# MONITORING AFFORESTATION WITH VERY HIGH RESOLUTION SATELLITE IMAGERY IN KENYA

## A METHOD TO DIFFERENTIATE *MELIA VOLKENSII*

Eva Forceville
Student number: 01605685

Promotor:  prof. dr. ir. Frieke Vancoillie
Tutor:  ir. Jasper Feyen

Master's Dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of Master of Science in Bioscience Engineering: Forest and Nature Management

Academic year: 2021 – 2022

UNIVERSITEIT
GENT

Ghent, 10 June 2022
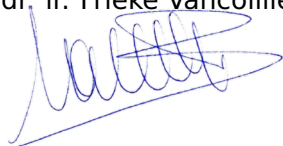

The author,


Eva Forceville


The promotor and tutor,


prof. dr. ir. Frieke Vancoillie                    ir. Jasper Feyen

# ACKNOWLEDGEMENT

After more than a year of blood, sweat, and tears, my Master's Dissertation is finally ready. First of all, I want to thank my promotor prof. dr. ir. Frieke Vancoillie and my tutor ir. Jasper Feyen for the excellent coaching and constructive feedback during the whole period. I also want to thank Jan and Samuel of Better Globe Forestry for the opportunity of working on this project and the confidence to bring it to the best possible end. Furthermore, I want to thank Halima and Kelvin for the data collection.

Besides, I want to thank my parents for all the opportunities and for being my biggest supporters, together with my sister, during my education at Ghent University. My boyfriend and my roomies deserve also a big thank you for always listening to all my Random Forest drama and dealing with my stress during this period. Last but not least, I want to thank Ellen Ghyselbrecht for the friendship starting from the first practical of Chemistry 1 up until now where we partially collaborated for this Master's Dissertation.

# ABSTRACT

Better Globe Forestry (BGF), a company active in East Africa, collaborates with local farmers to plant *Melia volkensii* trees in their fields in an agroforestry layout. These plantations appear to be highly successful as a result of which BGF is now in search of a monitoring tool to remotely supervise their efforts of afforestation over extended areas. In this study, it is researched if it is possible to differentiate *Melia volkensii* from its surroundings and other tree species at the individual tree crown level in the region of Kiambere.

Very high resolution (VHR) satellite imagery has proven to be promising for classification of tree species over broad spatial extents. Due to the high spectral variability associated with VHR data and the promising results compared to pixel-based approaches, an object-based approach was used. First, the image was segmented using the multiresolution segmentation algorithm after optimisation of the key parameters, namely image layer weights, shape (versus colour) factor, and scale. Then, the candidate tree crowns (CTC) were masked out from other land-cover classes using a rule-based classification, based on literature and expert knowledge, followed by a Random Forest classification with recursive feature elimination. Finally, the resulting CTC mask was further classified into *Melia volkensii* and other tree species with Random Forest. The classification results were not satisfactory (Cohen's Kappa coefficient of 0.118), whereby the obtained model is not suitable for large-scale application without some major adjustments such as the use of a VHR satellite with strategically located wavelength bands, the availability of ground truth data regarding the other tree species class, or the integration of height information.

# SAMENVATTING

Better Globe Forestry (BGF), een bedrijf dat actief is in Oost-Afrika, werkt samen met lokale boeren om *Melia volkensii* in hun velden aan te planten in een agroforestry lay-out. Deze aanplantingen blijken zeer succesvol te zijn. BGF is daardoor op zoek is naar een monitoringtool om deze inspanningen voor bebossingen over uitgestrekte gebieden vanop afstand op te volgen. In deze studie wordt de mogelijkheid onderzocht om *Melia volkensii* te onderscheiden van zijn omgeving en andere boomsoorten op het individuele boomkroonnniveau in de regio van Kiambere.

*Very high resolution* (VHR) satellietbeelden hebben bewezen veelbelovend te zijn voor de classificatie van boomsoorten in uitgestrekte gebieden. Vanwege de hoge spectrale variabiliteit geassocieerd met VHR data en de veelbelovende resultaten in vergelijking met pixel-gebaseerde benaderingen werd een object-gebaseerde benadering gebruikt. Eerst werd het beeld gesegmenteerd met behulp van het *multiresolution segmentation* algoritme, voorafgegaan door optimalisatie van de belangrijkste parameters, namelijk *image layer weights*, *shape (versus colour) factor* en *scale*. Daarna werden potentiële boomkronen onderscheiden van andere bodembedekkingsklassen met behulp van classificatieregels, gebaseerd op literatuur en vakkennis, gevolgd door een *Random Forest* classificatie met *recursive feature elimination*. Ten slotte werden de potentiële boomkronen verder geclassificeerd in *Melia volkensii* en andere boomsoorten met *Random Forest*. De classificatieresultaten waren ontoereikend (Cohen's Kappa coëfficiënt van 0,118) waardoor het verkregen model niet geschikt is voor grootschalige toepassing zonder enkele grote aanpassingen zoals het gebruik van een VHR-satelliet met strategisch geplaatste golflengtebanden, de beschikbaarheid van velddata omtrent andere boomsoorten of de integratie van hoogte informatie.

# **GLOSSSARY**

| | |
|---|---|
| 3D | three-dimensional |
| AM | before noon |
| Ang. 2nd moment | angular second moment |
| ASAS | African satellite augmentation system |
| B | blue |
| BGF | Better Globe Forestry |
| BSh | hot semi-arid climate zone |
| CAVIS | clouds, aerosols, vapors, ice & snow |
| CB | coastal blue |
| CHM | canopy height model |
| CI | confidence interval |
| CRS | coordinate reference system |
| CTC | candidate tree crown |
| CV | coefficient of variation |
| DSM | digital surface model |
| DTM | digital terrain model |
| ESP-tool | 'Estimate Scale Parameter' tool |
| EVI | enhanced vegetation index |
| FWHM | full width at half maximum |
| G | green |
| GLCM | grey-level co-occurrence matrix |
| GLDV | grey-level difference vector |
| GPS | global positioning system |
| LiDAR | light detection and ranging |
| Max. diff. | maximum difference |
| MDA | Mean Decreasing Accuracy |
| MDG | Mean Decrease Gini |
| mtry | number of input features used at each tree node in Random Forest |
| NDVI | normalised difference vegetation index |
| NDWI | normalised difference water index |
| NIR | near infrared |
| NTC | non-tree crown |
| nDSM | normalised digital surface model |
| ntree | number of decision trees for executing Random Forest |
| OBIA | object-based image analysis |
| OOB | out-of-bag |
| PAN | panchromatic |
| PC | principal component |
| PCA | principal component analysis |

| | |
|---|---|
| R | red |
| RE | red edge |
| ROC-LV | rate of change of local variance |
| ROI | region of interest |
| RVI | ratio vegetation index |
| SAVI | soil-adjusted vegetation index |
| SDI | shadow detector index |
| StdDev | standard deviation |
| SWIR | shortwave infrared |
| TOA | top-of-atmosphere |
| UTM | Universal Transverse Mercator |
| VHR | very high resolution |
| VIS | visible |
| VNIR | visible near infrared |
| WGS | worldwide geodetic system |
| Y | yellow |

# CONTENTS

# 1.  INTRODUCTION

Better Globe Forestry (BGF) is a company active in Kenya since 2004 with a vision to eradicate poverty and corruption in Africa. They finance a sustainable implementation by planting as many trees as there are people on this planet through 'Social Entrepreneurship'. Its expertise lies in commercial afforestation of drylands, particularly with *Melia volkensii* (locally known as mukau) in an agroforestry layout, to anticipate and mitigate effects of climate change in East Africa. In 2009, they expanded their activities to Uganda (Better Globe Forestry, nd).

BGF has pioneered the use of mukau, an indigenous mahogany species, on an industrial scale. Moreover, they introduced and popularised the planting of mukau in an authentic and productive agroforestry layout for tens of thousands of farmers. The afforestation can be divided into two categories. On the one hand classical tree plantations on leased land, and on the other hand cooperation with local farmers in an agroforestry layout with wide spacing to allow intercropping (Better Globe Forestry, nd). These plantations are widely applied and appear to be very successful. Due to the extended area, BGF is now in search of a sustainable monitoring tool to remotely supervise their efforts of afforestation.

In this explorative study, a prototype of the monitoring tool will be developed based on very high resolution (VHR) satellite imagery. The development consists of two parts:

- counting of individual tree crowns,

- differentiating *Melia volkensii* from other tree species and its surroundings.

The first part will be explored in the Master's Dissertation of Ellen Ghyselbrecht, whereas the second part is the scope of this Master's Dissertation. The final goal in the long-term is to combine both studies and make an estimation of the total number of *Melia volkensii* trees in each afforestation layout.

The main research question of this Master's Dissertation is: Is it possible to differentiate *Melia volkensii* accurately from (1) its surrounding, and (2) other tree species at the individual tree crown level in an agroforesty layout? Additionally, it is researched which combination of features (i.e. spectral, texture, or shape) enables the classification of *Melia volkensii*. The method to differentiate *Melia volkensii* is automated as much as possible to make it reproducible for other areas where BGF might be interested in. The objectives are likely to be challenging due to some limitations mainly associated with the research set-up. The study is based on Pléiades satellite imagery because of budgetary considerations although the spectral and spatial resolution are known to be suboptimal for tree species classification. Similarly, a single date classification is performed due to the high costs associated with multiple acquisitions. There was also the intention to acquire aerial imagery with the aim

of obtaining more accurate localisation of the ground truth data and performing a preliminary study regarding the integration of height information. However, this was not possible due to national policy. Therefore, the main goal of this study is to advice on the potential use of VHR optical satellite imagery for the differentiation of *Melia volkensii* in an agroforesty layout in the region of Kiambere. Based on the results, some adjustments will be recommended for optimisation of the methodology and a potential large-scale application.

In Chapter 2 of this Master's Dissertation, a literature review of tree species classification based on VHR imagery will be made. In Chapters 3 and 4, some of the discussed methods will be applied and adjusted for Pléiades imagery in the region of Kiambere, followed by a discussion of the results. Finally, some recommendations, based on the results and literature, (Chapter 5) for further development will be made that might improve the performance of the classification.

# 2. [TREE SPECIES CLASSIFICATION BASED ON VHR SATELLITE IMAGERY](#)

In 2007, the first commercial satellite with a spatial resolution of half a metre has been launched, namely WorldView-1 (Blaschke, 2010). Since that moment, many other very high resolution (VHR) satellites have followed (e.g. GeoEye-1, Pléiades, WorldView-2 and -3). The accessibility of commercial VHR satellites increases the amount of information on land cover at local to national scales, providing amazing details of the earth surface (Frauman and Wolff, 2005). The spatial resolution of these images varies from some tens of centimeters to 3 or 4 m resulting in both an increasing number of pixels and a lower proportion of mixed pixels (i.e. pixels consisting of two or more classes (Hsieh et al., 2001)). Accordingly, it is very likely that neighbouring pixels belong to the same land cover class as the pixel under consideration leading to within-class spectral variation, thus, one single pixel is probably not representative for the spectral characteristics of an individual tree. In addition, VHR images provide increased textural information regarding texture and shape of ground features (Van Coillie et al., 2007; Pu et al., 2018).

Pixel-based classification approaches focus only on spectral values of each pixel and each pixel is dealt with in isolation from its neighbours. After applying classification algorithms, neighbouring pixels are often assigned to different classes despite being similar. This results in salt-and-pepper effects and therefore high classification errors (Kelly et al., 2011). Many studies have suggested that object-based image analysis (OBIA) is superior to pixel-based approaches because they reduce these salt-and-pepper effects, hence increase classification accuracies (e.g. Drăgut et al. (2010); Immitzer et al. (2012); Ng et al. (2017); Pu et al. (2018)).

## 2.1 Object-based image analysis (OBIA)

OBIA is an alternative analytical framework that can mitigate the deficiencies associated with pixel-based approaches (Adam et al., 2016). According to Blaschke (2003), the core concept of OBIA is that important information necessary to interpret an image is not represented in a single pixel, but in meaningful image objects and their mutual relationship. It aims to delineate readily usable objects from imagery in order to use spectral and contextual information in an integrative way. The increased spectral information of objects compared to single pixels is one of the main advantages of OBIA, while the additional spatial information is probably of an even greater advantage (Blaschke, 2010). This enables the

extraction of spectral and textural features, not available for pixel-based approaches (Ng et al., 2017). Benz et al. (2004) has mentioned three additional advantages (1) meaningful statistics and texture calculation, (2) increased uncorrelated feature space using shape and topological features and (3) close relation between real-world objects and image objects which improves the value of the final classification. Nonetheless, the quality of the final classification depends largely on the experimental objects used (Adam et al., 2016).

OBIA has become more popular in recent years due to the availability of commercial eCognition® software, launched in 2000 by Definiens Imaging GmbH (Lewinski et al., 2010). However, the idea of incorporating contextual information in classification can be tracked back to the 1970s. In literature it is generally agreed upon that OBIA builds on older segmentation, edge-detection, feature extraction and classification concepts used in remote sensing image analysis for decades (Blaschke, 2010). The application of OBIA has extended in various fields, especially forest mapping and land use land cover classification (Lewinski et al., 2010). Recently there exists a multitude of implemented segmentation algorithms for remote sensing tasks, but only some of them are commercially available (Meinel and Neubert, 2004). An overview of available segmentation software has been provided by Hossain and Chen (2019).

During OBIA, a three-step object-oriented classification process is implemented. In the first and most crucial step, a pixel-based image is divided into relatively homogeneous image objects with regard to spatial and spectral characteristics. Consequently, within-class spectral variation is reduced (Blaschke, 2010; Drăgut et al., 2010) (Section 2.1.1). After segmentation, image objects are characterised using various features followed by extraction of the most significant features for optimal class separation (Nussbaum et al., 2006; Van Coillie et al., 2007) (Section 2.1.2). During the final step, image objects are assigned to classes by a classification algorithm or based on expert knowledge (Nussbaum and Menz, 2008) (Section 2.1.3).

## 2.1.1 Image segmentation

Pal and Pal (1993) has defined image segmentation as the process of partitioning an image into non-intersecting regions (i.e. image objects) such that each region is homogeneous and the union of no two adjacent regions is homogeneous. It has been suggested in literature that image analysis leads to meaningful objects only when the image is segmented into 'homogeneous' regions or 'relatively homogeneous' regions, of which the latter term reflects better the near-decomposability of natural systems. The key is that the internal heterogeneity of each image object is lower than the heterogeneity compared with its neighbouring image objects (Blaschke et al., 2004). Thus, segmentation has two desirable properties: (1) each of the resulting image objects should be internally homogeneous, and (2) should be distinguishable from its neighbours (Espindola et al., 2006). Segmentation can be approached as an optimisation process during which image objects of minimum heterogeneity given certain criteria have to be found (Benz et al., 2004). These criteria include colour (i.e. spectral values of the pixels forming the object) and shape (i.e. the result of texture) of an image object (Munyati, 2018). The resulting image objects are the basic processing units of OBIA, holding more real-world values than pixels alone, and may be considered as object primitives. These object primitives do not necessarily have any meaning, but they should

4

maintain some important properties, such as low spectral variance, similar sizes for the different objects included in the same land cover, and absence of under-segmentation. Object primitives are usually a crucial step before objects of interest, matching real-world objects, can be found by further processing (Benz et al., 2004; Frauman and Wolff, 2005; Blaschke, 2010; Drăgut et al., 2010; Cánovas-García and Alonso-Sarría, 2015). The process of endowing the object primitives with meaning is a complex one and usually takes place in the classification step (Drăguţ et al., 2014). In fact, image segmentation is a pre-classification step creating more compact and fewer object primitives (Blaschke et al., 2004).

Segmentation of an image into objects is a problem with a huge number of possible solutions. Segmentation algorithms often respond sensitively to insignificant variations such as slight parameter changes, the order of segmentation hierarchical approaches, or the image data itself. Thus, users are confronted with high degrees of freedom which should be minimised to satisfy the given requirements (Blaschke, 2003; Meinel and Neubert, 2004). Furthermore, segmentation needs to address a certain scale (Blaschke, 2003), which controls the average size of image objects (Kim et al., 2008). The average size of image objects critically impacts classification accuracy (Drăguţ et al., 2014). Therefore, segmentation can increase classification and statistical accuracy when conducted at an appropriate scale (Drăgut et al., 2010). The importance of an optimal scale is different between feature extraction and wall-to-wall classification. For feature extraction, the optimal scale is not as important as for classification, as it is often sufficient to obtain a preliminary approximation of the scale parameter (Drăgut et al., 2010). The optimal scale depends on the targeted ground features during analysis and occurs at the value where the objects of interest are not over- or under-segmented. In an over-segmented image, adjacent objects are on average somewhat similar, while in an under-segmented image, the objects are too large and therefore tend to lose their spectral homogeneity (Kim et al., 2008). Since landscapes typically consist of different types of land cover that vary in size (e.g. trees, roads, buildings), and meaningful structures appear at different scales (e.g. stand level at coarser scales, and individual trees at finer scales), single-scale segmentation methods may not be able to properly segment all objects of interest without the occurrence of over- or under-segmentation. This is why using a multi-scale approach often may be preferable. However, the use of multiple segmentation scales is typically based on extensive knowledge of the area under investigation (Trias-Sanz et al., 2008; Johnson and Xie, 2011)). Therefore, the selection of accurate scale parameters is a crucial decision during segmentation of an image (Drăguţ et al., 2014). Unfortunately, there is no recognisable relationship between the scale parameter, which is unitless, and spatial measures, such as area, specific for image objects composing a scene (Hay et al., 2005). This makes it very challenging to define appropriate segmentation parameters to generate objects that satisfy user requirements without performing subjective trial-and-error experimentation, leading to serious limitations for some users (Hay et al., 2005; Drăgut et al., 2010). Hence, the integration of instruments for the evaluation of segmentation quality appears desirable (Meinel and Neubert, 2004). Assessing the quality of segmentation is difficult because currently no standard evaluation methods exist (Drăguţ et al., 2014). Segmentation results can be evaluated visually, based on expert knowledge, because the human eye is acknowledged as a strong and experienced source for the evaluation of segmentation algorithms. Additionally, quantitative accuracy assessment of the segmentation results can be performed during the classification step (Benz et al., 2004; Drăgut et al., 2010).

Many segmentation techniques have been developed in the last decades. From an algorithmic perspective, Pal and Pal (1993) has grouped these techniques into four categories (Van Coillie et al., 2007; Blaschke, 2010):

1. Point-based: separates pixels of an image into different objects by thresholding (e.g. grey-level thresholding) (Tian and Chen, 2007).

2. Edge-based: identifies edges and closes them by using contouring algorithms. It assumes that between edges, the pixel properties change abruptly (e.g. edge detection) (Hossain and Chen, 2019).

3. Region-based: detects regions that meet certain homogeneity criteria (e.g. split and merge) (Tian and Chen, 2007).

4. Combined.

A review of segmentation techniques has been provided by Haralick and Shapiro (1985), Pal and Pal (1993), and more recently by Hossain and Chen (2019). They all address salt-and-pepper effects, even though many suffer from major drawbacks such as over- and under- segmentation or not being useful at all scales (Frauman and Wolff, 2005). This is why only a few of the existing segmentation techniques are widely available in commercial software packages and lead to qualitatively convincing results while being robust and operational (Blaschke, 2003). For instance, eCognition® Developer version 10.2 contains 11 segmentation algorithms. In the User Guide eCognition® Developer (2019), segmentation is defined as any operation that creates new image objects or alters the morphology of existing image objects according to specific criteria. The segmentation algorithms are classified according to two basic principles, namely a top-down or a bottom-up strategy. During the former, large objects are cut into smaller pieces, and during the latter, small pieces are merged to obtain larger objects. Table 2.1 gives an overview of the segmentation algorithms implemented in eCognition® Developer version 10.2.

Table 2.1: Segmentation algorithms implemented in eCognition® Developer version 10.2 (User Guide eCognition® Developer, 2019).

| Segmentation algorithm | Description |
| --- | --- |
| **Top-down** | |
| Chessboard | Cuts the image into equal squares of a given size. |
| Contrast filter | Uses pixel filters to detect potential image objects by contrast and gradient and creates suitable object primitives. |
| Contrast split | Segments an image into dark and bright image objects based on a threshold value that maximises the contrast between them. |
| Multi-threshold | Splits the image into image objects based on user-defined pixel threshold values. |
| Quadtree based | Creates squared image objects with different sizes. |
| Standard deviation split | Applies a contrast split algorithm based on conditions using the standard deviation of an input layer. |
| Vector-based | Creates or converts an image object level to reflect the content of a vector layer. |
| **Bottom-up** | |
| Multiresolution | Consecutively merges pixels or existing image objects, based on relative homogeneity criteria. |
| Spectral difference | Merges neighbouring image objects if the difference between their layer mean intensities is below the value given by the maximum spectral difference. |
| Superpixel | Starts from a grid of superpixels and then refines the boundaries by proposing updates of edges boundaries in iterations. |
| Watershed | Grows image objects from a seed (local intensity minima) until they touch image objects growing from neighboring seeds. |

**Multiresolution segmentation**

*Concept*

Multiresolution segmentation, introduced by Baatz and Schäpe (2000), is one of the most widely used and successful image segmentation algorithms and has been shown to produce image objects that closely resemble field reality (Munyati, 2018). It is a bottom-up region-merging technique, embedded in eCognition® Developer, starting with one-pixel objects, which are iteratively merged with their neighbours into larger objects. The process continues until a homogeneity threshold, defined by the scale parameter, of the image object is exceeded (Hay et al., 2005; Drăguţ et al., 2010; Wang et al., 2018). This homogeneity threshold is a combination of colour and shape, weighted by the user. The higher the shape factor is weighted, the lower the influence of colour in the segmentation, and vice versa. For the shape factor, it is also necessary to define compactness and smoothness values, which are inversely related to each other. Compactness targets objects with well-defined edges, while smoothness defines objects that have more transitional edges. The scale parameter is the most important determining parameter. It is the maximum standard deviation of the homogeneity threshold, regarding the weighted image layers, for the resulting image objects. A higher scale parameter leads to larger and less homogeneous image objects by decreasing the homogeneity threshold per object (Drăguţ et al.,

2010; Adam et al., 2016; Munyati, 2018). In summary, the multiresolution segmentation algorithm requires the specification of four key parameters prior to segmentation, namely shape (versus colour), compactness (versus smoothness), image layer weights and scale, which are often set using trial-and-error experimentation (Munyati, 2018).

After running the multiresolution segmentation algorithm, the resulting image objects do not yet correspond to real elements in the landscape and are called object primitives. These object primitives are often smaller than manually digitised objects. Hence, there is a need for other fusion algorithms to merge these object primitives into a set of larger objects that are more related to landscape elements. One of the most frequently used fusion algorithms in eCognition® Developer is the spectral difference segmentation algorithm, which refines the segmentation results. This algorithm improves the spatial correspondence between object primitives and the real-world landscape objects by favouring the generation of large, homogeneous objects to prevent over-segmentation, and thereby reducing the total number of objects (Cánovas-García and Alonso-Sarría, 2015).

*Parameter optimisation*

Pre-segmentation image analysis methods for reducing subjectivity during the above described procedure have been proposed in literature (e.g. Kim et al. (2008); Drăgut et al. (2010); Drăguţ et al. (2014); Munyati (2018)). The drawback is that most of these methods require stand-alone software modules, making it difficult to integrate them into existing image segmentation software packages (Munyati, 2018). Many of the methods use image texture operators based on variance for the optimisation of the multiresolution segmentation algorithm (Munyati, 2018). Woodcock and Strahler (1987) has used average local variance graphs for understanding how the spatial structure of an image changes with the pixel size. They have calculated average local variance by computing the standard deviation of the spectral reflectance or brightness values in a 3 x 3 kernel, subsequently, the average for the entire scene has been graphed as a function of the associated pixel size. This value is an indicator of the local variability in the image. The application of the concept of local variance exploits spatial autocorrelation, which is an inherent feature of remote sensing data and a reliable indicator of statistical separability between spatial objects (Espindola et al., 2006; Kim et al., 2008; Drăgut et al., 2010). Kim et al. (2008) has introduced the concept of local variance to OBIA by exploring the relationship between image object variance and spatial autocorrelation at different scale parameters to define the optimal scale for forest stand segmentation (Drăgut et al., 2010; Drăguţ et al., 2014). Image object variance represents the internal homogeneity of the image objects, whereas spatial autocorrelation is a measure of the external heterogeneity between image objects. According to Cánovas-García and Alonso-Sarría (2015) neither of them is able to identify the optimal segmentation scale without the other. Kim et al. (2008) has conducted a series of segmentations with altering scale parameters to estimate the optimal object size. For each segmentation, average local variance and spatial autocorrelation of the image objects have been calculated and graphed as a function of the associated scale parameter. Additionally, the repeatability of the segmentation process has been assessed. After repeating the process five times, there has been no observable difference in the number or average size of the image objects, as well as the calculation of local variance and spatial autocorrelation. They have concluded that the optimal segmentation occurs at the scale just before flattening of the average local variance curve, and at the scale associated with the lowest,

and even negative, autocorrelation between image objects (Kim et al., 2008). This method has focused on one optimal scale, which is appropriate for simple scene models (Drăgut et al., 2010). Even though, the analyses have not agreed on the exact same optimal segmentation scale, they have narrowed the wide range of possible scales. Users often do not know the initial order of magnitude to begin for determining the segmentation scale. First, they can perform a rough cut of segmentation at a few scales between a wide range by graphing local variance and spatial autocorrelation. Afterwards, researchers can target specific scales, which are most likely associated with the optimal image object sizes, and avoid wasting time for segmentation at non-optimal scales (Kim et al., 2008).

Building on the results of Kim et al. (2008), Drăgut et al. (2010) has used the concept of local variance, proposed by Woodcock and Strahler (1987), to develop the 'Estimate Scale Parameter' (ESP) tool. They have extended this concept into multiscale analysis since many environmental problems cannot be handled at a single scale of observation. Multiscale analysis requires more than one suitable scale parameter to account for different levels of organisation in a landscape structure (e.g. forest, stand level, tree level) or to include different categories of objects with different sizes in complex scenes. The ESP-tool, implemented in the eCognition$^{\circledR}$ Developer environment, automatically generates image objects at multiple scale levels in an iterative way and calculates the local variance of each scale as the mean of the standard deviation of the image objects obtained through segmentation (Drăgut et al., 2010). Drăgut et al. (2010) has established that object variance increases with object scale (Munyati, 2018), and thus local variance alone is not suitable for a multiscale approach to determine at which scale meaningful objects emerge. Therefore, the rate of change of local variance (ROC-LV) has been introduced as a measure of the dynamics of local variance across scales. The graph of ROC-LV enables the detection of multiple scale parameters. Peaks in the ROC-LV graph indicate the scale levels at which the image can be segmented in the most appropriate manner, relative to the data properties at the scene level. The tool can be used for the analysis of a single layer of image data or other continuous data (e.g. digital surface models). The application on a single layer makes it independent of specific sensors or parameterisation issues (Drăgut et al., 2010). In order to take full advantage of multispectral information, segmentation on multiple layers is desirable (Drăguţ et al., 2014). Drăguţ et al. (2014) has introduced a fully automated methodology for the selection of suitable scale parameters, considering multiple layers, and using a local variance based algorithm (Drăguţ et al., 2014; Munyati, 2018). This approach is basically an automation of the ESP-tool, by implementing a three-level hierarchy concept, where the production of a graph has been replaced by an iterative procedure that segments an image at the first thresholds that occur in the local variance graph. The average local variance of the objects in each image layer is calculated and serves as a condition for stopping the iterations (Drăguţ et al., 2014). The optimal scale is found when the average local variance of the objects in all layers is equal to or lower than the previous segmentation iteration (Munyati, 2018). Subsequently, objects segmented in the previous scale level are retained. All of the calculated local variance values are recorded in a table that can be exported and processed with freely available stand-alone software. Consequently, a user can assess the scales that are detected by the automated ESP-tool or can simply choose other representative scales, based on the resulting graph, according to the specific goal of their project. Adapting the procedure to perform multi-layer segmentation has resulted in the challenge of specifying the number of layers considered in the segmentation, as well as

in the calculation of the average local variance values. To address this issue, an index has been implemented that counts the total number of layers added to eCognition® Developer and considers them all in processing. This makes the tool independent of a specific sensor and allows the integration of multiple data sets (e.g. ancillary data). The tool has a significant potential of increasing the objectivity and automation in OBIA applications. However, the tool produces statistically significant segmentations, which do not necessarily belong to a semantically relevant category of objects (Drăguţ et al., 2014). Considering this limitation, Drăguţ et al. (2014) has not expected successful automation in any possible case, especially when targeting semantically complex categories of image objects. Nonetheless, making at least first approximations of scales that exist within the data are possible (Drăguţ et al., 2014).

All of the above mentioned methods focus only on defining the optimal scale parameter. However, multiresolution segmentation requires the specification of three additional parameters, namely image layer weights, shape (versus colour), and compactness (versus smoothness). Munyati (2018) has proposed a method to determine the image layer weights, as well as a method to determine the shape (versus colour) factor for mapping vegetation communities in South Africa. Principal component analysis (PCA) of the image data has been performed in order to determine band (i.e. image layer) weight values (Munyati, 2018). PCA is a technique allowing to identify the bands that contain most of the information present in the image data by indicating whether or not there is hierarchy in the band data variance. Indication of the band hierarchy can be obtained through the dimensionality shown by the data variance in the principal components (Lillesand et al., 2015; Munyati, 2018). For the specification of the shape (versus colour) factor, Munyati (2018) has used texture images of each band generated by a variance texture enhancement operator (Equation 2.1) using 3 x 3 kernels.

$$variance = \frac{\Sigma(x_{ij} - M)^2}{n - 1} \tag{2.1}$$

where $x_{ij}$ is the reflectance value of pixel $ij$, $M$ is the kernel mean reflectance, and $n$ is the number of pixels in the kernel. The smallest kernel size has been preferred in order to preserve the spatial detail of the vegetation cover, which would have been lost with larger kernel sizes. Variance enhances texture because it increases when the pixel grey-level values differ from their mean in the kernel. Afterwards, the texture images have been stacked into one texture colour composite image. For both the untransformed and the texture colour composite image, descriptive statistics, such as mean and coefficient of variation (CV), of the reflectance data of vegetation communities have been calculated. Finally, the images have been compared using the CV of the reflectance data, and the shape (versus colour) factor has been determined based on the ratio of the range in mean values of both images (Munyati, 2018).

### 2.1.2  Image feature analysis

Determining a suitable classification feature space is a crucial step before image classification. Object features are created for each image object in order to improve object separability, ensuring that classes are discriminated effectively and with sufficiently high

accuracy (Mishra and Crews, 2014; Adam et al., 2016). Many object features have been proposed in literature:

- Image object statistics: within an image object, all kinds of statistics based on a single spectral band or combinations of spectral bands can be computed (Benz et al., 2004) such as band means, standard deviations, brightness, maximum difference, and ratios (Van Coillie et al., 2007; Mishra and Crews, 2014).

- Image object texture: texture is the tonal variation in an image and may be a good proxy for vegetation structure, since these variations arise from canopy biophysical properties such as architecture of the crowns, as well as their arrangement of leaves and branches (Thierion et al., 2014; Ferreira et al., 2019). Various texture measures exist, but they are typically derived using a moving window or kernel method of a fixed size, shape and (limited) orientation(s), the most common being co-occurrence matrices of the image grey-levels (Van Coillie et al., 2007; Blaschke et al., 2014). First and second order texture measures derived from the grey-level co-occurrence matrix (GLCM) after Haralick et al. (1973) as well as the grey-level difference vector (GLDV), including homogeneity, contrast, dissimilarity, entropy, angular second moment, mean, standard deviation, and correlation can be calculated (Van Coillie et al., 2007).

- Image object shape: shape refers to the general form or outline of individual image objects (Blaschke et al., 2014). The closer the object primitives are to real-world objects, the more object features such as shape index, size, length, and number of edges can be used as uncorrelated object features (Benz et al., 2004; Van Coillie et al., 2007).

- Topological object features: these features describe the spatial context of the image objects. In eCognition® Developer image objects on different scales form a hierarchical network where each object knows its intrinsic relation to its super-object (is within) and its sub-objects (are contained), as well as relations to neighbouring objects at the same scale (Blaschke, 2003). This object network provides additional object features: (1) characterisation of an image object based on its sub-objects using texture analysis, line analysis, or class-related features; (2) characterisation of an image object based on its super-object (Benz et al., 2004).

- Semantic features: these higher order features are available after a first classification of image objects. Semantic features reduce ambiguities, allow landuse classification in addition to pure landcover classification and thus lead to a first step of scene understanding (Benz et al., 2004). These features require the knowledge of the expert on the landscape in question (Blaschke, 2003).

The above described features fall into two groups (Lillesand et al., 2015). The first group contains image object statistics, texture, and shape. They can be regarded as intrinsic features, since they are available for each image object (Benz et al., 2004). The second group, containing topological and semantic features, describes the relationship among objects (Lillesand et al., 2015).

According to Lillesand et al. (2015) image characteristics such as texture, shape, and pattern are the most important features in the visual interpretation of remotely sensed im-

agery. Image texture has been recognised as an important source of information when interpreting forest vegetation (Van Coillie et al., 2007; Mallinis et al., 2008). The improvement of classification accuracy by adding textural features has been demonstrated by numerous authors (Thierion et al., 2014). For instance, Mishra and Crews (2014) has determined that next to spectral features, object-level texture features are also important for distinguishing savanna vegetation morphology types, because vegetation morphology types in semi-arid systems are often marked by a subtle difference in vegetation physiognomy. On top of this, species composition has relatively similar spectral characteristics. Yet, it remains difficult to successfully automate the recognition of objects solely based on textural features. Using them in combination with spectral characteristics might be advantageous (Van Coillie et al., 2007).

The availability of hundreds of features leads to a high-dimensional predictor feature space. Therefore, classification algorithms might overfit, especially when they face a sparse number of training samples compared to the number of predictors. Although, the gravity of this problem depends partly on the algorithm used (Van Coillie et al., 2007; Fassnacht et al., 2016). To overcome this issue, it might be beneficial to reduce the input dimensionality by an appropriate selection of features for class description. This might even lead to an increase in classification accuracy (Van Coillie et al., 2007; Adam et al., 2016). The selection of features can be based on past experience and user knowledge, or it can use feature reduction algorithms (Adam et al., 2016; Wang et al., 2018). Feature reduction algorithms can be subdivided into feature extraction (e.g. PCA) and feature selection (e.g. manual band selection, Random Forest) algorithms. Feature extraction algorithms calculate new predictor variables that typically summarise the content of several original predictors, while feature selection algorithms select a subset of the original predictor variables. In many cases, feature selection algorithms are less efficient than feature extraction algorithms. Nevertheless, they have the advantage of enabling a meaningful interpretation of the remaining feature space and thereby increase the understanding of what exactly drives the discrimination of, for example, tree species, which is often not the case for the newly composed feature space in a feature extraction approach (Fassnacht et al., 2016).

### 2.1.3 Tree species classification

The classification step assigns image objects to semantic classes (i.e. high-level concepts such as vegetation) (Arvor et al., 2013). Objects can be assigned to these classes by a classifier or based on expert knowledge (Nussbaum and Menz, 2008). For classifying tree species this is usually done in a supervised way, having some field data and relating them to the information provided by remote sensing data. These data are usually spectral data, as different tree species are characterised by different spectral signatures (Dalponte et al., 2019). Nonetheless, the spectral signature of tree crowns across species often have high spectral similarity as well as significant intra-species variability, making tree species classification a challenging task (Zou et al., 2019).

Supervised machine learning algorithms are well suited for species classification of individual tree crowns (Sumsion et al., 2019). In general, supervised machine learning algorithms can be divided into two steps, namely training and testing. During the training step, a set of training data and the associated output labels are used to learn a mapping from the

input training data to the desired output labels. This mapping is then used to predict the species classes for unseen, unlabeled test data during the testing step (Zou et al., 2019). Fassnacht et al. (2016) has provided an overview of the most frequently applied classification algorithms and some of their advantages and disadvantages. They have divided these algorithms into two categories: (1) parametric approaches, and (2) non-parametric approaches. Some studies using mixed sets of input features (e.g. spectral, texture, shape) have preferred the use of non-parametric machine learning algorithms like Random Forest or Support Vector Machine (Fassnacht et al., 2016).

**Random Forest**

Random Forest is a non-parametric ensemble (i.e forest) learning algorithm. During the training stage, several decision trees are constructed using a random sample with replacement (i.e. bootstrapped sample) of the reference data set (Immitzer et al., 2012; Karlson et al., 2016). Roughly two thirds of the data are sampled with replacement, while one third are withheld from the tree construction (i.e. out-of-the-bag (OOB) samples) (Mishra and Crews, 2014). Additionally, the algorithm searches at each tree node across a randomly selected subset of predictor features for establishing binary splits. The split determination is based on the Gini index of node impurity. The best-splitting feature is identified by maximising the index, and therefore the classification accuracy (Gislason et al., 2006; Immitzer et al., 2012; Karlson et al., 2016). Trees are split in a manner that reduces the uncertainty present in the data and hence the probability of misclassification. The ideal splitting feature, or a Gini value of zero, occurs when only one class is represented at each terminal node (Watts and Lawrence, 2008). The number of features included in the subset is a user-defined parameter, but the algorithm is not sensitive to it. The value is often set to the square root of the total number of predictors (Gislason et al., 2006; Karlson et al., 2016). In summary, two parameters have to be set to run the classifier: (1) the number of decision trees for executing the classification (i.e. number of bootstrap iterations), and (2) the number of input features used at each tree node (Immitzer et al., 2012; Pu et al., 2018). Due to the dual randomisation, the correlation between decision trees is decreased and the overall computational complexity of the algorithm is reduced (Mishra and Crews, 2014). For classification, each tree in the Random Forest casts a unit vote for the most popular class. The final output of the classifier is determined by a majority vote of all trees. The trees in Random Forest are not pruned. Pruning is not needed as each classification is produced by a final forest that consists of independently generated trees created though a random subset of the data, further reducing the computational load (Gislason et al., 2006; Watts and Lawrence, 2008).

The algorithm has an internal accuracy measure, making it unnecessary to partition the reference data into training and test data. The accuracy of the reference data set is estimated by running the OOB samples (i.e. the remaining reference set samples that are not in the bootstrap for a particular tree) of each tree down through the tree as a form of cross-validation. The results are combined with a majority vote and then compared with the true class label giving an estimation of the misclassification error. Since each tree can only use the reference samples that were OOB to estimate the accuracy, the classification error estimation is lower than the estimation made by an independent test set (Gislason et al., 2006; Watts and Lawrence, 2008; Immitzer et al., 2012). Fassnacht et al. (2016) has

recommended the application of an iterative data-splitting approach in combination with a completely independent test set as a gold standard for tree species classification studies. Random Forest calculates estimates of feature importance as Mean Decreasing Accuracy (MDA). MDA is automatically generated within Random Forest by calculating the decrease in classification accuracy that results from randomly permuting the predictors in the OOB samples (Karlson et al., 2016; Ng et al., 2017). If the outcome is unaffected, the predictor has low feature importance, whereas an increased classification error is an indication of the importance of the feature (Gislason et al., 2006; Smith, 2010). The MDA values can be used for feature ranking and feature selection. By performing feature selection the model is optimised leading to an improved classification accuracy as well as preventing overfitting of the model (Karlson et al., 2016; Ng et al., 2017). Karlson et al. (2016) has proposed a backward feature elimination procedure to identify the most efficient and parsimonious classification model. The procedure starts by building the Random Forest model using the full predictor data set. The lowest-ranked predictors are then removed successively until the mean square error of the model, calculated from the OOB samples, is minimised (Karlson et al., 2016). Feature selection makes it possible to add a large number of features because they are screened for their respective discriminative power (Ng et al., 2017).

Random Forest is often selected over conventional parametric classifiers and other machine learning algorithms since it possesses several properties that are favourable for remote sensing-based tree species classification (Karlson et al., 2016). First of all, Random Forest makes no assumptions about the data distribution. In tree species classification the robustness against non-normally distributed data sets can be helpful, for example, if only the most relevant tree species are mapped individually while pooling the secondary tree species in a single class (Immitzer et al., 2012). Additionally, the algorithm is effective in separating spectrally similar classes (e.g. tree species) characterised by high intra-class variability since it does not require that classes have a common covariance matrix, which is often not the case in tree species classification (Immitzer et al., 2012; Karlson et al., 2016; Pu et al., 2018). Moreover, it is an easily implemented technique because of the small number of parameters that needs to be adjusted, and the low tendency to overfit the data due to the self-validation mechanism (Karlson et al., 2016; Pu et al., 2018). The measure of feature importance is also recognised as one of the main advantages of Random Forest since it has been demonstrated highly useful for identifying the most effective features for developing simplified classification models (Karlson et al., 2016; Pu et al., 2018).

## 2.2 Image data

According to Fassnacht et al. (2016), spatially explicit information of tree species plays an important role in a wide variety of applications in forestry, regardless of the forest or vegetation type. The only way to obtain this information over an extended area is by remote sensing (Ferreira et al., 2019). A review of studies on tree species classification from remotely sensed data has been provided by Fassnacht et al. (2016). Various types of active and passive sensors have been employed so far and the highest classification accuracies have been obtained using airborne hyperspectral data or combining hyperspectral with light detection and ranging (LiDAR) data. Nonetheless, hyperspectral and LiDAR sensors involve high costs and therefore motivate the use of more affordable remote sensing data such

as these obtained by spaceborne platforms (Ferreira et al., 2019), however, satellite-based tree species classification has previously been limited by the spectral and spatial resolution of the available sensor systems (Karlson et al., 2016). The launch of VHR spaceborne platforms has proven to be a promising technology for mapping tree species over broad spatial extents. These satellites acquire images featuring metric and sub-metric resolutions (Ferreira et al., 2019), hence providing sufficient spatial detail to achieve an accurate classification at the species level (Deur et al., 2021). The latest generation of VHR sensor systems (e.g. WorldView-3) has approached the level of hyperspectral systems by acquiring spectral data in strategically located and relatively narrow wavelength bands with high relevance for tree species classification (Karlson et al., 2016), but they are still limited by a small swath area, large data volumes, low temporal frequency, and high data costs (Mishra and Crews, 2014). In general, the higher the spatial, spectral and/or temporal resolution, the higher the data costs. For example, Pléiades has a moderate spatial resolution, low spectral resolution, and high temporal resolution, resulting in a data cost ranging between 1.40 and 1.80 EUR/km$^2$ with a minimum of 100 km$^2$, without VAT and processing costs. However, these prices are exclusive for the so called CAT-1 user community (i.e. authorised users of countries that partially funded the Pléiades programme), and are much lower than the prices for commercial use. The prices are also much lower compared to other commercial VHR satellite imagery. An overview of VHR satellites launched after 2007 is given in Table 2.2.

Most of these satellites collect data in both a panchromatic band with high spatial resolution and in several multispectral bands with low spatial resolution, so there exists a trade-off between spatial and spectral resolution. For obtaining improved spatially explicit information, the spatial resolution of the multispectral bands can be upgraded by incorporating information from the panchromatic band (Ibarrola-Ulzurrun et al., 2017; Deur et al., 2021). The process of generating synthetic high-resolution multispectral imagery by combining panchromatic and multispectral information while maintaining spatial and spectral fidelity is a special type of image fusion known as pansharpening or resolution merge. The process can be interpreted in two ways: (1) it increases the spectral resolution of the panchromatic image while enhancing interpretability, and (2) it enhances spatial resolution of the multispectral images while preserving spectral fidelity (Ghosh and Joshi, 2013). Resulting from this, pan-sharpened multispectral imagery with an effective spatial resolution of the panchromatic band is obtained (Lillesand et al., 2015). Several studies have analysed the influence of image fusion on tree species classification and indicated an increased classification accuracy when using the fused imagery (Pu et al., 2018; Deur et al., 2021). Even though tree species classification based on VHR satellite imagery seems to achieve promising results, several studies have researched the influence of the use of additional data (e.g. multi-seasonal imagery or height information such as a canopy height model or digital surface models) on the classification accuracy (Deur et al., 2021).

Table 2.2: Overview of very high resolution (VHR) satellites launched after 2007. For each satellite the spatial resolution or ground sampling distance (GSD) (m), number of spectral bands (i.e. spectral resolution) in the visible near infrared (VNIR) (coastal blue (CB), blue (B), green (G), yellow (Y), red (R), red edge (RE), and near infrared (NIR)), shortwave infrared (SWIR), and clouds, aerosols, vapours, ice & snow (CAVIS) spectral range, and the revisit time (i.e. temporal resolution) are listed.

| Satellite | Spatial resolution (m) | Spectral resolution | Temporal resolution | Source |
|---|---|---|---|---|
| Cartosat-3 | PAN: 0.28 VNIR: 1.12 | | | Indian Space Research Organisation (nd) |
| GeoEye-1 | PAN: 0.41 VNIR: 1.64 | B, G, R, NIR | 3 days | European Space Agency (nd) |
| Pléiades 1A & 1B | PAN: 0.50 VNIR: 2.00 | B, R, G, NIR | daily | Satellite Imaging Corporation (nd) |
| SkySat-1 & 2 | PAN: 0.86 VNIR: 1.00 | B, G, R, NIR | 4 - 5 days constellation sub-daily | European Space Agency (nd) |
| SkySat 3 to 15 | PAN: 0.65 VNIR: 0.81 | B, G, R, NIR | 4 - 5 days constellation sub-daily | European Space Agency (nd) |
| SkySat 16 to 21 | PAN: 0.57 VNIR: 0.75 | B, G, R, NIR | 4 - 5 days constellation sub-daily | European Space Agency (nd) |
| SuperView-1 | PAN: 0.50 VNIR: 2.00 | B, G, R, NIR | 2 days constellation sub-daily | Satellite Imaging Corporation (nd) |
| WorldView-1 | PAN: 0.50 | | 1.7 - 5.9 days | European Space Agency (nd) |
| WorldView-2 | PAN: 0.46 VNIR: 1.80 | CB, B, G, Y, R, RE, NIR, NIR2 | up to 1.1 days | European Space Agency (nd) |
| WorldView-3 | PAN: 0.31 VNIR: 1.24 SWIR: 3.70 CAVIS: 30 | CB, B, G, Y, R, RE, NIR, NIR2 8 bands 12 bands | 1 - 4.5 days | European Space Agency (nd) |

## 2.2.1 Multi-seasonal imagery

A critical issue for tree species classification based on remote sensing is to determine when image data should be acquired during the year for optimising the classification accuracy. Knowledge about when the spectral signature of the tree species is most distinct, and thus when the potential for accurate tree species identification is the greatest, can help plan the image acquisition (Karlson et al., 2016). According to Madonsela et al. (2017), tree species classification can be enhanced by using time-series covering different periods during the phenological cycle (i.e. seasonal variation in plant characteristics such as changes in the colour of the leaves due to senescence or flowering, and fruiting events (Ferreira et al., 2019)). Changes in phenology occur throughout the growing season at different rates amongst species. Data that capture these changes may amplify the spectral variability between deciduous species in relation with intra-species variability (Madonsela et al., 2017). Hence, selecting an ideal period in which target species show peculiar characteristics (e.g. flowers or senescent leaves), can potentially favour their detection (Ferreira et al., 2019). In other words, it is desirable to align the time of image acquisition with the phenological cycle. A combination of two images, one during spring time (green-up) and one during autumn (senescence), supports high classification accuracies, but it is recommended to avoid late fall imagery after most leaves have fallen. The influence from the background will increase during this period, leading to a more problematic separation of deciduous species. Furthermore, it has to be noted that temperature buffering effects of large water bodies can cause senescence gradients in the study area which may hamper the classification, similar effects can be expected for altitudinal gradients (Fassnacht et al., 2016).

Despite the fact that tree species mapping obtain more accurate results using multi-seasonal imagery, it is also important to consider the higher costs and processing requirements associated with multiple acquisitions. These considerations might become a restriction of phenological approaches, and thus a single image acquired in the optimal season may be more important (Fassnacht et al., 2016; Karlson et al., 2016; Pu et al., 2018). Karlson et al. (2016) has established that dry season imagery provide better classification results for mapping West African agroforestry tree species than imagery acquired in the wet season. During the dry season some tree species are undergoing senescence and leaf shedding earlier than others, or begin to develop new leaves, which tend to have rather divergent colours. These phenological stages cause tree species to be more separable from each other. In contrast, peak productivity occurs during the wet season leading to a lower spectral variability, therefore reducing the classification accuracies. Additionally, frequent cloud cover hinders image acquisition, favouring dry season imagery even more (Karlson et al., 2016).

## 2.2.2 Height information

The use of multispectral and hyperspatial imagery products have traditionally enabled the identification of individual tree species, but it is evident from various structural remote sensing studies (e.g. Cho et al. (2012); Holmgren et al. (2008)) that structural information, especially tree height, plays an important role in assisting or being solely used in vegeta-

tion cover and tree species classification and mapping. Structural variables may help to reduce spectral confusion, for instance, when a particular tree species possesses spectral properties similar to the underlying grass layer. However, the cost of additional data sets is likely to challenge budget constraints (Kato et al., 2009; Naidoo et al., 2012; Vaughn et al., 2012).

LiDAR remote sensing was found to capture variations in the three-dimensional (3D) structure of vegetation (Sarrazin et al., 2011), particularly vegetation height. LiDAR is an active sensing technology using a laser to transmit a light pulse towards a target and a receiver to measure the backscattered or reflected light from that target. The distance to the object is determined by recording the time interval between the transmitted and backscattered pulses (Cho et al., 2012). These data allow for measuring structural attributes and physical parameters of vegetation at the individual tree crown and canopy levels. These structural and biophysical measures may include tree height, above ground biomass, forest stand density, basal area, and leaf area index at both the individual tree and stand level (Pu and Landry, 2020). The position of individual trees can also be measured with sub-metre accuracy (Holmgren et al., 2008). Generally, tree height and stem location are derived from a normalised digital surface model (nDSM) or canopy height model (CHM). A CHM represents the difference between the top canopy surface and the underlying ground topography by taking the difference between a digital surface model (DSM) and a digital terrain model (DTM), and contains all information necessary to determine the vegetation height above the ground level. The resulting raster images allow to detect treetops and delineate individual tree crowns (Kato et al., 2009; Panagiotidis et al., 2017). There exist two kind of LiDAR-systems, on the one hand discrete-return devices, on the other hand full wave recording devices. The former captures returns from individual structures, which can provide more direct measurements of individual tree crown parameters, whereas the latter focuses primarily on plot-level estimations of canopy structures (Kato et al., 2009). For tree species classification, several studies (e.g. Vaughn et al. (2012)) have reported improvements in classification accuracy when applying waveform data sets, because they provide an increased number of echoes that can be extracted compared to typical discrete return data acquisitions (Fassnacht et al., 2016). Airborne LiDAR data sets are rare and expensive to collect, while spaceborne LiDAR-systems are inappropriate for tree height mapping at the individual level due to the wide gaps between laser shots and laser footprints. However, other satellite-based approaches are being developed which have good potential to produce VHR (ca. 2 m) tree height models (i.e. interferometry and Tandem-X, or stereo mapping) (Cho et al., 2012).

Fassnacht et al. (2016) has stated that canopy height per se is an illogical predictor to classify tree species, because the absolute height of a tree species mainly varies with age, site conditions, and competition and only to a minor degree with species. Consequently, height information rather leads to confusion than to improving the discriminative power in presence of several age classes of a single species (Fassnacht et al., 2016), but it was found to improve classification accuracies of smaller trees (Naidoo et al., 2012). Alternatively, the combination of structural and spectral variables may help to overcome the high intra-species spectral variability, while taking advantage of known differences in structure between species (Naidoo et al., 2012). It has been demonstrated that combining VHR satellite imagery with airborne LiDAR data can be an effective approach for monitoring forest stands, mapping individual tree crowns, and classifying tree species and species compo-

sition. This is attributed to a synergy of VHR data providing sufficient spatial and textural information, and LiDAR data providing vertical profile and structural information, which together is beneficial for tree species classification (Pu and Landry, 2020). Cho et al. (2012) has shown that the difference in overall accuracy when using a combined spectral and height data set versus a spectral only data set was small. Even though, the advantages of using an integrated approach could be traced in the steps leading to the classification, and in the classification output at the individual species level. A tree height mask allows to easily extract tree spectral information from the image, which would be a challenging task with spectral data only. Cho et al. (2012) has established that another advantage of using height information is the decrease in misclassification of short trees as tall trees and vice versa, but this advantage was balanced by an increased confusion between trees of the same height category. In conclusion, LiDAR-derived or stereo image-based vegetation height can be used to separate non-canopy from canopy elements using height thresholds, followed by a species classification based on passive optical information (Fassnacht et al., 2016).

# 3. MATERIAL AND METHODS

The concepts and methods discussed in Chapter 2 were applied and modified for the differentiation of *Melia volkensii* in an agroforestry layout in the region of Kiambere based on Pléiades satellite imagery. An overview of the study area and the data used, including image data as well as ground truth data, is provided, followed by a detailed description of the method developed to differentiate *Melia volkensii*.

## 3.1 Study area

The study focused on the region of Kiambere, situated in Kitui County in central Kenya, where both a plantation and cooperation with local farmers in an agroforestry layout occur (Figure 3.1). The plantation borders on Lake Kiambere, which is artificially formed after building a dam on Tana River, and has an area of approximately 300 ha. The area is mainly planted with mukau (*Melia volkensii*), but a small part is also provided for neem (*Azadirachta indica*) and acacia (*Acacia* sp.). These trees have different planting years, with the oldest planted in 2007 and the youngest in 2020. The farms are located east of Lake Kiambere, 38 of them are enrolled in the *Seven Forks Farmers Programme*, where BGF collaborates with smallholder farmers to plant mukau trees in their fields. These trees are all planted between 2015 and 2019 (Better Globe Forestry, nd). Next to *Melia volkensii*, there are some trees planted by the farmers themselves such as mango (*Mangifera indica*), and acacia, but they also planted some additional mukau trees that are no property of BGF.
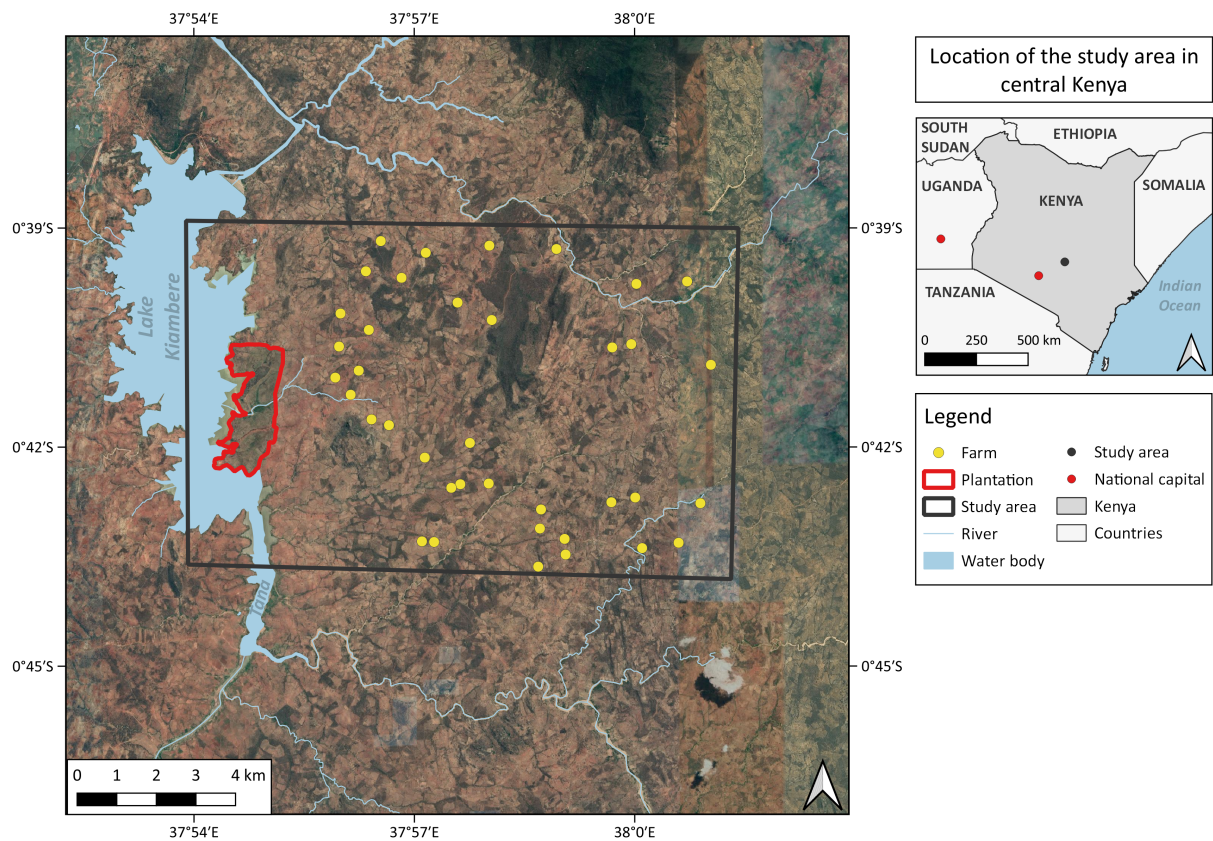
Figure 3.1: Location of the study area in central Kenya, with indication of the plantation (red line), bordering on Lake Kiambere, and the 38 farms (yellow dots) (open Africa, nd; OpenStreetMap, nd).

According to the Köppen-Geiger climate classification system, Kiambere is situated in a hot semi-arid (BSh) climate zone. These climates tend to have hot summers and warm to cool winters, with minimal to some precipitation (Beck et al., 2018). Kitui county experiences two wet seasons, a long rainy season from March until May and a short rainy season from October until December. The mean annual precipitation varies between 400 and 1000 mm, with an average of 750 mm. The mean annual temperature ranges between 21 and 31 °C (Figure 3.2) (The Ministry of Agriculture, Livestock, Fisheries and Co-operatives (MoALFC), 2021). The soil type in the region of Kiambere is classified as well-drained, deep to very deep, dark reddish brown to yellowish red, friable, sandy clay to clay (i.e. luvisols or acrisols) (European Soil Data Centre (ESDC), nd).
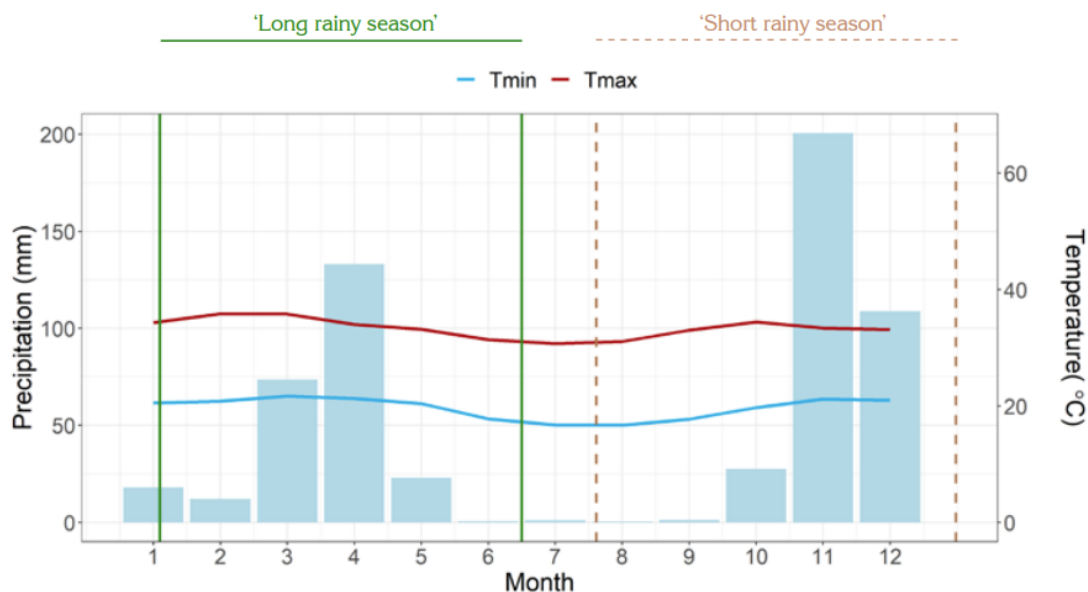
Figure 3.2: Historical monthly mean temperature (°C) and precipitation (mm) (average 1985-2015) for Kitui County. Bars represent total monthly precipitation, while lines represent minimum (blue line) and maximum (red line) monthly mean temperatures (The Ministry of Agriculture, Livestock, Fisheries and Co-operatives (MoALFC), 2021).

## 3.2 Data collection

### 3.2.1 Image data

Pléiades satellite imagery was selected for use in this study. The image was acquired on 2 June 2021 (at 07:52 AM local time), covering an area of approximately 14 x 9 km (Figure 3.3). Pléiades collects data in one panchromatic band and four multispectral bands (i.e. red (R), green (G), blue (B), and near infrared (NIR)) (Table 3.1) (Coeurdevey and Fernandez, 2012). The image was reprojected to the Universal Transverse Mercator (UTM) coordinate reference system (CRS) (datum: Worldwide Geodetic System 84 (WGS 84), zone 37S) and digital numbers were converted to top-of-atmosphere (TOA) reflectance values, resulting in a 16-bit image at pre-processing level 1C.

A 0.50 m resolution pan-sharpened Pléiades image was produced by fusing the 2 m resolution multispectral imagery with the 0.50 m resolution panchromatic imagery using the SPEAR pan-sharpening tool in ENVI 5.6.1 with the Gramm-Schmidt spectral sharpening method. Some studies have demonstrated that a pan-sharpened image might improve classification results of forests (Section 2.2). Therefore, all further analysis including image segmentation, feature calculation and tree species classification were conducted based on the 0.5 m resolution pan-sharpened Pléiades image (Pu et al., 2018).
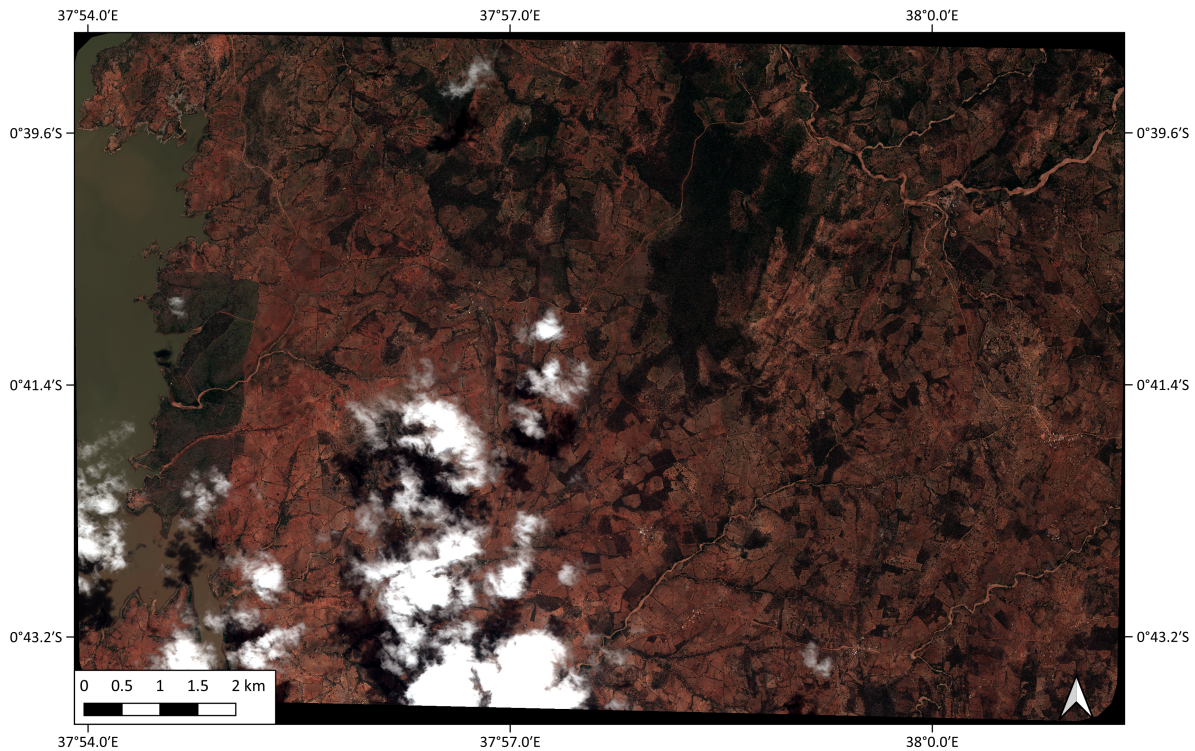
Figure 3.3: True colour composite of the Pléiades imagery with a spatial resolution of 2 m (Airbus, 2021).

Table 3.1: Spectral domain (nm), according to the standard expression Full Width at Half Maximum (FWHM), and spatial resolution (m) for each band (Coeurdevey and Fernandez, 2012)

| Band | Imin FWHM (nm) | Imax FWHM (nm) | Spatial resolution (m) |
|------|----------------|----------------|------------------------|
| Panchromatic | 490 | 810 | 0.5 |
| Blue (B) | 460 | 530 | 2 |
| Green (G) | 520 | 600 | 2 |
| Red (R) | 610 | 690 | 2 |
| Near Infrared (NIR) | 770 | 890 | 2 |

## 3.2.2 Ground truth data

A field inventory was conducted between October 2021 and January 2022 where a total of 64 square inventory plots (32 m x 32 m) in the plantation and 17 farms were collected. Information regarding their location in the study area can be found in respectively Table A.1 and Table A.2 in Appendix A. BGF divided the plantation into grid cells of 10 ha each. In each grid cell an inventory plot was randomly located where all trees were measured. In some grid cells with a tree coverage of more than 5 ha, an extra inventory plot was provided. Grid cells covered by clouds on the Pléiades imagery were not visited. A full inventory of the *Melia volkensii* trees belonging to BGF in the farms was conducted. Farms with less

than 30 standing seedlings and farms covered by clouds on the Pléiades imagery were not visited, consequently, data was collected in 17 out of the 38 farms. The following variables were surveyed in the plantation, as well as in the farms:

1. Location: X, Y coordinates using a handheld global positioning system (GPS) (Garmin 60scx). The coordinates were recorded at the north side of the trunk and notated in CRS WGS 84 / UTM zone 37S.

2. Species: identification in the field.

3. Diameter at breast height (cm): tree diameter measured 1.3 m above surface.

4. Crown diameter (cm): average of crown diameters measured in the North-South and East-West direction.

5. Height (m): total tree height measured with the phone application Arboreal - Tree height.

6. Tree age (year): in the plantation, the age was notated on the trunk of each tree. In the farms, the age was taken note of when it was recorded during planting.

7. Soil degradation: visual assessment based on soil erosion by water, since this is one of the main causes of soil degradation in Kenya (Luvai et al., 2020). Three types of soil loss were considered, namely sheet erosion, rill erosion, and gully erosion (Omuto, 2008).

8. Tree vitality: visual assessment based on expert knowledge. The trees were given a score ranging from 1 (perfect) to 5 (almost dead).

A summary of the descriptive statistics of *Melia volkensii* in the plantation as well as in the farms is given in Table 3.2. A similar table for data collected of neem, acacia, and baobab in the plantation can be found in Appendix B.

Table 3.2: Descriptive statistics of *Melia volkensii* in the plantation as well as in the farms. The trees are divided in classes of 5 cm diameter at breast height (DBH) (cm). For each class, the average crown diameter (cm), height (m), age (year), and vitality (1 to 5) are calculated. Additionally, the number of trees (%) in each class is represented.

| DBH (cm) | Crown diameter (cm) | Height (m) | Tree age (year) | Tree vitality | Number of trees (%) |
|---|---|---|---|---|---|
| **Plantation** | | | | | |
| 0.0 - 5.0 | 131.57 | 3.87 | 6.12 | 1.97 | 1.78 |
| 5.0 - 10.0 | 262.00 | 7.33 | 8.34 | 1.65 | 14.59 |
| 10.0 - 15.0 | 410.98 | 9.73 | 9.43 | 1.28 | 46.98 |
| 15.0 - 20.0 | 521.19 | 11.03 | 10.15 | 1.12 | 30.76 |
| 20.0 - 25.0 | 625.71 | 12.31 | 10.61 | 1.18 | 4.99 |
| 25.0 - 30.0 | 644.64 | 10.83 | 13.82 | 2.09 | 0.58 |
| 30.0 - 35.0 | 676.70 | 10.54 | 13.80 | 1.20 | 0.26 |
| 35.0 - 40.0 | *NA* | *NA* | *NA* | *NA* | 0.00 |
| 40.0 - 45.0 | *NA* | *NA* | *NA* | *NA* | 0.00 |
| 45.0 - 50.0 | 1150.00 | 12.20 | 15.00 | 1.00 | 0.05 |
| | | | | | |
| **Farms** | | | | | |
| 0.0 - 5.0 | 163.43 | 12.2 | 2.85 | 1.71 | 16.82 |
| 5.0 - 10.0 | 300.70 | 5.35 | 3.36 | 1.57 | 45.60 |
| 10.0 - 15.0 | 441.33 | 7.63 | 3.96 | 1.28 | 28.25 |
| 15.0 - 20.0 | 590.69 | 9.11 | 4.78 | 1.20 | 6.83 |
| 20.0 - 25.0 | 745.11 | 9.80 | 5.00 | 1.36 | 1.84 |
| 25.0 - 30.0 | 718.90 | 12.20 | 5.00 | 3.00 | 0.66 |

Since the location of single trees was the most important variable of this study, extra attention was paid to it during data processing. Firstly, all trees with incorrect X and/or Y coordinates were removed from the data set. Besides, some trees were recorded in the wrong CRS, namely WGS 84. After some control measurements, where the same trees were measured in CRS WGS 84 as well as CRS WGS 84 / UTM zone 37S, there was neither an observable relation nor a large distinction between both CRS. Therefore, reprojection of the location of these trees to CRS WGS 84 / UTM zone 37S using QGIS 3.16.16 was sufficiently accurate. Additionally, the GPS had an accuracy of 3 m, hence, based on the image data, data points in the farms were as much as possible manually relocated to the centre of the tree crown. In consequence, this may affect the final classification results. This was not possible in the plantation due to the sometimes high density resulting in overlapping crowns. Furthermore, trees with a crown area less than 0.25 m$^2$ were removed from the data set, since they were not detectable as a result of the spatial resolution of the image data. This was also done for trees with considerable crown areas according to the field inventory, yet undetectable on the image data. Finally, the ground truth data set included 2408 theoretically detectable trees in the plantation and 765 in the farms. Table 3.3 gives an overview of the frequency of each tree species in the data set per afforestation layout. Since Fassnacht et al. (2016) has recommended the use of a completely independent test set for tree species classification studies, the final data set was randomly split into two thirds training data and one third test data (Section 4.2.2).

Table 3.3: Frequency (%) of each tree species in the data set per afforestation layout.

| Species | | Frequency in field data (%) | |
| --- | --- | --- | --- |
| Scientific name | Common name | Plantation | Farms |
| *Melia volkensii* | Mukau | 78.91 | 100.00 |
| *Azadirachta indica* | Neem | 17.64 | 0.00 |
| Acacia sp. | Acacia | 3.32 | 0.00 |
| Adansonia sp. | Baobab | 0.12 | 0.00 |

## 3.3 A method to differentiate *Melia volkensii*

An object-based classification approach was used to differentiate *Melia volkensii* from its surroundings and other tree species at the individual tree crown level in an agroforesty layout, based on the pan-sharpened Pléiades imagery. Object-based image analysis (OBIA) was chosen because it can deal with the high spectral variability associated with very high resolution (VHR) data (Benz et al., 2004). Moreover, this spectral variability is often augmented when the image objects to be classified are larger than the pixel size, which is frequently the case when using VHR imagery for tree species classification (Karlson et al., 2016). Additionally, OBIA has produced significantly more accurate tree species classification results compared to pixel-based approaches, especially when applied in regions with small-scale farming (Cánovas-García and Alonso-Sarría, 2015), due to the reduced salt-and-pepper effects (Immitzer et al., 2012), and the availability of shape and topological features on top of more meaningful spectral and textural features (Benz et al., 2004) (Section 2.1).

The first step of differentiating *Melia volkensii* in the farms concentrated on masking out the candidate tree crown (CTC) class from other land cover classes (i.e. non-tree crown (NTC) class). Several sub-steps were considered to mask the CTC class with a minimum error in inclusion of all CTC class objects and exclusion of the regions belonging to the other classes (Vahidi et al., 2018). After extracting most of the NTC class objects, supervised classification with Random Forest was conducted. The classifier was applied twice. First for extracting the remaining NTC class objects from the CTC class, and second for differentiating *Melia volkensii* from other tree species. An overview of this workflow is given in Figure 3.4.
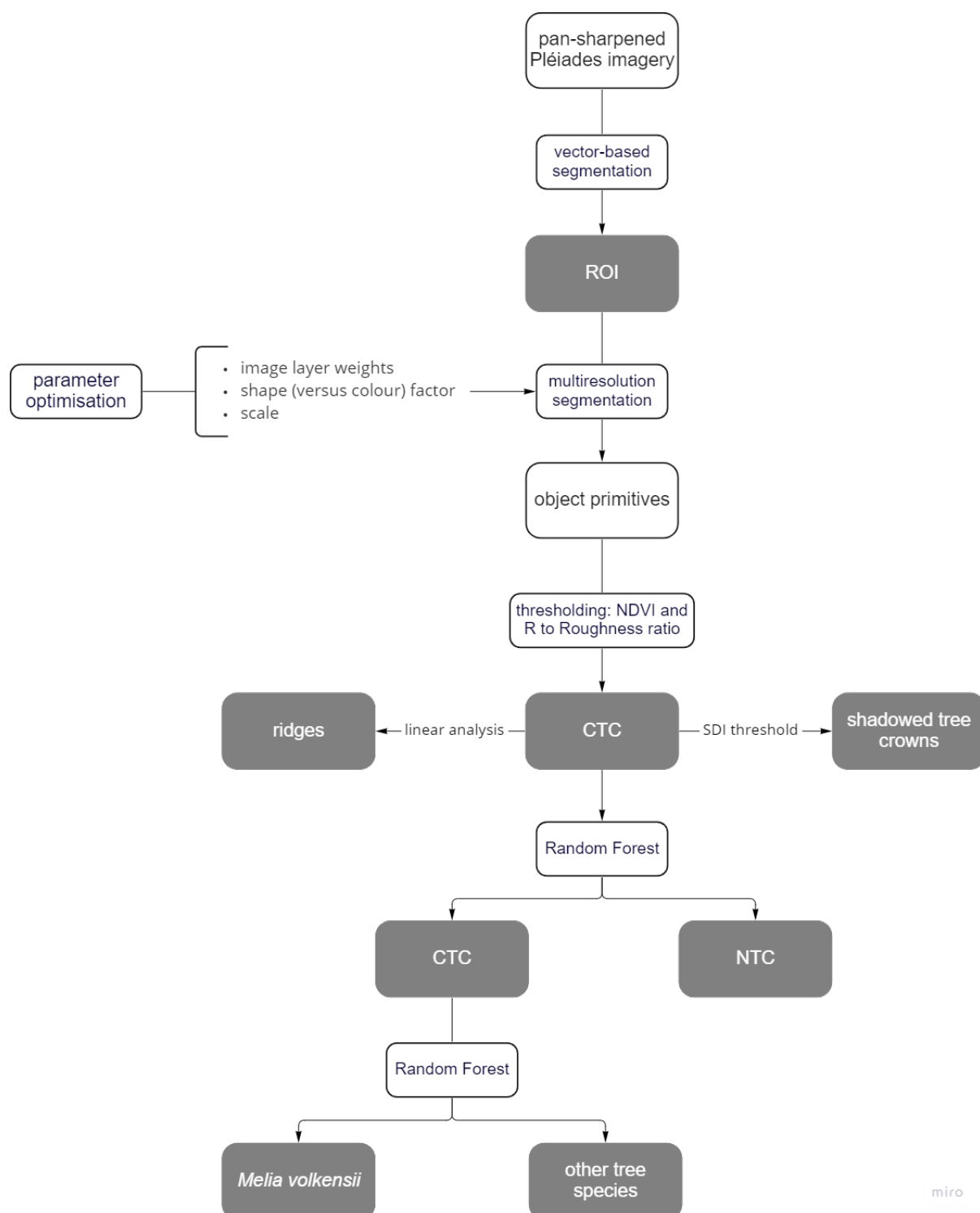
Figure 3.4: Workflow of the methodology used to differentiate *Melia volkensii*. The classes used during the process are shown in grey.

### 3.3.1  Extracting the non-tree crown (NTC) class

A CTC mask within the farms was created by extracting NTC class objects in several steps. Firstly, the regions of interest (ROI) (i.e. the 17 farms where ground truth data was collected) were extracted from the image data. Secondly, multiresolution segmentation was performed after optimisation of the segmentation parameters. Finally, CTC class objects were masked out from NTC class objects in several sub-steps.

**Masking regions of interest (ROI)**

During this step, clouds were extracted from the satellite imagery using a cloud mask that was delivered together with the Pléiades imagery, but it was manually refined in QGIS. Besides, vector layers of the ROI were created and merged with the cloud mask. The resulting vector layer was used for a vector-based segmentation in eCognition® Developer version 10.2 to create an object level containing all farms without cloud coverage.

**Multiresolution segmentation**

Multiresolution segmentation requires the specification of four key parameters, namely image layer weights, shape (versus colour), compactness (versus smoothness), and scale. These parameters are often set using trial-and-error experimentation and so rely on a high degree of subjectivity. In this study, some pre-segmentation analysis methods, based on literature, were tested to reduce subjectivity (Section 2.1.1).

A principal component analysis (PCA) of the image data was executed in Matlab 9.11.0 to determine the image layer weights. Furthermore, the shape (versus colour) factor was optimised using the variance texture enhancement operator (Equation 2.1), following Munyati (2018), to determine whether the different tree species were more separable by reflectance (i.e. colour) or by texture (i.e. shape). In the farms there was only data collected of one tree species, namely *Melia volkensii*. Because of that, the shape factor could not be determined based on the variance texture enhancement operator. Nonetheless, considering that the climatic and edaphic characteristics of the farms are similar to these of the plantation (Section 3.1), the shape factor was optimised based on the ground truth data collected in the plantation. The original spectral bands were uploaded into Matlab to calculate the operator in 3 x 3 kernels, resulting in a texture image for each spectral band. Afterwards, the reflectance values of the training data were determined in QGIS for both the untransformed and the texture-enhanced images. The resulting shapefile was used to generate descriptive statistics such as mean and coefficient of variation (CV) in Matlab. This was calculated for every untransformed as well as texture-enhanced image, and for every tree species. These statistics were then used to compare both images and derive the shape (versus colour) factor based on the average range in reflectance mean values of all bands for both images. Besides, the compactness (versus smoothness) factor was set to 0.4 as vegetation has in general more transitional edges, but due to the sometimes large variation in crown density, this factor was not weighted too large.

The 'Estimate Scale Parameter' (ESP) tool developed by Drăguţ et al. (2014) allows an automated estimation of the optimal scale parameter. However, two important shortcomings were identified while running this tool: (1) a long processing time for extended areas, and (2) not being able to target specific areas of interest. Instead, the optimal scale parameter was estimated by maximising the internal homogeneity of the image objects, calculated as the average local variance, and the external heterogeneity between the image objects, calculated as the spatial autocorrelation, using the methodology proposed by Kim et al. (2008). Given that this method is only applicable at one single layer simultaneously, the NIR band was selected over the visible bands (i.e. B, G, and R) to perform the calculations, because visible radiation is predominantly absorbed by leaves. On the contrary, NIR can reveal vegetation differences as it can transmit through the top layer of the canopy, reflect off lower layers, and then transmit back through the canopy to the sensor, revealing the variability below the top of the canopy (Campbell and Wynne, 2011). First, a rough cut of scale parameters (15 - 20 - 25 - 30 - 35 - 40 - 45 - 50) was made. These scales were chosen after some preliminary segmentations, because it was established that their range extend from over- to under-segmentation. For reducing processing time, the segmentations were executed in a hierarchical manner, which means that the information obtained in segmentations with lower scale parameters were used for segmentations with higher scale parameters (Cánovas-García and Alonso-Sarría, 2015). After running the multiresolution segmentation algorithm in eCognition® Developer with these scale parameters, the obtained image objects were uploaded into QGIS. For each image object, the mean and standard deviation of their spectral reflectance values were computed. Afterwards, the average local variance was calculated as the average of the standard deviation of all image objects and graphed as a function of the associated scale parameters. Additionally, spatial autocorrelation of the image objects was computed based on the global Moran's I index. This index measures, on average, how similar an image object is to its neighbours, hence it is a reliable indicator of statistical separation between spatial objects (Fotheringham et al., 2000). The index was calculated from the mean values using ArcMap 10.8.1. The contiguity matrix was calculated from the squared inverse Euclidean distance between image object centroids. The resulting indices were also graphed as a function of the associated scale parameters. Eventually, the optimal scale should occur just before flattening of the average local variance curve and at the scale associated with the lowest, and even negative, values of the Moran's I index (Kim et al., 2008).

After running the multiresolution segmentation algorithm with the optimised parameter set, object primitives were obtained, which did not yet have any meaning. In the following steps, the object primitives were assigned to more semantically relevant classes in order to match them with real-world objects of interest, namely individual tree crowns.

**Masking the candidate tree crown (CTC) class**

The first step of endowing the object primitives with meaning consisted of masking out the CTC class objects from NTC class objects, particularly understory vegetation (i.e. grass and shrubs), using several sub-steps following Vahidi et al. (2018). In absence of height information, the elevated vegetation layer (rough texture) was separated from the understory vegetation (smooth texture) by performing edge extraction Lee Sigma filtering for detecting both bright and dark edges in the R band (sigma value = 5). A new band (LeeSigmaSum)

was generated by adding the bright edge Lee Sigma band into the dark edge Lee Sigma band. Then, the Gaussian smoothing filter with a kernel size of 25 x 25 pixels was applied to the LeeSigmaSum band, resulting in the Roughness band. The CTC class was separated from the NTC class by thresholding on both the R to Roughness ratio and the normalised difference vegetation index (NDVI) (Equation 3.1) (Vahidi et al., 2018).

$$NDVI = \frac{NIR - R}{NIR + R} \qquad (3.1)$$

High NDVI values result from a combination of a high reflectance in the NIR and lower reflectance in the R band, which is typical for the spectral signature of vegetation, while non-vegetated areas, such as bare soil and open water, will have much lower NDVI values. This is why NDVI is widely used for vegetation monitoring and assessment (Lillesand et al., 2015). The threshold values for both the R to Roughness ratio and the NDVI were determined with the automatic threshold algorithm in eCognition$^{\circledR}$ Developer. This algorithm detects the best threshold value where the histogram of the layer can be split into dark and bright parts. The weighted standard deviation for both the dark and bright part is calculated and compared with the overall standard deviation. The maximum value found represents the threshold value associated with the best differentiation between these dark and bright parts (User Guide eCognition$^{\circledR}$ Developer, 2019). Finally, the image objects obtained after applying the multiresolution segmentation algorithm were assigned to the CTC class when they met both conditions in Table 3.4. However, after applying these conditions, the CTC mask was not yet precise enough because it still included some ridges, shadows, and bare soil. Therefore, additional conditions were implemented for further refinement of the mask.

Ridges were extracted using a linear analysis based on the object shape features *compactness* and *length/width*. The threshold values for both features were set after visual assessment (Table 3.4). However, some tree crowns also included linear objects. Therefore, additional features were added to reassign these objects to the CTC class. For this, the topological features *relative border to CTC* and *relative area of CTC* were chosen. Their threshold values were also set after visual assessment (Table 3.4).

Shadow areas appear with different intensity, brightness, and lower contrast compared to sunlit regions, as well as unwanted contrast in the borders of shadow overcast areas (Tatar et al., 2018). In view of this, shadow was masked, because it may severely affects visual interpretation and automatic identification of ground features on VHR satellite imagery, hence it may also affect the quality and the accuracy of the classification results (Duan et al., 2013; Mostafa and Abdelhafiz, 2017). First, pixel-level shadow detection was performed using multi-threshold segmentation based on the shadow detector index (SDI) (Equation 3.2). This index was chosen because Mostafa and Abdelhafiz (2017) has established that the rate of truly detected pixels is higher compared to other state-of-the-art indices, such as the $C_3$-index. Besides, the index shows stability with vegetation regions (Mostafa and Abdelhafiz, 2017).

$$SDI = \frac{(1 - PC1) + 1}{((G - B) * R) + 1} \qquad (3.2)$$

with *G*, *B*, and *R* are normalised components of respectively the G, B, and R band, normalisation of these bands was performed in Matlab. Furthermore, *PC 1* is a normalised

component of the first principal component (PC), wherefore all four spectral bands were employed. The listed bands were used in a way to maximise the fraction numerator and to minimise the fraction denominator. Finally, large values of the index were obtained where shaded regions occur (Mostafa and Abdelhafiz, 2017). The threshold value of the SDI index was determined with the automatic threshold algorithm (Table 3.4). Afterwards, CTC class objects were assigned to the shadow class based on the topological feature *majority vote area*. This feature classifies a super-object based on the classification majority of his sub-objects, whereby the majority class is returned based on the largest area of the sub-objects (User Guide eCognition® Developer, 2019).

Table 3.4: Indices and features with their corresponding threshold values used to separate the candidate tree crown (CTC) class from ridges and shadow.

| Class | Index or feature | Threshold |
|---|---|---|
| CTC | R to Roughness | < 55 |
|  | NDVI | > 0.24 |
|  |  |  |
| Ridges | Compactness | < 0.4 |
|  | Length/width | > 3 |
|  | Relative border to | > 0.6 |
|  | Relative area of | > 0.95 |
|  |  |  |
| Shadow | SDI | > 3.01 |

During the final step of masking the CTC class, the object primitives were merged using the spectral difference segmentation algorithm in order to improve their spatial correspondence with real landscape objects (Cánovas-García and Alonso-Sarría, 2015). Based on visual assessment, the maximum spectral difference was set to 70 because it was established that this value attenuates most of the over-segmentation obtained after running the multiresolution segmentation algorithm. Because of this, the number of image objects was reduced, and these newly formed image objects were more likely to represent actual tree crowns. The resulting image objects with an area of more than 1000 pixels, corresponding to approximately 250 m$^2$, were removed because the objects of interest will probably not have such large areas, for instance, the largest measured tree crown had an area of approximately 94 m$^2$. Considering the fact that some tree crowns might be merged into one image object, the threshold value was set sufficiently high so these trees would not be missed.

### 3.3.2 Supervised classification

The supervised machine learning algorithm Random Forest was applied for the classification of the CTC mask, obtained at the end of the previous section. After visual assessment of the mask, it was decided to run the Random Forest algorithm twice. First for an additional refinement of the mask by extracting the last remaining NTC class objects as much as possible, and second for classifying the resulting CTC objects into *Melia volkensii* and other tree species. In both cases, the algorithm was applied three times using different feature sets: (1) spectral features only, (2) a combination of spectral and textural features, and (3) a combination of spectral, textural, and shape features. Lastly, the accuracies of the best performing classifiers were assessed.

Reference data is needed to run the Random Forest algorithm. The field inventory provided data concerning the occurrence of *Melia volkensii* (Section 3.2.2), but reference data for the remaining NTC class objects and other tree species were missing. Therefore, additional reference data were manually selected within the CTC mask, based on the Pléiades satellite imagery in QGIS. The NTC class included ground features such as bare soil, grass, and small bushes, whereas the other tree species class included several tree species and some larger bushes that were difficult to distinguish from trees on the satellite imagery. It has to be noted that only *Melia volkensii* trees that are property of BGF were included in the ground truth data set obtained after the field inventory. However, local farmers planted some *Melia volkensii* trees themselves, so there exists a significant possibility that some of these trees were included in the class of the other tree species as they are not visually distinguishable based on the satellite imagery. Finally, 139 NTC class objects and 180 objects belonging to the other tree species class were included in the reference data set.

**Feature calculation**

Prior to classification, the following object features were generated in eCognition$^{\circledR}$ Developer, only intrinsic object features were calculated since these are available for each image object:

- Spectral features

  - Mean band value: the mean value represents the mean brightness of an image object within a single band (Mishra and Crews, 2014). This feature is generated for every spectral band (i.e. B, G, R, and NIR).

  - Standard deviation (StdDev): standard deviation of all pixels which form an image object within a band (Mishra and Crews, 2014). This feature is also generated for every spectral band (i.e. B, G, R, and NIR).

  - Brightness: this value represents the mean value of the spectral mean values of the four spectral bands (i.e. B, G, R, and NIR) of an image object (Van Coillie et al., 2007).

  - Maximum difference (max. diff.): minimum value of an image object subtracted from its maximum value. The means of all bands belonging to an image object are compared with each other and the result is divided by the brightness (Mishra and Crews, 2014).

  - Band ratios

    * Enhanced vegetation index (EVI) (Equation 3.3): this index was developed to optimise the vegetation signal with improved sensitivity in high biomass regions and improved vegetation monitoring through a decoupling of the canopy background signal and a reduction in atmospheric influences.

    $$EVI = G * \frac{NIR - R}{NIR + (C1 * R) - (C2 * B) + L} \tag{3.3}$$

    with *G* is a gain term, *C1* and *C2* are the coefficients of the aerosols resistance term, which uses the B band to correct for aerosol influences in the R band,

and *L* is a soil-adjustment factor (Jiang et al., 2008). The parameters *G*, *C1*, *C2*, and *L* were set to their default values, respectively 2.5, 6, 7.5, and 1.

 ∗ NDVI (Equation 3.1)
 ∗ Normalised difference water index (NDWI) (Equation 3.4): this index, proposed by McFeeters (1996), was designed to highlight open water while eliminating the presence of soil and terrestrial vegetation. The former returns positive values, whereas the latter has zero or negative values (McFeeters, 1996).

$$NDWI = \frac{G - NIR}{G + NIR} \tag{3.4}$$

 ∗ Ratio vegetation index (RVI) (Equation 3.5): this index was developed for the estimation and monitoring of vegetation cover. It enhances the contrast between the ground and vegetation, but the index is sensitive to atmospheric effects and its discriminating power is weak when the vegetation cover is sparse, while the performance of the index increases in densely vegetated areas (Bannari et al., 1995).

$$RVI = \frac{NIR}{R} \tag{3.5}$$

 ∗ Soil-adjusted vegetation index (SAVI) (Equation 3.6): this index was designed to compensate for a soil background in sparsely vegetated areas (Lillesand et al., 2015).

$$SAVI = \frac{(1 + L)(NIR - R)}{(NIR + R + L)} \tag{3.6}$$

with *L* is a soil-adjustment factor. This factor was set to 0.5 for it was found to minimise soil brightness variations and eliminate the need for additional calibration to adjust for different soils (Huete, 1988).

- Textural features: first and second order texture measures, including angular second moment (ang. 2nd moment), contrast, correlation, dissimilarity, entropy, homogeneity, mean, and standard deviation, in four spatial directions (i.e. 0°, 45°, 90°, and 135°) derived from:

   – Grey-level co-occurrence matrix (GLCM): describes how different combinations of pixel values occur within an image object (Mishra and Crews, 2014).
   – Grey-level difference vector (GLDV): sum of the diagonals of the GLCM (Mishra and Crews, 2014).

- Shape features: shape index (i.e. the border length of an image object divided by four times the square root of its area (Van Coillie et al., 2007)), length/width, area (pixels), length (pixels), and number of edges.

In total, 68 features were calculated of which 15 spectral features, 48 textural features, and 5 shape features.

**Random Forest**

First, it was tried to apply Random Forest in eCognition® Developer. The classifier is embedded in the supervised classification algorithm and allows to specify several parameters

including the number of decision trees, the number of input features, and the maximum depth of the forest. However, these hyperparameters (i.e. parameters that do not learn from the data (Pramoditha, 2020)) can only be tuned through some trial-and-error experimentation. Moreover, eCognition$^®$ Developer does not allow the application of an iterative data-splitting approach such as K-fold cross-validation (i.e. one of the best re-sampling approaches with limited input data, ensuring that every observation from the original dataset has the chance of appearing in the training and test set (Sanjay, 2018)). Lastly, feature selection has to be performed manually and is a very time-consuming process. This is why the Random Forest classification was eventually implemented using the caret package in RStudio 4.1.3.

A grid search (i.e. a hyperparameter optimisation technique) was performed to tune the two most important hyperparameters of Random Forest, namely the number of decision trees for executing the classification (ntree), and the number of input features used at each tree node (mtry). During this procedure, an optimal combination of hyperparameter values was searched so the best performing model would be obtained. The Random Forest classifier was trained for each combination of hyperparameter values and then evaluated through repeated K-fold cross-validation with ten folds and three repeats. K-fold cross-validation split the data set randomly into ten independent folds without replacement, of which nine folds were used for model training and one fold for evaluation of the performance. The model's performance was evaluated as the Cohen's Kappa coefficient (see Accuracy assessment). This procedure was run ten times and a performance estimate for each iteration was obtained (Pramoditha, 2020). During repeated K-fold cross-validation, the cross-validation procedure was repeated an additional three times and the mean performance across all folds from all runs was reported (Brownlee, 2020). The grid search stored all performance estimates after taking all combinations of hyperparameter values, and selected the optimal combination (Pramoditha, 2020). For mtry, each value between one and 15 was tested in steps of one, while for ntree the values were set to 10, 25, 50, 100, 200, and 500. This procedure was executed for all three feature sets. Finally, the feature set with the highest performance after the grid search was chosen for further analysis.

As described in Section 2.1.2 the classification algorithm might overfit when it faces a sparse number of training samples compared to the number of predictors. Feature selection may prevent this and even improve classification accuracy. Therefore, recursive feature elimination was performed on the feature set selected after the grid search. Recursive feature elimination is a backward selection procedure to identify the optimal combination of features. The Random Forest model was built based on all features and the importance of each feature was calculated as Mean Decrease Gini (MDG). This measure is calculated within the model and indicates how each feature contributes to the homogeneity of the nodes and leaves in the resulting Random Forest. The higher the value, the higher the importance of the feature in the model (Martinez-Taboada and Redondo, 2020). During the procedure, all features were rank-ordered and the lowest-ranked features were successively removed based on the model's performance, evaluated as the Cohen's Kappa coefficient. This process continued until a smaller subset of features with the highest performance was retained (Bulut, 2021). The initial ranking was used throughout all of the iterations to avoid model overfitting which may arise when feature importance is recalculated after each iteration (Svetnik et al., 2004; Karlson et al., 2016). Additionally, repeated K-fold

cross-validation with ten fold and three repeats was used to improve the performance of the feature selection procedure (Bulut, 2021).

This procedure was applied to the CTC mask, obtained at the end of Section 3.3.1, for extracting the remaining image objects representing the NTC class. The optimal model was trained using a subset of the reference data containing 93 NTC objects and 104 CTC objects of which the latter is a combination of data concerning both *Melia volkensii* and other tree species. The resulting CTC class was further classified into the classes *Melia volkensii* and other tree species by applying the same procedure. For this classification, 300 *Melia volkensii* objects and 110 objects of the other tree species class were used during the training stage.

**Accuracy assessment**

After fitting the optimised model to the training data, the classification accuracy was assessed based on a confusion matrix generated using an independent test set that was not used during the training stage of the Random Forest model. This was done to test the suitability of the proposed methodology for application in other areas, unknown by the classifier, where BGF might be interested in. The independent test set for the evaluation of the CTC/NTC classification contained 46 NTC objects and 58 CTC objects. The test set for evaluating the model's performance for the differentiation of *Melia volkensii* contained 126 *Melia volkensii* objects and 47 objects of the other tree species class. The confusion matrix compared the relationship between these known test set objects and the corresponding results of the model (Lillesand et al., 2015). Based on this matrix, several accuracy measures were calculated:

- Producer's accuracy: indicates how well test set objects of the given class are classified. It is calculated by dividing the number of correctly classified image objects in each class (i.e. major diagonal) by the number of test set objects used for that class (i.e. column total) (Lillesand et al., 2015). In other words, the producer's accuracy describes how well a specific area can be mapped (Story and Congalton, 1986).

- User's accuracy: indicates the probability that an image object classified into a given class actually represents that class on the ground (Lillesand et al., 2015), so the user's accuracy describes how well a map represents what is really on the ground, and thus is a measure of its reliability (Story and Congalton, 1986). This measure is calculated by dividing the number of correctly classified image objects in each class (i.e. major diagonal) by the total number of image objects that were classified in that class (i.e. row total) (Lillesand et al., 2015).

- Overall accuracy: this measure is calculated by dividing the total number of correctly classified image objects (i.e. major diagonal) by the total number of test set objects (Lillesand et al., 2015).

- Cohen's Kappa coefficient: measure of the difference between the actual observed agreement between reference data and an automated classifier and the chance agreement between the reference data and a random classifier. This coefficient is an indicator of the extent to which percentage of the correct classified image objects are

due to "true" agreement versus "chance" agreement (Lillesand et al., 2015). Cohen's Kappa is a standardised value and thus is interpreted the same across multiple studies (McHugh, 2012). Landis and Koch (1977) has suggested benchmarks for the qualitative interpretation of the relative strength of agreement: values lower than or equal too 0 as indicating no agreement, 0.01.-0.20 as none to slight, 0.21-0.40 as fair, 0.41-0.60 as moderate, 0.61-0.80 as substantial, and 0.81-1.00 as almost perfect agreement (McHugh, 2012). McHugh (2012) has argued that this interpretation allows very little agreement to be described as substantial, and that accepting a Cohen's Kappa coefficient of 0.40 as moderate, and thus an adequate agreement, is not acceptable. Therefore, another interpretation of Cohen's Kappa has been suggested (Table 3.5), keeping in mind that any agreement less than perfect is also a measure of disagreement. A simplified interpretation of Cohen's Kappa, according to McHugh (2012), is that any Cohen's Kappa coefficient below 0.60 indicates inadequate agreement and therefore little confidence in the study results should be placed.

Table 3.5: Interpretation of the Cohen's Kappa coefficient according to McHugh (2012).

| Value of Kappa | Level of agreement | Data that are reliable (%) |
| --- | --- | --- |
| 0.00 - 0.20 | None | 0 - 4 |
| 0.21 - 0.39 | Minimal | 4 - 15 |
| 0.40 - 0.59 | Weak | 15 - 35 |
| 0.60 - 0.79 | Moderate | 35 - 63 |
| 0.80 - 0.90 | Strong | 64 - 81 |
| above 0.90 | Almost perfect | 82 - 100 |

# 4.  <u>RESULTS AND DISCUSSION</u>

In this chapter, the results of the different steps of the methodology developed to differentiate *Melia volkensii* from its surroundings and other tree species will be presented and discussed. First, the optimisation of the parameters of the multiresolution segmentation algorithm, namely image layer weights, shape (versus colour), and scale, is discussed. Then, a description of the candidate tree crown (CTC) mask and a few of its shortcomings is provided. Finally, the results of the supervised classification with Random Forest are discussed to provide an answer on the research questions stated in Chapter 1.

## 4.1  Optimisation of the multiresolution segmentation algorithm

### 4.1.1  Determining image layer weight values

Principal component (PC) 1 explained most of the data variance (> 98 %), and thus possessed most of the information. All spectral bands loaded into PC 1 with high absolute values (> 0.99), meaning that this PC contained a lot of information from all original bands (i.e. red (R), blue (B), green (G), and near infrared (NIR)) (Table 4.1). Moreover, all bands were highly correlated with each other (> 0.97) (Table 4.2). Based on these results, the image layer weights for each original spectral band were set to one.

Table 4.1: Loadings of each spectral band (i.e. blue (B), green (G), red (R), and near infrared (NIR)) in the principal components and the percentage (%) data variance they represent.

| | Principal component | | | |
|---|---|---|---|---|
| **Band** | **1** | **2** | **3** | **4** |
| B | -0.9958 | 0.0826 | 0.0402 | 0.0028 |
| G | -0.9996 | 0.0180 | -0.0151 | -0.0136 |
| R | -0.9987 | 0.0302 | -0.0391 | 0.0089 |
| NIR | -0.9912 | -0.1315 | 0.0142 | 0.0020 |
| | | | | |
| % data variance | 99.1234 | 0.8245 | 0.0389 | 0.0025 |

Table 4.2: Correlations between each spectral band (i.e. blue (B), green (G), red (R), and near infrared (NIR)).

| Band | B | G | R | NIR |
|------|------|------|------|------|
| B | 1 | 0.9962 | 0.9955 | 0.9767 |
| G | 0.9962 | 1 | 0.9994 | 0.9882 |
| R | 0.9955 | 0.9994 | 1 | 0.9854 |
| NIR | 0.9767 | 0.9882 | 0.9854 | 1 |

## 4.1.2 Determining shape and colour weight values

Table 4.3 contains the reflectance statistics, mean and coefficient of variation (CV), for both the untransformed as well as the texture-enhanced image. The range in reflectance mean values within all four bands was largest for the untransformed image. This implicated that the different tree species (i.e. *Melia volkensii* (mukau), *Azadirachta indica* (neem), and *Acacia* sp. (acacia)) were more separable by reflectance (i.e. colour) than by texture (i.e. shape). According to Munyati (2018), higher CV values represent higher within-class variability. The CV values of the untransformed image were higher than these of the texture-enhanced image, so this was an additional argument to weigh colour higher than shape. Moreover, Pu et al. (2018) has pointed out that several existing studies have demonstrated that a higher weight coefficient for the colour criterion is frequently used to extract more meaningful objects. The average of the range in reflectance mean values of the untransformed image (859.2) was 36 % higher than the average of the range of the texture-enhanced image (311.7). As a result, the colour weight value was set to 0.73. Consequently, the shape factor was equal to 0.27, because these values are inversely related to each other and sum to one.

Table 4.3: Top-of-atmosphere (TOA) reflectance statistics, mean and coefficient of variation (CV) of each spectral band (i.e. blue (B), green (G), red (R), and near infrared (NIR)), and each tree species (i.e. mukau (*Melia volkensii*), neem (*Azadirachta indica*), and acacia (*Acacia* sp.)) for both the untransformed image as well as the texture-enhanced image. The range in reflectance mean values within one band is also calculated.

| | | TOA reflectance statictics: mean and CV | | | | | | | |
|---------|------|--------|--------|-------|--------|--------|--------|--------|--------|
| | | Untransformed image | | | | Texture-enhanced image | | | |
| Species | | B | G | R | NIR | B | G | R | NIR |
| Mukau | Mean | 1050.2 | 1015.6 | 765.2 | 2401.7 | 2680.2 | 2348.1 | 2038.0 | 3436.0 |
| | CV | 579.7 | 731.4 | 635.0 | 974.4 | 90.1 | 93.6 | 96.7 | 84.9 |
| Neem | Mean | 832.3 | 856.3 | 650.4 | 2464.7 | 5073.4 | 4546.7 | 4065.9 | 5947.5 |
| | CV | 328.9 | 466.4 | 387.6 | 526.5 | 111.3 | 109.7 | 108.6 | 110.6 |
| Acacia | Mean | 834.7 | 916.2 | 718.5 | 2525.1 | 1812.0 | 1332.1 | 1222.5 | 1760.9 |
| | CV | 637.1 | 1076.3 | 917.7 | 1768.5 | 129.5 | 136.0 | 131.5 | 103.6 |
| Range | Mean | 792.5 | 843.1 | 850.0 | 951.1 | 357.2 | 293.0 | 300.7 | 296.1 |

## 4.1.3 Determining the scale parameter

Figure 4.1 (top) was obtained after determining the average local variance and the spatial autocorrelation associated with the scales varying from 15 to 50 in steps of five. When

the scale parameter increased, the average local variance tended to increase, while the spatial autocorrelation tended to decrease. This means that as the average size of the image objects increased, the image objects tended to consist of more dissimilar pixels, and became less similar to neighbouring image objects (Johnson and Xie, 2011). Following the methodology of Kim et al. (2008), it was expected that the optimal scale would occur just before flattening of the average local variance curve, and at the scale associated with the lowest spatial autocorrelation. However, neither of these requirements was met. Therefore, scales varying from 50 to 400 in steps of 50 were tested (Figure 4.1 (bottom)). The average local variance curve flattened around the scale of 300, whereas the spatial autocorrelation still continued to decrease at the scale of 400, and thus did not reach its lowest value yet. According to Johnson and Xie (2011), the spatial autocorrelation will likely continue to decrease until the image objects become large enough to contain a mixture of many different land cover types. Once all image objects become mixed, the spatial autocorrelation should start to increase again. Despite this, the image was already visually under-segmented at the tested scale parameters. The failure of the methodology proposed by Kim et al. (2008) in this area is probably caused by the variability in land cover classes. The study area investigated by Kim et al. (2008) was a more homogeneous forest scene, in which spectral values for the different land cover types were quite similar, whereas the farms in the region of Kiambere were more heterogeneous resulting in different spectral characteristics for different land cover classes (Johnson and Xie, 2011). According to Cánovas-García and Alonso-Sarría (2015), the validity of internal homogeneity (i.e. average local variance) and external heterogeneity (i.e. spatial autocorrelation) criteria to optimise the scale parameter depends heavily on having a uniform study area for the scale parameter to be representative for the whole area.
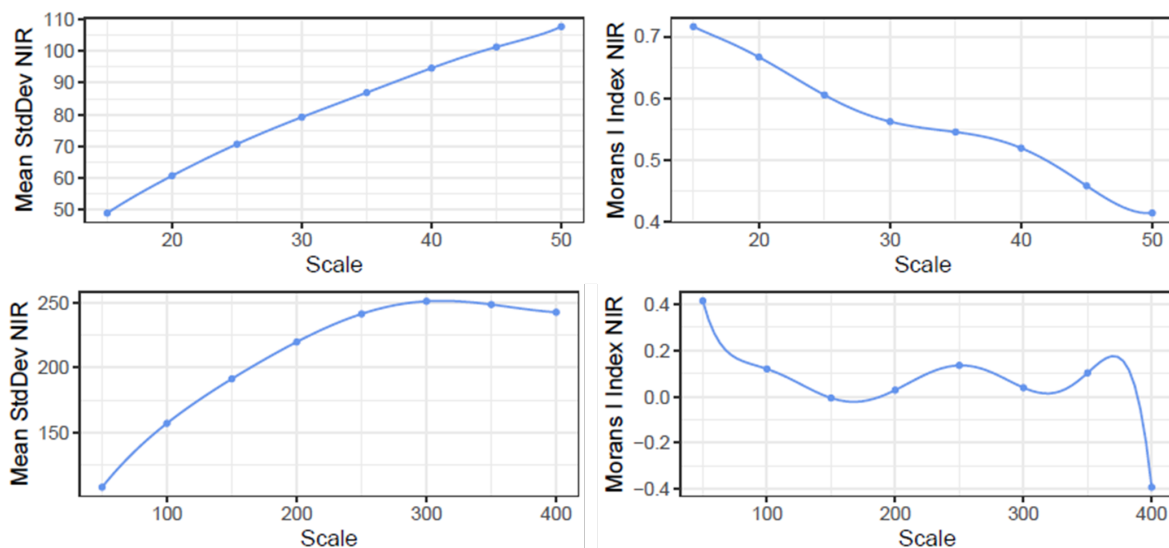


Figure 4.1: Average local variance (left), calculated as the mean standard deviation (Std-Dev) of the near infrared (NIR) band of all image objects, as well as the spatial autocorrelation (right), calculated as the global Moran's I index of the NIR band, for both the scale values between 15 and 50 (top) and the scale values ranging from 50 to 400 (bottom).

Because of the failure of the methodology of Kim et al. (2008), the scale parameter was set to 25 after visual assessment as increasingly more trees, mostly with small crowns, were included in mixed objects when higher scale values were used. At this scale parameter, the in-

dividual tree crowns were in most cases over-segmented, depending on the tree crown size and variation of spectrum/structure within the crown (Pu et al., 2018). Over-segmentation was preferred over under-segmentation, because under-segmented image objects hold little value since they likely contain more than one land cover class, while over-segmentation can easily be handled in the following classification steps (Johnson and Xie, 2011; Cánovas-García and Alonso-Sarría, 2015). Despite the presence of over-segmentation, some trees were still not properly identified at a scale of 25 because they were too small for the spatial resolution of the imagery, or too similar to the neighbouring objects for the radiometric resolution of the sensor (i.e. the ability of a sensor to distinguish between grey-scale values while acquiring an image (Lillesand et al., 2015)). The success of the segmentation of individual trees in an agroforestry layout depends heavily on the reflectivity of bare soil, which is related to the agricultural practices carried out. When there is another crop in the soil underneath or a dense cover of weeds, worse results will be obtained. Additionally, developed canopies without gaps between them make it impossible to separate individual trees and thus the quality of the segmentation may be variable (Cánovas-García and Alonso-Sarría, 2015). This is why, next to individual trees, some crown clusters were included in all further analyses. The scale value of 25 was used for each farm, but, according to Cánovas-García and Alonso-Sarría (2015), a global approach requires some degree of uniformity in the image. Therefore, Cánovas-García and Alonso-Sarría (2015) has suggested the use of uniform spatial units, resulting from pre-existing land use maps, for the optimisation of local scale parameters, so in future research it might be preferable to consider each farm as a different spatial unit as some differences between them probably occur. Nonetheless, the spatial uniformity within the farms is rather low so the success of a local approach is not guaranteed either. Besides, the searching and manually digitising of uniform spatial units within the farms would be time-consuming and run counter to the idea of a semi-automatic method to differentiate *Melia volkensii* (Cánovas-García and Alonso-Sarría, 2015). The use of a multi-scale segmentation approach might be a viable solution for this, in particular when the area under consideration is clearly composed of structures appearing at different scales (e.g. stand level and tree level) (Trias-Sanz et al., 2008). Delineating more or less homogeneous regions using a segmentation algorithm at a coarser scale, followed by a delineation of individual tree crowns at finer scales may improve the segmentation results. One of the prerequisites of the use of a multi-scale approach is the presence of somewhat uniform regions in the area of interest, so it may be easier to perform a segmentation within properly managed units which are clearly delimited on site.

It is important to note that scale values may not be generalisable to other images, since they depend on the radiometric resolution of the sensor used (Cánovas-García and Alonso-Sarría, 2015). In future research, it might be necessary to find more automated solutions for determining the scale parameter, and by extension for the delineation of individual tree crowns, in particular when Better Globe Forestry (BGF) wants to apply the proposed methodology over more extended areas. Automated delineation of individual tree crowns may be a challenging task, complicated by high tree densities and interlocking crowns, or when tree canopies are spectrally similar to the background, but they are essential for large-scale application (Immitzer et al., 2012; Karlson et al., 2016). In literature, there exists a multitude of methods proposed to define the optimal scale parameter with a minimum of trial-and-error experimentation, some more complex than others and some more automated than others. Especially further elaboration of the 'Estimate Scale Parameter'

(ESP) tool, developed by Drăguţ et al. (2014) might be viable, in particular making the tool applicable on a preference object level might be beneficial for tree species classification studies. Besides, in future research, seeded region growing algorithms, instead of the multiresolution segmentation algorithm, might be more suitable for the delineation of individual tree crowns. Seeded region growing, proposed by Adams and Bischof (1994), is a classic image processing method used in tree crown delineation that typically requires treetops as an initial input. For this, the results of Ghyselbrecht (2022) can be used. The algorithm starts from a set of seed points and grows regions until some stop criteria are satisfied (Zhen et al., 2014). Various extensions and modifications have been proposed in literature to improve the performance of seeded region growing (Fan and Lee, 2015), however, since both Master's Dissertations were developed simultaneously, the seeded region growing algorithm was not further explored in this study.

## 4.2 Candidate tree crown (CTC) mask

The candidate tree crown (CTC) mask was composed of two major steps, each consisting of several sub-steps. Firstly, classification rules were drawn up in eCognition® Developer based on literature and expert knowledge. After applying these rules, it was visually established that the CTC mask was not yet precise enough, mostly due to some remaining bare soil objects, so a Random Forest classification was performed for further refinement.

### 4.2.1 Rule-based classification

During the first step of the rule-based classification, most of the CTC objects were masked out from non-tree crown (NTC) objects by thresholding on both the normalised difference vegetation index (NDVI) and the R to Roughness ratio. However, after visual assessment, three shortcomings of the mask were identified, which may influence the final classification results. Firstly, some trees with smaller crowns were not properly identified by the multiresolution segmentation algorithm as described in Section 4.1.3, and therefore were included in mixed objects with bare soil. As a result, these trees did not always meet the conditions to be included in the CTC mask. Secondly, some smaller trees, that were included in pure CTC objects, met either the NDVI condition or the R to Roughness ratio condition, but not both. Consequently, they were also not included in the CTC mask. The criteria were not altered based on expert knowledge because this would make the classification procedure less automated, and thereby make it less reproducible in other areas of interest. The third shortcoming is related to the inclusion of objects belonging to other land cover classes. This error was mainly obviated by implementing extra conditions for the extraction of ridges, shadow, and some remaining bare soil.

Ridges were probably masked because they were vegetated with crops, grasses, and small bushes resulting in higher NDVI values. Moreover, differences in height may have caused a lower R to Roughness ratio as a result of which they were identified as CTC objects. In contrast, ridges have distinctively different shapes than the surrounding tree crowns, namely linear instead of circular, so linear analysis ensured the extraction of these NTC objects. Nonetheless, it also caused some trees, located on the ridges, to be removed because they

were included in the same image object. Furthermore, shadows were extracted because the different intensity, brightness, and lower contrast might hamper the final classification. By only considering the sunlit region of tree crowns, within-species variability was not increased by varying illumination conditions within the tree crown (Immitzer et al., 2012). As a result, many tree crowns were partly removed from the CTC mask. In future research, it would be recommended to apply light normalisation on the shaded tree crowns to preserve the shape of the crown, which may be beneficial in the following processing steps. Pu et al. (2018) has developed two linear algorithms to normalise the spectra in shaded regions to be similar to the spectra in sunlit regions. They have established that these algorithms might have a great potential to mitigate the negative effects of shade on tree species classification. Training samples of both shaded and sunlit image objects within the same tree crown are required to apply this method, but manually selecting these samples may be a time-consuming process. Moreover, this process would have to be repeated when applied to other images because light normalisation algorithms may not be generalisable due to varying illumination conditions during image acquisition. Further research regarding the automatisation of shadow detection and normalisation may be necessary. After eliminating shadows, the remaining bare soil was removed using the machine learning algorithm Random Forest (Section 4.2.2).

## 4.2.2 Classification with Random Forest

The removal of the remaining bare soil objects was executed using a CTC/NTC Random Forest classification. After performing optimisation of the hyperparameters (i.e. the number of input features at each node (mtry) and the number of decision trees (ntree)), with a grid search, for all three feature sets (i.e. (1) spectral features only, (2) a combination of spectral and textural features, and (3) a combination of spectral, textural, and shape features), the optimal combinations were selected and stored in Table 4.4 with their model's performance. The combination of all 68 features reached the highest average Cohen's Kappa coefficient (0.699) after repeated K-fold cross-validation, as well as the lowest out-of-bag (OOB) error (16.24 %) and thus was selected for further analysis. All combinations of hyperparameters and their associated performance, for the selected feature set, were stored in Figure 4.2.

Table 4.4: Optimal combination of hyperparameters (i.e. the number of input features at each node (mtry) and the number of decision trees (ntree)) and the associated model's performance, measured as the average Cohen's Kappa coefficient as well as the out-of-bag (OOB) error (%), obtained after repeated K-fold cross-validation. These values resulted from a grid search of the CTC/NTC Random Forest classification performed for all three feature sets (i.e. (1) spectral features only, (2) a combination of spectral and textural features, and (3) a combination of spectral, textural, and shape features)

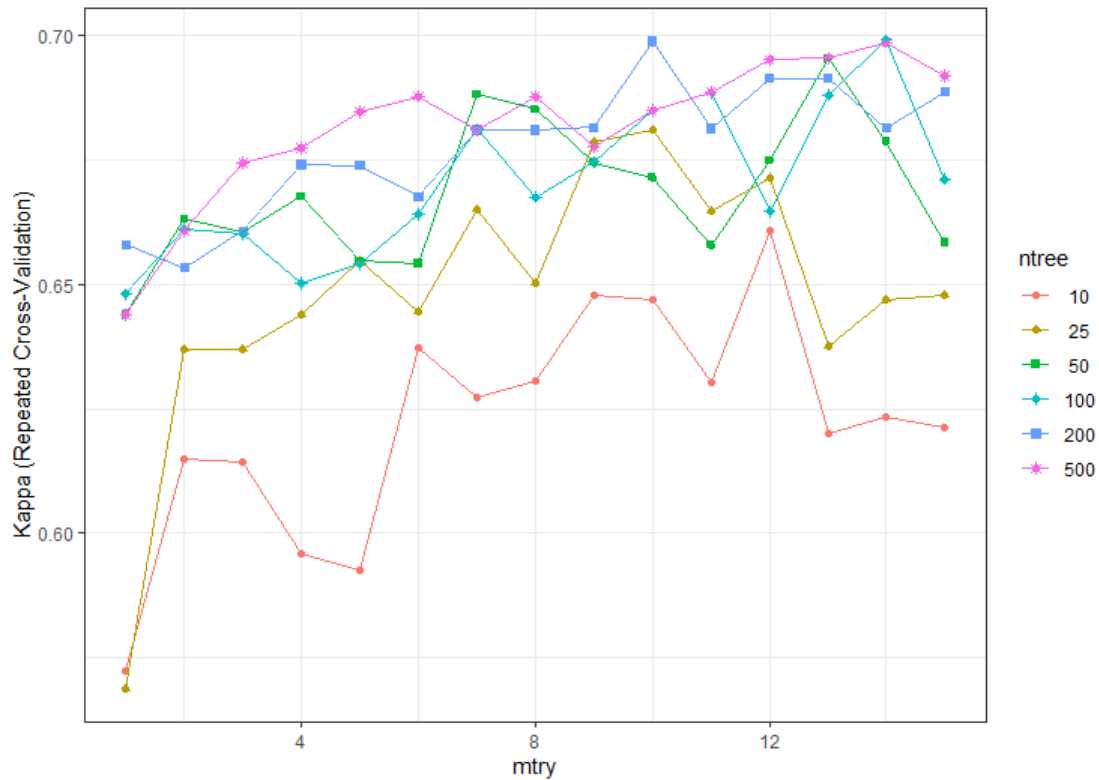| Feature set | ntree | mtry | Cohen's Kappa | OOB error (%) |
|---|---|---|---|---|
| Spectral | 25 | 7 | 0.676 | 18.27 |
| Spectral + texture | 10 | 15 | 0.697 | 18.78 |
| Spectral + texture + shape | 100 | 14 | 0.699 | 16.24 |

Figure 4.2: All combinations of hyperparameters (i.e. the number of input features at each node (mtry) (dots) and the number of decision trees (ntree) (lines)) with their associated model's performance, measured as the Cohen's Kappa coefficient, obtained after repeated K-fold cross-validation. These values resulted from a grid search of the CTC/NTC Random Forest classification performed for a combination of spectral, textural, and shape features.

Recursive feature elimination was performed 68 times for all features present in the spectral, textural, and shape feature set. For each number of features, the model's performance was stored in Figure 4.3 as the average Cohen's Kappa coefficient, obtained after repeated K-fold cross-validation. The highest Cohen's Kappa coefficient (0.707) was reached when 27 features were retained in the model. The recursive feature elimination procedure did not allow for the specification of the hyperparameters and therefore the Random forest was built with the default values, namely 500 decision trees (ntree), and the square root of the total number of features included in the model (mtry). Consequently, mtry was set to five. Based on Figure 4.2, there is little difference in the model's performance between this hyperparameter set (0.685) and the optimal set obtained after the grid search (0.699), hence the model with the default hyperparameters, instead of the optimal combination, is not likely to have a major influence on the final classification results.
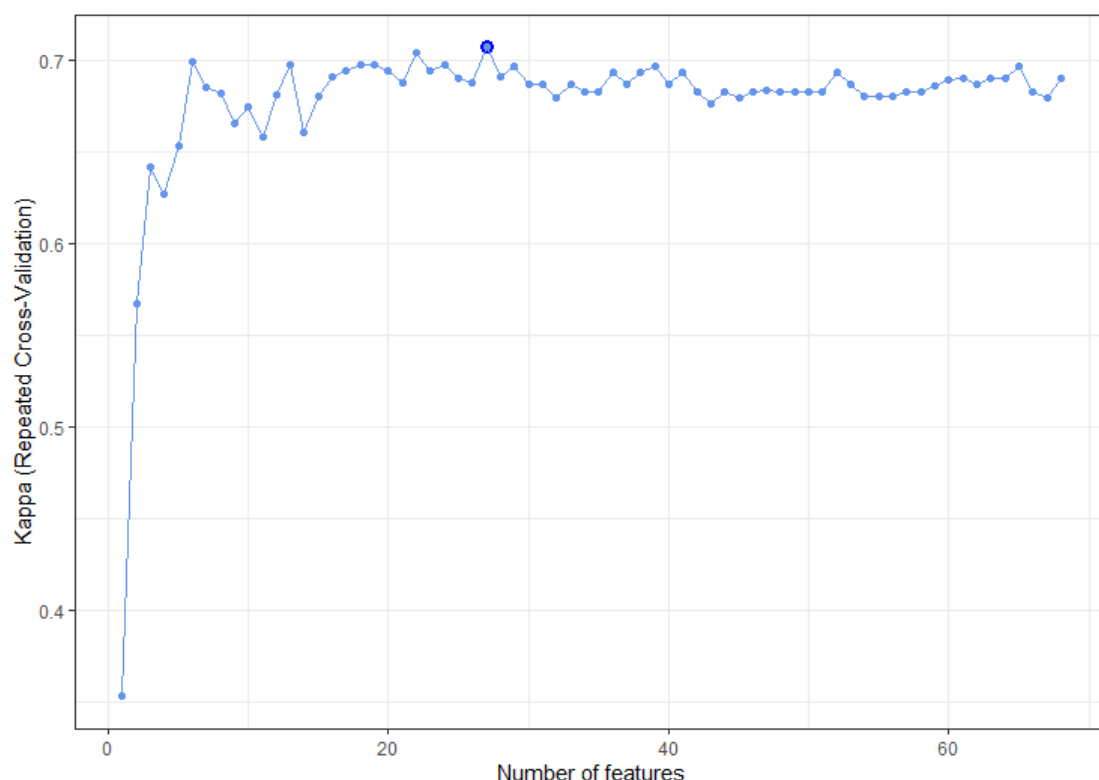
Figure 4.3: Number of features in the Random Forest model retained after recursive feature elimination, with the associated model's performance, measured as the average Cohen's Kappa coefficient, obtained after repeated K-fold cross-validation. The bold dot represents the number of features with the highest performance.

After feature selection, the model was applied to an independent test set for assessment of the classification accuracy. As a result, the confusion matrix in Table 4.5 was obtained as well as the overall accuracy (0.837) with its 95 % confidence interval (CI) (0.752, 0.902), and the Cohen's Kappa coefficient (0.672). Based on the confusion matrix, the producer's and user's accuracy of both classes were calculated (Table 4.5). The CTC class reached a producer's accuracy of 81.0 %, whereas the NTC class reached an accuracy of 87.0 %. This means that NTC class objects of the independent test set were more correctly classified than CTC class objects. The CTC class obtained a higher user's accuracy (0.887) which indicates that 88.7 % of the objects classified as CTC actually represent the CTC class on the ground. This value is lower for the NTC class (0.784), so the classification of CTC objects is more reliable. All things considered, the overall accuracy as well as the Cohen's Kappa coefficient scored relatively high. According to McHugh (2012), the Cohen's Kappa coefficient represents a moderate agreement, and it is likely that 35.0 % to 63.0 % of the data are reliable. As a result, it may be possible to relatively accurately differentiate *Melia volkensii*, included in the CTC mask, from its surroundings.

Table 4.5: Confusion matrix of the CTC/NTC Random Forest classification after feature selection, with calculation of the producer's and user's accuracy.

|                      | Reference data | | |
| --- | --- | --- | --- |
|                      | CTC | NTC | Total |
| **Classification data** | | | |
| CTC                  | 47  | 6   | 53    |
| NTC                  | 11  | 40  | 51    |
| Total                | 58  | 46  |       |
|                      |     |     |       |
| Producer's accuracy  | 0.810 | 0.870 |     |
| User's accuracy      | 0.887 | 0.784 |     |

A few remarks on the CTC mask can be made. Most of the misclassifications were related to trees with smaller crowns. Firstly, some trees were missed during the rule-based classification and therefore withheld from the classification with Random Forest. These errors were not included in the accuracy assessment of the CTC mask, so it is likely that an over-estimation of the correctly classified CTC objects occur. According to Cho et al. (2012), trees with small crowns have a lower ratio "pure" inner crown versus "mixed" outer crown pixels, making their detection more difficult unless spectrally very distinct. Therefore, they have suggested that classification should only consider trees greater than or equal to three times the pixel size of the image to mitigate the effects associated with false pixels in the border of a tree crown and background effects in general (Cho et al., 2012). For this study this would mean that only trees with a crown area greater than 2.25 m$^2$ should be considered. Secondly, some smaller trees, still included in the CTC mask after the rule-based classification, were part of mixed objects with bare soil. When these objects corresponded to ground truth data, biases in the reference data set were induced, often leading to misclassifications. This might also be partly anticipated by removing trees with a crown area smaller than 2.25 m$^2$. Alternatively, the availability of height information is likely to prevent most of these misclassifications by providing thresholds to separate NTC from CTC class objects. It is expected that this may critically increase the accuracy of the CTC mask and thus have a positive influence on the final classification results. The ground truth data was inspected to relate the described errors to the observations. However, it was found that sometimes trees with large crowns were measured while they were not detectable on the satellite imagery, conversely, some larger trees on the satellite imagery were identified as trees with small crown areas during the field inventory. It has to be noted that there was a relatively long period between image acquisition and the start and end of the field inventory, namely approximately four to eight months, so some changes might have happened. The imagery was acquired in the transition period between the long rainy season and the long dry season, while data collection started at the beginning of the short rainy season and the last trees were measured around the end of the second dry season (Section 3.1). During this period, trees may have grown extensively since *Melia volkensii* is a fast growing species. Additionally, some trees may have died or chopped down, followed by replanting or resprouting, hence resulting in smaller tree crowns on the satellite imagery. The differences between the ground truth data and the satellite imagery may also be due to some errors made during the field inventory. This made it almost impossible to determine which variables (Section 3.2.2) were responsible for the detection of individual trees.

Table 4.6 gives an overview of the descriptive statistics of the *Melia volkensii* trees that were not detected by the CTC mask.

Table 4.6: Descriptive statistics of *Melia volkensii* trees that were not included in the candidate tree crown mask. The trees are divided in classes of 5 cm diameter at breast height (DBH) (cm). For each class, the average crown diameter (cm), height (m), age (year), and vitality (1 to 5) are calculated. Additionally, the number of trees (%) in each class is represented.

| DBH (cm) | Crown diameter (cm) | Height (m) | Tree age (year) | Tree vitality | Number of trees (%) |
|---|---|---|---|---|---|
| 0.0 - 5.0 | 151.26 | 3.15 | 2.80 | 1.72 | 19.08 |
| 5.0 - 10.0 | 307.22 | 4.91 | 3.49 | 1.47 | 44.08 |
| 10.0 - 15.0 | 448.02 | 6.60 | 3.62 | 1.31 | 27.63 |
| 15.0 - 20.0 | 571.75 | 9.20 | 4.50 | 1.40 | 6.58 |
| 20.0 - 25.0 | 806.25 | 9.70 | *NA* | 1.75 | 2.63 |
| 25.0 - 30.0 | *NA* | *NA* | *NA* | *NA* | 0.00 |

For determining which features were responsible for the distinction between the CTC and NTC class, the selected features were ranked in descending order based on the Mean Decrease Gini index (Figure 4.4). The top 10 highest-ranked features mostly consisted of spectral features, including mean band values of R, G, and B, brightness, the standard deviation of NIR and R, and band ratios such as the enhanced vegetation index (EVI), the ratio vegetation index (RVI), and the NDVI. Additionally, one object shape feature, namely length, was included in the top 10. The other 17 features included mostly textural features (11), five additional spectral features, and one shape feature. For visualisation of the separability between CTC and NTC class objects, density functions of the classified objects were plotted for the 20 highest-ranked features (Figure 4.5 and Figure 4.6). Lower-ranked features showed an increasingly overlap, and thus similarity between both classes. These features may even not be able to separate CTC class objects from NTC class objects. Despite these findings, these features were still selected by the recursive feature elimination procedure. When uninformative features correlate with the highest-ranked features, they will also be retained by the selection procedure. This indicates the sensitivity of recursive feature elimination in finding important features and eliminating less relevant features. A preliminary elimination of features with low correlations to the desired classes, might increase classification accuracies (Bulut, 2021).
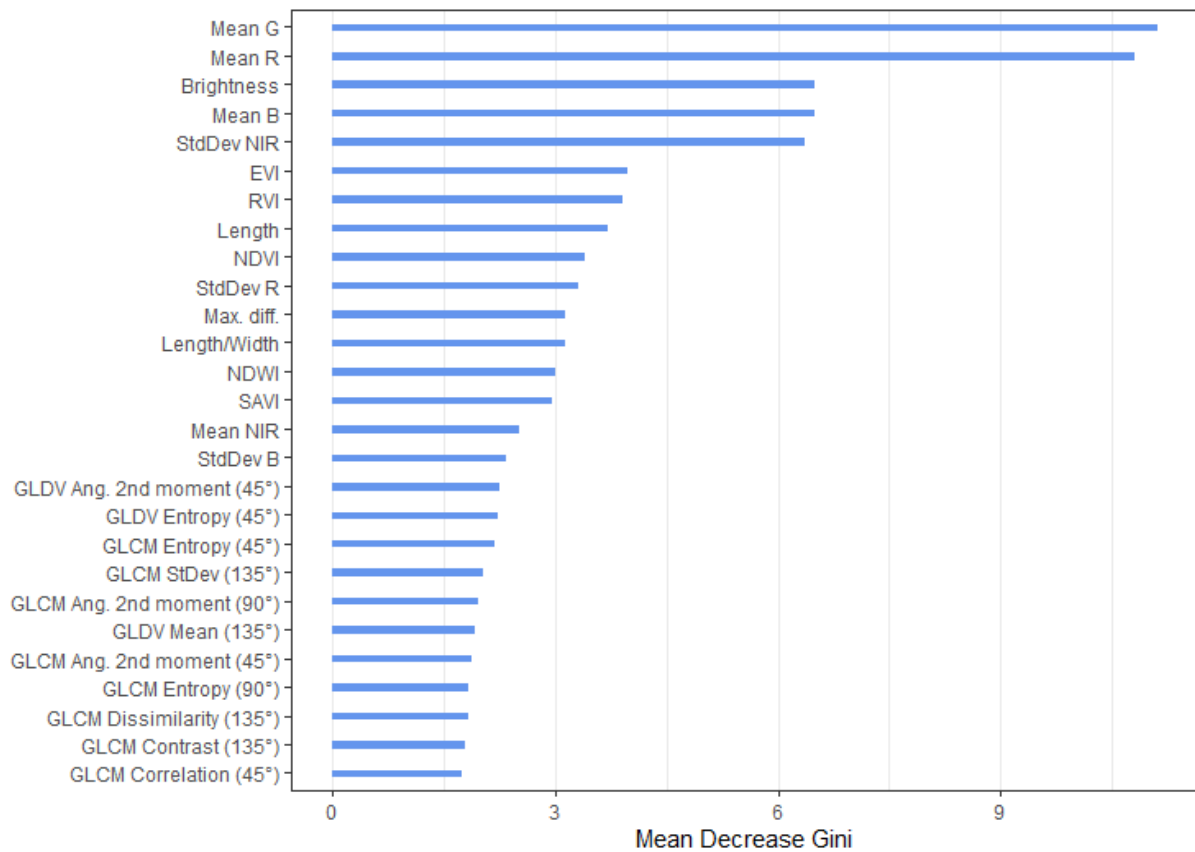
Figure 4.4: Ranking of the selected features in descending order based on the Mean Decrease Gini index after running the CTC/NTC Random Forest classification. Abbreviations of the features are explained in Section 3.3.2.
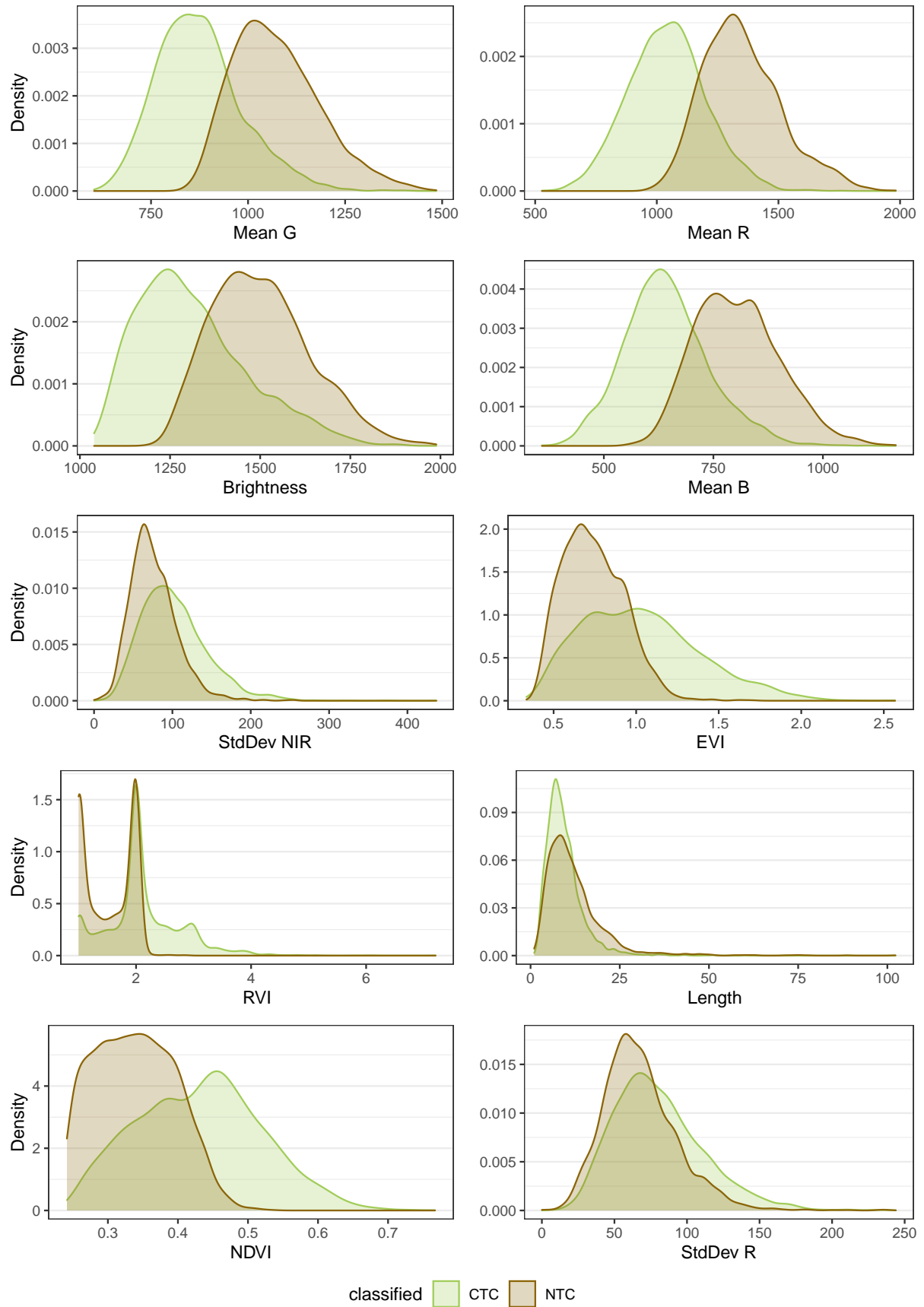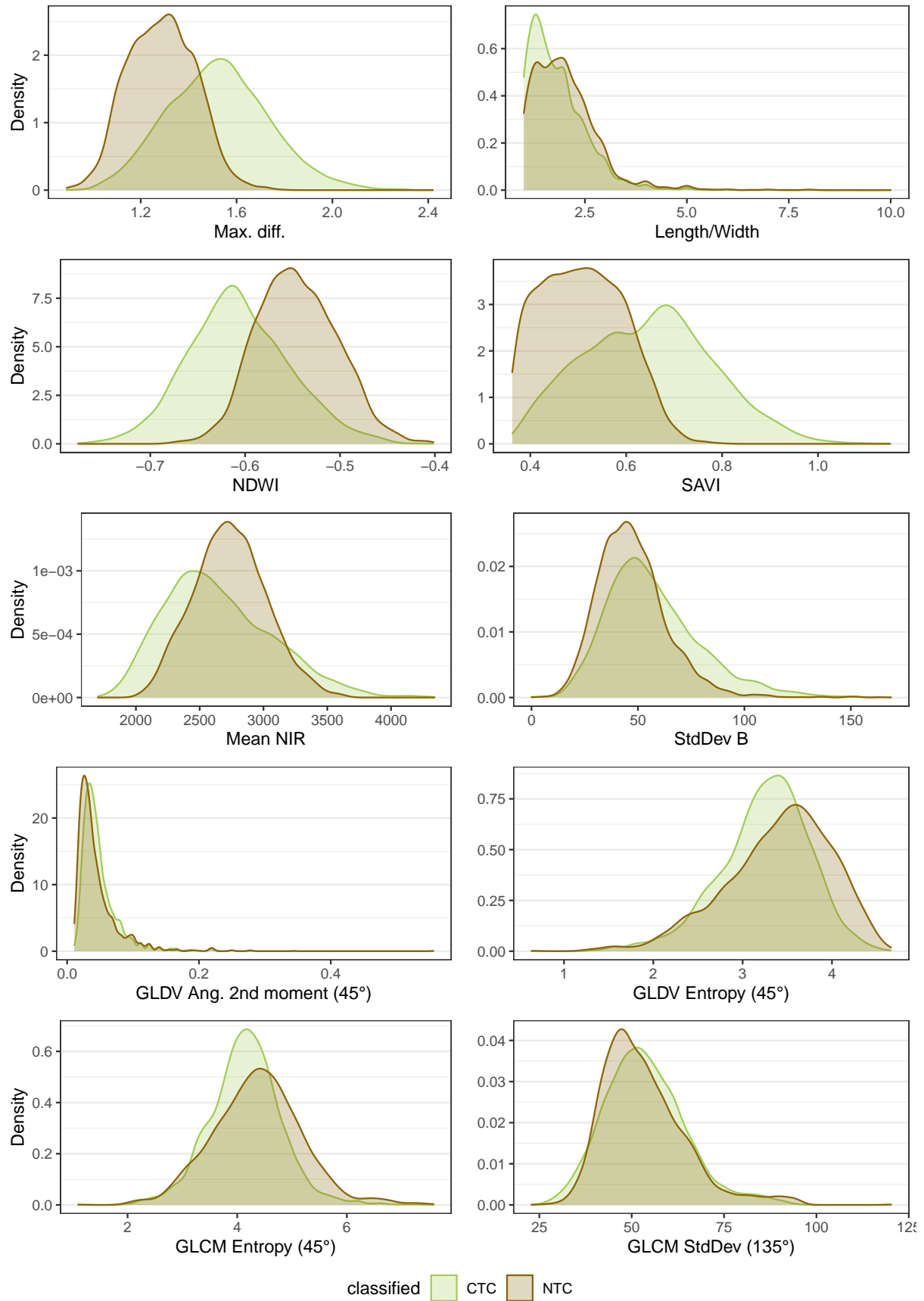
Figure 4.5: Density functions of the classified CTC as well as NTC objects of the 10 highest-ranked features. Abbreviations of the features are explained in Section 3.3.2.

Figure 4.6: Density functions of the classified CTC as well as NTC objects of the 11 to 20 highest-ranked features. Abbreviations of the features are explained in Section 3.3.2.

## 4.3  Differentiation of *Melia volkensii*

The CTC mask, obtained after the CTC/NTC Random Forest classification, was further classified into *Melia volkensii* and other tree species. The optimal hyperparameter set was also determined with a grid search for all three sets of features. The optimal combinations were selected and stored in Table 4.7 with the associated model's performance.  The feature set with only spectral features reached the highest performance with an average Cohen's Kappa coefficient of 0.392, obtained after repeated K-fold cross-validation, but the OOB error (23.90 %) of this feature set did not reach the lowest value.  Nonetheless, the spectral feature set was selected for further analysis as the Cohen's Kappa coefficient is a more qualitative measure of the model's performance. All combinations of hyperparameters and their associated average Cohen's Kappa coefficient were stored in Figure 4.7.

Table 4.7: Optimal combination of hyperparameters (i.e.  the number of input features at each node (mtry) and the number of decision trees (ntree)) and the model's performance, measured as the average Cohen's Kappa coefficient as well as the out-of-bag (OOB) error (%), obtained after repeated K-fold cross-validation.  These values resulted from a grid search of the differentiation of *Melia volkensii* with Random Forest for all three feature sets (i.e. (1) spectral features only, (2) a combination of spectral and textural features, and (3) a combination of spectral, textural, and shape features)

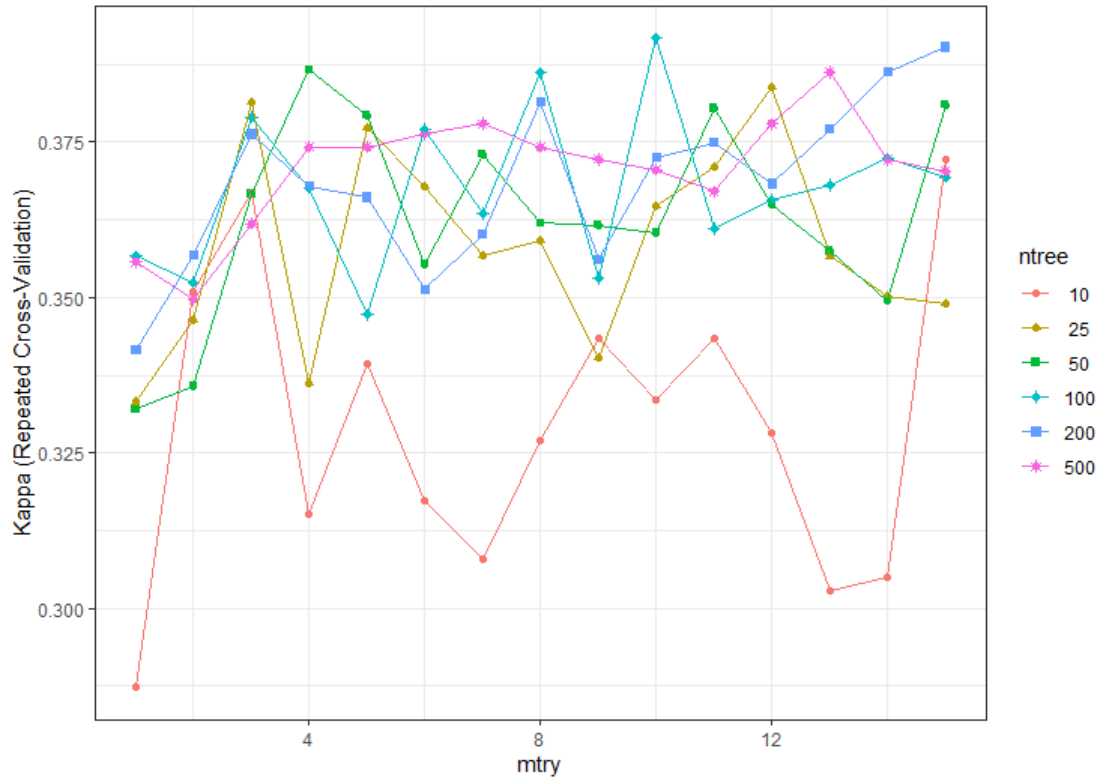| Feature set | ntree | mtry | Cohen's Kappa | OOB error (%) |
|---|---|---|---|---|
| Spectral | 100 | 10 | 0.392 | 23.90 |
| Spectral + texture | 500 | 15 | 0.378 | 22.93 |
| Spectral + texture + shape | 25 | 12 | 0.369 | 23.17 |

Figure 4.7: All combinations of hyperparameters (i.e. i.e. the number of input features at each node (mtry) (dots) and the number of decision trees (ntree) (lines)) with their associated model's performance, measured as the average Cohen's Kappa coefficient, obtained after repeated K-fold cross-validation. These values resulted from a grid search of the differentiation of *Melia volkensii* with Random Forest performed with the spectral feature set.

Recursive feature elimination was performed on the spectral feature set. For each elimination of a feature, the model's performance was stored in Figure 4.8 as the average Cohen's Kappa coefficient, obtained after repeated K-fold cross-validation. The highest Cohen's Kappa coefficient (0.351) was reached when nine features were included in the model. The Random Forest was built with the default parameters of three input features at each node (mtry) and 500 decision trees (ntree). Based on Figure 4.7, there is little difference in the model's performance between this hyperparameter set and the optimal set, so it is expected that this would not have a major influence on the final classification results.
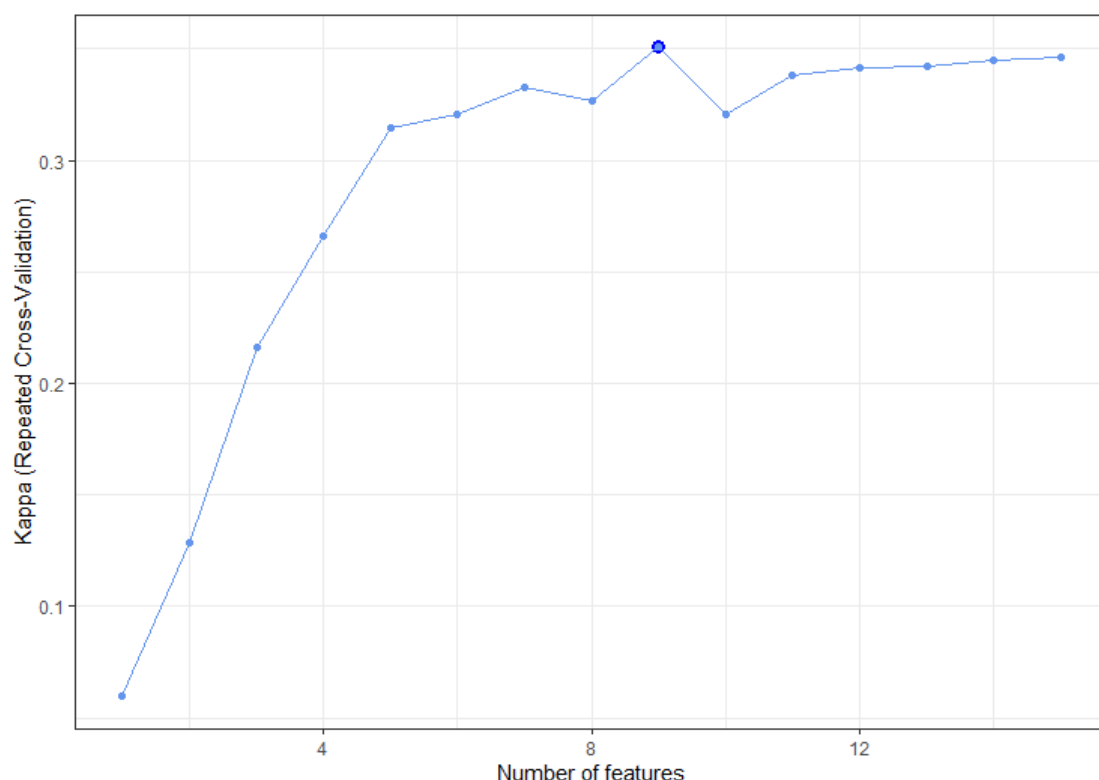
Figure 4.8: Number of features included in the Random Forest model retained after recursive feature elimination, with the associated model's performance, measured as the average Cohen's Kappa coefficient, obtained after repeated K-fold cross-validation. The bold dot represents the number of features with the highest performance.

The model with the selected features was applied to an independent test set for assessment of the classification accuracy. As a result, the confusion matrix in Table 4.8 was obtained, as well as the overall accuracy (0.665) with its 95 % CI (0.589, 0.735) and the Cohen's Kappa coefficient (0.118). Based on the confusion matrix, the producer's and user's accuracy were calculated and stored in Table 4.8. Most of the other tree species (i.e. other) test set objects were assigned to the *Melia volkensii* (i.e. mukau) class, namely 32 out of 47. This resulted in both a low producer's (0.319) and user's accuracy (0.366) for the other tree species class, meaning that the classification cannot be considered as reliable. *Melia volkensii* obtained a higher producer's (0.794) and user's accuracy (0.758). These differences in accuracy might be due to among other things, an imbalance in the number of test objects between both classes, 126 for *Melia volkensii* compared to 47 for the other tree species class. Overall, the Cohen's Kappa coefficient scored very low and corresponded to no agreement according to McHugh (2012). Data collected under conditions of such disagreement are not meaningful, and are more like random data (McHugh, 2012). Consequently, the obtained model has a very low predictive power and is not suitable for application without some major adjustments (Chapter 5). Maps resulting from the classification are included in Appendix C.

Table 4.8: Confusion matrix of the Random Forest classification for differentiating *Melia volkensii* (mukau) from other tree species (other) after feature selection, with calculation of the producer's and user's accuracy.

| | Reference data | | |
|---|---|---|---|
| | **Mukau** | **Other** | **Total** |
| **Classification data** | | | |
| Mukau | 100 | 32 | 132 |
| Other | 26 | 15 | 41 |
| Total | 126 | 47 | |
| | | | |
| Producer's accuracy | 0.794 | 0.319 | |
| User's accuracy | 0.758 | 0.366 | |

When comparing this study with literature, rather low classification results were achieved. For example, Pu et al. (2018) has performed a similar study for classifying urban tree species in the City of Tampa, Florida. They have reached Cohen's Kappa coefficients ranging from 0.334 to 0.514, corresponding to a minimal to weak agreement (McHugh, 2012). Their findings have reached a higher classification accuracy, but are not satisfactory either, so little confidence in the study results should be placed (McHugh, 2012). The main difference with this study is that they have used light normalisation to mitigate the disadvantages of shade for classification, thereby preserving the shape of the crown. Wang et al. (2018) have tried to classify artificial mangrove tree species, and reached Cohen's Kappa values up to 0.747, corresponding to a moderate agreement (McHugh, 2012). In contrast to this study, they have not performed the classification on the individual tree crown level, but they have rather conducted a land cover classification. The presence of monocultures in their study area is also beneficial for the delineation of the regions to be classified. Nonetheless, comparison between studies is not straightforward and several factors need to be considered, including the total number of classified tree species, classification algorithm, reference data characteristics and the method used for accuracy assessment (Karlson et al., 2016). In general, the highest accuracies have been achieved when only a small number of tree species have been analysed and/or when additional input data (e.g. height information) have been used (Immitzer et al., 2012).

The selected features were ranked in descending order based on the Mean Decrease Gini index (Figure 4.9). Furthermore, density functions of the training data for the selected features were plotted to visualise why *Melia volkensii* is not distinguishable from other tree species using these features (Figure 4.10). Based on the graphs, it can be visually determined that both classes show significant overlap for all features, therefore they were difficult to separate. The limited reference data set may be one of the possible causes of the poor classification results. The data was split into training and test data, as a result of which some information comprised within the data may be missed during the training stage, leading to the low predictive power of the model. Furthermore, there was no ground truth data available regarding the class representing other tree species. By manually selecting the reference data, there existed a significant possibility that some *Melia volkensii* trees were included in the other tree species class leading to erroneous reference samples. According to Immitzer et al. (2012), a large sample size per class may have a positive influence on the classification results.
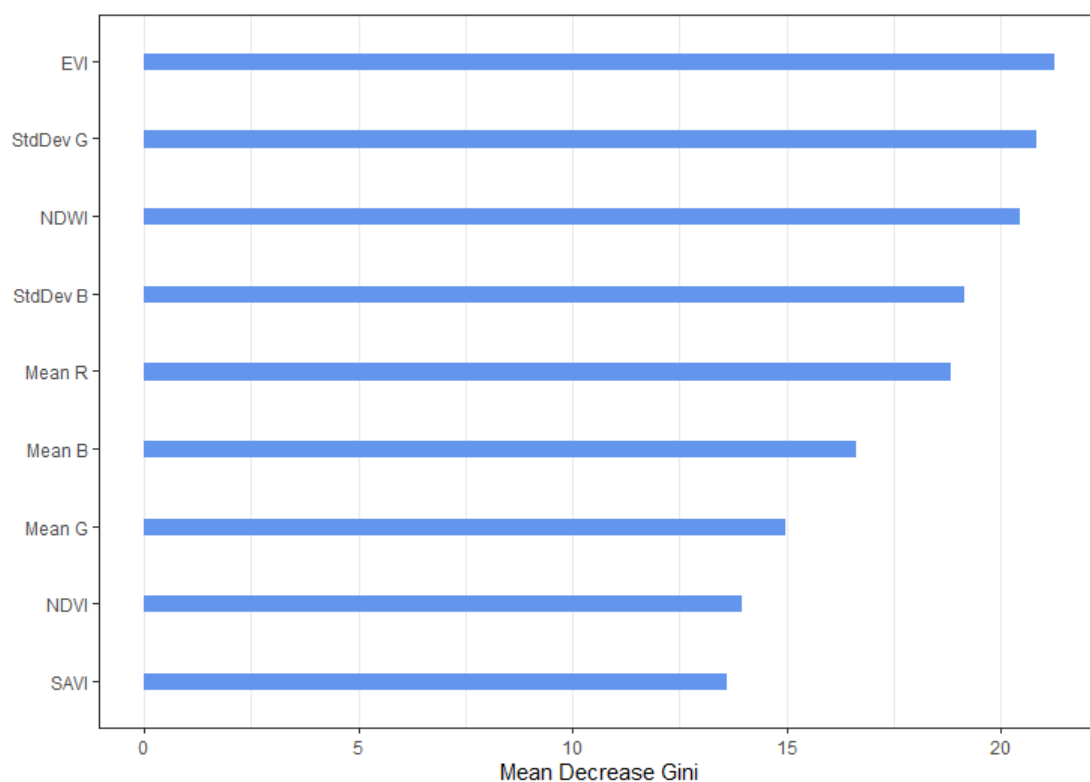
Figure 4.9: Ranking of the selected features in descending order based on the Mean Decrease Gini index after running the Random Forest classification for the differentiation of *Melia volkensii*. Abbreviations of the features are explained in Section 3.3.2.
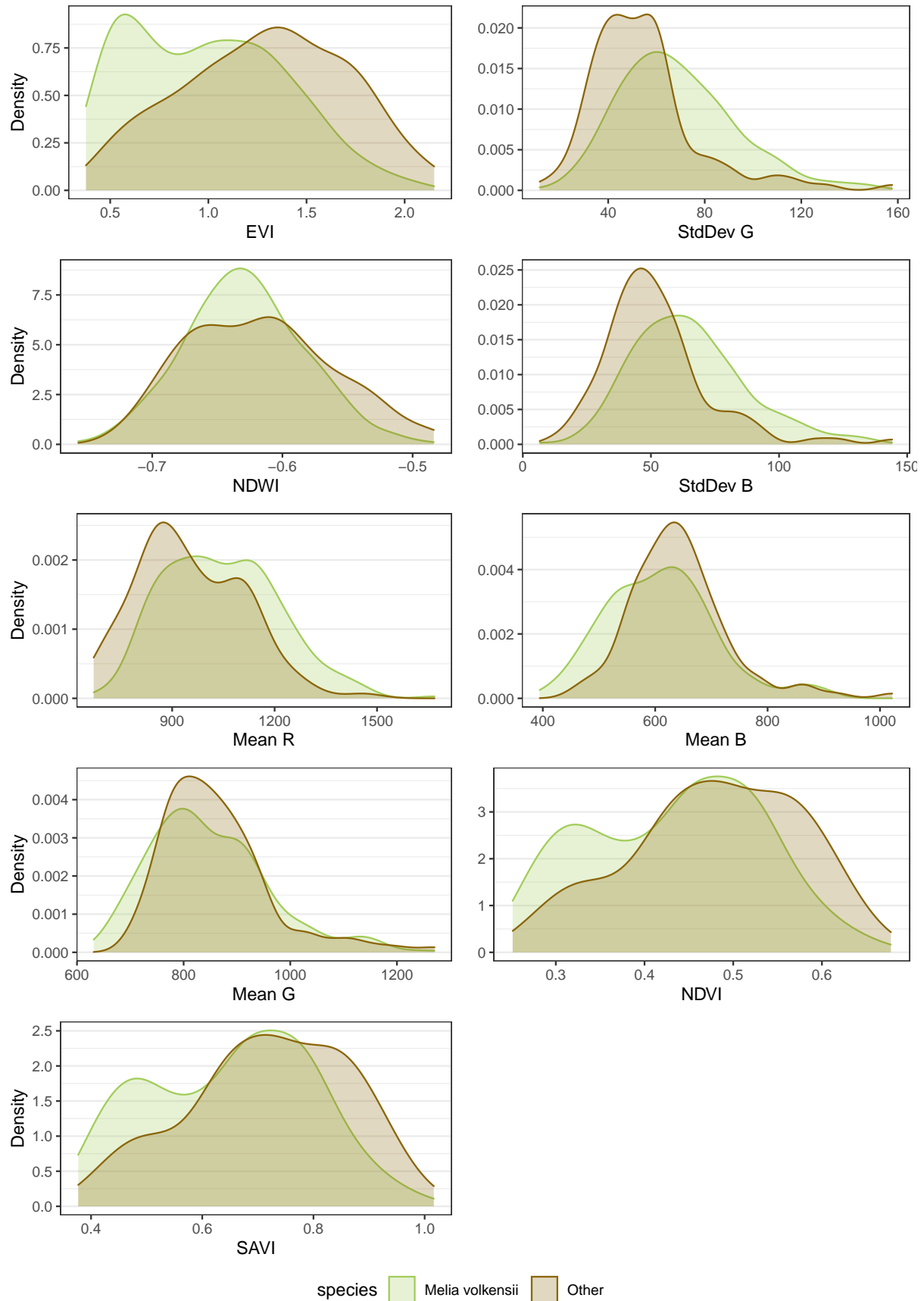
Figure 4.10: Density functions of *Melia volkensii* as well as the other tree species class of the selected features. Abbreviations of the features are explained in Section 3.3.2.

# 5. RECOMMENDATIONS

The classification of the candidate tree crown (CTC) mask in *Melia volkensii* and other tree species reached a very low accuracy (Cohen's Kappa coefficient of 0.118), wherefore little confidence in the results should be placed (McHugh, 2012). In this chapter, some recommendations based on literature will be made, that might enhance the accuracy of the tree species classification. Most of these recommendations are related to data collection, especially image acquisition.

## 5.1 Sensor

According to Cho et al. (2012), there is strong evidence that narrow wavelength bands are necessary to resolve the subtle spectral differences between canopies and crowns of different tree species in a landscape. These spectral signatures are linked to leaf and canopy biochemical and biophysical properties, including photosynthetic pigments, foliar nutrients, leaf morphology, canopy structure, and tree size compared to neighbouring trees (Cho et al., 2012; Fassnacht et al., 2016; Madonsela et al., 2017). However, these properties vary not only with species, but also with vertical leaf area density, leaf age, and health status. Furthermore, additional influences on the amount of reflected radiation are caused by background signals related to, for instance, bare soil or neighbouring trees (Fassnacht et al., 2016). In other words, the spectral reflectance of one species compared with another will never be identical, nor will the spectral reflectance of trees of the same species be exactly equal (Lillesand et al., 2015). The main drivers of discriminating tree species in the visible (VIS) - shortwave infrared (SWIR) region are related to differences in plant chemical and physical properties between species, causing different levels of reflectance (amplitude) (Fassnacht et al., 2016). Especially the yellow (Y) and/or red edge (RE) band are sensitive to subtle differences in carotenoid and chlorophyll pigments amongst species, and therefore are useful for enhancing tree species classification (Madonsela et al., 2017).

The latest generation of very high resolution (VHR) satellite systems designed with strategically located and relatively narrow wavelength bands in the absorption features of plant biochemicals and biophysicals offer promising solutions for widespread remote sensing of plant species (Cho et al., 2012; Karlson et al., 2016). In 2009, the VHR satellite WorldView-2 was launched by DigitalGlobe with four additional spectral bands (i.e. coastal blue (CB), Y, RE, and near infrared 2 (NIR2)) on top of the four standard bands (i.e. blue (B), green (G), red (R), and near infrared 1 (NIR1)). According to the data provider, the four new bands are all related to vegetation properties which enables successful tree species classification in various ecosystems, particularly when tree species show substantial spectral overlap, or when a larger number of species has to be separated. However, the added value of these bands seems to be species-dependent and play a minor role when a small number

of species has to be classified (Immitzer et al., 2012; Karlson et al., 2016). Considering the high costs related to the four additional bands, the user needs to evaluate if the new bands are beneficial in relation to these extra costs (Immitzer et al., 2012). Compared to WorldView-2, WorldView-3 provides both a higher spatial as well as spectral resolution. WorldView-3 includes eight additional SWIR bands, which makes it the only VHR satellite with SWIR sensing capabilities. These SWIR wavelength bands have previously shown to be useful for tree species classification. Therefore, WorldView-3 may thus enable even higher classification accuracies (Karlson et al., 2016; Ferreira et al., 2019). Because of the low classification accuracy in this study, it might be worth trying to differentiate *Melia volkensii* based on WorldView-2 or -3 satellite imagery instead of Pléiades imagery, despite the higher costs. It is expected that the additional spectral bands will critically improve the classification accuracy of the last step implemented in the methodology used in this study (Section 4.3). Alternatively, if there are budget constraints, using the free-of-charge Sentinel-2 data might be a viable solution as the increased spectral resolution (i.e. 13 spectral bands) compensates for the lack of spatial resolution (10, 20 or 60 m depending on the wavelength band) (Ng et al., 2017), however, differentiation of *Melia volkensii* at the individual tree crown level will no longer be possible. Ng et al. (2017) has established that the higher spectral resolution has a positive influence on classification accuracies, partly compensating the pixel size effects, and so contributes to improved detection and differentiation of vegetation. These accuracies can even be further improved by using temporal information available at five-day intervals (Ng et al., 2017).

## 5.2 Acquisition date

As discussed in Section 2.2.1, the time of image acquisition should be aligned with the phenological cycle of the species to be classified, in particular with periods in which target species show peculiar characteristics (e.g. flowers or senescent leaves). This may potentially favour the differentiation between the species of interest as changes in phenology occur at different rates amongst species (Madonsela et al., 2017; Ferreira et al., 2019). *Melia volkensii* is a fast-growing, deciduous tree species native in Kenya. It sheds its leaves twice a year during the dry seasons. Almost all leaves are shed during the long dry season. As a result, this period is characterised by high leaf fall and low leaf cover. Towards the end of the dry season, trees start flushing new leaves before the onset of the ensuing short rainy season. After flushing, leaf fall ceases for short periods, but thereafter it increases sharply during the later stages of the rainy season (Broadhead et al., 2003). Flowers and fruits are also produced twice a year, with fruits becoming ripe at the end of the dry season as the new leaves emerge. Stages of fruit development lack a seasonal pattern as a result of which it is not uncommon that trees at the same site flower or fruit at different times of the year (Orwa et al., 2009).

The optimal season to acquire imagery with the purpose of differentiating *Melia volkensii* is likely to occur in the transition period between dry and wet seasons, including May, October, December, and March (Section 3.1), because trees start to show specific characteristics during these periods by shedding or flushing their leaves. The early dry seasons (May-June and December-January) might be the best choice as spectral differences between tree species are likely to be higher than in the wet season. Besides, frequent cloud cover might

hinder image acquisition during the wet season (Karlson et al., 2016). For determining the optimal acquisition date, the phenological cycle of other tree species occurring in the area of interest should be examined, so an ideal period can be selected when the differences between the species are maximised. Furthermore, the possibility of background effects should be considered when selecting the optimal time of image acquisition. Background reflectance, originating from crops, or the underlying grass layer, might induce spectral confusion between trees, especially when trees are undergoing senescence accompanied by leaf shedding (Cho et al., 2012). Therefore, higher classification accuracies are expected when the image is acquired after crop harvesting or grass mowing. Nevertheless, there would always be some disadvantages and advantages when using data acquired during any particular season, for example, there could be less within-species variability in the spectral data at peak productivity, but the variability between species may also be limited during this period (Cho et al., 2012). Multi-seasonal imagery, acquired in two distinctive periods, namely one during green-up and one during senescence (Fassnacht et al., 2016), may also be beneficial for tree species classification. Considering the higher costs, it is important to evaluate if the use of multiple acquisitions would contribute to a significant increase in classification accuracy.

## 5.3 Field inventory

The location of individual trees was the most important variable used in this study (Section 3.2.2), hence, it greatly influenced the classification results. Unfortunately, the handheld global positioning system (GPS) used (i.e. Garmin 60 scx) reached an accuracy of only 3 m, consequently, a relatively high deviation compared to the true location of these trees was induced. As a result, questionable reference objects were obtained in some cases, for example, bare soil objects or other tree species were assigned to the *Melia volkensii* class, and vice versa, resulting in erroneous reference samples. Therefore, every reference sample in the farms was manually checked, and relocated if necessary to reduce these errors as much as possible, however, this was a very time-consuming process and may not have mitigated all errors or even induced new ones. The use of GPS devices with higher accuracy is probably the most important prerequisite to obtain reliable classification results, even if the classification is based on the best performing sensor, or the image is acquired in the most optimal season. It is expected that the African satellite augmentation system (ASAS) will be operational in Kenya between 2024 and 2030 (Durban University of Technology, 2015). This network is developed to augment the GPS, with the goal of improving accuracy, integrity, and availability (National Coordination Office for Space-Based Positioning, Navigation, and Timing, 2022). In the future, using GPS devices compatible with this system is expected to critically improve the positioning of individual trees during the field inventory. Alternatively, light detection and ranging (LiDAR) data can be employed to measure the location of individual trees with sub-metre accuracy (Holmgren et al., 2008).

Next to optimising the positioning accuracy of individual trees, some smaller actions can be taken to improve the ground truth data set. Collecting data of more trees to obtain a larger sample size per class is expected to positively affect the final classification (Immitzer et al., 2012), especially data concerning the other tree species class will have a great influence. Furthermore, aligning the field inventory with the time of image acquisition would

be beneficial because changes might happen when there is an extended period between both data collections. When both data sets are collected in approximately the same period, trees are more likely to appear in similar conditions in both data sets and it is less likely that either of them would miss trees that were, for instance, chopped down or replanted.

## 5.4   Height information

The availability of structural information, especially tree height, may also improve the classification results. In this study, trees were separated from understory vegetation by performing edge extraction Lee Sigma filtering, which resulted in the Roughness band (Section 3.3.1). This was a viable solution in absence of height information but there were still non-tree crown (NTC) objects included in the CTC mask after the rule-based classification (Section 4.2.1), so the CTC class was further separated into CTC and NTC class objects with Random Forest (Section 4.2.2) before differentiating *Melia volkensii*. The Random Forest classification reached a moderate agreement according to McHugh (2012). This indicates that the tree species of interest may partly possess similar properties to the underlying grass layer (Kato et al., 2009). According to Kato et al. (2009), structural variables may help to reduce this spectral confusion. The availability of a tree height mask would facilitate the extraction of spectral information regarding the CTC class (Cho et al., 2012), which proved to be challenging task with several sub-steps in this study (Section 4.2). By providing height information, most of these sub-steps could be omitted, thereby reducing a large amount of expert knowledge used during these steps.

Tree height is generally obtained from auxiliary three-dimensional (3D) information, for example, from canopy height models (CHM) (Kato et al., 2009; Immitzer et al., 2012). Such data sets can be derived from LiDAR data but these are commonly expensive to acquire, especially in Africa. A potential alternative is 3D data derived from the VHR satellite itself, providing that the area of interest is covered spectroscopically (Immitzer et al., 2012; Karlson et al., 2016; Piermattei et al., 2019). The long acquisition times of airborne remote sensing present a major challenge to its use in monitoring afforestation at the landscape scale. Full coverage of extended areas cannot be achieved in a single airborne inventory, and logistic challenges are associated with multi-phase inventories. In comparison, VHR stereo satellite imagery, which can now acquire images with sub-metre spatial resolution, has the benefit of wider area coverage and spatially/temporally more homogeneous image content with short repetition intervals. This makes VHR stereo satellite imagery a valuable solution for monitoring afforestation at the national or sub-national level (Piermattei et al., 2019). By combining the resulting digital surface model (DSM) with a pre-existing digital terrain model (DTM), CHMs can be obtained. For example, Piermattei et al. (2019) has used Pléiades satellite imagery, acquired in the stereo and tri-stereo mode, in combination with two DTMs that were generated earlier. They have established that the accuracy obtained with Pléiades stereo images are comparable to aerial image matching when the forest composition guarantees a certain amount of foliage, and shadows cast by trees or the terrain are minimal. Therefore, it may be worth investing in a DTM of the areas of interest so the afforestation can be monitored on a more frequent basis by only investing in VHR stereo satellite imagery. Three-dimensional information can also be used for accurate delineation of individual tree crowns (Karlson et al., 2016).

# 6. __CONCLUSION__

In this study, a methodology was proposed to differentiate *Melia volkensii* from its surrounding and other tree species at the individual tree crown level in an agroforestry layout. The methodology consisted of several sub-steps, with two major parts. The first part focused on separating the candidate tree crown (CTC) class, which included *Melia volkensii*, from the non-tree crown (NTC) class (i.e. the surroundings). The delineation of individual tree crowns played a major role in the possibility to differentiate *Melia volkensii* at the tree crown level. A more automated procedure for tree crown delineation will be essential for large-scale application of the proposed methodology. Due to background effects of crops, grasses, and bare soil, or the similarity between overlapping crowns, not every CTC object was delineated at the individual tree crown level. Therefore, crown clusters were also included in the procedure. The Random Forest classification reached an accuracy with moderate agreement, so the differentiation of *Melia volkensii* from its surroundings was relatively reliable, however, an overestimation of the accuracy was probably made due to the omission of trees during the rule-based classification. The availability of height information, LiDAR-derived or stereo-image based, can improve the classification results by enabling more accurate tree crown delineation and providing height thresholds to accurately separate CTC class objects from NTC class objects. During the second step, the obtained CTC mask was further differentiated into *Melia volkensii* and other tree species. The classification reached low accuracies as a result of which the model is not suitable for application without major adjustments. Due to the unsatisfactory classification results, no statements can be made about which combination of features enabled the differentiation of *Melia volkensii*. The use of very high resolution (VHR) satellites with narrow wavelength bands located in the absorption spectrum of plant biochemicals and biophysicals, and the availability of ground truth data regarding the other tree species class are likely to increase the accuracy of the classification. Additionally, some smaller adjustments such as acquiring the image in the most optimal season, or the use of multi-seasonal imagery may also positively influence the classification accuracy.

The proposed methodology was automated as much as possible to make it applicable to other areas where BGF might be interested in. However, some sub-steps (e.g. extracting ridges) of the first part will be hard to reproduce, mostly because they involve a certain amount of expert knowledge. These shortcomings are likely to be solved when there is height information available for the delineation and extraction of candidate tree crowns. The most important adjustment for both improving the classification accuracy as well as for automation of the methodology, is probably increasing the positioning accuracy of GPS devices used to locate the individual trees. In future research, extra attention should be paid to this during the field inventory.

# **BIBLIOGRAPHY**

Adam, H., Csaplovics, E., and Elhaja, M. (2016). A comparison of pixel-based and object-based approaches for land use land cover classification in semi-arid areas, sudan. In *IOP Conference Series: Earth and Environmental Science*, volume 37, page 012061. IOP Publishing.

Adams, R. and Bischof, L. (1994). Seeded region growing. *IEEE Transactions on pattern analysis and machine intelligence*, 16(6):641–647.

Airbus (2021).

Arvor, D., Durieux, L., Andrés, S., and Laporte, M.-A. (2013). Advances in geographic object-based image analysis with ontologies: A review of main contributions and limitations from a remote sensing perspective. *ISPRS Journal of Photogrammetry and Remote Sensing*, 82:125–137.

Baatz, M. and Schäpe, A. (2000). Multi resolution segmentation: an optimum approach for high quality multi scale image segmentation. *Beutrage zum AGIT-Symposium. Salzburg, Heidelberg, 2000*, pages 12–23.

Bannari, A., Morin, D., Bonn, F., and Huete, A. (1995). A review of vegetation indices. *Remote sensing reviews*, 13(1-2):95–120.

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., and Wood, E. F. (2018). Present and future köppen-geiger climate classification maps at 1-km resolution. *Scientific data*, 5(1):1–12.

Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I., and Heynen, M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for gis-ready information. *ISPRS Journal of photogrammetry and remote sensing*, 58(3-4):239–258.

Better Globe Forestry (n.d.). `https://www.betterglobeforestry.com/index.html`. Consulted on 2 January 2022.

Blaschke, T. (2003). Object-based contextual image classification built on image segmentation. In *IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data, 2003*, pages 113–119. IEEE.

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS journal of photogrammetry and remote sensing*, 65(1):2–16.

Blaschke, T., Burnett, C., and Pekkarinen, A. (2004). Image segmentation methods for object-based analysis and classification. In *Remote sensing image analysis: Including the spatial domain*, pages 211–236. Springer.

Blaschke, T., Hay, G. J., Kelly, M., Lang, S., Hofmann, P., Addink, E., Feitosa, R. Q., Van der Meer, F., Van der Werff, H., Van Coillie, F., et al. (2014). Geographic object-based image analysis–towards a new paradigm. *ISPRS journal of photogrammetry and remote sensing*, 87:180–191.

Broadhead, J., Ong, C., and Black, C. (2003). Tree phenology and water availability in semi-arid agroforestry systems. *Forest Ecology and Management*, 180(1-3):61–73.

Brownlee, J. (2020). `https://machinelearningmastery.com/repeated-k-fold-cross-validation-with-python/`. Consulted on 11 May 2022.

Bulut, O. (2021). `https://towardsdatascience.com/effective-feature-selection-recursive-feature-elimination-using-r-148ff998e4f7`. Consulted on 12 May 2022.

Campbell, J. B. and Wynne, R. H. (2011). *Introduction to remote sensing*. Guilford Press.

Cánovas-García, F. and Alonso-Sarría, F. (2015). A local approach to optimize the scale parameter in multiresolution segmentation for multispectral imagery. *Geocarto International*, 30(8):937–961.

Cho, M. A., Mathieu, R., Asner, G. P., Naidoo, L., Van Aardt, J., Ramoelo, A., Debba, P., Wessels, K., Main, R., Smit, I. P., et al. (2012). Mapping tree species composition in south african savannas using an integrated airborne spectral and lidar system. *Remote Sensing of Environment*, 125:214–226.

Coeurdevey, L. and Fernandez, K. (2012). Pléiades imagery - user guide.

Dalponte, M., Frizzera, L., and Gianelle, D. (2019). Individual tree crown delineation and tree species classification with hyperspectral and lidar data. *PeerJ*, 6:e6227.

Deur, M., Gašparović, M., and Balenović, I. (2021). An evaluation of pixel-and object-based tree species classification in mixed deciduous forests using pansharpened very high spatial resolution satellite imagery. *Remote Sensing*, 13(10):1868.

Drăguţ, L., Csillik, O., Eisank, C., and Tiede, D. (2014). Automated parameterisation for multi-scale image segmentation on multiple layers. *ISPRS Journal of photogrammetry and Remote Sensing*, 88:119–127.

Drăgut, L., Tiede, D., and Levick, S. R. (2010). Esp: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science*, 24(6):859–871.

Duan, G., Gong, H., Zhao, W., Tang, X., and Chen, B. (2013). An index-based shadow extraction approach on high-resolution images. In *Proceedings of the International Symposium on Satellite Mapping Technology and Application, Nanjing, China*, pages 6–8. Citeseer.

Durban University of Technology (2015). `https://www.dut.ac.za/wp-content/uploads/2015/02/African-Satellite-Augmentation-System-ASAS-African-SBAS-2.pdf`. Consulted on 28 May 2022.

Espindola, G., Câmara, G., Reis, I., Bins, L., and Monteiro, A. (2006). Parameter selection for region-growing image segmentation algorithms using spatial autocorrelation. *International Journal of Remote Sensing*, 27(14):3035–3040.

European Soil Data Centre (ESDC) (n.d.). `https://esdac.jrc.ec.europa.eu/content/machakos-kitui-embu-soil-survey-area`. Consulted on 4 April 2022.

European Space Agency (n.d.). `https://earth.esa.int/eogateway/missions`. Consulted on 20 March 2022.

Fan, M. and Lee, T. C. (2015). Variants of seeded region growing. *IET image processing*, 9(6):478–485.

Fassnacht, F. E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L. T., Straub, C., and Ghosh, A. (2016). Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment*, 186:64–87.

Ferreira, M. P., Wagner, F. H., Aragão, L. E., Shimabukuro, Y. E., and de Souza Filho, C. R. (2019). Tree species classification in tropical forests using visible to shortwave infrared worldview-3 images and texture analysis. *ISPRS journal of photogrammetry and remote sensing*, 149:119–131.

Fotheringham, A. S., Brunsdon, C., and Charlton, M. (2000). *Quantitative geography: perspectives on spatial data analysis*. Sage.

Frauman, E. and Wolff, E. (2005). Segmentation of very high spatial resolution satellite images in urban areas for segments-based classification. In *Proceedings for 3rd International Symposium Remote Sensing and Data Fusion Over Urban Areas. Tempe, Arizona*.

Ghosh, A. and Joshi, P. (2013). Assessment of pan-sharpened very high-resolution worldview-2 images. *International journal of remote sensing*, 34(23):8336–8359.

Ghyselbrecht, E. (2022). Monitoring afforestation with very high resolution satellite imagery in kenya. methods for individual tree counting.

Gislason, P. O., Benediktsson, J. A., and Sveinsson, J. R. (2006). Random forests for land cover classification. *Pattern recognition letters*, 27(4):294–300.

Haralick, R. M., Shanmugam, K., and Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6):610–621.

Haralick, R. M. and Shapiro, L. G. (1985). Image segmentation techniques. *Computer vision, graphics, and image processing*, 29(1):100–132.

Hay, G. J., Castilla, G., Wulder, M. A., and Ruiz, J. R. (2005). An automated object-based approach for the multiscale image segmentation of forest scenes. *International Journal of Applied Earth Observation and Geoinformation*, 7(4):339–359.

Holmgren, J., Persson, Å., and Söderman, U. (2008). Species identification of individual trees by combining high resolution lidar data with multi-spectral images. *International Journal of Remote Sensing*, 29(5):1537–1552.

Hossain, M. D. and Chen, D. (2019). Segmentation for object-based image analysis (obia): A review of algorithms and challenges from remote sensing perspective. *ISPRS Journal of Photogrammetry and Remote Sensing*, 150:115–134.

Hsieh, P.-F., Lee, L. C., and Chen, N.-Y. (2001). Effect of spatial resolution on classification errors of pure and mixed pixels in remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 39(12):2657–2663.

Huete, A. R. (1988). A soil-adjusted vegetation index (savi). *Remote sensing of environment*, 25(3):295–309.

Ibarrola-Ulzurrun, E., Gonzalo-Martín, C., and Marcello, J. (2017). Influence of pansharpening in obtaining accurate vegetation maps. *Canadian Journal of Remote Sensing*, 43(6):528–544.

Immitzer, M., Atzberger, C., and Koukal, T. (2012). Tree species classification with random forest using very high spatial resolution 8-band worldview-2 satellite data. *Remote sensing*, 4(9):2661–2693.

Indian Space Research Organisation (n.d.). Cartosat-3. `https://www.nrsc.gov.in/sites/default/files/pdf/training/C3_BROCHURE_JAN2021_modified.pdf`. Consulted on 20 March 2022.

Jiang, Z., Huete, A. R., Didan, K., and Miura, T. (2008). Development of a two-band enhanced vegetation index without a blue band. *Remote sensing of Environment*, 112(10):3833–3845.

Johnson, B. and Xie, Z. (2011). Unsupervised image segmentation evaluation and refinement using a multi-scale approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(4):473–483.

Karlson, M., Ostwald, M., Reese, H., Bazié, H. R., and Tankoano, B. (2016). Assessing the potential of multi-seasonal worldview-2 imagery for mapping west african agroforestry tree species. *International Journal of Applied Earth Observation and Geoinformation*, 50:80–88.

Kato, A., Moskal, L. M., Schiess, P., Swanson, M. E., Calhoun, D., and Stuetzle, W. (2009). Capturing tree crown formation through implicit surface reconstruction using airborne lidar data. *Remote sensing of Environment*, 113(6):1148–1162.

Kelly, M., Blanchard, S. D., Kersten, E., and Koy, K. (2011). Terrestrial remotely sensed imagery in support of public health: New avenues of research using object-based image analysis. *Remote Sensing*, 3(11):2321–2345.

Kim, M., Madden, M., and Warner, T. (2008). Estimation of optimal image object size for the segmentation of forest stands with multispectral ikonos imagery. In *Object-based image analysis*, pages 291–307. Springer.

Landis, J. R. and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, pages 159–174.

Lewinski, S., Bochenek, Z., and Turlej, K. (2010). Application of object-oriented method for classification of vhr satellite images using rule-based approach and texture measures. *Geoinformation Issues*, 2(1):19–26.

Lillesand, T., Kiefer, R. W., and Chipman, J. (2015). *Remote sensing and image interpretation*. John Wiley & Sons.

Luvai, A., Obiero, J., and Omuto, C. (2020). Methods for erosion estimates in assessment of soil degradation: A review for catchments in kenya. *Int. J. Eng. Res.*.

Madonsela, S., Cho, M. A., Mathieu, R., Mutanga, O., Ramoelo, A., Kaszta, Ż., Van De Ker-chove, R., and Wolff, E. (2017). Multi-phenology worldview-2 imagery improves remote sensing of savannah tree species. *International journal of applied earth observation and geoinformation*, 58:65–73.

Mallinis, G., Koutsias, N., Tsakiri-Strati, M., and Karteris, M. (2008). Object-based classifi-cation using quickbird imagery for delineating forest vegetation polygons in a mediter-ranean test site. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63(2):237–250.

Martinez-Taboada, F. and Redondo, J. I. (2020). Variable importance plot (mean decrease accuracy and mean decrease gini).

McFeeters, S. K. (1996). The use of the normalized difference water index (ndwi) in the delineation of open water features. *International journal of remote sensing*, 17(7):1425–1432.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*, 22(3):276–282.

Meinel, G. and Neubert, M. (2004). A comparison of segmentation programs for high resolu-tion remote sensing data. *International Archives of Photogrammetry and Remote Sensing*, 35(Part B):1097–1105.

Mishra, N. B. and Crews, K. A. (2014). Mapping vegetation morphology types in a dry savanna ecosystem: Integrating hierarchical object-based image analysis with random forest. *International Journal of Remote Sensing*, 35(3):1175–1198.

Mostafa, Y. and Abdelhafiz, A. (2017). Accurate shadow detection from high-resolution satel-lite images. *IEEE Geoscience and Remote Sensing Letters*, 14(4):494–498.

Munyati, C. (2018). Optimising multiresolution segmentation: Delineating savannah vege-tation boundaries in the kruger national park, south africa, using sentinel 2 msi imagery. *International Journal of Remote Sensing*, 39(18):5997–6019.

Naidoo, L., Cho, M. A., Mathieu, R., and Asner, G. (2012). Classification of savanna tree species, in the greater kruger national park region, by integrating hyperspectral and lidar data in a random forest data mining environment. *ISPRS journal of Photogrammetry and Remote Sensing*, 69:167–179.

National Coordination Office for Space-Based Positioning, Navigation, and Timing (2022). `https://www.gps.gov/systems/augmentations/`. Consulted on 28 May 2022.

Ng, W.-T., Rima, P., Einzmann, K., Immitzer, M., Atzberger, C., and Eckert, S. (2017). As-sessing the potential of sentinel-2 and pléiades data for the detection of prosopis and vachellia spp. in kenya. *Remote sensing*, 9(1):74.

Nussbaum, S. and Menz, G. (2008). ecognition image analysis software. In *Object-based image analysis and treaty verification*, pages 29–39. Springer.

Nussbaum, S., Niemeyer, I., and Canty, M. (2006). Seath-a new tool for automated feature extraction in the context of object-based image analysis. In *1st International Conference on Object-based Image Analysis (OBIA). Salzburg: Austria*.

Omuto, C. (2008). Assessment of soil physical degradation in eastern kenya by use of a sequential soil testing protocol. *Agriculture, ecosystems & environment*, 128(4):199–211.

open Africa (n.d.). `https://africaopendata.org/organization/code-for-kenya`. Consulted on 2 March 2022.

OpenStreetMap (n.d.). `https://download.geofabrik.de/africa/kenya.html`. Consulted on 2 March 2022.

Orwa, C., Mutua, A., Kindt, R., Jamnadass, R., and Anthony, S. (2009). Agroforestree Database: a tree reference and selection guide version 4.0 (`http://www.worldagroforestry.org/sites/treedbs/treedatabases.asp`). Consulted on 26 May 2022.

Pal, N. R. and Pal, S. K. (1993). A review on image segmentation techniques. *Pattern recognition*, 26(9):1277–1294.

Panagiotidis, D., Abdollahnejad, A., Surovỳ, P., and Chiteculo, V. (2017). Determining tree height and crown diameter from high-resolution uav imagery. *International journal of remote sensing*, 38(8-10):2392–2410.

Piermattei, L., Marty, M., Ginzler, C., Pöchtrager, M., Karel, W., Ressl, C., Pfeifer, N., and Hollaus, M. (2019). Pléiades satellite images for deriving forest metrics in the alpine region. *International Journal of Applied Earth Observation and Geoinformation*, 80:240–256.

Pramoditha, R. (2020). `https://towardsdatascience.com/k-fold-cross-validation-explained-in-plain-english-659e33c0bc0`. Consulted on 11 May 2022.

Pu, R. and Landry, S. (2020). Mapping urban tree species by integrating multi-seasonal high resolution pléiades satellite imagery with airborne lidar data. *Urban Forestry & Urban Greening*, 53:126675.

Pu, R., Landry, S., and Yu, Q. (2018). Assessing the potential of multi-seasonal high resolution pléiades satellite imagery for mapping urban tree species. *International Journal of Applied Earth Observation and Geoinformation*, 71:144–158.

Sanjay (2018). `https://towardsdatascience.com/why-and-how-to-cross-validate-a_-model-d6424b45261f#:~:text=K%2DFolds%20Cross%20Validation%3A&text=Because%20it%20ensures%20that%20every,have%20a%20limited%20input%20data`. Consulted on 11 May 2022.

Sarrazin, M., Van Aardt, J., Asner, G., McGlinchy, J., Messinger, D., and Wu, J. (2011). Fusing small-footprint waveform lidar and hyperspectral data for canopy-level species classification and herbaceous biomass modeling in savanna ecosystems. *Canadian Journal of Remote Sensing*, 37(6):653–665.

Satellite Imaging Corporation (n.d.). `https://www.satimagingcorp.com/satellite-sensors/superview-1/`. Consulted on 20 March 2022.

Smith, A. (2010). Image segmentation scale parameter optimization and land cover classification using the random forest algorithm. *Journal of Spatial Science*, 55(1):69–79.

Story, M. and Congalton, R. G. (1986). Accuracy assessment: a user's perspective. *Photogrammetric Engineering and remote sensing*, 52(3):397–399.

Sumsion, G. R., Bradshaw, M. S., Hill, K. T., Pinto, L. D., and Piccolo, S. R. (2019). Remote sensing tree classification with a multilayer perceptron. *PeerJ*, 7:e6101.

Svetnik, V., Liaw, A., Tong, C., and Wang, T. (2004). Application of breiman's random forest to modeling structure-activity relationships of pharmaceutical molecules. In *International workshop on multiple Classifier systems*, pages 334–343. Springer.

Tatar, N., Saadatseresht, M., Arefi, H., and Hadavand, A. (2018). A robust object-based shadow detection method for cloud-free high resolution satellite images over urban areas and water bodies. *Advances in Space Research*, 61(11):2787–2800.

The Ministry of Agriculture, Livestock, Fisheries and Co-operatives (MoALFC) (2021). Climate risk profile for kitui county. *Kenya County Climate Risk Profile Series*.

Thierion, V., Alleaume, S., Jacqueminet, C., Vigneau, C., Renaud, J., Michel, K., Breton, V., and Luque, S. (2014). The potential of pléiades imagery for vegetation mapping: an example of grasslands and pastoral environments. *Revue Française de Photogrammétrie et de Télédétection*, (208):105–110.

Tian, J. and Chen, D.-M. (2007). Optimization in multi-scale segmentation of high-resolution satellite images for artificial feature recognition. *International Journal of Remote Sensing*, 28(20):4625–4644.

Trias-Sanz, R., Stamon, G., and Louchet, J. (2008). Using colour, texture, and hierarchial segmentation for high-resolution remote sensing. *ISPRS Journal of Photogrammetry and remote sensing*, 63(2):156–168.

User Guide eCognition® Developer (2019). `https://docs.ecognition.com/v9.5.0/Page%20collection/eCognition%20Suite%20Dev%20UG.htm#Page%20collection/eCognition%20Suite%20Dev%20UG.htm%3FTocPath%3DeCognition%2520Suite%7CDeveloper%2520User%2520Guide%7C_____0`. Consulted on 25 January 2022.

Vahidi, H., Klinkenberg, B., Johnson, B. A., Moskal, L. M., and Yan, W. (2018). Mapping the individual trees in urban orchards by incorporating volunteered geographic information and very high resolution optical remotely sensed data: A template matching-based approach. *Remote Sensing*, 10(7):1134.

Van Coillie, F. M., Verbeke, L. P., and De Wulf, R. R. (2007). Feature selection by genetic algorithms in object-based classification of ikonos imagery for forest mapping in flanders, belgium. *Remote Sensing of Environment*, 110(4):476–487.

Vaughn, N. R., Moskal, L. M., and Turnblom, E. C. (2012). Tree species detection accuracies using discrete point lidar and airborne waveform lidar. *Remote sensing*, 4(2):377–403.

Wang, D., Wan, B., Qiu, P., Su, Y., Guo, Q., and Wu, X. (2018). Artificial mangrove species mapping using pléiades-1: An evaluation of pixel-based and object-based classifications with selected machine learning algorithms. *Remote Sensing*, 10(2):294.

Watts, J. and Lawrence, R. (2008). Merging random forest classification with an object-oriented approach for analysis of agricultural lands. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37(B7):2008.

Woodcock, C. E. and Strahler, A. H. (1987). The factor of scale in remote sensing. *Remote sensing of Environment*, 21(3):311–332.

Zhen, Z., Quackenbush, L. J., and Zhang, L. (2014). Impact of tree-oriented growth order in marker-controlled region growing for individual tree crown delineation using airborne laser scanner (als) data. *Remote Sensing*, 6(1):555–579.

Zou, S., Gader, P., and Zare, A. (2019). Hyperspectral tree crown classification using the multiple instance adaptive cosine estimator. *PeerJ*, 7:e6405.

# APPENDIX A
# LOCALISATION OF THE COLLECTED DATA

Table A.1: Coordinates (CRS WGS 84 / UTM zone 37S) of the grid cells (No.) collected during the field inventory.

| No. | East (UTM) | South (UTM) | No. | East (UTM) | South (UTM) |
|-----|------------|-------------|-----|------------|-------------|
| 1 | 378648.255 | 9925151.187 | 23 | 379315.064 | 9924069.307 |
| 2 | 379216.417 | 9925145.643 | 23 | 379402.699 | 9924079.318 |
| 3 | 379353.029 | 9925153.344 | 24 | 378780.631 | 9923663.785 |
| 4 | 378832.919 | 9924920.164 | 24 | 378808.354 | 9923739.868 |
| 4 | 378703.547 | 9925004.565 | 25 | 378982.699 | 9923712.607 |
| 5 | 379034.833 | 9924932.024 | 25 | 379224.965 | 9923860.000 |
| 5 | 379115.537 | 9924931.099 | 26 | 379320.455 | 9923706.447 |
| 6 | 379408.320 | 9924933.410 | 26 | 379470.466 | 9923782.838 |
| 6 | 379607.000 | 9925038.448 | 27 | 378861.797 | 9923560.440 |
| 7 | 379745.922 | 9924933.102 | 28 | 379192.468 | 9923520.550 |
| 8 | 378439.256 | 9924812.354 | 28 | 379011.038 | 9923494.060 |
| 10 | 379097.518 | 9924818.514 | 29 | 379512.512 | 9923444.775 |
| 10 | 378929.641 | 9924767.843 | 29 | 379313.986 | 9923498.834 |
| 11 | 379619.860 | 9924678.745 | 30 | 378545.065 | 9923266.117 |
| 11 | 379312.600 | 9924677.821 | 31 | 378925.868 | 9923148.835 |
| 12 | 379825.394 | 9924693.685 | 31 | 379278.409 | 9923332.498 |
| 13 | 378832.765 | 9924508.173 | 32 | 379429.883 | 9923188.031 |
| 13 | 378690.609 | 9924478.757 | 34 | 378795.340 | 9922944.302 |
| 14 | 379133.095 | 9924439.175 | 35 | 379009.883 | 9923043.180 |
| 14 | 379047.771 | 9924513.410 | 36 | 379631.489 | 9923008.604 |
| 15 | 379492.952 | 9924524.422 | 36 | 379580.818 | 9923005.369 |
| 15 | 379491.874 | 9924596.963 | 37 | 378339.685 | 9922728.912 |
| 16 | 379740.223 | 9924596.501 | 38 | 378576.484 | 9922785.281 |
| 17 | 378879.278 | 9924186.590 | 38 | 378531.203 | 9922809.154 |
| 18 | 379248.992 | 9924251.661 | 39 | 378924.712 | 9922713.356 |
| 18 | 379039.300 | 9924147.932 | 40 | 379453.293 | 9922811.233 |
| 19 | 379329.080 | 9924293.861 | 41 | 378476.759 | 9922418.879 |
| 19 | 379405.009 | 9924236.722 | 42 | 378728.189 | 9922552.256 |
| 20 | 379725.823 | 9924356.776 | 45 | 378438.640 | 9922260.012 |
| 21 | 378882.897 | 9923940.242 | 46 | 378568.013 | 9922269.253 |
| 22 | 379004.877 | 9924051.287 | 46 | 378600.125 | 9922196.250 |
| 22 | 379097.287 | 9924064.532 | 49 | 378596.737 | 9922076.272 |

Table A.2: Coordinates (CRS WGS 84 / UTM zone 37S) of the farms collected during the field inventory, with indication of the village where the farms are situated and the name of the farmer.

| No. | East (UTM) | South (UTM) | Village | Name of farmer |
|---|---|---|---|---|
| 1 | 381295.869 | 9925051.410 | Katithini | Stephene Mulwa Mukiti |
| 2 | 381118.480 | 9924378.304 | Katithini | Peter Muimi Muthui |
| 3 | 381653.000 | 9923907.733 | Katithini | Peter Mutua Gitune |
| 4 | 384629.458 | 9922789.833 | Mwangu | Muteti Muthui |
| 5 | 386408.941 | 9921059.196 | Maskanioni | Bonface Mutua Musyoka |
| 6 | 386404.878 | 9919557.809 | Masukanioni | Peter Munyoki Muthui |
| 7 | 386873.694 | 9920308.766 | Maskanioni | Peter Kimanzi Makau |
| 8 | 388171.355 | 9921223.529 | Maskanioni | Peter mwendwa Francis mutemi |
| 9 | 386361.811 | 9920473.338 | Maskanioni | Samuel musili munyoki |
| 10 | 390655.912 | 9924837.998 | Koriro | Benson Kimanzi Mwendwa |
| 11 | 388753.026 | 9925307.343 | Kalatine | Mboli Ngukuni |
| 12 | 388313.971 | 9925142.750 | Kalatine | Martha Mwendwa |
| 13 | 384217.805 | 9926323.223 | Tondora | Kilonzi Kimwele Ndatya |
| 14 | 383473.495 | 9927602.109 | Tondora | Muthakye Muteti |
| 15 | 382890.209 | 9926853.099 | Katithini | Peter Mutinda Muthengi |
| 16 | 381305.399 | 9925882.234 | Katithini | Muthui Masyuko |
| 17 | 381754.064 | 9924579.663 | Katithini | Mutuku Kitune |

# APPENDIX B
# DESCRIPTIVE STATISTICS

Table B.1: Descriptive statistics of neem (*Azadirachta indica*), acacia (*Acacia* sp.), and baobab (*Adansonia* sp.) in the plantation. The trees are divided in classes of 5 cm diameter at breast height (DBH) (cm), except for baobab as there were only three trees measured of this species. For each class, the average crown diameter (cm), height (m), age (year), and vitality (1 to 5) are calculated. Additionally, the number of trees (%) in each class is represented.

| DBH (cm) | Crown diameter (cm) | Height (m) | Tree age (year) | Tree vitality | Number of trees (%) |
|---|---|---|---|---|---|
| ***Azadirachta indica*** | | | | | |
| 0.0 - 5.0 | 155.46 | 3.34 | 14.00 | 2.60 | 13.87 |
| 5.0 - 10.0 | 251.61 | 5.39 | 14.00 | 1.94 | 46.72 |
| 10.0 - 15.0 | 337.15 | 7.90 | 13.90 | 1.42 | 32.36 |
| 15.0 - 20.0 | 386.88 | 9.52 | 13.85 | 1.58 | 6.33 |
| 20.0 - 25.0 | 595.17 | 9.57 | 14.00 | 1.00 | 0.73 |
| | | | | | |
| ***Acacia* sp.** | | | | | |
| 0.0 - 5.0 | 397.75 | 2.50 | 10.00 | 2.00 | 2.50 |
| 5.0 - 10.0 | 462.26 | 6.84 | 10.00 | 1.96 | 6.57 |
| 10.0 - 15.0 | 612.06 | 10.75 | 10.00 | 1.09 | 8.27 |
| 15.0 - 20.0 | 664.66 | 11.44 | 10.00 | 1.00 | 3.89 |
| 20.0 - 25.0 | 776.50 | 9.60 | 10.00 | 1.00 | 0.24 |
| | | | | | |
| ***Adansonia* sp.** | | | | | |
| 0.0 - 50.0 | 280.00 | 4.00 | 30.00 | 1.00 | 66.67 |
| 350.0 - 400.0 | 2377.50 | 21.30 | 600.00 | 1.00 | 33.33 |

# APPENDIX C
# CLASSIFICATION MAPS

In this appendix, the classification maps of the 17 farms, where ground truth data was collected, are shown. It has to be noted that the classification reached a very low accuracy (Cohen's Kappa coefficient of 0.118), so these maps are not representative for the actual conditions on the ground, but are rather a visualisation of the classification results.
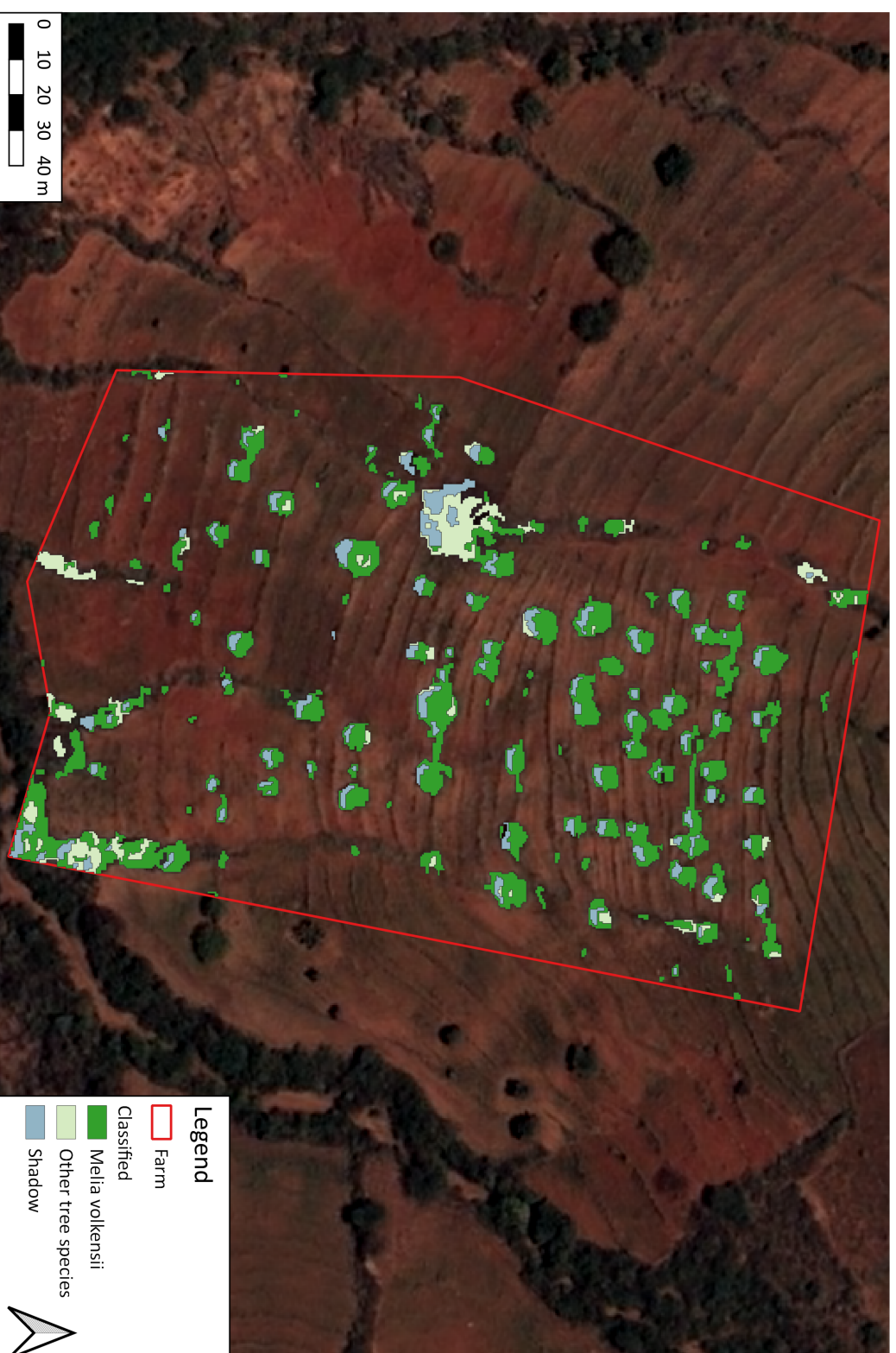
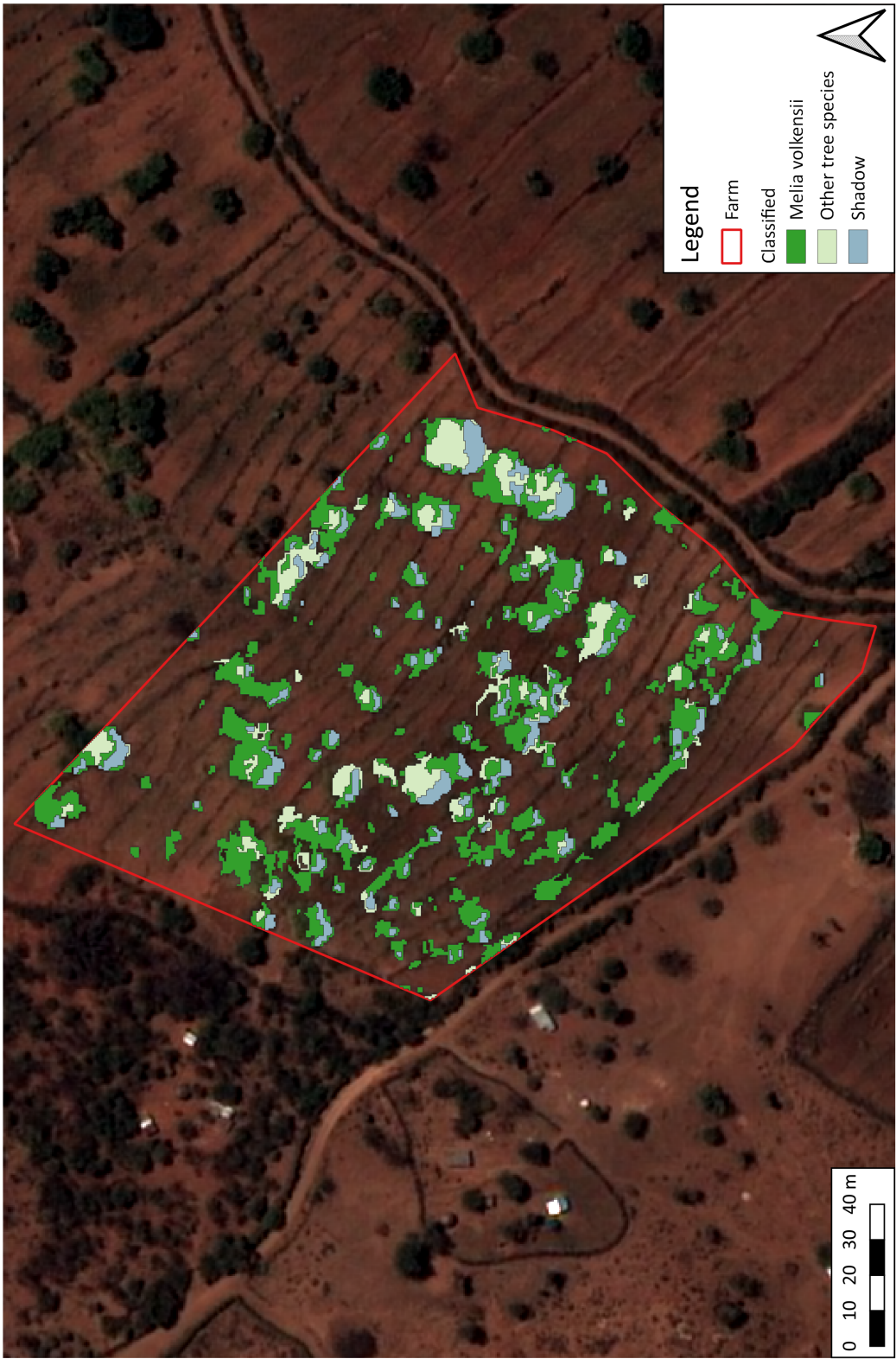Figure C.1: The classification result of the proposed method to differentiate *Melia volkensii* in farm 1 (Table A.2).
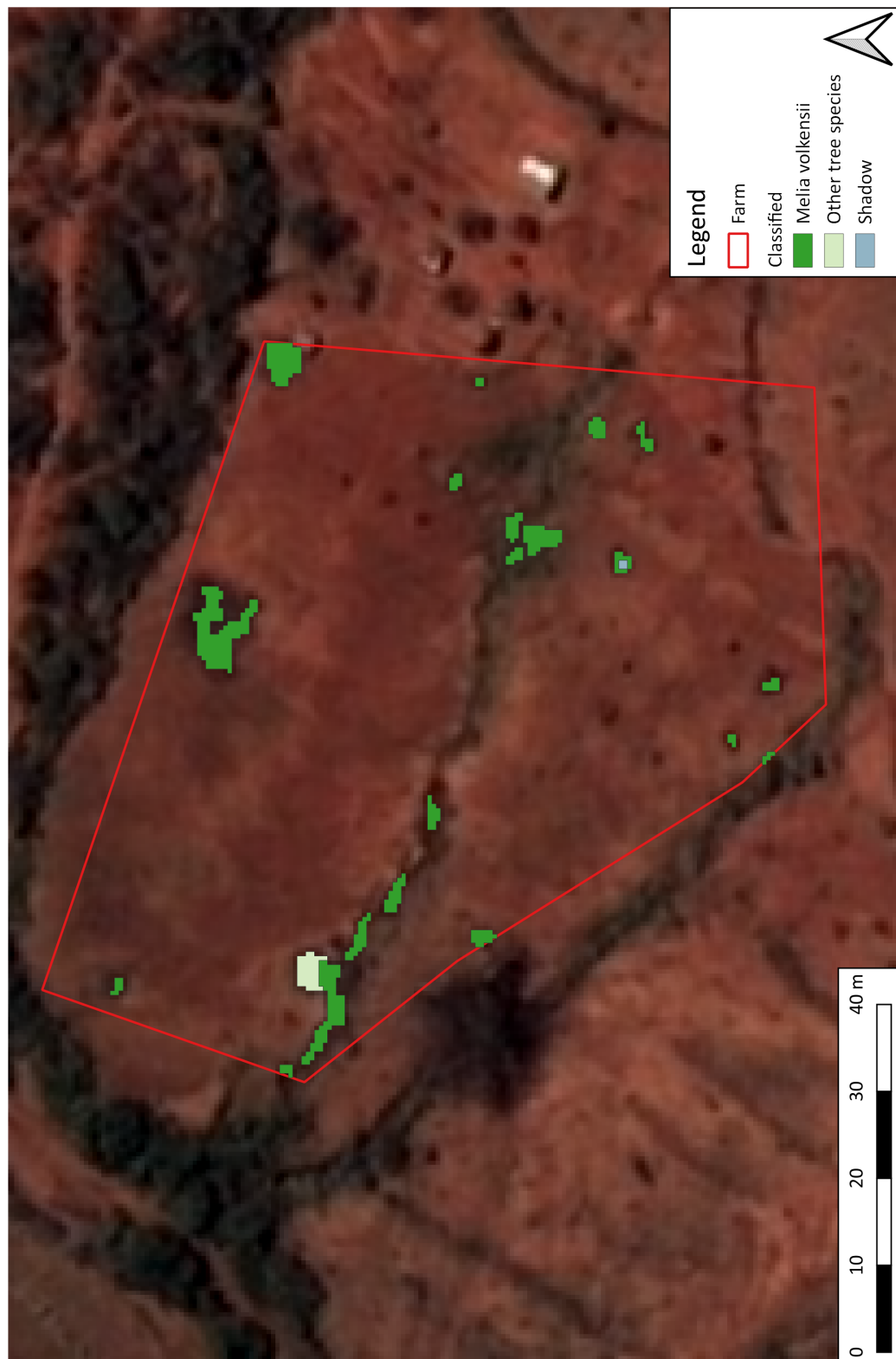
Figure C.2: The classification result of the proposed method to differentiate *Melia volkensii* in farm 2 (Table A.2).
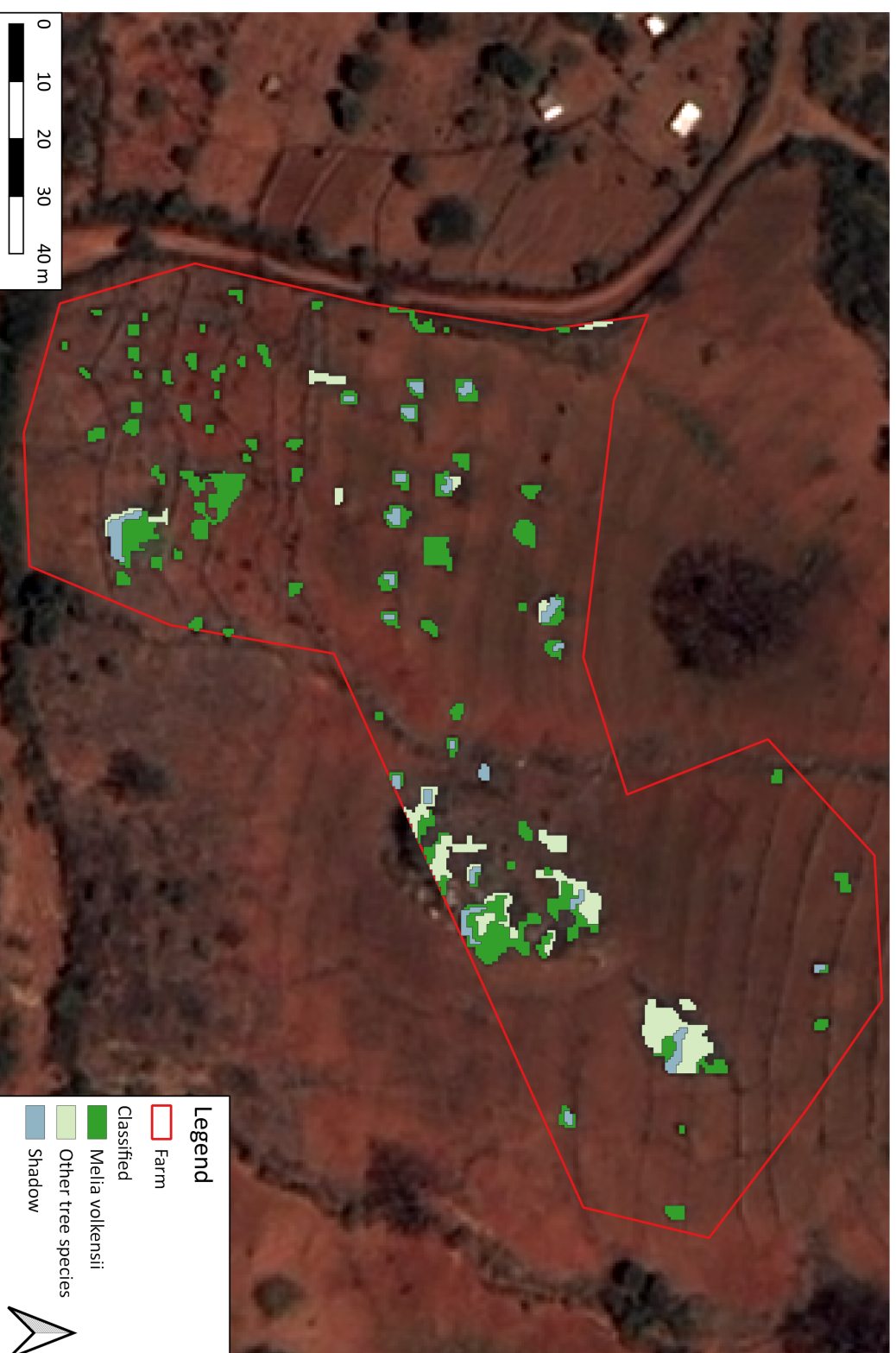
Figure C.3: The classification result of the proposed method to differentiate *Melia volkensii* in farm 3 (Table A.2).
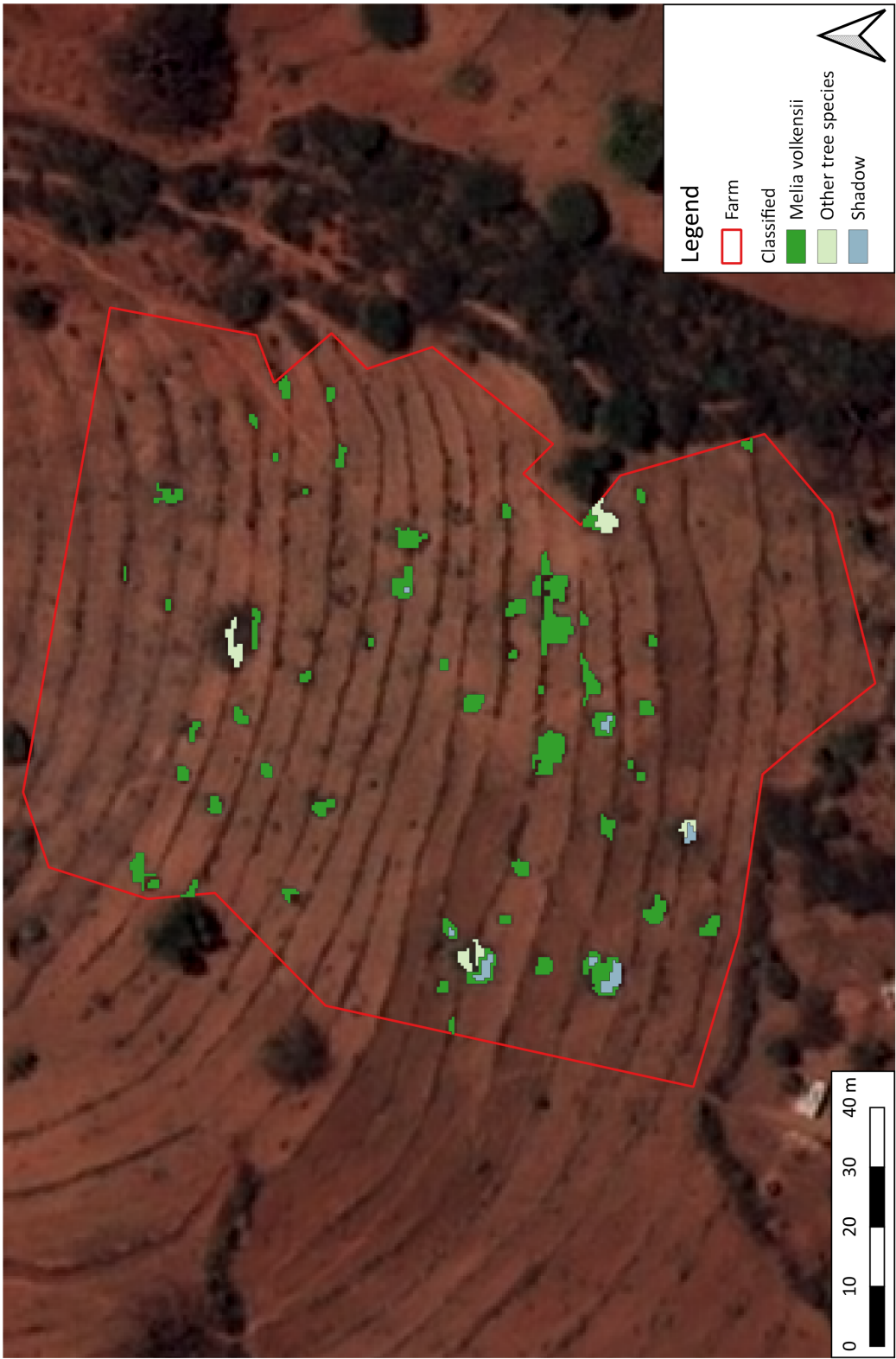
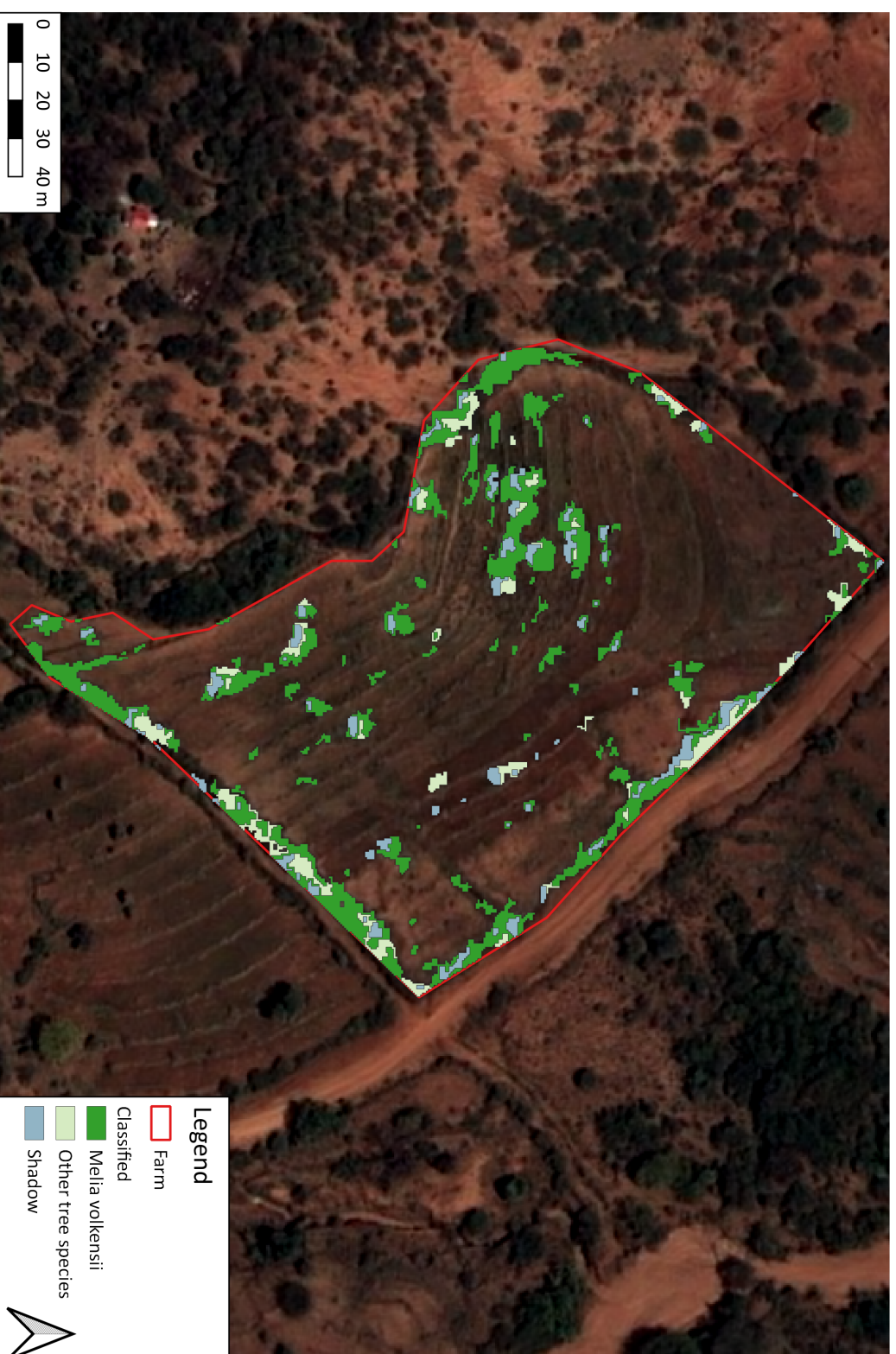Figure C.4: The classification result of the proposed method to differentiate *Melia volkensii* in farm 4 (Table A.2).
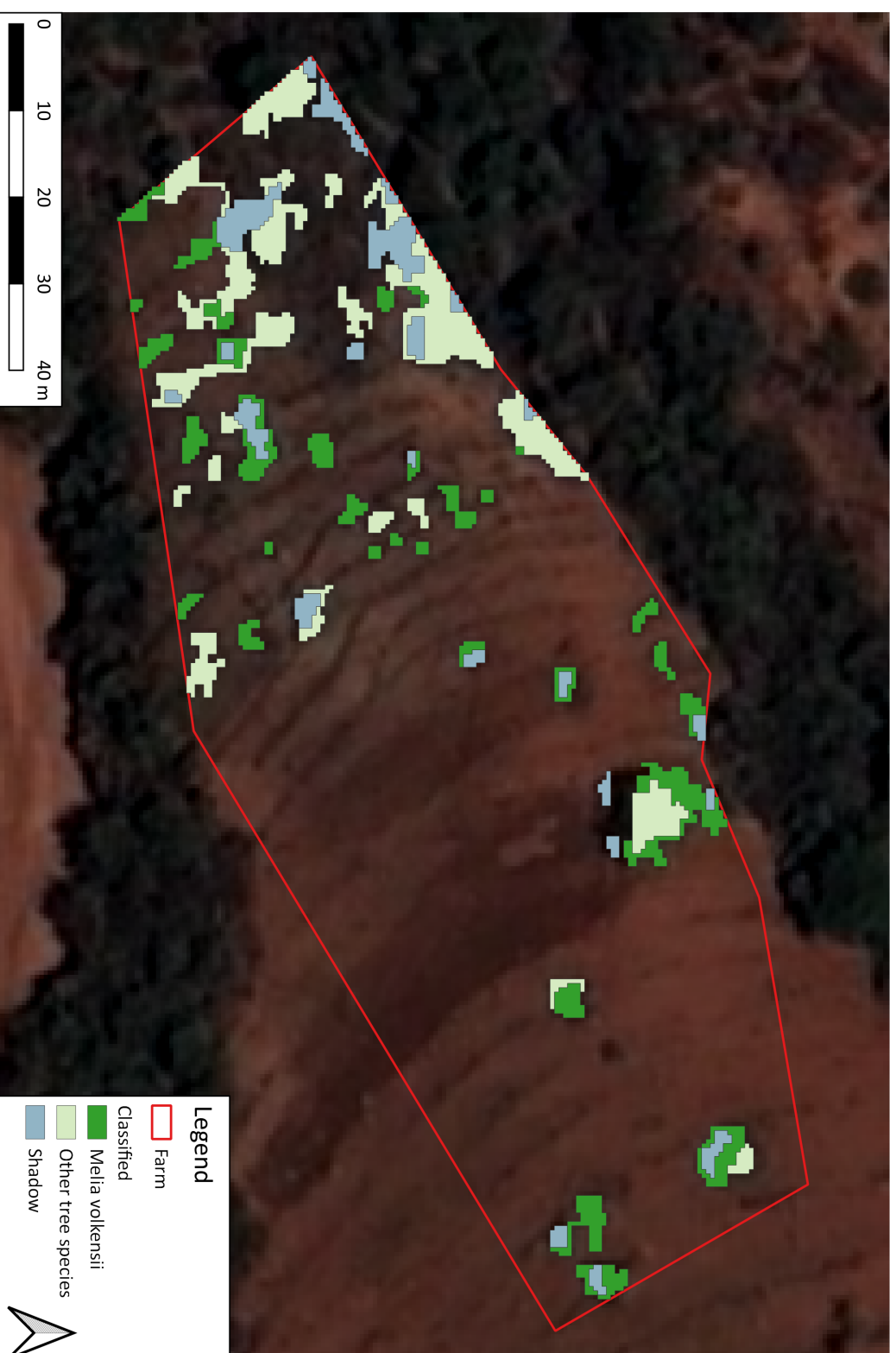
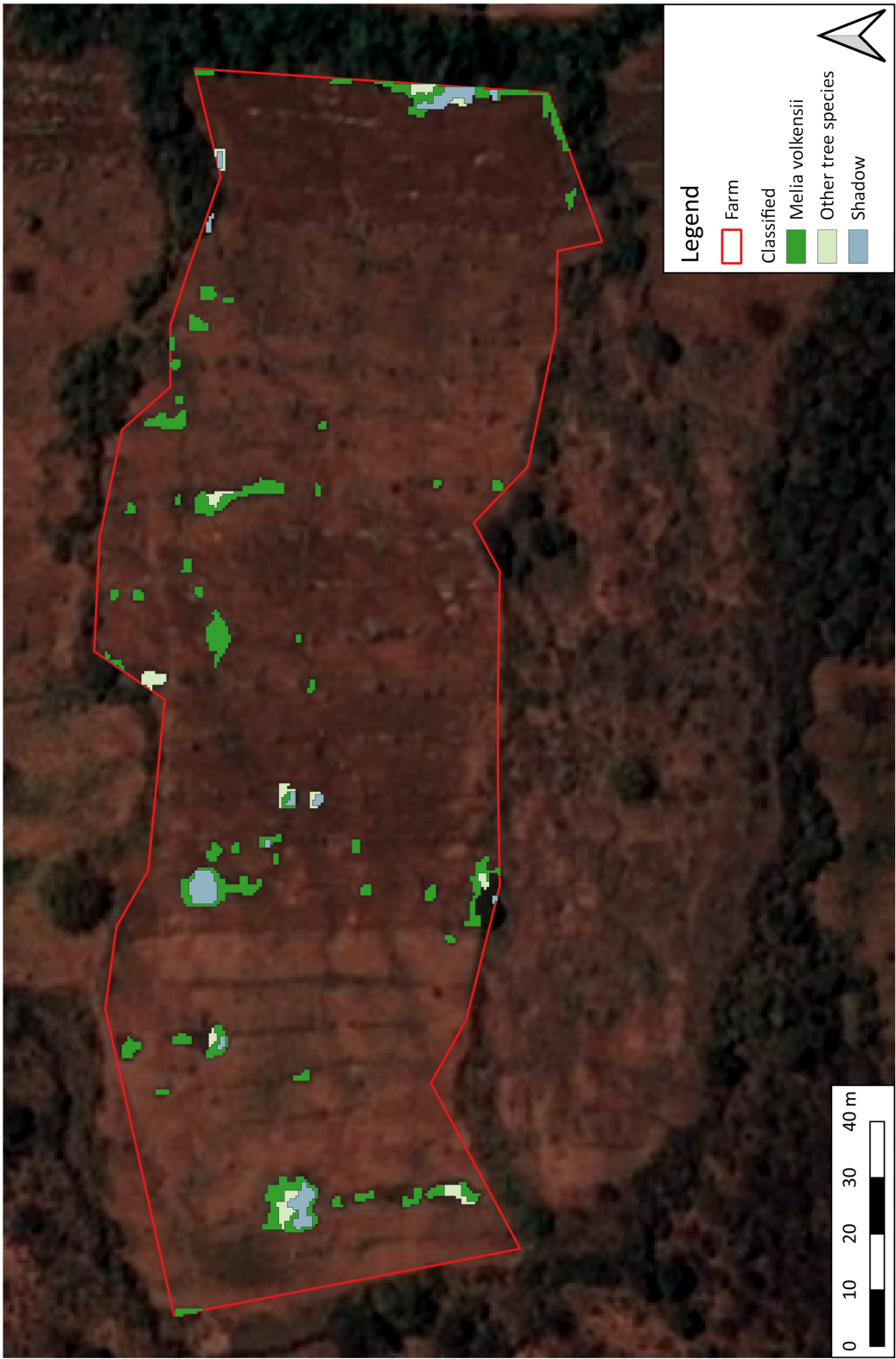Figure C.5: The classification result of the proposed method to differentiate *Melia volkensii* in farm 5 (Table A.2).
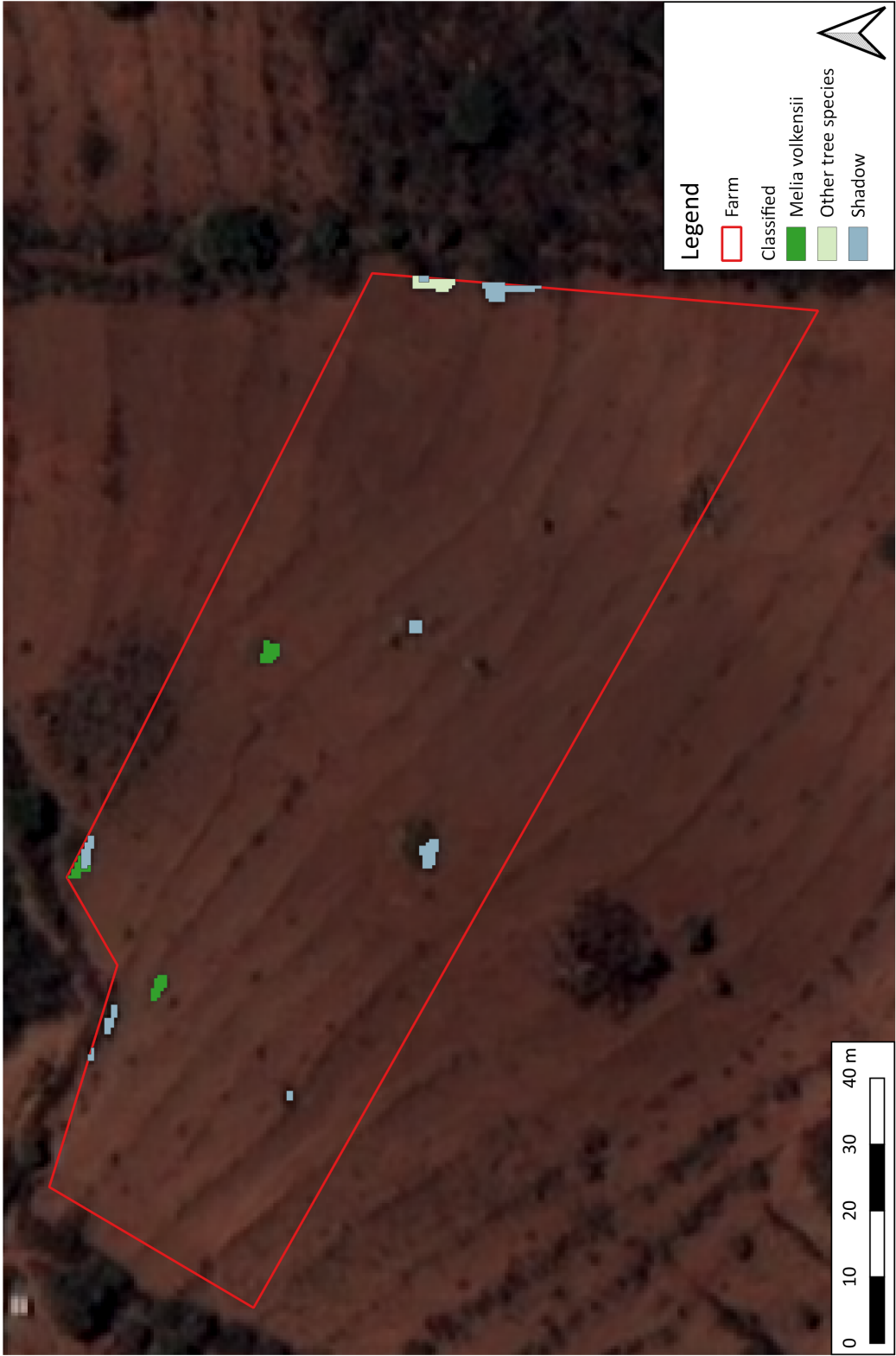
Figure C.6: The classification result of the proposed method to differentiate *Melia volkensii* in farm 6 (Table A.2).
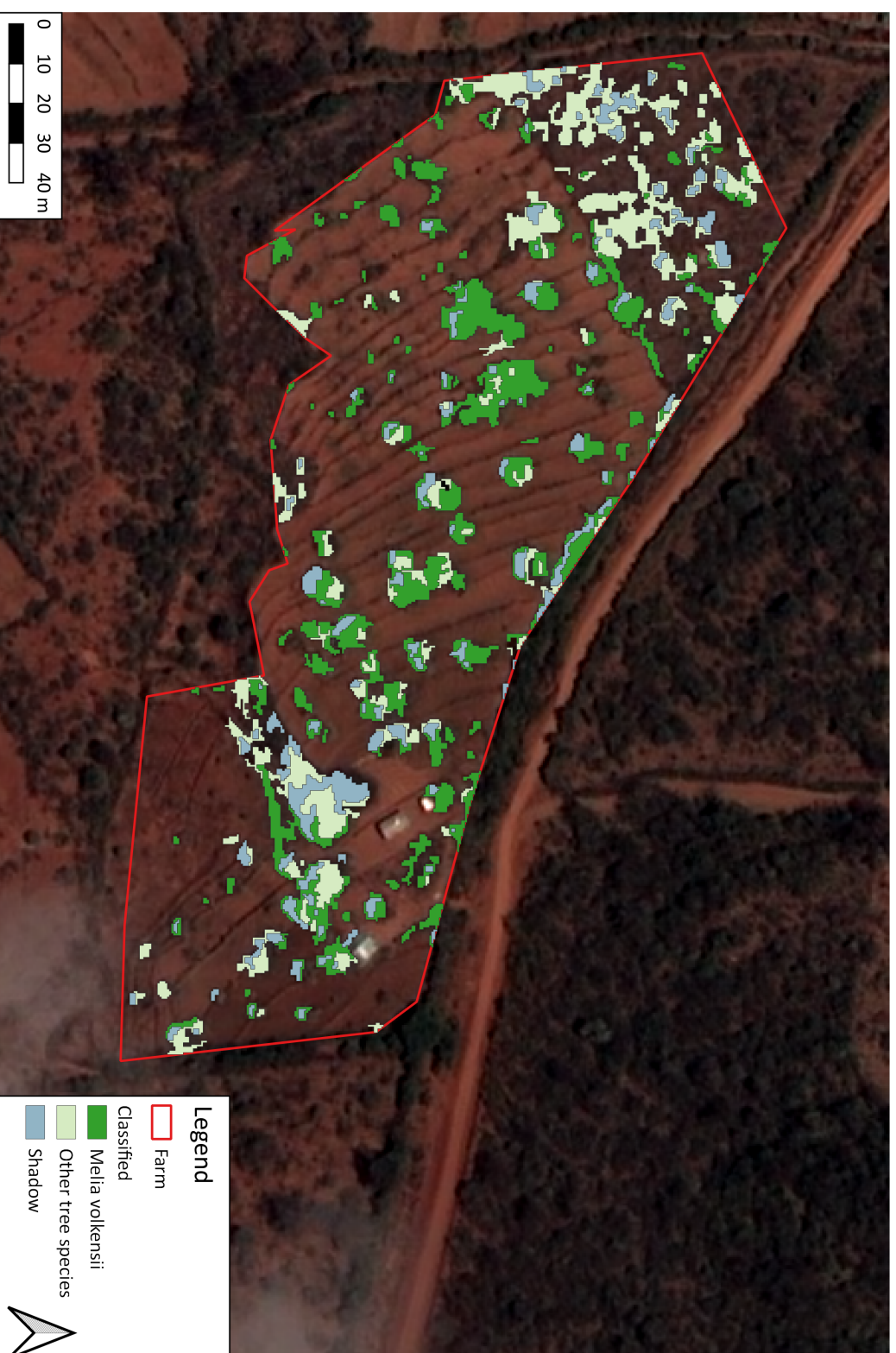
Figure C.7: The classification result of the proposed method to differentiate *Melia volkensii* in farm 7 (Table A.2).

Legend

Farm
Classified
Melia volkensii
Other tree species
Shadow

0  10  20  30  40 m
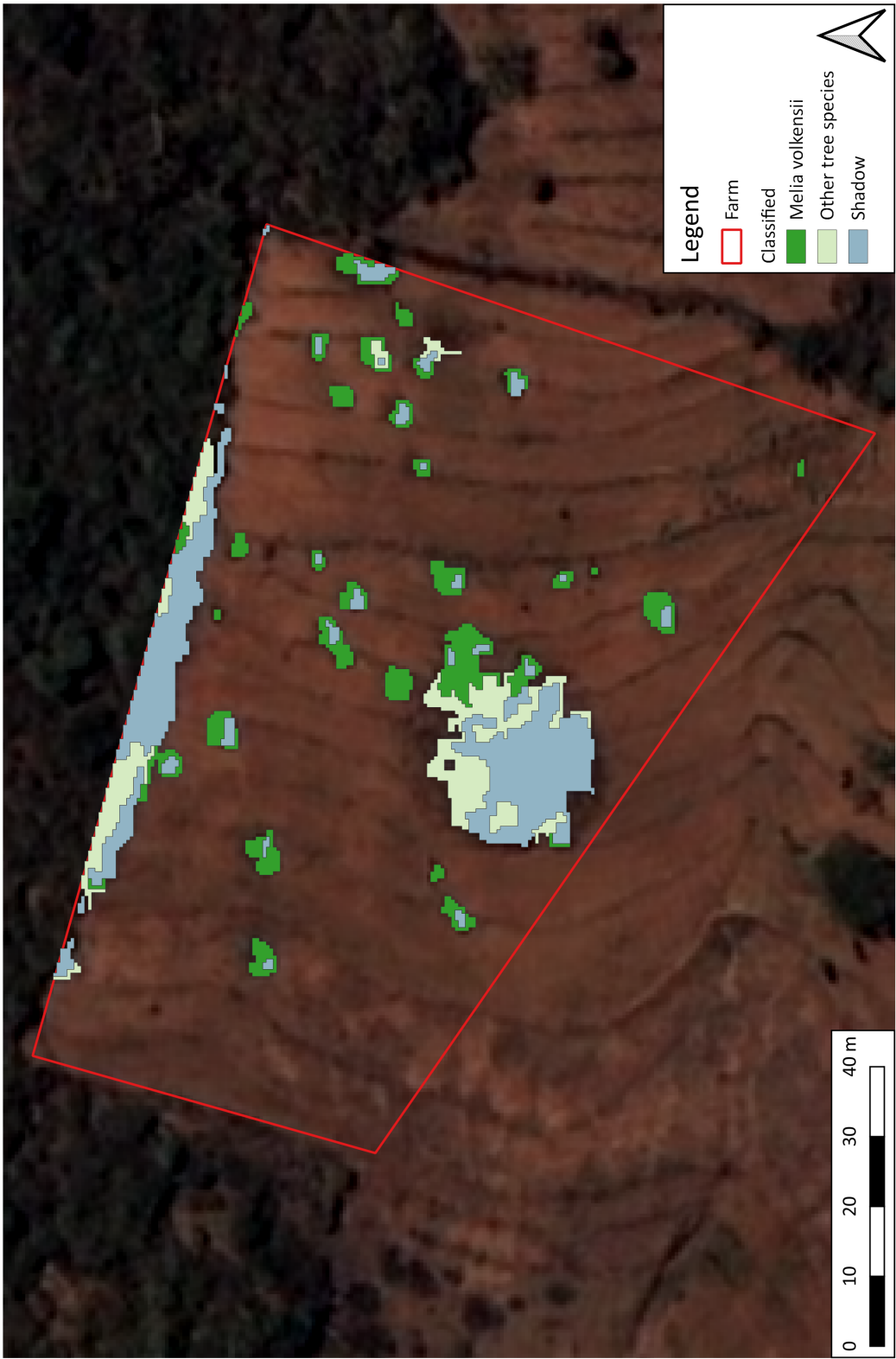
Figure C.8: The classification result of the proposed method to differentiate *Melia volkensii* in farm 8 (Table A.2).

Figure C.9: The classification result of the proposed method to differentiate *Melia volkensii* in farm 9 (Table A.2).

Figure C.10: The classification result of the proposed method to differentiate *Melia volkensii* in farm 10 (Table A.2).
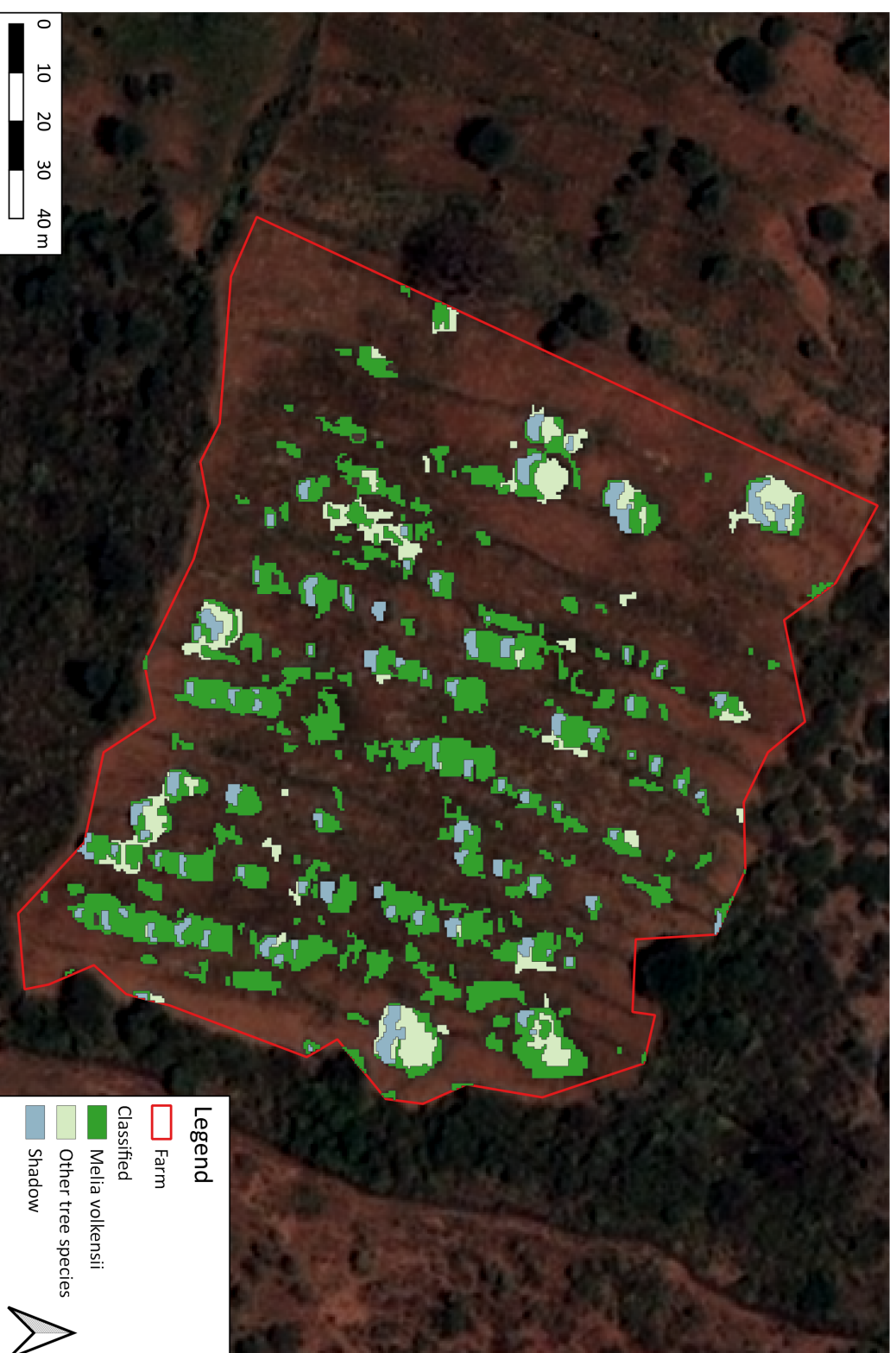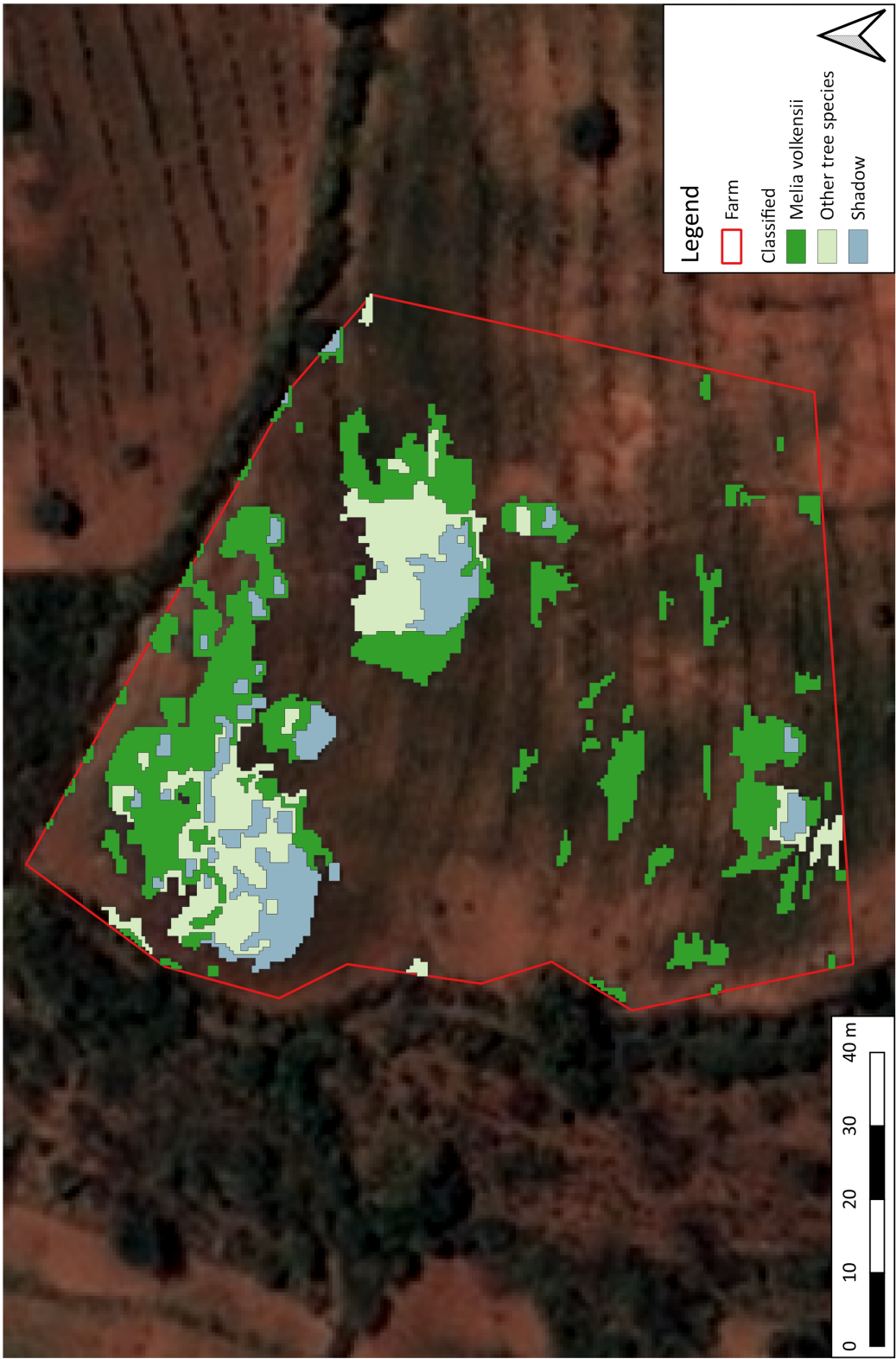
Figure C.11: The classification result of the proposed method to differentiate *Melia volkensii* in farm 11 (Table A.2).
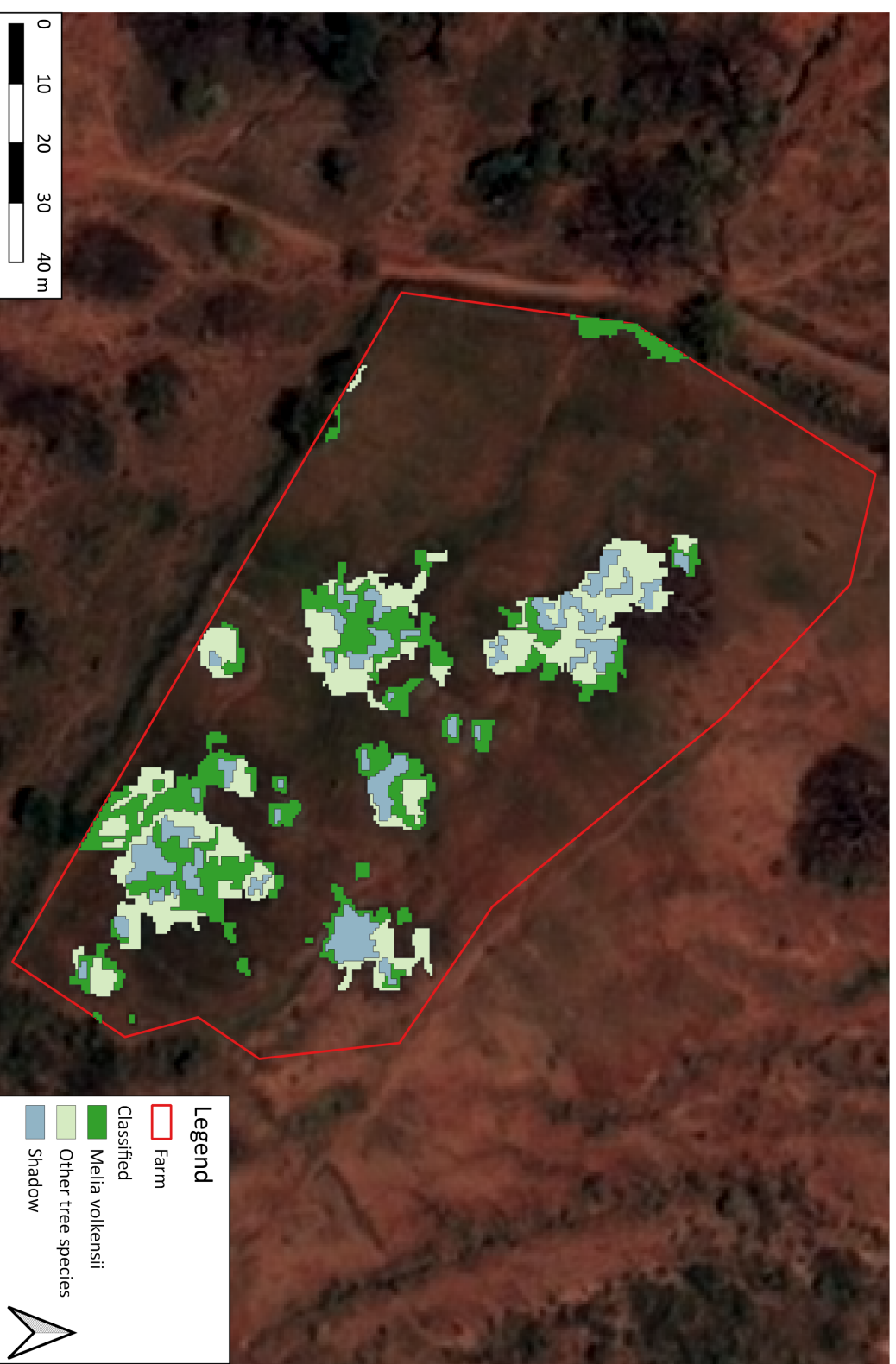
Figure C.12: The classification result of the proposed method to differentiate *Melia volkensii* in farm 12 (Table A.2).

Figure C.13: The classification result of the proposed method to differentiate *Melia volkensii* in farm 13 (Table A.2).

Figure C.14: The classification result of the proposed method to differentiate *Melia volkensii* in farm 14 (Table A.2).

Figure C.15: The classification result of the proposed method to differentiate *Melia volkensii* in farm 15 (Table A.2).
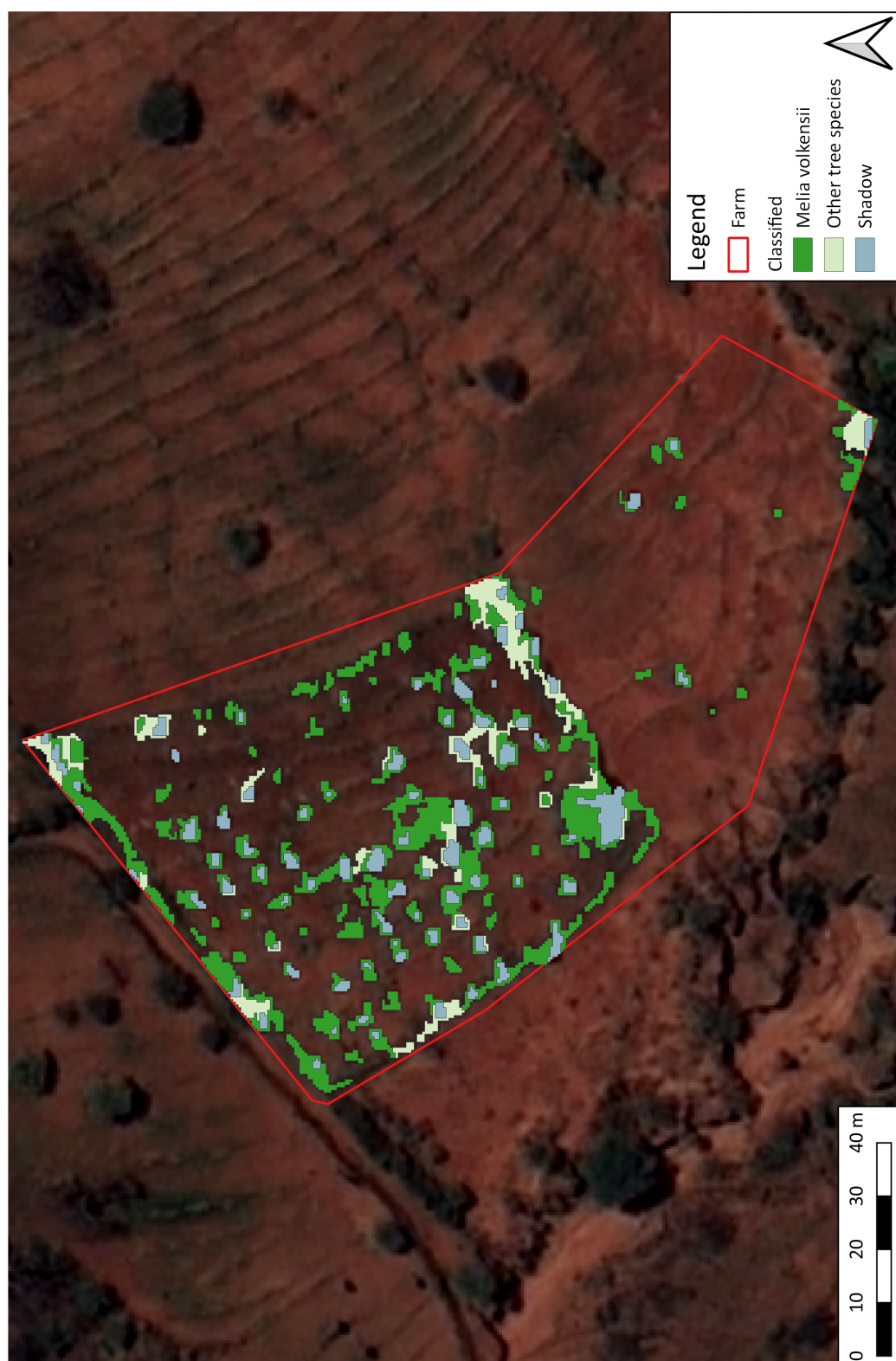
Figure C.16: The classification result of the proposed method to differentiate *Melia volkensii* in farm 16 (Table A.2).
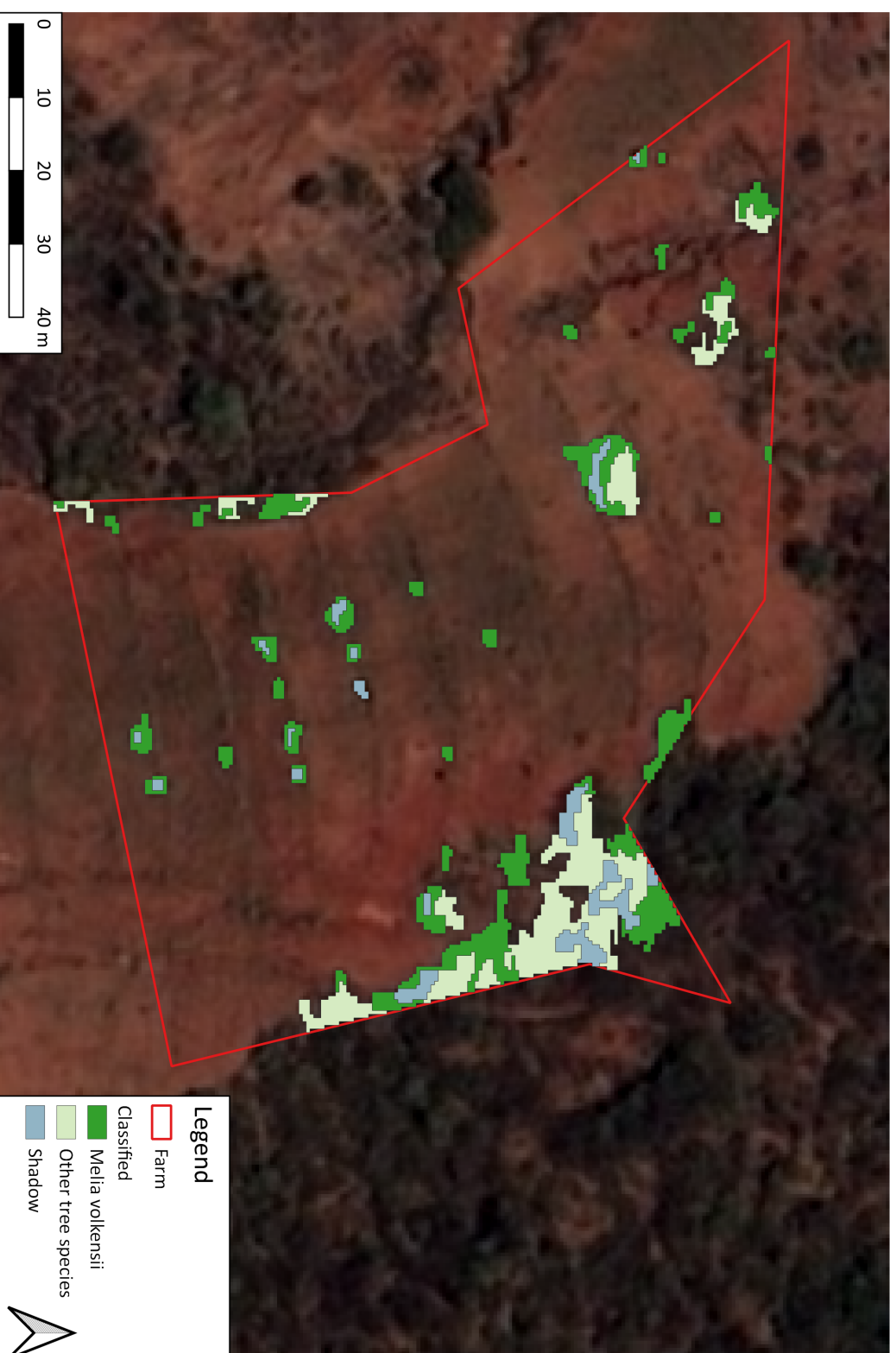
Figure C.17: The classification result of the proposed method to differentiate *Melia volkensii* in farm 17 (Table A.2).

Legend

Farm
Classified
  Melia volkensii
  Other tree species
  Shadow

0  10  20  30  40 m