The estimation of nutrient content using multispectral image analysis in palm oil (Elaeis guineensis Jacq)

R Budiman¹, K B Seminar² and Sudradjat³

¹Computer Sciences Department, IPB University, Bogor, Indonesia
²Mechanical and Biosystems Engineering Department, IPB University, Bogor, Indonesia
³Agronomy and Horticulture Department, IPB University, Bogor, Indonesia

Corresponding author email: rhavif.budiman@gmail.com

Abstract. Indonesia is the largest producer and exporter of palm oil in the world. 50% -70% of operational costs and around 25% of total production costs in oil palm plantations are only for fertilizer. The requirements for effective fertilizer management are meeting plant nutrition requirements and preventing nutrient deficiencies. This study aims to estimate the nitrogen, phosphorus and potassium nutrients contained in palm oil trees using multispectral cameras taken with drones. The method used was divided into data preparation and leaf sampling and taking pictures with drones. Second, pre-data processing, the things that are done was photo stitching, georeferencing and digitizing the land and oil palm canopy. Third, the data analysis stage found the correlation between nutrition and image analysis using regression analysis with the backward method. The results of this study are that the nitrogen model has correctness of 95.11%, the phosphorus model has correctness of 95.38%, and the potassium model has correctness of 88.65%.

Keywords: drone, nutrient estimation, multispectral analysis, palm oil, precision agriculture

1. Introduction
Palm oil is one of the plantation commodities that has an important role in the Indonesian economy. Indonesia is the largest producer and exporter of palm oil in the world. Indonesia's palm oil production in 2016 reached 31.40 million tons. Riau Province is Indonesia's largest palm oil-producing province, with 8.09 million tons or 25.7% of all palm oil production in Indonesia [1]. Increasing holes in oil palm have caused the plantation area in Indonesia to expand. The area of oil palm plantations in Indonesia before 2016 for the last five years tends to show an increase, around 5.38% to 10.96% per year. In 2011 Indonesia's oil palm plantation area was recorded at 9.13 million hectares, increasing to 10.75 million hectares in 2015 or 25.80 percent [1].

The rapid increase in the development of oil palm land in Indonesia has negative impacts, one of which is the ecological impact. In contrast, the increasing demand for palm oil must be met. The solution to overcome this problem is to enhance the productivity level of oil palm d. The way to enhance the productivity of oil palm in the field is good fertilizer management. The aim of fertilizer management is for every oil palm job with adequate nutrition and balanced proportions to guarantee health and growth of the oil palm well and optimal yields. The benefits of good fertilization management are in maintaining and healthy yields of oil palm and are also prerequisites for sustainable oil palm [2].

The supply of fertilizer in large quantities is not very good because the fertilizer price is not low, and many farmers who provide fertilizer are not under the circumstances and recommendations. The
application of a high amount of fertilizer can reduce the quality of soil and plants. In contrast, applying a low amount of fertilizer will not have an optimal impact on plant growth and production [3]. Fertilizers in operational costs are around 50% - 70% and about 25% of the total production costs [4].

Researchers [5] has studied several methods of detecting nutrients in oil palm, namely the Kjeldahl method, Soil Plant Analysis Development (SPAD), chlorophyll meter, Spectroradiometer, and the use of satellites. The first method is the Kjeldahl method; this method has been widely used to detect nitrogen nutrients in plants. The Kjeldahl method is easy but time-consuming, destructive and expensive for large oil palm plantations. The second method is the Soil Plant Analysis Development (SPAD) chlorophyll meter. The working principle of this method is a linear relationship between chlorophyll and nutrients. The SPAD method is simpler and more accessible, but this method has the disadvantage of being carried in the weather. The third method is to use a Spectroradiometer sensor. This tool supports the identification of palm oil but this sensor is expensive and requires special skills to use this tool. The fourth method is the use of satellites to determine the nutrition of oil palm. The detection of oil palm nutrients using satellites has been carried out by [6] using Sentinel-2A satellite imagery. The advantages using satellite is wide area coverage but biggest obstacle to using the satellite method is cloud cover and low resolution.

This research was conducted by taking information about the nutritional content of oil palm using images with better resolution and avoiding foreign objects when taking images using drones. The images studied with oil palm nutrients are nitrogen, phosphorus and potassium contained in oil palm leaves. The comparison results are then searched for the correlation between the images obtained and the nutritional conditions in oil palm.

2. Methodologies

The material used in this study was the 17th midrib of oil palm which was taken three pairs of leaves in the middle of the oil palm. This frond will be analyzed at the Testing Laboratory of the Department of Agronomy and Horticulture. The tools used in this research are Drone, Multispectral Camera, Computer, Tab phone, Pix4DCapture, Minitab 18, Pix4DMapper, GPS handheld, QGIS 2.18.

2.1. Research procedure

This research begins with the data preparation stage, which consists of leaf sampling and drone imagery. The leaf samples taken were recorded with the coordinates of the leaf samples and then analyzed in Testing Laboratory for nutrient analyses. After that, the analysis results were used to analyze the data. The drone image data obtained must be run in the pre-processing stage of the data. This stage consists of photo stitching, georeferencing and limiting the experimental location. Drone images that have gone through the pre-processing stage of data. The results of the laboratory analysis and the results of drone imagery were analyzed using regression analysis. After that, the regression results were selected using the backward method. All procedure workflow can be seen in Figure 1.

![Figure 1. Research procedure.](image-url)
2.2. Leaf data retrieval
Based on the guidelines for leaf sampling on oil palm plants [7], the leaves used as sample subjects must meet several conditions. Namely, they usually grow standard, not inserts, are not located adjacent to roads or rivers and are not invaded by pests or diseases. Leaf sampling was run in a distributed manner where sample plants were determined reasonable. Examples of leaves taken are from the 17th midrib.

Research conducted by [8] said that the 17th frond was the most sensitive because it showed the most significant difference in nutrient levels. Nutrient status on the 17th midrib had a better relationship between planting production when compared to other younger leaves and vice versa. Six leaflets were taken from the 17th leaf (at the meeting point of the two sides of the leaf midrib point, leaf take three on the left and three on the right).

The leaves were then dried in an oven at 60°C until they had a constant weight. After that, the oil palm leaves were crushed with a blender to a size of 0.5 mm for analysis of N, P, and K. The selected sample trees were tied with rope to mark the tree as a sample tree. The sample is marked on the GPS also for georeferenced between the map and the drone image results.

2.3. Drone data retrieval
Image capture by a drone using a multispectral camera using a Phantom 4 drone embedded Parrot Sequoia camera. The Phantom 4 drone does not have a slot to attach a Parrot Sequoia camera, so there must be modifications to the drone mounting. The iPad Air is used for automatic drone flight planning using the Pix4DMapper application, which is modified to the area of the pickup area and the hours that the drone can travel. Image capture is carried out in the morning from 8 to 11 in the afternoon.

2.4. Photo Stitching
Photo stitching is the process of combining one image with another. There are three steps in the photo stamping process: image calibration, image registration, and image blending [9]. Image calibration aims to minimize the difference between the ideal lens and the 9 lens combinations used. This difference causes differences between images resulting in optical defects and distortion. Image registration is the process of aligning two or more images taken from different points of view. The purpose of image registration is to add geometric values to the captured images. Image blending is a process used to make the transitions between images smoother by removing the merge between two or more images.

2.5. Georeferencing
Georeferencing is a way to get the coordinates of raster data on the earth's surface. Georeferencing of drone images is carried out by equalize images and POS (Positioning and Orientation System) images and data [10]. The georeferencing process is carried out simultaneously with the photo stamping process in the Pix4DMapper application. Files generated by multispectral cameras already have coordinates. The reference coordinates obtained from the drone image are WGS 84 / UTM zone 48S at EPSG: 32748.

2.6. Digitizing oil palm land and canopy
Bordering is the process of making garden boundaries by digitizing the image at the boundary points of the garden. The digitization is carried out according to the boundaries of the oil palm plantations per block. The digitization process without knowing the coordinates can be carried out if the position of the garden in the landscape in the surrounding area is known. The source image is obtained from google maps plugged in with QGIS to be overlaid with other images. Bordering results from digitization are stored in shapefile (SHP) format. After that, the sample palm trees were digitized and labelled according to the sample name. According to the palm tree canopy, the digitized results will later extract the average digital number value in the digitized results.

2.7. Regression analysis with the backward method
This stage begins after getting the average value of the digitizing attribute of the drone image. Regression analysis was conducted by correlating the average drone yield and actual oil palm nutrition
obtained from the 17 oil palm midrib analyses. Regression analysis using multiple linear regression as follows:

\[ y = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n \]

The dependent variable \( y \) represents the value of the dependent variable or the variable to be predicted. In this case the value of \( y \) is nitrogen, phosphor and potassium and for \( x_1, x_2, \) and \( x_n \) are independent variables with coefficients \( b_0, b_1, b_2, \) and \( b_n. \) The backward method is a step back where all X variables or dependent variables are regressed with Y or independent variables. The elimination of the X variable is based on the smallest F(partial) value and whether or not the X variable is included in the model is also determined by the F(table) value.

The backward method is a suitable regression method because, in this method, the behaviour of the response variable is explained as well as possible by selecting the explanatory variable from the many explanatory variables available in the data [11]. After obtaining the selected model, then the relationship between the variable and the independent variable is seen. Using \( R^2 \) is used to see how significant the proportion or presentation of the influence of the independent variable on the dependent variable is. The more value of \( R^2 \) to one is better and vice versa. The more independent variables influence the dependent variable [12]. Table 1 is the interpretation of \( R^2 \) based on [13].

### Table 1. \( R^2 \) Interpretation.

| Coefficient interval | Affect level |
|---------------------|-------------|
| 0.00 – 0.19         | Very low    |
| 0.20 – 0.39         | Low         |
| 0.40 – 0.59         | Medium      |
| 0.60 – 0.79         | High        |
| 0.80 – 1.00         | Very high   |

The backwards method's regression analysis model was validated using the Mean Absolute Percentage Error (MAPE). MAPE measures relative error and expresses the error percentage in the estimation results against actual demand over a certain period [14]. The MAPE formulation is as follows:

\[
MAPE = \frac{100}{N} \times \sum_{i=1}^{N} \left| \frac{x_i - \hat{x}_i}{x_i} \right|
\]

\( N \) is the number of data, \( x_i \) is the actual observed value, and \( \hat{x}_i \) is the predicted value. The actual value is the nutritional value of the leaves as a result of laboratory analysis, and the predicted value is the nutritional value calculated by the model. Table 2 is the basis used to determine the predictive ability of an estimator model.

### Table 2. MAPE Interpretation.

| MAPE (%) | Prediction level |
|----------|------------------|
| > 50     | Not good         |
| 20 – 50  | Decent           |
| 10 – 20  | Good             |
| < 10     | Very Good        |

Correctness is the percentage of accuracy obtained from reducing errors. Correctness formulation is as follows:

\[
Correctness = 1 - MAPE
\]
3. Result and discussion

3.1. Leaf data retrieval

IPB-Cargill Oil Palm Education Garden, commonly called IPB-Cargill Oil Palm Teaching Farm, is located in Singasari, Jonggol, Bogor, West Java with coordinates 06° 28’ 31.9” South, 107° 01’103” east and is at 116 m above sea level [15]. This area is divided into 5 plantation blocks with 59 hectares, and all oil palm plantations were planted in 2012. Based on the guidelines for leaf sampling on oil palm plants [7], the leaves used as sample subjects must meet several conditions. Namely, they usually grow standard, not inserts, are not located adjacent to roads or rivers and are not invaded by pests or diseases. Leaf sampling was run in a distributed manner where sample plants were determined reasonable. Examples of leaves taken are from the 17th midrib.

The picture was taken using a DJI Phantom 4 drone with a modified camera mount to hold a Parrot Sequoia multispectral camera. The Parrot Sequoia multispectral camera has 5 different bands: green, near-infrared, red, red edge, and RGB. The drone is controlled automatically using the Pix4dMapper app according to the flight plan. Before the drone flies, it is necessary to ensure that all drone equipment is installed correctly so that no problems will occur while flying.

Drone flight settings use standard planning under the recommendations that the Parrot Sequoia camera provides. With a flying height of 110 meters above ground level and a drone speed of 10 m/s, the image capture time is 2.1 seconds per image. Implementation of flight planning using an application called Pix4dCapture installed on the Ipad Air. Pix4dCapture has several features, such as the drone can return if the battery used will run out to a specific limit and continue to fly back where the drone left off. The number of flight plans is carried out according to the drone's battery capacity, which is around 18 minutes for a land area of about 650 m x 650 m with a height of 110 meters and a speed of about 10 m/s.

3.2. Data preprocessing

The image obtained using a multispectral camera produces 5 bands, namely green, NIR, red, red edge and RGB. The multispectral camera image obtained after flying the drone needs to be selected. First, the selected photos are unnecessary and interfere with the stitching and georeferencing process, such as images during camera checks, take-off, and landing. After that, stitching and georeferencing processes can be performed.

The stitching and georeferencing processes are automatically performed in the Pix4dMapper software. The advantage of this software compared to other software is that this software has collaborated with the Parrot Sequoia company so that the results from multispectral cameras can be directly stamped and georeferencing. Make a map's border is carried out by manual digitization according to the palm oil blocks produced pattern. This map-making is to limit which palm oil lands are not. In this research, manual digitization of the palm oil canopy at each sample coordinate point extracted the average digital number value for each pixel in the processed layer of drone image stitching.

3.3. Data processing

The palm oil digital number value is obtained from the oil palm canopy polygon and then averaged. All data extracted at correspond polygons with the actual position of the sample in the field will be linked to the 17th(midrib nutrition from the laboratory for regression analysis. Thirty samples will be divided into two groups, 23 samples are used to build the regression model using the backward method, and 7 samples are used to validate the model generated by the regression analysis using the backward method. The model was generated using the statistical application "Minitab 18" using linear regression with the backward method. The backward method includes all variables and then removes them according to the level of influence. When all variables have no effect, the process ends [16].

The backward method is a suitable regression method because in this method, the behaviour of the response variable is explained as well as possible by selecting the explanatory variable from the many explanatory variables available in the data [11]. $R^2$ to see how big the proportion or presentation of the
influence of the independent variable on the dependent variable. The closer the $R^2$ value is to 1, the more the independent variable affects the dependent variable [12].

The criteria for determining a good estimation model are models with a small MAPE value. A small MAPE means that the model error is small and vice versa. Oil palm nutrition is the dependent variable, and the value of the digital number is the independent variable. Each nutrient has its model. The model made was carried out with two different treatments. Treatment one model, namely the digital number value derived from the extraction of multispectral image values. Treatment model 2, namely the digital number value raised to the power of two to get the results of the polynomial model as done by [6]. The regression model for treatment one with the backward method is shown in Equation 1, Equation 2, and Equation 3, while the regression model for treatment two with the backward method can be seen in Equation 4, Equation 5 and Equation 6 (Table 3). The overall summary of the model can be seen in Table 4.

Table 3. Regression model for treatment.

| Nutrient | Model                                                                 |
|----------|------------------------------------------------------------------------|
| N1       | $N_1 = 1.777 + 0.00659 \text{ green}$                                  |
| P1       | $P_1 = 0.1063 + 0.000866 \text{ green} - 0.000764 \text{ red}$         |
| K1       | $K_1 = 0.073 - 0.00615 \text{ red} + 0.00989 \text{ red edge}$         |
| N2       | $N_2 = -11.09 + 0.2035 \text{ red edge} + 0.000039 \text{ green}_2 - 0.000787 \text{ red edge}_2$ |
| P2       | $P_2 = -1.021 + 0.01728 \text{ red edge} + 0.000004 \text{ green}_2 - 0.000005 \text{ red}_2 - 0.000065 \text{ red edge}_2$ |
| K2       | $K_2 = 17.23 - 0.0670 \text{ green} - 0.1921 \text{ NIR} - 0.01192 \text{ red} + 0.01431 \text{ red edge} + 0.000327$ |

Table 4. Model summary with $R^2$, MAPE and Correctness.

| Nutrition | $R^2$ (%) | Adjusted $R^2$ (%) | MAPE (%) |
|-----------|-----------|--------------------|-----------|
| N1        | 44.27     | 41.61              | 3.17      |
| P1        | 46.65     | 41.31              | 4.72      |
| K1        | 15.04     | 6.54               | 13.22     |
| N2        | 54.03     | 46.77              | 4.96      |
| P2        | 59.14     | 50.06              | 4.83      |
| K2        | 51.51     | 33.32              | 16.54     |

Table 4 shows that the correctness of the estimator models N, P and K for treatment one and treatment two have values that are not too significantly different, namely 1% - 2%. In comparison, for $R^2$ and Adjusted $R^2$. The comparison difference of up to 10%, because model estimator of the second treatment, the regression line will increasingly follow the distribution pattern of variables. Adjusted $R^2$ from the model shows the relationship between the variable and independent variable is not good, but to predict nutrition is excellent. The model chosen is a model with treatment. There is no polynomial function and fewer variables because the model with fewer variables will speed up the calculation process and create a nutrient distribution map.

3.4. Model validation

All models generated using regression analysis with the backward method in treatment one was validated using the second data group, which amounted to 7 pieces. The nutritional results from the model are then compared with the actual nutrients from the results of the laboratory nutrition analysis. The validation results can be seen in Table 5.
The validation results show excellent results for measuring the accuracy of the nitrogen and phosphorus models, while the potassium estimation results produce good results.

### 3.5. Nutrient distribution layer

The nutrient distribution layer comprises raster images whose values are converted using the obtained model and modified using the raster calculator tool. After the raster layer is obtained, the land bordering results intersect with the hexagon grid layer. The radius of the hexagon grid created is 9.2 meters which are obtained from the average plant spacing. After the polygon image is created, the polygon intersects with the land border layer, and then a polygon with a hexagon grid will be obtained. The average values in the hexagon grid are then classified based on the criteria [17] in Table 6. The distribution of nitrogen, phosphorus and potassium nutrients can be seen in Figure 2, Figure 3 and Figure 4.

**Table 5. Model summary with R², MAPE and Correctness**

| Nutrition | R² (%) | Adjusted R² (%) |
|-----------|--------|-----------------|
| N2        | 44.27  | 41.61           |
| P2        | 59.14  | 50.06           |
| K2        | 51.51  | 33.32           |

**Table 6. Concentration of nutrients in the 17th midrib at the age of oil palm more than 7 years.**

| Nutrition | Unit | Deficiency | Optimum  | Excess |
|-----------|------|------------|----------|--------|
| N         | %DM  | < 2.3      | 2.40 – 2.80 | > 3.00 |
| P         | %DM  | < 0.14     | 0.15 – 0.18 | > 0.25 |
| K         | %DM  | < 0.75     | 0.90 – 1.20 | > 1.60 |

**Figure 2.** Image of nitrogen nutrient distribution layer.

**Figure 3.** Image of phosphorus nutrient distribution layer.
Figure 4. Image of potassium nutrient distribution layer.

4. Conclusion
The model chosen is the model with treatment one because the model has fewer variables which will speed up the process of making nutrient distribution maps and MAPE and correctness, which are not significantly different between the two treatments. MAPE validation results on the N model of 4.89% and P of 4.62% indicate an excellent category for estimation power. The K model of 11.35% indicates a suitable category for the model's predictive power based on the category of the model's predictive power. The correctness of the validation results obtained is 95.11% for the N estimator model, 95.38% for the P estimator model, and 88.65% for the K estimator model.

References
[1] BPS B P S 2019 Statistik Kelapa Sawit Indonesia 2018
[2] Goh K J, Teo C B, Chew P S and Chiu S B 1998 Fertiliser Management in Oil Palm - Agronomic Principles and Field Practices
[3] Safuan L O, Rembon F S and Syaf H 2013 Evaluasi Status Hara Tanah dan Jaringan Sebagai Dasar Rekomendasi Pemupukan N,P, dan K pada Tanaman Kelapa Sawit Agriplus23 154–62
[4] Comte I, Colin F, Whalen J K, Grünberger O and Caliman J-P 2012 Agricultural Practices in Oil Palm Plantations and Their Impact on Hydrological Changes, Nutrient Fluxes and Water Quality in Indonesia Advances in Agronomy vol 116 (Elsevier) pp 71–124
[5] Rendana M, Rahim S A, Lihan T, Idris W M R and Rahman Z A 2015 A Review of Methods for Detecting Nutrient Stress of Oil Palm in Malaysia Journal of Applied Environmental and Biological Sciences5 60–4
[6] Kaliana I 2018 Development of a decision support system for oil palm fertilizer requirement based on precision agriculture (Bogor[ID]): IPB University)
[7] Winarna, Sutarta E S and Sugiyono 2005 Seri Buku Saku 07 Pedoman Pengambilan Contoh Daun dan Tanah pada Tanaman Kelapa Sawit (Medan: Pusat Penelitian Kelapa Sawit)
[8] Chapman G W and Gray H M 1949 Leaf Analysis and the Nutrition of the Oil Palm (Elaeis guineensis Jacq.) Annals of Botany13 415–33
[9] Lyu W, Zhou Z, Chen L and Zhou Y 2019 A survey on image and video stitching Virtual Reality & Intelligent Hardware1 55–83
[10] Wang S, Ding L, Chen Z and Dou A 2018 A Rapid UAV Image Georeference Algorithm Developed for Emergency Response Journal of Sensors2018 1–10
[11] Samosir N, Siagian P and Bangun P 2014 Analisa metode backward danmetode forward untuk menentukan persamaan regresi linear berganda studi kasus: jumlah kecelakaan lalu lintas di Kotamadya Medan Jurnal Sainsia Matematika2 345–60
[12] Gujarati D N and Porter D C 2013 Dasar-dasar Ekonometrika (Jakarta: Salemba Empat)
[13] Sugiyono 2007 Metode Penelitian Kuantitatif Kualitatif dan R&D (bandung: Alfabet)
[14] Widiyarini 2016 Penggunaan metode peramalan dalam produksi kayu untuk penentuan total permintaan (konsumen) Jurnal SOSIO-E-KONS8 54–61
[15] Sudradjat, Saputra H., and Sudirman Yahya 2015 Optimization of NPK Compound Fertilizer Package Rate on One Year Old Oil Palm (Elaeis guineensis, Jacq) Trees International Journal of Sciences 20 8

[16] Meyer F 2019 Spaceborne Synthetic Aperture Radar: Principles, Data Access, and Basic Processing Techniques (Alabama(US): National Space Science and Technology Center)

[17] Uexküll H R and Fairhurst T H 1991 Fertilizing for High Yield and Quality (Switzerland: International Potash Institute)