorithms with higher correlation and lower error between measured PNA and estimated PNA. The RF is always the optimal algorithm in both wet and dry season.

Table 11: R², RMSE and NRMSE for all prediction methods (SR, SMLR, PLSR, SVR and RF) used to estimate PNA of rice crop in 2016EWS.

	2016EWS				
	SR	SMLR	PLSR	SVR	RF
R²	0.81	0.8	0.8	0.85	0.9
RMSE	10.4	11.16	10.8	9.94	8.37
NRMSE (%)	13.5	14.5	14	12.9	10.9

Table 12: R², RMSE and NRMSE for all prediction methods (SR, MLR, PLSR, SVR and RF) used to estimate PNA of rice crop in 2017DS.

	2017DS				
	SR	SMLR	PLSR	SVR	RF
R²	0.89	0.89	0.91	0.88	0.93
RMSE	13.05	12.37	10.65	13.39	9.93
NRMSE (%)	10.7	10.1	8.7	11	8.1

4.5. Mapping N status for 2016 wet season and 2017 dry season

Figure 21a and Figure 22a show the estimated PNA maps in subplots using the RF as the most accurate method in 2016EWS and 2017DS with four classes. To compare the difference between the estimated and measured PNA, difference maps of PNA were generated using the estimated PNA minus measured PNA in 2016EWS and 2017DS (Figure 23). In order to compare the difference between two seasons, the difference percentage was calculated using the difference value divided by measured PNA for two seasons The difference percentage value was classified into four categories to assess the estimation result, the closer the percentage is to zero, the better the estimation result.
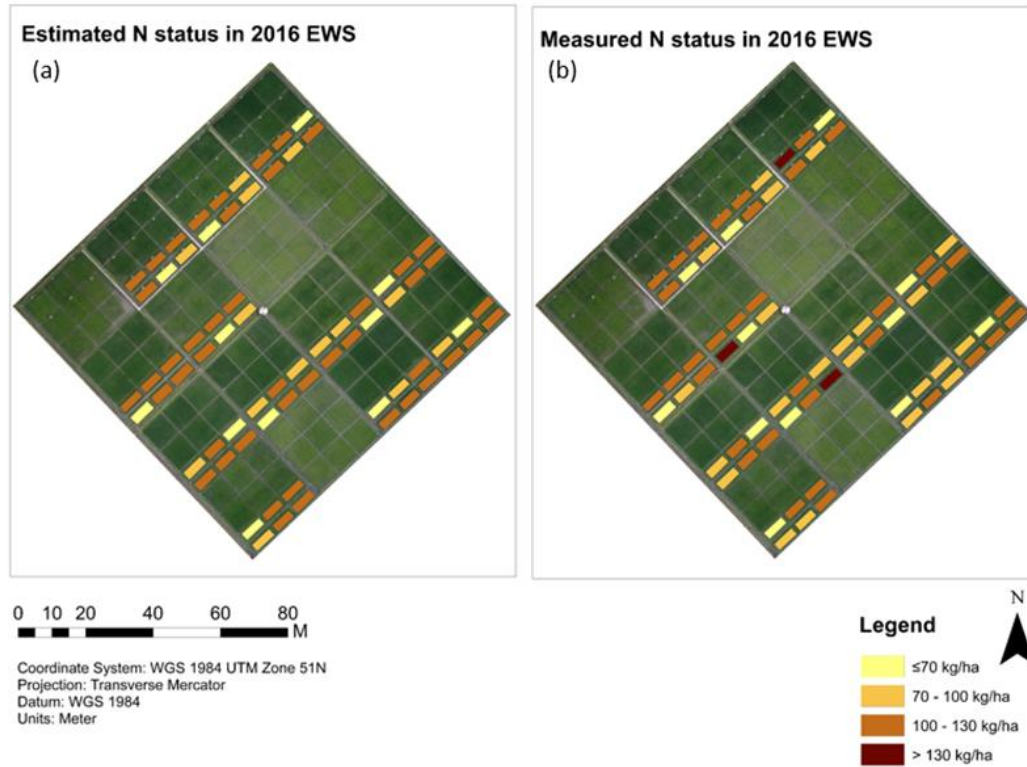

Figure 21: Estimated N status map(a) and measured N status map(b) on subplots on August 11th (DAT99), 2016 on LTCCE field in the IRRI Zeigler Experiment Station, Los Baños, the Philippines.
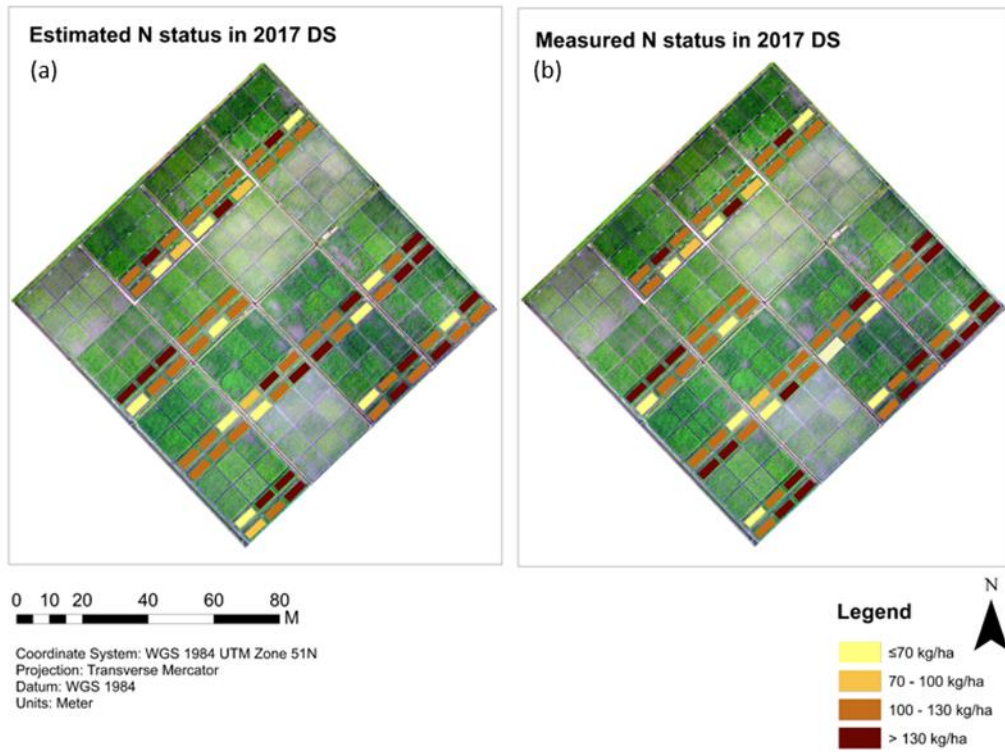
Figure 22: Estimated N status map(a) and measured N status map(b) for subplots on April 18th (DAT104), 2017 on LTCCE field in the IRRI Zeigler Experiment Station, Los Baños, the Philippines.
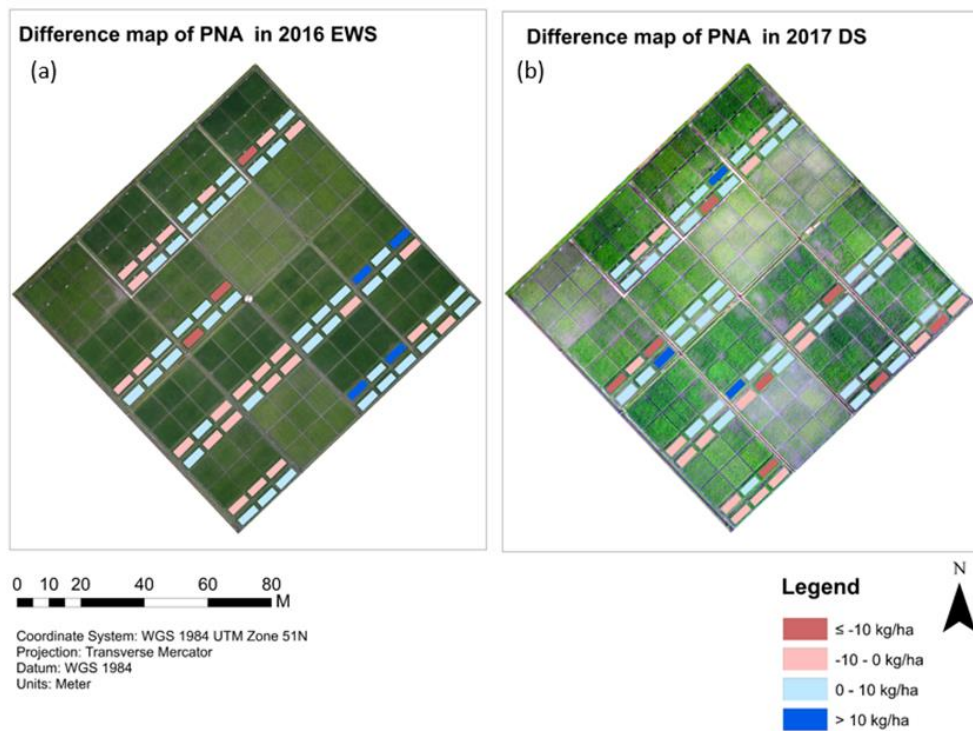

Figure 23: Difference map between estimated PNA and measured PNA in 2016EWS(a) and 2017DS(b) on LTCCE field in the IRRI Zeigler Experiment Station, Los Baños, the Philippines.
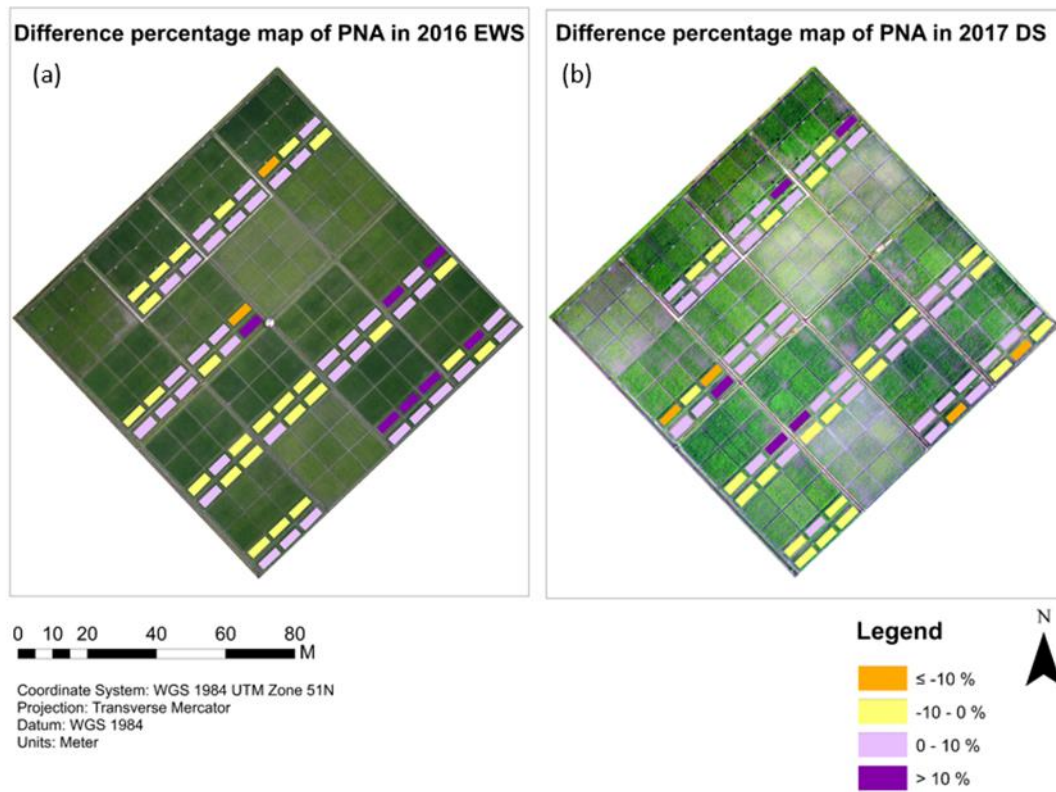
Figure 24: Difference percentage maps between estimated PNA and measured PNA in 2016EWS(a) and 2017DS(b) on LTCCE field in the IRRI Zeigler Experiment Station, Los Baños, the Philippines.
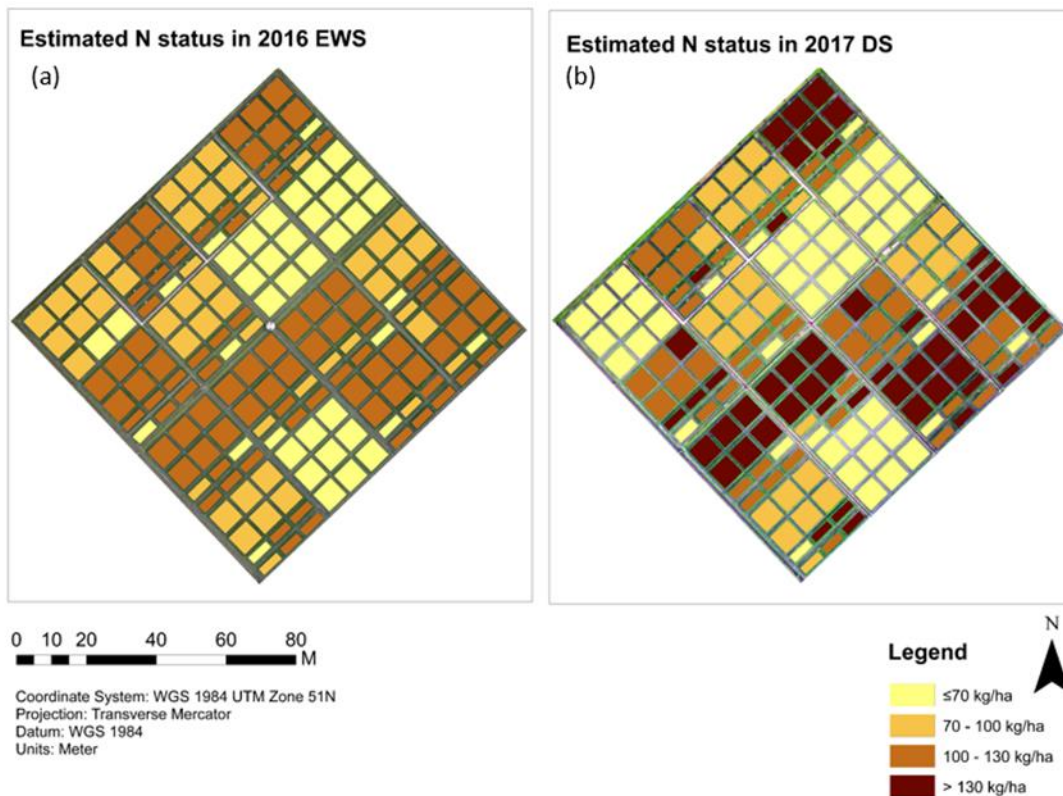

Figure 25: Estimated rice PNA maps for LTCCE field in the IRRI Zeigler Experiment Station, Los Baños, the Philippines at harvest time in 2016EWS(a) and 2017DS(b).

5.   DISCUSSION

The study aimed to accurately estimate nitrogen accumulation (PNA) of rice crops using time-series UAV multispectral images obtained for two growing seasons in IRRI Zeigler Experiment Station, Los Baños, the Philippines. Vegetation indices (VIs), linear regression, multivariate models (stepwise and partial least square regressions (PLSR)) as well as machine learning algorithms, including support vector regression (SVR) and random forest (RF) were used to estimate PNA of rice crops for 2016 early wet season (EWS) and 2017 dry season (DS). The model with the highest accuracy (lowest RMSE and highest R^2) was used for mapping PNA for both rice seasons. The results showed that the RF model had the highest accuracy in estimating PNA on rice and therefore was used for mapping PNA for both growing seasons.

5.1.   Relationships among filed agronomic parameters

The UAV imagery and field data offered by Dr Roland Buresh and Steve Klassen from IRRI organization were well organized. Strong correlations were found among the SPAD values, leaf colour chart readings (LCC) and PNA in the rice panicle initiation and heading stages. Suresh et al. (2017) also showed that SPAD and LCC measured in rice crops 56 days after transplanting (DAT) on rice had a significant and positive correlation with leaf N content on rice.

5.2.   Linear regression models using GNDVI for PNA estimation

Correlation coefficients between the time-series VIs and PNA were calculated for selection of the best performing VI in 2016EWS and 2017DS, and the dates in which PNA and VIs had highest correlations were selected as the most crucial days (rice panicle initiation and heading stages) for PNA estimation in Table 8 and Table 9. Among the studied VIs, GNDVI always had the highest correlation coefficient with SPAD, CCC and PNA during the whole growing season. Earlier findings by Gitelson et al. (1996) showed that GNDVI is useful for assessing the canopy chlorophyll content (CCC) in rice crops. The CCC has been confirmed as an indicator for reflecting the N content on winter wheat by Delloye et al. (2018).

As can be observed from results in Table 8 to Table 10, 56 and 63 DAT in 2016EWS, as well as 57 and 68 DAT in 2017DS are crucial dates for PNA estimation, therefore, simple linear regression models were built using GNDVI calculated for those days and field measurements of PNA for both 2016EWS and 2017DS (Figure 10). The R^2 value and RMSE (R^2=0.78-0.89, RMSE=10.4-13.75, NRMSE=10.7%-14.9%) of simple regression models illustrated the significance of using GNDVI in crucial days for PNA estimation. The feasibility of using VI in regression models for N status estimation has been shown in many studies (Noureldin et al., 2013; Maki et al., 2014). In addition, the stepwise multiple linear regression models were built using GNDVI time-series data to check if the PNA estimation would improve using more dates GNDVI data in the regression models (Figure 13). We found that the model performance of stepwise multiple linear regression model is quite similar to the simple linear regression model with R^2=0.8 and RMSE=11.16 in 2016EWS, R^2= 0.89 and RMSE=12.37 in 2017DS (Table 11 and Table 12).

5.3. Estimating and mapping PNA of rice crops using multivariate methods and machine learning models on rice in 2016EWS and 2017DS

Comparing multivariate method (PLSR) and machine learning algorithms (SVR, RF) performance for PNA estimation was an important part of this research. In general, machine learning algorithms improved PNA estimation accuracy compared to simple linear and stepwise regression models (Table 11 and 12). The PLSR models for PNA estimation performed well in both rice seasons (R^2=0.8-0.91. RMSE=10.65-10.8, NRMSE=8.7%-14%). In PLSR models, the more important input variables were the red band reflectance and GNDVI value in panicle initiation stage (Figure 18). Most of the input variables importance were larger than 50% in PLSR models. As for SVR models, it also showed good performance for PNA estimation (R^2=0.85-0.88. RMSE=9.94-13.39, NRMSE=11%-12.9%). Maimaitijiang et al. (2020) proved that SVR model has great potential in predicting crop above ground biomass and LAI using UAV data. The RF models performed as the most accurate one among all studied machine learning algorithms in both the 2016EWS and 2017DS (R^2=0.9-0.93, RMSE=8.37-9.93, NRMSE=8.1%-10.9%). Relatively more important features in the RF models (Figure 19) were the red band reflectance and GNDVI value in panicle initiation stage, which shows similarity to the PLSR models, but the features in RF models with high importance percentage is less than in PLSR models. The capability of using RF models to estimate N status on rice using GNDVI and individual bands reflectance is also demonstrated in the study by Sun et al. (2017). In recent years, the RF algorithm has gradually become popular for different applications such as crop parameter estimation, due to the stability with a varying number of base learners (Genuer et al., 2010). As such, Wang et al. (2016) showed that using different VIs derived from China's environmental satellite (HJ) charged-coupled device (CCD) images in RF regression models produced the most accurate estimates of wheat biomass in comparison with SVR and artificial neural network (ANN).

As can be observed from the PNA map generated by RF models (Figure 21a and Figure 22b), PNA values larger than 130g/ha do not exist in 2016EWS, which means the PNA value in 2017DS was generally higher than 2016EWS. It can also be observed from the estimated and measured PNA maps in Figure 21 and Figure 22 that PNA is more affected by N application and less by the varietal differences. The subplots with no N treatment always presented a PNA value less than 70kg/ha. Then, PNA started to increase when the total N application is 90kg/ha in 2016EWS and 130kg/ha in 2017DS. Since the total N application amount in 2017DS was higher than 2016DS, the number of subplots that had more than 130kg/ha of PNA in 2017EWS were higher than in 2016 DS. The PNA difference (residual) maps showed that for most of the subplots, the difference between the estimated and measured PNA value were closed to zero (Figure 23), which shows the reliability of the estimated PNA values using RF in both seasons. According to difference percentage maps for both seasons, the residuals between most measured and estimated subplots were less than 10% (Figure 24). Besides the comparison between the measured and estimated PNA, the RF models were further applied to the whole LTCCE site to map the PNA distribution (Figure 25). The produced PNA maps could be used for better N treatment and for actual field instruction.

5.4. Recommendations

Field measurements were presented as an average value at the subplots level, however, smaller-scale field measurements based on rice hills level could be taken inside the subplots for more accurate measurements. Due to the very high resolution of UAV images, there is also scope for further improvement on PNA estimation, for example, detailed PNA estimation based on rice hills level could be studied. As such, the variation inside the subplots could be studied.

The PNA destructive measurements were only made one time at the harvest time, it is suggested to take multiple measurements during the whole growing season same as other non-destructive methods. As such, the relationships between destructive N measurements and non-destructive measurements in different rice-growing stages could be investigated. The PNA value could also been estimated in each rice growth stage instead of just once at the harvest time by applying multivariate models.

The study only used linear and stepwise regression models, PLSR, SVR and RF algorithms for the selection of the best model for PNA estimation. More algorithms could be examined to select a convinced and robust algorithm in future works. Moreover, different scenarios could be tested by applying different inputs in machine learning algorithms. For examples, different VI or derivatives of reflectance can be used as input to train the models. The remotes sensing data combining with different field measurements such as the time-series SPAD value could also been tested as model input for PNA estimation. Based on the results of linear and stepwise regressions, it is suggested that instead of using the remote sensing data and VI for the whole rice-growing season, only selecting the crucial day's data as input could be sufficient for PNA estimation in the future studies.

6. CONCLUSION

In this research, the potential of UAV images was examined in agricultural fields experiments and the application for larger-scale agricultural management. The use of traditional vegetation indices to estimate PNA were evaluated for different seasons. Multivariate methods and machine learning algorithms including PLSR, SVR and RF were also evaluated and compared for PNA estimation. The PNA of rice crops under dry and wet seasons were finally mapped using the most accurate machine learning algorithm at the end of the study.

The field measurements of SPAD, LCC, and PNA are highly correlated during the rice panicle initiation and heading stages. Different VIs related to structure and chlorophyll is selected for rice PNA estimation at the harvest time. All the chosen VIs indicated good correlations with PNA in both growing seasons. The best performance VI among all VIs was the Green Normalized Difference Vegetation Index (GNDVI). It was used in simple linear regression models using crucial days (DAT 56 and DAT63 in 2016EWS, DAT57 and DAT68 in 2017DS) to estimate PNA is suitable and convincing in both rice-growing seasons. In addition, the time-series GNDVI data for the whole growing season were applied in the stepwise linear regression models for PNA estimation. Moreover, multivariate methods and machine learning algorithms (PLSR, SVR and RF) were applied using individual bands reflectance and GNDVI time-series data for PNA estimation. Among different machine learning algorithms in this study, the RF as the most accurate model (determined by R^2, RMSE) was used for rice PNA mapping in 2016EWS and 2017DS. Further efforts can be made for more accurate N status estimation, including applying the RF model in rice hill level to examine the N variation inside each subplot.

LIST OF REFERENCES

Ahmed, M., Rauf, M., Mukhtar, Z., & Saeed, N. A. (2017). Excessive use of nitrogenous fertilizers: an unawareness causing serious threats to environment and human health. *Environmental Science and Pollution Research*, *24*(35), 26983–26987. https://doi.org/10.1007/s11356-017-0589-7

Aho, K., Derryberry, D., & Peterson, T. (2014). Model selection for ecologists: The worldviews of AIC and BIC. *Ecology*. Ecological Society of America. https://doi.org/10.1890/13-1452.1

Almergren, J., Alves, T. P., Anderssen, E., & Araújo, M. C. U. (2009). *Comprehensive Chemometrics: Chemical and Biochemical Data Analysis*. (S. D. Brown, R. Tauler, & B. Walczak, Eds.). Retrieved from https://www.sciencedirect.com/referencework/9780444527011/comprehensive-chemometrics

Awad, M., Khanna, R., Awad, M., & Khanna, R. (2015). Support Vector Regression. In *Efficient Learning Machines* (pp. 67–80). Apress. https://doi.org/10.1007/978-1-4302-5990-9_4

Bansal, R., Gaur, N., & Singh, S. N. (2016). Outlier Detection: Applications and techniques in Data Mining. In *Proceedings of the 2016 6th International Conference - Cloud System and Big Data Engineering, Confluence 2016* (pp. 373–377). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/CONFLUENCE.2016.7508146

Bijay, S., & Yadvinder, S. (2002). Chlorophyll Meter– and Leaf Color Chart–Based Nitrogen Management for Rice and Wheat in Northwestern India. *Agronomy Journal*. Retrieved from https://www.academia.edu/6760580/Chlorophyll_Meter_and_Leaf_Color_Chart_Based_Nitrogen _Management_for_Rice_and_Wheat_in_Northwestern_India

Blackmer, T. M., Schepers, J. S., Varvel, G. E., & Walter-Shea, E. A. (1996). Nitrogen deficiency detection using reflected shortwave radiation from irrigated corn canopies. *Agronomy Journal*, *88*(1), 1–5. https://doi.org/10.2134/agronj1996.00021962008800010001x

Brinkhoff, J., Dunn, B. W., Robson, A. J., Dunn, T. S., & Dehaan, R. L. (2019). Modeling Mid-Season Rice Nitrogen Uptake Using Multispectral Satellite Data. *Remote Sensing*, *11*(15), 1837. https://doi.org/10.3390/rs11151837

Byju, G., & Haripriya Anand, M. (2009). Leaf color chart and chlorophyll-meter-based leaf nitrogen estimation and their threshold values for real-time nitrogen management in cassava. *Communications in Soil Science and Plant Analysis*, *40*(17–18), 2816–2832. https://doi.org/10.1080/00103620903173830

Cao, Q., Miao, Y., Wang, H., Huang, S., Cheng, S., Khosla, R., & Jiang, R. (2013). Non-destructive estimation of rice plant nitrogen status with Crop Circle multispectral active canopy sensor. *Field Crops Research*, *154*, 133–144. https://doi.org/10.1016/J.FCR.2013.08.005

Carneiro, F. M., Furlani, C. E. A., Zerbato, C., Menezes, P. C. de, Gírio, L. A. da S., Carneiro, F. M., … Gírio, L. A. da S. (2019). CORRELATIONS AMONG VEGETATION INDICES AND PEANUT TRAITS DURING DIFFERENT CROP DEVELOPMENT STAGES. *Engenharia Agrícola*, *39*(SPE), 33–40. https://doi.org/10.1590/1809-4430-eng.agric.v39nep33-40/2019

Cawley, G. C., & Talbot, N. L. C. (2010). On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research*, *11*, 2079–2107.

Chen, P. F., Tremblay, N., Wang, J. H., Vigneault, P., Huang, W. J., & Li, B. G. (2010). New index for crop canopy fresh biomass estimation. *Guang Pu Xue Yu Guang Pu Fen Xi/Spectroscopy and Spectral Analysis*, *30*(2), 512–517. https://doi.org/10.3964/j.issn.1000-0593(2010)02-0512-06

Chlingaryan, A., Sukkarieh, S., & Whelan, B. (2018). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Computers and Electronics in Agriculture*, *151*, 61–69. https://doi.org/10.1016/j.compag.2018.05.012

Conforti, P. (2011). *Looking ahead in worLd food and agricuLture: Perspectives to 2050*. (P. Conforti, Ed.).

Dabi, T., & Khanna, V. (2018). Effect of Climate Change on Rice. *Agrotechnology*, 7. https://doi.org/10.4172/2168-9881.1000181

Delloye, C., Weiss, M., & Defourny, P. (2018). Retrieval of the canopy chlorophyll content from Sentinel-2 spectral bands to estimate nitrogen uptake in intensive winter wheat cropping systems. *Remote Sensing of Environment*, *216*, 245–261. https://doi.org/10.1016/j.rse.2018.06.037

Drucker, H., Burges, C. J. C., Kaufman, L., Smola, A., & Vapoik, V. (n.d.). *Support Vector Regression Machines*.

Duy, T. K., & Huyen, T. T. N. (2019). A Rice Leaf Color Chart using Low-Cost Visible Spectro Sensor. *European Journal of Agronomy*, *8*(7C2), 27–30. https://doi.org/10.1016/s1161-0301(14)80141-x

FAO. (2017). *THE STATE OF FOOD SECURITY AND NUTRITION IN THE WORLD*. United Nations. Retrieved from http://www.fao.org/3/a-i7695e.pdf

Farifteh, J., Van der Meer, F., Atzberger, C., & Carranza, E. J. M. (2007). Quantitative analysis of salt-affected soil reflectance spectra: A comparison of two adaptive methods (PLSR and ANN). *Remote Sensing of Environment*, *110*(1), 59–78. https://doi.org/10.1016/j.rse.2007.02.005

Geladi, P., & Kowalski, B. R. (1986). Partial least-squares regression: a tutorial. *Analytica Chimica Acta*, *185*(C), 1–17. https://doi.org/10.1016/0003-2670(86)80028-9

Genuer, R., Poggi, J.-M., & Tuleau-Malot, C. (2010). *Variable selection using Random Forests*. Retrieved from http://www.r-project.org/

Gitelson, A. A. (2005). Remote estimation of canopy chlorophyll content in crops. *Geophysical Research Letters*, *32*(8), L08403. https://doi.org/10.1029/2005GL022688

Gitelson, A. A., Kaufman, Y. J., & Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS- MODIS. *Remote Sensing of Environment*, *58*(3), 289–298. https://doi.org/10.1016/S0034-4257(96)00072-7

Glass, A. D. M. (2003). Nitrogen Use Efficiency of Crop Plants: Physiological Constraints upon Nitrogen Absorption. *Critical Reviews in Plant Sciences*, *22*(5), 453–470. https://doi.org/10.1080/07352680390243512

Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., … Toulmin, C. (2010). Food Security: The Challenge of Feeding 9 Billion People. *Science*, *327*(5967), 812–818. https://doi.org/10.1126/science.1185383

Halinski, R., & Feldt, L. (n.d.). The Selection of Variables in Multiple Regression Analysis. *National Council on Measurement in Education*.

Hijmans, R. J., & Etten, J. van. (2020). *Package "raster."* Retrieved from https://github.com/rspatial/raster/issues/

Ho, T. K. (1995). Random decision forests. *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*, *1*, 278–282. https://doi.org/10.1109/ICDAR.1995.598994

Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(8), 832–844. https://doi.org/10.1109/34.709601

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

IRRI. (n.d.). Importance of Rice. Retrieved October 20, 2019, from http://www.knowledgebank.irri.org/ericeproduction/Importance_of_Rice.htm

Ju, X.-T., Xing, G.-X., Chen, X.-P., Zhang, S.-L., Zhang, L.-J., Liu, X.-J., … Zhang, F.-S. (2009). Reducing environmental risk by improving N management in intensive Chinese agricultural systems. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(9), 3041–3046. https://doi.org/10.1073/pnas.0813417106

Kim, N., & Lee, Y. W. (2016). Machine learning approaches to corn yield estimation using satellite images and climate data: A case of Iowa State. *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, *34*(4), 383–390. https://doi.org/10.7848/ksgpc.2016.34.4.383

Kuhn, M., Wing, J., & Weston, S. (2020). *Package "caret."*

Lang, C. A. (1958). Simple Microdetermination of Kjeldahl Nitrogen in Biological Materials. *Analytical Chemistry*, *30*(10), 1692–1694. https://doi.org/10.1021/ac60142a038

Liaw, A., & Wiener, M. (2002). *Classification and Regression by randomForest* (Vol. 2). Retrieved from http://www.stat.berkeley.edu/

Maimaitijiang, M., Sagan, V., Sidike, P., Daloye, A. M., Erkbol, H., & Fritschi, F. B. (2020). Crop Monitoring Using Satellite/UAV Data Fusion and Machine Learning. *Remote Sensing, 12*(9), 1357. https://doi.org/10.3390/rs12091357

Maki, M., & Homma, K. (2014). Empirical Regression Models for Estimating Multiyear Leaf Area Index of Rice from Several Vegetation Indices at the Field Scale. *Remote Sensing, 6*, 4764–4779. https://doi.org/10.3390/rs6064764

Manglapus, R. S. (1974). *The Philippine Archipelago. Worldview* (Vol. 17). https://doi.org/10.1017/s0084255900032721

Meyer, D., Dimitriadou, E., & Hornik, K. (2019). *Package "e1071."*

Mitchell, T., & Hill, M. (1997). *Machine Learning*. Retrieved from http://www.cs.cmu.edu/~tom/mlbook.html

Muñoz-Huerta, R. F., Guevara-Gonzalez, R. G., Contreras-Medina, L. M., Torres-Pacheco, I., Prado-Olivarez, J., & Ocampo-Velazquez, R. V. (2013). A review of methods for sensing the nitrogen status in plants: Advantages, disadvantages and recent advances. *Sensors (Switzerland)*, *13*(8), 10823–

10843. https://doi.org/10.3390/s130810823

Noureldin, N. A., Aboelghar, M. A., Saudy, H. S., & Ali, A. M. (2013). Rice yield forecasting models using satellite imagery in Egypt. *Egyptian Journal of Remote Sensing and Space Science*, *16*(1), 125–131. https://doi.org/10.1016/j.ejrs.2013.04.005

Peng, S., Garcia, F. V., Laza, R. C., Sanico, A. L., Visperas, R. M., & Cassman, K. G. (1996). Increased N-use efficiency using a chlorophyll meter on high-yielding irrigated rice. *Field Crops Research*, *47*(2–3), 243–252. https://doi.org/10.1016/0378-4290(96)00018-4

Philippine Statistics Authority. (2019). *RICE and CORN Situation and Outlook*.

Randolph Barker, Robert W. Herdt, & Beth Rose. (1985). *The rice economy of Asia*. Retrieved from http://books.irri.org/0915707152_content.pdf

Shao, Y., Zhao, C., Bao, Y., & He, Y. (2012). Quantification of Nitrogen Status in Rice by Least Squares Support Vector Machines and Reflectance Spectroscopy. *Food and Bioprocess Technology*, *5*(1), 100–107. https://doi.org/10.1007/s11947-009-0267-y

Shibayama, M., Sakamoto, T., Takada, E., Inoue, A., Morita, K., Yamaguchi, T., … Kimura, A. (2012). Estimating Rice Leaf Greenness (SPAD) Using Fixed-Point Continuous Observations of Visible Red and Near Infrared Narrow-Band Digital Images. *Plant Production Science*, *15*(4), 293–309. https://doi.org/10.1626/pps.15.293

Sun, J., & Yang, J. (2017). Estimating Rice Leaf Nitrogen Concentration: Influence of Regression Algorithms Based on Passive and Active Leaf Reflectance. *Remote Sensing*, 15.

Suresh, M., Balaguravaiah, D., Jayasree, G., & Balachennaiah, Y. (2017). Relation between Leaf N Content, LCC and SPAD Values on Yield in Rice. *Int.J.Curr.Microbiol.App.Sci*, 7. Retrieved from http://www.ijcmas.com

Tan, K., Wang, S., Song, Y., Liu, Y., & Gong, Z. (2018). Estimating nitrogen status of rice canopy using hyperspectral reflectance combined with BPSO-SVR in cold region. *Chemometrics and Intelligent Laboratory Systems*, *172*, 68–79. https://doi.org/10.1016/J.CHEMOLAB.2017.11.014

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, *8*(2), 127–150. https://doi.org/10.1016/0034-4257(79)90013-0

Uddling, J., Gelang-Alfredsson, J., Piikki, K., & Pleijel, H. (2007). Evaluating the relationship between leaf chlorophyll concentration and SPAD-502 chlorophyll meter readings. *Photosynthesis Research*, *91*(1), 37–46. https://doi.org/10.1007/s11120-006-9077-5

Ustin, S. (1999). *Geobotany: Vegetation Mapping for Earth Sciences*. Retrieved from https://www.researchgate.net/publication/271904290_Geobotany_Vegetation_Mapping_for_Earth_Sciences

Wang, L., Zhou, X., Zhu, X., Dong, Z., & Guo, W. (2016). Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. *Crop Journal*, *4*(3), 212–219. https://doi.org/10.1016/j.cj.2016.01.008

World Weather Online. (n.d.). Retrieved June 15, 2020, from https://www.worldweatheronline.com/los-banos-weather-averages/laguna/ph.aspx

Wu, X., Kumar, V., Ross, Q. J., Ghosh, J., Yang, Q., Motoda, H., … Steinberg, D. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*, *14*(1), 1–37. https://doi.org/10.1007/s10115-007-0114-2

Yoshida, S. (1981). Fundamentals of rice crops science. Growth and development of the rice plant. *Fundamentals of Rice crops Science*.

Zahir, S. (2014). Role of Nitrogen Fertilizer in Crop Productivity and Environmental Pollution. *International Journal of Agriculture and Forestry*, 201–206. https://doi.org/10.5923/j.ijaf.20140403.09

Zhang, J. H., Wang, K., Bailey, J. S., & Wang, R. C. (2006). Predicting nitrogen status of rice using multispectral data at canopy scale. *Pedosphere*. https://doi.org/10.1016/S1002-0160(06)60032-5

Zheng, H., Li, W., Jiang, J., Liu, Y., Cheng, T., Tian, Y., … Yao, X. (2018). A comparative assessment of different modeling algorithms for estimating leaf nitrogen content in winter wheat using multispectral images from an unmanned aerial vehicle. *Remote Sensing*, *10*(12). https://doi.org/10.3390/rs10122026

APPENDICES

Figure A1: Variation of spectral reflectance per variety during the rice-growing season in 2016EWS. ID represents the subplot number, and five subplots were selected for each variety for reflectance checking.
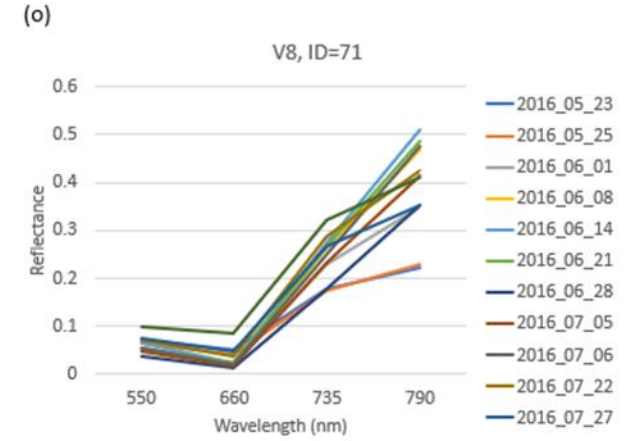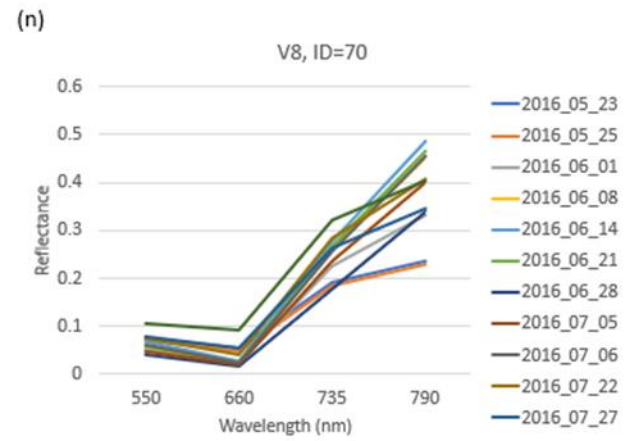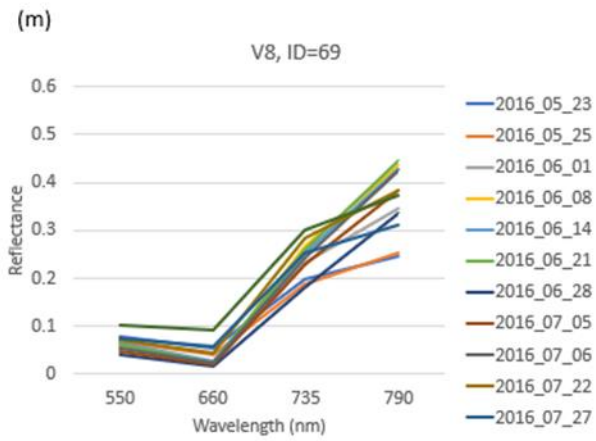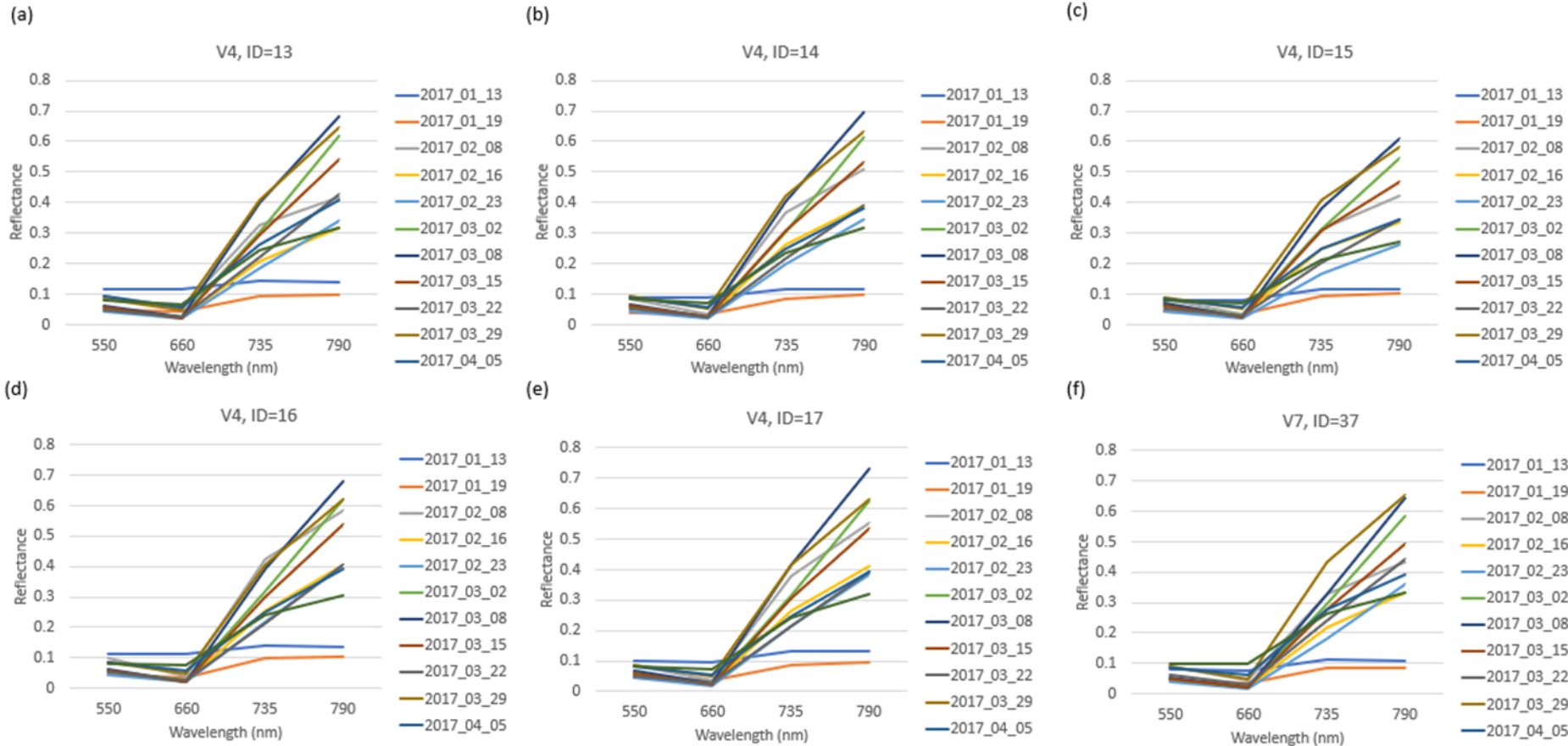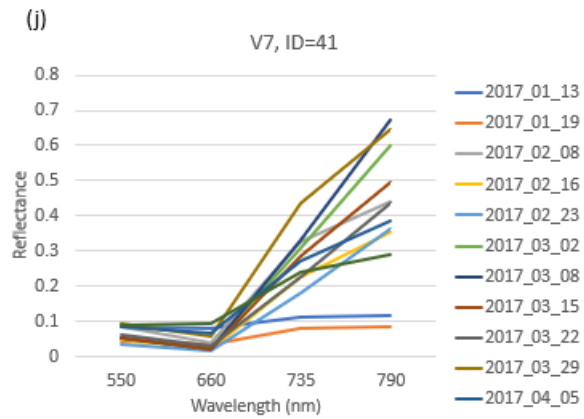
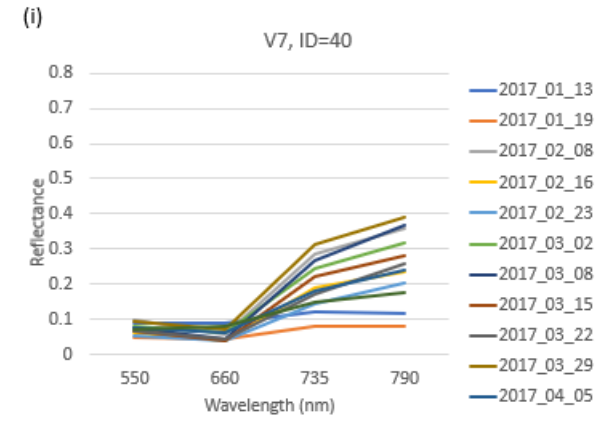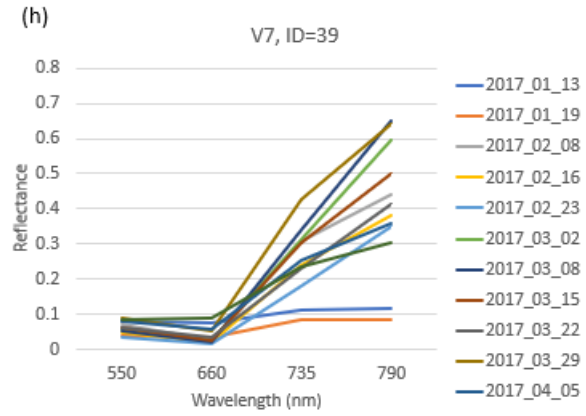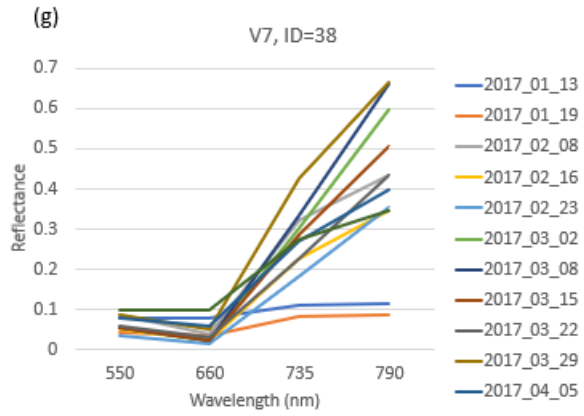(m) V8, ID=69

(n) V8, ID=70

(o) V8, ID=71
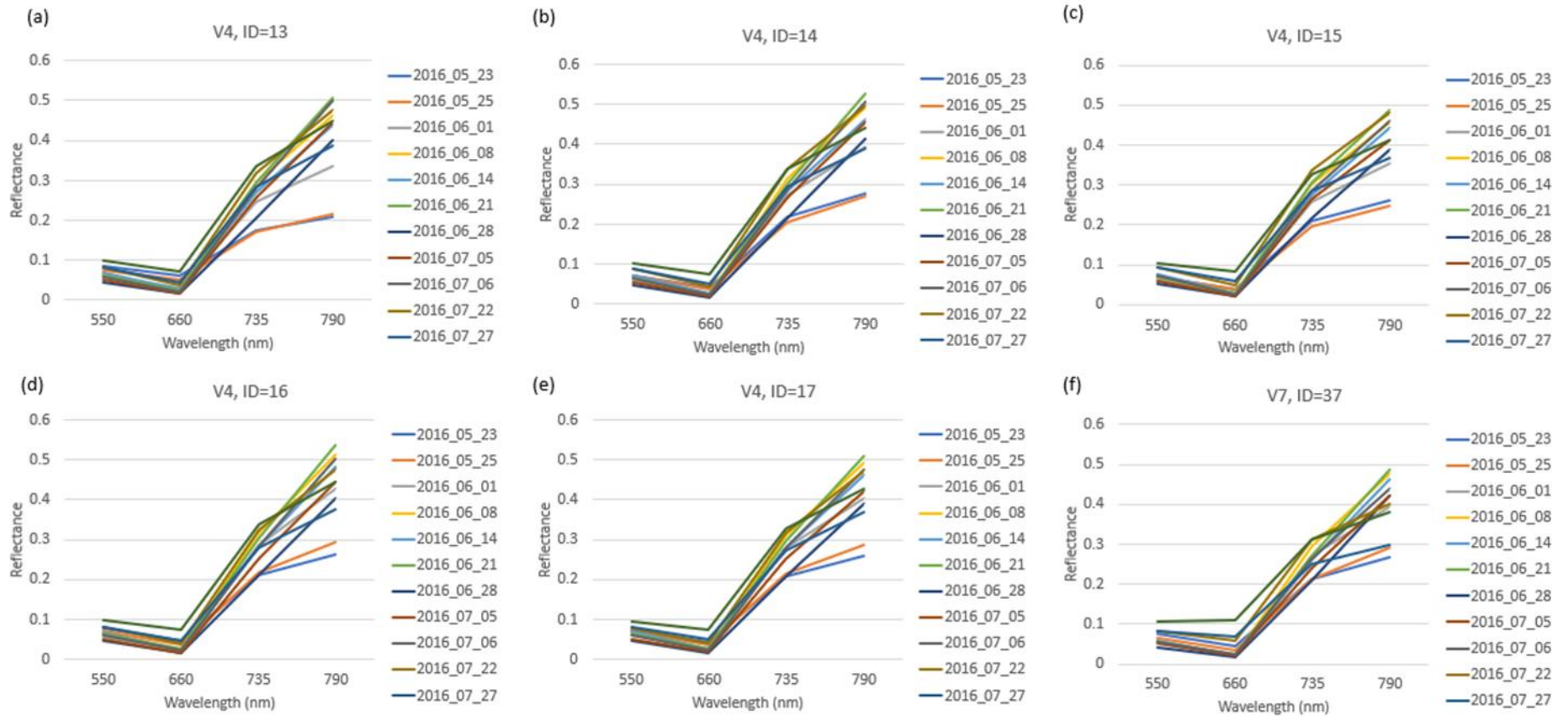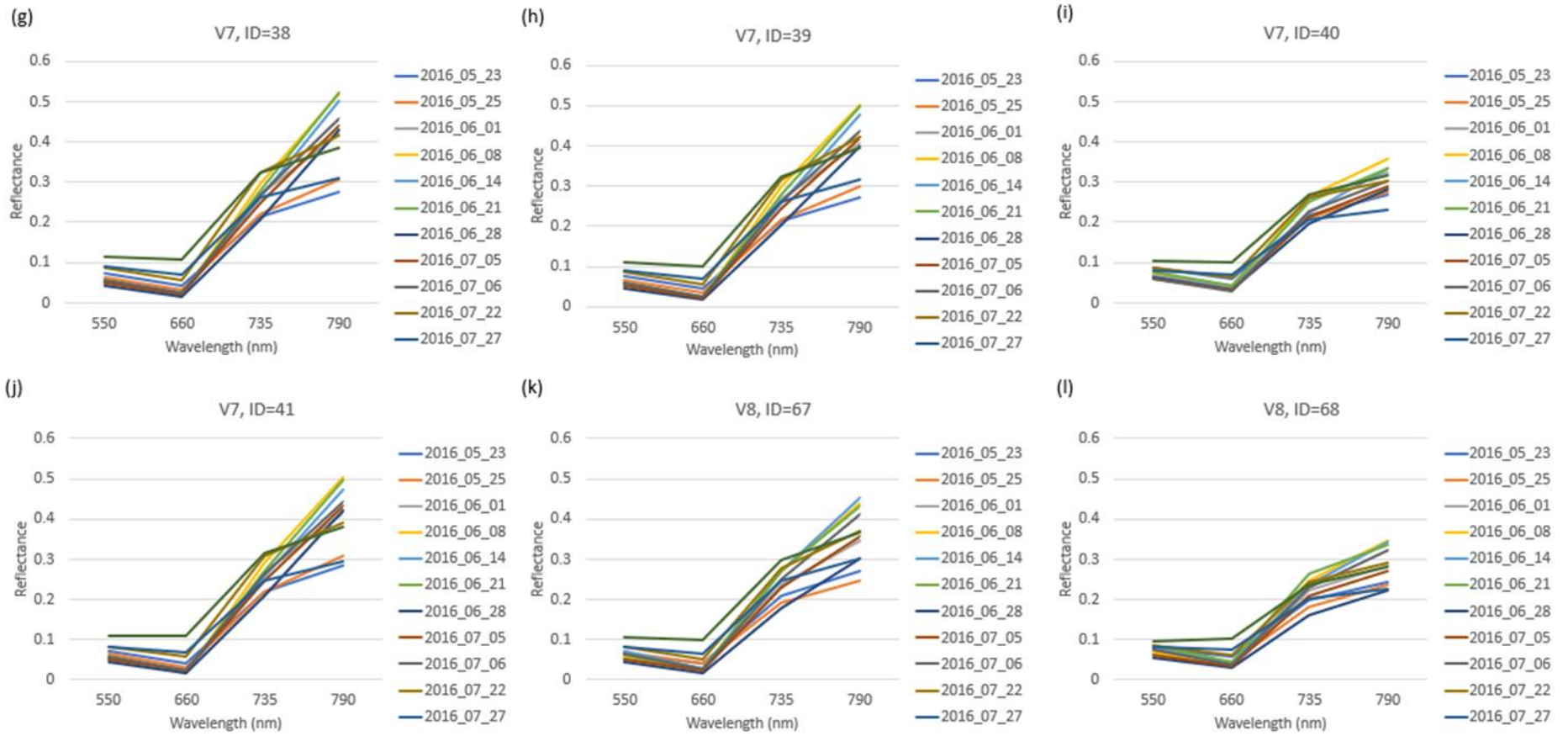
Figure A2: Variation of spectral reflectance per variety during the rice-growing season in 2017DS. ID represents the subplot number, and five subplots were selected for each variety for reflectance checking.
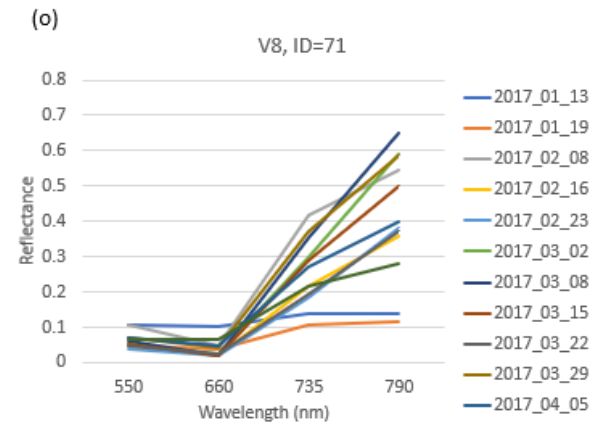
(g) V7, ID=38

(h) V7, ID=39

(i) V7, ID=40

(j) V7, ID=41

(k) V8, ID=67

(l) V8, ID=68

(m) V8, ID=69

(n) V8, ID=70

(o) V8, ID=71
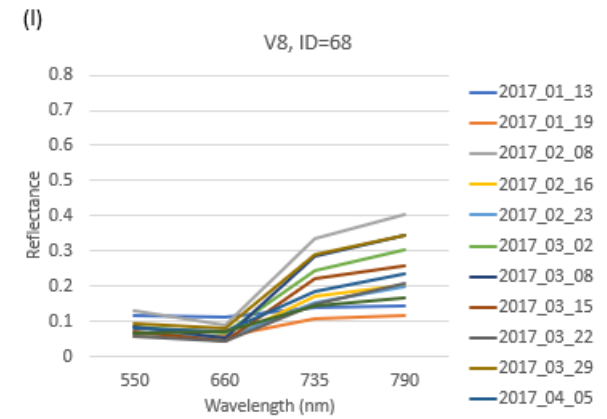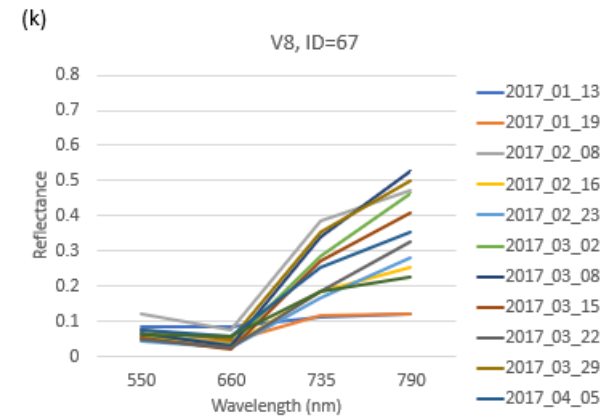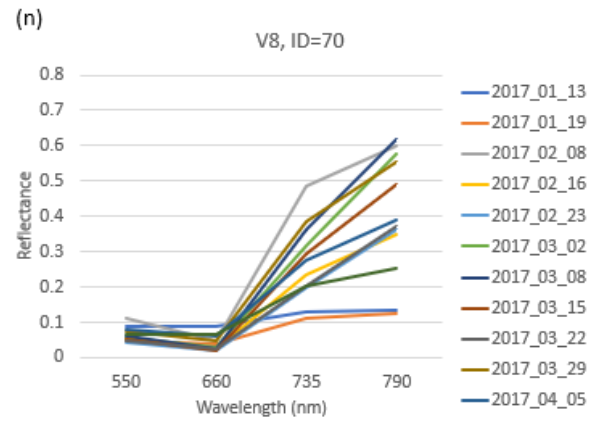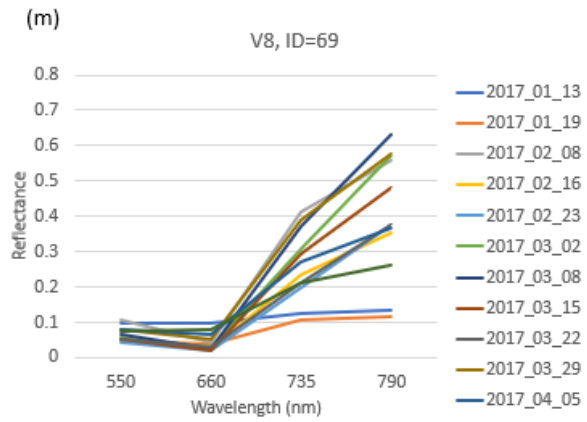
Figure A3: Field measurements (SPAD, LCC and LAI) exploration for trend analysis and outlier detection by variety (n=24) using line charts in 2016EWS and 2017DS. The SPAD value for V8 on DAT49 in 2016EWS showed an abnormally high value. Other detected outliers were the LAI value in 2017DS for V4 in DAT 43 and DAT 68. In DAT 43, the maximum LAI value reached to 10.1, which is not realistic comparing with the nearest measuring dates. In DAT 68, the minimum value is larger than the minimum value in DAT 43, which does not accord with the law of LAI changes. Therefore, these outliers were removed from further analysis.

(g)

V4 LCC 2016

DAT	22	29	36	42	49	56	63	71	78
Mean	3.3	3.3	3.2	3.3	3.2	3.2	3.0	3.0	2.7
Min	3	2.6	2.3	2.2	2	2.2	2	2.5	2.2
Max	3.5	3.5	3.5	4.2	3.5	3.5	3.4	3.4	3.2

(h)

V7 LCC 2016

DAT	22	29	36	42	49	56	63	71	78
Mean	3.4	3.4	3.3	3.6	3.3	3.3	3.2	3.1	2.4
Min	3.1	2.5	2.5	2.3	2.3	2.3	2.1	2.5	2
Max	3.5	3.5	3.5	4.8	3.5	3.5	3.5	3.5	3

(i)

V8 LCC 2016

DAT	22	29	36	42	49	56	63	71	78
Mean	3.5	3.4	3.3	3.3	3.3	3.1	3.1	3.1	2.6
Min	3.2	2.6	2.3	2	2.3	2.1	2.2	2.5	2
Max	3.5	3.5	3.5	4.4	3.5	3.5	3.5	3.5	3.3

(j)

V4 LCC 2017

DAT	23	29	36	43	50	57	64	71	78
Mean	2.9	3	3	2.9	3.3	3.2	2.9	2.8	2.7
Min	2.5	2.5	2.5	2.3	2.4	2.3	2.2	2.5	2.1
Max	3.4	3.5	3.3	3.4	3.5	3.5	3.4	3.2	3.2

(k)

V7 LCC 2017

DAT	23	29	36	43	50	57	64	71	78
Mean	3	3.2	3	3	3.3	3.2	3	3.1	2.9
Min	2.7	2.6	2.5	2.5	2.5	2.3	2.2	2.5	2.3
Max	3.4	3.4	3.5	3.3	3.5	3.5	3.5	3.5	3.5

(l)

V8 LCC 2017

DAT	23	29	36	43	50	57	64	71	78
Mean	3.1	3.2	3.1	3.1	3.3	3.2	3.0	3.0	2.9
Min	2.7	2.7	2.4	2.4	2.5	2.4	2.2	2.5	2.3
Max	3.5	3.5	3.4	3.4	3.5	3.5	3.4	3.5	3.3

(m)

V4 LAI 2016

	22	42	55	62
Mean	1.1	3.3	4.3	3.8
Min	0.7	2.2	2.2	1.7
Max	1.4	4.2	6	5.4

DAT

(n)

V7 LAI 2016

	22	42	55	62
Mean	1.1	3.6	4.7	4.3
Min	0.8	2.3	2.7	2.2
Max	1.5	4.8	6.3	6.1

DAT

(o)

V8 LAI 2016

	22	42	55	62
Mean	1.0	3.3	4.3	4.2
Min	0.7	2	2.3	1.8
Max	2.2	4.4	5.9	5.8

DAT

(p)

V4 LAI 2017

	23	43	57	68
Mean	0.4	2.6	4.4	3.8
Min	0.1	1.1	1.8	0.5
Max	0.6	10.1	6	7.2

DAT

(q)

V7 LAI 2017

	23	43	57	68
Mean	0.4	2.3	4.8	4.2
Min	0.2	1.5	1.7	1.7
Max	0.7	3.6	7.7	6

DAT

(r)

V8 LAI 2017

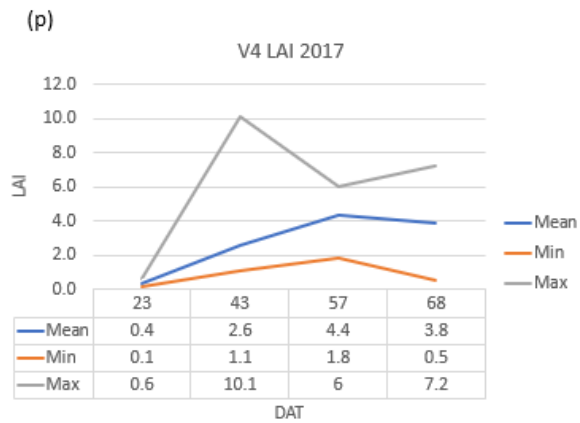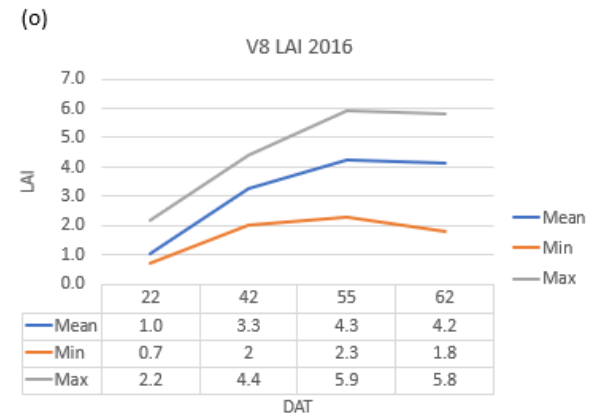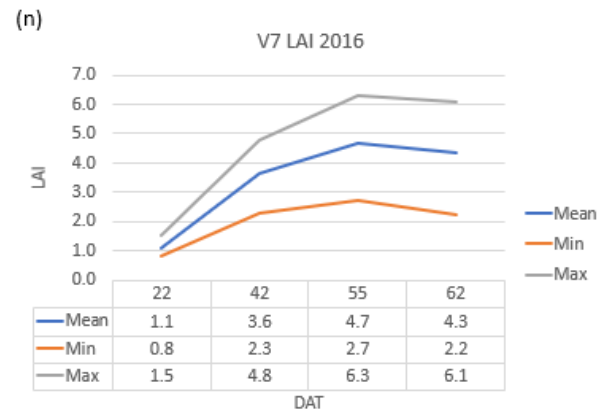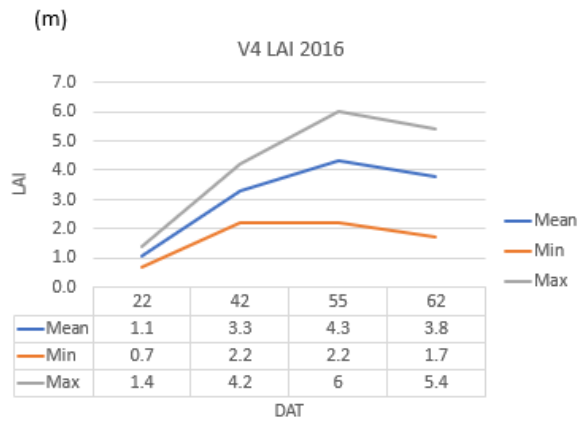|	23	43	57	68
Mean	0.4	1.8	3.7	3.7
Min	0.2	1	1.3	1.5
Max	0.6	2.8	7.6	5.2

DAT

Figure A4: Field measurements (SPAD, LCC and LAI) exploration for trend analysis and outlier detection by varieties(n=24) using boxplots in 2016EWS and 2017DS.

(m) V4 LAI value in 2016

(n) V7 LAI value in 2016

(o) V8 LAI value in 2016

(p) V4 LAI value in 2017

(q) V7 LAI value in 2017

(r) V8 LAI value in 2017