

SUPER RESOLUTION MAPPING OF TREES FROM SATELLITE IMAGES AT DIFFERENT SCALES

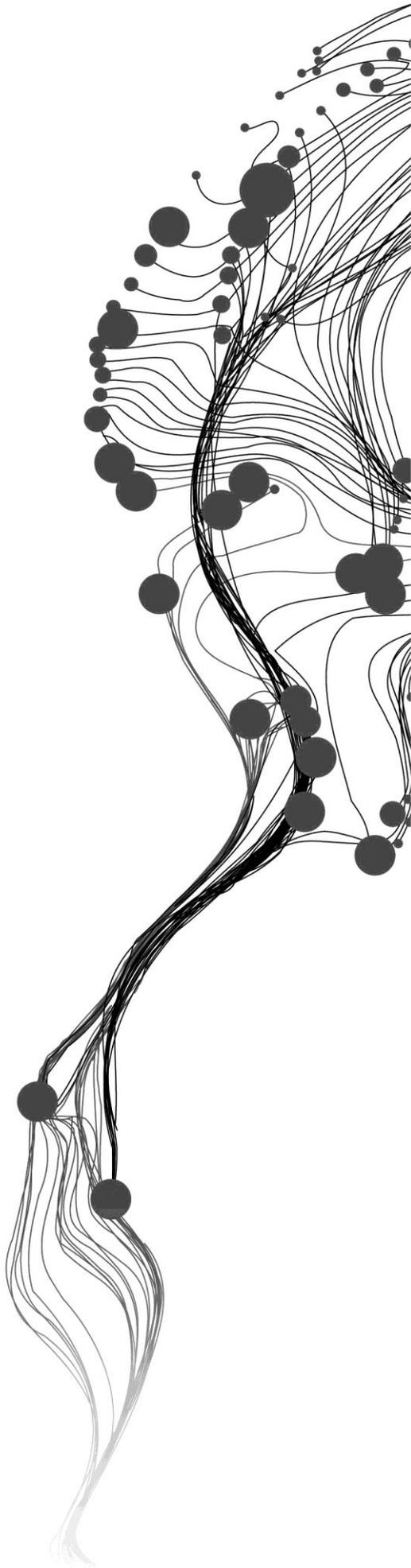
NDUJI NWANNEBUIKE NIXON

March, 2017

SUPERVISORS:

Dr. V. A. Tolpekin

Prof. Dr. Ir. A. Stein



SUPER RESOLUTION MAPPING OF TREES FROM SATELLITE IMAGES AT DIFFERENT SCALES

NDUJI NWANNEBUIKE NIXON

Enschede, The Netherlands, March, 2017

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.

Specialization: Geoinformatics

SUPERVISORS:

Dr. V. A. Tolpekin

Prof. Dr. Ir. A. Stein

THESIS ASSESSMENT BOARD:

Prof. Dr. Ir. M. G. Vosselman (Chair)

Dr. M. Oesting (External Examiner, University of Siegen)

DISCLAIMER

This document describes work undertaken as part of a programme of study at the Faculty of Geo-Information Science and Earth Observation of the University of Twente. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the Faculty.

ABSTRACT

Extraction of individual trees in urban environment and developed cities using satellite imageries has been quite challenging with many approaches to it. However, much more challenging is the mapping of this information in rural areas and developing countries. A number of reasons could be attributed to this ranging from indiscriminate planting of trees dispersed within heterogeneous and indistinct fields, to widespread shortage of up to date agricultural data which is partly due to lack of funds or limited government capacity and finally poor government policies and restrictions to finer resolution imageries.

With the launching of Sentinel satellite missions by the European space agency (ESA); medium and coarse resolution imageries have become freely and readily available. Most developing countries will like to leverage on this opportunity for up to date mapping and monitoring of their environment. A general limitation however exists in the spatial distribution of landcover data sets derived from most of these imageries due to varying contrast sensitivity. At coarse resolution and depending on the area of study, a row of trees may appear as individual trees, while single trees are much more visible at finer scale. This effect made some researchers to suggest that coarse resolution images are improper for tree detection. Although finer resolution images have improved the visibility and identification of both spectral and spatial information, it also has some limitations. The main limitation is due to the sensors large dynamic range which captures complex models making it difficult for spectral classifiers to deal with spectral separability between tree objects and other similar background classes. All these challenges further limit efforts to map or monitor individual trees in this data-sparse regions of the world.

This research is therefore focused on quantifying and comparing the detection of tree objects at different scale (coarse and fine) using super resolution mapping (SRM). To achieve this, we applied the pixel based MRF-SRM and partitioned the classified SRM thematic results into objects (segments) using region growing segmentation algorithm as post a classification procedure. The limitation of spectral information available in finer resolution images was addressed by the contextual approach of MRF; while the spatial resolution limitation for coarse images was reduced using SRM. The aim is to substantiate the reliability of individual tree crowns detected from coarse resolution images against the backdrop that suggests it's improper for tree detection. Also, the study will assist efforts to improve the capacity of monitoring agricultural change in data sparse regions and developing countries of the world. The results of our method extended beyond the norm of existing spatial dependence principle evident in pixel based SRM analysis, identified and validated individual trees taking cognisance of its range of characteristics.

Key words: *Extraction of individual trees, Super resolution mapping, Markov random fields, Class separability, Simulated annealing, Object-based image analysis, Image segmentation.*

ACKNOWLEDGEMENTS

All glory to my Lord and Saviour Jesus Christ for grace to accomplish this thesis.

I will like to express my profound gratitude to my first supervisor, Dr V.A. Tolpekin for encouraging me to work hard, think critically and his continuous support with ideas in the course of this research. God bless you so much. I also want to acknowledge my second supervisor, Prof. Dr. Ir. A. Stein, for his guidance and fruitful contribution to this work. May God bless you too.

My deepest appreciation goes to my parents Mr and Mrs Festus Nduji and family members, for their continuous prayers and encouragements throughout my study here. Thank you so much and am so proud of you all.

Special thanks goes to my boss Mr Chiemeka Ngwu and the Surveyor General of Plateau State, Nigeria, Surv. Tongkam Shenin Dasihit for their support and immense contribution towards achieving this educational pursuit. I sincerely appreciate and may God bless you.

Finally, I will like to acknowledge all my friends and classmates in GFM department, my colleagues in other departments and the wonderful people of Enschede, Netherlands. May God almighty bless you all.

TABLE OF CONTENTS

1.	INTRODUCTION.....	1
1.1.	Motivation and problem statement.....	1
1.2.	Research identification.....	3
1.2.1.	Research objectives.....	3
1.2.2.	Research questions.....	3
1.3.	Innovation.....	3
2.	LITERATURE REVIEW.....	2-5
2.1.	Object-based image analysis.....	2-5
2.2.	Image segmentation.....	2-5
2.3.	Maximum likelihood classification.....	2-6
2.4.	Super resolution mapping.....	2-6
2.5.	Detection of trees from remote sensing imagery.....	2-7
3.	STUDY AREA AND DATA.....	3-8
3.1.	Study area.....	3-8
3.2.	Available data.....	3-9
3.3.	Reference data.....	3-10
3.4.	Software.....	3-10
4.	METHODOLOGY.....	4-11
4.1.	Spectral class definition.....	4-11
4.2.	Markov random field.....	4-12
4.3.	Pixel based MRF-SRM.....	4-13
4.3.1.	Prior energy function.....	4-14
4.3.2.	Likelihood energy function.....	4-14
4.3.3.	Posterior energy function.....	4-15
4.4.	Parameter estimation.....	4-15
4.4.1.	Smoothness parameter.....	4-15
4.4.2.	Simulated annealing.....	4-16
4.5.	Object based image analysis of SRM result.....	4-16
4.5.1.	Segmentation of (SR map) c.....	4-16
4.6.	Accuracy assessment.....	4-17
4.6.1.	Pixel based accuracy assessment.....	4-17
4.6.2.	Object based accuracy assessment.....	4-17
5.	RESULTS.....	5-20
5.1.	Section - 1: Sentinel-2 multispectral image.....	5-20
5.1.1.	Spectral class definition.....	5-20
5.1.2.	Parameter estimation.....	5-22
5.1.3.	Classification results.....	5-23
5.2.	Section - 2: Worldview-3 multispectral image.....	5-27
5.2.1.	Spectral class definition.....	5-27
5.2.2.	Parameter estimation.....	5-29
5.2.3.	Classification results.....	5-30
	Section - 3 Comparison of the results at different scale.....	5-34
6.	DISCUSSION.....	6-36
6.1.	Section 1: Parameter estimation.....	6-36
6.1.1.	Sentinel-2 multispectral image.....	6-36

6.1.2. Worldview-3 multispectral image	6-37
6.2. Section 2: Results and comparison of detected tree objects	6-37
6.2.1. Results of Sentinel-2 multispectral image	6-37
6.2.2. Results of Worldview-3 multispectral image	6-39
6.2.3. Comparison of detected tree objects	6-40
7. CONCLUSIONS AND RECOMMENDATIONS	7-42
7.1. Conclusions	7-42
7.2. Recommendations	7-42

LIST OF FIGURES

Figure 3.1: Overview of study area (yellow boundary) using Sentinel-2 multispectral image displayed in NIR (band8), Green (band3) and Blue (band2).	3-8
Figure 3.2: Sentinel - 2 multispectral image subset of the study area displayed in NIR (band8), Green (band3) and Blue (band2).	3-9
Figure 3.3: Reference vectors of tree crowns (yellow) extracted from Worldview-3 panchromatic image of the study area.	3-10
Figure 4.1: Example of Sentinel-2 multispectral image showing the study area (yellow rectangle) and training sets (green polygons and points).	4-11
Figure 4.2: Showing (a) first order neighbourhood system; (b) second order neighbourhood system; (c) cliques corresponding to neighbourhood system.	4-12
Figure 4.3: Example of accuracy measure of tree crown objects; (A) existential accuracy measure, (B) extensional accuracy measure (C) positional accuracy measure.	4-19
Figure 5.1: Showing feature space plots of Sentinel-2 multispectral image.	5-21
Figure 5.2: Showing gradual reduction of mean energy using Sentinel-2 multispectral image (A) Tupd parameter is plotted against mean energy (B) T0 parameter is plotted against mean energy. The red line shows the mean energy and the bars show the standard deviation from the mean.	5-22
Figure 5.3: Showing subset of (A) Sentinel-2 multispectral image and (B) MLC result of Sentinel-2 multispectral image.	5-24
Figure 5.4: Showing subset of (A) Sentinel-2 multispectral image and (B) pixel-based MRF-SRM classification result of Sentinel-2 multispectral image.	5-25
Figure 5.5: Showing subset of (A) Sentinel-2 multispectral image and (B) object-based image analysis SRM result using Sentinel-2 multispectral image.	5-26
Figure 5.6: Showing the accuracy map for object-based image analysis of SRM classification result using the Sentinel-2 multispectral image.	5-26
Figure 5.7: Showing feature space plot using the Worldview-3 multispectral image.	5-28
Figure 5.8: Showing gradual reduction of mean energy using Worldview-3 multispectral image (A) Tupd parameter is plotted against mean energy (B) T0 parameter is plotted against mean energy. The red line shows the mean energy and the bars show the standard deviation from the mean.	5-29
Figure 5.9: Showing subset of (A) Worldview-3 multispectral image and (B) MLC result of Worldview-3 multispectral image.	5-31
Figure 5.10: Showing subset of (A) Worldview-3 multispectral image and (B) pixel-based MRF classification result using the Worldview-3 multispectral image.	5-32
Figure 5.11: Showing subset of (A) Worldview-3 multispectral image and (B) object-based image analysis of MRF result using the Worldview-3 multispectral image.	5-33
Figure 5.12: Showing the accuracy map of object-based image analysis of MRF classification results using the Worldview-3 multispectral image.	5-33
Figure 5.13: Showing the scatter plots of overestimation and underestimation for (A) Sentinel-2 multispectral image and (B) Worldview-3 multispectral image.	5-34

LIST OF TABLES

Table 5.1: Showing the total number of training samples for each class using the Sentinel-2 multispectral image.....	5-20
Table 5.2: Showing the contingency matrix for the Sentinel-2 multispectral image.....	5-20
Table 5.3: Showing the class separability for the Sentinel-2 multispectral image using Jeffries-Matusita distance.....	5-22
Table 5.4: Showing the updating temperature parameter T_{upd} estimation for Sentinel-2 multispectral image.....	5-23
Table 5.5: Showing the initial temperature parameter T_0 estimation for Sentinel-2 multispectral image.....	5-23
Table 5.6: Showing the smoothness parameter estimation using the Sentinel-2 multispectral image.....	5-23
Table 5.7: Showing the confusion matrix of MLC result using the Sentinel-2 multispectral image with accuracy assessment.	5-24
Table 5.8: Showing the confusion matrix of pixel-based MRF-SRM classification result using the Sentinel-2 multispectral image with accuracy assessment.	5-25
Table 5.9: Showing results of object based accuracy measure applied on the Sentinel-2 multispectral image (total detection, false negative, positional accuracy, local area difference, overestimated, underestimated total detection error).....	5-27
Table 5.10: Showing the total number of training samples for each class using the Worldview-3 multispectral image.....	5-27
Table 5.11: Showing the contingency matrix of the Worldview-3 multispectral image.....	5-28
Table 5.12: Showing the class separability of the Worldview-3 multispectral image using Jeffries-Matusita distance.....	5-29
Table 5.13: Showing the initial temperature parameter estimation using the Worldview-3 multispectral image.	5-30
Table 5.14: Showing the updating parameter T_{upd} estimation using the Worldview-3 multispectral image.....	5-30
Table 5.15: Showing the smoothness parameter estimation using the Worldview-3 multispectral image.....	5-30
Table 5.16: Showing the confusion matrix of MLC result using the Worldview-3 multispectral image with accuracy assessment.	5-31
Table 5.17 : Showing the confusion matrix of the Worldview-3 multispectral image with accuracy assessment.	5-32
Table 5.18: Showing results of object based accuracy measure applied on the Worldview multispectral image (total detection, false negative, positional accuracy, local area difference, overestimated, underestimated total detection error).....	5-34
Table 5.19: Showing the comparison of results of objects based accuracy assessment between both images (total detection, false negative, positional accuracy, local area difference, overestimated, underestimated total detection error).	5-35

1. INTRODUCTION

1.1. Motivation and problem statement

Extraction of individual trees in urban environment and developed cities using satellite imageries has been quite challenging and a subject of active research with many approaches (Ardila et al., 2012). There are so many reasons that can be attributed to this depending on the image scale (coarse or fine), algorithm used and the different types of tree in question (Ardila et al., 2012). However, much more challenging is the mapping of this information in developing countries and rural areas. There are quite a number of reasons for this. In Africa and most developing countries, individual trees are dispersed at random within heterogeneous and indistinct fields (Debats et al., 2016), while in developed countries and cities they are planted in fields which are organised in rows and specific patterns. The second reason is the widespread shortage of up to date agricultural data in these areas (World Bank, 2013). This is partly due to lack of funds and limited government capacity. The third reason is that most fine resolution satellite imageries are restricted and unmanned aerial vehicles (UAV) are prohibited by law in most of these regions. These challenges greatly limit efforts to map or monitor individual trees in this data-sparse regions of the world.

With the launching of Sentinel satellite missions by the European space agency (ESA); medium and coarse resolution imageries have become freely and readily available. Therefore, most developing countries will like to leverage on this opportunity as satellite imageries still remains the alternative option to mapping agricultural information. A general limitation however exists in the spatial distribution of landcover data sets derived from most of these imageries due to varying contrast sensitivity (Frampton et al., 2013). At coarse resolution and depending on the area of study, a row of trees may appear as individual trees, while single trees are much more visible at finer scale. The effects of individual tree spatial patterns (shape, size, location etc.) from coarse remote sensing images has made it very difficult and challenging to detect tree crown objects. This has made some researchers to suggest that coarse resolution images are improper for tree detection (Ardila et al., 2012). Although finer resolution remotely sensed images have improved the visibility and identification of both spectral and spatial information, it also has some limitations. Tolpekin et al., (2010) identified three fundamental factors that prevent the effective identification of tree crowns from very high resolution (VHR) imagery. First, they identified the variation between the Low and High resolution scene models with respect to the size of the tree objects. Secondly, they identified the limitations of classifiers in dealing with spectral separability between tree objects and other similar background classes e.g. shrubs, grasses etc. This is also affected by the sun illumination angle which projects shadows of tall buildings. Finally, some sensors have large dynamic range which captures complex models resulting in incoherent pixel-based classification outputs that doesn't represent the object of interest.

Super resolution mapping (SRM) refers to a technique for classification where hard classification maps are produced which have a resolution finer than that of the original image (Foody & Atkinson, 2002). The concept was introduced on the assumption that pixels are made up of matrices of sub-pixels whose location can be predicted based on the spatial dependency between them (Atkinson, 2004). In this process, a fraction of classes is generated with soft classification method and subsequently a spatial optimization model is used to predict the location of those class fractions producing a fine resolution map from coarse input images (Atkinson, 2009). There are two modes of application of SRM methods; first is on high resolution (HR) and second is on low resolution (LR) images in relation to the size of objects. For the HR scenario (where pixels are smaller than object of interest), the goal is to maximize

the spatial correlation between neighbouring pixels; while for the LR scenario (where pixels are much larger than object of interest), the goal is to match the spatial correlation to some prior model (Atkinson, 2009). When applying SRM for the HR case, the objective is to increase the accuracy of prediction on a sub-pixel by sub-pixel basis which in turn increases the accuracy and spatial resolution of the output map; while for the LR case, the objective is to predict the pattern of the object of interest which are smaller than pixels (Atkinson, 2009).

Several algorithms have been proposed in literature for SRM, adapted to various applications and work well in pixel based image analysis. For example, Jin et al. (2012) proposed geostatistical based method for SRM using local indicator variogram, Atkinson developed spatial pixel swapping method (Atkinson, 2005), (Sriwilai et al., 2013) developed an SRM algorithm based on level set method and (Kasetkasem et al, 2005) proposed Markov random field (MRF) based SRM with simulated annealing. Markov random field (MRF) has developed and become a very useful tool to characterize contextual information using prior and conditional distribution of original image (Geman & Geman, 1984). Ardila et al., (2011) “identified the attractiveness of MRF approach for image analysis as it addresses two main issues; first, it produces classification maps with finer resolution than the input image and secondly, incorporating contextual information using MRF approach produces results which overcome the mixed pixel effect and within-class spectral variation evident with some other techniques.” Markov random field (MRF) based SRM has been applied to remote sensing imageries (Kasetkasem et al., 2005); the goal however is not limited only to obtaining class proportion, but to model image spectral and land cover spatial information’s respectively (Li et al., 2016). Although incorporating MRF with SRM has improved the quality of classification outputs, most features in reality are still not well represented as they still rely on information from pixels alone.

Over the years, the concept of Object based image analysis (OBIA) has evolved and widely suggested as probably one of the ways to handle the limitations and challenges of pixel based image analysis. In this process, remote sensing imageries or thematic outputs are partitioned into regions or objects and subsequently their characteristics are analysed at various scales (Hay & Castilla, 2006). This technique takes cognisance of the temporal, spectral and spatial characteristics of objects during processing while also providing the ability to query and link individual objects (Blaschke, 2010). Where spectral properties are not unique, but shape and neighbourhood relations are distinct, OBIA technique is very useful in capturing reliable information (Blaschke et al., 2014). This is very useful in overcoming ambiguities caused by limited evidence during classification.

In this research, we applied SRM to quantify and compare tree detection using remote sensing imageries at different scales (coarse and high resolution). To achieve this, we applied the pixel based MRF-SRM of (Tolpekin & Stein, 2009), and partitioned the classified SRM thematic results into objects (segments) using region growing segmentation algorithm as post a classification procedure. At this stage and depending on size, some segments (objects) may get bigger, smaller or even disappear. Finally, we analysed the resulting thematic object characteristics using object based image analysis. The limitation of spectral information available in finer resolution images was addressed by the contextual approach of MRF; while the spatial resolution limitation for coarse images was reduced using SRM (Tolpekin et al., 2010). The novelty in this research is the comparison of the detected tree objects at different scale and spatial resolution using our method. The aim is to substantiate the reliability of individual tree crowns detected from coarse resolution images against the backdrop that suggests it’s improper for tree detection. Again, the study will assist efforts to improve the capacity of monitoring agricultural change in data sparse regions and developing countries. The results of our method extended beyond the norm of existing spatial dependence principle evident in pixel based SRM analysis, identified and validated image objects taking cognisance of its range of characteristics.

1.2. **Research identification**

The primary focus in this research can be identified through the following research objectives and research questions stated below;

1.2.1. **Research objectives**

The main objective of this research is to quantify and compare tree detection from satellite images (medium and high resolution) at different scale by applying SRM in HR scenario. Below are the subsequent sub-objectives;

- To perform super resolution mapping in HR scenario for tree detection.
- To analyse the characteristics of the detected tree object using object based image analysis.
- To compare the performance of the applied method at different scale and spatial resolution (Sentinel-2 MSI at 10m and Worldview-3 MSI at 2m).

1.2.2. **Research questions**

The research questions are outlined below;

- How can we apply SRM in HR scenario to detect tree objects at different scale?
- How can one determine the most appropriate parameters for optimizing the SRM model for each scale?
- How can we evaluate and quantify the detected tree objects?
- How should the assessment of the results be performed?

1.3. **Innovation**

The novelty of this research is;

- Quantifying and comparing the detected tree objects at different scale and spatial resolution by applying SRM in HR scenario.

2. LITERATURE REVIEW

2.1. Object-based image analysis

Over the years, the concept of Object based image analysis (OBIA) has evolved and widely suggested as probably one of the ways to handle the limitations and challenges of pixel based image analysis. The concept was introduced mainly as a result of the increasing commercially available high resolution imageries, ever-increasing sophisticated user needs which has to be met and the existing limitations with pixel-based image analysis (Hay & Castilla, 2006). OBIA builds on older techniques of remote sensing image analysis such as segmentation, feature extraction and edge detection. In this process, remote sensing imageries or thematic outputs are partitioned into regions or objects and subsequently their characteristics are analysed at various scales (Hay & Castilla, 2006). This technique takes cognisance of the temporal, spectral and spatial characteristics of objects during processing while also providing the ability to query and link individual objects (Blaschke, 2010). When applying object based image analysis, the image-object is the central methodological element and the object of investigation, therefore image context is well documented for object recognition (Baglioni et al., 2009). This is very useful in overcoming ambiguities caused by limited evidence during classification. Where spectral properties are not unique, but shape and neighbourhood relations are distinct, OBIA technique is very useful in capturing reliable information (Blaschke et al., 2014). The major aim is to incorporate the concept of human visual interpretation of objects during or after classification to obtain more accurate representation that will result in increased repeatability, while reducing labour subjectivity and time cost (Hay & Castilla, 2006).

2.2. Image segmentation

Segmentation refers to the partitioning or grouping of pixels into connected regions based on homogeneity criteria (Hay et al., 2005). A good segmentation is identified by uniform and homogenous or adjacent region which is clearly separated by a simple boundary that has a contrast with respect to their characteristics (Fleet & Jepson, 2007). Image segmentation plays an important role in OBIA where features are extracted with respect to their various characteristics to distinguish one region from another. "This is because, segmentation is the customary technique used to derive initial units within OBIA, at least from the point of view of the segmentation algorithm (Baglioni et al., 2009)." As important as segmentation is to OBIA, it cannot be completely relied on alone because of its major limitation of over segmentation and under segmentation. Furthermore, the traditional image segmentation algorithm works on the rigid grid of pixels which produces homogenous segmented regions whose boundaries are unrealistic as it appeals to the structure of the square grid cells and mostly does not conform to real object boundaries (Baglioni et al., 2009). This has prompted so many researchers to seek for more realistic and probabilistic techniques to improve segmentation such as super pixel methods (Achanta et al., 2012).

Within the last few decades, there has been an increasing adoption of super pixel based methods in image processing (Ladický et al., 2014). These tend to perform *a priori* segmentation of the images before subsequently been applied to various image analyses such as classification. Super pixels algorithms group pixels into meaningful regions which can be used to replace the rigid structure of pixel grid (Achanta et al., 2012). They rely on an initial partitioning of an image capturing redundancy and provide a primitive to compute various image features taking cognisance of spatial location and colour/texture of distribution (Sriwilai et al., 2013). Various super pixel algorithms applied on an image produces different

results. This depends on the category of algorithm method used (Graph-based or Gradient-based) and the purpose for which the algorithm is developed (adherence to image boundary, speed and impact on segmentation performance) (Achanta et al., 2012). When choosing super pixel algorithm it is desired that it should at least adhere to image boundary and reduce computational complexity.

2.3. Maximum likelihood classification

Maximum likelihood is a classification procedure which has been in existence for decades now. It is based on the assumption that using the probability distribution in spectral space, one can be able to describe the distribution of pixels in a given class or group of classes (Richards & Jia, 2006). To achieve this, the classifier assigns pixels to the class with highest probability of membership, while assuming the multidimensional normal distribution for all classes (Strahler, 1980). Here the normal or Gaussian distribution is mostly used because its properties are well known for any dimensionality and its parameters are easily estimated (Maselli et al., 1994). To model the normal distribution for each class's dispersion of pixels, one has to estimate the mean and covariance matrix. This is because the normal distribution is specified by its mean vector and covariance matrix. It is expected that pixels in spectrally distinct cluster or class should lie around the centre and less likelihood to be far away where pixels do not cluster (Richards & Jia, 2006). Therefore, if the mean and covariance matrix is known or estimated correctly for all classes, then a set of probabilities can be computed which shows the likelihood of a pixel or pattern in space belonging to a particular class (Richards & Jia, 2006).

2.4. Super resolution mapping

Super resolution mapping (SRM) refers to a technique for classification where hard classification maps are produced which have a resolution finer than that of the original image (Foody & Atkinson, 2002). The concept was introduced on the assumption that pixels are made up of matrices of sub-pixels whose location can be predicted based on the spatial dependency between them (Atkinson, 2004). In this process, a fraction of classes is generated with soft classification method and subsequently a spatial optimization model is used to produce a fine resolution map from coarse input images (Atkinson, 2009).

There are two modes of application of SRM methods; first is on high resolution (HR) and second is on low resolution (LR) images in relation to the size of objects. For the HR scenario (where pixels are smaller than object of interest), the goal is to maximize the spatial correlation between neighbouring pixels; while for the LR scenario (where pixels are much larger than object of interest), the goal is to match the spatial correlation to some prior model (Atkinson, 2009). When applying SRM for the HR case, the objective is to increase the accuracy of prediction on a sub-pixel by sub-pixel basis which in turn increases the accuracy and spatial resolution of the output map; while for the LR case, the objective is to predict the pattern of the object of interest which are smaller than pixels (Atkinson, 2009). Along and within the extremes of each scenario, lie a particular SRM goal and its application.

Several algorithms have been proposed in literature for SRM, adapted to various applications and work well in pixel based image analysis. For example, Jin et al. (2012) proposed geostatistical based method for SRM using local indicator variogram, Atkinson developed spatial pixel swapping method (Atkinson, 2005) and (Sriwilai et al., 2013) developed an SRM algorithm based on level set method. Previous researches has demonstrated that using different SRM methods may produce varied land cover representations from the same coarse input image and therefore the identification of an optimal SRM method in advance is difficult (Li et al., 2016). This is because each SRM algorithm has its own strengths and weaknesses depending on the technique used to develop it. Some has suggested the combination and of use multiple SRM algorithms to utilize information from each method while also addressing the drawbacks in using an individual method.

Markov random field (MRF) based SRM has been applied to remote sensing imageries (Kasetkasem et al., 2005); the goal however is not limited only to obtaining class proportion, but to model image spectral and land cover spatial information's respectively (Li et al., 2016). Also, MRF based SRM has been applied to quantify the accuracy of land cover class separability and the accuracy of such methods is highly influenced by the scale factor used and class separability in the image (Tolpekin & Stein, 2009). An advantage of MRF based SRM is that it does not strictly rely on soft classification results and can allow one represent spectral variation of class separability values. However, it has also been identified to have some spectral and land cover spatial constraints. The spectral constrain is the assumption that coarse pixel has a spectral response generated from combined spectra of classes in sub-pixel map, while the land cover spatial constrain assumes that neighbouring sub-pixels are more likely to affect a certain land cover class than a different class which is farther apart (Tolpekin & Stein, 2009). Incorporating MRF with SRM has improved classification output as well as overcome the problem of mixed pixel which is a limitation of some other techniques.

2.5. Detection of trees from remote sensing imagery

Remotely sensed imageries and remote sensing technology afford us the possibility of monitoring our planet at various scales. Various researches have been conducted using different methods and techniques developed to effectively extract useful and reliable information from these imageries which are increasingly available on daily basis. Extraction of individual trees from remotely sensed imageries has been quite a challenging subject of research for many decades now. There are so many reasons that can be attributed to this depending on the image scale (coarse or fine), algorithm used and the different types of tree in question (Ardila et al., 2012). At coarse resolution and depending on the area of study, a row of trees may appear as individual trees, while single trees are much more visible at finer scale. The effects of individual tree spatial patterns (shape, size, location etc.) from coarse remote sensing images has made it difficult and very challenging to detect tree crown objects. Based on these, some researchers suggests that it is improper to use coarse resolution images to identify tree crown boundaries (Ardila et al., 2012).

Although finer resolution (VHR) remotely sensed images have improved the visibility and identification of both spectral and spatial information, it also has some limitations. Tolpekin et al., 2010 identified three fundamental factors that prevent the effective identification of tree crowns from VHR imagery. First, they identified the variation between the Low and High resolution scene models with respect to the size of the tree objects. Secondly, they identified the limitations of classifiers in dealing with spectral separability between tree objects and other similar background classes e.g. shrubs, grasses etc. This is also affected by the sun illumination angle which projects shadows of tall buildings. Finally, some sensors have large dynamic range which captures complex models resulting in incoherent pixel-based classification outputs that doesn't represent the object of interest. To address some of these issues above, Tolpekin et al., (2010) extended the contextual MRF based SRM method for multispectral image to incorporate a panchromatic band with higher spatial resolution. They applied these to identify and extract tree from VHR imageries in urban areas. Ardila et al., (2012) employed context-sensitive extraction of tree crown objects from VHR images in urban scenes. Also, Wolf & Heipke, (2007) developed an approach for automatic tree detection from remote sensing data using scale-invariance and a digital surface model (DSM). With continuous technological advancements, different datasets have become readily available ranging from active sensor data, UAV data and point cloud datasets. The application and use of these datasets require a different approach, technique and methodology for identification and detectability of various objects including trees. They are also not without their various limitations. Whichever methods is used, detection of tree object from remote sensing imagery or other datasets still remain an interesting sphere of research in the field of remote sensing as improvement is still needed to overcome the challenges with existing techniques.

3. STUDY AREA AND DATA

This chapter provides information about the study area, available data and proposed software that will be used in this research. Below are the details;

3.1. Study area

The study area for this research is a part of Sukumba village located at latitude $12^{\circ}12'55.45''$ and longitude $-05^{\circ}13'04.82''$ in Koutiala district, Mali. It is a rectangular area with dimension of 1312 m x 777 m. The settlement is characterized by small holder farms and heterogeneous fields, and in most places trees partially cover the area. Kofa Village is built on a rocky terrain, with a number of large valleys. The general planting season is around May to Sep/Nov and most cropping system is mono-cropping. The trees here are evergreen, most of them being Shea trees which grow up to 35–40 m tall; with a crown radius between 8-10 m. Mango trees are also found in most places within the fields. These Shea trees were of interest in this research because of their economic importance. Their fruits are principal exported variety and mostly serve as food and source of revenue for occupants in the area (Vayssières et al., 2007).

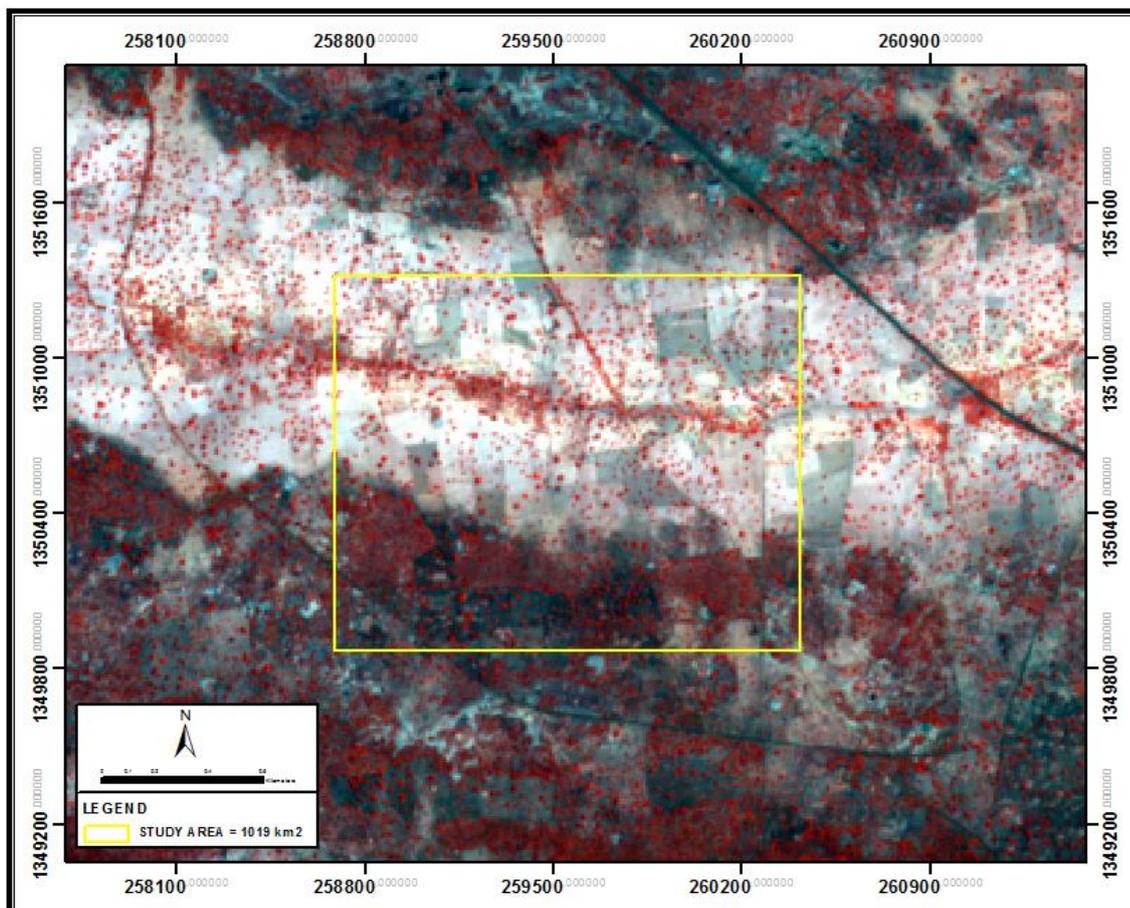


Figure 3.1: Overview of study area (yellow boundary) using Sentinel-2 multispectral image displayed in NIR (band8), Green (band3) and Blue (band2).

3.2. Available data

The available data for this research project is outline below;

- 10m Sentinel-2 multispectral image of the study area captured on the 8th of May 2016. This was used for generating the training and test sets.
- 2m Worldview-3 multispectral image acquired between 8th of May 2015 to 21st November 2015. This was used for generating the training and test sets.
- 0.5m Worldview-3 panchromatic image acquired between 8th of May 2015 to 21st November 2015. This was used for generating the reference data for validation.

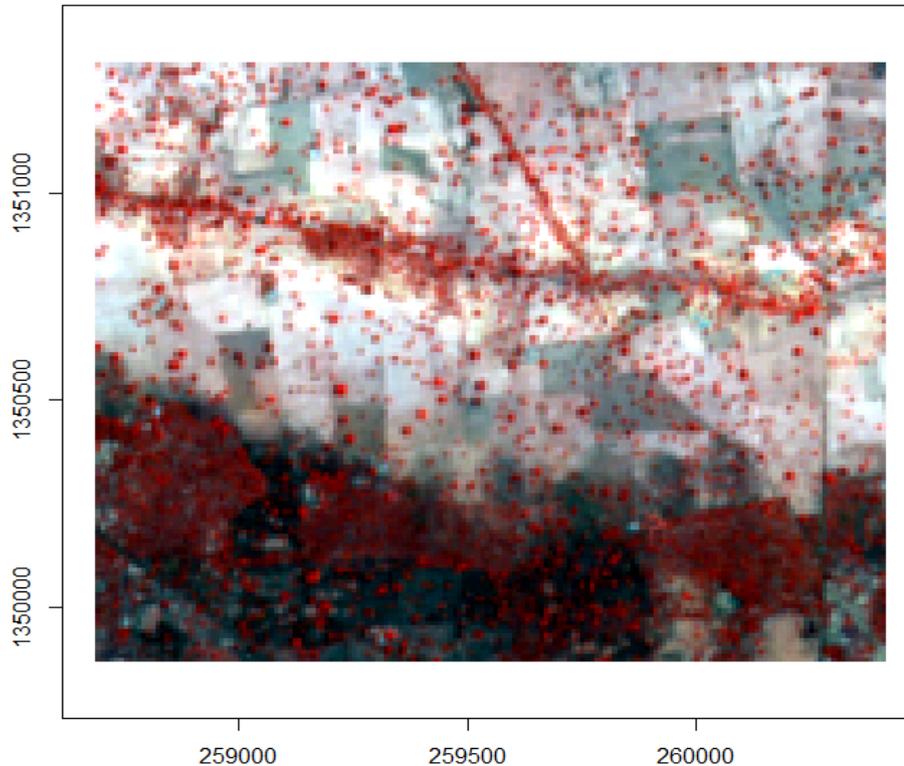


Figure 3.2: Sentinel - 2 multispectral image subset of the study area displayed in NIR (band8), Green (band3) and Blue (band2).

The Sentinel-2 multispectral image subset used in this research has 10m spatial resolution, with 4 bands (near infra-red, red, green and blue) of false colour composite. The Worldview-3 image on the other hand has 2m spatial resolution, with 8 bands (coastal, blue, green, yellow, red, red-edge, near infra-red_1 and near infra-red_2). These images capture both man-made and natural features which contain mixed pixels making SRM an ideal technique for application and interesting for my study. Here, the tree object is the main subject of investigation and as such, other information in the image is of less importance. Therefore, what is not a tree object is heterogeneous and as such will not be considered in the course of the research. The Sentinel-2 multispectral image contains 1.953% cloud cover and is assumed to have been geometrically corrected of all possible distortions during acquisition and therefore suitable for classification.

3.3. Reference data

Reference data is required in this research to implement accuracy assessment and validate the final classification result. The accuracy assessment is a statistical measure of how well the applied method performs compared to the reference data. The reference data is a vector layer of tree objects derived from digitizing the tree crowns objects from 0.5 m panchromatic Worldview-3 image of the study area. This reference vector was derived only on standalone trees within the fields. All other trees that are found around similar background classes like shrubs and rocks were ignored.

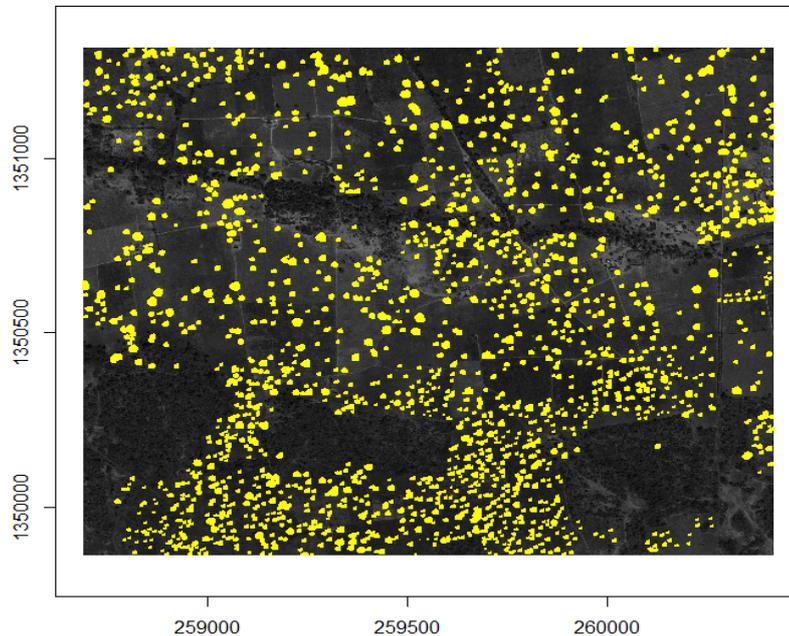


Figure 3.3: Reference vectors of tree crowns (yellow) extracted from Worldview-3 panchromatic image of the study area.

3.4. Software

In this research, different software and algorithm was used for applying the proposed methodology and data visualization which are;

- ArcGIS 10.4.1
- Erdas Imagine 2016.
- R version 3.2.2.
- MRF-SRM algorithm which was optimized using C++ library.
- Region-growing segmentation algorithm.

Both ArcGIS 10.4.1 and ERDAS Imagine 2016 are well known commercial software for GIS and remote sensing. R software on the other hand is open source software that is able to compute and visualize statistical data developed by R Core Team (2016). It has some packages which were used in this research which are outline below;

- rgl developed by Adler et al., (2016)
- rgdal developed by (Bivand et al., 2016)
- e1071 developed by (Meyer et al., 2015)
- kernlab developed by Karatzoglou et al., 2010)
- MASS developed by (Venables and Ripley, 2003)
- Rcpp developed by (Eddelbuettel, 2013)

4. METHODOLOGY

4.1. Spectral class definition

When performing supervised classification, user defined spectral classes are required to estimate statistics and extract reliable information from the remotely sensed image. These classes are defined based on a visual interpretation of pixel spectral reflectance that is identified which corresponds to information in reality. Pixel values were modelled with normal distribution and labelled to represent the various classes. Optimal estimation for class mean and covariance matrix requires sufficient number of training samples across the image with respect to the number of bands in the image. Visual interpretation is one of the key ways of identifying and defining land cover classes in remote sensing image analysis. This technique was employed in this research to identify various classes from our image subset. To prevent overfitting, training sets were generated outside the image subset of the study area, while the test sets were generated within the image subset of the study area. The training sets were used to estimate the class mean and covariance matrix values from the image, while test sets were used for statistical analysis of the classification accuracy.

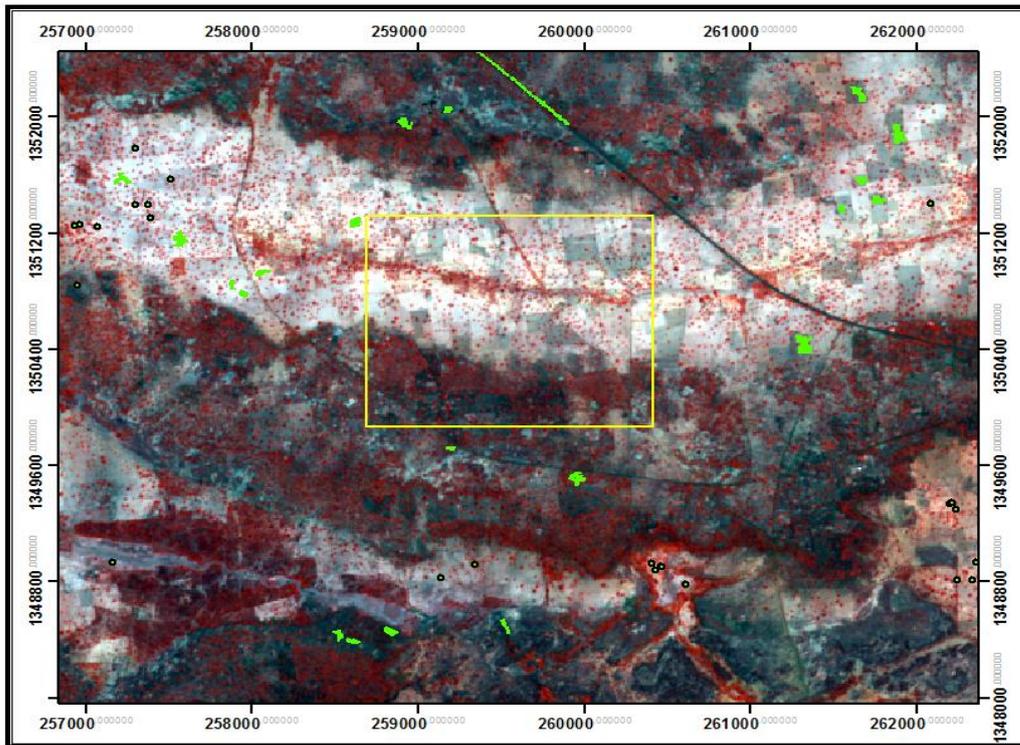


Figure 4.1: Example of Sentinel-2 multispectral image showing the study area (yellow rectangle) and training sets (green polygons and points).

The land cover classes defined with respect to the spectral values of pixels in the images were;

- Sentinel-2 multispectral image – The landcover classes are *trees*, *rock*, *barsoil_1*, *barsoil_2* and *shrub*.
- Worlview-3 multispectral image – The landcover classes are *trees*, *road*, *built_up*, *field_1*, *field_2*, *field_3*, *barsoil*, and *shrub*.

All these classes were defined because of the spectral classifier we are using which requires sufficient training samples to estimate the mean and covariance matrix of the distribution. If we have only one class, it is not possible to compute the mean, covariance matrix or invert of the matrix. This will make it impossible to apply our classification or to assign class labels to pixels in the image. For Sentinel-2 multispectral image, the total number of pixels used for training is 1011 pixels and 1010 for test. Also, for the Worlview-3 multispectral image, the total number of pixels used for training is 4271 and 1739 for test. Both training and test samples were sufficient enough for estimation of class mean and covariance matrix and provided a homogenous representation for the entire image subsets.

To evaluate the data value of one band plotted against another, feature space was used. It is basically represented or displayed as a 2D or 3D raster image of scatterplot with a dot for every pixel in an entire image or subset. For the 2D feature space, the pixel position represented shows the spectral values of the two chosen bands. The feature space image has a colour associated with each pixel dot plotted. The colours of the feature space represents the cumulative frequency (i.e. the number of pixels in the original image which have the given 2D or 3D combination). The 3D feature space is a graphical representation of three bands of image spectral values. It is dynamic and each class is represented with a coloured ellipse or clustered points. When analysing the feature space plot, it is expected that each ellipse is distinct from one another to achieve high classification accuracy. If the classes overlap, there is spectral confusion and the accuracy of the classification will subsequently be low. However, overlapping ellipse may not necessarily mean that training samples are wrong, sometimes it is natural and expected because some landcover spectral classes are highly correlated and overlap in reality.

4.2. **Markov random field**

When interpreting remote sensing image using visual interpretation, context is very important. In this concept, pixels are treated in relation with their neighbours and not in isolation. Markov random field (MRF) is a useful tool to characterize contextual information using prior and conditional distribution of original image (Geman & Geman, 1984). It allows for elimination of ambiguities caused by limited evidence, eliminates possible errors and generates smooth image classification output. Let a set of random variable $t = \{t_1, t_2 \dots t_m\}$ be defined on the set S containing m number of sites in which each random variable $t_r (1 \leq i \leq m)$ takes a label from a set of labels L . The set S is equivalent to an image containing m pixels; t is a set of pixel DN values, and the label set L depends upon the application. A random field that takes cognizance of its neighbours is a Markov random field if its probability density function satisfies the following properties;

- 1) Positivity: $P(t) > 0$ for all possible configurations of t ,
- 2) Markovianity: $P(t_r | t_{S-r}) = P(t_r | t_{N_r})$ and
- 3) Homogeneity: $P(t_r | t_{N_r})$ is the same for all sites r .

From the equations above, $S - r$ is the set difference (i.e., all pixels in the set S excluding r), t_{S-r} denotes the set of labels at the sites in $S - r$, and N_r denotes the neighbors of site r . Positivity property of MRF is usually satisfied and sustains the local conditional properties which is determined by the joint probability $P(t)$. Markovianity deals with the dependency of pixels on its neighbourhood with respect to the whole image. Homogeneity specifies the conditional probability for the label of a site r , given the labels of the neighboring pixels, regardless of the relative position of site r in S . In image analysis, neighbourhood system plays a vital role. The first order neighbourhood has four pixels which shares sides with the given pixel. The second order neighbourhood shows four pixels having corner boundaries with the pixel of interest. Subsequently, higher order neighbourhood can be formulated by extending the second order neighbours in the same manner. A clique is a subset of mutual neighbours for all pair of sites. The sites can be single, double, triples and so forth. As the number of cliques increases as a result of an increase in the neighbourhood system, the computational complexity increases also. The first order neighbourhood system was applied for our study.

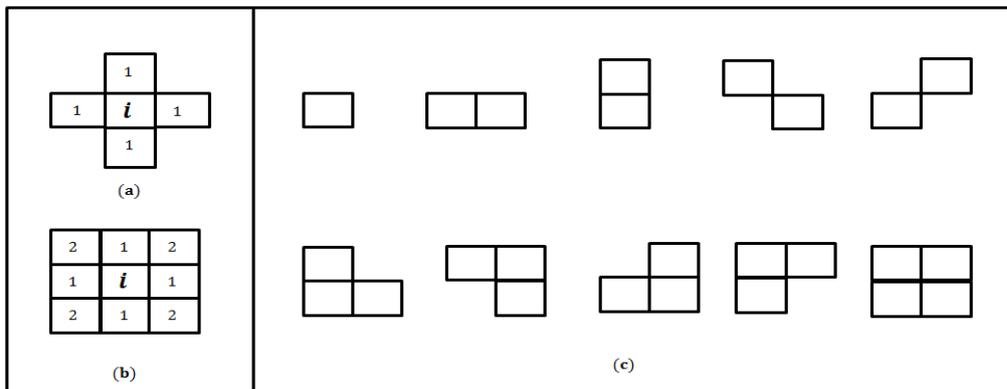


Figure 4.2: Showing (a) first order neighbourhood system; (b) second order neighbourhood system; (c) cliques corresponding to neighbourhood system.

4.3. Pixel based MRF-SRM

Pixel based MRF-SRM model adopted by (Tolpekin & Stein, 2009) was selected. Here, SRM image classification of a multispectral image y that consists of K spectral bands and R spatial resolution which corresponds to a square area of R^2 on the ground and a scale factor S . We denote the pixel location as $b_i \in B$, where B is the $M_1 \times M_2$ pixel matrix. The output of SRM is a fine resolution classified map c and it has a spatial resolution $r < R$, where R is for coarse resolution and r for fine resolution image, respectively. It is assumed that the scale factor $S=R/r$ is an integer. The area of ground S^2 covered by pixel b_i has a finer resolution a_{ij} and the pixel matrix $A = \{a_{ij}\}$ has a size of $(SM_1 \times SM_2)$. The (SR map) c is defined on a pixel location A which has the same area as the multispectral image y . For our study, a scale factor of $S = 5$ was used for Sentinel-2 multispectral image to obtain an (SR map) c and a scale factor of $S = 1$ for Worldview-3 multispectral image.

A multispectral image x is assumed which has the same spectral bands k as image y and acquired at the same spatial resolution c . Image y which is directly observed by satellite is assumed to be a degraded observation of image x which is not observed by satellite. Again, we further assume that each pixel in image x can be assigned to a unique class $C(a_{ij}) = \alpha \in \{1,2 \dots L\}$, $a_{ij} \in A$. A degradation model (b_i) is used to establish a relationship between y at pixel b_i and image x ;

$$y(b_i) = \frac{1}{S^2} \sum_{j=1}^{S^2} x(a_{ij}) \quad (4.4)$$

A symmetric neighbourhood $N(a_{ij})$ on A is defined for each pixel a_{ij} . This is the set of all pixels inside a square window with centre on the pixel a_{ij} , excluding pixels for itself. Therefore, a classified image c is modelled as an MRF with neighbourhood system $N(a_{ij})$, where the prior probability for the (SR map) c is $P(c)$, the conditional probability that image y is observed, given the c is $P(y|c)$, and the posterior probability for the (SR map) c is $P(c|y)$, given that image y is observed;

$$P(c) = \frac{1}{Z_o} \exp \left[-\frac{U(c)}{T} \right] \quad (4.5)$$

$$P(y|c) = \frac{1}{Z_p} \exp \left[-\frac{U(y|c)}{T} \right] \quad (4.6)$$

$$P(c|y) = \frac{1}{Z_q} \exp \left[-\frac{U(c|y)}{T} \right] \quad (4.7)$$

From equations (4.5, 4.6 and 4.7);

$U(c)$ = Prior energy function of (SR map) c

$U(y|c)$ = Likelihood energy function to observe image y given the true (SR map) c

$U(c|y)$ = Posterior energy function of (SR map) c given the observed image y

$Z_o, Z_p,$ and Z_q are normalizing constraints, while T is the constant temperature

Based on Bayes theorem:

$$P(c|y) \propto P(c)P(y|c) \quad (4.8)$$

Therefore, the energy function is expressed as:

$$U(c|y) = U(c) + U(y|c) \quad (4.9)$$

4.3.1. Prior energy function

When we consider the pair of sites interactions, the prior energy can be modeled as follows;

$$\begin{aligned}
 U(c) &= \sum_{i,j} U(c(a_{j|i})) \\
 &= \sum_{i,j} \sum_{l \in N(a_{j|i})} w(a_l) \delta(c(a_{j|i}), c(a_l)).
 \end{aligned} \tag{4.10}$$

From equation (4.10);

$U(c(a_{j|i}))$ = the local contribution of the prior energy from pixel $c(a_{j|i})$.

$w(a_l)$ = the weight of contribution from neighbouring pixels $al \in N(a_{j|i})$.

$N(a_{j|i})$ = the neighbourhood system of pixels $a_{j|i}$.

$\delta(c_1, c_2)$; takes the value 0 if $c_1 = c_2$ and 1 if otherwise.

The weight of the contribution to the equation (4.10) above can be modeled as;

$$w(a_l) = q \cdot \phi(a_l)$$

The overall magnitude of all weights and the prior energy term is controlled by $\sum_{l \in N(a_{j|i})} \phi(a_l) = 1$ and $0 \leq q < \infty$. For $\phi(a_l)$ an isotropic expression is employed and which depends on the distance $d(a_{j|i}, a_l)$ between pixels $a_{j|i}$ and a_l . Again, as the value of q increases, a smoother result is achieved which reduces the variation between different classes in the neighbourhood system.

4.3.2. Likelihood energy function

The configuration of the likelihood model is dependent on the pixel composition. For our study, we treat image y as consisting of mixed pixels. Therefore, a single pixel b_i which has a value $y(b_i)$ is assumed to be normally distributed with mean vector μ_i and covariance matrix C_i . Both are defined below;

$$\mu_i = \sum_{\alpha=1}^L \theta_{\alpha i} \mu_{\alpha} \tag{4.11}$$

$$C_i = \frac{1}{S^2} \sum_{\alpha=1}^L \theta_{\alpha i} C_{\alpha} \tag{4.12}$$

From equations (4.11 and 4.12), $\theta_{\alpha i}$ is the proportion of the class α in the pixel b_i . It is assumed that the spectral value $x(a_{j|i})$ of the S^2 fine resolution pixels $a_{j|i}$ are independent and identically distributed based on Gaussian normal distribution. This means that spectral values of neighbouring pixels which are of the same class are not spatially correlated. But the model of our study as seen in equation (4.10) which describes the prior energy allows fine resolution pixels to be spatially correlated and this is the case in real images where spectral values of pixels are dependent on each other. The likelihood probability of $y(b_i)$, given $c(a_{j|i})$ is expressed as;

$$\begin{aligned}
 P(y|c) &= \prod_{i,j} P(y(b_i) | c(a_{i|j})) \\
 &= \prod_{i,j} \frac{1}{(2\pi)^{K/2} |C_i|^{1/2}} \times \exp\left(-\frac{1}{2}(y(b_i) - \mu_i)' C_i^{-1} (y(b_i) - \mu_i)\right)
 \end{aligned} \tag{4.13}$$

The corresponding likelihood energy function is expressed as;

$$\begin{aligned}
 U(y|c) &= \sum_{i,j} U(y(b_i)|c(a_{j|i})) \\
 &= \sum_{i,j} \left[\frac{1}{2} (y(b_i) - \mu_i)' C_i^{-1} (y(b_i) - \mu_i) + \frac{1}{2} \ln |C_i| \right]
 \end{aligned} \tag{4.14}$$

where; $U(y(b_i)|c(a_{j|i}))$ is the local contribution to the likelihood energy from pixel $c(a_{j|i})$.

4.3.3. Posterior energy function

Most labelling problems in computer vision are formulated as a pairwise random field (a probabilistic model which combines the unary and pairwise potential). For MRF pixel class labelling, it's a sum or combination of the prior and likelihood models. $P(c|y)$ is the posterior probability distribution of MRF pixel class labelling whose energy function is expressed as;

$$U(c|y) = q \cdot \sum_{i,j} \sum_{l \in N(a_{j|i})} \phi(a_l) \cdot \delta(c(a_{j|i}), c(a_l)) + U(y|c) \tag{4.15}$$

From equation (4.15), the maximum *a posteriori* (MAP) probability solution for (SR map) c is that which minimizes the posterior energy function $U(c|y)$. Dividing equation (4.16) by a factor of $1 + q$ and defining the smoothness parameter $\lambda = q/(1 + q)$, $0 \leq \lambda < 1$, the posterior energy is expressed further as;

$$U(c|y) \propto \lambda \sum_{i,j} \sum_{l \in N(a_{j|i})} \phi(a_l) \cdot \delta(c(a_{j|i}), c(a_l)) + (1 - \lambda) U(y|c) \tag{4.16}$$

The smoothness parameter λ in equation (4.16) controls the contribution from the prior and likelihood models. Assigning a value $\lambda = 1$, neglects the likelihood term in the equation and the resulting minimal posterior energy is obtained where all pixels is assigned to one class. For our study, the (SR map) c that corresponds to the MAP of the posterior energy is our interest; therefore the expression on the right hand side of equation (4.17) is optimized.

4.4. Parameter estimation

4.4.1. Smoothness parameter

The maximum *a posteriori* (MAP) probability solution for the SRM problem is achieved by minimizing the posterior energy of (Fig 4.16). The smoothness parameter lambda (λ) in this equation is what controls the contribution from the prior and conditional distribution. This parameter which is scale dependent, ranges from 0 – 1 and needs to be estimated before the method can be applied. If we assign a value $\lambda = 1$, the likelihood term in the equation is neglected and the resulting minimal posterior energy is obtained where all pixels is assigned to one class (Tolpekin & Stein, 2009). Optimal selection of λ for a given image subset used for MRF-based analysis is achieved by several experiments (trial and error); this is computationally expensive and time consuming. Both visual inspection of results as well as the value of kappa k was used to make the choice for optimal λ in this research.

4.4.2. Simulated annealing

In remote sensing image analysis, optimization by simulated annealing (SA) is a stochastic relaxation process that generates a sequence of images which in essence converges to the MAP estimate (Geman & Geman, 1984). Because of the large configuration space of image labeling problems, achieving a global energy minimization is not easily obtainable (Tolpekin & Stein, 2009). However, in simulated annealing, Geman & Geman, (1984) identified a solution to the MAP estimate of optimal image labeling problem by employing the Metropolis sampler algorithm to sequentially decrease the temperature of the solution. Here, the annealing schedule is the sequence and time spent as temperature decreases until it reaches equilibrium. The idea or concept originates from the cooling, solidification and crystallization of hot metals and some liquids. A random starting point is required when implementing the annealing schedule; however this slows down the process (Tolpekin & Stein, 2009). Kirkpatrick et al., (1983) implemented a faster annealing schedule after studying its computational efficiency and introduced a solution showing that a good estimate for the starting point will improve the quality of the solution. Two parameters initial temperature (T_0) and updating parameter (T_{upd}) control the energy optimization in simulated annealing process. The cooling schedule is T_v where the temperature at the iteration v is changed according to $T_v = T_{v-1} \times T_{upd}$. Updating parameter $T_{upd} \in (0; 1)$ controls the rate of temperature decrease and the annealing process stops after all pixels have been updated. The value of λ also needs to be carefully defined as this affects the final classification outcome.

4.5. Object based image analysis of SRM result

In pixel based image analysis, MRF has been successfully applied to model contextual information using prior and conditional distribution of original image. Therefore, pixels of the same class are grouped together and directly assigned specific information classes at this level. However, in object based image analysis, the image-object is the central methodological element and the object of investigation. Image segmentation plays an important role in OBIA where features are extracted with respect to their various characteristics to distinguish one region from another. In most OBIA procedures, remote sensing imageries or thematic outputs are partitioned into regions (objects) and subsequently their characteristics are analysed at various scales (Hay & Castilla, 2006). Spectral classification applied on pixel basis using remote sensing images tends to ignore useful spatial information between pixels (Lucieer & Stein, 2002). This becomes a cause for concern in sub-pixel level classification like SRM whose output is also probabilistic. Segmentation however addresses this concern as it helps to extract spatial objects from image scene. Therefore, the classified SRM thematic results, was partitioned into raster objects (segments) using region growing segmentation algorithm as post a classification procedure and the resulting thematic object characteristics are analysed. At this stage and depending on size, some segments (objects) may get bigger, smaller or even disappear with respect to the neighbouring classified landcover information.

4.5.1. Segmentation of (SR map) c

Segmentation refers to the partitioning or grouping of pixels into connected regions based on homogeneity criteria. Using region growing segmentation algorithm, we partitioned the classified SRM thematic results into objects (segments). The segmentation technique chosen for any analysis depends on the type of data and the application of interest. We chose region growing because the location of the initial seed is automatic and free of any parameter tuning or optimization; this does not affect the outcome of the segmentation result. Region growing segmentation is a simple and effective segmentation process that partitions images into regions based on attributes such as area, shape, statistical parameters and texture (Bins et al., 1996). In this process, a list of regions is first created which is composed of only one pixel (seed). By adding pixels to the seed which is the starting point, each region is compared with its neighbour to determine if they are similar or not, joins and label similar pixels into one region and leave out dissimilar ones. As regions are merged to each other, they are removed from the list. This process iteratively continues from one region to another until there is no joinable region.

When segmentation is applied to a classified image, the initial class labels are changed and a homogenous region which represents objects of interest is formed. Local changes may change the size of segments and isolated pixels might disappear. The neighbourhood system determines the neighbouring pixels that will be taken into consideration and also the computational time required for the iteration to complete. The size of the image subset and spatial resolution of the image that is analysed also affects the computational time (Fan et al., 2005). Again, the size of the object of interest is also considered and affects the neighbourhood system chosen. For this research, since the tree object is of interest with an approximate crown radius of 8-10 m, a first order neighbourhood system (Fig 4.2a) was chosen. This is good because the tree objects are isolated randomly within the fields and a first order neighbourhood system considers only a small area surrounding the pixel of interest.

4.6. Accuracy assessment

Accuracy assessment in image analysis (classification) is a measure of reliability of the results and which indicates if the objectives of the analysis have been achieved (Richards & Jia, 2006). We performed both pixel based and object based accuracy assessment in this research

4.6.1. Pixel based accuracy assessment

There are various methods of quantifying pixel based accuracy of thematic maps based on various statistics but the most widely used is the *error matrix* (confusion or contingency matrix) (Jong & Meer, 2006). The aim of the error matrix is to estimate the number of correctly mapped pixels in an image. This is achieved by comparing the classified thematic map against the reference (ground truth or verification). The authenticity of the reference information used to perform the accuracy assessment determines the reliability of the classified map. When applying the error matrix, the reference data classes are listed by column, while the classes indicating the thematic map are listed by row. The diagonal of the matrix indicates agreement (i.e. the number of correct class label between the two data sets). For all classes, the ratio of the sum of the correctly classified pixels to the sum of the total pixels tested is determines the *overall accuracy* (Richards & Jia, 2006). A large number within the diagonal of the matrix shows good class agreement between both data and leads to a high accuracy; however, the reverse is the case with poor class agreement. Again, using the error matrix, one can also compute the *kappa k*, *error of commission* and *error of omission* respectively. Kappa statistics k is a measure of the quality of the thematic output (Congalton, 1991). The error of commission refers to pixels that belong to the reference classes which the classifier placed in their respective class of interest, while the error of omission corresponds to those pixels belonging to the reference classes which the classifier failed to correctly recognize (Richards & Jia, 2006). By dividing the total counts in the cell by the column sum, all the errors were converted to percentage. The error matrix was the *pixel based accuracy assessment measure* employed in this research. For both images, the tree crown classes as well as other classes identified were assessed against the reference classes generated using sampled test sets. Their results were compared against each other respectively to ascertain the performance of the applied method in different scale.

4.6.2. Object based accuracy assessment

In addition to the error matrix, another accuracy assessment measure performed in this research was on *object level validation of the tree crowns* because the tree crowns were of interest in our study. To achieve this, first we converted the segmented thematic (SR map) c into polygon shapefiles (vector). Then we isolated the tree shapefiles from other background classes. With both detected tree object and reference data (Fig 3.3) available as vector, we establish a one-to-one comparison between the identified tree crowns and reference tree objects respectively using three accuracy measures; (a) existential (b) extensional (c) positional accuracy measures (Mahour et al., 2016). All three accuracy measures are essential to evaluate the quality of our detection.

- a) **Existential accuracy:** The existential accuracy helps to ascertain the presence of the object (if they exist). Two accuracy indicators were estimated that concerns the existence of tree objects;
- i. Total detection or True positives (TP) – This is where detected trees exist in the polygon reference layer.
 - ii. False negatives (FN) – This is where there is failure to detect tree objects in the reference layer.

Note: The false positive (this concerns improper detection trees which does not exist in the polygon reference layer) was ignored because our interest was on individual tree objects. We therefore generated our reference from standalone trees and left out trees that are found around similar background classes like shrubs and rocks (Fig 3.3). Any attempt to compute the false positives will give a bias result.

- b) **Extensional accuracy:** The extensional accuracy determines the spatial extent of the existing tree object. We adopted a method developed by Clinton et al., (2010) for assessing the goodness of spatial objects segmented from satellite sensor imagery. To achieve this, first we used a spatial join overlay tool in ArcGIS to intersect both detected tree object and reference data (Fig 3.3). Secondly we computed the area of both detected tree object and reference tree crown. Finally, we exported them in excel and calculated the overestimation, underestimation and total detection errors using equations (4.18, 4.19 and 4.20) respectively. The quality of detection depends on the extent of agreement between the detected tree object and reference data.

For a given (SR map) c , the local area tree object difference (A_i);

$$\Delta A_i = A_{Ri} - A_{Di} \quad (4.17)$$

where; i = measure for a single tree object.

A_{Di} = area of the identified tree object.

A_{Ri} = area of the reference tree objects.

ΔA_i = difference in the local area tree object.

Below is an error matrix indicator which was proposed by (Clinton et al., 2010) and used in this study to evaluate the extential or spatial extent accuracy of identified tree crown objects;

$$\text{Overestimated } (A_{Di}) = 1 - \frac{\sum \text{area}(A_{Di} \cap A_{Ri})}{\sum \text{area}(A_{Di})} \quad (4.18)$$

$$\text{Underestimated } (A_{Di}) = 1 - \frac{\sum \text{area}(A_{Di} \cap A_{Ri})}{\sum \text{area}(A_{Ri})} \quad (4.19)$$

With the index ϵ_{Di} also ranging between 0 and 1, the total detection error can be identified as;

$$\text{Total detection error } (\epsilon_{Di}) = \sqrt{\frac{\text{Overestimated}(A_{Di})^2 + \text{Underestimated}(A_{Ri})^2}{2}} \quad (4.20)$$

Note: Overestimation and underestimation values range between 0 and 1. A good match is said to be achieved between the reference and the identified tree crowns if the values of equations (4.24 and 4.25) are close to zero. The reference tree object (R_i) is overestimated if $\Delta A_i > 0$, ie $A_{Ri} > A_{Di}$ and underestimated if $\Delta A_i < 0$.

- c) **Positional accuracy:** The positional accuracy measures the variation of the detected tree object centroid from the reference tree objects. Again, the spatial join overlay tool in ArcGIS was used to

overlay both detected tree object and reference data (*Fig 3.3*). Acceptance of agreement is based on having the centroid in between identified tree crown object and reference tree objects. We evaluated the distance between two centroids of reference tree crown $P = X_p, Y_p$ and detected tree objects $O = X_o, Y_o$. Their Euclidean distance and positional error is computed as;

$$\text{Euclidean distance } (E) = \sqrt{(X_o - X_p)^2 + (Y_o - Y_p)^2} \quad (4.21)$$

$$\text{Positional error } (P\epsilon) = \sum_{i=1}^N E \quad (4.22)$$

where; E = the Euclidean distance for one individual tree object.

$P\epsilon$ = the average positional error for all detected tree objects.

N = the total number of tree objects.

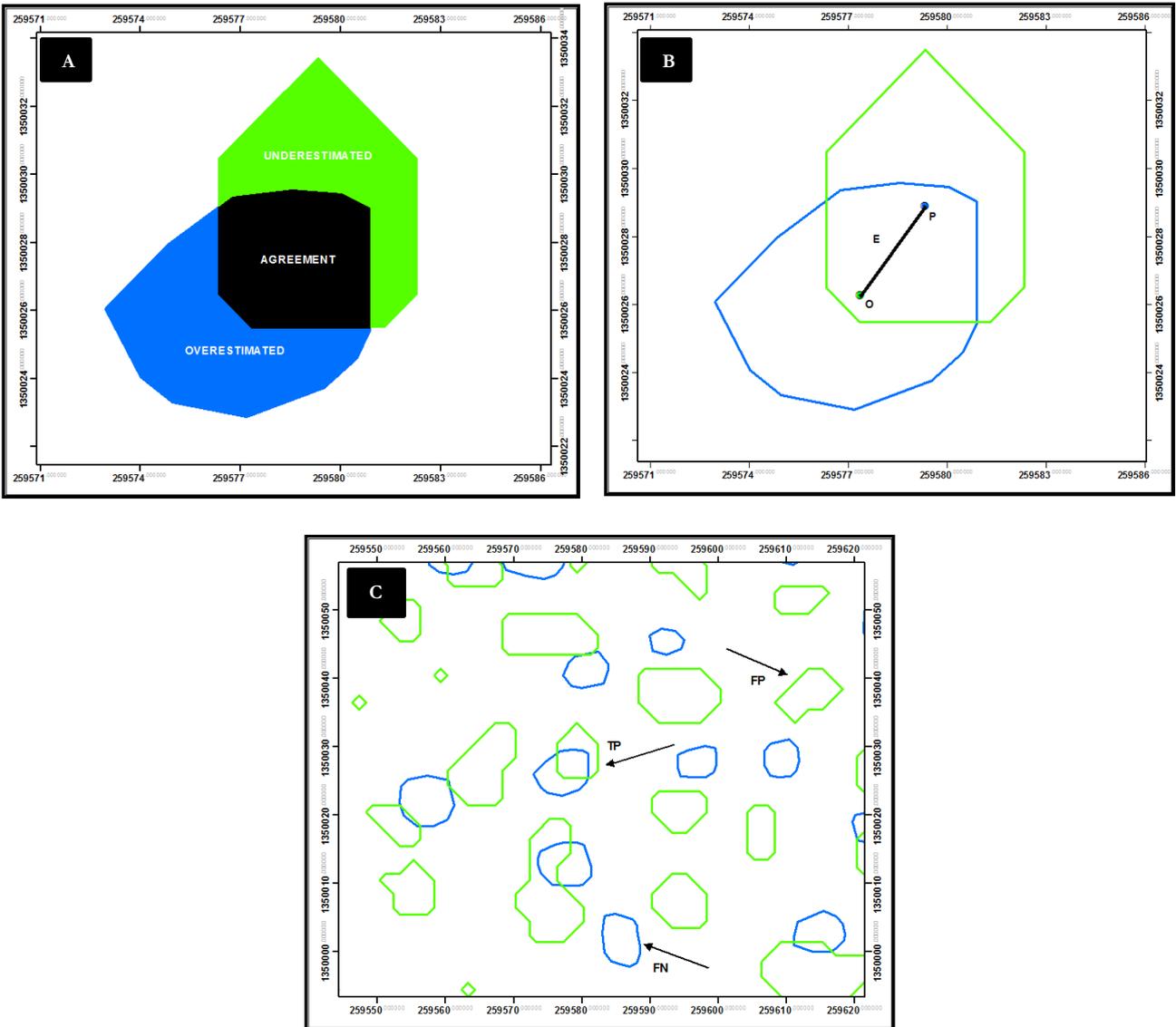


Figure 4.3: Example of accuracy measure of tree crown objects; (A) existential accuracy measure, (B) extensional accuracy measure (C) positional accuracy measure.

5. RESULTS

This chapter shows the results of our method as implemented in this research. It is divided into three sections; the first section shows the results of application on Sentinel-2 multispectral image, the second section shows the results of application on Worldview-3 multispectral image and the third section is the comparison of the results at different scale and spatial resolution.

5.1. Section - 1: Sentinel-2 multispectral image

5.1.1. Spectral class definition

We defined four (4) landcover classes for the Sentinel-2 multispectral image; which are *trees*, *rock*, *barsoil_1*, *baresoil_2* and *shrub*. To get a reliable estimate for the class conditional mean and co-variance matrix for this classification, a sufficient number of training samples is required for each class. For an N dimension spectral space with mean vector of N elements and a symmetric covariance matrix of size $N \times N$, the minimum number of independent training samples is $10N$ and the maximum is $100N$ and above if possible (Richards & Jia, 2006). This rule of thumb was applied in this research to generate the number of training samples for all the classes. For the tree class, point shapefile was generated, while polygon shapefile was used for other classes. The total number of pixels used for training is 1011 pixels and the individual number of pixels for each class is shown in Table 5.1;

Table 5.1: Showing the total number of training samples for each class using the Sentinel-2 multispectral image.

Class Name	Tree	Rock	Barsoil_1	Baresoil_2	Shrub
No. of pixels for training	413	215	113	221	49

A. Contingency matrix

Table 5.2 shows the contingency matrix of the Sentinel-2 multispectral image. The contingency (error matrix) in supervised classification is achieved by comparing the classified thematic map against the reference. The reference data classes are listed by column, thematic map classes are listed by row and the diagonal of the matrix indicates agreement. The number of training samples selected and the classification method applied influences the result of the contingency matrix. The correctly classified pixels are quite high for each class as seen from the diagonal matrix. There is also confusion between the classes showing that some pixels are misclassified. The reason for this is that the *tree* class and other background class like *shrubs* are spectrally similar; to distinguish them is a challenge for spectral classifiers. Again, because of the coarse resolution of this image, we couldn't get sufficient training samples for some classes (*rock*, *baresoil_1*, *baresoil_2*, and *shrub*); this made it difficult for the classifier to clearly distinguish one spectral class from another.

Table 5.2: Showing the contingency matrix for the Sentinel-2 multispectral image.

	Tree	Rock	Barsoil_1	Baresoil_2	Shrub
Tree	734	5	10	60	11
Rock	104	1200	11	688	84
Barsoil_1	26	34	2772	850	52
Baresoil_2	6	3	0	3679	14
Shrub	15	3	13	216	1396

B. Feature space

Fig 5.1 shows the results of both 2D and 3D feature space plots of the Sentinel-2 multispectral image. The 2D feature space of figure 5.1(A), show a plot of near infra-red and green band. The classes are clearly separated from each other. However, within class variance of *tree* class is quite high. The reason is that most pixel reflectance value from coarse resolution images forms a spectral mixture with background classes. This produces mixed pixels and spectrally distinct tree samples of which some were selected for training. To compensate for this, the covariance matrix was divided by a factor of 2 so as to reduce the within class variance while still maintaining sufficient cluster to estimate the distribution for the classifier. Fig 5.1(B), show the feature space plot in near infra-red and blue band. It is similar to fig 5.1(A). The position of the tree class is lower here because the spectral reflectance of the blue band has a shorter wavelength than that of the green band in the electromagnetic spectrum. Fig 5.1(C) shows the feature space plot in green and blue band. There is an overlap between class *tree*, *rock* and *shrub*. *Tree* class is lowest in band 3 because blue band has a shorter wavelength than the green band in the electromagnetic spectrum. Other classes' *rock* and *shrubs* is also closely overlapped with the *tree* class. The 3D feature space plot of figure 5.1(D) is similar to that of figure 5.1(A) with classes clearly separated.

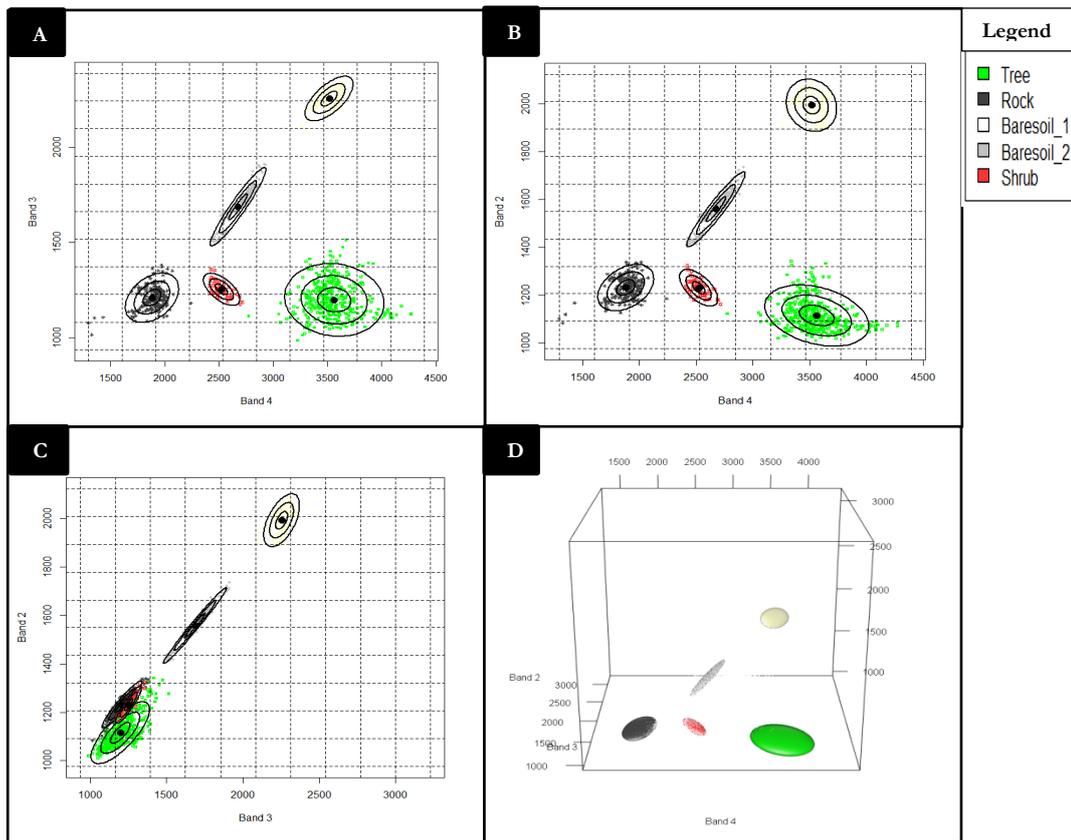


Figure 5.1: Showing feature space plots of Sentinel-2 multispectral image.

C. Class separability

Class separability is a measure of how well classes can be separated from each other. It is determined using the class mean and covariance matrix estimated from training samples for classifiers that estimates class probability distribution. Table 5.3, shows the class separability of Sentinel-2 multispectral image using Jeffries-Matusita distance. The class with the lowest class separability is between *rock* and *shrub* with a value of 1.996. This is followed by class *rock* and *baresoil_2* with a value of 1.998 and finally class *tree* and *shrub* with a value of 1.999. The highest class separability value is 2 showing that those classes are clearly separated. All these class separability values are quite high and close to perfect. Class separability in SRM is dependent on

the class spectral variation, which also depends on scale (Tolpekin & Stein, 2009). Because the resolution of this image is quite coarse, we couldn't generate sufficient number of training pixels in some classes (*rock*, *baresoil_1*, *baresoil_2*, and *shrub*) due to poor class spectral variation. This is why class separability values are high and close to perfect.

Table 5.3: Showing the class separability for the Sentinel-2 multispectral image using Jeffries-Matusita distance.

	Tree	Rock	Barsoil_1	Baresoil_2	Shrub
Tree	0	2.000	2.000	2.000	1.999
Rock	2.000	0	2.000	1.998	1.996
Barsoil_1	2.000	2.000	0	1.999	2.000
Baresoil_2	2.000	1.998	1.999	0	1.999
Shrub	1.999	1.996	2.000	1.999	0

5.1.2. Parameter estimation

A. Simulated annealing

Two parameters initial temperature (T_0) and updating parameter (T_{upd}) control the energy optimization in simulated annealing process (4.5.2). The complexity of SRM problem which depends on scale factor and class separability affects the annealing schedule (Tolpekin & Stein, 2009). As the value of T_{upd} increases, the annealing schedule increases as it takes more time to update. Figure 5.2, shows the plot of T_{upd} and T_0 parameter tuning for Sentinel-2 multispectral image. We plot the mean energy against the T_{upd} in Fig 5.2 (A) and mean energy against T_0 in Fig 5.2 (B). Both plots show a gradual reduction of the mean energy. If the final energy is lowest, we obtain a global minimum. Also if the standard deviation is low, it shows stability and that results are reproducible. In Fig 5.2 (A), we see that the mean energy changes as we experimented with varying the T_{upd} parameter (Table 5.4). The lowest mean energy was the basis for selecting our optimal T_{upd} at 0.8. The standard deviation also varies as well, but the difference between their standard deviations is insignificant.

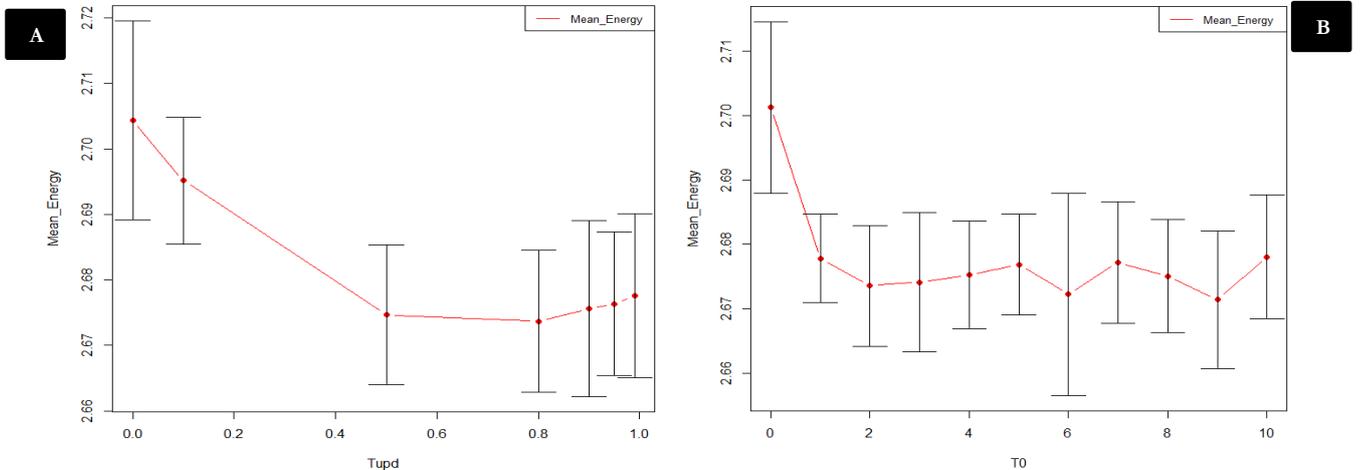


Figure 5.2: Showing gradual reduction of mean energy using Sentinel-2 multispectral image (A) T_{upd} parameter is plotted against mean energy (B) T_0 parameter is plotted against mean energy. The red line shows the mean energy and the bars show the standard deviation from the mean.

In Fig 5.2 (B), we also see that the mean energy changes as we experimented with varying the T_0 parameter (Table 5.5). The standard deviation also varies as well. A high T_0 value is required as to achieve the required global minimum of the solution. A lower T_0 value will achieve a local minimum which is not optimal. Between T_0 values 1 – 10, we observe that the mean energy values are consistently stable with little variation. Any value for T_0 is optimal within this range. Optimal T_0 value is therefore chosen at 4. The difference between their standard deviations is also insignificant.

Table 5.4: Showing the updating temperature parameter T_{upd} estimation for Sentinel-2 multispectral image.

T_{upd}	Mean energy	σ energy
0	2.70	0.015
0.1	2.695	0.009
0.5	2.675	0.010
0.8	2.673	0.010
0.9	2.675	0.013
0.95	2.676	0.010
0.99	2.677	0.012

Table 5.5: Showing the initial temperature parameter T_0 estimation for Sentinel-2 multispectral image.

T_0	Mean energy	σ energy
0	2.701	0.013
1	2.677	0.006
2	2.673	0.009
3	2.674	0.010
4	2.674	0.008
5	2.666	0.007
10	2.678	0.009

B. Smoothness parameter

Both visual inspection of results as well as the value of k was used to make the choice for optimal λ in this research. The optimal λ in SRM problem depends on scale factor. For the Sentinel-2 multispectral image, we used a scale factor of $S=5$. We obtained less noisy results as we applied λ value ranging from 0.50 – 0.55 (Table 5.6). At a higher λ value from 0.56 and above, the mean value of k increased, but the results (based on visual inspection) of the solution was too smooth. While other parameters like ($T_0 = 4$ and $T_{upd} = 0.8$) remained fixed, the optimal λ was selected at $\lambda = 0.55$ having a k value of 0.792 (Table 5.6).

Table 5.6: Showing the smoothness parameter estimation using the Sentinel-2 multispectral image.

λ	0.50	0.51	0.52	0.53	0.54	0.55	0.56
k	0.764	0.770	0.772	0.769	0.767	0.792	0.802

5.1.3. Classification results

Here, we present the results in three parts. The first part is the result of maximum likelihood classification, the second part is the results of pixel based MRF-SRM and the third part shows the result of object based image analysis.

A. Maximum likelihood classification

Fig 5.3 shows (A) Sentinel-2 multispectral image subset and (B) result of maximum likelihood classification of the Sentinel-2 multispectral image subset. Based on visual inspection, the result looks clumsy; *tree* class is clustered and smaller trees were merged with background class like *shrub*. The reason for this is that maximum likelihood assigns pixels to a class with the highest probability of membership. Looking at (Table 5.1), we see that only the *tree* class has sufficient number of training samples while other classes had insufficient training samples. This made it difficult for the classifier to clearly distinguish one spectral class from another. There is confusion between *tree* and *shrub* classes (Table 5.7). Also some confusion exists between *rock* and *shrub* classes. The reason for this is that *tree* class and background class like *shrubs* are spectrally similar; to distinguish them is a challenge for spectral classifiers.

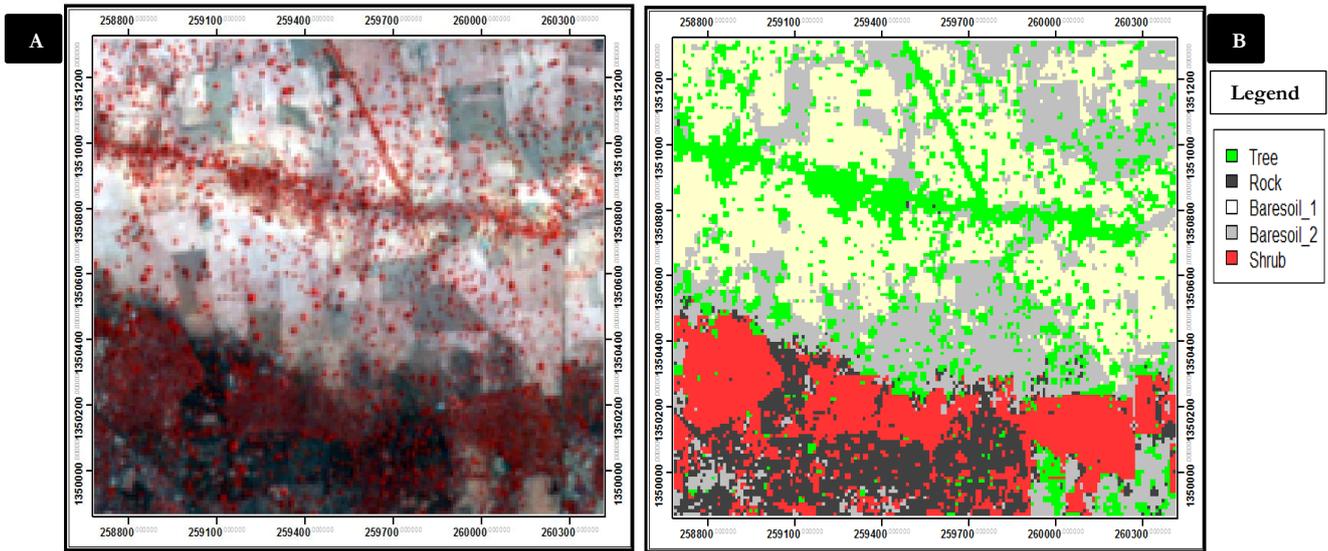


Figure 5.3: Showing subset of (A) Sentinel-2 multispectral image and (B) MLC result of Sentinel-2 multispectral image.

Table 5.7 shows the confusion matrix of the MLC classification with various accuracy measures. Other parts of error matrix accuracy measure include *producer accuracy*, *user accuracy*, *error of omission* and *error of commission* respectively. The classes with has the highest producer accuracy is *tree*, *rock*, *baresoil_1* and *baresoil_2* at 100%, while *baresoil_1*, *baresoil_2* and *shrub* has the highest user accuracy at 100%. Also, class *shrub* has the highest error of omission at 14.28%, while the class *tree* has the highest error of commission at 15%. The overall accuracy of the classification is 98% while the *k* value is 0.976.

Table 5.7: Showing the confusion matrix of MLC result using the Sentinel-2 multispectral image with accuracy assessment.

	Tree	Rock	Barsoil_1	Baresoil_2	Shrub	Error of Commission (%)	User Accuracy (%)
Tree	35	0	0	0	6	15%	85%
Rock	0	71	0	0	3	4%	96%
Barsoil_1	0	0	130	0	0	0%	100%
Baresoil_2	0	0	0	249	0	0%	100%
Shrub	0	0	0	0	54	0%	100%
Error of Omission (%)	0%	0%	0%	0%	14.28%		
Producer Accuracy (%)	100%	100%	100%	100%	86%		

B. Pixel based MRF-SRM

Fig 5.4 shows (A) Sentinel-2 multispectral image subset and (B) the result of pixel based MRF-SRM of the Sentinel-2 multispectral image subset. Based on visual inspection, the result looks smooth and shows a good representation of the input image than the maximum likelihood result. Both big and small trees were well identified. The smoothness of the result is because of the incorporation of λ in the solution (4.17). The initial spatial resolution is 10m. The final output reduced to 2m after we applied a scale factor of $S=5$ based on the degradation model (4.4). There are however some misclassifications between all the classes. Because the image has a coarse resolution, we couldn't get sufficient training samples for some classes (*rock*, *baresoil_1*, *baresoil_2*, and *shrub*); this made it difficult for the classifier to clearly distinguish one spectral class from another. This led to the misclassification between these classes. Again, the *tree* class and background class like *shrubs* are spectrally similar; to distinguish them is a challenge for spectral classifiers.

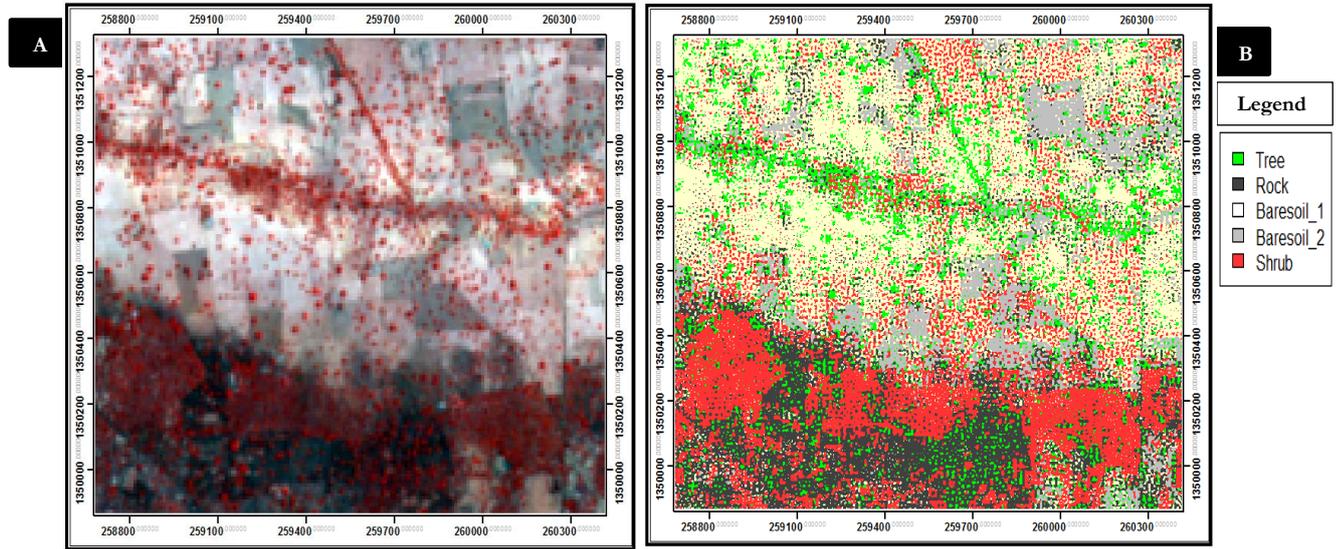


Figure 5.4: Showing subset of (A) Sentinel-2 multispectral image and (B) pixel-based MRF-SRM classification result of Sentinel-2 multispectral image.

Table 5.8 shows the confusion matrix of the classification with various accuracy measures. Other parts of error matrix accuracy measure include *producer accuracy*, *user accuracy*, *error of omission* and *error of commission* respectively. The class with has the highest producer accuracy is *baresoil_1* at 99%, while *baresoil_2* has the highest user accuracy at 99%. Also, the class *rock* has the highest error of commission at 43%, while the class *baresoil_2* has the highest error of omission at 33%. The overall accuracy of the classification is 82% while the k value is 0.79.

Table 5.8: Showing the confusion matrix of pixel-based MRF-SRM classification result using the Sentinel-2 multispectral image with accuracy assessment.

	Tree	Rock	Barsoil_1	Baresoil_2	Shrub	Error of Commission (%)	User Accuracy (%)
Tree	734	5	10	60	11	10%	90%
Rock	104	1200	11	688	84	43%	57%
Barsoil_1	26	34	2772	850	52	26%	74%
Baresoil_2	6	3	0	3679	14	1%	99%
Shrub	15	3	13	216	1396	15%	85%
Error of Omission (%)	17%	4%	1.2%	33%	10%		
Producer Accuracy (%)	83%	96%	99%	67%	90%		

C. Result of object based image analysis

Fig 5.5 shows the result of applying object based image analysis on the output of pixel based MRF-SRM using the Sentinel-2 multispectral image. Based on visual inspection, the result looks smooth and shows a good representation of the input image. Both big and small trees were well identified. However, because of the segmentation applied, some of the smaller *trees* have disappeared and are merged into larger homogenous segments. Other classes like *rock*, *baresoil_1*, *baresoil_2*, and *shrub* were also well identified.

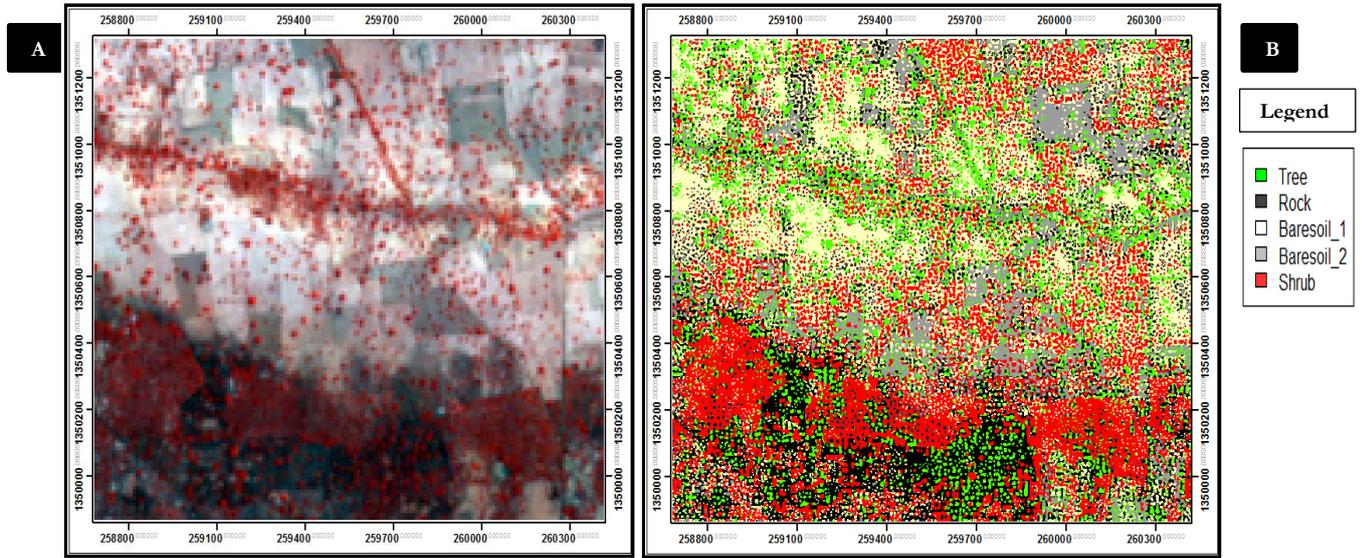


Figure 5.5: Showing subset of (A) Sentinel-2 multispectral image and (B) object-based image analysis SRM result using Sentinel-2 multispectral image.

Fig 5.6 we show the accuracy map after applying object-based image analysis on the output of pixel based MRF-SRM using the Sentinel-2 multispectral image. Fig 5.6 (B) shows overestimated, underestimated and agreement between the identified tree crowns and the reference tree crown overlaid on the Sentinel-2 multispectral image. Fig 5.6 (C) shows only overestimated, underestimated and agreement between the identified tree crowns and the reference tree crown. It can be seen that the identified tree crown matched the reference tree object in some areas. There are also areas where both the identified tree crowns and reference do not agree and areas where they partly agree. The spatial join overlay served as a basis for computing the object based accuracy assessment.

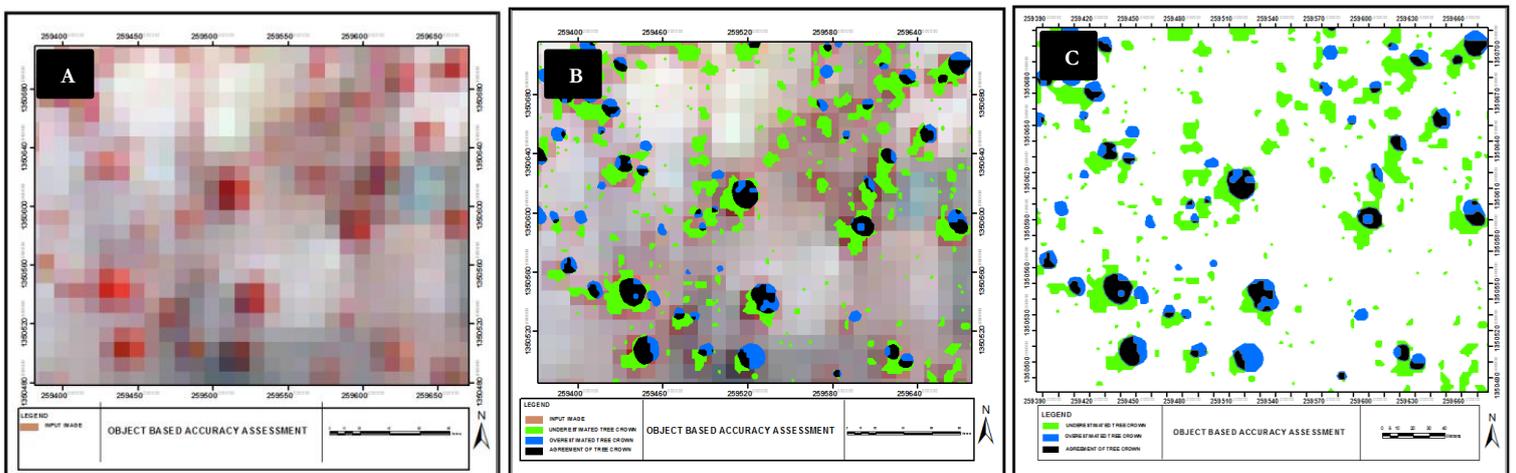


Figure 5.6: Showing the accuracy map for object-based image analysis of SRM classification result using the Sentinel-2 multispectral image.

Table 5.9 shows the results for all three accuracy measures (existential, extensional and positional) for validating detected tree objects in Sentinel-2 multispectral image. A total of 1214 (TP) individual tree crown objects were successfully detected out of 1787 reference objects, while a total of 573 (FP) were not detected out of the reference objects. The reason for missing some tree objects which were not detected in the reference could be because those trees are smaller in size. The positional accuracy between the detected tree objects and reference tree objects is 6.504 m. This is less than one pixel considering the spatial resolution of the input image which is 10 m. The table also shows the results of local area tree difference, overestimation, underestimation and total detection error respectively. Recall that the reference tree objects (R_i) is overestimated if $\Delta A_i > 0$ and underestimated if $\Delta A_i < 0$. Therefore the reference tree object (R_i) is underestimated here because the local area tree difference is -0.5061 which is less than zero. The reason for underestimation may be because of the sun illumination angle during image acquisition and considering that trees are elevated objects; these can cause occlusion casting shadow and misplacement of the position of the trees. This may be verified by visual interpretation of the results (Fig 5.6). Another reason is that most pixel reflectance value from coarse resolution images forms a spectral mixture with background classes. This produces the mixed pixel effect which might cause some misplacement in the actual position of the trees. The total detection error at 0.4669 is also greater than zero. This also signifies that the reference tree object (R_i) is underestimated. Finally, the detection accuracy is still acceptable at 68%. This shows good percentage detection between the reference tree object and the detected tree crown object.

Table 5.9: Showing results of object based accuracy measure applied on the Sentinel-2 multispectral image (total detection, false negative, positional accuracy, local area difference, overestimated, underestimated total detection error).

Image	Total detection (%)	False negative (%)	Positional accuracy (m)	Local area difference	Overestimated	Underestimated	Total detection error
<i>Sentinel-2</i>	1214 (68%)	573 (32%)	6.504	- 0.5061	0.6455	0.1394	0.4669

5.2. Section - 2: Worldview-3 multispectral image

5.2.1. Spectral class definition

We defined eight (8) landcover classes for the Worldview-3 multispectral image; which are *trees*, *road*, *built_up*, *field_1*, *field_2*, *field_3*, *baresoil* and *shrub*. The total number of pixels used for training is 4271 pixels and shown in Table 5.10.

Table 5.10: Showing the total number of training samples for each class using the Worldview-3 multispectral image.

Class Name	Tree	Road	Built_up	Field_1	Field_2	Field_3	Baresoil	Shrub
No. of pixels for training	1192	65	74	1158	960	254	363	205

A. Contingency matrix

Table 5.11 shows the contingency matrix of Worldview-3 multispectral image. The correctly classified pixels are quite high for each class as seen from the diagonal matrix. The spatial resolution of the image is high and most landcover classes were spectrally distinct. This availed us to get sufficient number of training samples in classes' *tree*, *field_1* and *field_2* (Table 5.10). There is however confusion between classes *tree*, *shrub* and *field_3*. The reason is that the *tree* class and these background classes like *shrubs*, *field_3* are spectrally similar; to distinguish them is a challenge for spectral classifiers. Some confusion also exists between class *road* and *baresoil* which still show spectral similarity; this is challenging to distinguish.

Table 5.11: Showing the contingency matrix of the Worldview-3 multispectral image.

	Tree	Road	Built_up	Field_1	Field_2	Field_3	Baresoil	Shrub
Tree	419	0	0	0	0	0	0	0
Road	0	41	0	0	0	0	77	0
Built_up	0	0	88	0	0	0	0	0
Field_1	0	0	0	386	0	0	0	0
Field_2	0	0	0	0	495	2	0	0
Field_3	0	0	0	1	0	37	0	0
Baresoil	0	36	0	0	1	0	11	0
Shrub	2	0	0	0	0	60	0	83

B. Feature space

Fig 5.7 shows the feature space plots of the Worldview-3 multispectral image. The 2D feature space of Fig 5.7(A), show a plot of near infra-red and green band. The dots are denser here for the *tree* class which also has a large variance than in the Sentinel-2 multispectral image. This is as a result of high spatial resolution and large number of the training samples (Table 5.10). The *built_up* and *road* classes also have a large variance. But this is because our training samples for both classes were spectrally distinct and insufficient (Table 5.10). There is overlap between class *tree*, *shrub*, *field_3* and *field_1* classes. This is because *tree* class and background classes like *shrubs*, *field_3* and *field_1* are spectrally similar; to distinguish them is a challenge for spectral classifiers. Again, this confusion also exists between class *field_2*, *baresoil* and *road* which still show similar spectral characteristics. Fig 5.7(B), show the feature space plot in near infra-red and blue band. It is similar to Fig 5.7(A) and the position of the tree class remained almost the same. Fig 5.7(C) shows the feature space plot in green and blue band. Most of the classes overlap among themselves. *Tree* class is lowest here because the blue band has a shorter wavelength than green band in the electromagnetic spectrum. The 3D feature space plot of figure 5.6(D) is similar to that of figure 5.6(A) with the ellipse of class *tree*, *shrub*, *field_3* and *field_1* overlapping as well as the ellipse of classes' *field_2*, *baresoil* and *road*.

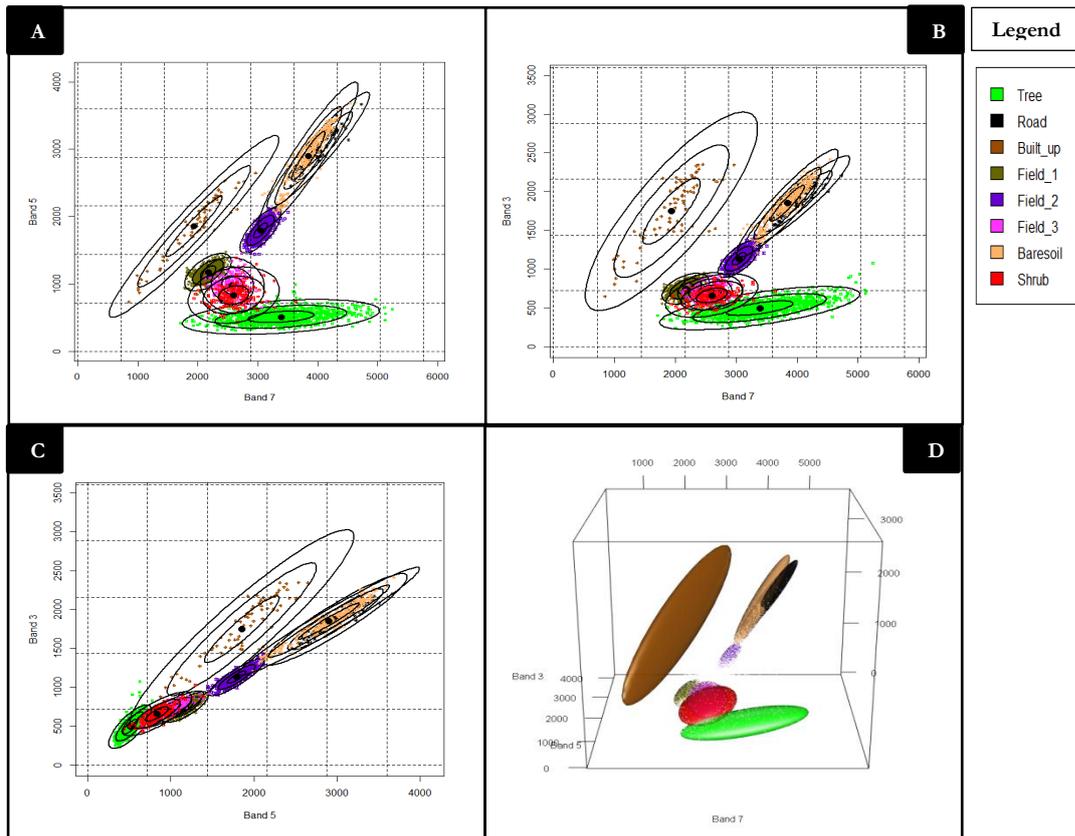


Figure 5.7: Showing feature space plot using the Worldview-3 multispectral image.

C. Class separability

We show the class separability of the Worldview-3 multispectral image in Table 5.12 using Jeffries-Matusita distance. The class with the lowest class separability is between *field_3* and *shrub* with a value of 1.230. This is followed by class *road* and *baresoil* with a value of 1.460 and finally class *field_1* and *field_3* with a value of 1.990. High class separability values are because most of the classes are spectrally distinct. We also had insufficient number of training samples for classes' *road*, *built_up*, *field_3*, *baresoil* and *shrub* (Table 5.10) leading to high class separability value. The insufficient number of training samples is due to the large configuration of image space and spectral dissimilarity between these classes. The reason for a lower separability is because spectral classifiers find it difficult to distinguish between spectral similar classes (*tree*, *shrub*, *field_3* and *field_1*). The highest class separability value is 2.

Table 5.12: Showing the class separability of the Worldview-3 multispectral image using Jeffries-Matusita distance.

	Tree	Road	Built_up	Field_1	Field_2	Field_3	Baresoil	Shrub
Tree	0	2.000	2.000	2.000	2.000	2.000	2.000	2.000
Road	2.000	0	2.000	2.000	2.000	2.000	1.460	2.000
Built_up	2.000	2.000	0	2.000	2.000	2.000	2.000	2.000
Field_1	2.000	2.000	2.000	0	2.000	1.990	2.000	2.000
Field_2	2.000	2.000	2.000	2.000	0	2.000	2.000	2.000
Field_3	2.000	2.000	2.000	1.990	2.000	0	2.000	1.230
Baresoil	2.000	1.460	2.000	2.000	2.000	2.000	0	2.000
Shrub	2.000	2.000	2.000	2.000	2.000	1.230	2.000	0

5.2.2. Parameter estimation

A. Simulated annealing

The two parameters which control the energy optimization in simulated annealing process are initial temperature (T_0) and updating parameter (T_{upd}). Fig 5.8 shows the plot of T_{upd} and T_0 parameter tuning for Worldview-3 multispectral image. We plot the mean energy against the T_{upd} in Fig 5.8 (A) and mean energy against T_0 in Fig 5.8 (B). Both plots show a gradual reduction of the mean energy. If the final energy is lowest, we obtain a global minimum. Also if the standard deviation is low, it shows stability and that results are reproducible. In Fig 5.8 (A), we see that the mean energy changes as we experimented with varying the T_{upd} parameter (Table 5.13). The lowest mean energy and standard deviation was the basis for selecting our optimal T_{upd} at 0.99.

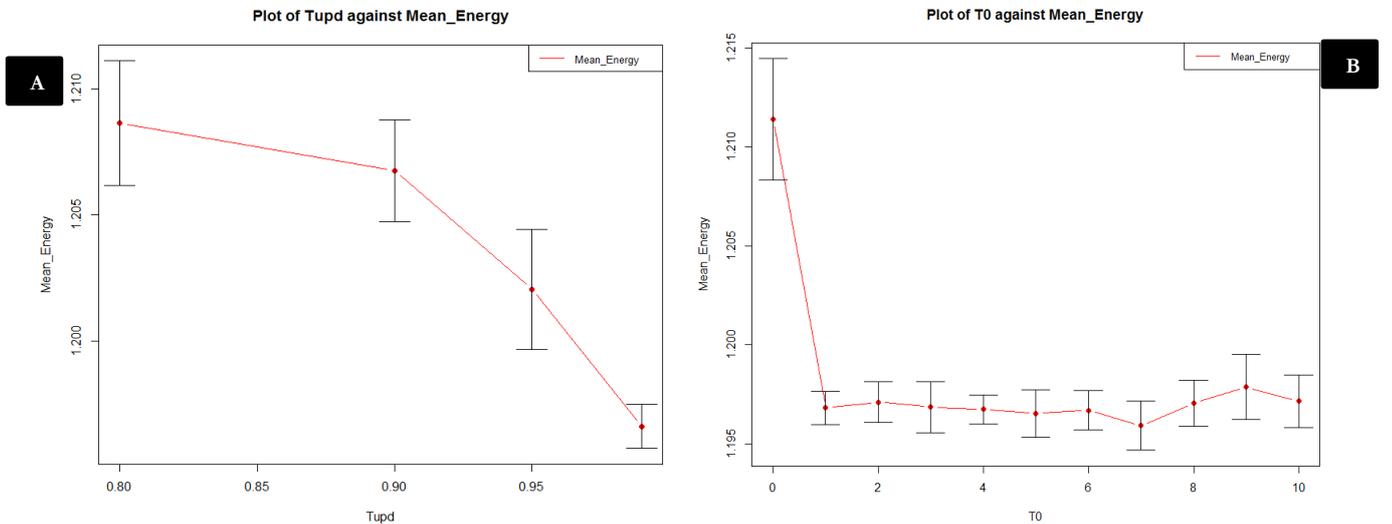


Figure 5.8: Showing gradual reduction of mean energy using Worldview-3 multispectral image (A) T_{upd} parameter is plotted against mean energy (B) T_0 parameter is plotted against mean energy. The red line shows the mean energy and the bars show the standard deviation from the mean.

In Fig 5.8 (B), we also see that the mean energy changes as we experimented with varying the T_0 parameter (Table 5.14). The standard deviation also varies as well. A high T_0 value is required as to achieve the required global minimum of the solution. A lower T_0 value will achieve a local minimum which is not optimal. Again, between T_0 values 1 – 10, we observe that the mean energy values are consistently stable with little variation. Any value for T_0 is optimal within this range. Optimal T_0 value is therefore chosen at 4. The difference between their standard deviations is also insignificant.

Table 5.13: Showing the initial temperature parameter estimation using the Worldview-3 multispectral image.

T_{upd}	Mean energy	σ energy
0.8	1.208	0.0024
0.9	1.206	0.0020
0.95	1.202	0.0023
0.99	1.196	0.0008

Table 5.14: Showing the updating parameter T_{upd} estimation using the Worldview-3 multispectral image.

T_0	Mean energy	σ energy
0	1.211	0.0030
1	1.196	0.0007
2	1.197	0.0010
3	1.196	0.0013
4	1.196	0.0007
5	1.196	0.0012
10	1.197	0.0013

B. Smoothness parameter

Again, both visual inspection of results as well as the value of k was used to make the choice for optimal λ . We obtained less noisy results as we applied λ value ranging from 0.86 – 0.95 (Table 5.15). At a higher λ value from 0.91 and above, the mean value of k increased, but the results (based on visual inspection) of the solution was poor and over smoothed. While other parameters like ($T_0 = 4$ and $T_{upd} = 0.99$) remained fixed, the optimal λ for the Worldview-3 multispectral image was selected at $\lambda = 0.90$ having a k value of 0.874 (Table 5.15).

Table 5.15: Showing the smoothness parameter estimation using the Worldview-3 multispectral image.

λ	0.85	0.86	0.87	0.88	0.89	0.90	0.91	0.92	0.93	0.94	0.95
k	0.878	0.876	0.876	0.873	0.875	0.874	0.877	0.862	0.880	0.871	0.886

5.2.3. Classification results

Here, we present the results in three parts. The first part is the result of maximum likelihood classification, the second part is the results of pixel based MRF and the third part shows the result of object based image analysis.

A. Maximum likelihood classification

Fig 5.9 shows (A) Worldview-3 multispectral image subset and (B) result of maximum likelihood classification of the Worldview-3 multispectral image subset. Based on visual inspection, the result looks noisy, but most classes can still be distinguished. The noise which exists mainly between classes' *tree*, *shrubs* and *field_1* is because these classes are spectrally similar and challenging to distinguish for spectral classifiers. From (Table 5.10), we see that the *road*, *built_up*, *field_3*, *baresoil* and *shrub* classes have insufficient number of training

samples. This made it difficult for the classifier to clearly distinguish them. There is confusion between class *tree*, *shrub*, *field_3* and *field_1* classes (Table 5.16). This is because *tree* class and background classes like *shrubs*, *field_3* and *field_1* are spectrally similar; to distinguish them is a challenge for spectral classifiers. Again, this confusion also exists between class *field_2*, *baresoil* and *road* which still show similar spectral characteristics.

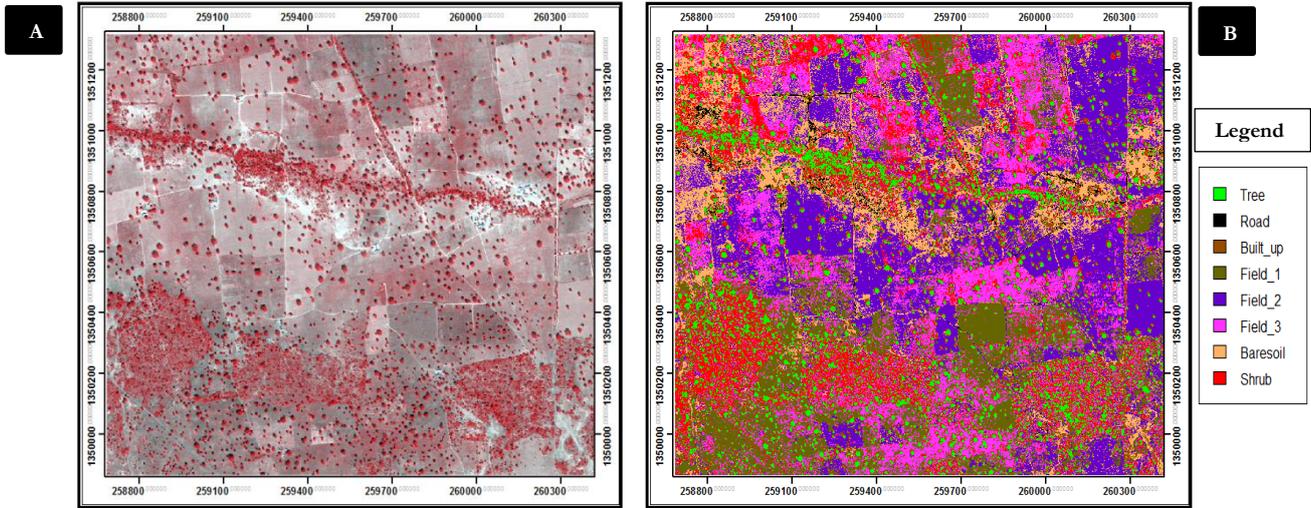


Figure 5.9: Showing subset of (A) Worldview-3 multispectral image and (B) MLC result of Worldview-3 multispectral image.

Table 5.16 shows the confusion matrix of the MLC classification with various accuracy measures. Other parts of error matrix accuracy measure include *producer accuracy*, *user accuracy*, *error of omission* and *error of commission* respectively. The classes with has the highest producer and user accuracy is *built_up* at 100%. The class *shrub* has the highest error of omission at 14.63%, while the class *road* has the highest error of commission at 35.2%. The overall accuracy of the classification is 96.3% while the k value is 0.844.

Table 5.16: Showing the confusion matrix of MLC result using the Worldview-3 multispectral image with accuracy assessment.

	Tree	Road	Built_up	Field_1	Field_2	Field_3	Baresoil	Shrub	Error of Commission (%)	User Accuracy (%)
Tree	1176	0	0	0	0	4	0	7	0.92%	99.1%
Road	0	57	0	0	0	0	31	0	35.2%	64.8%
Built_up	0	0	74	0	0	0	0	0	0%	100%
Field_1	0	0	0	1134	0	4	0	3	0.62%	99.4%
Field_2	0	0	0	0	958	1	1	0	0.21%	99.8%
Field_3	1	0	0	20	0	210	0	20	16.33%	83.7%
Baresoil	0	8	0	0	2	0	331	0	2.93%	97.1%
Shrub	15	0	0	4	0	35	0	175	23.58%	76.4%
Error of Omission (%)	1.34%	12.31%	0%	2.07%	0.21%	17.32%	8.82%	14.63%		
Producer Accuracy (%)	98.6%	87.7%	100%	97.9%	99.8%	82.7%	91.2%	85.4%		

B. Pixel based MRF

Fig 5.10 shows (A) Worldview-3 multispectral image subset and (B) the result of pixel based MRF of the Worldview-3 multispectral image subset. Based on visual inspection, the result looks smooth and shows a good representation of the input image than the maximum likelihood result. Both big and small trees were

well identified. The smoothness of the result is because of the incorporation of λ in the solution (4.17). The spatial resolution is 2m and to obtain the nominal scale for the results, we applied a scale factor of $S=1$ based on the degradation model (4.4). There is confusion between class *tree*, *shrub*, *field_3* and *field_1* classes (Table 5.16). This is because *tree* class and background classes like *shrubs*, *field_3* and *field_1* are spectrally similar; to distinguish them is a challenge for spectral classifiers. Again, this confusion also exists between class *field_2*, *baresoil* and *road* which still show similar spectral characteristics.

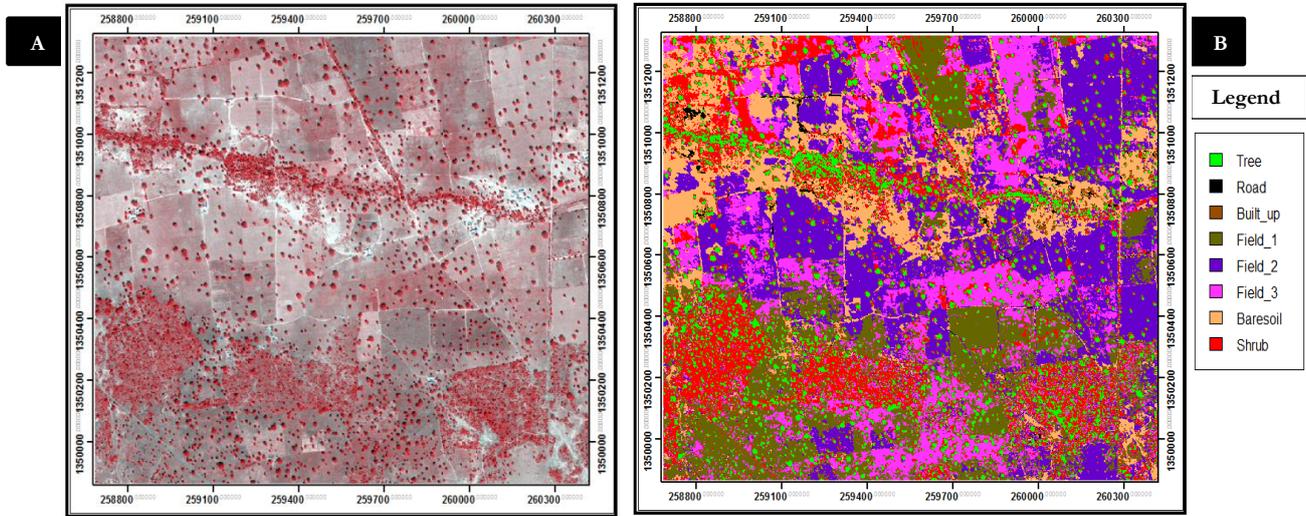


Figure 5.10: Showing subset of (A) Worldview-3 multispectral image and (B) pixel-based MRF classification result using the Worldview-3 multispectral image.

Table 5.17 shows the contingency matrix of the classification with various accuracy measures. The accuracy measure includes *producer accuracy*, *user accuracy*, *error of omission* and *error of commission* respectively. The classes with has the highest producer accuracy is *built_up* and *shrub* at 100%, while *tree*, *built_up* and *field_1* has the highest user accuracy at 100%. Also, the class *baresoil* has the highest error of omission and commission at 87.5% and 77% respectively. The overall accuracy of the classification is 90% while the k value is 0.87.

Table 5.17 : Showing the confusion matrix of the Worldview-3 multispectral image with accuracy assessment.

	Tree	Road	Built_up	Field_1	Field_2	Field_3	Baresoil	Shrub	Error of Commission (%)	User Accuracy (%)
Tree	419	0	0	0	0	0	0	0	0%	100%
Road	0	41	0	0	0	0	77	0	65.3%	34.8%
Built_up	0	0	88	0	0	0	0	0	0%	100%
Field_1	0	0	0	386	0	0	0	0	0%	100%
Field_2	0	0	0	0	495	2	0	0	0.40%	99.7%
Field_3	0	0	0	1	0	37	0	0	2.63%	97.4%
Baresoil	0	36	0	0	1	0	11	0	77%	22.9%
Shrub	2	0	0	0	0	60	0	83	42.8%	57.2%
Error of Omission (%)	0.47%	46.8%	0%	0.25%	0.20%	62.6%	87.5%	0%		
Producer Accuracy (%)	99.5%	53.2%	100%	99.7%	99.7%	37.4%	12.5%	100%		

C. Results of object based image analysis

Fig 5.11 (B) shows the result of applying object based image analysis on the output of pixel based MRF classification using the Worldview-3 multispectral image. Both big and small trees were well identified. However, because of the segmentation applied, some of the smaller *trees* have disappeared and are merged into larger homogenous segments. All other classes were also well identified.

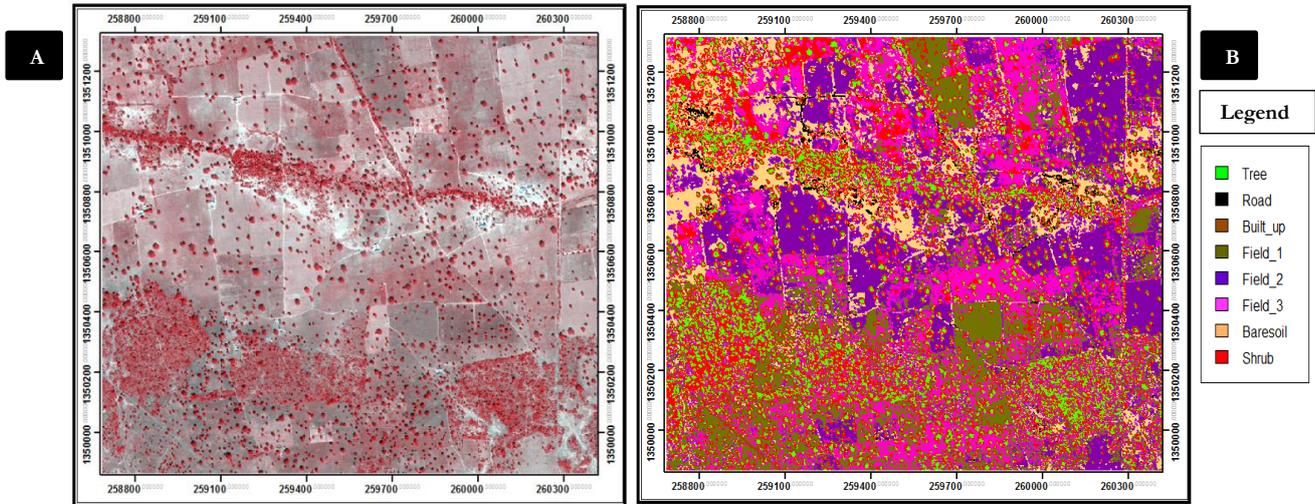


Figure 5.11: Showing subset of (A) Worldview-3 multispectral image and (B) object-based image analysis of MRF result using the Worldview-3 multispectral image.

Fig 5.12 we show the accuracy map after applying object-based image analysis on the output of pixel based MRF using the Worldview-3 multispectral image. Fig 5.12 (B) shows overestimated, underestimated and agreement between the identified tree crowns and the reference tree crown overlaid on the Worldview-3 multispectral image. Fig 5.12 (C) shows only overestimated, underestimated and agreement between the identified tree crowns and the reference tree crown. It can be seen that the identified tree crown correctly matched the reference tree object in all areas. There are however small areas where both the identified tree crowns and reference do not agree and areas where they partly agree. The spatial join overlay served as a basis for computing the object based accuracy assessment.

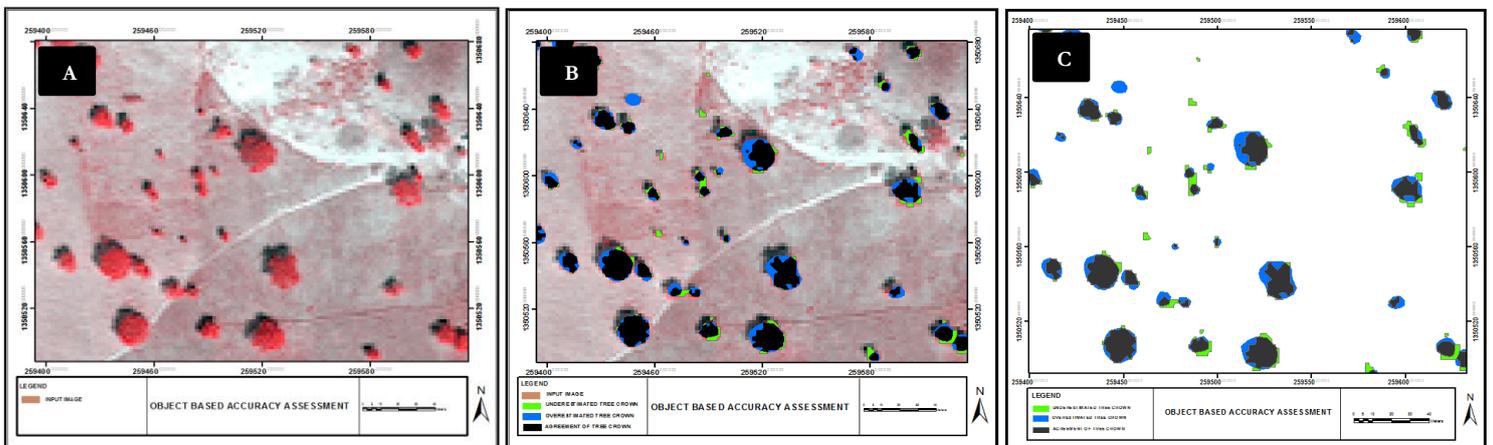


Figure 5.12: Showing the accuracy map of object-based image analysis of MRF classification results using the Worldview-3 multispectral image.

Table 5.18 shows the results for all three accuracy measures (existential, extensional and positional) for validating detected tree objects in Worldview-3 multispectral image. A total of 1772 (TP) individual tree crown objects were successfully detected out of 1787 reference objects, while a total of 15 (FP) were not

detected out of the reference objects. The positional accuracy between the detected tree objects and reference tree objects is 1.714 m. This is less than one pixel considering the spatial resolution of the input image which is 2 m. The table also shows the results of the local area tree difference, overestimation error, underestimation error and total detection error respectively. Recall that the reference tree objects (R_i) is overestimated if $\Delta A_i > 0$ and underestimated if $\Delta A_i < 0$. Therefore the reference tree object (R_i) is underestimated here because the local area tree difference is -0.1314 which is less than zero. The reason for underestimation may be because of the sun illumination angle during image acquisition and considering that trees are elevated objects; these can cause occlusion casting shadow and misplacement of the position of the trees. This may be verified by visual interpretation of the results (Fig 5.12). The total detection error at 0.3179 is also greater than zero. This also signifies that the reference tree object (R_i) is underestimated. Finally, the detection accuracy is much higher here at 99%. This shows very high percentage detection between the reference tree object and the detected tree crown object.

Table 5.18: Showing results of object based accuracy measure applied on the Worldview multispectral image (total detection, false negative, positional accuracy, local area difference, overestimated, underestimated total detection error).

Image	Total detection (%)	False negatives (%)	Positional accuracy (m)	Local area difference	Overestimated	Underestimated	Total detection error
Worldview-3	1772 (99%)	15 (1%)	1.714	- 0.1314	0.3767	0.2453	0.3179

Section - 3 Comparison of the results at different scale

Here, the results of our method as applied in different scale and spatial resolution is compared. Fig 5.13 shows the scatter plot of overestimation and underestimation of both (A) Sentinel-2 and (B) Worldview-3 multispectral images respectively. The scatter plot of Fig 5.13 (A) shows that most of the overestimated and underestimated values cluster (agree) around 0. The range of value between the lowest and highest overestimation error is very large (0 - 80). This confirms the high overestimation error (Table 5.9). Also, the range of value between the lowest and highest underestimation error is smaller (0 - 23). This confirms the lower underestimation error (Table 5.9). The scatter plot of Fig 5.12 (B) is similar to that of Fig 5.12 (A). The difference here is that we have a lower range of value for overestimation (0 - 90) and a higher value of underestimation (0 - 25). This difference in their respective value ranges confirms the overestimation and underestimation errors respectively (Table 5.18).

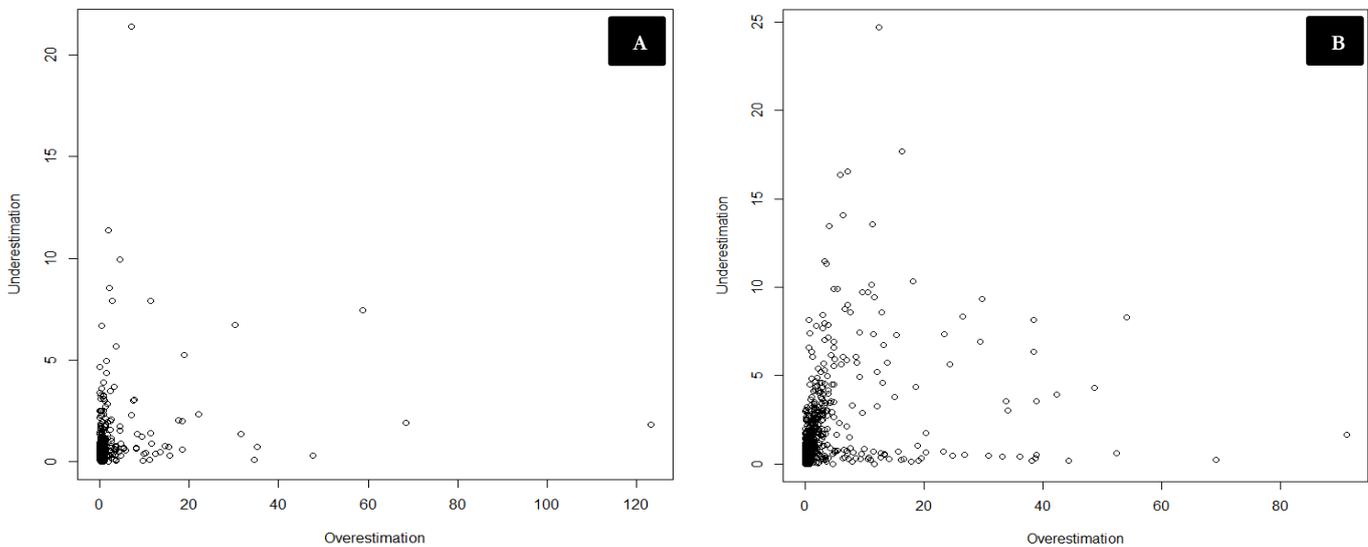


Figure 5.13: Showing the scatter plots of overestimation and underestimation for (A) Sentinel-2 multispectral image and (B) Worldview-3 multispectral image.

In Table 5.19, we compare the detection, positional accuracy, local area difference, overestimated, underestimated and total detection error respectively for both images. The positional accuracy between the detected tree objects and reference tree objects are 6.504 for Sentinel-2 multispectral image and 1.714 for Worldview-3 multispectral image respectively. This is less than one pixel considering the spatial resolution of the input images which is 10 m and 2 m respectively. It signifies that both images have a good positional accuracy. We see that the Worldview-3 multispectral image has a higher detection, a lower local area tree difference and a lower total detection error. However, both images show that the reference tree object (R_i) is underestimated because their respective local area tree difference at -0.5061 for Sentinel-2 multispectral image and -0.1314 for Worldview-3 multispectral image are both less than zero. Finally, both images have high percentage detection accuracy at 68% for Sentinel-2 multispectral image and 99% for Worldview-3 multispectral image respectively. This shows that they are both suitable for tree object detection using super resolution mapping in HR scenario.

Table 5.19: Showing the comparison of results of objects based accuracy assessment between both images (total detection, false negative, positional accuracy, local area difference, overestimated, underestimated total detection error).

<i>Image</i>	Total detection (%)	False negatives (%)	Positional accuracy (m)	Local area difference	Overestimated	Underestimated	Total detection error
<i>Sentinel-2</i>	1214 (67.94%)	573 (32%)	6.504	- 0.5061	0.6455	0.1394	0.4669
<i>Worldview-3</i>	1772 (99%)	15 (1%)	1.714	- 0.1314	0.3767	0.2453	0.3179

6. DISCUSSION

Here, the results obtained in preceding chapters are discussed in two sections. The first section covers the parameter estimation for both Sentinel-2 and Worldview-3 multispectral images, while the second section covers the results and comparison of detected tree objects.

6.1. Section 1: Parameter estimation

The objective of parameter estimation is to determine the optimal values for each of the parameters employed. Both images used in this research showed a distinct behaviour for different parameter combinations. This is discussed in two sub sections.

6.1.1. Sentinel-2 multispectral image

The simulated annealing process is stochastic and requires multiple experiments for optimal choice. We therefore conducted a systematic experiment and analysed the results. For each T_0 and T_{upd} parameter, 10 equal number of runs were made and we recorded the respective result of energy. As we vary the parameters, we repeated the same experiment. At the end, the mean as well as the standard deviation of energy was computed and plotted against the respectively parameters (Fig 5.2). Statistically, more than 10 runs are required for stability, but we were constrained to only 10 runs because of time.

The results of Fig 5.2 (A and B) show a gradual reduction of the mean energy which changes as we experimented with varying the T_0 and T_{upd} parameters. If the final energy is lowest, we obtain a global minimum. Also if the standard deviation is low, it shows stability and that results are reproducible. As the value of T_0 and T_{upd} increases, the annealing schedule increases as it takes more time to update. A high T_0 value is required as to achieve the required global minimum of the solution. A lower T_0 value will achieve a local minimum which is not optimal. It is expected that as we vary (increase) the T_0 and T_{upd} parameters, the energy should decrease. But in Fig 5.2 (A), we observe a different trend because the T_{upd} parameter decreases from 0 up to 0.8 and then begins to increase as we moved to 1.0. This is an interesting finding which may be attributed to the high contrast sensitivity of this image and/or the complexity of the SRM problem here compared to nominal resolution classification. In Fig 5.2 (B) and between T_0 values 1 – 10, we observe that the mean energy values are consistently stable with little variation. Any value for T_0 is assumed to be optimal within this range. An increase or decrease in the values of T_0 and T_{upd} parameters did not affect result of k . The difference between their standard deviations is also insignificant (very low) and does not affect value of k . Our choice for optimal T_{upd} was 0.8, while T_0 value was chosen at 4.

The smoothness parameter λ greatly affects the output of the results. We selected our optimal choice for lambda in this image by trial and error experiment. This is not sufficient as it is required to at least conducted a systematic experiment and analyse the results. But we were constrained because of time. However, depending on our scale factor, we obtained less noisy results as we applied λ value ranging from 0.50 – 0.55 (Table 5.6). A larger range of λ value may show a different trend but was not tested. As the value of λ increases, the value of k also increases as well and the result based on visual inspection becomes smoother. This also increases the computation time. However, a higher value of λ does not necessarily give the best output, rather we neglect the likelihood term in the equation (4.16). This also increases the contextual dependency of pixel and affects the spectral contribution to the solution. The effect is that the resulting minimal posterior energy is obtained where all pixels is assigned to one class leading to an over smooth output. A lower value of λ allows more isolated pixels to be identified, but the result based on visual inspection becomes noisy. Based on visual inspection, the optimal λ for this image was selected at 0.55, while other parameters like ($T_0 = 4$ and $T_{upd} = 0.8$) remained fixed.

6.1.2. Worldview-3 multispectral image

For the simulated annealing parameters using this image, we repeated the same experiment as we did with the Sentinel-2 multispectral image. The results of Fig 5.8 (A and B) also show a gradual reduction of the mean energy which changes as we experimented with varying the T_0 and T_{upd} parameters. If the final energy is lowest, we obtain a global minimum. Also if the standard deviation is low, it shows stability and that results are reproducible. As the value of T_0 and T_{upd} increases, the annealing schedule increases as it takes more time to update. Again, a high T_0 value is required as to achieve the required global minimum of the solution. A lower T_0 value will achieve a local minimum which is not optimal. It is also expected that as we vary (increase) the T_0 and T_{upd} parameters, the energy should decrease. From Fig 5.8 (A), we observe this expected trend because the T_{upd} parameter decreases as we varied from 0.8 up to 1.0. In Fig 5.8 (B) and between T_0 values 1 – 10, we observe that the mean energy values are consistently stable with little variation. Any value for T_0 is assumed to be optimal within this range. An increase or decrease in the values of T_0 and T_{upd} parameters did not affect result of \mathcal{K} . The difference between their standard deviations is again insignificant (very low) and does not affect value of \mathcal{K} . Our choice for optimal T_{upd} was 0.99, while T_0 value was chosen at 4.

The smoothness parameter λ greatly affects the output of the results. Again, we selected our optimal choice for lambda in this image by trial and error experiment. This is not sufficient as it is required to at least conducted a systematic experiment and analyse the results. But we were constrained because of time. Having applied a nominal scale for this image, we obtained less noisy results as we applied λ value ranging from 0.86 – 0.95 (Table 5.15). As the value of λ increases, the value of \mathcal{K} also increases and also the results based on visual inspection becomes smoother. This also increases the computation time. Again, a higher value of λ does not necessarily give the best output, rather we neglect the likelihood term in the equation (4.16). This also increases the contextual dependency of pixel and affects the spectral contribution to the solution. The effect is that the resulting minimal posterior energy is obtained where all pixels is assigned to one class leading to an over smooth output. Again, a lower value of λ allows more isolated pixels to be identified, but the result based on visual inspection becomes noisy. Based on visual inspection, the optimal λ for this image was selected at 0.90, while other parameters like ($T_0 = 4$ and $T_{upd} = 0.99$) remained fixed.

6.2. Section 2: Results and comparison of detected tree objects

6.2.1. Results of Sentinel-2 multispectral image

We applied the maximum likelihood assumption which is based on the Gaussian normal distribution to train our classifier (2.3). This requires sufficient training samples to be able to estimate correctly the mean and covariance matrix for all classes of the distribution. However from (Table 5.1), we see that only the *tree* class had sufficient number of training samples, while other classes *rock*, *baresoil_1*, *baresoil_2*, and *shrub* has insufficient training samples. We couldn't get sufficient training samples for some classes because of the coarse resolution of this image. As a result it was difficult for the classifier to clearly distinguish one spectral class from another and affected all the results of our classification applied on this image.

Another challenge faced is that within class variance of *tree* class was quite high (Fig 5.1). The reason for this is that most pixel reflectance value from coarse resolution images forms a spectral mixture with background classes. This produced mixed pixels and spectrally distinct *tree* samples of which some were selected for training. However, because maximum likelihood does not deal with mixed pixels, most spectral information is lost in the process of transforming the remotely sensed data into thematic maps. The implication of this is that it reports a unique attribution of each pixel to a class with highest membership, while ignoring the others with low class membership properties. To compensate for this, the covariance matrix was divided by a factor of 2 so as to reduce the within class variance; while still maintaining sufficient cluster to estimate the distribution for the classifier. The expectation of reduced variance is that classes should be clearly separable from each other.

The choice of neighbourhood system is another point of emphasis. This system determines the surrounding pixels that will be taken into consideration during analysis. For our study, we applied the first order neighbourhood system for both classification and segmentation (Fig 4.2). This considers only four pixels which shares sides with a given pixel and has effect on the objects detected. Tolpekin & Stein, (2009) identified that a large window is essential for result quality. We couldn't consider the second order neighbourhood system because of time constrains. It will be wise to see if this improves our detection of tree objects for this image.

In Fig 5.3 (B), we show the result of maximum likelihood classification of the Sentinel-2 multispectral image subset. Based on visual inspection, spectral details from input image were not properly distinguished. Most of the *trees* were clustered and smaller *trees* were merged with background class like *shrub*. Only a few standalone trees were clearly identified. Because only the *tree* class had sufficient number of training samples, it is assumed that most of the pixels in the image might have been wrongly assigned to it. Table 5.7 shows the confusion matrix of the MLC classification with various accuracy measures. Despite the visual appearance of the classification result, the overall accuracy and k value of the classification was quite high at 98% and 0.976 respectively. This suggests that most of the spectral classes with insufficient training samples were assigned to a wrong class because this classifier assigns pixels to the class with highest probability of membership. The inspection of the maximum likelihood result shows that it's not good for tree detection using this image.

In Fig 5.4 (B), we show the result of pixel based MRF-SRM of the Sentinel-2 multispectral image subset. Based on visual inspection, the result looks smooth and much better representation of the input image (Fig 5.4) than the maximum likelihood result (Fig 5.3). The smoothness of the result is because of the incorporation of λ in the solution (4.17). SRM helped to solve the mixed pixel problem and also reduced the spatial resolution of the output to 2 m. This helped to properly identify both big and small trees respectively. Table 5.8 shows the confusion matrix of the classification with various accuracy measures. The overall accuracy and k value of the classification was still substantially high at 82% and 0.79 respectively. The inspection of the results shows a better detection of trees objects for this image.

In Fig 5.5 (B), we show the outcome of post processing the result of pixel based MRF-SRM using the Sentinel-2 multispectral image. Based on visual inspection, the result looks similar to Fig 5.4 (B). However, because of the segmentation applied, some of the smaller *trees* have disappeared and are merged into larger homogenous segments. Also, both big and small trees as well as other classes like *rock*, *baresoil_1*, *baresoil_2*, and *shrub* were also well identified.

Table 5.9 shows the results for all three accuracy measures (existential, extensional and positional) for validating the detected tree objects in Sentinel-2 multispectral image. A total of 1214 (TP) individual tree crown objects were successfully detected out of 1787 reference objects, while a total of 573 (FN) were not detected out of the reference objects. The reason for missing some tree objects which were not detected in the reference could be because those trees are smaller in size. Again, considering the temporal variation between detection image (Fig 3.2) and validation image (Fig 3.3), there is a possibility that those trees has been cut down. The positional accuracy between the detected tree objects and reference tree objects is 6.504 m. This is less than one pixel considering the spatial resolution of the input image which is 10 m. The table also shows the results of local area tree difference, overestimation, underestimation and total detection error respectively. Recall that the reference tree objects (R_i) is overestimated if $\Delta A_i > 0$ and underestimated if $\Delta A_i < 0$. Therefore the reference tree object (R_i) is underestimated here because the local area tree difference is -0.5061 which is less than zero. The first reason for underestimation may be because of the sun illumination angle during image acquisition and considering that trees are elevated objects; these can cause occlusion casting shadow and misplacement of the position of the trees. This may be verified by visual interpretation of the results (Fig 5.6). The second reason is that most pixel reflectance value from coarse resolution images

forms a spectral mixture with background classes. This produces the mixed pixel effect which might cause some misplacement in the actual position of the trees. The third reason is that we are dealing with vague objects here whose boundaries are not specific. To distinguish these boundaries with spectral classifiers is challenging. Finally, considering the temporal variation between detection image (Fig 3.2) and validation image (Fig 3.3), there is a possibility that the tree crowns has changed with respect to time leading to underestimation. The total detection error at 0.4669 is also greater than zero. This also signifies that the reference tree object (R_i) is underestimated. Finally, the detection accuracy is still acceptable at 68%. This indicates good percentage detection between the reference tree object and the detected tree crown object.

6.2.2. Results of Worldview-3 multispectral image

We also applied the maximum likelihood assumption for training our classifier (2.3). However from (Table 5.10), we see that classes' *tree*, *field_1* and *field_2* had sufficient number of training samples, while other classes classes' *road*, *built_up*, *field_3*, *baresoil* and *shrub* has insufficient training samples. The insufficient number of training samples is due to the large configuration of image space and spectral dissimilarity between these classes. As a result it was also challenging for the classifier to clearly distinguish one spectral class from another and also affected all the results of our classification applied on this image. However, because the resolution of this image was finer, spectral details could still be distinguished.

A second challenge faced here was the large class variance (Fig 5.7). The implication of this is that we have spectral confusion in classes that overlap leading to some misclassifications (Fig 5.9). This general limitation with spectral classifiers was addressed using MRF approach.

Again, the choice of neighbourhood system is another point of emphasis as it determines the surrounding pixels that will be taken into consideration during analysis. We also applied the first order neighbourhood system for both classification and segmentation (Fig 4.2). This considers only four pixels which shares sides with a given pixel and has effect on the objects detected. We couldn't consider the second order neighbourhood system because of time constrains. It will be wise to see if this improves our detection of tree objects for this image.

In Fig 5.9 (B), we show the result of maximum likelihood classification using the Worldview-3 multispectral image subset. Based on visual inspection, the result looks noisy. However, most classes can still be distinguished and trees were identified. The noise which exists mainly between classes' *tree*, *shrubs* and *field_1* is because these classes are spectrally similar and challenging to distinguish with spectral classifiers. From (Table 5.10), we see that some classes have insufficient number of training samples. Because this classifier assigns pixels to the class with highest probability of membership, it suggests that most of the spectral classes with insufficient training samples were assigned to a wrong class. As a result, the identification of trees with maximum likelihood result alone using this image is not satisfactory. Table 5.16 shows the confusion matrix of the MLC classification with various accuracy measures. The overall accuracy of the classification is 96.3% while the k value is 0.844.

In Fig 5.10 (B), we show the result of pixel based MRF of the Worldview-3 multispectral image subset. Based on visual inspection, the result looks smooth and shows a better representation of the input image (Fig 5.10) than the maximum likelihood result (Fig 5.9). The contextual approach of MRF helped to address the limitation of spectral information available in finer resolution images. Both big and small trees were well identified and the noise was reduced. The smoothness of the result is because of the incorporation of λ in the solution (4.17). The inspection of the results shows a better detection of trees objects for this image. Table 5.17 shows the confusion matrix of the MLC classification with various accuracy measures. The overall accuracy of the classification is 90% while the k value is 0.87.

In Fig 5.11 (B), we show the outcome of post processing the result of pixel based MRF classification using the Worldview-3 multispectral image. Again, based on visual inspection, the result looks similar to Fig 5.9 (B). Also, because of the segmentation applied, some of the smaller *trees* have disappeared and are merged into larger homogenous segments. Also, both big and small trees as well as other classes were also well identified.

Table 5.18 shows the results for all three accuracy measures (existential, extensional and positional) for validating the detected tree objects using Worldview-3 multispectral image. A total of 1772 (TP) individual tree crown objects were successfully detected out of 1787 reference objects, while a total of 15 (FP) were not detected out of the reference objects. The positional accuracy between the detected tree objects and reference tree objects is 1.714 m. This is less than one pixel considering the spatial resolution of the input image which is 2 m. The table also shows the results of the local area tree difference, overestimation error, underestimation error and total detection error respectively. Recall that the reference tree objects (R_i) is overestimated if $\Delta A_i > 0$ and underestimated if $\Delta A_i < 0$. Therefore the reference tree object (R_i) is underestimated here because the local area tree difference is -0.1314 which is less than zero. The reason for underestimation may be because of the sun illumination angle during image acquisition and considering that trees are elevated objects; these can cause occlusion casting shadow and misplacement of the position of the trees. This may be verified by visual interpretation of the results (Fig 5.12). Another reason is that we are dealing with vague objects here whose boundaries are not specific. To distinguish these boundaries with spectral classifiers is challenging. The total detection error at 0.3179 is also greater than zero. This also signifies that the reference tree object (R_i) is underestimated. Finally, the detection accuracy is much higher here at 99% and implies very high percentage detection between the reference tree object and the detected tree crown object. This agrees with the high work of Mahour et al., (2016) who also recorded high detection rate between reference object and detected tree crown object using finer resolution multispectral image .

6.2.3. Comparison of detected tree objects

In Fig 5.6, we show the accuracy map after post processing the output of pixel based MRF-SRM using the Sentinel-2 multispectral image. Fig 5.6 (B) shows overestimated, underestimated and agreement between the identified tree crowns and the reference tree crown overlaid on the Sentinel-2 multispectral image. Fig 5.6 (C) shows only the overestimated, underestimated and agreement between the identified tree crowns and the reference tree crown. Likewise in Fig 5.12, we show the accuracy map after post processing the output of pixel based MRF using the Worldview-3 multispectral image. Fig 5.12 (B) shows overestimated, underestimated and agreement between the identified tree crowns and the reference tree crown overlaid on the Worldview-3 multispectral image. Fig 5.12 (C) shows only overestimated, underestimated and agreement between the identified tree crowns and the reference tree crown. In both images, we could see that there are agreements in some area, overestimation and underestimation respectively. However, the major difference is that the identified tree crowns correctly matched the reference tree object in almost all areas for the Worldview-3 multispectral image than the Sentinel-2 multispectral image which had a lower match. This shows the variation in detection accuracy and other accuracy measure applied (Table 5.19).

In Fig 5.13, we show the scatter plot of overestimation and underestimation of both (A) Sentinel-2 and (B) Worldview-3 multispectral images respectively. The scatter plot of Fig 5.13 (A) shows that most of the overestimated and underestimated values cluster (agree) around 0. The range of value between the lowest and highest overestimation error is very large (0 - 80). This is confirmed by the high overestimation error (Table 5.9). Also, the range of value between the lowest and highest underestimation error is smaller (0 - 23). This is confirmed by the lower underestimation error (Table 5.9). The scatter plot of Fig 5.12 (B) is similar to that of Fig 5.12 (A). The difference here is that we have a lower range of value for overestimation (0 - 90) and a higher value of underestimation (0 - 25). This variation in the range of values for overestimation and underestimation can be linked to the detectability of tree objects at different scales with respective spatial

resolutions of the images. Also, the difference in their respective value ranges confirms the overestimation and underestimation errors respectively (Table 5.18).

In Table 5.19, we compare the total detection, positional accuracy, local area difference, overestimated, underestimated and total detection error respectively for both images. The positional accuracy between the detected tree objects and reference tree objects are 6.504 for Sentinel-2 multispectral image and 1.714 for Worldview-3 multispectral image respectively. This is less than one pixel considering the spatial resolution of the input images which is 10 m and 2 m respectively. It signifies that both images have a good positional accuracy. We also see that the Worldview-3 multispectral image has a higher detection, a lower local area tree difference and a lower total detection error. However, both images show that the reference tree object (R_i) is underestimated because their respective local area tree difference at -0.5061 for Sentinel-2 multispectral image and -0.1314 for Worldview-3 multispectral image are both less than zero. Finally, both images have high percentage detection accuracy at 68% for Sentinel-2 multispectral image and 99% for Worldview-3 multispectral image respectively. Based on the above experiments and analysis conducted, we conclude that both images are suitable for tree object detection using super resolution mapping in HR scenario.

7. CONCLUSIONS AND RECOMMENDATIONS

7.1. Conclusions

- How can we apply SRM in HR scenario to detect tree objects at different scale?
With respect to the size of individual tree object, we applied MRF based SRM on both images and at different scales. The goal of incorporating MRF is to maximize the contextual dependency between neighbouring pixels, while SRM reduced spatial resolution limitation for coarse images.
- How can one determine the most appropriate parameters for optimizing the SRM model for each scale?
We determined the most appropriate parameter for optimizing the MRF based SRM model for each scale by conducting a systematic experiment and analysed the results (repeated experiments with equal number of runs). This was done for both images.
- How can we evaluate and quantify the detected tree objects?
We evaluated and quantified the detected tree objects using object based image analysis procedure. The spatial join overlay tool in ArcGIS was used to compute the area of tree crown objects.
- How should the assessment of the results be performed?
We performed both pixel based and object based accuracy assessment. Pixel based assessment was done using error matrix to determine the value of k and overall accuracy respectively. Object based accuracy assessment was done by applying three accuracy measures; (a) existential (b) extensional (c) positional accuracy measures.

7.2. Recommendations

- In this research, insitu data was not used because of the scope of the work and limited time available. However, insitu data or field information is required to improve the quality of training samples and proper verification of test samples for better accuracy assessment. I therefore recommend that these should be added to improve the quality of the analysis.
- Sun illumination angle during image acquisition leads to occlusion, thereby casting shadow and misplacement of the position of the trees. Again considering that trees are elevated objects, it is recommended to add a digital surface model (DSM) to this research. This can be an interesting way to improve the quality of detection and analysis of results.
- For transferability of our method to urban or developed cities using Sentinel-2 multispectral image, I recommend that a context-sensitivity approach should be considered as part of further studies. This approach applies some rules to model the spatial distribution of tree objects taking cognisance of the urban elements in the surrounding. This might improve the identification of tree objects and quality of results.

LIST OF REFERENCES

- Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., & Süsstrunk, S. (2012). SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *34*(11), 2274–2281. <https://doi.org/10.1109/TPAMI.2012.120>.
- Adler, A. D., Mur-, D., & Murdoch, M. D. (2016). rgl : 3D Visualization Using OpenGL. R package version 0.95.1441. <https://CRAN.R-project.org/package=rgl>.
- Ardila, J. P., Bijker, W., Tolpekin, V. A., & Stein, A. (2012). Context-sensitive extraction of tree crown objects in urban areas using VHR satellite images. *International Journal of Applied Earth Observation and Geoinformation*, *15*(1), 57–69. <https://doi.org/10.1016/j.jag.2011.06.005>.
- Ardila, J. P., Tolpekin, V. A., Bijker, W., & Stein, A. (2011). Markov-random-field-based super-resolution mapping for identification of urban trees in VHR images. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*(6), 762–775. <https://doi.org/10.1016/j.isprsjprs.2011.08.002>
- Atkinson, P. M. (2004). Resolution Manipulation and Sub-Pixel Mapping. In: S.M. de Jong & Van der Meer (Eds), *Remote Sensing Image Analysis: including the spatial domain*. KluwerAcademics, Dordrecht, (pp. 51–70). https://doi.org/10.1007/978-1-4020-2560-0_4.
- Atkinson, P. M. (2005). Sub-pixel target mapping from soft-classified, remotely sensed imagery. *Photogrammetric Engineering and Remote Sensing*, *71*(7), 839–846. Retrieved from http://info.asprs.org/publications/pers/2005journal/july/2005_july_839-846.pdf.
- Atkinson, P. M. (2009). Issues of uncertainty in super-resolution mapping and their implications for the design of an inter-comparison study. *International Journal of Remote Sensing*, *30*(20), 5293–5308. <https://doi.org/10.1080/01431160903131034>.
- Baglioni, M., Antônio, J., Macêdo, F. De, & Renso, C. (2009). *Advances in GIScience*. (B. Thomas, H. Geoffrey, & S. Lang, Eds.), *Notes*. Berlin Heidelberg: <https://doi.org/10.1007/978-3-642-00318-9>
- Bins, L. S., Fonseca, L. M. G., Erthal, G. J., & Li, F. M. (1996). Satellite Imagery Segmentation: a region growing approach. *Anais VIII Simposia Brasileiro de Sensoriamento Remoto, Salvador, Brasil, 14-19 Abril 1996, INPE*, 677–680.
- Bivand, R., Keitt, T., Rowlingson, B., Pebesma, E., Sumner, M., Hijmans, R., & Rouault, E. (2016). rgdal: Bindings for the Geospatial Data Abstraction Library: R Package version 1.1-10. <https://CRAN.R-project.org/package=rgdal>.
- Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, *65*(1), 2–16. <https://doi.org/10.1016/j.isprsjprs.2009.06.004>.
- Blaschke, T., Hay, G. J., Kelly, M., Lang, S., Hofmann, P., Addink, E., ... Tiede, D. (2014). Geographic Object-Based Image Analysis - Towards a new paradigm. *ISPRS Journal of Photogrammetry and Remote Sensing*, *87*, 180–191. <https://doi.org/10.1016/j.isprsjprs.2013.09.014>.
- Clinton, N., Holt, A., Scarborough, J., Yan, L., & Gong, P. (2010). Accuracy assessment measures for object-based image segmentation goodness. *Photogrammetric Engineering & Remote Sensing*, *76*(3), 289–299. <https://doi.org/10.14358/PERS.76.3.289>.
- Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, *37*(1), 35–46. [https://doi.org/10.1016/0034-4257\(91\)90048-B](https://doi.org/10.1016/0034-4257(91)90048-B).
- Debats, S. R., Luo, D., Estes, L. D., Fuchs, T. J., & Caylor, K. K. (2016). A generalized computer vision approach to mapping crop fields in heterogeneous agricultural landscapes. *Remote Sensing of Environment*, *179*, 210–221. <https://doi.org/10.1016/j.rse.2016.03.010>.

- Eddelbuettel, D. (2013). Seamless R and C++ Integration with Rcpp (Google eBook). *Journal Of Statistical Software*, 40(8), 248. <https://doi.org/10.1007/978-1-4614-6868-4>.
- Fan, J., Zeng, G., Body, M., & Hacid, M. S. (2005). *Seeded region growing: An extensive and comparative study*. *Pattern Recognition Letters* (Vol. 26). <https://doi.org/10.1016/j.patrec.2004.10.010>.
- Fleet, D. J., & Jepson, A. D. (2007). Image Segmentation. *Computer Vision, Graphics, and Image Processing*. <https://doi.org/10.1016/B978-012264841-0/50037-8>.
- Foody, G. M., & Atkinson, P. M. (2002). *Uncertainty in Remote Sensing and GIS*. *Uncertainty in Remote Sensing and GIS*. J. Wiley. <https://doi.org/10.1002/0470035269>.
- Frampton, W. J., Dash, J., Watmough, G., & Milton, E. J. (2013). Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 82, 83–92. <https://doi.org/10.1016/j.isprsjprs.2013.04.007>.
- Geman, S., & Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6(6), 721–741. <https://doi.org/10.1109/TPAMI.1984.4767596>.
- Hay, G. J., & Castilla, G. (2006). Object-based image analysis: strengths, weaknesses, opportunities and threats (SWOT). *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 36, 4. Retrieved from http://www.isprs.org/proceedings/XXXVI/4-C42/Papers/01_Opening Session/ OBIA2006_Hay_Castilla.pdf.
- Hay, G. J., Castilla, G., Wulder, M. A., & 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. <https://doi.org/10.1016/j.jag.2005.06.005>.
- Jin, H., Mountrakis, G., & Li, P. (2012). A super-resolution mapping method using local indicator variograms. *International Journal of Remote Sensing*, (August), 37–41. <https://doi.org/10.1080/01431161.2012.702234>.
- Jong, S. M. de, & Meer, F. D. Van Der. (2006). *Remote Sensing Image Analysis: Including the Spatial Domain*. (S. M. De Jong & F. D. Van der Meer, Eds.), *The Photogrammetric Record* (Vol. 5). Dordrecht: Springer Netherlands. <https://doi.org/10.1017/S0032247400010123>.
- Karatzoglou, A., & Feinerer, I. (2010). Kernel-based machine learning for fast text mining in R. *Computational Statistics and Data Analysis*, 54(2), 290–297. <https://doi.org/10.1016/j.csda.2009.09.023>.
- Kasetkasem, T., Arora, M. K., & Varshney, P. K. (2005). Super-resolution land cover mapping using a Markov random field based approach. *Remote Sensing of Environment*, 96(3–4), 302–314. <https://doi.org/10.1016/j.rse.2005.02.006>.
- Kirkpatrick, S., Gelatt, C. D. J., & Vecchi, M. P. (1983). Optimization by Simulated Annealing. *Science*, 220(4598), 671–680. <https://doi.org/10.1126/science.220.4598.671>.
- Ladický, L., Russell, C., Kohli, P., & Torr, P. H. S. (2014). Associative hierarchical random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(6), 1056–1077. <https://doi.org/10.1109/TPAMI.2013.165>.
- Li, X., Ling, F., Foody, G. M., & Du, Y. (2016). Improving super-resolution mapping through combining multiple super-resolution land-cover maps. *International Journal of Remote Sensing*, 37(10), 2415–2432. <https://doi.org/10.1080/01431161.2016.1148288>.
- Lucieer, A., & Stein, A. (2002). Existential uncertainty of spatial objects segmented from satellite sensor imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 40(11), 2518–2521. <https://doi.org/10.1109/TGRS.2002.805072>.

- Mahour, M., Tolpekin, V., & Stein, A. (2016). Tree detection in orchards from VHR satellite images using scale-space theory, *10004*, 1–17. <https://doi.org/10.1117/12.2241529>.
- Maselli, F., Conese, C., & Petkov, L. (1994). Use of probability entropy for the estimation and graphical representation of the accuracy of maximum likelihood classifications. *ISPRS Journal of Photogrammetry and Remote Sensing*, *49*(2), 13–20. [https://doi.org/10.1016/0924-2716\(94\)90062-0](https://doi.org/10.1016/0924-2716(94)90062-0).
- Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., Chang, C., & Lin, C. (2015). e1071:Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. R package version 1.6-7. <https://CRAN.R-project.org/package=e1071>.
- Richards, J. a, & Jia, X. (2006). *Remote Sensing Digital Image Analysis. Methods*. <https://doi.org/10.1007/3-540-29711-1>.
- Sriwilai, S., Kasetkasem, T., Chanwimaluang, T., Srinark, T., & Isshiki, T. (2013). A super-resolution mapping algorithm based on the level set method. *2013 10th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, 1–6. <https://doi.org/10.1109/ECTICon.2013.6559559>.
- Strahler, A. H. (1980). The use of prior probabilities in maximum likelihood classification of remotely sensed data. *Remote Sensing of Environment*, *10*(2), 135–163. [https://doi.org/10.1016/0034-4257\(80\)90011-5](https://doi.org/10.1016/0034-4257(80)90011-5).
- Tolpekin, V. A., Ardila, J. P., & Bijker, W. (2010). Super-Resolution Mapping for Extraction of Urban Tree Crown Objects From Vhr Satellite Images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVIII-4/C7, 3*, 1–6.
- Tolpekin, V. A., & Stein, A. (2009). Quantification of the effects of land-cover-class spectral separability on the accuracy of markov-random-field-based superresolution mapping. *IEEE Transactions on Geoscience and Remote Sensing*, *47*(9), 3283–3297. <https://doi.org/10.1109/TGRS.2009.2019126>.
- Vayssières, J.-F., Sanogo, F., & Noussourou, M. (2007). Inventory of the fruit fly species (Diptera: Tephritidae) linked to the mango tree in Mali and tests of integrated control. *Fruits*, *62*(5), 329–341. <https://doi.org/10.1051/fruits:2007029>.
- Venables, W. N., & Ripley, B. D. (2003). *Modern Applied Statistics With S*. Fourth Edition. Springer, New York. *Technometrics*, *45*(1), 111–111. <https://doi.org/10.1198/tech.2003.s33>.
- Wolf, B. M., & Heipke, C. (2007). Automatic extraction and delineation of single trees from remote sensing data. *Machine Vision and Applications*, *18*(5), 317–330. <https://doi.org/10.1007/s00138-006-0064-9>
- World Bank. (2013). *Unlocking Africa's Agricultural Potential: An Action Agenda for Transformation*. *Sustainable Development Series*. Retrieved from <https://openknowledge.worldbank.org/bitstream/handle/10986/16624/769900WP0SDS0A00Box374393B00PUBLIC0.pdf?sequence=1>