



## JRC TECHNICAL REPORTS

# The Canopy Health Monitoring (CanHeMon) project

Pieter S. A. Beck, Laura Martínez Sanchez, Margherita di Leo, Yann Chemin, Giovanni Caudullo, Begoña de la Fuente, Pablo J. Zarco-Tejada

2019



This publication is a Technical report by the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policymaking process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use that might be made of this publication.

**Contact information**

Name: Pieter S. A. Beck  
Address: Via E. Fermi 2749 – TP 261, I-21027 Ispra (VA), Italy  
Email: [pieter.beck@ec.europa.eu](mailto:pieter.beck@ec.europa.eu)  
Tel.: +39 0332 78 3671

**EU Science Hub**

<https://ec.europa.eu/jrc>

JRC115400

EUR 29652 EN

PDF ISBN 978-92-79-99639-9 ISSN 1831-9424 doi:10.2760/38697

Luxembourg: Publications Office of the European Union, 2019

© European Union, 2019

The reuse policy of the European Commission is implemented by Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39). Reuse is authorised, provided the source of the document is acknowledged and its original meaning or message is not distorted. The European Commission shall not be liable for any consequence stemming from the reuse. For any use or reproduction of photos or other material that is not owned by the EU, permission must be sought directly from the copyright holders.

All content © European Union, 2019

How to cite this report: Beck, P. S. A, L. Martínez Sanchez, M. Di Leo, Y. Chemin, G. Caudullo, B. de la Fuente, P. J. Zarco-Tejada, *The Canopy Health Monitoring (CanHeMon) project*, EUR 29652 EN, Publications Office of the European Union, Luxembourg, 2019, ISBN 978-92-79-99639-9, doi:10.2760/38697, JRC115400

# Contents

Foreword .....	3
Acknowledgements .....	4
Abstract .....	5
1 Introduction .....	6
2 Remote sensing data.....	7
2.1 RGB-NIR images with 15 cm resolution near Castelo Branco -autumn 2015 .....	7
2.2 Hyperspectral images of 'hotspots' in Castelo Branco - spring 2016.....	7
2.3 RGB-NIR images with 15 cm resolution near Castelo Branco -autumn 2016 .....	7
2.4 RGB-NIR images with 30 cm resolution, entire buffer zone .....	8
3 Complementary geospatial data .....	10
3.1 Carta de Uso e Ocupação do Solo de Portugal Continental (COS) .....	10
3.2 RGB-NIR images with 50 cm resolution near Castelo Branco .....	10
4 Interpretation of data.....	11
5 Image interpretation algorithms .....	17
5.1 Algorithm aim and challenges.....	17
5.2 Image pre-processing .....	17
5.3 Algorithm choice.....	18
5.4 Algorithm input .....	18
5.5 Algorithm settings .....	18
5.6 Algorithm tuning.....	20
6 Image analysis hardware.....	21
7 Processing load and time .....	22
8 Accuracy assessment of image interpretation .....	23
8.1 Sample-based validation and iterative model calibration .....	23
8.2 Image analysis results.....	23
9 Mobile data exchange platform.....	34
9.1 Installation .....	35
9.2 Authentication – Login Page .....	36
9.3 Visualisation of different screens.....	37
9.3.1 Sequence of sliding pages .....	38
9.3.2 Select Geographical Zone (GZ) View .....	39
9.3.3 Select Area of Interest (AOI) View .....	40
9.3.4 Select AOI – Part 1 .....	41
9.3.5 Select AOI – Part 2 .....	42
9.3.6 Select AOI – Part 3 .....	42
9.3.7 Select AOI – Part 4 .....	43

9.3.8	AOI actions .....	44
9.3.9	AOI List view .....	45
9.3.10	Menu view .....	46
9.3.11	Menu actions .....	47
9.3.12	Profile view .....	48
9.3.13	Map view (Survey view) .....	49
9.3.14	Map view actions: load legend and change layer visibility .....	50
9.3.15	Map view actions: load information by clicking on data icon .....	51
9.3.16	Map view actions: go to Menu .....	51
9.3.17	Map view actions: add data to survey .....	52
9.3.18	Data views (multiple): List view .....	53
9.3.19	Data views (multiple): List view – Actions .....	55
9.3.20	Data views (multiple): Create view .....	56
9.3.21	Data views (multiple): Create view – Actions .....	57
9.3.22	Detail view .....	58
9.3.23	Delete view .....	59
9.4	External server .....	60
9.4.1	Data typology in application workflow .....	60
9.4.2	Data storage .....	64
9.4.3	Application-server communication .....	64
9.4.4	Remote Server and deployment .....	67
10	What to consider when developing similar systems .....	68
10.1	Every case is different .....	68
10.2	Platform choice .....	68
10.3	Metadata matters .....	68
10.4	Non-image data .....	69
10.5	Making use of existing data .....	69
10.5.1	Orthophoto archives .....	69
10.5.2	Land cover maps .....	69
10.6	New generation of image analysis algorithms .....	70
11	Meetings .....	72
	References .....	73
	List of figures .....	74
	List of tables .....	75

## **Foreword**

The Canopy Health Monitoring (CanHeMon) project ran at the Joint Research Centre of the European Commission from mid-2015 to mid-2018 and was funded by DG SANTE. This report provides technical details about the work carried out during the project, while a separate Science for Policy report provides a summary of the main findings and a synoptic outlook on the potential use of remote sensing to support plant health policy in the EU.

## **Acknowledgements**

We are grateful to the Instituto da Conservação da Natureza e das Florestas of Portugal (ICNF) for their feedback throughout the CanHeMon project and for access to an archive of aerial photographs from 2012. We thank the JRC Earth Observation Data and Processing Platform (JEODPP) team for their support in hosting and processing remote sensing data.

## ***Authors***

Pieter S. A. Beck, Laura Martinez Sanchez, Margherita di Leo, Yann Chemin, Giovanni Caudullo, Begoña de la Fuente Martín, Pablo J. Zarco Tejada

## **Abstract**

As part of the EU emergency measures against the pine wood nematode (PWN) (*Bursaphelenchus xylophilus*) laid down under Decision 2012/535/EU, Portugal should perform, outside and during the flight season of the vector, surveys of coniferous trees located in the 20 km wide buffer zone established along the Spanish border, with the aim to detect trees which are dead, in poor health or affected by fire or storm. These trees shall be felled and removed, as required by the Decision, to avoid that they act as attractants for the longhorn beetle (*Monochamus* species), the vector responsible for the spread of PWN. The CanHeMon project tasked the Joint Research Centre with analysing a portion of the buffer zone, using remote sensing data, to support detection on the ground of declining pine trees. During the project, a 400 km<sup>2</sup> area was imaged twice, in autumn 2015 and autumn 2016, at 15 cm resolution from aircraft, and individual declining tree crowns were detected using a MaxEnt-based iterative image analysis algorithm, the performance of which was gauged through visual photointerpretation. The scalability of the automated methods was then tested using an image mosaic of the entire buffer zone at 30 cm resolution. Finally, broad recommendations were formulated on the use of remote sensing for large-area surveys in the context of plant health emergencies.

## 1 Introduction

The Canopy Health Monitoring (CanHeMon) project ran at the Joint Research Centre of the European Commission from mid-2015 to mid-2018 and was funded by DG SANTE. DG SANTE is responsible, among other things, for the European Union's plant health legislation, which aims to put in place effective measures to protect the Union's territory and its plants, as well as ensuring that trade is safe and that the impacts of climate change on the health of EU crops and forests are mitigated.

For specific harmful organisms that threaten its crops and forests, the EU takes emergency control measures. The pine wood nematode (*Bursaphelenchus xylophilus*) is such a quarantine pest. It can kill European coniferous tree species and has been spreading through Portugal since the end of the 1990s.

As part of the emergency measures against the pine wood nematode (Decision 2012/535/EU), Portugal should perform, outside and during the flight season of the vector, surveys of the coniferous trees located in the 20 km buffer zone established along the Spanish border with the aim to detect trees which are dead, in poor health or affected by fire or storm. According to the Decision, these trees shall be felled and removed each winter lest they attract the longhorn beetle (*Monochamus* species) that is responsible for spreading the nematode. Monitoring such a large area (greater than 20 000 km<sup>2</sup>) for individual trees in declining health is a daunting task, particularly as large stretches of the area are hard to reach with vehicles or on foot. The aim of the CanHeMon project was to demonstrate how remote sensing could support this monitoring task, by collecting and analysing remote sensing data for the area to pinpoint individual coniferous trees in declining health.

The remote sensing work needed to adhere to a tight time schedule, as trees may not decline and attract longhorn beetles until late summer or early autumn, meaning remote sensing data should depict the state of the trees in the autumn. Then, the images had to be processed and analysed within weeks for the results to be useful for implementing the emergency measures. These measures namely stipulate that all declining trees must be removed by the end of April.

## **2 Remote sensing data**

To carry out the CanHeMon project, we acquired four airborne remote sensing datasets.

### **2.1 RGB-NIR images with 15 cm resolution near Castelo Branco - autumn 2015**

The core study area of CanHeMon was close to Castelo Branco, in a 400 km<sup>2</sup> portion of the buffer zone closest to the PWN outbreak at the time. RGB-NIR images with 15 cm spatial resolution were procured and acquired specifically for this project on 26, 27 and 28 November 2015 in order to image the entire study area. The contractor used an Intergraph DMC photogrammetric camera system, mounted on a gyrostabilised platform T-AS Mount on board a Cessna 402B. It was deployed alongside a Daedalus sensor which imaged at 180 cm ground resolution in 15 narrow optical bands (nine visible, four near-infrared, and two mid-infrared), as well as measuring thermal emissivity.

Covering the entire area required 26 flight lines, with the aircraft flying in a north-south direction, spaced to ensure sufficient side-by-side overlap between images of adjacent lines in order to ensure precise mapping. Images were only acquired under bright-sky conditions, and between 11.15 and 14.00 local time to ensure sufficient illumination. Images were ordered with absolute spatial accuracy (RMSE) of < 1.5 m, with 95 % of pixels located with less than 3.5 m error. Band-to-band co-registration in the images had to be less than 0.5 pixels.

Image mosaics were ordered as 4-band GeoTIFF files, as a mosaic of the entire study area split into 1 km by 1 km tiles, with the tile boundaries coinciding with the 1 km integer coordinates in the EPSG:32629 projection and the band order Blue-Green-Red-NIR, and projected in EPSG:32629.

The composite images created from the orthophotos comprised circa 140 GB. These data can be accessed through the web mapping service WMS (<sup>1</sup>).

### **2.2 Hyperspectral images of 'hotspots' in Castelo Branco - spring 2016**

To investigate 13 specific areas where RGB-NIR images indicated declining trees in November 2015, a small campaign was carried out on 18 May 2016 to acquire hyperspectral and multispectral images of declining trees.

The hyperspectral images covered 200 spectral bands of 10 nm full width at half maximum (FWHM) in the region 400-800 nm, with 65 cm spatial resolution.

The multispectral images had six bands centred at 450 nm, 490 nm, 550 nm, 670 nm, 710 nm and 800 nm, with 10 nm FWHM and 25 cm spatial resolution. The images were radiometrically calibrated and atmospherically corrected to surface reflectance.

### **2.3 RGB-NIR images with 15 cm resolution near Castelo Branco - autumn 2016**

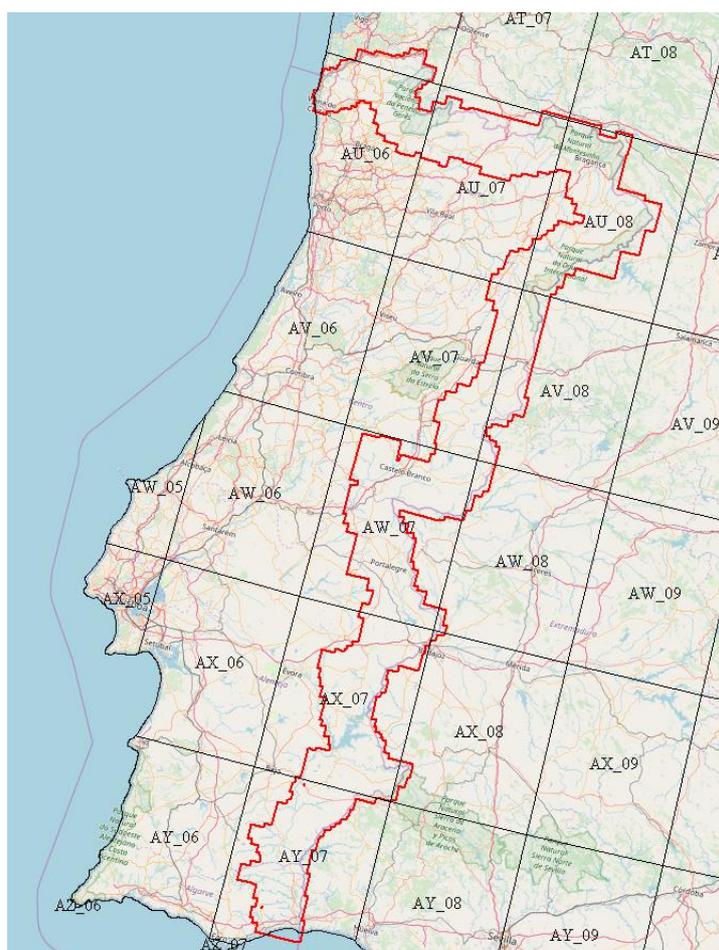
The RGB-NIR image collection of autumn 2015 over the Castelo Branco area was repeated in the autumn of 2016 using exactly the same acquisition scheme with flights.

These data can be accessed through the web mapping service WMS.

---

(<sup>1</sup>) <http://cidportal.jrc.ec.europa.eu/jeodpp/services/ows/wms/canhemon/portugal>

**Figure 1.** Map of Portugal, and outline in red of the 30 cm RGB-NIR images acquired by the project. The grid depicts the 'blocks' in which these data were generated, delivered, and analysed.



## 2.4 RGB-NIR images with 30 cm resolution, entire buffer zone

RGB-NIR orthophotos were procured for the entire PWN buffer zone (circa 23 000 km<sup>2</sup>). Given the size of the area, the short duration of the project, and the budget, it was impossible to carry out a dedicated airborne campaign to image the entire buffer zone. Instead, images from an existing campaign were sought. A mosaic of images with 30 cm spatial resolution, acquired between 7 December 2014 and 26 June 2016, was ultimately purchased. The data were acquired and processed, as part of a multinational aerial imaging campaign, in 1° x 1° 'blocks'.

**Table 1.** Start and end dates of the periods during which images were acquired to generate blocks of 30 cm RGB-NIR images.

Block	Start Capture [Date]	End Capture [Date]	Block	Start Capture [Date]	End Capture [Date]
AU_06	29/05/2015	29/05/2015	AW_08	30/05/2015	31/05/2015
AU_07	07/09/2015	15/10/2015	AX_05	16/05/2015	28/05/2015
AU_08	05/09/2015	26/06/2016	AX_06	06/12/2014	07/12/2014
AV_06	16/05/2015	17/05/2015	AY_06	01/12/2014	14/05/2015
AV_07	15/05/2015	16/05/2015	AY_07	07/12/2014	29/11/2015
AV_08	31/05/2015	23/07/2016	AZ_06	01/12/2014	14/05/2015
AW_05	16/05/2015	28/05/2015	AZ_07	07/12/2014	29/11/2015
AW_06	17/05/2015	24/08/2015	AX_07	08/12/2014	11/12/2014
AW_07	28/05/2015	05/09/2015			

The images were acquired using Leica ADS100 sensors which work on the principle of a push broom scanner and therefore record data continuously, while the aircraft is moving forward. The combination of the recorded lines (nadir, forward and backward) forms an image strip.

Images were produced following state-of-the-art photogrammetric procedures, to produce image mosaics with minimal spatial discontinuities, in colours that are unrelated to the land surface. This is challenging since extended image acquisition campaigns are needed to cover large areas such as the PWN buffer zone. As a result, the images can be from different seasons and years and thus have very different spectral characteristics.

Besides the standard image mosaics, which were produced in 8-bit, a 16-bit surface reflectance image mosaic was produced. This mosaic, which prioritises physical precision (quantifying the proportion of incoming electromagnetic radiation reflected by the land surface in  $W \cdot sr^{-1} \cdot m^{-2} \cdot nm^{-1}$ ), is not commonly produced in the photogrammetric community. However, it allows for a better comparison with satellite images.

Acquisition of the images, and associated intellectual property, was governed by service contract number 930705 between the EU and COWI A/S.

Given the large volume, the data were initially delivered in seven stages, with the final project data delivered to the JRC on 26 February 2018, after being released by the Portuguese authorities. Because errors in the data and metadata were detected by the JRC, as of July 2018, the data are being corrected by COWI A/S.

### 3 Complementary geospatial data

Complementary geospatial data can greatly aid the analysis of up-to-date remotely sensed images. They can provide contextual information that is not contained in, or easily extracted from, the remote sensing data. Here, two different types of data were used: one thematic and one image. The thematic map of land use in the study area allowed us to restrict the analysis to areas where coniferous trees may be present; this greatly improves processing speed, as less data need to be handled by the image interpretation. In some cases, it also significantly improves image classification, as the algorithm has fewer data combinations to learn to distinguish.

In addition, older orthophotos (in this case from 2012), where available, allowed us to identify many areas where coniferous trees were too young and small for reliable health evaluation through image analysis. These areas included plantations that were barren (unplanted) in 2012, and recognisable as such in the 2012 orthophotos, but densely planted with saplings in 2015 and 2016. As they are plantations, such areas were not excluded from analysis based on the COS (see below), and they do have abundant vegetation in the new images; however the small size of the trees, and sometimes the abundance of other vegetation, prevents reliable analysis. These areas were identified in the imagery by comparing vegetation indices (mathematical combinations of observations in different spectral bands that highlight vegetation) in different years; areas that had an extremely low vegetation index in 2012 (indicating they were barren), but a high one in 2016 or 2017, could not be covered by mature trees in 2016, and were excluded from the image classification. Below, we provide more details on the two auxiliary geospatial datasets.

Besides the aerial images analysed for this project, the following complementary geospatial data were used.

#### 3.1 Carta de Uso e Ocupação do Solo de Portugal Continental (COS)

The land cover and land use map of mainland Portugal (Carta de Uso e Ocupação do Solo de Portugal Continental, COS) at its highest level of detail (level 5) distinguishes circa 2 000 land cover classes with a minimum mapping unit of 1 ha [1] [2]. It was produced through the visual interpretation of orthophotos alongside ancillary data. Its estimated horizontal accuracy is 5.5 m. Further information about the COS is available at:

[http://www.dgterritorio.pt/dados\\_abertos/cos](http://www.dgterritorio.pt/dados_abertos/cos)

At its finest level, the COS distinguishes for example between '*Florestas de pinheiro bravo*' and '*Florestas de pinheiro manso*', which are *Pinus pinaster* and *Pinus pinea* forests, respectively. It also identifies multiple forest types with dominant, co-dominant, or secondary presence of coniferous tree species.

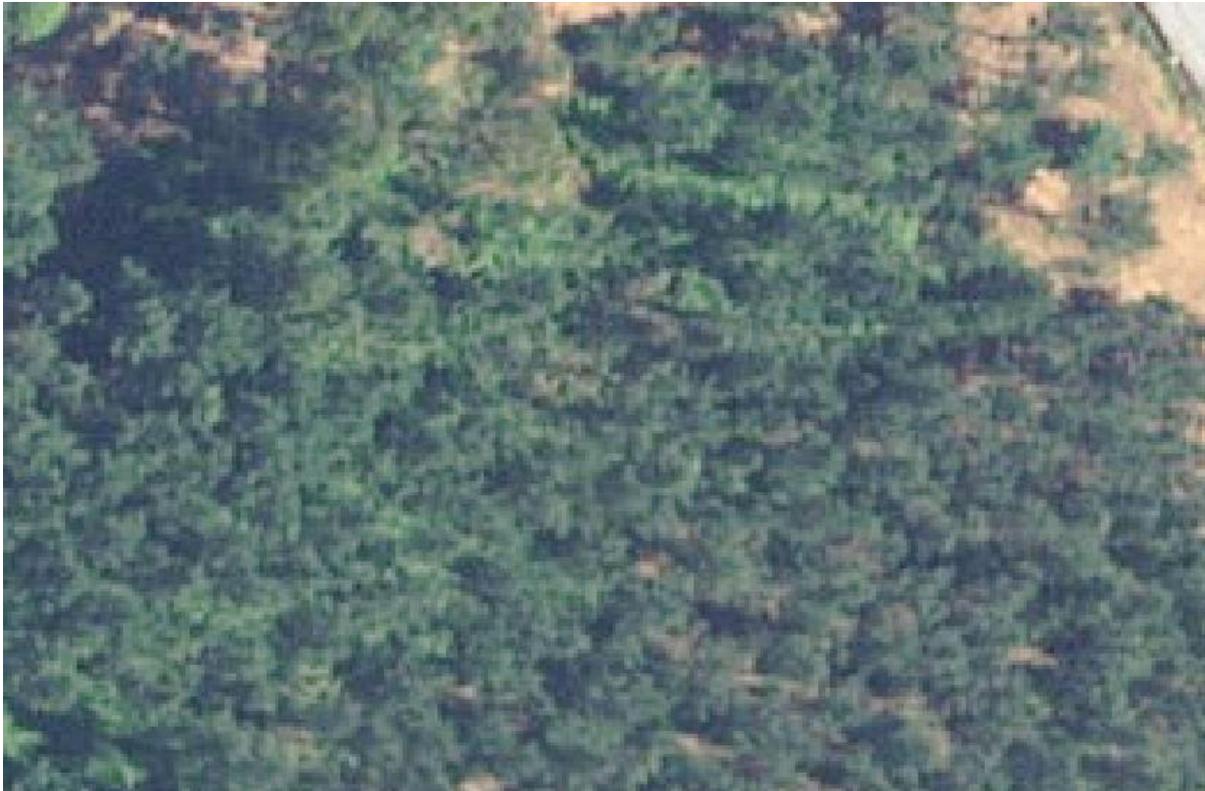
#### 3.2 RGB-NIR images with 50 cm resolution near Castelo Branco

RGB-NIR aerial images with 50 cm resolution near Castelo Branco were obtained from the ICNF for use in the CanHeMon project in June 2015. The data comprise circa 144 GB and the conditions for sharing them were set out in the ICNF letter 27935/2015/DGACPPF/DPFVAP.

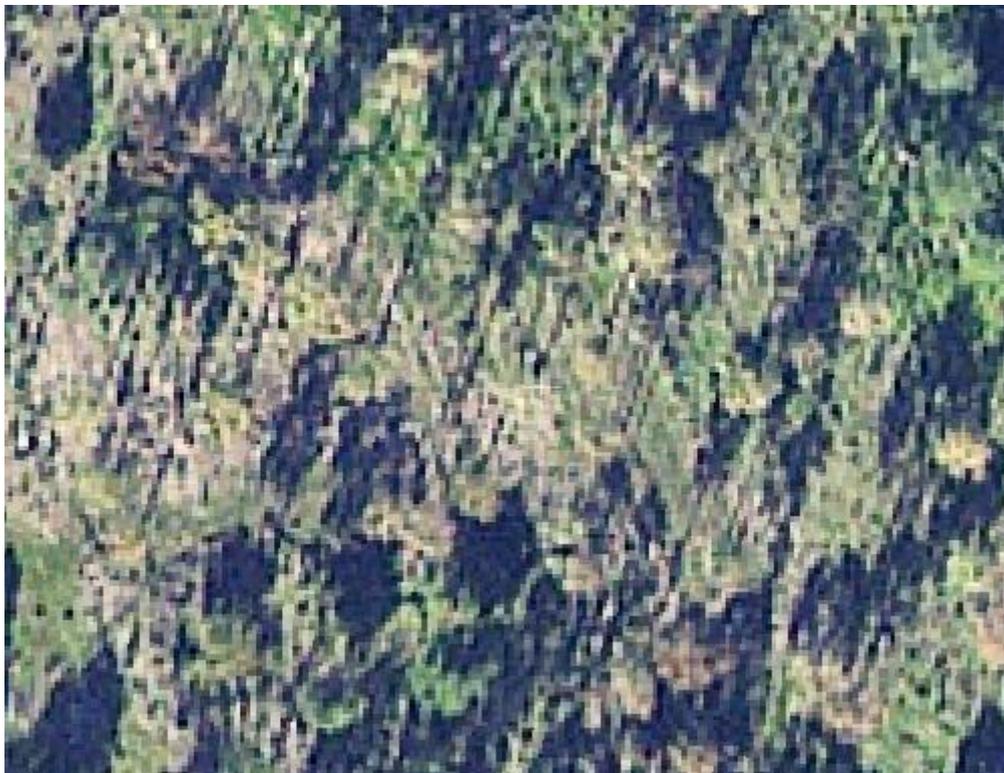
## 4 Interpretation of data

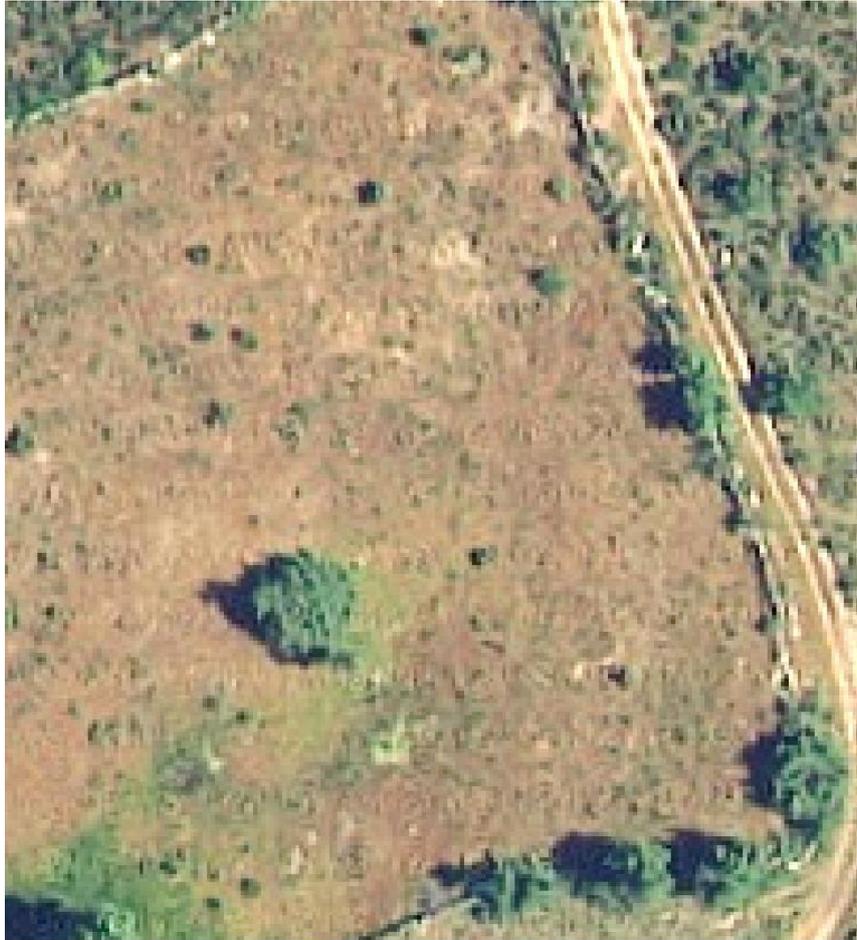
Before images were analysed, they were quality checked, paying particular attention to the following potential artefacts, which were remedied where possible:

- Blurriness in the images, stemming from movement and vibrations of the sensor.



- Cloud-cover or atmospheric haze in the images. These were not present, since the technical specifications in the procurement procedures requested that images only be acquired under clear-sky conditions.
- Image-processing artefacts. Errors at the sensor or elsewhere in the image-processing chain can cause anomalies in the images that hamper their interpretation or analysis.

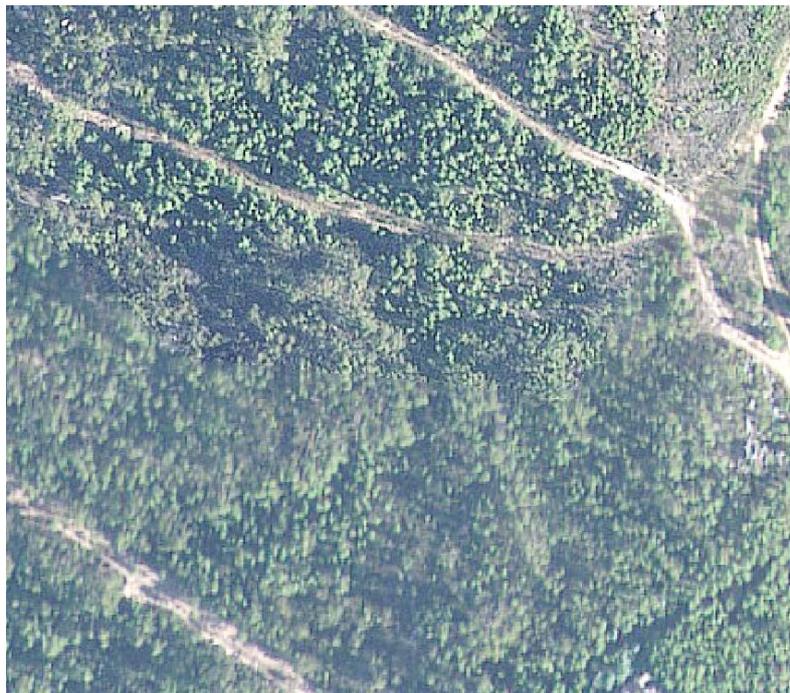




- Geo-registration quality. Geo-registration of the images was checked using the location of known features, and cartographic reference layers.
- Seamlines in the image mosaics. As the orthophoto mosaics delivered are composed of many individual images (whether individual frames or flight lines), they can display edge effects (i.e. seamlines) at the borders between frames or flight lines. These borders can reveal either faults in the geo-registration (bordering images don't align well), or - more commonly - spectral differences, owing to neighbouring images being acquired under different conditions. The differences in condition can be atmospheric (different illumination stemming from different time of day or flight direction), or surface-related (different season).
  - Seamline artefacts in orthophoto dataset.
  - Seamline between an image acquired on 15 October 2015 (top) and one acquired on 29 May 2015 (bottom).



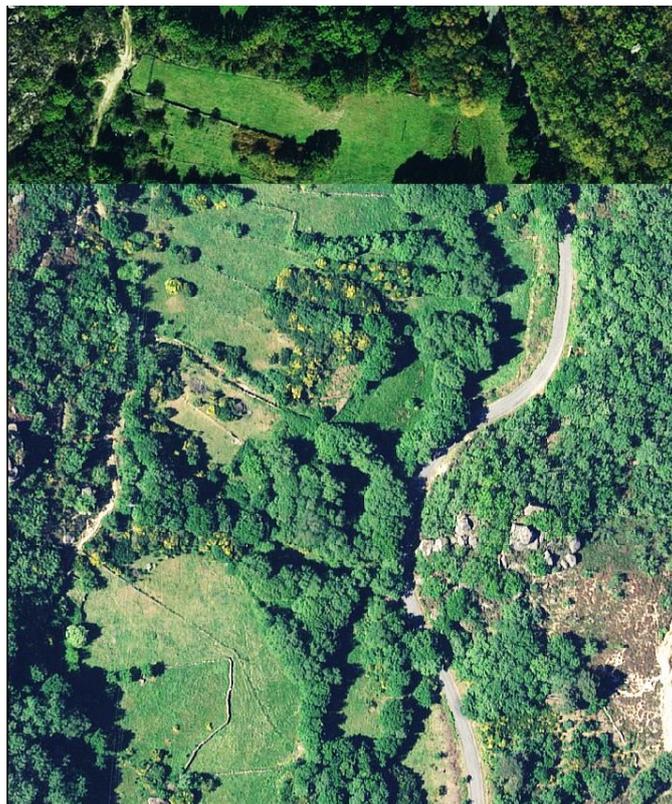
- Seamline below which image is blurred.



- Greatly different illumination on opposing sides of N-S seamline.



- Different illumination on opposing sides of W-E seamline.



- The original images, from which the large-area orthophoto mosaics were composed, were acquired in a wide range of seasons and conditions. We ensured this could be accounted for in processing the image mosaic, by requiring that the data be delivered with pixel-level metadata, describing, for individual pixels in the mosaic, exactly when they were imaged. As this is non-standard practice in the photogrammetry community, these metadata had to be carefully quality checked.
- Oblique viewing angles of the sensor can affect the interpretation of remote sensing data. This is the case for wide-angled cameras deployed on aircraft and for data from very high resolution satellite sensors that are able to acquire off-nadir images.

Particularly for forests, this can have dramatic effects on the way trees are seen in the images.

- The surface reflectance data were assessed for spectral accuracy, by checking that the magnitude and pattern of reflectance spectra for typical land surface features (water, soil, asphalt, healthy vegetation) were in line with expectations.

## 5 Image interpretation algorithms

### 5.1 Algorithm aim and challenges

We devised an image interpretation algorithm to highlight features in orthophotos that might represent coniferous trees with crowns showing signs of poor health. Several aspects made this challenging:

- There was no contemporary, authoritative database available to us, showing the precise location and time of observation of individual declining trees, with which to train such an algorithm.
- While relatively high in spatial detail, orthophotos are relatively poor in spectral detail; measuring light in only four, relatively broad, spectral bands (red, green, blue, near-infrared). The spectral signature of declining tree crowns in these bands can be very similar to that of, among others, bare soil, having a brownish colour in the visible wavelengths and low radiance in the near-infrared portion of the spectrum.

To tackle these challenges:

- We used algorithms that were not over-reliant on training data and could cope with uncertainty in these data.
- We made our algorithms iterative, so that their output could be evaluated and the results of the evaluation fed back to the algorithm for further refinement.
- We sought auxiliary geospatial data that would help distinguish areas with coniferous vegetation, where declining trees may occur, from other areas that might be spectrally similar. These data included the earlier COS land cover map, and older orthophotos.
- We based our image interpretation not only on the spectral characteristics of individual pixels in the images, but also on morphometric and textural aspects of the orthophotos.

### 5.2 Image pre-processing

Areas which, according to the COS land cover map, did not include any coniferous vegetation were masked from the images of the buffer zone and not analysed. To identify, and exclude from analysis, areas that had been cleared since the creation of the COS, we masked areas where either the 2012 orthophotos, or more recent orthophotos, had a particularly low normalised difference vegetation index (at 1 m resolution), which indicates the absence of any green vegetation.

The following 'texture layers' were created from the red and near-infrared bands of the orthophotos, using GRASS [3] (and more specifically its 'r.texture' function), and considered as potential predictor variables in the classification algorithm, to give more information about the spatial distribution of tonal variation within a band [4]:

- sum average (SA)
- entropy (ENT): this measure quantifies randomness. It is high when values within a moving window have similar values, and lower when they are more uniform.
- difference entropy (DE)
- sum entropy (SE)
- variance (VAR): a measure of grey tone variance within the moving window (second-order moment about the mean)
- difference variance (DV)
- sum variance (SV)
- angular second moment (ASM, also called Uniformity): this is a measure of local homogeneity and the opposite of Entropy. High values of ASM occur when the pixels in

the moving window are very similar. The square root of the ASM is sometimes used as a texture measure and is called Energy.

- inverse difference moment (IDM, also called Homogeneity): this measure relates inversely to the contrast measure. It is a direct measure of the local homogeneity of a digital image. Low values are associated with low homogeneity and vice versa.
- contrast (CON): this measure analyses the image contrast (local grey-level variations) as the linear dependency of grey levels of neighbouring pixels (similarity)
- correlation (COR): this measure analyses the linear dependency of grey levels of neighbouring pixels
- information measures of correlation (MOC)
- maximal correlation coefficient (MCC).

### 5.3 Algorithm choice

Here, we used the MaxEnt model, which was originally developed for modelling species niches and distributions by applying a machine learning technique based on maximum entropy [5]. A particularity of MaxEnt is that it is designed to be trained on presence observations only [6]. It does so by searching for the distribution with the maximum entropy for an input set of occurrences and a set of constraining variables [7], (which in our case are derived from the remote sensing data).

Entropy is a measure of unpredictability of a state or, equivalently, of its average information content [8]. Therefore, a distribution with higher entropy is less constrained and the maximum entropy principle seeks the maximum entropy probability distribution, subject to a constraint that each occurrence has the same mean under the observed distribution. Thus, the model output can be interpreted as predicted probability of presence, or as predicted local abundance [9]. Here, we selected the probability output of the model since we are considering the probability that a single pixel in an image belongs to a declining tree, rather than considering whether an individual may be present in the pixel (as is the case in traditional MaxEnt applications where pixels are larger than individuals).

The model was implemented in the programming language R [10], and its *dismo* package [11], initiated with a small set of declining crowns that were visually detected in the orthophotos. The algorithm then randomly selected locations from the background as 'pseudo-absences' [12] (i.e. points that are unlikely to be declining trees). The model was then calibrated, using as input predictor variables the RGB-NIR and texture bands of the orthophotos.

### 5.4 Algorithm input

The orthophotos are composed of four separate layers of information, corresponding to the blue, green, red, and near-infrared spectral bands. These individual layers, together with texture layers calculated from the red and near-infrared bands, were used as inputs (i.e. 'predictor variables') in the classification algorithm. The texture layers capture land surface properties that may relate to surface roughness and contrast - features which can help distinguish crowns from other features in the landscape.

### 5.5 Algorithm settings

The number of 'pseudo-absences' used by the algorithm can be set by the user. A number that is too high will slow down the algorithm; a number that is too low will compromise detection accuracies, as it prevents the algorithm from describing all the landscape features that are not declining trees. Optimising the number of background points thus involves determining the minimum number of points required to obtain near-optimal algorithm performance. To determine this, we ran a range of tests with different sample sizes of background points, relying on the area under the curve (AUC) and Cohen's kappa statistics

to determine algorithm performance. The AUC is used to estimate the predictive accuracy of a distributional model derived from presence-absence observations. Since our model is based on presence-only data, this metric can help to see the predictive power of the model, but it cannot be taken as an objective measure of the performance of the algorithm. The AUC namely weights omission and commission errors equally, while in presence-only models the omission errors are weighted more strongly than the commission errors. The kappa statistic measures inter-rater agreement for categorical variables. It is generally considered a more robust measure than simple per cent agreement calculation (also termed 'overall accuracy'), as kappa accounts for the possibility of the agreement occurring by chance. A kappa value close to 1 thus indicates a good agreement between the metrics (in our case observations and model predictions) and one close to 0 indicates the agreement is no better than expected from random chance. These tests suggested that 10 000 background points per 1 km<sup>2</sup> image tile were appropriate for the algorithm training, ensuring kappa and AUC values above 0.98.

Once the minimal number of background points was set, we determined whether the set of input variables fed into the model could be reduced; many of the original ones might be correlated. While MaxEnt is relatively robust to correlation among input variables, eliminating superfluous ones would speed up the processing. Principal component analysis revealed that nearly 100 % of the variation in the texture layers could be described by three components. Hence, we retained as predictor variables only the texture layers that contribute most to the eigenvectors of those three principal components.

After this process the predictors selected for the model were:

- RGB-NIR bands
- contrast texture of red and NIR band
- sum average texture of red and NIR band
- sum variance texture of red and NIR band
- variation texture of band red and NIR band
- information measures of correlation of first order texture of red band
- correlation texture of NIR band.

The graph shows multiple eigenvectors overlapping with respect to variability within the distribution, indicating they are correlated. So as not to carry this correlation into the variables fed into the model, and to reduce the dimensionality of the input variables, we selected layers that were less correlated and reduce the information content only minimally.



## 6 Image analysis hardware

We sought an image analysis platform that could efficiently handle and parallelise the computations on the large (terabyte) volumes of image data in this project. The JRC Earth Observation Data and Processing Platform (JEODPP), which was developed in parallel with the CanHeMon project, increasingly met these needs over the course of the project. Being an inhouse service of the EC, it facilitates processing of the data for which the licensing does not permit public distribution. It is a versatile platform that brings the users to the data through web access and allows for large-scale batch processing of scientific workflows, remote desktop access for fast prototyping in legacy environments, and interactive data visualisation/analysis with JupiterLab.

At the basis of the JEODPP are storage and processing nodes in an infrastructure based on commodity hardware and the CERN EOS distributed file system, which provides a disk-based, low-latency storage service. With its highly scalable hierarchical namespace, and data access possible by the XROOT protocol, it was initially used for physics data storage. The main target area for the service is scientific data (in CERN's case, physics data) analysis, which is characterised by many concurrent users, a significant fraction of random data access, and a large file-open rate. For user authentication, EOS supports Kerberos (for local access) and X.509 certificates for grid access. To ease experiment workflow integration, SRM as well as GridFTP access are provided. EOS further supports the XROOT third-party copy mechanism from/to other XROOT enabled storage services. As of summer 2018, the storage capacity of the EOS distributed file system of the JEODPP is 14 PiB, which is set up to store 7 PiB of data (with the other 7 PiB set up to ensure redundancy).

The JEODPP's processing servers contain circa 12 000 cores, with 10 to 19 GB RAM per core. It has 10 servers for data services, such as web mapping and database hosting. It was also recently extended with two GPU servers, with two processing units each.

On this hardware, large-scale batch processing is executed through lightweight Docker virtualisation and management with HTCondor. Docker is a computer program that performs operating-system-level virtualisation, also known as containerisation. Containers are isolated from each other and use their own set of tools and libraries; they can communicate through well-defined channels. All containers use the same kernel and are therefore more lightweight than virtual machines. Containers are created from 'images' which specify their precise contents. Images are often created by combining and modifying standard images downloaded from repositories. In our case, we use a Debian image with all the libraries needed to run the code - mainly R and GDAL libraries.

## 7 Processing load and time

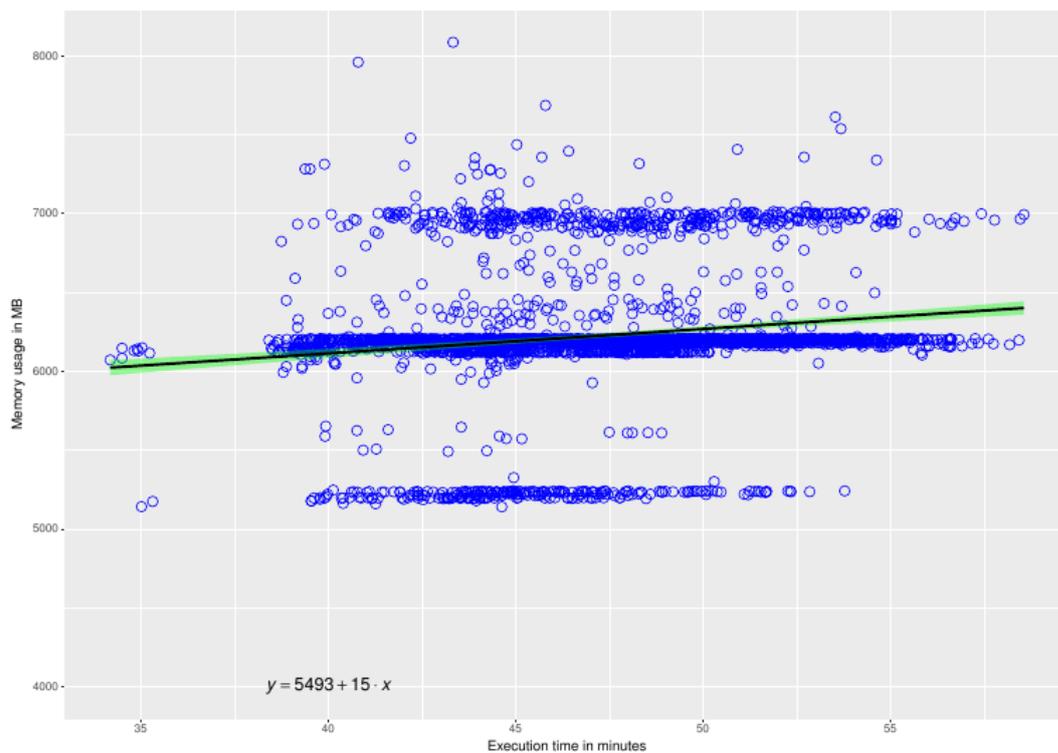
Covering the entire PWN buffer zone with 4-band images of 30 cm resolution, stored in 8-bit, generates 2.4 TB of data. Generating the associated texture layers used in the analyses here adds an additional 50 TB. The data were delivered and processed in 24 904 tiles measuring 1 km by 1 km.

Processing a single tile in each iteration takes 40 to 55 minutes, with a memory usage of 5-7 GB on a regular CPU (with 2 to 8 cores). Processing the entire buffer zone on a single CPU would thus take more than a year.

Assigning all of the 12 000 cores of the JEODPP to the task would see it completed in less than two hours. In practice, between 100 and 500 cores of the JDEOPP were used in the processing at any one time.

The figure below shows, for a subset of the dataset, the actual processing times and memory usage by the JDEOPP.

**Figure 3.** Processing and memory usages for jobs executed on the JDEOPP



## 8 Accuracy assessment of image interpretation

### 8.1 Sample-based validation and iterative model calibration

To facilitate the validation of image classification outputs, and more precisely the status and location of declining coniferous trees detected, we built a dedicated web page accessible on the JRC network. This page allowed an operator to visualise randomly selected declining trees, according to the image classification, or objects that were not classified as declining trees and judge their status visually, based on the available images. By assigning the images presented with the label 'declining tree', 'not declining tree', or 'possibly declining tree', using simple keyboard strokes, a set of validation records were generated (see screen capture of the site below).

**Figure 4.** Screen capture of the internal website to support the collection of validation data on visual image inspection



Using the records generated through the site, confusion matrices could be produced to estimate different performance statistics for the algorithm, including user accuracy, producer accuracy, and Cohen's kappa.

To facilitate visual inspection and validation of the declining tree detections stemming from the image classification, we generated a dedicated web application. The application loads the locations of the detections, commonly generated in shapefile format, into a database and generates a sample from them, as well as from points that were not classified as declining trees. It then displays the points in this sample on top of matching images, for an operator to visually classify them. The operator's classifications are stored in the database and used to generate accuracy statistics for the automated image classification.

Moreover, the collected records served as further training data to refine calibration of the model in an iterative way; i.e. after the initial model was run to classify the images and had detected declining trees, results from the validation were included as training data to improve the model. This iterative run-validate-recalibrate cycle could be repeated until the performance statistics of the classification saturate.

### 8.2 Image analysis results

After three iterations of the algorithm, we identified 19 656 trees in the Castelo Branco area that had considerably declined in health between November 2015 and November 2016.

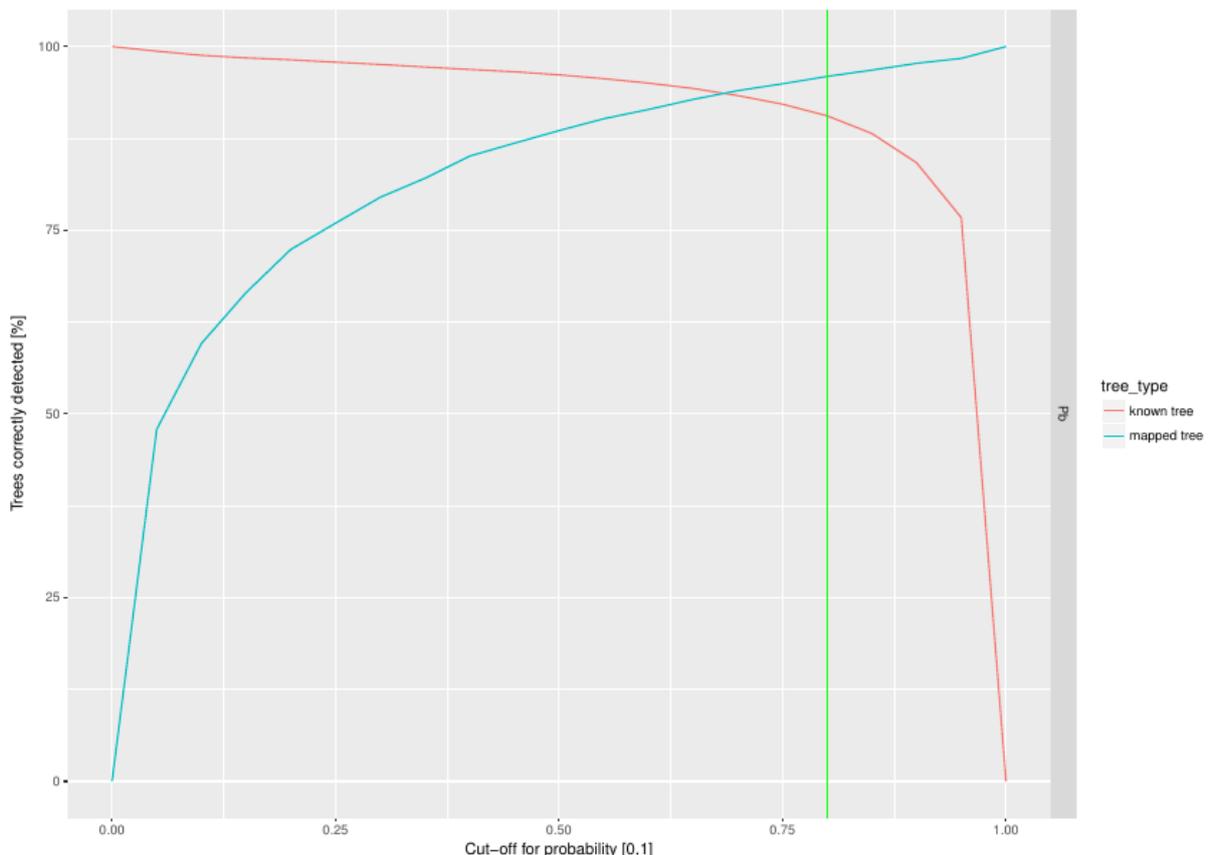
The algorithm was initiated with a limited set of declining trees visually detected in the images during quality checking, and then run in a first iteration that randomly selected points in the images that represented objects other than declining trees. The MaxEnt algorithm was then calibrated and left to run over the entire area. A sample of the objects classified after this first iteration as declining trees was then visually inspected. This yielded many newly detected declining trees, as well as objects that were wrongly detected as declining trees. Both sets of objects were fed back to the algorithm for recalibration, with the first group allowing it to better identify declining trees, and the latter group allowing it to better distinguish them from objects with which they could be confused. This process was repeated over three iterations.

To display how the performance of the declining tree detection algorithm is evaluated in different stages of the individual iterations, we use a series of graphs.

Since MaxEnt assigns to each cell in an image the probability that it belongs to a declining crown, to binarise its output a suitable cut-off probability needs to be determined. To do this, we considered a range of probability cut-offs and determined, for each one, the resulting tree classification accuracy.

The following graph shows the tree detection ratio (as a function of this cut-off probability), based on the fraction of known declining trees that are detected, and the fraction of detected trees that turn out to be correct.

**Figure 5.** Tree detection ratio as a function of the cut-off used to classify the probabilities generated by the MaxEnt model into presence/absence



As the cut-off value increases, a greater proportion of the mapped declining trees are indeed known declining trees. However, at the same time, fewer declining trees are actually mapped and more of the known declining are therefore missed.

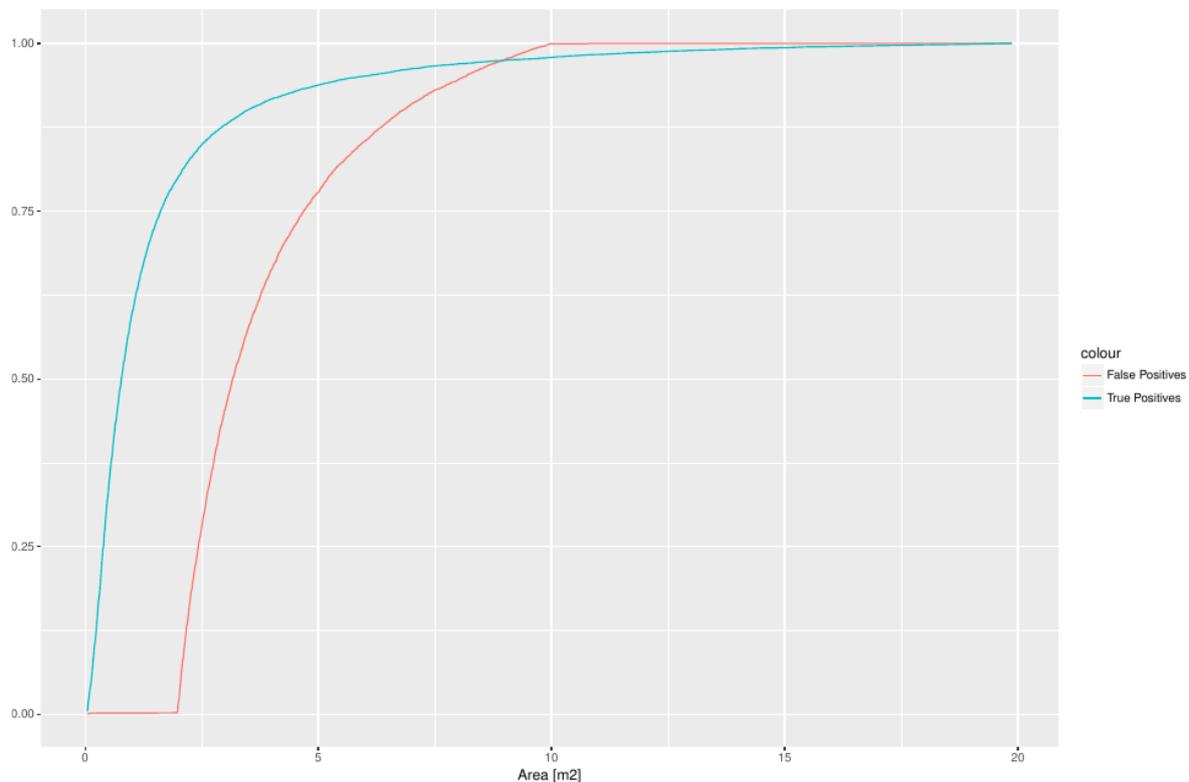
The green line indicates the cut-off probability below which less than 10 % of the known declining trees are missed by the algorithm.

The determined cut-off probability is then used to binarise all the pixels in the MaxEnt output, assigning those below the cut-off the value 0, and those above the value 1.

The binarised maps allow the delineation of objects (clusters of pixels) considered to be declining tree crowns. We checked how the size of these objects related to their likelihood of indeed representing tree crowns, as both extremely large as well as very small objects are unlikely to actually be tree crowns.

Therefore, only considering very small objects is likely to lead to declining tree detections. Also considering objects with a crown-like size is likely to greatly increase the number of declining trees detected, while including very large objects is unlikely to increase this number further. The following graph shows this for the actual results.

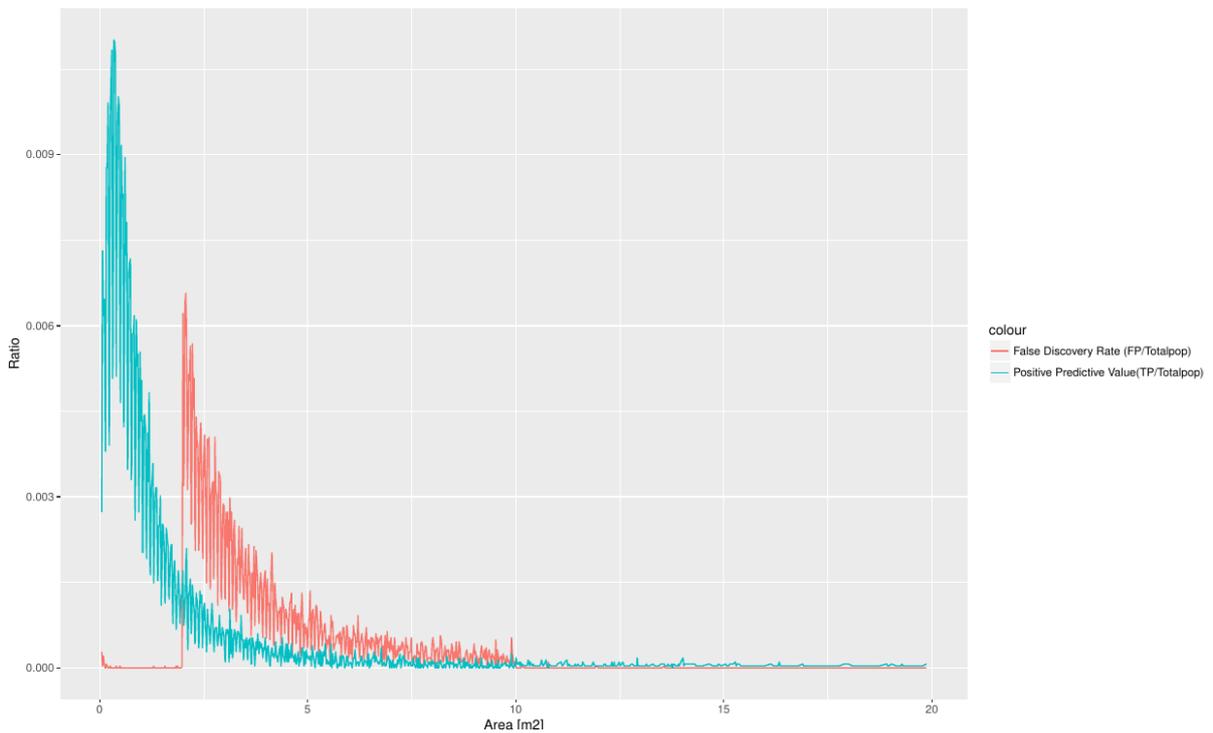
**Figure 6.** Cumulative ratio of trees retrieved as a function of object size cut-off, whereby all objects under a certain size cut-off (displayed on the x-axis) are retained.



This illustrates how, by selecting objects with an area between 0.1 m<sup>2</sup> and 3.5 m<sup>2</sup>, we detect almost 80 % of declining trees without any errors of commission. In other words, we achieve an accuracy of 100 % with a 20 % detection error.

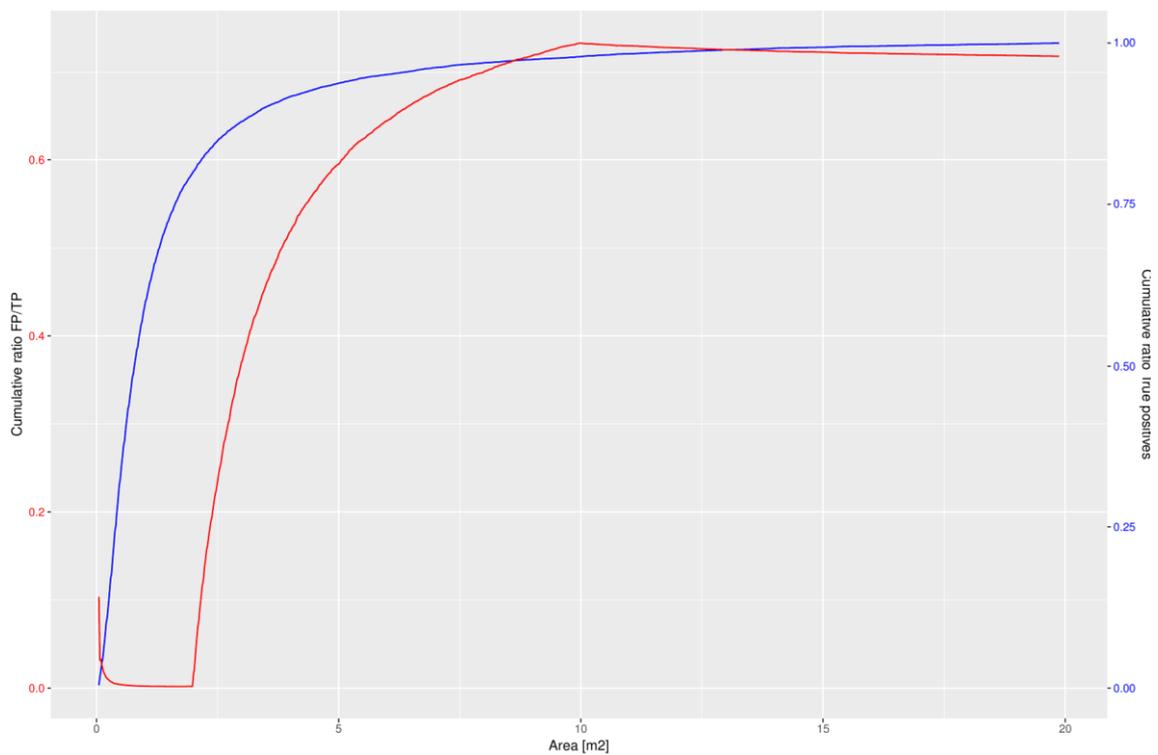
Indeed, the objects that correspond to known declining trees cluster in an area range between 0.1 m<sup>2</sup> and 4 m<sup>2</sup>. The majority of false positives, however, have an area between 2.5 m<sup>2</sup> and 6 m<sup>2</sup>, as evidenced by the graph below.

**Figure 7.** False discovery rate and positive predicted rate, as a function of object size



Since the algorithm aimed to detect nearly all declining trees, and detected trees need to be inspected, as a final step a visual check was performed on the detected trees. Here, a choice must be made, balancing the likelihood that an actually declining tree is detected, with the number of detections that need to be rejected during visual inspection because they are wrong.

**Figure 8.** Ratio of false positive to true positive detections and cumulative ratio of true positives, as a function of object size cut-off



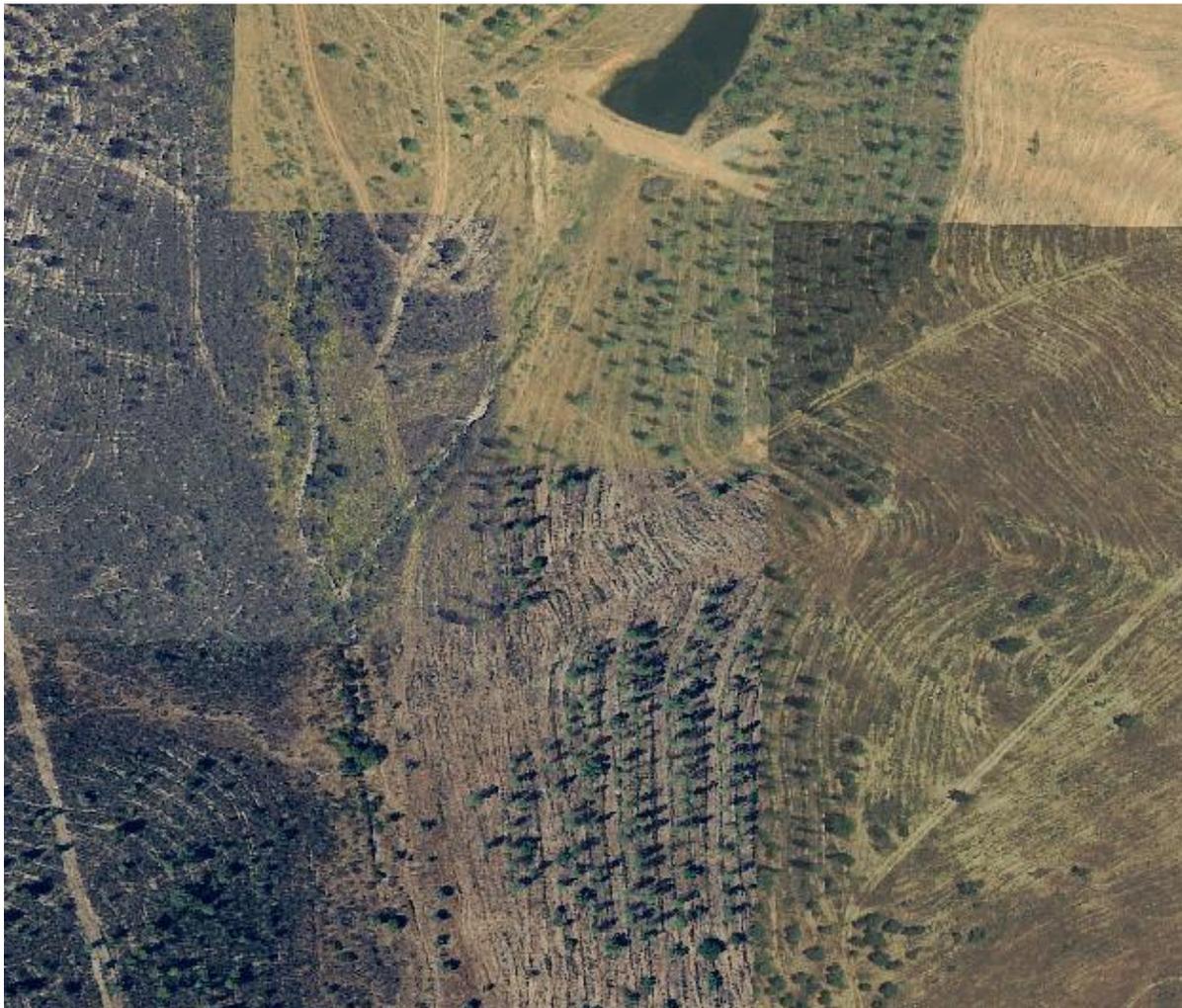
If it is paramount that all declining trees are detected, the algorithm is bound to have a larger number of 'false positives' that need to be filtered out in a final, manual step. If the final visual inspection needs to be kept to a minimum, then all detected objects can be guaranteed to be declining trees, but many actual declining trees will be missed. This choice is ultimately based on risk assessment and resources, and its implications are illustrated below using our actual results.

If it is paramount that absolutely all declining trees are detected, then for every 10 correctly identified trees, circa 7 need to be rejected during a final visual inspection of the output.

This effort can be reduced to circa 5 rejections per 10 correct identifications if it is considered sufficient that > 95 % of declining trees are detected.

The images over the Castelo Branco area were purpose-collected for this project. The experiment to scale up the analysis to the entire buffer zone was added later and necessarily relied on existing data which were not purpose-collected for this project. These data were mostly collected in summer, which is not the best period to assess decline in the context of the pine wood nematode outbreak. Moreover, they were collected over a long time period, with images of adjacent areas collected months or years apart. This difference in acquisition date and image calibration meant that the buffer zone had to be analysed in 'chunks' of data acquired in the same time window.

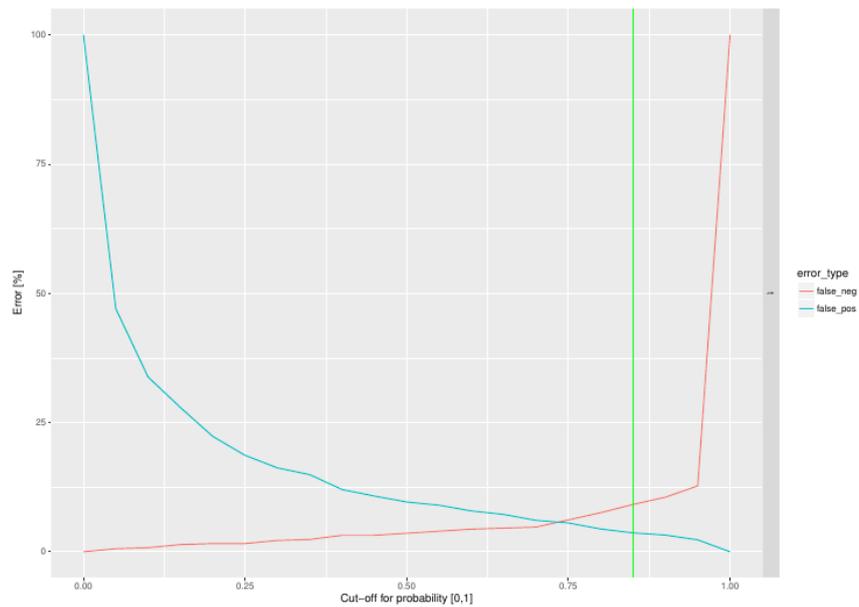
**Figure 9.** Example of seamlines in an image mosaic generated from images acquired on different dates



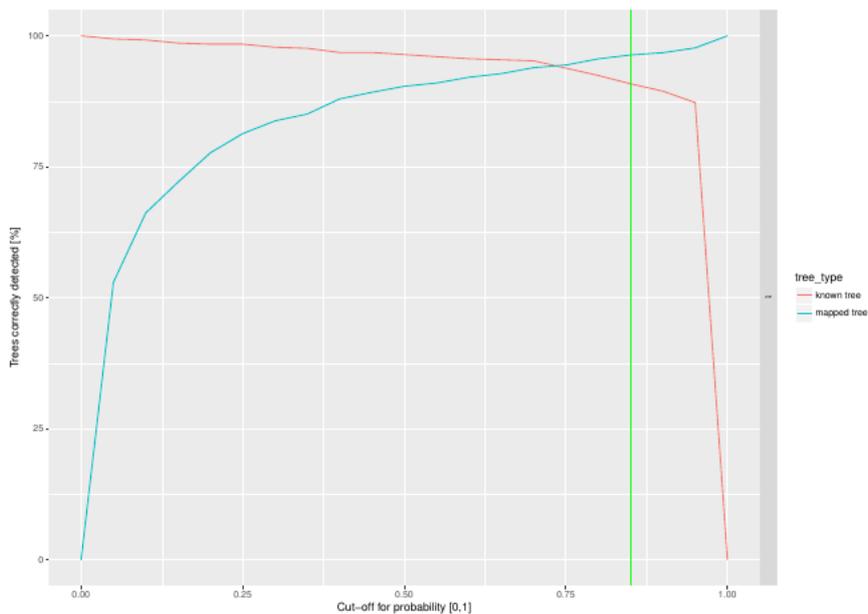
This implies that the algorithm had to be calibrated individually for different image 'chunks'. This was not possible in sections of the buffer zone in which few or no declining trees could be visually identified in the available images, as there was no material available to initiate the algorithm calibration iterations. The extremely low prevalence of visibly declining trees in the available buffer zone images severely hampered the ability to detect further such trees.

In concrete terms, the difference in acquisition dates and pre-processing of the data means that the crowns of declining trees show up in very different 'colours' depending on where they are. This makes it extremely challenging for a single parameterisation of the image analysis algorithm to successfully analyse the entire area.

**Figure 10.** Tree detection error ratio, as a fraction of the cut-off used to classify the probabilities generated by the MaxEnt model into presence/absence

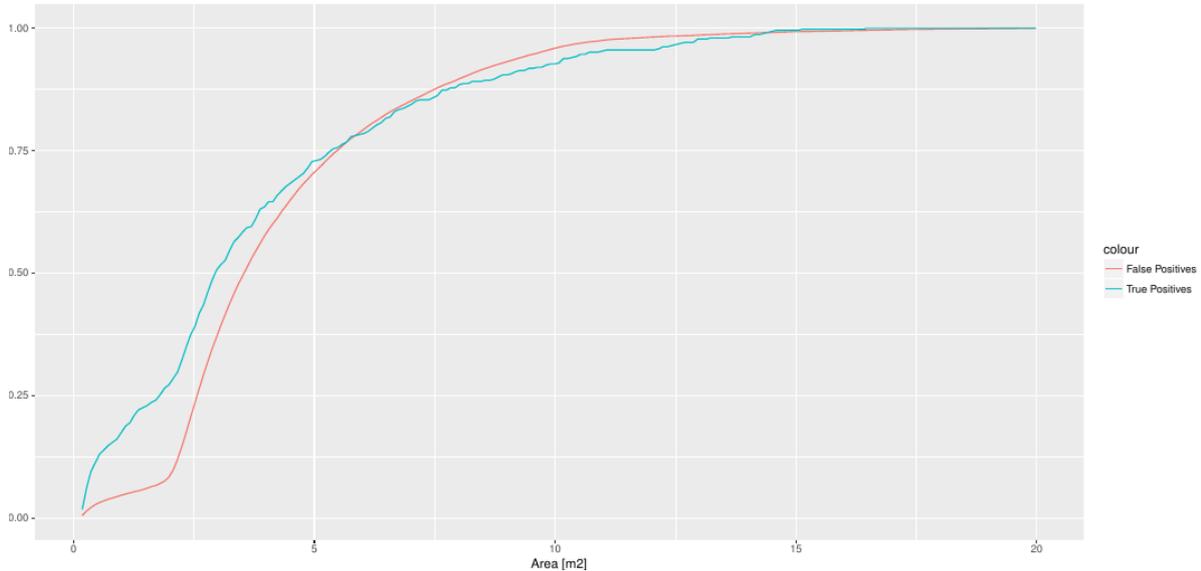


**Figure 11.** As Figure 5.



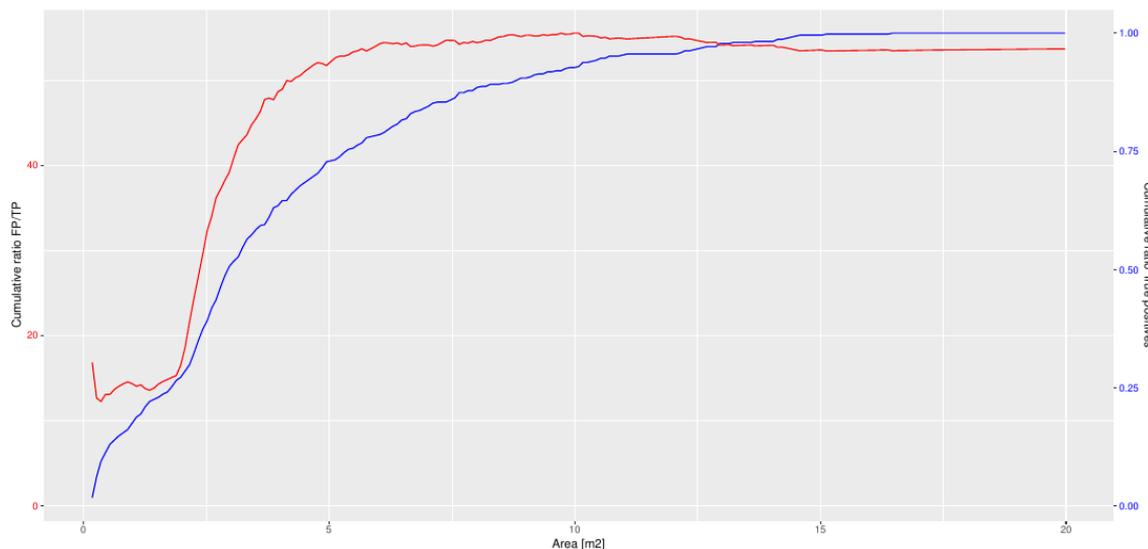
While acceptable error rates are achieved, objects misclassified as declining trees are less easily identified based on their size, than in the higher resolution purpose-collected images. While small objects are more likely to be false positive detections than larger objects, erasing them from the output also leads to the loss of correct detections; e.g. erasing detected objects smaller than 2.5 m<sup>2</sup> avoids 25 % of the false positives, but ensures only 65 % of the true positives are retained.

**Figure 12.** As Figure 5.



As a result, to ensure that more than 80 % of known declining trees were detected in the images, final visual inspections of the output would have to reject 50 to 60 wrongly detected crowns for each correct one.

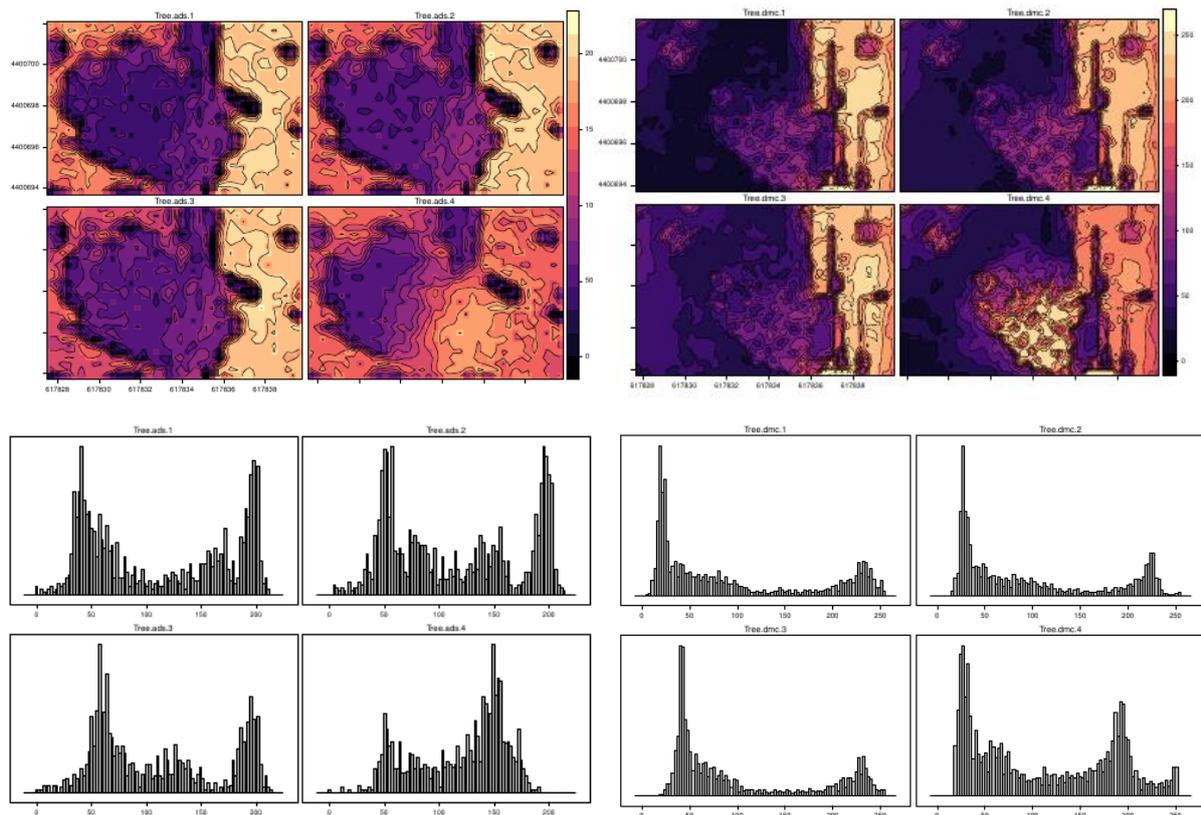
**Figure 13.** As Figure 8.



This discrepancy between the results that can be obtained with the purpose-collected and the off-the-shelf images has several causes. Firstly, the purpose-collected images were acquired in autumn, when more trees were in a state of decline, providing the algorithm with a larger target sample. Secondly, the purpose-collected images had a spatial resolution of 15 cm, as opposed to 30 cm for the off-the-shelf images available. Thirdly, the off-the-shelf images had poorer spectral definition than the purpose-collected images.

The figures below show the same healthy tree crown, as seen in the four bands (red, green, blue, NIR) of the purpose-collected DMC images (on the left), and in the off-the-shelf ADS100 images (on the right). The tree is bordered by vegetation on the left, and by a road on the right. The histograms on the bottom show the distribution of the values in each of the channels. It illustrates how the DMC images have a better contrast between crown and background, and make better use of the full data range in their description of the crown. The ADS100 data instead cluster in narrow data ranges.

**Figure 14.** Blue, Green, Red, and Near-Infrared views (clockwise from top left) of a single tree crown, as described by the DMC images (two left columns) and ADS100 images (two right columns).

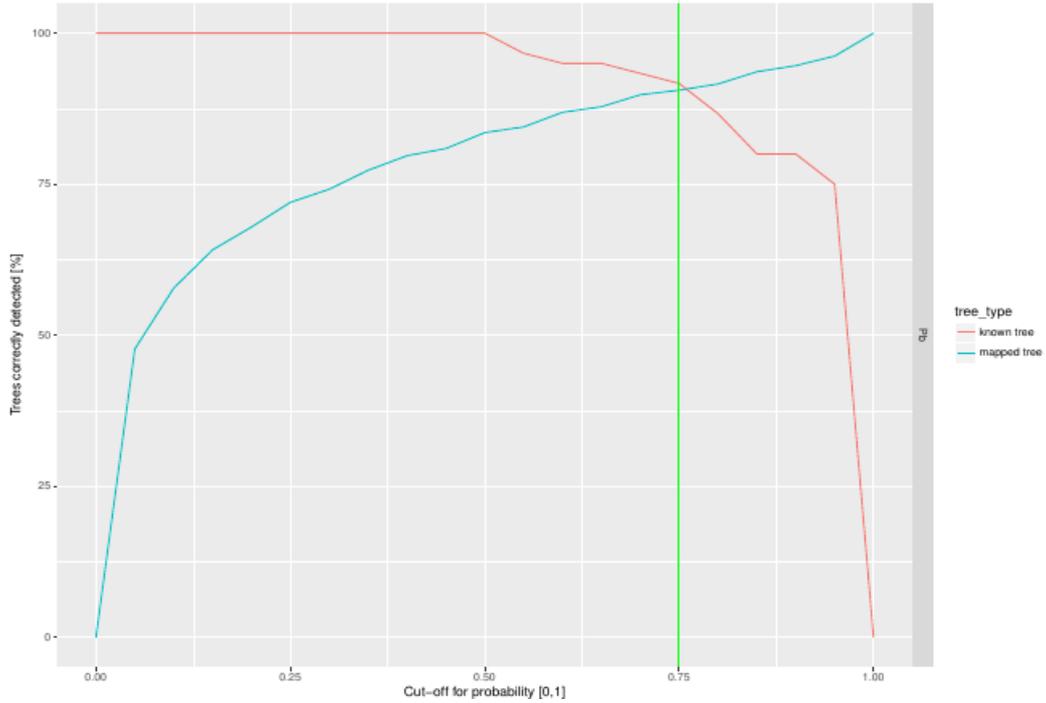


Finally, the ADS100 data contained many more artefacts from acquisition and pre-processing than the DMC data, which are practically free from artefacts.

During processing of the images of the entire buffer zone, the algorithm failed to calibrate when areas were large and extremely few training points were available (AU\_06, AU\_07, AU\_08, AV\_07, AY\_07). In other areas, where the model did calibrate, it did not detect any declining trees in the ADS100 images (AV\_08 and AW\_08).

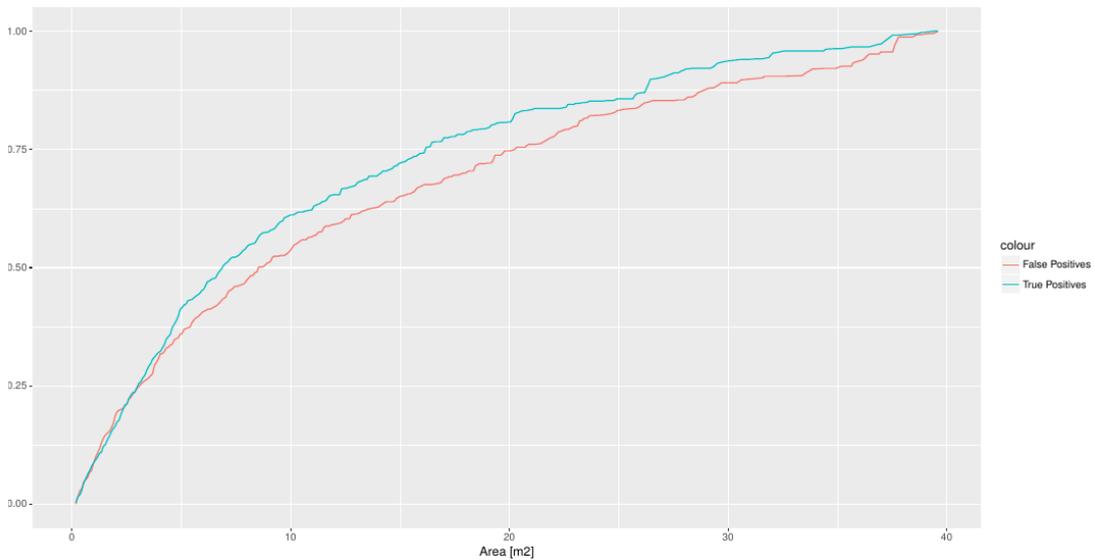
In yet other areas, e.g. AT\_06, which measures 488 km<sup>2</sup> in the north west of the buffer zone, detections were made but not as efficiently as with the purpose-collected imagery; selecting a probability cut-off of 0.75 ensures 91 % of the trees were correctly detected.

**Figure 15.** As Figure 5.



The selected cut-off probability was 0.75, with 91 % of the trees correctly detected. The circa 900 declining crowns correctly identified outnumber the objects wrongly identified as declining trees. This true-false ratio could not be further improved based on the sizes of the objects, as both showed a similar distribution along an object-size gradient. In total, 2 882 trees were flagged as declining in this area, of which 855 were confirmed upon visual inspection.

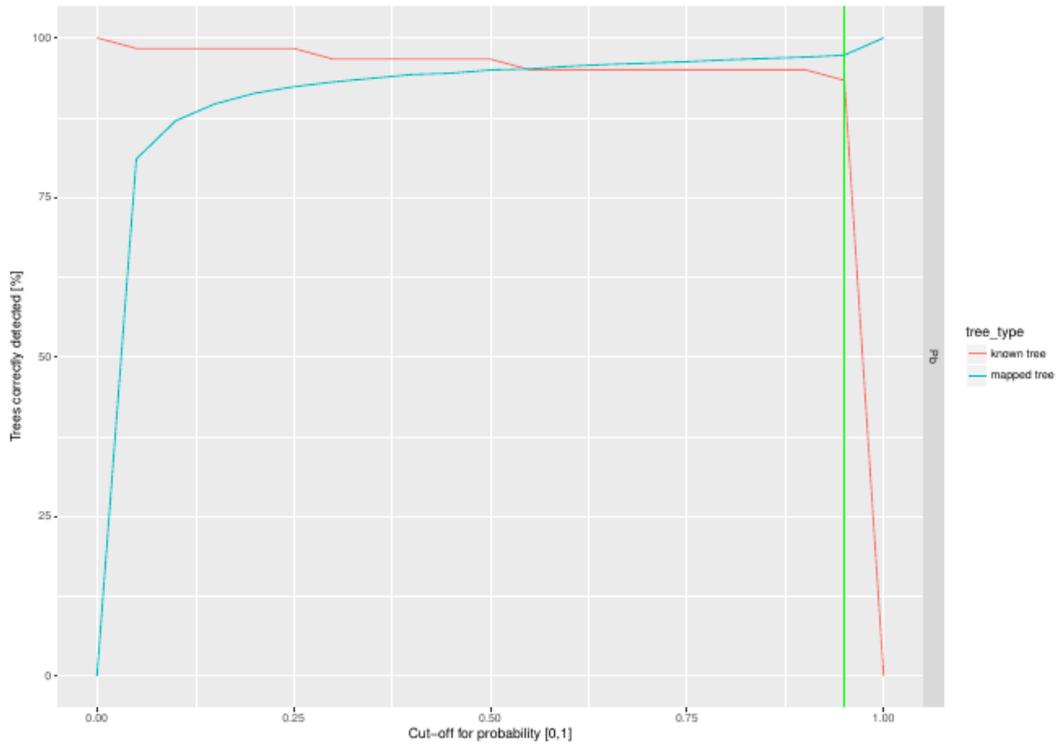
**Figure 16.** As Figure 6.



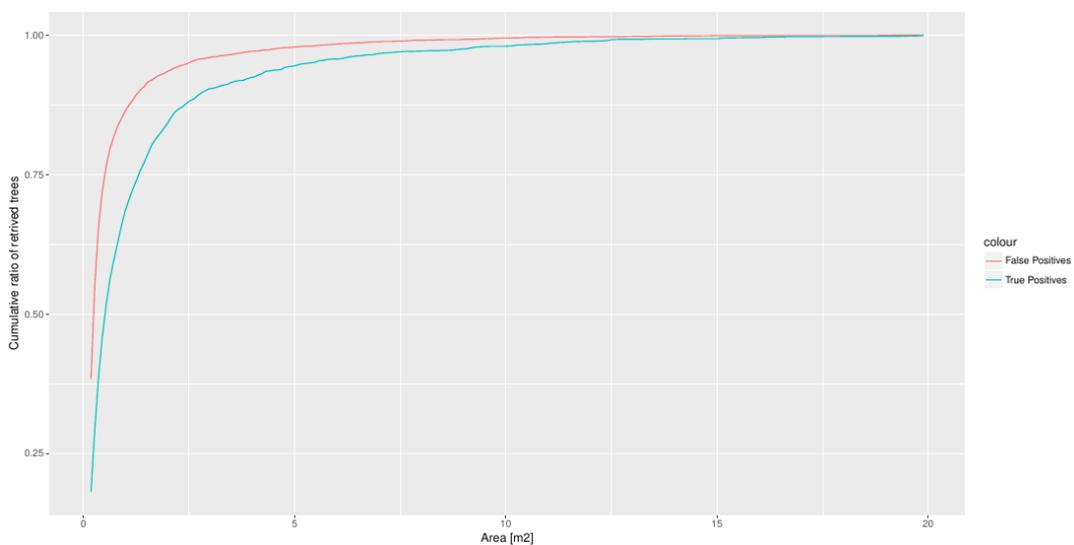
A final visual inspection of the output in this area would have to reject circa 8 detected for every 10 correct ones. If this ratio is to be brought down, smaller objects could be removed from the output; however this will be at the cost of losing about a quarter of correct detections.

For comparison, in the 4 269 km<sup>2</sup> AX\_07 area in the south of the buffer zone, we finally obtained lower rates of true positives to false positives, but had greater opportunity to trade off detection probability and final visual inspection effort. For this chunk of data, the selected cut-off probability was 0.95, which ensured 93 % of the declining trees were detected.

**Figure 17.** As Figure 5.



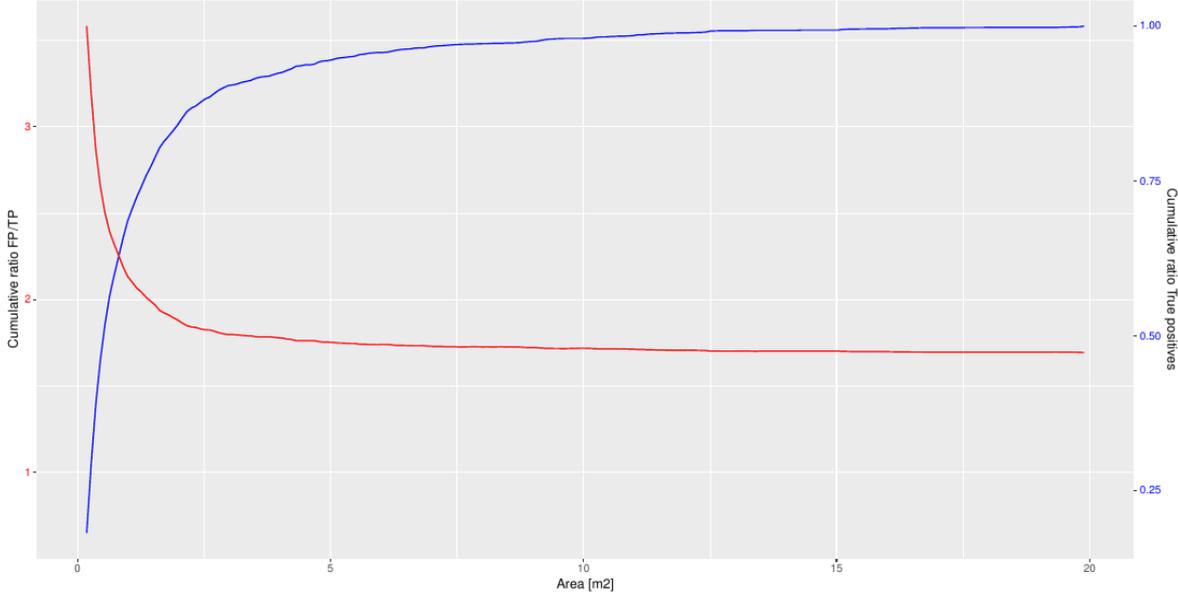
**Figure 18.** As Figure 6.



Among the objects detected, false positive detections are particularly prevalent among smaller sized objects. Consequently, rejecting these smaller objects (e.g.  $< 1 \text{ m}^2$ ) would eliminate circa 75 % of the false positive detections, but only 30 % of the true positive detections. Should the analyses be continued into a new iteration to further calibrate the model, this could be a good course of action.

Until then, for objects smaller than  $1 \text{ m}^2$ , circa 4 wrong detections for every correct detection need to be rejected during a final visual inspection, compared with a ratio of 17 to 10 for larger objects. Overall, 90 306 objects were flagged as declining trees in this area, of which we estimate 23 014 are correct identifications.

**Figure 19.** As Figure 8.



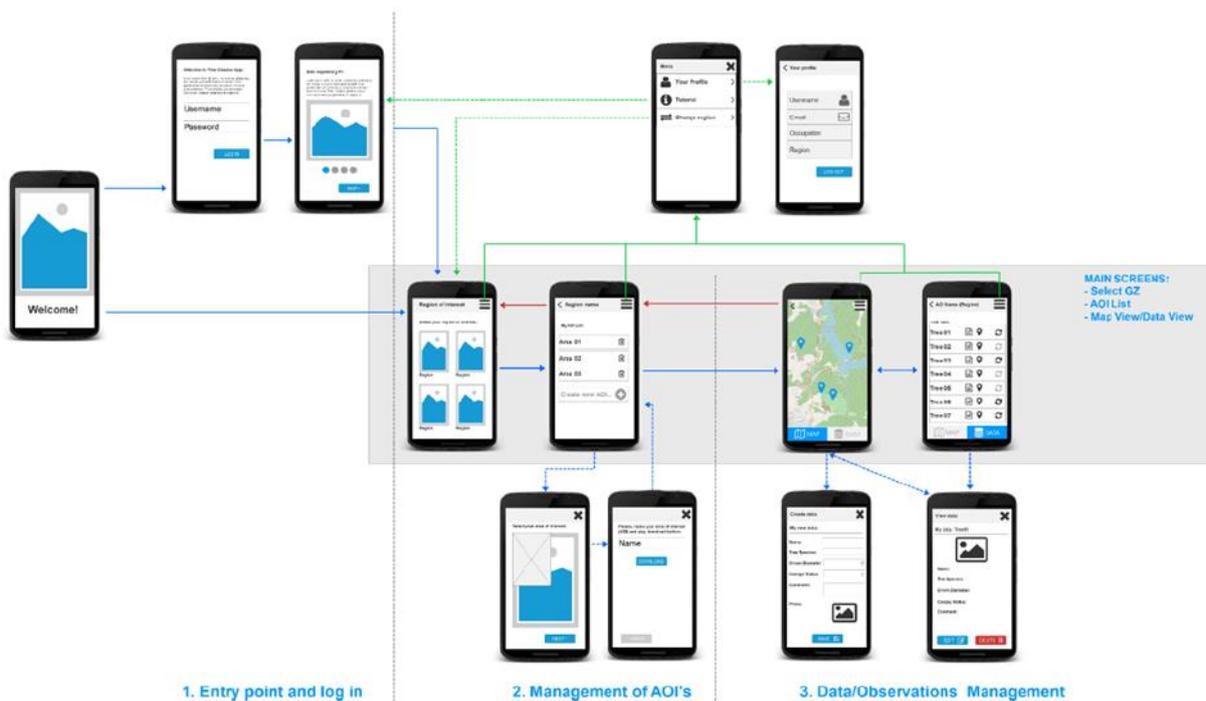
## 9 Mobile data exchange platform

We developed a mobile application (working title: TreeChecker) which facilitates the exchange of remote sensing data, and results stemming from their analysis, and in situ data collections. The application and its back-end are queued for a security assessment by the JRC Local Information Security Officer (LISO), without which it cannot be released.

The application allows the interactive visualisation of remote sensing data hosted on a server, and the download of these data to the device so that they can be shown in a dynamic map, even when the device is offline. This feature is critical, as mobile networks are often weak in remote areas. Data on individual trees can then be collected in the field (including tabular data, free text, form data and photographs). The user can directly link these data to what is seen in the images, or image-derived layers, meaning the geolocation of the observation is not limited to the accuracy of the device's positioning. Once the device comes back online on a Wi-Fi network, the collected data are then automatically synchronised with the server.

The platform can host data for different regions, which may correspond to different quarantine pest outbreaks, and set which users have access to data from which regions. This ensures that sensitive data (whether licensed remote sensing data or sensitive field observations) can be protected.

**Figure 20.** Overview of navigation in the TreeChecker mobile application



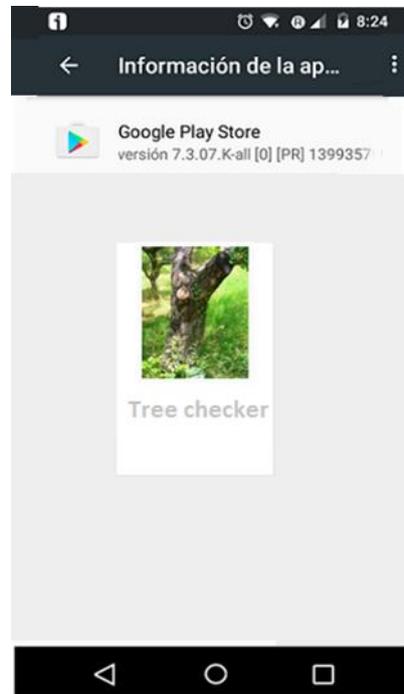
The following describes use cases for the Treechecker application, and its main functionality. It also describes the workflow of a user when interacting with the application.

The descriptions of the use cases consist of:

- A mockup of a mobile device for the different interactive screens envisaged in the final application. Note that the application should function on multiple screen sizes.
- A context that explains through which screen and/or action the user has reached this particular point in the workflow. The context can be considered as the pre-condition.
- A user-system interaction to describe in detail how the user can communicate with the application.

- A consequence to describe the final outcome(s) of the user-system interactions. The consequence can be considered as post-conditions.
- Observations which provide additional information.

## 9.1 Installation



### Context:

- The user downloads and installs the application after searching by its name: Tree Checker.

### User-system interaction:

- The user locates the application.
- The user gets a short description of the application and additional metadata such as developer, license, objectives and eventually user comments and user ratings.
- The user downloads and installs the mobile application on their terminal.

### Consequences:

- The terminal shows an icon as a bookmark link that provides access to the mobile application.

### Observations:

- The mobile application is compatible with Android versions  $\geq 5.0$ .

## 9.2 Authentication – Login Page



**Context:**

- The user has installed the application and has pressed the application’s icon.
- AND
- The user explicitly accepts the terms and conditions for the mobile application (exact wording to be provided by the JRC).
- AND
- The user has selected the language for the application. English shall be provided as the only option.

**User-system interaction:**

- The application presents the user with a login page.
- The user fills in the corresponding text boxes with personal info: email address and password.
- The user presses the Log in button.

**Consequences:**

- The user is authenticated against the application.
- The user is ready to browse through different mobile pages.
- The user is permanently logged in to the application, meaning that login session does not expire.

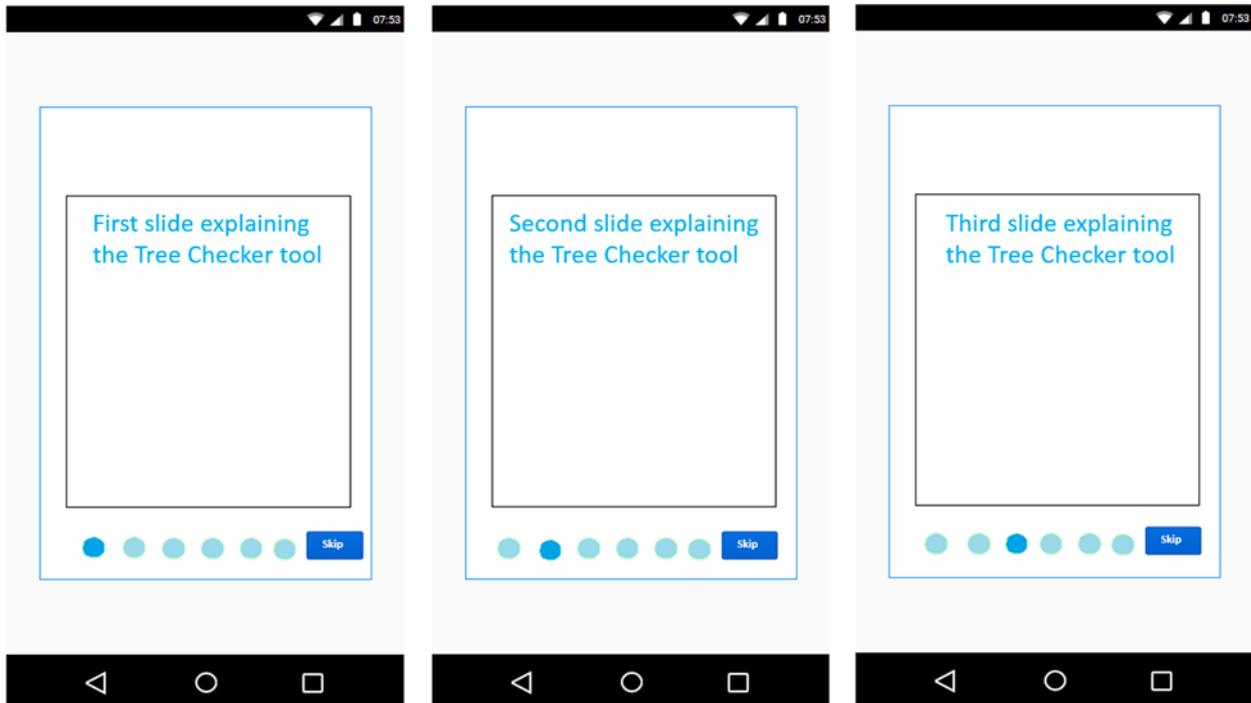
**Observations:**

- The user has internet connection.
- The mobile application does not offer options in multiple languages at present.
- The mobile application does not allow registration of new users.
- Only known pre-approved users can access the application.

- Different permissions have previously been established for different users.
- If the user deletes local data in their mobile storage settings, they must re-enter their credentials.

### 9.3 Visualisation of different screens

Explanation pages:



Context:

- The user has entered their credentials on the login page and has pressed the Log in button.

User-system interaction:

- The application presents the user with a sliding page (see Sequence of sliding pages).
- The user can browse from one slide to another by swiping their finger horizontally on the screen.
- In future, the user can skip this initial explanation by clicking on the Skip button.

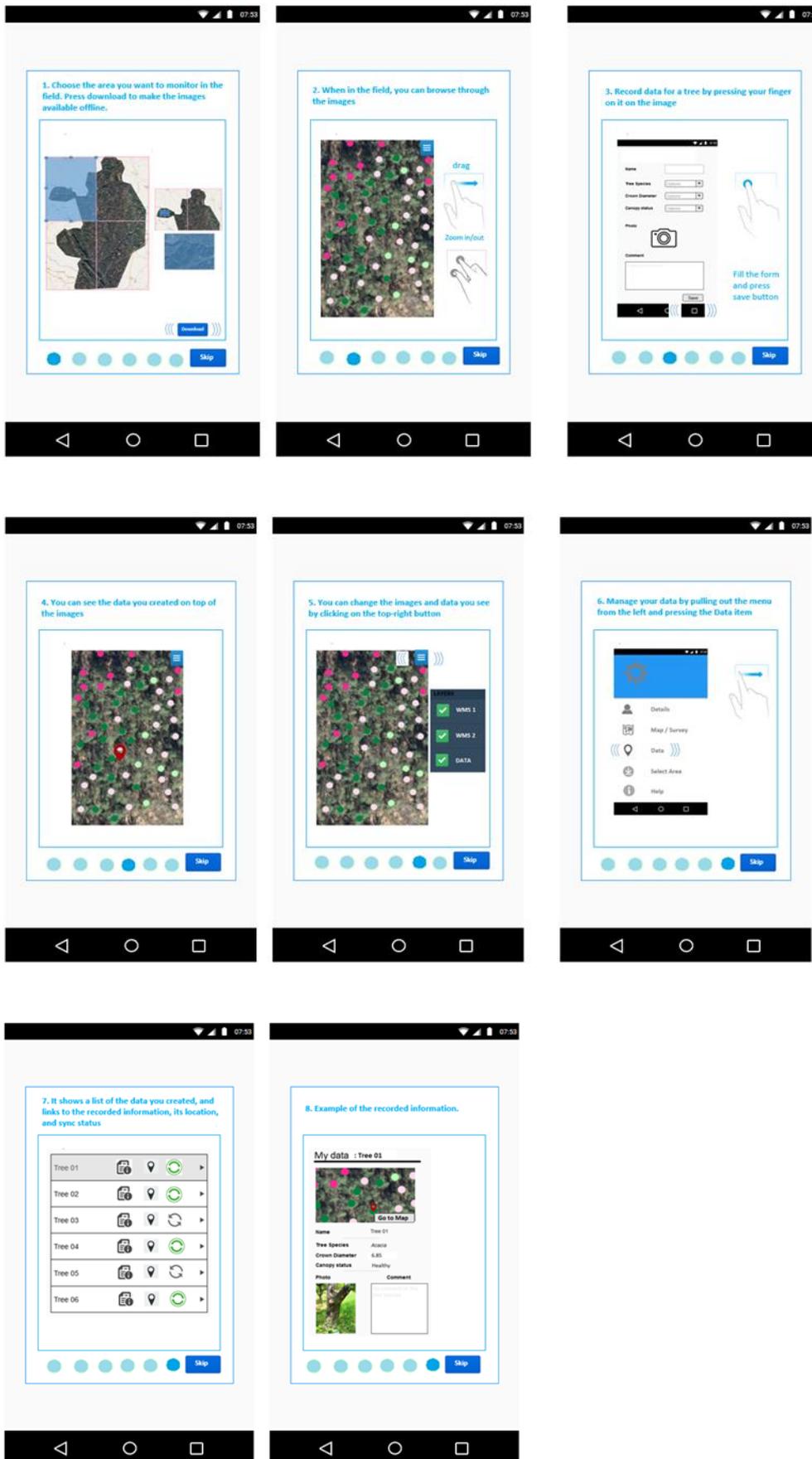
Consequences:

- The user reaches the first access point to the mobile application.

Observations:

- This page corresponds to a page slider in which the user can browse a few pages visually describing the main functionality of the application.
- This page is shown when the application is opened for the first time after installation, or when the user has clicked on the Info menu item in Menu View.
- Main menu contains a Help button.

### 9.3.1 Sequence of sliding pages



### 9.3.2 Select Geographical Zone (GZ) View



#### Context:

- The user has read the explaining pages or has skipped them.

OR

- The user has pressed the Select Area item in Menu view.

#### User-system interaction:

- The application presents the user with a page to select one of the geographical zones (GZ).
- Depending on user permissions, one or more GZs are available to explore.
- The user presses image-like button.

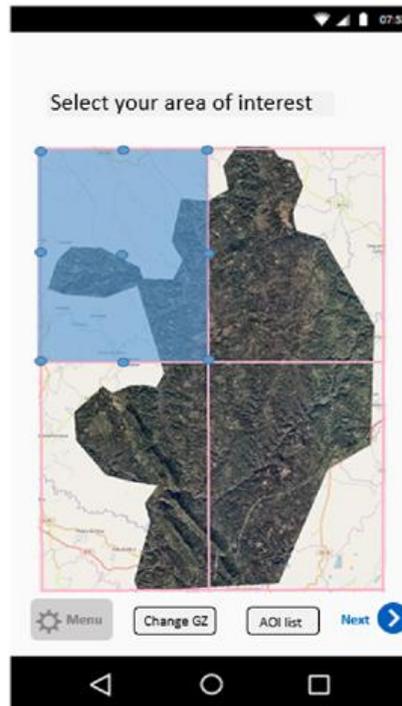
#### Consequences:

- The application goes directly to Select Area view to download images.
- The user selection is saved so it can be recovered in case of application closure/crash.

#### Observations:

- Interaction occurs with back-end via API. Available GZ for logged user is fetched from the back-end. Per-user zone information is available even if offline. It is thus pre-fetched when the user is logged in.
- Each GZ corresponds to different layers from a WMS service. The mobile application displays a screenshot of the GZ.
- If the user has permissions to a GZ, its image-like button is enabled; if the user does not have permission to GZ its image-like button is disabled.

### 9.3.3 Select Area of Interest (AOI) View



#### Context:

— The user has selected a desired GZ.

AND

— The user has internet connection.

#### User-system interaction:

— The application presents the user with a page to create the area of interest (AOI) to survey over a background image. The user can zoom the area to explore in the GZ.

— The pre-selected area is shown as a blue coloured semi-transparent rectangle that can be easily altered:

- using points to move the corners and change the extent around a central point,
- by dragging the rectangle to move it.

— The maximum size of the AOI should be at least 15 km<sup>2</sup>.

— The user gives a name for the new AOI (e.g. AOI 01).

— The user presses the Download button.

— The application shows the progress of the download process.

#### Consequences:

— The application redirects the user to the Menu view when the user presses the Menu button.

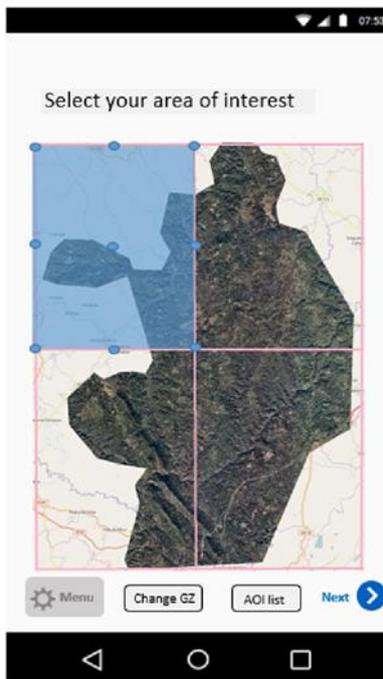
— The user can decide to select another AOI by clicking the Select Area item in Menu view.

- The user can check all AOIs created by clicking on AOI List button (and eventually delete them).
- If the selected AOI is larger than a predefined threshold, a warning message appears, and images are prevented from being downloaded.
- If no Wi-Fi connection is available, a warning message informs the user that downloading is recommended over a Wi-Fi connection.
- If offline, download is not possible, and a full screen message informs the user that Wi-Fi connection is recommended to download images for the AOI.

**Observations:**

- The application downloads images (background images and points from WMS server) falling within the selected extent and makes them later available on the terminal regardless of internet connection.
- The application grabs additional information like AOI extent (BBOX).
- Interaction with back-end via API is required.
  - AOIs created by the logged in user will be fetched from the back-end to display in AOI List View.
  - The application will push AOI data to the server.

**9.3.4 Select AOI – Part 1**



**User-system interaction:**

- The user selects the desired extent to explore by changing the blue coloured transparent rectangle, zooming and/or moving it.
- The user presses the Next button to go to the next view (Select AOI – Part 2).
- The user can press the Menu button to go to Menu view.
- The user can press the AOI list button to go to AOI List View.
- The user can press the Change GZ button to go to Select GZ View.

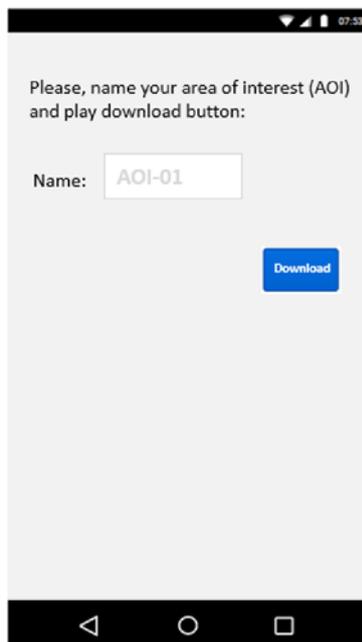
### 9.3.5 Select AOI – Part 2



User-system interaction:

- The application displays the selected area as described in the previous step.
- The user presses the Next button to go to next view (Select AOI – Part 3).
- The user can press the Menu button to go to Menu view.
- The user can press the AOI list button to go to AOI List View.
- The user can press the Change GZ button to go to Select GZ View.

### 9.3.6 Select AOI – Part 3



User-system interaction:

- The application displays a form with a text box.
- The user names the created AOI.
- The user presses the Download button to download images from WMS servers falling within that area.

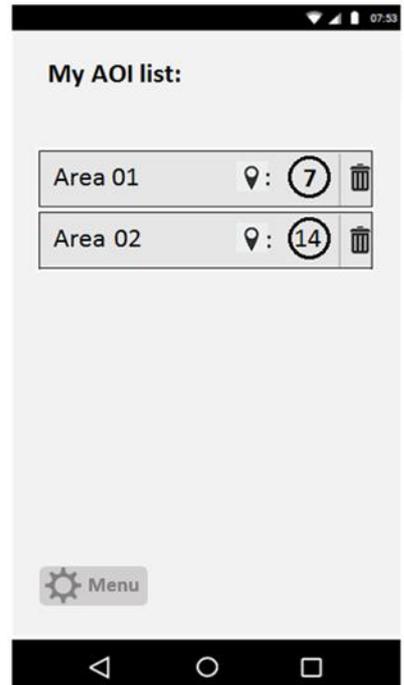
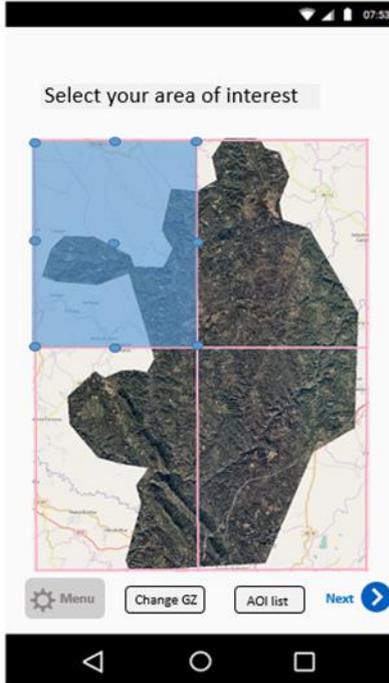
### 9.3.7 Select AOI – Part 4



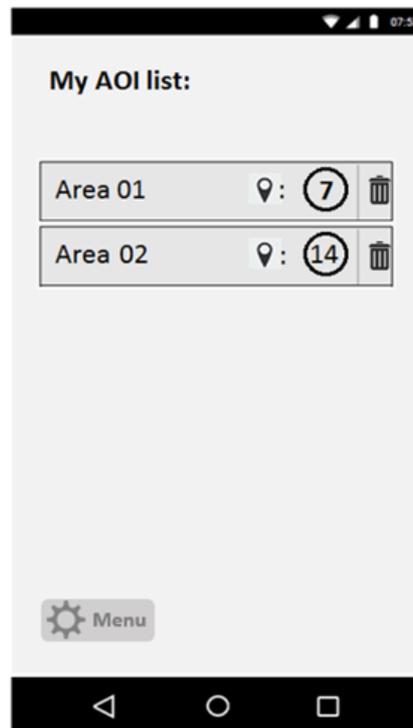
User-system interaction:

- The application shows information about the download process.
- The application redirects to AOI List View.

### 9.3.8 AOI actions



### 9.3.9 AOI List view



**Context:**

- The user has clicked on the Area List button in (Data) List view.

**User-system interaction:**

- The application presents a list of all AOIs created and downloaded by the user in the Select AOI view.
- A box displays the number of survey data falling within that AOI.
- There is also a trash button to delete that AOI.

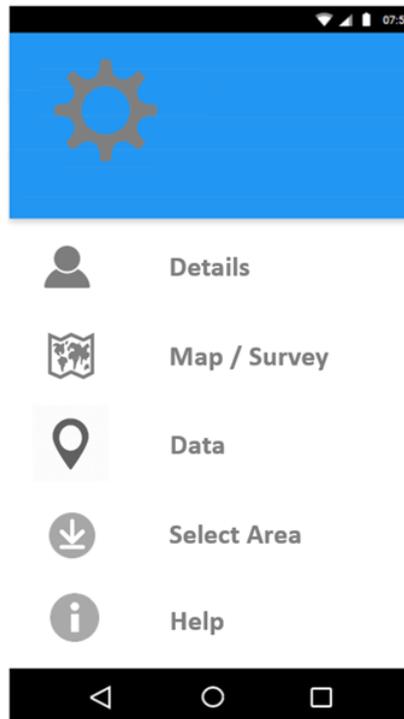
**Consequences:**

- The application provides a list of all AOI areas explored for a given GZ.

**Observations:**

- The user can delete AOI items created but cannot create one from this page. To create a new AOI, they need to return to the menu.
- AOI data deleted in the application does not imply AOI data or survey data will be deleted in the external database.
- Interaction occurs with back-end via API.

### 9.3.10 Menu view



#### Context:

— The user has clicked on the top left menu in Map view.

OR

— The user has clicked the settings button from Select AOI view, Profile, AOI List.

#### User-system interaction:

— The application presents the user with five clickable items:

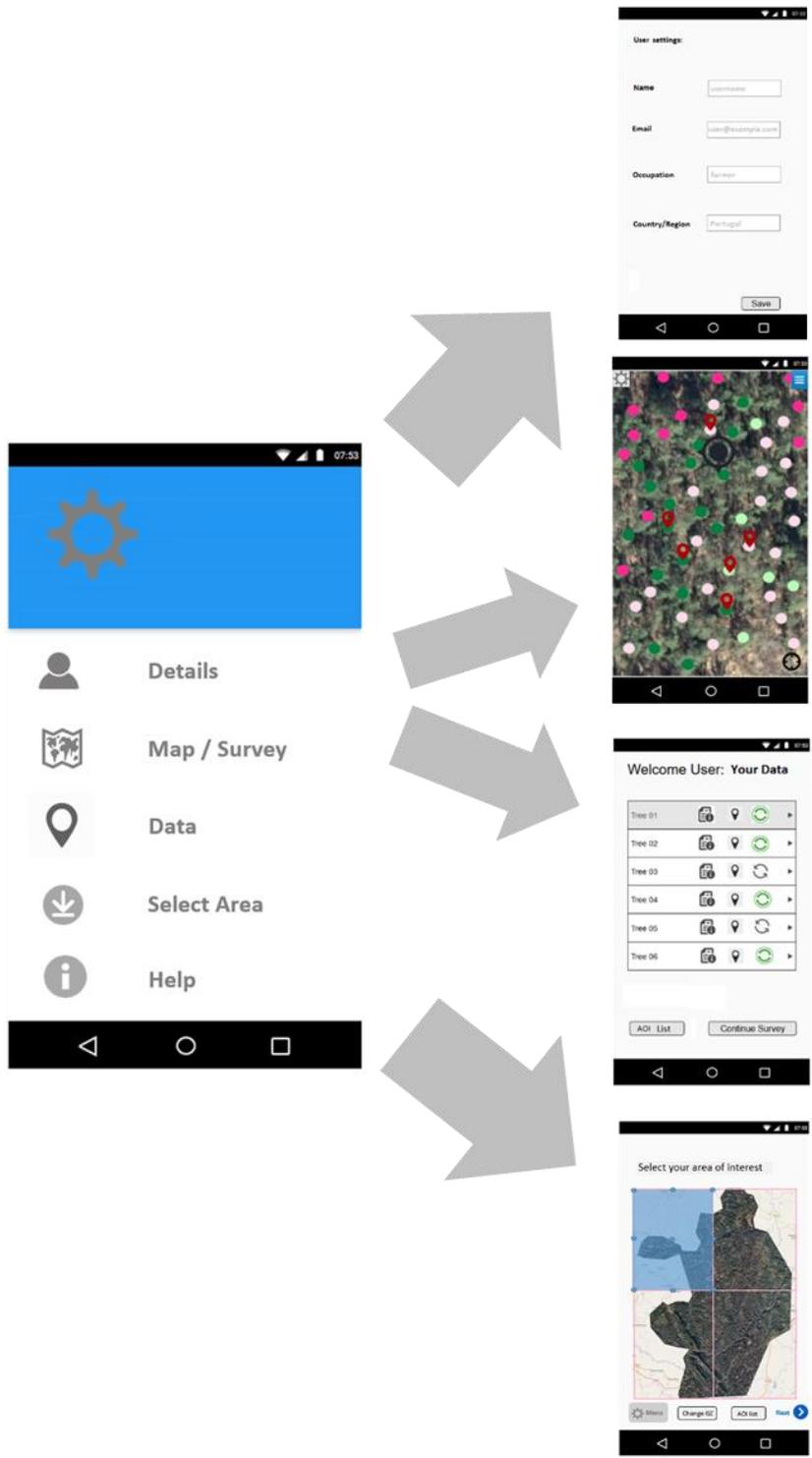
- When the user clicks the Details item, the menu disappears to yield to Profile view.
- When the user clicks the Map/Survey item, the menu disappears to yield to Map view.
- When the user clicks the Data item, the menu disappears to yield to Data view (List View).
- When the user clicks the Select Area item, the menu disappears to yield to Select AOI view.
- When the user clicks the Help item, the menu disappears to yield to explaining page.

#### Consequences:

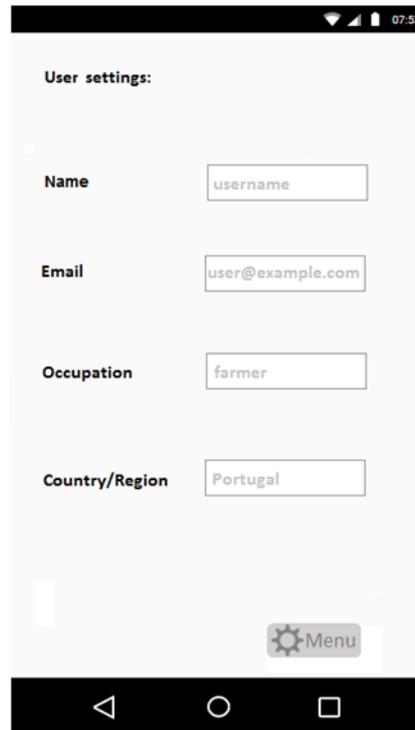
- The user can see (but not edit) information about their profile (as registered by administrators).
- The user can see information about the AOIs they have downloaded when online.
- The user can see information about the data they have created during ongoing or past surveys.

— The user can see information about maps served through an embedded external WMS service.

### 9.3.11 Menu actions



### 9.3.12 Profile view



#### Context:

- The user has opened the main menu and clicked the Details item.

#### User-system interaction:

- The application presents the user with a Profile view.
- The user can slide through Profile view to see, but not edit, information about the user:
  - Name
  - Email
  - Occupation
  - Country/Region
- The user can return to the menu screen by pressing Back button or/and Menu button.

#### Consequences:

- The application provides the user with information about their profile.

#### Observations:

- Items are not clickable. The user cannot change profile settings.
- Interaction with back-end via API is required.

### 9.3.13 Map view (Survey view)



#### Context:

— The user has selected an AOI and downloaded images for this area.

OR

— The user has access to the Map view from different views: Menu view, List view and Detail view (Data).

#### User-system interaction:

— The application displays an interactive map centred either on the selected AOI or on the location detected by the device's GPS sensor (and/or other location detectors).

— The user's location and direction in the field is shown on top of the map with a GPS/compass icon.

— The user can browse through the map by touching the screen:

- one-finger or two-finger swipe to drag the map,
- two-finger pinch for zoom in/out.

— The user can see the map legend and change the visibility of different layers by clicking on the menu icon in the top right corner of the screen.

— The user can re-centre the view on the location identified by the device by clicking on the bottom right button.

— The user can see the application main menu by clicking on the top left corner menu.

#### Consequences:

— This is the landing page that the user arrives on when launching the mobile application.

- The user can navigate and zoom to the desired location, at several zoom levels, to look for features or add new data.
- The application can display additional spatial information (on top of third-party images, if online):
  - WMS layer showing different layers (images + points),
  - data layer showing the data recorded by the user.

Observations:

- Map controls such as zoom in/out and drag are done using touchscreen gestures.
- The application shows WMS images (background and point feature) as previously downloaded.
- The application shows red locator icons showing the data the user has recorded in the area covered by the device screen. The data layer is not visible in the map in two situations: 1) when the user has not yet created any data, and 2) when the created data does not fall within the current map extent.
- If no images fall within the user location and the user is online, OpenStreetMap is shown.
- This map view is available even if the user has no internet connection.

### 9.3.14 Map view actions: load legend and change layer visibility



User-system interaction:

- The user clicks on the blue menu icon in the top right in Map View.
- The application automatically shows a floating panel on the top right side with map legend information. Basically, the legend shows the various colours with the respective attributes for WMS point features and user-created geodata on top of the downloaded background images.
- The user toggles the visibility of a layer (WMS layer/data layer) by clicking on the checkbox for the legend item.
- The user clicks again on the top right menu icon to make the legend panel disappear.

Observations:

- Points are not clickable; the user cannot get information about points beyond their legend.
- Geodata (MY DATA in the legend) are clickable.

### 9.3.15 Map view actions: load information by clicking on data icon



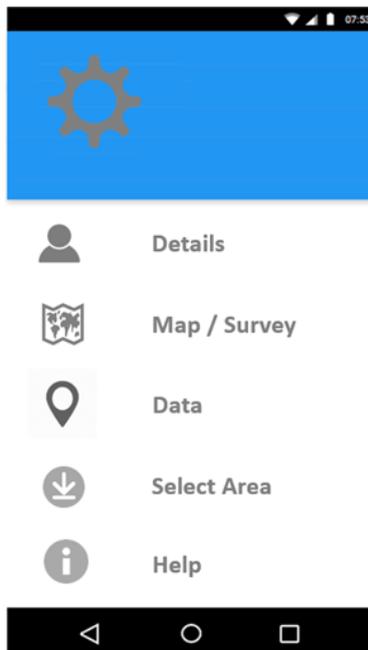
User-system interaction:

- The user clicks on a geodata icon.
- The application automatically shows a pop-up message located around that pressed icon with the name of the corresponding tree species.
- User can select another geodata icon.
- When another location with a MY DATA record is pressed, a new pop-up message appears and the earlier pop-up message disappears.

Observations:

- The user cannot get information about individual points (circles in the mock-up) for which they have not recorded data.
- Geodata are clickable.

### 9.3.16 Map view actions: go to Menu



User-system interaction:

- The user clicks on the top left menu icon in Map view.
- The application automatically shows the Menu view.
- The user can recover the Survey view by clicking on Map/Survey item.

Observations:

- See more information in Menu View.

### 9.3.17 Map view actions: add data to survey

The screenshot shows a mobile application interface with a title bar at the top displaying signal strength, Wi-Fi, and battery icons, and the time 07:53. The main content area is titled "My data" and contains several input fields: "Name" with the text "Tree 01", "Tree Species" (empty), "Crown Diameter" with a dropdown menu showing "Options", and "Canopy status" with a dropdown menu showing "Options". Below these is a "Photo" section with a camera icon, and a "Comment" section with a text area containing the placeholder "Write your comment". A "Save" button is located at the bottom right of the form. The bottom of the screen shows the standard Android navigation bar with back, home, and recent apps icons.

#### User-system interaction:

- The user is browsing in Map view.
- The user wants to create a geodata entry.
- The user long-presses with their finger in the location where the data should be assigned.
- The application presents a form-based page with various fields to be completed by the user.
- The user enters the fields, with choices from dropdown menus, options to take photographs using the mobile device, and option to add free text in the comment box, and presses the Save button.
- The application returns to Map View. Map view now also shows the new georeferenced data record with a red coloured locator icon.

#### Observations:

- See more information in Create View.

### 9.3.18 Data views (multiple): List view



#### Context:

- The user has pressed the Data button in the main menu (Settings view).

#### User-system interaction:

- The application presents a list of named items for all data the user has created (for any geographical area).
- The user can scroll through the list, if it extends beyond the screen, by sliding their finger up and down.
- The user can click on the first (i.e. left) icon for each item to see more information.
- The user can click on the second (i.e. middle) icon for each item to go directly to a map view centred on this data record.
- The third (i.e. right) icon shows the user whether the data item is in sync with the back-end.
- The user can click on Continue Survey button and go to Map View.
- The user can click on AOI List button and go to AOI List View.

#### Consequences:

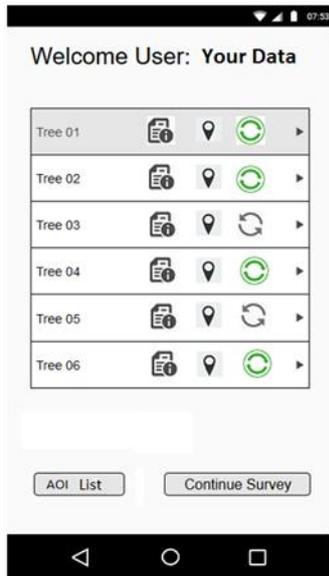
- The user can see all in-situ data they have created on this screen.
- The user can get more information on individual data entries by clicking on the first icon.

- The user can go directly to a map view of any data record by clicking on the second icon.
- The user can check whether a data entry is in sync with an external server.

Observations:

- Items created by the user are stored in the local storage.
- Interaction with back-end via API is required. In case of no connectivity, all survey data will be pre-fetched via API after the user logs in, and displayed with sync status when List view is loaded.
- If online, data synchronisation with the external server should take place in the foreground. If offline, data synchronisation is not achieved.
- Each item is shown with the name given to it by the user in Edit View.
- Interactions via API, such as pushing newly created survey data and pulling/fetching survey data from the server when online.

### 9.3.19 Data views (multiple): List view – Actions



List View



Detail View



Map view



### 9.3.20 Data views (multiple): Create view

The screenshot shows a mobile application interface for creating a data entry. At the top, the title 'My data' is displayed. Below the title, there are several input fields: 'Name' with the text 'Tree 01', 'Tree Species' (an empty text box), 'Crown Diameter' (a dropdown menu showing 'Options'), and 'Canopy status' (a dropdown menu showing 'Options'). Below these fields is a 'Photo' section with a camera icon. At the bottom, there is a 'Comment' section with a text area containing the placeholder 'Write your comment'. A 'Save' button is located at the bottom right of the form. The interface is framed by a black header and footer, with the footer containing standard Android navigation icons.

#### Context:

- The user is browsing the map and wants to create a data item in the desired location.
- AND
- The user has long-pressed with their finger in that location.

#### User-system interaction:

- Application presents a form-based page with various fields.
- The user fills the various form fields.
- Both Tree Species and Comment are text boxes that the user can fill freely. When the user enters text in the Tree Species field, the application provides auto-complete suggestions.
- Photo button allows the user to either go to Camera to take picture(s) or go to User Photo Gallery to select a photo there. After selection, a thumbnail for the image is attached and, if desired, the user can add more images.
- Crown Diameter and Canopy status are select boxes with predefined values and, if pressed, a full-screen view with a list of possible values is displayed. After selection, the system is backed to edit the form.

#### Consequences:

- When the Save button is pressed, data is stored locally on the mobile device.
- The application redirects to List view, with the message 'Your data was correctly saved'.
- Data synchronisation with the external server is achieved only when Wi-Fi connectivity is available.

- Once data is uploaded to the external server, the system will present the data item in the List view with a synced icon.
  - If data is stored but not yet uploaded, the system will present the data item in the List view with an unsynced icon.
- The application captures additional information:
- creation date timestamp (DD:MM:YY – hh:mm:ss),
  - spatial coordinates for the location pressed (longitude/latitude),
  - compass direction for the user location (degrees relative to north),
  - internal identifier for the recorded item.

**Observations:**

- Only the name attribute has a default value: *Tree XX* where XX is an auto-incremental value when subsequent data are created. User can decide whether to change this when creating or updating the data.
- Except for the Comment field, all information fields in this form are compulsory.

**9.3.21 Data views (multiple): Create view – Actions**



### 9.3.22 Detail view



#### Context:

- The user has clicked on the Edit button within the Detail view.

#### User-system interaction:

- The application presents a list of changeable items for all data as created by user.
- The user can click on the corresponding attribute fields: Name, Tree Species, Crown Diameter, Canopy status.
- The user can click on the photo box.
- The user can click on the Comment field.
- The user can click and move the red pin in the top browsable map.
- The user can click on the Save button to save changes in Edit view.

#### Consequences:

- The user can see on the screen all the data recorded in situ for a given location.
- The application displays an Edit form, allowing the user to change the attributes registered for the location.
- The application displays a Select form for changing Crown Diameter and Canopy status (see Create View).
- The application displays the camera options, to change photo or upload a new one.

#### Observations:

- When changes are made in individual Edit/Select forms, the application goes to Edit view.
- When changes are saved by pressing the Save button, the application goes to List View, with the message 'Your data were correctly modified'.
- The application retains additional information about modification and stores the date and timestamp of the update (DD:MM:YY – hh:mm:ss).

### 9.3.23 Delete view



#### Context:

- The user has clicked on the Delete button within the Detail view.

#### User-system interaction:

- The application presents an information message, asking them to confirm deletion of the recorded data item.
- The user can click on Yes button.
- The user can click on No button.

#### Consequences:

- When No button is pressed, the application takes them to Edit view.
- When Yes button is pressed, the application takes them to Detail view.

## 9.4 External server

The application is a survey tool, to both visualise external background images and to create and manage georeferenced data from different users. All data created are stored locally on the user device, to reduce dependence on network connection during data collection. Once an internet connection is available, the data are synchronised. This synchronisation includes the uploading of all data collected to an external server.

The mobile application currently displays pages only in English but the development should have potential to jump to multilingual later.

### 9.4.1 Data typology in application workflow

The data collected, as well as all other information, will eventually be stored on an external server, not only for proper functioning of the application itself, but also for further processing tasks and analysis.

User permissions allow the user to visualise different spatial layers, called **geographical zones** (GZs) throughout this document. Depending on a given user's membership of pre-defined user groups, they can access the content of one, several, or all available GZs.

After login, the application should display the information the user has provided for:

- geographical zones, available according to the user's role/group,
- area of interest created by the user when online,
- survey data created by the user (for that particular AOI) in the field.

For a given GZ, the user can access, download and query distinct **areas of interest** (AOIs). For a given GZ, the user can manage all AOIs downloaded. For each AOI, the user can create new survey data from the application's Map view, as well as manage data created in previous survey sessions from the application's List Detail views.

All this information and settings should be:

- fetched/pulled from the external database to be correctly displayed and managed by the application for all different logged in users,
- sent/pushed to the external database when the user selects/creates/modifies/deletes any of those pieces of information.

The types of information (typology) to be saved in the remote database encompass:

- users
- geographical zones
- areas of interest
- survey data
- data label (Tree species, Crown diameter, Canopy status)
- photo.

### Users

Relevant data to be saved (minimum):

- user\_id
- name
- username
- password

- email address
- occupation
- country
- language

### **Groups/Roles**

Relevant data to be saved (minimum):

- group\_id
- group\_name
- role

### **Country**

A list of all countries that encompass the geographical zones. These will be identified by country code; for instance, ES = Spain, IT = Italy.

### **Geographical zone (GZ)**

Relevant data to be saved for each accessible GZ (minimum):

- gz\_id
- gz\_name
- country
- layer\_name
- wms\_url
- proj

'layer\_name' is the name of the layer as published on the WMS server.

'wms\_url' is the URL to access the WMS server.

'proj' is the projection of the layer.

Observation: GZs are connected to the Group table through an M2M relationship.

### **Area of interest (AOI)**

Relevant data to be saved for each AOI created by the user (minimum):

- aoi\_id
- aoi\_name
- Xmin
- Xmax
- Ymin
- Ymax
- gz\_id
- creation\_date
- is\_deleted

'aoi\_name' is auto-generated by the application (e.g., AOI-01), but can be changed by the user during AOI creation.

'Xmin', 'Xmax', 'Ymin', 'Ymax' is the BBOX for the area the user wants to download. Units are related to the projection of the GZ.

'creation\_date' is the date timestamp for the AOI (format: DD:MM:YY - hh:mm:ss).

'is\_deleted' is a Boolean variable that indicates whether an AOI has been deleted by the user.

'gz\_id' refers to the GZ associated with that AOI.

### **Tree species**

— tree\_specie\_id

— tspec\_name

'tspec\_name' will be from a list of pre-selected names for tree species, to be provided by the JRC. The application uses this list to provide the user with auto-complete functionality, when the Tree Species field is filled out in Data Create view. Additional data can be added by users through the application.

### **Crown diameter**

— crown\_diameter\_id

— cdiameter\_name

'cdiameter\_name' will be numerical values between 0.1 m and 30 m, with 0.25 m increments.

### **Canopy status**

— canopy\_status\_id

— cstatus\_name

'cstatus\_name' can only be these text values: Healthy, Stressed, Sick, Dead.

### **Photo**

— photo\_id

— sd\_id

— compass

— longitude

— latitude

'sd\_id' refers to the survey data table. More than one photo can be associated with a given record in the survey data.

'longitude' and 'latitude' is the pair of spatial coordinates that determine the geolocation of a point in relative projection. These are detected by the mobile terminal, with precision of up to six decimals.

'compass' is the orientation of the geolocated photo, relative to north (perhaps in a metadata jpg file).

## **Survey data (SD)**

Relevant data to be saved for each SD created by the user in a survey session:

- sd\_id
- sd\_name
- tree\_specie\_id
- crown\_diameter\_id
- canopy\_status\_id
- comment
- aoi\_id
- gz\_id
- user\_id
- creation\_date
- update\_date
- longitude
- latitude
- compass
- is\_deleted

"sd\_name" would be auto-generated by the application (e.g., Tree-01), but can be changed by the user during SD creation.

'comment' will be text filled by the user, in the form displayed in Data Create view.

'crown\_diameter' and 'canopy\_status' will be one of the following multiple-option values:

- Crown diameter → numerical values between 0.1 m and 30 m, with 0.25 m increments,
- Canopy status → text values: Healthy, Stressed, Sick, Dead.

'tree\_specie\_id' refers to the tree species covered by this SD.

'photo' refers to image/images (jpg); the value is the path to the image saved on the server.

'aoi\_id' refers to the AOI with which this SD is associated.

'gz\_id' refers to the GZ with which this SD is associated.

'user\_id' refers to the user who created the SD.

'creation\_date' is the date timestamp for the SD (format: DD:MM:YY - hh:mm:ss).

'update\_date' is the date timestamp for the SD (format: DD:MM:YY - hh:mm:ss).

'longitude' and 'latitude' is the pair of spatial coordinates that determine the geolocation of a point in relative projection. These are detected by the mobile terminal, with precision up to six decimals.

'compass' is the orientation of the user device. Units are degrees relative to north.

'is\_deleted' is a Boolean variable that indicates whether an SD has been deleted by the user.

### 9.4.2 Data storage

All these data are saved in permanent storage such as a relational database management system, essentially PostgreSQL or MySQL. All pieces of information should be stored in various tables inside a database, maintaining the correspondences between them, to guarantee the integrity and updating of those data. No database schema is imposed, but we can at least state that it could contain the following tables:

- USER
- GROUP
- COUNTRY
- GEOGRAPHICAL\_ZONE
- AREA\_OF\_INTEREST
- SURVEY\_DATA
- TREE\_SPECIES
- STATUS\_CANOPY
- CROWN\_DIAMETER
- PHOTO
- METADATA

### 9.4.3 Application-server communication

Communication between the mobile application and the external server is done via HTTP protocol. Data coming from the server will be managed by the application for different purposes, basically editing and visualisation.

Spatial layers, referring to spatial data such as background images and points (GZs) are published on WMS servers and are accessible via HTTP endpoints.

To establish communication between the mobile application and the available external server, a REST API (Representational State Transfer / Application Programming Interface) should be implemented, to be consumed by the mobile application. The mobile application will act as consumer of the API, connecting transparently (push/pull) to objects on the server via HTTP in order to perform create, update and delete (CRUD) operations on data. This process should ensure synchronisation between the client device and the remote server.

It should be pointed out that permanent network connectivity is required to push data to the remote server/pull data from the server. The mobile application should be aware of this limitation, pre-fetching data needed to be completely operational during application usage. Business logic should be correctly implemented to only select AOI images when the terminal is online, and to be available when the user is not online and entering survey data.

Both user profile information and user permissions have been stored to guarantee that a certain number of users can log in and access relevant information. An authentication mechanism should be implemented; for instance, tokens or sessions via API to authorise access to the various content.

This REST API should be well documented, providing among other things information about URL endpoints, CRUD type operation on different data, HTTP verbs to access/modify resources on demand, and output format.

The following collections could be defined (minimum):

- gzone: geographical zone
- aoi: area of interest

- sdata: survey data
- user: user
- label: label
- photo: photo
- settings: settings

If we hold Tree-checker as the domain name for the external server, the following would be an example of some possible endpoints:

- for 'gzone' collection

URL endpoints	CRUD	HTTP verbs	Description	Output
<a href="http://domain/gzone/&lt;gzone_id&gt;">http://domain/gzone/&lt;gzone_id&gt;</a>	Read	GET	gzone_id is the unique identifier of the geographical zones you want to retrieve  Return gzone with the specified gzone_id, or else 404 response	JSON
<a href="http://domain/gzone">http://domain/gzone</a>	Read	GET	Return a list of geographical zones	JSON

- for 'aoi' collection

URL endpoints	CRUD	HTTP verbs	Description	Output
<a href="http://domain/aoi/&lt;aoi_id&gt;">http://domain/aoi/&lt;aoi_id&gt;</a>	Read	GET	aoi_id is the unique identifier of the AOI you want to retrieve  Return AOI with the specified aoi_id, or else 404 response	JSON
<a href="http://domain/aoi">http://domain/aoi</a>	Read	GET	Return a list of AOI	JSON

— for 'sdata' collection

URL endpoints	CRUD	HTTP verbs	Description	Output
<a href="http://domain/sdata/&lt;sdata_id&gt;">http://domain/sdata/&lt;sdata_id&gt;</a>	Read	GET	sdata_id is the unique identifier of the sdata you want to retrieve  Return sdata with the specified sdata_id, or else 404 response	GeoJSON
<a href="http://domain/sdata">http://domain/sdata</a>	Read	GET	Return a list of sdata	GeoJSON
<a href="http://domain/sdata/&lt;sdata_id&gt;">http://domain/sdata/&lt;sdata_id&gt;</a>	Create	POST	sdata_id is the unique identifier of the sdata you want to retrieve  Return 201 response with sdata with the specified sdata_id created	GeoJSON
<a href="http://domain/sdata/&lt;sdata_id&gt;">http://domain/sdata/&lt;sdata_id&gt;</a>	Update	PUT	sdata_id is the unique identifier of the sdata you want to update  Return sdata with the specified sdata_id, or else 404 response	GeoJSON
<a href="http://domain/sdata/&lt;sdata_id&gt;">http://domain/sdata/&lt;sdata_id&gt;</a>	Delete	DELETE	sdata_id is the unique identifier of the sdata you want to delete  Return 404 response if sdata_id does not exist	GeoJSON

— for 'user' collection

URL endpoints	CRUD	HTTP verbs	Description	Output
<a href="http://domain/user/&lt;user_id&gt;">http://domain/user/&lt;user_id&gt;</a>	Read	GET	user_id is the unique identifier of the user you want to retrieve	JSON

			Return user with the specified user_id, or else 404 response	
--	--	--	--	--

#### 9.4.4 Remote Server and deployment

CentOS is the operative system used in our back-end servers to deploy various Python-based solutions and web applications.

This external server should be live and will act as a web server to deliver web services through API, as outlined in this document.

These are the components of the system to be provided as a final product:

- database schema,
- REST API,
- source code for the mobile application.

A script or other utility tool should be provided for the deployment of this infrastructure.

## **10 What to consider when developing similar systems**

### **10.1 Every case is different**

When thinking about remote sensing-based support for tree pest management, it is crucial to recognise that every case is different and there is no 'one size fits all'. In the case of the pine wood nematode, the timing of image acquisition is critical as it links to the biology of the species, and the timing of required intervention. The aim of identifying individual trees, rather than stands of trees, dictates the image required. The fact that trees with obvious visible signs of decline, regardless of the cause, needed to be detected made high spectral resolution less important and meant that 4-band broadband sensors should be suitable, particularly given the size of the area. These considerations - and the associated trade-offs - differ between pest outbreaks, and recommendations should therefore be made on a case-by-case basis. The following is true where individual trees with canopies in clearly declining health need to be detected.

### **10.2 Platform choice**

The need for individual tree detection limits the choice of platforms from which data can be used, as it constrains spatial resolution. None of the satellites from the Copernicus programme, for example, provide images at high enough spatial resolution to distinguish individual trees. Some commercial satellites do acquire images with the necessary resolution, but for large areas, their cost is on a par with airborne campaigns. Moreover, they tend to have less flexibility in deployment than aircraft, as one depends on satellite overpass time and competitive scheduling between acquisition needs.

The acquisition of orthophotos from aircraft has a history going back further than that of satellite-based imaging. As historically such photographs were used for visual rather than quantitative analysis, orthophotos, unlike passive optical satellite images, are rarely processed to express physical units such as surface radiation or surface reflection, which hampers their automated analysis.

More recently, drones have gained a lot of attention for image acquisition; however, their autonomy, and particularly the legal restrictions on their deployment, means they are not yet suitable to collect imagery over very large areas.

Smallsats provide a new platform for image acquisition; they rely on off-the-shelf technology to keep the cost of individual satellites down, and on mass deployment to ensure short revisit times. However, their cost savings on the instrumentation side mean that, for the time being, their geometric and spectral accuracy remains low.

In the coming decades, it is likely that high altitude pseudo-satellites will provide an additional unmanned platform for high-resolution image acquisition.

### **10.3 Metadata matters**

There are a number of opportunities to take aerial photographs, as they are commonly acquired during photogrammetric campaigns, which are more useful for plant health monitoring. Historically, photogrammetric campaigns aimed to generate large-image mosaics to assess the state of the landscape or territory in a given period. Hence, such collections are referred to by the year of acquisition (e.g. [13]). To assess plant health status, more temporal detail is however required; phenology is very important to both pest biology and the assessment of plant health, particularly for deciduous trees, but also for non-deciduous ones. Moreover, this time-variant role in biological process is closely intertwined with opportunities for pest management efforts. The timing of the data acquisition should thus strike a balance between the optimal date for measurement based on epidemiology and physiology, and maximum usefulness for phytosanitary interventions, also allowing enough time for image processing and analysis.

When acquiring large collections of orthophotos, it is important that the acquisition date and time be included in the metadata, ideally on a per pixel basis. In the ADS100 data acquired for this project, the metadata corresponded instead to the original flight lines from which data were drawn for the final image mosaic, and contained the start and end date and time for each flight line.

Standard photogrammetric cameras have only four broad bands; when deployed in a purpose-designed campaign, this proved sufficient to detect obvious signs of decline in coniferous trees. As a general rule, automated algorithms analysing such images will not detect features that cannot be distinguished by the human eye in three-band composites made from these images. In many other pest management settings, it is paramount that much earlier stages of decline, indicating potential infection rather than suitability for the vector, are detected. In such cases, photogrammetric cameras are not adequate for the task, and hyperspectral cameras, or multispectral cameras with a carefully selected bandset, should be deployed [14].

Analysis of very high resolution images over large areas requires human resources combining expertise in remote sensing, big data handling, and plant health. Furthermore, adequate computational resources are required, so that data can be analysed quickly enough for the results to be relevant and timely enough for interventions on the ground. In most scenarios, such analysis greatly depends on the quality and availability of detailed georeferenced ground observations for calibration and validation of results.

Recent developments in image analysis techniques include methods that harness deep learning neural networks, such as TensorFlow developed by Google, to detect objects in images. These machine learning techniques can be powerful with sufficient training, which can be expedited by 'transfer learning', where a neural network for a new task leverages networks performing similar tasks. Several deep learning neural networks are available as open source libraries and will gain traction for image analysis over the coming years.

## **10.4 Non-image data**

Remote sensing analysis is unlikely to replace field and laboratory inspections for pest management in the foreseeable future. Instead, it is important that these different data sources are analysed in conjunction. In the scientific world this is commonly done, or even expected, through coordinated field and airborne imaging campaigns. Coordination and data flows between field, laboratory, and remote sensing operations will greatly improve the contribution of the latter in any operational deployment of remote sensing in plant health management.

## **10.5 Making use of existing data**

### **10.5.1 Orthophoto archives**

Particularly when using spectrally poor data, such as orthophotos, it is crucial to fully exploit the temporal dimensions of the data, when available, to overcome some of the shortcomings of the single-date imagery. In this project, for example, we used older orthophotos together with the newer ones, to detect areas that had recently been logged or planted (and were not distinguished as such in the land cover maps). These detections allowed us to mask such areas from analysis, as they are highly unlikely to contain fully-grown crowns. This masking allowed the image classification algorithms to be trained more quickly, as they did not have to 'learn' to distinguish such areas as not containing declining trees. As alternatives, LiDAR datasets - where available, up-to-date, and sufficiently detailed - could be used to distinguish areas with and without trees, or even individual tree crowns.

### **10.5.2 Land cover maps**

High resolution land cover maps can perform a similar function, allowing remote sensing analysis to focus on areas of interest and ignore those that do not contain the host species

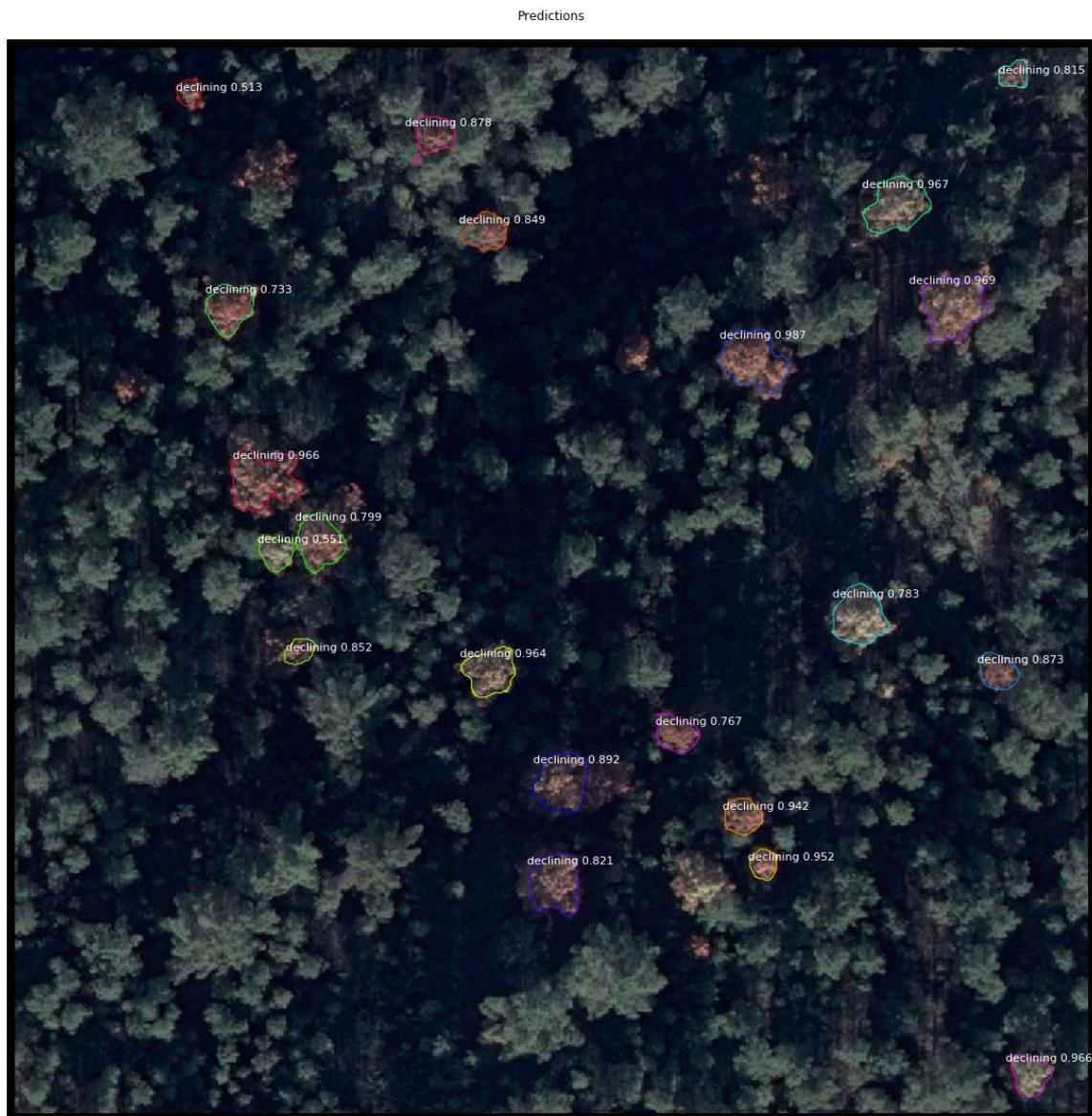
of the pest in question. Land cover maps are not always up to date. However, they are usually younger than trees with fully developed crowns, meaning they can be relied upon to identify areas with mature forest, but might not distinguish recently cleared or replanted forests.

## **10.6 New generation of image analysis algorithms**

A new generation of deep learning algorithms is providing new opportunities to accelerate and generalise image analysis algorithms. At the end of the project, we performed tests with neural networks to detect declining trees in the multispectral images (15 cm DMC) acquired over the study area. We chose and implemented a neural network based on instance segmentation, which combines object detection and pixel-wise segmentation to not only detect specific objects in an image, but also to outline them individually.

We used a training dataset of 350 crowns, detected in different zones of the study area through the methods described in this report. We trained the neural network through 80 epochs of 50 iterations each (where an 'epoch' describes the number of times the algorithm has seen the *entire* dataset and an 'iteration' describes the number of times a *sample* of the data passes through the algorithm). This model training phase was fairly limited, as the loss function indicated that model performance would improve with more epochs and iterations. Nonetheless, preliminary results show near-complete detection and delineation of known declining trees, and a very low false positive rate, in par with those achieved at the end of the MaxEnt iterations (with cases mainly occurring in bare soil areas with similar spectral characteristics to the declining crowns). Based on these preliminary results, further research is ongoing into the use and development of this new generation of image analysis algorithms for plant health applications.

**Figure 21.** A subset of declining tree crowns, as detected and outlined using a neural network based on image segmentation and trained on the 15 cm aerial photographs acquired in the Castelo Branco area in November 2016



## 11 Meetings

The CanHeMon work was presented at multiple internal and external meetings, including:

- October 2015: EFSA's Second Scientific Conference – 'Shaping the Future of Food Safety, Together', Milan, Italy
- March 2016: DG SANTE – JRC CanHeMon progress meeting
- May 2016: First meeting of the EC Expert Group on Forest Health and Pests
- June 2016: Standing Committee on Plants, Animals, Food and Feed, section Plant Health
- March 2017: DG SANTE – JRC CanHeMon progress meeting
- April 2017: ICNF Portugal – JRC meeting about CanHeMon, Lisbon, Portugal
- May 2017: JRC – Forest Europe meeting, Ispra, Italy
- May 2017: MAGRAMA Spain – JRC meeting about CanHeMon, Madrid, Spain
- May 2017: Unmanned Aircraft Systems for Remote Sensing Conference, Hobart, Tasmania
- June 2017: JRC – University of Birmingham meeting on forest monitoring, Ispra, Italy
- June 2017: Standing Committee on Plants, Animals, Food and Feed, section Plant Health
- August 2017: JRC – European Forest Institute meeting, Bonn, Germany
- December 2017: Visit from Commissioner Andriukaitis to the JRC, Ispra, Italy
- February 2018: COST Action ES1309 OPTIMISE Final Conference, Sofia, Bulgaria

## References

- [1] Especificações técnicas da Carta de uso e ocupação do solo de Portugal Continental for 1995, 2007, 2010 and 2015, Relatório Técnico, Direção-Geral do Território, 2018.
- [2] Carta de uso e ocupação do solo de Portugal Continental [Dataset], Version 2010v1, Direção-Geral do Território, Lisbon, Portugal, 2017. Accessed 13 August 2018. <http://mapas.dgterritorio.pt/wms-inspire/cos2010v1>
- [3] GRASS Development Team, Geographic Resources Analysis Support System (GRASS) Software, Version 7.4, Open Source Geospatial Foundation, 2018. Electronic document: <http://grass.osgeo.org>
- [4] Haralick, R. M., Shanmugam, K. and Dinstein, I., 'Textural Features for Image Classification' *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-3, Issue 6, 1973, pp. 610-621. <https://doi.org/10.1109/TSMC.1973.4309314>
- [5] Phillips, S.J., Anderson, R.P. and Schapire, R.E., 'Maximum entropy modeling of species geographic distributions', *Ecological Modelling*, Vol. 190, Issues 3-4, 2006, pp. 231-259, ISSN 0304-3800. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>.
- [6] Gormley, A.M., Forsyth, D.M., Griffioen, P., Lindeman, M., Ramsey, D.S., Scroggie, M.P. and Woodford, L., 'Using presence-only and presence-absence data to estimate the current and potential distributions of established invasive species', *Journal of Applied Ecology*, Vol. 48, Issue 1, 2011, pp. 25-34. <https://doi.org/10.1111/j.1365-2664.2010.01911.x>
- [7] Phillips, S.J., Anderson, R.P., Dudík, M., Schapire, R.E. and Blair, M.E., 'Opening the black box: an open-source release of Maxent', *Ecography*, Vol. 40, Issue 7, 2017, pp. 887-893. <http://doi.org/10.1111/ecog.03049>
- [8] Shannon, C.E., 'A Mathematical Theory of Communication', *Bell System Technical Journal*, Vol. 27, Issue 3, 1948, pp. 379-423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- [9] Phillips, S.J. and Dudík, M., 'Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation', *Ecography*, Vol. 31, Issue 2, 2008, pp. 161-175. <https://doi.org/10.1111/j.0906-7590.2008.5203.x>
- [10] R Core Team, 'R: A language and environment for statistical computing', R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org>
- [11] Hijmans, R.J., Phillips, S., Leathwick, J. and Elith, J., 'dismo: Species Distribution Modeling', R package version 1.1-4, 2017. <https://CRAN.R-project.org/package=dismo>
- [12] Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y.E. and Yates, C.J., 'A statistical explanation of MaxEnt for ecologists', *Diversity and Distributions*, Vol. 17, Issue 1, 2011, pp. 43-57. <https://doi.org/10.1111/j.1472-4642.2010.00725.x>
- [13] Ortofoto 2014 Agea (50cm), Agenzia per l'Italia Digitale, 2018. <https://www.dati.gov.it/dataset/ortofoto-2014-agea-50cm> Accessed 1 May 2018.
- [14] Zarco-Tejada, P.J., Camino, C., Beck, P.S.A., Calderon, R., Hornero, A., Hernandez-Clemente, R., Kattenborn, T., Montes-Borrego, M., Susca, L., Morelli, M., Gonzalez-Dugo, V., North, P.R.J., Landa, B.B., Boscia, D., Saponari, M., Navas-Cortes, J.A., 'Previsual *Xylella fastidiosa* infection revealed in spectral plant-trait alterations', *Nature Plants*, Vol. 4, 2018, pp. 432-439. <https://doi.org/10.1038/s41477-018-0189-7>

## List of figures

<b>Figure 1.</b> Map of Portugal, and outline in red of the 30 cm RGB-NIR images acquired by the project. The grid depicts the 'blocks' in which these data were generated, delivered, and analysed. ....	8
<b>Figure 2.</b> Principle component analysis of the initial set of image-based predictor variables .....	20
<b>Figure 3.</b> Processing and memory usages for jobs executed on the JDEOPP .....	22
<b>Figure 4.</b> Screen capture of the internal website to support the collection of validation data on visual image inspection.....	23
<b>Figure 5.</b> Tree detection ratio as a function of the cut-off used to classify the probabilities generated by the MaxEnt model into presence/absence.....	24
<b>Figure 6.</b> Cumulative ratio of trees retrieved as a function of object size cut-off, whereby all objects under a certain size cut-off (displayed on the x-axis) are retained. ....	25
<b>Figure 7.</b> False discovery rate and positive predicted rate, as a function of object size .26	
<b>Figure 8.</b> Ratio of false positive to true positive detections and cumulative ratio of true positives, as a function of object size cut-off.....	26
<b>Figure 8.</b> Example of seamlines in an image mosaic generated from images acquired on different dates .....	27
<b>Figure 10.</b> Tree detection error ratio, as a fraction of the cut-off used to classify the probabilities generated by the MaxEnt model into presence/absence.....	28
<b>Figure 11.</b> <i>As Figure 5.</i> .....	28
<b>Figure 12.</b> <i>As Figure 5.</i> .....	29
<b>Figure 13.</b> <i>As Figure 8.</i> .....	29
<b>Figure 14.</b> Blue, Green, Red, and Near-Infrared views (clockwise from top left) of a single tree crown, as described by the DMC images (two left columns) and ADS100 images (two right columns). ....	30
<b>Figure 15.</b> <i>As Figure 5.</i> .....	31
<b>Figure 16.</b> <i>As Figure 6.</i> .....	31
<b>Figure 17.</b> <i>As Figure 5.</i> .....	32
<b>Figure 18.</b> <i>As Figure 6.</i> .....	32
<b>Figure 14.</b> <i>As Figure 8.</i> .....	33
<b>Figure 20.</b> Overview of navigation in the TreeChecker mobile application .....	34
<b>Figure 21.</b> A subset of declining tree crowns, as detected and outlined using a neural network based on image segmentation and trained on the 15 cm aerial photographs acquired in the Castelo Branco area in November 2016 .....	71

**List of tables**

**Table 1.** Start and end dates of the periods during which images were acquired to generate blocks of 30 cm RGB-NIR images. .... 8

## **GETTING IN TOUCH WITH THE EU**

### **In person**

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: [https://europa.eu/european-union/contact\\_en](https://europa.eu/european-union/contact_en)

### **On the phone or by email**

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696, or
- by electronic mail via: [https://europa.eu/european-union/contact\\_en](https://europa.eu/european-union/contact_en)

## **FINDING INFORMATION ABOUT THE EU**

### **Online**

Information about the European Union in all the official languages of the EU is available on the Europa website at: [https://europa.eu/european-union/index\\_en](https://europa.eu/european-union/index_en)

### **EU publications**

You can download or order free and priced EU publications from EU Bookshop at: <https://publications.europa.eu/en/publications>. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see [https://europa.eu/european-union/contact\\_en](https://europa.eu/european-union/contact_en)).

## The European Commission's science and knowledge service

Joint Research Centre

### JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.



**EU Science Hub**

[ec.europa.eu/jrc](https://ec.europa.eu/jrc)



@EU\_ScienceHub



EU Science Hub - Joint Research Centre



Joint Research Centre



EU Science Hub



Publications Office

doi:10.2760/38697

ISBN 978-92-79-99639-9