HOKKAIDO UNIVERSITY

Title	Assessing oil palm growth condition in Indonesia using remote sensing techniques
Author(s)	Marimun, Heri Santoso
Citation	北海道大学. 博士(農学) 甲第13144号
Issue Date	2018-03-22
DOI	10.14943/doctoral.k13144
Doc URL	http://hdl.handle.net/2115/73369
Type	theses (doctoral)
File Information	Heri_Santoso_Marimun.pdf

Instructions for use

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

Assessing Oil Palm Growth Condition in Indonesia Using Remote Sensing Techniques

(リモートセンシング技術を用いたインドネシアにおけるアブラヤシの生育状態の評価)

Hokkaido University Graduate School of Agriculture
Division of Bio-Systems Sustainability Doctor Course

Heri Santoso Marimun

Table of Contents

List of Tables

List of Figures

Chapter 1. General Introduction

1.1. Oil palm plantations in Indonesia

In 2015, Indonesia had more than 11 million hectares of oil palm (*Elaeis guineensis*, Jacq.) plantations, owned by the government (6%), private companies (53%), and smallholders (41%). The oil palm plantations in Indonesia are mainly distributed on five islands: Sumatra (63.18%), Kalimantan (32.21%), Sulawesi (3.28%), Maluku and Papua (1.04%), and Java (0.3%) (Ditjenbun 2015). Oil palm plantations have contributed to regional development in Indonesia. Today, oil palm is an economic crop and the industry has become a key part of economic policy in Indonesia. In 2015, the average production of Indonesian palm oil (crude palm oil (CPO) and palm kernel oil (PKO)) was 3.3 ton/ha (Ditjenbun 2015), which is about half of the potential production of 6–9 ton/ha/year (Barcelos et al. 2015; Morcillo et al. 2013; Murphy 2009). Although there is clearly an oil yield gap, the current oil yield from oil palm is greater than that from any other oil crop worldwide (Barcelos et al. 2015). Oil palm is the most productive oil crop and is capable of meeting the global demand for vegetable oils (Corley and Tinker 2003; Corley 2009; Barcelos et al. 2015). The oil yield from oil palm is three- to eight- times greater than that from any other temperate or tropical oil crop. For example, in 2012, 56.2 million tons of oil was produced from 17.24 million ha of oil palm, while only 23.6 million tons of oil was produced from 36.4 million ha of soybean (Barcelos et al. 2015).

The factors affecting the site yield potential of oil palm include the planting material, light distribution, soil, and climate. The actual oil yield from oil palm is affected by agronomic practices, management practices, economic factors, and social and environmental factors (Goh, Chew, and Teo 1994, as cited in Corley and Tinker 2003). In Indonesia, 41% of oil palm plantations are owned by smallholders. In general, smallholders do not use the best planting materials, nor do they apply the best agronomic and management practices. The most common oil palm variety in smallholder plantations is Dura, which has an oil extraction ratio of less than 18%. A better variety is the hybrid Tenera, which has an oil extraction ratio of up to 26%. Smallholders also use poor management and agronomic practices because of a lack of knowledge and limited capital for operational costs. These are some of the reasons for the large gaps between the potential and actual oil yields from oil palm in Indonesia. Currently, the Indonesian government is developing and implementing programs to improve oil yields from smallholder oil palm plantations. These aims of these programs are to improve knowledge about good agronomic and management practices, to supply the Tenera variety as the planting material for new oil palm plantations and replanted areas, to provide financial support for replanting programs, and to improve the institutional strength of smallholder plantation associations (Haryana, Indarto, and Avianto 2010).

1.2. Main challenges in oil palm cultivation

Sustainable production is one of the main goals of oil plantation management in Indonesia. The Indonesian Ministry of Agriculture has issued a regulation (number 11/Permentan/OT.140/3/2015, Indonesian Sustainable Palm Oil; ISPO) that is mandatory for all plantations in Indonesia. There are several principles and criteria of the ISPO that must be followed by oil palm plantations, including pest and disease control, and the optimization of fertilizer inputs based on soil and leaf nutrient analyses.

The main disease of oil palms in Indonesia and Malaysia is basal stem rot disease (BSR), which is caused by *Ganoderma orbiforme* (= *G. boninense*). There is no effective treatment for BSR, and all currently applied treatments only prolong the life of the oil palm (Hushiarian, Yusof, and Dutse 2013; Susanto et al. 2013; Priwiratama, Prasetyo, and Susanto 2014; Priwiratama and Susanto 2014). Fertilizer is an important factor affecting oil palm production, and accounts for 40–50% of the total costs of field upkeep (Ng 2002). The fertilizer requirements of oil palm are determined based on analyses of leaf nutrient contents, like in other agricultural crops (Pritts and Heidenreich 2012; Memon, Memon, and Hassan 2005). Leaf samples are analyzed each year to optimize fertilizer inputs (Chapman and Gray 1949; Fairhurst and Mutert 1999).

This study had two main goals: to develop a fast and effective method to identify and monitor BSR disease, and to devise an inexpensive, rapid, and accurate method to monitor leaf nutrient status to optimize the fertilizer dosage. Remote

sensing is a potential technique to achieve these goals and is already used in several agricultural practices.

1.3. Scientific problem

The major disease of oil palm is BSR, for which there is no effective treatment. The early detection of BSR in oil palm plantations is one strategy to control this disease (Hushiarian, Yusof, and Dutse 2013). The detection method should be able to identify and map BSR disease quickly and inexpensively. Remote sensing is an appropriate strategy for such detection, and there is potential to improve the accuracy of remote sensing by combining it with machine-learning algorithms. For leaf nutrient monitoring, it would be advantageous to have an improved method that is simple, inexpensive, and accurate. Remote sensing techniques to estimate leaf nutrient status based on the reflectance of leaves have great potential in this regard.

The results of this study contribute to the development of new methods for early detection of BSR disease and for monitoring of leaf nutrient content in oil palm plantations. The main outcomes of this study are as follows:

- A method for the detection and segmentation of individual oil palm trees from high-resolution multispectral satellite imagery using a remote sensing technique.
- A method to extract the pixel values for each oil palm tree from high-resolution multispectral satellite imagery using a remote sensing technique.

- Classification of BSR disease using remote sensing data processed by machine-learning algorithms.
- Mapping of the distribution of BSR disease based on the best machine-learning classification results.
- A method to monitor leaf nutrient contents based on leaf reflectance values measured by a spectroradiometer.

1.4. Objectives

The overall aims of this research were to assess oil palm growth conditions using remote sensing, in some cases combined with R software to extract pixel values, to develop a classification method for BSR using a machine-learning model, and to apply a statistical approach. The proposed new methods were compared with established remote sensing methods.

The specific objectives were as follows:

1. To evaluate the potential of machine-learning algorithms to predict BSR disease in oil palm plantations.
2. To produce a distribution map of BSR disease in the study area using the best machine-learning model.
3. To explore suitable wavelengths for leaf reflectance measurements to predict leaf nutrient contents, especially of N, P, K, Ca, Mg, B, Cu and Zn, which are important for assessing fertilizer requirements in oil palm plantations.

1.5. Thesis structure

This study describes the assessment of oil palm growth conditions in Indonesia using a remote sensing technique, focusing on the identification and classification of BSR disease and the determination of leaf nutrient contents. The general introduction (Chapter 1) explains the oil palm plantation conditions in Indonesia and the main challenges in oil palm cultivation. Chapter 2 explains in detail the detection and segmentation of individual oil palm trees using high-resolution multispectral satellite imagery (QuickBird imagery). This chapter has been published in the International Journal of Remote Sensing (http://dx.doi.org/10.1080/01431161.2016.1226527). From the results in Chapter 2, pixel values of individual oil palm trees were extracted from four bands of QuickBird imagery. The pixel values of individual oil palm trees were used to classify BSR disease using machine-learning algorithms, which are explained in detail in Chapter 3. Chapter 3 presents a map of BSR disease distribution based on the best machine-learning classification results, using data extracted using the methods described in Chapter 2. Chapter 3 has been published in the International Journal of Remote Sensing (http://dx.doi.org/10.1080 /01431161.2017.1331474). Chapter 4 explains the implementation of the machine-learning algorithms to classify BSR severity at several levels. WorldView-3 imagery, which has eight bands as variables for machine-learning algorithms, was used to classify BSR disease into four levels: healthy (H); unhealthy and initially symptomatic (UH1); unhealthy and

moderately symptomatic (UH2); and unhealthy and severely symptomatic (UH3). Chapter 4 will submit to the journal of GIScience & Remote Sensing.

The prediction of oil palm leaf nutrient contents based on spectroradiometer data is explained in Chapter 5. The new vegetation indices proposed in this study are compared with existing vegetation indices, and the prediction of leaf nutrient contents by multivariate analysis is explained. Chapter 5 has been submitted to the Precision Agriculture Journal. Finally, Chapter 6 summarizes the main findings of this study and their importance, and provides recommendations for further research.

Chapter 2. A Simple Method for Detection and Counting of Oil Palm Trees Using High Resolution Multispectral Satellite Imagery

2.1. Introduction

Oil palm density, age, management, and parent material all affect oil palm production. Oil palm tree density is one of the parameters used to determine if replanting is needed and affects oil palm production (Kee 1972; Corley 1973; Breure, Menendez, and Powell 1990; Corley and Tinker 2003; Huth et al. 2014; Bonneau et al. 2014). Therefore, one way to increase oil palm production has been to develop new varieties that are suitable for planting at high densities (Breure 2010). Managers of oil palm plantations typically measure oil palm density manually every year. This important data can be used to estimate oil palm productivity, the amount of fertilizer needed, periodic weeding costs, and the number of workers needed, and is related to other activities (Kee 1972; Kiama, Raman, and Patrick 2014). Of course, counting oil palm trees manually requires many workers and has the potential problem of inaccurate counting.

In the past few years, several studies have dealt with the monitoring and mapping of oil palm using remote sensing in the various contexts. The Moderate Resolution Imaging Spectroradiometer (MODIS) has been used for studies on the primary productivity of oil palm trees as it is related to the global carbon balance (Cracknell et al. 2013; Tan, Kanniah, and Cracknell 2014; Cracknell et al. 2015). Kian Pang Tan, Kanniah, and Cracknell (2013) studied the age of oil palm trees using the Disaster Monitoring Constellation 2 and Advanced Land Observing Satellite

phased array L-band synthetic aperture radar. Moreover, remote sensing has been applied for the identification and monitoring of BSR (Lelong et al. 2010; Helmi Z.M. Shafri et al. 2011; Liaghat et al. 2014; H.Z.M. Shafri and Hamdan 2009; Santoso et al. 2011); several research studies have used remote sensing for oil palm tree detection and counting. Jusoff and Pathan (2009) used airborne hyperspectral sensing with linear spectral mixture analysis, along with a mix to a pure converter and Euclidean norm techniques to map individual oil palm trees. Helmi Z. M. Shafri, Hamdan, and Saripan (2011) achieved 95% accuracy using high spatial resolution airborne imagery with several steps to discriminate oil palms from non-oil palms using spectral analysis, texture analysis, edge enhancement, a segmentation process, morphological analysis and blob analysis. Kattenborn et al. (2014) using unmanned aerial vehicle imagery with point cloud classification (reconstruction of oil palm trees using 3D processing) and achieved a mapping accuracy to 86.1% for the entire study area and 98.2% stands of dense oil palm. Srestasathiern and Rakwatin (2014) using QuickBird imagery with 90% accuracy. They used a vegetation index for distinguishing oil palms from the background, and then applied rank transformation, a non-maximal suppression algorithm and semi-variogram analysis to determine the appropriate window size. In addition, Wong-in et al. (2015) achieved 90% accuracy using aerial images with several steps such as removing non-tree components from an image, distinguishing oil palms from other components using a low-pass filter and normalized cross correlation, identifying individual oil palm trees, and counting the number of oil palms. Korom et al. (2014)

segmented the shape of oil palm canopies/crowns using WorldView-2 imagery based on watershed segmentation and achieved an accuracy of about 77%.

Based on previous research, oil palm plantation managers (oil palm practitioners) need a simple method that is easy to implement for counting oil palm trees in plantations. The main objective of this study is to build a robust and easy-to-use method for counting oil palms using remote sensing.

2.2. Methods

This study analysed oil palms at an oil palm estate in Medan, North Sumatra, Indonesia (3°1.58′–3°13.92′N and 99°5.49′–99°13.44′E; Figure 2.1), with several oil palm densities. Basal stem rot disease in trees caused the observed heterogeneity of oil palm density. QuickBird imagery archived on 4 August 2008 was used in this study with bands consisting of visible red (630–690 nm), green (520–600 nm) and blue (450–520 nm), NIR (near infrared, 760–900 nm) and panchromatic (450–900 nm). Figure 2.2 provides a map of the oil palm density conditions in this study area. Study sites 1, 2, 3 and 4 had 16 (Figure 2.2(*a*)), 21 (Figure 2.2(*b*)), 15 and 18 (Figure 2.2(*c*)), and 10 (Figure 2.2(*d*)) year-old oil palms, respectively. Various conditions surrounding oil palm trees, such as the density of grass and other weeds, will affect oil palm tree detection and counting (Jusoff and Pathan 2009; Helmi Z. M. Shafri, Hamdan, and Saripan 2011; Kattenborn et al. 2014; Wong-in et al. 2015). Therefore, the main target in the present study is to build a robust method for oil palm detection and counting under any oil palm conditions and densities.

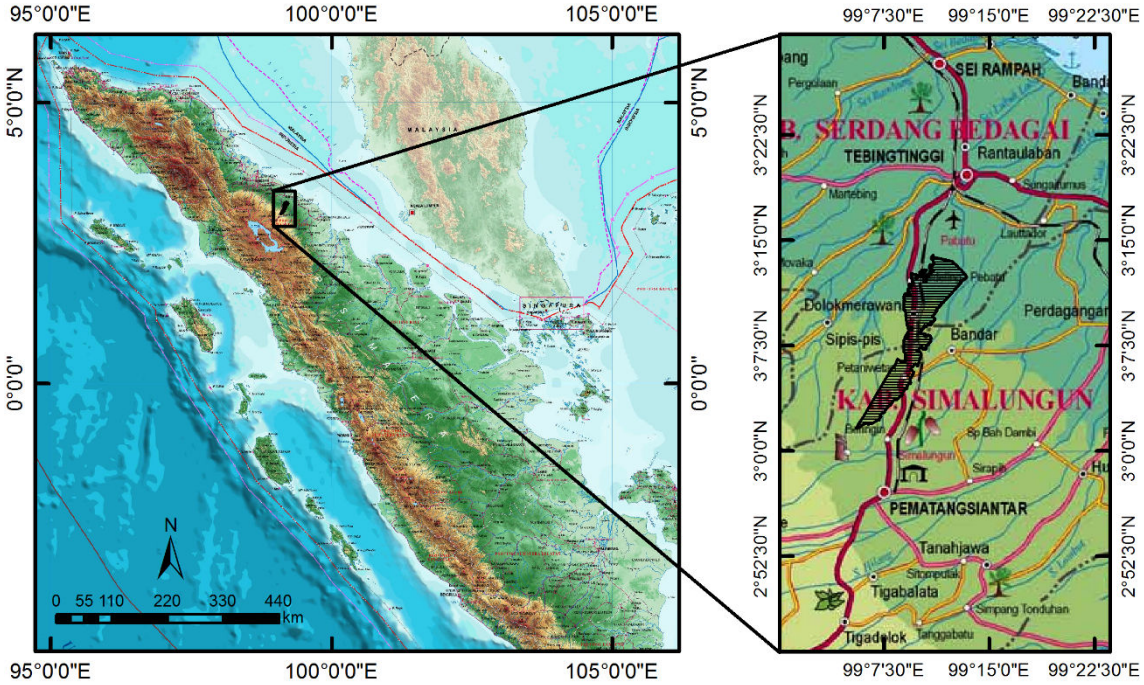


Figure 2.1. The study site is on an oil palm plantation in North Sumatra, Indonesia; A) vicinity map showing the vicinity of the study site on the island of Sumatra; 2) map of the study site (Source map: Indonesian Geospatial Board-http://www.bakosurtanal. go.id/download/).

Figure 2.3 shows the flow of the three-part method proposed in the present study: 1) oil palm tree detection, 2) delineation of the oil palm area, and 3) oil palm tree counting and accuracy assessment. ENVI 5.2, ERDAS Imagine 2015, and ArcGIS 10.2.2 were the types of software used in this study. As a common standard for digital image processing, converting a digital number into radiance was done using ENVI 5.2. In this research study, pansharpening was used for oil palm detection because this allows the shape of the oil palm canopy to be detected into detail. Johnson, Tateishi, and Hoan (2012) reported that a pansharpened image has more spatial information and minimizes the distortion of spectral information of a multispectral image with lower resolution. Lin et al. (2015) reported that

pansharpening can improve and increase the classification accuracy of land use and land cover using WorldView-2 imagery.

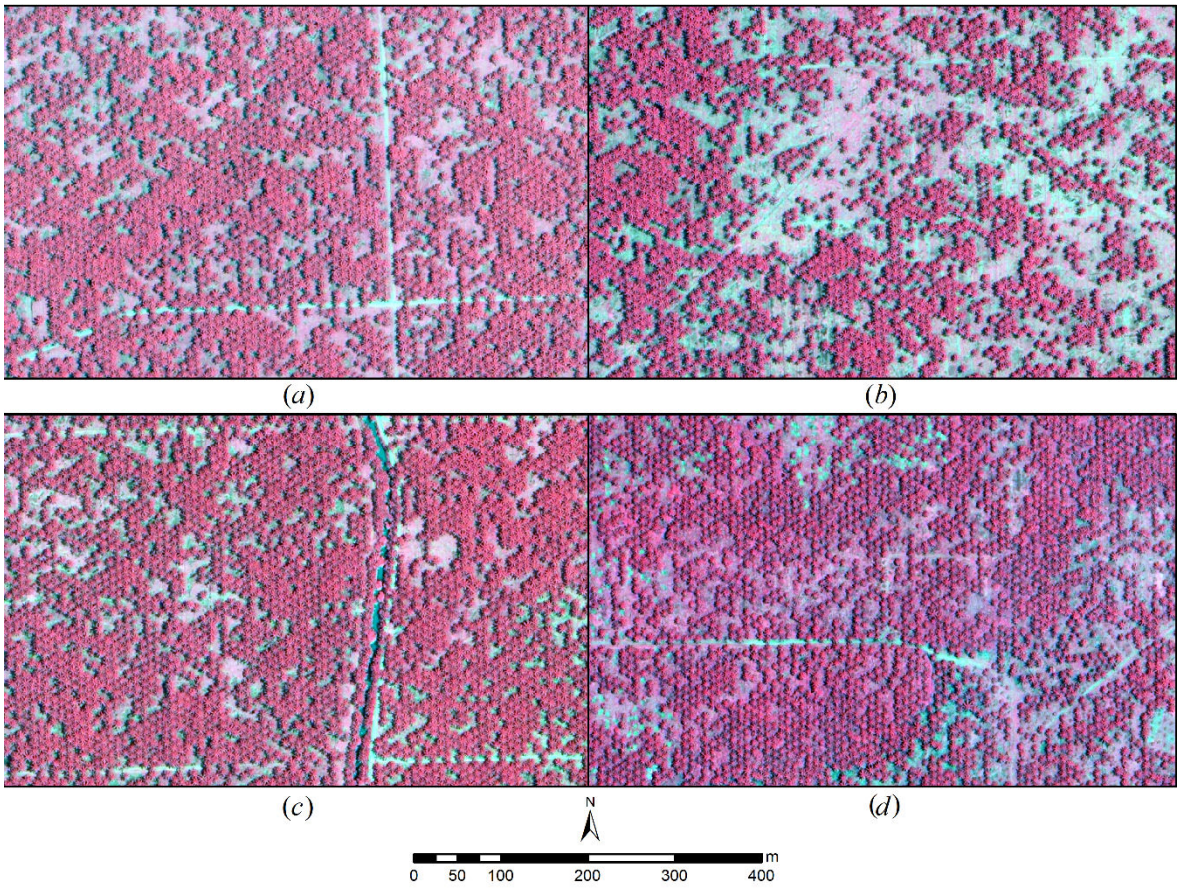


Figure 2.2. Oil palm density condition in this study showing parts of (a) study site 1 (150.16 ha); (b) study site 2 (139.31 ha); (c) study site 3 (130.49 ha); (d) study site 4 (128.99 ha).

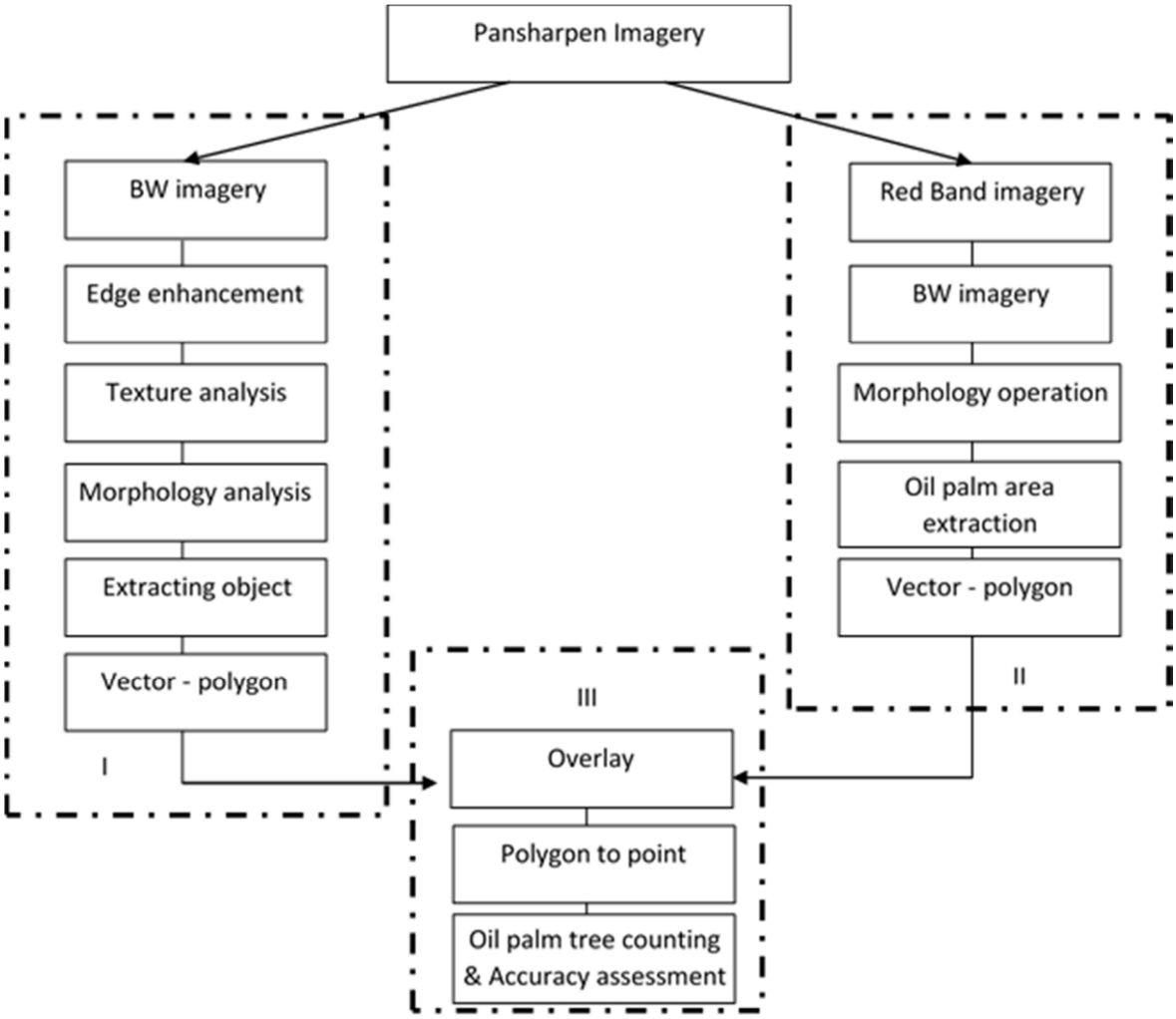


Figure 2.3. The study flow chart: (I) oil palm tree detection; (II) delineation of oil palm area, and (III) oil palm tree counting and accuracy assessment.

Several pansharpening methods were used and compared with each other in the present study to determine the best pansharpening method for oil palm detection. Panchromatic imagery was also used in this study. Colour Normalized (Brovey), Gram-Schmidt Spectral Sharpening, and PC Spectral Sharpening methods were processed with ENVI 5.2. Meanwhile, Subtractive Resolution Merge, Modified HIS Resolution Merge, and Wavelet Resolution Merge were processed with ERDAS Imagine 2015.

2.2.1. Oil palm detection

First, a pre-processing step was employed. All processing steps in the first part were completed using ENVI 5.2. The false colour composite RGB 432 (NIR-R-Green) images were built from pansharpening imagery and were then converted to a grayscale image (Korom et al. 2014). The pre-processing step continued using ENVI colour tables with a grayscale image as an input file. We chose a black and white linear colour table and then adjusted the stretch bottom slider to the right side and stretch top slider to the left until 70–75% of a wide slider (ENVI 2014). The image became light in the background, black in the foreground (oil palm canopy), and then the new image was saved (Guijarro et al. 2011; Müller-Linow et al. 2015). The further steps conducted after pre-processing are described below:

- Edge enhancement. The oil palm canopy shape was detected using a Sobel edge detector with approximately 75–85% of an image added back.
- Texture analysis in this study used co-occurrence with a mean parameter, 5 × 5 of processing windows and 3 × 3 of co-occurrence shift.
- Morphology analysis in this study consisted of two filters, a dilate filter with a 3 × 3 kernel size and an erode filter with a 5 × 5 kernel size.
- In the extracting objects step, a high-pass filter with a 9 × 9 kernel size was used and then we continued using an opening filter with a 5 × 5 kernel size. When the detected object was an oil palm tree, it was converted to a polygon.

2.2.2. Oil palm area delineation

The second part of the delineation of the oil palm area, a pre-processing step, was needed; this was done in the same way as the first part, but using the red band (Santoso et al. 2011) from pansharpening imagery as an input file. The image had become lighter in the foreground (oil palm area) and darker (black) in the background. After pre-processing, the morphology operation in this part was a closing filter with a 5 × 5 or a 7 × 7 kernel size. The object that was detected was the oil palm area and should be converted to polygon format as a masking polygon. In the final part in this study, oil palm tree counting was done in ArcGIS 10.2.2 using an overlay of both polygons from first and second parts; then the conversion processes were completed using polygons from overlay results as point features.

2.3. Results and discussion

The proposed methods for oil palm tree detection and counting in the present study were tested in several areas with different oil palm ages and densities. Figure 2.4 shows an example of the results of each process in the proposed method that was used in the present study. Figure 2.4(*a*) shows the shape of the oil palm canopy as dark-black with the light areas indicating the background. An oil palm canopy takes on a star-like shape; there are outer angular shapes around the centre of the canopy. The Sobel edge detector, texture analysis, and dilate filter were applied to minimize the outer shape of the oil palm canopies (see Figure 2.4(*b*), 2.4(*c*) and 2.4(*d*)). The erode filter produced an oil palm canopy outline as a square shape (Figure 2.4(*e*))

and the high-pass filter sharpened this square shape (Figure 2.4(*f*)). Moreover, the opening filter converted the square shape into the object that was selected and the background is erased or left empty (Figure 2.4(*g*)).

During oil palm tree extraction, several oil palm canopy/trees were not extracted and some objects were also extracted that were not oil palms (Figure 2.4(*i*)). This is caused when some of the vegetation that surrounds oil palm trees has the same radiance values as the radiance values of oil palms. Moreover, some areas were classified as noise. In a further step, the noise was reduced using the oil palm area extraction areas with the red band. Meanwhile, a part of the oil palm canopy was not extracted; this was caused when the oil palm canopy had some different pixel radiance values while only a small part of a particular tree had the same values with oil palm radiance; however, in these cases, the majority of the oil palms had the same values as the surrounding radiance values. Moreover, the small part of the pixel values of oil palm radiance were erased by some kernel/window size areas in filters applied in this study (Hosoi et al. 2012; Warner, McGraw, and Landenberger 2006). In addition, this change affected the accuracy assessment. The oil palm area extraction results will affect the accuracy of oil palm tree counting because they are used to reduce noise during oil palm tree extraction.

2.3.1. Oil palm tree counting and accuracy assessment

After using a red band to reduce noise during oil palm area extraction, oil palm trees could be counted from the results of the converted polygon of oil palm extraction

to create a point layer. The present study encompassed a total area of 548.95 ha for testing the proposed methods, which were divided into four different study sites areas to determine oil palm density and age. Tables 2.1–2.4 show the ability of the method to count oil palm trees and the count accuracy assessment for study sites 1–4. We assumed the manually counted values were accurate because we used QuickBird pansharpening imagery to manually count trees based on screen digitized imagery of oil palm trees in five square areas where each area covered 10,000 m^2. Each area and pansharpening method used in this study resulted in different oil palm tree counts. The research results from Wyczałek and Elżbieta (2013) produced similar results while using a different method of pansharpening that will affect the classification results. Tables 2.1–2.4 show the potential of the four pansharpening methods for counting oil palm trees using PC Spectral Sharpening, Colour Normalized (Brovey), Subtractive Resolution Merge, and Modified Intensity-Hue-Saturation Resolution Merge, at study sites 1–4, respectively.

The potential of using the pansharpening method for counting oil palm trees in different areas is shown in Table 2.5. Study sites 1, 2, 3 and 4 had 16, 21, 15/18 and 10-year-old oil palm trees, respectively. Study sites 1, 2, 3 and 4 had rather high, low, rather high, and low oil palm density, respectively. With these conditions, the best pansharpening methods for study sites 1, 2, 3 and 4 were Modified IHS Resolution Merge pansharpening (100% accuracy), Colour Normalized (Brovey) pansharpening (99.5%), Subtractive Resolution Merge pansharpening (99.8%), and PC Spectral Sharpening (99.3%), respectively. Two pansharpening methods

employed for oil palm tree extraction and counting in this study had 98.4% and 98.9% accuracy and were Wavelet Resolution Merge and Colour Normalized (Brovey), respectively (Table 2.6; Figure 2.5).

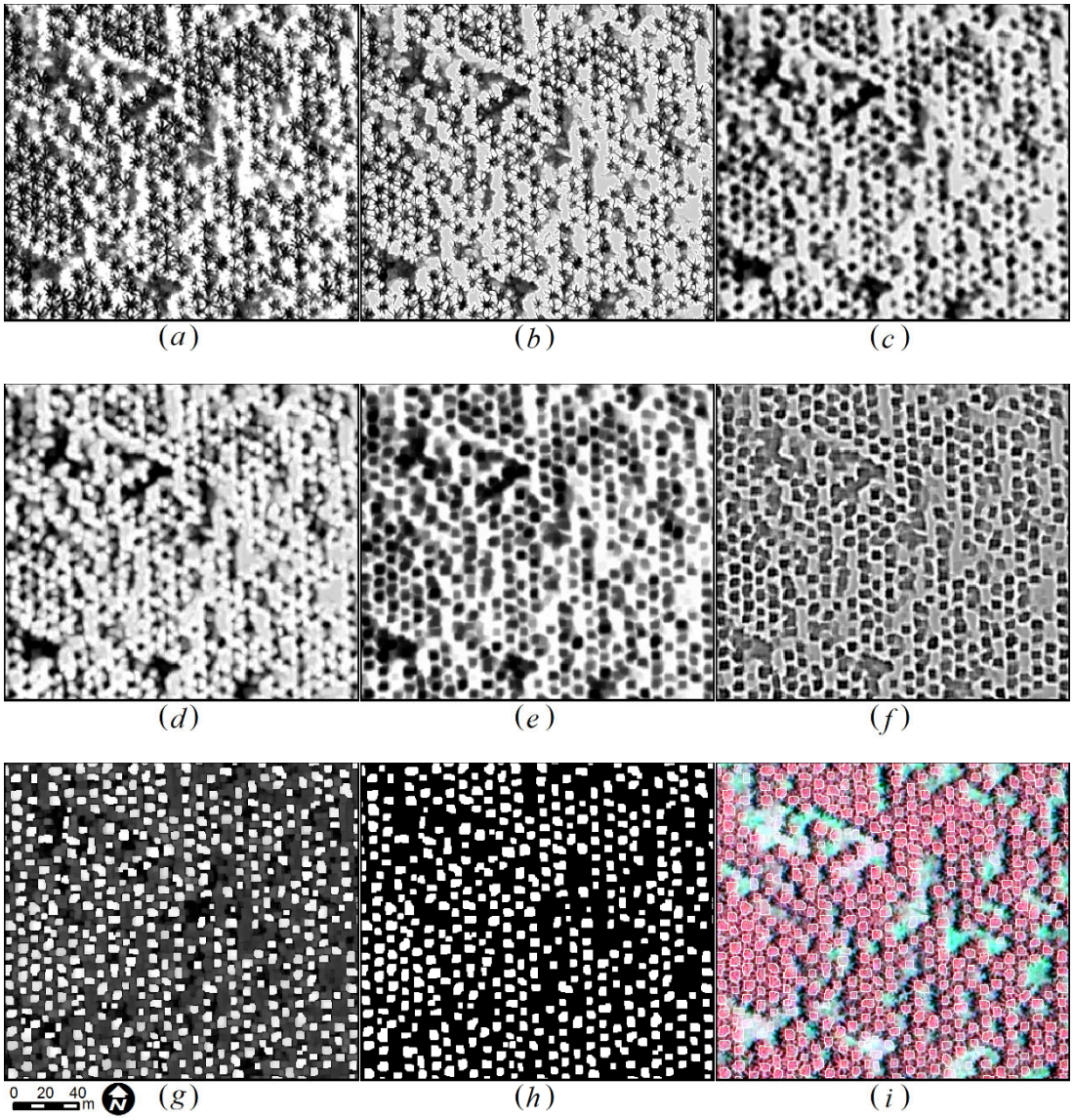


Figure 2.4. Example results of the proposed methods in this study: (a) black and white imagery; (b) image from Sobel applied; (c) image after texture analysis; (d) image after dilate filter; (e) image after erode filter; (f) image after high-pass filter; (g) image after opening filter; (h) oil palm object extraction; (i) polygon of oil palm extraction with pansharpening imagery as background.

Table 2.1. Capability and accuracy assessment for study site 1.

Method	Area (ha)	Total of oil palm extraction	Polygon test (ha)	Oil palm counted manually	Oil palm extraction	Difference	Accuracy (%) (compared with counted manually data)
Panchromatic	150.16	15,421	5	514	529	(15)	97.2
Gram Schmidt Spectral Sharpening	150.16	13,306	5	514	483	31	94.0
PC Spectral Sharpening	150.16	12,125	5	514	439	75	85.4
Colour Normalized (Brovey)	150.16	14,107	5	514	485	29	94.4
Subtractive Resolution Merge	150.16	13,216	5	514	475	39	92.4
Modified Intensity-Hue-Saturation Resolution Merge	150.16	14,471	5	514	514	-	100.0
Wavelet Resolution Merge	150.16	13,601	5	514	486	28	94.6

Table 2.2. Capability and accuracy assessment for study site 2.

Method	Area (ha)	Total of oil palm extraction	Polygon test (ha)	Oil palm counted manually	Oil palm extraction	Difference	Accuracy (%) (compared with counted manually data)
Panchromatic	139.31	11,263	5	400	433	(33)	92.4
Gram Schmidt Spectral Sharpening	139.31	9,604	5	400	370	30	92.5
PC Spectral Sharpening	139.31	9,564	5	400	375	25	93.8
Colour Normalized (Brovey)	139.31	10,703	5	400	398	2	99.5
Subtractive Resolution Merge	139.31	9,668	5	400	369	31	92.3
Modified Intensity-Hue-Saturation Resolution Merge	139.31	10,957	5	400	415	(15)	96.4
Wavelet Resolution Merge	139.31	10,111	5	400	392	8	98.0

Table 2.3. Capability and accuracy assessment for study site 3.

Method	Area (ha)	Total of oil palm extraction	Polygon test (ha)	Oil palm counted manually	Oil palm extraction	Difference	Accuracy (%) (compared with counted manually data)
Panchromatic	130.49	12,869	5	471	495	(24)	95.2
Gram Schmidt Spectral Sharpening	130.49	11,629	5	471	454	17	96.4
PC Spectral Sharpening	130.49	12,261	5	471	465	6	98.7
Colour Normalized (Brovey)	130.49	13,000	5	471	495	(24)	95.2
Subtractive Resolution Merge	130.49	12,297	5	471	472	(1)	99.8
Modified Intensity-Hue-Saturation Resolution Merge	130.49	12,817	5	471	493	(22)	95.5
Wavelet Resolution Merge	130.49	12,399	5	471	469	2	99.6

Table 2.4. Capability and accuracy assessment for study site 4.

Method	Area (ha)	Total of oil palm extraction	Polygon test (ha)	Oil palm counted manually	Oil palm extraction	Difference	Accuracy (%) (compared with counted manually data)
Panchromatic	128.99	13,183	5	449	514	(65)	87.4
Gram Schmidt Spectral Sharpening	128.99	11,414	5	449	464	(15)	96.8
PC Spectral Sharpening	128.99	11,137	5	449	446	3	99.3
Colour Normalized (Brovey)	128.99	12,127	5	449	476	(27)	94.3
Subtractive Resolution Merge	128.99	11,272	5	449	444	5	98.9
Modified Intensity-Hue-Saturation Resolution Merge	128.99	12,722	5	449	499	(50)	90.0
Wavelet Resolution Merge	128.99	11,579	5	449	458	(9)	98.0

Table 2.5. The potential of the pansharpening method for oil palm tree extraction in different areas.

Method	Accuracy (%) (compared with counted manually data)			
	Area 1	Area 2	Area 3	Area 4
Panchromatic	97.2	92.4	95.2	87.4
Gram Schmidt Spectral Sharpening	94.0	92.5	96.4	96.8
PC Spectral Sharpening	85.4	93.8	98.7	**99.3**
Colour Normalized (Brovey)	94.4	**99.5**	95.2	94.3
Subtractive Resolution Merge	92.4	92.3	**99.8**	98.9
Modified Intensity-Hue-Saturation Resolution Merge	**100.0**	96.4	95.5	90.0
Wavelet Resolution Merge	94.6	98.0	99.6	98.0

2.3.2. Practical implications

When compared with the accuracy of oil palm tree counting in previous research studies that used several methods, the methods proposed in the present study are potentially useful for oil palm planters and plantation managers and could be adopted and implemented as part of oil palm management practices. This proposed method is simple and all the steps and filter are available in ENVI software. In addition, it is possible to use different image processing software that has a grayscale convert function, a foreground and background stretching function, a Sobel edge detector, and texture analysis co-occurrence, as well as dilate, erode, high-pass, and opening filters.

We believe the accuracy achieved using the methods employed in the present study is adequate for estate managers to use so they can employ remote sensing to rapidly determine oil palm density. In oil palm management, if the estate has

planting records, especially stand maps of oil palm plantations, monitoring oil palm trees individually in the field is very easy. However, sometimes oil palm plantation managers need a second opinion and can accurately monitor oil palm density using imagery based on remote sensing.

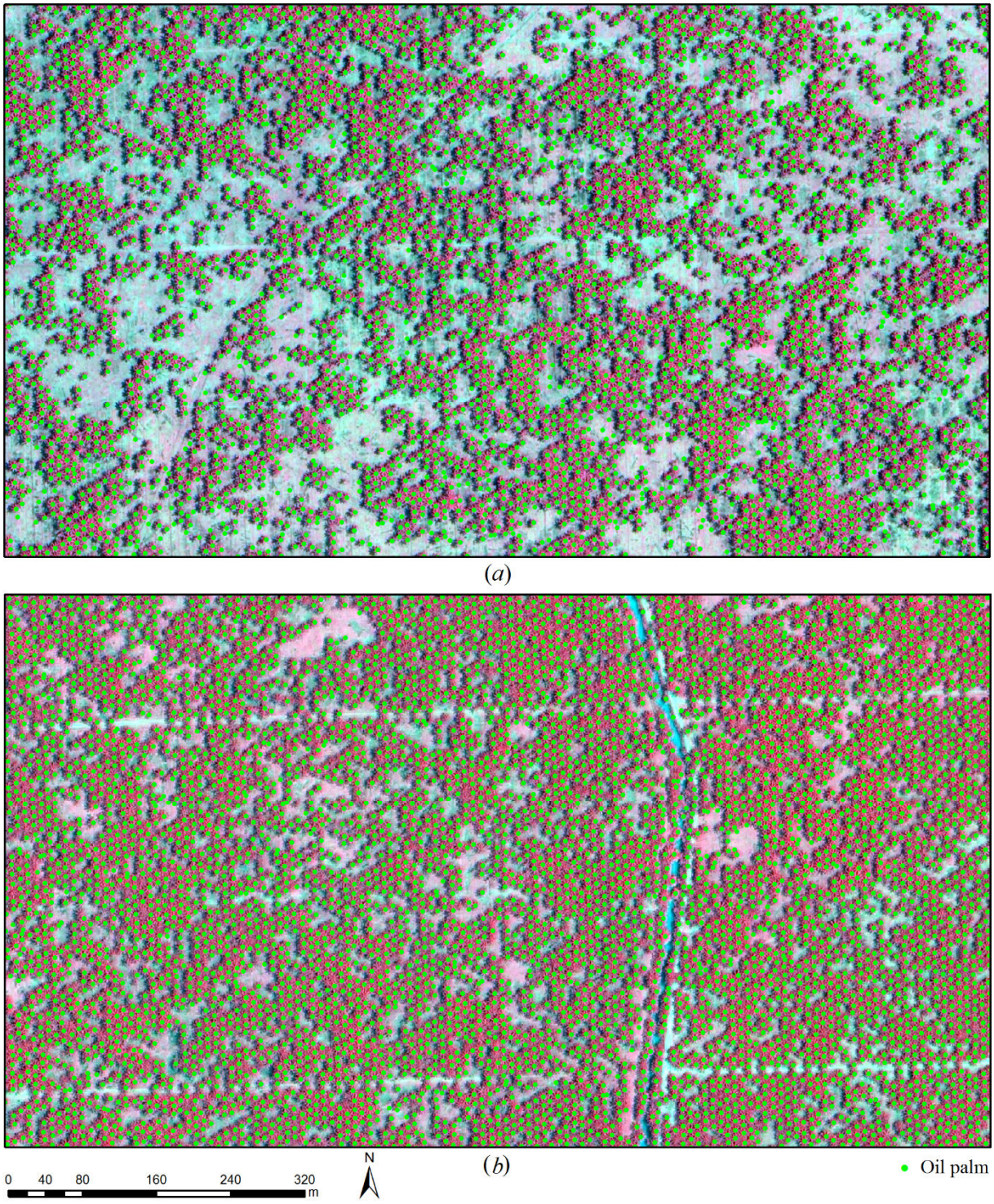


Figure 2.5. Oil palm extraction and counting using Colour Normalized (Brovey) in study site 2 (a) and using Wavelet Resolution Merge in study site 3 (b).

Table 2.6. Accuracy assessment of oil palm tree extraction.

Method	Area (ha)	Total of oil palm extraction	Polygon test (ha)	Oil palm counted manually	Oil palm extraction	Difference	Accuracy (%) (compared with counted manually data)
Panchromatic	548.95	52,736	20	1,834	1,971	(137)	93.0
Gram Schmidt Spectral Sharpening	548.95	45,953	20	1,834	1,771	63	96.6
PC Spectral Sharpening	548.95	45,087	20	1,834	1,725	109	94.1
Colour Normalized (Brovey)	548.95	49,937	20	1,834	1,854	(20)	98.9
Subtractive Resolution Merge	548.95	46,453	20	1,834	1,760	74	96.0
Modified HIS Resolution Merge	548.95	50,967	20	1,834	1,921	(87)	95.5
Wavelet Resolution Merge	548.95	47,690	20	1,834	1,805	29	98.4

2.4. Conclusions

The pansharpening method used with multispectral satellite imagery such as a QuickBird image has potential for use in oil palm tree detection and counting. The different pansharpening methods provided different results for oil palm tree counting. The methods proposed in the present study exhibited about 98% accuracy and provide potential methods for oil palm tree detection and counting as part of oil palm management. The methods proposed in the present study are easy to use with ENVI or other image processing software that have a grayscale conversion function, stretching foreground and background functions, a Sobel edge detector, texture analysis co-occurrence, as well as dilate, erode, high-pass, and opening filters. The proposed methods will be tested using images from unmanned aerial vehicles that many oil palm companies currently operate for surveillance and monitoring of oil palm coverage.

Chapter 3. Random Forest Classification Model of Basal Stem Rot Disease in Oil Palm Plantations

3.1. Introduction

Oil palm is an economically important plant in Indonesia and Malaysia (Basiron 2007; World Growth 2011) and has a productive life of up to 25 years. All parts of the oil palm plant can be used such as for vegetable oil (potential oil extraction of 6–9 ton ha^{-1}), food, energy (bio fuel, biogas and electricity generated from fibre and empty bunches), medicine, a blending material for the pulp industry, and bio fertiliser (from on- and-off farm waste). Additionally, oil palm fronds and by-products from oil palm kernel milling can be processed into animal feed for cattle and sheep (Zahari, Alimon, and Wong 2012; Hayashi 2007; Sarmidi, Enshasy, and Hamid 2009; Subramaniam et al. 2010; Abdullah and Sulaiman 2013; Embrandiri, Ibrahim, and Singh 2013; Mayulu 2014; Mba, Dumont, and Ngadi 2015). Oil palm production is affected by environmental and plant-related factors, including planting material, meteorological conditions, soil conditions and nutrition, pests and diseases, and on- and off-farm management (Corley and Tinker 2003). The fungus *Ganoderma orbiforme* (*G. boninense*) causes basal stem rot (BSR) disease in oil palm and has become a major disease in oil palm plantations in both Indonesia and Malaysia. No effective cure for BSR exists, and current treatments only prolong the life of oil palms (Idris et al. 2000; Hushiarian, Yusof, and Dutse 2013; Susanto, Prasetyo, and Wening 2013; Priwiratama, Prasetyo, and Susanto 2014; Priwiratama and Susanto 2014).

One strategy to control BSR is early detection BSR symptoms (Hushiarian, Yusof, and Dutse 2013). BSR symptoms can be identified in the field by observing the condition of the fronds and trunk. Commonly, BSR disease symptoms include more than two of the youngest leaves not opening and necrosis of older leaves. Other symptoms include more than three of the youngest leaves not opening, fractured old fronds and fungal fruiting bodies on the oil palm trunk (Kee 1972; Corley, J.J., and B.J. 1976; Turner 1981; Corley and Tinker 2003). Knowing the status of oil palm health and the spread of BSR in oil palm plantations is important for early disease control through plantation management. Monitoring the disease will hopefully prolong the life of oil palms and raise productivity (Priwiratama, Prasetyo, and Susanto 2014). Using BSR symptoms, many researchers have applied remote sensing techniques for early detection and mapping of BSR disease in oil palms. Furthermore, the primary productivity of oil palm trees has been studied in various contexts by monitoring and mapping plantations using remote sensing technology. Cracknell et al. (2013), Tan, Kanniah, and Cracknell (2013), and Cracknell et al. (2015) used moderate resolution imaging spectroradiometer to study global carbon balance. Tan, Kanniah, and Cracknell (2013) studied the age of oil palm trees using the Disaster Monitoring Constellation 2 and Advanced Land Observing Satellite phased array L-band synthetic aperture radar.

Three kinds of remote sensing techniques have been used to study BSR. First, BSR has been detected using a hyperspectral spectrometer: Lelong et al. (2010) obtained 94% overall accuracy (OA); Helmi Z.M. Shafri et al. (2011) achieved 82%

net accuracy and a kappa value of 0.72%; and Liaghat et al. (2014) achieved 97% OA using a *k*-nearest neighbour (*k*NN)-based classification model. Second, BSR has been detected using airborne hyperspectral imagery (H.Z.M. Shafri and Hamdan 2009) with vegetation indices and the red edge technique to achieve 73%–84% accuracy; the highest accuracy was obtained using the Lagrangian interpolation technique. Third, Santoso et al. (2011) mapped and identified BSR disease using QuickBird multispectral imagery with 69%–85% interpretation accuracy and 32%–67% mapping accuracy. However, it was difficult to individually distinguish healthy and unhealthy oil palms from the map generated by Santoso et al. (2011) because both healthy and unhealthy classes had more than one pattern of pixel composition and the individual oil palms were drawn with different pixel amounts. Therefore, there is potential to improve the accuracy of multispectral imagery and relevant techniques to identify, classify, and map the distribution of infected oil palm trees in the study area.

One potential method for improving the study of BSR is using machine learning algorithms. Machine learning algorithms are techniques that use automated computation to estimate a function with input variables (x) that best predicts the output variable (y) (Ayodele 2010; Brownlee 2016; Burrell 2016). Supervised learning algorithms have been used frequently in classification-related medical studies (Bind, Tiwari, and Sahani 2015; Khalilia, Chakraborty, and Popescu 2011; Furlanello et al. 2003; Kononenko 2001), crop classification (Sonobe et al. 2014a; Sonobe et al. 2014b; Watts and Lawrence 2008), and land cover mapping

(Stefanski, Mack, and Waske 2013; Jhonnerie et al. 2015; Lowe and Kulkarni 2015). Learning algorithms have also been applied to the study of plant diseases. Liu, Kelly, and Gong (2006) monitored disease spread in forests; Römer et al. (2011) detected wheat leaf rust; and El-telbany and Warda (2016) classified Egyptian rice disease, all using learning algorithms. Therefore, there is a challenge to implement machine learning algorithms for BSR disease prediction using multispectral (QuickBird) imagery data with the four bands as variables. The main objectives of the current study were to evaluate the potential of machine learning algorithms for identifying BSR disease in oil palm plantations and to produce distribution maps of BSR disease in the study area using the best-fitting machine learning algorithms.

3.2. Data and Methods

3.2.1. Study area

The study area and data were adopted from Santoso et al. (2011). The study area was an oil palm estate in North Sumatra, Indonesia, located within 3°1.580'–3°13.920' N and 99°5.490'–99°13.440' E (Figure 3.1). This study used QuickBird data archived on 4 August 2008 with the following bands: visible red (630–690 nm), green (520–600 nm), blue (450–520 nm), near infrared (NIR; 760–900 nm), and panchromatic (450–900 nm). There were four sites with different oil palm ages, severities of disease, and oil palm densities. Site 1 had 16-year-old oil palms covering 150.16 ha; site 2 had 21-year-old palms covering 139.31 ha; site 3 had 15- and 18-year-old palms covering 130.49 ha; and site 4 had 10-year-old palms

covering 128.99 ha (Figure 3.2).

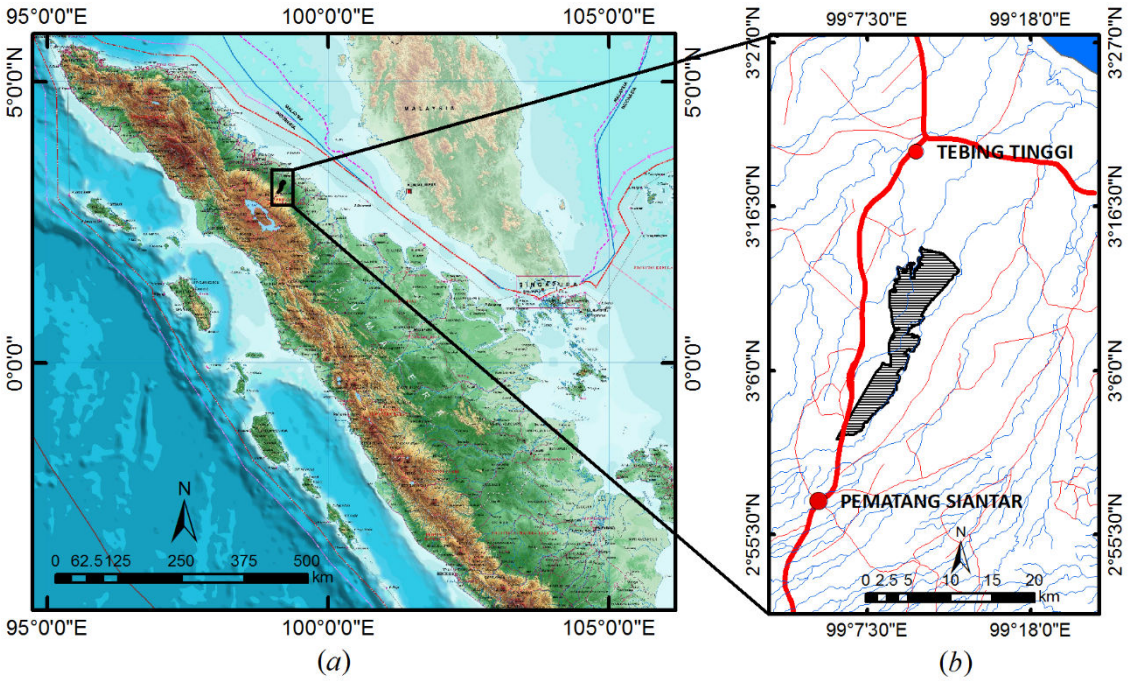


Figure 3.1. The vicinity of the study area on Sumatra island (a); the map of the study area (b) (source: BIG 2015).

3.2.2. Data analysis and evaluation

There were four activities in this study (Figure 3.3): (1) field observations; (2) segmentation and extraction of all the pixel values for oil palms in the study sites; (3) applying the learning model in R software; and (4) producing distribution maps of healthy and unhealthy oil palms based on the best classifier model.

First, the sample data for healthy and unhealthy palms were collected. The data used in this study were adopted from Santoso et al. (2011), and only healthy and unhealthy data were used for the learning model. There were 144 data consisted of 99 healthy and 45 unhealthy oil palm samples. Second, we carried out segmentation and extraction. The segmentation process involves detecting and

drawing out each oil palm object in all the study sites and the extraction process involves obtaining the pixel values for each oil palm in the study as variable values for the learning algorithm. We applied the Segmentation step according to the methods of Santoso, Tani, and Wang (2016); we used Brovey pansharpening of QuickBird imagery and applied filters including Sobel, texture co-occurrence, dilation, erode, high pass, and opening. Segmentation was processed in ENVI 5 and ArcGIS 10.2.2, and the output was the points of all oil palm objects in the study area. After segmentation, we extracted the pixel values of all bands of the QuickBird multispectral images, based on the points of all oil palms that were segmented. Extraction of pixel values was carried out in RStudio software (RStudio 2015) using scripts adopted from Neon (2014), NCEAS (2014), and Stackoverflow (2014). To extract the pixel values, the points of all oil palms from the segmentation step were converted into square polygons of nine pixels each with a length of 3.6 m from the centre of the polygon. The value 3.6 m was calculated based on the diameter of the oil palm canopy and the pixel size of the QuickBird imagery. The formula used to determine this was 2.4 m QuickBird pixel size multiplied by three pixels and divided by 2. The mean pixel value was then calculated from the nine pixel values inside each polygon representing a single oil palm object.

The third step was the learning model. In this step, the mean pixel values based on field observations were used for the learning model. The mean pixel values from the four bands of QuickBird imagery were used as variables in this study. The other mean values of polygons were used for the fourth step. The

learning model was processed in RStudio using the Caret package (Kuhn 2015a) for data splitting, preprocessing, tuning and validating data, and testing the model. The data were randomized with stratified random sampling based on the amounts of each class in the sampling data and separated into training data and testing data. To split the data, we used the ratio of Liaghat et al. (2014), in which 75% of data was used for training and 25% for testing. A total of 144 data points were used in the current study, split into 109 data points for training and 35 for testing (Table 3.1). The learning algorithms applied were the support vector machine (SVM), random forest (RF), and classification and regression tree (CART) models. These three learning algorithms are often used for classification studies in the fields of environmental science, agriculture and ecology (Karatzoglou, Meyer, and Hornik 2006; Prasad, Iverson, and Liaw 2006; Wiesmeier et al. 2011; Du et al. 2014; Sonobe et al. 2014b).

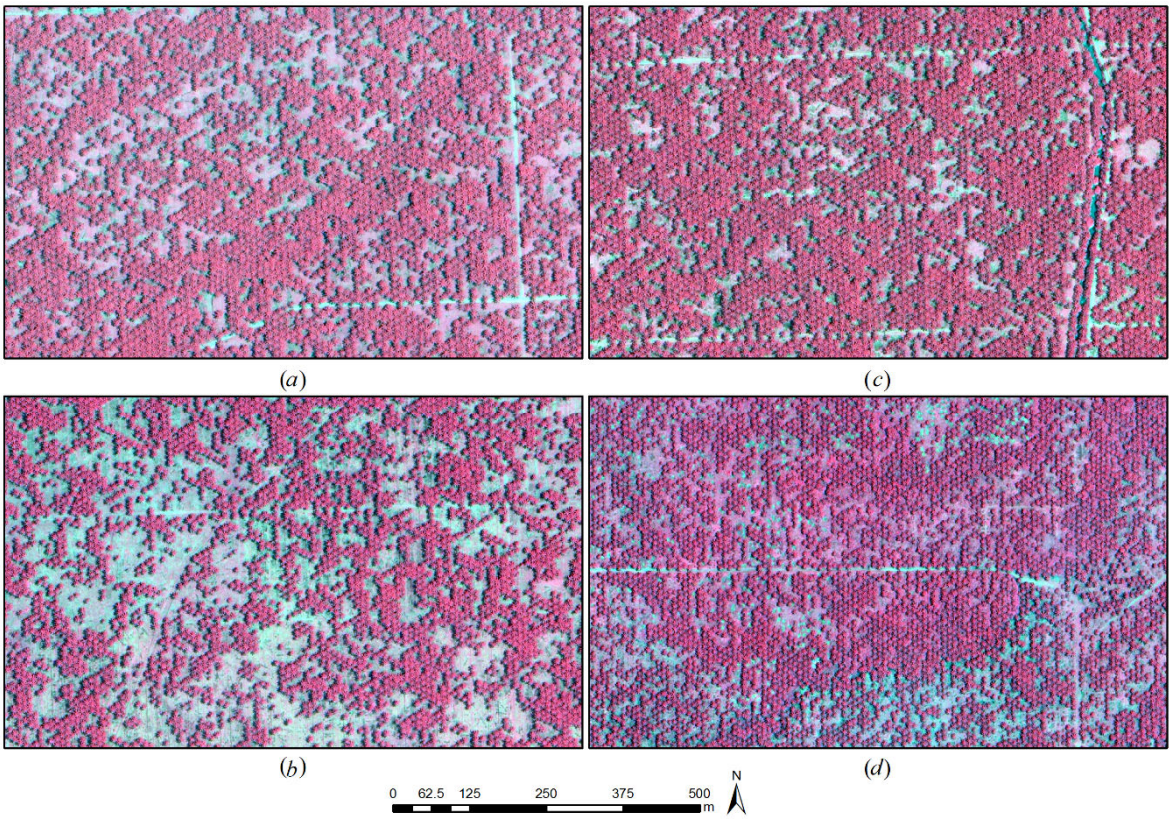


Figure 3.2. Oil palm conditions for study sites 1–4; shown as (a), (b), (c) and (d), respectively). The red points in the QuickBird images (RGB = 432) are oil palms.

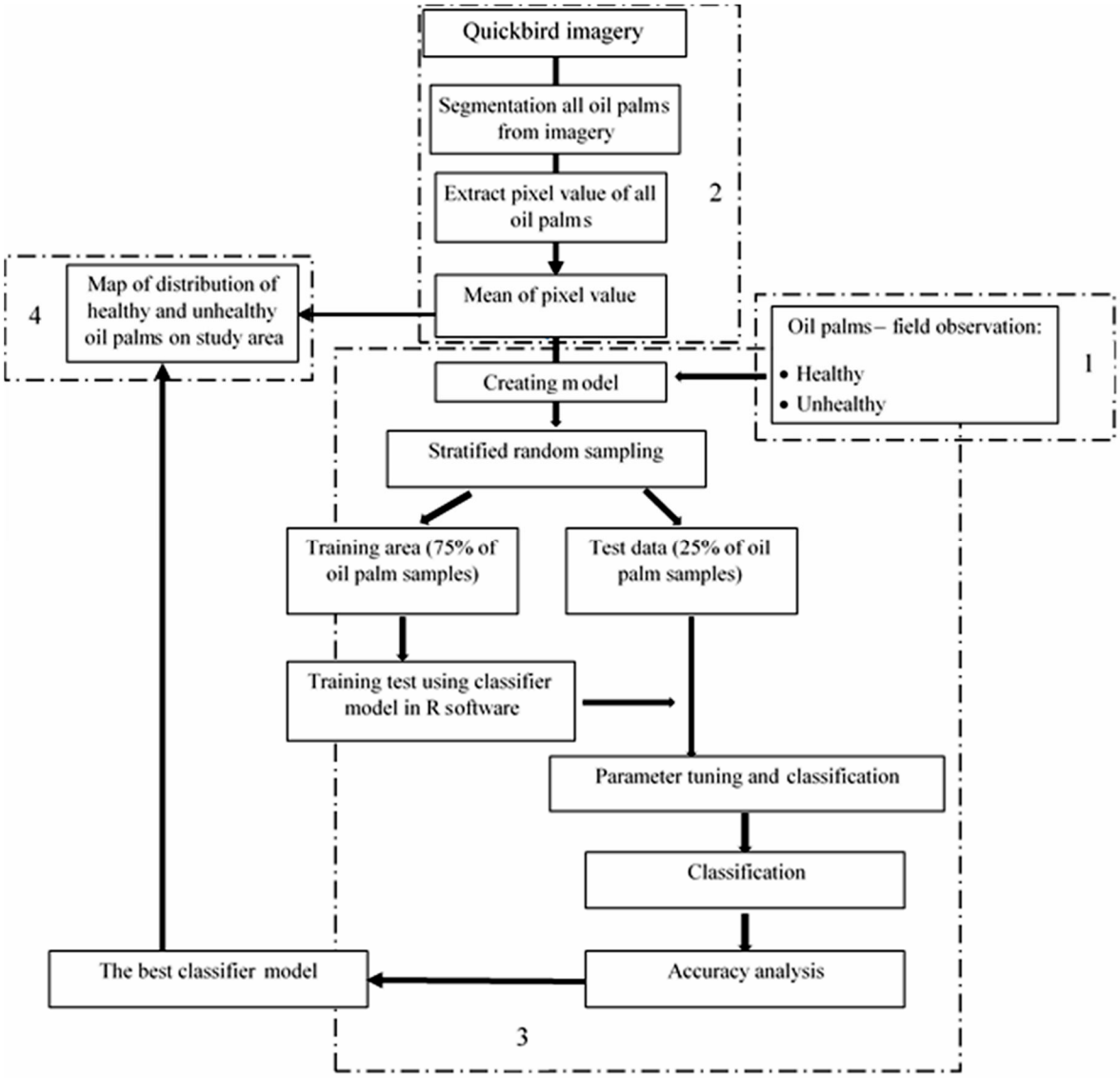


Figure 3.3. Study flow: (1) field observations; (2) segmentation and extraction of all pixel values for oil palms in the study area; (3) applying the learning model in R software; and (4) producing distribution maps of healthy and unhealthy oil palms based on the best classifier model.

Table 3.1. Data splitting.

Splitting data	Percentage (%)	Healthy	Unhealthy	Total
Training	75	74	35	109
Testing	25	25	10	35
Total	100	99	45	144

In this study, all classifier models used training control with 10 repeated cross-validations (kFold = 10), repeated 10 times for resampling data. All classifier models applied preprocessing with principal component analysis (PCA) to fit the best parameters to be used in model prediction. Using PCA for preprocessing means that the training data values will be centred and scaled (Kuhn 2016). According to Kuhn (2008) and Kelly (2014a), the SVM classifier model with the SVM radial method requires the packages kernlab, *sigma*, and *cost* (*C*) as tuning parameters. The RF classifier model requires the packages randomForest and *mtry* as tuning model parameters (Breiman 2004; Kuhn 2008; Kuhn 2015a; Kelly 2014b). The CART classifier model requires the packages rpart and *complexity parameter* (CP) as tuning parameters (Kuhn 2015b). The models were evaluated using a confusion matrix, and the kappa value was used to measure the classifier model performance (Short 1982; Congalton 1991; Congalton 2005; J. Cohen 1960; Viera and Garrett 2005). The fourth step was producing a distribution map of oil palms infected by BSR using the best model according to the confusion matrix. To construct the map, all pixel values of oil palms from the segmentation step were processed in R software, excluding the 144 data points used for building the classifier model. Finally, the map was further processed for layout in ArcGIS 10.2.2.

3.3. Results and discussion

The tuning parameters were selected from the best parameters using the Caret package based on the highest accuracy values. The results of the tuning parameters

for the SVM, RF, and CART classifier models are shown in Figure 3.4–3.6, respectively. The best tuning parameter for the SVM classifier model were *sigma* value = 0.5036928 and *C* = 4, the best parameter for RF was *mtry* value = 2, and the best parameter for CART was CP value = 0.01.

The accuracy assessment of each model is shown in Table 3.2. The RF classifier model had higher OA of 91% and kappa value of 0.81 compared with the CART (OA = 80%, kappa value = 0.57) and SVM models (OA = 77%, kappa value = 0.52). The kappa values indicated that the CART and SVM classifier models had moderate agreement and the RF classifier model had almost perfect agreement (Viera and Garrett 2005). Producer's accuracy (PA) and user's accuracy (UA) values showed that the RF classifier model had high accuracy of classification for both healthy and BSR-infected oil palms. The SVM and CART classifier models gave low PA in the healthy class and this was correlated with their low UA values in the unhealthy class.

Compared with our previous work (Santoso et al. 2011), we used the same data for identifying and mapping oil palms infected by *G. boninense* with several vegetation indices; the present study results showed improved accuracy especially with the RF classifier model for mapping healthy and infected palms. In Santoso et al. (2011), the mapping accuracy was 63–67% for healthy oil palms and 50–53% for unhealthy palms caused by BSR. However, the present study had 88% mapping accuracy for healthy and 77% for unhealthy oil palms.

The steps of the learning model in this study involved all bands of the QuickBird imagery: NIR and visible bands blue, green, and red. The importance of variables in building the RF classifier model was examined. The red band of QuickBird imagery was the most important variable for healthy classification and the NIR band the most important for unhealthy classification (Table 3.3), which was the same as found in Santoso et al. (2011). The most important variable affecting accuracy, based on mean decrease in accuracy, was the red band. The mean decrease in Gini coefficient (Table 3.3) showed that there were high to low contributions to the homogeneity of nodes and leaves in the resulting RF classifier model according to the following order of variables: red, NIR, green, and blue QuickBird imagery bands (Liaw and Wiener 2002). In the final RF classifier model, there were 500 trees; three variables were attempted at each split; the out-of-bag error rate was 33%.

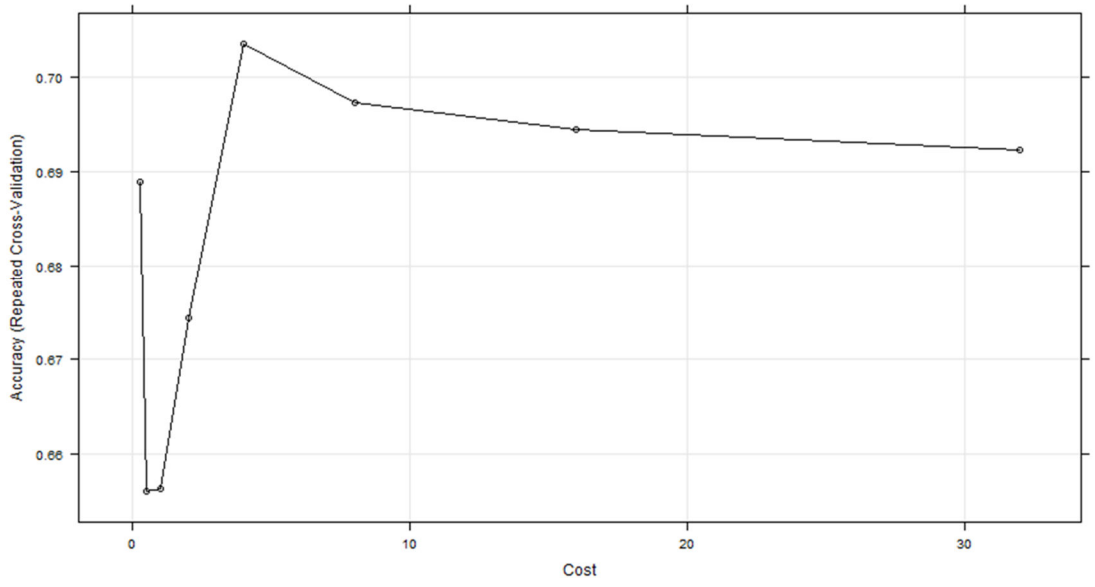


Figure 3.4. Tuning parameter results for the SVM classifier model.

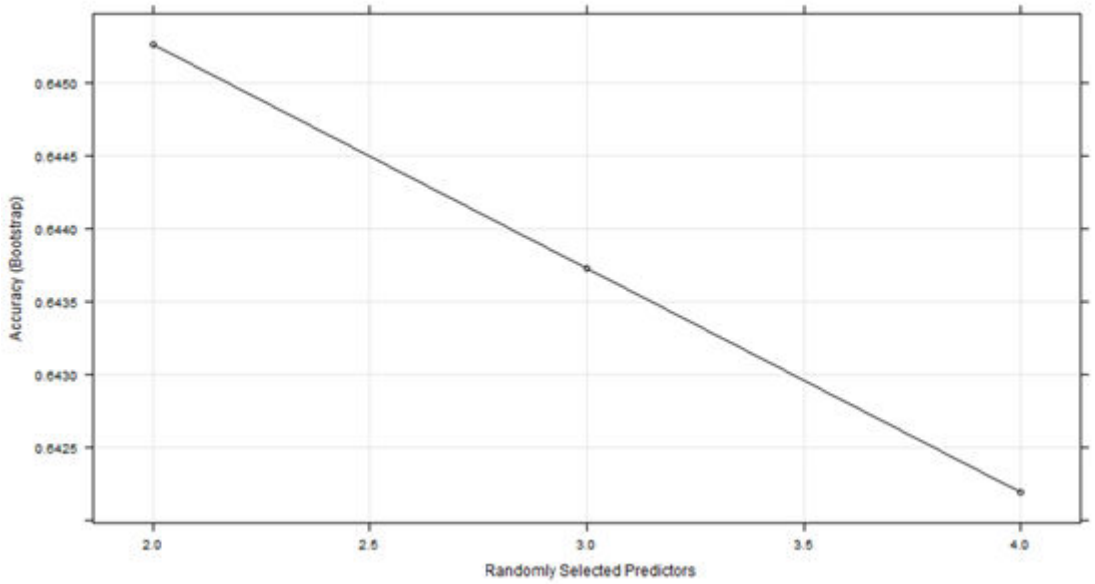


Figure 3.5. Tuning parameter results for the RF classifier model.

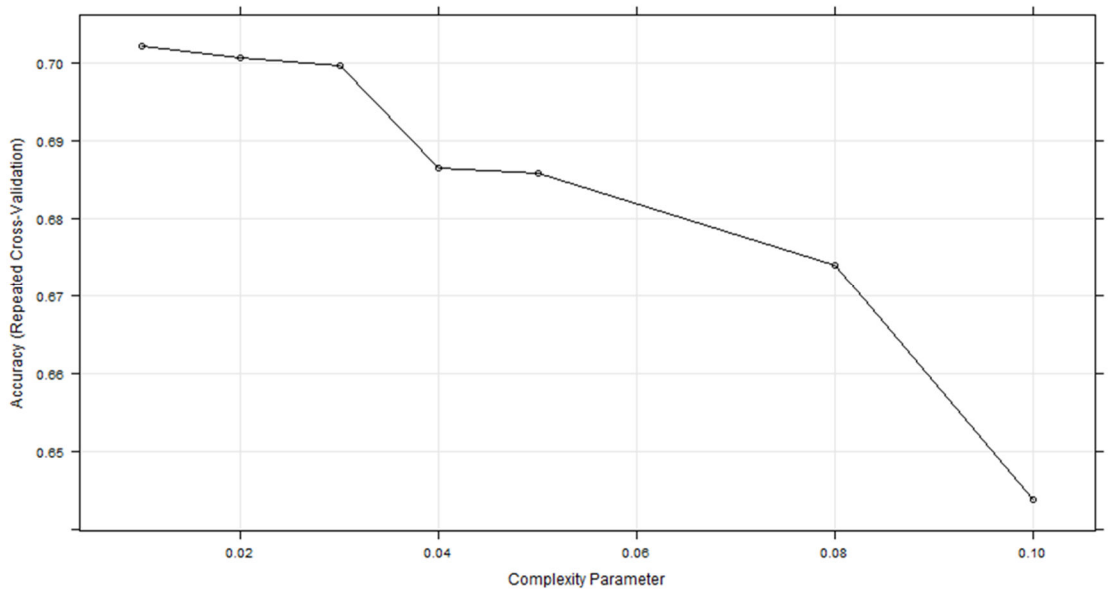


Figure 3.6. Tuning parameter results for the CART classifier model.

Table 3.2. Accuracy results for the SVM, RF, and CART classifier models.

	SVM			RF			CART		
Actual	Healthy	Unhealthy	All	Healthy	Unhealthy	All	Healthy	Unhealthy	All
Healthy	18	2	20	22	1	23	19	2	21
Unhealthy	6	9	15	2	10	12	5	9	14
All	24	11	35	24	11	35	24	11	35
Producer's accuracy									
(%)	75	82		92	91		79	812	
User's accuracy (%)	90	60		96	83		91	64	
Mapping Accuracy (%)	69	53		88	77		73	56	
Overall Accuracy (%)		77			91			80	
Observed Agreement		0.77			0.91			0.80	
Expected Agreement		0.53			0.56			0.54	
Kappa value		0.52			0.81			0.57	

The final output in this study was the distribution map of healthy and unhealthy oil palms produced using the best classifier model. In this step, the RF classifier model was used to produce the map. The model applied in R software using 49,793 palms from the segmentation process gave results of 37,518 palms classified as healthy and 12,275 palms classified as unhealthy. Furthermore, 37,617 (75%) healthy and 12,320 unhealthy oil palms (25%) were used in building the healthy and unhealthy tree map. The maps of distributions of healthy and unhealthy palms in study sites 1–4 are shown in Figures 3.7 and 3.8. There were fewer infected palms in study sites 1 and 3 than in sites 2 and 4 (Table 3.4). A pattern of disease occurred surrounding dead palms previously infected by BSR, consistent with the conclusion of Susanto, Prasetyo, and Wening (2013) that BSR could spread by direct root contact. Therefore, the RF classifier model has potential for detecting and mapping BSR disease caused by *G. orbiforme*. The blank areas in Figure 3.7 and 3.8 were open spaces in fields following death of oil palms infected by BSR. The greatest area infected by BSR disease was site 2 with 40% of oil palms infected (Table 3.4). All study sites had oil palm densities of less than 100 plants ha^{-1}, caused by deaths following disease (Santoso et al. 2011).

Machine learning algorithms used in this study

The learning algorithms applied in this study will contribute to the implementation of machine learning algorithms in oil palm management, especially for studying and monitoring the spread of disease. In this study, learning algorithms based on decision tree models (RF and CART) had better results than

that of the SVM model. In this case, we used 75% of the observation data for the learning model and 25% of the observation data for the testing model. Beleites et al. (2013) stated that it is not sufficient to build a good model but the performance must also be proved. Therefore, in this study we used accuracy and kappa values as performance indicators.

Table 3.3. Importance of variables in the RF classifier model.

QuickBird imagery band	Healthy	Unhealthy	Mean decrease accuracy	Mean decrease Gini
Blue (b1)	1.048	1.631	1.938	9.897
Green (b2)	5.179	1.576	5.327	10.730
Red (b3)	13.249	4.718	14.157	13.160
NIR (b4)	10.354	8.720	12.690	12.520

Table 3.4. Healthy oil palms and BSR-diseased oil palms in each study site according to the RF classifier model.

Study site	Area (ha)	Healthy	Percentage (%)	Unhealthy	Percentage (%)	Total
1	150.16	12,018	85	2089	15	14,107
2	139.31	6380	60	4323	40	10,703
3	130.49	11,200	86	1800	14	13,000
4	128.99	8019	66	4108	34	12,127
Total	548.95	37,617	75	12,320	25	49,937

We ran the model using three different ratios of training and testing data and found that a ratio of 75:25 training to testing data was best in terms of its accuracy and kappa value (Table 3.5). Table 3.5 shows that the kappa values of models with ratios of 25:75 and 50:50 (as used in Viera and Garrett 2005) have fair agreement (kappa value = 0–0.32); the exception was in the RF model, which had a kappa value showing moderate agreement (0.45) and 77% accuracy. Another way to test the

learning performance in learning algorithms using the number of observations in a training set is with the learning curve (Mukherjee et al. 2003; Figueroa et al. 2012; Beleites et al. 2013; Bischl et al. 2016). Figure 3.9 shows that the learning curve in this study, obtained with the machine learning in R (mlr) package (Bischl et al. 2016), is in accordance with the results. Increasing the number of observations in the training set improved the accuracy in the RF model as well as in the CART and SVM models. However, the maximum accuracy peaked at approximately 0.82 in the CART model and at approximately 0.80 in the SVM. Therefore, RF was the best-fitting learning algorithm.

RFs are learning algorithms based on learning ensembles of decision trees that consist of a large number of trees; for each decision tree an individual bootstrap sample is drawn from the original data set by sampling with replacement (Gall, Razavi, and Van Gool 2012; Immitzer, Atzberger, and Koukal 2012). We surmise that this is one of the reasons that the RF was the best model here. In the current study, there were imbalanced data in each class. According Khalilia, Chakraborty, and Popescu (2011), the ensemble learning in RF is suitable for imbalanced data. QuickBird imagery bands 2, 3, and 4 contributed highly to building the decision trees and voting the correct classes in RF. In our previous study (Santoso et al. 2011), we found that band 4 has a tendency to characterize oil palms as healthy and bands 2 and 3 have a tendency to characterize oil palms as unhealthy. Furthermore, a method was proposed in this study to detect, segment, and extract pixel values of

oil palm objects from QuickBird imagery combined with machine learning algorithms that can be applied in precision agriculture in oil palm plantations using high spatial resolution imagery.

Table 3.5. Relationship between the ratio of training and testing data and learning performance.

Ratio training and test data	RF Performance		SVM Performance		CART Performance	
	Accuracy (%)	Kappa value	Accuracy (%)	Kappa value	Accuracy (%)	Kappa value
25:75	77	0.45	69	0.00	62	0.26
50:50	66	0.19	72	0.25	70	0.32
75:25	91	0.81	77	0.52	80	0.57

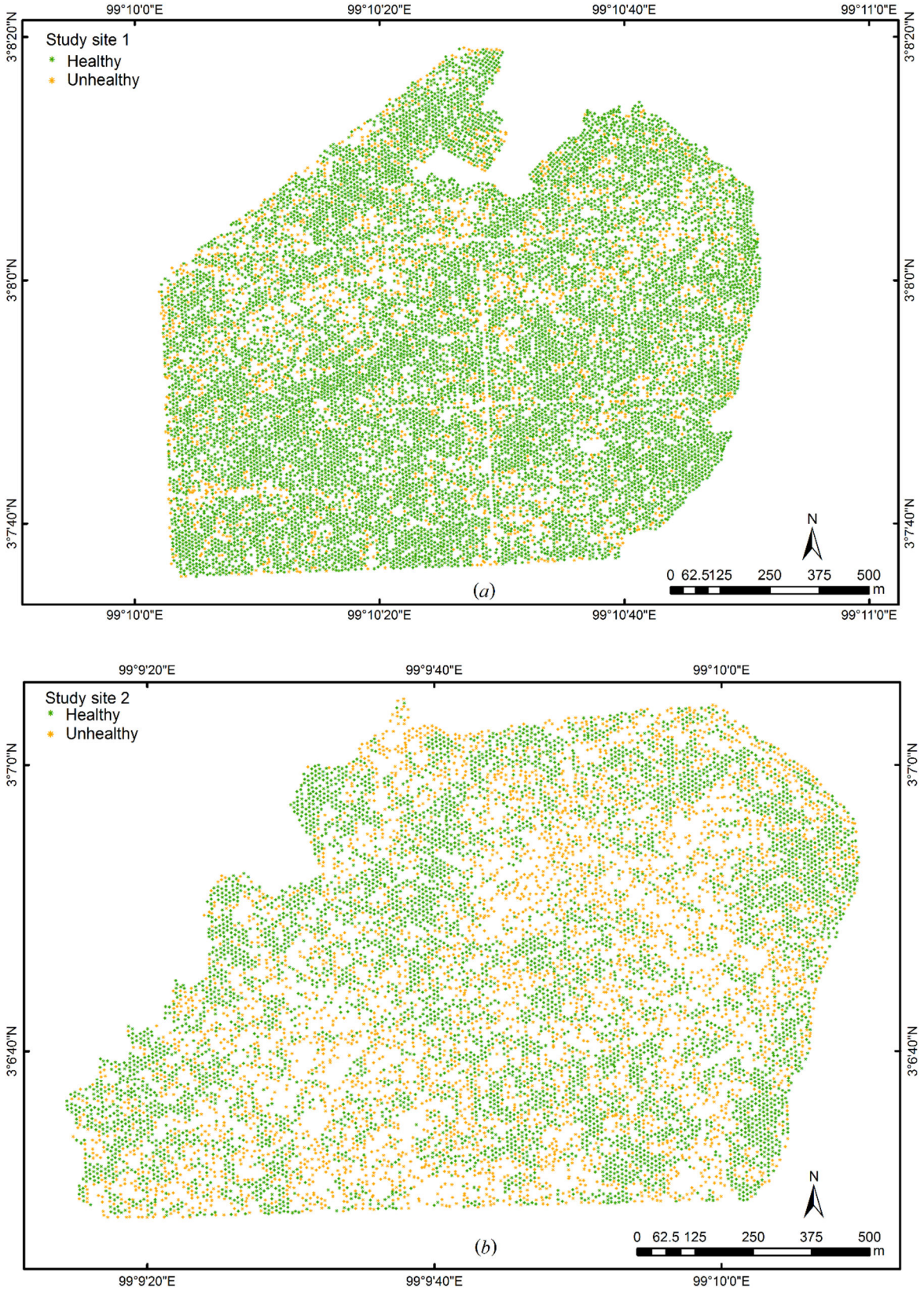


Figure 3.7. Sample map of the distribution of healthy and unhealthy oil palms in sites 1 (a) and 2 (b).

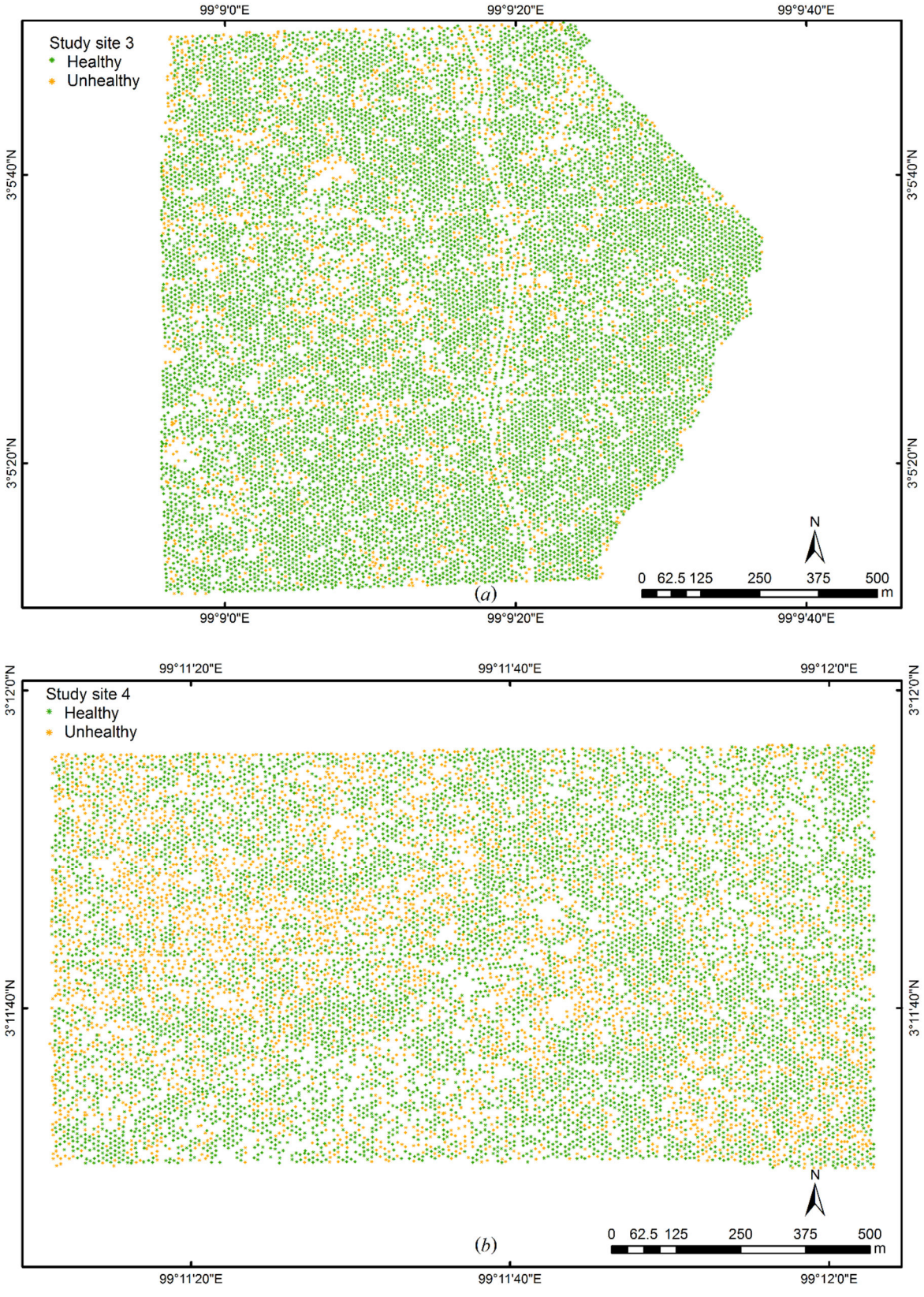


Figure 3.8. Sample map of the distribution of healthy and unhealthy oil palms in sites 3 (a) and 4 (b).

Implementation of the findings in oil palm management

The distribution of BSR-infected oil palms is important information for oil palm management in the field and will be the basis for treatment to prolong the life of infected oil palms and avoid infection of healthy palms. Surgery and soil mounding treatment can prolong the life of infected oil palms by several years. The isolation trench treatment is suitable for preventing BSR infection of healthy oil palms (Priwiratama, Prasetyo, and Susanto 2014). Knowing the number of infected oil palms is one consideration for calculating the potential decrease in production and estimating suitable times for replanting and any treatments to avoid disease infection. Treatments to prevent BSR in oil palms at replanting include sanitation and removal of diseased material (all parts of oil palms including trunk, leaves, and roots), ploughing, and harrowing. In addition, fallowing, using the "big-hole" planting technique, applying bioagents, planting a legume cover crop, and adding organic matter as compost derived from oil palm empty fruit bunches can be useful (Hushiarian, Yusof, and Dutse 2013; Priwiratama, Prasetyo, and Susanto 2014).

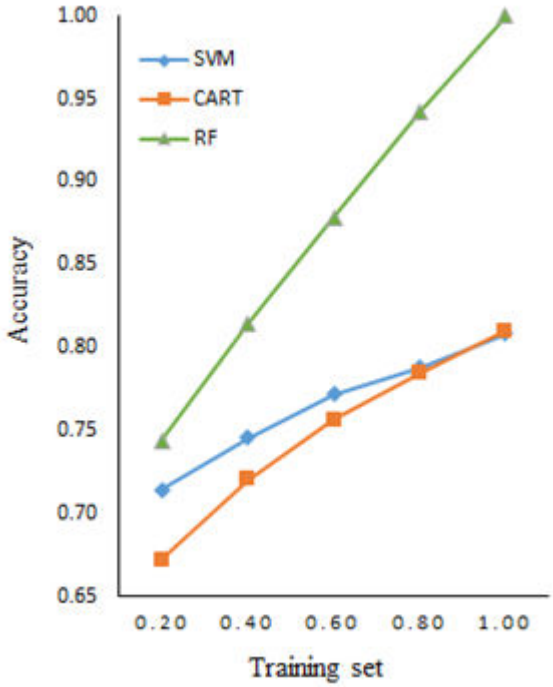


Figure 3.9. Learning curves showing the relationship between performance and the training set.

3.4. Conclusions

QuickBird imagery archived on 4 August 2008 was used to generate a machine-learning model for prediction of BSR disease, and three classifier models were applied: SVM, RF, and CART. The RF model was the best for prediction, classification, and mapping of oil palm disease caused by *G. boninense* in terms of OA, PA, UA, and kappa value. Using 75% of the data for training and 25% for testing, the RF model achieved 91% OA. Applying this model separated the healthy and unhealthy oil palms in the study area into 37,617 (75%) and 12,320 (25%) trees, respectively.

The RF classifier model has potential for predicting and classifying oil palm disease caused by *G. boninense*. Compared with the results of our previous study

(Santoso et al. 2011), which used the same data for identifying and mapping oil palms infected by *G. boninense* with several vegetation indices, the results of the current study had greater mapping accuracy for both healthy and unhealthy oil palms (88% and 77%, respectively). The early detection of this disease is very important for prolonging the life of oil palms. Therefore, in future studies, the potential of the RF classifier model for early detection of BSR and distinguishing the different levels of disease severity in oil palms will be tested using unmanned aerial vehicles with multispectral camera.

Chapter 4. Classifying Severity of Basal Stem Rot Disease in Oil Palm Plantation Using WorldView-3 Imagery and Machine Learning Algorithms

4.1. Introduction

Basal stem rot (BSR) disease, a major disease in oil palm plantations, is caused by *Ganoderma orbiforme* (= *G. boninense*). *G. orbiforme* infection causes losses in oil palm production because there is no effective fungicide to control this disease. BSR disease is found not only in old plantations (2nd or 3rd generations) and old oil palms but also in new oil palm plantations and young oil palms, and *G. orbiforme* has been detected in both mineral and peatland. As a result of this disease, the oil palm density in a plantation can decrease to less than half the original population (Paterson 2007; Idris et al. 2000; Hushiarian, Yusof, and Dutse 2013; Susanto et al. 2013; Priwiratama, Prasetyo, and Susanto 2014). Early detection of BSR infection is a key to controlling this disease, although all current treatments can only prolong the life of an oil palm tree (Hushiarian, Yusof, and Dutse 2013; Priwiratama, Prasetyo, and Susanto 2014). In oil palm practices, each oil palm tree is monitored manually for symptoms every six months, especially in areas known to be infected by BSR. If the infected area is large the manual monitoring is not effective. Hence, techniques that can easily, quickly, and with high accuracy monitor the dispersion of BSR disease severity are urgently required. One such potential technique is remote sensing (Liaghat et al. 2014; Helmi Z.M. Shafri et al. 2011; Santoso et al. 2011; H.Z.M. Shafri and Hamdan 2009; Lelong et al. 2010).

Several researchers have applied remote sensing to determine BSR disease severity defined under various classes. By defining two classes (healthy and unhealthy), Shafri and Hamdan (2009) achieved 73–84% accuracy with airborne hyperspectral data, Santoso et al. (2011) achieved 69.0–84.91% accuracy with Quickbird imagery, and Santoso, Tani, and Wang (2017) achieved 77-91% accuracy with Quickbird imagery. Subsequently, using three classes (healthy, unhealthy with mild symptoms, and unhealthy with severe symptoms), Helmi Z.M. Shafri et al. (2011) achieved 82% accuracy with spectroradiometer data on oil palm seedlings. Using four classes (healthy, unhealthy with mild symptoms, unhealthy with moderate symptoms, and unhealthy with severe symptoms), Lelong et al. (2010) achieved 94% accuracy with spectroradiometer data on oil palm canopy in the field and Liaghat et al. (2014) achieved 97% accuracy with spectrometer data on leaf surface of leaf number 17. Leaf number 17 was identified by defining the youngest fully open leaf as leaf number 1, and counting backwards down the trunk. Therefore, there is a need to further study the classification of BSR disease using multispectral data and four classes.

Supervised learning algorithms are commonly used for classification using remote sensing data. Supervised learning algorithms need learner classifiers to learn a function that can map output variables based on input variables (Ayodele 2010; Brownlee 2016). Wu et al. (2008) noted that, in December 2006, the top 10 learners identified by the IEEE International Conference on Data Mining were C4.5, k-Means, support vector machine (SVM), Apriori algorithm, Expectation-

Maximization, PageRank, AdaBoost, k-nearest neighbor (kNN), Naïve Bayes, and Classification and Regression Trees (CART). Rumpf et al. (2010) applied a SVM using hyperspectral reflectance to discriminate diseased from non-diseased sugar beet leaves with 97% accuracy, and differentiated the type and stage of Cercospora leaf spot disease based on leaf rust and powdery mildew symptoms with 65–90% accuracy. A non-linear SVM using remote sensing data (high-resolution hyperspectral imagery) was applied successfully to distinguish healthy trees from trees with incipient infection of almond red leaf blotch (López-López et al. 2016). Pujari, Yakkundimath, and Byadgi (2015) suggested that SVM classifiers were more suitable than other classifiers for classifying fungal disease symptoms that affect the leaves of cereal crops. Römer et al. (2011) achieved 93% accuracy using a SVM for the detection of pre-symptomatic wheat leaf rust.

Decision tree (DT) algorithms are often used in classification studies and as comparison learners (Rumpf et al. 2010; Römer et al. 2011). El-telbany and Warda (2016) achieved 97.57% accuracy using a DT to classify Egyptian rice disease. The 90% of accuracies were achieved by a DT that was used to build an automated knowledge base for a rice plant agriculture expert system that included pest and disease information (Joy and K 2015). While the usage for coffee plant disease detection was achieved 85% accuracy (Suhartono et al. 2013). The random forest (RF) algorithm is also often used in classification studies. This algorithm was based on a simple structure of forest of trees where the training set is a bootstrap sample from the original training set and an integer number of variables randomly sampled

as candidates at each split (*mtry*) is tunable (Breiman 2001). At each node, *mtry* variables are selected randomly and the node is split on the best split based on the selected *mtry* (Breiman 2004). Besides the DT, El-telbany and Warda (2016) also used RF to classify Egyptian rice disease with high accuracy (95.63%). Chaudhary, Kolhe, and Kamal (2016) proposed an improved RF classifier for multi-class classification of groundnut disease that increased the accuracy from 80% (common RF algorithm) to 98%. An RF algorithm achieved 93% accuracy in predicting the risk of Stagonospora nodorum blotch in winter wheat (Mehra et al. 2016).

Supervised learning algorithms can be used to identify and classify BSR disease in oil palm trees. Santoso, Tani, and Wang (2017) improved the classification accuracy of healthy and unhealthy oil palms (two classes) using SVM, CART, and RRF compared with their previous results (Santoso et al. 2011), using the same data. Santoso, Tani, and Wang (2017) also applied the algorithms to all oil palm trees in a study area that was identified by segmentation processes from Santoso, Tani, and Wang (2016), to produce a map of the outspread of healthy and infected oil palm trees with QuickBird imagery. Therefore, in this study, we used SVM, DT, and RF learners to classify BSR disease severity using WorldView-3 imagery.

WorldView-3 is a high-resolution commercial satellite that was launched on 13 August 2014. It has 1.24 m multispectral resolution and 31 cm panchromatic resolution, 8 multispectral bands, 8 SWIR bands, 12 CAVIS (clouds, aerosols, vapors, ice, and snow) bands, and a revisit frequency of less than 1 day if 1 m GSD

(ground sample distance) and 4 to 5 days at 20° off-nadir or less (DigitalGlobe 2014). Multispectral imagery has not been used previously to identify and classify four classes of BSR disease. We took into consideration the availability of imagery, price, spatial resolution, and technical image processing to select the multispectral satellite imagery that was most suitable for identification and classification into four disease severity. Therefore, the objectives of this study were to predict disease severity with WorldView-3 imagery using supervised learning algorithms, and to describe the characteristic symptoms of BSR disease in oil palm at different disease severity that could be identified by WorldView-3 imagery.

4.2. Data and Methods

4.2.1. Study area

The study area was an oil palm estate in North Sumatra, Indonesia, located within 3°9′50.34" – 3°10′11.66" N and 99°27′39.728" – 99°28′8.674" E (Figure 4.1). The oil palm estate was a planted block with flat topography and 32 ha of coverage area. According to the estate records, the oil palm trees, which were planted from 2004 till early August 2016, were infected by BSR disease with 32% of oil palm population death caused by this disease. In the study area, rainfall is evenly distribution throughout the years (Figure 4.2), which makes it very suitable for oil palm growth and production.

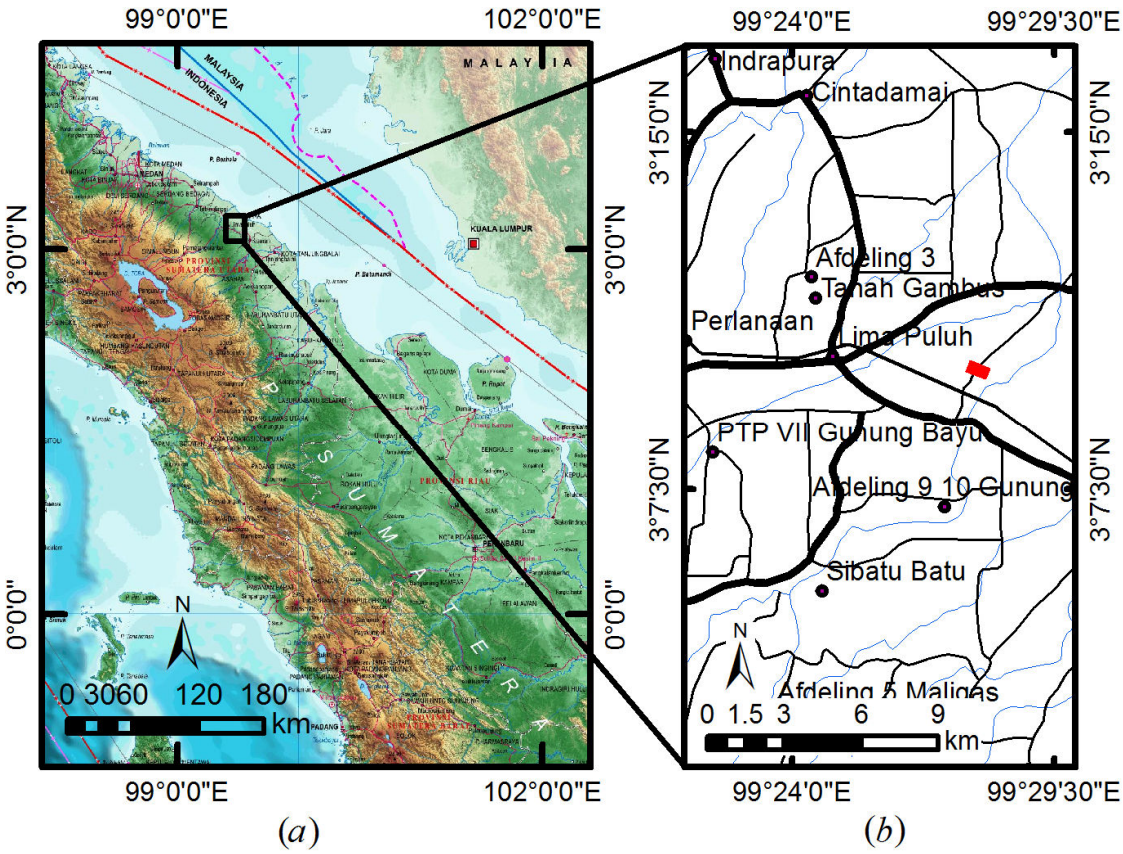


Figure 4.1. Vicinity of the study area on the island of Sumatra, Indonesia (a); Zoom in of the study area (red rectangle) (b) (Source map: Indonesian Geospatial Board – http:// www.bakosurtanal.go.id/download/).

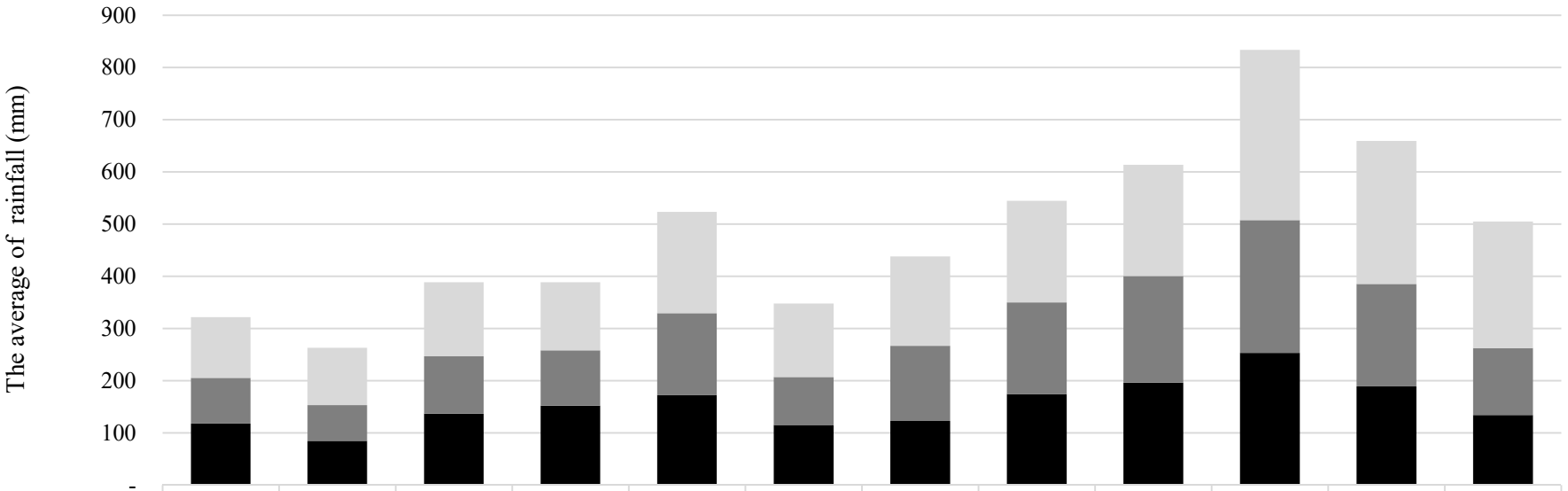
4.2.2. Image Acquisition and Preprocessing

In this study, we used WorldView-3 imagery archived on 6 August 2016 with the following bands: coastal (400–450 nm), blue (450–510 nm), green (510–580 nm), yellow (585–625 nm), red (630–690 nm), red edge (705–745 nm), near infrared 1 (NIR1: 770–895 nm), near infrared 2 (NIR2: 860–1040 nm), and panchromatic (450–800 nm). We applied the geometric and radiometric correction to the WorldView-3 imagery. ENVI 5.2 (ENVI 2014) was used to preprocess the WorldView-3 imagery by applying the geometric and radiometric correction. The geometric correction was applied by rational polynomial coefficient (RPC) methods, while the radiometric

correction was applied according to a technical note of WorldView-3 (Kuester 2016). The reflectance data from the eight WorldView-3 bands were used as variables in this study. The WorldView-3 imagery of the study area is shown in Figure 4.3.

4.2.3. Field Data Collection

Field data collection was done on 3 August 2016. We randomly selected oil palm trees and observed the canopy and oil palm trunk for BSR infection symptoms using the estate criteria as follows: (1) healthy (H) oil palm tree with no BSR infection symptoms (asymptomatic); (2) unhealthy oil palm tree initially symptomatic (UH1) having one of the following BSR disease symptoms, three unopened emerging young leaves (spears), yellowing leaves along with a wide range of necrosis, older leaves fractured, emerging fruiting bodies (mushroom), rotting in basal stem, and the oil palm still produces fruits; (3) unhealthy oil palm tree moderately symptomatic (UH2) having two of the listed BSR disease symptoms, and the oil palm still produce fruits; and (4) unhealthy oil palm tree severely symptomatic (UH3) having more than one of the BSR disease symptoms, and the oil palm tree does not produce fruits. We collected observation data for 1923 oil palm trees comprising 695 H (36%), 432 UH1 (23%), 348 UH2 (18%), and 448 UH3 trees (23%).



	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Station 3	116	110	142	130	194	141	171	195	214	326	274	242
Station 2	88	68	110	106	157	92	144	175	204	254	196	128
Station 1	118	85	137	152	173	115	123	175	196	253	189	134

Figure 4.2. Average of rainfall from 2004 to 2015 in study site. The data were obtained from the three stations around the study site.

4.2.4. Data Analysis

The corrected WolrdView-3 imagery was used to extract the mean of the nine pixel values with points from field observation. The 1923 observation data of oil palm status (H, UH1, UH2, and UH3) were converted to square polygons (4 × 4 m^2) in R software (RStudio 2015). Thereafter, the mean of the nine pixel-values of each point observation was extracted by the square polygon, with the result that every point had eight data of WorldView-3. The flow of the data analysis is shown in Figure 4.4.

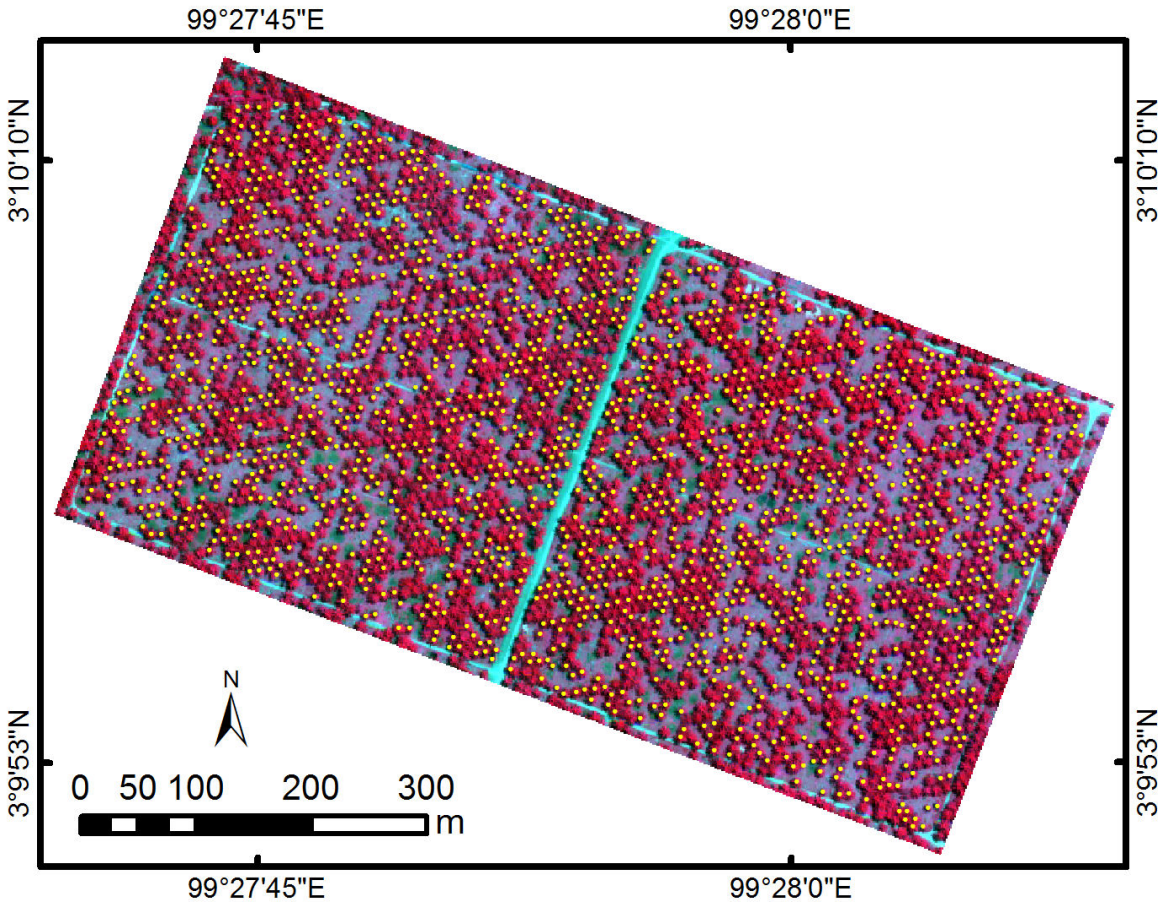


Figure 4.3. WorldView-3 imagery of the study area with RGB 753. Red indicates the oil palm trees and yellow indicates the reference data.

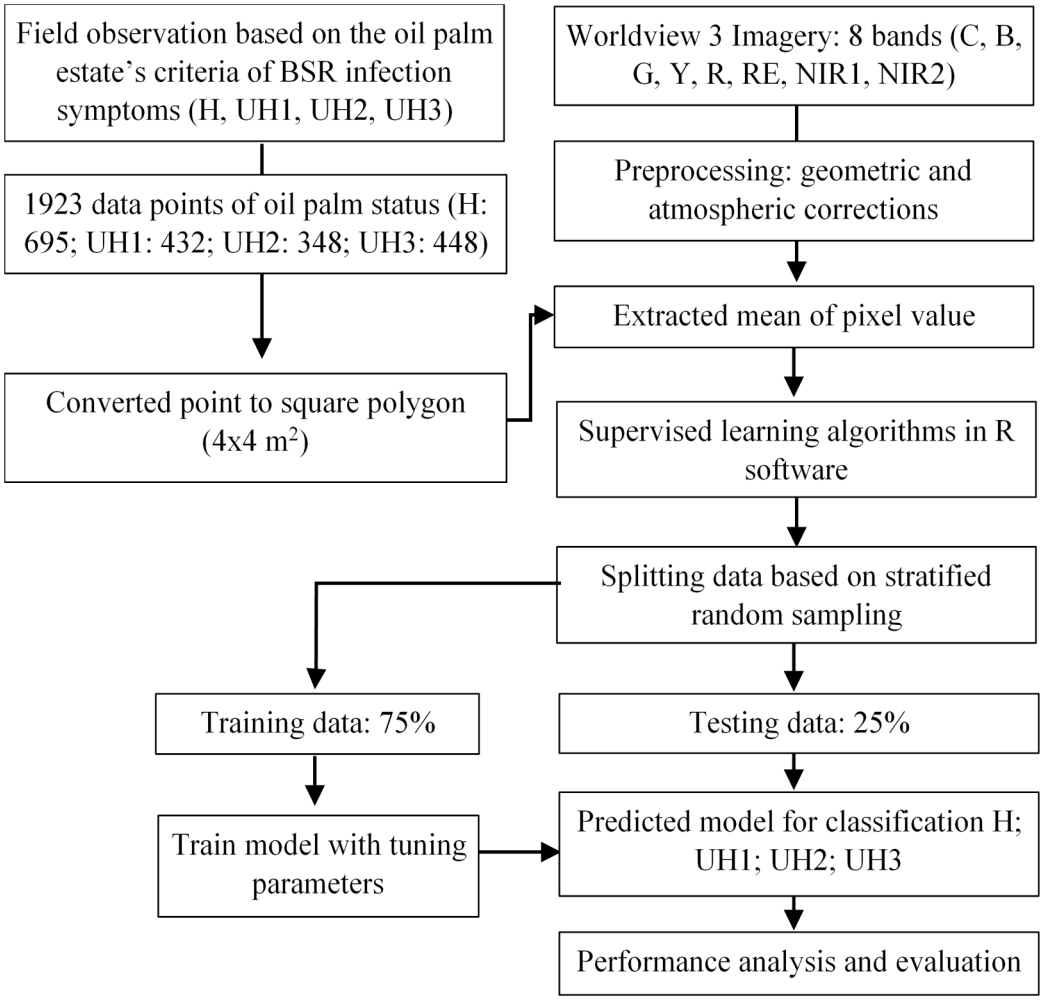


Figure 4.4. Data analysis flow chart.

4.2.5. Tuned Parameters

The supervised learning algorithms, SVM, DT, and RF, were applied in R software. The learning algorithms were used to evaluate the accuracy of the estate criteria that we used to classify oil palm status in the field observations for the identification and classification of BSR infection with the remote sensing technique. For this purpose, the observation data were split into 75% training data and 25% testing data (Santoso, Tani, and Wang 2017) with stratified random sampling. The numbers of samples in the training data were 522 of H, 324 UHI, 261 UH2, and 336 UH3,

whereas, in the testing data, there were 173 of H, 108 UH1, 87 UH2, and 112 UH3. The data were partitioned using the '*caret*' package (Kuhn 2016) in R software. The '*mlr*' package in R was used for normalization, resampling, and training and prediction (Bischl et al. 2016). All variables were normalized with a standardizing method (center and scale). The resampling technique was applied with repeat cross-validation and five times repetition.

Tuning parameters in association with training data were used to process the models. In the SVM classifier, we tuned the cost (c) in range 10-1000 and sigma in range 10^{-6}-100 and applied the grid search and 'kernlab' package (Karatzoglou, Smola, and Hornik 2015). The DT algorithm was tuned using the following parameters: minimum number of observations in a node (minsplit) in range 10-20; minimal number observations in a terminal node (minbucket) in range 5-30; and threshold complexity parameter (cp) in range 0.001-0.2 with the 'rpart' package (Therneau and Atkinson 2015).The RF algorithm was applied with tuned parameters: as follows: *mtry* in the range 3-10; number of trees (*ntree*) in the range 50-500; and number of observations in terminal nodes (nodesize) in the range 10-50 with the 'randomForest' package (Breiman et al. 2015; Breiman 2001). The tuned parameters were used in the prediction model to classify H, UH1, UH2, and UH3 in the testing data set.

4.2.6. Performance Analysis and Evaluation

Performance analysis was conducted using the confusion matrix and quantity disagreement (QD) and allocation disagreement (AD) (Pontius and Millones 2011)

and kappa value (J. Cohen 1960; Short 1982). If the accuracy was high, the criteria used in the field data collection proved to identify BSR disease with WorldView-3 imagery. If the performance was low, further analysis of the learning model and reevaluation of BSR symptoms were performed. The 'stepclass' algorithm (Price and Fischer 2014; Weihs et al. 2005; Szepannek and Weihs 2006; Dechaume-Moncharmont, Monceau, and Frank 2011; Roever et al. 2015) was used to obtain appropriate variables that could distinguish the four classes of oil palm status. The 'stepclass' is stepwise variable selection for classification.

The 'stepclass' algorithm in the 'klaR' package (Roever et al. 2015) with parameters set uses linear discriminant analysis in the backward and forward (both) directions. Based on the selected variables from the 'stepclass' algorithm, we used a box plot of the distribution of reflectance values of the selected variable to remove outliers. Outliers can affect the accuracy of the learning model according to Hosseinioun (2016); W. Li et al. (2015); Cao, Stojkovic, and Obradovic (2016); and Osborne and Overbay (2004). Then, the learning algorithms were rerun and the performance analysis was performed using the data with outliers removed. Reevaluation of BSR symptoms was performed based on a box plot of the distribution of reflectance values of the selected variable. The data were selected for H in the minimum range of 25th quantile values, for UH1 in the median range of 25th quantile values, for UH2 in the median range of 75th quantile values, and for UH3 in the maximum range of 75th quantile values. Further field observations were performed to redefine the BSR symptoms with selected data by observing the

canopy and oil palm trunk. Furthermore, the new BSR symptoms were described based on the symptoms in each class that were recorded in the further field observation step.

4.3. Results

4.3.1. Performance of Learning Algorithms

The distributions of reflectance values of WorldView-3 imagery for H, UH1, UH2, and UH3 are shown in Figure 4.5. The characteristics of the reflectance values in all the bands include the minimum, 25th quantile, median, 75th quantile, and maximum values of the oil palm status (H, UH1, UH2, and UH3). The reflectance values overlapped each other or did not separate clearly. The best tuned parameters are shown in Table 4.1. The supervised learning algorithms were applied with the best tuned parameters and the results are shown in Table 4.2. All the algorithms with all the percentage samples had low overall accuracy (OA) 47.5-51.7 % with fair agreement of kappa value (<0.40) and high of QD (14-22%) and AD (26-36%). The SVM has higher overall accuracy than RF and DT. Further analysis was performed based on these results. The 'stepclass' algorithm was applied using the observation data for the 1923 oil palms with eight bands of reflectance values of Worldview-3 imagery. The results showed only band 4 (yellow band) has the ability to separate four classes of BSR severity. The value of the ability to separate was 0.24 calculated by 'stepclass' analysis, which means the low ability. If the ability to separate is close to 1, the classification is very well done and if the value is close to 0, the classification is not good (Roever et al. 2015; Garczarek 2002).

Box plots of the distribution of reflectance values of band 4 before and after outliers were removed are shown in Figure 4.6. A total of 31 outliers were removed from the data and the remaining 1892 oil palm trees, comprising 682 H, 427 UH1, 342 UH2, and 441 UH3 trees, were used for the second learning algorithms, as described for the first learning algorithms. The best-tuned parameters are shown in Table 4.3 and the results are shown in Table 4.4. The accuracies from the three learners were improved by 1–5% (53.1–54.1%).

Table 4.1. The best tuned parameters with original observation data.

Model	Parameter	Tuned value
DT	minsplit	12
	minbucket	22
	cp	0.001
RF	*ntree*	252
	mtry	3
	nodesize	23
SVM	c	1000
	sigma	0.0001

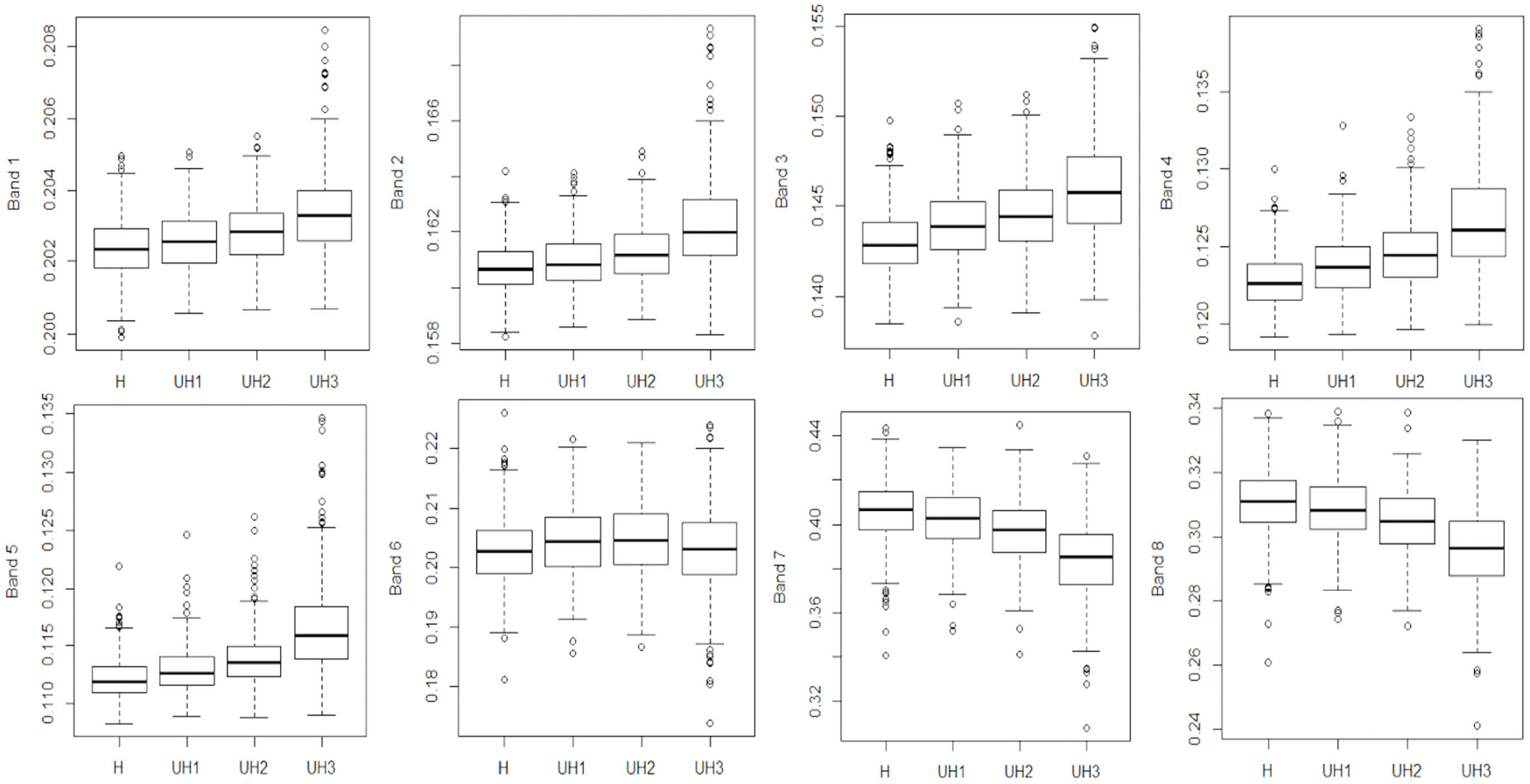


Figure 4.5. Box plots of the distribution of reflectance values of WorldView-3 imagery for H, UH1, UH2, and UH3. The line across the middle of each box is the median value, and the ends of the boxes are 25th and 75th quantiles.

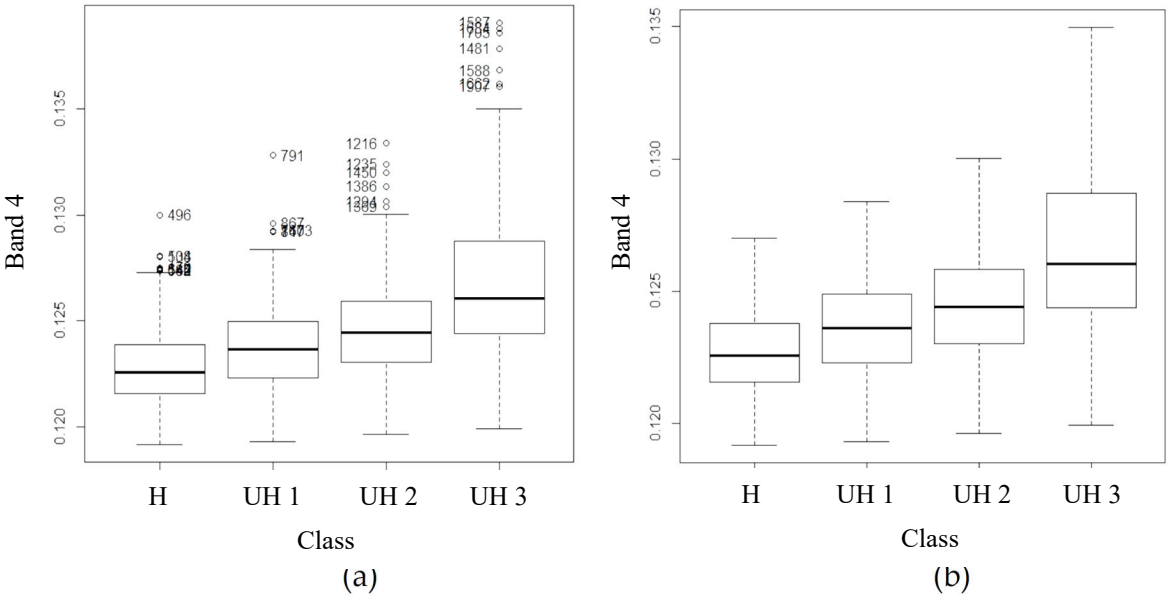


Figure 4.6. Box plots of the distribution of reflectance values of band 4 for H, UH1, UH2, and UH3. The line across the middle of each box is the median value, and the ends of the boxes are 25th and 75th quantiles: (a) before the outliers were removed; (b) after the outliers were removed.

Table 4.2. Results of the learning algorithms with tuned parameters and original observation data.

Model	OA (%)	kappa	QD (%)	AD (%)	Mapping accuracy (%) [a]				Producer's Accuracy [b]				User's Accuracy [c]			
					H	UH1	UH2	UH3	H	UH1	UH2	UH3	H	UH1	UH2	UH3
DT	47.5	0.257	16	36	42.91	12.00	13.16	44.22	0.73	0.19	0.17	0.58	0.51	0.24	0.36	0.65
RF	51.3	0.314	14	34	47.33	11.33	18.46	47.06	0.77	0.16	0.28	0.64	0.55	0.29	0.36	0.64
SVM	51.7	0.310	22	26	48.32	8.90	14.68	47.17	0.83	0.12	0.18	0.67	0.54	0.25	0.42	0.61

OA=overall accuracy, QD=quantity disagreement, AD=allocation disagreement, DT=decision tree, RF=random forest, SVM=support vector machine, H=healthy, UH1=unhealthy 1, UH2=unhealthy 2, UH3=unhealthy 3.

[a] Mapping accuracy is stated as the number of correctly classified divided by the total number of correctly and incorrectly classified.

[b] Producer's accuracy is related to the probability that a reference sample will be correctly classified.

[c] User's accuracy indicates the probability that a sample matches the reference data.

Table 4.3. The best tuned parameters with the data after outliers were removed.

Model	Parameter	Tuned value
DT	minsplit	14
	minbucket	30
	cp	0.001
RF	*ntree*	246
	mtry	5
	nodesize	26
SVM	c	1000
	sigma	0.0001

The SVM still had the highest accuracy with 54.1% overall accuracy, although the kappa values were still in the range of fair agreement (<0.40). These results are in accordance with previous research that found outliers affected accuracy (Hosseinioun 2016; W. Li et al. 2015; Cao, Stojkovic, and Obradovic 2016; Osborne and Overbay 2004). The results from the two learning algorithms revealed that the BSR severity criteria used to collect the observation data were not appropriate for identification and classification of BSR severity using WorldView-3 imagery. Therefore, we reevaluated the BSR symptoms based on the 'stepclass' algorithm results.

Table 4.4. Results of the learning algorithms with tuned parameters and the data after outliers were removed.

Model	OA (%)	kappa	QD (%)	AD (%)	Mapping accuracy (%)[a]				Producer's Accuracy[b]				User's Accuracy[c]			
					H	UH1	UH2	UH3	H	UH1	UH2	UH3	H	UH1	UH2	UH3
DT	53.3	0.348	12	34	49.62	14.58	20.45	47.74	0.59	0.36	0.36	0.62	0.76	0.20	0.32	0.67
RF	53.1	0.344	11	35	49.24	13.75	20.61	51.82	0.58	0.29	0.37	0.72	0.76	0.21	0.32	0.65
SVM	54.1	0.348	20	26	51.97	11.19	14.16	51.32	0.57	0.30	0.36	0.65	0.85	0.15	0.19	0.71

OA=overall accuracy, QD=quantity disagreement, AD=allocation disagreement, DT=decision tree, RF=random forest, SVM=support vector machine, H=healthy, UH1=unhealthy 1, UH2=unhealthy 2, UH3=unhealthy 3.

[a] Mapping accuracy is stated as the number of correctly classified divided by the total number of correctly and incorrectly classified.

[b] Producer's accuracy is related to the probability that a reference sample will be correctly classified.

[c] User's accuracy indicates the probability that a sample matches the reference data.

*4.3.2. Reevaluation of BSR Disease Symptom*s

The second objective of this chapter is to describe the characteristic symptoms of BSR disease severity that could be identified by WorldView-3 imagery. To obtained this objective, we have selected the oil palm trees from all data in the field data collection (1923 samples) based on the distribution of reflectance value of band 4 from the result of the 'stepclass' algorithm. As shown in Figure 4.7, the distribution of band 4 reflectance data had a minimum value of 0.1192, 25th quantile of 0.1223, median of 0.1238, 75th quantile of 0.1256, and a maximum value of 0.1391. Furthermore, we selected data for H in the minimum range of 25th quantile values (0.1192–0.1223), for UH1 in median range of 25th quantile values (0.1223–0.1238), for UH2 in the median range of 75th quantile values (0.1238–0.1256), and for UH3 in the maximum range of 75th quantile values (0.1256–0.1391). Finally, 781 data samples remained; 287 of H (37%), 140 of UH1 (18%), 109 of UH2 (14%), and 245 of UH3 (31%).

The scatterplot of band 4 and 8 shown each class of the BSR severity was easy to distinguish using 781 data (Figure 4.8). Hereafter, the second field observation was performed to reevaluate the BSR disease symptoms using 781 data on 22–23 December 2016. Although there is the time lag around four months between first and second observation, we believe the characteristics of oil palm canopy and the BSR disease symptoms do not change so much. All the selected data (781 samples)

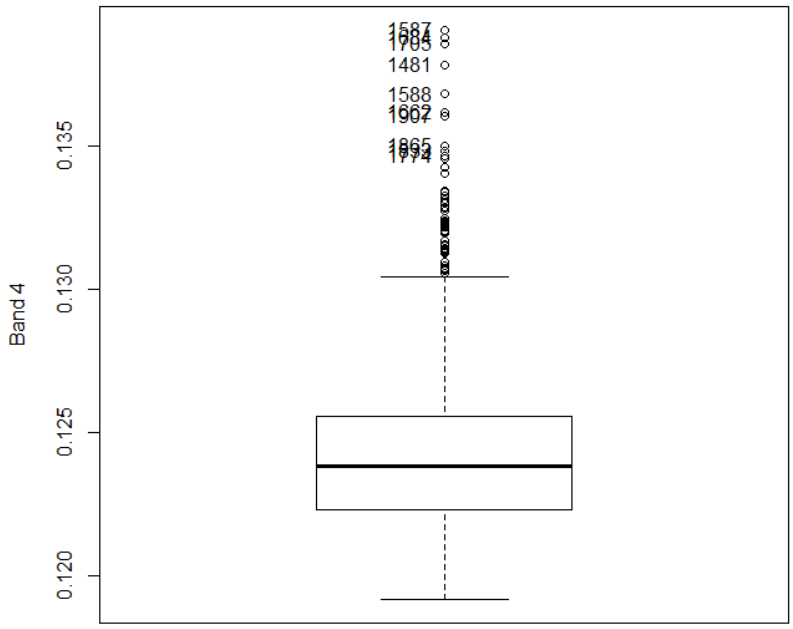


Figure 4.7. Box plot of the distribution of reflectance values of band 4. The line across the middle of each box is the median value, and the ends of the boxes are 25th and 75th quantiles.

were compared with the BSR disease symptoms observed on the canopy and trunks of the oil palm trees.

We described new BSR infection criteria based on the second field observation as follows: the new description of H oil palm trees as trees with perfect canopy condition. For UH1, the new description is trees with two unopened spears and emerging yellowish leaves and appearance of necrosis on older leaves or decreased leaflet size. For UH2, the description is trees with more than two unopened spears and yellowish leaves on middle to bottom part of the canopy with the appearance of necrosis or 1-2 older leaves fractured. For UH3, the description is trees with more than three unopened spears, almost all older leaves fractured,

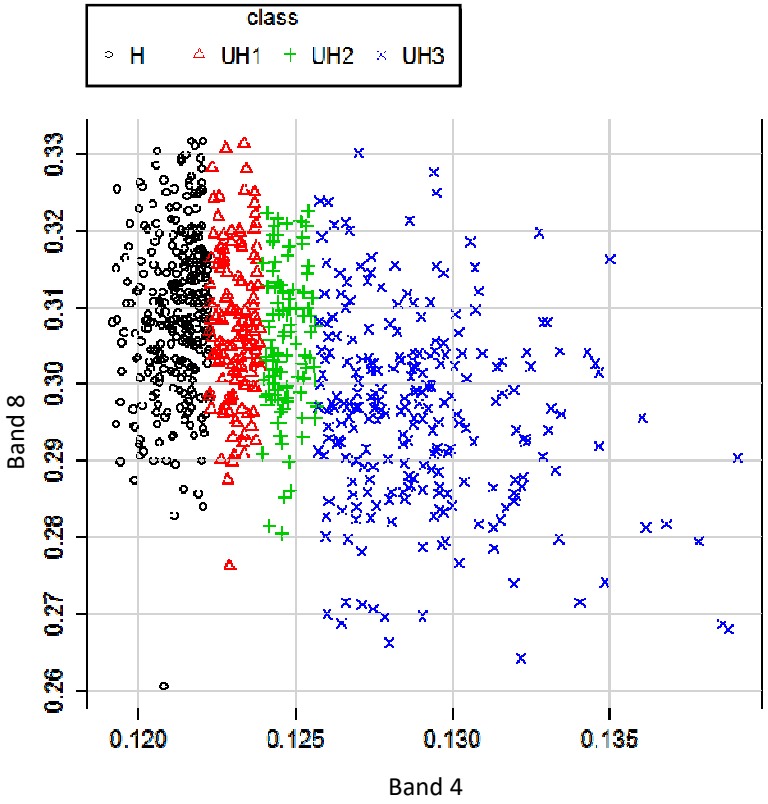


Figure 4.8. Scatter plot of the distribution of reflectance values of band 4 and 8 of WorldView-3 imagery for H, UH1, UH2, and UH3.

yellowish leaves along with wide range of necrosis, not producing new bunch/fruit, and having damaged parts in the basal stem (rotten and hole).

4.4. Discussion

WorldView-3 imagery has rarely been used to study vegetation, especially to investigate plant diseases. Some studies using WorldView-2, which has same bands used in this study, for vegetation and mapping the health status of trees have been reported (Immitzer, Atzberger, and Koukal 2012; Waser et al. 2014; Immitzer and Atzberger 2014; Fassnacht et al. 2017). In this study, we used the multispectral bands of WorldView-3 to classify multilevels of BSR disease severity in an oil palm

plantation. BSR disease can cause stem damage, which affects water uptake, and can also impact the tree canopy as unopened young leaves (Corley and Tinker 2003). The remote sensing technique has previously been used to monitor plant and biomass based on reflectance (Haboudane et al. 2002; Tan, Kanniah, and Cracknell 2013; Morel, Fisher, and Malhi 2012; Coops et al. 2006).

We found that the different canopy conditions caused by BSR disease infection affected different characteristics of the WorldView-3 spectra, as shown in Figure 4.5. In this study, we used homogeneous age, which meant, in general, the oil palm trees had the same crown size. The oil palm crown size has a correlation with the number of pixels used to calculate the mean value of the oil palm crown reflectance. Therefore, we used a square polygon inside the oil palm crown. Different square polygon sizes should be used for different oil palm crown sizes because the polygon size will influence the mean value of the oil palm crown reflectance and affect the classification results. Band 5 had the lowest reflectance values compared with the other bands because its wavelength range (630–690 nm) is used by chlorophyll for photosynthesis. Similarly, band 2 (450–510 nm) and band 4 (585–625 nm), which are within the visible spectrum, are used by chlorophyll for photosynthesis. Muharam et al. (2015) used the visible spectrum to calculate total chlorophyll content to estimate the nitrogen nutrient content in cotton.

Wavelengths in the range 550–750 nm have been found to be affected by chlorophyll content (Delalieux et al. 2008). In fresh tea leaves, chlorophyll a was found to absorb light in the range 400–500 nm while chlorophyll b absorbed light in

the range 660–690 nm (X.-L. Li and He 2008). Healthy oil palm trees are likely to have absorbed light in the spectrum ranges of bands 1 to 5 because of high photosynthesis, which would explain the low reflectance values for these bands. For UH1, UH2, and UH3, the photosynthetic rate would likely decrease in the increasingly unhealthy plants, which would explain the higher reflectance values for these bands. The reflectance values for bands 7 and 8 (infrared bands) ranged from high to low for the healthy to unhealthy oil palm trees. Healthy plants do not normally use wavelengths in this range for metabolism; indeed, a characteristic band in the near infrared is commonly used as an indicator of healthy plants in the remote sensing technique (Liaghat et al. 2014; Santoso et al. 2011; Helmi Zulhaidi M. Shafri, Hamdan, and Izzuddin Anuar 2012). For band 6 (red edge band), H and UH3 had almost the same reflectance values and so did UH1 and UH2.

The reflectance values (Figure 4.5) for each class of BSR disease severity symptoms were close to one another, which affected the accuracy of the classification with the supervised learning algorithms used in this study. The learning algorithms run with the original observation data achieved low accuracy and kappa value (overall accuracy <51.7 and kappa value <0.314). We identified problems with the BSR disease symptoms criteria that were used in field observation. The estate criteria allowed observers to easily classify the oil palm status in the field but potentially overlapped because the criteria had no clear borders between them.

The ‘stepclass’ algorithm results showed that band 4 as the variable that could separate the different classes in BSR disease severity. The band 4 wavelength is in the range 585–625 nm and so can be absorbed by chlorophyll for photosynthesis. BSR disease infection causes yellowish leaves, chlorosis, and decreased water uptake, which disturbs photosynthesis. Therefore, healthy oil palm to unhealthy with severe symptoms had reflectance values in band 4 that ranged from low to high.

A total of 32 outliers (13, 5, 6, and 7 from H, UH1, UH2, and UH3, respectively) were identified because the data points were inconsistent with the rest of the data in each class. Osborne and Overbay (2004) mentioned that outliers can have deleterious effects on statistical analysis. We think the outliers in this study arose from the first field observation because the BSR severity criteria did not have clear borders between classes. Although the learning algorithms using the data with outliers removed did not produce great results, the accuracy did improve by 1–5% and the number of variables tried at each split in RF increased from 3 to 5. We found that the deep overlap of the distribution of reflectance values in each class from the original data (Figure 4.5) was the main reason why the removal of outlier data did not greatly improve the accuracies of the learning algorithms.

We considered the performance achieved in this study as described by Beleites et al. (2013). Hence, we used the learning curve to test the learning performance in learning algorithms using the number of observations in a training set. The learning curves obtained with the ‘machine learning in R (mlr)’ package

(Bischl et al. 2016) are shown in Figure 4.9. Increasing the number of observations in the training set improved the accuracy in all learners, except SVM where the accuracy decreased when more than 80% of the observation data were used for the training set (Figure 4.9a). The SVM had higher accuracy than RF and CART in learning algorithms using all the original data and the data with outliers removed (Figure 4.9a, b). Therefore, we used 75% of the data for the training data and 25% for the testing data in the learning algorithms.

An important outcome in this study is the revised description of each class of BSR disease symptoms that can be used to identify and classify multilevels of BSR disease severity using remote sensing data. We compared the BSR disease symptoms described by Liaghat et al. (2014) for their study that used hyperspectral reflectance data with that obtained in this study, as shown in Table 4.5. The main differences are the number of unopened spears in UH1, and, from this study, a more detailed description regarding leaves fractured in UH2 and UH3, and additional symptoms that in UH3 new bunch/fruit are not produced and the basal stem had damage parts (rotten and hole). More than two unopened spears can also be caused by drought stress (Corley and Tinker 2003), but no drought conditions were recorded in the study site according to the average rainfall from 2004 to 2015 taken from the three stations around the study site (Figure 4.2).

We believe that our results will be useful for oil palm practices because the one key for controlling BSR disease is early detection (Hushiarian, Yusof, and Dutse 2013), although all treatments are generally applied for the purpose of prolonging

the lives of infected oil palm trees for several years (Priwiratama, Prasetyo, and Susanto 2014). BSR disease in oil palm is like a 'cancer', which is very difficult to detect at the early stages of infection (Naher et al. 2013), especially in oil palm with more than two generations. As a result, BSR disease is becoming a non-diagnostic disease. Oil palm trees that have a green canopy and several bunches of fruit can fall down because the basal stem has rotted inside. Rot inside the basal stem cannot be seen from outside and affected trees can potentially fall down in strong winds. To address some of the difficulties for early detection of BSR disease in oil palm plantation, our results can be applied for monitoring and classifying the oil palm status using WorldView-3 imagery and DT, RF, or SVM algorithms.

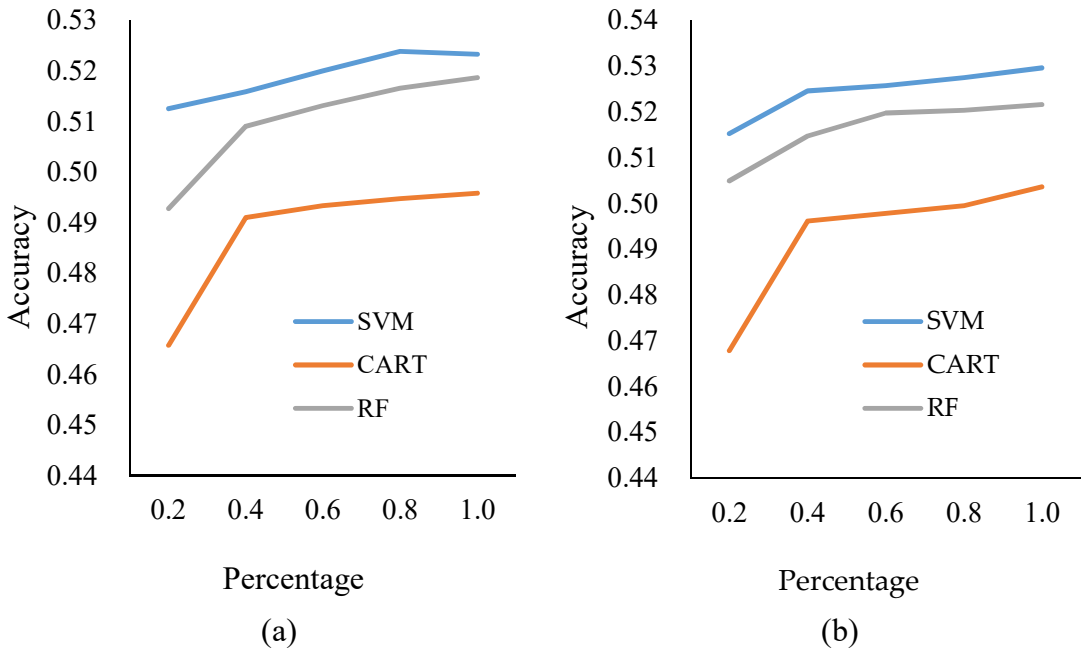


Figure 4.9. Learning curves showing the relationship between performance and size of the training set: (a) the original data; (b) the data with outliers removed.

Table 4.5. Comparison of basal stem rot disease infection symptoms.

Class	Liaghat et al. (2014)	This study
H	Healthy oil palm tree without any disease symptoms (asymptomatic)	Healthy oil palm with **perfect canopy condition**
UH1	Unhealthy oil palm initially symptomatic which have **one or two unopened spears**, yellowing leaves, small canopy, and without Ganoderma fruiting bodies (mushroom)	Unhealthy oil palm initially symptomatic which have **two unopened spears** and emerging yellowish leave and **appearance of necrosis on older leaves or the leaflet size decreased**
UH2	Unhealthy oil palm tree with moderately symptomatic which have **three to five unopened spears**, yellowing leaves along with the appearance of leaf necrosis, small canopy, and without Ganoderma fruiting bodies (mushroom)	Unhealthy oil palm tree with moderately symptomatic which **have more than 2 unopened spears** and yellowish leaves on middle to bottom part of canopy with the appearance of necrosis or **1-2 older leaves fractured**
UH3	Unhealthy oil palm tree with severely symptomatic which have **unopened spears**, yellowing leaves along with a wide range of necrosis, **small canopy**, and with **the existence of Ganoderma fruiting bodies (mushroom)**	Unhealthy oil palm tree with severely symptomatic which have **more than 3 unopened spears, almost older leave fractured**, yellowish leaves along with wide range of necrosis, **not produced new bunch/fruit, and have damage parts in basal stem (rotten and hole)**

The different criteria's showed in boldface

4.5. Conclusions

The eight bands of reflectance values from multispectral WorldView-3 imagery were examined for their ability to classify multilevels of BSR disease severity in an oil palm plantation by applying decision tree, random forest, and support vector machine as supervised learning algorithms. The results were affected by the criteria used to describe BSR infection symptoms in the observation step. The removal of outliers based on band 4 from the results of the 'stepclass' algorithm improved the accuracies of the learning algorithms by 1–5%. New criteria of BSR infection symptoms were described after selecting the data using threshold values that were determined based on the distribution of reflectance values of band 4. In future studies, we plan to use the new criteria for describing BSR infection symptoms to identify and classify BSR infection in oil palm plantations using an unmanned aerial vehicle with multispectral camera.

Chapter 5. Predicting Oil Palm Leaf Nutrient Contents in Kalimantan, Indonesia by Measuring Reflectance with a Spectroradiometer

5.1. Introduction

Oil palm (*Elaeis guineensis* Jacq) is an economic crop in Indonesia and Malaysia (World Growth 2011; Basiron 2007), which are now the largest exporters of vegetable oil in the world (OECD/Food and Agriculture Organization of the United Nations 2015). Fertilizer is an important factor affecting oil palm production and represents 40–50% of the total field upkeep costs (Ng 2002). Oil palm plantations use leaf nutrient contents to determine fertilizer requirements, as is common for many other crops (Pritts and Heidenreich 2012; Memon, Memon, and Hassan 2005). To determine leaf nutrient contents, leaf samples are collected for analysis annually (Chapman and Gray 1949; Fairhurst and Mutert 1999).

Leaf samples are taken annually from leaf number 17 for adult palms and leaf number 9 for young palms. Commonly, the sampled leaves are grouped in leaf sampling units (LSUs) that consist of 30–40 selected palms each. One LSU is assigned for every 30–40 ha (Foster and Choong 1976; Uexkull, Fairhurst, and von Uexkull 1991). A potential problem is the accuracy of leaf sampling from leaf number 17. To identify leaf number 17, workers must first identify leaf number 1, the youngest fully open leaf, and count backwards. However, there can be some ambiguity in determining the youngest fully open leaf. Therefore, an improved technique is needed that can be used to easily monitor leaf nutrients in oil palm

plantations. Remote sensing techniques based on the reflectance values of leaves have potential in this regard.

Several researchers have used remote sensing techniques based on reflectance values for detection and predicting plant nutrient status. Albayrak (2008) measured reflectance for determining nitrogen (N), phosphorus (P), potassium (K), acid detergent fiber (ADF), and neutral detergent fiber (NDF) contents in sainfoin pasture and found a strong relationship between plant nutrient content and canopy reflectance with coefficient of determination (r^2) values in the range of 0.68–0.93. Cohen et al. (2010) estimated leaf N in potato using spectral data and simulated bands from the VENμS satellite on N fertilizer treatment levels, and obtained r^2 = 0.95, 80.5% overall accuracy and a kappa value of 74%. Özyigit and Bilgen (2013) showed significant relationships between spectral reflectance and leaf nutrient contents of N, P, and K in rangeland plants with r^2 values of 0.85, 0.43, and 0.84, respectively. Menesatti et al. (2010) estimated plant nutritional status using visible to near infra-red (NIR) spectrophotometric analysis in orange leaves and obtained high r^2 values of 0.909, 0.429, 0.991, 0.947, 0.944, 0.917, 0.925, and 0.889 for N, P, K, calcium (Ca), magnesium (Mg), iron (Fe), manganese (Mn) and zinc (Zn), respectively.

Using NIR reflectance spectroscopy, Rothman et al. (2009) showed that wavelengths in the range of 1100–2498 nm had a strong relationship with the nutritional value of food eaten by mountain gorillas with r^2 of 0.73–0.99. Başayiğit, Dedeoğlu, and Akgül (2015) showed that wavelengths of 540–560 nm (green visible)

and 990–1010 nm (NIR) were correlated with active Fe levels in apple, cherry, and peach with coefficients of accuracy of 76.70, 75.28, and 78.69%, respectively. The coefficient of accuracy is a statistic parameter for methods comparison (L. Lin and Torbeck 1998). Stein et al. (2014) predicted macronutrient contents from loblolly pine reflectance. They found that the important wavelengths in the partial least squares regression reflectance model for leaf N status were visible, red edge, and NIR regions, while the visible and red-edge regions were the important wavelengths for determining leaf P, K, and Mg. Min and Lee (2005) identified wavelengths of 448, 669, 719, 1377, 1773, and 2231 nm as significant for N detection in citrus. Özyigit and Bilgen (2013) found that wavelengths of 609, 647, 651, 654, 669, 675, 676, 680, 721, 727, and 760 nm were suitable for estimating the N level of rangeland plants; those of 675 and 680 nm were suitable for estimating P level; and those of 410, 411, 417, 422, 460, 463, 468, 646, 651, 658, 669, 670, 674, 676, and 682 nm were suitable for estimating K level. Thus, reflectance values from spectrometer data at different suitable wavelengths can be used to determine leaf nutrient contents in different plants. This method is likely applicable to oil palm.

The objective of this study was to explore suitable wavelengths for reflectance to predict leaf nutrient contents, especially those of N, P, K, Ca, Mg, B, Cu, and Zn, which are important for assessing fertilizer requirements in oil palm plantations. To achieve these objectives, oil palm leaf reflectance data and leaf nutrient analyses were collected. Predicting leaf nutrient contents in oil palm using

reflectance data will contribute to precision agriculture in oil palm plantations, especially for monitoring leaf nutrient status to determine fertilizer requirements.

5.2. Materials and Methods

5.2.1. Study site and leaf sampling methods

The study was conducted in one oil palm plantation belonging to a company in Pundu, East Kotawaringin District, Central Kalimantan Indonesia (Figure 5.1). Leaf samples were collected during 14–15 March 2015 from three groups of oil palms of different ages: 5, 12, and 17 years. Symptoms of K and Mg deficiencies in oil palms show in older leaves, N deficiency symptoms occur in younger and older leaves, while micronutrient deficiencies commonly appear in younger leaves (Corley and Tinker 2003; H.R. Von Uexkull and T.H. Fairhurst 1999). Therefore, leaf samples were taken from leaf number 9 (young), 17 (middle), 25 (old) and 33 (older) in every group. A total of 42 leaf samples were collected for leaf nutrient analyses, which was performed at the Leaf Laboratory of the Indonesian Oil Palm Research Institute in Medan, North Sumatra, Indonesia. Before leaf samples were processed in the laboratory for analysis, they were cleaned or wiped using aquadest (distilled water) to remove dust. Leaf samples were analyzed for N, P, K, Ca, Mg, B, Cu, and Zn. Samples were taken from three parts of every leaf (bottom, middle, and upper) for reflectance measurements and each sample included three leaflets. Reflectance was measured on the surface of leaflets at 10 positions per leaflet (Photo 5.2).

5.2.2. Reflectance measurements of oil palm leaves

The reflectance of oil palm leaves was measured using a portable spectroradiometer (FieldSpec3; Analytical Spectral Devices [ASD], Inc., Boulder, CO, USA) with a plant probe in the wavelength range of 350–2500 nm. Reflectance measurements were periodically calibrated with a white reference panel or white reference standard. The spectral resolution output data from the ASD operating system was 1 nm along the whole spectrum. The digital number (DN) values from measurement were converted to spectral reflectance values using ViewSpecPro version 6.2.0 of ASD software. The mean values of reflectance with 1 nm resolution were set as dataset 1. Data selected every 3 nm for wavelengths in the range of 350–1000 nm and every 10 nm for wavelengths of 1001–2500 nm were set as dataset 2 (Hatchell

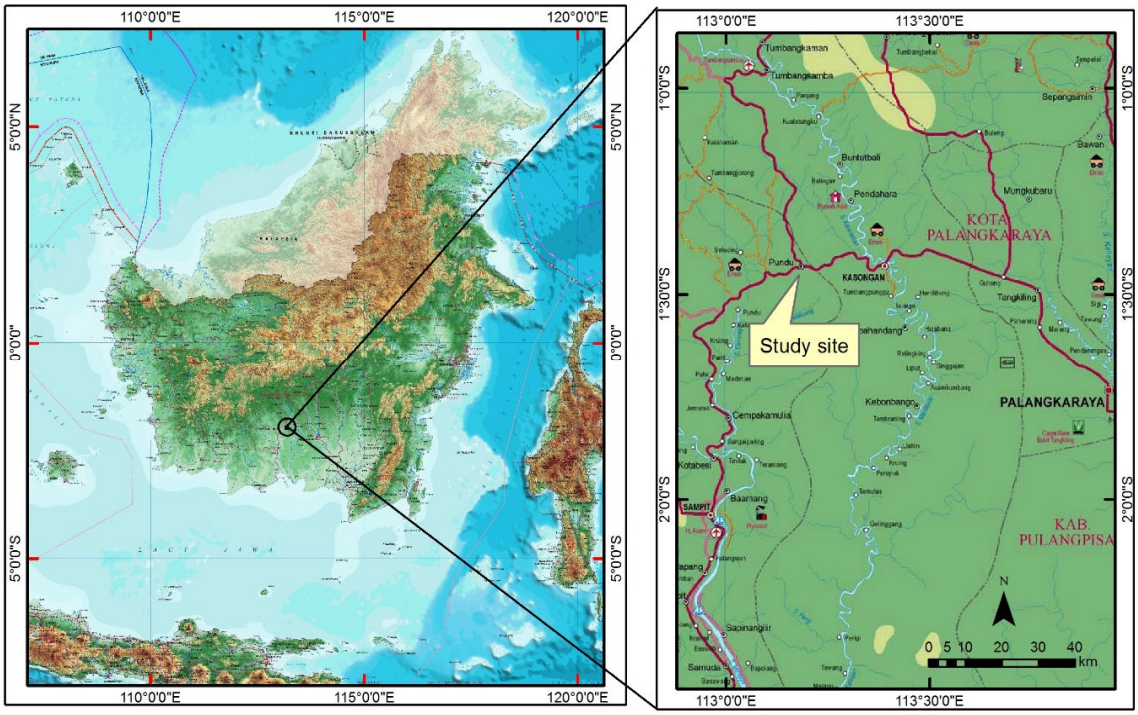


Figure 5.1. Map of study site (source map: Indonesian Geospatial Board, http://www.bakosurtanal.go.id/download/).

1999). Therefore, there were 42 data points of observation (leaf nutrient analysis), with 2151 spectral data as variables for dataset 1 and 367 spectral data as variables for dataset 2.

Figure 5.3 shows the mean reflectance values of oil palm leaves from dataset 2 with standard deviations. This reflectance pattern is a common vegetative spectral pattern (R M Hoffer 1984). The oil palm reflectance was low in visible wavelengths and high in NIR wavelengths; it started to decrease at approximately 1150 nm, and sharply decreased at approximately 1300 nm (Figure 5.3). Low reflectance in visible wavelengths is due to chlorophyll absorption and low reflectance at approximately 1450 and 1900 nm is from water absorption (R M Hoffer 1984).

Table 5.1. The wavelength ranges.

Groups	Symbol	Wavelengths (nm)
Ultra Violet A	G1	350–400
Blue (B)	G2	401–525
Green (G)	G3	526–605
Yellow (Y)	G4	606–655
Red (R)	G5	656–725
Far Red (Re)	G6	726–750
Short wave NIR (SW-NIR)	G7	751–1100
SW-NIR 1/NIR 1	G8	1101–1800
SW-NIR 2/NIR 2	G9	1801–2500

Photo 5.2. Photos of several activities in the study area: leaf sampling: leaf numbers 9, 17, 25 and 33 (upper left); leaf taken in field (upper right); spectral leaf measurement (bottom left); and wiped oil palm leaflets before leaf analysis (bottom right).

5.2.3. Data analysis

The wavelengths were divided into nine groups (Table 5.1) as described by Hatchell (1999). The analyses in this study were carried out using R software (RStudio 2015) as described below.

5.2.3.1 Variable screening for building normalized difference and simple ratio models

Dataset 1 was used for variable screening to build normalized difference (ND) and simple ratio (SR) formulae. The script adopted from Sonobe and Wang (2016) and used in this study explored and selected variables from among 2151 variables that

can be used to predict leaf nutrient contents based on r^2 values. The r^2 values calculated from predicted values of the ND and SR formulae were compared with leaf nutrient analysis (reference values). Thus, we found different suitable variables for each type of nutrient content in leaves.

5.2.3.2. Vegetation indices

Many previous studies have used vegetation indices for estimating and monitoring leaf nutrient contents in plants such as Albayrak (2008), Cohen et al. (2010), Munoz-Huerta et al. (2013), Mahajan et al. (2014), and Stein et al. (2014). The vegetation indices used in this study are shown in Table 5.2. The vegetation indices calculated from dataset 1 were used as variables for relationship analysis against the reference values. The goodness-of-fit parameter was r^2.

5.2.3.3. Variables selected using stepwise regression and the glm formula

Stepwise methods with the *glm* formula were applied to dataset 2 to determine the variables suitable for predicting leaf nutrient content. The stepwise method has been commonly used in research for selecting variables suitable for predicting or simulating leaf nutrient content in several plants (Starks et al. 2006; Albayrak 2008; Joffre et al. 1992; Başayiğit, Dedeoğlu, and Akgül 2015; Serusi et al. 2010; Min and Lee 2005; Basayigit and Senol 2009; Özyigit and Bilgen 2013). In this study, stepwise processing was performed with the MASS package in RStudio software (Ripley et al. 2015) using the *glm* formula (RStudio 2015).

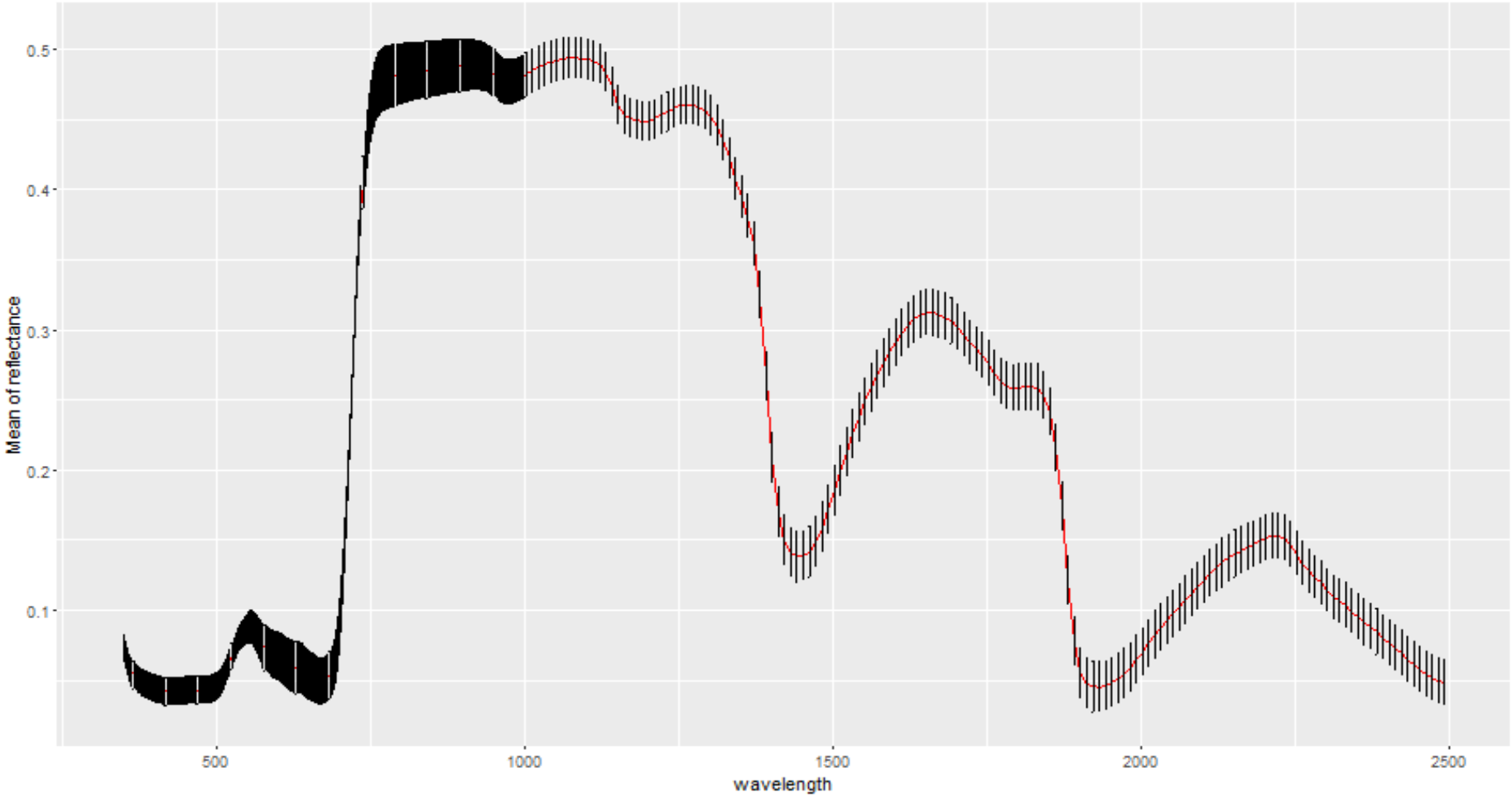


Figure 5.3. Mean of reflectance of oil palm leaf from spectral measurements with standard deviations.

Table 5.2. Vegetation indices used in this study.

Vegetation index	Formula [a]	Reference
Normalized difference vegetation index (NDVI)	$(R_{NIR} - R_{red})/(R_{NIR} + R_{red})$	Rouse et al. (1974)
Green normalized difference vegetation index (GNDVI)	$(R_{800–900} - R_{540–560})/(R_{800–900} + R_{540–560})$	Gitelson, Kaufman, and Merzlyak (1996)
Simple ratio (SR)	$R_{800–900}/R_{650–700}$	Birth, Gerald S.; McVey (1968)
Modified chlorophyll absorption reflectance index (MCARI)	$[(R_{700} - R_{670}) - 0.2(R_{700} - R_{550})] \times (R_{700}/R_{670})$	Daughtry (2000)
Transformed chlorophyll absorption reflectance index (TCARI)	$3[(R_{700} - R_{670}) - 0.2(R_{700} - R_{550})] \times (R_{700}/R_{670})$	Haboudane et al. (2002)
MCARI1	$1.2[2.5(R_{800} - R_{670}) - 1.3(R_{800} - R_{550})]$	Haboudane et al. (2004)
MCARI2	$1.2[2.5(R_{800} - R_{670}) - 1.3(R_{800} - R_{550})]/[((2 \times R800 + 1)^2 - (6 \times R_{800} - 5(R_{650})^{0.5}))^{0.5} - 0.5]$	Haboudane et al. (2004)
N870_1450	$(R_{870} - R_{1450})/(R_{870} + R_{1450})$	Pimstein et al. (2011)
N1645_1715	$(R_{1645} - R_{1715})/(R_{1645} + R_{1715})$	Pimstein et al. (2011)
P1080_1460	$(R_{1080} - R_{1460})/(R_{1080} + R_{1460})$	Mahajan et al. (2014)

[a] The letter "R" followed by a three or four-digit number stands for the wavelength of the respective reflectance value.

The parameters set for the *glm* formula were Gaussian for family, TRUE for model, and *glm.fit* for method. The stepwise processing was based on stepAIC with both (forward and backward) directions, k was set to 2, and trace was set to 1. The stepwise processing was applied to the nine wavelength ranges, and all variables selected as output from stepwise processing were tested for significant covariates. Therefore, the final variables selected were all significant at $p = 0.05$ for inclusion in the model to predict leaf nutrient contents (Bursac et al. 2008). The function of automated model selection based on *p*-value was adopted from Meys (2013). The parameters of goodness-of-fit were the adjusted r^2 and r^2 between the reference values and predicted values from the output of the model predicting nutrient content.

5.3. Results

5.3.1. Leaf nutrients analysis

Leaf nutrients had low to high variability in the order P, N, K, Ca, Zn, Mg, B, and Cu according to the leaf nutrient content analysis (Table 5.3). The leaf nutrient contents found in this study were highly variable; this was probably because samples were taken from young (leaf number 9), middle (number 17), old (number 25), and older leaves (number 33). Leaf age is one factor affecting leaf nutrient concentration (Fairhurst and Mutert 1999). The variability of leaf nutrient contents in this study was important for determining the relationship with leaf reflectance spectra (Starks et al. 2006).

Table 5.3. Summary of leaf nutrient content analysis results.

Nutrient	Min	Max	Mean	SD	CV (%)
N (%)	1.80	3.22	2.51	0.31	12.36
P (%)	0.16	0.20	0.17	0.01	6.63
K (%)	0.86	1.49	1.04	0.17	16.29
Ca (%)	0.42	0.93	0.67	0.12	18.11
Mg (%)	0.16	0.39	0.23	0.06	28.22
B (ppm)	7.52	25.82	16.10	4.60	28.58
Cu (ppm)	1.52	9.43	5.29	1.77	33.40
Zn (ppm)	7.98	26.97	14.17	3.98	28.05

Abbreviations: Min, minimum; Max, maximum; SD, standard deviation; CV, coefficient of variation

5.3.2. Normalized difference and simple ratio formulae

Explored and selected variables from the leaf reflectance spectra used to build the ND and SR formulae for predicting leaf nutrient contents are shown in Table 5.4. Both ND and SR formulae had the same variables selected (Table 5.4); N had the highest r^2 (0.534) and B the lowest (0.331). Similar to the results of Mukaka (2012), the r value in range of 0.70-0.90 that achieved in both ND and SR showed that spectral numbers X1423 and X1877 were highly positively correlated with N leaf nutrient analysis (reference value) and X1164 and X1238 were highly positively correlated with Ca. The equations of ND and SR for N and Ca leaf nutrient contents are as follows:

$$\text{N-ND}_{(X1423\_X1877)} = \frac{(X1423 - X1877)}{(X1423 + X1877)} \quad (1)$$

$$\text{N-SR}_{(X1423\_X1877)} = \frac{X1423}{X1877} \quad (2)$$

$$\text{Ca-ND}_{(X1164\_X1238)} = \frac{(X1164 - X1238)}{(X1164 + X1238)} \quad (3)$$

$$\text{Ca-SR}_{(X1164\_X1238)} = \frac{X1164}{X1238} \quad (4)$$

P, K, Mg, B, Cu, and Zn had moderate positive correlations with ND and SR formulae, and the variables selected are given in Table 5.4. The format of the equations for P, K, Mg, B, Cu, and Zn of ND were the same as those of Equations 1 and 3, and the structures of the equations for SR were the same as those of Equations 2 and 4. Different spectra were selected for the different variables: NIR 1 (G8) was used for determining P and Ca; NIR 1 (G8) and NIR 2 (G9) were used for N and B; SW-NIR (G7) was used for Mg and Cu; green (G3) and red (G5) were used for K; and far red (G6) and SW-NIR (G7) were used for Zn.

The wavelengths selected from the ND and SR models, including those for predicting N, P, Ca, B, and Cu contents, were in the range of 1005–1877 nm (NIR or SWIR to SWIR2 or NIR2). This is in accordance with the results of Pimstein et al. (2011) showing that N and K had correlations with the visible and SWIR wavelengths. They proposed a new ND equation for N with wavelength numbers X870 and X1450 for Equation 1 and X1645 and X1715 for Equation 2 that had better results than those of the common vegetation indices. Mahajan et al. (2014) proposed new a vegetation index for predicting P content using wavelength numbers X1080 and X1460. The wavelength number selected for building the ND and SR models to predict N and P contents was selected from the NIR region, which is in accordance with the results of Pimstein et al. (2011); they found interactions for both N and P

with wavelengths in the NIR region. Ca in oil palm leaves is correlated with the lignin contents in the leaves (Nur Sabrina, Sariah, and Zaharah 2012). In Yao et al. (2010), the wavelength numbers X1164 and X1238 in the ND and SR formulae used to predict Ca content in leaves were related to those in the NIR to determine lignin content in *Acacia* trees. B functions in oil palms in translocation of N and P and pollen viability. The Cu in oil palms is needed for N metabolism and pollen viability (Tengoua et al. 2015). Therefore, the variables selected for predicting B and Cu by ND and SR were similar to those selected for N and P.

The different wavelength numbers selected from ND and SR for K, Mg, and Zn were in the range of X530–X773. In oil palm, K is important for proper stomatal function in leaves, transport of assimilates from photosynthesis, enzyme activation, and oil synthesis. The Mg is important in chlorophyll and is essential for efficient photosynthesis (Rankie and Fairhurst 1999), while Zn has important roles in catalyzing activity of enzymes especially in oil synthesis (Ng 2002). In addition, deficiency symptoms for K, Mg, and Zn in the field can be seen in middle to older leaves. Indeed, the wavelength numbers selected were near to NIR, especially for Mg and Zn, while for K they were in the green and red spectra. For K, the wavelength number X530 was close to the peak of the green region and X707 occurred where reflectance started to increase, close to NIR. In addition, for Mg and Zn, the selected wavelength numbers were close to the peak of reflectance in the NIR region.

Table 5.4. The normalized difference (ND) and simple ratio (SR) models for dataset 1.

Leaf nutrients	Model	Variables [a]	r^2 [b]	r [c]
N	SR	X1423, X1877	0.53	0.73
	ND	X1423, X1877	0.53	0.73
P	SR	X1215, X1317	0.45	0.67
	ND	X1215, X1317	0.45	0.67
K	SR	X530, X707	0.38	0.62
	ND	X530, X707	0.38	0.62
Ca	SR	X1164, X1238	0.50	0.71
	ND	X1164, X1238	0.50	0.71
Mg	SR	X753, X773	0.43	0.66
	ND	X753, X773	0.43	0.66
B	SR	X1439, X1883	0.33	0.58
	ND	X1439, X1883	0.33	0.58
Cu	SR	X1005, X1033	0.49	0.70
	ND	X1005, X1033	0.49	0.70
Zn	SR	X744, X764	0.34	0.58
	ND	X744, X764	0.34	0.58

[a] Variables selected; the letter "X" followed by a three or four-digit number stands for the wavelength of the respective reflectance value.

[b] r^2 coefficient of determination.

[c] r coefficient of correlation.

5.3.3. Vegetation indices

The relationship of vegetative indices with leaf nutrient contents is shown in Table 5.5, with r^2 in the range of 0.000905–0.372752, and the highest r^2 for MCARI1 in predicting Ca leaf contents. The ND used spectral numbers 870 and 1450 (termed N870_1450) and predicted N, P, and K contents with r^2 = 0.347, 0.209, and 0.347, respectively. The GNDVI predicted B and Zn contents with r^2 = 0.296 and 0.086, respectively. MCARI1 predicted Ca contents with r^2 = 0.373, while MCARI predicted Mg contents with r^2 = 0.293, and TCARI predicted Cu content with r^2 = 0.244.

The results were in accordance with those of Pimstein et al. (2011) and Mahajan et al. (2014), especially in regard to the r^2 of vegetation indices for predicting leaf nutrient contents. The ND and SR results (Table 5.4) in this study had higher r^2 values than did common vegetation indices (Table 5.5). Different plant types respond different to the vegetation indices applied to predict leaf nutrient contents. Furthermore, leaf reflectance is affected by leaf structure, leaf conditions, moisture, maturity, and culture practices (Roger M Hoffer 1971; R M Hoffer 1984).

5.3.4. Multivariate analysis

Before multivariate analysis, variables were subjected to correlation analysis, which showed very high correlations between variables but very low correlations with leaf nutrient analyses. Therefore, the final variables selected in the multivariate analysis were all significant at p = 0.05 for inclusion in the model to predict leaf nutrient contents as described by Bursac et al. (2008). Multivariate analysis, using stepwise regression with the *glm* formula, improved r^2 and adjusted r^2 for common

vegetation indices and for the proposed ND and SR. No variable was selected from wavelength groups G2 and G7–G9 when applying stepwise regression (Table 5.6). In the stepwise regression, variables were added or deleted from the model one at a time until a stopping criterion (in this case the AIC value) was reached—wavelength numbers in G2 and G7–G9 were not selected to build the model for predicting leaf nutrient contents. The stepwise regression was ceased if the AIC value was very low at initial analysis.

The relationship between N content in leaves and leaf nutrient analysis (reference value) from stepwise regression results had adjusted r^2 values in the range of 0.27–0.82, with the higher adjusted r^2 values in wavelength group G3 (green). The relationship for P in leaves had adjusted r^2 values in the range of 0.18–0.74, with the highest r^2 in the G5 group (red). For K, Ca, Mg, B, Cu, and Zn, the adjusted r^2 values were in the ranges of 0.31–0.67, 0.41–0.66, 0.21–0.78, 0.30–0.63, 0.30–0.59, and 0.09–0.49, respectively. All variables selected were significant at $p < 0.05$.

Stepwise regression was applied to all variables selected from G1 and G3–G6 to produce a new model of the relationship between leaf nutrient contents and leaf reflectance. The r^2 and adjusted r^2 values were more strongly correlated in each wavelength group after stepwise regression. The adjusted r^2 values for N, P, K, Ca, Mg, B, Cu, and Zn were 0.99, 0.78, 0.89, 0.84, 0.94, 0.92, 0.64, and 0.76, respectively. Separating all the selected variables in the model from groups G1 and G3–G6 showed that wavelengths in the G3 (green) contributed most to the model. The

relative contributions of each wavelength group are shown in Table 5.7. Plant photosynthesis is strongly affected by chlorophyll. Gitelson et al. (1996) found that the green group had maximum sensitivity for a wide range of chlorophyll concentrations and proposed that green NDVI had a wider dynamic range than NDVI.

Table 5.5. Relationships (r2 values) between vegetative indices and leaf nutrient contents.

Vegetation	Leaf nutrient contents							
indice	N	P	K	Ca	Mg	B	Cu	Zn
NDVI	0.20	0.12	0.15	0.33	0.00	0.21	0.26	0.05
GNDVI	0.14	0.10	0.15	0.27	0.01	**0.30**	0.19	**0.09**
SR	0.20	0.07	0.07	0.32	0.00	0.16	0.31	0.07
MCARI	0.05	0.00	0.01	0.14	**0.293**	0.01	0.24	0.07
TCARI	0.05	0.00	0.01	0.14	0.291	0.02	0.24	0.07
MCARI1	0.27	0.13	0.12	**0.37**	0.01	0.15	0.34	0.02
MCARI2	0.27	0.10	0.08	0.36	0.01	0.13	**0.36**	0.04
N870_1450	**0.35**	**0.21**	**0.35**	0.25	0.02	0.02	0.10	0.03
N1645_1715	0.01	0.04	0.01	0.00	0.04	0.00	0.07	0.01
P1080_1460	0.31	0.18	0.31	0.22	0.03	0.02	0.08	0.02

Bold text shows the highest r^2 in each column.

Table 5.6. Results of stepwise regression.

G	N			P			K			Ca			Mg			B			Cu			Zn		
	1 [a]	*2* [b]	*3* [c]	*1*	*2*	*3*	*1*	*2*	*3*	*1*	*2*	*3*	*1*	*2*	*3*	*1*	*2*	*3*	*1*	*2*	*3*	*1*	*2*	*3*
G1	2	0.31	0.27	3	0.24	0.18	6	0.42	0.32	6	0.57	0.50	4	0.28	0.21	4	0.43	0.36	5	0.48	0.40	2	0.13	0.09
G2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G3	13	**0.88**	**0.82**	11	0.78	0.70	7	**0.72**	**0.67**	10	**0.74**	**0.66**	9	**0.83**	**0.78**	8	**0.71**	**0.63**	11	**0.70**	**0.59**	6	0.43	0.33
G4	6	0.68	0.63	5	0.67	0.62	3	0.46	0.42	7	0.56	0.47	5	0.73	0.69	5	0.51	0.44	2	0.41	0.38	5	0.49	0.42
G5	7	0.72	0.66	10	**0.80**	**0.74**	9	0.65	0.55	8	0.65	0.57	7	0.80	0.76	11	0.71	0.60	8	0.55	0.44	9	**0.60**	**0.49**
G6	3	0.32	0.27	4	0.29	0.22	4	0.38	0.31	5	0.48	0.41	5	0.63	0.57	4	0.37	0.30	3	0.35	0.30	2	0.33	0.29
G7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
G9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C*	22	*0.99*	*0.99*	13	*0.85*	*0.78*	17	*0.93*	*0.89*	20	*0.92*	*0.84*	23	*0.97*	*0.94*	17	*0.95*	*0.92*	11	*0.74*	*0.64*	13	*0.84*	*0.76*

*Stepwise regression applied to all selected variables of wavelength groups G1–G9; highlighted data are the results of the highest adjusted r^2 and r^2 values for the groups G1–G9 and italics represent the results of the adjusted r^2 and r^2 for all variables selected in G1–G9.

[a] number of variables selected.

[b] coefficient of determination (r^2) from linear relationships.

[c] adjusted r^2.

Table 5.7. Variable separation based on wavelength groups in the model C.

Nutrient	Number of variables	Number of variables selected from wavelength group					r^2 [a]	*Adjusted* r^2 [b]	RMSE [c]
		G1	G3	G4	G5	G6			
N	22	6	12	4	-	-	0.99	0.99	0.02
P	13	-	7	-	6	-	0.85	0.78	0.004
K	17	2	9	2	-	4	0.93	0.89	0.04
Ca	20	-	10	3	6	1	0.92	0.84	0.04
Mg	23	-	12	3	5	3	0.98	0.94	0.01
B	17	5	7	4	-	1	0.95	0.92	1.01
Cu	11	-	7	2	2	-	0.74	0.64	0.89
Zn	13	-	5	5	-	3	0.84	0.76	1.58

[a] r^2: coefficient of determination.

[b] adjusted r^2: adjusted of coefficient of determination.

[c] *RMSE*: root mean square error of prediction.

The equations from the models built from stepwise regression and based on the highest adjusted r^2 from groups G1–G9 (model 1: N1, P1, K1, Ca1, Mg1, B1, Cu1, and Zn1) and from group C (model 2: N2, P2, K2, Ca2, Mg2, B2, Cu2, and Zn2) are shown in Table 5.8. Model 2 had an improved root mean square error (RMSE), which was less than that for model 1. Predictions of B and Zn contents had RMSE values > 1, with 3.44, 1.01, 2.97, and 1.25 for B1, B2, Zn1, and Zn2, respectively. The RMSEs were 0.004–0.95 for the contents of all other leaf nutrients. An RMSE value close to zero is one indicator of good prediction in a model. These results are consistent with those of Albayrak (2008) showing that spectral reflectance in the visible to NIR range can be used for leaf nutrient prediction. Min and Lee (2005) found high correlation coefficients between absorbance at each wavelength number and leaf N concentration for wavelength numbers in the range of X553–X707. Stein et al. (2014) showed that N, P, K, Ca, and Mg contents in the leaves of loblolly pine were correlated with the visible spectrum (500–600 nm) and at approximately 700 nm.

5.4. Discussion

As described above, the variables selected in the ND and SR equations were in accordance with the results of previous research (Mahajan et al. 2014; Pimstein et al. 2011; Yao et al. 2010). The ND and SR equations in this study were able to predict leaf nutrient contents better than existing vegetation indices. The performance of predictions for leaf nutrient contents were compared with common vegetation

indices described in the "Vegetation Indices" section above. Validation is an important step before any model can be applied to the field. Regarding the wavelength number selected in the proposed ND and SR equations, the appropriate method for calibrating and selecting wavelength number that match with selected wavelength number in the proposed models if used imagery both hyperspectral or multispectral as data source. It caused the different available of wavelength numbers in spectrometer and optic sensors for hyperspectral or multispectral imagery.

Table 5.8. The final model of variables selected by stepwise regression to predict leaf nutrient contents.

Model[a]	Equation[b]	Adjust r^2	RMSE
N1	3.1024 + 467.2588 *X527* − 684.8238 *X533* + 3383.4250 *X554* − 3716.2002 *X557* − 1448.3226 *X563* + 3663.9700 *X566* − 1504.2360 *X569* − 2331.0380 *X572* + 4116.1217 *X575* − 1758.3488 *X581* − 3567.0650 *X596* + 4866.7420 *X602* − 1480.2157 *X605*	0.82	0.11
N2	0.9398 + 17.7735 *X350* − 48.8114 *X356* + 66.5315 *X359* − 49.29 *X362* − 23.2067 *X371* + 29.5811 *X386* − 140.9929 *X533* + 1092.5735 *X560* − 1734.083 *X563* + 1220.4245 *X566* + 2687.8925 *X572* − 5640.1083 *X575* + 4687.7969 *X581* − 6064.3569 *X584* + 3474.066 *X587* − 864.0621 *X593* + 3647.2296 *X596* − 1482.6584 *X602* − 4400.7149 *X614* + 6228.861 *X620* − 6652.4031 *X638* + 3951.2114 *X641*	0.99	0.02
P1	0.14225 + 61.41929 *X659* − 164.28227 *X665* + 173.01223 *X674* − 140.26707 *X683* + 62.41344 *X686* + 110.94254 *X695* − 209.82219 *X698* + 119.49512 *X701* − 14.03119 *X710* + 1.6346 *X722*	0.74	0.01
P2	0.19823 + 33.92419 *X527* −36.76529 *X530* + 82.71930 *X560* − 85.65716 *X563* + 32.01229 *X575* − 115.32474 *X584* + 82.28470 *X587* + 148.20318 *X662* −166.33245 *X665* + 29.98227 *X674* − 30.54503 X698 + 30.45886 *X701* − 4.60416 *X710*	0.78	0.004
K1	−1.261e−02 − 2.280e+02 *X545* + 2.961e+03 *X560* − 2.212e+03 *X563* − 1.826e+03 *X566* + 1.626e+03 *X572* − 1.581e+03 *X584* + 1.270e+03 *X587*	0.67	0.09
K2	0.3447 − 18.6181 *X353* + 23.0783 *X356* + 490.5255 *X542* − 821.7928 *X545* + 2794.2797 *X560* − 1411.0233 *X563* − 2414.2687 *X566* + 1564.4641 *X572* −1706.4495 *X584* + 1686.5171 *X587* − 829.7678 *X602* + 992.9684 *X608* − 337.0019 *X629* − 1668.2468 *X737* + 3210.4104 *X740* − 2952.7050 *X746* + 1405.8705 *X749*	0.89	0.04
Ca1	1.3847 − 740.5431 *X545* + 1449.708 *X548* − 1749.5866 *X557* + 692.9173 *X563* + 1325.2003 *X566* − 2672.9279 *X572* + 2078.7968 *X575* − 455.3902 *X590* − 987.0152 *X599* + 1059.245 *X602*	0.66	0.06
Ca2	2.367 +181.6897 *X527* − 487.654 *X536* + 1083.138 *X548* − 1781.1 *X557* + 834.1712 *X563* + 1390.809 *X566* − 3256.21 *X572* + 2143.295 *X575* − 975.563 *X599* + 986.4457 *X602* − 736.395 *X638* + 1398.383 *X644* − 840.926 *X647* + 867.3186 *X671* − 1454.06 *X674* + 620.4471 *X677* + 635.7719 *X713* − 2213.57 *X719* + 1733.621 *X722* − 130.24 *X737*	0.84	0.04
Mg1	0.43531 − 313.49378 *X545* + 706.5662 *X548* − 283.17626 *X551* − 639.89135 *X554* + 876.32033 *X557* − 442.20013 *X563* + 224.28792 *X590* +303.0187 *X602* − 431.43494 *X605*	0.78	0.03
Mg2	0.2826 − 362.1573 *X545* + 520.1222 *X548* − 216.3194 *X551* + 471.2007 *X557* − 462.811 *X563* − 497.4692 *X569* + 821.7298 *X575* − 300.6222 *X578* + 336.6243 *X590* − 772.0851 *X596* + 1184.2648 *X602* − 727.4807 *X605* + 201.404 *X620* − 408.3228 *X644* + 592.4485 *X650* − 507.8061 *X656* + 225.9669 *X677* − 111.4806 *X683* + 56.6352 *X704* − 429.0739 *X722* + 774.2846 *X728* − 436.1921 *X734* + 46.9323 *X749*	0.94	0.01
B1	31.726 + 13993.706 *X527* − 24088.922 *X530* + 23742.099 *X536* − 18657.087 *X539* + 15175.417 *X551* − 33592.466 *X563* + 28364.629 *X569* − 4906.555 *X593*	0.63	3.44
B2	56.78 + 589.16 *X377* + 1538.81 *X383* − 3068.97 *X386* − 1877.26 *X395* + 3340.11 *X398* + 19929.74 *X536* − 38269.38 *X539* + 48398.62 *X551* − 53011.61 *X563* + 21088.17 *X572* + 32501.82 *X599* − 2303.44 *X602* + 32411.17 *X620* − 41840.03 *X623* + 58478.05 *X638* − 47856.94 *X641* − 99.04 *X749*	0.92	1.01
Cu1	1.461 − 6023.451 *X530* + 8493.368 *X533* − 9279.56 *X542* + 8900.348 *X548* − 11739.268 *X572* + 16234.376 *X578* − 17964.269 *X584* + 28034.335 *X590* − 22428.39 *X593* + 19608.022 *X599* − 13923.307 *X602*	0.59	0.95
Cu2	−2.08E+01 – 6.61E+03 *X530* + 6.94E+03 *X533* − 1.14E+03 *X551* + 1.35E+04 *X590* − 1.75E+04 *X593* + 1.94E+04 *X599* − 1.52E+04 *X602* − 1.18E+04 *X632* + 1.23E+04 *X635* − 5.32E+02 *X722* + 5.71E+02 *X725*	0.64	0.89
Zn1	17.33 − 30405.63 *X668* + 74156.81 *X674* − 76207.32 *X680* + 65150.69 *X686* − 37410.03 *X689* + 41547.49 *X704* − 46237.16 *X707* + 25245.86 *X719* − 15708.24 *X722*	0.49	2.97
Zn2	8.635 – 14322.2 *X530* + 20355.71 *X533* + 50406.23 *X563* − 71712.8 *X602* − 47384.6 *X557* + 40321.98 *X608* + 97905.03 *X614* − 47719.9 *X617* − 79472.9 *X629* + 51667.07 *X635* − 6885.32 *X728* + 8627.049 *X731* − 1711.68 *X749*	0.76	1.58

[a] The equations from the models built using stepwise regression and based on the highest adjusted r^2 from wavelength groups: N1, P1, K1, Ca1, Mg1, B1, Cu1, and Zn1. The equations using variables selected from each wavelength groups in model 2 are as follows: N2, P2, K2, Ca2, Mg2, B2, Cu2, and Zn2. Leaf nutrient *N* nitrogen, *P* phosphorus, *K* potassium, *Ca* calcium, *Mg* magnesium, *B* boron, *Cu* copper, and *Zn* zinc.

[b] The letter "X" followed by a three-digit number stands for the wavelength of the respective reflectance value.

An important result from the multivariate analysis was that wavelengths in group G3 (green) were the most important in this study. Variables selected from G3 had strong effects in all proposed models for predicting leaf N, P, K, Ca, Mg, B, Cu, and Zn contents. A limitation of the model built using multivariate analysis in this study was that many variables were selected to create the model. Generally, a multivariate model has better prediction power compared with simple regression. This study showed that the model involving variables from visible wavelengths (G1 and G3–G6) performed well. The results of the present study are in accordance with previous results showing that visible wavelengths contribute highly to predicting leaf nutrient contents (Özyigit and Bilgen 2013; Stein et al. 2014; Min and Lee 2005; Albayrak 2008). The next challenge will be to implement the proposed multivariate model using hyperspectral imagery because of the different spectral resolutions in spectroradiometer and hyperspectral image data. Another challenge in producing new techniques for precision agriculture, especially for predicting leaf nutrient contents in oil palm plantations, is adapting robust models and methods to the specific characteristics of the plantations. Oil palm plantations are commonly very large; one estate may be 2000–3000 ha. Leaf nutrient monitoring to determine fertilizer needs in oil palm plantations is based on the leaf nutrient contents of LSUs that may each represent 30 ha. Furthermore, the characteristics of the leaf samples taken from the 30 selected trees in each LSU depend on the homogeneity of oil palm conditions (tree age, vigorous, soil type, and land topography). Thus, the proposed model provides an early version of the kind of robust model and method needed to

precisely calculate leaf nutrient contents based on spectral values from remote sensing.

5.5. Conclusions

The ND and SR models using wavelength numbers X1423 and X1877 had strong positive correlations with N leaf nutrient analysis ($r = 0.731$), while X1164 and X1238 had strong positive correlations with Ca leaf nutrient analysis ($r = 0.707$). The P, K, Mg, B, Cu, and Zn had moderate positive correlations, with r values in the range of 0.575–0.699. Several vegetation indices commonly used for predicting leaf nutrient contents had lower r values than the proposed ND and SR in this study.

Multivariate analysis using stepwise regression with the *glm* formula resulted in higher r^2 values than those for the proposed ND and SR and several existing vegetation indices. Wavelengths the in G2 (blue), G7 (SW-NIR), G8 (SWIR1), and G9 (SWIR2) groups did not contribute to the final models for predicting leaf nutrient content. For all leaf nutrient elements, models that involved variables selected from the G1 and G3–G6 groups had better performance than those that involved variables selected from each group of wavelengths individually. Variables from the G3 (green) group were always selected and contributed the most to building the final models. Variables from the G4 (yellow) group were the second most important group selected for building the models. The wavelengths in the proposed vegetation indices and the multivariate model proposed should be studied further using hyperspectral sensors or remote sensing satellites.

Chapter 6. Summary

1. The actual oil yield from oil palm is affected by agronomic practices, management practices, economic factors, and social and environmental factors. There were two main issues in the agronomic practices. First, the main disease of oil palms in Indonesia and Malaysia is basal stem rot disease (BSR), which is caused by *Ganoderma orbiforme*. There is no effective treatment for BSR, and all currently applied treatments only prolong the life of the oil palm. Second, fertilizer is an important factor affecting oil palm production, and accounts for 40–50% of the total costs of field upkeep. The fertilizer requirements of oil palm are determined based on analyses of leaf nutrient contents, like in other agricultural crops. Leaf samples are analyzed each year to optimize fertilizer inputs. This study had two main goals: to develop a fast and effective method to identify and monitor BSR disease, and to devise an inexpensive, rapid, and accurate method to monitor leaf nutrient status to optimize the fertilizer dosage. Remote sensing is a potential technique to achieve these goals and is already used in several agricultural practices.
2. The detection method for BSR should be able to identify and map BSR infections quickly and inexpensively. Remote sensing is an appropriate strategy for such detection, and there is potential to improve the accuracy of remote sensing by combining it with machine-learning algorithms. For leaf nutrient monitoring, it would be advantageous to have an improved method that is simple,

inexpensive, and accurate. Remote sensing techniques to estimate leaf nutrient status based on the reflectance of leaves have great potential in this regard.

3. In the past, oil palm density has been determined by manually counting trees every year in oil palm plantations. The measurement of density provides important data related to palm productivity, fertilizer needed, weed control costs in a circle around each tree, labourers needed and needs for other activities. Manual counting requires many workers and has potential problems related to accuracy. Remote sensing provides a potential approach for counting oil palm trees. The oil palm trees analysed in this study have different ages and densities. QuickBird imagery was applied with the six pansharpening methods and was compared with panchromatic QuickBird imagery. The black and white imagery from a false colour composite of pansharpening imagery was processed in three ways: (1) oil palm tree detection, (2) delineation of the oil palm area using the red band, and (3) counting oil palm trees and accuracy assessment. For oil palm detection, we used several filters that contained a Sobel edge detector, texture analysis co-occurrence, and dilate, erode, high-pass, and opening filters.
4. The results of this study improved upon the accuracy of several previous research studies that had an accuracy of about 90–95%. The results in this study show (1) modified intensity-hue-saturation (IHS) Resolution Merge is suitable for 16-year-old oil palm trees and have rather high density with 100% accuracy; (2) Colour Normalized (Brovey) is suitable for 21-year-old oil palm trees and have low density with 99.5% accuracy; (3) Subtractive Resolution Merge is

suitable for 15- and 18-year-old oil palm trees and have a rather high density with 99.8% accuracy; (4) PC Spectral Sharpening with 99.3% accuracy is suitable for 10-year-old oil palm trees and have low density; and (5) for all study object conditions, Colour Normalized (Brovey) and Wavelet Resolution Merge are two pansharpening methods that are suitable for oil palm tree extraction and counting with 98.9% and 98.4% accuracy, respectively. The 49,937 palms were identified by the Brovey method.

5. We applied remote sensing and machine learning to identify and classify the two classes (healthy and unhealthy) of BSR disease with QuickBird imagery. The machine-learning models were support vector machine, random forest (RF), and classification and regression tree models used for predicting BSR disease in oil palm plantations. QuickBird imagery archived on 4 August 2008 was applied in three classifier models. The RF model was best at predicting, classifying, and mapping oil palm BSR in terms of overall accuracy, producer accuracy, user accuracy, and kappa value. Using 75% of the data for training and 25% for testing, the RF classifier model achieved 91.43% overall accuracy. The RF classifier model used to produce outspread of the healthy and unhealthy oil palms using 49,937 data from the Brovey method and separated into 37,617 (75%) and 12,320 (25%) individuals, respectively.
6. WorldView-3 imagery has not been used to classify the four levels of BSR disease severity. The objectives were to predict the multilevels of BSR disease severity using the eight multispectral bands of WorldView-3 imagery and

supervised learning algorithms, and to describe the characteristics of the severity levels that can be identified by WorldView-3 imagery. Observation data were collected for 1923 oil palm trees with various levels of infection, including healthy trees, and unhealthy trees with three levels of symptoms from mild to severe. Decision tree, random forest, and support vector machine learning algorithms were applied.

7. The overall accuracy was low and the accuracy was improved by around 1–5% after outliers were removed from the data. The results were affected by the criteria used to describe the BSR disease symptoms in the observation step. New criteria of BSR disease symptoms that can be identify using WorldView-3 imagery were described after selecting the data using threshold values that were determined based on the distribution of reflectance values of band 4.
8. The new criteria of BSR disease symptoms based on this study were healthy oil palm trees (H) as trees with perfect canopy condition; unhealthy oil initially symptomatic (UH1) is trees with two unopened spears and emerging yellowish leaves and appearance of necrosis on older leaves or decreased leaflet size; unhealthy oil palm trees moderately symptomatic (UH2) is trees with more than two unopened spears and yellowish leaves on middle to bottom part of the canopy with the appearance of necrosis or 1-2 older leaves fractured; and unhealthy oil palm trees severely symptomatic (UH3) is trees with more than three unopened spears, almost all older leaves fractured, yellowish leaves along with wide range of necrosis, not producing new bunch/fruit, and having

damaged parts in the basal stem (rotten and hole).

9. In future studies, we plan to use the new criteria of BSR disease symptoms to identify and classify BSR disease in oil palm plantations using an unmanned aerial vehicle with multispectral camera.
10. We studied leaf nutrient contents based on spectral reflectance data to explore suitable wavelengths for predicting the content of the most important leaf nutrients: nitrogen, phosphorus, potassium, calcium, magnesium, boron, copper, and zinc. The samples were taken from one plantation belonging to an oil palm company in Pundu, Central Kalimantan, Indonesia. The proposed vegetation indices, several common vegetation indices, and stepwise regression were used to build models for predicting leaf nutrient contents.
11. The ND and SR models using wavelength numbers X1423 and X1877 had strong positive correlations with N leaf nutrient analysis (r = 0.731), while X1164 and X1238 had strong positive correlations with Ca leaf nutrient analysis (r = 0.707). The P, K, Mg, B, Cu, and Zn had moderate positive correlations, with r values in the range of 0.575–0.699. Several vegetation indices commonly used for predicting leaf nutrient contents had lower r values than the proposed ND and SR in this study.
12. For all leaf nutrient elements, models that involved all variables selected as significant through stepwise regression from the G1 (ultra violet A) and G3 (green) to G6 (far red) wavelength groups had better performance than models that involved variables selected from each wavelength group individually. For

predicting all leaf nutrient contents, variables from the G3 (green) wavelength group were always selected and these contributed more to building the models than any other groups. Our results indicate that our proposed vegetation indices and multivariate model can be used to predict leaf nutrient contents and determine the fertilizer requirements of oil palms.

13. The next challenge will be to implement the proposed multivariate model using hyperspectral imagery because of the different spectral resolutions in spectroradiometer and hyperspectral image data. And another challenge in producing new techniques for precision agriculture, especially for predicting leaf nutrient contents in oil palm plantations, is adapting robust models and methods to the specific characteristics of the plantations.

Acknowledgments

Praise be to Allah SWT, the Lord of the worlds that have been blessed me to reach these works.

I would like to express my deepest gratitude to my supervisor Associate Professor Hiroshi Tani, Laboratory of Environmental Informatics for all support, guidance, and suggestion for me to complete this research. I also thank to Prof. Norio Kondo, Prof. Noboru Noguchi and Associate Professor Xiufeng Wang for their valuable comments and suggestions on this dissertation. And many thanks to all member of the Laboratory of Environmental Informatics, Graduate School of Agriculture Hokkaido for their advice, support and friendship.

Special thanks to Director of Indonesian Oil Palm Research Institute for the scholarship and occasion for me to complete my Doctoral study at Hokkaido University, Japan. I also thanks to Dhimas Wiratmoko, Agus Eko Prasetyo, Panca Agus Priyanto, Niwanto, the KITA-GIS team, and the Plant Protection Team of Indonesian Oil Palm Research Institute for all helped in collecting data. I would like to thanks to Hendrik Segah, Ph.D, Dr. Witjaksana Darmosarkoro, Fredy Wijaya, and all Staff of Research Division PT Bumitama Gunajaya Agro (PT BGA) for their invaluable help and support during the leaf sampling and spectral measurement. And also thank to Dean of Agriculture Faculty of University of Palangka Raya (UPR) for allowing the utilization of FiledSpec3/ASD instruments in this study.

Thanks to all members of Indonesian Student Association of Hokkaido (PPI-Hokkaido), all members of Hokkaido Islamic Society, and all ITH members who always give me spirit and inspiration to reach these works.

My grateful to my loving wife, Dini Mufriah, my son, Rafif Herdin Mukti, and my daughter, Ratih Azzahra Santoso, for their supports and sacrifices. I also grateful to my father, my mother, my father in law, my mother in law, and my brother for their loves, prayers, encouragement and moral support that inspired me to complete this study.

References

Abdullah, N, and F Sulaiman. 2013. "The Oil Palm Wastes in Malaysia." *Biomass Now – Sustainable Growth and Use*, 75–100. doi:10.5772/55302.

Albayrak, Sebahattin. 2008. "Use of Reflectance Measurements for the Detection of N, P, K, ADF and NDF Contents in Sainfoin Pasture." *Sensors* 8 (11): 7275–7286. doi:10.3390/s8117275.

Ayodele, Taiwo Oladipupo. 2010. "Types of Machine Learning Algorithms." In *Types of Machine Learning Algorithms, New Advances in Machine Learning*, edited by Yagang Zhang, 19–49. Shanghai: InTech. doi:10.5772/9385.

Barcelos, Edson, Sara de Almeida Rios, Raimundo N. V. Cunha, Ricardo Lopes, SÃ©rgio Y. Motoike, Elena Babiychuk, Aleksandra Skirycz, and Sergei Kushnir. 2015. "Oil Palm Natural Diversity and the Potential for Yield Improvement." *Frontiers in Plant Science* 6 (March): 190. doi:10.3389/fpls.2015.00190.

Başayiğit, Levent, Mert Dedeoğlu, and Hüseyin Akgül. 2015. "The Prediction of Iron Contents in Orchards Using VNIR Spectroscopy." *Turkish Journal of Agriculture and Forestry* 39 (1): 123–134. doi:10.3906/tar-1406-33.

Basayigit, Levent, and Huseyin Senol. 2009. "Prediction of Plant Nutrient Contents in Deciduous Orchards Fruits Using Spectroradiometer." *International Journal of ChemTech Research* 1 (2): 212–224.

Basiron, Yusof. 2007. "Palm Oil Production through Sustainable Plantations." *European Journal of Lipid Science and Technology* 109 (4): 289–295. doi:10.1002/ejlt.200600223.

Beleites, Claudia, Ute Neugebauer, Thomas Bocklitz, Christoph Krafft, and Jürgen Popp. 2013. "Sample Size Planning for Classification Models." *Analytica Chimica Acta* 760 (June 2012). Elsevier B.V.: 25–33. doi:10.1016/j.aca.2012.11.007.

Bind, Shubham, Arvind Kumar Tiwari, and Anil Kumar Sahani. 2015. "A Survey of

Machine Learning Based Approaches for Parkinson Disease Prediction." *International Journal of Computer Science and Information Technologies* 6 (2): 1648–1655. http://www.ijcsit.com/docs/Volume 6/vol6issue02/ijcsit20150602163.pdf.

Birth, Gerald S.; McVey, George R. 1968. "Measuring the Color of Growing Turf with a Reflectance Spectrophotometer1." *Agronomy Journal* Vol.60 (6): 640–643. doi:10.2134/agronj1968.00021962006000060016x.

Bischl, Bernd, Michel Lang, Jakob Richter, Jakob Bossek, Leonard Judt, Tobias Kuehn, Erich Studerus, and Lars Kotthoff. 2016. "Mlr: Machine Learning in R." *Journal of Machine Learning Research* 17: 1–5. http://cran.r-project.org/package=mlr.

Bonneau, Xavier, Pieter Vandessel, Maxwell Buabeng, and Charles Erhahuyi. 2014. "Early Impact of Oil Palm Planting Density on Vegetative and Oil Yield Variables in West Africa." *Ocl* 21 (4): A401. doi:10.1051/ocl/2014009.

Breiman, Leo. 2001. "Random Forests." *Mach Learn* 45 (1): 5–32. doi:10.1023/A:1010933404324.

Breiman, Leo. 2004. "Consistency for a Simple Model of Random Forests." *Technical Report 670. Statistics Department, University of California at Berkeley*. https://www.stat.berkeley.edu/~breiman/RandomForests/consistencyRFA.pdf.

Breiman, Leo, Adele Cutler, Andy Liaw, and Matthew Wiener. 2015. "Random Forests for Classification and Regression." https://www.stat.berkeley.edu/~breiman/RandomForests/.

Breure, C. J. 2010. "Rate of Leaf Expansion: A Criterion for Identifying Oil Palm (Elaeis Guineensis Jacq.) Types Suitable for Planting at High Densities." *NJAS - Wageningen Journal of Life Sciences* 57 (2). Royal Netherlands Society for Agriculture Sciences: 141–147. doi:10.1016/j.njas.2010.03.001.

Breure, C. J., T. Menendez, and M. S Powell. 1990. "The Effect of Planting Density on the Yield Components of Oil Palm (Elaeis Guineensis)." *Experimental*

Agriculture 26 (1): 117–124. doi:10.1017/S0014479700015453.

Brownlee, Jason. 2016. "How Machine Learning Algorithms Work (They Learn a Mapping of Input to Output)." *Machine Learning Algorithms*. http://machinelearningmastery.com/how-machine-learning-algorithms-work/.

Burrell, Jenna. 2016. "How the Machine 'thinks': Understanding Opacity in Machine Learning Algorithms." *Big Data & Society* 3 (1): 2053951715622512. doi:10.1177/2053951715622512.

Bursac, Zoran, C Heath Gauss, David Keith Williams, and David W Hosmer. 2008. "Purposeful Selection of Variables in Logistic Regression." *Source Code for Biology and Medicine* 3: 17. doi:10.1186/1751-0473-3-17.

Cao, Xi Hang, Ivan Stojkovic, and Zoran Obradovic. 2016. "A Robust Data Scaling Algorithm to Improve Classification Accuracies in Biomedical Data." *BMC Bioinformatics* 17 (359). BMC Bioinformatics: 1–10. doi:10.1186/s12859-016-1236-x.

Chapman, G. W., and H. M. Gray. 1949. "Leaf Analysis and the Nutrition of the Oil Palm (Elaeis Guineensis Jacq.)." *Annals of Botany* 13 (4): 415–433. http://aob.oxfordjournals.org/content/13/4/415.short.

Chaudhary, Archana, Savita Kolhe, and Raj Kamal. 2016. "An Improved Random Forest Classifier for Multi-Class Classification." *Information Processing for Remote Sensing* 3: 215–222. doi:10.1016/j.inpa.2016.08.002.

Cohen, Jacob. 1960. "A Coefficient of Agreement for Nomial Scales." *Educational and Psychological Measurement* 20 (1): 37–46. doi:10.1177/001316446002000104.

Cohen, Y., V. Alchanatis, Y. Zusman, Z. Dar, D. J. Bonfil, A. Karnieli, A. Zilberman, et al. 2010. "Leaf Nitrogen Estimation in Potato Based on Spectral Data and on Simulated Bands of the VEN??S Satellite." *Precision Agriculture* 11 (5): 520–537. doi:10.1007/s11119-009-9147-8.

Congalton, Russell G. 2005. "Thematic and Positional Accuracy Assess- Ment of Digital Remotely Sensed Data." *2005 Proceedings of the Seventh Annual Forest Inventory and Analysis Symposium*, no. 1993: 149–154.

Congalton, Russell G. 1991. "A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data." *Remote Sensing of Environment* 37 (1): 35–46. doi:10.1016/0034-4257(91)90048-B.

Coops, Nicholas C., Matt Johnson, Michael A. Wulder, and Joanne C. White. 2006. "Assessment of QuickBird High Spatial Resolution Imagery to Detect Red Attack Damage due to Mountain Pine Beetle Infestation." *Remote Sensing of Environment* 103 (1): 67–80. doi:10.1016/j.rse.2006.03.012.

Corley, R. H. V. 1973. "Effects of Plant Density on Growth and Yield of Oil Palm." *Experimental Agriculture* 9 (2): 169. doi:10.1017/S0014479700005639.

Corley, R. H. V., Hardon J.J., and Wood B.J. 1976. *Oil Palm Research*. Amsterdam: Elsevier Scientific Publ.

Corley, R. H V. 2009. "How Much Palm Oil Do We Need?" *Environmental Science and Policy* 12 (2): 134–139. doi:10.1016/j.envsci.2008.10.011.

Corley, R H V, and P B Tinker. 2003. *The Oil Palm*. *Blackwell Science Ltd*. Fourth Edi. Malden, MA: Blackwell Science Ltd,. doi:10.1002/9780470750971.

Cracknell, Arthur Philip, K D Kanniah, Kian Pang Tan, and L Wang. 2013. "Evaluation of MODIS Gross Primary Productivity and Land Cover Products for the Humid Tropics Using Oil Palm Trees in Peninsular Malaysia and Google Earth Imagery." *International Journal of Remote Sensing* 34 (20): 7400–7423. doi:Doi 10.1080/01431161.2013.820367.

Cracknell, Arthur Philip, Kasturi Devi Kanniah, Kian Pang Tan, and Lei Wang. 2015. "Towards the Development of a Regional Version of MOD17 for the Determination of Gross and Net Primary Productivity of Oil Palm Trees." *International Journal of Remote Sensing* 36 (1): 262–289.

doi:10.1080/01431161.2014.995278.

Daughtry, C. 2000. "Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance." *Remote Sensing of Environment* 74 (2): 229–239. doi:10.1016/S0034-4257(00)00113-9.

Dechaume-Moncharmont, Francois-Xavier, Karine Monceau, and Cezilly Frank. 2011. "Sexing Birds Using Discriminant Function Analysis: A Critical Appraisal." *The Auk* 128 (1): 78–86. doi:10.1525/auk.2011.10129.

Delalieux, S, B Somers, S Hereijgers, W W Verstraeten, W Keulemans, and P Coppin. 2008. "A near-Infrared Narrow-Waveband Ratio to Determine Leaf Area Index in Orchards." *Remote Sensing of Environment*. doi:10.1016/j.rse.2008.05.003.

DigitalGlobe. 2014. "WorldView-3." *Data Sheet*, 1–2. doi:www.digitalglobe.com.

Ditjenbun. 2015. *Tree Crop Estate Statistics of Indonesia 2014-2016*. *Tree Crop Estate Statistics of Indonesia 2014-2016*. Ditjenbun. http://ditjenbun.pertanian.go.id/tinymcpuk/gambar/file/statistik/2016/SAWIT 2014-2016.pdf.

Du, Peijun, Pei Liu, Junshi Xia, Li Feng, Sicong Liu, Kun Tan, and Liang Cheng. 2014. "Remote Sensing Image Interpretation for Urban Environment Analysis: Methods, System and Examples." *Remote Sensing* 6: 9458–9474. doi:10.3390/rs6109458.

El-telbany, Mohammed E, and Mahmoud Warda. 2016. "An Empirical Comparison of Tree-Based Learning Algorithms : An Egyptian Rice Diseases Classification Case Study." *International Journal of Advanced Research in Artificial Intelligence* 5 (1): 22–26. www.ijarai.thesai.org.

Embrandiri, Asha, Mahamad H Ibrahim, and Rajeev P Singh. 2013. "Palm Oil Mill Wastes Utilization; Sustainability in the Malaysian Context." *International Journal of Scientific and Research Publications* 3 (1): 2250–3153. doi:2250-3153.

ENVI. 2014. "ENVI Classic Help." Boulder, USA: Exelis Visual Information

Solutions, Inc. doi:Exelis Visual Information Solutions, Inc.

Fairhurst, T, and E Mutert. 1999. "Interpretation and Management of Oil Palm Leaf Analysis Data." *Better Crops International* 13, No 1: 48–51.

Fassnacht, Fabian Ewald, Daniel Mangold, Jannika Schäfer, Markus Immitzer, Teja Kattenborn, Barbara Koch, and Hooman Latifi. 2017. "Estimating Stand Density, Biomass and Tree Species from Very High Resolution Stereo-Imagery – towards an All-in-One Sensor for Forestry Applications?" *Forestry: An International Journal of Forest Research*, 1–19. doi:10.1093/forestry/cpx014.

Figueroa, Rosa L, Qing Zeng-Treitler, Sasikiran Kandula, and Long H Ngo. 2012. "Predicting Sample Size Required for Classification Performance." *BMC Medical Informatics and Decision Making* 12: 8. doi:10.1186/1472-6947-12-8.

Foster, H L, and Chang Kwong Choong. 1976. "Palm Leaf Nutrient Levels." *MARDI Res. Bull.* 5, 2: 74–90.

Furlanello, C, M Neteler, S Merler, S Menegon, S Fontanari, A Donini, A Rizzoli, and C Chemini. 2003. "GIS and the Random Forest Predictor : Integration in R for Tick-Borne Disease Risk Assessment." In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, edited by Kurt Hornik, Friedrich Leisch, and Achim Zeileis, 1–11. Vienna, Austria: Technische Universitat Wien. http://www.r-project.org/conferences/DSC-2003/.

Gall, Juergen, Nima Razavi, and Luc Van Gool. 2012. "An Introduction to Random Forests for Multi-Class Object Detection." *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 7474 LNCS: 243–263. doi:10.1007/978-3-642-34091-8_11.

Garczarek, Maria Ursula. 2002. "Classification Rules in Standardized Partition Spaces." University of Dortmund. http://hdl.handle.net/2003/2789%0D.

Gitelson, Anatoly A., Yoram J. Kaufman, and Mark N. Merzlyak. 1996. "Use of a Green Channel in Remote Sensing of Global Vegetation from EOS-MODIS."

Remote Sensing of Environment 58 (3): 289–298. doi:10.1016/S0034-4257(96)00072-7.

Goh, K. J., P.S Chew, and C.B. Teo. 1994. "Maximising and Maintraining Oil Palm Yields on Comercial Scale in Malaysia. In: Management for Enhanced Profitability in Plantations (Ed. by Chee K.H.)." *Incorp. Soc. Planters*, 121–141.

Guijarro, M., G. Pajares, I. Riomoros, P. J. Herrera, X. P. Burgos-Artizzu, and A. Ribeiro. 2011. "Automatic Segmentation of Relevant Textures in Agricultural Images." *Computers and Electronics in Agriculture* 75 (1). Elsevier B.V.: 75–83. doi:10.1016/j.compag.2010.09.013.

H.R. Von Uexkull and T.H. Fairhurst. 1999. "Some Nutritional Disorder in Oil Palm." *Better Crop International* 13, No 1: 16–21.

Haboudane, Driss, John R. Miller, Elizabeth Pattey, Pablo J. Zarco-Tejada, and Ian B. Strachan. 2004. "Hyperspectral Vegetation Indices and Novel Algorithms for Predicting Green LAI of Crop Canopies: Modeling and Validation in the Context of Precision Agriculture." *Remote Sensing of Environment* 90: 337–352. doi:10.1016/j.rse.2003.12.013.

Haboudane, Driss, John R. Miller, Nicolas Tremblay, Pablo J. Zarco-Tejada, and Louise Dextraze. 2002. "Integrated Narrow-Band Vegetation Indices for Prediction of Crop Chlorophyll Content for Application to Precision Agriculture." *Remote Sensing of Environment* 81 (2–3): 416–426. doi:10.1016/S0034-4257(02)00018-4.

Haryana, Arif, Jarot Indarto, and Noor Avianto. 2010. *NASKAH KEBIJAKAN (POLICY PAPER): Kebijakan Dan Strategi Dalam Meningkatkan Nilai Tambah Dan Daya Saing Kelapa Sawit Indonesia Secara Berkelanjutan Dan Berkeadilan*. Edited by Wahyuningsih Darajati. Direktorat Pangan dan Pertanian, Badan Perencanaan Pembangunan Nasional (BAPPENAS). http://old.bappenas.go.id/files/1813/5182/6723/naskah-kebijakan-final-sawit__20110211150840__4.pdf.

Hatchell, D.C. 1999. "Analytical Spectral Devices, Inc. (ASD) Technical Guide. 3rd Ed." *Inc. ASD Technical Guide*.

Hayashi, Kiichiro. 2007. "Environmental Impact of Palm Oil Industry in Indonesia." *Proceedings of International Symposium on EcoTopia Science* 7: 646–651.

Hoffer, R M. 1984. "Remote Sensing of the Distribution and Structure of Vegetation-Chapter 5." In *The Role of Terrestrial Vegetation in the Global Carbon Cycle: Measurement by Remote Sensing*, edited by G.M. Woodwell, 131–159.

Hoffer, Roger M. 1971. *The Importance of "Ground Truth" Data in Remote Sensing*. LARS Print 120371. Sao Jose' dos Campos, Sao Paulo., Brazil. doi:Paper presented at the "United Nations Panel Meeting on the Establishment and Implementation of Research Programs in Remote Sensing. Held at the National Institute for Space Research. November 29 - December 10, 1971.

Hosoi, Fumiki, Hiroaki Matsugami, Kenichi Watanuki, Yo Shimizu, and Kenji Omasa. 2012. "Accurate Detection of Tree Apexes in Coniferous Canopies from Airborne Scanning Light Detection and Ranging Images Based on Crown-Extraction Filtering." *Journal of Applied Remote Sensing* 6 (1): 63502. doi:10.1117/1.JRS.6.063502.

Hosseinioun, Nargess. 2016. "Forecasting Outlier Occurrence in Stock Market Time Series Based on Wavelet Transform and Adaptive ELM Algorithm." *Journal of Mathematical Finance* 6 (February): 127–133.

Hushiarian, Roozbeh, Nor Azah Yusof, and Sabo Wada Dutse. 2013. "Detection and Control of Ganoderma Boninense: Strategies and Perspectives." *SpringerPlus* 2 (1): 555. doi:10.1186/2193-1801-2-555.

Huth, Neil I, Murom Banabas, Paul N Nelson, and Michael Webb. 2014. "Development of an Oil Palm Cropping Systems Model: Lessons Learned and Future Directions." *Environmental Modelling & Software* 62 (0). Elsevier Ltd: 411–419. doi:http://dx.doi.org/10.1016/j.envsoft.2014.06.021.

Idris, A. S., D Ariffin, T. R. Swinbuane, and T. A. Watt. 2000. "The Identity of Ganoderma Species Responsible for BSR Disease of Oil Palm in Malaysia - Pathogenicity Test." *MPOB Information Series* MPOB TT No: 1–4. doi:http://www.mpob.gov.my/tot/tt77b.pdf.

Immitzer, Markus, and Clement Atzberger. 2014. "Early Detection of Bark Beetle Infestation in Norway Spruce (Picea Abies, L.) Using WorldView-2 Data." *Photogrammetrie - Fernerkundung - Geoinformation*, no. 5: 351–367. doi:10.1127/1432-8364/2014/0229.

Immitzer, Markus, Clement Atzberger, and Tatjana Koukal. 2012. "Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band worldView-2 Satellite Data." *Remote Sensing* 4 (9): 2661–2693. doi:10.3390/rs4092661.

Jhonnerie, Romie, Vincentius P. Siregar, Bisman Nababan, Lilik Budi Prasetyo, and Sam Wouthuyzen. 2015. "Random Forest Classification for Mangrove Land Cover Mapping Using Landsat 5 TM and Alos Palsar Imageries." *Procedia Environmental Sciences* 24. Elsevier B.V.: 215–221. doi:10.1016/j.proenv.2015.03.028.

Joffre, R., D. Gillon, P. Dardenne, R. Agneessens, and R. Biston. 1992. "The Use of near-Infrared Reflectance Spectroscopy in Litter Decomposition Studies." *Annales Des Sciences Forestieres* 49 (5): 481–488. doi:10.1051/forest:19920504.

Johnson, Brian, Ryutaro Tateishi, and Nguyen Hoan. 2012. "Satellite Image Pansharpening Using a Hybrid Approach for Object-Based Image Analysis." *ISPRS International Journal of Geo-Information* 1 (3): 228–241. doi:10.3390/ijgi1030228.

Joy, Divya, and Sreekumar K. 2015. "Bridging Decision Tree Data Mining Model to the Automated Knowledge Base for Rice Plant Agriculture Expert System." *International Journal of Computer Applications* 121 (24): 975–8887.

Jusoff, Kamaruzaman, and Mubeena Pathan. 2009. "Mapping of Individual Oil Palm Trees Using Airborne Hyperspectral Sensing : An Overview." *Applied Physics Research* 1: 15–30.

Karatzoglou, Alexandros, David Meyer, and Kurt Hornik. 2006. "Support Vector Machines in R." *Journal of Statistical Software* 15 (9): 28. http://www.jstatsoft.org/v15/i09/paper.

Karatzoglou, Alexandros, Alex Smola, and Kurt Hornik. 2015. "Package ' Kernlab .'" https://cran.r-project.org/web/packages/kernlab/index.html. doi:https://cran.r-project.org/web/packages/kernlab/index.html.

Kattenborn, T., M. Sperlich, K. Bataua, and B. Koch. 2014. "Automatic Single Palm Tree Detection in Plantations Using UAV-Based Photogrammetric Point Clouds." *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* 40 (3): 139–144. doi:10.5194/isprsarchives-XL-3-139-2014.

Kee, N. S. 1972. *The Oil Palm, Its Culture, Manuring and Utilisation*. Switzerland: International Potash Institute. http://www.ipipotash.org/udocs/the_oil_palm-its_culture-manuring_and_utilisation.pdf.

Kelly, Ryan. 2014a. "Support Vector Machines." *RPubs*. http://rpubs.com/ryankelly/svm.

Kelly, Ryan. 2014b. "Bagging , Random Forests , Boosting." http://www.rmdk.ca/boosting_forests_bagging.html.

Khalilia, Mohammed, Sounak Chakraborty, and Mihail Popescu. 2011. "Predicting Disease Risks from Highly Imbalanced Data Using Random Forest." *BMC Medical Informatics and Decision Making* 11 (1). BioMed Central Ltd: 51. doi:10.1186/1472-6947-11-51.

Kiama, J W, V Raman, and T H H Patrick. 2014. "Low-Cost RFID-Based Palm Oil Monitoring System (PMS): First Prototype." *IOP Conference Series: Earth and*

Environmental Science 18: 12054. doi:10.1088/1755-1315/18/1/012054.

Kononenko, Igor. 2001. "Machine Learning for Medical Diagnosis: History, State of the Art and Perspective." *Artificial Intelligence in Medicine* 23 (1): 89–109. doi:10.1016/S0933-3657(01)00077-X.

Korom, Alexius, M-H Phua, Y Hirata, and T Matsuura. 2014. "Extracting Oil Palm Crown from WorldView-2 Satellite Image." *IOP Conference Series: Earth and Environmental Science* 18: 6. doi:10.1088/1755-1315/18/1/012044.

Kuester, Michele. 2016. "Radiometric Use of WorldView-3 Imagery." *Technical Note*, 12. https://dg-cms-uploads-production.s3.amazonaws.com/uploads/document/file/207/Radiometric_Use_of_WorldView-3_v2.pdf.

Kuhn, Max. 2008. "Building Predictive Models in R Using the Caret Package." *Journal Of Statistical Software* 28 (5): 1–26. doi:10.1053/j.sodo.2009.03.002.

Kuhn, Max. 2015a. "A Short Introduction to the Caret Package, R Technical Report," 1–10. https://cran.r-project.org/web/packages/caret/vignettes/caret.pdf.

Kuhn, Max. 2015b. "Caret." *The Comprehensive R Archive Network (CRAN)*. https://cran.r-project.org/web/packages/caret/caret.pdf.

Kuhn, Max. 2016. "'caret - Classification and Regression Training' R Package." https://github.com/topepo/caret/. https://github.com/topepo/caret/.

Lelong, Camille C D, Jean Michel Roger, Simon Brégand, Fabrice Dubertret, Mathieu Lanore, Nurul a. Sitorus, Doni a. Raharjo, and Jean Pierre Caliman. 2010. "Evaluation of Oil-Palm Fungal Disease Infestation with Canopy Hyperspectral Reflectance Data." *Sensors* 10: 734–747. doi:10.3390/s100100734.

Li, Weizhi, Weirong Mo, Xu Zhang, John J Squiers, Yang Lu, Eric W Sellke, and Wensheng Fan. 2015. "Outlier Detection and Removal Improves Accuracy of Machine Learning Approach to Burn Diagnostic Imaging." *Journal of Biomedical Optics* 20 (December): 1–10. doi:10.1117/1.JBO.20.12.XXXXXX.

Li, Xiao-Li, and Yong He. 2008. "Chlorophyll Assessment and Sensitive Wavelength Exploration for Tea (Camellia Sinensis) Based on Reflectance Spectral Characteristics." *HORTSCIENCE* 43 (5): 1586–1591.

Liaghat, Shohreh, Reza Ehsani, Shattri Mansor, Helmi Z.M. Shafri, Sariah Meon, Sindhuja Sankaran, and Siti H.M.N. Azam. 2014. "Early Detection of Basal Stem Rot Disease (Ganoderma) in Oil Palms Based on Hyperspectral Reflectance Data Using Pattern Recognition Algorithms." *International Journal of Remote Sensing* 35 (10): 3427–3439. doi:10.1080/01431161.2014.903353.

Liaw, Andy, and Matthew Wiener. 2002. "Classification and Regression by {randomForest}." *R News* 2 (3): 18–22. http://cran.r-project.org/doc/Rnews/.

Lin, Chinsu, Chao-Cheng Wu, Khongor Tsogt, Yen-Chieh Ouyang, and Chein-I Chang. 2015. "Effects of Atmospheric Correction and Pansharpening on LULC Classification Accuracy Using WorldView-2 Imagery." *Information Processing in Agriculture* 2 (1). China Agricultural University: 25–36. doi:10.1016/j.inpa.2015.01.003.

Lin, L, and L D Torbeck. 1998. "Coefficient of Accuracy and Concordance Correlation Coefficient: New Statistics for Methods Comparison." *PDA J Pharm Sci Technol* 52 (2): 55–59.

Liu, Desheng, Maggi Kelly, and Peng Gong. 2006. "A Spatial-Temporal Approach to Monitoring Forest Disease Spread Using Multi-Temporal High Spatial Resolution Imagery." *Remote Sensing of Environment* 101 (2): 167–180. doi:10.1016/j.rse.2005.12.012.

López-López, Manuel, Rocío Calderón, Victoria González-Dugo, Pablo Zarco-Tejada, and Elías Fereres. 2016. "Early Detection and Quantification of Almond Red Leaf Blotch Using High-Resolution Hyperspectral and Thermal Imagery." *Remote Sensing* 8 (4). Multidisciplinary Digital Publishing Institute: 276. doi:10.3390/rs8040276.

Lowe, Barrett, and Arun Kulkarni. 2015. "Multispectral Image Analysis Using Random Forest." *International Journal on Soft Computing* 6 (1): 1–14. doi:10.5121/ijsc.2015.6101.

Mahajan, G. R., R. N. Sahoo, R. N. Pandey, V. K. Gupta, and Dinesh Kumar. 2014. "Using Hyperspectral Remote Sensing Techniques to Monitor Nitrogen, Phosphorus, Sulphur and Potassium in Wheat (Triticum Aestivum L.)." *Precision Agriculture* 15 (5): 499–522. doi:10.1007/s11119-014-9348-7.

Mayulu, Hamdi. 2014. "The Nutrient Digestibility of Locally Sheep Fed with Amofer Palm Oil Byproduct-Based Complete Feed." *International Journal of Science and Engineering (IJSE)* 7 (2): 106–111. doi:10.12777/ijse.7.2.106-111.

Mba, Ogan I., Marie-Josée Dumont, and Michael Ngadi. 2015. "Palm Oil: Processing, Characterization and Utilization in the Food Industry – A Review." *Food Bioscience* 10 (JANUARY): 26–41. doi:10.1016/j.fbio.2015.01.003.

Mehra, Lucky K, Christina Cowger, Kevin Gross, and Peter S Ojiambo. 2016. "Predicting Pre-Planting Risk of Stagonospora Nodorum Blotch in Winter Wheat Using Machine Learning Models." *Frontiers in Plant Science* 7. Frontiers Media SA: 390. doi:10.3389/fpls.2016.00390.

Memon, N, K S Memon, and Zia-ul Hassan. 2005. "Plant Analysis as a Diagnostic Tool for Evaluating Nutritional Requirements of Bananas." *International Journal of Agriculture & Biology* 7 (5): 824–831.

Menesatti, P., F. Antonucci, F. Pallottino, G. Roccuzzo, M. Allegra, F. Stagno, and F. Intrigliolo. 2010. "Estimation of Plant Nutritional Status by Vis-NIR Spectrophotometric Analysis on Orange Leaves [Citrus Sinensis (L) Osbeck Cv Tarocco]." *Biosystems Engineering* 105 (4): 448–454. doi:10.1016/j.biosystemseng.2010.01.003.

Meys, Joris. 2013. "Stepwise Regression Using P-Values to Drop Variables with Nonsignificant P-Values." http://stackoverflow.com/questions/3701170/

stepwise-regression-using-p-values-to-drop-variables-with-nonsignificant-p-value.

Min, M, and WS Lee. 2005. "Determination of Significant Wavelengths and Prediction of Nitrogen Content for Citrus." *Transactions of the ASAE* 48 (1998): 455–461. doi:10.13031/2013.18308.

Morcillo, F, D Cros, N Billotte, G-F Ngando-Ebongue, H Domonhédo, M Pizot, T Cuéllar, et al. 2013. "Improving Palm Oil Quality through Identification and Mapping of the Lipase Gene Causing Oil Deterioration." *Nature Communications* 4: 2160. doi:10.1038/ncomms3160.

Morel, Alexandra C., Joshua B. Fisher, and Yadvinder Malhi. 2012. "Evaluating the Potential to Monitor Aboveground Biomass in Forest and Oil Palm in Sabah, Malaysia, for 2000–2008 with Landsat ETM+ and ALOS-PALSAR." *International Journal of Remote Sensing* 33 (11): 3614–3639. doi:10.1080/01431161.2011.631949.

Muharam, Farrah Melissa, Stephen J. Maas, Kevin F. Bronson, and Tina Delahunty. 2015. "Estimating Cotton Nitrogen Nutrition Status Using Leaf Greenness and Ground Cover Information." *Remote Sensing* 7 (6): 7007–7028. doi:10.3390/rs70607007.

Mukaka, M. M. 2012. "Statistics Corner: A Guide to Appropriate Use of Correlation Coefficient in Medical Research." *Malawi Medical Journal* 24 (3): 69–71.

Mukherjee, S, P Tamayo, S Rogers, R Rifkin, A Engle, C Campbell, T R Golub, and J P Mesirov. 2003. "Estimating Dataset Size Requirements for Classifying DNA Microarray Data." *Journal of Computational Biology : A Journal of Computational Molecular Cell Biology* 10 (2): 119–142. doi:10.1089/106652703321825928.

Müller-Linow, Mark, Francisco Pinto-Espinosa, Hanno Scharr, and Uwe Rascher. 2015. "The Leaf Angle Distribution of Natural Plant Populations: Assessing the Canopy with a Novel Software Tool." *Plant Methods* 11 (1): 11. doi:10.1186/s13007-015-0052-z.

Munoz-Huerta, Rafael F., Ramon G. Guevara-Gonzalez, Luis M. Contreras-Medina, Irineo Torres-Pacheco, Juan Prado-Olivarez, and Rosalia V. Ocampo-Velazquez. 2013. "A Review of Methods for Sensing the Nitrogen Status in Plants: Advantages, Disadvantages and Recent Advances." *Sensors (Basel, Switzerland)* 13 (8): 10823–10843. doi:10.3390/s130810823.

Murphy, Denis J. 2009. "Oil Palm: Future Prospects for Yield and Quality Improvements." *Lipid Technology* 21 (11–12): 257–260. doi:10.1002/lite.200900067.

Naher, Laila, Umi Kalsom Yusuf, Ahmad Ismail, Soon Guan Tan, and M M A Mondal. 2013. "Review Article Ecological Status of *Ganoderma* and Basal Stem Rot Disease of Oil Palms (Elaeis Guineensis Jacq.)." *Australian Journal of Crop Science* 7 (11): 1723–1727.

NCEAS. 2014. "Overlay Point on Satellite Image/extract Pixel Values. National Center for Ecological Analysis and Synthesis." https://www.nceas.ucsb.edu/scicomp/usecases/ExtractSatelliteDataAtPoints.

Neon. 2014. "Creating A Square Buffer Around a Plot Centroid in R." http://neondataskills.org/working-with-field-data/Field-Data-Polygons-From-Centroids/.

Ng, Sk. 2002. "Nutrition and Nutrient Management of Oil Palm-New Thrust for the Future Perspective." *In: Pasricha, N.S., Bansal, S.K. (Eds.), Potassium for Sustainable Crop Production. International Symposium on Role of Potassium in India. Potash Research Institute of India, and International Potash Institute, New Delhi.*, 415–429. http://www.ipipotash.org/udocs/Nutrition and Nutrient Management of the Oil Palm.pdf.

Nur Sabrina, A. A., M. Sariah, and A. R. Zaharah. 2012. "Suppression of Basal Stem Rot Disease Progress in Oil Palm (Elaeis Guineensis) after Copper and Calcium Supplementation." *Pertanika J. Trop. Agric. Sci.* 35 (S): 13–24.

http://www.pertanika.upm.edu.my/Pertanika PAPERS/JTAS Vol. 35 (S) Dec. 2012/02_page 13-24.pdf.

OECD/Food and Agriculture Organization of the United Nations. 2015. *OECD-FAO Agricultural Outlook 2015-2024*. OECD Publishing, Paris. doi:http://dx.doi.org/10.1787/agr_outlook-2015-en.

Osborne, Jason W., and Amy Overbay. 2004. "The Power of Outliers (and Why Researchers Should Always Check for Them)." *Practical Assessment, Research & Evaluation* 9 (6): 1–8. doi:10.1017/CBO9781107415324.004.

Özyigit, Y., and M. Bilgen. 2013. "Use of Spectral Reflectance Values for Determining Nitrogen, Phosphorus, and Potassium Contents of Rangeland Plants." *Journal of Agricultural Science and Technology* 15 (SUPPL): 1537–1545.

Paterson, R. R M. 2007. "Ganoderma Disease of Oil Palm-A White Rot Perspective Necessary for Integrated Control." *Crop Protection* 26 (9): 1369–1376. doi:10.1016/j.cropro.2006.11.009.

Pimstein, Agustin, Arnon Karnieli, Surinder K. Bansal, and David J. Bonfil. 2011. "Exploring Remotely Sensed Technologies for Monitoring Wheat Potassium and Phosphorus Using Field Spectroscopy." *Field Crops Research* 121 (1): 125–135. doi:10.1016/j.fcr.2010.12.001.

Pontius, Robert Gilmore, and Marco Millones. 2011. "Death to Kappa: Birth of Quantity Disagreement and Allocation Disagreement for Accuracy Assessment." *International Journal of Remote Sensing* 32 (15): 4407–4429. doi:10.1080/01431161.2011.552923.

Prasad, Anantha M., Louis R. Iverson, and Andy Liaw. 2006. "Newer Classification and Regression Tree Techniques: Bagging and Random Forests for Ecological Prediction." *Ecosystems* 9 (2): 181–199. doi:10.1007/s10021-005-0054-1.

Price, Tabitha, and Julia Fischer. 2014. "Meaning Attribution in the West African Green Monkey: Influence of Call Type and Context." *Animal Cognition* 17 (2):

277–286. doi:10.1007/s10071-013-0660-9.

Pritts, Marvin, and Cathy Heidenreich. 2012. "Leaf and Soil Analysis Special Edition." *New York Berry News. Cornell University. College of Agriculture and Life Sciences.* 11 (8).

Priwiratama, Hari, Agus Eko Prasetyo, and Agus Susanto. 2014. "Pengendalian Penyakit Busuk Pangkal Batang Kelapa Sawit Secara Kultur Teknis." *Jurnal Fitopatologi Indonesia* 10 (51): 1–7. doi:10.14692/jfi.10.1.1.

Priwiratama, Hari, and Agus Susanto. 2014. "Utilization of Fungi for the Biological Control of Insect Pests and Ganoderma Disease in the Indonesian Oil Palm Industry." *Journal of Agricultural Science and Technology* 4: 103–111.

Pujari, Jagadeesh D., Rajesh Yakkundimath, and Abdulmunaf S. Byadgi. 2015. "Image Processing Based Detection of Fungal Diseases in Plants." *Procedia Computer Science* 46 (Icict 2014): 1802–1808. doi:10.1016/j.procs.2015.02.137.

Rankie, Ian, and T.H. Fairhurst. 1999. "Management of Phosphorus, Potassium and Magnesium in Mature Oil Palm." *Better Crops International* 13, No 1: 10–15.

Ripley, Brian, Bill Venables, Douglas M Bates, Kurt Homik, Albrecht Gebhardt, and David Firth. 2015. "Package ' MASS ,'" 167.

Roever, Christian, Nils Raabe, Karsten Luebke, Uwe Ligges, Gero Szepannek, and Marc Zentgraf. 2015. "'klaR-Classification and visualization.' R Package." https://cran.r-project.org/web/packages/klaR/index.html. http://www.statistik.tu-dortmund.de.

Römer, Christoph, Kathrin Bürling, Mauricio Hunsche, Till Rumpf, Georg Noga, and Lutz Plümer. 2011. "Robust Fitting of Fluorescence Spectra for Pre-Symptomatic Wheat Leaf Rust Detection with Support Vector Machines." *Computers and Electronics in Agriculture* 79 (2): 180–188. doi:10.1016/j.compag.2011.09.011.

Rothman, Jessica M., Colin A. Chapman, Julie L. Hansen, Debbie J R Cherney, and Alice N. Pell. 2009. "Rapid Assessment of the Nutritional Value of Foods Eaten by Mountain Gorillas: Applying near-Infrared Reflectance Spectroscopy to Primatology." *International Journal of Primatology* 30 (5): 729–742. doi:10.1007/s10764-009-9372-z.

Rouse, J W, R H Haas, J A Schell, D W Deering, and J C Harlan. 1974. *Monitoring the Vernal Advancements and Retrogradation of Natural Vegetation*. *Texas A&M University. Remote Sensing Center. College Station*. Texas.

RStudio. 2015. "Integrated Development Environment (IDE) for R Code." Boston, MA: RStudio, Inc. https://www.rstudio.com/products/rstudio/.

Rumpf, T., A. K. Mahlein, U. Steiner, E. C. Oerke, H. W. Dehne, and L. Plümer. 2010. "Early Detection and Classification of Plant Diseases with Support Vector Machines Based on Hyperspectral Reflectance." *Computers and Electronics in Agriculture* 74 (1). Elsevier B.V.: 91–99. doi:10.1016/j.compag.2010.06.009.

Santoso, Heri, Totok Gunawan, Retnadi Heru Jatmiko, Witjaksana Darmosarkoro, and Budiman Minasny. 2011. "Mapping and Identifying Basal Stem Rot Disease in Oil Palms in North Sumatra with QuickBird Imagery." *Precision Agriculture* 12 (2): 233–248. doi:10.1007/s11119-010-9172-7.

Santoso, Heri, Hiroshi Tani, and Xiufeng Wang. 2016. "A Simple Method for Detection and Counting of Oil Palm Trees Using High Resolution Multispectral Satellite Imagery." *International Journal of Remote Sensing* 37 (21): 5122–5134. doi:10.1080/01431161.2016.1226527.

Santoso, Heri, Hiroshi Tani, and Xiufeng Wang. 2017. "Random Forest Classification Model of Basal Stem Rot Disease Caused by Ganoderma Boninense in Oil Palm Plantations." *International Journal of Remote Sensing* 38 (16). Taylor & Francis: 4683–4699. doi:10.1080/01431161.2017.1331474.

Sarmidi, Mohammad R, Hesham A El Enshasy, and Mariani Abdul Hamid. 2009.

"Oil Palm : The Rich Mine for Pharma , Food , Feed and Fuel Industries Chemical Engineering Pilot Plant (CEPP), Faculty of Chemical and Natural Resource Engineering , Department of Bioprocess Development , GEBRI , Mubarak City for Scientific Research ,." *American-Eurasian J. Agric. & Environ. Sci.* 5 (6): 767–776.

Serusi, Simona, W Dr. Verhoef, J Dr. Clevers, and M Dr. Schlerf. 2010. "Estimation of Leaves Nutrient Content in Seagrass Using Spectral Data The Case of Halodule Uninervis." *Master of Thesis. International Institute for Geo-Information and Earth Observation. Enschede, The Netherlands*. Master of Thesis. ITC. International Institute for Geo-Information and Earth Observation. Enschede, The Netherlands.

Shafri, H.Z.M., and Nasrulhapiza Hamdan. 2009. "Hyperspectral Imagery for Mapping Disease Infection in Oil Palm Plantation Using Vegetation Indices and Red Edge Techniques." *American Journal of Applied Sciences* 6 (6): 1031–1035. doi:10.3844/ajassp.2009.1031.1035.

Shafri, Helmi Z. M., Nasrulhapiza Hamdan, and M. Iqbal Saripan. 2011. "Semi-Automatic Detection and Counting of Oil Palm Trees from High Spatial Resolution Airborne Imagery." *International Journal of Remote Sensing* 32 (8): 2095–2115. doi:10.1080/01431161003662928.

Shafri, Helmi Z.M., M.I. Anuar, Idris A. Seman, and Nisfariza M Noor. 2011. "Spectral Discrimination of Healthy and Ganoderma-Infected Oil Palms from Hyperspectral Data." *International Journal of Remote Sensing* 32 (22): 7111–7129. doi:10.1080/01431161.2010.519003.

Shafri, Helmi Zulhaidi M., Nasrulhapiza Hamdan, and Mohamad Izzuddin Anuar. 2012. "Detection of Stressed Oil Palms from an Airborne Sensor Using Optimized Spectral Indices." *International Journal of Remote Sensing* 33 (14): 4293–4311. doi:10.1080/01431161.2011.619208.

Short, N.M. 1982. *The Landsat Tutorial Workbook-Basics of Satellite Remote Sensing*. Washington, D.C.: National Aeronautics and Space Administration, Scientific and Technical Information Branch.

Sonobe, Rei, Hiroshi Tani, Xiufeng Wang, Nobuyuki Kobayashi, and Hideki Shimamura. 2014a. "Random Forest Classification of Crop Type Using Multi-Temporal TerraSAR-X Dual-Polarimetric Data." *Remote Sensing Letters* 5 (2): 157–164. doi:10.1080/2150704X.2014.889863.

Sonobe, Rei, Hiroshi Tani, Xiufeng Wang, Nobuyuki Kobayashi, and Hideki Shimamura. 2014b. "Parameter Tuning in the Support Vector Machine and Random Forest and Their Performances in Cross- and Same-Year Crop Classification Using TerraSAR-X." *International Journal of Remote Sensing* 35 (23): 7898–7909. doi:10.1080/01431161.2014.978038.

Sonobe, Rei, and Quan Wang. 2016. "Assessing the Xanthophyll Cycle in Natural Beech Leaves with Hyperspectral Reflectance." *Functional Plant Biology*, 438–447. doi:10.1071/FP15325.

Srestasathiern, Panu, and Preesan Rakwatin. 2014. "Oil Palm Tree Detection with High Resolution Multi-Spectral Satellite Imagery." *Remote Sensing* 6 (10): 9749–9774. doi:10.3390/rs6109749.

Stackoverflow. 2014. "How Do I Extract Raster Values from Polygon Data Then Join into Spatial Data Frame ?" http://stackoverflow.com/questions/22333473/how-do-i-extract-raster-values-from-polygon-data-then-join-into-spatial-data-fra/2#2.

Starks, Patrick J., Duli Zhao, William A. Phillips, and Samuel W. Coleman. 2006. "Development of Canopy Reflectance Algorithms for Real-Time Prediction of Bermudagrass Pasture Biomass and Nutritive Values." *Crop Science* 46 (2): 927–934. doi:10.2135/cropsci2005.0258.

Stefanski, J, B Mack, and B Waske. 2013. "Optimization of Object-Based Image

Analysis With Random Forests for Land Cover Mapping." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 6 (6): 2492–2504. doi:10.1109/JSTARS.2013.2253089.

Stein, Beth R., Valerie a. Thomas, Laura J. Lorentz, and Brian D. Strahm. 2014. "Predicting Macronutrient Concentrations from Loblolly Pine Leaf Reflectance across Local and Regional Scales." *GIScience & Remote Sensing* 51 (3): 269–287. doi:10.1080/15481603.2014.912875.

Subramaniam, Vijaya, Choo Yuen May, Halimah Muhammad, Zulkifli Hashim, Yew Ai Tan, and Puah Chiew Wei. 2010. "Life Cycle Assessment of the Production of Crude Palm Oil (Part 3)." *Journal of Oil Palm Research* 22 (DECEMBER): 895–903. http://jopr.mpob.gov.my/life-cycle-assessment-of-the-production-of-crude-palm-oil-part-3/?v=true.

Suhartono, Derwin, Wahyu Aditya, Miranty Lestari, and Muhammad Yasin. 2013. "Expert System in Detecting Coffee Plant Diseases." *International Journal of Electrical Energy* 1 (3): 156–162. doi:10.12720/ijoee.1.3.156-162.

Susanto, Agus, Agus Eko Prasetyo, Hari Priwiratama, Sri Wening, and Surianto Surianto. 2013. "Ganoderma Boninense Penyebab Penyakit Busuk Batang Atas Kelapa Sawit." *Jurnal Fitopatologi Indonesia* 9 (4): 123–126. doi:10.14692/jfi.9.4.123.

Susanto, Agus, Agus Eko Prasetyo, and S Wening. 2013. "Laju Infeksi Ganoderma Pada Empat Kelas Tekstur Tanah." *Jurnal Fitopatologi Indonesia* 9 (2): 39–46. doi:10.14692/jfi.9.2.39.

Szepannek, Gero, and Claus Weihs. 2006. *Local Modelling in Classification on Different Feature Subspaces*. Technical Report / Universität Dortmund, SFB 475 Komplexitätsreduktion in Multivariaten Datenstrukturen, No. 2006,08.

Tan, Kian Pang, Kasturi Devi Kanniah, and Arthur Philip Cracknell. 2013. "Use of UK-DMC 2 and ALOS PALSAR for Studying the Age of Oil Palm Trees in

Southern Peninsular Malaysia." *International Journal of Remote Sensing* 34 (20): 7424–7446. doi:10.1080/01431161.2013.822601.

Tan, Kian Pang, Kasturi Devi Kanniah, and Arthur Philip Cracknell. 2014. "On the Upstream Inputs into the MODIS Primary Productivity Products Using Biometric Data from Oil Palm Plantations." *International Journal of Remote Sensing* 35 (6): 2215–2246. doi:10.1080/01431161.2014.889865.

Tengoua, F. F., M. M. Hanafi, A. S. Idris, and S. R. Syed-Omar. 2015. "Screening for Optimum Concentrations of Boron, Copper and Manganese for the Growth of Three-Month Old Oil Palm Seedlings in Solution Culture." *Pertanika Journal of Tropical Agricultural Science* 38 (1): 113–126.

Therneau, Terry M, and Elizabeth J Atkinson. 2015. "An Introduction to Recursive Partitioning Using the RPART Routines," 1–62.

Turner, P.D. 1981. *Oil Palm Disease and Disorders*. Oxford: Oxford University Press.

Uexkull, H.R. von, T.H. H Fairhurst, and H R von Uexkull. 1991. *Fertilizing for High Yield and Quality the Oil Palm*. *IPI-Bulletin*. Vol. 12. Bern/Switzerland: International Potash Institute.

Viera, Anthony J., and Joanne M. Garrett. 2005. "Understanding Interobserver Agreement: The Kappa Statistic." *Family Medicine* 37 (5): 360–363. doi:Vol. 37, No. 5.

Warner, Timothy A., James B. McGraw, and Rick Landenberger. 2006. "Segmentation and Classification of High Resolution Imagery for Mapping Individual Species in a Closed Canopy, Deciduous Forest." *Science in China, Series E: Technological Sciences* 49 (SUPPL. 1): 128–139. doi:10.1007/s11431-006-8114-0.

Waser, Lars T, Meinrad Küchler, Kai Jütte, and Theresia Stampfer. 2014. "Evaluating the Potential of WorldView - 2 Data to Classify Tree Species and Different Levels of Ash Mortality." *Remote Sens* 6 (5): 4515–4545.

doi:10.3390/rs6054515.

Watts, J.D., and R.L. Lawrence. 2008. "Merging Random Forest Classifciation with an Object-Oriented Approach for Analysis of Agricultural Lands." *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XXXVII (Pat B7): 2006–2009.

Weihs, Claus, Uwe Ligges, Karsten Luebke, and Nils Raabe. 2005. "klaR Analyzing German Business Cycles." In *Data Analysis and Decision Support*, edited by D Baier, R Becker, and L Schmidt-Thieme, 335–343. Berlin/Heidelberg: Springer-Verlag. doi:10.1007/3-540-28397-8_36.

Wiesmeier, Martin, Frauke Barthold, Benjamin Blank, and Ingrid Kögel-Knabner. 2011. "Digital Mapping of Soil Organic Matter Stocks Using Random Forest Modeling in a Semi-Arid Steppe Ecosystem." *Plant and Soil* 340: 7–24. doi:10.1007/s11104-010-0425-z.

Wong-in, Teerawut, Tonphong Kaewkongka, Nagul Cooharojananone, and Rajalida Lipikorn. 2015. "Automatic Oil Palm Detection and Identification from Multi-Scale Clustering and Normalized Cross Correlation." In , 403–410. Springer Berlin Heidelberg. doi:10.1007/978-3-662-47200-2_43.

World Growth. 2011. *The Economic Benefit of Palm Oil to Indonesia. A Report by World Growth February 2011. World Growth Palm Oil Green Development Campaign*. http://worldgrowth.org/site/wp-content/uploads/2012/06/WG_Indonesian_Palm_Oil_Benefits_Report-2_11.pdf.

Wu, Xindong, Vipin Kumar, Quinlan J. Ross, Joydeep Ghosh, Qiang Yang, Hiroshi Motoda, Geoffrey J. McLachlan, et al. 2008. "Top 10 Algorithms in Data Mining." *Knowledge and Information Systems* 14 (1): 1–37. doi:10.1007/s10115-007-0114-2.

Wyczałek, Ireneusz, and Wyczałek Elżbieta. 2013. "Studies on Pansharpening and Object-Based Classification of Worldview-2 Multispectral Image." *Archives of*

Photogrammetry, Cartography and Remote Sensing Vol. Spec. (Monograph): 109–117.

Yao, Sheng, Guofen Wu, Mian Xing, Shuke Zhou, and Junwen Pu. 2010. "Determination of Lignin Content in Acacia Spp. Using near-Infrared Reflectance Spectroscopy." *BioResources* 5 (2): 556–562.

Zahari, Mw, Ar Alimon, and Hk Wong. 2012. "Utilization of Oil Palm Co-Products as Feeds for Livestock in Malaysia." In *Biofuel Co-Products as Livestock Feed – Opportunities and Challenges*, edited by Harinder P.S. Makkar, 243–302. Rome: FAO. http://pef.czu.cz/~linhart/zpravy/FAO_2012_Biofuel_coproducts_as_livestock_feed.pdf#page=262.