

## Modeling Potato Yield Response to Different Nitrogen Application Rates Using Hyperspectral Data and PLS Regression

<sup>1</sup>Khaled Abutaleb and <sup>2</sup>Fatma Sayed Aboud

<sup>1</sup>National Authority for Remote Sensing and Space Sciences  
<sup>2</sup>Agriculture Research Center, Central Laboratory for Agriculture Climate

**Abstract:** This study was conducted to explore whether hyperspectral data could be used to predict the yield of potato crop under different rates of nitrogen application with the help of logistic regression. The field experiment was carried out under the open field conditions in the Central Laboratory of Agriculture Climate, Agriculture Research Center. Four different rates of nitrogen fertilizer in the form of Ammonium Nitrate ( $\text{NH}_4\text{NO}_3$ ) were used to test yield and spectral response of potato crop grown into sandy soil for the two seasons of 2018-2019 and 2019-2020. Treatments included four nitrogen application 75%, 100%, 125% and 150% from the recommended rate of 180 kg N fad<sup>-1</sup>. Spectral reflectance was measured using the ASD spectroradiometer with a range of 350 – 2500 nm on three different phenological stages during the growing seasons. Band-Band-r<sup>2</sup> analysis was carried out to identify which growing stage is useful to discriminate between N levels, this was followed by Stepwise Discriminant analysis to identify the highest 20 top ranked bands that discriminate application rate. The Partial Least Square (PLS) linear regression analysis was carried out using the 5 top ranked bands that correlates with yield data. BBr2 results indicated the mid-growth stage was found to be very useful to discriminate between different N application rates. SDA results indicated that reflectance in the red-edge and the red wavelengths ranked as the most important features to discriminate between the different N application rates specially reflectance at wavelengths at 730, 740, 750, 770 and 560. The established PLS model was capable to estimate potato yield (ton/feddan) with a mean absolute percentage error of 3.6% that give a yield error ranged between 300 kg/feddan to 600 kg/feddan in the two successive seasons.

**Key word's:** Hyperspectral Data • Potato • Nitrogen • Stepwise discriminant analysis • PLS

### INTRODUCTION

Potatoes are rich in antioxidants [1], it comes in fourth highest crop after wheat, rice and corn [2]. Potato is the first highest crop in energy production and the second in the production of proteins after soybean. Potatoes are rich in vitamin C, potassium and dietary fiber [3].

Remote sensing technologies has made tremendous enhancement in relation to spectral, spatial and temporal resolutions over the past decades. Currently, remote sensing is being used widely in many agricultural activities and applications [4]. These applications include but not limited to crop discrimination, crop monitoring and yield

mapping and forecasting. Hyperspectral remote sensing has been proven as a valuable tool in plant monitoring and yield forecasting. Spectral reflectance is defined as the ratio between the reflected to the incident radiation by a specific object. Spectral reflectance is a function of the object characteristics. Many use the large differences in the spectral characteristics of the soil and crop, especially at the 'red edge' the point where the electromagnetic spectrum changes from visual to near infra-red at a wavelength of approximately 700nm. The principle is that; plants need blue and red radiations to trigger the photosynthesis process. The majority of the red light is absorbed by the chlorophyll in the canopy and therefore little is reflected, in contrast a high proportion of the near

infrared light is reflected. As canopy green cover increases, either due to increasing crop density or chlorophyll content, the percentage of red reflectance decreases whilst the near infrared reflectance increases [5-8].

Chlorophyll which is the product of the photosynthesis process, reflects the near-infrared radiation. So, any changes in the reflected radiation in the blue, red and the near-infrared wavelengths can be attributed to plant health (i.e. plant health significantly affects the spectral response of the plant specially in the blue, red and the near infrared radiation). There are many vegetation indices were developed for the purpose of plant health monitoring such as the Normalized Difference Vegetation Index (NDVI) which is a key function of the plant health, the Normalized Difference Water Index (NDWI) which is widely used to monitor water stress in plants. Vegetation indices are defined as the ratio between two or more spectral bands. Many studies developed different vegetation indices that were correlated against to crop parameters e.g. canopy characteristics, soil properties and water content. These studies showed that use of the narrow band vegetation indices make significant improvements in detecting plant nutrient stress identifying differences in vegetation green cover, crop water stress, weed detection, soil properties and crop discrimination [5-9]. However, the interpretation of this hyperspectral data can be complicated by the inter-relationships between wavelength variables [9] and many statistical techniques have been utilized to analyse such data. For example, neural networks [10-12]; partial least-squares analysis [13]; fuzzy logic [14]; principle component analysis and stepwise multiple linear regression [15] have all been used.

Macronutrients are essential to plant growth as they are directly related to plant structure and development. Among them is Nitrogen (N) which is a key factor in plant growth and insufficient amounts of N can lead to dramatic decrease in plant development and crop yield. In the other side, excessive use of N is costly and causes environmental pollution [16]. Monitoring plant nitrogen status and proper N fertilizer management are essential to balance fertilizer cost, N crop demand and minimize environmental impacts [17]. Traditional methods of identifying N contents in plants are expensive specially with field variations, sample dependent and time consuming making it difficult to has a proper management

throughout the growing season. Many authors indicated that N monitoring is could be achieved using remote sensing data [8, 17]

This paper aimed to investigate the capabilities of using hyperspectral remote sensing data and Partial Least Square (PLS) regression as tools to monitor and predict yield of potato crop under different rates of nitrogen application.

## MATERIAL AND METHODS

**Field Experiment and Plant Materials:** Two field experiments were carried out in the experimental station at the Central Laboratory of Agriculture Climate. Potato tubers were cultivated in terraces filled with sandy soil for the two successive seasons of 2018-2019 and 2019-2020 (Table 1). Phosphorus, Potassium macronutrients fertilizers were supplied in adequate amounts according to the general nutrient status of the field as determined by soil analysis (Table 2): 24 kg ha<sup>-1</sup> P<sub>2</sub>O<sub>5</sub> and 125 kg ha<sup>-1</sup> K<sub>2</sub>O. Irrigation was adapted for both the concentration of salts through fertilization so the concentration of fertilizers do not negatively affect the plants and the quantity for the water requirements of potato so no water stress occurred.

Four nitrogen application rates were implemented in this study. These rates were 75%, 100%, 125% and 150% from the recommended rate which is 180 kg N ha<sup>-1</sup> (43 g/m<sup>2</sup>). The experimental design was a randomized complete block design with four different treatments and three replicates as illustrated in Figure 1. A number of eight plants were cultivated in each block.

**Plant Samples and Data Collection:** Plant samples were collected three times each season. The fourth leaf that intersects with the spectral measurements field of view were collected and send to the laboratory for total nitrogen, phosphorus and potassium determination.

**NPK Determination:** A dried sample of 0.1 g was taken in 500 ml keijldahl flask, then added 10 ml of conc. H<sub>2</sub>SO<sub>4</sub> and digested till colorless solution appeared. The content was cooled down and diluted to about 25 ml with distilled water. Total nitrogen was determined according to keijldahl method [18]. Potassium was determined by using flame photometer according to [18]. Phosphorus was determined colorimetrically according to [19], as modified by [20].

Table 1: Different activities during the two seasons

	First	Second
Planting Date	15, October, 2018	20, October, 2019
First Measurement	29, November, 2018	05, December, 2019
Second Measurement	30, December, 2018	05, January, 2020
Third Measurement	28, January, 2019	05, February, 2020
Harvest Date	14, February, 2019	20, February, 2020

Table 2: Soil physical and chemical properties of the experimental field at the beginning of the experiment

Soil physical properties		Soil chemical properties			
Particle size distribution (%)		Soluble cations (meq. L <sup>-1</sup> )		Available nutrients (mg kg <sup>-1</sup> )	
Sand	86.55	Ca <sup>++</sup>	1.12	N	25.5
Silt	8.56	Mg <sup>++</sup>	0.76	P	3.9
Clay	4.89	Na <sup>+</sup>	1.01	K	140
Soil texture	Sandy	K <sup>+</sup>	0.18	Fe	4.09
Physical analysis		Soluble anions (meq. L <sup>-1</sup> )		Mn	0.95
Bulk density g/l	1665	CO <sub>3</sub> <sup>-</sup>	-	Zn	1.11
Total pore space %	22	HCO <sub>3</sub> <sup>-</sup>	0.12	Calcium carbonate (%)	2.11
water-holding capacity %	18.7	Cl <sup>-</sup>	1.23	pH	8.25
Air porosity %	3.3	SO <sub>4</sub> <sup>-</sup>	1.76	EC (dS m <sup>-1</sup> )	0.33

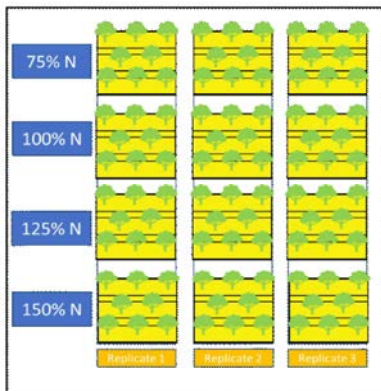


Fig. 1: Experimental design diagram as implemented in the field

**Tubers Characteristics and Yield Data Collection:** By the end of each season, potato plants were harvested and the potato tubers were collected and weighted for the three replicates for each treatment. Harvesting was carried out after 120 days from the cultivation date. Moreover, tubers weight, diameter, length and dry matter were measured and recorded.

**Spectral Data Collections and Preprocessing:** The ASD full range spectroradiometer was used to collect potato canopy spectral reflectance data. Data was collected on clear sky days when the sun is perpendicular to the earth surface to minimize shadows, atmosphere interference and solar radiation fluctuation effects. Spectral measurements were taken from the nadir position (I.e. the pistol was held perpendicular on the center of the plant). Measurements

were taken from a distance of 70 cm above the canopy with a fiberoptic lens of 25° viewing angle, resulting in a circle of 31.04 cm diameter field of view over the plant canopy. Spectral reflectance is calculated as the ratio of measured radiance to the radiance emitted from the white standard reference panel (Labsphere spectralon panel 99% reflectance). Immediately after the white reference measurement, three spectra of the potato top canopy were recorded and saved internally on the device memory for each replicate. Each of the recorded measurements is the arithmetic mean of 50 readings as configured in the ASD spectral measurement software.

Spectral data was collected three times during each season. These three times corresponding to different plant phenological stages. These three times were taken after 30, 60 and 90 days from plant emergence and forming true leaves. A total number of 36 spectra were collected each time. The mean of the three spectra was then calculated to provide a single spectrum for each replicate. Data had to be converted from binary format to a readable ASCII format. This was carried out using the ASD ViewSpec software. Once data are converted the following processing was carried out using the R software.

After removing the atmospheric window at 1340 -1440 due to the effect of water vapor on the reflected radiation. Moreover, bands beyond the 1800 nm wavelength were dropped from the data as it has many noises due to the atmospheric effect in these wavelengths and the weak response of the plant to these wavelengths as well. These bands were then grouped in 10 nm wavebands to reduce the number of bands to 135.

**Band–Band R<sup>2</sup> (BBR<sup>2</sup>):** Pearson r correlation test was carried out between each individual band and all other bands to test how bands are correlated to each other. The coefficient of determination (R<sup>2</sup>) was used instead of the correlation coefficient (r) to remove the negative signs of inverse correlations. This helped to determine bands that are rich in information and those that are redundant. High R<sup>2</sup> values between any two spectral bands indicate similar or redundant information while low R<sup>2</sup> values indicate that the two bands contain unique information about the measured target [21]. Wavelength combinations of the lowest values of R<sup>2</sup> from all the combinations were used as inputs for the following steps.

**Stepwise Discriminant Analysis:** The BBR<sup>2</sup> analysis was carried out to narrow down the number of spectral bands and remove bands with similar information. However, to test spectral response to different N application rates the Stepwise Discriminant Analysis (SDA) was required. Stepwise discriminant analysis was carried out using the R software “sda” package. Stepwise discriminant analysis is carried out to determine the optimal spectral bands to discriminate between different N levels using multivariate Wilks’ lambda measures of separability [22]. Wilks’ lambda is a multivariate test of significance, which ranges between 0 and 1. Values close to 0 indicate that group means are different and values close to 1 indicate that group means are not different and 1 indicates that all means are same.

Stepwise discriminant analysis was carried out on the spectral data from each of the four treatments different rates of N application. The full range of spectral bands was divided into 10 nm spectral bands to determine the spectral bands that discriminate the best.

**Yield Modelling and Validation:** The highest 20 spectral bands obtained from the SDA test were selected for correlation analysis with yield data. The highest five correlated bands with yield were then selected for Partial Least Squares (PLS) linear regression along with the treatment level of N applied as a quantity (ton/feddan). Regression models refer to a model that relates inputs to outputs by statistical means [23] as illustrated in the following equation (Eq. 1).

$$Yield = C_1 + C_2 N + \sum_{i=1}^{i=5} a_i B_x \quad \text{Eq. (1)}$$

where C<sub>1</sub>, C<sub>2</sub> are constants, N is the amount of nitrogen fertilizer in kg/feddan, a<sub>i</sub> is constant, B<sub>x</sub> is the reflectance value of the waveband at wavelength (x).

Model validation is necessary to determine how accurate is a model. Validation can be carried out through measuring the difference between model predicted values and the actual one in reality. Error measurement is usually used to compare models and test model accuracy. Many statistical methods were well known for error measurements, such as the root square mean error (RSME), the mean absolute error (MAE) and the mean absolute percentage error (MAPE). In this study, the MAPE was used to test and validate the LAI and the yield used models (Eq. (2)). The MAPE has the advantage of scale independence and gives the error as a percentage from the actual value, so it becomes easy to compare and test models. MAPE provides more of a standardized error measure.

$$MAPE = \left( \frac{1}{n} \sum \frac{|Y_a - Y_p|}{Y_a} \right) * 100 \quad \text{Eq. (2)}$$

where MAPE is the mean absolute percentage error, n is the sample size and Y<sub>a</sub> are the actual and Y<sub>p</sub> is the predicted yield.

## RESULT AND DISCUSSION

**NPK Levels in Plant Leaves:** Table (3a, b) illustrates the levels of N, P and K in potato leaves at each growing stage. Data revealed that there are no significant differences in the P and K levels in potato leaves during the growing season as these two elements were supplied by equally adequate amounts during the growing seasons. In contrast, N levels varied from treatment to another as the amount supplied to the plants were different. The heist significant difference N level was noticed in the treatment that received 150% of the recommended dose while the lowest level was recorded in case of 75% of the recommended N fertilizer dose. Table (3b) illustrates results of Duncan statistical test where treatments with similar letters indicates no significant difference between these treatments.

These results are in harmony with those reported by many authors. [24] reported increased plant leaf area with increasing nitrogen fertilizer application rates which consequently increased solar radiation interception and photosynthesis, decreasing days to flowering, days to maturity. Similarly, [8, 25, 26] stated that potato crop requires high amounts of nitrogen for optimum yields and they also reported that higher rate of nitrogen provides better growth, development and translocation of photosynthetic from leaves to tubers leading to higher yields.

Table 3a: NPK levels in potato leaves for the three different stages

Treatment	N (ppm)		P (ppm)		K (ppm)	
	1 <sup>st</sup> Season	2 <sup>nd</sup> Season	1 <sup>st</sup> Season	2 <sup>nd</sup> Season	1 <sup>st</sup> Season	2 <sup>nd</sup> Season
30 DAE						
75%	3.77 <sup>b</sup>	3.70 <sup>b</sup>	0.47 <sup>a</sup>	0.49 <sup>a</sup>	4.06 <sup>a</sup>	4.01 <sup>a</sup>
100%	4.32 <sup>ab</sup>	4.17 <sup>ab</sup>	0.48 <sup>a</sup>	0.47 <sup>a</sup>	4.04 <sup>a</sup>	3.94 <sup>a</sup>
125%	4.78 <sup>ab</sup>	4.74 <sup>ab</sup>	0.49 <sup>a</sup>	0.48 <sup>a</sup>	4.06 <sup>a</sup>	3.98 <sup>a</sup>
150%	5.23 <sup>a</sup>	5.30 <sup>a</sup>	0.50 <sup>a</sup>	0.50 <sup>a</sup>	4.08 <sup>a</sup>	4.02 <sup>a</sup>
60DAE						
75%	4.98 <sup>b</sup>	4.86 <sup>b</sup>	0.56 <sup>a</sup>	0.55 <sup>a</sup>	5.55 <sup>a</sup>	5.10 <sup>a</sup>
100%	5.77 <sup>ab</sup>	5.67 <sup>ab</sup>	0.56 <sup>a</sup>	0.55 <sup>a</sup>	5.55 <sup>a</sup>	5.08 <sup>a</sup>
125%	6.00 <sup>ab</sup>	5.90 <sup>ab</sup>	0.56 <sup>a</sup>	0.54 <sup>a</sup>	5.47 <sup>a</sup>	5.05 <sup>a</sup>
150%	6.22 <sup>a</sup>	6.12 <sup>a</sup>	0.55 <sup>a</sup>	0.54 <sup>a</sup>	5.39 <sup>a</sup>	5.02 <sup>a</sup>
90 DAE						
75%	4.81 <sup>b</sup>	4.77 <sup>b</sup>	0.50 <sup>a</sup>	0.50 <sup>a</sup>	5.15 <sup>a</sup>	5.10 <sup>a</sup>
100%	5.41 <sup>ab</sup>	5.39 <sup>ab</sup>	0.50 <sup>a</sup>	0.50 <sup>a</sup>	5.19 <sup>a</sup>	5.14 <sup>a</sup>
125%	5.73 <sup>ab</sup>	5.71 <sup>ab</sup>	0.51 <sup>a</sup>	0.50 <sup>a</sup>	5.47 <sup>a</sup>	5.18 <sup>a</sup>
150%	6.05 <sup>a</sup>	6.02 <sup>a</sup>	0.52 <sup>a</sup>	0.50 <sup>a</sup>	5.74 <sup>a</sup>	5.22 <sup>a</sup>

Table 3b: Duncan letters and means of the NPK for all treatments

Treatment	N (ppm)		P (ppm)		K (ppm)	
	1 <sup>st</sup> Season	2 <sup>nd</sup> Season	1 <sup>st</sup> Season	2 <sup>nd</sup> Season	1 <sup>st</sup> Season	2 <sup>nd</sup> Season
150%	5.83 <sup>a</sup>	5.81 <sup>a</sup>	0.52 <sup>a</sup>	0.51 <sup>a</sup>	5.07 <sup>a</sup>	4.75 <sup>a</sup>
125%	5.50 <sup>ab</sup>	5.45 <sup>ab</sup>	0.52 <sup>a</sup>	0.51 <sup>a</sup>	5.00 <sup>a</sup>	4.74 <sup>a</sup>
100%	5.17 <sup>ab</sup>	5.08 <sup>ab</sup>	0.51 <sup>a</sup>	0.51 <sup>a</sup>	4.93 <sup>a</sup>	4.72 <sup>a</sup>
75%	4.52 <sup>b</sup>	4.44 <sup>b</sup>	0.51 <sup>a</sup>	0.51 <sup>a</sup>	4.92 <sup>a</sup>	4.74 <sup>a</sup>
LSD	1.22	1.21	0.07	0.06	1.53	1.23

Table 4: Tubers physical characteristics and yield

Treatment	Tuber weight (g)		Tuber diameter (cm)		Tuber length (cm)		Tuber dry matter(g/100g)		Yield (ton/fed.)	
	1 <sup>st</sup> season	2 <sup>nd</sup> season	1 <sup>st</sup> season	2 <sup>nd</sup> season	1 <sup>st</sup> season	2 <sup>nd</sup> season	1 <sup>st</sup> season	2 <sup>nd</sup> season	1 <sup>st</sup> season	2 <sup>nd</sup> season
75%	66.2 <sup>c</sup>	59.44 <sup>b</sup>	3.7 <sup>b</sup>	3.96 <sup>a</sup>	6.86 <sup>b</sup>	8.27 <sup>c</sup>	11.53 <sup>b</sup>	11.39 <sup>b</sup>	11.03 <sup>b</sup>	10.4 <sup>b</sup>
100%	83.2 <sup>b</sup>	72.78 <sup>b</sup>	4.63 <sup>a</sup>	4.56 <sup>a</sup>	7.87 <sup>b</sup>	9.8 <sup>ab</sup>	12.71 <sup>ab</sup>	12.42 <sup>b</sup>	12.67 <sup>b</sup>	12.62 <sup>b</sup>
125%	78.4 <sup>bc</sup>	76.59 <sup>b</sup>	3.96 <sup>ab</sup>	4.11 <sup>a</sup>	7.92 <sup>b</sup>	8.57 <sup>bc</sup>	12.64 <sup>b</sup>	12.65 <sup>b</sup>	13.46 <sup>b</sup>	13.11 <sup>b</sup>
150%	104.56 <sup>a</sup>	100.11 <sup>a</sup>	4.48 <sup>a</sup>	4.47 <sup>a</sup>	10.49 <sup>a</sup>	10.13 <sup>a</sup>	14.11 <sup>a</sup>	14.26 <sup>a</sup>	17.21 <sup>a</sup>	17.53 <sup>a</sup>
LSD	15.76	19.09	0.71	0.61	1.56	1.49	1.45	1.54	2.94	3.18

**Tuber Characteristic and Yield:** Data illustrated in Table (4) showed that the highest tuber yield and best tuber physical quality was obtained in case of the highest nitrogen application rate (150% of the recommended dose). While the lowest values were obtained in the lowest N rate (75%). This holds true for the two successive seasons. Potato yield ranged from 10.4 and 17.53 ton/feddan. The difference in yield can be attributed mainly due to the different N fertilization rates as P and K were supplied by equally sufficient amounts. Moreover, all other parameters were similar over all treatments. This includes irrigation, pesticides and other farm management practices.

An increased potato yield is mostly associated with N application, whereas phosphorus (P) and potassium (K) without N have insignificant impact on yield [24]. [27]

confirmed that tuber yields and tuber characteristics such as tuber weight diameter and dry matter increased with nitrogen fertilizer rates. Moreover, [28] confirmed that A further increase in nitrogen fertilizer rate up to 140% of the recommended dose caused a significant increase in tuber yield.

**Spectral Signature Measurements:** Spectral curve for the three stages of potato crop are presented in Figure (2) below. The highest spectral curve was recorded after 60 days from plant emergence, while the lowest reflectance observed at the early development stage of the plant. After 90 days from plant emergence the spectral reflectance declined as plant tended to senescence. nitrogen is macro nutrient which is essential for plant development and related physiological processes.

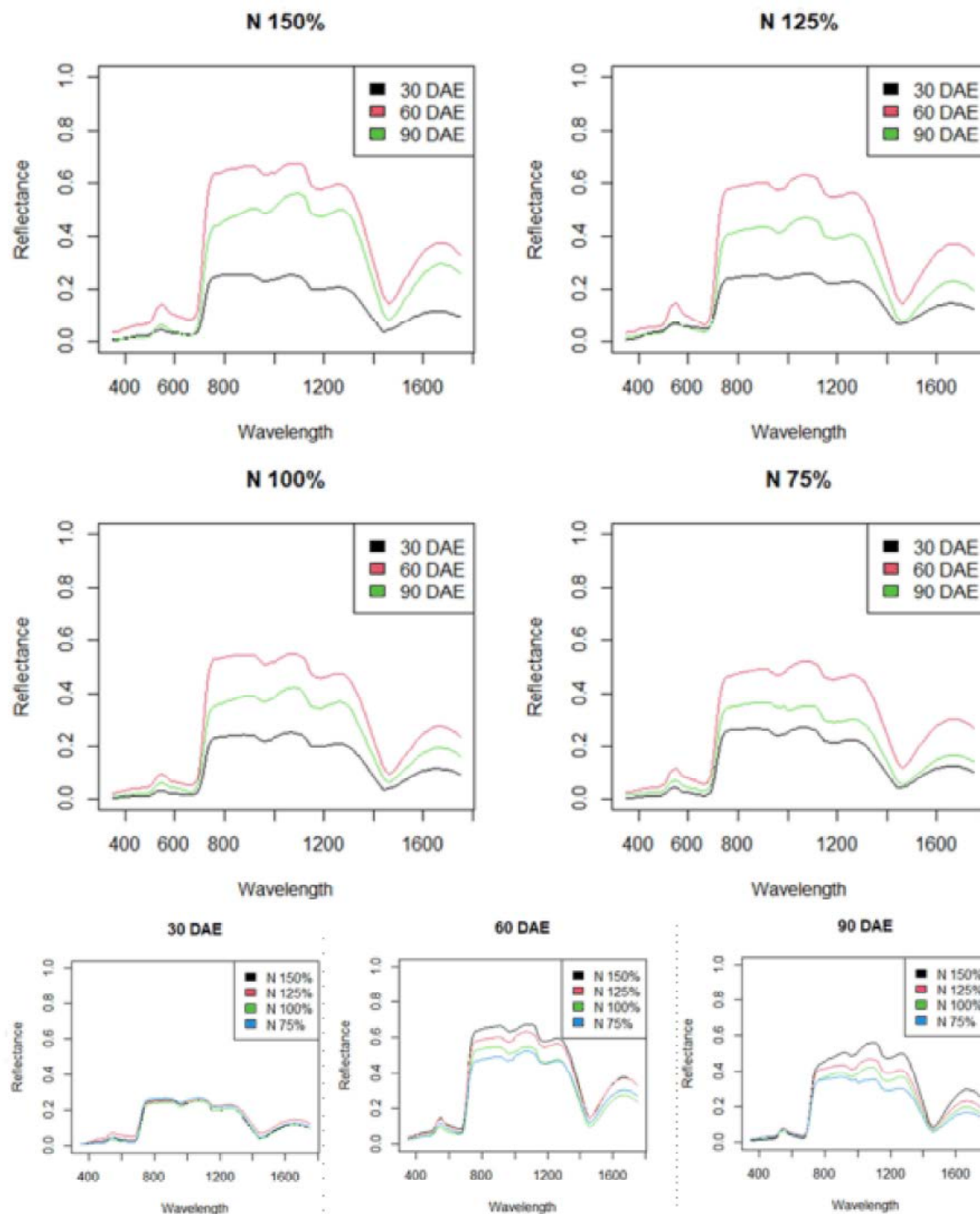


Fig. 2: Spectral response for potato crop under different N application rates and at different growing stages

In addition, nitrogen can affect the cellular structure and, consequently, reflectance of the incident radiation in the near-infrared spectrum.

It is evident from Figure (2) that, increasing the N fertilizer doses caused an increase in the reflectance in the near-infrared and the shortwave infrared regions.

Although the short-wave infrared region is sensitive to the changes in leaf water content [5, 6], but this increase in the spectral reflectance could be attributed to the promotive role of increasing nitrogen fertilizer which increase all plant physiological process [8, 24].

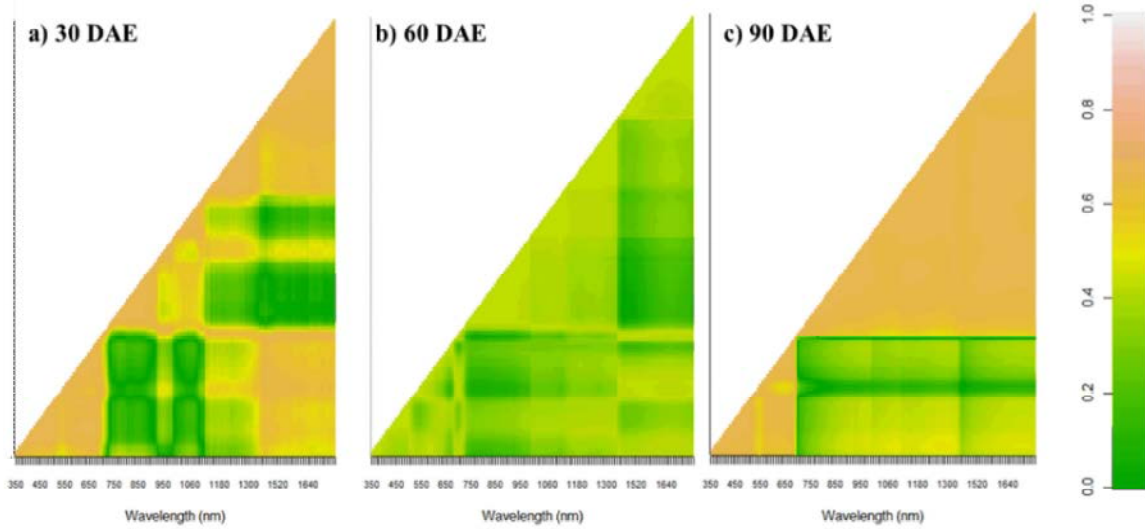


Fig. 3: Band-Band R2 correlation matrix for the three-development stage

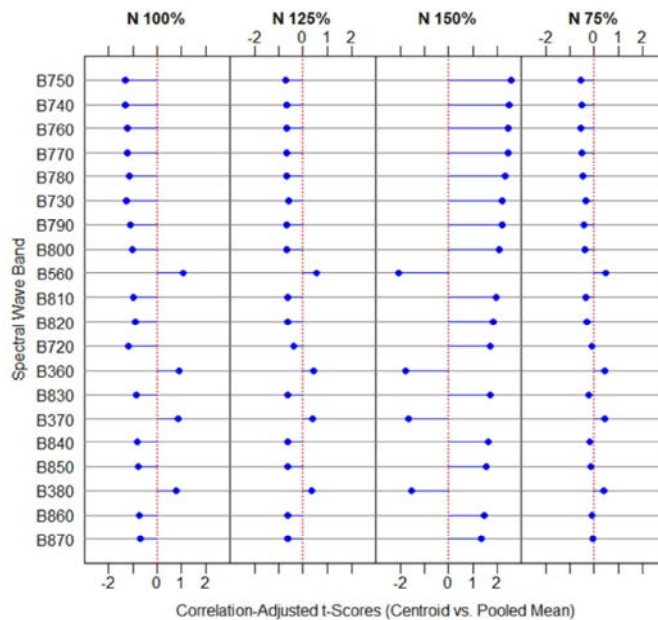


Fig. 4: Top ranked 20 bands according to their importance to discriminate N levels

**Band-Band R2 Results:** The BBr2 analysis results are presented in Figure 3. It is evident that data collected after 60 days from plant emergence were less correlated to each other's while those at the start and the end of the season were highly correlated. Moreover, it is evident that the red and red-edge region were rich in data about the crop status and have useful information regarding the N application rate [4, 5, 8].

**Stepwise Discriminant Analysis:** Results of the stepwise discriminant analysis is represented in Figure (4).

It is evident that the most important bands discriminate between different nitrogen levels were near the green peak, the red-edge and the near infrared wavelengths. Many authors reported that the green and red-edge spectral bands were sensitive to plant pigment variation. Therefore, reflectance at these spectral bands were found to be of great importance in modeling plant biophysical parameters [11, 12, 29].

The correlation-adjusted t (CAT) score is the product of the square root of the inverse correlation matrix with a vector of t scores. The CAT score describes the

Table 5: Summary statistics of the 20 top ranked bands

Band	CAT Score	CAT.N 100%	CAT.N 125%	CAT.N 150%	CAT.N 75%
B750	6.77	-1.30	-0.70	2.56	-0.55
B740	6.49	-1.32	-0.66	2.49	-0.50
B760	6.28	-1.26	-0.66	2.46	-0.55
B770	6.12	-1.22	-0.69	2.43	-0.53
B780	5.57	-1.16	-0.67	2.32	-0.49
B730	5.22	-1.29	-0.57	2.20	-0.35
B790	5.01	-1.10	-0.66	2.20	-0.44
B800	4.51	-1.05	-0.65	2.08	-0.39
B560	4.46	1.05	0.55	-2.08	0.48
B810	3.92	-0.98	-0.63	1.94	-0.34
B820	3.50	-0.92	-0.62	1.83	-0.29
B720	3.39	-1.20	-0.39	1.70	-0.11
B360	3.32	0.92	0.43	-1.79	0.45
B830	3.15	-0.88	-0.62	1.73	-0.23
B370	2.88	0.84	0.39	-1.67	0.44
B840	2.85	-0.84	-0.62	1.64	-0.18
B850	2.54	-0.79	-0.62	1.54	-0.13
B380	2.49	0.79	0.35	-1.55	0.41
B860	2.26	-0.74	-0.62	1.44	-0.09
B870	1.98	-0.69	-0.61	1.34	-0.05

Table 6: Top ranked bands correlation with the yield (ton/feddan)

Band	R	Band	r
B740	0.87	B800	0.72
B730	0.84	B810	0.65
B750	0.82	B380	0.60
B770	0.82	B820	0.60
B560	0.81	B830	0.55
B780	0.79	B720	0.50
B760	0.78	B840	0.50
B370	0.77	B850	0.44
B360	0.77	B860	0.36
B790	0.76	B870	0.30

Table 7: PLS model details

$C_1$	$C_2$	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$
8.39	-0.007	-19.81	23.85	-65.48	64.76	217.28

contribution of each individual feature in separating the two groups, after removing the effect of all other features. [30] reported that the CAT is useful criterion to rank features in the presence of correlation.

The overall CAT score of the top 20 ranked bands ranged between 6.77 in case of B750 and 1.98 in case of B870 as illustrated in Table (5). The overall CAT is equivalent to a weighted sum of squared CAT scores across the classes.

**Yield Modelling and Validation:** The top 20 ranked bands were selected and correlated against the yield and represented in Table (6) below as order from the highest r to the lowest one. The top 5 bands were chosen as inputs to the PLS model along with the N quantity in kg/feddan. These bands correlation with yield (kg/feddan)

ranged between 0.87 in case of B740 to 0.30 in case of B870.

The resulted model is expressed in equation (3), where the model details are illustrated in Table (7) below.

$$\text{Yield} = C_1 + C_2 N + a_1 B730 + a_2 B740 + a_3 B750 + a_4 B770 + a_5 B560$$

Eq. (3)

**Model Validation:** The MAPE was estimated to be 3.6% from the actual yield. This is about 0.3 to 0.6 ton/feddan in case of the minimum and the maximum yield obtained from the experiment respectively. The relationship between the predicted and the actual yield was found to be very strong with  $R^2=0.91$  and adjusted  $R^2 = 0.89$ .



## CONCLUSION

This experiment illustrated the potentiality of using hyperspectral data to monitor and predict yield of potato crop under different levels of N fertilizer. The mid growth stage was found to be very useful discriminating between different N application rates specially the wavelength in the red and the red-edge spectrum. SDA was found to be of great importance to reduce number of high dimensionality data and rank these data by its importance regarding to N rates. Estimating the yield using the spectral reflectance at 730, 740, 750, 770 and 560 wavelengths. PLS model was implemented in this study and showed a satisfactory result with a margin of error ranged between 0.3 and 0.6 ton/feddan in case of minimum and maximum yield.

## REFERENCES

1. Haan, S.D. and F. Rodriguez, 2016. Potato origin and production. In *Advances in Potato Chemistry and Technology*. J. Singh and L. Kaur, 1-32. 2<sup>nd</sup> ed. London, UK: Academic Press, Elsevier.
2. Chen, Q.J., S.U. Nandy and G. Kereliuk, 2007. Screening potato genotypes for antioxidant capacity and total phenolic.
3. Katherine A. Beals, 2019. Potatoes, Nutrition and Health. *American Journal of Potato Research*, 96: 102-110.
4. Weiss, M., F. Jacob and G. Duveiller, 2020. Remote sensing for agricultural applications: A meta-review. *Remote Sens. Environ.*, 236: 111402.
5. Cho, M. and S.K. Andrew, 2006. A new technique for extracting the red edge position from hyperspectral data: The linear extrapolation method. *Remote Sensing of Environment*, 101: 181-193.
6. Mutanga, O. and Andrew K. Skidmore, 2007. Red edge shift and biochemical content in grass canopies. *ISPRS Journal of Photogrammetry & Remote Sensing*, 62: 34-42.
7. Farifteh, J., F. Van Der Meer, C. Atzberger and E.J.M. Carranza, 2007. Quantitative analysis of salt-affected soil reflectance spectra: A comparison of two adaptive methods (PLSR and ANN). *Remote Sensing of Environment*, 110: 59-78.
8. Abutaleb, K.A., S.N. Abd El-Aziz, U.A. Sheta, M.H.A. El-Behairy and A.F. Hussain, Abou-Hadid, 2012. Effect of Nitrogen Application on the Spectral Signature of Potato Plants using One Way ANOVA Statistical Test. *Journal of Biological Chemistry & Environmental Sciences*, 7(1): 379-390.
9. Gcayi Siphokazi, R., J. Chirima George, A. Adelabu Samuel, Adam Elhadi and Abutaleb Khaled, 2019. Evaluating the Potential of Narrow-Band Indices to Predict Soybean (*Glycine Max L. Merr*) Grain Yield in The Free State and Mpumalanga of South Africa. *Open Acc. J. Envi. Soi. Sci.*, 3(1): 2019. OAJESS. M S . I D . 0 0 0 1 5 3 . D O I : 10.32474/OAJESS.2019.03.000153.
10. Goel, P.K., S.O. Prasher, J.A. Landry, R.M. Patel, A.A. Viau and J.R. Miller, 2005. Estimation of crop biophysical parameters through airborne and field hyperspectral remote sensing. *Transactions of the ASAE*, 46(4): 1235-1246.
11. Mutanga, O. and A.K. Skidmore, 2004. Narrow band vegetation indices overcome the saturation problem in biomass estimation. *International Journal of remote Sensing*, 25: 1-16.
12. Yao, X., X. Yao, W. Jia, Y. Tian, J. Ni, W. Cao and Y. Zhu, 2013. Comparison and intercalibration of vegetation indices from different sensors for monitoring above-ground plant nitrogen uptake in winter wheat. *Sensors*, 13(3): 3109-3130. DOI: 10.3390/s130303109.
13. Hansen, P.M. and J.K. Schjoerring, 2003. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sensing of Environment*, 86(4): 542-553.
14. Hemming, J. and T. Rath, 2001. Computer-vision-based weed identification under field conditions using controlled lighting. *Journal of Agricultural Engineering Research*, 78(3): 233-243.
15. Hummel, J.W., K.A. Sudduth and S.E. Hollinger, 2001. Soil moisture and organic matter prediction of surface and subsurface soils using an NIR soil sensor. *Computers and Electronics in Agriculture*, 32: 149-165.
16. Jaynes, D.B., T.S. Colvin, D.L. Karlen, C.A. Cambardella and D.W. Meek, 2001. Nitrate loss in subsurface drainage as affected by nitrogen fertilizer rate. *Journal of Environment Quality*, 30: 1305-1314.
17. Zhao, D., K.R. Reddy, V.G. Kakani, J.J. Read and S. Koti, 2005. Selection of optimum reflectance ratios for estimating leaf nitrogen and chlorophyll concentrations of field-grown cotton. *Agronomy Journal*, 97: 89-98.
18. A.O.A.C., 1990. Official methods of analysis of the association of Official analytical chemists 15<sup>th</sup> Ed. Published by the association of official analytical chemists, INC suite 400. 2200 Wilson Boulevard, Arlington, Virginia. 22201 USA.

19. Murphy, R.J. and J.P. Riley, 1962. A modified single solution method for determination of phosphate in natural waters. *Anal. Chim. Acta.*, 27: 31-36.
20. Watanabe, F.S. and S.R. Olsen, 1965. Test of an ascorbic acid method for determining phosphorus in water and NaHCO<sub>3</sub> extracts from soil. *Soil Sci. Soc. Am. Proc.*, 29: 677-678.
21. Thenkabail, P.S., E.A. Enclona, M.S. Ashton and B. Van Der Meer, 2004. Accuracy assessments of hyperspectral waveband performance for vegetation analysis applications. *Remote Sensing of Environment*, 91: 354-376.
22. Thenkabail, P.S., R.B. Smith and E.D. Pauw, 2002. Evaluation of narrowband and broadband vegetation indices for determining optimal hyperspectral wavebands for agricultural crop characterization. *Photogrammetric Engineering and Remote Sensing*, 68: 607-627.
23. Baker, R.E., J.M. Peña, J. Jayamohan and A. Jérusalem, 2018. Mechanistic models versus machine learning, a fight worth fighting for the biological community? *Biol. Lett.*, 14: 20170660.
24. Baniuniene, A. and V. Zekaite, 2008. The effect of mineral and organic fertilizers on potato tuber yield and quality. *Latv. J. Agron.*, 11: 202-206.
25. Fontes, P.C.R., H. Braun, C. Bussato and P.R. Cecon, 2010. Economic optimum nitrogen fertilization rates and nitrogen fertilization rate effects on tuber characteristics of potato cultivars. *Potato Res.*, 53: 167-179.
26. Goffart, J.P., N. Olivier and M. Frankinet, 2008. Potato crop nitrogen status assessment to improve N fertilization management and efficiency: Past –present – future. *Potato Res.*, 51: 355-383.
27. Antons Ruža, Ilze Skrabule and Aija Vaivode, 2013. Influence of nitrogen on potato productivity and nutrient use efficiency. *Proceedings of The Latvian Academy of Sciences. Section B*, 67(3) (684) 247-253.
28. Srek, P., M. Hejcman and E. Kunzova, 2010. Multivariate analysis of relationship between potato (*Solanum tuberosum* L.) yield, amount of applied elements, their concentrations in tubers and uptake in a long-term fertilizer experiment. *Field Crop Res.*, 118: 183-193.
29. Cammarano, D., G. Fitzgerald, R. Casa and B. Basso, 2014. Assessing the robustness of vegetation indices to estimate wheat N in Mediterranean environments. *Remote Sensing*, 6(4): 2827-2844. DOI: 10.3390/rs6042827.
30. Namrata Jain, Shibendu Shankar Ray, J.P. Singh and Sushma Panigrahy, 2007. Use of hyperspectral data to assess the effects of different nitrogen applications on a potato crop. *Precision Agric.* DOI 10.1007/s11119-007-9042-0.