






Article

Automated Production of a Land Cover/Use Map of Europe Based on Sentinel-2 Imagery

Radek Malinowski ^{1,*}, Stanisław Lewiński ¹, Marcin Rybicki ¹, Ewa Gromny ¹,
Małgorzata Jenerowicz ¹, Michał Krupiński ¹, Artur Nowakowski ¹, Cezary Wojtkowski ¹,
Marcin Krupiński ¹, Elke Krätzschar ² and Peter Schauer ²

¹ Centrum Badań Kosmicznych Polskiej Akademii Nauk (CBK PAN), Bartycka 18A, 00-716 Warszawa, Poland; stlewinowski@cbk.waw.pl (S.L.); mrybicki@cbk.waw.pl (M.R.); egromny@cbk.waw.pl (E.G.); mjenerowicz@cbk.waw.pl (M.J.); mkrupinski@cbk.waw.pl (M.K.); anowakowski@cbk.waw.pl (A.N.); cwojtkowski@cbk.waw.pl (C.W.); marcinkrupinski@cbk.waw.pl (M.K.)

² Industrieanlagen-Betriebsgesellschaft MBH (IABG), Hermann-Reichelt-Str. 3, 01109 Dresden, Germany; Kraetzschar@iabg.de (E.K.); Schauer@iabg.de (P.S.)

* Correspondence: rmalinowski@cbk.waw.pl

Received: 31 August 2020; Accepted: 24 October 2020; Published: 27 October 2020



Abstract: Up-to-date information about the Earth's surface provided by land cover maps is essential for numerous environmental and land management applications. There is, therefore, a clear need for the continuous and reliable monitoring of land cover and land cover changes. The growing availability of high resolution, regularly collected remote sensing data can support the increasing number of applications that require high spatial resolution products that are frequently updated (e.g., annually). However, large-scale operational mapping requires a highly-automated data processing workflow, which is currently lacking. To address this issue, we developed a methodology for the automated classification of multi-temporal Sentinel-2 imagery. The method uses a random forest classifier and existing land cover/use databases as the source of training samples. In order to demonstrate its operability, the method was implemented on a large part of the European continent, with CORINE Land Cover and High-Resolution Layers as training datasets. A land cover/use map for the year 2017 was produced, composed of 13 classes. An accuracy assessment, based on nearly 52,000 samples, revealed high thematic overall accuracy (86.1%) on a continental scale, and average overall accuracy of 86.5% at country level. Only low-frequency classes obtained lower accuracies and we recommend that their mapping should be improved in the future. Additional modifications to the classification legend, notably the fusion of thematically and spectrally similar vegetation classes, increased overall accuracy to 89.0%, and resulted in ten, general classes. A crucial aspect of the presented approach is that it embraces all of the most important elements of Earth observation data processing, enabling accurate and detailed (10 m spatial resolution) mapping with no manual user involvement. The presented methodology demonstrates possibility for frequent and repetitive operational production of large-scale land cover maps.

Keywords: land cover and use; CORINE land cover; High Resolution Layers (HRL); machine learning; multi-temporal; random forest; Sentinel-2

1. Introduction

Land cover mapping provides information on the Earth's surface characteristics. It is considered to be the most important variable in effective monitoring of natural and anthropogenic processes [1], as it supports environmental and socio-economic decision-making [2], and the implementation of various national and international programs and initiatives [3,4]. Continuous and reliable monitoring

of land cover/use and land cover/use changes is necessary due to ongoing climate change, human population growth and the resulting agricultural expansion. It is also helpful in monitoring natural resources, ecosystem conditions, biodiversity as well as the estimation of losses due to natural and human-induced disasters. Consequently, it has been selected as one of the terrestrial essential climate variables used to analyse the state and characteristics of the Earth's climate [4,5].

Earth Observation (EO) data are the most important source of information for land cover mapping due to the high spatial coverage, systematic monitoring and access to remote, inaccessible areas [6,7]. The first global land cover map was derived from EO data captured by the Advanced Very High Resolution Radiometer instrument flown on the National Oceanic and Atmospheric Administration's satellites (NOAA-AVHRR) [8]. Since then, imagery from many different satellites has been captured, and new image analysis techniques have been developed [6,9–13]. Land cover products have improved in terms of spatial extent (regional to global), level of detail (class range and composition, spatial resolution), update frequency, and the number of sub-products, among many other characteristics. Similarly, analysis techniques have improved, notably in terms of the level of automation, the sophistication of the method, and the amount of processed data. The present study, however, focuses on large-scale land cover mapping, and addresses research performed and tested on continental-to-global scale, with only exceptional cases of regional studies, highly relevant to the main subject.

The pioneering work of DeFries and Townshend [8] was followed by numerous projects that developed global land cover products mainly based on data collected by coarse resolution instruments including, among others, NOAA-AVHRR, Satellite Pour l'Observation de la Terre Vegetation (SPOT Vegetation), Envisat Medium Resolution Imaging Spectrometer (MERIS), and Moderate Resolution Imaging Spectroradiometer (MODIS) [14–18]. Although there were numerous applications of these products, they were characterised by significant inconsistencies and a low level of details [6,7,19]. At the same time, there was a high demand within the user community for large-area land cover datasets with higher spatial resolution, produced at higher temporal frequency [6,7]. The cost of developing such a product was, however, prohibitive until the Landsat archive was made available to all users, free of charge in 2008. This was followed by improvements in computing performance, and numerous global datasets with 30 m spatial resolution were produced including both single- and multi-class data [7,20–24]. Although Landsat data increased the spatial resolution and level of detail of newly-developed products, similar improvements in thematic accuracy have not been seen in all cases, compared to lower-resolution datasets. Currently, even higher resolution products are feasible due to the regular collection of EO data by Sentinel-2 and -1 satellites, and its provision under free and open access policies.

At the present time, there are several programs that manage the production of regularly-updated regional or continental databases [3,25–27]. Such products make it possible to directly monitor changes that appear between consecutive database releases, while maintaining the same class definitions. For example, the CORINE and Cover (CLC) database [25] is the flagship European database, with 44 land cover/use classes. The first version was produced for the reference year 1990 and have seen four updates covering the period between 1990 and 2018. However, as the workflow remains heavily dependent on manual work by dozens of national teams, the update frequency remains low. Furthermore, the CLC's minimum mapping unit (MMU) of 25 ha is another limitation to its usefulness, as many applications require a higher level of detail. Although other products, such as for instance, High Resolution Layers (HRLs) [26] and the products of the North American Land Change Monitoring System [3], are available at much higher spatial resolution (20 m and 30 m, respectively), the update frequency is low (i.e., 3–5 years). Additionally, HRLs only cover five land cover types (impervious areas, forests, grasslands, wetness and water, and small woody features), which limits the thematic information they provide. Considerable effort has also been made to produce the United State National Land Cover Database (NLCD), including updates that cover a time period between 1990 and 2016, that is similar to CLC data [28,29]. However, it should be noted that, although a comprehensive set of

products is provided, with this 30 m spatial resolution database covering the conterminous United States, its production is not fully-automated and requires access to extensive sets of reference data.

The performance of most classification techniques, especially supervised ones, relies on the quality and quantity of training samples [30]. The collection of a sample suitable for large-scale mapping is often seen as a bottleneck, and may require a tremendous amount of manual work [31,32]. For example, Gong et al. [20,33] and Li et al. [34] produced large-scale, high resolution land cover databases, however they required the enormous and usually cost-intensive manual visual image interpretation that seriously hampers repetitive and regular (e.g., annual) map production or their update [1].

There is an increasing number of applications that require high spatial resolution products and more frequent updates [3,32,33]. Given the currently available satellite missions, the annual production of a high spatial resolution land cover/use database is feasible, but requires a very high level of automation. Therefore, there is a clear need for an operational method that can be applied at local to global scales that will enable efficient, reliable, and automated processing of high resolution EO data.

One automated way to collect training data for large-scale mapping is to extract samples from existing land cover/use databases. There have already been several promising attempts to test such an approach [1,3,32,35,36]. In most cases, however, EO imagery has been classified at similar or lower spatial resolution than the reference source. High spatial resolution mapping can only be achieved by using local databases, which are necessarily spatially-limited and insufficient for large-scale continental or global mapping. It therefore seems useful to test the potential effectiveness of training samples derived from databases that have lower spatial resolution than the satellite data being classified. Such an approach has not been widely evaluated in large-scale land cover mapping.

This challenge was addressed within the Sentinel-2 Global Land Cover (S2GLC) project [37], funded by the European Space Agency. The project ran from 2016 to 2019 and aimed to develop a land cover/use classification methodology ready for implementation on a global scale. The goal was to develop a method based on satellite imagery acquired by Sentinel-2 sensors, and deliver a product with 10-m spatial resolution. Extensive analyses were carried out in test areas located in various parts of the world, including large regions of Africa, Asia, Europe, and South America. Among the various approaches that have been tested, a supervised method based on training data extracted from existing low-resolution databases was selected as the most appropriate, as it enables global land cover mapping without manual the collection of training data [38]. This paper presents the results of the operational implementation of the developed approach to a wide area of the European continent.

The overall goal of this study was to develop a highly-automated and ready for operational use, land cover/use classification methodology that could be applied on a large-scale, from regional to global. The more detailed objectives were:

- To develop a method to generate a unified, reference training dataset from existing land cover databases.
- To develop an automated classification workflow based on multi-temporal Sentinel-2 imagery, resulting in a final product with high spatial resolution (10-m) and thematic accuracy.
- To optimise data processing by using the Data and Information Access Services (DIAS) cloud-based computing environment.

A key assumption was that information contained in land cover databases used as the reference, and source satellite images were sufficient to produce reliable land cover/use mapping on a large scale. We also assumed that Sentinel-2 data with spatial resolution up to 10 m would enable a detailed analysis of heterogeneous areas, and provide more accurate estimates of class coverage [33].

2. Materials and Methods

Figure 1 presents the overall structure of our automated land cover/use classification procedure. In general, it follows a supervised remote sensing methodology. The workflow starts with the selection of

a set of Sentinel-2 images from a time series covering the whole of the year 2017. Then, training samples are selected for classification using a random forest machine learning algorithm. Next, the results of the classification of individual images are aggregated, and post-processed. Finally, the results are validated. All calculations and data processing was carried out on CREODIAS [39], one of five cloud-based DIAS computing platforms deployed by the European Commission [40].

The following sections provide details of each step in the procedure. It should be noted that, with the exception of the last step, all data analyses and processing was performed on a tile-by-tile basis.

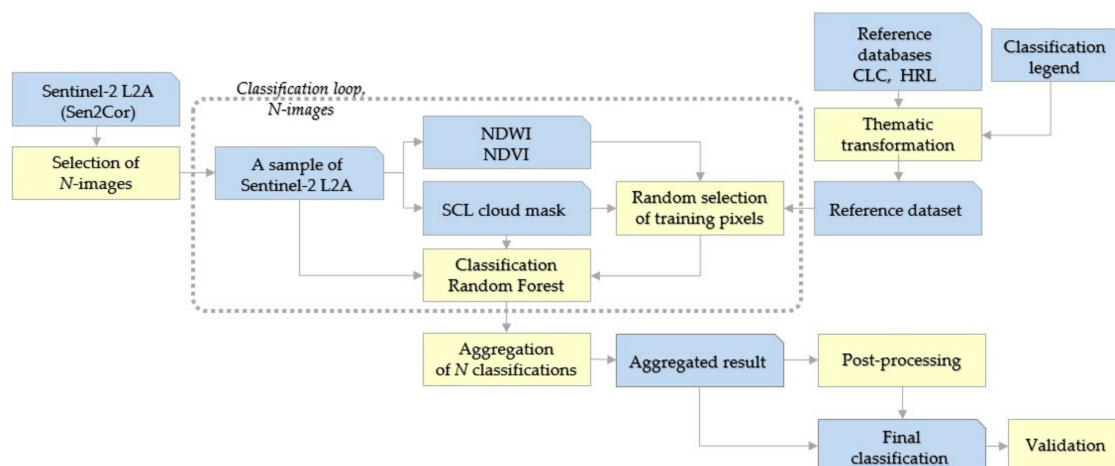


Figure 1. Land cover map production workflow for a single Sentinel-2 tile. Blue boxes indicate input data and the results of analyses, while yellow boxes represent data processing steps.

2.1. Study Area

The main study area covers 27 European Union Member States together with Liechtenstein, Norway, Switzerland, the United Kingdom, the Western Balkans and Iceland. It also encompasses the Mediterranean islands (the Greek Islands, Cyprus, Malta, Sicily, Sardinia, Corsica, the Balearic Islands), and islands surrounding the United Kingdom and Ireland (Figure 2). This geographical zone is similar to the CLC database [25], with the exception of Turkey (Figure 3). The analysed area corresponds to 791 standard Sentinel-2 tiles (granules) [41]. The resulting land cover/use map covers 34 countries with climates that vary from dry and hot (the Mediterranean, southern Europe) to cold and boreal (the far north).

The initial, development phase that preceded classification of the whole study area aimed to design and test individual parts of the classification methodology. These tests were performed on selected sites representing different European climatic regions. The division into climatic classes was based on the Köppen–Geiger climate classification system, updated by Peel et al. [42], and the stratification method proposed by Olofsson et al. [43]. Consequently, the area was divided into seven regions: boreal forest, continental forest, marine west coast, Mediterranean, steppe, temperate evergreen forest, and tundra (Figure 3). A test site was selected for each region, while the two biggest regions, i.e., continental forest and marine west coast, were represented by two sites. Each test site corresponds to a single Sentinel-2 tile, covering an area of 110×110 km.

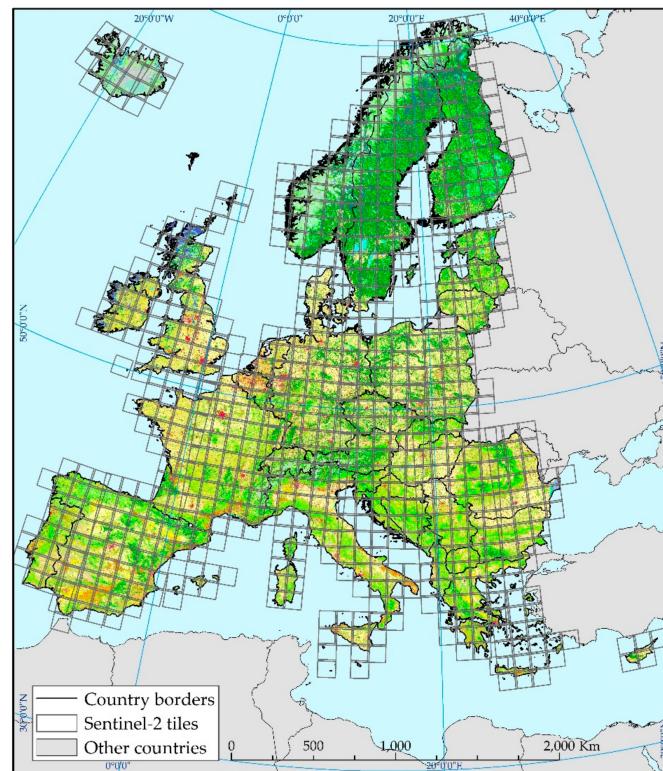


Figure 2. The study area showing boundaries of Sentinel-2 tiles. The background map presents the CORINE Land Cover (CLC) 2012 dataset.



Figure 3. Location of the nine test sites distributed over the seven climatic regions of Europe, identified by the modified Köppen–Geiger climate classification system.

2.2. Input Data

2.2.1. Sentinel-2 Imagery

We used multi-temporal Sentinel-2 data originating from the Copernicus program. These data are particularly suited to operational land cover mapping at regional to global scale. Sentinel-2 sensors provide 13 spectral band imagery, with spatial resolution ranging from 10 m to 60 m, covering visible, near-infrared and short-infrared parts of the electromagnetic spectrum [41]. Additionally, Sentinel-2 data are the only free of charge, spaceborne imagery produced at such high spatial resolution. We selected 2017, as this was the first year in which both Sentinel-2 satellites (Sentinel-2A and Sentinel-2B) operated. Sentinel-2A has been collecting data since the second half of 2015, while Sentinel-2B started operating in the middle of 2017. The operation of both sensors increased the revisit period to five days, and improved the availability of cloud-free data.

In order to avoid classifying all available Sentinel-2 images, and to reduce computation requirements, we developed a data selection strategy. The aim was to select, for every Sentinel-2 tile in the study area, the two, least-cloudy images for each month in the growing season (April to October) and one for each of the remaining months (January–March and November–December). Specifically, over 50% of the image area had to be cloud-free (cloud, cloud shadows and areas with no data making up the remainder). Therefore, the method assumes the use of 19 images per tile, in ideal conditions. If, in a given month, no image met the criteria, an image from another month with the closest acquisition date was used. These rules were modified for mountainous tiles, tiles with diverse relief, and those with steep, long slopes. The deep shadows in such areas, for images acquired between October and March, led to them being excluded from the analysis. In a few cases, data selection criteria were not met, and less than 19 images were classified.

Sentinel-2's L2A products were accessed from the CREODIAS platform, after atmospheric correction performed with the Sen2Cor processor [41], and representing Bottom-Of-Atmosphere surface reflectance values. Data analyses were only performed on 10 m and 20 m spatial resolution bands, with the latter resampled to 10 m resolution by the nearest neighbour method. Spectral bands with 60 m spatial resolution were excluded, due to their limited usefulness for land surface analyses [44]. The L2A product also includes a scene classification (SCL) map, which provides information on, among other things, the extent of different types of cloud and cloud shadow [41]. The SCL map was used during image processing and classification to mask image pixels and reference data covered by cloud and cloud shadow.

2.2.2. Classification System

Drawing upon the available reference data sources, that include the CLC and HRLs (more details provided in Section 2.2.3), we developed a classification legend composed of 13 classes, and an additional class representing surfaces permanently covered by cloud (Table 1). The overall goal was to develop relatively general, common classes that would make it possible to automate the selection of reference samples and data classification. A second requirement was that the selected classes should be highly homogenous and represent relatively similar characteristics across Europe. Therefore, most of the selected classes represent the physical characteristics of surface cover (land cover), and not how the land is used (land use). Only two classes (artificial surfaces and cultivated areas) can represent different land cover types, as they are managed and used for various purposes. Thus, for simplicity, we use the term land cover in the remaining part of the manuscript for all the mapped classes.

Table 1. Land cover classification nomenclature and description used in the current study.

S2GLC ¹ Class Name	Description
Artificial surfaces	All surfaces where the landscape has been changed by, or is under the influence of human activities that replace natural surfaces with artificial abiotic constructions or artificial materials.
Natural material surfaces (consolidated and un-consolidated)	Any kind of unvegetated surface material including consolidated and un-consolidated surfaces. Any surface with loose mineral particles of any size e.g. mountain slope debris, glacier moraines, river pebble banks, beaches, sand dunes (unvegetated) extraction sites or quarries.
Broadleaf tree cover	Land covered with broadleaved tree canopy that loses leaves seasonally, regardless of the plant height.
Coniferous tree cover	Land covered with needle-leaved tree canopy that do not lose needles seasonally, regardless of the plant height.
Herbaceous vegetation	Land covered by herbaceous vegetation, including both natural, low productivity grassland and managed grassland, used for grazing and/ or mowing.
Moors and heathland	Low growing vegetation with closed cover, and predominately shrub and bushy vegetation (limited herbaceous species allowed).
Sclerophyllous vegetation	Bushy sclerophyllous vegetation characteristic of the Mediterranean climate, including maquis and garrigue. May exist in both closed and discontinuous cover.
Cultivated areas	Cultivated areas managed by humans that include non-irrigated and irrigated arable land with different crops, and land under rice cultivation. It also includes temporary bare soils (e.g. fallow lands).
Vineyards	Areas planted with vines.
Marshes	Low-lying areas covered with non-woody vegetation distinguished by the presence of water at the surface (waterlogged) either permanent or temporary, due to high precipitation rates or flooding by fresh or sea water.
Peatbogs	Peatlands with deposits of decomposed moss or other dead plant material (including exploited peatlands).
Water bodies	Water in a liquid state of aggregation regardless of location, shape, salinity and origin (natural or artificial).
Permanent snow cover	Snow cover that persists throughout the year, above the climatic snow line. Persistent ice cover formed by the accumulation of snow.
Surfaces permanently covered by cloud	Areas where no land cover interpretation is possible due to obstruction caused by cloud and their shadows, smoke or haze.

¹ Sentinel-2 Global Land Cover.

2.2.3. Training Data Sources

The output of a supervised classification relies on the quality and quantity of training samples. Therefore, their appropriate preparation is a crucial step in the workflow. However, the collection of the qualitatively homogenous training dataset on a continental or global scale is a very demanding task [34,45]. Field surveys are infeasible due to the high cost, the required time and workload. Visual interpretation of high resolution EO data is also time-consuming and requires access to a huge set of aerial photography, satellite imagery, or maps. Therefore, the optimal solution is to use existing land cover databases as the data source. Our method drew upon two databases (the CLC and HRLs) to provide complementary information on land cover across the whole study area.

The CLC database is one of the most popular and widely used pan-European land cover databases [25]. Composed of 44 classes, it provides rich and harmonised information on land cover, land use, and changes that occurred between updates, starting from 1990. Produced by national teams, it is integrated and coordinated at the European level by the European Environment Agency. CLC is mainly produced by visual interpretation of EO data (e.g., Landsat, SPOT, LISS III and, more recently,

Sentinel) with a limited level of automation. The nominal scale is 1:100 000 with a MMU of 25 ha for a single areal object, and a minimum width of 100 m for linear objects. The 2012 version of the CLC database was used to produce the training dataset, as it was the most up-to-date version available at the time of our data processing.

HRLs were developed under the Copernicus program. These databases represent information on several land cover types. Each layer shows the distribution of an individual class, together with its characteristics. HRLs are available at 20 m or 100 m resolution and, therefore, outperform CLC in most cases, notably in terms of the detailed outline of certain objects, and more homogenous classes. The method used to produce HRLs combines automated and interactive rule-based classification of EO images, for both optical and radar data [26]. In this study, we used 20 m resolution HRL imperviousness, grassland, and forest layers products from the year 2015.

HRL Imperviousness represents sealed areas as a degree of imperviousness, ranging from 0–100% [46], reflecting built-up areas and other artificial cover. The database was produced using an automated procedure, based on calibrated normalized difference vegetation index (NDVI) values, with overall accuracy (OA) estimated to be above 90% [47]. HRL grassland represents the spatial distribution of grassland and non-grassland areas. This grassy and non-woody vegetation product encompasses natural grassy vegetation, semi-natural grasslands, and managed grasslands, but excludes crop areas [48]. The OA of the grassland layer used in this study is estimated to be approximately 85% [48,49]. HRL Forest consists of the following products: tree cover density (TCD) representing the per-pixel crown coverage in the range 0–100%; and dominant leaf type (DLT) which provides information on the dominant leaf type: broadleaved or coniferous [50]. The TCD product is characterized by high user's and producer's accuracy of above 85% ([51] p. 52), with accuracies of DLT product being in the range of 70–88% for broadleaved and coniferous leaf types ([51] p. 59).

Finally, the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (GDEM) version 2 was used to correct the land cover map in the post-processing phase. ASTER GDEM is a global elevation dataset that is produced jointly by the Japanese Ministry of Economy, Trade, and Industry, and the United States National Aeronautics and Space Administration. GDEM horizontal resolution is 2.4 arc-seconds (72 m), with 12.6 m standard deviation in elevation error [52].

2.2.4. Training Data Preparation

Both CLC and HRL datasets were respectively re-projected, rasterised and resampled to fit into a pixel of Sentinel-2 data. Reference databases were subdivided into parts corresponding to individual Sentinel-2 tiles. Tiles were stored using the Universal Transverse Mercator projection, the original projection of Sentinel-2 product.

The strategy to select training samples from available databases was investigated and developed in the mentioned developing phase. The performed tests focused on the use of different combinations of reference datasets and spectral indices, and helped to design rules for the selection of samples for each land cover class. The aim was to develop rules that were transparent, easy to apply, able to be fully automated, and that did not require the application of any other data sources beyond CLC, HRLs and Sentinel-2. Selection of optimal rules for particular classes resulted from analyses performed within the predefined test sites. A detailed summary of the composition of the training dataset and filtration rules for all land cover classes is presented in Table 2.

Table 2. Filtration rules, and the source of training data. Numbers shown in brackets in the column ‘Training source’ indicate the original numbers of classes in the CLC nomenclature.

Land Cover Class	Training Source	Filtration Rules				
		NDWI ¹	HRL Imp ³	HRL TCD ⁴	HRL Grassland	NDVI ⁵
Artificial surfaces	HRL Imp > 70%	<0.0		<10%	+	
Cultivated areas	CLC (2.1.1., 2.1.2., 2.1.3.)	<0.0	<30%	<10%	+	
Vineyards	CLC (2.2.1.)	<0.0	<30%	<10%	+	
Herbaceous vegetation	HRL Grassland	<0.0	<30%	<10%		
Broadleaf tree cover	Intersection of HRL DLT ⁶ and CLC (3.1.1.)	<0.0	<30%		+	
Coniferous tree cover	Intersection of HRL DLT and CLC (3.1.2.)	<0.0	<30%		+	
Moors and heathland	CLC (3.2.2.)	<0.0	<30%	<10%		
Sclerophyllous vegetation	CLC (3.2.3.)	<0.0	<30%	<10%	+	
Natural material surfaces	CLC (1.3.1., 3.3.1., 3.3.2.)	<0.0	<30%	<10%	+	<0.25
Permanent snow	CLC (3.3.5.)	<0.0	<30%	<10%	+	
Marshes	CLC (4.1.1., 4.2.1.)	<0.0	<30%	<10%	+	
Peatbogs	CLC (4.1.2.)	<0.0	<30%	<10%	+	
Water bodies	NDWI and CLC (5.1.1, 5.1.2., 5.2.1., 5.2.2., 5.2.3.)	>0.2				<0.25

¹ Normalized Difference Water Index, ² High Resolution Layer, ³ Imperviousness, ⁴ Tree Cover Density, ⁵ Normalized Difference Vegetation Index, ⁶ Dominant Leaf Type.

In most cases (cultivated areas, vineyards, sclerophyllous vegetation, marshes, peatbogs, natural material surfaces, permanent snow) training datasets were taken from the corresponding CLC classes (Table 2). They were, however, filtered using the Normalized Difference Water Index (NDWI), HRL Imperviousness, HRL TCD and HRL Grassland, and empirically-estimated thresholds (Table 2). This step was intended to reduce mislabelling resulting from the generalisation of the CLC dataset, and its MMU. Additionally, the natural material surfaces class was filtered with the NDVI, as there should be no vegetation in this class. Samples for artificial surfaces and herbaceous vegetation were taken from HRL Imperviousness and HRL Grassland layers, respectively, and were also filtered using the NDWI, and the other HRL Imperviousness, HRL TCD and HRL Grassland datasets.

For broadleaf and coniferous tree cover classes, training data were prepared based on the spatial intersection of the HRL DLT dataset and the CLC classes broadleaved forest (3.1.1.) and coniferous forest (3.1.2.), respectively. These classes were also filtered using the NDWI, HRL Imperviousness and HRL Grassland data. The moors and heathland class were similarly filtered using the NDWI, HRL imperviousness, and HRL TCD.

Finally, frequent changes in the extent of water surfaces meant that training data for the water bodies class required a different approach. They were derived by combining NDWI and CLC water-related classes. The water mask resulting from NDWI thresholding (> 0.2) was subdivided into inland waters (5.1.1., 5.1.2.) and marine waters (5.2.1., 5.2.2., 5.2.3.) with the appropriate CLC classes. In tiles where both sub-classes occurred, training samples for the water bodies class were selected equally from both. This approach ensured the uniform representation of all types of water surfaces in the classification model [45]. Training samples for the water bodies class were also filtered with an NDVI threshold of <0.25 to exclude pixels with a vegetation component.

It is important to note that filtration with the NDVI and the NDWI was performed separately for the classification of each Sentinel-2 image. This approach helps to clean artefacts resulting from the low resolution of reference databases, or due to changing environmental conditions (e.g., floods). The main purpose of the filtration rules was to filter out false positives in a given land cover class, although it should be noted that they may also have excluded some true positives. Nevertheless, in most cases, dozens of thousands of samples could be selected for each class, for each tile. The size of this dataset meant that the required subset of samples could be selected for multiple classifications.

2.3. Classification Procedure

2.3.1. Classification Features and Algorithm

The core element in the procedure is the pixel-based classification of Sentinel-2 images. Images in every tile were independently analysed. The random forest classifier was applied. This ensemble-based algorithm generates numerous decision trees from a randomly-selected subset of both training samples and classification features [53]. The characteristics of the random forest classifier makes it more resistant to overfitting, and erroneous training samples than other machine learning algorithms [54]. The final decision is based on classification results from all decision trees, using majority voting. The two main parameters of random forest are the number of trees to be generated and the number of features selected for each split in the tree. We used a relatively small set of trees (50) to limit the calculation time, as increasing the number of trees has only a minor influence on the final classification result [55,56]. The second parameter was set to the square root of the number of input features, which also helped to limit the computational complexity of the classifier [57].

For every tile, each Sentinel-2 image was classified separately at pixel level, with a new set of training samples selected randomly from the prepared dataset (Section 2.2.4.). The number of training samples was set to 1000 for every class, with a minimum of 50 (for very cloudy images). Pixels representing cloud and cloud shadow were masked using information from the SCL map.

It is common practice to use, besides the spectral bands of satellite imagery, other explanatory variables (classification features), such as spectral indices derived from the analysed images [34,58–61]. Therefore, to enhance recognition of land cover classes, we fed the classifier with similar features. However, instead of calculating predefined, commonly-used indices such as NDVI, we calculated all possible combinations of Sentinel-2 spectral bands following Equation (1), namely the normalised difference between two spectral bands:

$$\text{Spectral index} = (\text{BandX} - \text{BandY})/(\text{BandX} + \text{BandY}), \quad (1)$$

where BandX and BandY are any two bands of a satellite image. A total of 100 features were available for classification for every analysed Sentinel-2 image.

2.3.2. Aggregation

After classification of all Sentinel-2 data, the results from individual images for every tile were combined into a single output. This procedure, called aggregation, was developed as an alternative to the analysis of multi-temporal data [62]. The method is based on two sets of input data, both resulting from the application of the random forest classifier. The first is the classification result, in the form of a thematic land cover map. The second, often called the posteriori class probability [11,63], is a raster image that presents information on the number of trees voting for the final class found in the ensemble, for each classified pixel, divided by the total number of trees [3]. In the aggregation process, posteriori probability values are summed, for each land cover class found in the investigated pixel, from all images of a time series. Then, the sum is divided by the number of all valid occurrences (not cloudy) in a time series, regardless of the class label. The method assumes that individual classifications are not weighted by acquisition date, cloud coverage or thematic accuracy. The only weighting factor is the posteriori probability, which is applied to each pixel individually, as explained above. Figure 4 presents an example of the aggregation process for a single pixel.

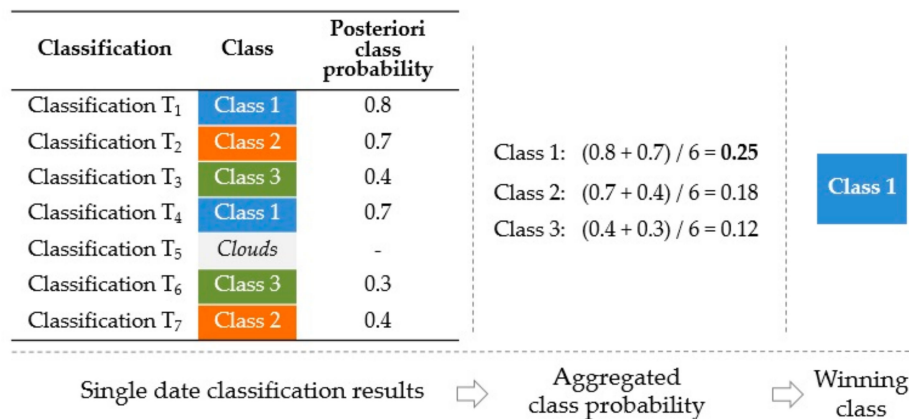


Figure 4. An example of a class label estimation with the aggregation procedure for a single pixel from seven multi-temporal images. Class 1 is the winner, and is selected for use in the resulting land cover map.

2.3.3. Post-Processing

Post-processing aimed to refine the classification result, by reducing post-classification errors. It was applied to the map resulting from the aggregation process. Classification errors are typically due to spectral similarities between classes, or spectral values that are influenced by terrain characteristics such as slope, altitude or sun illumination [22]. A commonly-used method is to replace noisy, isolated pixels (the salt-and-pepper effect) using median filters [11], resulting in a smoothed thematic map. However, all small objects represented by single pixels, including those that have been correctly classified, are removed. Our procedure only modifies pixels with a relatively low posteriori class probability obtained from the random forest classifier, and modified during aggregation. In addition, for each given pixel, information about its surrounding pixels is analysed, along with slope and altitude characteristics derived from the digital elevation model.

All five post-processing steps concern the artificial surfaces class. Built-up areas are characterised by a wide range of spectral values, and are spectrally similar to other land cover types such as bare soil, rock, and other bright surfaces. Consequently, the class is frequently overestimated, and our refinement aims to reduce misclassification. The following steps were designed to be performed in sequence, while parameter and threshold values were derived during the development phase of the study:

- Pixels in the class artificial surfaces with posteriori probability values below 0.35 are reclassified as belonging to the neighbouring class that they have the longest border with.
- Pixels located on the edges of water surfaces (rivers, lakes, seas) are reclassified as water bodies if they are initially classified as artificial surfaces with a posteriori probability value below 0.48, and if the adjacent water body covers over 5 ha.
- Pixels in the artificial surfaces class are reclassified as natural material surfaces if the posteriori probability value is below 0.48, and if they are completely surrounded by an area, classified as natural material surfaces larger than 1 ha.
- Pixels classified as artificial surfaces, located in high mountain regions, above a certain altitude or on steep slopes are reclassified as the next class given by ranking of posteriori probability values. Thresholds are adjusted to fit the characteristics of different mountainous areas in Europe. The rule was applied to 171 mountainous tiles.
- The last step was related to the erroneous cloud mask originating from the Sen2Cor processor [64]. Some areas with high spectral reflection values, such as built-up areas, are misclassified in the cloud mask as cloud. Consequently, in some cases, all cities, especially in the south of Europe, are permanently masked out. To minimise this problem, for every tile, the least-cloudy image from the predefined time series was selected and classified without applying the cloud mask.

In this way, incorrectly masked-out pixels from the result of the aggregation were replaced with the result of the classification of the least-cloudy image.

2.4. Validation of the Land Cover Map

We validated the thematic accuracy of the land cover map product, consisting of 13 classes (Table 1) and cloud, which was the result of the post-processing of the aggregated time series classifications. To ensure that the assessment was unbiased and reliable, a fully-independent reference dataset was created.

The selected validation sites were distributed by stratified random sampling with proportional allocation over Europe, with country borders representing stratum in the dataset. The sampling strategy was based on a single Sentinel-2 tile covered by at least 75% land. The accuracy assessment was performed for approximately 10% of all Sentinel-2 tiles in each country, except for Andorra, the Vatican, San Marino, and Lichtenstein (Figure 5). These sites were named “base sites”. Luxemburg was included in one of three sites selected for Germany (Figure 5). For some countries, the selected Sentinel-2 tiles reached near-zero values. In these cases, to assure the accuracy assessment on a country level, it had been decided to randomly select a single tile, named “added sites” (Figure 5). Sites located at national borders were merged with sites selected for the neighbouring country, and are represented in the summary as validation sites for the two countries. This was the case for Albania and Greece, and Croatia and Slovenia. Table 3 presents a summary of validation sites and sample selection.

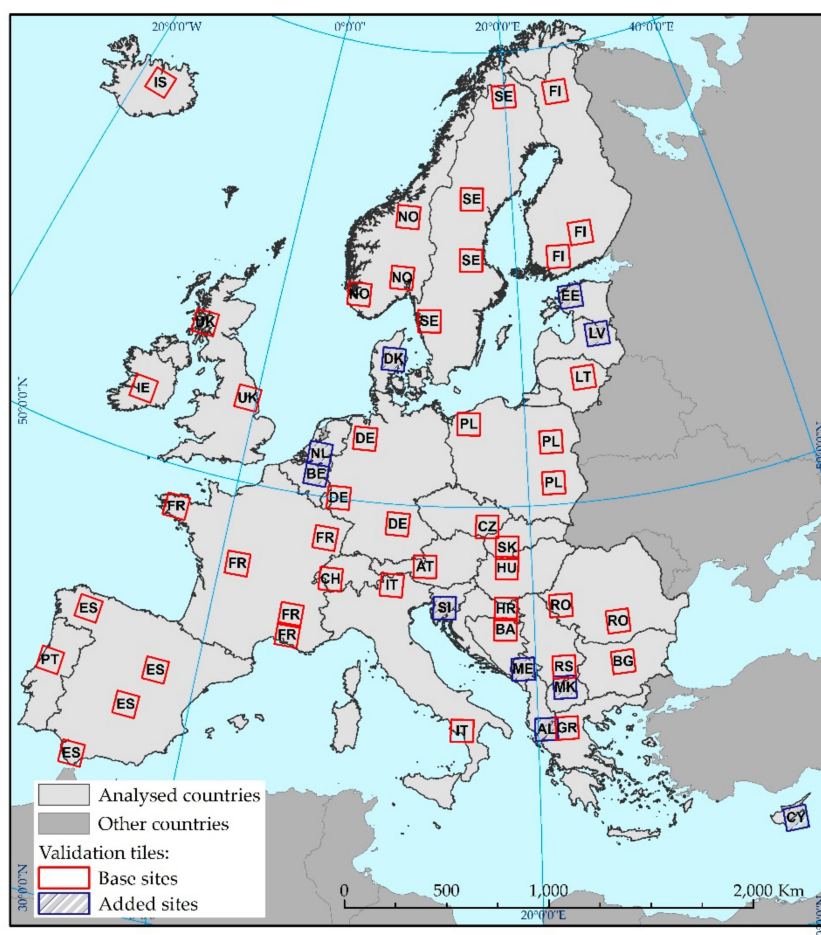


Figure 5. Map of randomly-selected validation sites showing country codes. Base sites (shown in red) represent sites that compose 10% of tiles in a given country. Added sites (shown in blue) represent cases where a random tile was selected, because 10% of tiles had a near-zero value.

Table 3. Summary of the randomly-selected validation sites.

No. of Base Sites (10% of Tiles)	No. of Added Sites	Total No. of Sites	Mean Country Area Coverage
45	10	55	21%

The stratification of validation samples was based on CLC reference data. We selected the appropriate land cover class using our legend (Table 1). An additional class was created by merging all the remaining CLC classes. Samples were selected randomly from all classes within the strata, as a function of the proportion of a given class within a specific Sentinel-2 tile, and were submitted for verification without the predefined land cover class given in the attribute table. Following the assumption of the selected samples clear identification in respect to the assessed outcome and reference data, the number of selected samples was increased from pre-estimated 800–1000 per tile. This approach was driven by the fact that some samples needed to be eliminated due to their location on a border between land cover classes or, in specific cases, the inability to identify a class.

The following conditions were applied to the stratification process:

- Land cover classes covering less than 0.95% of a tile were excluded.
- For land cover classes covering between 0.95–2% of a tile, a minimum of 20 validation samples was set.
- For artificial surfaces and natural material surfaces classes, a default minimum of 20 validation samples was set, even if they covered less than 0.95% of a tile.

This strategy resulted in a set of 55,000 initial samples for the 55 validation sites.

Verification of the initial sample dataset was based on reference data. The latter had to meet rigorous criteria, in order to compensate for missing ‘ground truth’ found in in-situ surveys. The main objective was to verify samples against real-world conditions carried out as virtual truthing with EO data. Therefore, the applicable EO data had to have the same or preferably higher spatial resolution, and represent the same time span as the data used to produce the land cover map. Planet Scope imagery, in combination with Sentinel-2 was chosen, representing a similar time range. This latter criterion is necessary to ensure the correct analysis of samples representing classes with seasonal changes, where multi-temporal analysis is necessary.

Verification took the form of a stable, manual, but efficient workflow. The area was divided into zones representing different climate regions, with each zone coordinated by an expert in the area. The process resulted in a very rich set of 51,926 samples that could be used to validate the thematic maps. The minimum number of 800 samples was met for all sites. The final stage was to compare the reference data with the results of the automatic classification. We developed metrics that indicated thematic accuracy in terms of the effectiveness and reliability of the information content group. Using validation samples, confusion matrices were calculated for each country, and for the entire study area. In addition to the overall classification accuracy, the confusion matrix presents individual accuracies for each land cover category, along with both user’s accuracy (commission errors) and producer’s accuracy (omission errors). Finally, the F1 score was calculated to indicate the balance between commission and omission errors.

3. Results

Figure 6 illustrates the results of the land cover classification performed with our methodology, for the year 2017, at a spatial resolution of 10 m. The map is composed of 791 separately classified Sentinel-2 tiles, trimmed to take account of coastlines. Most of the sea area in the water bodies class, along with areas in external countries, was excluded. This map can be downloaded from the S2GLC project website and the CREODIAS platform [37]. Our classification found that the most widespread classes were cultivated areas (nearly 20% of the study area), broadleaf tree cover (19%), herbaceous vegetation (19%) and coniferous tree cover (over 16%). The remaining nine classes sum up to only 25%

of the study area. Table 4 presents the confusion matrix; accuracy measures are derived from samples for all validation sites. On a continental level, our classification achieved 86.1% OA, with a kappa coefficient of 0.83, and a weighted average F1 score of 0.86. At the land cover class level, water bodies and coniferous tree cover were identified most accurately (F1 score 0.96), followed by broadleaf tree cover (0.95) and cultivated areas (0.91). Accuracy was lowest for permanent snow (0.16), sclerophyllous vegetation (0.30) and marshes (0.30).

The F1 score was above 0.91 for most of the classes with the widest coverage. The exception was herbaceous vegetation, with a score of 0.77. Here, grassy vegetation was confused with cultivated areas, and moors and heathland (Table 4). Among the less-frequent classes, the F1 score was highest for water bodies, followed by artificial surfaces (0.80) and natural material surfaces (0.70). The two last classes were also mixed often with each other. Peatbogs (F1 score 0.66) were often confused with moors and heathlands (F1 score 0.45), which, in turn, were frequently misclassified as herbaceous vegetation. Accuracy was moderate (F1 score 0.60) for vineyards, despite a high producer's accuracy (89%). This result was due to their misclassification as cultivated areas.

At the country level, OA values range from 67% to over 95% for all 32 countries (or pairs of countries) (Table 5). For all countries, average OA is high (86.5%) and the average kappa coefficient is 0.82. Table 5 shows that OA for 12 of the 32 countries exceeded a high value of 90%, while 14 countries fall in the range 80–90%. For the other five countries, OA was in the range 70–80%, and was below 70% for only one country. OA scores were highest for the Germany/ Luxembourg pair (95.8%), Slovakia (95.4%) and Ireland (95.0%), and scores were lowest for Portugal (67.0%). From a climatic perspective, accuracy was relatively evenly distributed, although values were slightly lower for Mediterranean countries (Table 5).

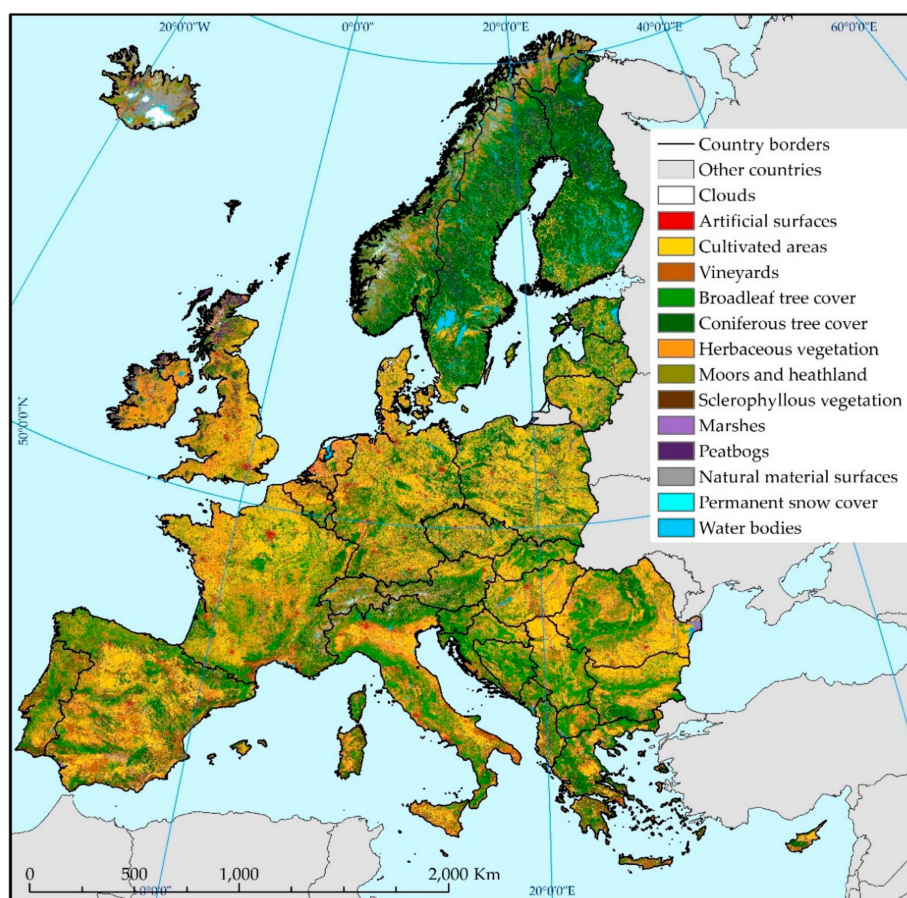


Figure 6. Land cover map of Europe for the year 2017 produced by the new methodology, showing the division into 13 land cover classes and cloud.

The post-processing procedure only increased OA by 0.3% (from 85.8% to 86.1%) on a continental scale. However, a greater improvement was observed for the artificial surfaces class. Although producer's accuracy decreased by 3% (to 85.1%), user's accuracy increased by 13%, and the F1 score by 7%. Accuracies of most of the remaining classes did not change considerably. The extent of areas affected by post-processing was a function of the local land cover type, and reached a maximum of 5.7% for one tile. On a continental scale, 1.2% of the entire study area (60 633 km²) was modified by post-processing.

Table 4. Land cover classification confusion matrix for the entire study. The diagonal numbers shown in bold represent the number of correctly-classified pixels of each class.

Classification Data	Reference Data													UA (%) ¹
	1	2	3	4	5	6	7	8	9	10	11	12	13	
1 Artificial surfaces	1548	65	1	7	0	0	5	2	430	0	0	0	6	75.0
2 Cultivated areas	37	11,774	26	574	32	5	12	23	46	0	4	0	1	93.9
3 Vineyards	8	426	445	58	12	1	10	26	7	0	0	0	1	44.8
4 Herbaceous vegetation	15	886	20	5592	127	6	689	289	72	7	88	13	0	71.7
5 Broadleaf tree cover	5	44	1	40	10,436	202	165	113	4	0	12	5	1	94.6
6 Coniferous tree cover	8	16	2	12	171	8374	82	60	1	0	3	6	2	95.8
7 Moors and heathland	10	65	1	249	53	24	760	50	49	0	5	58	2	57.3
8 Sclerophyllous vegetation	3	26	0	39	54	10	53	181	3	0	4	0	0	48.5
9 Natural material surfaces	96	73	2	86	2	0	38	64	1319	7	1	2	7	77.7
10 Permanent snow	4	0	0	0	0	0	0	0	73	9	0	0	0	10.5
11 Marshes	37	76	2	55	40	3	58	7	18	0	140	83	79	23.4
12 Peatbogs	5	17	0	62	16	4	197	0	5	0	61	578	55	57.8
13 Water bodies	42	1	0	4	2	1	2	0	63	6	6	0	3558	96.6
Producer's accuracy (%)	85.1	87.4	89.0	82.5	95.3	97.0	36.7	22.2	63.1	31.0	43.2	77.6	95.9	
F1-score per class	0.80	0.91	0.60	0.77	0.95	0.96	0.45	0.30	0.70	0.16	0.30	0.66	0.96	
Overall accuracy	86.1													
Kappa coefficient	0.83													
Overall F1-score	0.86													

¹ User's accuracy.

Table 5. Country-level accuracy statistics for the new land cover map.

Country	No. of Validation Sites	% Country Area	No. of Validation Samples	Overall Accuracy	Kappa Coefficient
Albania, Greece	2	15%	1935	70.8	0.64
Austria	1	14%	926	84.5	0.78
Bosnia and Herzegovina	1	24%	890	92.7	0.89
Belgium	1	39%	979	87.6	0.84
Bulgaria	1	11%	989	89.2	0.84
Switzerland	1	29%	883	90.1	0.88
Cyprus	1	46%	845	83.5	0.77
Czechia	1	15%	951	94.1	0.91
Germany, Luxemburg	3	10%	2908	95.8	0.94
Denmark	1	28%	978	94.0	0.91
Estonia	1	27%	929	89.0	0.87
Spain	4	10%	3499	78.7	0.75
Finland	3	11%	2764	91.4	0.88
France	5	9%	4854	88.5	0.86
Croatia, Slovenia	2	32%	1829	92.6	0.90
Hungary	1	13%	909	93.0	0.90
Ireland	1	17%	929	95.0	0.89
Iceland	1	12%	953	82.7	0.67
Italy	2	8%	1727	88.0	0.85
Lithuania	1	19%	978	82.5	0.76

Table 5. Cont.

Country	No. of Validation Sites	% Country Area	No. of Validation Samples	Overall Accuracy	Kappa Coefficient
Latvia	1	19%	910	84.8	0.81
Montenegro	1	87%	940	77.4	0.68
North Macedonia	1	48%	959	77.3	0.69
Netherlands	1	32%	904	87.6	0.84
Norway	2	7%	2721	74.3	0.70
Poland	2	8%	2661	92.9	0.90
Portugal	1	13%	890	67.0	0.62
Romania	2	10%	1956	89.1	0.83
Serbia	1	14%	1019	92.0	0.85
Sweden	4	11%	3843	80.0	0.74
Slovakia	1	25%	930	95.4	0.92
United Kingdom	2	10%	1882	85.7	0.82

4. Discussion

This study aimed to develop a complete, ready for operational use automated land cover classification workflow based on multi-temporal Sentinel-2 imagery. The proposed procedure meets all predefined objectives. It demonstrates that it is possible to fully automate all of the steps required to process large datasets, from training data preparation, access to Sentinel-2 imagery, to data classification and post-processing. The high automation of the method enables straightforward processing of new training and EO data, and frequent and repetitive large-scale production of land cover maps. Furthermore, the 10-m spatial resolution ensures a very high level of detail of the landscape (Figure 7).

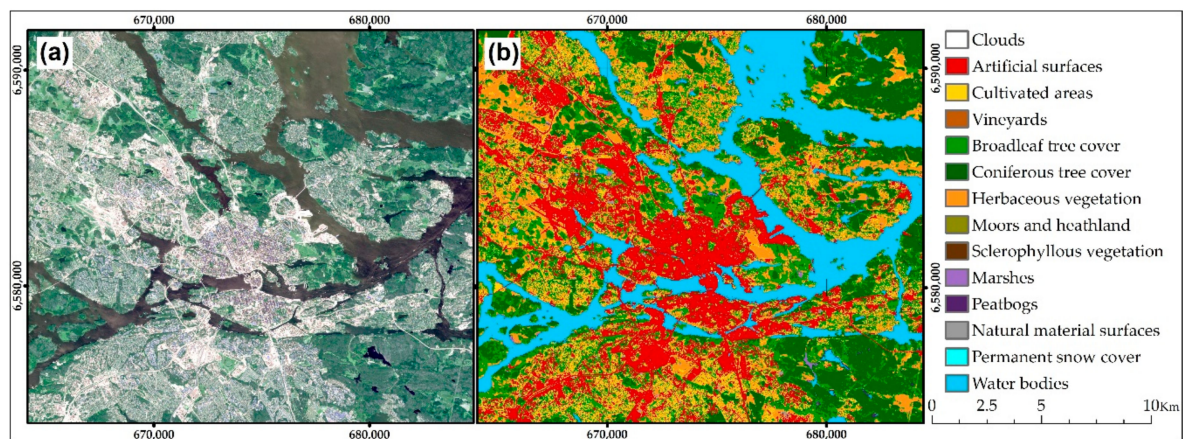


Figure 7. A subset of the tile 34VCL illustrating the detailed land cover classification of the city of Stockholm and its neighbourhood: (a) a Sentinel-2 image displayed in natural colours; (b) the resulting land cover classification.

The operability of our classification system is enhanced by the application of the CREODIAS computing environment. On the one hand, CREODIAS supplies the system with Sentinel-2 data and, on the other hand, enables the efficient analysis of huge datasets that remarkably streamlines overall data processing.

Our analyses of Sentinel-2 data uses a cloud mask originating from the Sen2Core processor, regardless of its drawback related to the misinterpretation of high-reflectance areas [64,65]. However, application of ready-to-use datasets, available from the Copernicus Scientific Hub or CREODIAS, allowed avoiding additional testing and the implementation of third-party algorithms and processing tools. This strategy increases the operational usefulness of our method, which relies entirely on open-source components.

4.1. Classification Accuracy

OA of 86.1% is very high for an automated method, and a promising indication for the future. Classification accuracy for individual land cover classes varies, with the most accurate results obtained for the most highly-represented classes of cultivated areas, broadleaf tree cover, coniferous tree cover, herbaceous vegetation and water bodies [33]. Almost all producer's and user's accuracies of these classes were well above 80%, and often exceeded 90% for both or one of them (Table 4). Additionally, good results were obtained for the artificial surfaces class, despite the much lower spatial coverage. It was very well recognised by the classifier (omission errors < 15%) and is characterised by a relatively low level of commission error (approx. 25%).

Other classes, with relatively low spatial coverage, were less-well classified. Examples include shrubland classes (moors and heathland, sclerophyllous vegetation) and wetland classes (marshes and peatbogs) (Table 4). These groups appear to be very difficult to map accurately. However, our results are consistent with other studies, which find that accuracy is lower for wetlands and shrublands than other vegetation types [20,33,58,59,66]. Therefore, the poor recognition of these classes in our study is neither related to the high level of automation, nor to the large scale of the mapping. Instead, it can be attributed to the spectral similarity of these classes both with other vegetation classes, and between each other [67].

The very low accuracy for the permanent snow cover class may be explained by outdated CLC reference training data, and shrinking permanent snow cover and glaciers in the Alps, Scandinavia, and Iceland [68,69]. Another reason is that the high reflectance of this class is similar to the natural material surfaces class. This similarity, and the natural coexistence of these classes, increases misclassification. This class was also represented by the fewest validation samples, which may have further reduced accuracy.

Undoubtedly, further research is needed to improve the quality of recognition of all of the poorly-classified classes in our study. The analysis of our classification results suggests that the legend could be modified to combine selected classes, based on their thematic and spectral similarity. The results of our study suggest that vineyards could be combined with cultivated areas, moors and heathland could be combined with herbaceous vegetation, and marshes could be combined with peatbogs. These combinations would increase OA from 86.1% to 89.0%. A classification based on 10 rather than 13 classes would provide improved and more homogeneous classification results for nearly all classes (except permanent snow and glaciers, and sclerophyllous vegetation).

Another issue that needs to be addressed in further research is a problem appearing in several locations on tiles' boundaries. In some cases, artefacts can be found on the tile's edges that show disagreement in the classification between adjacent tiles. Further development, therefore, is planned to address this issue and improve the final land cover product.

An analysis of differences in accuracy between neighboring countries was beyond the scope of this study. Such analysis would require additional effort and more detailed reference datasets. However, the existing differences occur, most probably due to the varying proportion of problematic classes that are difficult to map, such as sclerophyllous vegetation, moors and heathland, and wetlands classes.

4.2. Training Data Preparation

The training samples used in our methodology are taken from existing land cover databases, which makes the automation of the classification process easier. There is no need to organise field campaigns, or prepare training samples through the visual interpretation of EO. Indeed, such processes are prerequisites in numerous projects [33,34]. The joint application of the complementary CLC and HRL datasets, along with spectral indices derived from satellite imagery, help to improve the quality of the training set by filtering out erroneous samples. Another potential source of inaccuracy results from differences in scale and the MMU in training sources compared to Sentinel-2 data. This is minimised by the use of the random forest classifier, which is known to be less sensitive to the quality of input training data than other machine learning algorithms [33,53,54,70]. Another advantage of the

application of existing databases is that they cover large areas (e.g., the whole of Europe), and provide huge sets of samples. This is important for the individual classification of satellite scenes with often highly-mosaicked cloud coverage. Finally, a rich set of training data enables the application of an independent set of samples to the classification of individual images from a time series; this, in turn, ensures a more objective analysis that is not fitted to a specific sample set.

The accuracy assessment performed with independent validation dataset showed high values for most accuracy measures, although the accuracy of classification of individual land cover classes are at different levels. Nonetheless, the study shows that the application of our automated method to select training samples is highly efficient, and enables accurate land cover classification. A direct comparison with other studies is questionable due to differences among classification systems, the MMU, geographical extent, and validation approaches [6]. It is clear, however, that our method outperforms or at least achieves similar thematic accuracy to those based on the manual collection of training samples [7,20,33,34,71].

4.3. Direct Comparison with HRLs and Other Land Cover Datasets

In addition to checks of accuracy, we ran a quantitative comparison of our results with comparable HRL datasets. The comparison covered four classes including artificial surfaces, broadleaf tree cover, coniferous tree cover and herbaceous vegetation (Table 6). The selected HRL databases are most similar, in terms of thematic category, to the respective classes in our classification. Furthermore, they were the most up-to-date, continental-scale dataset available, with only two years between the compared databases.

Table 6. Statistical comparison of experimental land cover map classes and HRL databases for the classified area of Europe. Percentage cover refers to the total coverage of the S2GLC database.

S2GLC ¹ Classes	Km ²	%	HRL Classes	Km ²	%
Artificial surfaces	113,489.7	2.2	IMP ²	123,301.9	2.4
Both tree types	1,789,169.0	35.5	TCD ³	2,001,695.2	39.7
Broadleaf tree cover	965,083.3	19.1	DLT ⁴	1,214,105.5	24.1
Coniferous tree cover	824,085.7	16.3	DLT	787,589.7	15.6
Herbaceous vegetation	957,028.3	19.0	GRA ⁵	655,272.9	13.0
Sclerophyllous vegetation	132,200.5	2.6			
Total	5,046,219.2	100.0			

¹ Sentinel-2 Global Land Cover, ² Imperviousness, ³ Tree Cover Density, ⁴ Dominant Leaf Type, ⁵ Grassland.

The presented statistics (Table 6) show relatively good agreement between both databases. The difference is smallest for artificial surfaces and coniferous tree cover, 0.2% and 0.7%, respectively. Broadleaf tree cover differs by 5%, with greater coverage in the HRL. However, the HRL broadleaf class also contains sclerophyllous vegetation [50], which we classified separately. When sclerophyllous vegetation and broadleaf tree cover were combined, they summed up to 21.7% (Table 6), reducing the difference between S2GLC and HRL data to 2.4%. The greatest difference in coverage was found for herbaceous vegetation. The area covered by grassy vegetation was found to be smaller by 6% in the HRL database. Although some differences between the databases were expected, this difference was much bigger than for other classes. The discrepancy may result from the slight overestimation of grassland in our classification, and the misclassification of cultivated areas, moors and heathland, and broadleaf tree cover (Table 4). The latter areas are often confused with each other, due to their similar spectral response [3,58,66]. Finally, both user's and producer's accuracies of the HRL Grassland dataset are estimated to be around 55% [49]. Thus, it is rather difficult to draw reliable conclusions from the comparison with this dataset. Compared to other methods, which are characterised by a high level of automation of both training data collection and the overall workflow, our approach appears to be more accurate, faster and simpler to use. For example, Pflugmacher et al. [58] presented an approach based on Landsat imagery and training samples derived from the European Land Use/Cover Area

frame Survey (LUCAS) point-based dataset. On the one hand, their reference data is more up-to-date, as the update period is three years, compared to six years for the CLC data used in our study. On the other hand, however, the LUCAS dataset is spatially limited to EU countries, and the number of surveyed samples is limited, which may be problematic for the classification of imagery with greater cloud cover. Although user's and producer's accuracies for particular classes fluctuate more in our approach, OA is better, by over 10%. Another important advantage is the fact that our methodology is based on one year time series, while Pflugmacher et al. [58] used imagery covering a period of up to three years to achieve the mentioned accuracies.

Inglada et al. [59] proposed a fully-automated methodology and demonstrated its performance on a country scale, providing a land cover map of France. Like our approach, training data are based on the CLC data, but unlike our method, they draw up on other, national datasets that are not publicly available for use in other EU Member States. This reduces the transferability of their method to larger areas. Although the accuracy of their method is similar to ours, its performance on a larger scale (e.g., continental) has not been tested and results may differ substantially when applied to different landscapes and climate zones.

Our study reports levels of accuracy that are similar to the latest land cover map found in NLCD products [27,29]. This Landsat-based dataset corresponds spatially to Europe, it includes numerous sub-products and dates back to the early 1990s [28]. However, it is characterized by a highly complex data preparation and classification process, requiring numerous input from experts and manual tasks. Additionally, training data are derived from a comprehensive list of ancillary datasets that would be difficult to collect for a similar area composed of separate countries, such as the EU.

The above comparisons highlight that our methodology is a valuable alternative to other classification workflows and programs, and ensure fast, automated data processing providing a reliable land cover product that can be easily used for annual updates. While there are numerous studies that describe land cover mapping methodologies, they are either based on manually collected training data or provide a single-class product, and are therefore beyond the scope of our analysis. Typically, they either provide a thematically-limited product that is difficult to compare to multi-class databases, or the processing workflow is based on manual data analysis, which considerably limits their application to large-scale land cover mapping.

4.4. Post-Processing Performance

The post-processing procedure is an integral part of the workflow, and was designed to correct common classification errors. It was applied primarily to the artificial surfaces class, and thus most changes were found for this class, simultaneously increasing its mapping accuracy (i.e., an increasing of F1 score by 7%). Although our automated rules corrected a number of classification errors, the difference in OA with and without post-processing is not significant (an increase of 0.3%). However, visually the final classification map has been improved, especially in urban and mountainous areas. Figure 8 presents a comparison of results before and after post-processing for a mountainous area in Norway, and illustrates the reduction in the number of misclassified pixels for the artificial surfaces class. This figure also highlights the problem of correct land cover class recognition in mountains. The spectral similarity between rock and artificial surfaces, along with numerous shadows, rapid changes in altitude, and the aspect of slopes, make mountain areas one of the most difficult regions to map accurately [59]. Figure 9 illustrates the positive changes resulting from the application of post-processing to a data subset representing urban areas of the city of Athens.

The development of our post-processing rules required the determination of exact thresholds for the posteriori class probability and elevation. This was based on the manual verification of values in numerous locations across Europe, which aimed to find thresholds that would best-fit the entire study area. Therefore, it would be necessary to determine new thresholds before applying the method to other areas or continents. Once specified, however, these thresholds can be applied automatically to all analysed tiles, improving the classification.

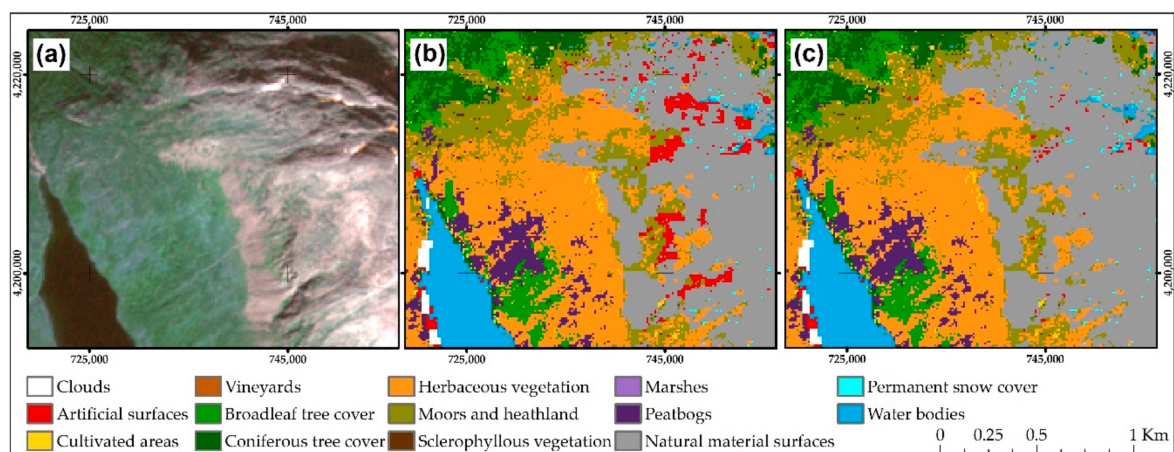


Figure 8. An example of the application of the post-processing procedure: (a) a cloud-free Sentinel-2 image in natural colours; (b) classification results before post-processing; note the extent of pixels representing artificial surfaces (in red) in rocky areas; (c) classification results after post-processing; most of the incorrectly-classified artificial areas have been reclassified as natural mineral surfaces (in grey).

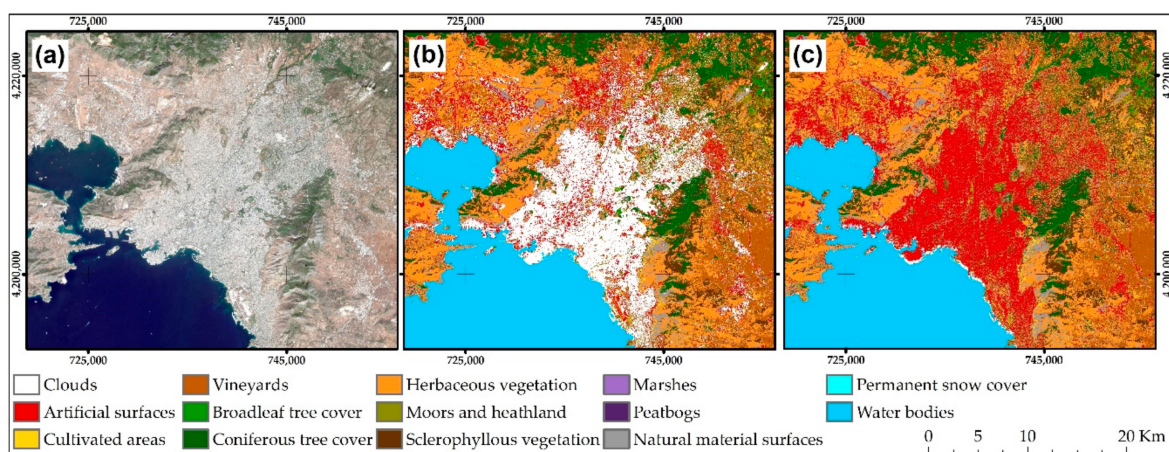


Figure 9. An example of a built-up area misclassified as cloud in the area of the city of Athens: (a) a cloud-free Sentinel-2 image in natural colours; (b) classification results before post-processing; note that an enormous part of the urbanised area is masked out as cloud (in white); (c) classification results after post-processing; cloudy pixels have been reclassified as artificial surfaces (red).

Finally, we found that posteriori class probability data generated by the random forest classification was very useful in improving land cover classification during post-processing [11]. The application of information originating from the posteriori class probability allowed us to focus only on certain, potentially erroneous areas (pixels), and prevented the modification of correctly classified pixels.

4.5. Regional Class Differences

The spectral signature of some land cover classes may vary in different regions of an analysed area [1,59]. This particularly concerns vegetation classes that may be composed of different plant species (e.g., different broadleaf tree species), or be maintained under different land management systems (e.g., agricultural areas) [58,72]. To address this issue, and account for geographic variation in land cover class signatures, our approach performs locally adaptive classification [3,58] through the individual analysis of Sentinel-2 images, on a tile-by-tile basis. Here, we assume that a single tile (110 × 110 km) is sufficient to account for changing climatic conditions and phenological differences. This approach forces the classifier to learn locally, without aggregating class characteristics from

large geographic regions. It also helps to reduce discrepancies in atmospheric correction between neighbouring areas (tiles), and to prevent the transfer of training-related classification errors from one area (Sentinel-2 tile) to other regions. Phenological differences are also taken into consideration by the classification of individual time series images, which approach enable and enforce, in fact, application of multi-seasonal training set. Such a strategy can account for the current state of vegetation, and avoids the need to generalise or average phenological characteristics.

5. Conclusions

This study presents details of a large-scale land cover classification workflow, and illustrates its application to a study area covering the majority of the European continent. It demonstrates that the thematic accuracy of our method is at least comparable to (or higher than) current techniques that use a training dataset derived from the visual interpretation of remote sensing data. It shows that even relatively outdated databases, with coarser resolution than the image being classified, may be an effective source of training samples, and enable mapping with high thematic accuracy. The method is enhanced by the efficient random forest machine learning classifier, which is highly resistant to noise in training data. Moreover, the selection of an independent training set for each classified image eliminates the need for training characteristics to be transferrable between areas that may have different spectral signatures for the same class.

Our method aims to fill a research gap addressing the problem of automatic, large-scale land cover classification, with an emphasis on the fully automatic generation of a training dataset. Furthermore, we demonstrated the application of an alternative method for the analysis of multi-temporal data.

The satisfactory performance of our method was confirmed using an appreciable validation dataset that made it possible to carry out a reliable assessment on both continental and country levels. Like in other studies, accuracy varies, with some classes having an F1 score above 0.9, while others perform below expectations. The latter requires further investigation, and improvements are needed to create an even more robust classification workflow.

An important aspect of our methodology is that it combines all data processing steps, enabling detailed mapping with no need for user intervention. The operability of the method is enhanced by the application of the DIAS computing environment, which facilitates both access to Sentinel-2 imagery and data processing.

The methodology presented in our study could be developed into a fully automated classification service, able to provide maps composed of basic land cover classes at a relatively high (e.g., annual) frequency. Such an automated and computationally efficient service could supplement existing land monitoring services and provide valuable input when analyzing environmental changes.

Author Contributions: Conceptualization and methodology, S.L., R.M., M.R., E.G., M.K. (Michał Krupiński), and A.N.; software, M.R., C.W. and M.K. (Marcin Krupiński) validation, M.J., E.G., E.K. and P.S.; investigation, S.L., M.R., E.G., R.M., M.K. (Michał Krupiński) and C.W.; writing—original draft preparation, R.M., M.J., S.L., E.G. and M.K. (Michał Krupiński); supervision and project leading S.L.; All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the European Space Agency under the Scientific Exploitation of Operational Missions (SEOM) element—S2-4Sci Land and Water, as the project Sentinel-2 Global Land Cover (S2GLC), contract number 4000116197.

Acknowledgments: We wish to thank CREODIAS' staff for the help with using CREODIAS platform and implementing our tools. We are also very grateful for making our products accessible via CREODIAS platform to a wide user community. Besides, we appreciate all the support provided by Espen Volden and Diego Fernandez-Prieto, from the European Space Agency, during the execution of the S2GLC project.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Radoux, J.; Lamarche, C.; Van Bogaert, E.; Bontemps, S.; Brockmann, C.; Defourny, P. Automated training sample extraction for global land cover mapping. *Remote Sens.* **2014**, *6*, 3965–3987. [[CrossRef](#)]

2. Herold, M.; Mayaux, P.; Woodcock, C.E.; Baccini, A.; Schmullius, C. Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets. *Remote Sens. Environ.* **2008**, *112*, 2538–2556. [\[CrossRef\]](#)
3. Latifovic, R.; Pouliot, D.; Olthof, I. Circa 2010 land cover of Canada: Local optimization methodology and product development. *Remote Sens.* **2017**, *9*, 1098. [\[CrossRef\]](#)
4. Bojinski, S.; Verstraete, M.; Peterson, T.C.; Richter, C.; Simmons, A.; Zemp, M. The concept of essential climate variables in support of climate research, applications, and policy. *Bull. Am. Meteorol. Soc.* **2014**, *95*, 1431–1443. [\[CrossRef\]](#)
5. Hollmann, R.; Merchant, C.J.; Saunders, R.; Downy, C.; Buchwitz, M.; Cazenave, A.; Chuvieco, E.; Defourny, P.; De Leeuw, G.; Forsberg, R.; et al. The ESA climate change initiative: Satellite data records for essential climate variables. *Bull. Am. Meteorol. Soc.* **2013**, *94*, 1541–1552. [\[CrossRef\]](#)
6. Ban, Y.; Gong, P.; Giri, C. Global land cover mapping using Earth observation satellite data: Recent progresses and challenges. *ISPRS J. Photogramm. Remote Sens.* **2015**, *103*, 1–6. [\[CrossRef\]](#)
7. Chen, J.; Chen, J.; Liao, A.; Cao, X.; Chen, L.; Chen, X.; He, C.; Han, G.; Peng, S.; Lu, M.; et al. Global land cover mapping at 30m resolution: A POK-based operational approach. *ISPRS J. Photogramm. Remote Sens.* **2015**, *103*, 7–27. [\[CrossRef\]](#)
8. DeFries, R.S.; Townshend, J.R. NDVI-derived land cover classifications at a global scale. *Int. J. Remote Sens.* **1994**, *15*, 3567–3586. [\[CrossRef\]](#)
9. Phiri, D.; Morgenroth, J. Developments in Landsat land cover classification methods: A review. *Remote Sens.* **2017**, *9*, 967. [\[CrossRef\]](#)
10. Ma, L.; Li, M.; Ma, X.; Cheng, L.; Du, P.; Liu, Y. A review of supervised object-based land-cover image classification. *ISPRS J. Photogramm. Remote Sens.* **2017**, *130*, 277–293. [\[CrossRef\]](#)
11. Wulder, M.A.; Coops, N.C.; Roy, D.P.; White, J.C.; Hermosilla, T. Land cover 2.0. *Int. J. Remote Sens.* **2018**, *39*, 4254–4284. [\[CrossRef\]](#)
12. Congalton, R.G.; Gu, J.; Yadav, K.; Thenkabail, P.; Ozdogan, M. Global land cover mapping: A review and uncertainty analysis. *Remote Sens.* **2014**, *6*, 12070–12093. [\[CrossRef\]](#)
13. Yu, L.; Liang, L.; Wang, J.; Zhao, Y.; Cheng, Q.; Hu, L.; Liu, S.; Yu, L.; Wang, X.; Zhu, P.; et al. Meta-discoveries from a synthesis of satellite-based land-cover mapping research. *Int. J. Remote Sens.* **2014**, *35*, 4573–4588. [\[CrossRef\]](#)
14. Loveland, T.R.; Zhu, Z.; Ohlen, D.O.; Brown, J.F.; Reed, B.C.; Yang, L. An analysis of the IGBP global land-cover characterization process. *Photogramm. Eng. Remote Sens.* **1999**, *65*, 1021–1032.
15. Bartholomé, E.; Belward, A.S. GLC2000: A new approach to global land cover mapping from earth observation data. *Int. J. Remote Sens.* **2005**, *26*, 1959–1977. [\[CrossRef\]](#)
16. Friedl, M.A.; McIver, D.K.; Hodges, J.C.F.; Zhang, X.Y.; Muchoney, D.; Strahler, A.H.; Woodcock, C.E.; Gopal, S.; Schneider, A.; Cooper, A.; et al. Global land cover mapping from MODIS: Algorithms and early results. *Remote Sens. Environ.* **2002**, *83*, 287–302. [\[CrossRef\]](#)
17. Arino, O.; Bicheron, P.; Achard, F.; Latham, J.; Witt, R.; Weber, J.-L. GlobCover: The most detailed portrait of Earth. *Eur. Sp. Agency Bull.* **2008**, *2008*, 24–31.
18. Hansen, M.C.; Sohlberg, R.; Defries, R.S.; Townshend, J.R.G. Global land cover classification at 1 km spatial resolution using a classification tree approach. *Int. J. Remote Sens.* **2000**, *21*, 1331–1364. [\[CrossRef\]](#)
19. Giri, C.; Pengra, B.; Long, J.; Loveland, T.R. Next generation of global land cover characterization, mapping, and monitoring. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *25*, 30–37. [\[CrossRef\]](#)
20. Gong, P.; Wang, J.; Yu, L.; Zhao, Y.; Zhao, Y.; Liang, L.; Niu, Z.; Huang, X.; Fu, H.; Liu, S.; et al. Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data. *Int. J. Remote Sens.* **2013**, *34*, 2607–2654. [\[CrossRef\]](#)
21. Yu, L.; Wang, J.; Clinton, N.; Xin, Q.; Zhong, L.; Chen, Y.; Gong, P. FROM-GC: 30 m global cropland extent derived through multisource data integration. *Int. J. Digit. Earth* **2013**, *6*, 521–533. [\[CrossRef\]](#)
22. Townshend, J.R.; Masek, J.G.; Huang, C.; Vermote, E.F.; Gao, F.; Channan, S.; Sexton, J.O.; Feng, M.; Narasimhan, R.; Kim, D.; et al. Global characterization and monitoring of forest cover using Landsat data: Opportunities and challenges. *Int. J. Digit. Earth* **2012**, *5*, 373–397. [\[CrossRef\]](#)
23. Sexton, J.O.; Song, X.-P.; Feng, M.; Noojipady, P.; Anand, A.; Huang, C.; Kim, D.-H.; Collins, K.M.; Channan, S.; DiMiceli, C.; et al. Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of

- MODIS vegetation continuous fields with lidar-based estimates of error. *Int. J. Digit. Earth* **2013**, *6*, 427–448. [CrossRef]
24. Giri, C.; Ochieng, E.; Tieszen, L.L.; Zhu, Z.; Singh, A.; Loveland, T.; Masek, J.; Duke, N. Status and distribution of mangrove forests of the world using earth observation satellite data. *Glob. Ecol. Biogeogr.* **2011**, *20*, 154–159. [CrossRef]
 25. Büttner, G. CORINE land cover and land cover change products. In *Land Use and Land Cover Mapping in Europe. Remote Sensing and Digital Image Processing*; Manakos, I., Braun, M., Eds.; Springer: Dordrecht, Germany, 2014; Volume 18, pp. 55–74.
 26. HRL. Available online: <https://land.copernicus.eu/pan-european/high-resolution-layers> (accessed on 22 April 2020).
 27. Yang, L.; Jin, S.; Danielson, P.; Homer, C.; Gass, L.; Bender, S.M.; Case, A.; Costello, C.; Dewitz, J.; Fry, J.; et al. A new generation of the United States National Land Cover Database: Requirements, research priorities, design, and implementation strategies. *ISPRS J. Photogramm. Remote Sens.* **2018**, *146*, 108–123. [CrossRef]
 28. Vogelmann, J.E.; Howard, S.M.; Yang, L.; Larson, C.R.; Wylie, B.K.; Van Driel, N. Completion of the 1990s National Land Cover Data set for the conterminous United States from Landsat thematic mapper data and ancillary data sources. *Photogramm. Eng. Remote Sens.* **2001**, *67*, 650–662.
 29. Jin, S.; Homer, C.; Yang, L.; Danielson, P.; Dewitz, J.; Li, C.; Zhu, Z.; Xian, G.; Howard, D. Overall methodology design for the United States national land cover database 2016 products. *Remote Sens.* **2019**, *11*, 2971. [CrossRef]
 30. Foody, G.M.; Arora, M.K. An evaluation of some factors affecting the accuracy of classification by an artificial neural network. *Int. J. Remote Sens.* **1997**, *18*, 799–810. [CrossRef]
 31. Gómez, C.; White, J.C.; Wulder, M.A. Optical remotely sensed time series data for land cover classification: A review. *ISPRS J. Photogramm. Remote Sens.* **2016**, *116*, 55–72. [CrossRef]
 32. Colditz, R.R.; Schmidt, M.; Conrad, C.; Hansen, M.C.; Dech, S. Land cover classification with coarse spatial resolution data to derive continuous and discrete maps for complex regions. *Remote Sens. Environ.* **2011**, *115*, 3264–3275. [CrossRef]
 33. Gong, P.; Liu, H.; Zhang, M.; Li, C.; Wang, J.; Huang, H.; Clinton, N.; Ji, L.; Li, W.; Bai, Y.; et al. Stable classification with limited sample: Transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017. *Sci. Bull.* **2019**, *64*, 370–373. [CrossRef]
 34. Li, Q.; Qiu, C.; Ma, L.; Schmitt, M.; Zhu, X.X. Mapping the land cover of Africa at 10 m resolution from multi-source remote sensing data with google earth engine. *Remote Sens.* **2020**, *12*, 602. [CrossRef]
 35. Xian, G.; Homer, C.; Fry, J. Updating the 2001 National Land Cover Database land cover classification to 2006 by using Landsat imagery change detection methods. *Remote Sens. Environ.* **2009**, *113*, 1133–1147. [CrossRef]
 36. Mùcher, S.; Steinnocher, K.; Champeaux, J.-L.; Griguolo, S.; Wester, K.; Heunks, C.; Van Katwijk, V. Establishment of a 1-km pan-European land cover database for environmental monitoring. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. ISPRS Arch.* **2000**, *33*, 702–709.
 37. S2GLC. Available online: <http://s2glc.cbk.waw.pl/> (accessed on 20 May 2020).
 38. Nowakowski, A.; Lewiński, S.; Rybicki, M.; Malinowski, R.; Gromny, E.; Krupiński, M.; Kraetzschmar, E.; Bielski, C.; Fernandez-Prieto, D. The use of low resolution databases for Sentinel-2 land cover classification. *Eur. J. Remote Sens.* **2020**. under review.
 39. CREODIAS. Available online: <https://creodias.eu/> (accessed on 21 April 2020).
 40. DIAS. Available online: <https://www.copernicus.eu/en/access-data/dias> (accessed on 21 April 2020).
 41. Gascon, F.; Bouzinac, C.; Thépaut, O.; Jung, M.; Francesconi, B.; Louis, J.; Lonjou, V.; Lafrance, B.; Massera, S.; Gaudel-Vacaresse, A.; et al. Copernicus Sentinel-2A calibration and products validation status. *Remote Sens.* **2017**, *9*, 584. [CrossRef]
 42. Peel, M.C.; Finlayson, B.L.; McMahon, T.A. Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci.* **2007**, *11*, 1633–1644. [CrossRef]
 43. Olofsson, P.; Stehman, S.V.; Woodcock, C.E.; Sulla-Menashe, D.; Sibley, A.M.; Newell, J.D.; Friedl, M.A.; Herold, M. A global land-cover validation data set, part I: Fundamental design principles. *Int. J. Remote Sens.* **2012**, *33*, 5768–5788. [CrossRef]
 44. Ettehadi Osgouei, P.; Kaya, S.; Sertel, E.; Alganci, U. Separating Built-Up Areas from Bare Land in Mediterranean Cities Using Sentinel-2A Imagery. *Remote Sens.* **2019**, *11*, 345. [CrossRef]

45. Midekisa, A.; Holl, F.; Savory, D.J.; Andrade-Pacheco, R.; Gething, P.W.; Bennett, A.; Sturrock, H.J.W. Mapping land cover change over continental Africa using Landsat and Google Earth Engine cloud computing. *PLoS ONE* **2017**, *12*, e0184926. [[CrossRef](#)]
46. Langanke, T. *Copernicus Land Monitoring Service—High Resolution Layer Imperviousness, Product Specifications Document*; Copernicus Team at EEA; EEA: Copenhagen K, Denmark, 2016.
47. Smith, G. GMES Initial Operations/Copernicus Land monitoring services—Validation of products Validation Services for the geospatial products of the Copernicus land Continental and local components HRL IMPERVIOUSNESS DEGREE 2015 VALIDATION REPORT. 2018.
48. Langanke, T. *Copernicus Land Monitoring Service—High Resolution Layer Grassland, Product Specifications Document*; Copernicus Team at EEA; EEA: Copenhagen K, Denmark, 2016.
49. Weirather, M.; Zeug, G. GMES Initial Operations/Copernicus Land monitoring services—Validation of products Validation Services for the geospatial products of the Copernicus land Continental and local components, HRL GRASSLAND 2015 VALIDATION REPORT. 2018.
50. Langanke, T.; Moran, A.; Dulleck, B.; Schleicher, C. *Copernicus Land Monitoring Service—High Resolution Layer Forest, Product Specifications Document*; Copernicus Team at EEA; EEA: Copenhagen K, Denmark, 2017.
51. Pennec, A. GMES Initial Operations/Copernicus Land monitoring services—Validation of products Validation Services for the geospatial products of the Copernicus land Continental and local components HRL FOREST 2015 FINAL VALIDATION REPORT. 2018.
52. Tachikawa, T.; Hato, M.; Kaku, M.; Iwasaki, A. Characteristics of ASTER GDEM version 2. In Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS), Vancouver, BC, Canada, 24–29 July 2011; pp. 3657–3660.
53. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
54. Belgiu, M.; Drăguț, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [[CrossRef](#)]
55. Topouzelis, K.; Psyllos, A. Oil spill feature selection and classification using decision tree forest on SAR image data. *ISPRS J. Photogramm. Remote Sens.* **2012**, *68*, 135–143. [[CrossRef](#)]
56. Du, P.; Samat, A.; Waske, B.; Liu, S.; Li, Z. Random Forest and Rotation Forest for fully polarized SAR image classification using polarimetric and spatial features. *ISPRS J. Photogramm. Remote Sens.* **2015**, *105*, 38–53. [[CrossRef](#)]
57. Gislason, P.O.; Benediktsson, J.A.; Sveinsson, J.R. Random Forests for land cover classification. *Pattern Recognit. Lett.* **2006**, *27*, 294–300. [[CrossRef](#)]
58. Pflugmacher, D.; Rabe, A.; Peters, M.; Hostert, P. Mapping pan-European land cover using Landsat spectral-temporal metrics and the European LUCAS survey. *Remote Sens. Environ.* **2019**, *221*, 583–595. [[CrossRef](#)]
59. Inglada, J.; Vincent, A.; Arias, M.; Tardy, B.; Morin, D.; Rodes, I. Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series. *Remote Sens.* **2017**, *9*, 95. [[CrossRef](#)]
60. Estoque, R.C.; Murayama, Y. Classification and change detection of built-up lands from Landsat-7 ETM+ and Landsat-8 OLI/TIRS imageries: A comparative assessment of various spectral indices. *Ecol. Indic.* **2015**, *56*, 205–217. [[CrossRef](#)]
61. Dash, J.; Mathur, A.; Foody, G.M.; Curran, P.J.; Chipman, J.W.; Lillesand, T.M. Land cover classification using multi-temporal MERIS vegetation indices. *Int. J. Remote Sens.* **2007**, *28*, 1137–1159. [[CrossRef](#)]
62. Lewiński, S.; Nowakowski, A.; Malinowski, R.; Rybicki, M.; Kukawska, E.; Krupiński, M. Aggregation of Sentinel-2 Time Series Classifications as a Solution for Multitemporal Analysis. In Proceedings of the SPIE 10427, Image and Signal Processing for Remote Sensing XXIII, 104270B. Warsaw, Poland, 4 October 2017.
63. Polikar, R. Ensemble based systems in decision making. *IEEE Circuits Syst. Mag.* **2006**, *6*, 21–45. [[CrossRef](#)]
64. Kukawska, E.; Lewinski, S.; Krupinski, M.; Malinowski, R.; Nowakowski, A.; Rybicki, M.; Kotarba, A. Multitemporal Sentinel-2 data—Remarks and observations. In Proceedings of the 2017 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), Brugge, Belgium, 27–29 June 2017; pp. 2–5.
65. Baetens, L.; Desjardins, C.; Hagolle, O. Validation of copernicus Sentinel-2 cloud masks obtained from MAJA, Sen2Cor, and FMask processors using reference cloud masks generated with a supervised active learning procedure. *Remote Sens.* **2019**, *11*, 433. [[CrossRef](#)]

66. Mack, B.; Leinenkugel, P.; Kuenzer, C.; Dech, S. A semi-automated approach for the generation of a new land use and land cover product for Germany based on Landsat time-series and Lucas in-situ data. *Remote Sens. Lett.* **2017**, *8*, 244–253. [[CrossRef](#)]
67. Heymann, Y.; Steenmans, C.; Croisille, G.; Bossard, M.; Lenco, M.; Wyatt, B.; Weber, J.-L.; O'Brian, C.; Cornaert, M.-H.; Sifakis, N. CORINE Land Cover: Technical Guide. In *Environment, Nuclear Safety and Civil Protection Series*; Commission of the European Communities, Office for Official Publications of the European Communities: Luxembourg, 1994; p. 144.
68. Paul, F.; Kääb, A.; Haeberli, W. Recent glacier changes in the Alps observed by satellite: Consequences for future monitoring strategies. *Glob. Planet. Chang.* **2007**, *56*, 111–122. [[CrossRef](#)]
69. Li, Y.-J.; Ding, Y.-J.; Shangguan, D.-H.; Wang, R.-J. Regional differences in global glacier retreat from 1980 to 2015. *Adv. Clim. Chang. Res.* **2019**, *10*, 203–213. [[CrossRef](#)]
70. Rodriguez-Galiano, V.F.; Ghimire, B.; Rogan, J.; Chica-Olmo, M.; Rigol-Sanchez, J.P. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens.* **2012**, *67*, 93–104. [[CrossRef](#)]
71. Yu, L.; Wang, J.; Gong, P. Improving 30 m global land-cover map FROM-GLC with time series MODIS and auxiliary data sets: A segmentation-based approach. *Int. J. Remote Sens.* **2013**, *34*, 5851–5867. [[CrossRef](#)]
72. Fernandes, R.; Fraser, R.; Latifovic, R.; Cihlar, J.; Beaubien, J.; Du, Y. Approaches to fractional land cover and continuous field mapping: A comparative assessment over the BOREAS study region. *Remote Sens. Environ.* **2004**, *89*, 234–251. [[CrossRef](#)]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).