JNCC Report
No. 655

Review of image segmentation algorithms for analysing Sentinel-2 data over large geographical areas

Sideris, K., Colson, D., Lightfoot, P., Heeley, L. & Robinson, P.

July 2020

ISSN 0963-8091

In partnership with Natural England

For further information please contact:

Joint Nature Conservation Committee
Monkstone House
City Road
Peterborough PE1 1JY
www.jncc.gov.uk

This report should be cited as:

Sideris, K., Colson, D., Lightfoot, P., Heeley, L. & Robinson, P. 2020. Review of image segmentation algorithms for analysing Sentinel-2 data over large geographical areas. JNCC Report No. 655. JNCC, Peterborough, ISSN 0963-8091.

JNCC EQA Statement:

This report is compliant with the JNCC Evidence Quality Assurance Policy

https://jncc.gov.uk/about-jncc/corporate-information/evidence-quality-assurance/

Acknowledgements

We would like to thank our colleagues at JNCC for their support and input into the project. We would also like to thank the Rural and Environmental Science & Analytical Services (RESAS) team in the Scottish Government for providing access to data sources. Finally, we would like to extend our thanks and acknowledge the Evidence Earth Observation Service (EEOS) in Natural England for their support throughout the project.

Executive summary

Earth Observation (EO) has been extensively used to provide a synoptic view of land use, cover and change at a variety of scales. In 2017 Natural England published the methodology for a satellite-based habitat mapping project, building on proposals and lessons learnt in the Making Earth Observation Work for UK Biodiversity (MEOW) programme. The resulting project and methodology were called Living Maps (Kilcoyne *et al.* 2017). Subsequently this methodology has been employed by organisations across the public and commercial sector (e.g. Colson & Robinson 2019; Jones *et al.* 2019; Kilcoyne *et al.* 2019). An integral part of Living Maps is a reliable spatial framework, referred to as a segmentation, to act as the basis of the classification.

Traditionally EO analysis and classifications have been performed on a per-pixel basis. However, accuracy improves when spatially and spectrally homogeneous units are identified and then grouped, classifying areas of similar properties; a segmentation. This method is referred to as object-based image analysis (OBIA). Comparisons between OBIA methods have been explored in academic settings, however this has not been applied extensively in an operational public sector environment where there are often different constraints.

This paper outlines the process of deriving OBIA from open-source and proprietary packages currently available. Three packages have been explored; OrfeoToolBox (OTB), Geographic Resources Analysis Support System (GRASS) Geographic information System (GIS) and eCognition.

Sentinel-2 imagery was processed by JNCC for an area around Kelso, Scotland, where there was an existing spatial framework with a high degree of accuracy that could be used for object validation. This area contains a reasonably large variety of upland and lowland land cover types within a small area. Processed imagery was then segmented into objects using the three software packages.

Results were compared with field validation polygons. Discrepancies in the detail, statistical and visual accuracy of the polygons were assessed. eCognition produced fewer, larger segments than the other software with the input parameters. In comparison, the OTB and GRASS GIS outputs produced smaller segments using as similar parameters as possible. This indicates that while eCognition is flexible and suitable for both broad and fine scale segmentations, including combining both in a nested hierarchy, OTB and GRASS GIS are more suited for smaller input images. These two pieces of software have less user-friendly control over the number of outputted segments. Their algorithms can be scaled up for use with larger input images and mosaics, but their processing times increase considerably; in order to reduce this, powerful CPUs are needed. eCognition is a more flexible piece of software given the current constraints.

This study has demonstrated that each software package is suitable for generating acceptable segmentations; each package performed well. However, some tools are optimised for certain applications given computational restraints. Recommendations and future directions are outlined.

Contents

1 Introduction

Earth Observation (EO) has been extensively used to provide a synoptic view of land use, cover and change at a variety of scales. New sensors are being developed and launched at an increasing rate, with some missions making data openly accessible (such as the Copernicus Programme's Sentinel data). EO is a valuable resource even when no other data are available but is most powerful when combined with field data and a variety of other data sources to create products that provide critical information, particularly for decision making.

Approaches to analyse the imagery into distinct classes are perhaps the most frequently used methods employed to extract this important information from EO data (Gitas *et al.* 2009). Various methods of classification are available, traditionally these were pixel-based unsupervised and supervised methods. Unsupervised classification involves statistical clustering of data to identify natural groupings of pixels or objects, the number of groups being specified in advance by the analyst. In a supervised classification, the analyst defines the classes, then assigns objects to those classes using classification rules or machine learning (Foody 2002).

There has been a recent move towards object-based classifications (Blaschke 2010), in which the data products are presented as polygons representing individual features of the landscape, such as fields or forestry blocks. Object-based image analysis (OBIA) has been advocated as a suitable approach to creating maps from EO (Gitas *et al.* 2009; Shepherd *et al.* 2019; Stoian 2019), with many researchers reporting that it improves classification accuracy (Blaschke 2010). The principal of OBIA is to group spectrally similar pixels together to form objects, and then classify these objects based on their spectral and geometric properties, relationship to other objects and/or to contextual data (Blaschke 2010; Teodoro & Araujo 2016; Torres-Sánchez *et al.* 2015).

OBIA has several advantages over per-pixel classification for habitat and land cover mapping. It avoids the 'salt-and-pepper' or 'speckle' effect, reduces within-class spectral variability, generates a vast number of features which can be used in classification, and enables analysts to use their ecological knowledge and contextual data in the classification process (Blaschke 2010; Blaschke *et al.* 2014; De Luca *et al.* 2019; Gitas *et al.* 2009).

The aim of image segmentation is to create objects which represent real-world geographic or ecological features; these objects should be as large as possible to reduce computer processing time, but as small as necessary to be ecologically meaningful. To optimise segmentation, the analyst must decide which imagery layers or derived products to use, what weighting to accord to each layer, and then define parameters of scale and homogeneity. Current projects that use this technique in the United Kingdom are Natural England's Living England (habitat map of England), the crop map of Scotland (Scottish Government & Edinburgh University 2019) and the Habitat Map of Northern Ireland (produced by JNCC and the Northern Ireland Environment Agency).

1.1 Software

The need for image segmentation has resulted in the production of several toolboxes and software solutions. Currently, the most widely used software for segmentation is eCognition (Witharana *et al.* 2014; Baraldi *et al.* 2018). However, the software is proprietary, and the annual licence cost is about £12,000 making it a considerable barrier to its use. There are also several open-source alternatives available, including OrfeoToolBox (Teodoro & Araujo 2016; Grizonnet *et al.* 2017) and GRASS (Lennert *et al.* 2019). The aim of this piece of work was to investigate these two free-to-use alternatives and compare them to the market leader, eCognition.

1.1.1 eCognition 9.2.1

eCognition is a suite of software tools and algorithms for image analysis and classification developed by Trimble Geospatial Inc. It focuses on object-based analysis and data fusion and integrated analyses. (Trimble Geospatial Inc. 2020). It includes eleven different segmentation algorithms. The algorithm used for this project was the Multiresolution Segmentation algorithm (Trimble Germany GmbH 2016). The Multiresolution Segmentation algorithm is a region growing algorithm which creates objects by starting with a single pixel, known as a 'seed' pixel, and merging it with neighbouring pixels until a threshold value of maximum heterogeneity is reached. This algorithm was chosen because it works well with multiband images, e.g. optical satellite images such as those produced by Copernicus Sentinel-2.

1.1.2 Orfeo ToolBox (OTB) 6.6.1

OTB is a library designed for remote sensing image processing, which was created by the French space agency (CNES) in 2006 and continues to be developed. The software is available to download from the OTB website[1]. OTB can be accessed via the command line, a graphical interface, Python and as a QGIS plugin. The toolbox provides several algorithms for satellite image processing and a segmentation tool that offers two different segmentation algorithms. The first is a watershed algorithm and the second is the mean shift algorithm. The mean shift algorithm was used as it is more versatile than the watershed algorithm allowing more user defined thresholds. In addition, the availability of the same algorithm from GRASS GIS provides a good opportunity for comparative evaluation.

1.1.3 GRASS 7.6.1 GIS

Geographic Resources Analysis Support System (GRASS), version 7.6.1, Geographic Information System (GIS), commonly referred to as GRASS, is a suite of software used for geospatial data management and analysis, image processing, graphics and maps production, spatial modelling, and visualisation. Initially released in 1982 by the U.S Government, it has been further developed by a team of researchers and scientists. The software is available as a standalone application from the GRASS GIS[2] website however it is more commonly used from within QGIS, where the software is in the standard QGIS core[3]. The software provides a tool called i.segment which uses two algorithms, the mean shift algorithm and the region growing algorithm, to perform segmentation of images (Fukunaga & Hostetler 1975). The mean shift algorithm was used as it provided a clearer comparison with OTB.

[1] https://www.orfeo-toolbox.org/CookBook-6.6.1/Installation.html
[2] https://grass.osgeo.org/
[3] https://docs.qgis.org/3.10/en/docs/user_manual/grass_integration/grass_integration.html

2 Methods

2.1 Input data

In order to obtain an accurate comparison of outputs from the different software suites the input data used was the same 10-band Sentinel-2 satellite image from the European Space Agency (ESA European Space Agency 2019) Copernicus programme.

Sentinel-2 is a high-resolution, multi-spectral imaging mission. The mission consists of twin satellites that are flying in the same orbit but phased at 180° which enables a high re-visit frequency of five days at the equator. At higher latitudes, the revisit time is more frequent. Both satellites carry optical instruments that can sample 13 spectral bands at different resolutions. There are four bands at 10m, six bands at 20m and three bands at 60m spatial resolution. Sentinel-2 products are provided in granules of fixed size within a single orbit, with each granule 100km x 100km. Table 1 indicates the data products that are available to users.

Table 1. Sentinel-2 product types (ESA European Space Agency 2019).

Name	High Level Description	Production and Distribution	Data Volume
Level-1C	Top-Of-Atmosphere reflectance in cartographic geometry	Systematic generation and online distribution	~600 MB (each 100km x 100km)
Level-2A	Bottom-Of-Atmosphere reflectance in cartographic geometry	Systematic and on-user side (Sentinel-2 Toolbox)	~800 MB (each 100km x 100km)

Working with Defra's Earth Observation Centre of Excellence and liaising with other devolved governments across the UK, JNCC has defined and agreed a standard set of processing steps for the raw data from the EU Copernicus satellites Sentinel-1 and Sentinel-2 which would underpin most of the potential uses within the natural environment sector (Jones *et al.* 2017; Minchella 2018). To fully exploit the valuable information contained within Copernicus data, users are required to undertake a series of complex pre-processing steps to turn the data from a 'raw' unprocessed format into a state that can be analysed.

To enable wider use and exploitation of EO data, JNCC are promoting the systematic and regular provision of Analysis Ready Data (ARD). This aligns with the Committee on Earth Observation Satellites' (CEOS) work on facilitating access to satellite data through the international CEOS Analysis Ready Data for Land (CARD4L) project. This notion of accepted standards is recognised by JNCC, the devolved governments and the wider CEOS community as a vital step for repeatable and comparable analytical work. Availability of ARD allows end users to use data immediately for visualisation or analysis, without needing to carry out complex pre-processing themselves. The transformations include geo-positioning the data, removing the effects of the atmosphere on the signals detected by the satellite sensors and masking out clouds (Jones *et al.* 2017). Data created using this standard for the project are ARD products generated specifically for the UK using higher resolution elevation data than is used in the ESA products. They have been processed to surface reflectance using the Atmospheric and Radiometric Correction of Satellite Imagery (ARCSI) software, v3.1.6[4] and reprojected using GDAL v2.2.4 using a singularity image based on docker image jncc/s2-ard-processor:0.0.0.45 produced by JNCC.

[4] http://www.rsgislib.org/arcsi

2.2 Study area

The study area for this project was Kelso in South East Scotland (Figure 1).

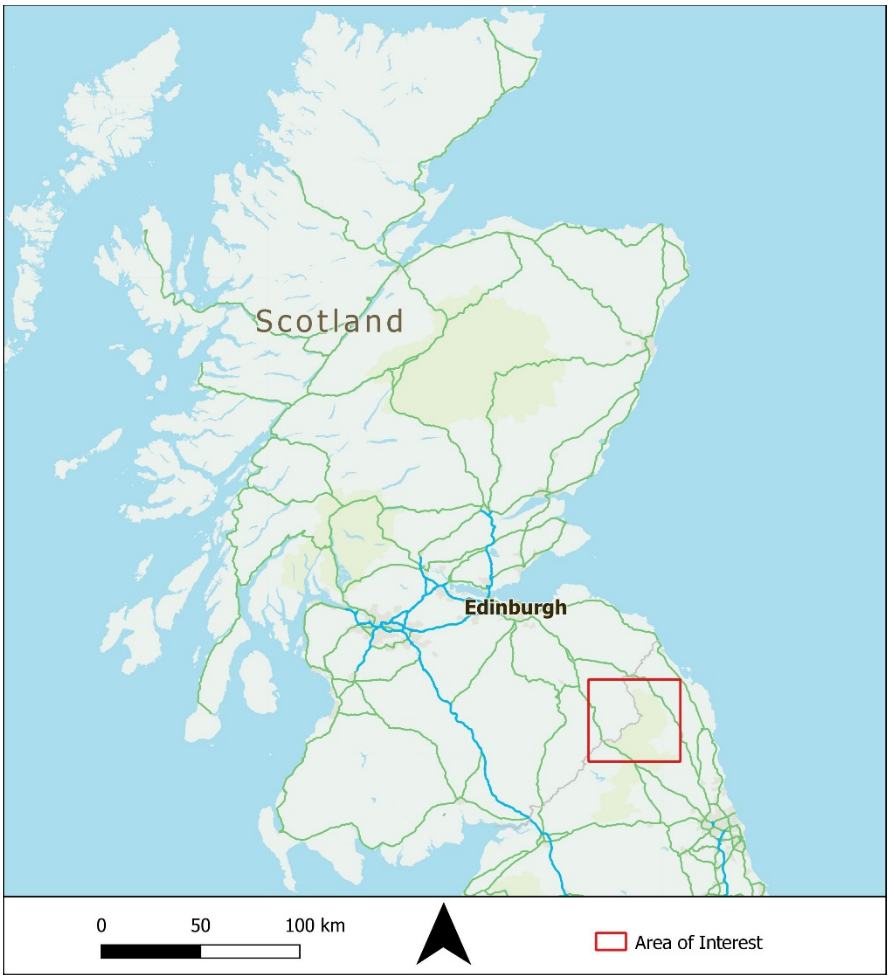

Figure 1. Map of the area of interest (highlighted red). Contains OS data © Crown copyright and database right 2019.

This area was chosen because it contains a large variety of land cover types within a relatively small area, and because ground truth data were available to evaluate the segmentation outputs. The ground truth polygons were provided by the Rural and Environmental Science and Analytical Services (RESAS) in the Scottish Government (Figure 2). This dataset is produced by the annual land survey of agricultural fields conducted by RESAS to quantify agricultural use in Scotland. The measurements of these fields are extremely accurate; they are surveyed using high accuracy GPS devices and then often snapped to Ordnance Survey MasterMap boundaries.

Figure 2. Field boundaries used as ground truth data in the northern agricultural lowland area of the AOI.

The Sentinel-2 granule used in the work is shown in Figure 3. The imagery was captured on 24 June 2018. It was processed by JNCC to ARD as described above (file name S2B_20180624_lat55lon213_T30UWG_ORB037_utm30n_osgb). The image was cropped to the test area using GDAL, a translator library for raster and vector geospatial data formats.

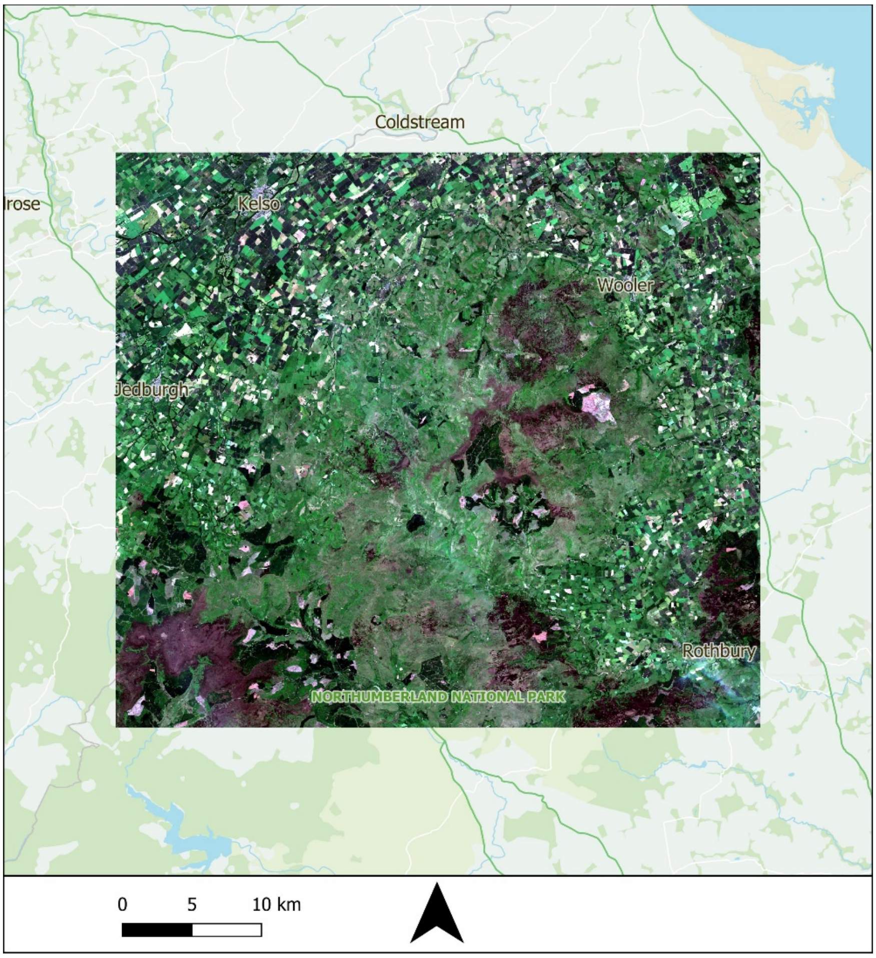

Figure 3. Sentinel-2 input image (acquired 24 June 2018), displayed in true colour. Contains OS data © Crown copyright and database right 2019.

2.3 Segmentation parameters

Each piece of software enables the user to influence the output by setting segmentation parameters. However, because the two open source tools use a different type of algorithm to eCognition, they offer a different set of user-defined parameters. The choice of parameters offered by each software are shown in Appendix 1.

To ensure a fair comparison, attempts were made to use parameter settings which were as similar as possible within each of the three pieces of software. The chosen parameters were based on those used for the creation of Natural England's Living England project using eCognition (Kilcoyne *et al.* 2019). Some experimentation was therefore required to identify the parameter settings which most closely resembled those used in the Living Maps method. The OTB mean shift algorithm was run with 29 different sets of parameters to identify those which produced the most similar results to the Living England segmentation. As GRASS GIS utilises the same algorithm as OTB the parameters identified for use in OTB were also used in GRASS. The parameters used in each piece of software for this study are provided in more detail below.

2.3.1 eCognition

Users can influence the size of objects by defining a scale parameter which determines the maximum allowable heterogeneity of pixel values in an object. The larger the scale parameter, the larger the objects. The scale parameter operates in combination with weighting criteria defined by the user to determine the degree to which segmentation is influenced by homogeneity of shape versus colour, and by object compactness versus smoothness. For example, if a user selects a shape weighting of 0.2, the colour weighting will be 0.8. A compactness weighting of 0.7 results in a smoothness weighting of 0.3. The input parameters available with the multiresolution algorithm of eCognition are described in Table 7.1 in Appendix 1.

The scale parameter and weighting criteria can be applied to multiple layers of imagery simultaneously. The user can select which bands are used in the segmentation and what weighting they should carry in the calculation of heterogeneity of pixel values.

Six segmentations of the Kelso test area were carried out using eCognition. One segmentation used a multi-level hierarchical approach adapted from Natural England's Living Map methodology (Kilcoyne *et al.* 2017). This is potentially a very efficient method because it takes advantage of eCognition's ability to segment different parts of the imagery at different scales, for example using a large-scale parameter for large homogeneous objects such as water bodies, and a smaller scale parameter for fine-scale patchy habitats such as scrubby grassland. The segmentation was carried out in three stages:

1. **Identify large homogeneous objects**
   Segmentation was carried out using a scale parameter of 100 with shape weighting of 0.2 and compactness weighting of 0.3 applied to the following bands: Green (weighted 2), NIR (weighted 1) and Red (weighted 1). Objects were classified as 'Large Homogeneous Objects' if they were larger than 50ha and if the standard deviation of NIR values within the object was lower than 15.

2. **Identify medium homogeneous objects and fields**
   Segmentation was then carried out on the remaining unclassified areas of imagery using a scale parameter of 75 with shape weighting of 0.1 and compactness weighting of 0.6 using the same bands and weightings as above. Objects were classified as 'Fields' if they had a rectangular fit value greater than 0.8 (an object with a rectangular fit value of 1.0 would be a perfect square) and if the standard deviation of NIR values within the object was lower than 35. Objects were classified as 'Medium Homogeneous Objects' if they had a rectangular fit value smaller than 0.8 and if the standard deviation of NIR values within the object was lower than 30. All remaining objects were classified as 'Non-Homogeneous Objects' and copied to a new level below.

3. **Refine non-homogeneous object segmentation**
   Non-Homogeneous Objects were segmented using a scale parameter of 50 with shape weighting of 0.1 and compactness weighting of 0.8 applied to the following bands: Blue (weighted 1), Green (weighted 1), NIR (weighted 5), Red (weighted 5), Red Edge 5 (weighted 1), Red Edge 6 (weighted 1), Red Edge 7 (weighted 1), Red Edge 8a (weighted 5), Short-Wave Infrared 1 (weighted 2) and Short-Wave Infrared 2 (weighted 2).

The other five segmentations were carried out at a single level, i.e. applying the same set of parameters throughout the imagery. The parameters and layer weightings used in these five segmentations are shown in Table 8.1 in Appendix 2.

Segmentations produced by eCognition can be exported as vector shapefiles. If desired, they can be exported with geometric and spectral attribute data selected by the user.

2.3.2 Orfeo ToolBox

The mean shift algorithm is nonparametric and iterative, creating a density gradient estimation using a generalized kernel approach to perform segmentation of satellite images (Fukunaga & Hostetler 1975). The mean shift approach is a flexible clustering technique that groups similar values together; values are represented in a space and through an iterative technique the segmentation is achieved. In OTB the segmentation with the mean shift algorithm is performed in two steps:

1. **Dividing**
   In this step the input image is divided into tiles and then the mean shift segmentation is performed within each tile. This is done by grouping together neighbouring pixels whose range distance is below the range parameter and (optionally) spatial distance is below the spatial range parameter.

2. **Stitching**
   In this step the segmentation algorithm stitches together the segmentation generated in the previous step. There is also the option not to stitch back together the generated segments at the boundaries of the generated tiles. The resulting segmentation is stored as a geospatial layer defined by the user.

OTB can be used either from within its own user interface, via command line, as a python library or as a plug-in in QGIS. All the parameters and inputs are the same in all the formats. For this project the command line interface was used. A description of the input parameters for the mean shift algorithm of OTB are described in in Appendix 1. In total 29 segmentations were run in OTB to identify the most similar parameters to the Living Maps parameters. The parameter values for each of these 29 test runs are shown in Table 8.2 in Appendix 2.

In each segmentation, all 10 bands of the Sentinel-2 ARD were used and given equal weighting.

2.3.3 GRASS GIS

GRASS GIS i.segment tool uses the mean shift algorithm (Fukunaga and Hostetler 1975), and also consists of two steps but has a different approach:

1. **Anisotrophic filtering**
   In this first step new cell values are calculated from all pixels not farther than hs (spatial range bandwidth) pixels away from the current pixel of the input image and with a spectral difference no larger than hr (spectral range bandwidth). This means that pixels that are different from the current pixel in the input image are not taken into consideration in the calculation of the new pixel values.

2. **Clustering**
   In this step, cell values are iteratively recalculated until the maximum number of iterations are reached or until the largest shift is smaller than the threshold. The threshold takes values between 0 and 1, where 0 merges only identical segments and 1 merges all segments.

The resulting segmentation is stored as a raster image file (.tif). For the purposes of comparison with the outputs from the other software this needed to be converted to a vector format. The GRASS GIS algorithm r.to.vect was used which first traces the perimeter of each unique area in the raster layer and creates vector data to represent it. The cell category values for the raster layer are used to create attribute information for the resultant vector area edge data. A true vector tracing of the area edges might appear blocky, since the vectors outline the edges of raster data that are stored in rectangular cells, so the r.to.vect algorithm also smooths the corners of the vector data as they are being extracted.

A description of the input parameters for the mean shift algorithm of i.segment are described in Table 7.3 in Appendix 1. Four segmentations were run in GRASS GIS to identify the most similar parameters to the Living Maps parameters. The parameter values used for each of the four test runs are shown in Table 7.3 in Appendix 2.

In in the first three segmentations, all 10 bands of the Sentinel-2 ARD were used and given equal weighting. In the fourth segmentation (output name “NDVI”) the only input layer used was an NDVI layer created from bands 3 and 7 of the Sentinel-2 ARD using the formula:

$$\frac{NIR - Red}{NIR + Red}$$

2.4 Comparative evaluation

One segmentation output per software was selected for evaluation:

- eCognition multi-level hierarchical segmentation described in Section 2.3.1
- OTB segmentation produced by the parameters highlighted in Table 8.2 (output 24)
- GRASS GIS segmentation produced by the parameters highlighted in Table (output 24)

Three methods of evaluation were used to compare these segmentations: (1) visual comparison of the segmentation with the input satellite imagery and the ground truth polygons, (2) comparison of descriptive statistics derived from the polygon geometry of the segmentation and of the ground truth data, and (3) comparison of processing time needed to produce each segmentation.

2.4.1 Visual assessment

Visual assessment provided qualitative evaluation of the segmentation outputs. The segmentation vector files were symbolised in QGIS 3.10 so that only the boundaries of the polygons were visible. These were then overlain on the original Sentinel-2 imagery, enabling visual assessment of the segmentation’s delineation of landscape features visible in the imagery such as woodlands, hedgerows, water bodies, roads and built-up areas. Particular attention was paid to features with a small area and high spectral contrast, such as small groups of trees surrounded by grassland. Discrepancies in the detail and visual accuracy of the polygons were assessed, and checks were made for the presence of anomalies, artefacts and edge effects.

The segmentation polygons were then visualised over the ground truth polygons, enabling visual assessment of the segmentation’s delineation of agricultural field boundaries.

2.4.2 Polygon statistics

The total number of polygons produced for each segmentation was recorded, as this gives insight into the level of landscape detail likely to be captured by the segmentation. The number of polygons also influences the processing time required to produce the segmentation, and the time required for potential future processing or analytical steps, such as conversion of raster to vector outputs or the generation of zonal statistics.

The area of each polygon in hectares was calculated using QGIS ‘field calculator’ tool for the three segmentations and the ground truth data. A spatial query was carried out in QGIS to select all polygons that intersected the 413 ground truth polygons. The ‘basic statistics for fields’ tool in QGIS was used to calculate a set of descriptive statistics from the area in hectares of the ground truth polygons and of the polygons in each segmentation that intersected the ground truth data. The statistics generated were minimum value, maximum value, range, sum, mean value, median value, standard deviation, coefficient of variation, first quartile, third quartile and interquartile range.

The statistics generated from each segmentation were compared with the statistics generated from the ground truth data to provide a quantitative evaluation of the segmentation results.

2.4.3 Processing speed

All segmentations were carried out on laptop computers with a 1.99GHz Intel i7-855OU CPU and 16GB of RAM running Windows 10 64-bit operating system. The time taken to complete each segmentation of the Kelso test area was recorded.

To test the performance of the software on a larger area, segmentations were carried out using the parameters outlined in Section 2.4 above to segment a mosaic of Sentinel-2 imagery covering the whole of Scotland. The mosaic imagery was a 14GB raster file in TIF format. The total time taken by each piece of software to complete the segmentation was recorded, along with the number of polygons produced.

3 Results

3.1 Visual assessment

The multi-level segmentation produced by **eCognition** successfully delineated medium-scale landscape features such as fields, hedgerows and woodlands. Figure 4 shows a section of the eCognition segmentation visualised over the original Sentinel-2 imagery, demonstrating that the segmentation provides a reasonably detailed representation of the landscape.

However, the level of detail was not as high for certain features such as agricultural fields and buildings, and in some cases visually distinct areas had been grouped into a single object, as exemplified in Figure 5.

Comparison with the ground truth polygons reinforced these findings, showing that the eCognition segmentation successfully delineates field boundaries in many cases, but in some cases merges adjacent fields into a single polygon (Figure 6).

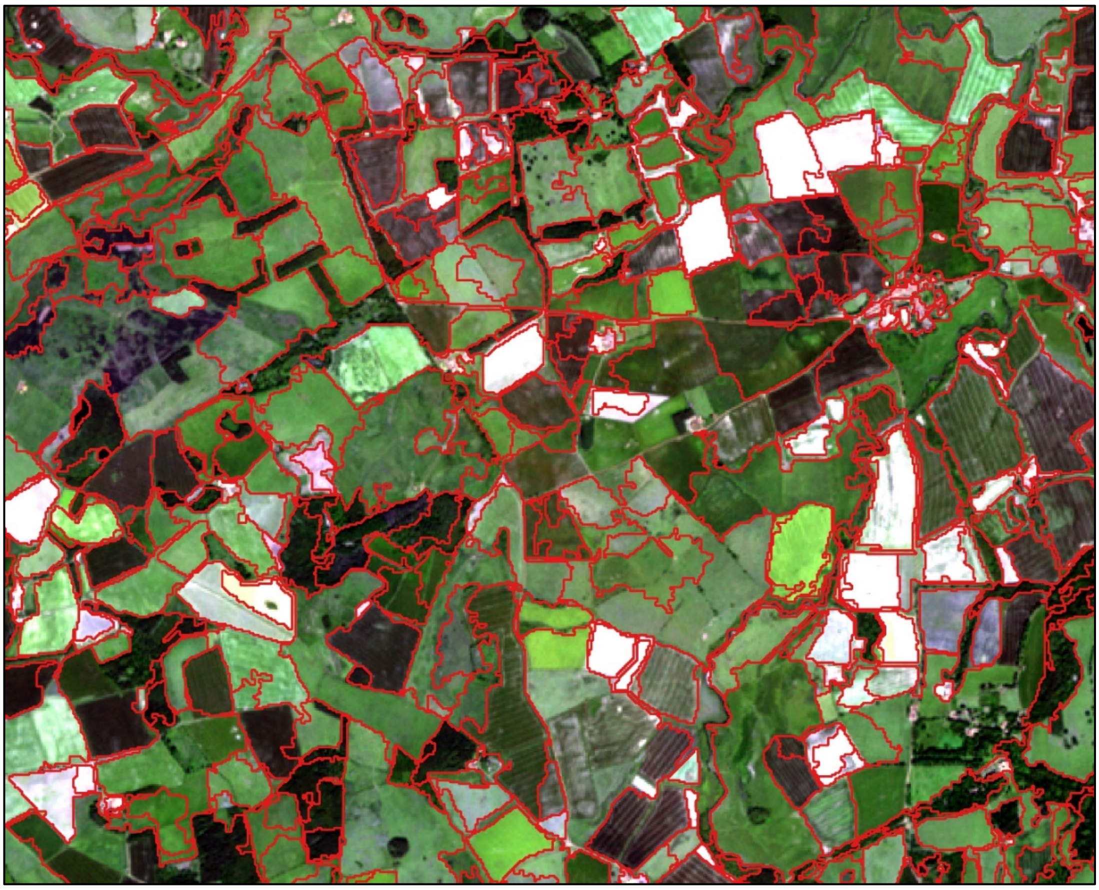

Figure 4. Section of eCognition multi-level segmentation over Sentinel-2 imagery. Image width ca. 6km.

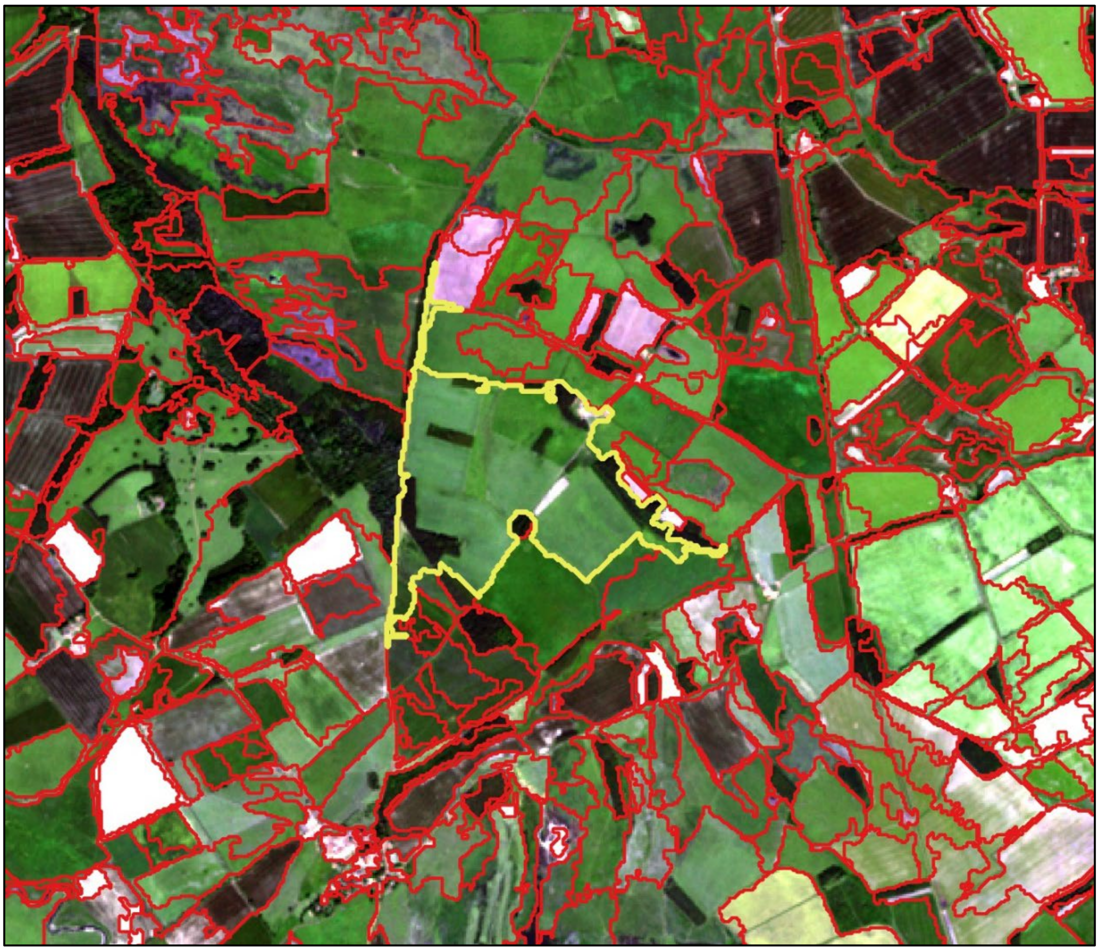

Figure 5. Example of failure to separate small but spectrally distinct areas in the eCognition multi-level segmentation output. Image width ca. 5.5km.

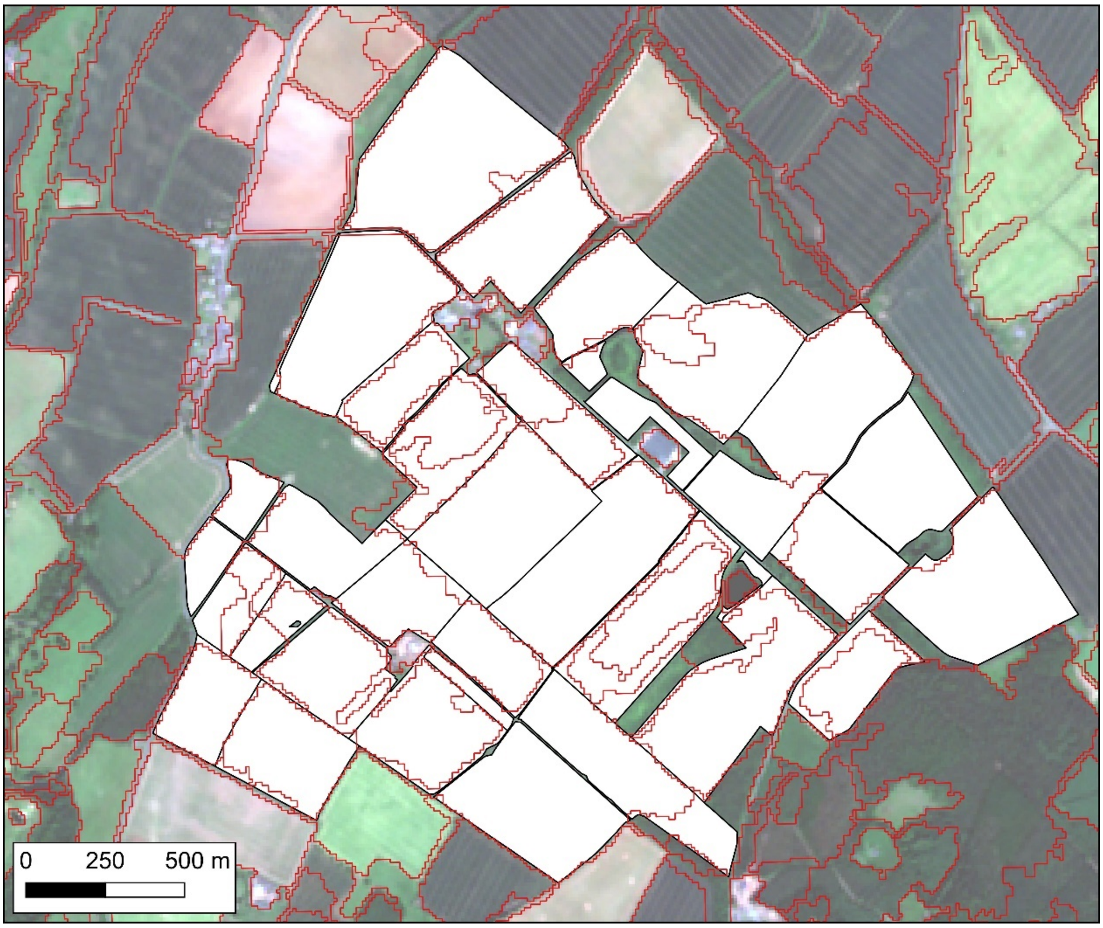

Figure 6. Example of visual comparison between section of eCognition segmentation and ground truth data.

The segmentation produced by **OTB** was extremely detailed, and often delineated objects based on spectral differences within individual fields, for example patches of drier ground. However, in many instances its split areas into separate segments where no discernible difference could be identified visually. Figure 7 shows a section of the OTB segmentation output visualised over the original Sentinel-2 imagery, demonstrating that seemingly homogeneous fields have been segmented into several polygons. Visual assessment also showed that linear artefacts are visible in the segmentation, having been produced by the stitching process after the area had been tiled for processing. These are perfectly straight horizontal or vertical lines intersecting spectrally homogeneous areas of the imagery, as exemplified in Figure 8.

Comparison with ground truth polygons also showed the high level of detail of this segmentation, with some fields being split into 10 or more polygons in the example shown in Figure 9. Vertical stitching artefacts are also visible in Figure 9.

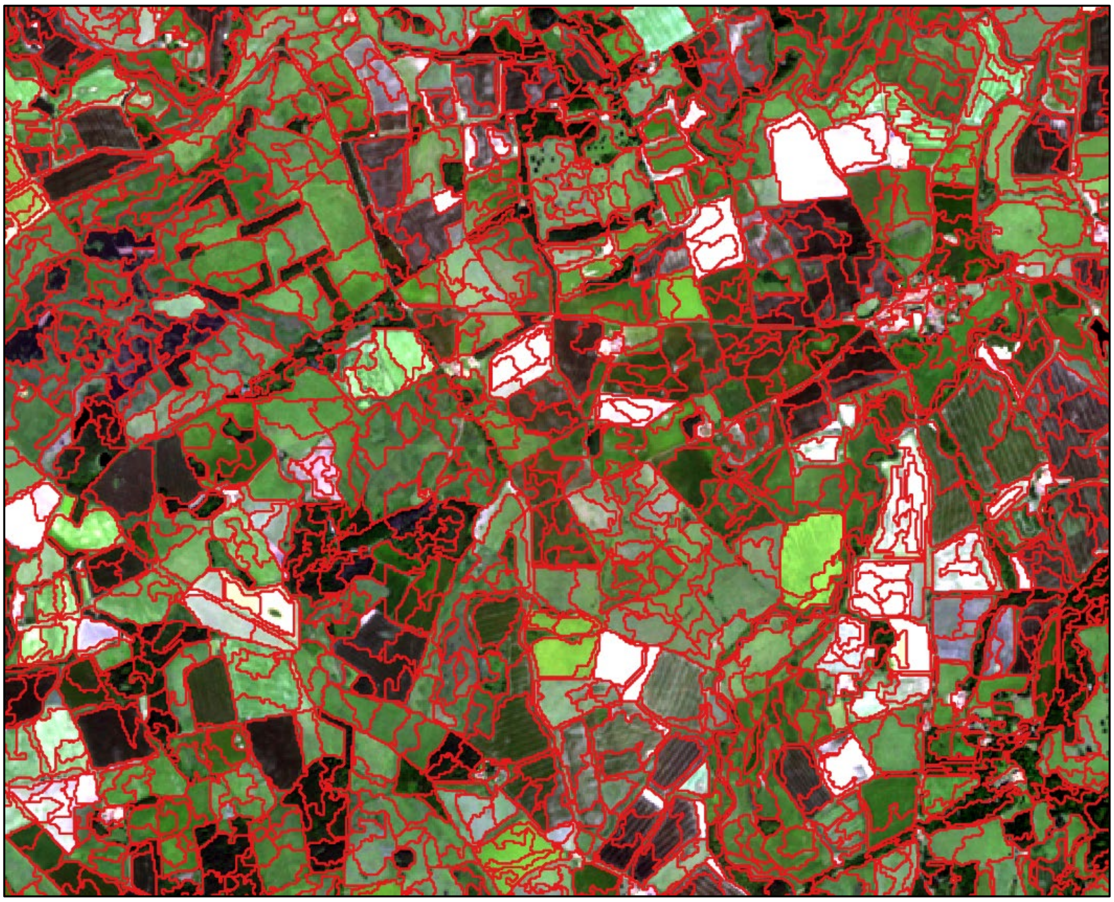

Figure 7. Section of OTB segmentation visualised over Sentinel-2 imagery. Image width ca. 6km.

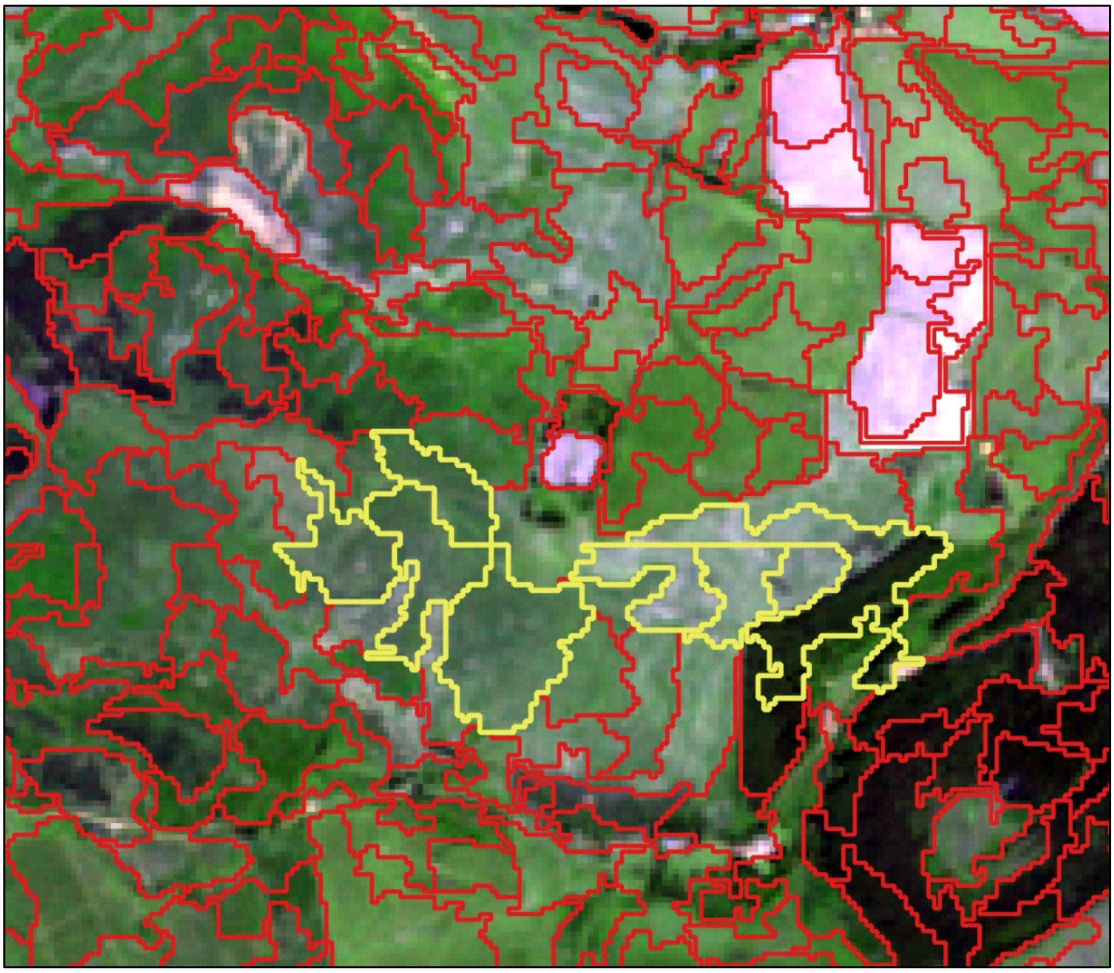

Figure 8. Example of linear artefacts produced by stitching of the OTB output.

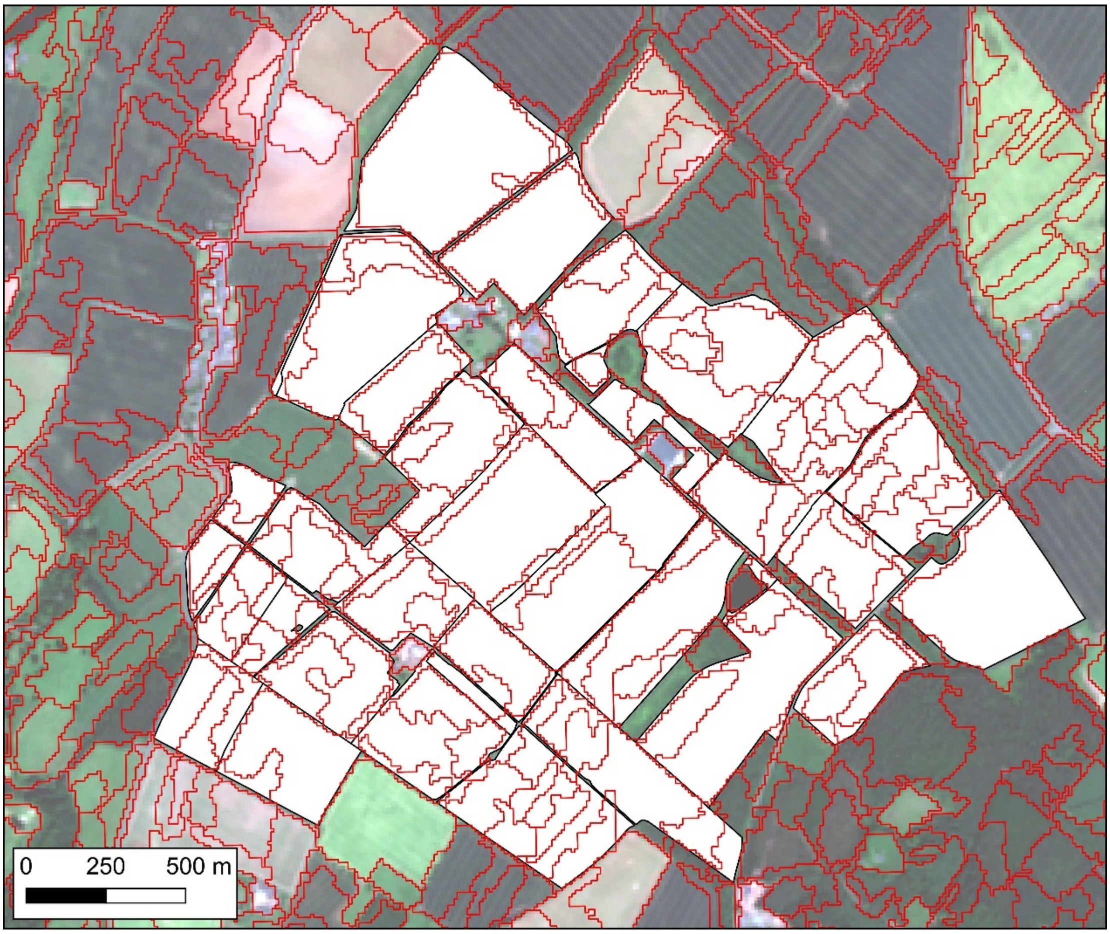

Figure 9. Example of visual comparison between section of OTB segmentation and ground truth data.

The segmentation produced by **GRASS GIS** was also highly detailed, but less detailed than the OTB segmentation. Like the OTB segmentation, it successfully delineated different land uses within fields, but also generated multiple polygons in many areas where there was no discernible visual difference. Figure 10 shows a section of the GRASS GIS segmentation results visualised over the original Sentinel-2 imagery.

Visual comparison with the ground truth data showed that the GRASS GIS segmentation successfully delineated the outer boundaries of agricultural fields (Figure 11). Figure 11 also shows that although the GRASS segmentation occasionally splits fields into multiple polygons, it does not produce as many small within-field polygons as the OTB segmentation.

GRASS does not tile imagery during processing and stitch the output, so it does not produce linear artefacts.

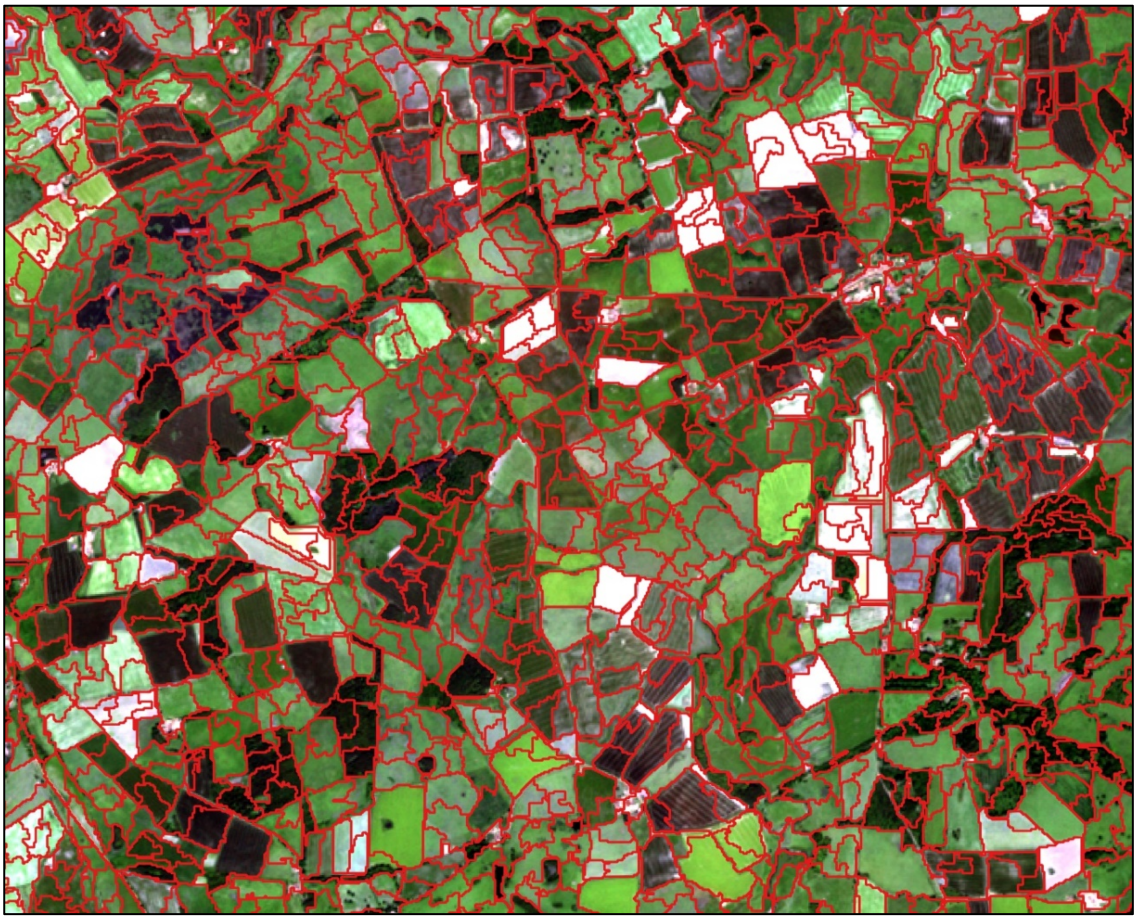

Figure 10. Section of the GRASS segmentation visualised over Sentinel-2 imagery. Image width ca. 6km.

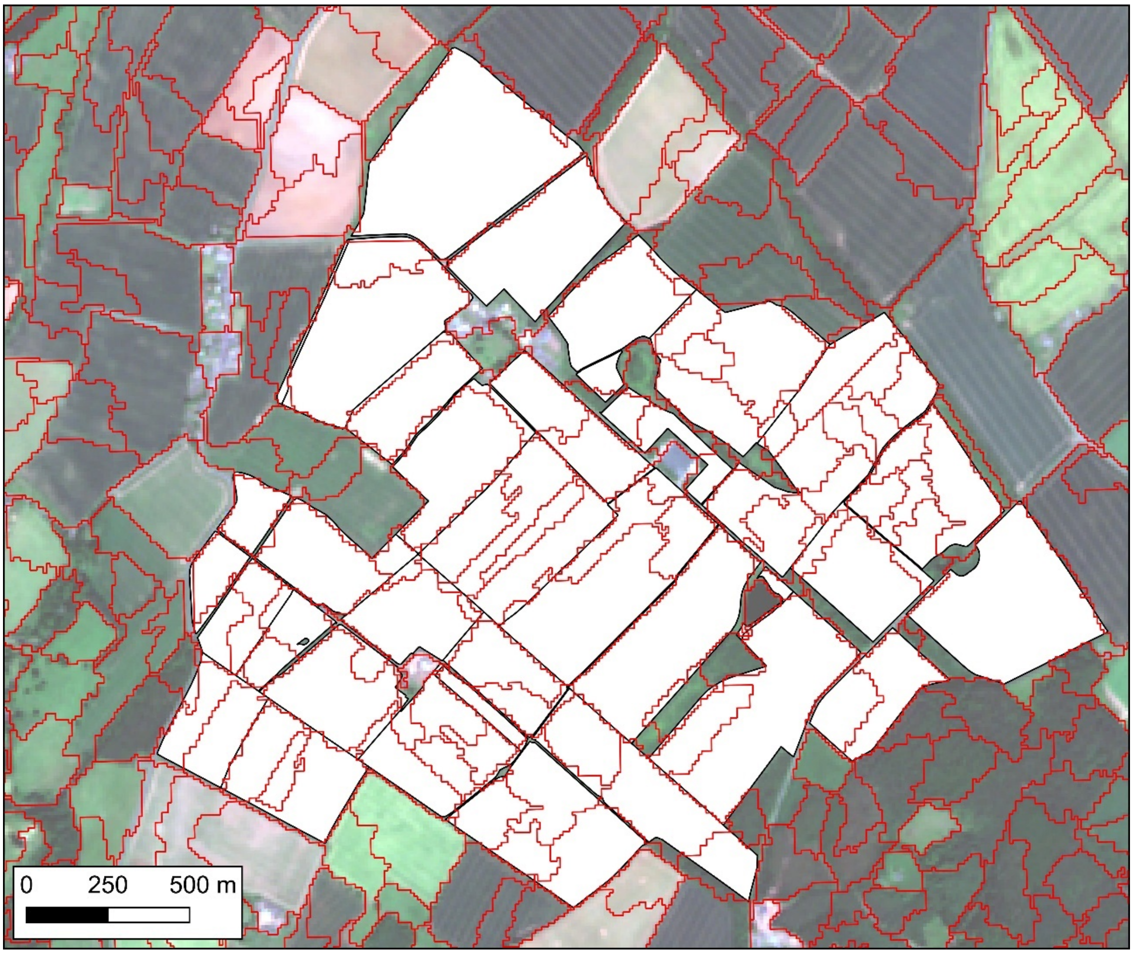

Figure 11. Example of visual comparison between GRASS segmentation and ground truth data.

3.2 Polygon statistics

As Table 2 shows, OTB produced the largest total number of polygons for the Kelso test area, followed by GRASS GIS, with eCognition producing the lowest number. The total number of polygons produced by OTB was consistently high for all of the 29 experimental segmentations, exceeding 90,000 objects in several of the outputs. The total number of polygons produced by the four GRASS GIS i.segment algorithm ranged from 48,705 to 65,093. The five single-level segmentations produced by eCognition generated a lower number of objects than the multi-level hierarchical segmentation, ranging from 6,978 to 9,490. These segmentations used scale parameter 100 or 120, while the multi-level segmentation used sequential scale parameters of 100, 75 and 50 as outlined in Section 2.3.1.

Table 2. Total number of polygons generated by each piece of software for the test area of Kelso, and number of polygons which intersect the ground truth data.

	Ground truth data	eCognition segmentation	OrfeoToolBox segmentation	GRASS GIS segmentation
Total number of polygons	n/a	18,494	83,705	54,053
No. of polygons intersecting ground truth data	413	756	2,062	1,381

The polygons from each segmentation which were selected for the generation of descriptive statistics are shown in Figure 12. This clearly shows that in each case, the area covered by the selected polygons is greater than the area covered by the ground truth polygons. The number of polygons intersecting the ground truth data is shown in Table 2 and their total area is shown in

Table 3. The eCognition segmentation has the largest total area and smallest number of polygons, while the OTB segmentation has the smallest total area and largest number of polygons. The spatial query selects all polygons that intersect the ground truth data, including polygons that just touch or slightly overlap the boundary of a ground truth polygon. This causes the difference between the total area of the ground truth data and that of the overlapping polygons, which is more pronounced in the eCognition segmentation because it contains larger polygons.

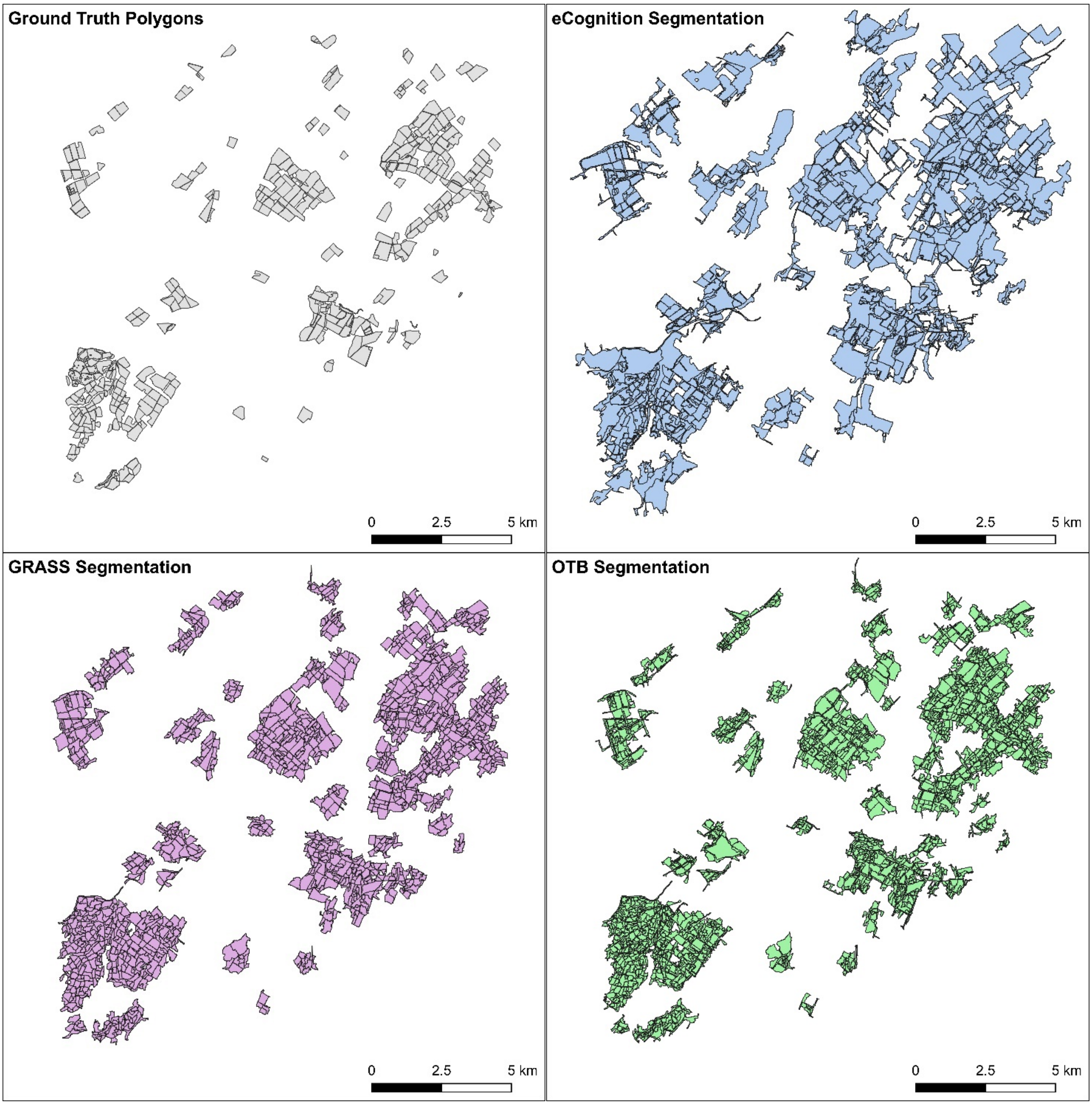

Figure 12. The polygons from each segmentation which were selected due to intersection with the ground truth data polygons (top left).

The descriptive statistics derived from the area in hectares of the ground truth data and the sections of the three segmentations that intersect the ground truth data are shown in Table 3. Mean, median and standard deviation of area in hectares are shown in Figure 13 for each segmentation compared with the ground truth data.

The segmentation produced by **eCognition** had the largest mean value of 11.61ha, which was larger than the mean value of the ground truth polygons. The eCognition segmentation also had by far the largest standard deviation in area values, indicating a wide range of polygon size produced by the three sequential segmentation scales. The median area of the eCognition segmentation is slightly lower than that of the ground truth data, which together with the high standard deviation suggests that a few very large polygons are skewing the mean area value.

The segmentation produced by **OTB** had the smallest mean area of 2.67ha and the lowest standard deviation, suggesting a more consistent segmentation. The statistics derived from the **GRASS GIS** segmentation are similar to those derived from the OTB segmentation, but the GRASS statistics are closer to those of the ground truth polygons. Of the three segmentations evaluated, the GRASS segmentation produced polygons with a mean area and standard deviation closest to those of the ground truth data.

These quantitative data are consistent with the qualitative visual assessment, which showed that at the scale parameters used, the eCognition segmentation occasionally merged spectrally distinct landscape features into single large polygons, while OTB and GRASS tended to split spectrally homogeneous features into several small polygons.

Table 3. Statistics derived from area in hectares of each segmentation output and the ground truth data.

Statistics from polygon area (ha)	Ground truth polygons	eCognition polygons	OrfeoToolBox polygons	GRASS GIS polygons
Minimum	0.51	0.23	0.81	1.01
Maximum	29.96	171.58	50.09	37.04
Range	29.45	171.35	49.28	36.03
Total	3,023	8,779	5,524	6,278
Mean	7.32	11.61	2.67	4.54
Median	6.45	5.84	1.87	3.47
Standard deviation	5.59	18.43	2.86	3.60
Coefficient of variation	0.76	1.587	1.07	0.79
First quartile	2.69	2.67	1.30	2.34
Third quartile	10.42	11.79	2.92	5.33
Interquartile range	7.73	9.02	1.61	2.99

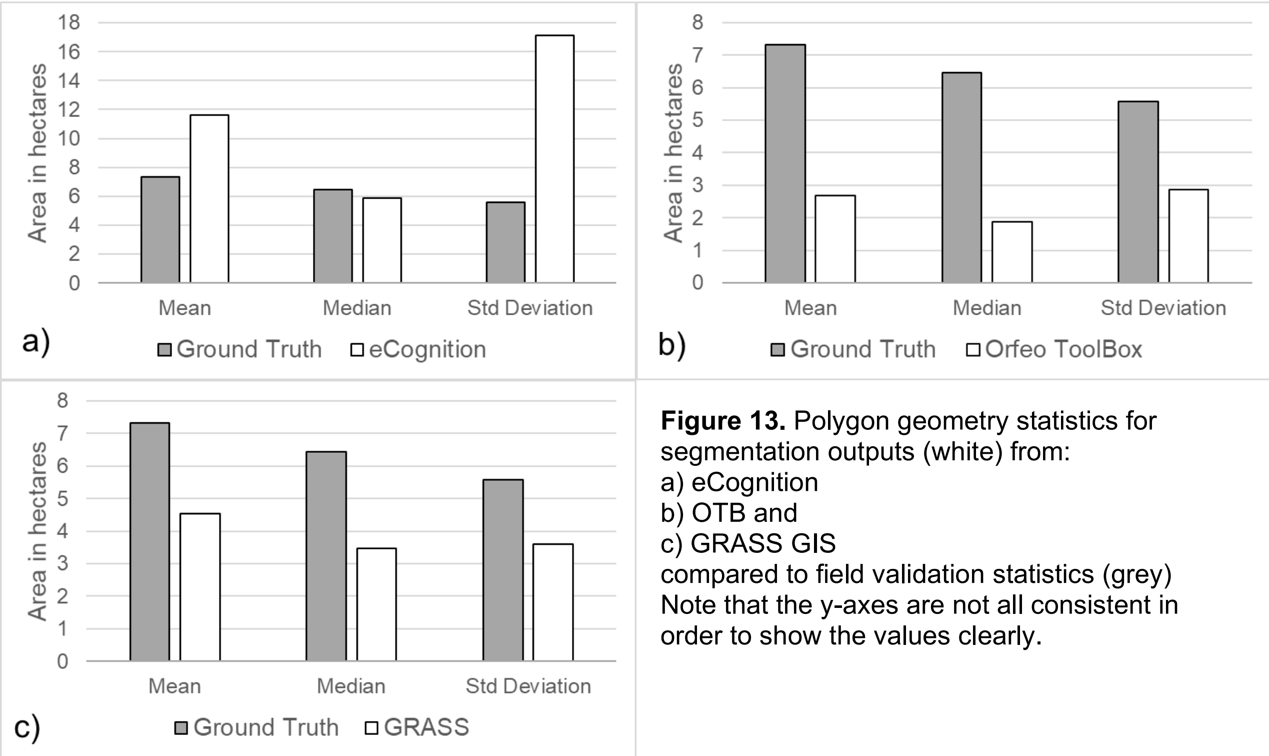

Figure 13. Polygon geometry statistics for segmentation outputs (white) from:
a) eCognition
b) OTB and
c) GRASS GIS
compared to field validation statistics (grey)
Note that the y-axes are not all consistent in order to show the values clearly.

3.3 Processing speed

The time taken to produce the three segmentations selected for comparative evaluation, both for the Kelso test area and for the whole of Scotland, are shown in Table 4. The eCognition segmentation was the quickest to run, taking only three minutes to segment the Kelso test area. This produced a smaller number of objects than the GRASS and OTB segmentations, which took longer to run, but it should also be noted that the eCognition workflow consisted of three sequential segmentations which were all completed in that time. The OTB mean shift algorithm took between four and nine minutes to segment the Kelso test area, depending on the parameters used. The GRASS GIS i.segment algorithm was the slowest to run, and extra time was needed to convert the segmented output from raster to vector format once segmentation was complete.

Table 4. Processing time and number of objects produced by OTB, GRASS and eCognition segmentations of the Kelso test area and the whole of Scotland.

	Kelso test area		Scotland	
	Time (mins:seconds)	**No. of objects**	**Time (to nearest hour)**	**No. of objects**
OTB	4:00 to 9:00	83,705	52	3,148,467
GRASS	11:49 (segment) 2:00 (vectorise)	54,053	48 (segment) 5 (vectorise)	3,957,220
eCognition	3:00	18,494	8	1,202,110

When applied to the mosaic imagery of Scotland, eCognition again performed the fastest segmentation, completing the process without any issues in around 8 hours. In this instance, OTB performed slightly slower than GRASS, taking 52 hours (to the nearest hour) due to the very high number of polygons generated. GRASS took 48 hours to complete the segmentation, but a further five hours to convert the output raster to vector format.

4 Discussion

4.1 Key findings

Each of the software tools produced a segmentation of the small test area and the mosaic imagery of Scotland. However, the tests conducted (visual checks, statistics, processing time) revealed several key observations in the outputs discussed below.

Similarity of results produced by OTB and GRASS GIS is unsurprising since they used the same segmentation algorithm however certain differences can be explained by the process. Whilst the mean-shift segmentation algorithm was employed, the automated steps prior to this are different, such as the tiling of the input image. GRASS GIS employs an anisotriphic filtering technique on the raster image before creating the segmentation. This highlights the subtle discrepancies noted despite the mean-shift segmentation algorithm used in both.

There were observed errors in the stitching of the OTB output (Figure 8). Whilst not consistent throughout the image, it was still prevalent and could produce a misclassification in further steps of an integrated workflow. This error appears to be due to the tiling instigated at the start of the process, where the raster image is divided into a number of tiles. Whilst the stitching option is supposed to remove these errors, some are still noticeable and likely due to a non-perfect histogram matching used by OTB.

Use of greater weighting on the SWIR band was successfully used in eCognition to distinguish woodland and grassland objects, this concurs with Kilcoyne *et al.* (2019) and Colson and Robinson (2019). This demonstrates the importance of the ability to give additional weighting given to bands used in the segmentation. Limitations of this for open-source software are discussed in Section 4.2 as it is not easily integrated for all users when compared to eCognition.

Including brightness and NDVI alongside other bands improved segmentation in eCognition but using NDVI instead of other bands decreased segmentation accuracy in GRASS. There is not a comparative way to seamlessly integrate brightness and NDVI in OTB and GRASS GIS. Limitations of this for open-source software are discussed in Section 4.2. Where eCognition allows for this via the software, if users wanted to achieve a similar workflow in OTB or GRASS GIS additional geospatial manipulation steps would be required. Some users may not feel confident utilising different packages and steps to achieve the same process that is readily available in eCognition.

The field validation data is intended to show field boundaries for agricultural purposes and delineation of crop type. This dataset is not designed to pick out within field features, such as dry patches or lone trees that have been picked up in some areas of each of the segmentations. Therefore, it is to be expected that the image segmentations will have more polygons than the ground truth data.

There is a trade-off between processing time and detail in the outputs. For an ideal segmentation, objects need to be as large as possible but as small as necessary; they must be suitable for the application. For OTB, the parameters of the best run were chosen. This was because although it took more processing time to complete, the result was more detailed, and the segments were more accurate in relation to the features in the input satellite and the field validation polygons. Runs of OTB that produced less detailed results were quicker, but the resulting objects were less detailed in comparison; the algorithm merged areas into large polygons that were not representative of the reality and of no practical use. Where landscape- or catchment-level of detail is required, OTB does not appear to be optimal. This indicates that OTB is better suited for finer mapping, either over large or small input areas of varying resolutions, to capture more subtle and nuanced features with less user manipulation.

4.2 Limitations of study

Living England (Kilcoyne *et al.* 2019) used summer and winter imagery in combination, while this study used only one summer image. In eCognition this approach of using imagery in combination is easily achieved through adding multiple images to a single “project” and running the segmentation on this. Image weightings can also be assigned. To replicate this in OTB and GRASS GIS, raster manipulation is required. There is no option to run a segmentation on multiple images, therefore an image stack needs to be produced. In this instance, a summer and winter 20 band raster dataset would need to be generated. To assign weightings to the assorted image layers in eCognition requires a user to input a figure, for example assigning an image weighting of 1 to the red band and 5 to the SWIR1 band. As stated in Section 4.1, increased weighting of the SWIR band is useful for distinguishing woodland and grassland objects. If the user wanted to assign weightings to an OTB or GRASS GIS segmentation, additional layers would need to be added to the raster stack. In the same example, a single red band would be included in the stack, however five copies of the SWIR1 band would be included. This greatly increases computational storage for running the same process in OTB/GRASS GIS, without guaranteeing the same level of results are generated.

The differences in processing speeds observed are due to different number of polygons created in the different segmentation scales. It should be noted that of the experimental segmentations for OTB that created fewer objects the processing time was much quicker and closer to that of eCognition. However, these segmentations did not visually or statistically compare to the best segmentation that was selected for comparative evaluation.

It was expected that the number of objects would differ between the segmentation outputs and the ground truth polygons as these are generated using different methods: the ground truth polygons were created from field measurements, i.e. they are based on the physical boundaries of the fields rather than the spectral difference as per the segmentations. For comparison only those segmented polygons that overlapped with the ground truth polygons were selected for the statistical analysis. The higher the number of objects generated the more detailed the resulting segmentation output.

4.3 Implications

There was difficulty in replicating segmentation parameters across the three software packages, and therefore it was not a perfectly fair test. Whilst this was attempted to the best of the authors’ abilities, there is no known direct translation available within the literature. Visual assessment showed that eCognition produced mixed objects, but a smaller segmentation scale would have reduced this.

eCognition is an image processing software that has been designed for OBIA; a complete system that has a consistent interface requiring a single installation for the program. The software can be installed on multiple machines with a single licence (as demonstrated by NE/JNCC colleagues), however only one instance of the program can be open at a time. Trimble state requirements for eCognition are a minimum 4GB RAM to run, with 16GB recommended. Optimal performance comes from machines with 32GB or more. Whilst machines with this specification are becoming more prevalent, they are not widely used in the public sector. Traditionally only available for machines running Windows OS, eCognition has recently been made available for machines running Linux OS.

In contrast, OTB and GRASS GIS do not require a licence to operate and are available across machines running Windows, Linux and Mac OS X binaries. Installation is available via a number of routes; standalone installers, integration with QGIS or compiling from source code. There are no specific requirements regarding memory listed for installing this software, however when executing processes (after installation) a message states there is a minimum requirement of 128MB to perform the task. There is no clear indication of requirements for a non-specialist user. Regarding processor optimisation, OTB is designed to natively tile scenes then stitch the tiles at the end of the process. If multiple cores are available on a machine, OTB will instinctively make use of this. eCognition is also able to utilise multiple cores for processing by tiling images, however Trimble state the minimum requirement for cores is a dual-core CPU with recommendations for a quad-core CPU. The software requirements demonstrate that each piece of software is deployable on local machines or via cloud-based virtual machines (VMs). This has been demonstrated by Natural England deploying eCognition on an Amazon Web Services (AWS) Windows-10 VM (Kilycoyne *et al.* 2019) and through JNCC operating OTB via AWS UNIX VMs and OTB/GRASS GIS via QGIS on a JASMIN VM[5].

All three pieces of software offer the ability to be operated from a Graphical User Interface (GUI), however eCognition is the only software designed to be operated instinctively from a GUI. As such, it uses the concept of "projects" as the file structure to open between machines if the workflow is to be shared. Although many of the functions in OTB and GRASS GIS are available through a GUI, through Monteverdi for OTB or QGIS for OTB and GRASS GIS, using a GUI is not required and the features described can be accessed through a command line interface or Python bindings. This presents its own opportunities and challenges. Whilst working in a GUI and being able to visually see workflows might fit their specific requirements better, some users may find there are a number of advantages to a workflow based on Python scripts. Working through Python scripts allows for easier sharing between colleagues if there is only a single commercial licence. For integrated workflows, OTB can be called via the command line or executed via Python. Through packages, such as Reticulate (Ushey *et al.* 2020), OTB can be integrated into R workflows.

Stitching artefacts were most prevalent in the OTB segmentation, followed by the GRASS GIS i.segment segmentation. If this segmentation is to be used in an OBIA classification and the objects are classified as the same habitat and subsequently merged, the artefact will disappear. However, if these segments are assigned a different target classification it will highlight discrepancies and users will not be confident in the output. A hierarchical approach may therefore be more suitable, but this would need to be optimised to the imagery.

Whilst multi-level and hierarchical segmentations are widely utilised in eCognition, for example the Living England method of identifying large homogeneous objects first, there is no like-for-like process available via OTB or GRASS GIS to a non-specialist user. The advantage of a multi-level approach is that it has the benefit of reducing processing time, as the fine scale segmentation only needs to be applied to the areas of imagery which require it. As mentioned, eCognition is an image processing software designed for OBIA and has this functionality readily available for multi-level segmentation. If a user wanted to undertake the same type of workflow in OTB or GRASS GIS, it would require a strong geospatial background and confidence in the toolsets available. The process would be less refined than in eCognition. The user would be required to run a segmentation across the whole image and mask out areas of a certain size iteratively through merging and masking tools, re-running segmentations on objects of certain sizes within the raster dataset and then merging the results from the various outputs at the end. Whilst this approach is valid, it would not necessarily be as efficient and could introduce a number of errors; a multi-level segmentation would be possible however it is not as straightforward or instinctively available as in eCognition.

Segmentations produced by OTB were judged likely to be too detailed for many uses due to the large number of objects created. From visual inspection of the outputs, subtle objects were noted, such as within field dryness. This is suitable for precision agriculture decisions being made at a sub-field level. Whilst this could be considered suitable for Site of Special Scientific Interest (SSSI) management and monitoring at a 'unit' level, other conservation decisions made at landscape or catchment scale do not require this level of detail. This level of detail would lead to slower processing time in later steps of a classification process, such as the Living Map method (Kilcoyne *et al.* 2017) where the assignment of each segment to the target land cover classification is achieved using a random forest machine learning algorithm.

---

[5] http://www.jasmin.ac.uk/services/jasmin-analysis-platform/

4.4 Future research

Further research is recommended to explore the differences in outputs between GRASS GIS and OTB, despite the same parameters being utilised. Discrepancies are likely due to the tiling and stitching process where OTB tiles the image into a number of subsets before mosaicking at the end of the process to produce a seamless output. The i.segment tool in GRASS GIS executes a similar but not exact process, producing a smaller number of image subsets before mosaicking at the end. An additional step is the anisotrophic filtering used by GRASS GIS. This is a step performed to filter the raster image before creating the segmentation. In OTB, the dividing and mosaicking step produced a higher number of stitching errors despite the finer segmentation. The noted discrepancies are likely due to this process however this warrants further research to help understand the intricacies of the process and raised an area of interest for future avenues of research.

The possibility of a hierarchical segmentation in the OS tools is feasible however would take more time than the same approach in eCognition, even with a geospatially confident user. An investigation into this approach is recommended.

Future studies should also evaluate ways to produce more quantitative assessments of the accuracy of segmentation outputs than the methods used in this study, such as the AssessSeg command line tool (Novelli *et al.* 2017).

5 Conclusions

In conclusion, all the software tools performed well and generated acceptable segmentations. However, there are some differences between the tools that must be considered:

eCognition performed well, producing fewer, larger segments than the other software which was closer to the overall number of objects in the ground truth data. This, coupled with the fast processing times, make it ideal for segmenting large input images or mosaics at broader scales. The larger objects produced by scale parameters 100-120 missed some in-field detail, so a smaller scale parameter, or a hierarchical combination of scale parameters, would be needed to produce a reliable segmentation of an area with fine-scale habitat heterogeneity. It should be noted that the high cost of acquiring a working licence of eCognition makes it a less viable solution for projects with limited budgets.

OTB also performed well producing more highly detailed segmentations and was able to process large input images and mosaics relatively well, although the processing time was slower and on occasions, the second part of the algorithm ("stitching") failed to complete or created errors (note that it is possible to deactivate this second part and run without, but another method to stitch the polygons would be required). The outputs were ideal for small-scale, highly detailed segmentation at no monetary cost. OTB is available through an independent GUI and a GUI from within QGIS. For integrated workflows, OTB can be called via the command line or executed via Python. Through packages, such as Reticulate (Ushey *et al.* 2020), OTB can be integrated into R workflows.

GRASS GIS identified the different objects well and produced a segmentation with a reasonable number of objects; more than eCognition and the ground truth data but a lot fewer than OTB. Overall, despite using the same algorithm as OTB, the visual and statistical results of the GRASS GIS output were closer to those of the ground truth polygons. Although the lack of stitching process was an advantage in terms of reliability and quality of the output the additional manual step of converting the resulting raster segmentation into a vector layer added to the processing time. Its free-to-use status makes it a useful option.

A hierarchical approach is often preferred when undertaking segmentations. This is easily developed in eCognition but presents a more challenging workflow in OTB and GRASS GIS. This hierarchical and multi-level approach is better suited to eCognition as it can be done in the software without additional geospatial tools. The hierarchical approach is perfectly suited for habitat mapping from satellite imagery in the UK where there is a mix of large homogeneous habitats and fine scale patchy mosaics. To achieve similar in OTB/GRASS GIS requires either a scripting or manually executed toolset requiring more input from the user, as discussed in Section 4.3. eCognition is much better suited to this commonly used approach.

All the tools performed well and delivered the expected results. eCognition produced fewer, larger segments than the other software with the input parameters. In comparison, the OTB and GRASS GIS outputs produced smaller segments using as similar parameters as possible. This indicates that while eCognition is flexible and suitable for both broad and fine scale segmentations, including combining both in a nested hierarchy, OTB and GRASS GIS are more suited for smaller input images. These two pieces of software have less user-friendly control over the number of outputted segments. Their algorithms can be scaled up for use with larger input images and mosaics, but their processing times increase considerably; in order to reduce this, powerful CPUs are needed. eCognition is a more flexible piece of software given the current constraints. The eCognition Multiresolution Segmentation algorithm performed better with large input images and mosaics making it ideal for projects that focus on covering large geographical areas.

Future work in this area should further investigate the segmentation algorithms now available with the RSGISLIB python library developed by the Department of Geography and Earth Sciences (DGES) at Aberystwyth University (Wales, UK). This library uses the Shepherd Segmentation algorithm (Shepherd *et al.* 2019) which produces a multiband raster segmentation layer. Initial testing and results of which were very promising. The work reported here reviewed the mean shift algorithm in GRASS as it provided a clearer comparison with OTB, but a further test should review the region growing algorithm, to compare with eCognition outputs.

6 References

BARALDI, A., LANG, S., TIEDED, D. & BLASCKE, T. (2018). Earth observation big data analytics in operating mode for GIScience applications. The (GE)OBIA acronym(s) reconsidered. In *GEOBIA 2018 – From pixels to ecosystems and global sustainability*.

BLASCHKE, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*. **65**, pp. 2-16

BLASCHKE, T., HAY, G.J., KELLY, M., LANG, S., HOFMANN, P., ADDINK, E., QUEIROZ FEITOSA, R., van der MEER, F., van der WERFF, H. & van COILLIE, F. (2014). Geographic Object-Based Image Analysis – Towards a new paradigm. *ISPRS Journal of Photogrammetric Remote Sensing*. **87**, pp. 180-191.

COLSON, D. & ROBINSON, P. (2019). Northern Ireland Mapping Project: Mapping Fermanagh, Tyrone and Londonderry. JNCC Report, JNCC, Peterborough.

DE LUCA, G., SILVA, J.M.N., CERASOLI, S., ARAUJO, J., CAMPOS, J., DI FAZIO, S. & MODICA, G. (2019). Object-Based Land Cover Classification of Cork Oak Woodlands using UAV Imagery and Orfeo ToolBox. *Remote Sensing*. **11 (10)**, pp. 1238-1260.

EUROPEAN SPACE AGENCY. (2019). *Sentinel-2 Overview.* [online] Available at: https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-2/overview [Accessed 03 Dec 2019]

FOODY, G. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment,* **80(1)**, pp. 185-201

GITAS, I., SANTIS, A. & MITRI, G. (2009). Remote Sensing of Burn Severity. In: E. Chuvieco, ed., *Earth Observation of Wildland Fires in Mediterranean Ecosystems*, 1st ed. Berlin: Springer-Verlag, pp.129-148

GRIZONNET, M., MICHEL, J., POUGHON, V., INGLADA, J., SAVINAUD, M. & CRESSON, R. (2017). Orfeo ToolBox: open source processing of remote sensing images. *Open Geospatial Data, Software and Standards*. **2 (15)**, pp. 1-8.

JONES, G., COLSON, D., ROBINSON, P., HARRIS, M. & SMITH, M. (2019). WP7 Sustainable Livelihoods: Peruvian Case Study: Mapping the Viru Valley. Project No 417000001416

JONES, T., WICKS, D., AGASS, S. & BUNTING, P. (2017). Developing standards and automated production for Sentinel-2 Analysis Ready Data. Evidence Project SD1707

FUKUNAGA, K. & HOSTETLER, L. (1975). The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Transactions on Information Theory,* **21 (1)**, pp. 32-40.

KILCOYNE, A.M., ALEXANDER, R., COX, P. & BROWNETT, J. (2017). *Living Maps: Satellite-based Habitat Classification*. *Evidence Project SD1705.* Available online at: http://randd.defra.gov.uk/Document.aspx?Document=14198_SD1705_FinalReport.pdf

KILCOYNE, A.M., PICTON PHILLIPPS, G.P., COX, P., BROWNETT, J., KEANE, R. & COLSON, D. (2019). *Living England: Satellite-based habitat classification. Phase I.* Natural England.

LILLESAND, T.R., KIEFER, R.W. & CHIPMAN, J.W. (2008) Remote Sensing and Image Interpretation, Sixth edn., New Jersey: John Wiley & Sons, Inc.

LENNERT, M., GRIPPA, T., RADOUX, J., BASSINE, C., BEAUMONT, B., DEFOURNY, P. & WOLFF, E. (2019). Creating Wallonia's new very high-resolution land cover maps: Combining

GRASS GIS OBIA and OTB Pixel Based Results. In: *FOSS4G 2019, 26-30 Aug 2019, Bucharest.* Bucharest, Remote Sensing and Spatial Information Sciences, pp. 151-157.

LOBO, A. (1997). Image segmentation and discriminant analysis for the identification of land cover units in ecology. *IEEE Transactions on Geoscience and Remote Sensing.* **35** (**5**), pp. 1136-1145.

MATHER, P. & KOCH, M. (2010). *Computer Processing of Remotely-Sensed Images*. Chichester: John Wiley & Sons.

MOMSEN, E. & METZ, M. (2012). *i.segment – GRASS GIS manual*. [online]. Available at: https://grass.osgeo.org/grass78/manuals/i.segment.html [Accessed 12 Nov 2019]

NOVELLI, A., AGUILAR, M.A., AGUILAR, F.J., NEMMAOUI, A. & TARANTINO, E. (2017) *AssessSeg – A command line tool to quantify image segmentation quality: a test carried out in Southern Spain from satellite imagery.* Remote Sensing **9 (1)**, pp. 40.

ORFEOTOOLBOX (2018). *The OrfeoToolBox Software Guide.* [online]. Available at: https://www.orfeo-toolbox.org/SoftwareGuide/SoftwareGuidech16.html#x45-24500016.2.1 [Accessed 03 Apr 2020]

SCOTTISH GOVERNMENT. (2019). *Scotland Crop Map.* [online]. Scottish Government and Edinburgh University. Available at: http://data.cropmap.edina.ac.uk/cropmap/ [Accessed 12 Apr 2020]

SHAPIRO, L. & ROSENFELD, A. (1992). *Computer vision and image processing*. San Diego: Academic Press.

SHEPHERD, J.D., BUNTING, P. & DYMOND, J.R. (2019). Operational Large-Scale Segmentation of Imagery Based on Iterative Elimination. *Remote Sensing*. 11 (6), pp. 658-678.

STOIAN, A., POULAIN, V., INGLADA, J., POUGHON, V. and DERKSEN, D. (2019). Land Cover Maps Production with High Resolution Satellite Image Time Series and Convolutional Neural Networks: Adaptations and Limits for Operational Systems *Remote Sensing*. 11 (1986), pp. 1-26.

TARABALKA, Y., BENEDIKTSSON, J. & CHANUSSOT, J. (2009). Spectral-Spatial Classification of Hyperspectral Imagery Based on Partitional Clustering Techniques. *IEEE Transactions of Geoscience and Remote Sensing*. **47 (8)**, pp. 2973-2987.

TEODORO, A.C. & ARAUJO, R. (2016). Comparison of performance of object-based image analysis techniques available in open source software (Spring and Orfeo ToolBox/Monteverdi) considering very high spatial resolution data. *Journal of Applied Remote Sensing.* **10 (1)**, pp. 11-16.

TORRES-SÁNCHEZ, J., LÓPEZ-GRANADOS, F. & PEÑA, J.M. (2015). An automatic object-based method for optimal thresholding in UAV images: Application for vegetation detection in herbaceous crops. *Computing and Electronic Agriculture*. **114**, pp.43-52.

TRIMBLE GEOSPATIAL INC. (2020). *What is eCognition?* [online] Available at: https://geospatial.trimble.com/what-is-ecognition [Accessed 12 Dec 2019]

TRIMBLE GERMANY GmbH. (2016) *eCognition Developer Reference Book*. Munich, Germany, Trimble Germany GmbH.

USHEY, K., ALLAIRE, J.J., TANG, Y., EDDELBUETTEL, D., LEWIS, B., KEYDANA, S., HAFEN, R. & GEELNARD, M. (2020). *Reticulate* (1.15). [Software]. [Accessed 28 April 2020]

WITHARANA, C., CIVCO, D. L. & MEYER, T.H. (2014). Evaluation of data fusion and image segmentation in earth observation based rapid mapping workflows. *ISRPS Journal of Photogrammetry and Remote Sensing*. **87**, pp. 1-18.

Appendix 1

Descriptions of each of the input parameters available in each piece of software.

Multiresolution Segmentation algorithm of eCognition.

Table 7.1. eCognition multiresolution algorithm description of parameters.

Parameter	Description
Image Layer Weights	Image layers can be weighted depending on their importance or suitability for the segmentation result.
Compatibility mode	Compatibility mode to previous software versions (eCognition Version 8.64 or earlier)
Thematic Layer Usage	Specify the thematic layers to be candidates for segmentation. Each thematic layer that is used for segmentation will lead to additional splitting of image objects while enabling consistent access to its thematic information.
Scale Parameter	The Scale Parameter is an abstract term that determines the maximum allowed heterogeneity for the resulting image objects.
Shape	The value of the Shape field modifies the relationship between shape and colour criteria.
Compactness	The compactness criterion is used to optimize image objects with regard to compactness. This criterion should be used when different image objects which are rather compact but are separated from non-compact objects only by a relatively weak spectral contrast

Mean shift segmentation algorithm parameters of OrfeoToolBox

Table 7.2. OrfeoToolBox mean shift algorithm description of parameters

Parameter	Description
Spatial radius	Spatial radius of the neighbourhood
Range radius	Range radius defining the radius (expressed in radiometry unit) in the multispectral space
Mode convergence threshold	Algorithm iterative scheme will stop if mean shift vector is below this threshold or if iteration number reached maximum number of iterations
Maximum number of iterations	Algorithm iterative scheme will stop if convergence hasn't been reached after the maximum number of iterations
Minimum region size	Minimum size of a region (in pixel unit) in segmentation. Smaller clusters will be merged to the neighbouring cluster with the closest radiometry.

i.segment mean shift segmentation parameters of GRASS GIS.

Table 7.3. GRASS GIS i.segment description of parameters.

Parameter	Description
group	Name of input imagery group or raster maps
output	Name for output raster map
band_suffix	Suffix for output bands with modified band values
threshold	Difference threshold between 0 and 1
radius	Spatial radius in number of cells
hr	Range (spectral) bandwidth [0, 1]
method	Segmentation method Options: *region_growing, mean_shift*
similarity	Similarity calculation method Options: *euclidean, manhattan*
minsize	Minimum number of cells in a segment
memory	Memory in MB
iterations	Maximum number of iterations
seeds	Name for input raster map with starting seeds
bounds	Name of input bounding/constraining raster map
goodness	Name for output goodness of fit estimate map

Appendix 2

Tables showing the values for each parameter used in the input for each of the software suites.

Table 8.1. eCognition values for each input parameter and layer weighting used in the single-level segmentations. The settings which produced the best of the single-level segmentations (based on visual assessment) are bordered in bold.

		Segmentation outputs				
		LM_small	LM_large	Test1	Test2	Test3
Weighting of layers used in segmentation	Blue	0	1	1	1	1
	Green	0	1	1	1	1
	Red	1	5	1	1	1
	Red Edge 704	1	1	1	1	1
	Red Edge 740	1	1	1	1	1
	Red Edge 780	1	1	1	1	1
	NIR	2	5	1	1	1
	Narrow NIR	1	5	1	1	1
	SWIR 1	3	2	10	10	10
	SWIR 2	3	2	10	10	10
	NDVI	0	0	5	5	5
	Brightness (RGB)	0	0	10		
	Brightness (RE-NIR)	0	0	5		
	Brightness (all)	0	0	0	10	10
Segmentation parameters	Scale	120	120	120	120	100
	Shape	0.3	0.1	0.1	0.1	0.1
	Compactness	0.2	0.8	0.5	0.5	0.5

Table 8.2. OrfeoToolBox values for each input parameter. The settings which produced the segmentation used in the comparative analysis are bordered in bold.

Segmentation outputs	Parameters					
	Spatial Radius	**Range Radius**	**Mode Conv. Thresh**	**Max no. Iterations**	**Min. Region Size**	**Stitch Polygons**
1	4	200	0.1	100	10	1
2	4	200	0.1	100	4	1
3	4	150	0.1	100	4	1
4	4	175	0.1	100	4	1
5	4	175	0.1	100	2	1
6	8	175	0.1	100	2	1
7	5	175	0.1	100	100	1
8	5	15	0.1	100	100	1
9	5	150	0.1	100	100	1
10	5	125	0.1	100	100	1
11	5	100	0.1	100	100	1
12	5	75	12	100	100	1
13	5	50	0.1	100	100	1
14	5	25	0.1	100	100	1
15	5	35	0.1	100	100	1
16	5	25	0.1	100	10	1
17	5	25	0.1	100	20	1
18	5	25	0.1	100	30	1
19	5	25	0.1	100	40	1
20	5	25	0.1	100	50	1
21	5	25	0.1	100	60	1
22	5	25	0.1	100	70	1
23	5	25	0.1	100	80	1
24	5	25	0.1	100	90	1
25	5	25	0.1	100	100	1
26	5	25	0.1	100	110	1
27	5	25	0.1	100	120	1
28	5	25	0.1	100	130	1
29	5	25	0.1	100	140	1

Table 8.3. GRASS GIS values of each input parameter tested. The settings which produced the segmentation used for comparative analysis are bordered in bold.

Segmentation outputs	Parameters					
	Spatial Radius	Range radius	Mode Conv. Thresh	Max no. Iterations	Min. Region Size	No. of objects
23	5	25	0.1	100	80	54,053
24	5	25	0.1	100	90	60,880
12	5	75	0.1	100	100	48,705
NDVI	5	25	0.1	100	80	65,093