Article

Remotely Piloted Aircraft and Random Forest in the Evaluation of the Spatial Variability of Foliar Nitrogen in Coffee Crop

Diego Bedin Marin [1,*], Gabriel Araújo e Silva Ferraz [1], Paulo Henrique Sales Guimarães [2], Felipe Schwerz [1], Lucas Santos Santana [1], Brenon Dienevam Souza Barbosa [1], Rafael Alexandre Pena Barata [1], Rafael de Oliveira Faria [1], Jessica Ellen Lima Dias [1], Leonardo Conti [3] and Giuseppe Rossi [3]

[1] Agricultural Engineering Department, Federal University of Lavras, Lavras 37200-000, Brazil; gabriel.ferraz@ufla.br (G.A.eS.F.); felipe.schwerz@ufla.br (F.S.); lucas.santana1@estudante.ufla.br (L.S.S.); bdiennevan@estudante.ufla.br (B.D.S.B.); rafael.barata@estudante.ufla.br (R.A.P.B.); rafael.faria@ufla.br (R.d.O.F.); jessica.dias@estudante.ufla.br (J.E.L.D.)
[2] Statistics Department, Federal University of Lavras, Lavras 37200-000, Brazil; paulo.guimaraes@ufla.br
[3] Department of Agriculture, Food, Environment and Forestry, University of Florence, Via San Bonaventura, 13-50145 Florence, Italy; leonardo.conti@unifi.it (L.C.); giuseppe.rossi@unifi.it (G.R.)
* Correspondence: diego-marin@ufla.br; Tel.: +55-991-284-644

Abstract: The development of approaches to determine the spatial variability of nitrogen (N) into coffee leaves is essential to increase productivity and reduce production costs and environmental impacts associated with excessive N applications. Thus, this study aimed to assess the potential of the Random Forest (RF) machine learning method applied to vegetation indices (VI) obtained from Remotely Piloted Aircraft (RPA) images to measure the N content in coffee plants. A total of 10 VI were obtained from multispectral images by a camera attached to a rotary-wing RPA. The RGB orthomosaic was used to determine sampling points at the crop area, which were ranked by N levels in the plants as deficient, critical, or sufficient. The chemical analysis of N content in the coffee leaves, as well as the VI values in sample points, were used as input parameters for the image training and its classification by the RF. The suggested model has shown global accuracy and a kappa coefficient of up to 0.91 and 0.86, respectively. The best results were achieved using the Green Normalized Difference Vegetation (GNDVI) and Green Optimized Soil Adjusted Vegetation Index (GOSAVI). In addition, the model enabled the evaluation of the spatial distribution of N in the coffee trees, as well as quantification of N deficiency in the crop for the whole area. The GNDVI and GOSAVI allowed the verification that 22% of the entire crop area had plants with N deficiency symptoms, which would result in a reduction of 78% in the amount of N applied by the producer.

Keywords: machine learning; vegetation indices; unmanned aerial vehicle; nitrogen management; RGB camera

1. Introduction

To succeed in coffee production, good fertility of the soil, along with an intense fertility care program is necessary [1]. Without the proper assessment of soil fertility, nutritional deficiencies will affect the survival and productivity of coffee plants [2,3]. Among the essential nutrients that are necessary for coffee crops, nitrogen (N) is considered the one that limits the development and productivity of the coffee [4]. The adequate administration of N promotes an increase of pairs of leaves and branches per node in the plants, which is immediately related to coffee [3]. Additionally, the N determines the plant settlement and root development, influencing numerous aspects of the plant's health [5].

Meanwhile, the proper management of N in coffee crops is still a challenging task for most producers. Cases of both excessive and deficient N application are problems in coffee production [6]. Excessive application of nitrogen fertilizers is a common practice among small and big producers, being the main cause of low-efficiency fertilization, reducing the

quality of the production and lowering the profit for farmers [7,8]. Furthermore, the excess of N can cause susceptibility for plague attacks on the plants [9,10]. Environmental damage has been reported in the literature caused by leaching and volatilization of excessive N. In contrast, N deficiency causes fall of leaves, fewer leafy plants, smaller sized fruits, and drying in branches [11,12].

Consequently, to achieve satisfactory levels of N in coffee plants, producers apply nitrogen fertilizers based on a calendar or samples occasionally collected from soil and leaves for laboratory analysis [5,6]. In extreme cases of nutrient deficiency, symptoms like the yellowing of leaves start to show and are used for reactive decision-making [5,13]. However, these methods are impractical in large areas, requiring a lot of human resources, materials, sample collection, laboratory testing, and data processing [14,15]. Additionally, only a few plants are randomly sampled, making it possible that it does not represent the spatial variability properly across the crops [16]. In the long term, these methods are not suitable for balancing the nitrogen need of the coffee plants, resulting in an unstable production with productiveness deficits [5].

As an alternative to those methods, remote sensing technologies have been used to estimate nitrogen content in coffee plants. For example, [5] used vegetation indices obtained from Sentinel-2 MultiSpectral Instrument (MSI) images and machine learning to model the leaf nitrogen content in coffee bushes. The results have shown precision up to $R^2 = 0.78$. The study performed by [8] obtained precision up to $R^2 = 0.81$ while estimating nitrogen levels in coffee plants for different growing stages and field conditions. However, these studies have not explored the potential of Remotely Piloted Aircraft combined with machine learning techniques.

In the past few years, remote sensing based on Remotely Piloted Aircraft (RPA) has rapidly developed, due to its low cost, operation facility, and wide field of view [17,18]. With a nondestructive approach, data from remote sensing obtained using multispectral images from RPAs is commonly used to monitor nitrogen content in vegetation [19–22]. Still, due to the amount of data that is generated, remote sensing combined with high-resolution images requires a robust technique, such as methods of machine learning [23]. These methodologies, when applied to agriculture, achieve advantages, like the ability to solve nonlinear problems using datasets from various sources [24–26] and discover information hidden on the data [27]. There are still few studies in the literature contemplating RPA, management of nitrogen fertilizers, and machine learning [23,28–30]. For coffee production, [31] study used RGB-based vegetation indices from RPA images with Random Forest to monitor the nitrogen status of the plants. However, the models developed in the study were not capable of explaining and predicting the spatial variability of the nitrogen in the coffee plants. Therefore, a solution to these unsatisfactory results is possibly to explore vegetation indices from multispectral images.

Thus, based on the hypothesis that the learning method Random Forest applied to vegetation indices obtained from RPAs can contribute to the more efficient management of the nitrogen content in coffee crops, the objectives of this study were (i) to map the spatial variability of the nitrogen content in the coffee plantations, (ii) quantify the deficiency of nitrogen in coffee plants, and (iii) determine the most efficient vegetation index to predict N content in coffee plants using the Random Forest (RF) machine learning method.

2. Materials and Methods

The methodology proposed in this study is briefly explained in Figure 1. In the first step, multispectral images of the study area were acquired by a multispectral camera coupled to an RPA. These images were processed, creating the orthomosaics, to then later calculate the vegetation indices. In the second step, the RGB mosaic composition was used to define sampling points in the study area, used for analyzing the chemical content of N in the coffee leaves. Finally, in the third step, values of the vegetation indices and chemical analysis of N in the leaves in each sample point were used as input parameters to calibrate the algorithm Random Forest and classify the images in three N content categories.

The accuracy of the classification was measured by the overall accuracy metrics, kappa coefficient, receiver operating characteristic (ROC) curve, and area under the curve (UAC). The methodology is further detailed in the following topics.
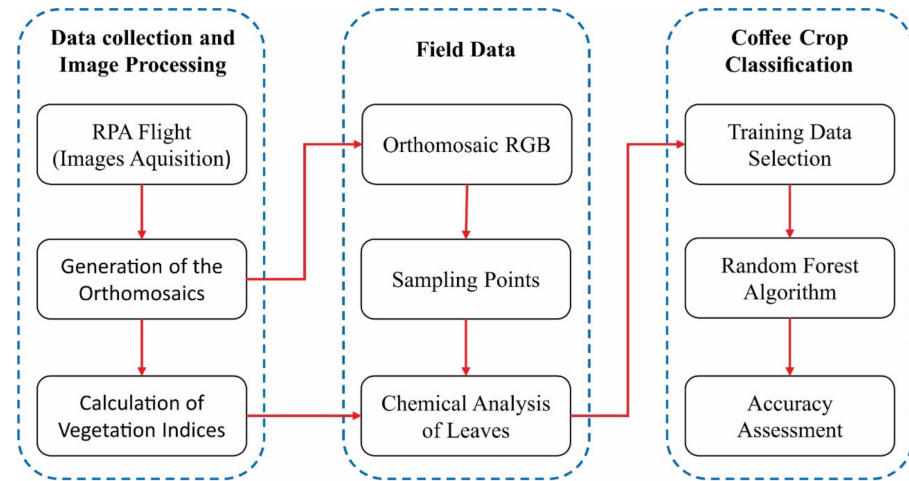
Figure 1. Workflow used in the methodology of this study.

2.1. Study Site

The study was carried out in a field located in the commercial farm Bom Jardim, in the city of Santo Antonio do Amparo, state of Minas Gerais, Brazil, geographic coordinates 21°01′09.69″S and 44°55′45.03″W, and the average altitude of 935 m. This plantation occupies an area of 1.5 ha cultivating coffee plants (*Coffea arabica* L.), Catucai Amarelo 2SL cultivar aging 3 years old, spaced by 3.5 m from centerlines, and 0.5 m plant to plant, adding 5700 plants ha^{-1} (Figure 2).
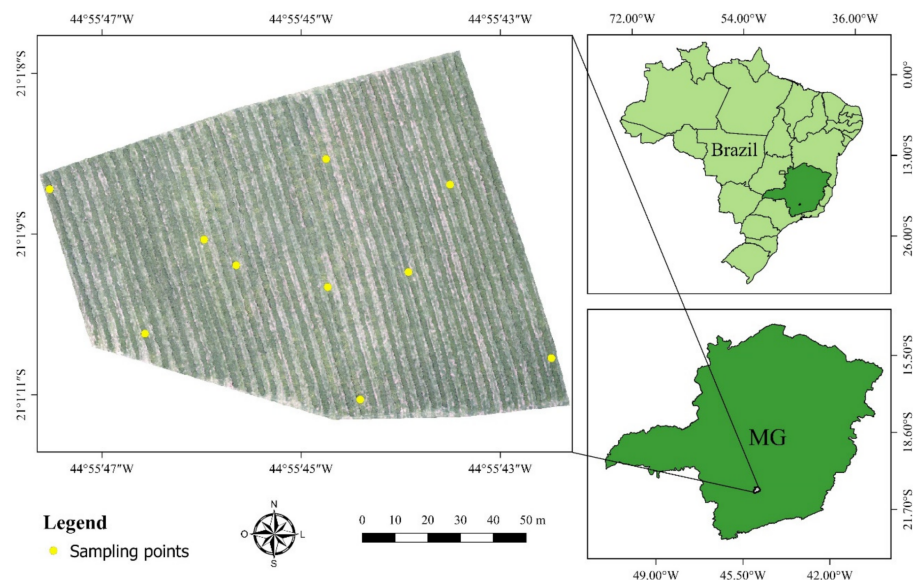
Figure 2. The geographic location of the study area. MG is the State of Minas Gerais, Brazil.

According to the Köppen–Geiser climate classification, the region has a Cwa climate, humid subtropical, with hot and humid summers and cold and dry winters, with an annual average air temperature of 19.8 °C and average annual total rainfall of 1670 mm [32]. The climatic data [33] of the study area from the period of the analysis is shown in Table 1.

Table 1. Climatic data of the region on 10 December 2018. City of Santo Antônio do Amparo, MG, Brazil.

Temp (°C)	RH (%)	Pressure (kPa)	WS (m.s^{-1})	Max Temp (°C)	Min Temp (°C)	RFE (mm)
19.73	73.31	91.6	1.51	25.6	13.28	0

Temp is temperature; RH is relative humidity; WS is wind speed; Max Temp is maximum temperature; Min Temp is minimum temperature; RFE is rainfall.

2.2. RPA Image Acquisition and Preprocessing

The images were captured using a commercial RPA 3DR Solo (3D Robotics, Berkeley, CA, USA), with four engines (quadcopter), powered by the automatic pilot system 3DR Pixhawk 2 and flight controller APM:Copter. A multispectral camera Parrot Sequoia with a focal distance fixed at 4.0 mm was coupled to the RPA. This camera has an RGB sensor with 16 megapixels resolution (4608 × 3456) and four extra sensors with 1.5 megapixels resolution (1280 × 960) in the spectral bands of green (550 nm BP 40), red (660 nm BP40), red-edge (735 nm BP 10), and near-infrared (790 nm BP 40). This camera is designed to map and monitor vegetation; it includes an upward-facing sunshine sensor, which allows the radiometric calibration of those 4 multispectral bands during image collection. The software Mission Planner from a complete ground station, open-source for RPA automatic piloting systems, was used to plan the flight mission. The flight altitude was fixed at 60 m above ground, and the flight speed was an average of 3 m s^{-1}. The images were captured every 1 s with a spatial resolution of 2.07 cm and overlapping the frontal position of 80% and lateral position of 75%.

The image processing was made using the software Agisoft PhotoScan® Professional, version 1.2.4 (Agisoft LLC, St. Petersburg, Russia). This software works in a three-step workflow. The first step was the alignment of the images by identifying corresponding resources. For executing the alignment of the image, the software calculates the parameters that orient the internal and external camera, including the radial nonlinear distortion. For this study, this task was executed with a high precision set. The results of this step are positions of the camera matching each image, representing the parameters for internal calibration, and the 3D coordinates of the sparse cloud of points in the terrain. In the second step, the sparse cloud is georeferenced in a local coordinates system (WGS 84—UTM Zone 23S), and the densified cloud of points is obtained using the heightfield method, which is based on paired depth map computation, resulting in a detailed 3D model. The third step applies a texture to the mesh obtained previously, generating the orthomosaic. This way, orthomosaics were created for each spectral band (green, red, red-edge, and near-infrared).

To minimize the effects of other targets in the spectral response of the coffee plants, such as soil and weeds, the mosaics were differentiated in the software eCognition version 9.0.1 (eCognition Developer, Munich, Germany). Two distinct classes were visually defined: coffee plants and noncoffee plants. From that, the classes were manually sorted for the orthomosaics of each spectral band.

2.3. Determining Leaf Nitrogen

Leaf nitrogen analyzes were performed on 10 December 2018. After obtaining the orthomosaics, the nitrogen level in the leaves of the coffee plants was visually evaluated, defining three distinct regions as deficient, critical, and sufficient N content (Figure 3). For that, we used the orthomosaic of the RGB composition (red, green, and blue) that allows a better perception than the human eye of vegetation. Thus, 10 targeted sampling points were chosen, with 2 sampling points for regions with N content in the plants considered sufficient, 4 points for regions of critical level, and 4 points for deficient regions (Figure 2).

Figure 3. Description of the classes of nitrogen level in the coffee leaves identified through RGB (red, green, and blue) composition in the orthomosaic images, to select sampling points in each level.

Each sample point was composed of 5 plants, 1 central plant and 2 plants oriented north-south from the central plant. From every plant, 3 leaves were collected from both sides of the planting line in three different canopy heights, totalizing 30 leaves collected for each sample point. After the collection, the leaves were sent to the laboratory of leaf analysis in the University of Lavras to quantify the nitrogen content using the Kjeldahl method. This method has been widely applied to determine nitrogen content, especially in the analysis of plant tissues. The Kjeldahl procedure involves three steps—destruction, distillation, and titration [8]. The leaf diagnosis for mature coffee plants proposed by [34] was used to classify the nitrogen content in the leaves as N-deficient (<2.5%), N-critical (2.5 to 3.0%), and N-sufficient (3.0 to 3.5%).

2.4. Vegetation Indices

The vegetation indices (Table 2) were chosen based on their capability to determine the nitrogen content in the coffee crops from remote sensing data [5,7,8].

2.5. Random Forest (RF) Classification

The Random Forest (RF) algorithm was used for classifying the N content in the leaves of the coffee plants from the vegetation indices proposed in this study (Table 2). The modeling was created in R [45], using the R Random Forest package [46]. For the training of the algorithm, the values of the vegetation indexes in the pixels referring to each of the 5 plants of each sampling point were used with the results of the chemical analysis of the N content in the leaves of the coffee plants of these sampling points. RF is one of the most successful classifiers based on learning strategies. The algorithm consists of a collection of three-based classifiers {h (x, Θk), k = 1, . . . }, where x is the input vector and {Θk} are the random vectors distributed identically and independently [47]. Each decision tree was constructed using a bootstrap deterministic algorithm, allowing the remaining data points to validate and issue a unitary vote to the most popular class. The RF uses the procedure of initialization with replacement to increase the diversity of the classifying

trees, which allocated pixels to one class, according to the maximum number of votes in the collection [48].

Table 2. Vegetation indices of multispectral images obtained using Remotely Piloted Aircraft (RPA).

Vegetation Indices	Equation	Reference
GNDVI (Green Normalized Difference Vegetation Index)	$\dfrac{\rho_{nir} - \rho_{green}}{\rho_{nir} + \rho_{green}}$	[35]
GOSAVI (Green Optimal Soil Adjusted Vegetation Index)	$(1 + 0.16)\dfrac{\rho_{nir} - \rho_{green}}{\rho_{nir} + \rho_{green} + 0.16}$	[36]
NDVI (Normalized Difference Vegetation Index)	$\dfrac{\rho_{nir} - \rho_{red}}{\rho_{nir} + \rho_{red}}$	[37]
SAVI (Soil Adjusted Difference Vegetation Index)	$\dfrac{(1+L)\,\rho_{nir} - \rho_{red}}{\rho_{nir} + \rho_{red} + L}$	[38]
MTCI (MERIS Terrestrial Chlorophyll Index)	$\dfrac{\rho_{nir} - \rho_{edge}}{\rho_{edge} - \rho_{red}}$	[39]
NDRE (Normalized Difference Red Edge)	$\dfrac{\rho_{nir} - \rho_{edge}}{\rho_{nir} + \rho_{edge}}$	[40]
EXR (Excessive Red)	$1.44 * \rho_{red} - \rho_{green}$	[41]
MPRI (Modified Photochemical Reflectance Index)	$\dfrac{\rho_{green} - \rho_{red}}{\rho_{green} + \rho_{red}}$	[42]
GRRI (Green–Red Ratio Index)	$\dfrac{\rho_{green}}{\rho_{red}}$	[43]
NDI (Normalized Different Index)	$\dfrac{\rho_{green} - \rho_{red}}{\rho_{green} + \rho_{red} + 0.01}$	[44]

ρ_{blue}: reflectance in the blue band; ρ_{green}: reflectance in the green band; ρ_{red}: reflectance in the red band; ρ_{nir}: reflectance in the near-infrared band; ρ_{edge}: reflectance in the red-edge band.

To execute the RF for classifying images, two parameters must be defined: the number of trees (ntree), and the number of predictors concerning the maximization of the model (mtry) [49,50]. In this study, the ntree was set to 1000 trees, while the mtry was set to equal to the square root of the total number of input resources.

2.6. Accuracy Assessment

To validate the results of the RF classification, a crossed validation approach was used 10 times [51]. This involved the random division of the reference objects in 10 joints of datasets, each one including around 15% of data from each class. In the steps of the assessment, the RF was tuned with 85% of the reference data and then applied to the other 15% left (that were the validation joint of data). This step was repeated ten times. In the end, the results were aggregated to one confusion matrix. The classification performance was assessed based on common statistical measures [52], derived from the confusion matrix. The select statistical measures included the overall accuracy (Equation (1)), kappa coefficient (Equation (2)), receiver operating characteristic (ROC curve), and area under the curve (UAC). The ROC curve was obtained by plotting a graph of sensitivity (true positive rate) versus specificity (false positive rate), and UAC was estimated using the method proposed by [53].

$$\text{Overall Accuracy} = \frac{\sum_{i=1}^{q} n_{ii}}{n} \times 100\% \tag{1}$$

$$\text{Kappa Coefficient} = \frac{n\sum_{i=1}^{q} n_{ii} - \sum_{i=1}^{q} n_{i+} n_{+i}}{n^2 - \sum_{i=1}^{q} n_{i+} n_{+i}} \times 100\% \tag{2}$$

where q is the number of classes, n represents the total number of considered pixels, n_{ii} are the diagonal elements of the confusion matrix, n_{i+} represents the marginal sum of the rows in the confusion matrix, and n_{+i} is the marginal sum of the columns in the confusion matrix.

3. Results and Discussion

3.1. Nitrogen Content in Coffee Leaves

First, as described in the Material and Methods section, one flight was performed in the area; then, from the RGB images, the orthomosaic was generated, which allowed the identification and establishment of the leaves sampling points. The results of the statistical descriptive analyses of the chemical analysis of nitrogen content in coffee leaves are shown

in Table 3. The areas in the crops with N levels considered sufficient, critical, and deficient are denoted with values within the nutritional scale established by [34]. However, as can be seen in Figure 2, regions with a critical level of N presented plants with a considerable amount of yellowish leaves, which is a symptom of nutritional deficiency. This may have occurred because the plants in this region had leaves with a medium nitrogen content (2.69%), with the content value close to the level considered nutritionally deficient (<2.5%).

Table 3. Results of the descriptive statistics of the chemical analysis of N content in plants located in the sampling points, considered as N-sufficient, -critical and -deficient.
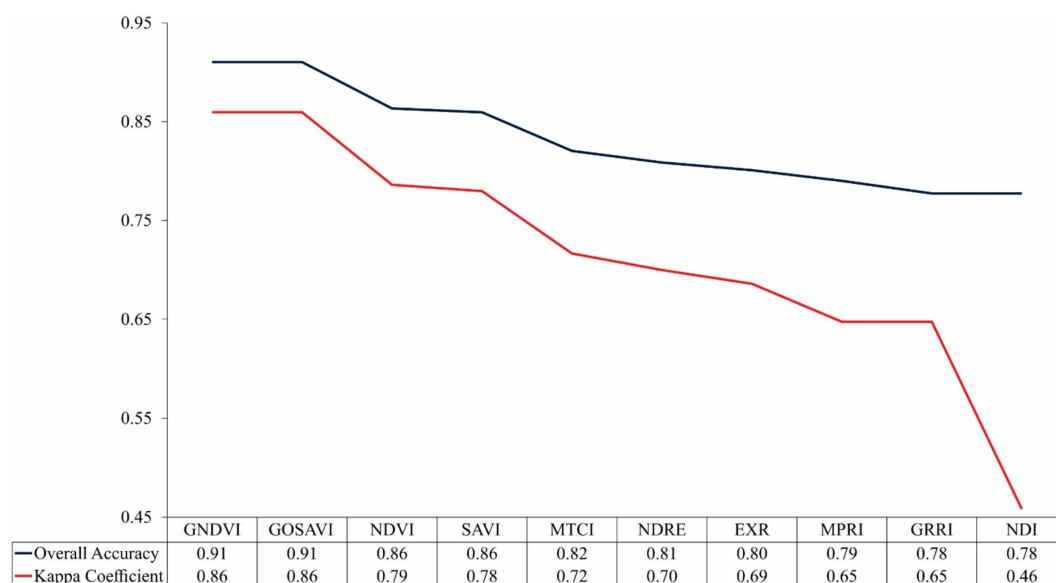
Nitrogen Levels	%			
	Min	Max	Mean	SD
Sufficient	3.00	3.11	3.05	0.08
Critical	2.51	2.85	2.69	0.18
Deficient	2.13	2.44	2.31	0.13

Min is minimum; Max is maximum; SD is the standard deviation.

The high-resolution images captured using RPA showed great potential in differentiating samples to evaluate the nitrogen content in the coffee leaves. According to [13], changes in the nutritional status of the plants have a direct impact on the canopy color of the coffee plantation. The color of leaves deficient in nitrogen is lighter, causing the color of the canopy to be green–yellowish, whereas leaves with sufficient nitrogen content reflect energy more intensely in the green wavelength than in the red wavelength. In contrast, leaves with N deficiency reflect energy intensely both in the green and in the red wavelengths, resulting in leaves with a yellow coloration [8,13].

3.2. Overall Accuracy and Kappa Performance

The results of the overall accuracy and the kappa coefficient of the images classified by RF from vegetation indices are shown in Figure 4. In general, the classification has presented a high performance in assessing the nitrogen content of the leaves in coffee plants: values of accuracy and kappa coefficient were from 0.78 to 0.91 and 0.46 to 0.86, respectively.

	GNDVI	GOSAVI	NDVI	SAVI	MTCI	NDRE	EXR	MPRI	GRRI	NDI
Overall Accuracy	0.91	0.91	0.86	0.86	0.82	0.81	0.80	0.79	0.78	0.78
Kappa Coefficient	0.86	0.86	0.79	0.78	0.72	0.70	0.69	0.65	0.65	0.46

Figure 4. Overall accuracy and kappa coefficient for image classification through Random Forest (RF) from vegetation indices.

It was possible to notice that vegetation indices that use a combination of the near-infrared band with bands in the visible spectral region presented better results. Between

those indices, the Green Normalized Difference Vegetation (GNDVI) and the Green Optimized Soil Adjusted Vegetation Index (GOSAVI) obtained the greatest values. These results are associated with the larger predominance of plants with sufficient N content in the crops and, consequently, higher chlorophyll content in the leaves. According to [54,55], the reflectance in the green band is greatly affected by the variations in the chlorophyll content, when compared to the red band reflectance. Thus, the spectral indices based on the near-infrared and green bands have proved more sensitive to changes in chlorophyll content than the indices with near-infrared and red bands [56,57]. The GNDVI and GOSAVI have also been reported by other authors as more suitable indices to evaluate N status in vegetation. Similarly to this study, [29] used an approach based on RF and vegetation indices from RPA images to describe the variation of plant N uptake (PNU), and N nutrition index (NNI) for rice crops in the northwest lowlands of China. The results, resembling this study, presented GOSAVI and GNDVI indices as having better performances. In [22], they applied the RF model to vegetation indices obtained from three sensors (RGB, color infrared, and multispectral), coupled to an RPA, for determining the N status in the rice crops, finding the GNDVI to have the better results. Therefore, the results of this study align with the literature and recent research.

For the Medium Resolution Imaging Spectrometer (MERIS) Terrestrial Chlorophyll Index (MTCI) and Normalized Difference Red Edge (NDRE) indices, the bandwidth of the red-edge band that is 4 times smaller than the green and red bands may have contributed to the worst performances among the indices that use the combination of the near-infrared band with visible bands. As demonstrated by [5], the width of the red-edge band can make indices less sensitive to the vegetation characteristics. However, better results were expected from them, since the relation between near-infrared and red-edge spectral bands with the N content in the plants is well established in the literature [5,6,22].

For the indices Excessive Red (EXR), Modified Photochemical Reflectance Index (MPRI), Green–Red Ratio Index (GRRI), and Normalized Different Index (NDI) that use only visible spectral bands, the results had inferior performances. The possible explanation for that is the lack of radiometric correction for RGB images, letting fewer alterations in the lighting conditions influence the precision of the color reproduction [58,59]. However, even with those indices showing worse performances, when compared to indices that combine near-infrared bands to visible bands, their results still express the potential for the application of RGB cameras. These types of cameras can be a viable alternative to evaluate the N status in plants, especially for medium and smaller farmers. RGB cameras are easy to operate and accessible for most researchers. Furthermore, the RGB images can be used right after downloading from the memory card, without preprocessing, for example, for band registry, and to process images, various sophisticated commercial platforms are now available, such as Agisoft Photoscan and Pix4D.

Regarding the RGB indices, the results obtained by this study diverge from the reported by [31]. The Random Forest models based on RGB indices from RPA that were developed by the authors did not succeed in evaluating and predicting the N content of the coffee plants. This may have occurred because the Parot Sequoia camera that was used in the present study presents superior techniques, compared to the RGB camera used by [31], for example, when calibrating the radiometric individual bands. Additionally, the focus of the sampling points in this study may have been crucial for more effective training samples and clearer results.

3.3. ROC Curve and AUC

Regarding the analysis of the image classification using the ROC curve and AUC, in general, all the vegetation indices presented excellent performances for assessing the nitrogen content in coffee plants (Figure 5).
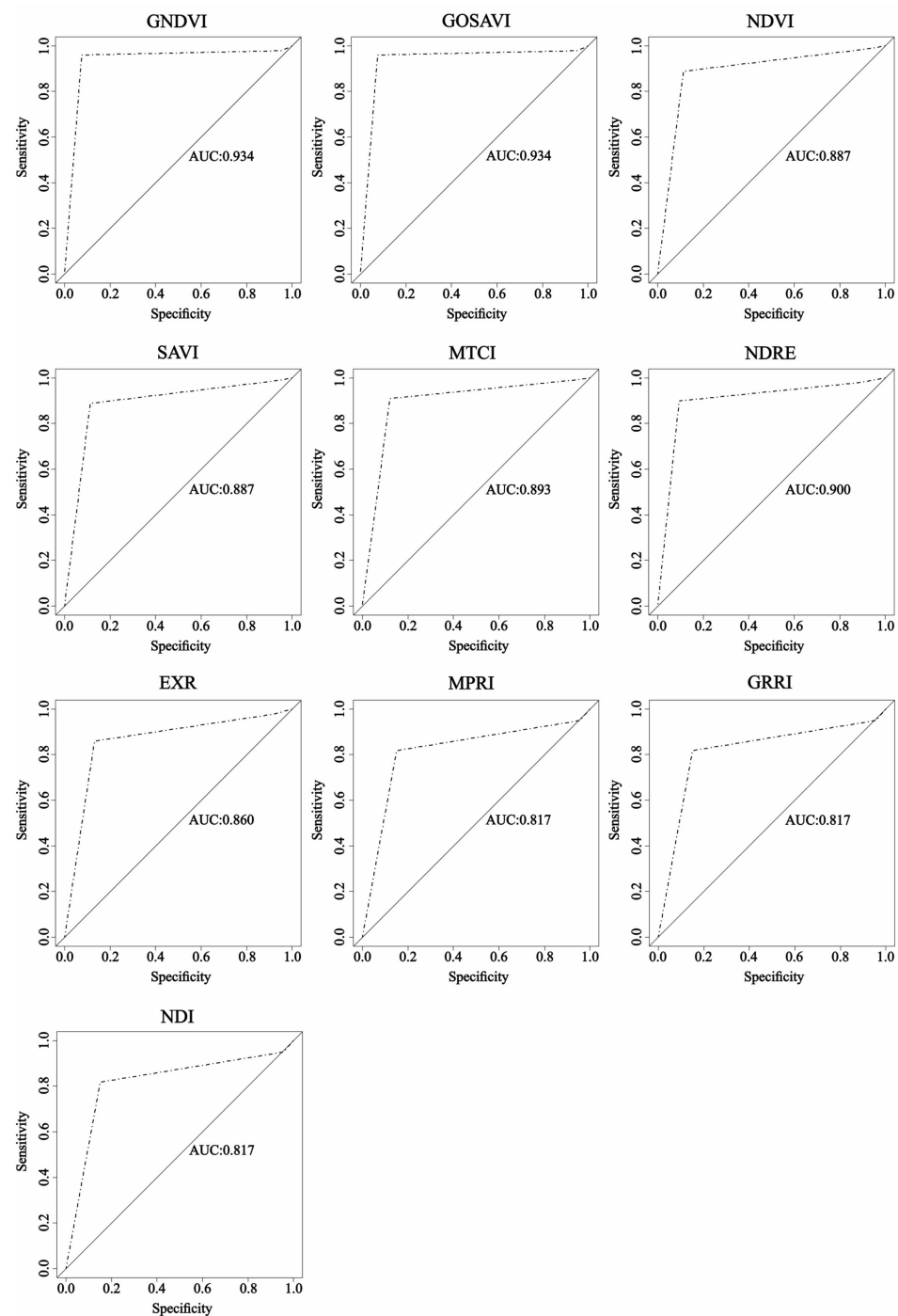
Figure 5. Receiver operating characteristic (ROC) curve and area under the curve (AUC) for image classification through RF from vegetation indices.

It was possible to observe all vegetation indices approximate to the point (0, 100%) in the ROC curve. According to [60], to rate a model as optimum, its analysis has to be as close as possible to the point (0, 100%). Moreover, [61] describes the ROC curve as a description of the relative compensation between benefits (true positives) and costs (false positives) of a classification. Thus, a point located in the upper left corner means a greater number of positive and negative examples were classified correctly, consequently creating a lower cost to the classification.

For the AUC, the values were from 0.817 to 0.934, where greater values came from the indices that combined near-infrared band to visible bands, matching what happened with overall accuracy and kappa coefficient. According to [62], the closer the AUC gets to

1, the better the overall test performance, meaning the AUC= 1 test is the perfectly accurate result. For the approach in this study, the indices GNDVI and GOSAVI had the greatest performances, while the indices MPRI, GRRI, and NDI presented the worse performances.

3.4. Mapping and Quantifying N Spatial Distribution in Coffee Leaves

Predictive maps for nitrogen spatial distribution in coffee leaves and quantification in the percentage of those maps, obtained from the classification of images using RF from vegetation indices data, are shown in Figure 6, Figure 7 and Table 4, respectively.
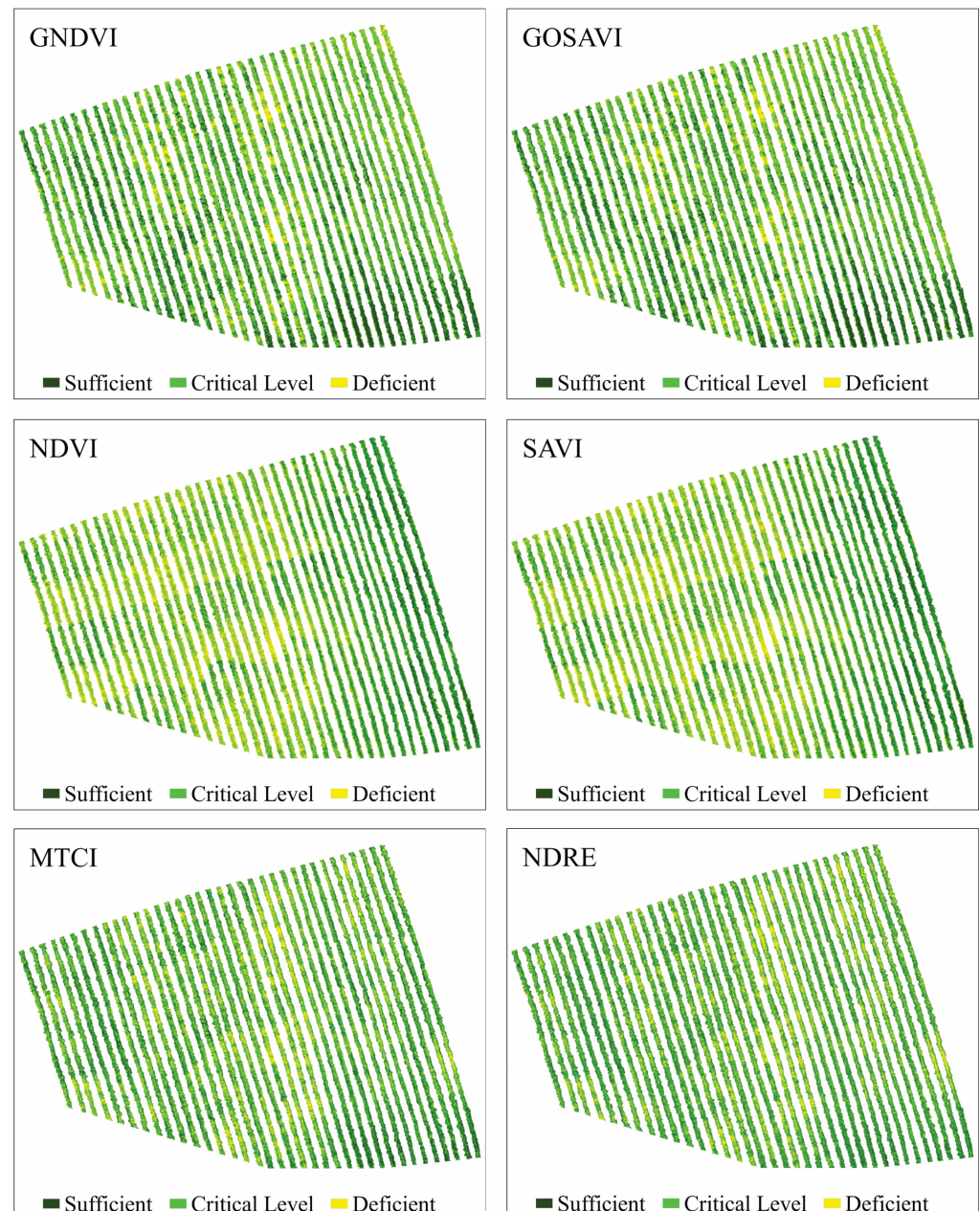
Figure 6. Predictive maps of the N content spatial distribution in coffee leaves, obtained from the classification of images using RF from vegetation indices Green Normalized Difference Vegetation (GNDVI), Green Optimized Soil Adjusted Vegetation Index (GOSAVI), Normalized Difference Vegetation Index (NDVI), Soil Adjusted Difference Vegetation Index (SAVI), Medium Resolution Imaging Spectrometer (MERIS) Terrestrial Chlorophyll Index (MTCI), and Normalized Difference Red Edge (NDRE).
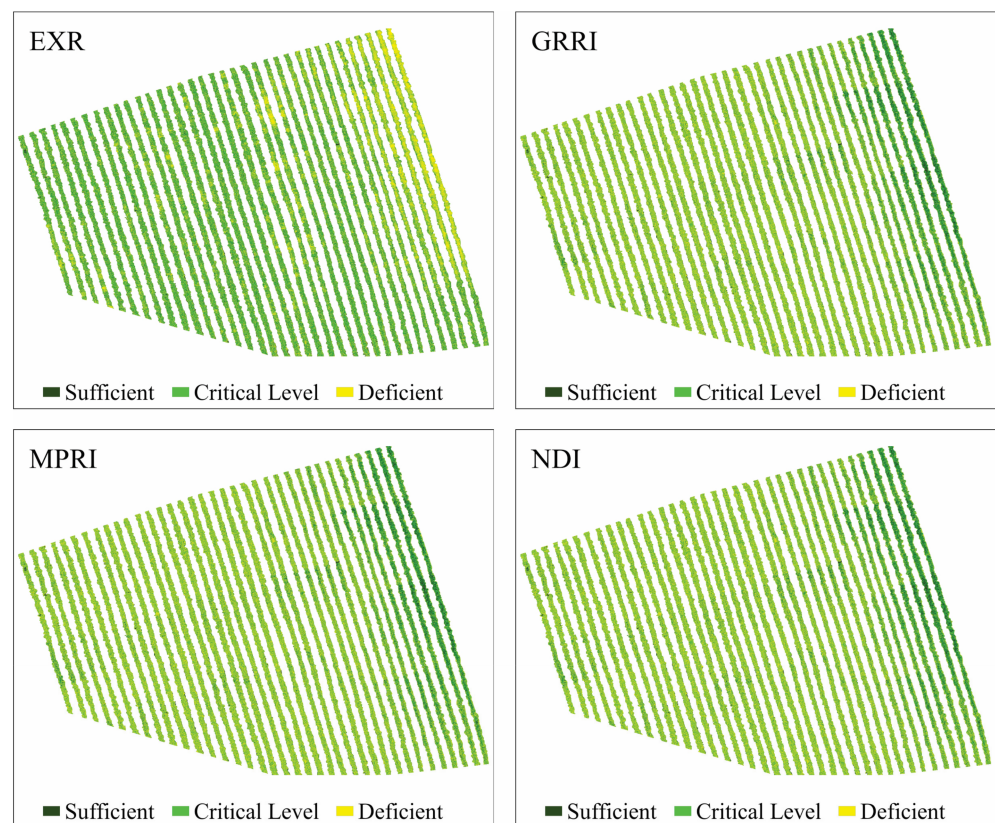
Figure 7. Predictive maps of the N content spatial distribution in coffee leaves, obtained from the classification of images using RF from vegetation indices Excessive Red (EXR), Green–Red Ratio Index (GRRI), Modified Photochemical Reflectance Index (MPRI), and Normalized Different Index (NDI).

Table 4. Percentage (%) of areas in the coffee plantation that had N content sufficient, critical, and deficient, according to the appraisal maps of the N content spatial distribution in coffee leaves, presented from the best performance to the worst performance.

Vegetation Indices	Area (%)			
	Sufficient	**Critical**	**Deficient**	**Total**
GNDVI	26	52	22	100
GOSAVI	26	52	22	100
NDVI	23	46	31	100
SAVI	23	46	31	100
MTCI	31	47	22	100
NDRE	31	46	22	100
EXR	25	45	31	100
MPRI	21	42	37	100
GRRI	21	41	38	100
NDI	21	41	38	100

GNDVI and GOSAVI indices had the best performances (Figure 6), presenting the best class definition. These maps had the largest area at the N-critical level (52%) and the smallest area for deficient N levels (22%). According to these maps, the plants in the central region of the crops showed a greater N deficiency. Similar to these indices, NDVI and SAVI maps indicated greater N deficiency in the central region of the crops, although they also indicated a large number of N-deficient plants in the western region, evinced by the 9% increase in the N-deficient areas, when compared to the GNDVI and GOSAVI maps. Regarding the MTCI and NDRE indices (Figure 6), they both presented the same percentage of N-deficient areas as the GNDVI and the GOSAVI, but they overestimated areas with

sufficient N content (31%). Additionally, in these maps, the area with N deficiency extended from the central region to the eastern region of the plantation. The indices that had worse performances were MPRI, GRRI, and NDI (Figure 7). They overestimated the percentage of N-deficient areas and underestimated areas with sufficient N levels (21%) as well as areas with critical N levels (41%). The maps for these indices showed mixed classes, especially between critical and deficient levels of nitrogen, showing deficiency in practically all regions of the crops.

These results confirmed the possibility of managing nitrogen locally using images obtained from RPA. Knowing the spatial distribution of areas with N deficiency becomes essential for managing nitrogen properly in the plantation. Overestimating or underestimating areas with N deficiency can increase production costs, reduce productivity, and cause environmental damage. In this context, the methodology proposed by this study, especially for GNDVI and GOSAVI indices, can promote a 78% reduction in nitrogen fertilizer treatments in the studied area, when compared to the conventional management and application methods. For this, applying fertilizers would be recommended only for plants showing N deficiency, which means 22% of the total area, unlike the conventional application of fertilizer, where producers apply it to the entire area of the plantation.

Based on information previously published by [34,63], it is possible to illustrate how the methodology proposed in this study could prove economically relevant to monitor the N status in coffee plantations. The data used in the simulation were as follows: study area: 1.5 ha; ammonium nitrate price for 50 Kg: $15.00; average amount of N applied by coffee producers: 300 kg ha^{-1} per year; N concentration in ammonium nitrate: 33%; total ammonium nitrate applied by coffee producers: 833 kg ha^{-1} per year; and reduction of the demand for N observed in this study, considering results from the GNDVI and GOSAVI indices: 78%.

To that end, the calculation of the total price per hectare of N was 833 kg ha^{-1} per year $\times$ $15.00 (50 Kg) = $249.90 kg ha^{-1}. Considering the 1.5 ha plantation area, it would have a cost of $374.85 per year for nitrogen fertilization. When considering the actual demand for N for this coffee field, which, in this study, would reduce by 78% of the total, the coffee producer cost could be reduced to $374.85 $\times$ 0.22 = $82.47 per year. In conclusion, the coffee producer could save $292.38 during the year with nitrogen fertilization. This was possible, considering the use of N only for plants that showed deficiency. This example is hypothetical. However, it shows the application and importance of carrying out properly the monitoring of N status in coffee plantations, which is especially relevant to minimize the use of N and reduce the costs of nitrogen fertilization, as well as its impacts.

In this study, the learning method of the RF algorithm applied to the vegetation indices from images obtained using a multispectral camera coupled to an RPA was used to evaluate the N content in a coffee plantation. The results proved that this methodology could be used to diagnose the nitrogen status in coffee plants and also orient producers on how to do the application of nitrogen fertilizers at a variable rate. However, other local conditions, such as density, canopy cover, age, species of plant, and spacing can reduce the accuracy of the model suggested. From this, to deal with the problem complexity, a systemic approach that simulates and predicts the impact of applying N fertilizers over space and time, affecting the cultivation yield, is necessary. To validate this relation, more studies using different conditions and plants are needed. In addition, the use of RPAs with greater flight autonomy and different multispectral cameras with better resolution can improve the model efficiency, as well as the size of the monitored area. To administer N fertilizers at a variable rate, the process used in the practical application needs to develop into management zones, pixel-based, and plant-based, to make easy the reduction and viable the application for producers.

4. Conclusions

The results of this study showed that the machine learning method Random Forest, applied to vegetation indices from multispectral images obtained by RPA, offers a very

promising approach to map and quantify nitrogen status in a coffee plantation and also indicates the possibility to apply fertilizers in a localized manner. The proposed model allowed the assessment of the N spatial distribution in the coffee leaves and the quantification of area presenting N deficiency. The model also indicated that the most efficient vegetation indices to evaluate the nitrogen status in coffee plants were GNDVI and GOSAVI. Additionally, the application of the methodology proposed in this study can contribute to a more rational management of N in the crops.

Author Contributions: Conceptualization, D.B.M. and G.A.eS.F.; Methodology, D.B.M., G.A.eS.F., P.H.S.G., L.S.S. and B.D.S.B.; Validation, D.B.M., G.A.eS.F., P.H.S.G. and F.S.; Formal Analysis and Investigation, D.B.M., G.A.eS.F., P.H.S.G., F.S., L.C., G.R., R.d.O.F., R.A.P.B. and J.E.L.D.; Resources, D.B.M., G.A.eS.F., L.C. and G.R.; Writing and Writing–Review, Editing and Visualization, all authors. Project Administration and Funding Acquisition, G.A.eS.F., L.C. and G.R. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare that they have no conflict of interest.

References

1. Chemura, A. Modelling Spatial Variability of Coffee (*Coffea Arabica* L.) Crop Condition with Multispectral Remote Sensing Data. Ph.D. Thesis, University of KwaZulu-Natal, Pietermaritzburg, South Africa, 2017.
2. Bote, A.D.; Ayalew, B.; Ocho, F.L.; Anten, N.P.R.; Vos, J. Analysis of coffee (*Coffea arabica* L.) performance in relation to radiation levels and rates of nitrogen supply I. Vegetative growth, production and distribution of biomass and radiation use efficiency. *Eur. J. Agron.* **2018**, *92*, 115–122. [CrossRef]
3. Nazareno, R.B.; Da Silva Oliveira, C.A.; Sanzonowicz, C.; Ramos Sampaio, J.B.; Pereira Da Silva, J.C.; Guerra, A.F. Crescimento inicial do cafeeiro Rubi em resposta a doses de nitrogênio, fôsforo e potássio e a regimes hídricos. *Pesqui. Agropecu. Bras.* **2003**, *38*, 903–910. [CrossRef]
4. Coste, R. *Coffee—The Plant and The Product*; MacMillan Press: London, UK, 1992.
5. Chemura, A.; Mutanga, O.; Odindi, J.; Kutywayo, D. Mapping spatial variability of foliar nitrogen in coffee (*Coffea arabica* L.) plantations with multispectral Sentinel-2 MSI data. *ISPRS J. Photogramm. Remote Sens.* **2018**, *138*, 1–11. [CrossRef]
6. Putra, B.T.W.; Soni, P.; Morimoto, E.; Pujiyanto, P. Estimating biophysical properties of coffee (*Coffea canephora*) plants with above-canopy field measurements, using CropSpec. *Int. Agrophys.* **2018**, *32*, 183–191. [CrossRef]
7. Putra, W.B.T.; Soni, P. Enhanced broadband greenness in assessing Chlorophyll a and b, Carotenoid, and Nitrogen in *Robusta coffee* plantations using a digital camera. *Precis. Agric.* **2018**, *19*, 238–256. [CrossRef]
8. Putra, B.T.W.; Soni, P. Improving nitrogen assessment with an RGB camera across uncertain natural light from above-canopy measurements. *Precis. Agric.* **2020**, *21*, 147–159. [CrossRef]
9. Lima, L.M.; Pozza, E.A.; Torres, H.N.; Pozza, A.A.A.; Salgado, M.; Pfenning, L.H. Relação nitrogênio/potássio com mancha de Phoma e nutrição de mudas de cafeeiro em solução nutritiva. *Trop. Plant Pathol.* **2010**, *35*, 223–228. [CrossRef]
10. Pérez, C.D.P.; Pozza, E.A.; Pozza, A.A.A.; Freitas, A.S.; Da Silva, M.G. Nitrogênio e potássio na intensidade da mancha aureolada do cafeeiro em solução nutritiva. *Coffee Sci.* **2017**, *12*, 60–68. [CrossRef]
11. Martinez, H.E.P.; de Souza, B.P.; Caixeta, E.T.; de Carvalho, F.P.; Clemente, J.M. Water deficit changes nitrate uptake and expression of some nitrogen related genes in coffee-plants (*Coffea arabica* L.). *Sci. Hortic.* **2020**, *267*, 109254. [CrossRef]
12. DaMatta, F.M.; Loos, R.A.; Silva, E.A.; Loureiro, M.E. Limitations to photosynthesis in *Coffea canephora* as a result of nitrogen and water availability. *J. Plant Physiol.* **2002**, *159*, 975–981. [CrossRef]
13. Feng, D.; Xu, W.; He, Z.; Zhao, W.; Yang, M. Advances in plant nutrition diagnosis based on remote sensing and computer application. *Neural Comput. Appl.* **2019**, 1–10. [CrossRef]
14. Ye, X.; Abe, S.; Zhang, S. Estimation and mapping of nitrogen content in apple trees at leaf and canopy levels using hyperspectral imaging. *Precis. Agric.* **2020**, *21*, 198–225. [CrossRef]
15. Mahajan, G.R.; Pandey, R.N.; Sahoo, R.N.; Gupta, V.K.; Datta, S.C.; Kumar, D. Monitoring nitrogen, phosphorus and sulphur in hybrid rice (*Oryza sativa* L.) using hyperspectral remote sensing. *Precis. Agric.* **2017**, *18*, 736–761. [CrossRef]

16. Fitzgerald, G.; Rodriguez, D.; O'Leary, G. Measuring and predicting canopy nitrogen nutrition in wheat using a spectral index-The canopy chlorophyll content index (CCCI). *Field Crop. Res.* **2010**, *116*, 318–324. [CrossRef]

17. Santos, L.M.; Ferraz, G.A.E.S.; Barbosa, B.D.S.; Andrade, A.D. Use of remotely piloted aircraft in precision agriculture: A review. *DYNA* **2019**, *86*, 284–291. [CrossRef]

18. Yang, G.; Liu, J.; Zhao, C.; Li, Z.; Huang, Y.; Yu, H.; Xu, B.; Yang, X.; Zhu, D.; Zhang, X.; et al. Unmanned aerial vehicle remote sensing for field-based crop phenotyping: Current status and perspectives. *Front. Plant Sci.* **2017**, *8*. [CrossRef]

19. Buchaillot, M.L.; Gracia-Romero, A.; Vergara-Diaz, O.; Zaman-Allah, M.A.; Tarekegne, A.; Cairns, J.E.; Prasanna, B.M.; Araus, J.L.; Kefauver, S.C. Evaluating maize genotype performance under low nitrogen conditions using RGB UAV phenotyping techniques. *Sensors* **2019**, *19*, 1815. [CrossRef]

20. Wang, H.; Mortensen, A.K.; Mao, P.; Boelt, B.; Gislum, R. Estimating the nitrogen nutrition index in grass seed crops using a UAV-mounted multispectral camera. *Int. J. Remote Sens.* **2019**, *40*, 2467–2482. [CrossRef]

21. Hunt, E.R.; Horneck, D.A.; Spinelli, C.B.; Turner, R.W.; Bruce, A.E.; Gadler, D.J.; Brungardt, J.J.; Hamm, P.B. Monitoring nitrogen status of potatoes using small unmanned aerial vehicles. *Precis. Agric.* **2018**, *19*, 314–333. [CrossRef]

22. Zheng, H.; Cheng, T.; Li, D.; Zhou, X.; Yao, X.; Tian, Y.; Cao, W.; Zhu, Y. Evaluation of RGB, color-infrared and multispectral images acquired from unmanned aerial systems for the estimation of nitrogen accumulation in rice. *Remote Sens.* **2018**, *10*, 824. [CrossRef]

23. Osco, L.P.; Paula, A.; Ramos, M.; Pereira, D.R.; Akemi, É.; Moriya, S.; Imai, N.N.; Matsubara, E.T. Predicting Canopy Nitrogen Content in Citrus-Trees Using Random Forest Algorithm Associated to Spectral Vegetation Indices from UAV-Imagery. *Remote Sens.* **2019**, *11*, 2925. [CrossRef]

24. Chlingaryan, A.; Sukkarieh, S.; Whelan, B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Comput. Electron. Agric.* **2018**, *151*, 61–69. [CrossRef]

25. Liakos, K.G.; Busato, P.; Moshou, D.; Pearson, S.; Bochtis, D. Machine learning in agriculture: A review. *Sensors* **2018**, *18*, 2674. [CrossRef]

26. Ali, I.; Greifeneder, F.; Stamenkovic, J.; Neumann, M.; Notarnicola, C. Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data. *Remote Sens.* **2015**, *7*, 16398–16421. [CrossRef]

27. Qiu, J.; Wu, Q.; Ding, G.; Xu, Y.; Feng, S. A survey of machine learning for big data processing. *EURASIP J. Adv. Signal Process.* **2016**, *67*, 1–16. [CrossRef]

28. Osco, L.P.; Junior, J.M.; Ramos, A.P.M.; Furuya, D.E.G.; Santana, D.C.; Teodoro, L.P.R.; Gonçalves, W.N.; Baio, F.H.R.; Pistori, H.; Junior, C.A.D.S.; et al. Leaf nitrogen concentration and plant height prediction for maize using UAV-based multispectral imagery and machine learning techniques. *Remote Sens.* **2020**, *12*, 3237. [CrossRef]

29. Zha, H.; Miao, Y.; Wang, T.; Li, Y.; Zhang, J.; Sun, W.; Feng, Z.; Kusnierek, K. Improving Unmanned Aerial Vehicle Remote Sensing-Based Rice Nitrogen Nutrition Index Prediction with Machine Learning. *Remote Sens.* **2020**, *12*, 215. [CrossRef]

30. Zheng, H.; Li, W.; Jiang, J.; Liu, Y.; Cheng, T.; Tian, Y.; Zhu, Y.; Cao, W.; Zhang, Y.; Yao, X. A comparative assessment of different modeling algorithms for estimating leaf nitrogen content in winter wheat using multispectral images from an unmanned aerial vehicle. *Remote Sens.* **2018**, *10*, 2026. [CrossRef]

31. Parreiras, T.C.; Lense, G.H.E.; Moreira, R.S.; Santana, D.B.; Mincato, R.L. Using unmanned aerial vehicle and machine learning algorithm to monitor leaf nitrogen in coffee. *Coffee Sci.* **2020**, *15*, 1–9. [CrossRef]

32. De Freitas, A.F.; Nadaleti, D.H.S.; Silveira, H.R.D.O.; Carvalho, G.R.; Venturin, R.P.; Silva, V.A. Productivity and beverage sensory quality of arabica coffee intercropped with timber species. *Pesqui. Agropecuária Bras.* **2020**, *55*. [CrossRef]

33. National Aeronautics and Space Administration—NASA. Power Data. Available online: https://power.larc.nasa.gov/data-access-viewer/ (accessed on 29 March 2021).

34. Martinez, H.E.P.; Neves, J.C.L.; Shuler, J. Mineral Nutrition and Fertilization. In *Coffee: Production, Quality and Chemistry*; Farah, A., Ed.; Royal Society of Chemistry: London, UK, 2019; pp. 163–201.

35. Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* **1996**, *58*, 289–298. [CrossRef]

36. Rondeaux, G.; Steven, M.; Baret, F. Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* **1996**, *55*, 95–107. [CrossRef]

37. Rouse, J.W.; Haas, R.H.; Deering, D.W.; Schell, J.A.; Harlan, J.C. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*; Type III Final Report; NASA/GSFC: Greenbelt, MD, USA, 1974.

38. Huete, A.R. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [CrossRef]

39. Dash, J.; Curran, P.J. The MERIS terrestrial chlorophyll index. *Int. J. Remote Sens.* **2004**, *25*, 5403–5413. [CrossRef]

40. Gitelson, A.; Merzlyak, M.N. Quantitative estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves. *J. Photochem. Photobiol. B Biol.* **1994**, *22*, 247–252. [CrossRef]

41. Meyer, G.E.; Hindman, T.W.; Lakshmi, K. Machine Vision Detection Parameters for Plant Species Identification. In Proceedings of the Precision Agriculture and Biological Quality, Bellingham, DC, USA, 14 January 1998; pp. 327–335. [CrossRef]

42. Yang, Z.; Willis, P.; Mueller, R. Impact of band-ratio enhanced AWIFS image to crop classification accuracy. In Proceedings of the 17th William Pecora Memorial Remote Sensing Symposium, Denver, CO, USA, 18–20 November 2008.

43. Gamon, J.A.; Surfus, J.S. Assessing leaf pigment content and activity with reflectometer. *New Phytol.* **1999**, *143*, 105–117. [CrossRef]

44. Mao, W.; Wang, Y.; Wang, Y. Real-time detection of between-row weeds using machine vision. In Proceedings of the ASAE Annual Meeting, Las Vegas, NV, USA, 27–30 July 2003; pp. 1–9. [CrossRef]
45. R Core Team. R: A Language and Environment for Statistical Computing, Vienna, Austria. 2013. Available online: http://www.R-project.org/ (accessed on 10 July 2020).
46. Liaw, A.; Wiener, M.; Breiman, L.; Cutler, A. Package "randomForest". 2018. Available online: https://cran.r-project.org/web/packages/randomForest/randomForest.pdf (accessed on 10 July 2020).
47. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
48. Son, N.T.; Chen, C.F.; Chen, C.R.; Minh, V.Q. Assessment of Sentinel-1A data for rice crop classification using random forests and support vector machines. *Geocarto Int.* **2018**, *33*, 587–601. [CrossRef]
49. Oliveira, S.; Oehler, F.; San-Miguel-Ayanz, J.; Camia, A.; Pereira, J.M.C. Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression and Random Forest. *For. Ecol. Manag.* **2012**, *275*, 117–129. [CrossRef]
50. Rodriguez-Galiano, V.F.; Chica-Olmo, M.; Abarca-Hernandez, F.; Atkinson, P.M.; Jeganathan, C. Random Forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. *Remote Sens. Environ.* **2012**, *121*, 93–107. [CrossRef]
51. Kohavi, R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. *Int. Jt. Conf. Artif. Intell.* **1995**, *2*, 1137–1143.
52. Foody, G.M. Status of land cover classification accuracy assessment. *Remote Sens. Environ.* **2002**, *80*, 185–201. [CrossRef]
53. Hand, D.J.; Till, R.J. A Simple Generalisation of the Area Under the ROC Curve for Multiple Class Classification Problems. *Mach. Learn.* **2001**, *45*, 171–186. [CrossRef]
54. Hunt, E.R.; Doraiswamy, P.C.; McMurtrey, J.E.; Daughtry, C.S.T.; Perry, E.M.; Akhmedov, B. A visible band index for remote sensing leaf chlorophyll content at the Canopy scale. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *21*, 103–112. [CrossRef]
55. Hunt, E.R.; Daughtry, C.S.T.; Eitel, J.U.H.; Long, D.S. Remote sensing leaf chlorophyll content using a visible band index. *Agron. J.* **2011**, *103*, 1090–1099. [CrossRef]
56. Gitelson, A.A.; Viña, A.; Ciganda, V.; Rundquist, D.C.; Arkebauer, T.J. Remote estimation of canopy chlorophyll content in crops. *Geophys. Res. Lett.* **2005**, *32*, 1–4. [CrossRef]
57. Nigon, T.J.; Mulla, D.J.; Rosen, C.J.; Cohen, Y.; Alchanatis, V.; Rud, R. Evaluation of the nitrogen sufficiency index for use with high resolution, broadband aerial imagery in a commercial potato field. *Precis. Agric.* **2014**, *15*, 202–226. [CrossRef]
58. Bendig, J.; Yu, K.; Aasen, H.; Bolten, A.; Bennertz, S.; Broscheit, J.; Gnyp, M.L.; Bareth, G. Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *39*, 79–87. [CrossRef]
59. Fernández, P.D.M.; Peña, F.A.G.; Ren, T.I.; Leandro, J.J.G. Fast and robust multiple ColorChecker detection using deep convolutional neural networks. *Image Vis. Comput.* **2018**, *81*, 15–24. [CrossRef]
60. Prati, R.C.; Batista, G.E.D.A.P.A.; Monard, M.C. Evaluating Classifiers Using ROC Curves. *IEEE Lat. Am. Trans.* **2008**, *6*, 215–222. [CrossRef]
61. Fawcett, T. An introduction to ROC analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [CrossRef]
62. Park, S.H.; Goo, J.M.; Jo, C.H. Receiver operating characteristic (ROC) curve: Practical review for radiologists. *Korean J. Radiol.* **2004**, *5*, 11–18. [CrossRef]
63. Fageria, N.K.; Baligar, V.C.; Li, Y.C. The role of nutrient efficient plants in improving crop yields in the twenty first century. *J. Plant Nutr.* **2008**, *31*, 1121–1157. [CrossRef]