information

Article

The Use of Random Forest Regression for Estimating Leaf Nitrogen Content of Oil Palm Based on Sentinel 1-A Imagery

Sirojul Munir [1], Kudang Boro Seminar [2,*], Sudradjat [3], Heru Sukoco [1] and Agus Buono [1]

[1] Department of Computer Science, Faculty of Mathematics and Natural Sciences, IPB University, Bogor 16680, Indonesia
[2] Department of Mechanical and Biosystem Engineering, Faculty of Agricultural Engineering & Technology, IPB University, Bogor 16680, Indonesia
[3] Department of Agronomy and Horticulture, Faculty of Agriculture, IPB University, Bogor 16680, Indonesia
* Correspondence: kseminar@apps.ipb.ac.id

Abstract: For obtaining a spatial map of the distribution of nitrogen nutrients from oil palm plantations, a quite complex Leaf Sampling Unit (LSU) is required. In addition, sample analysis in the laboratory is time consuming and quite expensive, especially for large plantation areas. Monitoring the nutrition of oil palm plants can be achieved using remote-sensing technology. The main obstacles of using passive sensors in multispectral imagery are cloud cover and shadow noise. This research used C-SAR Sentinel equipped with active sensors that can overcome cloud barriers. A model to estimate leaf nitrogen nutrient status was constructed using random forest regression (RFR) based on multiple polarization (VV-VH) and local incidence angle (LIA) data on Sentinel-1A imagery. A sample of 1116 LSU data from different islands (i.e., Sumatra, Java, and Kalimantan) was used to develop the proposed estimation model. The performance evaluation of the model obtained the averaged MAPE, correctness, and MSE of 9.68%, 90.32% and 11.03%, respectively. Spatial maps of the distribution of nitrogen values in certain oil palm areas can be produced and visualized on the web so that they can be accessed easily and quickly for various purposes of oil palm management such as fertilization planning, recommendations, and monitoring.

Keywords: C-SAR sentinel-1; leaf nitrogen estimation; palm oil; random forest regression

Citation: Munir, S.; Seminar, K.B.; Sudradjat; Sukoco, H.; Buono, A. The Use of Random Forest Regression for Estimating Leaf Nitrogen Content of Oil Palm Based on Sentinel 1-A Imagery. *Information* **2023**, *14*, 10. https://doi.org/10.3390/info14010010

Academic Editors: Eftim Zdravevski, Petre Lameski and Ivan Miguel Pires

Received: 14 October 2022
Revised: 16 December 2022
Accepted: 21 December 2022
Published: 26 December 2022

1. Introduction

The palm oil industry has become a global industrial commodity through processing Crude Palm Oil (CPO) as a raw material for various food industries, oleochemicals, solid fuels, and renewable energy [1–3]. The oil palm tree (*Elaeis guineensis*), which produces palm oil, has spread and developed quickly in Southeast Asia. There are several oil palm plantations in this area, where 85% of the world's palm oil production comes from Indonesia and Malaysia [4,5].

Oil palm needs nutrients, especially nitrogen, phosphor, and potassium, for optimum growth and yields. Hence, knowledge of leaf nutrients is essential for the practical application of fertilizers [6]. Fertilization has a vital role in the management of oil palm plantations; fertilization costs in oil palm plantations are needed for around 50–70% of the operational costs of maintaining the plantation and 25% of the overall production costs [7]. The availability of adequate macro-nitrogen nutrients and the balance of macro-nitrogen nutrients with other macro-elements play an essential role in the success of oil palm growth [8–10]. The key to growing high-yield oil palm is having a sufficient supply of nutrients, particularly nitrogen. In contrast to annual crops, determining the nitrogen status of tall perennial crops such as oil palm is complicated and challenging due to age and complex nitrogen partitioning [8]. In addition, the cost of analyzing leaves through a laboratory is time-consuming and expensive for a large coverage area [11].

Sensor technology is now widely used to monitor, control, and optimize industrial processes [12], including the use of remote-sensing applications in agriculture [13–15]. By delivering reliable data on crop status throughout the season at various scales, remote sensing can support the adaptive evolution of agricultural methods to meet this significant challenge [16]. It is possible to examine the physical, chemical, and biological characteristics of vegetation, water, and soil using satellite photos captured by passive sensors such as multispectral images and active sensors such as C-Synthetic Aperture Radar (C-SAR) images [17–19]. The biggest challenge of using passive sensors in multispectral image acquisition in oil palm plantation areas with tropical climates such as Indonesia and Malaysia is obtaining clean images from cloud and shadow barriers [20]. Therefore, passive sensors cannot accurately estimate oil palm leaf nutrition in real time.

On the other hand, the performance of active sensors is unaffected by the weather, and the active sensor C-SAR can avoid cloud obstructions and obtain clear images of cloud shadows. C-SAR is a valuable and significant technology for monitoring crops and other agricultural targets [21]. Sentinel-1 is a C-SAR instrument and all-weather satellite owned by the European Space Agency (ESA), which performs day and night radar imaging missions for land and sea services [22].

In agriculture, remote-sensing data and Random Forest Regression (RFR) algorithms are applied to estimate nitrogen content in plants. Modeling estimation of nitrogen status using the RFR algorithm with multispectral data was investigated in rice [23], potato [24], wheat [25,26], corn [27,28], and sugarcane [29,30] plants; the results of the estimation model of nitrogen status were used as a guide for fertilization applications. In experiments on the prediction of nitrogen status in plants [25,28,30], the MLR model and machine learning (RFR, Support Vector Machine) were evaluated, and the results showed that RFR had excellent performance.

Through the use of Sentinel-1 radar imagery data and machine learning, this study aimed to determine a technique for estimating the nutritional of nitrogen in palm oil plantation. This machine learning algorithm model used data samples from several oil palm plantations in Indonesia, precisely in Sumatra, Java, and Kalimantan, to estimate the nitrogen nutrient status of oil palm leaves.

2. Literature Review

2.1. Nitrogen in Palm Oil

The nutrient elements N, P, K, Mg, B, Cu, and Zn are essential to the growth of oil palm leaves [31]. Nitrogen is a macronutrient needed by oil palm plants for protein formation, chlorophyll synthesis, and metabolic processes [32,33]. The application of nitrogen fertilizer can support oil palm fruit production by increasing the number of leaves on oil palm trees [10,32].

The concentration of the critical nutritional value of oil palm leaves has been widely used in its application. Leaf concentration value results from agronomic laboratory analysis of leaf samples number 17 for mature trees and leaf number 9 for young trees [34]. Geostatistical procedures help describe the variability of nutritional status spatially in oil palm plantations and could be used to design site-specific application fertilizer strategies [35].

Table 1 shows ranges of nitrogen (%) values of deficiency, optimum, and excess conditions concentrations of nitrogen nutrients in for young and mature oil palm leaves. Leaf analysis has been used as a general method for estimating the nitrogen status of oil palm plantations, using the oil palm's 17th midrib (>2.5 years) to relate the actual nitrogen status of the plant in the laboratory with the predicted nitrogen status of the plant. The 17th leaf's midrib's nutritional condition correlates more with oil palm production than other early leaves [11]. Figure 1 shows an example of taking the 17th leaf and recording the sampling's GPS coordinates.

Table 1. Nutrient concentration of nitrogen (%) in leaf 17 is associated with deficiency, optimum, and excess in young and mature palms [34].

Name	Deficiency	Optimum		Excess	Age
Young palms	<2.50	2.60	2.90	>3.10	≤6 years
Mature palms	<2.30	2.40	2.80	>3.00	>6 years

Figure 1. Taking of Leaf Sampling Units (LSU) in oil palm plantations. (**a**) The 17th leaf midribs were chosen and collected from oil palm plants. (**b**) Taking of four leaves from the flat endpoint in the centerpiece of the palm midribs. (**c**) Recording the GPS coordinates of the location of the palm trees.

2.2. C-Synthetic Aperture Radar (C-SAR) Sentinel 1-A

Sentinel-1 is an all-weather satellite owned by the European Space Agency (ESA), which performs day and night radar imaging missions for land and sea services. Sentinel-1 orbits the earth every 12 days. Sentinel-1 has a C-Synthetic Aperture Radar sensor for SAR data with a spatial resolution that operates in the C-band (5405 GHz). Its objective is to gather photographs for use in mapping polar and sea ice areas, tracking the marine environment, keeping an eye on the risk of ground movement, and vast mapping areas, including woods, land, and agricultural lands [22].

The Sentinel-1 satellite C-SAR sensor produces image products named Level-0, Level-1 SLC, Level-1 GRD, and Level-2 OCN. All products feature single polarization (VV or HH) for Wave mode and dual polarization (VV + VH or HH + HV) or single polarization (HH or VV) for SM, IW, and EW methods. Most data users typically have access to level-1 data.

Level-1 comprises two products, Single Look Complex (SLC) and Detected Ground Range (GRD). The Level-1 GRD product consists of SAR data that are detected, focused, multi-visible, and projected to the ground range using an Earth ellipsoid model. This Level-1 product has several capabilities:

- Backscattering-type classification.
- Tracking natural rainfall and changes in vegetation growth based on long-term time-series earth observations.
- Plant or vegetation classification.

Combining the backscatter coefficient band of C-SAR as an experimental basis has been widely used for vegetation monitoring research [36–40]. Kim et al. [36] introduced the Radar Vegetation Index (RVI) with a value of RVI = 8HV/(HH + VV + 2HV) to calculate the water content of rice and soybeans. Gonenc, Yamada, and Dey [37] compared RVI with NDVI (Normalized Difference Vegetation Index) produced from multispectral satellite pictures to evaluate the vegetation index. In another study, Mandal [39] modified the RVI

for dual polarization VV + VH with a value of RVI = 4VH/(VV + VH) to monitor vegetation growth using Sentinel-1 image data.

There have been several studies using sentinel-1 imagery on oil palm plantation objects, such as research on mapping the classification of oil palm vegetation using the Decision Tree algorithm based on the analysis of the VH and VV backscatter attribute values from the SAR [41]. Carolita et al. [42] used regression analysis to monitor the growth of oil palm age by analyzing the pattern of backscatter attribute values HH and VH. Xu et al. [43] use the Random Forest algorithm to distinguish the age of mature and young oil palms by analyzing the backscattering values of the SAR, vegetation index, and texture features. Miettinen and Xu [41,43] used a combination of VH and VV backscattering attributes in the form of VV-VH, VV/VH, and Normalized Difference Index values (NDI = (VV − VH)/(VV + VH)) to detect oil palm trees. The backscattering characteristic of C-SAR is displayed in Table 2 and can be used as input for classification and regression machine learning models.

Table 2. C-SAR backscatter features used in oil palm research.

Index	Backscatter Attribute Combinations	References
Polarization VV	VV	[41,42]
Polarization VH	VH	[41,42]
Dual-Polarization difference	VV − VH	[41,43]
Dual-Polarization Ratio	VV/VH	[41,43]
Radar Vegetation Index (RVI)	$\frac{4VH}{VV+VH}$	[37,39]
Normalized Difference Index (NDI)	$\frac{VV-VH}{VV+VH}$	[41,43]

The C-SAR system has an angle of view between the radar line of sight and the vertical antenna to the ground (Figure 2). The incidence angle (θ) is the angle between the pulse of radar energy and the line perpendicular to the earth's surface it hits. The angle of incidence θ on a target with a flat ground equal ($\theta = 90 - \gamma$) of the angle of depression.

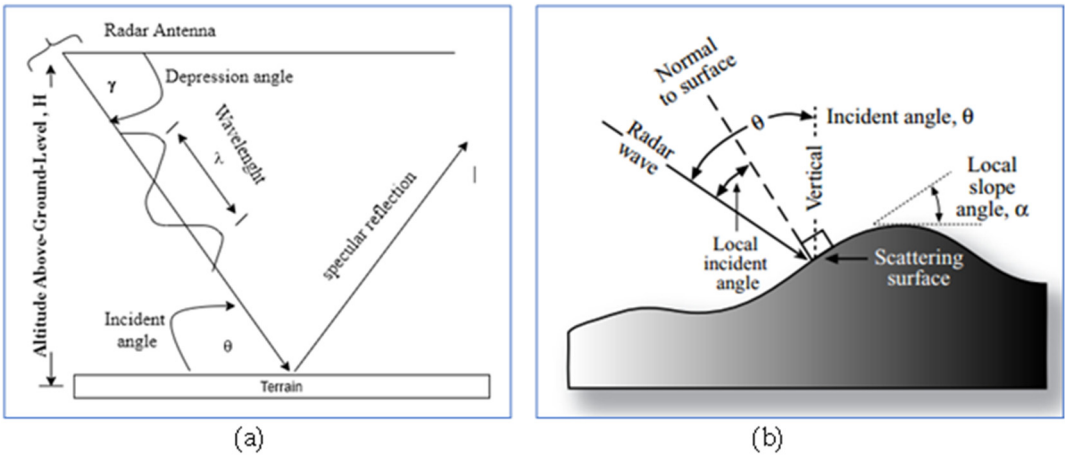

Figure 2. Targets on the surface of the earth with (**a**) flat terrain and (**b**) varied terrain that are exposed to radar wave radiation [44].

Since the data are directly derived from radar systems with various display geometries and changes with the local angle of incidence, normalization (normalized backscatter) is required to increase correctness and decrease bias [45,46].

2.3. Random Forest Regression

Random Forest is a supervised learning algorithm for solving classification and decision tree problems. The term "Random Forest" refers to a grouping of many decision trees

wherein each tree is dependent on the value of a random vector sampled separately and equally across all trees in the forest [47]. The random forest algorithm can significantly enhance the decision tree's predicting performance by mixing numerous decision trees [48].

This algorithm is random because of two factors [49], specifically: (1) a sample chunk of m variables from the original data set is chosen for each split node during decision tree formation, then the best one is used in that node; (2) each tree grows randomly on a different bootstrap sample from the training data. Random Forest is an effective machine learning tool for making predictions [47]. Harrison et al. [50] recommended the RFR Model for estimating nutrient concentrations using high-frequency sensor data. The technique is ideally suited to this application since it can be implemented for multivariate datasets that include highly skewed data, multi-collinearity between predictors, and nonlinear correlations between predictor variables and response variables. According to the study in [51], Random Forest Regression has the advantage of having a higher R-squared (R2) value than least squares regression.

2.4. Evaluation Model

The Mean Absolute Percentage Error (MAPE) and Mean Squared Error (MSE) evaluation models are often used in regression modeling that incorporates machine learning for model evaluation [52,53]. By computing the absolute percentage error of the mean, MAPE is used to measure the correctness of forecasts in the forecasting method statistically. In contrast, MSE is used to detect outlier data. Interpretation of MAPE and MSE values is best if it is close to 0, and vice versa is worse if it is close to $+\infty$.

In the formulas below, the X_i is the predicted value for *i*th, and the Y_i element is the actual value for *i*th. The regression method predicts the X_i element for the corresponding Y_i element of the ground truth dataset. The following equation represents MAPE and MSE [53]:

$$\text{MAPE} = \frac{1}{m}\sum_{i=1}^{m}\left|\frac{Y_i - X_i}{Y_i}\right| \quad (1)$$

$$\text{MSE} = \frac{1}{m}\sum_{i=1}^{m}(X_i - Y_i)^2 \quad (2)$$

Correctness is the inverse of MAPE, which is used to make it easier to understand the results of MAPE. The following is the correctness formula [54].

$$Correctness = 100 - \text{MAPE} \quad (3)$$

The K-Fold cross-validation method was applied in this study to evaluate the model data's level of quality and avoid overfitting the model. K-Fold Cross-validation randomly separates X data into K data, where each separated component has the same number. After randomizing the training and test data, a model was created from each K segment. The results of the evaluation of the quality of the best model were chosen as the best model [55].

3. Material and Methods

3.1. Research Design and Data

The research design (shown in Figure 3) was carried out in several major stages, including the steps of data collection, data preprocessing, data processing, a final evaluation of the prediction model, and making a map of the distribution of nitrogen status in oil palm plantation based on the resulting estimation model.

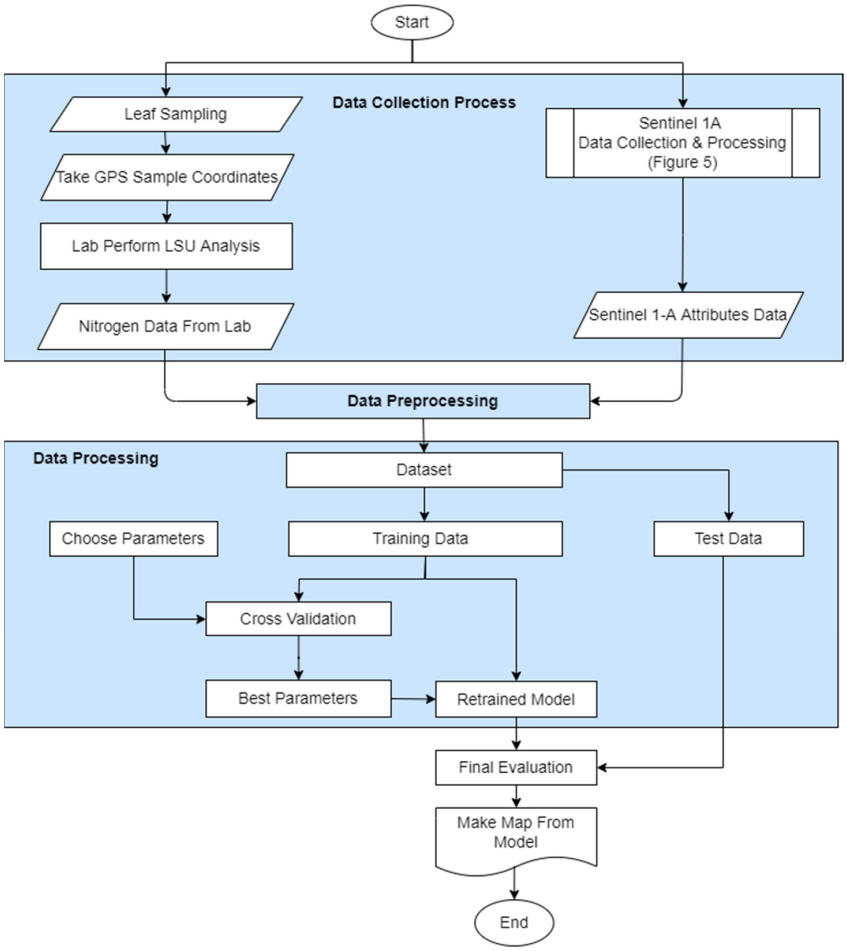

Figure 3. Research design.

At this step of the data collecting process, data was collected from two different sources: leaf sampling data from in situ and remote-sensing data obtained from sentinel satellite imaging data. During the leaf sampling data collection process, the GPS coordinates of the sampling locations were also recorded. This study collected 1116 LSU samples from various oil palm plantations in Sumatra, Java, and Kalimantan between 2018 to 2021, and all data were taken from oil palm plantations with mineral land types. Figure 4 shows the location for collecting the Leaf Sampling Unit (LSU) of oil palm leaves.

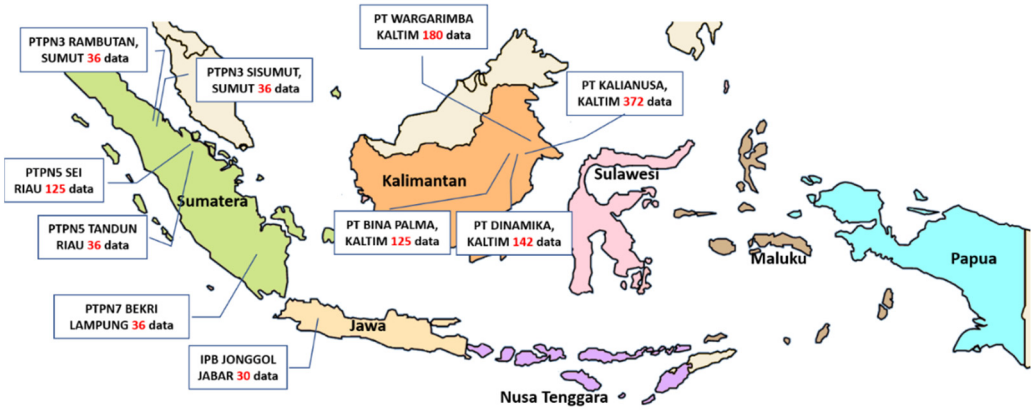

Figure 4. Map of Leaf Sampling Unit (LSU) locations in Sumatra, Java, and Kalimantan.

Then LSU was prepared and analyzed in the agronomy laboratory. The laboratory used is the Department of Agronomy and Horticulture of IPB University, accredited and certified by a national standard institution. The output of the agronomy laboratory is the nutritional value of nitrogen in oil palm leaves at the coordinate points in oil palm plantations. The Table 3 shows the sampling distribution of Mature Oil Palm Leaves for leaf midrib Number 17 for nitrogen nutrition.

Table 3. Example of LSU nitrogen status data from an agronomy laboratory.

ID Sample	In Situ Date	Latitude	Longitude	Nitrogen Total (%)
1	27 September 2021	0.68591	117.2197	2.66
2	27 September 2021	0.68591	117.2213	2.98
3	27 September 2021	0.67053	117.2430	2.15

At the stage of data collection through remote sensing, satellite image data were acquired following the collection period and the position of the GPS coordinates of leaf analysis data samples in the oil palm plot area. The primary data source for this study is the Sentinel-1A satellite imagery with Interferometric Wide (IW) observation acquisition mode from the GRD product features a wide swath (250 km) at a spatial resolution of 5×20 m. Table 4 displays the location and time information for the acquisition of Sentinel 1-A Imagery.

Table 4. Sentinel 1A image acquisition location and time.

Palm Plantations	Leaf Sampling Units	Image Acquisition Dates	Sentinel 1-A Image Files
PTPN 5 Sei, Riau	125	13 July 2018	S1A_IW_GRDH_1SDV_20180710T225636_20180710T225701_022738_0276E1_8857
PT. BPN, East Kalimantan	125	23 August 2018	S1A_IW_GRDH_1SDV_20180823T215033_20180823T215058_023379_028B27_2926
PTPN 5 Tandun, Riau	36	2 February 2019	S1A_IW_GRDH_1SDV_20190210T113323_20190210T113348_025866_02E10F_0E5C
PTPN 7 Bekri, Lampung	36	4 February 2019	S1A_IW_GRDH_1SDV_20190205T112333_20190205T112358_025793_02DE64_10F2
IPB Jonggol, Bogor	30	6 February 2019	S1A_IW_GRDH_1SDV_20190208T223343_20190208T223411_025844_02E035_470C
PTPN 3 Rambutan, North Sumatera	36	4 April 2019	S1A_IW_GRDH_1SDV_20190404T114214_20190404T114239_026640_02FD21_1430
PTPN 3 Sisumut, North Sumatera	36	5 April 2019	S1A_IW_GRDH_1SDV_20190411T113349_20190411T113414_026742_0300D4_BCE8
Kalianusa 1, East Kalimantan	50	2 March 2020	S1A_IW_GRDH_1SDV_20200302T215838_20200302T215903_031502_03A0D7_A2A7
Kalianusa 2, East Kalimantan	70	14 April 2020	S1A_IW_GRDH_1SDV_20200314T215838_20200314T215903_031677_03A6EA_5EE6
Kalianusa 1, East Kalimantan	50	28 November 2020	S1A_IW_GRDH_1SDV_20201128T215046_20201128T215111_035454_0424ED_2E00
Kalianusa 2, East Kalimantan	70	10 December 2020	S1A_IW_GRDH_1SDV_20201210T215046_20201210T215111_035629_042AF1_6E78
Kalianusa, Dinamika, Warga Rimba, East Kalimantan	454	24 September 2021	S1A_IW_GRDH_1SDV_20210924T215053_20210924T215118_039829_04B625_63CA

The following step involves feature extraction and data preprocessing on Sentinel 1 satellite radar image data collected of oil palm plantation plots. This process uses the Sentinel Application Platform (SNAP) software Sentinel-1 Toolbox and Quantum GIS (QGIS). The SNAP Sentinel-1 Toolbox application is a data processor, has a read and write a feature for image products, displays and analyzes data, and supports extensive data archives. It manages images obtained from Sentinel satellites.

SNAP's tasks include general raster, optical image processing, processing of radar data (calibration, speckle filtering, field correction), orthorectification, mosaic, data conversion, polarimetry, and interferometry. Feature extraction of sentinel-1A image data includes the stages of applying orbit file, radiometric calibration, geometric terrain correction, raster subset, and speckle filtering [56]. Quantum GIS is a geographic information system program that generates, manages, stores, displays, and analyzes spatial data. This study used QGIS Desktop to preprocess sentinel-1 satellite image data to obtain symbology maps, points, lines, polygons, categories, and labels.

The preprocessing workflow stage using the SNAP Toolbox consists of six processing steps (Apply Orbit File, Determine Subset Area, Perform Radiometric Calibrate, Perform Speckle Filtering, Perform Geometric Terrain Correction, and save into Geo TIFF format). The QGIS application consists of five processing steps (Create Layer from CSV File, Plantation Map, Latitude Longitude Coordinate Subset Area, Apply Layer Same CRS, Perform

Point Sampling Conversion to DB). Figure 5 shows the process of using SNAP and QGIS to feature extract and preprocess Sentinel 1-A image data.

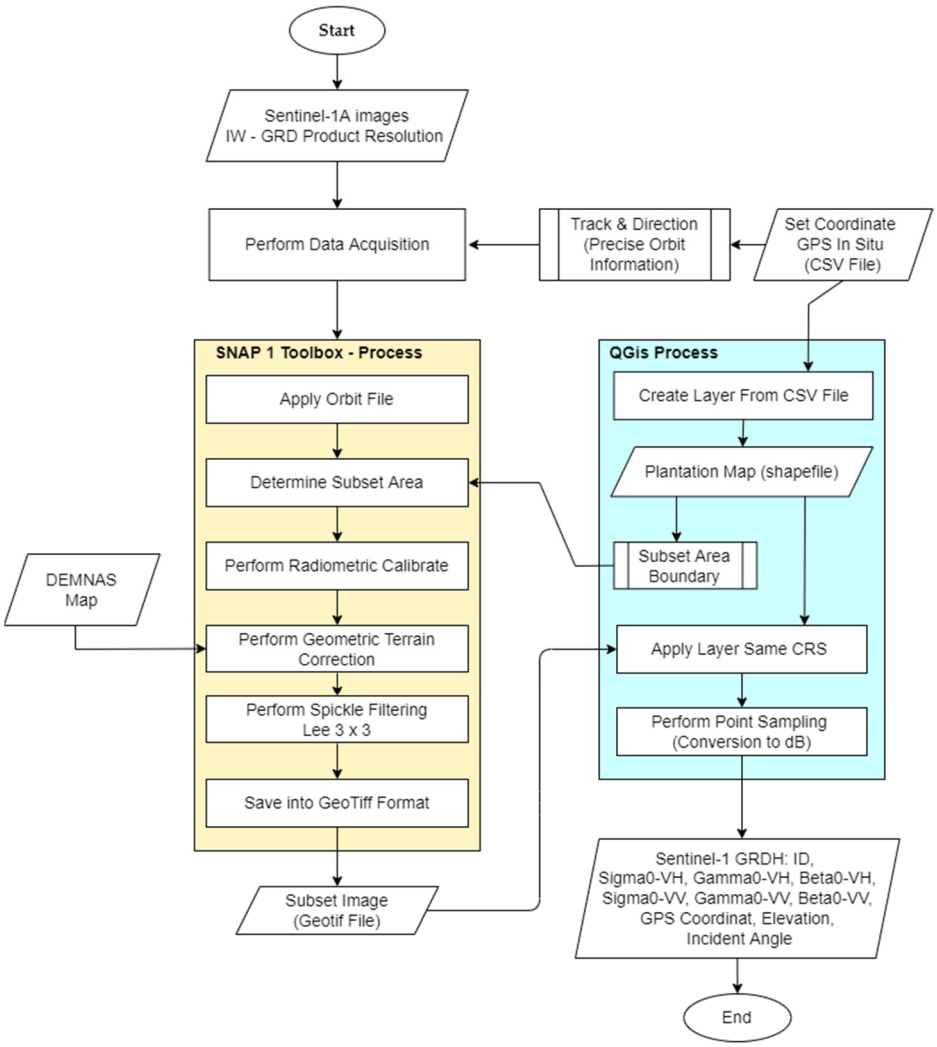

Figure 5. The steps of data preprocessing and feature extraction.

In this research, six backscatter coefficient bands were used as an experimental basis. The SNAP software tool that is for image processing produces six backscatter values consisting of band data values of VV (Sigma0_VH, Gamma0_VH, Beta0_VH) and VH (Sigma0_VV, Gamma0_VV, Beta0_VV), and six other additional attribute data consisting of variable data elevation, latitude, and longitude GPS coordinates, Local Incident Angle (LIA), Projected Local Incident Angle (PLIA), and Incident Angle Form Ellipsoid (IAFE). Table 5 provides examples of feature extraction operation of Sentinel 1A images.

Table 5. Example of backscatter data extracted from Sentinel 1A images.

ID Sample	Latitude	Longitude	Sigma0 VH	Gamma0 VH	Beta0 VH	Sigma0 VV	Gamma0 VV	Beta0 VV	Elevation	LIA	PLIA	IAFE
1	0.68591	117.2197	0.0261	0.0368	0.0371	0.1996	0.2813	0.2835	135.2062	54.3296	54.1659	44.7795
2	0.68591	117.2213	0.0491	0.0692	0.0698	0.2854	0.4020	0.4052	124.2401	36.6264	36.1343	44.7708
3	0.67053	117.2430	0.0283	0.0398	0.0404	0.3044	0.4278	0.4333	119.8694	44.4817	44.4398	44.6277

3.2. Data Preprocessing

Data transformation was carried out in the preprocessing stage to obtain variables required for the RFR models. The transformation method used is the brute force method, namely by trying various possible data transformations by performing a combination of mathematical operations from the Sentinel 1A data attribute. Some procedures and functions are quadratic operations, roots, cube roots, logarithmic operations, rounding operations, exponential and trigonometric functions. After the transformation process, of the 12 independent variables of the C-SAR attributes, 86 variables were obtained, which will be evaluated in the model using cross-validation. As indicated in Table 2, parameter determination also involved experimenting with different sentinel 1A backscatter index value combinations.

From the 86 variables, hyperparameter tuning [57] was carried out in Python using Scikit-Learn tools implementing a set of sensible default hyperparameters to select the most influencing variables. In the RFR, hyperparameter tuning includes the number of decision trees in the forest and the number of features considered by each tree when splitting a node. The parameters of a random forest are the variables and thresholds used to split each node learned during training. Hyperparameter tuning resulted in the five most-significant variables, namely Sigma0_VH, Gamma0_VH, Beta0_VH, RVI and LIA, as shown in Figure 6.

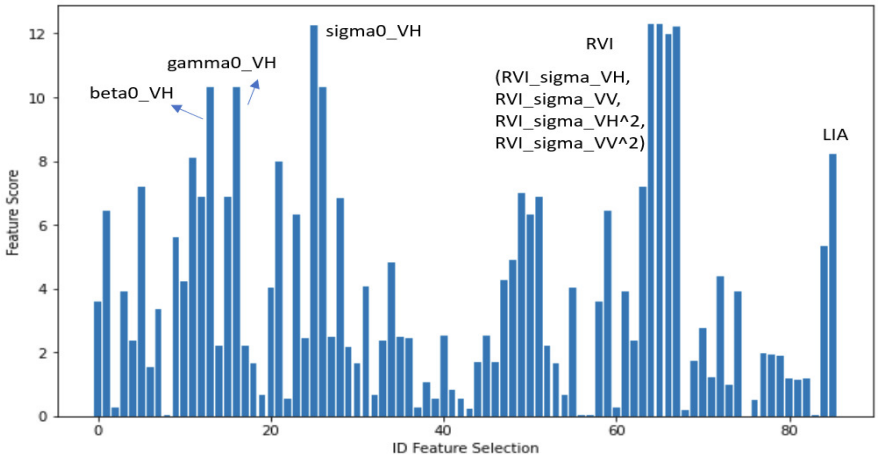

Figure 6. The feature scores obtained in hyperparameter tuning.

The dataset consists of 1116 data, divided into 80% training data and 20% basic test data. The built RFR algorithm was validated using K-Fold Cross-Validation, where all parts of the dataset can be used for training and testing, enabling a better evaluation of the performance of the proposed model. The first step of the validation method is to choose a value for k, then divide the dataset into several k folds, each of which represents a test dataset, and the other folds are used as training datasets. After that the model was trained on the training dataset and validated on the test dataset. This process was repeated k = 5 times. The table shows the validation results. In this study, the value of k = 5 was chosen referring to research in [54].

4. Results and Discussion

Table 6 shows the results of the k-fold validation (where k = 5) of the RFR model algorithm developed to estimate nitrogen nutrition in oil palm leaves. The validation produces 5 different correctness results when using correctness as a performance measure. The validation resulted in averaged MAPE, correctness, and MSE of 9.68%, 90.32%, and 11.03% respectively.

Table 6. The results of k-fold cross validation with RFR model training.

Fold	MAPE	Correctness	MSE	Fit Time (Seconds)	Score Time (Seconds)
1	8.36%	91.64%	7.51	1.31362	0.03392
2	7.33%	92.67%	5.75	1.25192	0.02194
3	7.78%	92.22%	7.55	1.31149	0.03619
4	11.31%	88.69%	19.52	1.19929	0.02798
5	13.59%	86.41%	14.81	1.34625	0.02798
Average	9.68%	90.32%	11.03	1.28451	0.02960

Figure 7 shows a comparison of the actual and predicted values of nitrogen in leaves using the base-test data. The *x*-axis represents the sample data IDs sorted by the original nitrogen values in increasing order, while the *y*-axis represents the nitrogen values.

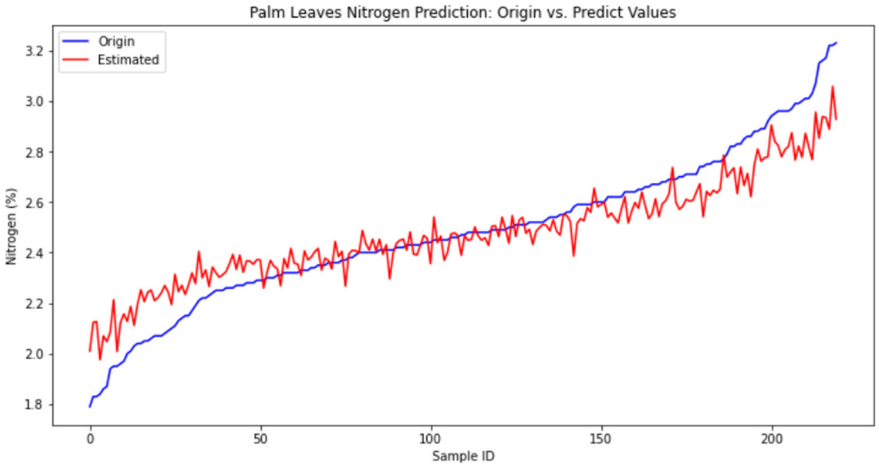

Figure 7. Line graph RFR model prediction: Origin vs. Estimated Values, with data sorted by the origin value of increasing order in the *x* axis.

After validation, the developed and validated model was tested on a completely new data set (130 data items) which was not in the previously sampled data set. Based on the results of this test, the performance of the model achieved correctness, MAPE, and MSE of 92.09%, 7.91%, and 6.80%, respectively. Note that the correctness obtained in this test is still within the range of the correctness values obtained by 5-fold validation (Table 6).

In this study, the RFR model was applied to estimate nitrogen concentrations. The model is used to visualize spatial maps using the QGIS application to represent the distribution (variation) of nutrients in an oil palm area. Oil palm plantations on the islands of Kalimantan and Java were used to demonstrate the visualization of Nitrogen distribution in the form of spatial maps, as shown in Figures 8–10.

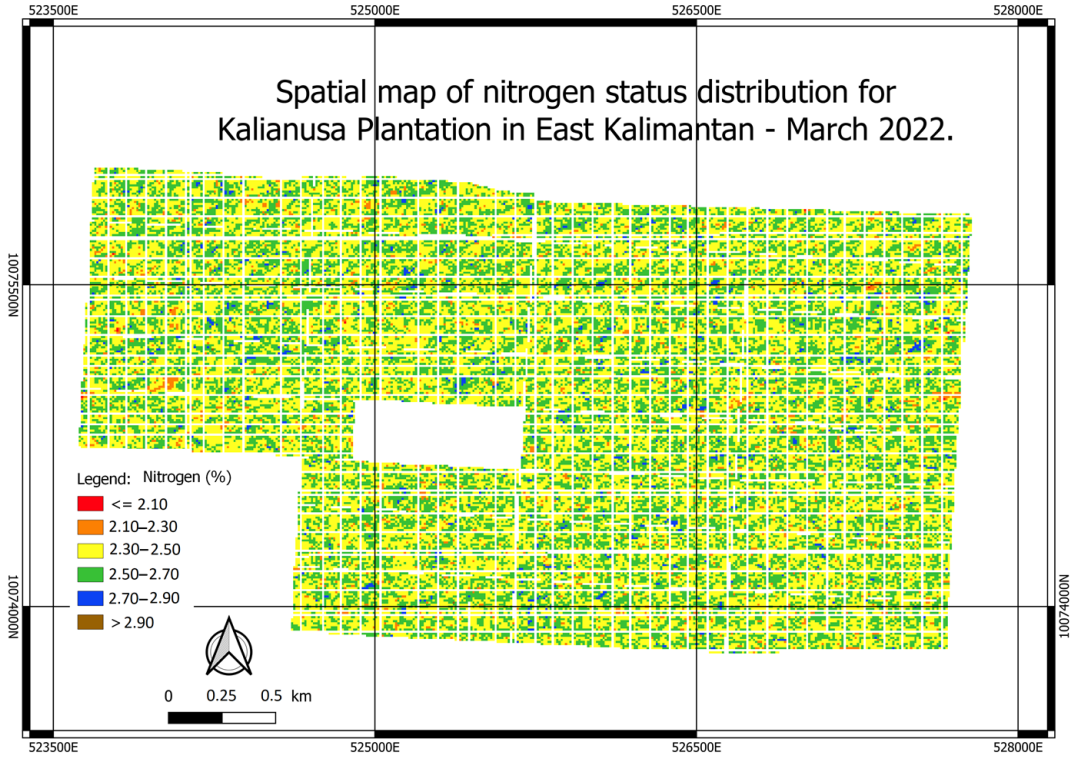

Figure 8. Spatial map of nitrogen status distribution for Kalianusa plantation in East Kalimantan March 2022.

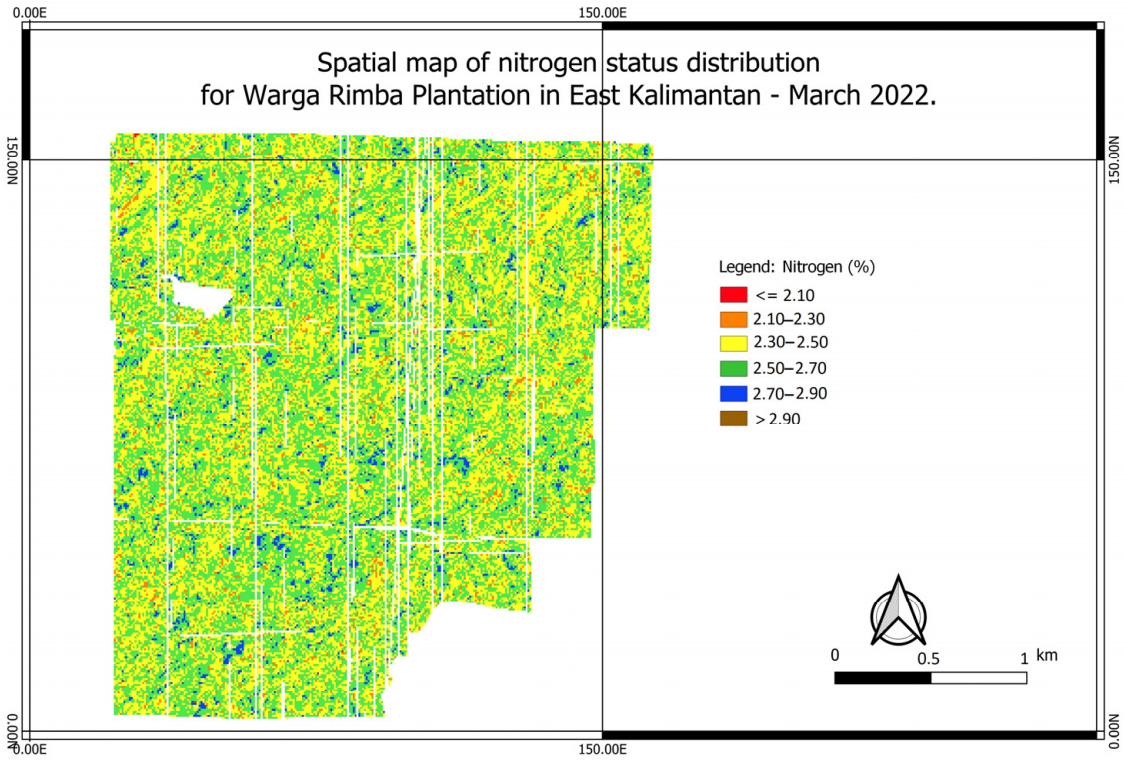

Figure 9. Spatial map of nitrogen status distribution for Warga Rimba plantation in East Kalimantan March 2022.

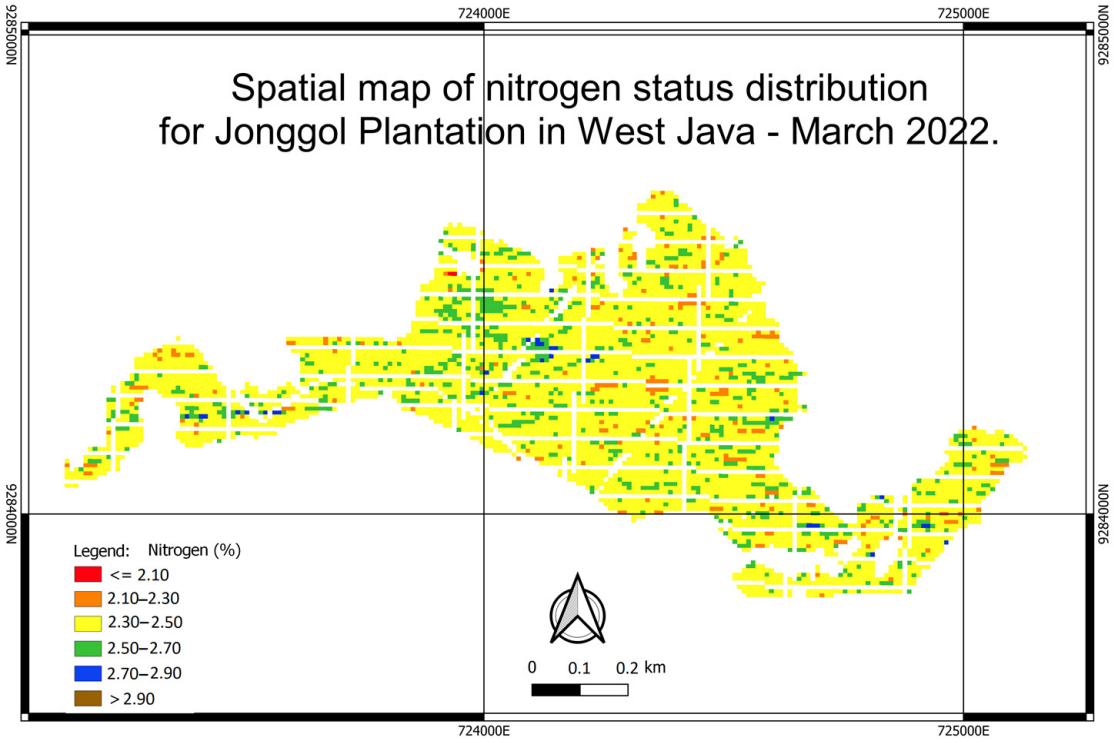

Figure 10. Spatial Map of nitrogen status distribution for Jonggol plantation in West Java March 2022.

Figure 8 is a map of the distribution of Nitrogen status in a demonstration plot in an oil palm plantation belonging to the Kalianusa Oil Palm Company located in District Rantau Pulung, East Kutai Regency, East Kalimantan. At this plantation location, oil palm trees have been produced (mature palms). Figure 8 shows that green and yellow colors dominate the map, meaning that the concentration of Nitrogen nutrients (%) in oil palm plantations is at the optimum value, in the 2.3–2.7% range.

Figure 9 is a map of the distribution of Nitrogen in the oil palm plantations belonging to the Warga Rimba palm oil company, located in District Karangan, East Kutai Regency, East Kalimantan. It can be seen that the map of oil palm plantations is dominated by green and yellow colors, which means that the nutrient concentration value of Nitrogen (%) is in the range of 2.3–2.7%, which is at optimum conditions.

Figure 10 is a map of nitrogen distribution in an oil palm plantation located in the Jonggol area, West Java. The Jonggol oil palm plantation's map is dominated by a yellow color, meaning that the concentration of Nitrogen (%) is 2.3–2.5%, which is still considered to have enough nitrogen levels for plants.

5. Conclusions

In this research, we developed and tested a model using machine learning for estimating oil palm leaf nitrogen status based on Sentinel-1 A imagery using Random Forest Regression. The number of data samples (LSU) used to develop the proposed model was 1116 LSUs taken from Sumatra, Java, and Kalimantan islands. The performance evaluation of the model obtained the averaged MAPE, correctness, and MSE of 9.68%, 90.32%, and 11.03 respectively. The validated model was also tested for new real data (130 data items) with the testing results with averaged MAPE, correctness, and MSE of 7.91%, 92.09%, and 6.80 respectively.

The results of the estimation model of the leaf nitrogen content of oil palm can be produced and visualized in the form of a spatial map displaying the distribution of nitrogen nutrients on the web platform, which can be easily and quickly accessed for various oil palm management purposes such as fertilization planning, recommendation, and monitoring. For future research, oil palm plantations in Indonesia are distributed in several geographical

areas consisting of mineral and peat soil; therefore, we need to develop an estimation model for the other macro and micronutrients and also for the peat soils of the oil palm plantation.

The future use of this proposed model is to enhance the application of precision agriculture for site-specific fertilization recommendations for oil palm. In addition, the spatial map visualized by the proposed system can be used to increase transparency and observation of oil palm for better field management tasks.

Author Contributions: Conceptualization, K.B.S. and S.M.; methodology, K.B.S. and S.M.; software, S.M.; validation, K.B.S., S., H.S. and A.B.; formal analysis, K.B.S., S., H.S. and A.B.; investigation, K.B.S. and H.S.; resources, S.M.; data curation, S.M.; writing—original draft preparation, S.M. and K.B.S.; writing—review and editing, S.M. and K.B.S.; visualization, S.; supervision, K.B.S.; project administration, K.B.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hamzah, N.; Tokimatsu, K.; Yoshikawa, K. Solid fuel from oil palm biomass residues and municipal solid waste by hydrothermal treatment for electrical power generation in Malaysia: A review. *Sustainability* **2019**, *11*, 1060. [CrossRef]
2. Ismail, M.I.; Yunus, N.A.; Hashim, H. The challenges and opportunities of solar thermal for palm oil industry in Malaysia. *Chem. Eng. Trans.* **2020**, *78*, 601–606.
3. Khatiwada, D.; Palmén, C.; Silveira, S. Evaluating the palm oil demand in Indonesia: Production trends, yields, and emerging issues. *Biofuels* **2021**, *12*, 135–147. [CrossRef]
4. Purnomo, H.; Okarda, B.; Dermawan, A.; Ilham, Q.P.; Pacheco, P.; Nurfatriani, F.; Suhendang, E. Reconciling oil palm economic development and environmental conservation in Indonesia: A value chain dynamic approach. *For. Policy Econ.* **2020**, *111*, 102089. [CrossRef]
5. Ayompe, L.M.; Schaafsma, M.; Egoh, B.N. Towards sustainable palm oil production: The positive and negative impacts on ecosystem services and human wellbeing. *J. Clean. Prod.* **2020**, *278*, 123914. [CrossRef]
6. Behera, S.K.; Suresh, K.; Rao, B.N.; Manoja, K.; Manorama, K. Soil nutrient status and leaf nutrient norms in oil palm (Elaeis Guineensis Jacq.) plantations grown in the west coastal area of India. *Commun. Soil Sci. Plant Anal.* **2016**, *47*, 255–262. [CrossRef]
7. Comte, I.; Whalen, J.K.; Gru, O. Agricultural practices in oil palm plantations and their impact on hydrological changes, nutrient fluxes and water quality in Indonesia: A review. *Adv. Agron.* **2012**, *116*, 71–124.
8. Amirruddin, A.D.; Muharam, F.M.; Paing, T.N.; Singh, D.S.K.; Yusoff, M.M. Nitrogen effects on growth and spectral characteristics of immature and mature oil palms. *Asian J. Plant Sci.* **2017**, *16*, 200–210. [CrossRef]
9. Pardon, L.; Ian Huth, N.; Netelenbos Nelson, P.; Banabas, M.; Gabrielle, B.; Bessou, C. Yield and nitrogen losses in oil palm plantations: Main drivers and management trade-offs determined using simulation. *Field Crop. Res.* **2017**, *210*, 20–32. [CrossRef]
10. Mohidin, H.; Man, S.; Hanafi, M.M.; Rufai, S.; Idris, J.; Fonguimgo, T.; Idris, A.S.; Rafii, M. Optimum levels of N, P, and K nutrition for oil palm seedlings grown in tropical peat soil. *J. Plant Nutr.* **2019**, *42*, 1461–1471. [CrossRef]
11. Rendana, M.; Abdul Rahim, S.; Lihan, T.; Mohd Razi Idris, W.; Ali Rahman, Z. A Review of Methods for Detecting Nutrient Stress of Oil Palm in Malaysia. *J. Appl. Environ. Biol. Sci.* **2015**, *5*, 60–64. Available online: http://www.textroad.com (accessed on 1 August 2022).
12. Jiang, Y.; Yin, S.; Dong, J.; Kaynak, O. A review on soft sensors for monitoring, control, and optimization of industrial processes. *IEEE Sens. J.* **2021**, *21*, 12868–12881. [CrossRef]
13. Karthikeyan, L.; Chawla, I.; Mishra, A.K. A Review of Remote Sensing Applications in Agriculture for Food Security: Crop Growth and Yield, Irrigation, and Crop Losses. *J. Hydrol.* **2020**, *586*, 124905. [CrossRef]
14. Radoglou-Grammatikis, P.; Sarigiannidis, P.; Lagkas, T.; Moscholios, I. A compilation of UAV applications for precision agriculture. *Comput. Netw.* **2020**, *172*, 107148. [CrossRef]
15. Sishodia, R.P.; Ray, R.L.; Singh, S.K. Applications of remote sensing in precision agriculture: A review. *Remote. Sens.* **2020**, *12*, 3136. [CrossRef]
16. Weiss, M.; Jacob, F.; Duveiller, G. Remote sensing for agricultural applications: A meta-review. *Remote. Sens. Environ.* **2020**, *236*, 111402. [CrossRef]
17. Chuvieco, E.; Huete, A. *Fundamentals of Satellite Remote Sensing*; CRC Press: Boca Raton, FL, USA, 2009; pp. 1–419.
18. Lillesand, T.; Kiefer, R.W.; Chipman, J. *Remote Sensing and Image Interpretation*; John Wiley & Sons, Inc: Hoboken, NJ, USA, 2015; p. 770.

19. Harry, I.; Seminar, K.B.; Hermawan, W.; Saptomo, S.K. Prediction of soil surface water content and soil workability on an unplanted sugarcane farm area using Sentinel-1A data. *Information* **2022**, *13*, 493.
20. Kaliana, I.; Seminar, K.B.S. Development of a Decision Support System for Oil Palm Fertilizer Requirement Based on Precision Agriculture, IPB University Scientific Repository. 2018. Available online: https://repository.ipb.ac.id/bitstream/handle/123456789/92621/2018ika.pdf (accessed on 1 August 2022).
21. Liu, C.-A.; Chen, Z.-X.; Shao, Y.; Chen, J.-S.; Hasi, T.; Pan, H.-Z. Research advances of SAR remote sensing for agriculture applications: A review. *J. Integr. Agric.* **2019**, *18*, 506–525. [CrossRef]
22. Vincent, P.; Bourbigot, M.; Johnsen, H.; Piantanida, R. Sentinel-1 Product Specification. 2020. Available online: https://sentinel.esa.int/documents/247904/1877131/Sentinel-1-Product-Definition (accessed on 1 August 2022).
23. Zha, H.; Miao, Y.; Wang, T.; Li, Y.; Zhang, J.; Sun, W. Sensing-based rice nitrogen nutrition index prediction with machine learning. *Remote Sens.* **2020**, *12*, 215. [CrossRef]
24. Peng, J.; Manevski, K.; Kørup, K.; Larsen, R.; Andersen, M.N. Random Forest regression results in accurate assessment of potato nitrogen status based on multispectral data from different platforms and the critical concentration approach. *Field Crop. Res.* **2021**, *268*, 108158. [CrossRef]
25. Liang, L.; Di, L.; Huang, T.; Wang, J.; Lin, L.; Wang, L.; Yang, M. Estimation of leaf nitrogen content in wheat using new hyperspectral indices and a random forest regression algorithm. *Remote Sens.* **2018**, *10*, 1940. [CrossRef]
26. Liu, Y.; Cheng, T.; Zhu, Y.; Tian, Y.; Cao, W.; Yao, X.; Wang, N. Comparative analysis of vegetation indices, non-parametric and physical retrieval methods for monitoring nitrogen in wheat using UAV-based multispectral imagery. In Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10–15 July 2016; pp. 7350–7353.
27. Zhao, B.; Duan, A.; Ata-Ul-Karim, S.T.; Liu, Z.; Chen, Z.; Gong, Z.; Zhang, J.; Xiao, J.; Liu, Z.; Qin, A.; et al. Exploring new spectral bands and vegetation indices for estimating nitrogen nutrition index of summer maize. *Eur. J. Agron.* **2018**, *93*, 113–125. [CrossRef]
28. Lee, H.; Wang, J.; Leblon, B. Using linear regression, random forests, and support vector machine with unmanned aerial vehicle multispectral images to predict canopy nitrogen weight in corn. *Remote Sens.* **2020**, *12*, 2071. [CrossRef]
29. Abdel-Rahman, E.M.; Ahmed, F.B.; Ismail, R. Random forest regression and spectral band selection for estimating sugarcane leaf nitrogen concentration using EO-1 Hyperion hyperspectral data. *Int. J. Remote Sens.* **2013**, *34*, 712–728. [CrossRef]
30. Soltanikazemi, M.; Minaei, S.; Shafizadeh-Moghadam, H.; Mahdavian, A. Field-scale estimation of sugarcane leaf nitrogen content using vegetation indices and spectral bands of Sentinel-2: Application of random forest and support vector regression. *Comput. Electron. Agric.* **2022**, *200*, 107130. [CrossRef]
31. Ng, S.K. Nutrition and nutrient management of the oil palm—New thrust for the future perspective. *Int. Symp.* **2002**, 415–429.
32. Corley, R.H.V.; Tinker, P.B. *The Oil Palm*; John Wiley & Sons: Hoboken, NJ, USA, 2016; p. 692.
33. Apichatmeta, K.; Sudsiri, C.J.; Ritchie, R.J. Photosynthesis of oil palm (Elaeis guineensis). *Sci. Hortic.* **2017**, *214*, 34–40. [CrossRef]
34. Fairhurst, B.T.H.; Mutert, E. Interpretation and management of oil palm leaf analysis data. *Better Crop. Int.* **1999**, *13*, 48.
35. Behera, S.K.; Suresh, K.; Ramachandrudu, K.; Manorama, K.; Rao, B.N. Mapping spatial variability of leaf nutrient status of oil palm (Elaeis guineensis Jacq.) plantations in India. *Crop. Pasture Sci.* **2016**, *67*, 109–116. [CrossRef]
36. Kim, Y.; Jackson, T.; Bindlish, R.; Lee, H.; Hong, S. Radar vegetation index for estimating the vegetation water content of rice and soybean. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 564–568.
37. Gonenc, A.; Ozerdem, M.S.; Acar, E. Comparison of NDVI and RVI vegetation indices using satellite images. In Proceedings of the 2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics), Istanbul, Turkey, 16–19 July 2019; p. 22.
38. Yamada, Y. Comparison radar vegetation index (RVI) with conventional methods for paddy rice field, lotus pond and soybean. *Int. Geosci. Remote Sens. Symp.* **2019**, *10*, 7247–7250.
39. Ratha, D.; Mandal, D.; Dey, S.; Bhattacharya, A.; Frery, A.; Rao, Y.S.; McNairn, H. New vegetation indices for full and compact polarimetric sar data: In preparation for the radarsat constellation mission (RCM). *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *IV-3/W2-20*, 41–46. [CrossRef]
40. Mandal, D.; Kumar, V.; Ratha, D.; Dey, S.; Bhattacharya, A.; Lopez-Sanchez, J.M.; McNairn, H.; Rao, Y.S. Dual polarimetric radar vegetation index for crop growth monitoring using sentinel-1 SAR data. *Remote Sens. Environ.* **2020**, *247*, 111954. [CrossRef]
41. Miettinen, J.; Liew, S.C.; Kwoh, L.K. Usability of sentinel-1 dual polarization C-band data for plantation detection in Insular Southeast Asia. In Proceedings of the ACRS 2015—36th Asian Conference on Remote Sensing: Fostering Resilient Growth in Asia, Manila, Philippines, 19–23 October 2015.
42. Carolita, I.; Darmawan, S.; Permana, R.; Dirgahayu, D.; Wiratmoko, D.; Kartika, T.; Arifin, S. Comparison of optic landsat-8 and sar sentinel-1 in oil palm monitoring, case study: Asahan, north Sumatera, Indonesia. *IOP Conf. Ser. Earth Environ. Sci.* **2019**, *280*, 012015. [CrossRef]
43. Xu, K.; Qian, J.; Hu, Z.; Duan, Z.; Chen, C.; Liu, J.; Sun, J.; Wei, S.; Xing, X. A new machine learning approach in detecting the oil palm plantations using remote sensing data. *Remote Sens.* **2021**, *13*, 236. [CrossRef]
44. Jensen, J.R. *Remote Sensing of the Environment: An Earth Resource Perspective Second Edition*; Pearson Education Limited: Harlow, UK, 2014; Volume 1, pp. 333–378.
45. Gouveia, N.D.A.; Alves, F.C.; Pereira, L.D.O. Pre-processing of Sentinel-1 C-band SAR images based on incidence angle correction for dark target detection. *Remote Sens. Lett.* **2019**, *10*, 939–948. [CrossRef]

46. Kaplan, G.; Fine, L.; Lukyanov, V.; Manivasagam, V.S.; Tanny, J.; Rozenstein, O. Normalizing the local incidence angle in sentinel-1 imagery to improve leaf area index, vegetation height, and crop coefficient estimations. *Land* **2021**, *10*, 680. [CrossRef]
47. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
48. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning with Applications in R*; Springer: New York, NY, USA, 2013.
49. Rodriguez-Galiano, V.F.; Chica-Olmo, M.; Chica-Rivas, M. Predictive modelling of gold potential with the integration of multisource information based on random forest: A case study on the Rodalquilar area, Southern Spain. *Int. J. Geogr. Inf. Sci.* **2014**, *28*, 1336–1354. [CrossRef]
50. Harrison, J.W.; Lucius, M.A.; Farrell, J.L.; Eichler, L.W.; Relyea, R.A. Prediction of stream nitrogen and phosphorus concentrations from high-frequency sensors using random forests regression. *Sci. Total Environ.* **2021**, *763*, 143005. [CrossRef]
51. Schonlau, M.; Zou, R.Y. The random forest algorithm for statistical learning. *Stata J.* **2020**, *20*, 3–29. [CrossRef]
52. De Myttenaere, A.; Golden, B.; Le Grand, B.; Rossi, F. Mean absolute percentage error for regression models. *Neurocomputing* **2016**, *192*, 38–48. [CrossRef]
53. Chicco, D.; Warrens, M.J.; Jurman, G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Comput. Sci.* **2021**, *7*, e623. [CrossRef] [PubMed]
54. Budiman, R.; Seminar, K.B. Sudradjat Development of Soil Nitrogen Estimation System in Oil Palm Land with Sentinel-1 Image Analysis Approach. 2021, pp. 153–165. Available online: https://link.springer.com/10.1007/978-3-030-88259-4_11 (accessed on 1 August 2022).
55. Rodríguez, J.D.; Pérez, A.; Lozano, J.A. Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 569–575. [CrossRef]
56. Filipponi, F. Sentinel-1 GRD preprocessing workflow. *Proceedings* **2019**, *18*, 11.
57. Schratz, P.; Muenchow, J.; Iturritxa, E.; Richter, J.; Brenning, A. Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. *Ecol. Model.* **2019**, *406*, 109–120. [CrossRef]