

Annual satellite-based NDVI-derived land cover of Europe for 2001–2019

Verhoeven, Vincent B.; Dedoussi, Irene C.

DOI

[10.1016/j.jenvman.2021.113917](https://doi.org/10.1016/j.jenvman.2021.113917)

Publication date

2022

Document Version

Final published version

Published in

Journal of Environmental Management

Citation (APA)

Verhoeven, V. B., & Dedoussi, I. C. (2022). Annual satellite-based NDVI-derived land cover of Europe for 2001–2019. *Journal of Environmental Management*, 302, Article 113917. <https://doi.org/10.1016/j.jenvman.2021.113917>

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.



Annual satellite-based NDVI-derived land cover of Europe for 2001–2019

Vincent B. Verhoeven, Irene C. Dedoussi*

Aircraft Noise and Climate Effects Section, Faculty of Aerospace Engineering, Delft University of Technology, 2629 HS, Delft, the Netherlands

ARTICLE INFO

Keywords:

NDVI
Land cover classification
Classification trees
Europe
MODIS
CORINE

ABSTRACT

Land cover plays an important role in the Earth's climate as it affects multiple biochemical cycles and is critical for food security and biodiversity. As land cover is continuously evolving, influenced by anthropogenic and other factors, the availability of temporally varying land cover data sets of large spatial domains is integral to understanding, monitoring, and informing environmental management efforts. Here we use classification trees to generate annual land cover maps of the European continent for 2001 to 2019 on a ~250 m resolution. The classification trees are trained using gap-filled and smoothed MODIS normalised difference vegetation index (NDVI) satellite data, as well as CORINE reference land cover data. We apply the bagging ensemble technique on oversampled NDVI data, with an additional majority vote for overlapping segments over the continent-wide domain. We distinguish between 39 land cover classes, with a total classification accuracy of 75% and average precision of 76%. The accuracy varies between the classes, with common classes (e.g. agricultural and forest classes) performing better than rarer ones (e.g. artificial land cover). Over the entire continent, we find that artificial land cover, wetlands, and forests have increased on average by 0.76, 0.50 and 0.22%/year respectively, while the agricultural area has decreased by 0.21%/year. We also quantify these changes in land cover on a national and metropolitan level. Given the near-real-time availability of global NDVI data, we note the potential of the presented approach for generating 'near-real-year' annual land cover data sets of large geographic domains, for the continuous monitoring of land cover change and the effects of interventions.

1. Introduction

Land cover and its dynamics play a key role in multiple global challenges. First, land cover is an integral part of the Earth system. Land cover, and changes therein, affect the global carbon cycle through the surface albedo, heat flux and atmospheric moisture (IPCC, 2019). Earth system models, employed to further our understanding of the evolving Earth system by numerically simulating the Earth's climate, require detailed land cover data among other inputs (Jeffries et al., 2015). Although progress has been made in improving our understanding of the interaction between land cover and the Earth system, land cover and land cover change still remain a significant source of uncertainty in the global carbon cycle (Houghton et al., 2012). Besides its importance in understanding the Earth system, land cover is one of the most widely used indicators for measuring pressure on ecosystems and biodiversity (OECD, 2018). Finally, land cover and associated dynamics are tightly related to the ability to produce the necessary food to sustain the global population.

Land cover evolves heterogeneously over time, depending on the

land cover type as well as location. It is estimated that only 13% of the global wetlands area that existed in 1700 remains today (Davidson, 2014). Furthermore, since the 1960s the global area used for cropland has increased by 15% and the area used for pastures by 8% to support the world's growing population (IPCC, 2019). Similarly, the global urban surface area has increased over the last few decades and is expected to triple between 2000 and 2030 (Martellozzo et al., 2015; d'Amour et al., 2017). The global area covered by forests has also declined by 3% between 1990 and 2015 and continues to decrease, with the largest losses in Brazil, Indonesia and Nigeria. In Europe, the forest area has steadily increased however (Keenan et al., 2015). Further land cover change is caused by desertification, which, despite national and international efforts, is continuing at accelerating rates (Abahussain et al., 2002).

To better monitor and understand the evolution of land cover over time and the effects that various natural and anthropogenic interventions have on it, land cover data sets of high temporal resolution and wide geographic spans are needed (An and Brown, 2008; Mena, 2008). A variety of data sets exist that provide fine resolution (~30 m)

* Corresponding author.

E-mail address: i.c.dedoussi@tudelft.nl (I.C. Dedoussi).<https://doi.org/10.1016/j.jenvman.2021.113917>

Received 18 October 2020; Received in revised form 22 September 2021; Accepted 6 October 2021

Available online 23 October 2021

0301-4797/© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

annual land cover maps, but they are restricted to small spatial domains (Mena, 2008; Braimoh and Vlek, 2005; Vågen, 2006). Besides the inherently large geographic spans of Earth system modelling, determining the land cover over a large geographical area also enables us to determine whether spatial patterns can be found in the land cover and its change. Additionally, land cover transitions affect the surrounding areas of where they occur (Mahmood et al., 2014). As such, the causes and effects of changes in land cover may be overlooked when a restricted area is investigated.

While continent-spanning, high resolution (10–100 m) land cover data sets exist, they lack a fine temporal resolution. For example, the CORINE database exists for 2000, 2006, 2012 and 2018 (European Environment Agency, 2017), Copernicus for 2015 (Buchhorn et al., 2019) and S2GLC for 2017 (Malinowski et al., 2019). As land cover change is often a nonlinear process (Braimoh and Vlek, 2005; Vågen, 2006), it cannot necessarily be interpolated or extrapolated using such existing land cover maps. Given the methodological and input data differences between the aforementioned land cover products, using multiple existing products to gain temporal resolution or spatial coverage would introduce uncertainties.

In this work we produce continent-spanning time series of land cover at continental spatial scales. To do this, we use remotely sensed normalised difference vegetation index (NDVI) data, for which global data are available spanning several decades (Kerr and Ostrovsky, 2003; Xie et al., 2008). NDVI data have previously been used to characterise the land cover of constrained spatial domains (Geerken et al., 2005; Pacheco et al., 2014; Lunetta et al., 2006; de Bie et al., 2011; Usman et al., 2015; Sheffield et al., 2015; Ali et al., 2013) as well as larger domains (De Fries and Townshend, 1994; Hansen et al., 2013; Estel et al., 2016) but with low temporal resolution. In some cases, these data products make use of multiple input data sets (e.g. multiple spectral indices). Here however, we classify land cover using solely NDVI data.

We produce and analyse annual land cover results of a large geographical area, specifically the European continent. The distribution of land cover and changes therein are quantified and discussed on several spatial levels, indicating the suitability of this approach for monitoring and understanding the implications of environmental management efforts. While there are both prior examples of the classification of a large region as well as prior classification studies producing annual results, to the best of our knowledge this work presents the first comprehensive analysis of recent annual land cover change over the European continent.

2. Data and study area

2.1. Study area

We classify the land cover of the continent of Europe. The spatial domain extends from 11°W to 32°E longitude, and 34.5°N to 71.5°N latitude. It contains a heterogeneous climate, including arid (e.g. Iberian peninsula), temperate (e.g. Mediterranean and western Europe), cold (e.g. eastern and northern Europe), and polar (parts of northern Europe) Köppen-Geiger climate zones (Beck et al., 2018). Within the domain, different landscape types exist. Examples are the Tabernas desert in southern Spain, the wetlands found in Scotland, the forests of Scandinavia, the grasslands of Great Britain, and the urban areas of various sizes predominantly located in western Europe.

2.2. NDVI data

We use Moderate Resolution Imaging Spectroradiometer (MODIS) MOD13Q1 NDVI data. In order to quantify long-term fluctuations in land cover, nineteen years of data are analysed between 2001 and 2019. The data sets are available every sixteen days, at a spatial resolution of 232 m × 232 m (Didan et al., 2015).

Remotely-sensed vegetation data contain inherent noise as the

Earth's surface and atmosphere obfuscate present patterns. The MODIS NDVI data product used here has been corrected for the atmospheric conditions as well as the different viewing angles (Brown et al., 2006; Didan et al., 2015). The quality of each NDVI pixel is indicated by its pixel reliability index. Of these, the Fill/No Data, Snow/Ice and Cloudy indices were considered invalid. Any such invalid values in each pixel's (grid-cell's) time series are gap-filled using weighted temporal interpolation, with a symmetrical window width of five timestamps.

We use the iterative Savitzky-Golay filter to further reduce the noise (Chen et al., 2004). The Savitzky-Golay filter is a moving average filter with weighting given as a polynomial. These weight coefficients when applied to a signal, perform a least-squares fit within the filter window (Savitzky and Golay, 1964). In this iterative approach, an initial filter is applied followed by the iterative filter which favours positive outliers (Chen et al., 2004). This enables the algorithm to compensate for the negative bias within the data set which is a result of atmospheric interference (Gutman, 1991; Jönsson and Eklundh, 2004). We use a window width corresponding to 208 days and a polynomial degree of 2 for both the initial, as well as the iterative Savitzky-Golay filter. An example of the initial noise within the data set, as well as the effects of interpolation and the iterative Savitzky-Golay filter are shown in Fig. 1.

2.3. Reference land cover data

In order to both train the supervised classification methodology and assess its effectiveness, reference land cover data are required. To this end, we use the openly available CORINE Land Cover database, provided by the Copernicus Global Land Service (European Environment Agency, 2019). This land cover database has been generated using high resolution satellite imagery from different satellite sensors and spectral ranges. NDVI data are not directly used in CORINE, but the near infrared and red spectra included in the NDVI definition are used to generate some of the CORINE data set years (European Environment Agency, 2002, 2007, 2014, 2017). The CORINE Land Cover database has been verified using expert assessment of aerial photographs, and has an overall classification accuracy of 85% (Büttner and Maucha, 2006).

CORINE distinguishes between 44 classes in total. These are grouped as subclasses of the artificial, agricultural, forest and semi-natural areas, as well as wetlands, and water bodies classes. CORINE is at a 100 m × 100 m resolution. We remap and regrid the CORINE dataset to match the projection and resolution of the ~ 250 m MODIS NDVI data used in this work. This is done by selecting the most common Corine pixel within the spatial bounds of each respective pixel of the MODIS data. The CORINE database is produced every six years, and four CORINE data sets are used here (2000, 2006, 2012 and 2018).

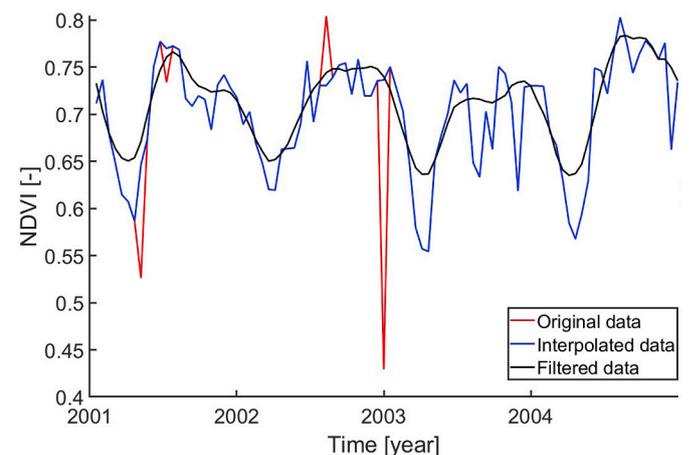


Fig. 1. An example time series of four years showing the effect of the pre-processing steps. Interpolation removed four invalid pixels, and the Savitzky-Golay filter smoothed the resulting time series.

Finally, we note that the CORINE reference land cover database does not cover the full spatial domain. This further illustrates the lack of universal land cover data, which together with the fact that these data are produced only once every six years, and with a lag of several years, motivates the use of continuously available NDVI data for land cover classification.

3. Methods

Fig. 2 presents an overview of the methods used for obtaining annual land cover data, using the aforementioned interpolated and smoothed NDVI time series and the four years of CORINE reference data.

3.1. Types of land cover classes and their classification profiles

Supervised classification methods require the definition of classification profiles that are representative of each pixel's time series. We use the following NDVI and normalised growing season metrics to characterise each pixel (Martínez and Gilabert, 2009; Reed et al., 1994):

- **NDVI metrics:** The mean, maximum and minimum, as well as the amplitude of the NDVI time series. The time at which the maximum occurs is also used, as well as the slope at the start and end of the growing season (rates of greenup and senescence respectively).
- **Growing season metrics:** The start, end and length of the growing season, as well as the area under the NDVI time series during the growing season (a proxy for the gross primary productivity) (Park et al., 2016).

To determine the growing season metrics we use the midpoint_{pixel} method, due to its good performance compared to other methods (White et al., 2009). The start of the growing season is set as the first time the average between the minimum and maximum annual NDVI is exceeded, with the end of the growing season being the last (White et al., 2009). Additionally, the start of the growing season has to occur prior to the annual NDVI maximum, and the slope has to be positive. The opposite criteria are used for the end of the growing season. This alleviates the issue that the growing season may be reversed around the Mediterranean (Estel et al., 2015). Similar to the other feature data, the growing season metrics are determined on an annual basis for each pixel's time series individually.

To further characterise each pixel, textural features are used to describe their spatial distribution and thus the landscape they represent. The textural features are computed using the mean NDVI as per Haralick et al. (1973). The homogeneity, contrast, correlation and angular second moment metrics were selected due to their prior use for land cover classification (Rodríguez-Galiano et al., 2012; Beekhuizen and Clarke, 2010). A square window of 15×15 pixels ($\sim 4 \text{ km} \times \sim 4 \text{ km}$) was found

to be optimal.

Finally, while CORINE has 44 classes, here we merge the water bodies classes. This was deemed necessary, as the MODIS NDVI data use a single value to indicate water pixels, and is thus not suited to distinguish between different types of water bodies. We further divide the mixed forest class between the coniferous and broad-leaved classes to avoid the mixed-pixel problem (Carreiras et al., 2006).

3.2. Classification trees

We use classification trees to generate a classifier that is applied to the NDVI data. Classification trees split the data into ever-more homogeneous groups. They start out at the 'root', which contains all the data. The data are then divided into 'branches', to create more homogeneous groups. Each division is determined with a decision stage, where the data either meet the given threshold for the specific feature or not. The 'leaves' are located at the top of the tree, and give the pixel its classification type. Decision stages represent an optimisation problem where the objective is to minimise impurity, or, in other words, maximise the separability (Safavian and Landgrebe, 1991).

This optimisation process favours common classes over rarer ones. For the European domain as a whole, the artificial, wetlands, and non-vegetated classes are usually rare compared to the agricultural, and grasslands and forests classes. To overcome the issue of rarity, we apply sampling techniques to adjust the distribution of the classes within the training data. When the training data distribution is different from the actual data distribution, the classifier becomes biased. To prevent this, the prior probabilities must be changed accordingly (Weiss and Provost, 2003). For this purpose, Laplace's law of succession is used, as it lacks the asymptotic behaviour for small sample sizes that the frequency based alternative has (Weiss and Provost, 2003). The fully balanced distribution was found to be the most suitable in this work. Four sampling techniques were considered: oversampling, undersampling and the synthetic sampling strategies SMOTE (Chawla et al., 2002) and Borderline-SMOTE (Han et al., 2005).

Improvements to the classifier itself were also considered in the form of the ensemble techniques random forests, bagging and boosting. Ensemble techniques work by using the results from multiple classifiers, an ensemble, to arrive at the final result. The error of a single classification is outweighed by the group through majority voting, where the mode of the classification results for a pixel is selected. Ensemble methods suffer less from over-fitting and thus enable the use of larger trees (Ghimire et al., 2012). This enables the classifier to also consider rarer cases, which could otherwise be ignored. All combinations between the aforementioned sampling and ensemble techniques were assessed, and the combination of oversampling with bagging was found to be the most optimal.

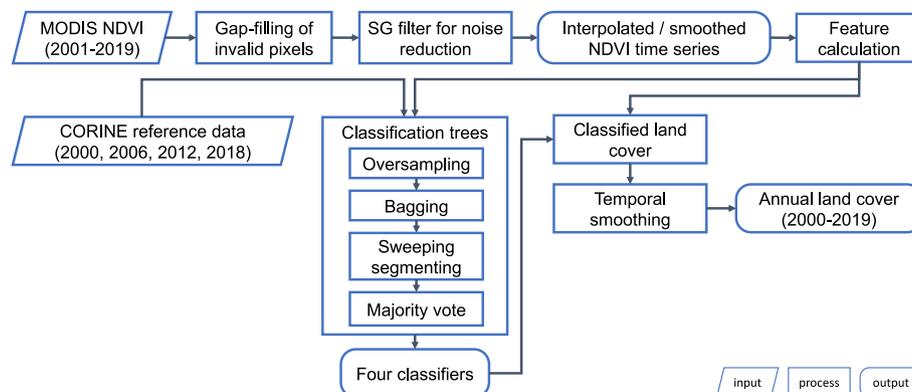


Fig. 2. Overview of the approach for producing the annual land cover, including all inputs, intermediate processes and (intermediate) outputs. The performance of the classified land cover is assessed as described in Section 3.4.

3.3. Segmenting the data

Large heterogeneous landscapes are challenging to classify as they have low inter-class separability and high intra-class variability (Ghimire et al., 2012). This makes it difficult to accurately classify a large geographic area, such as Europe, or even a smaller domain, such as Iberia. Aside from the water bodies class, the classes are difficult to separate and the values are generally lower than those found in literature (Shao et al., 2016). We overcome the low inter-class separability and high intra-class variability in the continent spanning NDVI data set by segmenting the spatial domain in 60×60 pixel sections and classifying each section at a time. This is selected as a trade-off between the ability to more accurately classify rarer classes, and classification bias and computational cost. The resulting improvements in the separability and variability are discussed in section C1 of the Supplementary Information (SI).

The use of classification trees yields inconsistency in the produced land cover type. To increase the classification consistency, a second layer of majority voting is introduced. Multiple iterations of the classification procedure described thus far are performed, with the most common classification (mode) for each pixel returned as the final result. The similarity between successive levels of majority voting approaches the asymptote of 100% in five to ten iterations.

Additionally, discontinuous classification results could be obtained when the data are segmented, as each segment is trained on entirely separate data. To ensure spatial consistency, as well as introduce the aforementioned second layer of majority voting, segmenting is applied in a sweeping manner horizontally as well as vertically (see Figure SI.1 in the SI). This smooths out resulting edges between sections, and, as these segments are layered on top of one another, this inherently introduces a degree of iteration. An overlap of 80% in both dimensions is used, resulting in a conservative ten iterations to ensure consistency.

3.4. Annual classification results

The CORINE land cover data used to train the classifier are only available for four years within the assessed nineteen year time window (European Environment Agency, 2019). As a result, four classifiers are produced. To produce land cover data at an annual temporal resolution, the classification trees generated using the nearest (in time) CORINE data set are applied to the corresponding year's NDVI data. The effects of this on the classification results are relatively minor (discussed in section C2 in the SI), indicating that the classification algorithm is robust to evolving feature data.

The classification algorithm can occasionally yield anomalous results, necessitating temporal smoothing. In this temporal smoothing, a classification is seen as erroneous if the classification land cover does not cover at least three successive years, and is not the majority land cover type within the twenty year window. The erroneous classification results are substituted for the most common class in a symmetrical window of five years for each pixel's land cover trajectory. If after this procedure any outliers remain, they are substituted for by the majority class within the entire time span. The effects of this temporal smoothing are discussed further in section C4 in the SI.

We assess the performance of the annual classification results by quantifying the accuracy and precision of the classified years for which reference land cover data is available. We use the ratio of correct predictions (true positives and true negatives) to total predictions to assess the classification accuracy, and the ratio of true positives to total predicted positives (true positives and false positives) to assess the classification precision. Besides CORINE, we also use the Copernicus land cover map data set for an additional classification performance assessment.

4. Results

4.1. Classification performance

The land cover of Europe in 2018 is given in Fig. 3, both as determined by the classifier (a) as well as the CORINE reference data set (b). The classified land cover captures all major patterns given by the CORINE land cover data, including all large urban and natural areas. Results differ where small areas are covered by a land cover type that is distinct from its surroundings. These are often no bigger than a pixel and the likely explanation for this is that these distinct areas cannot provide the classifier with sufficient training data for our 60×60 pixel segment size.

Fig. 4 presents the accuracy and precision of the classifier for the 39 classes, resulting from the test data used for the entire domain for the four years for which reference CORINE land cover data are available (2000, 2006, 2012, 2018). The full confusion matrix is shown in section B1 in the SI. Whilst differences exist between the obtained classification performance for the years for which this can be tested, they are small and do not affect the conclusions drawn from them.

The total classification accuracy and average precision are 75% and 76% respectively. This is comparable to the 74% accuracy of the MODIS land cover product at a $500 \text{ m} \times 500 \text{ m}$ resolution, however it should be noted that this product distinguishes between 16 rather than 39 classes (Sulla-Menashe et al., 2019). When our results are aggregated to the second tier of CORINE land cover classes with a more comparable 14 classes, an over-all accuracy of 79% is found. For the five tier 1 classes, this is further improved to 89%.

The highest accuracy is found for the most common agricultural classes, for instance non-irrigated (89%) and permanently irrigated (81%) agriculture, and natural classes, such as broad-leaved (78%) and coniferous (85%) forests. The classification accuracy for many of the artificial built-up classes is relatively low, ranging between 9% and 61%. The road/rail class performs the worst with an accuracy of 9%, likely due to the relatively large pixel resolution of $\sim 250 \text{ m}$ exceeding the general size patterns of this land cover type.

Fig. 4 further shows that for the rarer classes, the precision is considerably higher than the accuracy. This indicates that the classifier is conservative when allocating the rarer classes. It also explains the classifier's reduced ability to replicate individual artificial class pixels that are present in the reference land cover, but not the classified land cover shown in Fig. 3.

The results produced by the classifier for 2015 were also compared to the Copernicus land cover data set of 2015 (Buchhorn et al., 2019). When compared with the Copernicus land cover data set, the over-all accuracy is 71% and the average precision is 77%, excluding the water bodies class. The classification accuracy is high for the agricultural (89%) and water bodies classes (99%), with a lower accuracy of 69% for the forests class and 66% for the artificial class. The performance is worse for non/sparsely vegetated land cover, wetlands and shrubs and herbaceous vegetation (<32%), likely due to the discrepancies between the CORINE and Copernicus land cover maps for these classes. We provide further information and quantify the discrepancies between the CORINE and the Copernicus land cover data sets in section B2 in the SI.

4.2. Land cover changes

Determining the number of land cover classes to analyse is a trade-off between fidelity and accuracy. On the one hand, a useful portrayal of the ground cover diversity warrants a large number of classes. On the other hand, this reduces the accuracy of the produced land cover map, and makes the results more complex to analyse. As such, while we classify 39 land cover types, we choose to aggregate these into the following 7 classes for calculating the temporal trends: artificial (urban and built-up areas), agricultural, forests, shrubs and herbaceous vegetation, non-

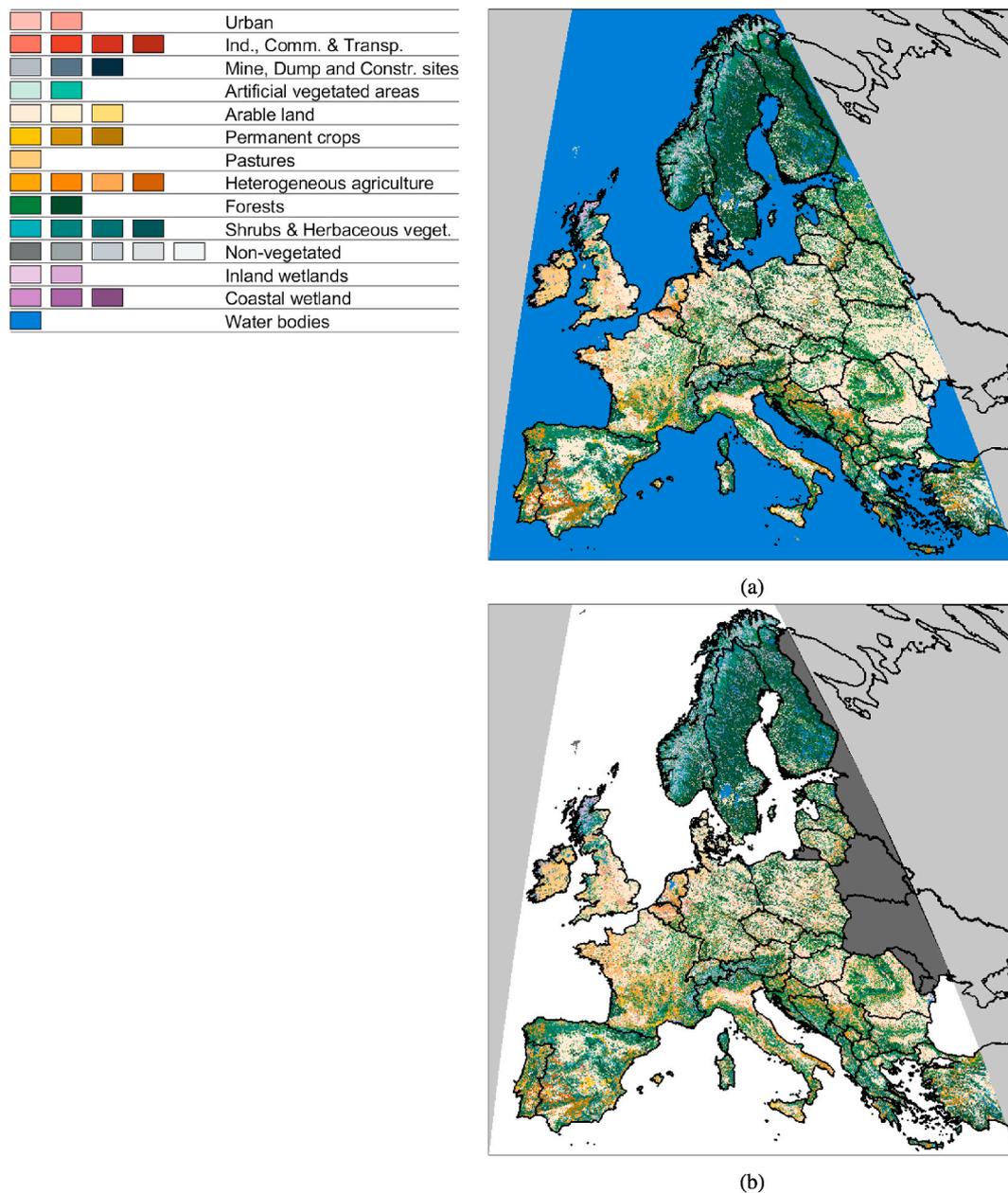


Fig. 3. Comparison between the (a) classified and the (b) reference land cover (CORINE) of the European domain in 2018. Pixels that are missing from the CORINE data are shown in white (water) and dark grey (land). Land outside of the domain is shown in light grey.

vegetated areas and wetlands. The over-all classification accuracy for this level of class fidelity is 85%. As the area covered by water bodies is near-constant, it is not examined further. For the exact grouping of these classes, the reader is referred to section D in the SI.

4.2.1. The land cover of the European continent

Table 1 presents the distribution of the land cover classes for the European continent in 2001 and 2019. Overall, we find that the distribution of classes is imbalanced, and that absolute changes in land cover are minor over this time period ($<40,000 \text{ km}^2$). Specifically, we find the greatest relative increase (5.9%) in artificial land cover, in line with European urban sprawl figures (Hennig et al., 2015, 2016). Meanwhile the agricultural class decreased in size by 1.3%, and is the only class to show a decrease in surface area. The surface area occupied by more natural vegetation increased in size, albeit by less than 1%. The forested area increased by 0.9% and the shrubs and herbaceous vegetation class shows an increase of 0.5%. The non-vegetated class also increased in size

by 0.13%. This is a minor change, but has previously been associated with the increase in global temperature (Martínez and Gilabert, 2009). The wetlands class increased in size by 5.0%, deviating from the negative rate of change previously reported for Europe (Davidson, 2014). We note that a high relative increase in the wetlands class was found originally due to the CORINE data set for the year 2000 not labelling the wetlands of northern Scotland, contrary to the data sets of 2006, 2012 and 2018. To overcome this, the classification results of Great Britain for the years 2001–2003 were substituted for those obtained for 2004.

The time series of these classes are shown in Fig. 5, along with the rates of change as determined by the Theil-Sen estimator, w.r.t. the size of the class itself. These time series show that, whilst only agricultural land cover decreased between 2001 and 2019 according to Table 1, the shrubs and herbaceous vegetation class also shows a negative trend ($-0.287\%/ \text{year}$). The forest trend shows the opposite behaviour of the agricultural class ($0.217\%/ \text{year}$). This loss rate closely resembles the $0.3\%/ \text{year}$ determined previously by the EEA (Büttner et al., 2019). The

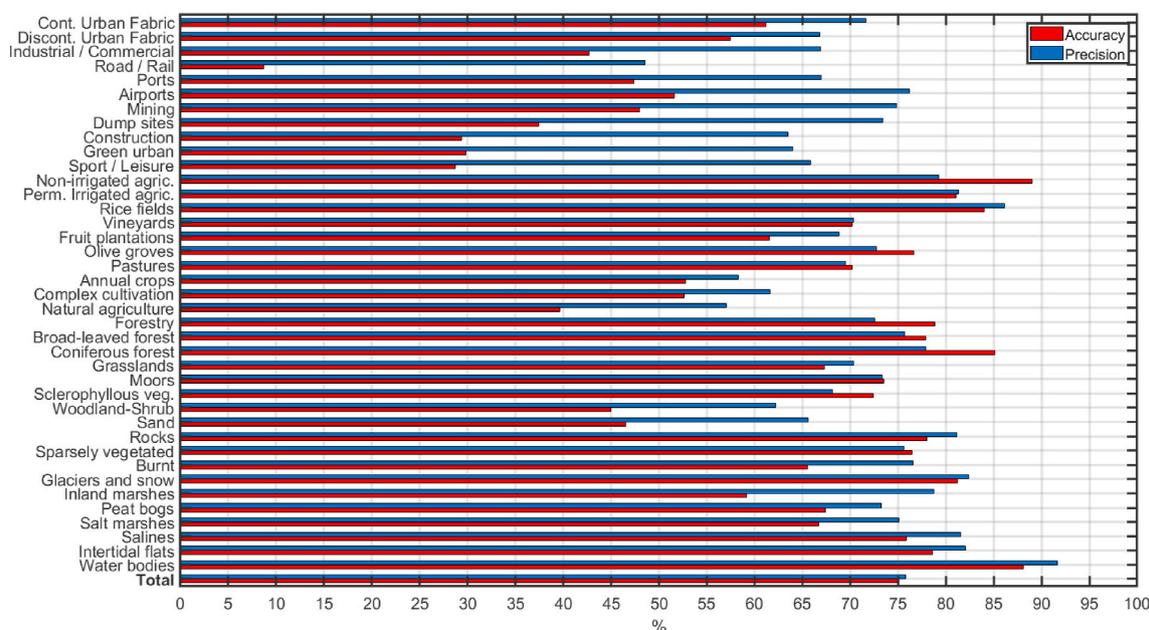


Fig. 4. The accuracy (red, bottom bar) and precision (blue, top bar) of the classifier for each class. Determined using the test data from 2000, 2006, 2012 and 2018.

Table 1

The distribution of land cover for the European continent in 2001 and 2019, including the relative change in class size as well as the absolute change in surface area between 2001 and 2019.

	Distribution 2001 [%]	Distribution 2019 [%]	Δ Class size [%] (2001–2019)	Δ Surface area [km ²] (2001–2019)
Artificial	2.83	3.00	5.90	9536
Agricultural	48.03	47.41	-1.30	-38,153
Forests	36.21	36.54	0.91	17,428
Shrubs & Herb. Vegetation	8.17	8.21	0.50	1971
Non-Vegetated	3.14	3.14	0.13	93
Wetlands	1.62	1.70	5.04	4652

wetlands and artificial classes increased most rapidly in size (0.498 and 0.763%/year respectively) and the latter’s increase is to be expected due to urban sprawl. Finally, the rate of change in non-vegetated land cover is near-zero (0.0165%/year). We note that the aggregated time series deviate from their respective trend lines, indicating that the intra-year variability is greater than the computed rates of change.

4.2.2. The land cover of European nations

We also characterise the land cover of the nations within the domain, along with changes in its distribution. The difference between the national land cover distribution in 2001 and 2019, relative to class size, is presented in Fig. 6. It should be noted that these are snapshots in time, and as Fig. 5 shows are thus not necessarily representative of the overall temporal trend. Whilst the common agricultural and forests classes show only moderate levels of change (average absolute change <5%). For the rarer classes significant change is found (average absolute change 15–45%). The obtained class distribution for the first year of the data set, 2001, as well as the changes for each nation are given in section E in the SI.

Nearly all countries within the domain show an increase in artificial land cover, with the greatest relative increase in south-eastern Europe (24%). These findings are in agreement with those reported by the European Environment Agency and the Federal Office for the Environment (Hennig et al., 2016). Furthermore, a high degree of urbanisation in

many former Eastern Bloc countries has also been previously determined (Radeloff and Gutman, 2017), although Romania, Bulgaria and Belarus show a decrease. For the latter, it is important to note that the classifier did not always classify urban areas not covered by training data correctly.

The agricultural class has changed less over the twenty year period, with small changes throughout the domain, with the exception of Russia and Latvia. Abandonment of farmland occurred in Eastern Europe after the dissolution of the Soviet Union, however the rate of recultivation in Eastern Europe is greater than the rate of abandonment (Estel et al., 2015; Radeloff and Gutman, 2017). Nevertheless, a slight negative change is determined for some of the countries. They further state that farmland abandonment has occurred in the mountainous regions of Europe, which is confirmed by the Alpine countries, however not for Norway.

Many of the north-western countries of Europe show increases in the area occupied by forests (5%), with the opposite in much of south-eastern Europe (-3%). Previous studies report that that the logging industries of the Baltic countries and Romania grew to unsustainable levels, and a decrease in forest cover is found for Romania (-2%) but not the Baltic countries (Radeloff and Gutman, 2017). In Portugal, a relatively large amount of forests was replaced with agriculture over the twenty year period, and the loss agrees with the numbers found by Hansen et al. (2013).

On average, the area covered by the shrubs and herbaceous vegetation types is near-constant around the Mediterranean (-0.2%), although it should be noted that this change is not uniform. In Scandinavia, there is a significant increase in Denmark (19%) and decrease in Finland (-28%). In Finland, the shrubs and herbaceous vegetation was replaced by forests.

Whilst few countries contain a non-negligible amount of non-vegetated land cover, the countries with the greatest proportion covered by mountains show a very slight increase (0.5%). Land degradation due to soil erosion has also been predicted for the Mediterranean and mountainous regions due to rainfall erosivity (Panagos et al., 2015). The Tabernas desert in south-eastern Spain slowly transformed into forests and grasslands in the early 2000’s according to the classifier and the CORINE land cover data, resulting in the large reduction in non-vegetated land cover in Spain.

The British Isles contain the greatest proportion of wetlands. Ireland

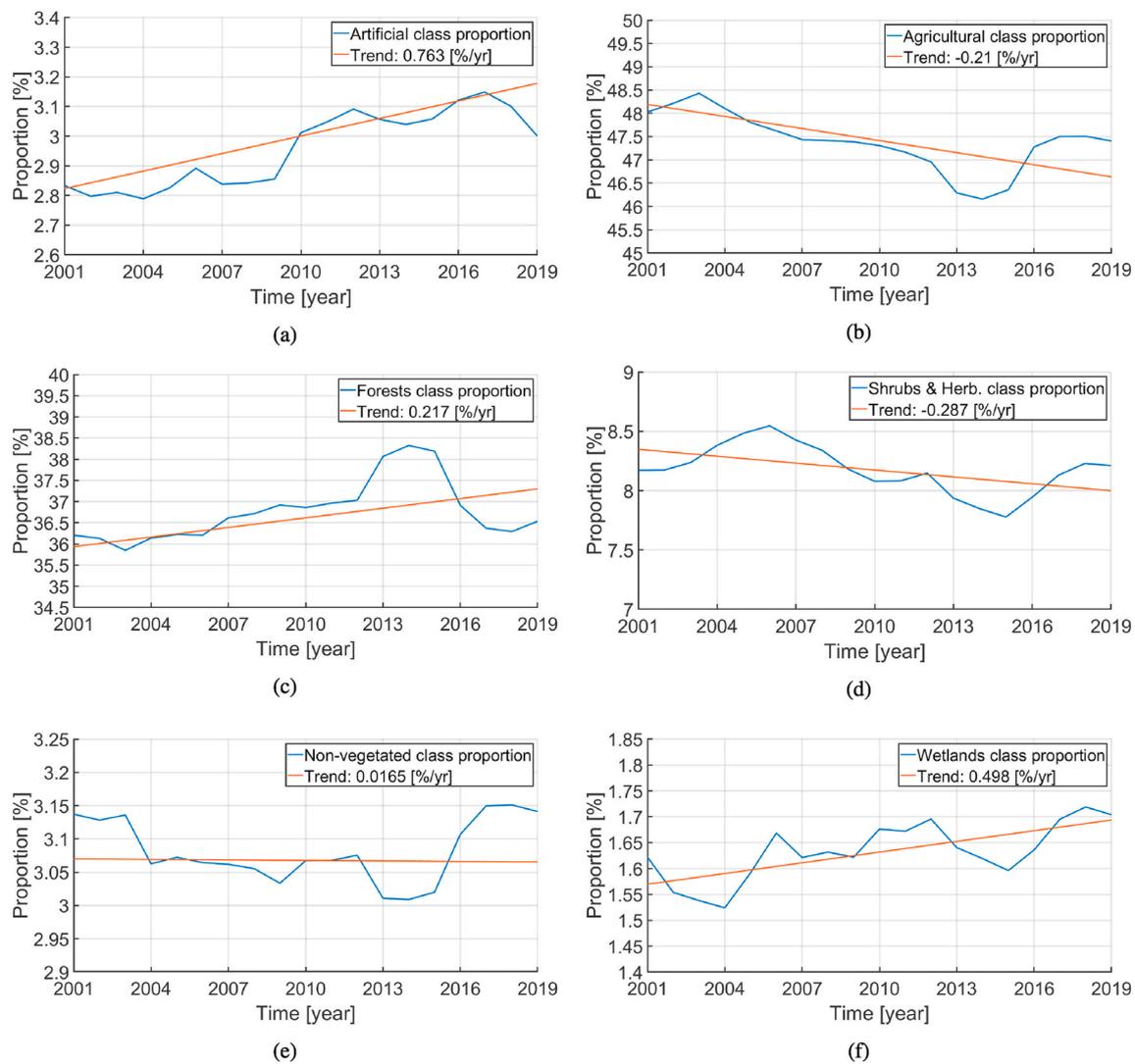


Fig. 5. Time series of the (a) artificial, (b) agricultural, (c) non-vegetated, (d) shrubs & herbaceous vegetation, (e) non-vegetated and (f) wetlands class distributions (blue), along with the computed rates of change (red). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

shows a moderate decrease (-8%) over the time period, whilst Great Britain shows an increase (11%). In northern Europe, the average change in wetlands is 2% , although not uniform in its sign.

4.2.3. The land cover of European cities

We also quantify the rates of change in land cover for the ten most populated cities within the domain covered by CORINE, using the Theil-Sen estimator. These are summarised in Table 2. The city domains are chosen such that they encompass the city and its surrounding regions, and vary drastically in size and land cover distribution. As such, only the rates of change relative to the size of the class itself are given in Table 2. The selected city domains are depicted in section F in the SI. Due to the different nature of these subsections of the domain, the artificial land cover is examined in more detail, resulting in the set of land cover classes listed in section D in the SI. While the accuracy of these classes is relatively low for the entire domain (30% – 59%), it is higher for these subsets due to its commonality there. The shrubs and herbaceous vegetation, non-vegetated and wetlands classes are not included, as many of the assessed regions do not contain them in sufficient number.

The area occupied by urban land cover (defined as the continuous and discontinuous urban fabric) decreased for the metropolitan areas of London, Madrid, Vienna and Hamburg (average of $-0.33\%/year$). The

same cities are the only ones that show an increase in agricultural land cover over this time period (average of $0.20\%/year$), and all apart from London show an increase in forested area (average of $0.40\%/year$). We note that here we quantify other urban aspects in separate classes. For example, we find the land covered by industrial, commercial and transportation areas has increased for all cities apart from Vienna (-0.38%), and the average rate of increase is 1.5% . Of the classes considered here, the mine, dump and construction sites class is the rarest, and only Istanbul contains more than 0.5% of this land cover type. Finally, the artificially vegetated area, containing amongst other aspects parks, shows an increasing rate for all cities apart from Hamburg and London, with an average rate of change of $2.22\%/year$ for the remaining cities. Finally, we note this land cover analysis does not reflect increases in urban density.

5. Discussion and conclusions

The type of land cover and how it evolves over time, driven by anthropogenic, natural and other processes, has implications for the local and wider environment, as well as the way it is managed. In addition, land cover change provides a good metric for measuring and monitoring pressures on ecosystems and biodiversity (OECD, 2018). As

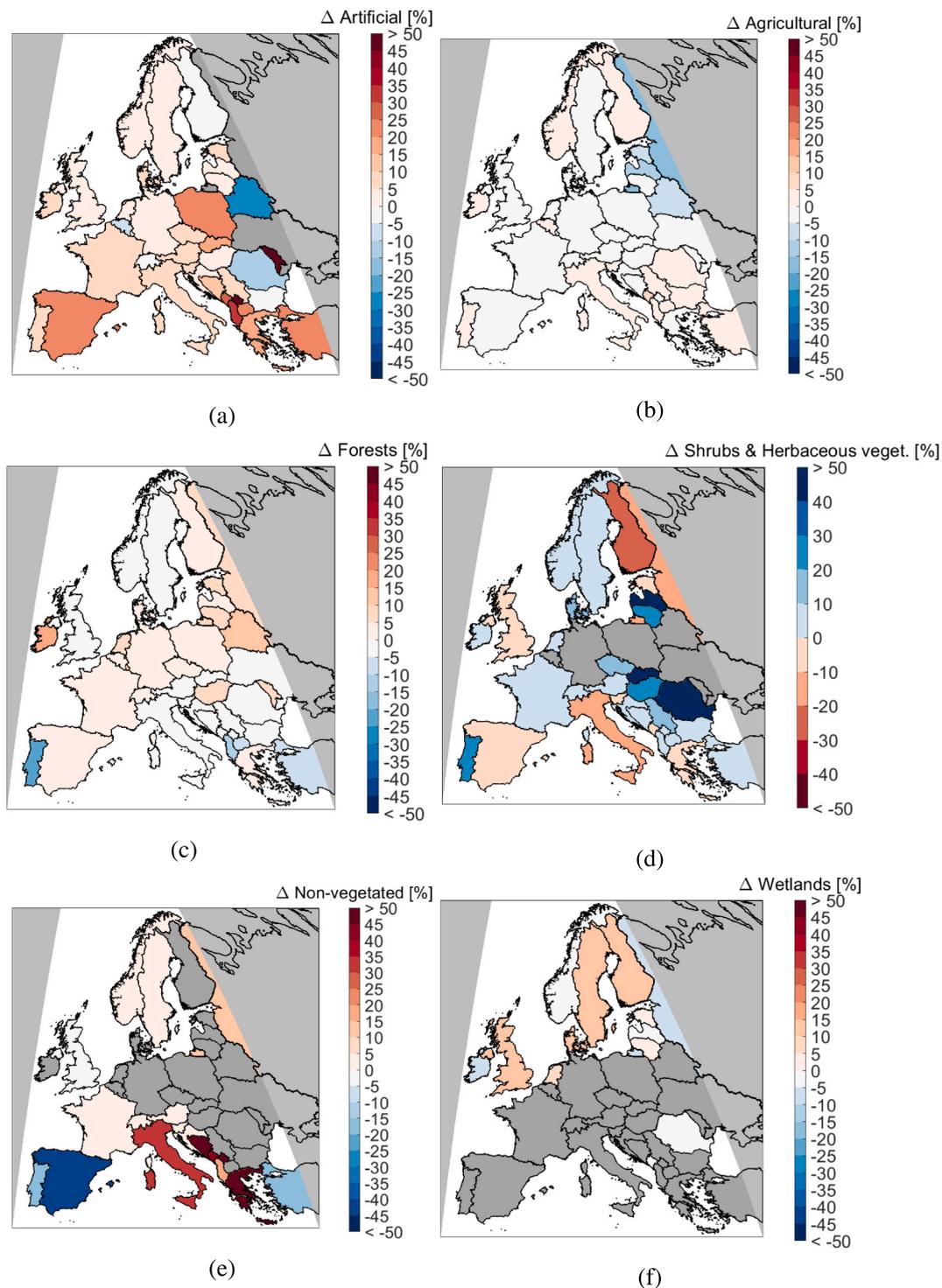


Fig. 6. Relative change in land cover with respect to the number of samples within the class itself, between 2001 and 2019. Countries where the class covers less than 0.5% of the land are shown in grey. The change is shown for the (a) artificial, (b) agricultural, (c) forests, (d) shrubs and herbaceous vegetation, (e) non-vegetated and (f) wetlands classes.

a result, determining land cover annually over a large geographical area can grant insight into when and how land cover is altered and any existing spatial patterns.

In this work we produce and analyse the annual European land cover between 2001 and 2019, both for the entire continent and on a national as well as metropolitan level, on a $\sim 250 \text{ m} \times \sim 250 \text{ m}$ resolution. Our work indicates that the production of NDVI-derived annual land cover classification of large spatial domains with a high class fidelity is possible. We find good agreement between the generated land cover and

reference data, with a total classification accuracy of 75% and precision of 76%. Critical to the ability to produce a continent-spanning land cover classifier is the sweeping segmentation, in which smaller overlapping segments of the continent-wide data (60×60 pixels) are classified at a time. However the classification performance varies between the 39 land cover types classified, with rarer classes (e.g. artificial) performing worse than more common ones (e.g. agricultural). This may be driven by the aforementioned rarity of these classes in the training set, but it could also be driven by the inherent suitability of NDVI

Table 2

The relative rates of change of the land cover types between 2001 and 2019 for the ten most populated cities within the domain, listed alphabetically.

City	Urban [%/yr]	Ind., Comm. & Transportation [%/yr]	Mine, Dump & Construction [%/yr]	Artificial Vegetated [%/yr]	Agricultural [%/yr]	Forests [%/yr]
Berlin	0.21	1.86	-3.30	2.97	-0.52	0.20
Bucharest	0.62	3.90	-	4.87	-0.72	0.42
Hamburg	-0.56	0.91	1.23	3.67	0.06	0.63
Istanbul	0.43	4.09	2.07	0.26	-0.35	-0.21
London	-0.32	0.77	-1.66	2.60	0.20	-1.41
Madrid	-0.19	1.28	-2.10	0.82	0.26	0.57
Paris	0.71	0.13	-3.59	1.79	-0.32	0.26
Rome	0.24	1.37	-5.58	-0.64	-0.11	-0.16
Vienna	-0.26	-0.38	2.38	0.81	0.27	0.01
Warsaw	0.66	1.17	0.00	-3.30	-0.82	-0.20

towards medium-highly vegetated areas (da Silva et al., 2020). While some of the individual classes exhibit similar NDVI behaviour, the use of segmentation enhances the inter-class separability. A smaller segment size could further increase the classification accuracy of rarer classes (e. g. artificial), at the expense of classification bias and spatial continuity. Overall, we find that the classification performance was most sensitive to changes in the segment size, with a lower sensitivity for the size and number of the used classification trees, and oversampling ratio.

Our results are in part limited by the reference land cover data set that the supervised classification algorithm is trained on. We note that an absolutely true reference land cover data set does not exist, and while the CORINE land cover database reports a total classification accuracy of 85%, errors or inconsistencies in the CORINE land cover labels result in reduced training data quality and subsequently reduced classification performance. For the classification performance assessment, we also compare our classification results with the Copernicus land cover map for 2015 and find a total classification accuracy of 71%. Our classified land cover resembles the Copernicus land cover map more closely than CORINE for the agricultural, forests and water bodies classes, but is less accurate for the artificial and wetlands classes.

Over the assessed time period we find that the artificial land cover class exhibits the highest growth rate, followed by the wetlands and forest classes. The agricultural, and the shrubs and herbaceous vegetation classes show an opposite trend. The non-vegetated area has remained roughly constant over the time period. These results could in future work be analysed to quantify the ‘conversions’ between the land cover classes, assessing the main conversion pathways between losses and gains in land cover for specific classes. In addition, we find that these land cover class changes are heterogeneous over the domain. While environmental management is often focused on a local scale, its implications can have wider impacts for which uniform data sets of a larger spatial domain are needed.

We also find that the land cover time series fluctuate over the nineteen year period. This highlights the importance of quantifying trends in land cover change over time, as opposed to differences between specific years. In addition, this indicates the need for fine temporal granularity of land cover data sets (such as the ones produced in this work) for monitoring and understanding the impacts of environmental management and other interventions. Finally we note, that while annual results are produced, and despite the temporal smoothing of the classification results, some temporal class inconsistencies still remain in the produced data set.

At the metropolitan level, we find mostly increasing trends for the urban class and the industrial, commercial and transportation related land cover, in line with increasing urbanisation. We also find increasing trends for the artificial vegetated class, possibly indicating the results of efforts to ‘green’ urban space. Nevertheless, decreases in urban land cover were predicted for four metropolitan areas. This is likely a result of underestimation by the classifier, rather than indicative of decreases in urban size. Furthermore, the used classification method is only able to detect changes in area, and is thus not well suited to capture other

changes in land cover, such as urban intensification (Hennig et al., 2016).

We expect that the $\sim 250 \text{ m} \times \sim 250 \text{ m}$ spatial resolution of our results make them less relevant for smaller spatial domains, and as such, the metropolitan results presented are likely to be the most sensitive to this. This is additionally because of the relatively low accuracy of the higher fidelity artificial classes (9%–61%), although it should be noted that the accuracies of these classes are higher for these metropolitan areas than the domain as a whole. We further note that we are able to capture small changes in land cover (e.g. the construction of new airports).

Improvements in the accuracy of the annual classification results are expected if finer resolution data are used, and if the NDVI and growing season metrics are determined using time series decomposed into trend and seasonal components. The accuracy of the land cover data set can be further improved by using multiple sources of data (e.g. multiple spectral indices (Inglada et al., 2017)). Finally, while we consider the computational cost of the classification to be reasonable (~ 1000 CPU-hours per classified year), the majority of this cost is attributed to the second layer of majority voting (sweeping segmenting). We note that the ten iterations used here can be considered conservative and a lower number of iterations can be selected, which would reduce the computational cost.

Our results and accompanying data sets can aid further studies that assess the impact of environmental management interventions on land cover change, or, vice versa, determine the implications of land cover change on water, agriculture, forestry and other management efforts. With the continuous availability of large scale NDVI data, the approach presented in this article shows promise for the creation of ‘near-real-year’ annual land cover results, eliminating the current lag in the availability of large scale land cover data sets. This could further contribute to the monitoring of land cover change and associated interventions, supporting efforts to measure progress towards UN’s Sustainable Development Goals and other targets (OECD, 2018). Finally, while the European domain has been analysed here, this approach could be expanded to any domain. A suitable reference land cover database is needed, but the classification method has shown that the reference data set does not need to be cover the full time span or spatial domain.

Author contributions

Vincent B. Verhoeven: Methodology, Software, Formal analysis, Data curation, Visualization, Writing–original draft preparation, Writing–review and editing. Irene C. Dedoussi: Conceptualisation, Methodology, writing–review and editing, Supervision, Funding acquisition.

Data availability

The produced annual land cover products are provided in DOI: <https://doi.org/10.4121/14478171>. The code is available at DOI: <https://doi.org/10.4121/16571516>.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was carried out on the Dutch national e-infrastructure with the support of SURF Cooperative. It was partly supported by the Dutch Research Council (NWO) under grant OCENW.XS.060.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jenvman.2021.113917>.

References

- Abahussain, A.A., Abdu, A.S., Al-Zubari, W.K., El-Deen, N.A., Abd-ur-Raheem, M., 2002. Desertification in the arab region: analysis of current status and trends. *J. Arid Environ.* 51, 521–545.
- Ali, A., de Bie, C., Skidmore, A.K., Scarrott, R.G., Hamad, A., Venus, V., Lymberakis, P., 2013. Mapping land cover gradients through analysis of hyper-temporal NDVI imagery. *Int. J. Appl. Earth Obs. Geoinf.* 23, 301–312.
- An, L., Brown, D.G., 2008. Survival analysis in land change science: integrating with giscience to address temporal complexities. *Ann. Assoc. Am. Geogr.* 98, 323–344.
- Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2018. Present and future köppen-geiger climate classification maps at 1-km resolution. *Scientific Data* 5.
- Beekhuizen, J., Clarke, K.C., 2010. Toward accurate land use mapping: using geocomputation to improve classification accuracy and reveal uncertainty. *Int. J. Appl. Earth Obs. Geoinf.* 12, 127–137.
- de Bie, C.A.J.M., Khan, M.R., Smakhtin, V.U., Venus, V., Weir, M.J.C., Smaling, E.M.A., 2011. Analysis of multi-temporal spot ndvi images for small-scale land-use mapping. *Int. J. Rem. Sens.* 32, 6673–6693.
- Braimah, A.K., Vlek, P.L.G., 2005. Land-cover change trajectories in northern Ghana. *Environ. Manag.* 36, 356–373.
- Brown, M.E., Pinzón, J.E., Didan, K., Morisette, J.T., Tucker, C.J., 2006. Spot-vegetation, seawifs, modis, and landsat etm+ sensors. *IEEE Trans. Geosci. Rem. Sens.* 44, 121–136.
- Buchhorn, M., Smets, B., Bertels, L., Lesiv, M., Tsendbazar, N.E., Li, L., 2019. Copernicus Global Land Operations “Vegetation and Energy”. Technical Report. VITO NV.
- Büttner, G., Maucha, G., 2006. The Thematic Accuracy of Corine Land Cover 2000. Technical Report 7. European Environment Agency.
- Büttner, G., Soukup, T., Kosztra, B., 2019. Climate Change Adaptation in the Agricultural Sector in Europe. Technical Report 4. European Environment Agency, Copenhagen, Denmark.
- Carreiras, J.M.B., Pereira, J.M.C., Shimabukuro, Y.E., 2006. Land-cover mapping in the brazilian amazon using spot-4 vegetation data and machine learning classification methods. *Photogramm. Eng. Rem. Sens.* 72, 897–910.
- Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P., 2002. Smote: synthetic minority over-sampling technique. *J. Artif. Intell. Res.* 16, 321–357.
- Chen, J., Jönsson, P., Tamura, M., Gu, Z., Matsushita, B., Eklundh, L., 2004. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. *Rem. Sens. Environ.* 91, 332–344.
- Davidson, N.C., 2014. How much wetland has the world lost? long-term and recent trends in global wetland area. *Mar. Freshw. Res.* 65, 934–941.
- Didan, K., Munoz, A.B., Solano, R., Huete, A., 2015. MODIS Vegetation Index User’s Guide (MOD13 Series) Version 3.00. Technical Report. The University of Arizona.
- d’Amour, C.B., Reitsma, F., Baiocchi, G., Barthel, S., Güneralp, B., Erb, K.H., Haberl, H., Creutzig, F., Seto, K.C., 2017. Future urban land expansion and implications for global croplands. *Proc. Natl. Acad. Sci. Unit. States Am.* 114, 8939–8944.
- Estel, S., Kuemmerle, T., Alcántara, C., Levers, C., Prishchepov, A., Hostert, P., 2015. Mapping farmland abandonment and recultivation across europe using modis ndvi time series. *Rem. Sens. Environ.* 163, 312–325.
- Estel, S., Kuemmerle, T., Levers, C., Baumann, M., Hostert, P., 2016. Mapping cropland use-intensity across europe using modis ndvi time series. *Environ. Res. Lett.* 11.
- European Environment Agency, 2002. Corine Land Cover Update 2000. Technical Report, Copenhagen, Denmark.
- European Environment Agency, 2007. CLC2006 Technical Guidelines. Technical Report, Copenhagen, Denmark.
- European Environment Agency, 2014. CLC2012 Addendum to CLC2006 Technical Guidelines. Technical Report.
- European Environment Agency, 2017. CLC2018 Technical Guidelines. Technical Report, Vienna, Austria.
- European Environment Agency, 2019. Corine Land Cover. URL: <https://land.copernicus.eu/pan-european/corine-land-cover>.
- De Fries, R.S., Townshend, J.R.G., 1994. NDVI-derived land cover classifications at a global scale. *Int. J. Rem. Sens.* 15, 3567–3586.
- Geerken, R., Zaitchik, B., Evans, J.P., 2005. Classifying rangeland vegetation type and coverage from ndvi time series using fourier filtered cycle similarity. *Int. J. Rem. Sens.* 26, 5535–5554.
- Ghimire, B., Rogan, J., Galiano, V.R., Panday, P., Neeti, N., 2012. An evaluation of bagging, boosting, and random forests for land-cover classification in cape cod, Massachusetts, USA. *GIScience Remote Sens.* 49, 623–643.
- Gutman, G.G., 1991. Vegetation indices from AVHRR: an update and future prospects. *Rem. Sens. Environ.* 35, 121–136.
- Han, H., Wang, W.Y., Mao, B.H., 2005. Borderline-smote: a new over-sampling method in imbalanced data sets learning. *International Conference on Intelligent Computing* 878–887.
- Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-resolution global maps of 21st-century forest cover change. *Science* 342, 850–853.
- Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics* 3, 610–621.
- Hennig, E.I., Schwic, C., Soukup, T., Orlitová, E., Kienast, F., Jaeger, J.A.G., 2015. Multi-scale analysis of urban sprawl in europe: towards a european de-sprawling strategy. *Land Use Pol.* 49, 483–498.
- Hennig, E.I., Jaeger, J.A.G., Soukup, T., Orlitová, E., Horáková, M., Schwic, C., Kienast, F., 2016. Urban Sprawl in Europe. Technical Report 11. EEA and FOEN. Denmark and Bern, Copenhagen. Switzerland.
- Houghton, R.A., House, J.I., Pongratz, J., van der Werf, G.R., DeFries, R.S., Hansen, M.C., Le Quéré, C., Ramankutty, N., 2012. Carbon emissions from land use and land-cover change. *Biogeosciences* 9, 5125–5142.
- Inglada, J., Vincent, A., Arias, M., Tardy, B., Morin, D., Rodes, I., 2017. Operational high resolution land cover map production at the country scale using satellite image time series. *Rem. Sens.* 95.
- IPCC, 2019. *Climate Change and Land*. <https://www.ipcc.ch/srccl/>. (Accessed 31 December 2019).
- Jeffries, M.O., Richter-Menge, J., Overland, J.E., 2015. Arctic Report Card 2015. Technical Report. NOAA.
- Jönsson, P., Eklundh, L., 2004. Timesat - a program for analyzing time-series of satellite sensor data. *Comput. Geosci.* 30, 833–845.
- Keenan, R.J., Reams, G.A., Achard, F., de Freitas, J.V., Grainger, A., Lindquist, E., 2015. Dynamics of global forest area: results from the fao global forest resources assessment 2015. *For. Ecol. Manag.* 352, 9–20.
- Kerr, J.T., Ostrovsky, M., 2003. From space to species: ecological applications for remote sensing. *Trends Ecol. Evol.* 18.
- Lunetta, R.S., Knight, J.F., Ediriwickrema, J., Lyon, J.G., Worthy, L.D., 2006. Land-cover change detection using multi-temporal modis ndvi data. *Rem. Sens. Environ.* 105, 142–154.
- Mahmood, R., Pielke, R.A., Hubbard, K.G., Niyogi, D., Dirmeyer, P.A., McAlpine, C., Carleton, A.M., Hale, R., Gameda, S., Beltrán-Przedkurt, A., Baker, B., McNider, R., Legates, D.R., Shepherd, M., Du, J., Blanken, P.D., Frauenfeld, O.W., Nair, U.S., Fall, S., 2014. Land cover changes and their biogeophysical effects on climate. *Int. J. Climatol.* 34, 929–953.
- Malinowski, R., Lewinski, S., Rybicki, M., Jenerowicz, M., Gromny, E., Krupinsky, M., ojtkoski, C., Krupinski, M., Günther, S., Krätzschmar, E., 2019. Final Report. Technical Report. Centrum Badań Kosmicznych PAN.
- Martellozzo, F., Ramankutty, N., Hall, R.J., Price, D.T., Purdy, B., Friedl, M.A., 2015. Urbanization and the loss of prime farmland: a case study in the calgary-edmonton corridor of alberta. *Reg. Environ. Change* 15, 881–893.
- Martínez, B., Gilabert, M.A., 2009. Vegetation Dynamics from Ndvi Time Series Analysis Using the Wavelet Transform. *Remote Sensing of Environment*, pp. 1823–1842.
- Mena, C.F., 2008. Trajectories of land-use and land-cover in the northern equatorial amazon: temporal composition, spatial configuration, and probability of change. *Photogramm. Eng. Rem. Sens.* 74, 737–751.
- OECD, 2018. *Monitoring Land Cover Change - Occd*. <https://www.oecd.org/env/indicators-modelling-outlooks/monitoring-land-cover-change.htm>. (Accessed 22 September 2021).
- Pacheco, C.E., Aguado, M.I., Mollicone, D., 2014. Identification and characterization of deforestation hot spots in Venezuela using modis satellite images. *Acta Amazonica* 44, 185–196.
- Panagos, P., Borrelli, P., Poesen, J., Ballabio, C., Lugato, E., Meusberger, K., Montanarella, L., Alewell, C., 2015. The new assessment of soil loss by water erosion in europe. *Environ. Sci. Pol.* 54, 438–447.
- Park, T., Ganguly, S., Tommervik, H., Euskirchen, E.S., Hogda, K.A., Karlsen, S.R., Brovkin, V., Nemani, R.R., Myneni, R.B., 2016. Changes in growing season duration and productivity of northern vegetation inferred from long-term remote sensing data. *Environ. Res. Lett.* 11.
- Radeloff, V.C., Gutman, G., 2017. *Land-Cover and Land-Use Changes in Eastern Europe after the Collapse of the Soviet Union in 1991*, 1 ed. Springer.
- Reed, B.C., Brown, J.F., VanderZee, D., Loveland, T.R., Merchant, J.W., Ohlen, D.O., 1994. Measuring phenological variability from satellite imagery. *J. Veg. Sci.* 5, 703–714.
- Rodriguez-Galiano, V.F., Chica-Olmo, M., Abarca-Hernandez, F., Atkinson, P.M., Jeganathan, C., 2012. Random forest classification of mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. *Rem. Sens. Environ.* 121, 93–107.
- Safavian, S.R., Landgrebe, D., 1991. A survey of decision tree classifier methodology. *IEEE Transactions on Systems, Man and Cybernetics* 21.
- Savitzky, A., Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* 36, 1627–1639.

- Shao, Y., Lunetta, R.S., Wheeler, B., Liames, J.S., Campbell, J.B., 2016. An evaluation of time-series smoothing algorithms for land-cover classifications using modis-ndvi multi-temporal data. *Rem. Sens. Environ.* 174, 258–265.
- Sheffield, K., Morse-McNabb, E., Clark, R., Robson, S., Lewis, H., 2015. Mapping dominant annual land cover from 2009 to 2013 across Victoria, Australia using satellite imagery. *Scientific Data* 2, 150069.
- da Silva, V.S., Salami, G., da Silva, M.I.O., Silva, E.A., Monteiro Junior, J.J., Alba, E., 2020. Methodological evaluation of vegetation indexes in land use and land cover (LULC) classification. *Geology, Ecology, and Landscapes* 4, 159–169.
- Sulla-Menashe, D., Gray, J.M., Abercrombie, S.P., Friedl, M.A., 2019. Hierarchical mapping of annual global land cover 2001 to present: the modis collection 6 land cover product. *Rem. Sens. Environ.* 222, 183–194.
- Usman, M., Liedl, R., Shahid, M.A., Abbas, A., 2015. Land use/land cover classification and its change detection using multi-temporal MODIS NDVI data. *J. Geogr. Sci.* 25, 1479–1506.
- Vågen, T.G., 2006. Remote sensing of complex land use change trajectories - a case study from the highlands of Madagascar. *Agric. Ecosyst. Environ.* 115, 219–228.
- Weiss, G.M., Provost, F., 2003. Learning when training data are costly: the effect of class distribution on tree induction. *J. Artif. Intell. Res.* 19, 315–354.
- White, M.A., de Beurs, K.M., Didan, K., Inouye, D.W., Richardson, A.D., Jensen, O.P., O'Keefe, J., Zhang, G., Nemani, R.R., van Leeuwen, W.J.D., Brown, J.F., de Wit, A., Schaepman, M., Lin, X., Dettinger, M., Bailey, A.S., Kimball, J., Schwartz, M.D., Baldocchi, D.D., Lee, J.T., Lauenroth, W.K., 2009. Intercomparison, interpretation and assessment of spring phenology in north America estimated from remote sensing for 1982-2006. *Global Change Biol.* 15, 2335–2359.
- Xie, Y., Sha, Z., Yu, M., 2008. Remote sensing imagery in vegetation mapping: a review. *J. Plant Ecol.* 1, 9–23.