Received: 7 January 2022 | Accepted: 7 June 2022

DOI: 10.1002/2688-8319.12162

Ecological Solutions and Evidence | BRITISH ECOLOGICAL SOCIETY

RESEARCH ARTICLE

An operational land cover and land cover change toolbox: processing open-source data with open-source software

Naomi Gatis[1] | Donna Carless[1] | David J. Luscombe[1] | Richard E. Brazier[1] | Karen Anderson[2]

[1]Centre for Resilience in Environment, Water and Waste, College of Life and Environmental Sciences, University of Exeter, Exeter, UK

[2]Environment and Sustainability Institute, University of Exeter Penryn Campus, Penryn, UK

Correspondence
Naomi Gatis, Centre for Resilience in Environment, Water and Waste, College of Life and Environmental Sciences, University of Exeter, Amory Building, Rennes Drive, Exeter, EX4 4RJ, UK.
Email: N.Gatis@exeter.ac.uk

Handling Editor: Daniel Bebber

Funding information
Dartmoor National Park Authority; Department for Environment, Food and Rural Affairs : Dartmoor Environmental Land Management Schemes Test & Trials; Natural Environment Research Council : NE/P011217/1; South West Partnership for Environmental and Economic Prosperity : SWEEP023

[Correction added on 28 July 2022, after first online publication: abbreviated author names were changed to their full names.]

Abstract

1. Accurate and up-to-date land cover maps are vital for underpinning evidence-based landscape management decision-making. However, the technical skills required to extract tailored information about land cover dynamics from these open-access geospatial data often limit their use by those making landscape management decisions.
2. Using Dartmoor National Park as an example, we demonstrate an open-source toolkit which uses open-source software (QGIS and RStudio) to process freely available Sentinel-2 and public LiDAR data sets to produce fine scale (10 m^2 grain size) land cover maps.
3. The toolbox has been designed for use by staff within the national park, for example, enabling land cover maps to be updated as required in the future.
4. An area of 945 km^2 was mapped using a trained random forest classifier following a classification scheme tailored to the needs of the national park.
5. A 2019 land cover map had an overall user's accuracy of 79%, with 13 out of 17 land cover classes achieving greater than 70% accuracy.
6. Spatially, accuracy was related via logistical regression to blue band surface reflectance in the spring and topographic slope derived from LiDAR (1 m resolution), with greater accuracy in steeper terrain and areas exhibiting higher blue reflectance.
7. Between an earlier (2017–2019) and later (2019–2021) time frame, 8% of pixels changed, most of the change by area occurred in the most common classes. However, the largest proportional increase occurred in Upland Meadows, Lowland Meadows and Blanket Bog, all habitats subject to restoration efforts. Identifying areas of change enables future field work to be better targeted.
8. We discuss the application of this mapping to land management within the Dartmoor national park and of the potential of tailored land cover and land cover change mapping, via this toolbox, to evidence-based environmental decision-making more widely.

KEYWORDS
habitat change, habitat classification, land management, random forest classifier, satellite imagery, Sentinel-2

1 | INTRODUCTION

Accurate and up-to-date land cover maps are vital for underpinning evidence-based landscape management decision-making. The availability and accuracy of these maps can impact on a range of policy areas, such as agri-environment schemes (e.g. Environmental Land Management Schemes (Department for Environment Food & Rural Affairs & Rural Payments Agency, 2021) and the Common Agricultural Policy); targeted Land-Use, Land-Use Change and Forestry (Watson et al., 2000) including peatland restoration and afforestation; conservation and urban development. A time series of land cover maps enables land cover change to be quantified which can highlight areas of concern and provide a means to assess the impacts of decisions within a dynamic landscape (Brown et al., 2020).

Earth Observation satellites have been routinely collecting image data of the Earth's surface since the launch of Landsat-1 in 1972. Their use for land cover mapping has increased over time as the spatial resolution of image pixels has increased, the return period between image collection dates has decreased and the costs of data access have reduced or become free (Woodcock et al., 2008). Meanwhile data are increasingly analysis ready, whilst computational power, image processing and analysis capabilities have improved in many settings (Wulder et al., 2018).

Landsat imagery has been used to make global (e.g. Chen et al., 2015; Gong et al., 2013; Liu et al., 2021), continental (e.g. CORINE Land Cover 2000 in Europe [Heymann et al., 1994]), national (e.g. Conterminous United States [Vogelmann et al., 2001]) and local (Yuan et al., 2005) scale land cover maps for target periods (usually 1–3 year window) for decades. Bringing advanced processing to the archive of Landsat imagery (1972-present), studies are now also looking at land use change (e.g. Souza et al., 2020). However, these have a minimum spatial resolution of 30 m.

Within the United Kingdom, the Centre for Ecology and Hydrology produced a series of national land cover maps (1990, 2000, 2007, 2015, 2017, 2018, 2019, 2020) using consistent broad habitat types. Originally using parcel-based thematic classification of Landsat data, since 2017 a Sentinel-2-based 20 m raster product has been produced (Morton et al., 2020a, 2020b). Using Sentinel-2 data to mask non-agricultural land, the CRop Map Of England (CROME) (Rural Payments Agency, 2019) uses weekly mosaiced Sentinel-1 data in a random forest classifier to output 4157 m^2 hexagonal vector tiles of 20 crop types. The recently released Living England (Kilcoyne et al., 2022) land cover map of England segments the landscape, with a 300 m^2 minimum mappable unit, before using Sentinel-1 and Sentinel-2 data with additional data in a supervised Random Forest Classifier to produce a vector map of England.

Commercial Very High-Resolution satellite constellations have the potential to map land cover at sub-metre spatial resolution. These data have been used to map, for example, urban areas (Aguilar et al., 2013; Hashim et al., 2019), crops (Esetlili et al., 2018; Moody et al., 2016) and coastal areas (Adam et al., 2014) with high accuracies. However, with finer spatial and temporal resolution come challenges as well as opportunities. Finer spatial resolution has been shown to result in increased within class spectral variability which reduces the between-class spectral discrimination (Woodcock & Strahler, 1987).It can also increase the size of datasets (e. g. Moody et al. (2016) had 4-bands of Planet Lab imagery, each with −110 million 3 m pixels to cover a 29 × 34 km2 area) which has led to an increase in the complexity of pre-processing and/or the algorithms used (e.g. 3D convoluted neural networks (Saralioglu & Gungor, 2022). Arguing that within-object variability is a potential source of information, Carleer and Wolf (2006) used object-based segmentation to derive spectral (e.g. mean blue), textural (e.g. homogeneity) and morphological (e.g. length) information from QuickBird and Ikonos imagery about segmented objects before classifying urban and rural areas within Belgium. The main limitation on these data is the cost to the user: a 38 × 38 km image (as required for this study) would cost between £1412 (Planet Labs 3 m resolution) and £32,367 (GeoEye 0.5 m resolution) per image (based on values in Sozzi et al., 2018).

The European Space Agency's Sentinel-2A and Sentinel-2B satellites (launched in 2015 and 2017, respectively) provide high temporal frequency (5-day), fine spatial scale (10 m^2), freely available data suitable for mapping land cover and land cover change at a scale and extent useful for local and regional land management decision-making. Workflows to produce local (e.g. Noi & Kappas, 2018), national (e.g. Close et al., 2018). This is ecologically unlikely given 7 % of broad land cover classes within DNP were mapped as changed between 1990-2015 (Morton et al., 2020b). and global (e.g. Karra et al., 2021) land cover maps are already being developed for these data with promising results.

To maximize the potential of these data, mapping needs to be tailored to the area of interest and the question at hand because ecological conditions vary geographically. Global models usually use broad land cover classes (Friedl et al., 2002). Other studies have mapped the presence/absence of a particular invasive species (Royimani et al., 2019) or habitat of interest, for example fenland (Foody et al., 2007). Where multiple land cover classes have been used, these are often tailored to the habitat being mapped, for example the UNESCO Vegetation Classification only maps natural vegetation (Küchler & Zonneveld, 1988), whilst the Urban Atlas focuses on land cover for large urban zones (https://land.copernicus.eu/local/urban-atlas). The technical remote sensing workflow required to extract tailored information about land cover dynamics from these open-access geospatial data

often precludes those who make landscape management decisions. It is vital that this gap is bridged, so that the process of land cover mapping, including choices about what habitats to map, can be placed in the hands of those making decisions (Saah et al., 2019).

To address this need, there is a growing trend towards open-source, open access processing chains which require limited technical skills. Grippa et al. (2017) used detailed annotations to guide the user through their Python code, whichlinks a series of GRASS GIS and R functions together to produce an object-basedclassification of urban areas, based on World-view-3 imagery and whereavailable LiDAR data. Murry et al. (2018) made the data storage and analysis capacities of Google Earth Engine available to non-specialists via a web-based platform (https://remap-app.org), which produces 30 m^2 land cover and land cover change (between 2003 and 2017) maps based on a user's supplied training data. Similarly Xing et al. (2021) have developed O-LC Mapping, although no link to the website was available. Offline, plugins for QGIS (Semi-Automatic Classification (Congedo, 2021) and dzetsaka (Karasiak, 2021)) enable users to classify remote sensing images through a click-button graphical users interface. Over time, updates have increased the number of tools available, including Random Forest classifier in Semi-Automatic Classification v7.0.0 (October 2020), improved reliability and reduced the required technical capacity of the user.

In this project, we set out to develop an open-source toolkit aimed at putting the remote sensing workflow in the hands of agencies leading, or advising on, land management decisions such as national park authorities, councils, landowners, farm advisers or wildlife charities. Using Dartmoor National Park as an example, we set out to test the toolkit using freely available Sentinel-2 data combined with freely available airborne LiDAR data with the goal of testing a workflow for producing robust and repeatable annual land cover and land cover change maps at a fine resolution (10 m^2) over a landscape extent (954 km^2).

2 | MATERIALS AND METHODS

2.1 | Study area

Dartmoor National Park, located in southwest England, covers 954 km^2 and ranges in elevation from 30 to 621 mASL. The western side of the park is dominated by upland moorland of which 315 km^2 has been mapped as underlain by peat (Gatis, Luscombe, et al., 2019) with some ecohydrological degradation (Carless et al., 2019). The eastern side is generally lower in altitude and dominated by grasslands.

Dartmoor National Park Authority have a dual mandate: to promote the economic and social well-being of local communities whilst also conserving and enhancing the natural beauty, wildlife and cultural heritage of the park and promoting opportunities for the public to understand and enjoy its special qualities (https://www.dartmoor.gov.uk/about-us/who-we-are). They have identified accurate and up-to-date land cover maps as an essential requirement to allow them to balance these management goals: to efficiently monitor and understand the effects of climate change, increase visitor numbers and new agri-environment schemes as well as assess the success or otherwise of conservation and enhancement schemes.

2.2 | Land cover mapping

2.2.1 | Classification scheme

For land cover mapping to be useful, it must map the land cover(s) of interest. Here, we used a classification scheme derived from the UK Habitat Classification system (www.ukhab.org) which has been designed so that all terrestrial and freshwater habitats found in the United Kingdom fit within a hierarchy that aligns with both the largescale Mapping and Assessment of Ecosystems (MAES) categories as well as the European Habitats Directive Annex 1 habitats at the small scale. It is the current standard survey method for Preliminary Ecological Appraisals and underlies the Department for Environment, Food and Rural Affairs (DEFRA) Biodiversity Metric, a means by which ecologists, planners and developers can assess biodiversity change resulting from development or changes in land management in the United Kingdom. Other classification systems, more appropriate to different geographical areas and policy decisions, should be derived and used as required.

The classification scheme created for this tool (Table 1) is derived from the UK Habitat Classification System to levels 2, 3 and 4 as appropriate with the exception of f1a5 Blanket Bog and f1a6 Degraded Blanket bog which were mapped to level 5 due to the extent and importance of this habitat type within the National Park. Degraded blanket bog was then subdivided based on the overlying vegetation type. Some UK Hab classes were not included in this tool; the reasons are given in Table 1.

As lowland and upland grasslands as well as lowland and upland heathland are not spectrally distinguishable in Sentinel-2 images, these were grouped together for classification. Post-classification pixels identified in either of these groups were allocated to upland or lowland classes based on the ground elevation (uplands extent layer derived from LiDAR). We tested a broader Acid grassland class including Purple moor grass and rush pastures and Fen marsh and swamp, but the increased within-class variability reduced overall accuracy and resulted in these important classes being omitted.

2.2.2 | Training and validation data

Using the open-source GIS software QGIS 3.16.2 Hannover (QGIS Development Team, 2020), training and validation points were selected by stratified random sampling using the Random Points inside Polygons tool based on landscape character types downloaded from https://www.devon.gov.uk/planning/planning-policies/landscape/devons-landscape-character-assessment. A greater density of pixels (10%) were selected in the less common landscape character types (e.g. Moorland edge slope) and a lower density (2%) in the more common (e.g. Farmed and Forested Plateau) in an effort to balance the sample sizes. For an area with no previous mapping, where there is

TABLE 1 Classification scheme used by this tool and how it relates to the UK Habitat Classification scheme

UK Hab level	UK Hab code	UK Hab label	Land Cover class code	Land Cover class label
4	g1a	Lowland dry acid grassland	1	Lowland acid grassland
4	g1b	Upland acid grassland	3	Upland acid grassland
4	g1c	Bracken	2	Bracken
3	g2	Calcareous grassland	Negligible/no extent of this habitat within DNP	
4	g3a	Lowland meadows	6	Lowland meadows
4	g3b	Upland hay meadows	7	Upland hay meadows
4	g3c	Other neutral grassland	Not distinguishable from upland hay and lowland meadows	
3	g4	Modified grassland	9	Modified grassland
4	w1a	Upland oakwood	10	Upland oakwood
3	w1	Broadleaved mixed and yew woodland	11	Other broadleaved, mixed and yew woodland
3	w2	Coniferous woodland	12	Coniferous woodland
4	h1a	Lowland Heathland	13	Lowland Heathland
4	h1b	Upland Heathland	14	Upland Heathland
4	h1c	Mountain heaths and willow scrub		
			Not distinguishable from upland and lowland heathland	
2	h2	Hedgerows	Not distinguishable from other broadleaves, mixed, yew or conifer woodlands	
4	h3e	Gorse scrub	18	Gorse scrub
5	f1a5	Blanket bog (H7130)	19	Blanket bog (H7130)
5	f1a6	Degraded blanket bog	303	Acid grass over degraded blanket bog
			314	Heathland over degraded blanket bog
			328	Unvegetated degraded blanket bog
4	f2b	Purple moor grass and rush pastures	23	Purple moor grass and rush pastures
3	f2	Fen marsh and swamp	25	Flushes, fens, marsh and swamp
2	c	Cropland	26	Cropland
2	u	Urban	27	Urban
3	s1	Inland rock	28	Inland rock
3	s2	Supralittoral Rock	Negligible/no extent of this habitat within DNP	
3	s3	Supralittoral Sediment		
2	r	Rivers and lakes	29	Rivers and lakes
2	t	Marine inlets and transitional waters	Negligible/no extent of this habitat within DNP	

no information to guide a random stratified pixel selection, selecting random points is recommended (Olofsson et al., 2014), and this can be done using the Random Points in Extent tool. Despite the deliberate sampling bias, some classes were still under-represented due to their limited spatial extent. Additional pixels were identified so that all classes had a minimum of 50 pixels following recommendations by Hay (1979) and Congalton (1991) being careful that pixels selected from small fragments were not auto-correlated. In total, there were 4140 pixels. Statistical methods (under-sampling more common classes, over sampling rarer classes and Synthetic Majority Oversampling [SMOTE]) to balance the sample size were found to reduce classification accuracy (details in Supporting Information S1).

Pixels were assigned a land cover class based on photo-interpretation of higher resolution aerial imagery via Google Earth and

TABLE 2 R packages used in the code

Package	Version	Reference
Caret	6.0–90	Kuhn, 2020
doParallel	1.0.16	Microsoft Coporation & Weston, 2020)
diffeR	0.0–6	Pontius Jr & Santacruz, 2019
e1071	1.7–9	Meyer et al., 2020
Here	1.0.1	Müller, 2020
MASS	7.3–54	Venables & Ripley, 2002
Packrat	0.5.0	Ushey et al., 2018
randomForest	4.6–14	Liaw & Weiner, 2002
Raster	3.5–9	Hijmans, 2020
Rgdal	1.5–28	Bivand et al., 2021
Rgeos	0.5–9	Bivand et al., 2020
Splitstackshape	1.4.8	Mahto, 2019
svDialogs	1.0.3	Grosjean, 2021

Google Street View. Pixels were divided up into 10 groups—these were allocated non-sequentially to four people, familiar with the land covers expected, in order to spread any individual's bias in identification geographically and across the different habitat classes, that is, one person did not photo-interpret all the pixels on the high moor and another all the farmland but each person covered a range of landscape character types.

The tool randomly assigns pixels to one of two groups with 75% assigned to training the random forest classifier and the remaining 25% to assess the accuracy of the classifier via a confusion matrix and to train the logistic regression model which spatially estimates classification accuracy.

2.2.3 | R-Code

All code was written in RStudio (v1.4.1106 [RStudio Team, 2020]) using R (v.4.1.2 [R Core Team, 2020]) using the packages outlined in Table 2. The full reproducible code and input data are available at https://zenodo.org/record/5797735 (Gatis et al., 2021). It has been written so that it can be implemented for a new area of interest, without editing the code, by supplying appropriate input files (shapefile outlining the area of interest; shapefile of peat extent [optional]; digital terrain map [optional]; spring and summer Sentinel images; shapefile of training/testing pixels and a table of class codes and labels). A user's guide is provided in Supporting Information S2 and with the code. Dependent on the user's computer and the areal extent, processing may take some time.

2.2.4 | Random forest classifier

Many algorithms are available to classify land cover; popular supervised classifiers include maximum likelihood classification, K-nearest neighbour, Support Vector Machine (Mountrakis et al., 2011) and Random Forests (Breiman, 2001). Random forest classifier was selected here as it is less sensitive to noise and overfitting than other non-parametric methods (Rodriguez-Galiano et al., 2012) with the additional benefit of being parallelizable via caret (Kuhn, 2020) with comparatively small computational demands (and therefore shorter processing times).

A random forest classifier randomly takes a subset of the training data and the features (Sentinel-2 bands, vegetation indices, aspect and slope). This subset of data is used to create a decision tree. Within the decision tree, the training data are sequentially split, each time based on a single feature until the data have been separated into classes. This is then repeated for a given number of trees (ntree), in this case 500 trees, with the final classification for each training pixel based on the majority vote (classit is most often classified as out of ntree (500) repeats). The number of features randomly sampled as candidates at each split (mtry) was tuned using a 10-folds cross-validation. This randomly splits the data into 10 subsets repeating the analysis with mtry varying from 1 to 15.

This random forest model, based on the pixels of known classes, is then applied to the whole area to predict the land cover class in unknown areas. Areal estimates for each land cover class were then adjusted for omission and commission errors following Olofsson et al. (2014).

2.2.5 | Accuracy assessment

The first accuracy assessment compares the predicted land cover class (via the random forest classifier) with the photo-interpreted classification to produce a confusion matrix. This provides both overall and per class user's accuracy: class I user's accuracy number of pixels correctly predicted as class I/total number of pixels predicted as class I; overall user's accuracy total number of correctly predicted pixels/total number of pixels.

One way of evaluating spatial patterns of accuracy is via per pixel accuracy assessment; therefore, a second spatially explicit accuracy assessment was carried out. The tool uses a global logistical regression to model accuracy (Khatami et al., 2017). The validation pixels were assigned 1 where correctly classified and 0 where incorrectly classified. A stepwise (forward and backward) multiple logistical regression model was run using the pixels as the dependent variable and using all input layers to the random forest classifier (standardized) as the independent variables. Global logistic (accuracy) model inputs were selected by minimizing the Akaike information criterion.

2.2.6 | Remote sensing data

Sentinel-2 imagery

Sentinel-2 data were selected as they are freely available and have global coverage and a fine spatial scale (10 m^2). The revisit period with 2 satellites is 5 days; this high temporal frequency increases the chance of cloud free images, enabling land cover to be mapped at a scale and extent useful for decision-making across a landscape.

Ideally, all satellite images would be free from contamination by cloud, cloud shadow and haze. One option to ensure this is to use a temporal composite with uncontaminated pixels selected within a time frame based on a set of rules. Maximum normalized difference vegetation index (Gutman et al., 1994) was popular, but with increased computing power other rules are being trialled (e.g. Azzari & Lobell, 2017; Lück & van Niekerk, 2016). There is still no universally agreed set of rules for creating composites, with different rules performing better for each satellite and ecosystem combination. Additionally, image compositing can induce spectral differences within a class across the composite image as not all pixels will be contemporaneous, and therefore differences due to phenology, land cover and radiation may occur. For example, bracken varies in the near-infrared, red and blue wavelengths over a growing season (Blackburn & Pitman, 1999). Carrasco et al. (2019) found a two-date composite more accurate than a seasonal median. For these reasons as well as processing simplicity, it was decided to focus on a limited number of cloud-free images and not to create a composite image.

Following visual inspection, cloud free images from 2019 were downloaded from https://www.sentinel-hub.com and the overall user's accuracies with different combinations were assessed. In a balance between processing time, accuracy and the likelihood of good-quality images in future years, it was decided two images—one from spring (1 January 2019 to 30April 2019) and one from summer (1 June 2019 to 30September 2019)—offered the best trade-off (see Supporting Information S1). Cole et al. (2014) also found better spectral discrimination between moorland species during spring (April) and summer (July) months.

Using Sentinel-hub EO Browser (https://www.sentinel-hub.com), all images within the time frame of interest (spring or summer) with <50% cloud were visually assessed and the best were shortlisted. Where multiple cloud free images were available, they were selected in order of priority from March > February > April > January for the spring and June > July > August > September for the summer. Bands 1 (coastal aerosol), 9 (water vapour) and 10 (cirrus) were removed due to their coarser spatial resolution (60 m). To reduce processing time, all images were clipped to the national park extent (shapefile provided by the national park) with a 20-m buffer to reduce edge effects.

TABLE 3 Vegetation Indices Used

Vegetation index	Formula	Reference
Excess of Green	(2 x Green) – Red – Green	Rouse et al., 1974
Difference Vegetation Index	NIR – Red	Jordan, 1969
Modified Green Red Vegetation Index	$(Green^2 - Red^2) / (Green^2 + Red^2)$	Bendig et al., 2015
Normalized Difference Water Index	(Green – NIR) / (Green + NIR)	McFeeters, 1996
Moisture Index	(Narrow NIR – Aerosol) / (Narrow NIR + Aerosol)	After Gao (1996)

Five vegetation indices were selected (Table 3) from 18 which increased overall users' accuracy (compared to no vegetation indices) and were within the top 20 most important variables for the random forest classifier (see Supporting Information S1). These were calculated for both the spring and the summer images separately.

LiDAR data

The tool derives slope and aspect from a digital terrain model, in this case 1 m resolution 2013 Tellus-SW LIDAR (www.tellusgb.ac.uk). Inclusion of these layers increased the overall user's accuracy by 0.3%.

As some of the habitat types are distinguished by elevation, for example upland heathland being above 300 m and lowland heathland below, a layer defining the extent of uplands is derived from the LiDAR based on a user supplied (via pop-up box) definition of uplands.

Peat extent

As degraded peat cannot be reliably identified by the overlying vegetation (Gatis, Luscombe, et al., 2019), it was decided to use previously mapped peat extent to define the area of blanket bog. The tool will run without this layer but this will impact the land cover accuracy for peat discrimination in particular—see Sections 4.1 and 4.3 for greater discussion.

2.3 | Land cover change mapping

2.3.1 | Pixel assessment

As the tool is designed to be used into the future, the photo-interpreted pixels will need to be kept up to date. Manually re-evaluating all 4140 pixels would be prohibitively time consuming. Instead, we have

TABLE 4 Sentinel-2 image acquisition dates for land cover mapping

Year	Spring	Summer
2017	25/01/2017 (L1C data)	24/07/2017
2018	24/02/2018	27/09/2018
2019	20/04/2019	04/07/2019
2020	25/03/2020	23/06/2020
2021	24/04/2021	13/06/2021

developed a workflow to shortlist pixels that may have changed land cover since the last classification. Sentinel-2 images from year 2 are classified using the random forest classifier defined for year 1. The predicted class for year 2 is then compared to the photo-interpreted class in year 1. The pixels where these are different are shortlisted. A random sample of 10% for each class is then output for re-evaluation.

2.3.2 | Change mapping

To detect change, transitional classes, for example cultivated land to forest, can be defined and used in a classifier (Zhang et al., 2019). The number of habitat classes, and therefore number of possible transitions, in this study (484) made the amount of training data required to capture all combinations unfeasible. Conversely, pre-classification methods such as image differencing, image rationing, image overlay and multidate principal component analysis do not require training data. However, post-classification change detection has been shown to be preferable for landscapes with multiple land covers (Chughtai et al., 2021). As such, land cover classification maps were created for the years 2017–2021 using the methods outlined above with Sentinel-2 imagery acquisition dates given in Table 4.

Initially, land cover change between individual years suggested around 20% of pixels were changing. This is ecologically unlikely given 7% of broad land cover classes within DNP were mapped as changed between 1990 and 2015 (Morton et al., 2020). It is more likely much of this change reflects uncertainties in the classification for each year. To reduce this noise, Zhou et al. (2008) used pixel histories over five time points to allocate pixels to unchanged, human or natural change, whilst Souza et al. (2020) added a temporal filter to identify and correct class transitions over a 3–5 year period. In this study, the modal class for two 3-year time frames (2017–2019 vs. 2019–2021) were compared. Where a modal value did not exist (i.e. different class for each year), a no data value was given and therefore no change recorded. The two modal maps were overlain, areas where land cover differed were identified and two maps produced. In the first map, pixels are assigned to the land cover predicted for time frame 2 (habitat gain); in the second map, the pixels are assigned to the land cover predicted for time frame 1 (habitat loss). The probability that the modal pixel value for a time frame is correct is given by the probability that years 1 and 2, or 2 and 3, or 1 and 3, or 1 and 2, and 3 are correct given they are not mutually exclusive. The resultant land cover change accuracy map was the product of the two modal land cover accuracy maps. The difference due to change compared to error was assessed following Fuller et al. (2003).

3 | RESULTS

3.1 | Mapped land cover

Figure 1 shows the land cover map produced using this tool. It covers the entire national park area (954 km^2) and has mapped land cover with a consistent method to a fine spatial resolution (10 m^2) shown in Figure 2.

Mapping suggests (Figure 1; Table 5) that the park is currently dominated by grasslands (>55%) with Modified (24%), Upland acid (18%) and Upland acid over degraded blanket bog (8%) the most common. The next most common land cover is woodland (16%) of which Other broadleaved, mixed and yew woodland (12%) is more common than Coniferous woodland (3%).

Typical of much of the lowland area, Figure 2a is dominated by Modified grassland. However, the land cover map shows patches of Lowland meadows (pale orange) as well as Broadleaf and Coniferous woodland both as contiguous woodland and hedgerows. Figure 2b shows an example of the complex mosaic of land covers often found in semi-natural environments. There is a predominance of Lowland heathland, Lowland Acid grassland and Bracken, but there is also Gorse and the roads (Urban class). The classification scheme used in this study subdivides degraded blanket bog by the overlying vegetation (acid grassland, heathland and unvegetated). Figure 2c shows a complex transition between Blanket bog mosaiced with Unvegetated degraded blanket Bog through Heathland and Acid grassland over degraded blanket bog to Acid grass and Upland Heathland. There are also areas of water and Flushes, Fens, Marsh and Swamp along the streams. The final panel (Figure 2d) shows a typical example of gorse and bracken encroachment around the moorland fringe.

3.2 | Accuracy of the mapped land cover

This tool achieved an overall user's accuracy of 79% with 13 out of 17 classes achieving greater than 70% accuracy and all but gorse achieving at least 60% accuracy (Table 6; Figure 3). Acid grassland and Neutral grassland were commonly confused with the spectrally similar and more geographically dominant Modified grassland (Table 6; Figure 3). Acid Grassland and Heathland, which form a vegetation continuum, were commonly confused when separated into discrete classes. We were asked to map Upland Oak separately as it is an important land cover for DNPA (due to its rarity and conservation status). It has a small training sample size (due to its limited geographical extent) and has very poor spectral distinction from Other broadleaved, mixed and yew woodland (Table 6) and was mis-classified as that for 17 of 18 pixels; however, where it was mapped it was correct.

TABLE 5 Pixel count, error adjusted pixel count and error adjusted area for each land cover class within the national park

Habitat class label	Pixel count	Error-adjusted pixel count	Error-adjusted area (km^2)	Area (%)
1 Lowland dry acid grassland	229,277	224,979	22.5	2
2 Bracken	570,677	568,864	56.9	6
3 Upland Acid grassland	1,754,756	1,721,862	172.2	18
6 Lowland meadows	14,862	239,725	24.0	3
7 Upland hay meadows	4965	80,086	8.0	1
9 Modified grassland	2,731,585	2,258,237	225.8	24
10 Upland oakwood	3892	119,720	12.0	1
11 Other broadleaved, mixed and yew woodland	1,369,494	1,166,275	116.6	12
12 Coniferous woodland	235,381	284,381	28.4	3
13 Lowland heathland	135,007	129,151	12.9	1
14 Upland heathland	581,180	555,970	55.6	6
18 Gorse	104,730	131,392	13.1	1
19 Blanket bog	62,000	89,347	8.9	1
23 Purple moor grass and rush pastures	91,792	116,618	11.7	1
25 Flushes, fens, marsh and swamp	53,952	139,901	14.0	1
26 Cropland	58,009	132,185	13.2	1
27 Urban	197,889	177,485	17.7	2
28 Inland rock	53,406	89,993	9.0	1
29 Rivers and lakes	29,861	77,559	7.8	1
303 Acid grass over degraded blanket bog	795,388	780,478	78.0	8
314 Heathland over degraded blanket bog	412,114	394,238	39.4	4
328 Unvegetated degraded blanket bog	62,019	73,792	7.4	1

TABLE 6 Confusion matrix showing number of pixels predicted by the random forest classifier for each class by the pixels photo-interpreted class

		Photo-interpreted as																	
		Acid grassland	Bracken	Neutral grassland	Modified grassland	Upland oakwood	Other broadleaved, mixed and yew woodland	Coniferous woodland	Heathland	Gorse	Blanket bog	Purple moor grass and rush pastures	Flushes, fens, marsh and swamp	Cropland	Urban	Inland rock	Rivers and lakes	Unvegetated degraded blanket bog	Class-based users accuracy (%)
Predicted as	Acid grassland	191	3	1	5	0	0	0	21	1	0	1	1	1	0	1	0	1	84.1
	Bracken	2	34	0	2	0	2	1	3	1	0	0	0	0	0	0	0	0	75.6
	Neutral grassland	0	0	4	2	0	0	0	0	0	0	0	0	0	0	0	0	0	66.7
	Modified grassland	22	5	34	250	0	3	0	0	1	0	1	0	5	2	0	0	0	77.4
	Upland oakwood	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	100
	Other broadleaved, mixed and yew woodland	0	1	1	6	17	162	6	1	1	0	3	0	0	1	0	2	0	80.6
	Coniferous woodland	1	0	0	0	0	2	37	0	0	0	0	0	0	0	0	0	0	92.5
	Heathland	12	2	0	0	0	0	0	51	2	3	1	6	1	0	0	1	0	64.6
	Gorse	0	1	0	0	0	0	1	2	4	0	0	0	0	0	0	0	0	50
	Blanket bog	0	0	0	0	0	0	0	1	1	9	0	0	0	0	0	0	1	75
	Purple moor grass and rush pastures	0	2	0	1	0	0	0	0	1	0	12	0	2	0	0	0	0	66.7
	Flushes, fens, marsh and swamp	0	0	0	0	0	0	0	2	0	0	0	7	0	0	0	0	0	77.8
	Cropland	0	0	0	1	0	0	0	0	0	0	0	0	11	0	0	0	0	91.7
	Urban	0	0	0	0	0	0	0	0	0	0	0	0	0	15	3	2	0	75
	Inland rock	0	0	0	0	0	0	0	0	0	0	0	0	0	1	9	0	0	90
	Rivers and lakes	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13	0	100
	Unvegetated degraded blanket bog	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	10	90.9
	Overall users accuracy (%)																		79.2

Note: Percentage class based and overall user's accuracies are also given. Correctly predicted pixels are shaded in grey.

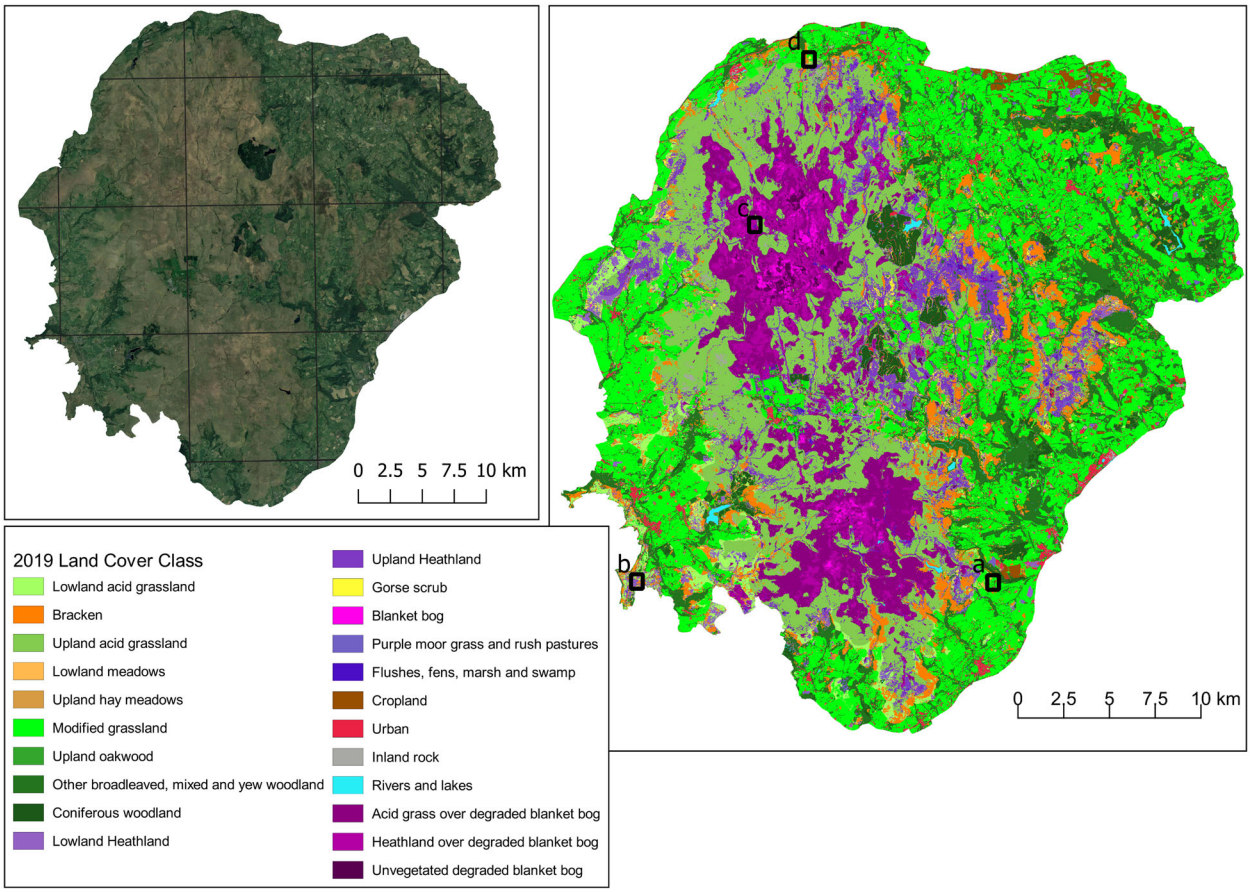

FIGURE 1 True colour aerial Imagery ©2021 Google 2019 and 2019 land cover map for the entire Dartmoor National Park area. Labelled boxes show the extent of the maps in Figure 2.

Accuracy was also assessed spatially using a global logistic regression. It was found that surface reflectance of the blue band in the spring and slope were the best predictors of spatial accuracy (Table 7). This results in the highest accuracies along the steeper banks of rivers, and generally better accuracy in the uplands (dominated by Acid grassland) than the lowlands (dominated by Modified and Neutral grassland) (Figure 4).

The three classes with the greatest User's accuracy (Rivers and lakes: −100%, Coniferous woodland: −92.5% and Cropland: 91.7%) also had the lowest median reflectance in the spring blue band for photo-interpreted pixels 182, 205 and 311, respectively. Whilst the class with the lowest User's accuracy (64.6 %), heathland, had higherreflectance (839).

The results were more mixed for slope with the median slope for the photo-interpreted pixels being both the flattest (1.3°) and steepest (19.4°) for classes with 100% user's accuracy (River and lakes and Upland Oakwood, respectively). However, the number of pixels in these classes is limited due to the limited geographical extent of these land cover types. Generally, for the more extensive land cover classes, higher accuracies occurred for land cover classes that occurred on steeper slopes, in particular woodlands (other broadleaved, mixed and yew woodland [users accuracy 80.6, slope 13.4°]) and coniferous woodland (users accuracy 92.5%, slope 11.5°), whilst land cover classes that occurred on flatter land had lower accuracies, for example heathland (users accuracy 64.6%, slope 5.2°) and neutral grassland (users accuracy 66.7%, slope 6.0°).

3.3 | Mapped land cover change

Comparing the time frames 2017–2019 to 2019–2021, 84% of pixels persisted, 8% changed and 8% were undefined (i.e. did not have a modal value for at least one of the time frames). The largest number of pixels changed in the most commonly occurring classes, that is Modified grassland. Much of the mapped change is single isolated pixels which are more likely to be incorrectly mapped. However, where clumps of pixels show change as in Figure 5, it is more likely to reflect real change. The largest proportional change was observed in Lowland meadows, Upland meadows and Blanket bog, all land cover classes that have been subject to restoration efforts. Figure 5 shows an example of change from Modified grassland to Lowland meadows following meadow restoration (https://moormeadows.org.uk/).

Using modal values over a 3-year time frame, the percent of correctly mapped pixels (as change/no change) was increased from ~50% for a year-on-year analysis to >73%.

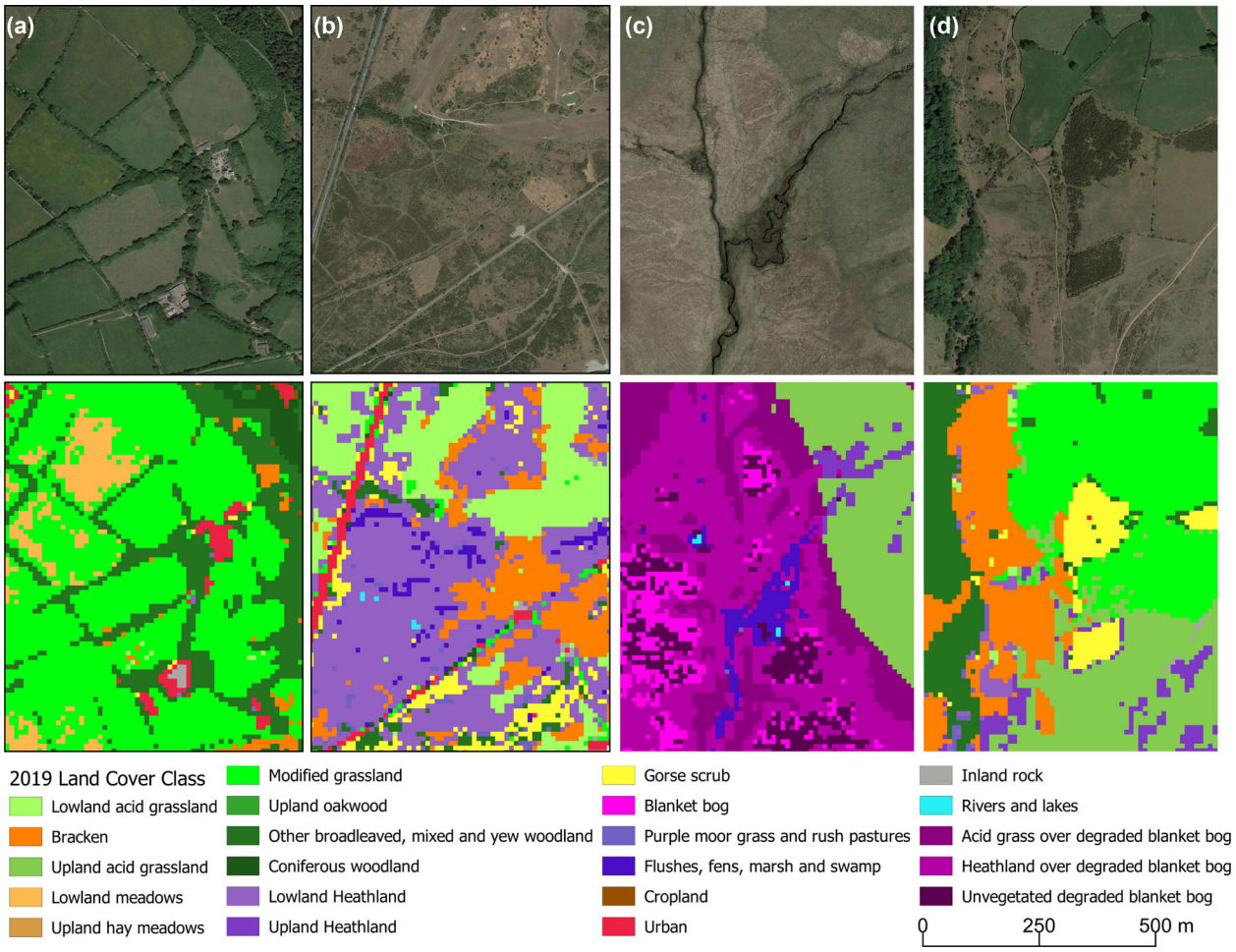

FIGURE 2 Smaller scale (1:10 000) examples from the 2019 land cover map for Dartmoor National Park showing (a) mapped lowland habitat types including Lowland meadows, (b) the complex mosaic of fragmented habitats captured in semi-natural areas, (c) the variation in vegetation cover overlying degraded blanket bog and (d) Gorse and Bracken encroachment on the moorland fridge. Aerial Imagery ©2021 Google

4 | DISCUSSION

4.1 | Mapped land cover

Due to the random nature of a random forest classifier (selection of training pixels, bands and variation in the number of bands selected at each node), the land cover map produced by each run of the tool may differ slightly (unless the random seed is fixed). However, given good quality training data, the tool is robust and consistent.

Figure 2b exemplifies the heterogeneity that can occur in semi-natural environments. Being able to map this complexity is both a strength and weakness. Land cover occurs in small, fragmented patches; the fine pixel size (10 m^2) of this land cover map can resolve land cover patches of this size. This results in a more detailed map; however, this detail can make the map harder to read and therefore decision-making more difficult. It was decided not to filter post-classification as filtering removes both erroneous pixels and correct but small land cover patches.

Blanket bog is a good example of the difficulty of using remote sensing to map land cover using a field survey scheme. Blanket bog is defined by the presence and thickness of soil (in this case, >50 cm peat) beneath the vegetation. In a hydrologically functioning state, blanket bog has distinct plant communities visible to remote sensors. However, in a degraded state blanket bog is overlain by non-peat-forming vegetation typical of other land cover classes. Using the overlying vegetation to define the class would have missed these areas overlying blanket bog, whilst classing these all as bog would have missed the variation in overlying vegetation which requires differing management strategies. In this study, we used peat depth and extent mapped via other remote sensing methods (Gatis, Luscombe, et al., 2019); if these data were not available, it would not be possible to map degraded blanket bog as accurately. In this case, we would suggest mapping the overlying vegetation classes rather than a Degraded blanket bog class.

Of the 133.8 km^2 mapped as blanket bog, a habitat identified as most threatened and requiring conservation (Bain et al., 2011; Gatis, Benaud, et al., 2019), only 8.9 km^2 is not degraded, the remaining is in a degraded state with 7.4 km^2 unvegetated. Although they show a similar geographical distribution this total extent is greater than the 3.6 km^2 of eco-hydrologically functioning peat and 0.9 km^2 of bare peat previously mapped for Dartmoor using higher resolution remote

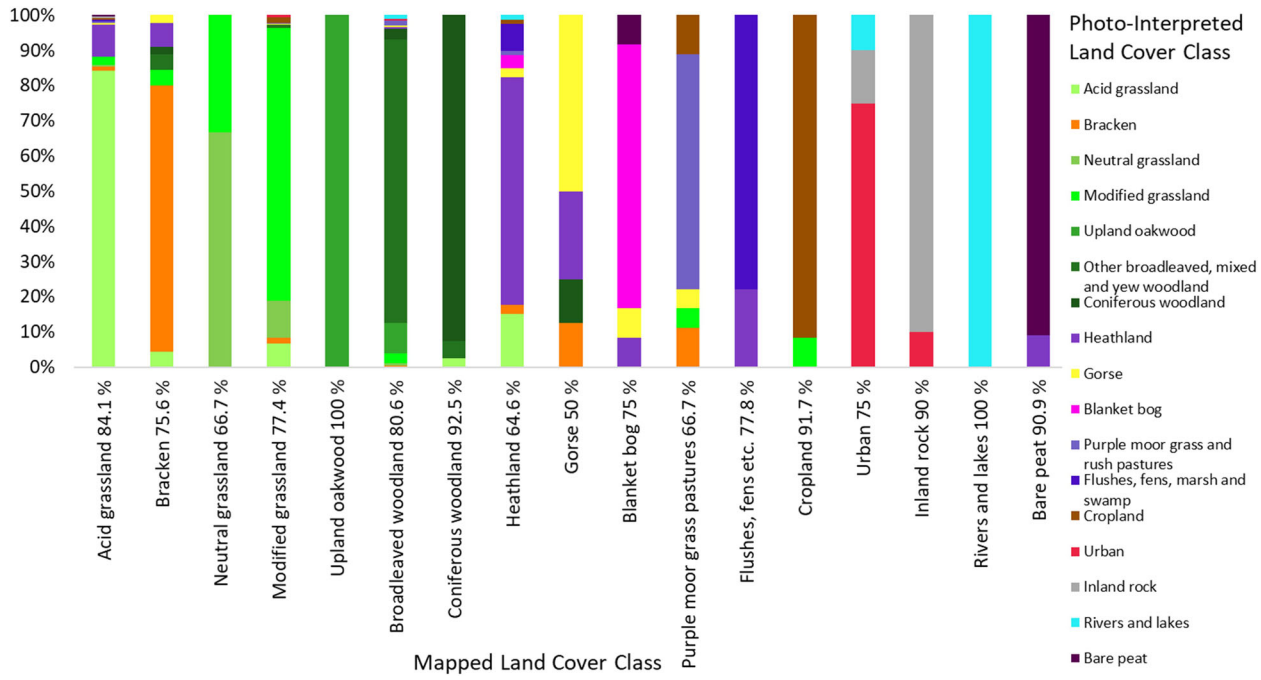

FIGURE 3 User's accuracy (%) for each mapped land cover, along the bottom, coloured by the actual (photo-interpreted) land cover class

TABLE 7 Best models for predicting correct and incorrectly mapped pixels and their Akaike information criterion (AIC) from multiple logistic regression

	AIC	Model
Forward	894.09	B2_Sp + B3_Sp + B4_Sp + B5_Sp + B6_Sp + B7_Sp + B8_Sp +B11_Sp + B8a_Sp + B2_Sum + B3_Sum + B4_Sum + B5_Sum+B6_Sum + B7_Sum + B8_Sum + B11_Sum + B8a_Sum + ExG_Sp +DVI_Sp + NDWI_Sp + MGVRI_Sp + ExG_Sum + DVI_Sum + NDWI_Sum + MGVRI_Sum + Slope + Aspect
Backward	892.17	B2_Sp + Slope
Stepwise	892.17	B2_Sp + Slope

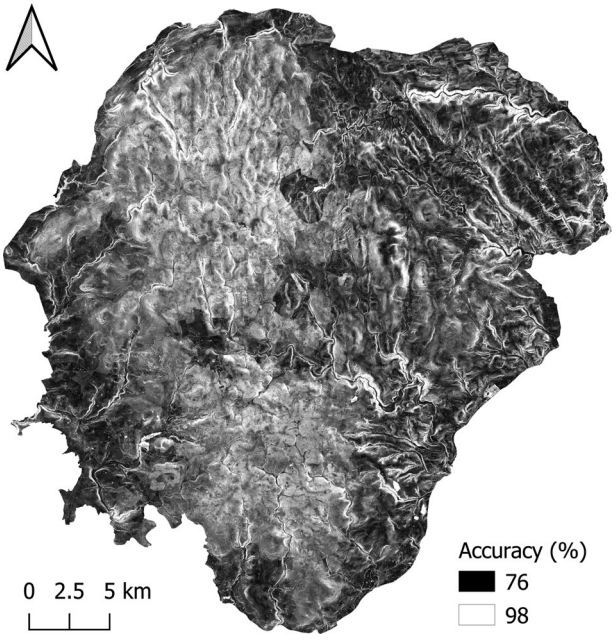

FIGURE 4 Accuracy of the 2019 land cover map (Figure 1)

sensing data sets (Carless et al., 2019), in part due to the larger pixel size. Knowledge of the extent and location of priority habitats will help in the targeting of restoration efforts and in the application to potential funders to support these efforts. Although the rate of change of vegetation post-restoration can be on decadal scales, it is hoped changes in land cover as a result of peatland restoration efforts (Grand-Clement et al., 2015; Parry et al., 2014) will, in time, be discernible with this tool.

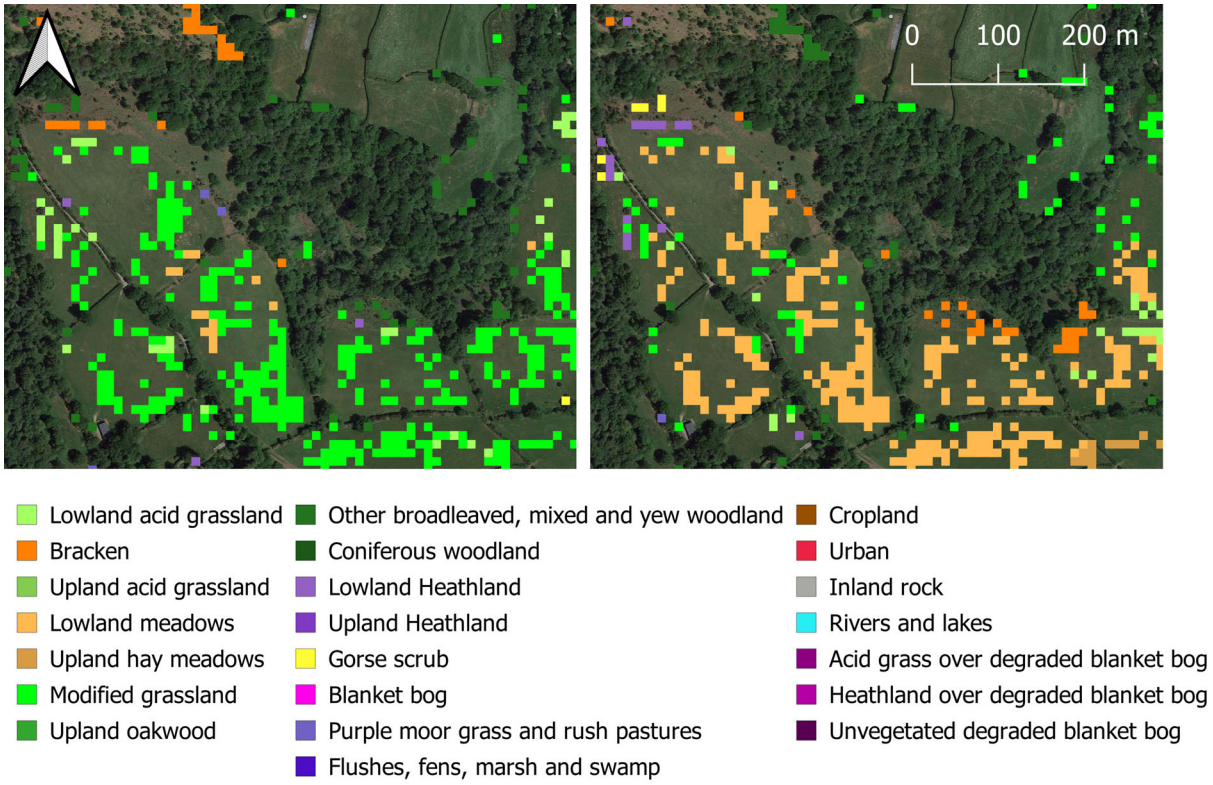

FIGURE 5 Mapped land cover change showing a decrease (on the left) from Modified grassland in the earlier time frame (2017–2019) changing to Lowland meadows (on the right) in the later time frame (2019–2021) in an area undergoing meadow restoration. Aerial Imagery ©2021 Google

Bracken encroachment, evidenced in Figure 2d, has long been a problem (Taylor, 1985) particularly for heathland and upland acid grassland communities. Bracken has a distinct seasonal phenology unveiling green fonds in the spring, rapid summer growth, then dying back to red/brown litter in the autumn/winter which persists into the following spring (Holland & Aplin, 2013). By using images from two seasons, the spring one before green-up and the summer one after, this phenology discriminates bracken from other vegetation. Pakeman et al. (1996) estimated that Dartmoor had 53.6 km^2 of bracken in the 1980s. The land cover map for 2019 estimates bracken to cover a similar area (56.9 km^2) suggesting current management is preventing rapid encroachment. Knowledge of the extent and location of bracken patches and how these change over time enables continued targeted management (cutting, rolling or herbicide application), or identifies potential areas for woodland regeneration, as bracken often grows in formerly wooded locations.

There are two species of gorse common on Dartmoor, European (*Ulex europaeus*) and Western (*Ulex gallii*). Western gorse typically grows to 30–40 cm tall and together with the dwarf-shrubs heather (*Calluna vulgaris*), cross-leaved heath (*Erica tetralix*) and bilberry (*Vaccinium myrtillus*) forms a mosaic of species within the rare and designated (EU Habitats Directive) Upland heathland habitats. In contrast, European gorse grows to 2–3 m tall forming dense, and often laterally extensive areas. Although European gorse is valuable to wildlife and is part of the heathland community, if not managed it may dominate an area. Due to the smaller size of the Western gorse (height and lateral extent) it is most likely that the gorse identified by this land cover map represents European gorse (covering 13.1 km^2). Being able to map the extent (and growth) of gorse patches, as in Figure 2d, enables management to be targeted to areas where gorse may be problematic.

For the most geographically limited classes, for example Upland oakwood and Gorse, classification accuracies are reduced due to the small training sample size. These rare features could be mapped separately (e.g. Foody, Boyd & Sanchez Hernandez, 2007) however this would likely end in over-prediction of these rarer habitats, multiple land cover map layers and for some areas potentially multiple land cover options. This was not considered desirable given the end-users of the maps being produced.

As outlined by the examples above, land management and the strategies required to obtain the desired land covers are unique to each area. The ability to select relevant land cover classes of interest, with some limitation discussed below, amplifies the usefulness and relevance of any land cover map produced.

4.2 | Cross comparison to other land cover maps

The Living Maps Landscape Pioneer (Kilcoyne et al., 2017) which covered Dartmoor National Park and the North Devon Biosphere Reserve achieved overall accuracies of 67% for detailed and 75% for broad

habitats, similar to the 70% achieved by this study (Table 6). They reported difficulties with mosaic or fragmented habitats such as those found in the upland heathland possibly because of object segmentation prior to classification. They reported 5 classes (Lowland Meadows, Scrub, Upland Flushes, Fens and Swamps, Upland Hay Meadows and Upland Heathland) with below 50% accuracy compared to one class (Gorse) in this study. Due to licensing restrictions, we were unable to obtain the data to enable a direct comparison of mapped land covers.

The subsequent national scale product, 'Living England' (Kilcoyne et al., 2022), had different land cover classes. It segmented the landscape, with a minimum mappable unit of 300 m^2, as such it underestimates smaller scale features especially Inland rock (0.1 km^2 compared to 9.0 km^2), Rivers and lakes (1.9 km^2 compare with 7.8 km^2), smaller urban settlements (5.9 km^2 compared with 17.7 km^2) and trees in hedgerows, resulting in a smaller extent of Broadleaved, mixed and yew woodland (89 km^2). See Supporting Information S1 for areal extents of all maps compared. Object based segmentation, oversimplifies the complex mosaic of habitats found in the upland moor. The Bog category has a minimum, mean and maximum object size of 1200, 41,030 and 577,600 m^2, respectively. The authors acknowledge Acid, Calcareous and Neutral Grassland is over-mapped (360 km^2) at the expense of Improved grassland, with 57 km^2 mapped compared to 226 km^2 of Modified grassland in this study. To maintain consistency between Government maps, the Arable and Horticultural land in the Living England map is based on the crop map of England discussed below. Despite these differences, the map produced here, and the Living England map broadly agree on the location and extents of land covers. The differences observed are primarily a result of the spatial scale at which they are intended to be used.

The crop map of England (CROME) (Rural Payments Agency, 2019) poorly represents land covers other than crops, it indicates 81% (783 km^2) of the extent is Grass. Focussing on the area mapped as crops (23.9 km^2) indicates this map under-represents cropland (13.2 km^2), with areas commonly mis-classified as Lowland heathland and Modified grassland. For land management associated with arable and horticultural land use, CROME would be better suited as this is its intended purpose.

LCM2019 (Morton et al., 2020) also uses Sentinel data and a random forest classifier. They acknowledge that their classification scheme is not the most suited to detection by remote sensors, however, it enables a direct comparison with previous mapping to evaluate land cover change. They report overall accuracy of 79.4% and class-based user's accuracies of between 41.4% (Heather grassland) and 97.8% (Fen, Marsh and Swamp) comparable to this study (Table 6). The two maps show similar areal extents (Table 8) and geographical locations of Modified/Improved grassland, Meadows/Neutral grassland and Urban/Urban and Suburban classes. They also show similar extents and geographical location of woodlands, however, the split between broadleaf and conifer is notably different.

This study illustrates a much larger extent of blanket bog (133 km^2 compared to 97 km^2). We believe this reflects the miss-classification of bog in the LCM2019 map due to non-peat forming vegetation, in particular Acid grass and Heath, overlying areas of peat.

Acid grassland is frequently confused with Heathland by this tool (Table 6; Figure 3). To minimize this effect LCM2019 (Morton et al., 2020) has an additional class 'Heath Grassland'. This may have improved classification but would also have required photointerpretation into an additional class. Even if Heath over degraded blanket bog is not included and Heather and Heather grassland are combined this study found a greater area of Heathland around the moorland fringe than the LCM2019 map (72 km^2 compared to 13 km^2). This is most likely due to different points that each classification scheme uses to designate the split between discrete heath/grass classes along the grassland to heathland continuum.

4.3 | Applicability to other studies

The tool has been developed and tuned to the study area to maximize the classification accuracy; lower accuracies may occur in other areas. However, the inputs used—Sentinel bands, LiDAR slope, aspect, elevation and vegetation indices—are common across multiple studies and so sufficient accuracies would be expected.

For areas with degraded blanket bog, it would be ideal to have a layer accurately defining the extent of peat >50 cm deep, however availability of this information is difficult to obtain, and it is not readily available for many sites. The classifier will function without this layer, but it is recommended to map these areas based on their overlying vegetation cover and for the user to be aware of this shortcoming.

The classifier used in this study has LiDAR derived metrics (upland extent, slope, aspect). The availability of LiDAR data is increasing all the time, however, if these are omitted from the tool it will still function. It is however expected that the classification accuracies would decrease.

Due to overestimation of change, a very high proportion of change is correctly mapped. This can guide us to understand appropriate and inappropriate uses of these data. For example, using these data to prioritize areas for field survey to assess areas mapped as bracken encroachment would result in some areas being assessed unnecessarily but would be unlikely to miss areas of encroachment. Conversely, using these data to quantify positive change, for example because of peatland restoration without further field survey, would give an overestimate of success.

This tool has been developed to be as user friendly as possible, requiring basic skills with GIS and a readiness to run, but not alter code, in R. The pre-processing can be done in open-source software, for example QGIS and the tool itself runs in RStudio, also open source. The inputs required are a shapefile outlining the area of interest; a shapefile of peat extent (optional); a digital surface model (optional); spring and summer sentinel images; a shapefile of training/testing pixels and a table outlining the class codes and labels.

It is envisaged that the accuracy and number of photo-interpreted training/testing pixels will be the main constraints on resultant habitat map accuracy. Errors within the training data and small training sample size will result in poorer predictive accuracy. As the tool produces accuracy metrics (confusion matrix and map) this data is available to

TABLE 8 Comparison of areal extents of land cover classes mapped by this study and the CEH LCM2019 (Morton et al., 2020)

This study		CEH LCM2019	
Class label	Areal extent (km^2)	Areal extent (km^2)	Class label
Lowland acid grassland	22.9	372.6	Acid Grassland
Upland acid grassland	57.1		
Bracken	175.5		
		0.1	Calcareous Grassland
Lowland meadows	1.5	1.9	Neutral Grassland
Upland hay meadows	0.5		
Modified grassland	273.2	271.3	Improved Grassland
Upland oakwood	0.4	122.3	Broadleaved woodland
Other broadleaved, mixed and yew woodland	136.9		
Coniferous woodland	23.5	34.5	Coniferous woodland
Lowland heathland	13.5	12.7	Heather
Upland heathland	58.1	0.1	Heather grassland
Gorse scrub	10.5		
Blanket bog	6.2	97.4	Bog
Acid grass over degraded blanket bog	79.5		
Heathland over degraded blanket bog	41.2		
Unvegetated degraded blanket bog	6.2		
Purple moor grass and rush pastures	9.2		
Flushes, fens, marsh and swamp	5.4	0.0	Fen, Marsh and Swamp
Cropland	5.8	22.9	Arable and Horticulture
Urban	19.8	0.8	Urban
	0.0	18.2	Suburban
Inland rock	5.3	0.1	Inland Rock
Rivers and lakes	3.0	0.7	Freshwater

the user to determine if the maps produced are sufficient for their intended use.

5 | CONCLUSION

Using Dartmoor National Park as an example, we demonstrate that it is possible to produce an open-source tool using freely available data to map land cover and land cover change at a fine resolution (10 m^2) over a landscape extent (954 km^2). The method is robust, repeatable and designed to be usable by non-specialists. The tool enables those taking land management decisions to choose the land cover classes that are important to their management objectives to maximize the potential of these data.

Access to detailed, up-to-date, accurate and tailored land cover maps will transform evidence-based landscape management decision-making where this was previously impossible due to the cost or technical workflows required to produce these maps.

ACKNOWLEDGEMENTS

The authors would like to thank Benjamin Beaumont and an anonymous reviewer for their thorough and helpful reviews. We would also like to thank Sara Zonneveld, Rebecca Abrahams and the wider SWEEP team for all the help throughout the project, as well as our partners within Dartmoor National Park Authority whose feedback helped guide the development of this tool. This work was supported by Dartmoor National Park Authority with funding from the Defra Environmental Land Management Schemes Test & Trials and the South West Partnership for Environmental and Economic Prosperity (SWEEP023) which was funded by the Natural Environment Research Council (NE/P011217/1).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

All authors conceived of the presented idea, discussed the results and contributed to the final manuscript. Namoi Gatis, Donna Carless and David Luscombe photo-interpreted pixels. Namoi Gatis wrote the R-code and analysed the data. Karen Anderson and Richard Brazier supervised the project. Richard Brazier was PI on the funding proposal which was also supported by David Luscombe, Naomi Gatis, Donna Carless and Karen Anderson.

DATA AVAILABILITY STATEMENT

The full reproducible code and input data are available at https://doi.org/10.5281/zenodo.5797735 (Gatis et al., 2021). Maps and data tables produced from this tool are openly available from the University of Exeter's institutional repository https://doi.org/10.24378/exe.4044 (Gatis et al., 2022). The availability and accuracy of these maps can impact on a range of policy areas, such as agri-environment schemes (e.g. Environmental Land Management Schemes (Department for Environment Food & Rural Affairs & Rural Payments Agency, 2021) and the Common Agricultural Policy); targeted Land-Use, Land-Use Change and Forestry (Watson et al., 2000) including peatland restoration and afforestation; conservation and urban development.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1002/2688-8319.12162.

ORCID

Naomi Gatis https://orcid.org/0000-0002-0996-5568

REFERENCES

Adam, E., Mutanga, O., Odindi, J., & Abdel-Rahman, E. M. (2014). Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: Evaluating the performance of random forest and support vector machines classifiers. *International Journal of Remote Sensing*, *35*(10), 3440–3458. https://doi.org/10.1080/01431161.2014.903435

Aguilar, M. A., Saldaña, M. M., & Aguilar, F. J. (2013). GeoEye-1 and WorldView-2 pan-sharpened imagery for object-based classification in urban environments. *International Journal of Remote Sensing*, *34*(7), 2583–2606. https://doi.org/10.1080/01431161.2012.747018

Azzari, G., & Lobell, D. B. (2017). Landsat-based classification in the cloud: An opportunity for a paradigm shift in land cover monitoring. *Remote Sensing of Environment*, *202*, 64–74. https://doi.org/10.1016/j.rse.2017.05.025

Bain, C. G., Bonn, A., & Chapman, R. (2011). *IUCN UK Peatland Programme*. IUCN UK Commission of Inquiry on Peatlands.

Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., Gnyp, M. L., & Bareth, G. (2015). Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *International Journal of Applied Earth Observation and Geoinformation*, *39*, 79–87. https://doi.org/10.1016/j.jag.2015.02.012

Bivand, R., Keitt, T., & Rowlingson, B. (2021). *rgdal: Bindings for the "Geospatial" Data Abstraction Library*. R package version 1.5-23. https://CRAN.R-project.org/package=rgda

Bivand, R., Rundel, C., Pebesma, E., Stuetz, R., Hufthammer, K. O., Giraudoux, P., & Santilli, S. (2020). *rgeos: Interface to Geometry Engine - Open Source ('GEOS'): Version 05.-5*. https://CRAN.R-project.org/package=rgeost

Blackburn, G. A., & Pitman, J. I. (1999). Biophysical controls on the directional spectral reflectance properties of bracken (*Pteridium aquilinum*) canopies: Results of a field experiment. *International Journal of Remote Sensing*, *20*(11), 2265–2282. https://doi.org/10.1080/014311699212245

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brown, J. F., Tollerud, H. J., Barber, C. P., Zhou, Q., Dwyer, J. L., Vogelmann, J. E., & Rover, J. (2020). Lessons learned implementing an operational continuous United States national land change monitoring capability: The Land Change Monitoring, Assessment, and Projection (LCMAP) approach. *Remote Sensing of Environment*, *238*, 111356. https://doi.org/10.1016/J.RSE.2019.111356

Carleer, A. P., & Wolff, E. (2006). *Region-based classification potential for land-cover classification with very high spatial resolution satellite data*. Proceedings of 1st International Conference on Obgect-Based Image Analysis (OBIA 2006). http://citeseerx.ist.psu.edu/viewdoc/download?rep=rep1&type=pdf&doi=10.1.1.222.4764

Carless, D., Luscombe, D. J., Gatis, N., Anderson, K., & Brazier, R. E. (2019). Mapping landscape-scale peatland degradation using airborne lidar and multispectral data. *Landscape Ecology*, *34*(6), 1329–1345. https://doi.org/10.1007/s10980-019-00844-5

Carrasco, L., O'Neil, A. W., Daniel Morton, R., & Rowland, C. S. (2019). Evaluating combinations of temporally aggregated Sentinel-1, Sentinel-2 and Landsat 8 for land cover mapping with Google Earth Engine. *Remote Sensing*, *11*(3), 288. https://doi.org/10.3390/rs11030288

Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., …, & Mills, J. (2015). Global land cover mapping at 30 m resolution: A POK-based operational approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, *103*, 7–27. https://doi.org/10.1016/J.ISPRSJPRS.2014.09.002

Chughtai, A. H., Abbasi, H., & Karas, I. R. (2021). A review on change detection method and accuracy assessment for land use land cover. *Remote Sensing Applications: Society and Environment*, *22*, 100482. https://doi.org/10.1016/j.rsase.2021.100482

Close, O., Benjamin, B., Petit, S., Fripiat, X., & Hallot, E. (2018). Use of Sentinel-2 and LUCAS database for the inventory of land use, land use change, and forestry in Wallonia, Belgium. *Land*, *7*(4), 154. https://doi.org/10.3390/LAND7040154

Cole, B., McMorrow, J., & Evans, M. (2014). Spectral monitoring of moorland plant phenology to identify a temporal window for hyperspectral remote sensing of peatland. *ISPRS Journal of Photogrammetry and Remote Sensing*, *90*(0), 49–58. https://doi.org/10.1016/j.isprsjprs.2014.01.010

Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, *37*(1), 35–46. https://doi.org/10.1016/0034-4257(91)90048-B

Congedo, L. (2021). Semi-Automatic Classification Plugin: A Python tool for the download and processing of remote sensing images in QGIS. *Journal of Open Source Software*, *6*(64), 3172. https://doi.org/10.21105/joss.03172

Department for Environment Food & Rural Affairs, & Rural Payments Agency. (2021). *Environmental land management schemes: Overview*. https://www.gov.uk/government/publications/environmental-land-management-schemes-overview/environmental-land-management-scheme-overview

Esetlili, M. T., Bektas Balcik, F., Balik Sanli, F., Ustuner, M., Kalkan, K., Goksel, C., & Kurucu, Y. (2018). Comparison of object and pixel-based classifications for mapping crops using Rapideye imagery: A case study of Menemen Plain, Turkey. *International Journal of Environment and Geoinformatics*, *5*(2), 231–243. https://doi.org/10.30897/ijegeo.442002

Foody, G. M., Boyd, D. S., & Sanchez-Hernandez, C. (2007). Mapping a specific class with an ensemble of classifiers. *International Journal of Remote Sensing*, *28*(8), 1733–1746. https://doi.org/10.1080/01431160600962566

Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., Strahler, A. H., & Schaaf, C. (2002). Global land cover mapping from MODIS: Algorithms and early results. *Remote Sensing of Environment*, *83*(1–2), 287–302. https://doi.org/10.1016/S0034-4257(02)00078-0

Fuller, R. M., Smith, G. M., & Devereux, B. J. (2003). The characterisation and measurement of land cover change through remote sensing: Problems in operational applications?. *International Journal of Applied Earth Observation and Geoinformation*, *4*(3), 243–253. https://doi.org/10.1016/S0303-2434(03)00004-7

Gao, B. C. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, *58*(3), 257–266. https://doi.org/10.1016/S0034-4257(96)00067-3

Gatis, N., Benaud, P., Ashe, J., Luscombe, D. J., Grand-Clement, E., Hartley, I. P., & Brazier, R. E. (2019). Assessing the impact of peat erosion on growing season CO_2 fluxes by comparing erosional peat pans and surrounding vegetated haggs. *Wetlands Ecology and Management*, *27*(2–3), 187–205. https://doi.org/10.1007/s11273-019-09652-9

Gatis, N., Carless, D., Luscombe, D. J., Brazier, R. E., & Anderson, K. (2021). *An operational land cover and land cover change toolbox; processing open-source data with open-source software (R Code)*. Zenodo. https://doi.org/10.5281/zenodo.5797735

Gatis, N., Carless, D., Luscombe, D. J., Brazier, R. E., & Anderson, K. (2022). *An operational land cover and land cover change toolbox; processing open-source data with open-source software (dataset)*. University of Exeter. https://doi.org/10.24378/exe.4044

Gatis, N., Luscombe, D. J., Carless, D., Parry, L., E, F., R, M., Harrod, T. R., & Anderson, K. (2019). Mapping upland peat depth using airborne radiometric and LiDAR survey data. *Geoderma*, *335*, 78–87. https://doi.org/10.1016/j.geoderma.2018.07.041

Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., & Chen, J. (2013). Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data. *International Journal of Remote Sensing*, *34*(7), 2607–2654. https://doi.org/10.1080/01431161.2012.748992

Grand-Clement, E., Anderson, K., Smith, D., Angus, M., Luscombe, D. J., Gatis, N., & Brazier, R. E. (2015). New approaches to the restoration of shallow marginal peatlands. *Journal of Environmental Management*, *161*, 417–430. https://doi.org/10.1016/j.jenvman.2015.06.023

Grippa, T., Lennert, M., Beaumont, B., Vanhuysse, S., Stephenne, N., & Wolff, E. (2017). An open-source semi-automated processing chain for urban object-based classification. *Remote Sensing*, *9*(4), 358. https://doi.org/10.3390/rs9040358

Grosjean, P. (2021). *SciViews-R*. UMONS. http://www.sciviews.org/SciViews-R

Gutman, G. G., Ignatov, A. M., & Olson, S. (1994). Towards better quality of AVHRR composite images over land: Reduction of cloud contamination. *Remote Sensing of Environment*, *50*(2), 134–148. https://doi.org/10.1016/0034-4257(94)90040-X

Hashim, H., Abd Latif, Z., & Adnan, N. A. (2019). Urban vegetation classification with NDVI threshold value method with Very High Resolution (VHR) Pleiades imagery. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *42*(4/W16), 237–240. https://doi.org/10.5194/isprs-archives-XLII-4-W16-237-2019

Hay, A. M. (1979). Sampling designs to test land-use map accuracy. *Photogrammetric Engineering and Remote Sensing*, *45*(4), 529–533.

Heymann, Y., Steenmans, C., Croiselle, G., & Bossard, M. (1994). *CORINE land cover: Technical guide*. European Commission, Directorate-General, Environment, Nuclear Safety and Civil Protection.

Hijmans, R. J. (2020). *raster: Geographic data analysis and modeling*. R package version 3.4-5. https://CRAN.R-project.org/package=raster

Holland, J., & Aplin, P. (2013). Super-resolution image analysis as a means of monitoring bracken (*Pteridium aquilinum*) distributions. *ISPRS Journal of Photogrammetry and Remote Sensing*, *75*, 48–63. https://doi.org/10.1016/J.ISPRSJPRS.2012.10.002

Jordan, C. F. (1969). Derivation of leaf-area index from quality of light on the forest floor. *Ecology*, *50*(4), 663–666. https://doi.org/10.2307/1936256

Karasiak, N. (2021). *Dzetsaka QGIS plugin*. https://github.com/nkarasiak/dzetsaka

Karra, K., Kontgis, C., Statman-Weil, Z., Mazzariello, J., Mathis, M., & Brumby, S. (2021). *Global land use/land cover with Sentinel-2 and deep learning*. IGARSS 2021-2021 IEEE International Geoscience and Remote Sensing Symposium, IEEE, Brussels.

Khatami, R., Mountrakis, G., & Stehman, S. V. (2017). Mapping per-pixel predicted accuracy of classified remote sensing images. *Remote Sensing of Environment*, *191*, 156–167. https://doi.org/10.1016/j.rse.2017.01.025

Kilcoyne, A., Clement, M., Moore, C., Phillipps, G. P., Keane, R., Potter, S., Stefaniak, A., & Trippier, B. (2022). *Living England: Satellite-based habitat classification*. Technical User Guide. Natural England.

Kilcoyne, A. M., Alexander, R., Cox, P., & Brownett, J. (2017). *Living Maps: Satellite-based habitat classification*. Evidence Project SD1705. Lincoln.

Küchler, A. W., & Zonneveld, I. S. (1988). The UNESCO classification of vegetation. In Küchler, A. W., & Zonneveld, I. S. (Ed.), *Vegetation mapping. Handbook of vegetation science* (pp. 531–549). Springer. https://doi.org/10.1007/978-94-009-3083-4_45

Kuhn, M. (2020). *caret: Classification and regression training*. R package version 6.0-86. https://CRAN.R-project.org/package=caret

Liaw, A., & Weiner, M. (2002). Classification and regression by randomForest. *R News*, *2*(3), 18–22.

Liu, H., Gong, P., Wang, J., Wang, X., Ning, G., & Xu, B. (2021). Production of global daily seamless data cubes and quantification of global

land cover change from 1985 to 2020 - iMap World 1.0. *Remote Sensing of Environment*, *258*, 112364. https://doi.org/10.1016/j.rse.2021.112364

Lück, W., & van Niekerk, A. (2016). Evaluation of a rule-based compositing technique for Landsat-5 TM and Landsat-7 ETM+ images. *International Journal of Applied Earth Observation and Geoinformation*, *47*, 1–14. https://doi.org/10.1016/j.jag.2015.11.019

Mahto, A. (2019). *splitstackshape: Stack and reshape datasets after splitting concatenated values*. R package version 1.4.8. https://CRAN.R-project.org/package=splitstackshape

McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, *17*(7), 1425–1432. https://doi.org/10.1080/01431169608948714

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. (2020). *e1071: Misc functions of the department of statistics, probability theory group (Formerly: E1071), TU Wien*. R package version 1.7-4. https://CRAN.R-project.org/package=e1071

Microsoft Corporation, & Weston, S. (2020). *doParallel: Foreach parallel adaptor for the "parallel" package*. R package version 1.0.16. https://CRAN.R-project.org/package=doParallel

Moody, D. I., Bauer, D. E., Brumby, S. P., Chisolm, E. D., Warren, M. S., Skillman, S. W., & Keisler, R. (2016). *Land cover classification in fused multisensor multispectral satellite imagery*. Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation, 2016-April, pp. 85–88. https://doi.org/10.1109/SSIAI.2016.7459181

Morton, R. D., Marston, C. G., O'Neil, A. W., & Rowland, C. S. (2020a). *Land Cover Map 2017 (25m Rasterised Land Parcels, GB)*. NERC Environmental Information Data Centre. https://doi.org/10.5285/499212CD-D64A-43BA-B801-95402E4D4098

Morton, R. D., Marston, C. G., O'Neil, A. W., & Rowland, C. S. (2020b). *Land Cover Change 1990–2015 (25m raster, GB)*. NERC Environmental Information Data Centre. https://doi.org/10.5285/07b6e5e9-b766-48e5-a28c-5b3e35abecc0

Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*, 247–259. https://doi.org/10.1016/j.isprsjprs.2010.11.001

Müller, K. (2020). *here: A simpler way to find your files*. R package version 1.0.1. https://CRAN.R-project.org/package=here

Murray, N. J., Keith, D. A., Simpson, D., Wilshire, J. H., & Lucas, R. M. (2018). Remap: An online remote sensing application for land cover classification and monitoring. *Methods in Ecology and Evolution*, *9*(9), 2019–2027. https://doi.org/10.1111/2041-210X.13043

Noi, P. T., & Kappas, M. (2018). Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery. *Sensors*, *18*(1), 18. https://doi.org/10.3390/S18010018

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., & Wulder, M. A. (2014). Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, *148*, 42–57. https://doi.org/10.1016/j.rse.2014.02.015

Pakeman, R. J., Marrs, R. H., Howard, D. C., Barr, C. J., & Fuller, R. M. (1996). The bracken problem in Great Britain: Its present extent and future changes. *Applied Geography*, *16*(1), 65–86. https://doi.org/10.1016/0143-6228(95)00026-7

Parry, L. E., Holden, J., & Chapman, P. J. (2014). Restoration of blanket peatlands. *Journal of Environmental Management*, *133*(0), 193–205. https://doi.org/10.1016/j.jenvman.2013.11.033

Pontius, R. G., Jr., & Santacruz, A. (2019). *diffeR: Metrics of difference for comparing pairs of maps or pairs of variables*. R package version 0.0-6. https://CRAN.R-project.org/package=diffeR

QGIS Development Team. (2020). *QGIS geographic information system*. Open Source Geospatial Foundation Project. http://qgis.osgeo.org

R Core Team. (2020). *R: A language and environment for statistical computing v4.1.2*. https://www.r-project.org

Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *67*(1), 93–104. https://doi.org/10.1016/j.isprsjprs.2011.11.002

Rouse, J. J. W., Haas, R. H., Deering, D. W., Schell, J. A., & Harlan, J. C. (1974). *Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation* (No. E75-10354). NASA.

Royimani, L., Mutanga, O., Odindi, J., Dube, T., & Matongera, T. N. (2019). Advancements in satellite remote sensing for mapping and monitoring of alien invasive plant species (AIPs). *Physics and Chemistry of the Earth, Parts A/B/C*, *112*, 237–245. https://doi.org/10.1016/J.PCE.2018.12.004

RStudio Team. (2020). *RStudio: Integrated development for R*. RStudio v1.4.1106. https://www.rstudio.com/

Rural Payments Agency. (2019). *Crop Map of England (CROME) 2019*. https://data.gov.uk/dataset/8c5b635f-9b23-4f32-b12a-c080e3f455d0/crop-map-of-england-crome-2019

Saah, D., Tenneson, K., Matin, M., Uddin, K., Cutter, P., Poortinga, A., … Chishtie, F. (2019). Land cover mapping in data scarce environments: Challenges and opportunities. *Frontiers in Environmental Science*, *7*, 150. https://doi.org/10.3389/FENVS.2019.00150

Saralioglu, E., & Gungor, O. (2022). Semantic segmentation of land cover from high resolution multispectral satellite images by spectral-spatial convolutional neural network. *Geocarto International*, *37*(2), 657–677. https://doi.org/10.1080/10106049.2020.1734871

Souza, C. M., Shimbo, J. Z., Rosa, M. R., Parente, L. L., Alencar, A. A., Rudorff, B. F. T., & Azevedo, T. (2020). Reconstructing three decades of land use and land cover changes in Brazilian biomes with landsat archive and earth engine. *Remote Sensing*, *12*(17), 2735. https://doi.org/10.3390/RS12172735

Sozzi, M., Marinello, F., Pezzuolo, A., & Sartori, L. (2018). *Benchmark of satellites image services for precision agricultural use*. Agricultural Engineering Conference. https://www.researchgate.net/publication/326417596

Taylor, J. A. (1985). Bracken encroachment rates in Britain. *Soil Use and Management*, *1*(2), 53–56. https://doi.org/10.1111/j.1475-2743.1985.tb00656.x

Ushey, K., McPherson, J., Cheng, J., Atkins, A., & Allaire, J. (2018). *packrat: A dependency management system for projects and their R package dependencies*. R package version 0.5.0. https://CRAN.R-project.org/package=packrat

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). Springer.

Vogelmann, J. E., Howard, S. M., Yang, L., Larson, C. R., Wylie, B. K., & Van Driel, N. (2001). Completion of the 1990s National Land Cover Data set for the conterminous United States from Landsat thematic mapper data and ancillary data sources. *Photogrammetric Engineering and Remote Sensing*, *67*(6), 650–662.

Watson, R. T., Noble, I. R., Bolin, B., Ravindranath, N. H., Verardo, D. J., & Dokken, D. J. (2000). *Land use, land-use change and forestry: A special report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

Woodcock, C. E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., & Wynne, R. (2008). Free access to landsat imagery. *Science*, *320*, 1011. https://doi.org/10.1126/science.320.5879.1011a

Woodcock, C. E., & Strahler, A. H. (1987). The factor of scale in remote sensing. *Remote Sensing of Environment*, *21*(3), 311–332. https://doi.org/10.1016/0034-4257(87)90015-0

Wulder, M. A., Coops, N. C., Roy, D. P., White, J. C., & Hermosilla, T. (2018). Land cover 2.0. *International Journal of Remote Sensing*, *39*, 4254–4284. https://doi.org/10.1080/01431161.2018.1452075

Xing, H., Hou, D., Wang, S., Yu, M., & Meng, F. (2021). O-LCMapping: A Google Earth Engine-based web toolkit for supporting online land cover classification. *Earth Science Informatics*, *14*(1), 529–541. https://doi.org/10.1007/s12145-020-00562-6

Yuan, F., Sawaya, K. E., Loeffelholz, B. C., & Bauer, M. E. (2005). Land cover classification and change analysis of the Twin Cities (Minnesota) Metropolitan Area by multitemporal Landsat remote sensing. *Remote Sensing of Environment*, *98*(2–3), 317–328. https://doi.org/10.1016/J.RSE.2005.08.006

Zhang, J., Zhang, W., Mei, Y., & Yang, W. (2019). Geostatistical characterization of local accuracies in remotely sensed land cover change categorization with complexly configured reference samples. *Remote Sensing of Environment*, *223*, 63–81. https://doi.org/10.1016/j.rse.2019.01.008

Zhou, Q., Li, B., & Kurban, A. (2008). Trajectory analysis of land cover change in arid environment of China. *International Journal of Remote Sensing*, *29*(4), 1093–1107. https://doi.org/10.1080/01431160701355256

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Table 1 Users Accuracy for different numbers and combinations of images from 2019 used as inputs to the random forest classifier

Table 2 Vegetation indices and Sentienl-2 bands trialed as inputs to the random forest classifier with the resultant Overall User's Accuracy and parameter importance (scaled as a percent of the most important parameter)

Table 3 Vegetation indices (and Sentienl-2 bands used) trialed as inputs to the random forest classifier with the resultant Overall User's Accuracy and parameter importance (scaled as a percent of the most important parameter)

Table 4 Overall User's accuracy results for the random forest classifier with sampling techniques applied to account for sample imbalance

Figure 1 Ranked Users Accuracy (by year) based on the month of image capture of the spring (left) and summer (right) images.

Figure 2 Effect on overall user's accuracy of including aspect and/or slope as inputs to the random forest classifier

Figure 1 Search page of Sentinel-hub

Figure 2 Close up of results panel with information on the Data (Sentinel-2 L2A), image date (2020-04-24), image time (11:27:02 UTC), cloud cover (23.8% and tile (30UVB) and the link button.

Figure 3 Examples of good (2020-03-25), moderate (2020-04-09) and too cloudy (2020-01-15, 2020-04-04) images.

Figure 4 Link button to reveal SciHub link.

Figure 5 Selecting all layers as Input Layers for the GDAL Merge tool

Figure 6 Input parameters for the Merge tool when creating a multi-band image from the 9 layers

Figure 7 Input parameters for the Merge tool when mosaicking the two merged multi-band images

Figure 8 Installing packrat package

Figure 9 Code pasted into console

Figure 10 LandCoverClassTool folder with project highlighted

Figure 11 Habitat Classification Tool open in RStudio, source button is highlighted

Figure 12 Adding the Pixels to Check delineated text file

Figure 13 Selecting pixels in the photo-interpreted pixels layer based on the location of pixels in the pixels to check layer

Figure 14 Pixel to check, was classed as 26 Cropland, still Cropland so no amendment required, moving to next pixel

Table 2 Example of csv table defining class codes (first column) and class names (second column)

How to cite this article: Gatis, N., Carless, D., Luscombe, D. J., Brazier, R. E., & Anderson, K. (2022). An operational land cover and land cover change toolbox: processing open-source data with open-source software. *Ecological Solutions and Evidence*, *3*, e12162. https://doi.org/10.1002/2688-8319.12162