Tree Cover Mapping Based on Sentinel-2 Images Demonstrate High Thematic Accuracy in Europe

Thor-Bjørn Ottosen[a,*], Geoffrey Petch[a], Mary Hanson[a], Carsten Skjøth[a]

[a] School of Science and The Environment, University of Worcester, Worcester, UK

*Corresponding author, E-mail address: t.ottosen@surrey.ac.uk (T.-B. Ottosen)

Abstract

The spatial and temporal distribution of trees has a large impact on human health and the environment through contributions to important climate mechanisms as well as commercial, recreational and social activities in society. A range of tree mapping methodologies has been presented in the literature, but tree cover estimates still differ widely between the individual datasets, and comparisons of the thematic accuracy of the resulting tree maps are rather scarce. The Copernicus Sentinel-2 satellites, which were launched in 2015 and 2017, have a combination of high spatial and temporal resolution. Given that this is a new satellite, a substantial amount of research on development of tree mapping algorithms as well as accuracy assessment of said algorithms have to be done in the years to come. To contribute to this process, a tree map produced through unsupervised classification was created for six Sentinel-2 tiles. The agreement between the tree map and the corresponding national forest inventory, as a function of the band combination chosen, was analysed and the thematic accuracy was assessed for two out of the six tiles. The results show that the highest agreement between the present tree map and the national forest inventory was found for bands 2, 3, 6 and 12. The present tree map has a relative difference in tree cover between 8% and 79% compared to previous estimates, but results are characterised by large scatter. Lastly, it is shown that the overall thematic accuracy of the present map is up to 90%, with the user's accuracy ranging from 34.85 % to 92.10 %, and the producer's accuracy ranging from 23.80 % to 97.60 % for the various thematic classes. This demonstrates that tree maps with high thematic accuracy can be produced from Sentinel-2. In the future the thematic accuracy can be increased even more through the use of temporal averaging in the mapping procedure, which will enable an accurate estimate of the European tree cover.

Keywords

Sentinel-2, band combinations, tree cover, unsupervised classification, Copernicus

Highlights

- A tree map was created using unsupervised classification for six Sentinel-2 tiles.
- The combination with highest agreement with the NFI was the bands 2, 3, 6 and 12.
- Good qualitative agreement between the present map and the NFI.
- The map shows between 8% and 79% greater tree cover compared to previous estimates.
- The overall accuracy of the present map was assessed to be up to 90 %.

1. Introduction

Trees serve as a major carbon pool contributing to important feedback mechanisms to the earth's climate (Bonan, 2008). Likewise, trees are known to release gases, such as Biogenic Volatile Organic Compounds (BVOC) (Kesselmeier and Staudt, 1999), and Primary Biological Aerosols (PBA), such as pollen (Pauling et al., 2012) and fungal spores (Sadys et al., 2014), to the atmosphere. Repeatedly, it has been demonstrated that location and abundance of trees are important in relation to the release of VOC (Arneth et al., 2011; van Meeningen et al., 2016) and their contribution towards production of secondary organic aerosols (Oderbolz et al., 2013; Tchepel et al., 2014) or PBA (Hernandez-Ceballos et al., 2011; Pauling et al., 2012). Furthermore, the spatial and temporal distribution of trees is known to be important for commercial, recreational and social activities in society (FAO, 2015) as well as the ecological or biodiversity functionality of the landscape (e.g. Ren et al., 2013; Schindler et al., 2013). It is thus evident that the spatial distribution of trees and changes in the spatial distribution of trees over time has a large impact on human health and the environment.

A range of tree mapping methodologies has been presented in the literature. Focus in this section will be on mapping of trees in the United Kingdom (UK) due to the scarcity of map comparisons for other countries. Skjøth et al. (2015) assessed the accuracy of Corine Land Cover (Bossard et al., 1994) and Globcover (Bicheron et al.,

2009) against the National Forest Inventory (NFI) over the UK (Forestry Commission, 2001). Despite reported high thematic accuracy for Corine Land Cover (Büttner and Maucha, 2006; Caetano et al., 2006) and Globcover (Defourny et al., 2009), large biases were found in these compared to the national dataset. Similar map comparison exercises have been carried out at European scale (Seebach et al., 2011a; Seebach et al., 2011b). In old cultural landscapes, like the UK, many trees are located in smaller patches, such as hedgerows, or in urban areas (McInnes et al., 2017). A remote sensing approach was used in Kempeneers et al. (2011), who mapped European scale tree cover as presence/absence, and Hansen et al. (2013), who mapped global tree cover as a percentage, both with a spatial resolution of 25 m – 30 m. An estimate by the authors shows a relative difference of 26% in the area of the total UK tree cover between Hansen et al. (2013), Kempeneers et al. (2011) and Forestry Commission (2011), which indicates a considerable uncertainty relating to the total tree cover in the UK. Moreover, Hansen et al. (2013) does not distinguish between broadleaved and coniferous trees, a minimum requirement in a range of scientific applications such as air quality modelling (Oderbolz et al., 2013; Steinbrecher et al., 2009), dynamic vegetation modelling (Hickler et al., 2012), the generation of national forest inventories (Paivinen et al., 2009) and modelling of climate change including future woodland changes (e.g. Jones et al., 2009). As seen, there is a general lack of consensus on mapping methodologies (Hansen and Loveland, 2012) and comparisons of the thematic accuracy of the resulting tree maps are rather scarce.

The Copernicus Sentinel-2 satellites, which were launched in 2015 and 2017, have four bands with a spatial resolution of 10 m and a total band combination of 13 bands, with spatial resolutions ranging from 10 m to 60 m, specifically designed for vegetation monitoring (Drusch et al., 2012). The high spatial, temporal and radiometric resolution of data from this satellite should enable the creation of tree cover maps with a higher thematic accuracy than previously achieved, and recent examples include Grabska et al. (2019) and Korhonen et al. (2017). Given that this is a new satellite, a substantial amount of research on development of tree mapping algorithms as well as accuracy assessment of said algorithms have to be done in the years to come. To contribute to this process, a tree map for six selected Sentinel-2 tiles was created, the optimal choice of spectral bands as input to the map was analysed, and the accuracy of this map was assessed.

Section 2.1 describes the creation of the tree map using unsupervised classification. The data included in the respective analyses is specified in section 2.2, section 2.3 describes the analysis of the optimal choice of spectral bands and section 2.4 details how the accuracy assessment was performed. The study area is described in section 2.5 as a foundation for the discussion of the map accuracy. The results and discussion are presented in section 3 and the conclusion in section 4.

2. Methods

2.1. Tree cover mapping

In order to support monitoring processes, reduce the cost of the development and increase the production speed, the mapping methodology should proceed without analyst interference (Hansen and Loveland, 2012). The tree map in the present study was therefore created using an unsupervised classification approach. The tree mapping algorithm consisted of a number of steps:

1. Removing pixels with cloud cover, defect pixels, no-data pixels, and saturated pixels using the accompanying masks for the individual Sentinel scene.
2. Resampling all bands to 10 m x 10 m using nearest neighbour interpolation.
3. Normalizing the bands using mean centring and division by the standard deviation to remove effects of different scale of reflectance in the images obtained in the different bands following the approach of e.g. Nguyen et al. (2018). Tests showed that the accuracy of the mapping procedure increased considerably through adding this step.
4. Classification of the satellite image using unsupervised *k*-means classification within R. The *k*-means algorithm in R was very time consuming on the 5GB tiles from Sentinel-2. The approach was therefore improved numerically by using Intel® Data Analytics Acceleration Library (DAAL) (https://software.intel.com/en-us/intel-daal) linked directly within R. The unsupervised classification was performed with 25 classes, based on the authors experience with similar classification exercises, and with a maximum of 20 iterations to limit calculation time. This number is higher than previous identified optimum number of 12 classes in specific Landsat scenes (Yıldırım, 2014) and therefore ensures sufficient

number of classes without compromising quality. Sensitivity tests showed that the mapping algorithm was not particularly sensitive to these choices.

5. The classified image was filtered to remove non-vegetation pixels by calculating the Normalized Difference Vegetation Index (NDVI) (Tucker, 1979) for the entire image by using band number 4 (red) and 8 (Near-Infrared) and setting a lower threshold. The lower NDVI threshold for vegetation was found by analysing the distribution of NDVI values in the image that was mapped as forest in Corine Land Cover. The assumption was that pixels with an NDVI less than the median minus the distance from the median to the 95-percentile of the distribution were non-vegetation pixels (e.g. buildings, roads or lakes found in forests), that according to the definition can be expected to be present in forested areas identified by the Corine Land Cover. The removal of these pixels also has the effect of removing clouds, shadows and other artefacts not included in the accompanying mask files.
6. The classes belonging to respectively coniferous and broadleaved trees were labelled using the forests classes from Corine Land Cover as training data but with error pixels and non-vegetation removed (step 1 and step 5). Broadleaved and conifer forests in Corine Land Cover can contain up to 25% of other land cover types. Moreover, Corine Land Cover has a minimum mapping unit of 25 ha (Bossard et al., 1994). These two properties introduce noise in the training data. To circumvent this problem, an iterative procedure, with the aim of finding the dominating classes, from the classification performed in step 4, for respectively broadleaved and coniferous forests, was developed:
    a. Within each Sentinel-2 scene, the polygons for respectively coniferous forest and broadleaved forest from Corine Land Cover were sorted in descending order as a function of the area. The iterations proceeded from the largest polygons to the smallest based on the assumption that the uncertainty was largest on the smallest polygons in Corine Land Cover, an assumption that was confirmed during the algorithm development phase.
    b. The largest polygon for broadleaved and coniferous trees was then masked out from the classified image (step 4) after the filtering (step 5) and the proportion of pixels in the respective

classes defined by the *k*-means algorithm was calculated separately for broadleaved trees and coniferous trees.

c. This procedure was then repeated for the second largest polygon for respectively broadleaved and coniferous, and the pixels from the new polygon added to the distribution created in step b).

d. Convergence was checked by comparing the percentage change in each class in the distribution between iterations, and convergence was reached when the largest change in a class was less than 1%. If convergence was not reached step c) was repeated with the next largest polygon until convergence. As the polygons are getting smaller and smaller, convergence will eventually be achieved in this way.

e. All 25 classes from the *k*-means algorithm applied on the entire image and extracted within the Corine Land Cover forest areas without non-vegetation classes were then labelled as either *mostly broadleaved* or *mostly coniferous* trees based on which category had the largest proportion of the selected class.

f. Subsequently, for respectively the *mostly broadleaved* classes and the *mostly coniferous* classes, a *k*-means clustering was applied to divide the distribution into two classes: Dominating and non-dominating. This resulted in a subset of the 25 classes where the forest type could be identified.

g. The dominating classes were then labelled as either broadleaved or coniferous forest. The remaining classes were labelled non-forest and the separation of the 25 classes into three categories was applied on the entire image.

In this way, a tree map was created without analyst interference. To test the sensitivity of the method to the use of Corine Land Cover as training data, the tile 30UWC from 19.07.2016 was classified using Globcover (Bicheron et al., 2009) as training and the results compared with the result using Corine Land Cover. The details of replacing Corine Land Cover with Globcover and the results are described in Appendix A.

2.2. Data

Six Sentinel-2 single tile images were downloaded as L1C data from United States Geological Survey (USGS) earthexplorer (https://earthexplorer.usgs.gov/) for the creation of a tree map. The level 1C processing includes radiometric and geometric correction using ground control points and a digital elevation model to correct for parallax error (Drusch et al., 2012). L1C data provide top of atmosphere reflectances, and thus no further preprocessing was applied to the images. The tiles were selected to cover the summer period (June-August) 2016 and to have as small a cloud cover as possible. The tile 30UWC was selected since it covers Worcester, UK an area familiar to the authors. Two images of this tile were downloaded to elucidate seasonal differences. The tile 30VUH was selected to cover an area of Scotland, which has a much larger fraction of coniferous trees compared with most of England and thus provides a different type of landscape to the analysis. The tile 32VNH covers an area in western Denmark and the tile 33VUC covers an area in eastern Denmark and southern Sweden. These were selected since high resolution tree cover maps (www.kortforsyningen.dk, www.lantmateriet.se) used by national forest inventories (Nord-Larsen et al., 2016) were available for these countries for the testing of band combinations (described in section 2.3) and because the areas are familiar to the authors. The available data is here available as final classified data sets delivered in the form of shape files, where the central input data for providing the tree cover maps in all regions are based on a combination of high-resolution aerial photography and administrative records combined with sites visits all with a spatial accuracy much higher than the 10m resolution provided by Sentinel-2.

The tile 30TWN was selected as a blind test of the forest mapping methodology in Southern Europe, since tile 30UWC was used during the development of the algorithm. The tile covers an area in Northern Spain selected to both have a large urban fraction and substantial tree cover, to allow the accuracy assessment using Google Earth. The algorithm was applied to one image at a time, to better analyse the performance of the algorithm, to keep the data and calculation requirements small for the present study, and to limit the study scope. Future work should aim at analysing the impact of the input data on the accuracy of this algorithm as well as related

algorithms. Sentinel-2 provides a new opportunity for methods development in land cover analysis by providing a large number of images over the same area taken within a short time span. This enables new possibilities for land cover analysis and the associated error assessment by taking into account multiple images within the area of interest. Such improvements are likely to remove the occasional errors caused by outliers in the data set, thereby increasing the accuracy of the final map. Red, green, blue images of the Sentinel-2 scenes can be seen in Figure 1 in the supplementary material, the location of the individual tiles can be seen in Figure 2 in the supplementary material and the properties of the individual tiles are summarized in Table 1.

Table 1 Properties of the Sentinel-2 tiles used in the present study.

Tile code:	Location:	Date:	Cloud Cover (%):	Solar Zenith Angle (°)
30UWC	Worcester, UK	19.07.2016	0.03	32.5
30UWC	Worcester, UK	15.08.2016	0.61	39.7
30VUH	Scotland	24.08.2016	6.48	46.1
32VNH	West Denmark	24.07.2016	4.89	37.9
33VUC	East Denmark and Southern Sweden	24.07.2016	1.62	37.4
30TWN	Spain	16.07.2016	2.53	25.9

2.3. Testing band combinations

To determine whether all 13 bands from the Sentinel-2 satellite were needed in the algorithm described in section 2.1, or whether some bands made the classification more noisy, the algorithm was run for all band combinations of 3 to 13 bands. To avoid subjective assessments of which bands to include and which to leave out, all 13 bands were included in this part of the analysis. This was done for the five Northern European images due to the availability of recent high resolution tree cover maps as described in section 2.2. This summed to a total of 8100 combinations. For each classification the wall-to-wall kappa coefficient (Cohen, 1960; Congalton et al., 1983) between the national forest inventory and the tree map was calculated. The kappa coefficient is a popular

189 approach to map comparison in remote sensing (Foody, 2006), since a visual comparison of the National Forest
190 Inventories with the red, green, blue image of the corresponding satellite image showed that these also contained
191 errors. The kappa coefficient should not be used for accuracy assessment (Pontius and Millones, 2011) (the
192 details of this analysis are described in section 2.4) but can be used to assess "inter-rater agreement" (Foody et
193 al., 2013). This choice is also based on that the present study only analyses the difference in kappa coefficient for
194 the respective band combinations, which removes the risk associated with using one specific kappa coefficient.

195

196 The satellite-based tree map was filtered to remove small patches of trees before the calculation of the kappa
197 coefficient to make it comparable with the corresponding national forest inventory. This resulted in a minimum
198 mapping unit of 0.5 ha for images 30 UWC and 30VUH, 0.25 ha for 32VNH and 0.01 ha for 33VUC, since the
199 Swedish data are made in a way that does not operate with a minimum mapping unit. The kappa coefficients
200 were summed across the five images, different approaches to select an optimal (based on the images in the
201 present analysis) band combination were explored and an optimal band combination, conditional on the present
202 algorithm and input data, was chosen to produce an automated tree cover map using Sentinel-2.

203 ## 2.4. Accuracy assessment of forest map

204 The accuracy of the map resulting from the analysis described in section 2.1 was assessed at tiles 30UWC and
205 30TWN to cover both Northern and Southern Europe. No filtering was applied to the map in this part of the
206 analysis, and the minimum mapping unit is therefore 0.01 ha. The accuracy assessment needed reference data
207 which were derived from Google Earth as described in section 2.4.2. High resolution images from Google Earth
208 are available for the entire 30UWC tile and areas close to the larger cities for the 30TWN tile. Reference data
209 points therefore cover the entire tile 30UWC and within 10km of the four cities Bilbao, Vitoria, Logrono and
210 Pamplona in tile 30TWN. The accuracy assessment for both images followed the sampling design, response design
211 and analysis methodology of Stehman and Czaplewski (1998).

2.4.1. Sampling design

To test the thematic accuracy of the map, an accuracy assessment dataset was produced. To generate this dataset, 999 pixels were extracted from the image. The sampling was made using stratified (broadleaved trees, coniferous trees and no trees) random sampling (Stehman, 2009) with equal sample size for each strata , since the area covered by the no trees category naturally will be much larger than the area covered by the two forest categories for both images. This ensured 333 pixels in each strata, which exceeds the 100 pixel threshold, which according to Stehman (2001), is required to obtain a standard error of 0.05 on the overall accuracy almost regardless of the sample size.

2.4.2. Response design

Stehman and Wickham (2011) discuss the use of pixels, blocks of pixels and polygons as the spatial unit for accuracy assessment based on the recommendations in Congalton and Green (2009). They show, through a numerical example, that the effect of moving from pixels to block of pixels to polygons has a small effect on the overall accuracy of the map. It was therefore decided to stick with 10 m x 10 m pixels as the spatial unit for the accuracy assessment, an approach also used by e.g. Feng et al. (2016) and Wickham et al. (2017).

Each 10 m x 10 m pixel was assigned a primary land cover class and eventually a minor land cover class if this was present following similar approaches as Benza et al. (2016); Shubho et al. (2015); Wickham et al. (2017); Yan and Roy (2016).

The collection of reference labels was done by three interpreters within the study group. To enhance consistency among interpreters, a written guide to the classification procedure was produced and 99 points, selected using the sampling design described in section 2.4.1, for both the tile 30UWC and 30TWN, were classified by all interpreters. The interpreter did not have access to the forest map from the satellite during classification to avoid biasing the manual classification (blind interpretation). Each interpreter was supplied a Google Earth KML file containing the sample pixels for overlay on Google Earth imagery. The interpreter selected the Google Earth image with an image date as close as possible to the date of the satellite image and with good visibility and subsequently decided the most appropriate land cover category. The interpreter could select among the three

categories from the tree map plus “unclassified trees” and “unclassified” for images and points where a distinct category could not be determined. Pixels in the last two categories were subsequently excluded from the analysis and the initial number of 333 sampling points in each category thereby ensured that the total number of pixels is substantially above the minimum number of 100 according to Stehman (2001). The number of remaining pixels can be found in the Results section (Section 3.3).

2.4.3. Analysis

The reference dataset based on Google Earth was used to produce a confusion matrix for the two classified Sentinel-2 images covering three classes (broadleaved trees, coniferous trees and non-trees) and two classes (trees and no trees), by merging the two tree-classes to one. Following recommended “good practice” in accuracy assessment (Olofsson et al., 2014; Stehman and Foody, 2019), the error matrix was reported in terms of estimated area proportions $\hat{p}_{ij}$:

$$\hat{p}_{ij} = W_i \frac{n_{ij}}{n_{i+}} \quad (1)$$

Where W_i is the proportion of area mapped as class i, n_{ij} is the sample counts of pixels mapped as class i which belong to class j, and n_{i+} is the sample size from stratum i. The user’s accuracy, producer’s accuracy, overall accuracy, plus the proportion of area in each class based on the reference classification along with their corresponding standard errors were calculated using the formulas from Olofsson et al. (2014); Stehman and Foody (2019). The confusion matrix for the two-class case was made using the indicator functions described in Stehman (2014).

2.5. Study Area

Figure 1a Map of the study area in the North. Data sources: Counties, Urban areas, Geographical areas, rivers (https://www.ordnancesurvey.co.uk/business-and-government/products/strategi.html), Surface water (Corine Land Cover), Forest areas (Morton et al., 2011). The forest polygons with an area < 1.5 ha have been filtered away to increase map readability. Map is produced by the authors.

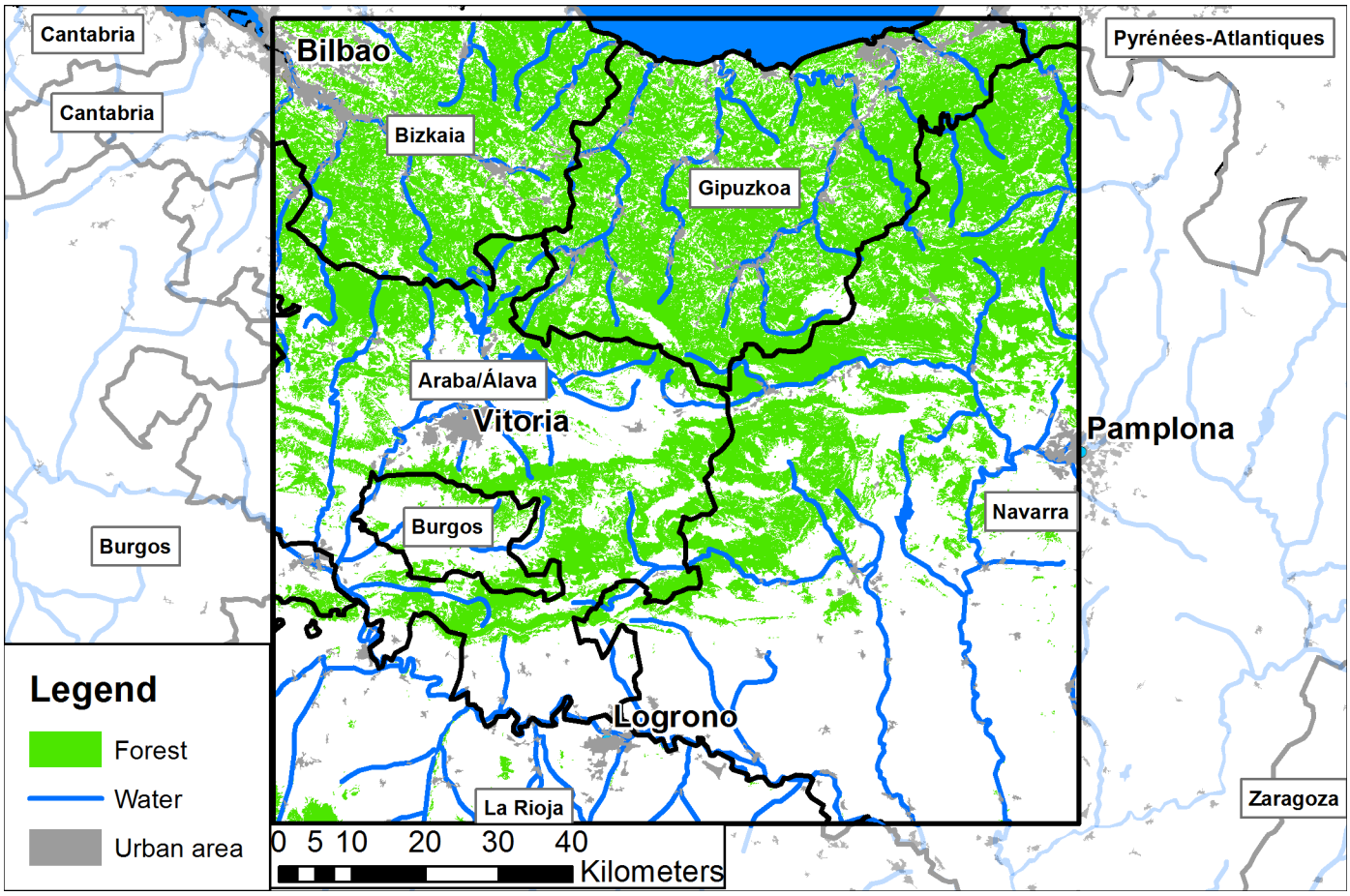


Figure 1b Map of the study area in the south. Data sources: Counties (Eurostat NUTS, https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units/nuts), Urban areas (Bossard et al., 1994) (data from Corine Land Cover 2012), rivers and surface water (Digital Chart of the World, http://www.soest.hawaii.edu/wessel/dcw/), Forest areas (Hansen et al., 2013) reclassified with forests containing more than 50% trees. The forest polygons with an area < 1.5 ha have been filtered away to increase map readability. Map is produced by the authors

The study area in the North with reference data, tile 30UWC, is centred on the city of Gloucester (Figure 1a), encompassing Gloucestershire and parts of 9 other counties located in the Midlands, England. The relief of the landscape is marked by the Severn Valley in the centre and associated tributaries with a uniform low level terrain between Gloucester and Worcester and the Bristol Channel to the Southwest (e.g. Sadys et al., 2014). No large upland areas occur within the area, but the land rises towards the Birmingham plateau in the north and towards the massifs of mid-Wales in the west. Nevertheless some prominent hills exist; the Malverns (peak height 425 metres), Bredon Hill (293 metres), the Cotswold range (up to 300 metres) and the Black Mountain (550 metres) as

seen in Figure 3 in the supplementary material. The area has one large woodland in the Forest of Dean (Forestry Commission, 2011) and numerous small woodlands and groups of trees (Forestry Commission, 2017; Skjøth et al., 2015), distributed approximately homogeneously across the area and located in both the rural and urban areas. According to Forestry Commission (2011) the area covered by forests amounts to 8.32 %. The area is dominated by privately owned woodlands, where broadleaved trees are the most abundant tree type (Forestry Commission, 2002). The broadleaved part is typically dominated by *Quercus sp*, *Fraxinus sp* and *Fagus sp*, while the coniferous part often consists of a broad range of unclassified species complemented by *Picea abies* and *Pinus Sylvestris* (Skjøth et al., 2008). The rest of the landscape covers urban areas and in particular agricultural areas used for annual crops within rotation systems and permanent pastures (Sadyś et al., 2015), but also with significant areas for fruit production (e.g. Sadys et al., 2014). The climate of the region is relatively uniform and characterized as maritime and cold temperate (UK Met Office, n.d.-a) with mild winters and warm summers, an annual mean temperature around 10 degrees, and regular rainfall throughout the year ranging from about 600 mm/year to more than 800 mm/year (e.g. Sadys et al., 2014; UK Met Office, n.d.-b).

The study area in the South with reference data, tile 30TWN, is bordered by the cities of San Sebastian, Bilbao, Logrono and Pamplona (Figure 1b). The region encompasses the three regions of Gipuzkoa, Vizcaya, and La Rioja and partly covers several other regions, located in the most Northern parts of Spain towards the Bay of Biscay. The central part of the region is covered by the Cantabrian Mountains with elevation up to 1500m (as seen in Figure 4 in the supplementary material), contrasted by the large Ebro Valley and the Ebro River in the southern part of the domain. The area has numerous larger woodlands, in particular in the mountainous part but also in lower areas to the North, while the valleys such as the Ebro Valley are mainly covered by agricultural land, therefore containing very few trees. The total tree cover of the region is, according to Hansen et al. (2013), 41.8%. The coniferous part of the woodland is dominated by various types of pinus species such as *Pinus sylvestris*, *Pinus halepensis* and *Pinus nigra*, while the broadleaved part is dominated by *Fagus sylvatica* and several *Quercus* species such *Quercus ilex*, *Quercus robur* and *Quercus faginea* (Skjøth et al., 2008). The climate of the region varies substantially due to the large variations in elevation and is, according to generalised maps for the global climate (UK Met Office, n.d.-a), in a region partly covered by temperate and partly by Mediterranean climate. This

means that it is a region where winters tend to be warm and wet while summers are dry with little or no rainfall, here considerably modified by the presence of mountains. This has the effect that the annual average rainfall in the region can be below 400mm/year or above 700mm/year and that mean annual temperatures can be higher than 15 degrees Celsius in the Ebro Valley and lower than 12 degrees in the nearby elevated terrain (e.g. Vicente-Serrano et al., 2003).

3. Results and discussion

3.1. Testing band combinations

The calculation of the wall-to-wall kappa coefficient with the corresponding national forest inventory for all 8100 band combinations shows that the highest summed kappa coefficients generally are 2.7 to 2.8, where the theoretical maximum is 5.0 and the highest kappa-coefficient is found when using a combination of four bands (Table 2). Typically, the coefficients vary from 2-2.8, where the highest abundance is in the range 2.4-2.6 as seen in Figure 2, which displays the kappa-coefficients for the images applying combinations of four bands. Similar results were obtained for band combinations of other lengths. It is evident that there is a very large scatter between the band combinations with some having very high kappa coefficients and others having very low kappa coefficients. This means, that the driver of the mapping performance with respect to identifying forests in the five examples is not the number of bands, but the choice of bands.

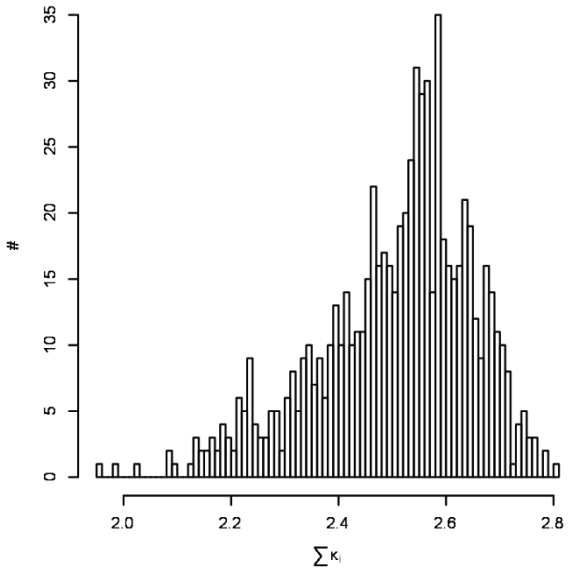


Figure 2 Histogram of sum of kappa coefficients across all five images for all band combinations with n=4.

The maximum kappa coefficients using combinations of between four and seven bands are almost equal. It is evident that bands 2, 3, 6 and 12 appear in many of the combinations. Band 2 is the blue band (496.6 nm, 10 m), band 3 is the green band (560.0 nm, 10 m), band 6 is a red-edge band (740.2 nm, 20 m), and band 12 is a short-wave infrared band (2202.4 nm, 20 m) and this combination is also the highest scoring combination of all bands (Table 2). Using USGS Spectral Characteristics Viewer (https://landsat.usgs.gov/spectral-characteristics-viewer), it can be seen that these bands are particularly suitable to separate different types of vegetation. It is natural that band 4 and band 8 will not contribute much to the classification, since these two bands are already included in the analysis through the NDVI-filter. Columns two and three in Table 2 show that the difference between the individual combinations' performance in each image is larger than the difference between the performances of the individual combination, which indicates that the highest agreement is achieved by a different band combination for each of the Sentinel images. This result is also seen in Figure 2, where up to 35 band combinations have a performance differing by less than 1%. This makes it difficult to choose the optimum band combination.

Table 2 Combinations with the highest summed κ as a function of number of bands (#). κ_i is the kappa coefficient for image i. The maximum value of $\sum \kappa_i$ is 5.000 (1.000 for each of the five images). Columns 2 and 3 are respectively the minimum and maximum difference in κ

between the best performing combination across all five images and the best performing combination for the individual image for the same n.

#	$\sum \kappa_i$	$\text{Min}(\kappa_{\max} - \kappa)$	$\text{Max}(\kappa_{\max} - \kappa)$	Combination:
3	2.791	0.012	0.068	2, 5, 6
4	2.803	0.010	0.042	2, 3, 6, 12
5	2.799	0.014	0.042	3, 5, 6, 11, 12
6	2.797	0.006	0.050	2, 3, 4, 5, 6, 11
7	2.800	0.016	0.041	1, 3, 4, 5, 6, 11, 12
8	2.784	0.011	0.055	2, 5, 6, 7, 8a, 9, 11, 12,
9	2.787	0.005	0.052	3, 4, 5, 6, 7, 8a 9, 11, 12
10	2.740	0.008	0.062	1, 3, 4, 5, 6, 7, 8a 9, 11, 12
11	2.738	0.012	0.071	1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 12
12	2.734	0.019	0.103	1, 2, 3, 4, 5, 6, 8, 8a 9, 10, 11, 12
13	2.680	0.021	0.072	1, 2, 3, 4, 5, 6, 7, 8, 8a, 9, 10, 11, 12

As a way to overcome this problem, the band combinations within 5% of the best performing combination for each image was selected, and the band combinations appearing in the top 5% for all the images are tabulated in Table 3. It is again evident, that the bands 2, 3, 6 and 12 appear in many of the combinations.

Table 3 Combinations appearing in the top 5% of each image

$\sum \kappa_i$	Combination
2.755	1, 2, 3, 4, 5, 7, 9, 12
2.803	2, 3, 6, 12
2.774	1, 3, 5, 6, 12
2.771	1, 2, 3, 5, 6, 12

2.800	1, 3, 4, 5, 6, 11, 12
2.797	2, 4, 5, 6, 12
2.799	3, 5, 6, 11, 12
2.797	2, 3, 4, 5, 6, 11
2.757	3, 4, 5, 6, 7, 9, 11, 12

As can be seen from Table 3, the best performing combination with four bands, which is also the best performing combination for the entire dataset, is among the top 5% combinations for each image. It is therefore selected for the following accuracy assessment. Given that this analysis shows a negligible small difference in agreement with the NFIs between the selected band combination and a large number of other band combinations, this choice of band combination must be considered a provisional result. Nevertheless, the results clearly illustrate that using all available bands in Sentinel-2 for this type of land cover analysis does not provide the best results. Future work should therefore aim at arriving at a more definitive answer to the question of choice of bands e.g. through performing this analysis on a larger and more variable set of Sentinel-2 tiles.

3.2. Tree mapping

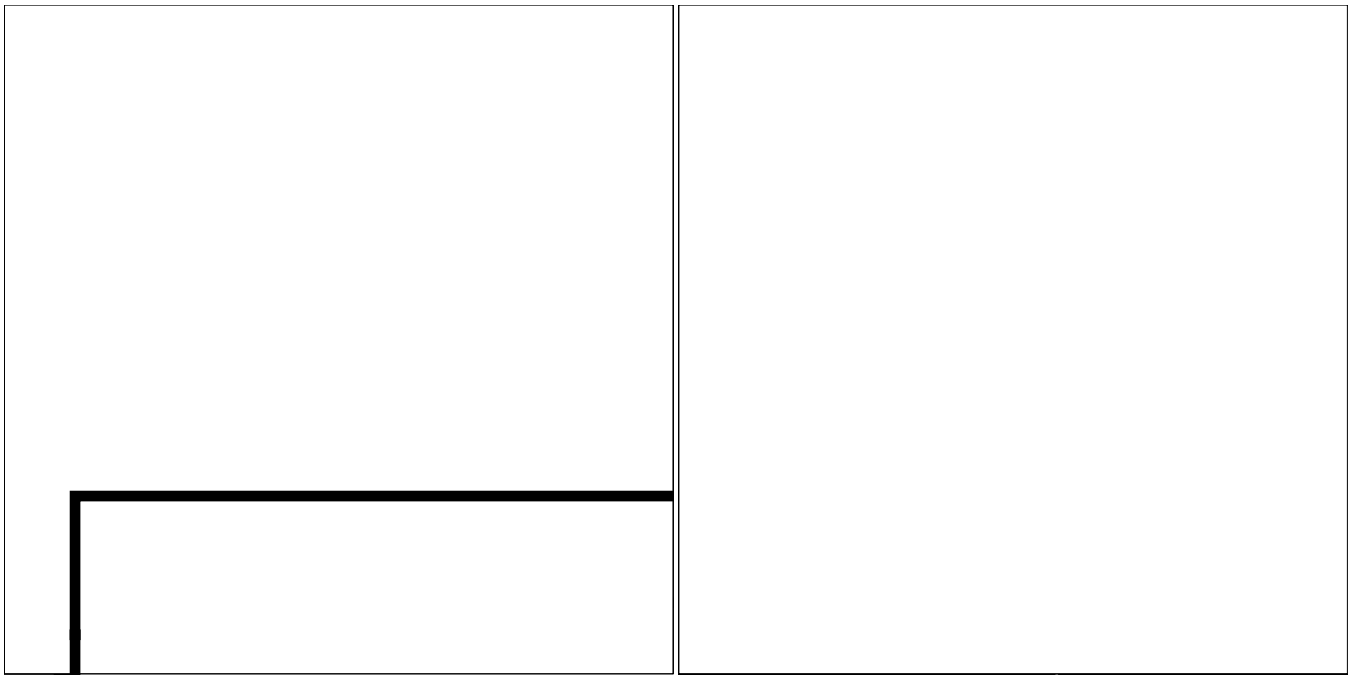
Maps of the tree cover or broadleaved trees or coniferous trees based on the satellite or the respective national forest inventory are shown in Figure 3 a-n. A visual inspection of the raw data reveals a number of interesting features. For tile 33VUC there is generally good agreement between the NFI and the satellite derived map in areas with a high forest density. On the righthand side of Figure 3b there are a number of white areas caused by clouds in the satellite image. In the lower left part of the figure and centrally in the picture, the satellite derived map predicts more trees than the NFI. These areas include, according to Corine Land Cover, large amounts of urban residential areas (Corine Land Cover code 112) and sport & leisure facilities (Corine Land Cover code 142), where the latter has actually vast areas covered by summer houses. Local knowledge by the authors established the fact that in particular the summer house areas contain large amounts of trees. However, from a land cover perspective these areas are not forests and do therefore not appear in either the national forest inventories or

land cover data sets like Corine Land Cover. Nevertheless, these areas contribute substantially to the tree cover in these regions. A secondary effect in this region is minor woodlands and hedges found throughout the part of the region that is designated as agricultural landscape found in both Denmark and Sweden. In this case these minor woodlands are not found in the national forest inventory or the Corine Land Cover. It is therefore a potential source of error, if the Corine classes are used as training element as the classes are known to be neither spectrally pure or unique (e.g. Pekkarinen et al. (2009)). We have here solved that issue by two steps: 1) filtering the Corine classes by removing the part of the area that causes the problems with spectral confusion and 2) by using only the most pure fraction of the entire data set as training elements. Some of the difference can also be explained by the algorithm confusing the spectral signal from trees and other types of land cover (e.g. green fields). However, this effect is considered to be of minor importance compared to the very large tree cover found in urban areas, residential areas (summer houses) and the agricultural landscape.

The tile 30UWC (Figure 3c,d,e,f) generally shows good agreement between the NFI and the satellite derived map in the areas with a high woodland density. However, the satellite based map shows a higher amount of broadleaved trees in areas with low tree density (Figure 3c and e) and a slightly smaller amount of conifers in the picture from 19.07.2016 (Figure 3f) compared to the picture from 15.08.2016 (Figure 3h), where the last picture has the closest resemblance to the national forest inventory (Figure 3d). This suggests that it will be an advantage to take several scenes into account over the same area if the purpose is to create very accurate inventories by using Sentinel-2 images. The higher amount of broadleaved trees could be related to orchards as the area is well known for its cider production. Orchards are technically considered a part of the agricultural landscape and therefore not included in the NFI. However, using remote sensing they will be identified either as grassland (the underlying vegetation) or as tree cover – depending on the density of the fruit trees. In any case, this type of vegetation contributes to the overall tree cover. The upper left corners of Figure 3g and h are missing data in the satellite image.

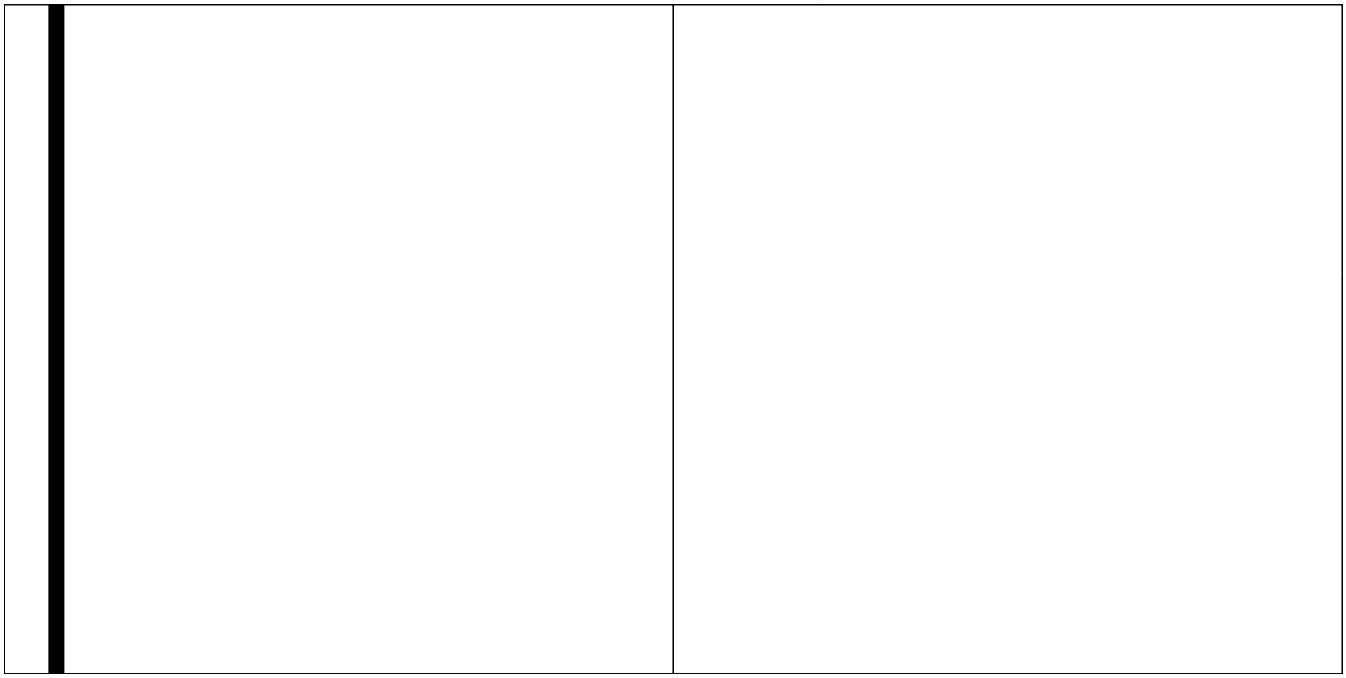
For tile 30VUH the coniferous forest compares reasonably well with the NFI whereas the broadleaved forest overestimates the tree cover. This is likely related to the spectral signature of these trees not being significantly different from their surroundings.

For tile 32VNH there is good qualitative agreement between the satellite-derived map and the NFI. However, the satellite derived image shows regions with somewhat higher amount of trees compared to the NFI. According to Corine Land Cover this region also contains substantial areas of urban land cover (in particular cities of Aarhus, Silkeborg, Randers and Horsens) and these areas have in the remote sensing picture been classified with substantial woodland cover whereas the NFI does not include those regions. As such, part of the difference is actual trees not included in the NFI whereas another part of the difference is spectral confusion between trees and green fields/grass.



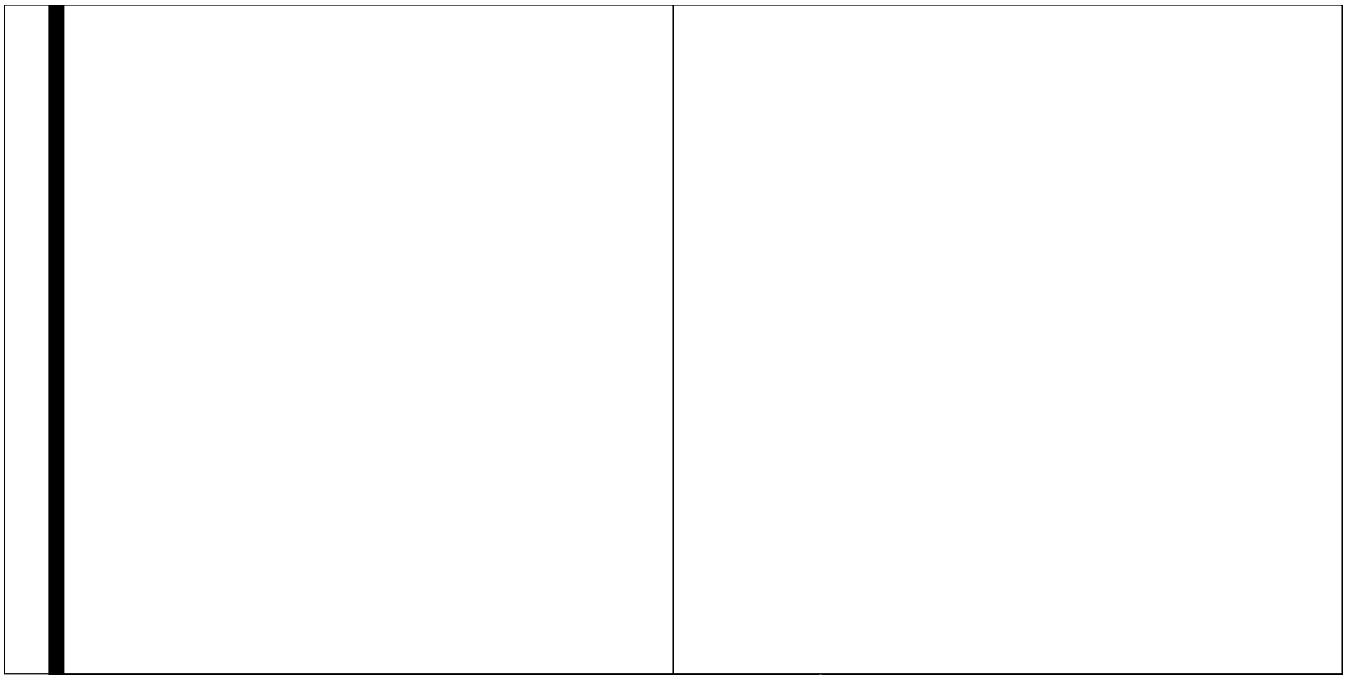
a) National forest inventory for 33VUC

b) Satellite based forest map for 33VUC



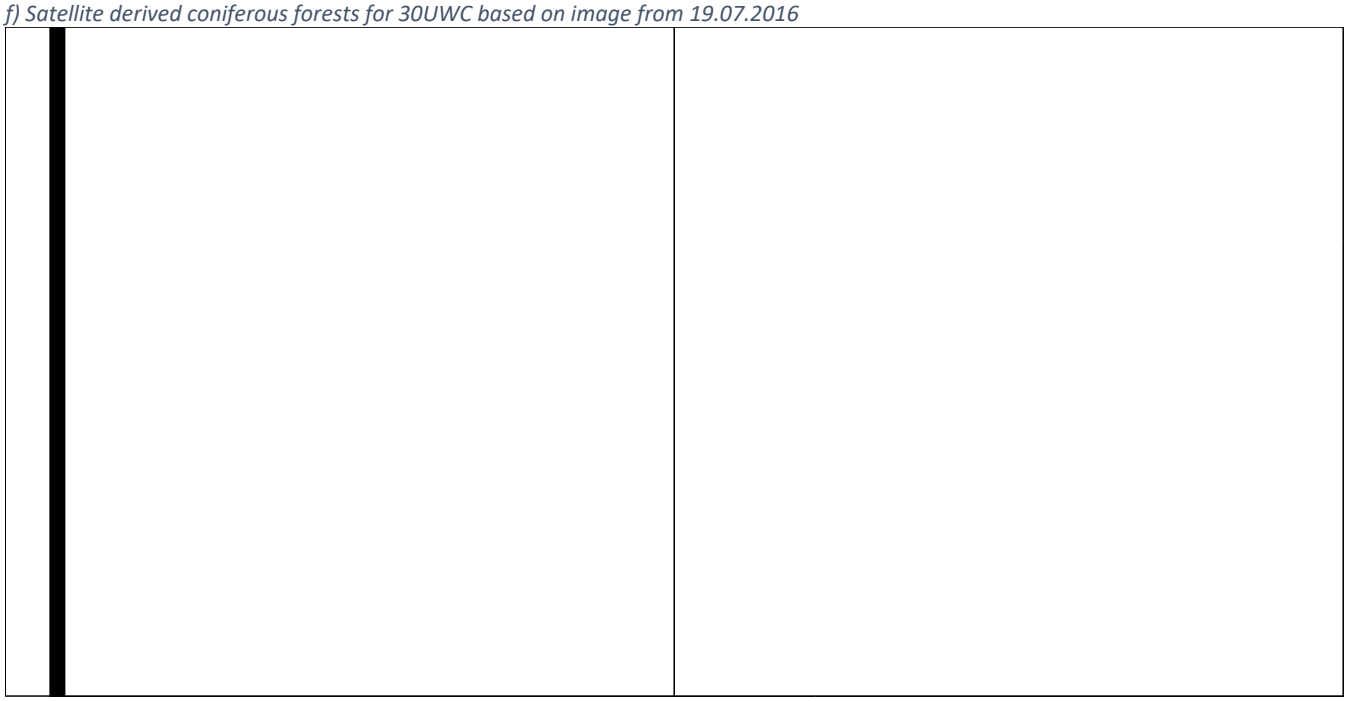
c) Broadleaved forests in NFI for 30UWC

d) Coniferous forests in NFI for 30 UWC



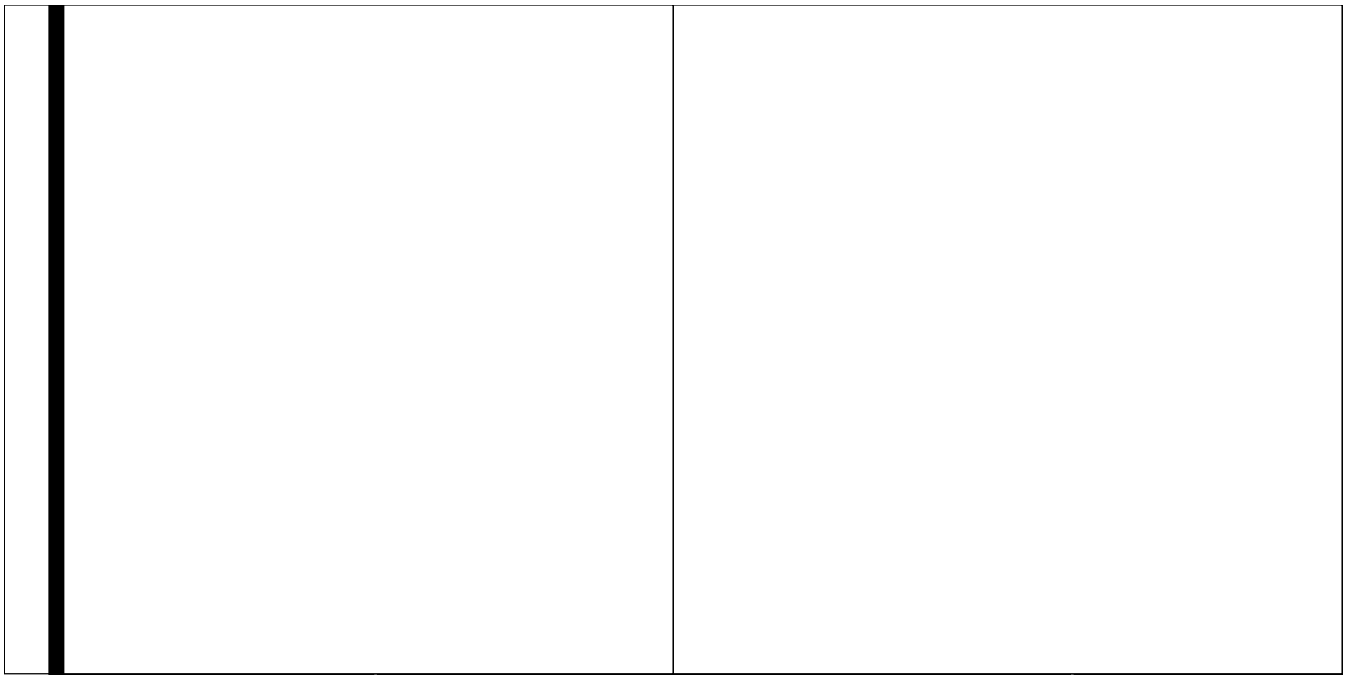
e) Satellite derived broadleaved forests for 30UWC based on image from 19.07.2016

f) Satellite derived coniferous forests for 30UWC based on image from 19.07.2016



g) Satellite derived broadleaved forests for 30UWC based on image from 15.08.2016

h) Satellite derived coniferous forests for 30UWC based on image from 15.08.2016

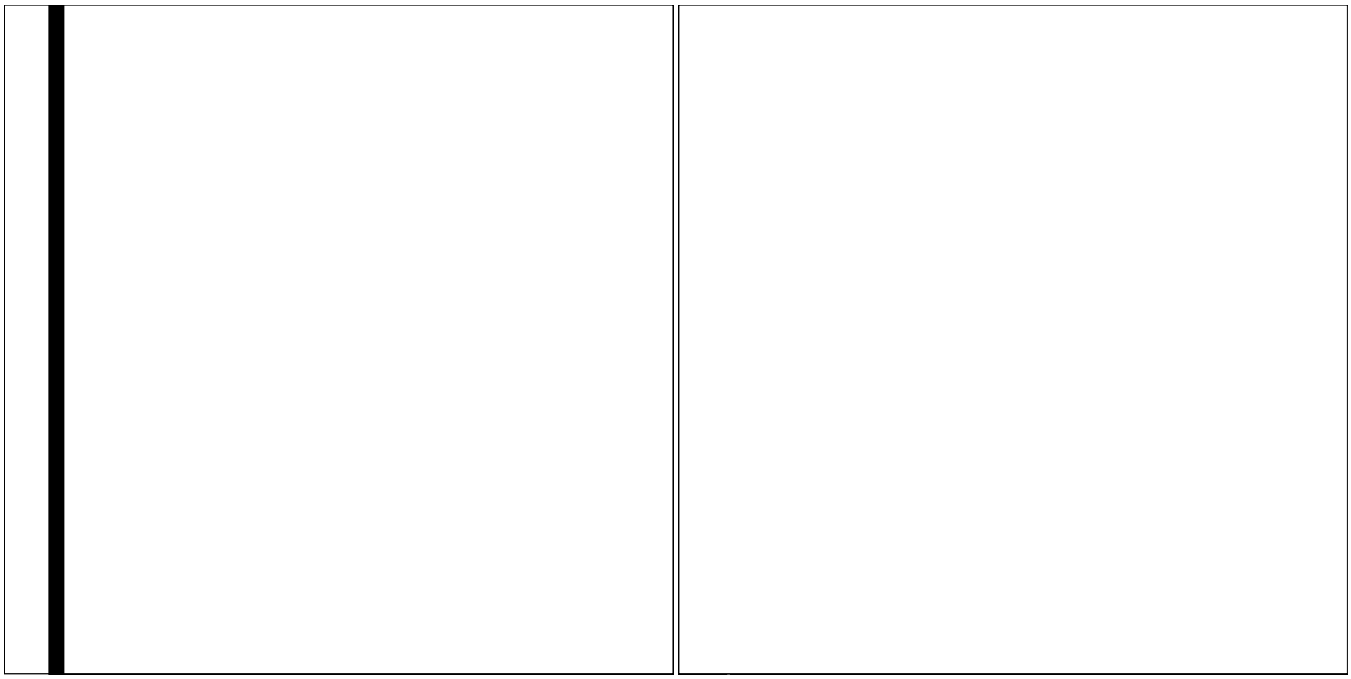


i) Broadleaved forests in NFI for 30VUH

j) Coniferous forests in NFI for 30 VUH

k) Satellite derived map of broadleaved forest for 30VUH.

l) Satellite derived map of coniferous forest for 30VUH.



m) National forest inventory for 32VNH *n) Satellite derived forest cover for 32VNH.*

Figure 3 Maps of the fractional cover of respectively the forest/non-forest or broadleaved or coniferous forest either as a satellite based map or based on the corresponding national forest inventory. All maps use the same legend as Figure 3a, and all maps are aggregated to 500 m x 500 m. The sea is marked with blue.

The summary statistics for the present study and the respective NFI plus the statistics for the studies by Kempeneers et al. (2011) and Hansen et al. (2013) are presented in Table 4 to Table 7. Tile 33VUC is the only image where the present study yields a smaller total tree cover compared with the other datasets. As described above, this is caused by cloudy areas in the image, not included in the accompanying cloud mask, but removed by the NDVI-filter of the present algorithm. As seen from section 2.1, the present algorithm does not distinguish between clouds and non-forest pixels. The reason is that this distinction is complicated and thus beyond the scope of the present study (see Deng et al. (2019); Li et al. (2019); Sui et al. (2019) for some recent examples). The present approach is designed for cloud free or almost cloud free images and is designed with computational efficiency in mind. Besides that, incorporating multiple images over the same area in a subsequent study is expected, to some extent, to alleviate this problem. Despite this bias, the total tree cover is quite close to the other estimates, which indicates that the remaining areas have a larger tree coverage than previously thought.

Part of this tree cover is technically not accounted for in the NFI as the land use is either agricultural (e.g. orchards), urban (e.g. low density residential) or recreational (summer cottages).

For the remaining images, the relative difference between the tree cover area of the present study and the previous studies is between 8 % and 79 %. As can also be seen, there is a large variation between the previous estimates of the tree cover for the respective image, which contributes to the large variation in the relative difference between the previous estimates and the present study. For tile 33UWC the image dated 15.08.2016 always has a smaller tree cover compared with the image 15.07.2016 due to the smaller area covered by the satellite on this particular day. As can be seen, even though the total area does not change much, the distribution between broadleaved and coniferous trees changes significantly – a result also seen in Figure 3 e – h. The image from 19.07.2016 has a larger cover of broadleaved trees and a smaller cover of coniferous trees compared with the national forest inventory and vice versa for the image from 15.08.2016. The correct result is probably somewhere between the two estimates, which underlines that temporal averaging or other approaches that utilize several images in order to create accurate tree maps would yield a higher accuracy.

For tile 33 VUH the cover of coniferous trees is in reasonably agreement with the national forest inventory, whereas the cover of broadleaved trees is much larger. A visual inspection of the image reveals that the area of broadleaved trees in Corine Land Cover in this image is much smaller than the area of coniferous trees in Corine Land Cover. This means that there is a larger probability that clouds and other artefacts can influence the training data and thus introduce noise in the labelling procedure. This is also seen in that approximately six times as many pixels are used for the labelling of conifer trees compared to the labelling of broadleaved trees. Future work should aim at reducing this effect.

For 32VNH the present algorithm also finds a considerably larger tree cover compared with the previous studies, but given the large variation in the previous estimates, it is difficult to conclude on the validity of this estimate. However, it is known that this particular area contains a substantial amount of land cover that technically is not part of the national forest inventories (e.g. urban land) and that the tree density in these areas requires at least 10 m spatial resolution in order to be accurately mapped (Uuemaa et al., 2013). This suggests that the true tree cover in those regions likely to be better mapped with the Sentinel-2 satellite compared to previous estimates.

Table 4 Summary statistics for tile 33VUC.

Tile:	33VUC		
Dataset:	Broadleaved Trees (km^2):	Coniferous Trees (km^2):	Total tree cover (km^2):
Present study	866	1309	2175
Kempeneers et al., (2011)	1100	1422	2522
National forest inventory	-	-	2676
Hansen et al., (2013)	-	-	2334

Table 5 Summary statistics for tile 30UWC

Tile:	30UWC					
Dataset:	Broadleaved Trees (km^2):		Coniferous Trees (km^2):		Total tree cover (km^2):	
Date:	19.07.2016	15.08.2016	19.07.2016	15.08.2016	19.07.2016	15.08.2016
Present study	1047	666	101	320	1148	986
Kempeneers et al., (2011)	806	737	103	86	909	823
National forest inventory	796	738	208	175	1004	913

Hansen et al., (2013)	-	-	868	783

458

459 *Table 6 Summary statistics for tile 30VUH*

Tile:	30VUH		
Dataset:	Broadleaved Trees (km^2):	Coniferous Trees (km^2):	Total tree cover (km^2):
Present study	1330	875	2204
Kempeneers et al., (2011)	94	1194	1288
National forest inventory	357	1078	1435
Hansen et al., (2013)	-	-	1488

460

461 *Table 7 Summary statistics for tile 32VNH*

Tile:	32VNH		
Dataset:	Broadleaved Trees (km^2):	Coniferous Trees (km^2):	Total tree cover (km^2):
Present study	1892	496	2387
Kempeneers et al., (2011)	237	798	1035
National forest inventory	-	-	1673
Hansen et al., (2013)	-	-	1149

462

463 3.3. Accuracy assessment

464 The results of the accuracy assessment for the primary land cover class for tile 30UWC can be found in Table 8 for
465 three categories and in Table 10 for two categories. As can be seen it was not possible to manually classify 58
466 pixels. Since this corresponds to approximately 5% of the data this is not assumed to influence the results, which
467 can also be seen on the low standard errors on all the accuracies. The overall accuracy for three classes is 90%
468 with a standard error of 1.35 %. Comparing this to the commonly used accept criterion of 85%, the accuracy of

the map is high, even though this acceptance criterion has been questioned (Foody, 2006). The producer's and user's accuracies are quite low for coniferous trees, even though the area comparisons are quite close to each other. It is natural that the accuracy of the map will be better between the no trees category and the two tree categories, compared to between the two tree categories due to the spectral similarity between different types of trees. The significant fraction mapped as broadleaved trees being no trees is most likely green fields which have a similar spectral signature. A part of the pixels mapped as coniferous trees being broadleaved trees is due to shadows e.g. at forest roads or forest edges, where the shadows cause the trees to appear darker and thus fall in the coniferous category. Future work should aim at reducing this effect. When classifying the map in two classes (trees/no trees) the percent correctly classified is 90 % again being above the accept criterion. It is noteworthy, that even though the present map has reported substantially more tree cover compared to previous maps, as shown in section 3.2, the accuracy assessment indicates that the tree cover based on the reference data is actually substantially higher – especially for broadleaved trees. The actual tree cover might therefore be substantially higher.

Table 8 Accuracy assessment for tile 30UWC in percentages of area. The table includes the user's accuracy (User) and the producer's accuracy (Prod), standard errors are presented in parentheses along with the number of pixels in each category (n). Estimated overall accuracy is 89.97 % with a standard error of 1.35 %.

	Map					
Reference	No trees	Broadleaved	Coniferous	Total	Prod (SE)	n
No trees	82.89	1.98	0.06	84.92 (1.87)	97.60(0.25)	388
Broadleaved	6.29	6.73	0.59	16.61 (1.21)	49.42 (4.65)	433
Coniferous	0.82	0.30	0.35	1.46 (0.45)	23.80 (8.15)	120
Total	90.00	9.00	1.00	100.00		
User (SE)	92.10 (1.49)	74.75 (2.49)	34.85 (2.72)			
N	329	305	307			941

Table 9 Accuracy assessment for tile 30TWN in percentages of area. The table includes the user's accuracy (User) and the producer's accuracy (Prod) , standard errors are presented in parentheses along with the number of pixels in each category (n). Estimated overall accuracy is 83.43% with a standard error of 1.64%.

	Map					
Reference	No trees	Broadleaved	Coniferous	Total:	Prod:	n
No trees	76.11	0.80	0.39	77.29 (2.23)	98.47 (0.34)	317
Broadleaved	10.13	4.40	2.20	16.72 (1.58)	36.31 (2.66)	418
Coniferous	2.76	0.27	2.92	5.95 (0.86)	49.03 (7.69)	191
Total	89.00	5.47	5.50	100.00		
User	85.52 (2.07)	80.50 (2.21)	53.04 (2.83)			
N	290	323	313			926

Table 10 Accuracy assessment for tile 30UWC in percentages of area for two categories. The table includes the user's accuracy (User) and the producer's accuracy (Prod), standard errors are presented in parentheses along with the number of pixels in each category (n). Estimated overall accuracy is 90.43% with a standard error of 1.38%.

	Map				
Reference	No trees	Trees	Total:	Prod:	n
No trees	82.39	1.96	84.35 (1.89)	97.68 (0.18)	388
Trees	7.61	8.04	16.65 (1.32)	51.37 (4.02)	579
Total	90.00	10.00	100.00		
User	91.54 (1.53)	80.41 (1.57)			
N	331	636			967

Table 11 Accuracy assessment for tile 30TWN in percentages of area for two categories. The table includes the user's accuracy (User) and the producer's accuracy (Prod), standard errors are presented in parentheses along with the number of pixels in each category (n). Estimated overall accuracy is 85.43% with a standard error of 1.87%.

	Map				
Reference	No trees	Trees	Total:	Prod:	n
No trees	75.59	1.16	76.74 (2.24)	98.49 (0.30)	317
Trees	13.41	9.84	23.26 (1.78)	42.33 (1.97)	632
Total	89.00	11.00	100.00		
User	84.93 (2.09)	89.49 (1.20)			
N	292	657			949

The results of the accuracy assessment for three land cover classes for tile 30TWN are shown in Table 9 and for two classes in Table 11. The overall accuracy is 83.43 % with a standard error of 1.64 % for three classes, and as such, a little bit lower than for tile 30UWC. This is expected, since tile 30UWC has been part of the development process for the algorithm. In particular the separation between broadleaved and coniferous trees is better for this tile. The accuracy for the no trees category is slightly lower compared to tile 30UWC. A part of this can be explained by 41 pixels where the manual interpreter could not determine whether the pixel was showing orchards of young trees or fruit bushes, where the tree map has classified it as no trees. This is one of the explanations why the number of unclassified pixels is slightly higher for this image, and the standard error therefore slightly larger. As stated in section 2.4.1, pixels for this analysis were only sampled within a 10 km radius of the four cities in the image. This means that the actual accuracy for the entire image is likely to be higher, since the land cover will be more homogeneous in the rural areas. Reducing the number of classes to two (trees/no trees) gives an overall accuracy of 85.43 % with a standard error of 1.87 %. The phenomenon that the area of tree cover estimated from the reference data is substantially higher than the mapped area is likewise found for this tile.

4. Conclusion

Tree maps with high thematic accuracy can be produced from Sentinel-2. The high spatial resolution of this satellite means that a larger tree cover is generally found compared with previous estimates (on average 36%), for the five Sentinel-2 tiles in the present study and in particular a large tree cover is found in regions officially classified as urban landscapes. The performance of the present map compared to the respective national forest inventory does not depend on the number of bands included in the analysis, but on the choice of bands, with the band combination 2, 3, 6 and 12 as the best performing combination in the present study. Likewise, the difference in performance for the individual band combination is larger for the different images compared with between the band combinations. With a few exceptions, the present tree map agrees well with the corresponding national forest inventory, and add to this the non NFI tree resource. This non NFI resource can in some regions be substantial. The thematic accuracy, for the two tiles where accuracy assessment was performed, was above the commonly applied 85% threshold for three land cover classes (non-forest, broadleaved trees, and coniferous trees) at a resolution of 10 m × 10 m.

Acknowledgements

The authors acknowledge financial support from the project “New approaches for the early detection of tree health pests and pathogens” funded by BBSRC (ProjectID: BB/L012286/1).

Declarations of interest: none.

Appendix A Using Globcover as training data

To test the sensitivity of the map to the use of Corine Land Cover as training data for the labelling procedure and to test the applicability of the procedure outside Europe, the tile 30UWC was classified using Globcover as training data. Tile 30UWC was selected since the thematic accuracy of the produced tree map was known for this tile. The coniferous trees was represented by class 70 *closed needle leaved evergreen forest*. A natural choice for broadleaved trees would be class 50 *closed broadleaved deciduous forest* or class 40 c*losed to open broadleaved evergreen or semi-deciduous forest.* However, tile 30UWC contains hardly any pixels classified as class 50 and

none classified as class 40. Instead, class 110 *Mosaic Forest-Shrubland/Grassland* was used for broadleaved forest, knowing well that this will introduce considerable noise in the training data.

The classification algorithm was run as described in section 2.1 with the exception that the distribution for the NDVI filter (step 5) was only based on the polygons classified as class 70 in Globcover, since class 110 contained significant non-forest elements.

A visual inspection of the resulting map shows good agreement between the two maps, with the map based on Globcover resulting in 639 km^2 broadleaved trees and 236 km^2 conifer trees. The map using Globcover as training is thus probably reasonably accurate with respect to coniferous trees, but underestimates the number of broadleaved trees. These numbers compare well with the figures in Table 5, which is further underlined by the wall-to-wall kappa coefficient between the map based on Globcover and the map based on Corine Land Cover being 0.58. In this way, the applicability of Globcover as training data is shown.

References

Arneth, A., Schurgers, G., Lathiere, J., Duhl, T., Beerling, D.J., Hewitt, C.N., Martin, M., Guenther, A., 2011. Global terrestrial isoprene emission models: sensitivity to variability in climate and vegetation. Atmospheric Chemistry and Physics 11, 8037-8052.

Benza, M., Weeks, J.R., Stow, D.A., López-Carr, D., Clarke, K.C., 2016. A pattern-based definition of urban context using remote sensing and GIS. Remote Sens. Environ. 183, 250-264.

Bicheron, P., Defourny, P., Brockmann, C., Schouten, L., Vancutsem, C., Huc, M., Bontemps, S., Leroy, M., Achard, F., Herold, M., Ranera, F., Arino, O., 2009. GLOBCOVER 2009 Products Description and Validation Report.

Bonan, G.B., 2008. Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests. Sci 320, 1444 LP-1449.

Bossard, M., Feranec, J., Otahel, J., 1994. CORINE Land Cover. Technical Guide.

Büttner, G., Maucha, G., 2006. The thematic accuracy of Corine land cover 2000. Assessment using LUCAS (land use/cover area frame statistical survey). European Environment Agency.

Caetano, M., Mata, F., Freire, S., 2006. Accuracy assessment of the Portuguese CORINE Land Cover map, in: Marcal, A. (Ed.), Proceedings of the 25th EARSeL Symposium. Millpress, Porto, Portugal, pp. 459-467.

Cohen, J., 1960. A Coefficient of Agreement for Nominal Scales. Educ. Psychol. Meas. 20, 37-46.

Congalton, R.G., Green, K., 2009. Assessing the accuracy of remotely sensed data : principles and practices. CRC Press/Taylor & Francis.

Congalton, R.G., Oderwald, R.G., Mead, R.A., 1983. Assessing Landsat Classification Accuracy Using Discrete Multivariate-Analysis Statistical Techniques. Photogramm. Eng. Remote Sensing 49, 1671-1678.

Defourny, P., Schouten, L., Bartalev, S., Bontemps, S., Caccetta, P., Bella, C.D., Gond, V., Hazeu, G.W., Heinimann, A., Herold, M., Knoops, J., Jaffrain, G., Latifovic, R., Lin, H., Nonguierma, A., Bogaert, E.V., Vancutsem, C., Bicheron, P., Leroy, M., Arino, O., 2009. Accuracy Assessment of a 300 m Global Land Cover Map: The GlobCover Experience. 33rd International Symposium on Remote Sensing of Environment, Sustaining the Millennium Development Goals, 1-5.

572 Deng, C.W., Li, Z., Wang, W.Z., Wang, S.G., Tang, L.B., Bovik, A.C., 2019. Cloud Detection in Satellite Images Based
573 on Natural Scene Statistics and Gabor Features. Ieee Geoscience and Remote Sensing Letters 16, 608-612.
574 Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P.,
575 Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., Bargellini, P., 2012. Sentinel-2: ESA's Optical High-
576 Resolution Mission for GMES Operational Services. Remote Sens. Environ. 120, 25-36.
577 FAO, 2015. Global Forest Resources Assessment 2015. Food and Agricultural Organization of the United Nations.
578 Feng, M., Sexton, J.O., Huang, C., Anand, A., Channan, S., Song, X.-P., Song, D.-X., Kim, D.-H., Noojipady, P.,
579 Townshend, J.R., 2016. Earth science data records of global forest cover and change: Assessment of accuracy in
580 1990, 2000, and 2005 epochs. Remote Sens. Environ. 184, 73-85.
581 Foody, G.M., 2006. What is the difference between two maps? A remote senser's view. J. Geograph. Systems 8,
582 119-130.
583 Foody, G.M., See, L., Fritz, S., Van der Velde, M., Perger, C., Schill, C., Boyd, D.S., 2013. Assessing the Accuracy of
584 Volunteered Geographic Information arising from Multiple Contributors to an Internet Based Collaborative
585 Project. Transactions in GIS 17, 847-860.
586 Forestry Commission, 2001. National Inventory of Woodland and Trees. Forestry Commission.
587 Forestry Commission, 2002. National Inventory of Woodland and Trees - Regional Report for West Midlands.
588 Forestry Commission.
589 Forestry Commission, 2011. NFI 2011 woodland map GB. Forestry Commission.
590 Forestry Commission, 2017. Tree cover outside woodland in Great Britain. National Forest Inventory.
591 Grabska, E., Hostert, P., Pflugmacher, D., Ostapowicz, K., 2019. Forest Stand Species Mapping Using the Sentinel-2
592 Time Series. Remote Sens. 11, 1197.
593 Hansen, M.C., Loveland, T.R., 2012. A review of large area monitoring of land cover change using Landsat data.
594 Remote Sens. Environ. 122, 66-74.
595 Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V.,
596 Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-
597 Resolution Global Maps of 21st-Century Forest Cover Change. Sci 342, 850 LP-853.
598 Hernandez-Ceballos, M.A., Garcia-Mozo, H., Adame, J.A., Dominguez-Vilches, E., Bolivar, J.P., De la Morena, B.A.,
599 Perez-Badia, R., Galan, C., 2011. Determination of potential sources of Quercus airborne pollen in Codoba city
600 (southern Spain) using back-trajectory analysis. Aerobiologia 27, 261-276.
601 Hickler, T., Vohland, K., Feehan, J., Miller, P.A., Smith, B., Costa, L., Giesecke, T., Fronzek, S., Carter, T.R., Cramer,
602 W., Kühn, I., Sykes, M.T., 2012. Projecting the future distribution of European potential natural vegetation zones
603 with a generalized, tree species-based dynamic vegetation model. Global Ecol. Biogeogr. 21, 50-63.
604 Jones, C., Lowe, J., Liddicoat, S., Betts, R., 2009. Committed terrestrial ecosystem changes due to climate change.
605 Nature Geoscience 2, 484-487.
606 Kempeneers, P., Sedano, F., Seebach, L., Strobl, P., San-Miguel-Ayanz, J.s., 2011. Data fusion of different spatial
607 resolution remote sensing images applied to forest-type mapping. ITGRS 49, 4977-4986.
608 Kesselmeier, J., Staudt, M., 1999. Biogenic Volatile Organic Compounds (VOC): An Overview on Emission,
609 Physiology and Ecology. JAtC 33, 23-88.
610 Korhonen, L., Hadi, Packalen, P., Rautiainen, M., 2017. Comparison of Sentinel-2 and Landsat 8 in the estimation
611 of boreal forest canopy cover and leaf area index. Remote Sens. Environ. 195, 259-274.
612 Li, Z.W., Shen, H.F., Cheng, Q., Liu, Y.H., You, S.C., He, Z.Y., 2019. Deep learning based cloud detection for medium
613 and high resolution remote sensing images of different sensors. ISPRS J. Photogramm. Remote Sens. 150, 197-
614 212.
615 McInnes, R.N., Hemming, D., Burgess, P., Lyndsay, D., Osborne, N.J., Skjøth, C.A., Thomas, S., Vardoulakis, S.,
616 2017. Mapping allergenic pollen vegetation in UK to study environmental exposure and human health. Sci. Total
617 Environ. 599-600, 483-499.
618 Morton, D., Rowland, C., Wood, C., Meek, L., Marston, C., Smith, G., Wadsworth, R., Simpson, I.C., 2011. Final
619 Report for LCM2007 - the new UK Land Cover Map. Centre for Ecology & Hydrology.
620 Nguyen, C., Starek, M.J., Tissot, P., Gibeaut, J., 2018. Unsupervised Clustering Method for Complexity Reduction of
621 Terrestrial Lidar Data in Marshes. Remote Sens. 10, 133.
622 Nord-Larsen, T., Johannsen, V.K., Riis-Nielsen, T., Thomsen, I.M., Suadicani, K., Vesterdal, L., Gundersen, P.,
623 Jørgensen, B.B., 2016. Skove og plantager 2015: Forest statistics 2015.

Oderbolz, D.C., Aksoyoglu, S., Keller, J., Barmpadimos, I., Steinbrecher, R., Skjøth, C.A., Plaß-Dülmer, C., Prévôt, A.S.H., 2013. A comprehensive emission inventory of biogenic volatile organic compounds in Europe: improved seasonality and land-cover. Atmospheric Chemistry and Physics 13, 1689-1712.
Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sens. Environ. 148, 42-57.
Paivinen, R., Van Brusselen, J., Schuck, A., 2009. The growing stock of European forests using remote sensing and forest inventory data. Forestry 82, 479-490.
Pauling, A., Rotach, M.W., Gehrig, R., Clot, B., 2012. A method to derive vegetation distribution maps for pollen dispersion models using birch as an example. Int. J. Biometeorol. 56, 949-958.
Pekkarinen, A., Reithmaier, L., Strobl, P., 2009. Pan-European forest/non-forest mapping with Landsat ETM+ and CORINE Land Cover 2000 data. ISPRS J. Photogramm. Remote Sens. 64, 171-183.
Pontius, R.G., Millones, M., 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. Int. J. Remote Sens. 32, 4407-4429.
Ren, Y., Wei, X., Wang, D., Luo, Y., Song, X., Wang, Y., Yang, Y., Hua, L., 2013. Linking landscape patterns with ecological functions: A case study examining the interaction between landscape heterogeneity and carbon stock of urban forests in Xiamen, China. For. Ecol. Manage. 293, 122-131.
Sadys, M., Skjøth, C.A., Kennedy, R., 2014. Back-trajectories show export of airborne fungal spores (Ganoderma sp.) from forests to agricultural and urban areas in England. Atmos. Environ. 84, 88-99.
Sadyś, M., Skjøth, C.A., Kennedy, R., 2015. Determination of Alternaria spp. habitats using 7-day volumetric spore trap, Hybrid Single Particle Lagrangian Integrated Trajectory model and geographic information system. Urban Climate 14, 429-440.
Schindler, S., Von Wehrden, H., Poirazidis, K., Wrbka, T., Kati, V., 2013. Multiscale performance of landscape metrics as indicators of species richness of plants, insects and vertebrates. Ecol. Indicators 31, 41-48.
Seebach, L.M., Strobl, P., San Miguel-Ayanz, J., Bastrup-Birk, A., 2011a. Identifying strengths and limitations of pan-European forest cover maps through spatial comparison. Int. J. Geogr. Inf. Sci. 25, 1865-1884.
Seebach, L.M., Strobl, P., San Miguel-Ayanz, J., Gallego, J., Bastrup-Birk, A., 2011b. Comparative analysis of harmonized forest area estimates for European countries. Forestry: An International Journal of Forest Research 84, 285-299.
Shubho, M.T.H., Islam, S.R., Ayon, B.D., Islam, I., 2015. An improved semiautomatic segmentation approach to land cover mapping for identification of land cover change and trend. Int. J. Environ. Sci. Technol. (Tehran) 12, 2593-2602.
Skjøth, C.A., Baker, P., Sadyś, M., Adams-Groom, B., 2015. Pollen from alder (Alnus sp.), birch (Betula sp.) and oak (Quercus sp.) in the UK originate from small woodlands. Urban Climate 14, 414-428.
Skjøth, C.A., Geels, C., Hvidberg, M., Hertel, O., Brandt, J., Frohn, L.M., Hansen, K.M., Hedegaard, G.B., Christensen, J.H., Moseholm, L., 2008. An inventory of tree species in Europe—An essential data input for air pollution modelling. Ecol. Model. 217, 292-304.
Stehman, S.V., 2001. Statistical rigor and practical utility in thematic map accuracy assessment. Photogramm. Eng. Remote Sensing 67, 727-734.
Stehman, S.V., 2009. Sampling designs for accuracy assessment of land cover. Int. J. Remote Sens. 30, 5243-5272.
Stehman, S.V., 2014. Estimating area and map accuracy for stratified random sampling when the strata are different from the map classes. Int. J. Remote Sens. 35, 4923-4939.
Stehman, S.V., Czaplewski, R.L., 1998. Design and Analysis for Thematic Map Accuracy Assessment: Fundamental Principles. Remote Sens. Environ. 64, 331-344.
Stehman, S.V., Foody, G.M., 2019. Key issues in rigorous accuracy assessment of land cover products. Remote Sens. Environ. 231, 111199.
Stehman, S.V., Wickham, J.D., 2011. Pixels, blocks of pixels, and polygons: Choosing a spatial unit for thematic accuracy assessment. Remote Sens. Environ. 115, 3044-3055.
Steinbrecher, R., Smiatek, G., Köble, R., Seufert, G., Theloke, J., Hauff, K., Ciccioli, P., Vautard, R., Curci, G., 2009. Intra- and inter-annual variability of VOC emissions from natural and semi-natural vegetation in Europe and neighbouring countries. Atmos. Environ. 43, 1380-1391.
Sui, Y.L., He, B., Fu, T.J., 2019. Energy-based cloud detection in multispectral images based on the SVM technique. Int. J. Remote Sens. 40, 5530-5543.

Tchepel, O., Ferreira, J., Martins, H., Silveira, C., Miranda, A.I., Borrego, C., 2014. Investigating the Contribution of Biogenic Emissions to the Formation of Secondary Pollutants in Portugal, in: Steyn D., R., M. (Eds.), Air Pollution Modeling and its Application XXIII. Springer, Cham, pp. 121-127.
Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ. 8, 127-150.
UK Met Office, n.d.-a. Climate - Fact sheet 16 — World climates.
UK Met Office, n.d.-b. UK Climate.
Uuemaa, E., Mander, Ü., Marja, R., 2013. Trends in the use of landscape spatial metrics as landscape indicators: A review. Ecol. Indicators 28, 100-106.
van Meeningen, Y., Schurgers, G., Rinnan, R., Holst, T., 2016. BVOC emissions from English oak, Quercus robur and European beech, Fagus sylvatica along a latitudinal gradient. Biogeosciences 13, 6067-6080.
Vicente-Serrano, S.M., Saz-Sánchez, M.A., Cuadrat, J.M., 2003. Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): application to annual precipitation and temperature. Clim. Res. 24, 161-180.
Wickham, J., Stehman, S.V., Gass, L., Dewitz, J.A., Sorenson, D.G., Granneman, B.J., Poss, R.V., Baer, L.A., 2017. Thematic accuracy assessment of the 2011 National Land Cover Database (NLCD). Remote Sens. Environ. 191, 328-341.
Yan, L., Roy, D.P., 2016. Conterminous United States crop field size quantification from multi-temporal Landsat data. Remote Sens. Environ. 172, 67-86.
Yıldırım, A., 2014. Unsupervised classification of multispectral Landsat images with multidimensional particle swarm optimization. Int. J. Remote Sens. 35, 1217-1243.

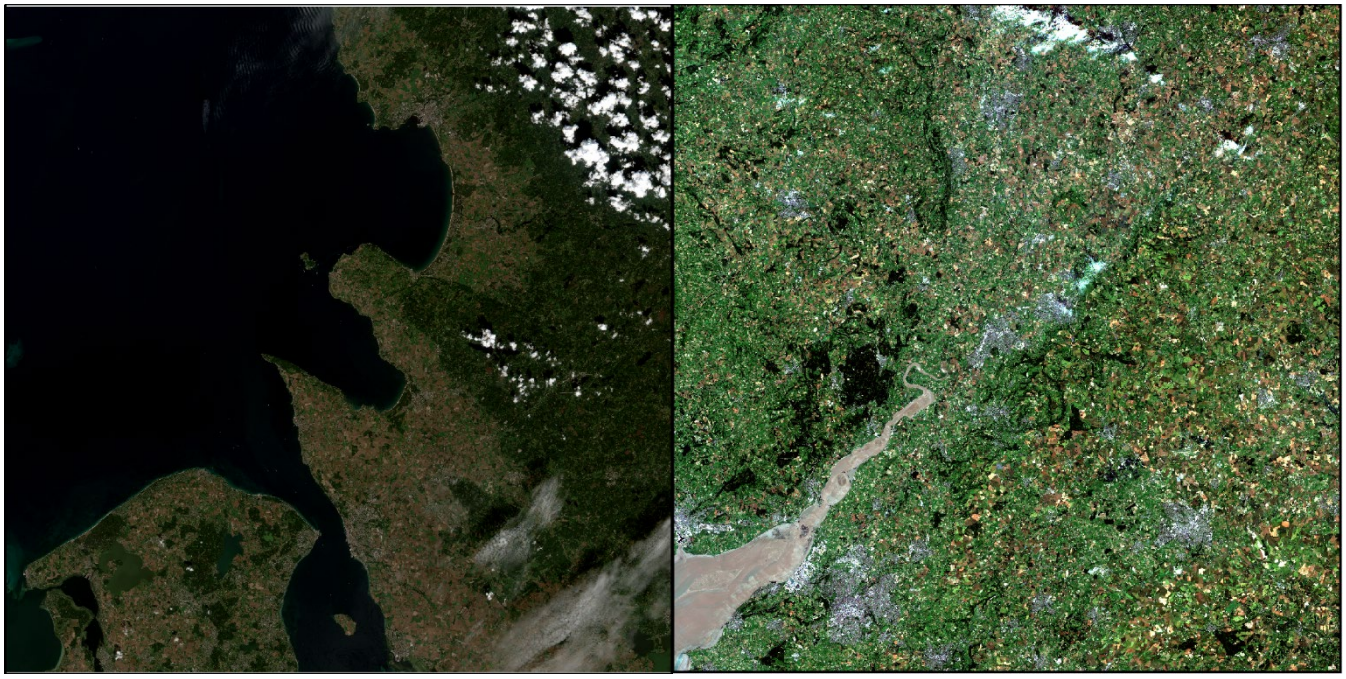
Supplementary material for Tree Cover Mapping Based on Sentinel-2 Images Demonstrate High Thematic Accuracy in Europe

Thor-Bjørn Ottosen[a,*], Geoffrey Petch[a], Mary Hanson[a], Carsten Skjøth[a]

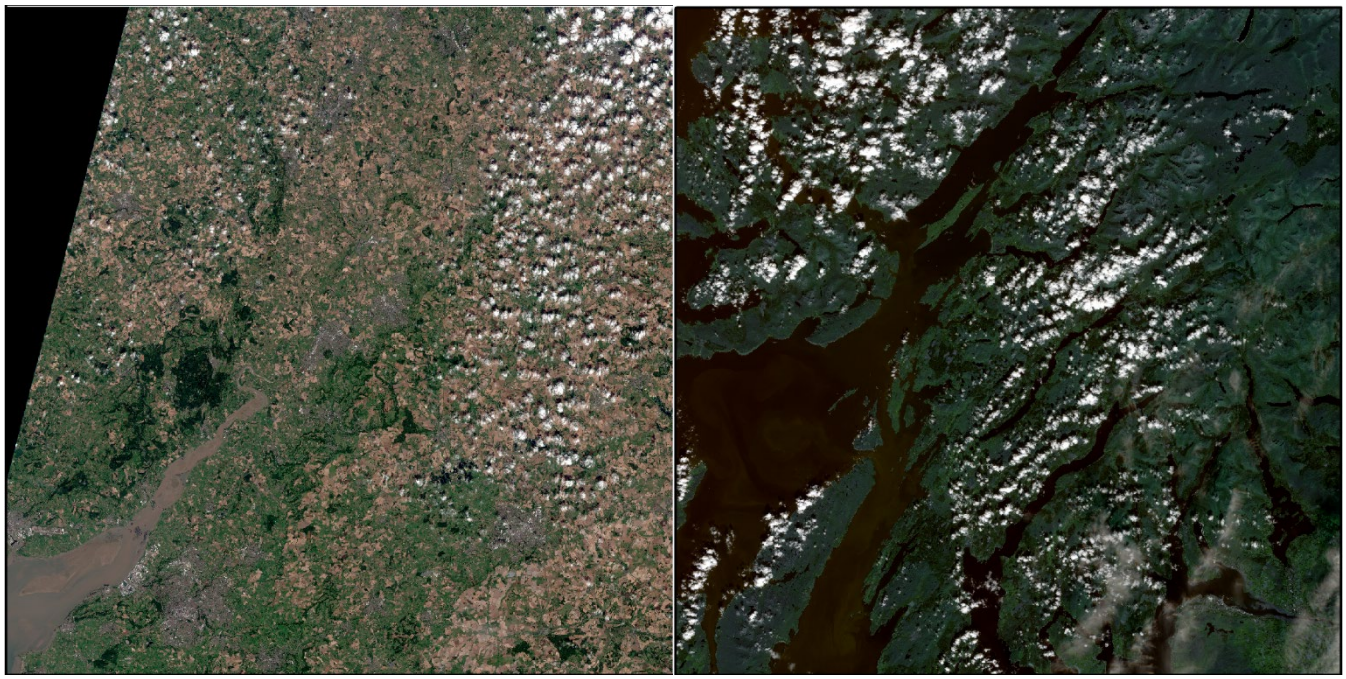
[a] School of Science and The Environment, University of Worcester, Worcester, UK

*Corresponding author, E-mail address: t.ottosen@surrey.ac.uk (T.-B. Ottosen)

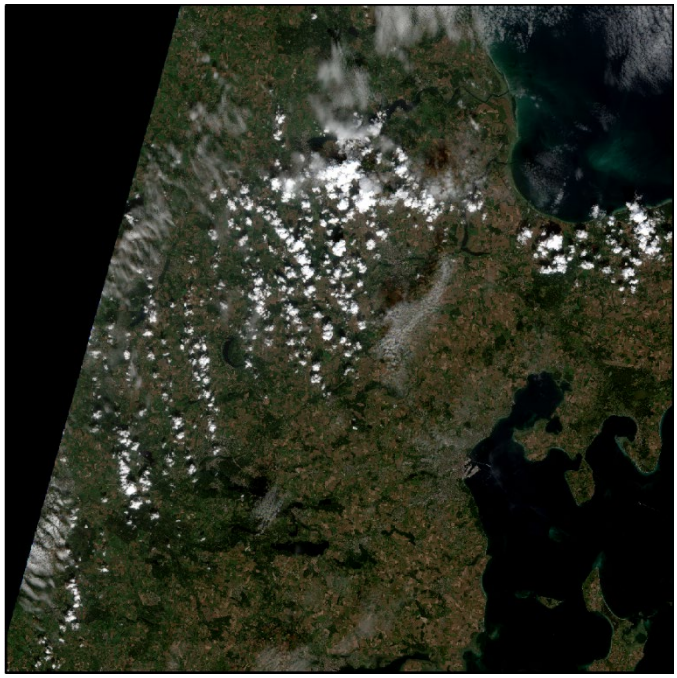
1. RGB images of the Sentinel-2 tiles



a) *RGB image of tile 33VUC* *b) RGB Image of tile 30UWC from 19.07.2016*



c) RGB image of tile 30UWC from 15.08.2016 *d) RGB image of tile 30 VUH*



e) RGB Image of tile 32VNH.

Figure 1 RGB images of the five Sentinel-2 images analysed in the present study.

2. Map of the location of the tiles

Figure 2 Location of the tiles used in the present study.

3. Elevation map of tile 30UWC

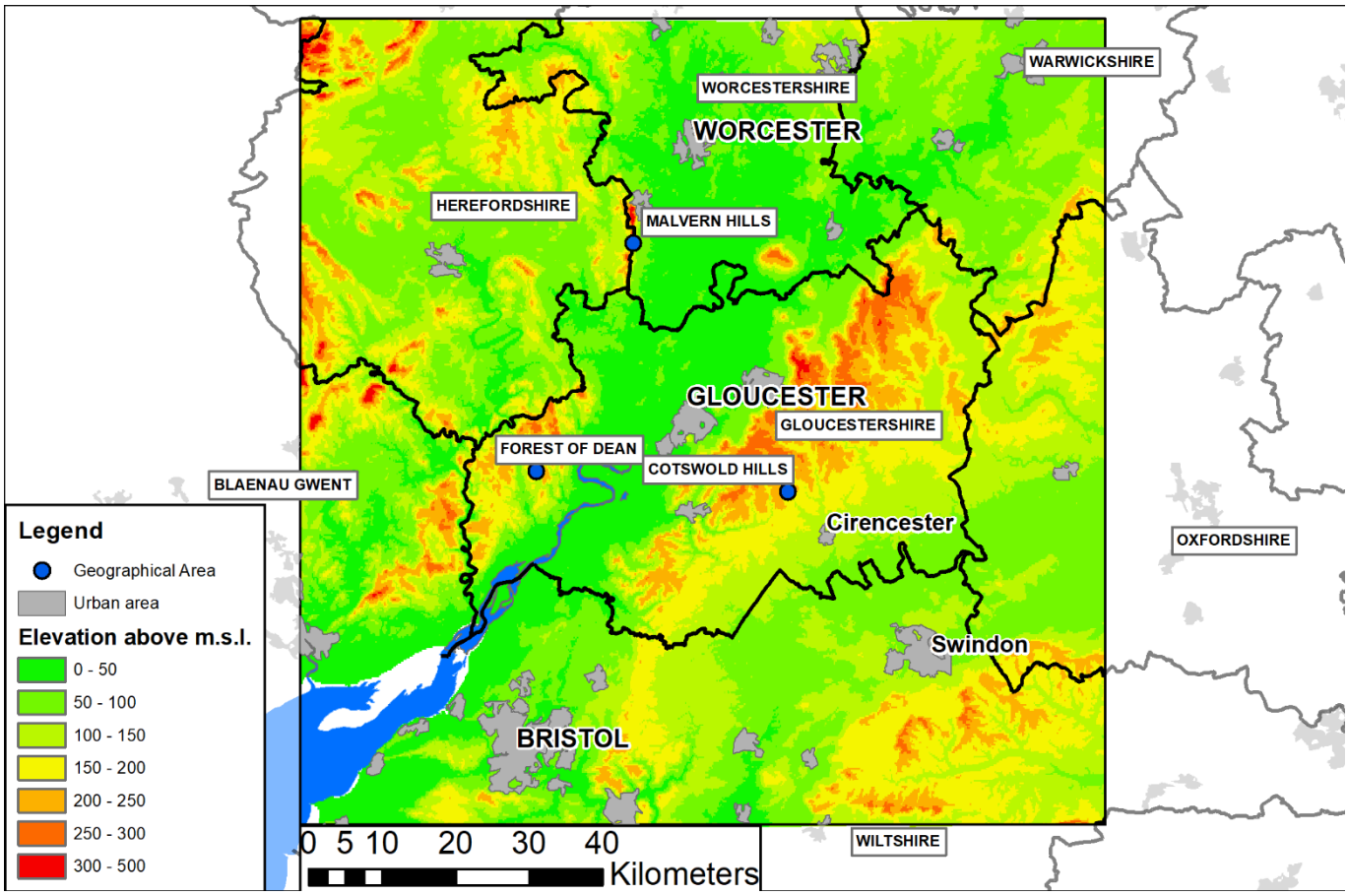


Figure 3 Elevation map of the tile 30UWC. Data sources: Counties, Urban areas, Surface water (CLC), elevation (the SRTM mission) (Reuter et al., 2007).

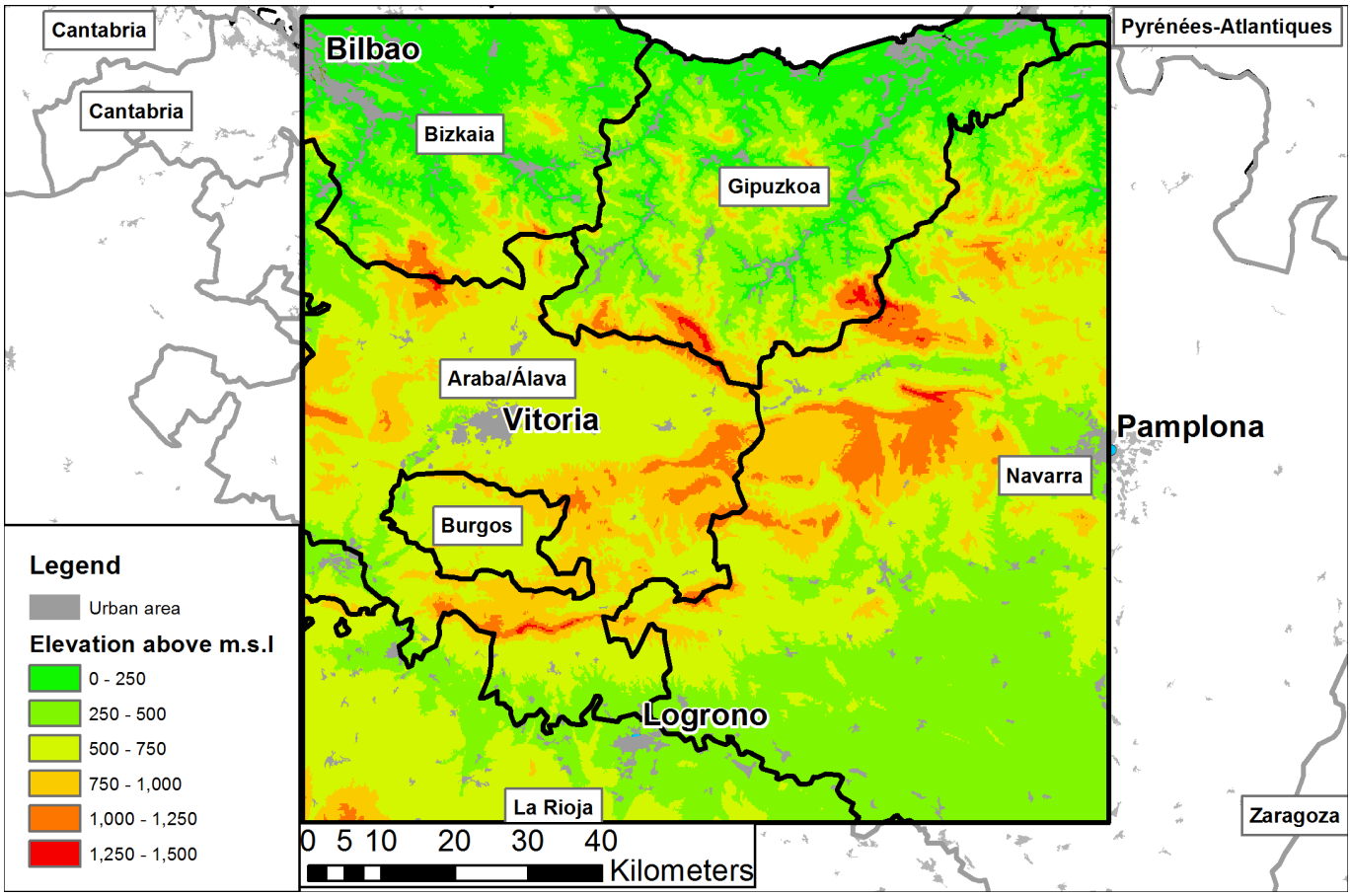


Figure 4 Elevation map of the tile 30TWN. Data sources: Counties (Eurostat NUTS, *https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units/nuts*), Urban areas (Bossard et al. 1994), *elevation (the SRTM mission)(Reuter et al. 2007).*

References

Bossard, M., Feranec, J., & Otahel, J. (1994). *CORINE Land Cover. Technical Guide*.

Reuter, H.I., Nelson, A., & Jarvis, A. (2007). An evaluation of void-filling interpolation methods for SRTM data. *International Journal of Geographical Information Science, 21*, 983-1008