

Cadastral mapping with flying sensors and satellite imagery in Zambezia, Mozambique



REPORT

220

CLIENT **Terra Firma**

AUTHORS **Jonna van Opstal
Jack Beard**

DATE **February 2021**

Cadastral mapping with flying sensors and satellite imagery in Zambezia, Mozambique

FutureWater Report 220

Client

Terra Firma

Authors

Jonna van Opstal – Senior Water Productivity Expert (j.vanopstal@futurewater.nl)

Jack Beard – Hydrologist and Programming Expert (j.beard@futurewater.nl)

Date

February 2021

FutureWater

ADDRESS

**FutureWater B.V.
Costerweg 1V
6702 AA Wageningen
The Netherlands**

TELEPHONE

+31 317 460 050

WEBSITE

www.futurewater.eu

Content

1	Introduction	5
1.1	Background	5
1.2	Project description	5
1.3	Reading guide	5
2	Study area and imagery acquisition	6
2.1	Location of flights	6
2.2	Flying sensor imagery	6
2.2.1	Flying Sensor Operators	6
2.2.2	Technical description	7
2.2.3	Overview flights	7
2.2.4	Imagery processing	8
2.3	Satellite imagery	8
2.3.1	Sentinel imagery	8
2.3.2	Planet imagery	9
2.3.3	Digital Globe imagery	10
2.3.4	Google Earth imagery	10
3	Boundary delineation methodology	12
3.1	Image pre-processing	12
3.2	R Packages	15
3.3	ilastik	18
3.4	QGIS Segmentation (GRASS)	20
3.5	Manual boundary delineation	20
4	Field boundary results	21
4.1	Flying Sensors	21
4.1.1	R packages	21
4.1.2	QGIS GRASS segmentation	23
4.2	Satellite imagery	24
4.3	Discussion	25
4.3.1	Segmentation methodologies	25
4.3.2	Satellite and flying sensor data	27
5	Concluding remarks	28

Tables

Table 2-1 Number of flights and total area in hectares for each flight location	7
Table 4-1 Overview of the pros and cons of the different methodologies as used for flying sensor imagery.....	25
Table 4-2 Overview of the pros and cons of using flying sensor and satellite imagery for boundary delineation	27

Figures

Figure 2-1 Location of the three flight areas close to the city Quelimane in Mozambique	6
Figure 2-2 ThirdEye flying sensor operators at location with Terra Firma representative.....	7
Figure 2-3 Photo of the Mavic Pro in action.....	7
Figure 2-4 Results ortho-mosaics of the flying sensor imagery	8
Figure 2-5 Sentinel 2 image (false colour) for 26 th December 2020 indicating (pinpoint) the study location	9
Figure 2-6 Planet data platform indicating the study area and data availability	9
Figure 2-7 Digital Globe data platform indicating data availability for Area 2.....	10
Figure 2-8 Google Earth Satellite image for the study area	11
Figure 2-9 Detailed zoomed in image of Google Earth image for Area 1 (left) and Area 2 (right).....	11
Figure 3-1. Kmeans showing clear “tree” class (number 2 in this case)	13
Figure 3-2. Before and after removal of trees from image	13
Figure 3-3. Green Leaf Index of image	14
Figure 3-4. “Linear stretch” composite of image (right) compared with original image (left).	14
Figure 3-5. Optimised Kmeans: showing 4 classes used on the masked GLI image	15
Figure 3-6. Optimised canny edge approach applied to linear stretch RGB image with trees masked.	16
Figure 3-7. Best results from superpixels approach applied to linear stretch RGB image with trees masked.....	17
Figure 3-8. Best results from contours approach applied to GLI image with trees masked.	18
Figure 3-9 Image segmentation workflow from previous FutureWater report	18
Figure 3-10 Training in ilastik indicating various labels for different classes.....	19
Figure 3-11 Pixel classification result of ilastik using four labels.....	19
Figure 3-12 Results of the GRASS QGIS segmentation for two different threshold settings	20
Figure 4-1 Flying sensor image of selected block in Area 3	21
Figure 4-2 Results of kmeans step using the Green Leaf Index (GLI)	22
Figure 4-3 Results for the canny edges algorithm	22
Figure 4-4 Results for the super pixels algorithm	23
Figure 4-5 Results of GRASS segmentation algorithm.....	24
Figure 4-6 Detailed zoom of the segmentation results from GRASS (threshold = 0.1), converted to polygons	24
Figure 4-7 Results comparing the Flying Sensor (drone) imagery with satellite Sentinel and Google imagery.....	25
Figure 4-8 Example of classification for Bekaa Valley (Lebanon) using multi-date imagery. Top: layer stack of three NDVI images at three different dates during the growing season. Bottom: Google satellite image.....	26
Figure 4-9 Comparison of flying sensor image (resized to 0.4m pixel resolution) and Google Satellite	27

1 Introduction

1.1 Background

Cadastre mapping is the practice of determining the field boundaries, which are then used for recording the land ownership of the parcel of land. Remote sensing is a powerful tool for supporting cadastre mapping, because it covers large areas and reduces the labour required for performing a land survey. High resolution satellite imagery and flying sensor (drone) imagery are suitable to be used for cadastre mapping. These are then used to manually determine the field boundaries. Alternatively, algorithms are rapidly developing and adopted for classification or segmentation of the imagery, which automatically determines field boundaries in varying degrees of accuracy.

The practice of using remote sensing imagery is becoming more widespread. However, the suitability of satellite or flying sensor imagery needs to be evaluated by location. Satellite imagery is available at different price ranges and is fixed in terms of spatial and temporal resolution. Flying sensor imagery is adaptable and can be deployed at any requested time. The suitability of these remote sensing approaches is piloted in this study for a small-scale agricultural area in Mozambique.

1.2 Project description

Terra Firma is an organization in Mozambique with the task to map and document land rights. In this project, they hired FutureWater¹ (Wageningen, The Netherlands) and ThirdEye Limitada² (Chimoio, Mozambique) to acquire flying sensor imagery over a pilot area near Quelimane, Mozambique. The objective of this pilot is to determine the suitability of using flying sensor imagery for cadastre mapping in an area of small-scale agriculture in Mozambique. The flying sensor imagery was acquired over the period of a few days in December 2020, for a total area of 1,120 hectares.

This imagery was used as input for various algorithms that can be suitable for classification and segmentation. This study shows some initial results of using flying sensor imagery in combination with these algorithms. In addition, comparison is made with satellite imagery to indicate the differences in results.

1.3 Reading guide

In chapter two a description is provided of the study area of this project. In addition, a technical description is provided of the flying sensor technology used and the flights acquired for this study area. Chapter three provides an overview of the various algorithms that are applicable for field boundary mapping. These are organized into R packages and readily available software programs such as ilastik and GRASS (in QGIS). Chapter four discusses the suitability of these algorithms over a selected block in the study area to indicate the applicability of these algorithms for small-scale agriculture. In addition, satellite imagery is added to this analysis to indicate the differences that exist in results between flying sensor and satellite imagery. The last chapter provides some conclusions and recommendations following the findings from this pilot study.

¹ <https://www.futurewater.eu/>

² <http://www.thirdeyewater.com/>

2 Study area and imagery acquisition

2.1 Location of flights

A study area was selected by TerraFirma in the vicinity of the city Quelimane located at the Eastern coast of Mozambique in Zambezia province, see Figure 2-1. The study area was divided in three locations for performing the flying sensor flights. The three locations together have a total area of 1,120 ha. The study area was flown during the rainfed season; therefore, the main crop being cultivated is maize. The flights were performed early in the growing season, so the vegetation cover varied between fields. Small villages and settlements are also part of the study area and a few roads and tree rows. Overall, the study area is a suitable pilot location for studying cadastre mapping over small-scale agriculture.

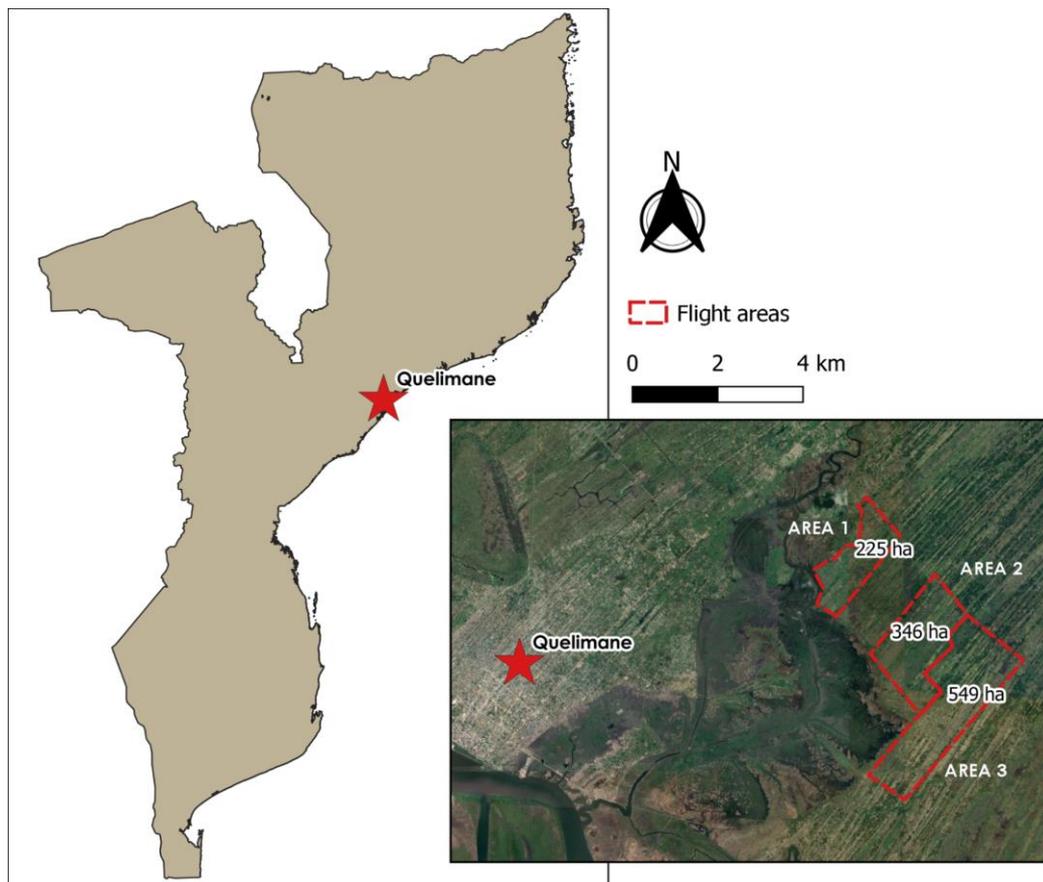


Figure 2-1 Location of the three flight areas close to the city Quelimane in Mozambique

2.2 Flying sensor imagery

2.2.1 Flying Sensor Operators

ThirdEye Limitada located in Chimoio, Mozambique performed the flying sensor flights at the location with supervision from experts of HiView¹ and FutureWater. Due to the Covid-19 travel restrictions, the supervision of the experts was provided by online meetings in preparation of the flights and on location. The operators team performed the flights in December 2020, see Figure 2-2 for a photo of the activities.

¹ <https://www.hiview.nl/>



Figure 2-2 ThirdEye flying sensor operators at location with Terra Firma representative

2.2.2 Technical description

The Flying Sensor equipment used for this project is a Mavic Pro drone with a DJI camera. Figure 2-3 shows a photo of the Flying Sensor. The camera makes RGB (red-green-blue) images, similar to visual images as seen with the human eye.

Advantages of using Flying Sensors are the flexibility for imagery capture and the high-spatial resolution of the acquired imagery. The flying sensors can make flights when required at the desired intervals. The spatial resolution of the imagery is 4-8 cm, providing sufficient detail to capture the spatial variation of small holder agriculture that is typical in several regions of Mozambique



Figure 2-3 Photo of the Mavic Pro in action

2.2.3 Overview flights

The flying sensor flights were performed early December 2020, which is the start of the rainfed growing season. The conditions were generally suitable for flying on most dates, except the first day when rain showers delayed the flights somewhat. An overview of the flights performed for the three different areas (as indicated in Figure 2-1) is presented in Table 2-1. The number of flights per location and per date are listed. At the end of each day of performing flights, a quality check was performed on the images acquired to assess if a second flight is required. On the last day, 11th December 2020, extra flights were performed, for the locations that needed an extra flight.

Table 2-1 Number of flights and total area in hectares for each flight location

	Area 1	Area 2	Area 3
7 th December 2020	8 flights		
8 th December 2020	2 flights	11 flights	
9 th December 2020		5 flights	7 flights
10 th December 2020			15 flights
11 th December 2020 (extra flights)		(3 flights)	(1 flight)
Total flights taken	10 flights	16 flights	22 flights
Total area covered	225 ha	346 ha	549 ha

Flights were performed at 120m height which gave 4 cm pixel resolution. Flights were not performed very early or late in the day to reduce the occurrence of shadow in the images with the low sun elevation. The software for the flight plans were done with Pix4D with settings at 60% speed and 80% overlap.

2.2.4 Imagery processing

The single images acquired by the flying sensors during the flights, were stitched to create georeferenced ortho-mosaics. This processing step was performed using the Agisoft Metashape software and the computing power of a 'supercomputer' at the office in The Netherlands. A stitch was created for all the three areas individually, thus creating one image that stitched all flights together for that location. For example, Area 1 consisted of 10 flights, where the images of all flights were given as input for providing one full ortho-mosaic for Area 1.

The results of the imagery processing were delivered by SD card to TerraFirma to conserve the high resolution quality. Figure 2-4 displays the results of the ortho-mosaics of the three locations. Some contrast is noticeable because several flights are stitched into one raster image. The detailed zoom in Figure 2-4 demonstrates the amount of detail that is achieved with flying sensor imagery. The individual fields and even furrows within fields are noticeable.



Figure 2-4 Results ortho-mosaics of the flying sensor imagery

2.3 Satellite imagery

2.3.1 Sentinel imagery

High resolution satellite imagery is available at different prices for this location. The imagery available by Sentinel 2¹ is for free and has a pixel resolution of 10m. The challenge for the rainfed season is the issue

¹ <https://apps.sentinel-hub.com/eo-browser/>

with cloud cover. An image was selected with the least amount of cloud cover for the study area, which also coincides reasonably well with the dates of the flying sensor flights: 26th December 2020. Postprocessing of the Sentinel imagery was performed with the SCP tool (Semi-Automatic Classification Plug-in) of QGIS¹, which translates the digital numbers of the imagery to reflectance values.

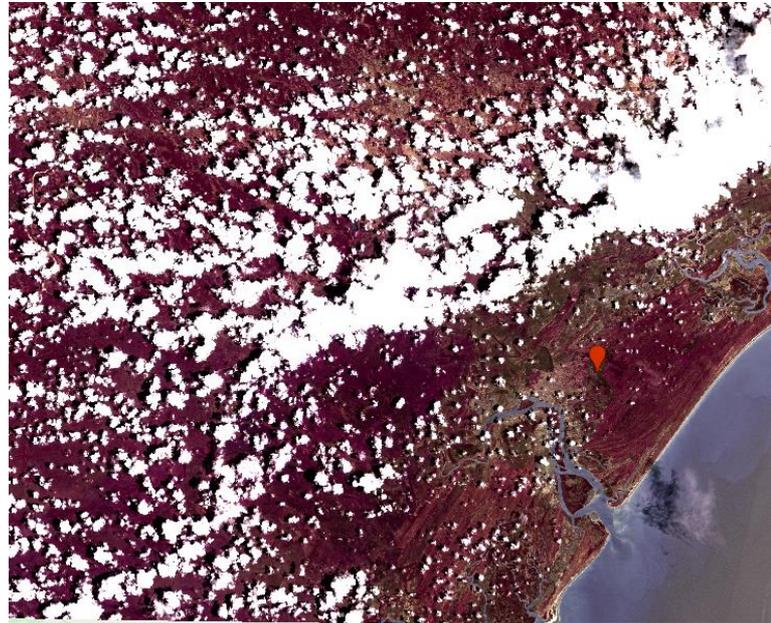


Figure 2-5 Sentinel 2 image (false colour) for 26th December 2020 indicating (pinpoint) the study location

2.3.2 Planet imagery

Planet² provides imagery at high pixel resolution, namely from 3-5 m (PlanetScope and RapidEye) or 0.5m (SkySat). These images are available for purchase through their data platform and are not used in this study. An example of the data available for the study location is shown in Figure 2-6.



Figure 2-6 Planet data platform indicating the study area and data availability

¹ <https://plugins.qgis.org/plugins/SemiAutomaticClassificationPlugin/>

² <https://www.planet.com/products/planet-imagery/>

2.3.3 Digital Globe imagery

Digital Globe¹ is another commercial platform for providing satellite data at high resolution. The highest resolution they offer is 30-60 cm. pixel resolution. An example image for the study area is shown in Figure 2-7. This imagery is available for purchasing and is not used in the further analysis of this pilot study.

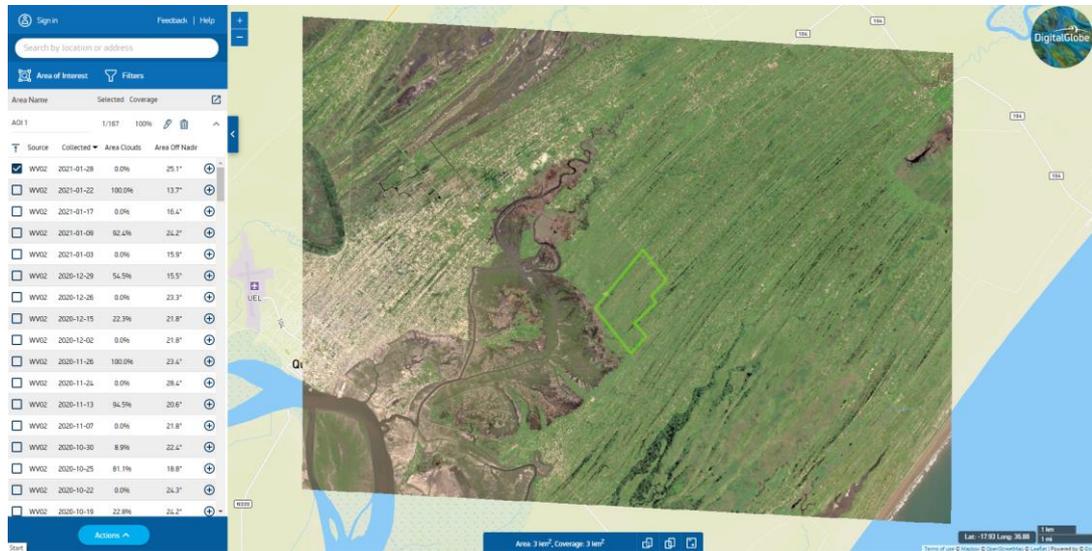


Figure 2-7 Digital Globe data platform indicating data availability for Area 2

2.3.4 Google Earth imagery

Google Earth imagery² is freely available through Google Earth platform or as basemap layer in QGIS. The imagery provided is at high pixel resolution and usually presents a certain date that is representative for that year or period. The Google Earth image is not downloadable as GeoTiff multi-band raster image, therefore cannot be used for further analysis in algorithms. An example of the Google Earth image for the study location is presented in Figure 2-8 and two detailed zoomed in images in Figure 2-9.

¹ <https://discover.digitalglobe.com/>

² <https://www.google.com/earth/>



Figure 2-8 Google Earth Satellite image for the study area



Figure 2-9 Detailed zoomed in image of Google Earth image for Area 1 (left) and Area 2 (right)

3 Boundary delineation methodology

Various algorithms can be used to delineate the field boundaries, which assists the cadastre mapping process. In the following sections, the various methods piloted in this study are presented. These are tested for a small excerpt of the full study area to demonstrate the effectiveness of the various methods. This “test” image or excerpt included the various land covers that are typical for the study area: small farm fields both with vegetation and bare soil, roads, trees, and settlements. The RGB image of this test image is shown in Figure 3-2 on the right.

3.1 Image pre-processing

Some pre-processing steps were deemed necessary to prevent the following:

- Smaller features or details were adding extra “noise” to images which complicates segmentation
- Inclusion of extraneous objects (e.g. trees, houses) which could be misclassified by segmentation algorithms

Additionally, the fact that RGB images are by nature split into 3 bands (red-green-blue) means that for some image processing algorithms, this needed to be simplified into one raster image for input. The following pre-processing steps were used to assist the segmentation or boundary delineation algorithms.

Resolution

The effects of resolution on the effectiveness of different methodologies was found to be substantial. Keeping the images at higher resolutions tended to pick up furrows in fields which was then isolated by the segmentation algorithms, hence, reducing the accuracy with which field boundaries were delineated. When aggregated to a lower resolution, this averaged out the smaller details in images and gave a clearer differentiation between fields. Images were finally aggregated from 4cm to 40cm resolution.

Removing Trees

It was found that trees have a disruptive effect on all the segmentation algorithms as they added clear features to be segmented and therefore took away from some of the variation between fields. Areas of the imagery that were found to contain trees were therefore masked (values set to NA) to eliminate this effect. Masking was done using a preliminary kmeans¹ approach (described below) to identify areas of trees, which were then buffered to a wider area to create the mask. This process and its results are shown in Figure 3-1 and Figure 3-2. Adding this extra step is usually referred to as a two-step classification method, thereby first distinguishing major classes and then continuing with a selected class for detailed classification.

¹ <https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/kmeans>

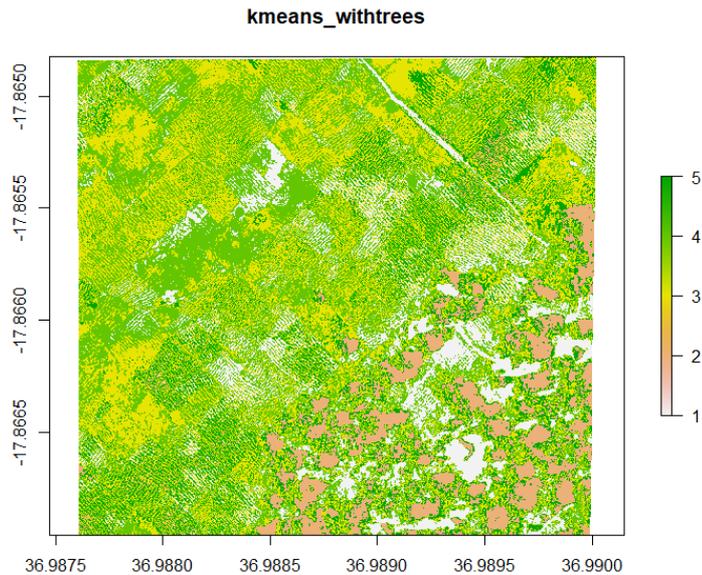


Figure 3-1. Kmeans showing clear “tree” class (number 2 in this case)



Figure 3-2. Before and after removal of trees from image

Combining bands

The kmeans algorithm described below only takes as input single band images. It was therefore necessary to either a) combine the 3 bands which make up each RGB image to a single raster image or b) take one of the three bands. After some trial and error, Green Leaf Index (GLI), as defined by Louhaichi et al (2001)¹, was decided to be the most effective way to combine bands, prioritizing the green band as the most important to conserve information within the image.

The Green Leaf Index (shown in Figure 3-3) was calculated as such:

$$GLI = \frac{2Green - Red - Blue}{2Green + Red + Blue}$$

¹ Louhaichi, Mounir & Borman, Michael & Johnson, Douglas. (2001). Spatially Located Platform and Aerial Photography for Documentation of Grazing Impacts on Wheat. Geocarto International. 16. 10.1080/10106040108542184.

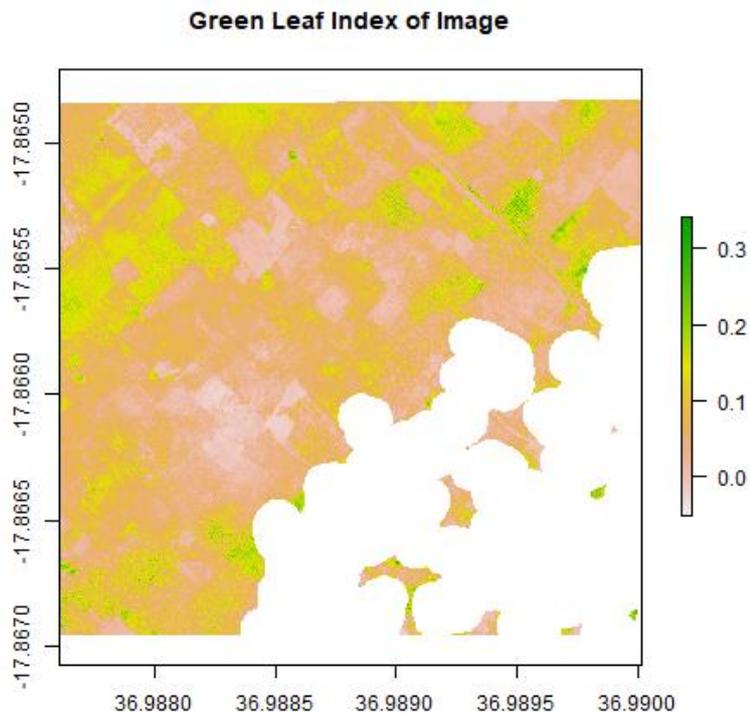


Figure 3-3. Green Leaf Index of image

Other R packages (Magick, ImageR) can take as input a 3-band image. The original image was therefore tested as input, but it was found that the contrast between fields could be increased in the image by reading it into R and using the “plotRGB” function from the raster package¹ with linear stretching applied to each band. This was recommended on a technical note on image segmentation in R, with the differences between the original image and the outputted “linear stretch” RGB shown in Figure 3-4.



Figure 3-4. “Linear stretch” composite of image (right) compared with original image (left).

¹ <https://www.rdocumentation.org/packages/raster/versions/3.4-5/topics/plotRGB>

3.2 R Packages

A variety of R packages were piloted to delineate field boundaries. R¹ is useful as a programming tool for this purpose as it has a variety of openly available, well documented packages for image processing and deals with spatial and raster data in a simple way.

Kmeans

The Kmeans methodology groups pixels of similar values to a given number of classes specified by the user. A kmeans algorithm from the “stats” package² was applied to the imagery to differentiate fields from one another by grouping them in a number of classes. This was applied to both single band images and the compiled GLI image. For this algorithm the best results were found using the following:

- Input image: GLI
- Resolution: 20cm
- Trees: Masked
- Parameterisation: Classes = 4

The results of this were fairly successful in terms of picking out some fields that contained significantly different values in terms of GLI (Figure 3-5). The weaknesses of this approach are likely that if similar crops are grown in adjacent fields, these will be grouped together rather than delineated as separate. In this way, the methodology relies more on different colours resulting from different crops or bare earth/panted distinctions. This approach could therefore be particularly useful to delineate fields growing a certain type of crop at a given stage in its growth cycle in which it appears quite distinct to other crop types.

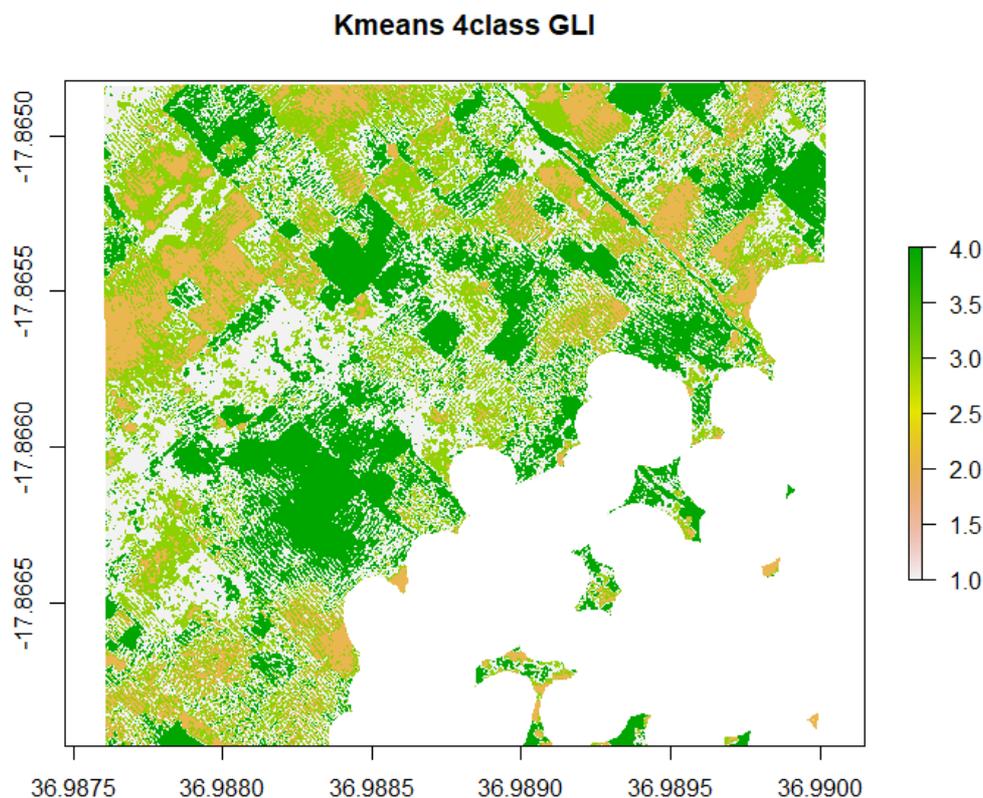


Figure 3-5. Optimised Kmeans: showing 4 classes used on the masked GLI image

¹ <https://cran.r-project.org/>

² <https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/kmeans>

Canny edges

Canny edge processing works via detecting sharp gradients in values in images. It is therefore highly useful in detecting clear edges or outlines in optical imagery. The algorithm implemented here comes from the “magick” package in R¹, which also gives an interface to allow users to arrive at an optimal parameterisation of the algorithm through some trial and error. This was applied to both “linear stretch” RGB images and the compiled GLI image. For this algorithm the best results were found using the following:

- Input image: Linear stretch RGB image
- Resolution: 40cm
- Trees: Masked
- Parameterisation: Radius = 4, Sigma = 1, Lower % = 10, upper % = 30

The results of this were clearly successful in picking out some field boundaries, especially between very different coloured fields (Figure 3-6). It also succeeds in picking out the path feature in one area of the image, suggesting it would also pick out linear features such as fencing or paths between fields well. The weaknesses of this approach are that it struggles to pick out fields which do not have clear boundaries or are of similar colour. It is therefore likely that if similar crops are grown in adjacent fields with no separating boundary, these will be grouped together rather than delineated as separate. This approach could therefore be particularly useful to delineate fields with clear boundary features between them or with very different crop types adjacent to each other.

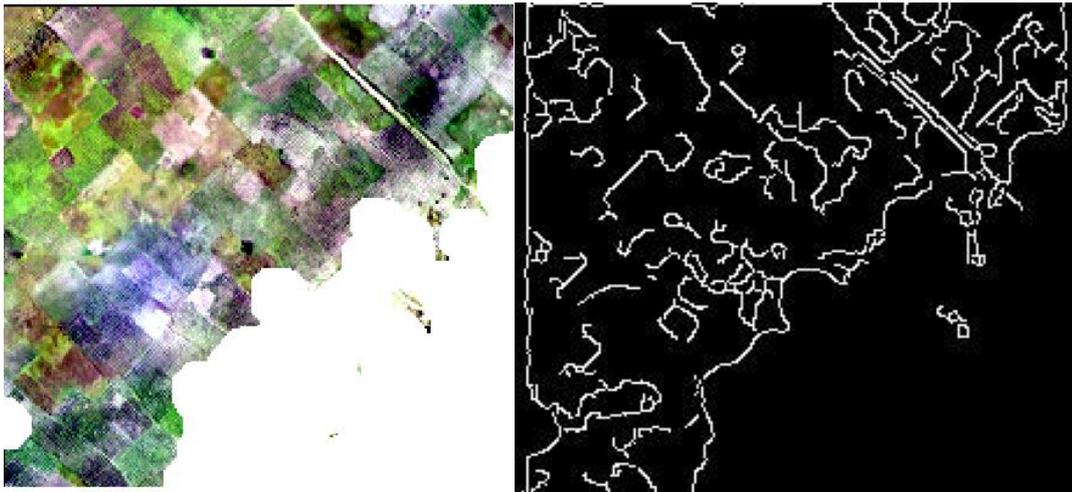


Figure 3-6. Optimised canny edge approach applied to linear stretch RGB image with trees masked.

Superpixels

The superpixels approach works via creating a mesh over the image of a given size (representing “superpixels”) then deforming this mesh to fit features within the image within the grid of this mesh. The algorithm implemented here comes from the “OpenImageR” package², with lots of guidance coming in the form of a technical note on image segmentation³. This was applied to both “linear stretch” RGB images and the compiled GLI image. For this algorithm the best results were found using the following:

- Input image: Linear stretch RGB image
- Resolution: 40cm
- Trees: Masked
- Parameterisation: Method = “slic”, Superpixel = 220

¹ https://rdr.io/cran/magickGUI/man/interactive_canny.html

² <https://www.rdocumentation.org/packages/OpenImageR/versions/1.1.7/topics/superpixels>

³ <https://www.r-bloggers.com/2020/03/analyzing-remote-sensing-data-using-image-segmentation/>

The results of implementing this algorithm were fairly successful when qualitatively compared with the original image (Figure 3-7). Some fields with clearer boundaries or a very different colour to surrounding fields were successfully picked out by the algorithm, but it is also clear that in other areas boundaries delineated did not correspond to those in the original image. The weaknesses of this approach are therefore that the mesh may create false field boundaries which, unless carefully reviewed, may be taken as accurate. This algorithm is likely to be most useful in areas with fairly homogenous field sizes and varied crop types or land covers.

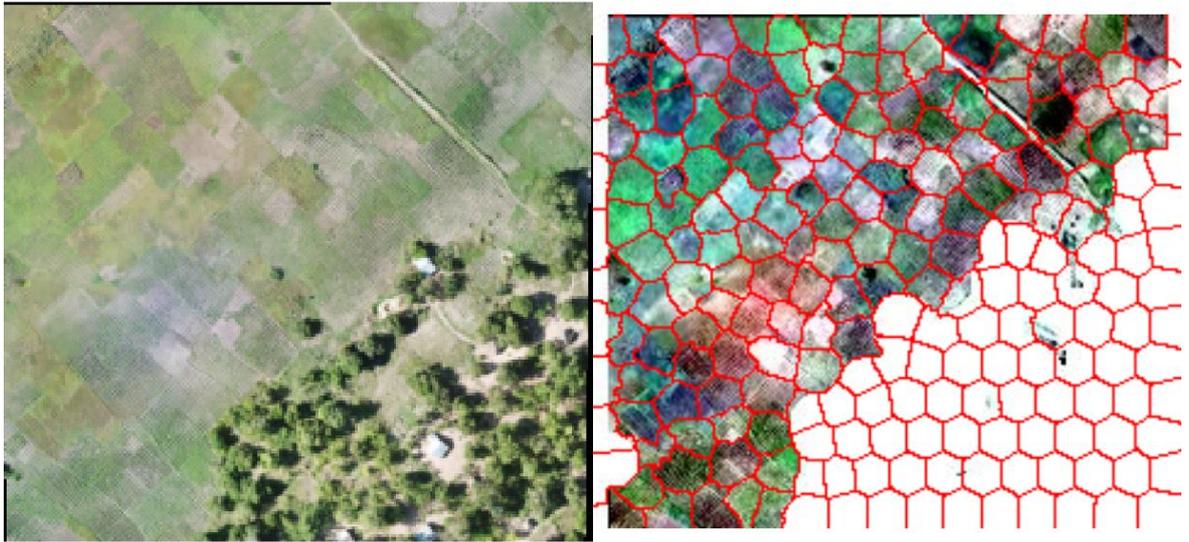


Figure 3-7. Best results from superpixels approach applied to linear stretch RGB image with trees masked.

Contours

The contours approach works via mapping contours around the values of an image, hence outlining areas of similar values. The algorithm implemented here comes from the “imager” package¹. This was applied to both “linear stretch” RGB images and the compiled GLI image, after first converting them to grayscale using the same package. For this algorithm the best results were found using the following:

- Input image: GLI image
- Resolution: 40cm
- Trees: Masked
- Parameterisation: nlevels = 2

The results of implementing this algorithm were fairly successful in picking out field boundaries, especially those which contained very different (higher) values in the grayscale image (Figure 3-8). The algorithm has less success, however, in picking out the fields represented by lower values in the grayscale image (lighter grey in Figure 3-8). The weaknesses of this approach are therefore likely to be that it seems accurate at picking out only some types of fields. An iterative process of masking out fields picked out and then applying contours to the remaining image may help with this. This approach is likely useful in picking out fields with higher crop cover after extracting a parameter such as GLI or NDVI.

¹ <https://rdr.io/cran/imager/man/contours.html>

nlevels= 2



Figure 3-8. Best results from contours approach applied to GLI image with trees masked.

3.3 ilastik

The delineation of field and farm boundaries can be performed using the *ilastik* software (Sommer et al., 2011; Berg et al., 2019, <https://www.ilastik.org/>). *ilastik* is an easy-to-use interactive tool that brings machine-learning-based (bio)image analysis to end users without substantial computational expertise. Moreover, this approach has been successfully applied in a previous FutureWater project on the detection of on-site farm reservoirs¹, an example of the workflow used in that project is shown in Figure 3-9.

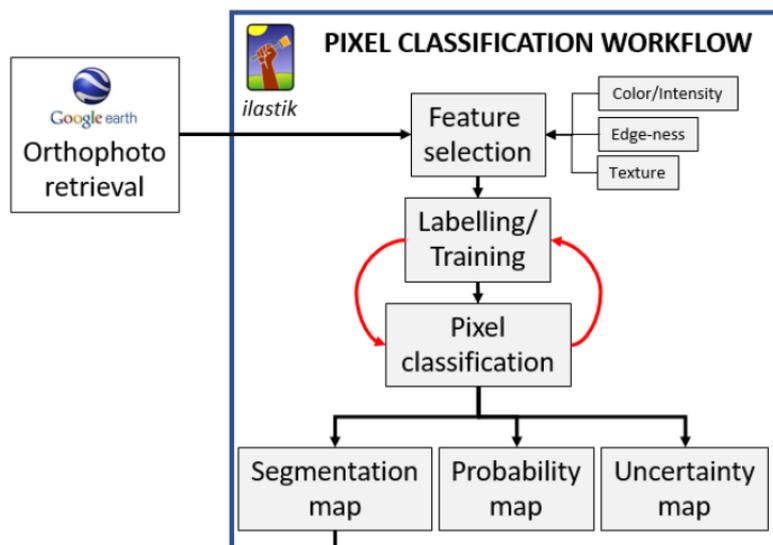


Figure 3-9 Image segmentation workflow from previous FutureWater report

¹ <https://www.futurewater.eu/projects/detection-of-on-farm-reservoirs-in-irrigated-areas/>

In Ilastik labelling or training was performed using the tools in the program. This requires some manual labour and insight to correctly label the image. A screenshot of this process is shown in Figure 3-10. The results of the pixel classification process in Ilastik is shown in Figure 3-11. The result indicates that four labels is not sufficient for finding distinct segments of fields. In addition, a two step classification could benefit this process and remove the trees in the first step. This is likely available as the software offers an 'Autocontext (2-stage) method. In addition, it also offers assistance with selecting features to label. These functionalities require more exploration to be able to get the best result from Ilastik.

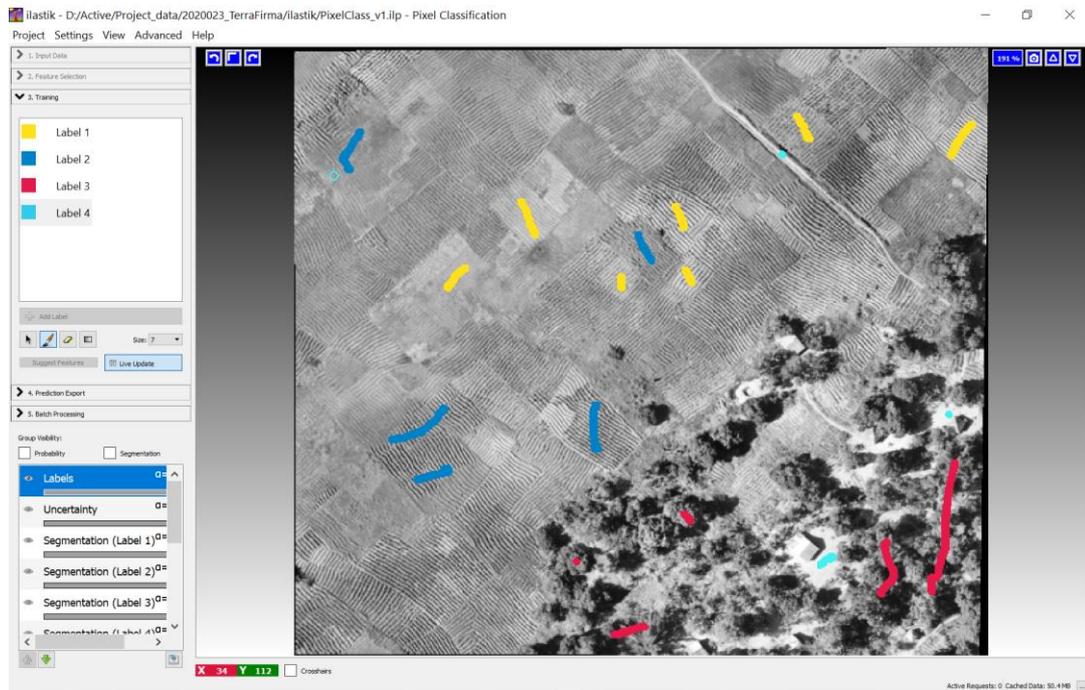


Figure 3-10 Training in Ilastik indicating various labels for different classes

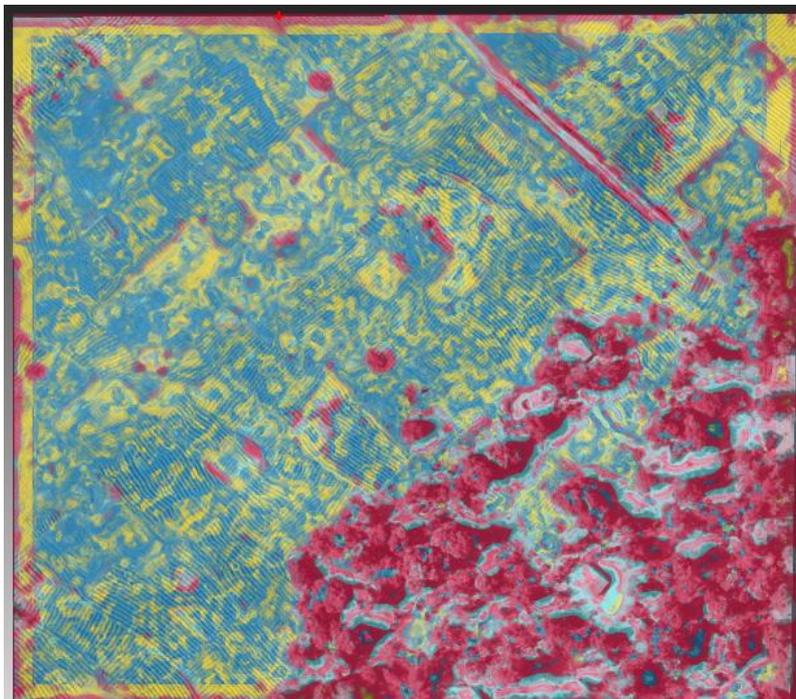


Figure 3-11 Pixel classification result of Ilastik using four labels

3.4 QGIS Segmentation (GRASS)

QGIS provides a segmentation algorithm in the GRASS plugin¹, which has been well-documented and is applicable for multi-band data. This algorithm was tested for two different threshold levels in the plugin: 0.1 and 0.3 as shown in Figure 3-12. A higher threshold indicates that more pixels will be agglomerated to one class thus creating larger segments. The results show reasonably good results however the lines of the furrow are still prominent especially in the results with threshold 0.1. Further processing and piloting this tool could improve the results.



Figure 3-12 Results of the GRASS QGIS segmentation for two different threshold settings

3.5 Manual boundary delineation

The previous methods mentioned all indicate automatic algorithms for processing imagery to support boundary delineation. However, in practice these algorithms have varying levels of accuracy. Therefore, frequently boundary delineation is performed manually, thus by mapping the boundaries by hand. The flying sensor imagery indicates that the parcels are clearly visible, which can assist this process. It can also be selected to perform a hybrid approach, thereby first running an automatic algorithm and then making final corrections by hand. This can lead to best accuracy but will require more time and effort. As an estimate this “test” area of 6 hectares can cost likely around one hour of labour to draw polygons around each individual field, which are on average 0.05 hectares each. By adopting the pre-processing steps of removing trees and creating more contrast with the plotRGB tool, this manual boundary delineation can be assisted. If additionally, one of the algorithms is used to create preliminary segments, it can be estimated that the manual labour required for the process can be halved.

¹ <https://grass.osgeo.org/grass76/manuals/i.segment.html>

4 Field boundary results

4.1 Flying Sensors

To get a better idea of how the algorithms described in chapter 3 perform on a larger scale, a large area was considered. This area represents around 80 ha and is made up of a large number of stitched drone images from flights in the same day for Area 3. The image, shown in Figure 4-1, is made up of a number of field types and also contains roads, trees and buildings. As mentioned in chapter 3, some pre-processing steps were conducted such as removing trees, reducing the pixel resolution, and translating the image to single band image.



Figure 4-1 Flying sensor image of selected block in Area 3

4.1.1 R packages

The results of applying the kmeans, canny edges, superpixels and contours algorithms are found in Figures 4-2, 4-3, and 4-4. These are interesting as they show that for the larger area, different approaches are more successful. The following observations can be made:

- Kmeans still has a fair amount of success in resolving different fields, but may need some smoothing and further postprocessing to get a clearer distinction
- Canny edges approach has more success in picking out the edges of fields for a larger image, but still does not show field boundaries clearly
- Superpixels approach in this case fails to pick out almost all fields
- Contours approach was unsuccessful in resolving field boundaries for the larger area

Kmeans 4class GLI



Figure 4-2 Results of kmeans step using the Green Leaf Index (GLI)

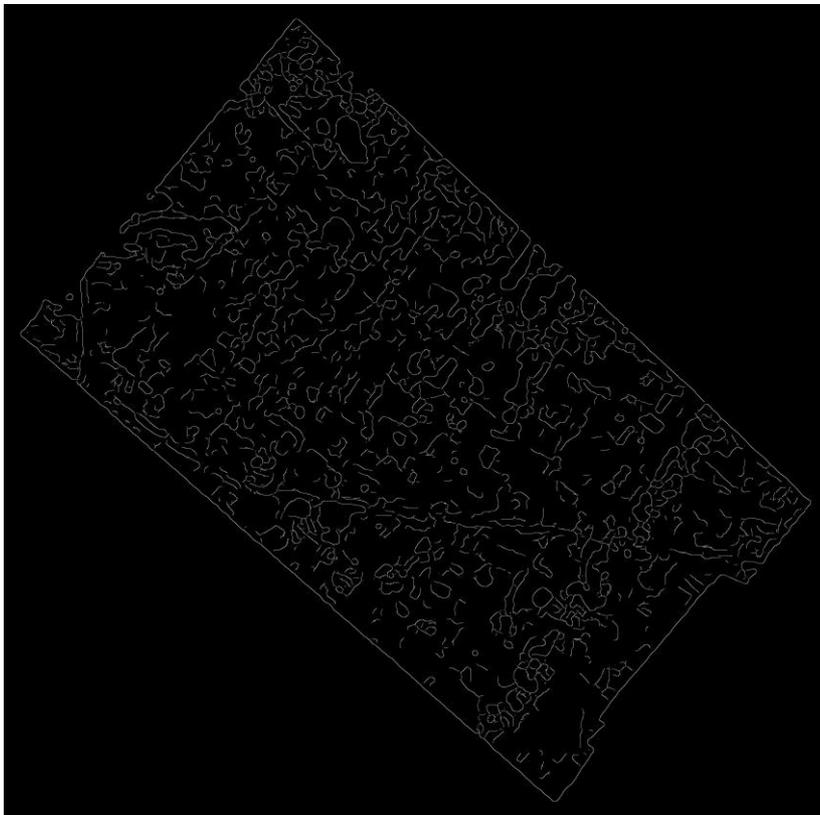


Figure 4-3 Results for the canny edges algorithm

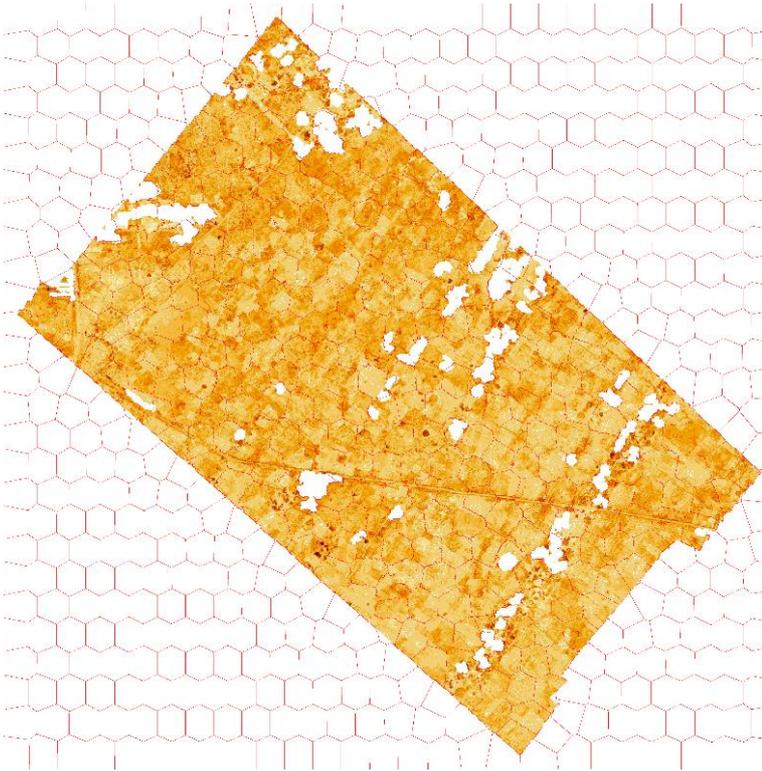


Figure 4-4 Results for the super pixels algorithm

Overall, the algorithms were less successful in resolving field boundaries when a larger image was taken as input. This therefore shows a need to either split larger images to more specific areas of interest when applying these procedures, or to apply a moving window approach to processing. This pilot study suggests a moving window approach with a window of size 10x10 average field size would be more successful in general.

The image hereby selected is also limited as it shows crops in similar growth stages. It is therefore also likely that the ability of these algorithms to detect field boundaries would increase if multitemporal images that showed fields at differing growth stages were used. This would likely help differentiate between crop types and also perhaps show dividing features between fields better at bare earth stages.

4.1.2 QGIS GRASS segmentation

The QGIS GRASS algorithm was also piloted over this large scale area, with the results presented in Figure 4-5. The results show that the 0.1 or 0.3 thresholds gave reasonable results in distinguishing fields. Classes were segmented indicating the fields, although the field boundaries are not aligned perfectly with the visual image. However, the roads are indicated well in the segments and trees are separately correctly in different classes. The parcels of the fields are similar therefore some fields are agglomerated in one class. At large scale the effects of the furrows are less pronounced which was an issue in the “test” image of Chapter 3. A detailed zoom of the results in Figure 4-6 shows the varying results of the field boundaries with some fields having good alignment and others being mis-classified. In general, the result shows good potential as an automated approach for defining field boundaries with flying sensor imagery.

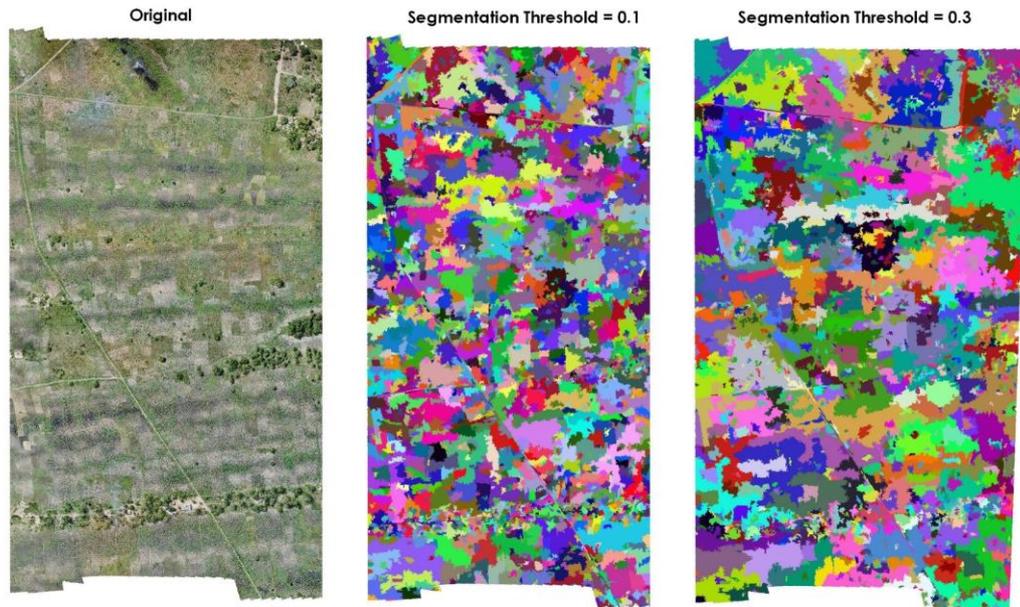


Figure 4-5 Results of GRASS segmentation algorithm



Figure 4-6 Detailed zoom of the segmentation results from GRASS (threshold = 0.1), converted to polygons

4.2 Satellite imagery

The algorithms as presented in chapter 3 could also be performed on satellite imagery. However, the available satellite imagery did not permit this. The imagery available through Digital Globe and Planet are available against a price. The Google Image provides high resolution data but is not downloadable as a Geotiff multi-band dataset, therefore cannot be used for further processing in algorithms. Lastly, the Sentinel image can be used in the algorithms, but the resolution was insufficient for distinguishing fields.

A comparison is made visually of the satellite imagery available and the drone image, which is presented in Figure 4-7. This comparison shows an area in the study location including two smaller details. The red box indicates a field with bare soil which is large enough to be detected also by the Sentinel image. As shown the surrounding fields are all green in the Sentinel image, whilst the drone image indicates the variety of vegetation cover in the surrounding fields thus providing more detailed information. The blue box indicates an area of the tree row, which is detected by the Sentinel image as dark green. The details of this tree row are only visible in the drone image and the Google satellite image. The Google image however is a fixed point in time, which is a disadvantage when land cover is continuously changing.

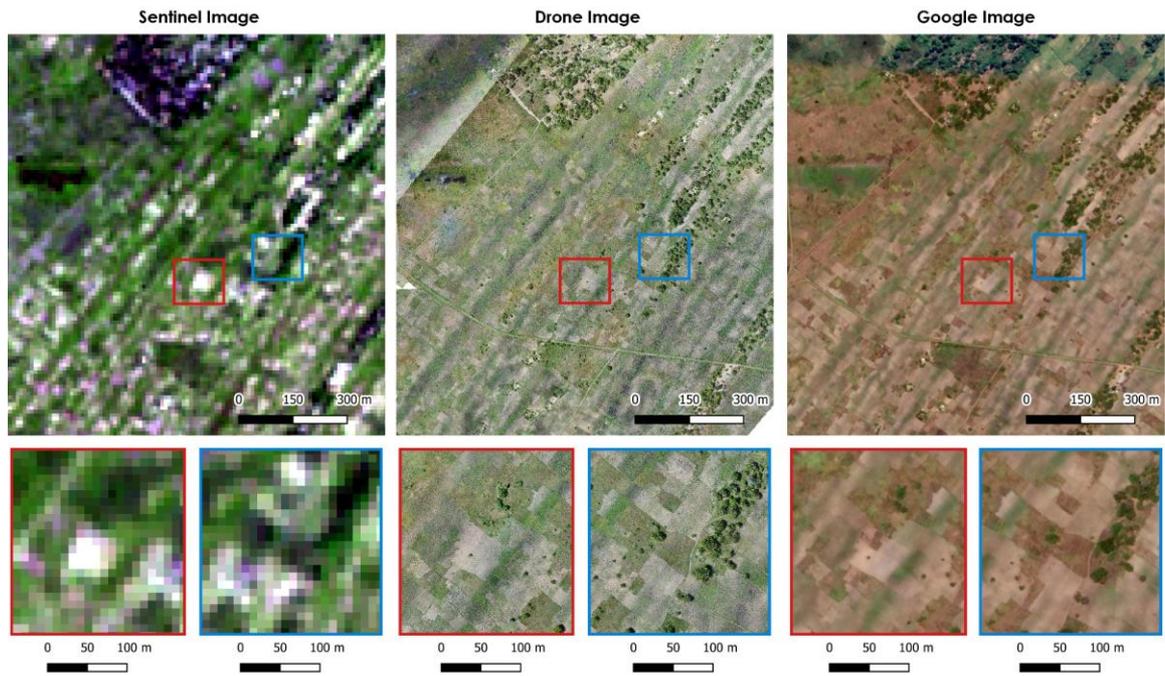


Figure 4-7 Results comparing the Flying Sensor (drone) imagery with satellite Sentinel and Google imagery

4.3 Discussion

4.3.1 Segmentation methodologies

In this chapter, the several methods for classification and segmentation are piloted. These performed in varying degrees and the pros and cons of these methods are demonstrated in the results of the flying sensor images. There does not seem to be a 'one size fits all' solution, with results that give the field boundaries accurately. However, this is a challenge especially in the situation of small-scale agriculture with parcels of land smaller than a hectare. The various pros and cons are summarized in Table 4-1.

Table 4-1 Overview of the pros and cons of the different methodologies as used for flying sensor imagery

	kmeans	Canny edges	Superpixels	Contours	ilastik	GRASS
Pros	<ul style="list-style-type: none"> Effective in removing trees Performed well in distinguishing fields Easy to run if experienced in R 	<ul style="list-style-type: none"> Distinguishes contrasting edges well such as roads and trees Easy to run if experienced in R 	<ul style="list-style-type: none"> Creates polygons that can be adjusted manually Easy to run if experienced in R 	<ul style="list-style-type: none"> Ability to distinguish field boundaries with sufficient contrast Easy to run if experienced in R 	<ul style="list-style-type: none"> User friendly interface is provided 	<ul style="list-style-type: none"> User friendly interface is provided in QGIS Performs well in distinguishing between fields
Cons	<ul style="list-style-type: none"> Requires further smoothing for better result 	<ul style="list-style-type: none"> Contrast between fields was insufficient to be detected 	<ul style="list-style-type: none"> Requires further improvement to distinguish between fields 	<ul style="list-style-type: none"> Less successful in distinguishing fields with similar cover 	<ul style="list-style-type: none"> Requires effort and expertise for adding labels 	<ul style="list-style-type: none"> Requires adjustments of parameters for best result

The various methods show good potential and this pilot indicated recommendations for improving this work further. Firstly, the input parameters for the algorithms can be optimized further to improve results. The results in this pilot are preliminary and with expert knowledge and optimization results are expected to improve.

Secondly, additional processing steps can improve the results. The addition of the step for removing trees, thus a two-stage classification process, gave better results. Another processing step that can be added is the moving window for the analysis. This can smoothen the contrast within fields thus laying focus on the contrast between fields.

Thirdly, a dataset of images for different (two or three) dates throughout the growing season is expected to improve the classification significantly. The flying sensor images show distinction between fields but several fields have similar vegetation cover thus the algorithms cannot detect the field boundaries. If data was also provided at the middle and end of the season, the different timing of planting and peak vegetation will give more contrast between fields. As an example, the work of a previous project in classification in the Bekaa Valley of Lebanon shows the result that is achievable by using multi-date images, shown in Figure 4-8. Both larger fields and the smaller agricultural fields are clearly distinguished.

Lastly, a hybrid approach of using automated algorithms in combination with some manual corrections of the field boundaries is recommended. The algorithms, especially kmeans and the GRASS segmentation algorithm, show that it is able to distinguish several fields, thus saving time in the manual drawing of field boundaries.

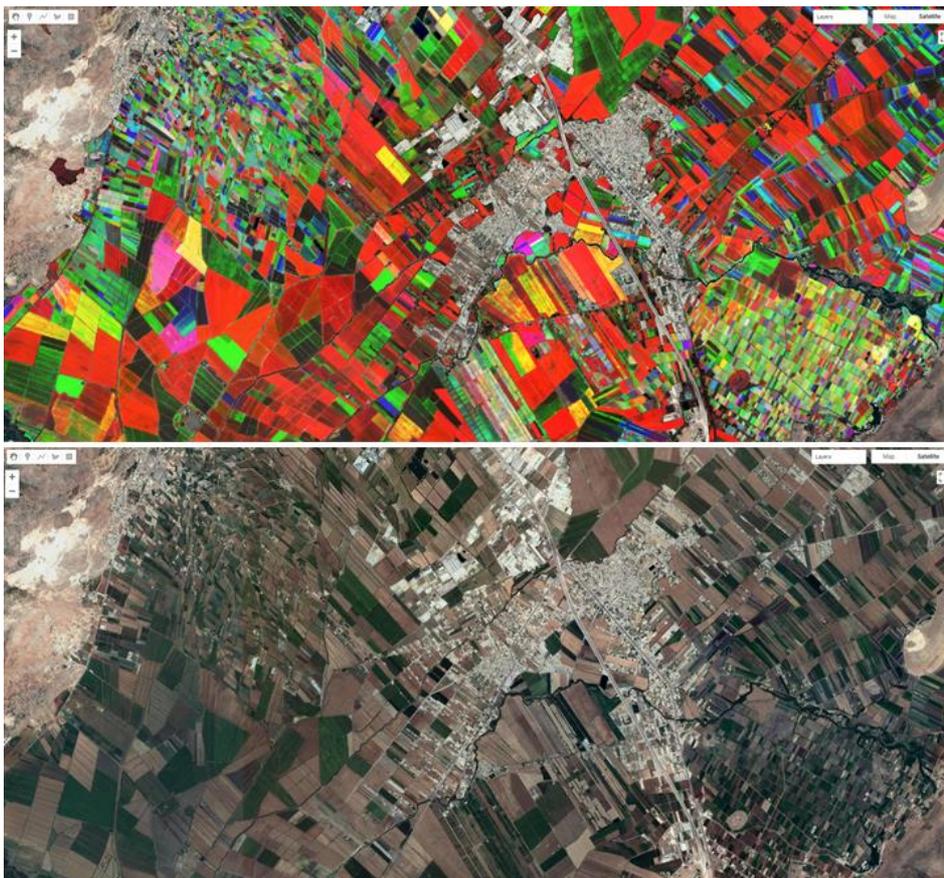


Figure 4-8 Example of classification for Bekaa Valley (Lebanon) using multi-date imagery. Top: layer stack of three NDVI images at three different dates during the growing season. Bottom: Google satellite image

4.3.2 Satellite and flying sensor data

Cadastre mapping can make use of either flying sensor data or satellite data. The suggested methods for segmentation and classification were not successfully applied on the available satellite data due to inavailability of suitable images for the methods. However, some general comparisons can be made between the different types of imagery available. Table 4-2 presents an overview of some key pros and cons typical for the different data products.

Table 4-2 Overview of the pros and cons of using flying sensor and satellite imagery for boundary delineation

	Flying Sensors	Google Earth	Sentinel 2	Planet or Digital Globe
Pros	<p>High pixel resolution of 4 cm achieved, which is suitable for indicating individual parcels in small-scale agriculture</p> <p>Imagery acquired on demand for desired time and date</p> <p>Cloud cover does not obstruct the image</p>	<p>High resolution imagery available, suitable for indicating individual fields</p> <p>Imagery freely made available</p> <p>Coverage of large area</p>	<p>Imagery freely made available</p> <p>Images available at 5-day time steps</p> <p>Coverage of large area</p>	<p>Imagery available at reasonably high resolution of 0.3 to 5m</p> <p>Imagery available at daily (or multi-day) time steps</p> <p>Coverage of large area</p>
Cons	<p>Requires post processing to create ortho-mosaic images</p>	<p>Imagery is not suitable for further analysis in segmentation algorithms</p> <p>Timing of the satellite image is not flexible</p>	<p>Pixel resolution is 10m which is not suitable for distinguishing individual fields in small-scale agriculture</p> <p>Cloud cover obstructs images and limits the availability of images</p>	<p>Imagery available after purchase</p> <p>Cloud cover obstructs images</p>

Notably, flying sensor imagery presents the highest resolution dataset compared to the satellite data products. In the pre-processing sometimes the flying sensor imagery was down-scaled to a more coarse resolution to achieve better results. However, the amount of detail obtained by having original images of high resolution provides a better image quality than images with the same pixel resolution as the satellite data. As example the flying sensor “test” image which was processed to 0.4m pixel resolution is compared with the Google satellite image (of a different date) in Figure 4-9.



Figure 4-9 Comparison of flying sensor image (resized to 0.4m pixel resolution) and Google Satellite

5 Concluding remarks

This study provides the findings from piloting several classification and segmentation algorithms for delineating field boundaries using flying sensor imagery. The suitability of using flying sensor and satellite imagery for boundary delineation in small-scale agriculture is discussed. The study area used in this pilot study is in the vicinity of Quelimane, Mozambique located in the Zambezia province. The flying sensor imagery was for 1,120 ha total, which was acquired in December 2020. Satellite imagery from Google, Planet, DigitalGlobe, and Sentinel 2, was acquired for the same location and used as comparison.

The methods piloted for segmentation and classification to support the field boundary delineation process was categorized into: 1) packages available in R; and 2) readily available software. The R packages that were used in this study are: kmeans, canny edges, contours, and superpixels. The readily available software used in this study are ilastik and GRASS (QGIS).

These algorithms gave mixed levels of success on the flying sensor imagery. The R package kmeans was successful in classifying trees from other land surfaces. It also performed well with distinguishing vegetated fields from non-vegetated. However, the contrast between vegetated fields was not sufficient to be classified as separate classes. The other R packages performed reasonably well for the “test” image but was less successful for a larger image. Overall, it can be suggested that different algorithms are applicable for different cases. For example, canny edges can be successful if the fields have clear dividing features. The GRASS image segmentation tool provided in QGIS showed good results for distinguishing individual fields.

Overall, some of the algorithms especially GRASS and kmeans show potential for being used for field boundary delineation, with several fields being delineated correctly. Issues occurred with small fields and fields with unclear dividing features. Piloting these algorithms shows that an automatic method will not be fully successful for defining fields in small-scale agriculture with perfect accuracy. However, a hybrid approach can be a good benefit to the field delineation work. Several processing steps using the piloted algorithms can be taken to provide a preliminary field boundary delineation. Thereafter, the boundaries can be manually corrected if necessary. It is estimated that the time required for field boundary delineation can be halved compared to drawing all field boundaries manually.

Further recommendations for improving results are to adopt a moving window approach for running algorithms on larger images. Some algorithms had better success on a small image (6 ha) compared to the larger image (80ha). A multi-date image can also assist the classification process. By adding a second and third image to the dataset, more contrast between fields will become evident because crops will peak in vegetation at different times.

The use of flying sensor or satellite imagery for field boundary delineation is compared. The flying sensor imagery provides the highest pixel resolution, which is necessary for identifying the individual parcels. In addition, flying sensor imagery is flexible in timing and does not have issues with cloud cover. Other satellite products, such as Planet, DigitalGlobe and Sentinel have images which are obstructed by cloud cover. However, the frequency of imagery acquisition is high (daily up to 5-day). The high-resolution imagery of Planet and DigitalGlobe are sold at commercial prices. Sentinel 2 and Google Satellite provide imagery for free. However, the 10m pixel resolution of Sentinel 2 is not suitable for identifying fields. Google Satellite provides a high-resolution dataset but is not suitable for further use in algorithms.

In conclusion, this pilot study demonstrates the benefits of using flying sensor imagery for field boundary delineation. In addition, preliminary analysis with various algorithms for segmentation and classification, show that an automatic approach can assist the field boundary delineation process but is not fully successful with defining fields with perfect accuracy. This will require additional manual corrections or adding imagery from multiple dates throughout the growing season.