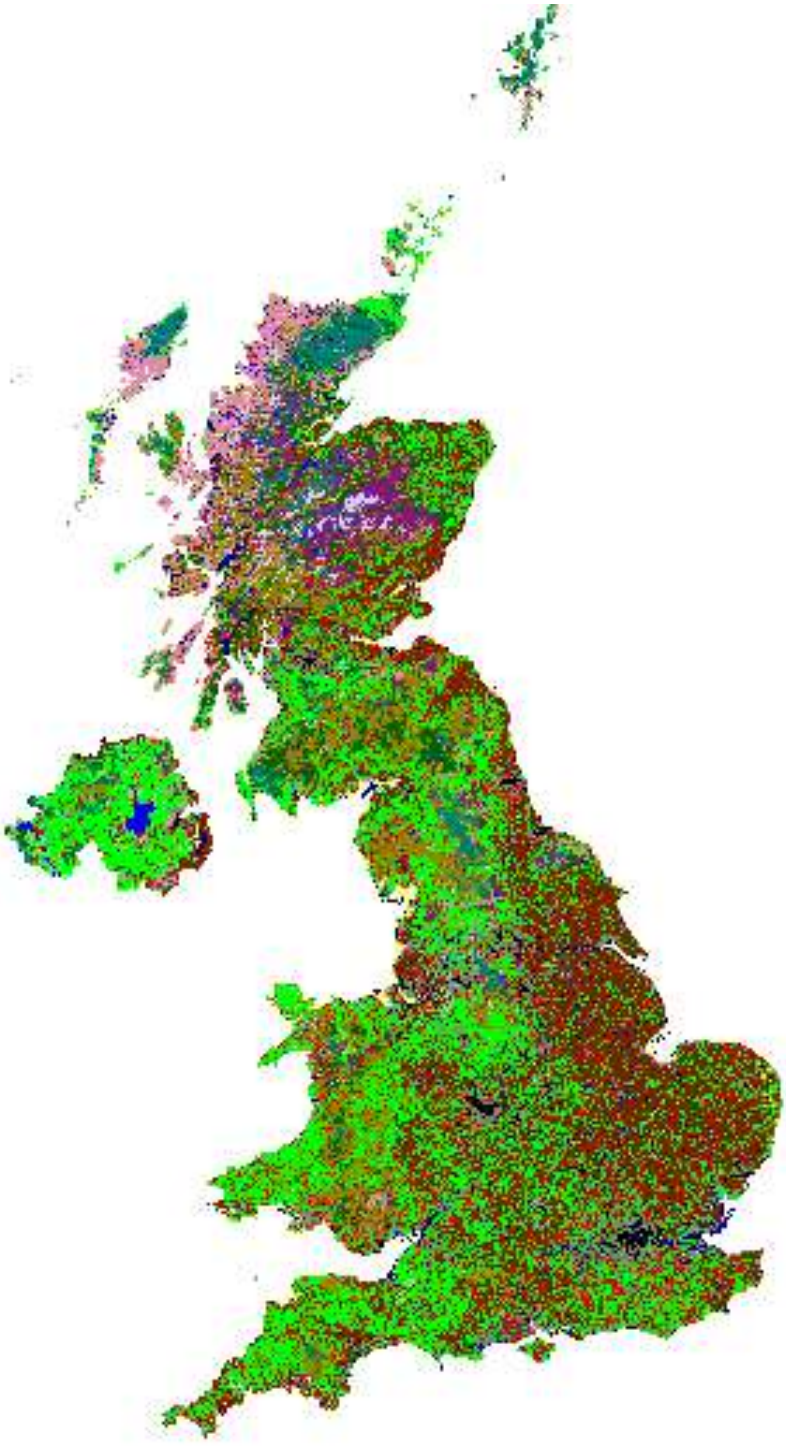
The UKCEH Land Cover Maps for 2017, 2018 and 2019



v1.5.1	30/06/2020

Purpose

This is the user guide for three new UKCEH Land Cover Maps (LCMs): LCM2017, LCM2018 and LCM2019. Each new UKCEH LCM represents a range of geospatial datasets. The content and structure of each are explained. Appropriate use of complex data requires a deeper understanding than structure and content. We therefore describe how the data were produced and validated, some of the reasons behind production decisions and outline some future plans. Our goal is to help users make informed decisions regarding the application UKCEH LCM data in their current and future work.

Contents

Glossary

Bootstrap Training	A method that automatically selects training observations from an historical land cover map. Bootstrap Training datasets are used by a Random Forest to classify Classification Scenes.
Context Raster:	A multi-band raster with contextual information, such as terrain, coastal and urban proximity. Context Rasters are combined with Seasonal Composite Images to create Classification Scenes.
Classification Scene:	A multi-layer raster comprising spectral information from a Seasonal Composite Image and Context Rasters. Classification Scenes are classified to yield a pixel land cover classification.
Seasonal Composite Image	A multi-band raster representing spectral observations across four seasons. These are designed to provide seasonal phenology which a classifier can use to differential vegetation types.
UKCEH Aggregate Land Cover Classes	A set of land cover classes that generalise the detail provided in UK CEH Land Cover Classes.
UKCEH Land Cover Classes	A set of land cover classes defined by UKCEH, derived from the UK Biodiversity Action Plan Broad Habitats (Jackson *et. al*. 2000)
UKCEH Land Parcel Spatial Framework	A polygon database of land parcel objects derived from generalising national cartography in order to remove unnecessary detail. All UKCEH LCM Land Parcel datasets use this spatial framework.

Introduction

UKCEH has previously released four national land cover maps (LCMs): The Land Cover Map of Great Britain for 1990 (usually referred to as LCM1990) and UK land cover maps LCM2000, LCM2007 and LCM2015. This document accompanies the release of three new UKCEH LCMs to extend this series: LCM2017, LCM2018 and LCM2019. Land cover in the new products is given as 21 UKCEH Land Cover Classes based upon Biodiversity Action Plan (BAP) Broad Habitats (Jackson, 2000) and match classes used for UKCEH LCM2015. UKCEH LCMs 2017, 2018 and 2019 were created automatically using a technique we call Bootstrap Training combined with a Random Forest classifier. For each new map we classified Sentinel-2 Seasonal Composite Images generated using the Google Earth Engine representing median reflectance per season. Before classification these images were combined with ten Context Layers, which helped to reduce spectral confusion, to create Classification Scenes. The new automatic methods enable us to produce UK-wide land cover maps very rapidly. Historically, because of high costs associated with labour intensive image pre-processing and gathering training data, there has been lag of approximately 10 years between UKCEH LCMs. Going forward we intend to release a new LCM every year.

Errors occur in all land cover products and for earlier UKCEH LCMs we performed significant manual corrections in regions where classification errors were obvious or prolific. There were no manual corrections this time. Classification is automatic and we cannot achieve annual maps if we undertake significant manual steps. Annual maps are important to understand land cover change. Classification errors are usually randomly distributed in space, due for example to vagaries in image quality or viewing conditions, so the same errors should not occur in the same location year after year. Therefore, as the time series of maps develops real land cover changes should persist and be visible against background noise.

Despite the lack of manual accuracy corrections be assured that visual checks and formal validation exercises were performed on the new land cover maps (Appendix 4). These indicate that the new maps are of similar quality to the most recent predecessor, LCM2015.

In this text we have introduced some UKCEH terminology and have adopted the convention of capitalising their first letters. This is so you know we are referring to something defined and specific rather than a more general use of language. We do not guarantee that these will have been explained before they are first encountered but in most cases this should not cause confusion. However, if at any point you come across capitalised terminology and you are not sure exactly what we mean please be reassured and read on; it should become clear. We have also provided a glossary for a quick refresh.

UK CEH Land Cover Classes and BAP Broad Habitat Classes

LCM2017, LCM2018 and LCM2019 map 21 UKCEH Land Cover Classes. These are the same classes used in LCM2015 and a near exact match to LCM2007 (Morton *et. al.*, 2011) and LCM2000 classes. They are based on UK Biodiversity Action Plan (BAP) Broad Habitats (Jackson et. al., 2000). But they are not UK BAP Broad Habitats. They are similar. BAP Broad Habitats were designed for field-based detection by botanists, not remote sensors orbiting the Earth at an altitude of circa 800km. We cannot therefore detect them exactly, but we do our best.

In the early 2000s conservation and regulatory agencies had reporting obligations under the UK BAP and were co-funding stakeholders of LCM2000 and LCM2007. For this reason LCM2000 and LCM2007 were based on BAP. New land cover and habitat descriptions have become available in recent years and some of these are certainly better suited for detection by remote sensors. Moreover, in 2012 the UK government published a new biodiversity strategy and the BAP became obsolete. Despite this, we have stayed close to the BAP. We are not resistant to change. But when it comes to changing the way we describe land cover we need to be careful not to devalue the sequence for land cover change detection. It is easiest to compare like with like.

[For readers interested in land cover change: UKCEH have recently created a revised UK-wide 1990 Land Cover Map to support comparisons with LCM2015 and newer products. With the revised LCM1990 we also provide quarter-Century, UK-wide change datasets. These and supporting documents became available simultaneously with the products described here. Please visit https://catalogue.ceh.ac.uk/eidc/c0078881-7d5a-4641-91e2-c271426bc8a1 for more information]

Describing complex land cover and habitat types that share similar nomenclatures and similar (but not equal) meanings precisely with words can be confusing for the writer, so definitely the readers. In most cases we could refer to UKCEH Land Cover Classes and BAP Broad Habitats as if they are one without causing confusion but sometimes we need to contrast them and it can get messy. To reduce ambiguity regarding land cover and habitat descriptions we therefore italicise UK BAP Broad Habitats when explicitly referring to these and at all times when referring to a defined class (UKCEH or UK BAP) we will begin each element with a capital letter. For example: Improved Grassland refers to the UKCEH Land Cover Class, *Improved Grassland* the BAP Broad Habitat.

Some users do not require the thematic detail of UKCEH Land Cover Classes so we also provide generalised land cover, UK CEH Aggregate Classes. Table 1 shows the relationship between UKCEH Land Cover Classes and Aggregate Classes with the UK BAP Broad Habitats. Usually there is a one-to-one relationship between UK BAP Broad Habitats and derived UKCEH Land Cover Classes. However, the UK BAP *Standing water and Canals* and the UK BAP *Rivers and Streams* are represented by a single UKCEH Land Cover Class, Freshwater. There is not a UK BAP for saltwater, but we separate saltwater and freshwater when possible, so include a UKCEH LCM Saltwater class. In three cases, we have found it appropriate to refine UK BAP Broad Habitats. The UK BAP *Dwarf Shrub and Heath* class is split into UKCEH Land Cover Classes Heather, and Heather Grassland. The UK BAP *Built up Areas and Gardens* class is split into UKCEH Land Cover Classes Urban and Suburban. For UK BAP *Littoral Sediment* we have been able to detect Saltmarsh, which matches a UK BAP Priority Habitat subclass of *Littoral Sediment.* Fuller details of class derivations, motivations for these and issues associated with their detection are given by Appendix 1 and 2.

Table 1 Relationship between UK CEH Aggregate classes, UK BAP Broad Habitats (Jackson 2000) and UKCEH Land Cover class and associated integer identifiers. Additional notes on the on the relationships between UKCEH Land Cover Classes and UK BAP Broad Habitats are given in Appendices 1 and 2 together with satellite-based mapping considerations.

UK CEH Aggregate Class (AC)	AC Identifier	UK BAP Broad Habitat	UKCEH Land Cover Class	LC Identifier
Broadleaf woodland	1	***Broadleaved mixed and yew woodland***	Deciduous woodland	1
Coniferous woodland	2	***Coniferous woodland***	Coniferous woodland	2
Arable	3	***Arable and horticulture***	Arable	3
Improved grassland	4	***Improved grassland***	Improve grassland	4
Semi-natural grassland	5	***Neutral grassland***	Neutral grassland	5
		Calcareous grassland	Calcareous grassland	6
		Acid grassland	Acid grassland	7
		Fen marsh and swamp	Fen	8
Mountain, heath and bog	6	***Dwarf shrub and heath***	Heather	9
			Heather grassland	10
		Bog	Bog	11
		Inland rock	Inland rock	12
Saltwater	7	-	Saltwater	13
Freshwater	8	***Standing open water and canals***	Freshwater	14
		Rivers and streams		
Coastal	9	***Supralittoral rock***	Supralittoral rock	15
		Supralittoral sediment	Supralittoral sediment	16
		Littoral rock	Littoral rock	17
		Littoral sediment	Littoral sediment	18
			Saltmarsh	19
Built-up areas and gardens	10	***Built-up areas and gardens***	Urban	20
			Suburban	21

Table 2. UKCEH LCM Land Parcel Products – additional metadata

	Great Britain, GB	Northern Ireland, NI
Coordinate System	British National Grid. EPSG: 27700	TM 75 Irish Grid. EPSG: 29903
Number of parcels	6736652	902248
Number of classes	21	21

Table 3. UKCEH LCM 8-bit Raster Products – geographic extents

	Great Britain, GB	Northern Ireland, NI
Coordinate System	British National Grid. EPSG: 27700	TM 75 Irish Grid. EPSG: 29903
Extent, metres	Min East 0 Max East 700000 Min North 0 Max North 1300000	Min East 180000 Max East 400000 Min North 300000 Max North 500000 (Truncated Irish Grid)

Table 4 UKCEH LCM 8-bit Raster Products – additional metadata

	Classified Pixels	Rasterised Land Parcels	Percent Cover	Percent Aggregate Cover	Dominant Cover	Dominant Aggregate cover
Pixel size	20m	25m	1km	1km	1km	1km
Number of classes	21	21	21	10	21	10
Number of bands	2	3	21	10	1	1

Dataset descriptions

The package of datasets for LCM2017, 2018 and 2019 follows that of UKCEH LCM2015, but with an additional dataset. This time we include classified pixels (at 20m resolution) from which all the remaining products are derived. The ancestry of datasets is shown by Figure 1, with additional detail and metadata provided by tables 2, 3 and 4. For each dataset there is a Great Britain version in the British National Grid (EPSG:27700) and a Northern Ireland Version in the Irish National Grid (EPSG:29903). This product suite is duplicated for each year giving a total 42 datasets (see Appendix 5 for the complete list). Text subsections below give more detailed descriptions for each dataset and highlight minor differences from LCM2015.

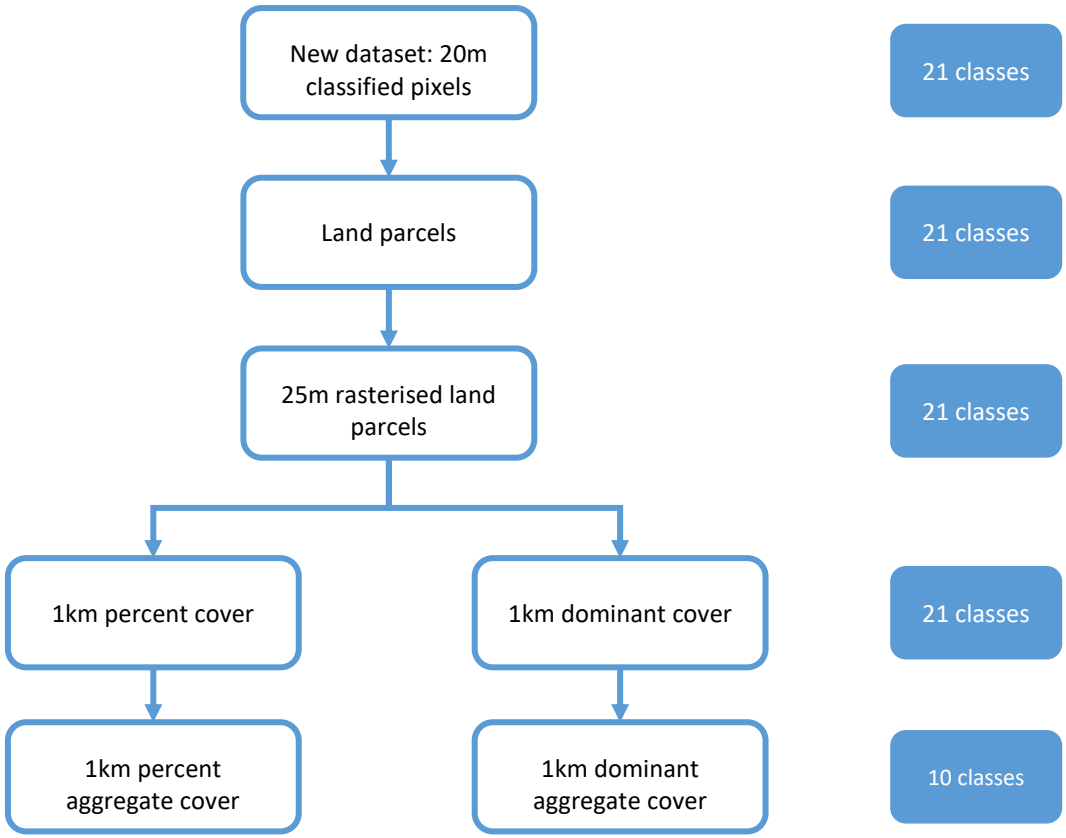


Figure 1. UKCEH LCM dataset structure within this product release and at each level the number of classes (21 or UKCEH Land Cover Classes, 10 for UKCEH Aggregate Classes) represented by the dataset.

20m Classified Pixels

These were not present in LCM2015 or earlier products. They are the new addition to the UKCEH LCM product suite. They are Random Forrest (RF) classification results from the 20m pixel Sentinel-2 Classification Scenes. All remaining products were derived from the 20m Classified Pixels datasets by summarising pixel information using the UKCEH Land Parcel Spatial Framework.

The 20m Classified Pixels datasets are provided as 2-band, 8-bit integer rasters. The RF classifier assigns each pixel a probability of membership for each of the 21 UKCEH Land Cover Classes. The nominate land cover, band 1, is the class with the highest membership probability. Band 2 is this probability, but rescaled and rounded giving an integer value over the range of 0 to 100 to allow 8-bit integer representation. This gives an indication of per-pixel classification confidence (uncertainty); high values equate to high confidence low uncertainty.

Unlike pixels of the 25m Rasterised Land Parcels datasets the 20m Classified Pixels have not been generalised by combination with the UKCEH Land Parcel Spatial Framework. This preserves intricate features of the landscape such as narrow linear features and small patches of habitat that fall below the 0.5 hectare minimum mappable unit (MMU). We anticipate that this extra detail will be of interest and useful.

Land Parcels

The Land Parcels datasets are the result of intersecting the 20m Classified Pixels datasets with the UKCEH Land Parcel Spatial Framework to generate land parcel attributes, see Table 5. There have been some minor attribute changes since LCM2015, these are also described. NOTE: Depending on how you receive the Land Parcels dataset, there may be an additional numeric identifier field. This will have been imposed by the specification of the file format in which data were supplied to you, and has no relevance. We only describe the attributes that we have derived.

Table 5. UKCEH LCM Land Parcel dataset attributes for LCMs 2017-2019 and the corresponding UKCEH LCM2015 attributes in brackets for comparison.

UKCEH LCM attribute name	**Description**	**LCM2015 attribute name***
gid	Unique identifier for the land parcel geometry maintained in UK CEH Land Parcel Spatial Framework	gid
_hist	The frequency of 20m pixels per UK CEH land cover class encountered in a land parcel given as a list of tuples <left>:<right> separated by commas. The figure to the left of the colon per tuple represents the UKCEH Land Cover Class identifier, to the right the frequency of occurrence. For example, **`2:3, 11:361`** means we predicted 3 pixels of Coniferous Woodland (2) and 361 of Bog (11).	Pix_dist
_mode	The most frequently occurring (modal) land cover type.	Mode
_purity	The percentage of the modal land cover class over the total number of pixels (_n) encountered.	Modal_prop
_conf	This, when combined with *_purity*, helps to understand classification confidence. It is the mean value of the class membership probability for each pixel, rescaled to the range of 0 to 100.	Unc
_stdev	The standard deviation of *_conf*	Unc_stdev
_n	The total number 20m pixels intersecting the land parcel	npix

* *The attribute names of the new products do not match those of LCM2015. The attribute names were changed to align with naming conventions used in our production software and internal documents. This was done to simplify production. We hope it does not cause user inconvenience.*

gid The *gid* is a unique identifier for each land parcel and its value is unchanged between products. However, *gid* values of the new products will not match those of the LCM2015 land parcels. The reasons for this is given in the section describing the UKCEH Land Parcel Spatial Framework.

_hist The presentation of the histogram data, *_hist*, has been improved. Previously it was given as list of 21 counts, one for each land cover class, inclusive of zeros. This was untidy and very difficult to read compared to the new tuple structure.

_purity The *_purity* attribute was previously given as a proportion, instead of a percentage. *_purity* is carried into the 25m Rasterised Land Parcel dataset, which is an 8-bit integer raster. A floating value over the range of 0-1 would degrade to binary information when converted into an integer, so we now represent this as a percentage.

_conf In LCM2015 *_conf* was scaled over 0 to 255, it is now scaled 0 to 100 as this is more intuitive to most users.

The LCM2015 Land Parcel dataset included an extra *BHAB* attribute, which is a textual description of the habitat. This was an unnecessary duplication of information. The dominant per-parcel UKCEH Land Cover Class is provided by the *_mode*, so a straightforward lookup can be used for corresponding text descriptions.

The LCM2015 Land Parcels dataset included an attribute called *Composite*, which identified the main twin-date composite image from which the land parcel's class was derived. Since the new products are derived from seamless composite images produced by the Google Earth Engine this attribute no longer has relevance and has been removed.

25m Rasterised Land Parcels

These are the result of rasterising the Land Parcels datasets into 25m pixels, with the pixel origin matching the origin of the British National Grid or the Irish National Grid. Three attributes are carried from the Land Parcels datasets, giving a 3-band raster. Band 1 is the dominant land cover *_mode*; band 2 is *_conf*; and band 3 _purity.

Differences since UKCEH LCM2015. In the LCM2015 25m Rasterised Land Parcel dataset only the *_mode* and *_conf* were rasterised, to give a 2-band raster. We consider that the *_purity* band provides useful information relating to the classification confidence so we include this as a third band in the new product. We expect that users will find this useful.

1km Rasters

A 1km grid, with an origin matching the origin of British National Grid or the Irish National Grid was used to intersect the 25m Rasterised Land Parcels. All 25m pixels within each 1km square were accumulated to derive the following 1km Raster datasets:

1km Percent Cover.	This is a 21-band, 8-bit integer raster giving the percentage cover of each UKCEH Land Cover class per 1km.
1km Percent Aggregate Cover	This is a 10-band, 8-bit integer raster product giving the percentage cover of each UKCEH Aggregate Land Cover Class per 1km.
1km Dominant Cover.	This is a single band, 8-bit integer raster giving the dominant UKCEH Land Cover Class per 1km grid square.
1km Dominant Aggregate Cover	This is a 1-band, 8-bit integer raster giving the dominant UKCEH Aggregate Land Cover Class per 1km grid square.

Dataset examples

Here we provide figures to clarify dataset explanations. UKCEH Land Cover Classes are coloured according to our recommended scheme given in Appendix 3.

Figure 2 gives example datasets of the 20m Classified Pixels, Land Parcels and 25m Rasterised Land Parcels. The 20m Classified Pixel dataset preserves detail lost in the Land Parcels and 25m Rasterised Land Parcels. Land Parcel boundaries are dissolved in 25m dataset. The 20m Classified Pixel dataset has two bands: band 1 gives the most likely UKCEH Land Cover Class; band 2 the membership probability of this class over a grey-scale: 0 black, 100 white. The Land Parcel product has a range of attributes (see Table 5) within a single example given. The 25m Pixel Rasterised Land Parcels have 3 bands: band 1 the modal UKCEH Land Cover Classes from the Land Parcels dataset (*_mode*); band 2 the parcel-averaged pixel membership probability (the *_conf* attribute from the Land Parcels); and band 3 the parcel purity (the *_purity* attribute from the Land Parcels dataset).

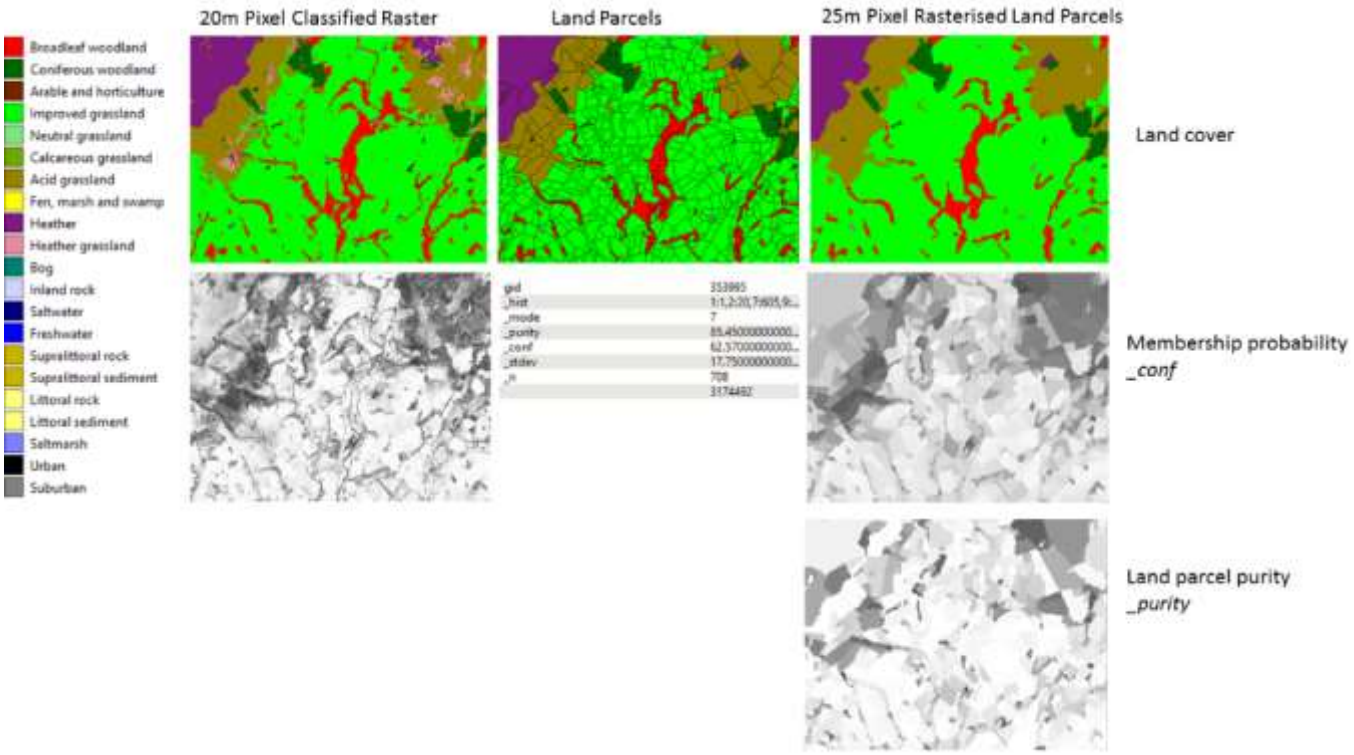


Figure 2. Examples of 20m Classified Pixels, Land Parcels and 25m Rasterised Land Parcels. Note the extra granularity and detail of the 20m Classified Pixels over the 25m Rasterised Land Parcels. The Land Parcel dataset highlights the parcel boundaries, which are dissolved in the 25m Rasterised Land Parcel dataset.

It is not informative to display the full list of raster bands for 1km datasets, nor all datasets. The text descriptions combined with the Figure 3 should be sufficient. Figure 3 juxtaposes the 1km Dominant Cover with the 25m Rasterised Land Parcels showing UKCEH Land Cover Classes to provide an example of the generalisation effect of the 1km dataset. Also provided is and a single band representing Improved Grassland of the 21-band 1km Percent Cover raster. Please refer to the colour key in Figure 2 the UKCEH Land Cover Classes. The 1km Percent Cover is on a grey-scale: 0 black, 1 white.

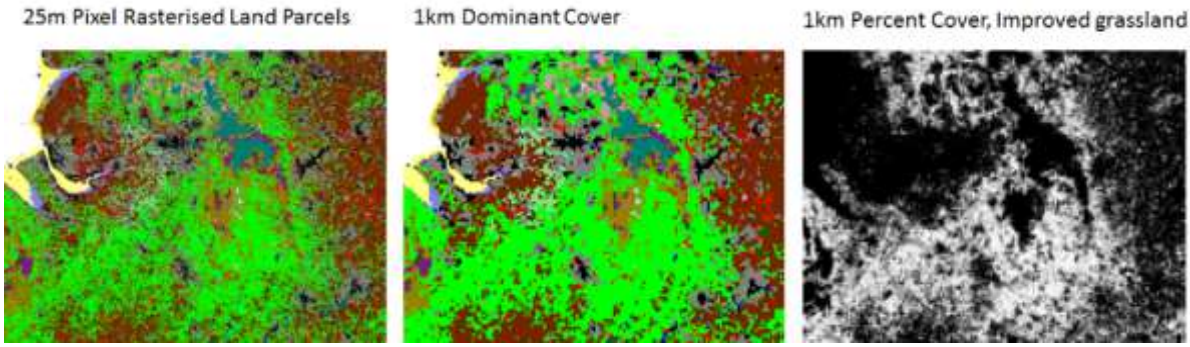


Figure 3. Juxtaposed 25m a 1km products to show the generalisation effect of the 1km Dominant Cover product. Also a 1-band example of the 21-band Percent Cover product (0 black, 1 white).

Material and Methods

Seasonal Composite Images

Seasonal Composite Images for classification were derived from Google Earth Engine. Median top of atmosphere reflectance (TOA) values resampled to 20m pixel resolution were computed for each season: winter (January-March), spring (April-June), summer (July-September), and autumn (October-December) using nine Sentinel-2 bands 2, 3, 4, 5, 6, 7, 8, 11, and 12 (after Carrasco *et. al.*, 2019) identified by Table 6. There were some seasonal gaps in the Seasonal Composite Images, because of excessive cloud, and these were represented by null data. Fortunately, this never occurred across all seasons at a single location, so at least some Sentinel-2 satellite data were available for all locations. Our classification algorithm will tolerate partially incomplete spectral information, so we were able to produce land cover for the whole of the UK without the need to manually fill gaps (manual gap filling was necessary for all previous UKCEH land cover maps). It is possible that full 4-season gaps at a single location will occur in the future (but rarely based on current experience) and we are exploring the opportunities for filling these gaps with classifications of alternative optical sources and multi-temporal Sentinel-1 Synthetic Aperture Radar, which is relatively insensitive to clouds and weather.

Table 6 Sentinel-2 spectral bands and spatial resolutions

Sentinel-2 Bands	Central Wavelength (µm)	Resolution (m)
Band 1 - Coastal aerosol	0.443	60
Band 2 - Blue	0.490	10
Band 3 - Green	0.560	10
Band 4 - Red	0.665	10
Band 5 - Vegetation Red Edge	0.705	20
Band 6 - Vegetation Red Edge	0.740	20
Band 7 - Vegetation Red Edge	0.783	20
Band 8 - NIR	0.842	10
Band 8A - Vegetation Red Edge	0.865	20
Band 9 - Water vapour	0.945	60
Band 10 - SWIR - Cirrus	1.375	60
Band 11 - SWIR	1.610	20
Band 12 - SWIR	2.190	20

The Google Earth Engine now provides Sentinel-2 data as land surface reflectance (SR) and coverage of this collection is expanding. However, during production these data were not fully available across all years for the whole UK, so after assessing the potential benefits of SR over TOA we chose to use TOA, since SR brought no improvements when classifying to UKCEH Land Cover Classes. Whether this relationship will hold when seeking higher thematic detail, we do not know. In the future we are likely to increase the range of land surface information products so we will continuously review satellite input choices.

Context Rasters

Spectral confusion can occur between different land cover types that have similar spectral properties. For example, bare rocks in the littoral coastal zone lack significant vegetation, so too do exposed mountain rocks and sealed urban surfaces. Spectrally these surfaces can appear very similar when viewed from space. Contextual information provides additional clues for the RF's learning algorithm to separate spectrally similar land cover types and for the above example coastal proximity, urban proximity and terrain context would clearly help. We used 20m Context Rasters to resolve a range of confusion types. For Great Britain the 20m Context Rasters were:

1. Height, derived from the NEXTMap® terrain product from Intermap® Solutions
2. Aspect, derived from the NEXTMap® terrain product from Intermap® Solutions
3. Slope, derived from the NEXTMap® terrain product from Intermap® Solutions
4. Distance from the nearest building, derived from Ordnance Survey open data
5. Distance from road, derived from Ordnance Survey open data
6. Distance from sea, derived from Ordnance Survey open data
7. Distance from freshwater, derived from Ordnance Survey open data
8. A foreshore binary mask, derived from Ordnance Survey open data
9. A tidal water binary mask, derived from Ordnance Survey open data
10. A woodland binary mask, derived from Ordnance Survey open data.

For Northern Ireland 20m Context Rasters were:

1. Height, derived from the NEXTMap® terrain product from Intermap® Solutions
2. Aspect, derived from the NEXTMap® terrain product from Intermap® Solutions
3. Slope, derived from the NEXTMap® terrain product from Intermap® Solutions
4. An urban binary mask derived from open data of the Northern Ireland Statistics and Research Agency
5. A distance to coast layer derived from the Ordnance Survey or Northern Ireland open data
6. A distance to freshwater layer derived from the Ordnance Survey or Northern Ireland open data
7. A distance to road layer derived from the Ordnance Survey or Northern Ireland open data.

Classification Scenes

For Great Britain a patchwork of overlapping 100x100km tiles was created (Figure 4). These were used to select and extract 36-band Seasonal Composite Images from the Google Earth Engine. These were then combined with the Context Rasters to give a set of 100x100km 46-band GB Classification Scenes. In total we classified 74 overlapping Classification Scenes whose combined area significantly exceeds the total area of the GB land surface. Each Classification Scene was trained and classified independently. Classification Scenes were overlapped to ensure that each had a good range and balance of training observations to maximise classification accuracy. In regions where overlaps occurred the same region will have been classified multiple times. Where this occurred there will have been slight variations in results because of different training. Visual inspection was used to determine precedence so that the best results made it into the final cut; this is the only manual judgement in the GB UK CEH LCM production process. The 100x100km tile size was chosen because it provides a manageable size for processing. Moreover, if regions were much larger phenological variation due to climatic difference across a Classification Scene could begin to degrade results.

For Northern Ireland, because of its much smaller area, it was possible (and much easier) to use a single 43-band (36 spectral and 7 context layers) Classification Scene for each year determined by the minimum bounding rectangle of the Northern Ireland land mass (Figure 4).

The UKCEH Land Parcel Spatial Framework

The UK Land Parcel Spatial Framework used for LCM2017, LCM2018 and LCM2019 was originally developed for LCM2007 (Morton *et. al.,* 2011) and a modified version was used for LCM2015. For the new products we had to make further minor changes. The total number of land parcels and their geometries are unchanged, they are an exact match to those used in LCM2015, but we have re-ordered database storage and defined new indices to enable faster processing. The consequence of this is that the unique identifiers (the *gid* attribute) for each land parcel do not match those provided in the LCM2015 Land Parcels dataset. For most users this will have no significance, but for users who wish to compare the new land cover with LCM2015 using parcel identifiers, instead of spatial overlap, this will not be possible. We apologise if this causes inconvenience. LCM2017 to LCM2019 land parcel comparisons can be made using the *gid.*

Land parcels in the UKCEH Land Parcel Spatial Framework have a minimum area of approximately 0.5 hectares, the minimum mappable unit, MMU. The land parcels were derived by generalising national cartography (Morton et. al., 2011; Smith et. al., 2007) and are designed to represent discrete real-world units of land such as fields, parks, urban areas, woodlands, lakes and so forth. It is usual, but not universal, that the land parcels are dominated by a single land cover type. Therefore organising the 20m Classified Pixels into land parcels helps to reduce classification noise to provide a clean, easier to use product. It also provides a convenient fixed structure for comparing land cover over time for change detection.

The UKCEH Land Parcel Spatial Framework is not the only structure by which the 20m Classified Pixels dataset could be summarised. Through providing the 20m Classified Pixels in this release we give users the option to employ their own parcel structure to summarise land cover. For example, these might be catchment boundaries, administrative boundaries, nature reserves and so forth. In the future we may refine the UK CEH Land Parcel Spatial Framework as it was originally optimised for ~30m pixel resolution satellite data used for UKCEH LCM2007. Current trends are towards higher spatial resolution, so a finer framework with a reduced MMU will soon become more appropriate.

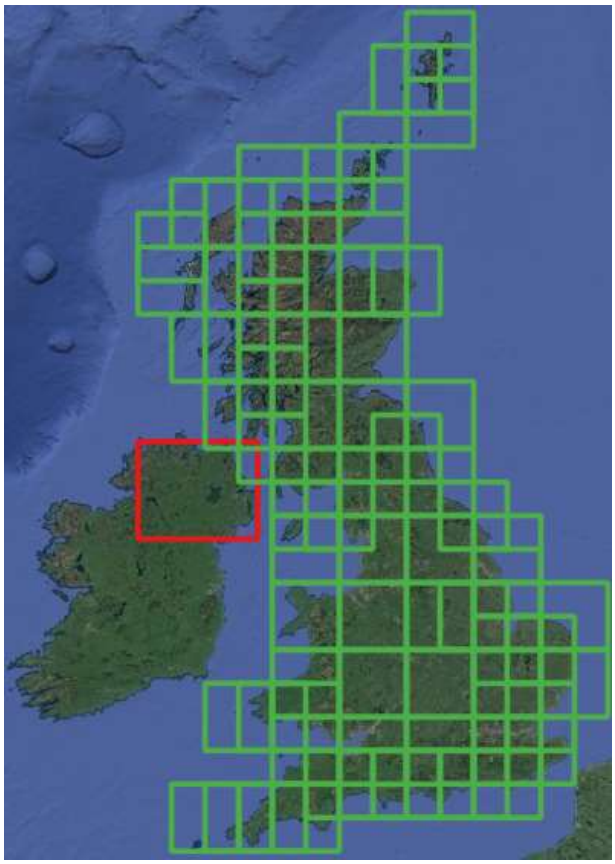


Figure 4. 100x100km tiles used for selecting Sentinel-2 Seasonal Composite Images for Great Britain (green) and the single tile for Northern Ireland (red).

Bootstrap Training

Bootstrapping is used to refer to a self-starting process that proceeds without external input. UKCEH have developed a fully automatic training process for land cover/habitat classification that does not require a fresh collection of (expensive) field-gathered data, so we have named the process Bootstrap Training. Land cover and habitat change is usually gradual. Transitions from one land cover or habitat to another typically occur over a number of years. Therefore recent habitat/land cover maps can be a valuable source of training data for a new map if the original maps are accurate and the update interval of the new map is short relative to target (land cover class) dynamics. Land cover observations from the historic maps can be used to sample the current satellite image to produce spectral training observations. These are then used by a RF classifier to yield a classification image, which contributes to the bootstrap for the next map and so forth. Because the historic UKCEH LCMs give coast-to-coast coverage they provide can provide a very large number of training observations, which is the key to learning success. Machine-learning algorithms, such as RF, rely on the majority signal to assign class membership, so when the Bootstrap Training set is very large it doesn't matter if a minor proportion have changed class over the refresh interval since these will have little influence on the dominant signal.

The Bootstrap Training for UKCEH LCMs 2017, 2018 and 2019 all came from UKCEH LCM2015, and LCM2015 was derived from a bootstrap sample from stable land cover observations from 1990 to 2007 combined with additional manual observations. To generate the Bootstrap Training set for LCM2017-2019 we selected all land parcels from UKCEH LCM2015 with greater than or equal to 99% purity (the *_purity* attribute). As stated earlier, our methods are evolving and Bootstrap Training strategies are likely change. Our current intention is that the LCM2020, planned for release in 2021, will use a bootstrap from the majority signal across 2017 to 2019; similarly the bootstrap for 2021 will use the 3-year majority signal over 2018 to 2020. Figure 5 shows an example of a Bootstrap Training dataset and resultant RF classification result.

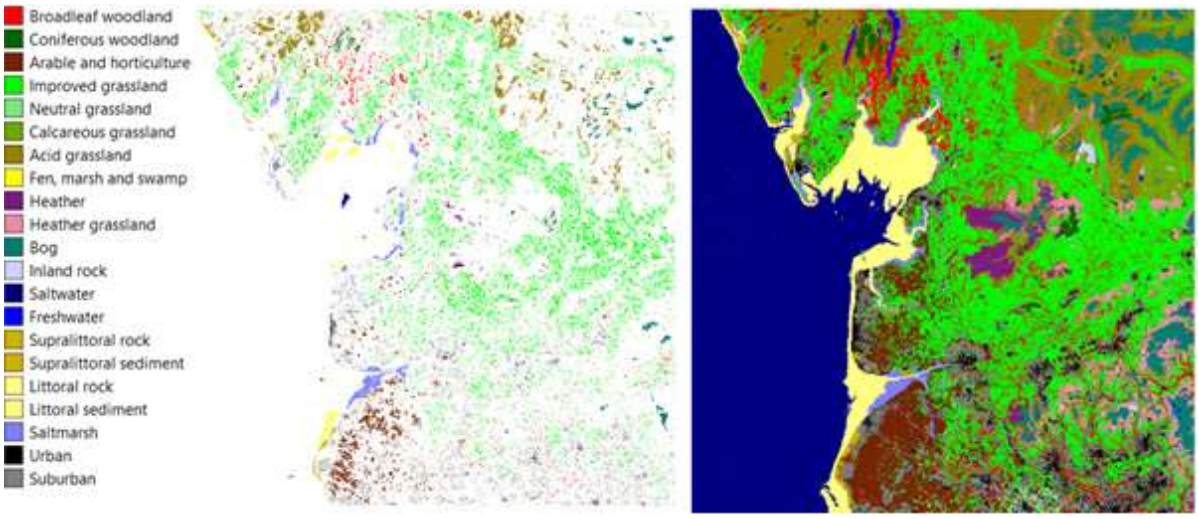


Figure 5. UKCEH Land Cover Classes (left), a Bootstrap Training set derived from UKCEH LCM2015 (centre) and the resultant RF land cover classification (right).

Random Forest classification

Random Forest (RF) classification (Breiman 2001) is a supervised learning technique that uses a training set of known observation to derive an empirical relationship that is used to predict the membership of unknown observations. Bootstrap Training data derived from UKCEH LCM2015 were used to sample all underlying pixels from the Classification Scenes for 2017, 2018 and 2019. Pixels were placed into labelled bags and from each bag 10,000 samples, with replacement were drawn to train the RF classifier. The RF classifier subsequently yields the 20m Classified Pixel product (for example, Figure 3, right). Sampling with replacement ensures that all land cover classes have an equal number of pixel observations for training, which balances learning. Without balanced learning the signal of rarer classes will be weak and susceptible to domination from commoner classes, causing misclassification.

The classification software used for UKCEH LCMs 2017 to 2019 is bespoke and was developed by UKCEH scientific staff. It integrates the Weka (Frank *et. al.* 2016) machine learning suite with a PostGIS geospatial database and gdal tools (https://gdal.org/). These are all open source technologies.

Product validation

LCM2017, 2018 and 2019 at UK scale were validated by comparison with observations derived from the GB countryside survey 2019 data, open source National Forest Inventory data, IACS data a set of bespoke LCM validation points generated from manual image interpretation. In total, this generated 22,325 point locations. These were intersected with UKCEH LCM Land Parcel datasets to determine correspondence. The overall accuracy figures from this validation exercise are:

- LCM2017: 78.6%;
- LCM2018: 79.6%;
- LCM2019: 79.4%.

Please note that the same validation data were used for all three products and the currency of validation points, relative to each product is therefore variable. Validation observations should not be considered as absolute truth, but the best available indicators of reality. A proportion of validation observations will be incorrect. Moreover, the validation sample was derived from data sources with a different purpose and class descriptions that did not match exactly to UKCEH Land Cover Classes. They needed converting to UKCEH Land Cover Classes and schematic conversions can be subjective. Given this information ~80% correspondence for UK-wide 21-class automatically produced land cover maps is very impressive. We anticipate higher levels of accuracy as methods mature. We shall also continue to seek improved validation resources. Full correspondence tables from the validation are given in Appendix 4.

References

Breiman, L., 2001. Random forests. *Machine learning*, *45*(1), pp.5-32.

Carrasco, L., O'Neil, A.W., Morton, R.D. and 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), p.288.

Frank, E., Hall, M. A., and Witten I. H. (2016). The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.

Jackson D.L., (2000), Guidance on the interpretation of the Biodiversity Broad Habitat Classification (terrestrial and freshwater types): Definitions and the relationship with other classifications, JNCC Report 307, 73 pages, ISSN 0963 8091(http://www.jncc.gov.uk/page-2433 accessed 2020-06-01)

Morton, D., Rowland, C., Wood, C., Meek, L., Marston, C., Smith, G., Wadsworth, R. and Simpson, I., 2011. Final Report for LCM2007-the new UK land cover map. Countryside Survey Technical Report No 11/07.

Smith, G., Beare, M., Boyd, M., Downs, T., Gregory, M., Morton, D., Brown, N. and Thomson, A., 2007. UK land cover map production through the generalisation of OS MasterMap®. *The Cartographic Journal*, *44*(3), pp.276-283.

Appendix 1. Notes on UKCEH Land Cover Classes

(see also Appendix 2 for a summary of UK BAP Broad Habitat definitions)

UKCEH Land Cover Class	**Notes**
Broadleaved woodland	In the UK BAP *Broadleaved, mixed and yew woodland* the broad leaved woodlands are characterised by stands >5 m high with tree cover >20%. Scrub (<5 m) requires a cover >30% for inclusion. Such fine distinctions cannot be made through remote sensing. Open-canopy woodland (stands with trees <50%) is a particular problem, albeit occurring relatively rarely. These are likely to be confused with other classes due to the dominance of the non-woodland understory. Broadleaved evergreen trees rarely occur in stands >0.5 hectares; an area large enough to create suitable training areas appropriate for classification. As a consequence the classifier would struggle with this land cover. It is likely they will be classified as Conifer because of the full-year chlorophyll signal. Mixed woodland stands of broad-leaved or evergreen trees exceeded the minimum mappable unit, they were treated as separate blocks within the woodland; in many parts of the UK, truly 'mixed woodlands' as opposed to those with mosaic-blocks of broadleaved and coniferous trees, are unusual. Stands with near-closed canopies can be interpreted easily in the field and pure examples can normally be found for training the classifier.
Coniferous woodland	The UK BAP *Coniferous Woodland* class includes semi-natural stands and plantations, with cover >20%. Classification of coniferous woodland is generally straightforward, but rare examples of open canopy semi-natural pinewoods are likely to be classified according to the dominant understorey class. The UK BAP includes new plantation and recently felled areas. These are land use, not land cover. Newly felled areas are often dominated by grass, heather and encroaching vegetation and more likely to be classified as these, instead of coniferous woodland. Deciduous larch has potential for confusion with broadleaved deciduous woodland but is generally correctly identified.

UKCEH Land Cover Class	**Notes**
Arable and Horticulture	The BAP Broad Habitat *Arable and Horticulture* includes annual crops, perennial crops such as berries and orchards and freshly ploughed land. This is a very broad class and as a consequence has large potential for spectral confusion with non-arable surfaces. The main confusion between arable and other classes occurs between arable land and improved grassland. This is especially likely when grassland is managed by cutting, followed by periods of low growth and reflectance from chlorophyll. When this happens the observed seasonal reflectance pattern can be similar to graminid crops, such as wheat and barley. Indeed grass managed in this way is technically a crop, so an arable classification isn't necessarily wrong.
Improved Grassland	Improved grassland is distinguished from semi-natural grasslands based on its higher productivity, lack of winter senescence, location and/or context. Grasslands lie on a continuum, so some confusion with other grassland types is inevitable. Confusion with grass-like crops will also occur.
Neutral Grassland	The UK BAP Broad Habitat *Neutral Grassland* we expected to be troublesome for satellite-based classification, and we had considered removing it from our scheme. BAP *Neutral Grassland* is defined by botanical composition and includes semi-improved grasslands managed for silage, hay or pasture (Jackson, 2000). There isn't generally an obvious spectral difference between these and other productive grass types. However, the inclusion of Context Rasters for slope and distance to rivers appear to have helped greatly with Neutral Grassland detection. We were surprised: validation results are better than expected for this class.
Calcareous Grassland	Calcareous Grassland class is mapped spectrally. However, the inclusion of ancillary layers for slope is expected to improve results. We do not have free access to a highly resolved soil PH/soil type layer, which we would expect to help further. For regions know to contain substantial coverage of Calcareous Grassland, for example Limestone Dales of Derbyshire and North Yorkshire, the South Downs and Salisbury Plain our results match expectations.

UKCEH Land Cover Class	Notes
Acid Grassland	The UK BAP *Acid Grassland* can be spectrally variable, depending on dominant species composition. Deciduous Acid grassland, dominated by *Molinea caerula* has a distinct signal from more species diverse acid grasslands with mixtures of grasses, rushes, mosses, herbs and sedges. In other work we have been able to refine this class successfully. However, we did not make this separation in historical maps, so we are not able to retrieve suitable observations from Bootstrap Training. Bracken has a very distinctive spectral signal, but only at certain times of the year when its foliage begins to dominate its grassland understory. Historically, with restricted availability of satellite images we could not reliably separate the UK BAP *Bracken* class from *Acid Grassland* so we combined these into a single UKCEH Land Cover Class. With the greater image frequency and therefore better access to seasonal signals it may now be possible to overcome this limitation, but to do this we will need novel training data as we will not be able to retrieve a signal from Bootstrap Training as we do not mapped Bracken in recent UKCEH LCMs.
Heather; and Heather grassland	For LCM2007 we refined the BAP *Dwarf Shrub and Heath* into two classes, depending on the density of heather, producing the UKCEH Heather and Heather Grassland classes (It is heather when there is greater than 25% Heather Cover). This was to retain some consistency between the LCM1990 and LCM2000 classes Open Shrub Heath and Dense Shrub Heath. In some parts of the UK, significant areas of low lying non-heather shrubs occur. For example, Gorse can form a dominant shrub layer. Here we would the UKCEH Heather class, although it is a misnomer and perhaps the original LCM1990 and LCM2000 names would be better. Note: the Land Cover Maps typically show confusion over Heather and Heather grassland (and Bog too). However, they are often difficult to separate in the field. It is not easy to accurately estimate coverage above and below the defining threshold.
Fen, Marsh and Swamp	The UK BAP *Fen, Marsh and Swamp* includes fen, fen meadows, rush pasture, swamp, flushes and springs. From a remote sensing perspective Fen, Marsh and Swamp is problematic as it is can be comprised of a wide range of vegetation types and many patches are below the MMU of the UKCEH Land Parcel Spatial Framework. The small size of many Fen, Marsh and Swamp patches, plus their typically mosaic nature make it difficult to find reliable training data. Consequently, Fen, Marsh and Swamp is likely to be underestimated in some regions. However, substantial areas of contiguous reed dominated Fenland appear to be well detected.

UKCEH Land Cover Class	Notes
Bog	The UK BAP *Bog* includes ericaceous, herbaceous and mossy swards in areas with a peat depth > 0.5 m. We cannot detect peat depth from satellites. Vegetation on deep peat soils represent a continuum involving acid grassland, dwarf shrub heath and some types of fen, marsh and swamp and the separation of continuously varying land cover into discrete types can be difficult, especially when they exist in a complex small patch mosaic and their definitions are vague. We retain the Bog class to maintain consistency with historical UKCEH LCM products and the RF classifier learns Bog presence based on training data automatically generated from these. The predicted distribution occurs in regions where it is expected, so is a good indicator of where Bog is likely to be occurring. However, Bog and the range of upland vegetation classes expected to occur on peaty soils (Acid grassland, Fen marsh and swamp, Heather, and Heather grassland) show a cluster of interclass confusion (Appendix 4). This is partly due fine-scale variation but largely an effect of ambiguous definitions. UK BAP Broad Habitats (on which UKCEH Land Cover Classes are based) were not defined with satellite remote sensing in mind. A refined set of upland vegetation types for suitable for satellite detection would be preferable. In other work UKCEH have found it possible to separate a revised set upland vegetation types with higher reliability, but the necessary training data is restrictive. Mapping upland peatland vegetation with higher accuracy is desirable, as peatland vegetation has a key role in the carbon-cycle, water storage and flood management. UKCEH remote sensing staff are engaged in upland habitat studies and intend to bring the benefits of this to future UKCEH LCMs if we can find adequate training samples.
Saltwater	Saltwater is rarely different spectrally from freshwater, and the Saltwater distribution predicted by the RF classifier is determined by coastal Context Rasters in Classification Scenes. There will be some confusion between Saltwater and Freshwater in tidal rivers, but not substantial. Occasionally, Saltwater is confused with non-vegetated surfaces close to the coast and this happens because the automatically generated Saltwater training classes coincide with the tide being out in the satellite view. The effect has so far been trivial but the result is that we predict Saltwater with slightly lower accuracy than Freshwater. Our main goal is to map land cover so coastal water and intertidal regions are not high priority

UKCEH Land Cover Class	**Notes**
Freshwater	The UK CEH Freshwater class comes from merging two BAP BHs (*Standing Open Water and Canals,* and *Rivers and Streams*) since they cannot be separated by spectra. In many cases small and/or narrow water bodies fall below the MMU of the UKCEH Land Parcel Spatial Framework so effectively disappear into the dominant surrounding vegetation. Where these features are appropriately aligned with the satellite swath and sufficiently wide relative to pixel resolution they may be detected and if so will be visible in the 20m Classified Pixels datasets. Water bodies > 0.5 ha and wider than 40m are mapped with very high accuracy. The exceptions are temporary water bodies and quarries. Water in some quarries is strongly affected by the minerals in the rock and can result in atypical colours and misclassification.
Inland Rock	The BAP Broad Habitat *Inland Rock* covers both natural and artificial exposed rock surfaces which are >0.25ha, such as inland cliffs, caves, screes and limestone pavements, as well as various forms of excavations and waste tips such as quarries and quarry waste. Opportunistic vegetation is common amongst rocky landscapes. We will classify Inland Rock if rock has the dominant signature.
Urban; and Suburban	Within the *Built-up Areas and Gardens* BAP Broad Habitat we can reliably separate two UKCEH Land Cover Classes: Urban and Suburban. Urban includes dense urban, such as town and city centres, where there is little, if any, vegetation. Urban also includes areas such as dock sides, car parks and industrial estates. It is sometimes confused with other non-vegetated surfaces; for example open cast quarries or more rarely coastal rocks or ploughed fields. Suburban includes suburban areas where the spectral signature is a mix of urban and vegetation signatures. Suburban and Urban lie on a continuum and confusion is expected.
Supralittoral Rock	Features that may be present in this coastal class include vertical rock, boulders, gullies, ledges and pools generally forming a narrow band when viewed from above. Only limited areas can be mapped using satellite remote sensing.
Supralittoral Sediment	This class includes sand-dunes, which are reliably detected. Areas of coastal sand may be confused between this class and the Littoral Sediment class. Supralittoral sediments can stabilise and from increasing volumes of vegetation. Heavily vegetated Supralittoral Sediment is likely to be assigned to a vegetation class.
Littoral Rock	This class occurs in the maritime zone on a rocky coastline. It is generally more extensive than Supralittoral Rock and thus more readily detected using satellite images.

UKCEH Land Cover Class	Notes
Littoral Sediment; and Saltmarsh	The BAP Broad Habitat *Littoral Sediment* has a subclass, the BAP Priority Habitat *Saltmarsh*. Saltmarsh is generally distinct from nearby vegetation and only occurs near the coast. As a consequence we can map this well with remote sensing. We therefore include a UKCEH Land Cover Class, Saltmarsh. The Saltmarsh class is occasionally subject to commission error, when we mistake other vegetation in the coastal zone (mainly Arable) as Saltmarsh. This is a result of the coastal context layers being a powerful effect on Saltmarsh detection, but they solve far more problems than they create. The Littoral Sediment is sometimes confused with the Supralittoral Sediment class.

Appendix 2: Biodiversity Action Plan (BAP) Broad Habitats

This appendix provides a brief summary of the JNCC definitions of the Broad Habitats and is based on Jackson (2000). There is some duplication of material Appendix 2 with Appendix 1. However in combination they provide users with key information to understand UKCEH Land Cover Classes and the Broad Habitat definitions they are based on. Note: class numbers here are JNCC's Broad Habitat class numbers, not UKCEH Land Cover Class numbers.

The text in this Appendix has been copied from Morton *et. al.*, 2011, with very minor updates to remove obsolete cross-references to historical projects and documents, plus some corrections. Some new comments are included regarding BAP Broad Habitat *Linear Features* and the new 20m Classified Pixel datasets.

1. Broadleaved, Mixed and Yew Woodland

This Broad Habitat is characterised by vegetation dominated by trees >5m high when mature, with tree cover >20%. Scrub (<5 m) requires cover >30% for inclusion in this Broad Habitat. It includes stands of both native and non-native broadleaved trees and yew. Woodlands dominated by coniferous species but with >20% cover by deciduous species are included in this category. Areas of fen woodland dominated by species such as willow (*Salix* spp.), alder (*Alnus glutinosa)* or birch (*Betula* spp.) are also included.

2. Coniferous Woodland

This Broad Habitat is characterised by vegetation dominated by trees >5m high when mature, which forms a canopy having a cover of >20%. *'Coniferous Woodland'* includes semi-natural stands and plantations and includes both native and non-native coniferous trees.

3. Boundaries and Linear Features

This Broad Habitat type covers a range of linearly arranged landscape features such as hedgerows, lines of trees, walls, stone and earth banks, grass strips and dry ditches. These features are not included as a UKCEH Land Cover Class as they are generally too narrow to be reliably captured from the satellite images. However, linear features are of general interest and an important component of the landscape. Some linear features will be visible in the 20m Classified Pixel dataset. In future UKCEH LCMS we are likely to resolve to 10m and more of these features will be visible.

However, regardless of whether or not we can see linear features, we cannot classify them using our current methods. Linear features are in fact just thin fragments of the various land cover types that we already map. Membership in this class is defined by shape, not spectra. We don't have tools for classifying shapes. New tools for narrow feature finding within the 20m Classified Pixels products could help with their detection. Line finding algorithms are well developed and commonly used in other areas of remote sensing, mainly for military applications.

4. Arable and Horticulture

This Broad Habitat includes annual crops, perennial crops, woody crops, intensively managed commercial orchards, commercial horticultural land (such as nurseries, commercial vegetable plots and commercial flower growing areas), freshly-ploughed land, annual leys, rotational set-aside and fallow.

5. Improved Grassland

'Improved Grassland' is characterised by vegetation dominated by a few fast-growing grasses such as *Lolium* spp*.,* and also white clover (*Trifolium repens),* on fertile, neutral soils. Improved Grasslands are typically either managed as pasture or mown regularly for silage production or in non-agricultural contexts for recreation and amenity purposes.

6. Neutral Grassland

This Broad Habitat type is characterised by vegetation dominated by grasses and herbs on a range of neutral soils usually with a pH of between 4.5 and 6.5. It includes enclosed dry hay meadows and pastures, together with a range of grasslands which are periodically inundated with water or permanently moist.

7. Calcareous Grassland

Calcareous Grassland is characterised by vegetation dominated by grasses and herbs on shallow, well-drained soils which are rich in bases (principally calcium carbonate) formed by the weathering of chalk and other types of limestone or base-rich rock. Soil pH tends to be high (>6) although it may be as low as 5.

8. Acid Grassland

Acid Grassland is characterised by vegetation dominated by grasses and herbs on a range of lime-deficient soils which have been derived from acidic bedrock or from superficial deposits such as sands and gravels. Such soils usually have a low base status, with a pH of <5.5.

9. Bracken

This Broad Habitat type covers areas dominated by a continuous canopy cover of bracken (*Pteridium aquilinum*) at the height of the growing season. It does not include areas with scattered patches of bracken or areas of bracken which are less than 0.25 ha. These are included in the Broad Habitat type with which they are associated.

10. Dwarf Shrub Heath

'Dwarf Shrub Heath' is characterised by vegetation that has >25% cover of plant species from the heath family (ericoids) or dwarf gorse *Ulex minor*. It generally occurs on well-drained, nutrient-poor, acid soils. This habitat type does not include dwarf shrub dominated vegetation in which species characteristic of peat-forming vegetation such as cotton-grass *Eriophorum* spp. and peat-building sphagna are abundant, or that occurs on deep peat (> 0.5 m) as these are included in the '*Bog*' Broad Habitat type.

11. Fen, Marsh and Swamp

This habitat includes fen, flushes, springs, fen meadows, rush pasture and swamp. Fens are peatlands which receive water and nutrients from groundwater and surface run-off, as well as from rainfall. Flushes are associated with lateral water movement, and springs with localised upwelling of water. Marsh is a general term usually used to imply waterlogged soil; it is used more specifically here to refer to fen meadows and rush-pasture communities on mineral soils and shallow peats. Swamps are characterised by tall emergent vegetation. Reedbeds (i.e. swamps dominated by stands of common reed *Phragmites australis*) are also included in this type. Apart from rush pasture, examples of this Broad Habitat are relatively rare.

12. Bog

This Broad Habitat type covers wetlands that support vegetation that is usually peat-forming and which receive mineral nutrients principally from precipitation rather than ground water. This is referred to as ombrotrophic (rain-fed) mire. The Bog Broad Habitat includes ericaceous, herbaceous and mossy swards in areas with a peat depth >0.5m.

13. Standing Open Water and Canals

This Broad Habitat type includes natural systems such as lakes, meres and pools, as well as man-made waters such as reservoirs, canals, ponds and gravel pits.

14. Rivers and Streams

The '*Rivers and Streams*' Broad Habitat type covers rivers and streams from bank top to bank top, or where there are no distinctive banks or banks are never overtopped, it includes the extent of the mean annual flood.

15. Montane

The *'Montane Habitats'* category includes a range of vegetation types that occur exclusively in the montane zone such as prostrate dwarf shrub heath, snow-bed communities, sedge and rush heaths, and moss heaths. The distinction between the sub-montane and montane zone is often blurred and the two usually merge through a band of transitional vegetation.

16. Inland Rock

This Broad Habitat type covers both natural and artificial exposed rock surfaces which are >0.25ha, such as inland cliffs, caves, screes and limestone pavements, as well as various forms of excavations and waste tips such as quarries and quarry waste.

17. Built-Up Areas and Gardens

This Broad Habitat type covers urban and rural settlements, farm buildings, caravan parks and other man-made built structures such as industrial estates, retail parks, waste and derelict ground, urban parkland and urban transport infrastructure. It also includes domestic gardens and allotments. This type does not include amenity grassland which should be included in the '*Improved Grassland*' category.

18. Supralittoral Rock

This habitat occurs above the high water mark, in areas influenced by wave splash and sea-spray. Features that may be present include vertical rock, boulders, gullies, ledges and pools, depending on the wave exposure of the site and its geology.

19. Supralittoral Sediment

This habitat occurs above the high water mark, but in areas influenced by wave splash and sea-spray. It includes shingle beaches, sand dunes and machair.

20. Littoral Rock

The geology and wave exposure of the shore influence the form of Littoral Rock habitats, which can be as varied as vertical rock, shore platforms, boulder shores, or rocky reefs surrounded by areas of sediment. In general, *Littoral Rock* tends to be colonised by algae in wave-sheltered conditions, and by limpets, barnacles and mussels as wave-exposure increases.

21. Littoral Sediment

Areas of *Littoral Sediment* are widespread around the UK forming features such as beaches, sand banks, and intertidal mudflats. A large proportion of this habitat occurs in estuaries and inlets where it can cover extensive areas. Saltmarsh is included within this Broad Habitat.

22. Inshore Sublittoral Sediment

All areas of sea and estuary class are assumed to be *Inshore Sublittoral Sediment*. It is defined as within six nautical miles of the shoreline by JNCC.

Appendix 3: Recommended RGB colour recipe for displaying UKCEH Land Cover Classes

UKCEH Land Cover class number	UKCEH Land Cover class		Red	Green	Blue
1	Broadleaved woodland		255	0	0
2	Coniferous Woodland		0	102	0
3	Arable and Horticulture		115	38	0
4	Improved Grassland		0	255	0
5	Neutral Grassland		127	229	127
6	Calcareous Grassland		112	168	0
7	Acid grassland		153	129	0
8	Fen, Marsh and Swamp		255	255	0
9	Heather		128	26	128
10	Heather grassland		230	140	166
11	Bog		0	128	115
12	Inland Rock		210	210	255
13	Saltwater		0	0	128
14	Freshwater		0	0	255
15	Supralittoral Rock		204	179	0
16	Supralittoral Sediment		204	179	0
17	Littoral Rock		255	255	128
18	Littoral sediment		255	255	128
19	Saltmarsh		128	128	255
20	Urban		0	0	0
21	Suburban		128	128	128

Appendix 4. Confusion matrices for UKCEH LCMs 2017, 2018 and 2019

Confusion matrix for LCM2017

	Broadleaved woodland	Coniferous woodland	Arable	Improved grassland	Neutral grassland	Calcareous grassland	Acid grassland	Fen, Marsh, Swamp	Heather	Heather grassland	Bog	Inland Rock	Saltwater	Freshwater	Supralittoral rock	Supralittoral sediment	Littoral rock	Littoral sediment	Saltmarsh	Urban	Suburban	Total	User's Accuracy
Broadleaved woodland	1032	203	15	77	2	1	31	9	22	14	9	5		5	2					2	21	1450	**71.2**
Coniferous woodland	53	660	4		2	2	4		16	17	1			2		1						762	**86.6**
Arable	10	3	1893	250	9	18	3	3	1			18		2		1		1	2	12	11	2237	**84.6**
Improved grassland	42	3	354	3561	188	65	243	12	18	72	48	3				21	2			9	39	4680	**76.1**
Neutral grassland	6		19	120	401		3	21	2	2				3							6	583	**68.8**
Calcareous grassland	11		4	7	3	855	4			3		10										897	**95.3**
Acid grassland	13	2	2	138	6	107	1340		138	238	97	15		7						1	3	2107	**63.6**
Fen, Marsh, Swamp	3		3	6				534			3			1								550	**97.1**
Heather	5	3			1		29	2	716	64	108	1			1							930	**77.0**
Heather grassland	23	4	3	5		2	126	2	137	358	179	6				1						846	**42.3**
Bog		5		1			42	6	34	65	910			6								1069	**85.1**
Inland Rock			1	1	1	6	7		7	3	4	144								3	1	178	**80.9**
Saltwater													83	2				11				96	**86.5**
Freshwater	3			1			4	6		1		2		508				1			1	527	**96.4**
Supralittoral rock				1			5		1		1			1	34	6	16	8				73	**46.6**
Supralittoral sediment			1	5	3		3				1				1	164		6			1	185	**88.6**
Littoral rock												1	1	1	21	12	76	38		2		152	**50.0**
Littoral sediment													11	4		14	10	183	2			224	**81.7**
Saltmarsh	3		3	19				37						4	1	11		7	164	1	2	252	**65.1**
Urban	2		4		1		1				2	9		2				1		1407	291	1720	**81.8**
Suburban	22		19	44	1				1	1		1				3				184	2531	2807	**90.2**
Total	1228	883	2325	4236	618	1056	1845	632	1093	838	1363	215	95	548	60	234	104	256	168	1621	2907	22325	
Producer's Accuracy	**84.0**	**74.7**	**81.4**	**84.1**	**64.9**	**81.0**	**72.6**	**84.5**	**65.5**	**42.7**	**66.8**	**67.0**	**87.4**	**92.7**	**56.7**	**70.1**	**73.1**	**71.5**	**97.6**	**86.8**	**87.1**	**Accuracy**	**78.6**

Confusion matrix for LCM2018

	Broadleaved woodland	Coniferous woodland	Arable	Improved grassland	Neutral grassland	Calcareous grassland	Acid grassland	Fen, Marsh, Swamp	Heather	Heather grassland	Bog	Inland Rock	Saltwater	Freshwater	Supralittoral rock	Supralittoral sediment	Littoral rock	Littoral sediment	Saltmarsh	Urban	Suburban	Total	User's Accuracy
Broadleaved woodland	1044	195	10	43	2	1	31	4	15	15	3	1		4	2			1		2	14	1387	**75.3**
Coniferous woodland	40	669	4			1	5		18	11	2	2										752	**89.0**
Arable	6	3	1879	235	6	4		4	3			16		1		2		1	2	9	9	2180	**86.2**
Improved grassland	49	3	362	3602	168	102	255	16	4	62	29	2				22	1			6	39	4722	**76.3**
Neutral grassland	4	2	23	125	431		3	22	2	2				3						1	4	622	**69.3**
Calcareous grassland	15		1	7	1	845	4			5		11										889	**95.1**
Acid grassland	19	3	6	135	5	87	1305		62	235	81	3		1							1	1943	**67.2**
Fen, Marsh, Swamp	3		1	14				536			3											557	**96.2**
Heather		2		1	1		20	1	814	60	101	1		1	1							1003	**81.2**
Heather grassland	18	2	10	5		8	163	3	137	375	210	5		1		1						938	**40.0**
Bog		3		1			41	4	33	71	927											1080	**85.8**
Inland Rock			1			7	8		4	2	3	154		2						4	1	186	**82.8**
Saltwater													80	2			2	10				94	**85.1**
Freshwater	3			1			3	5				2		522				1			1	538	**97.0**
Supralittoral rock	1			1			5				3			1	32	3	16	3				65	**49.2**
Supralittoral sediment			1	6	3	1	2				1	1			2	171		7			1	196	**87.2**
Littoral rock												1	1	3	23	5	75	37	1	2	1	149	**50.3**
Littoral sediment													14	3		14	10	189	2			232	**81.5**
Saltmarsh	3		4	17				37						3		12		5	163			244	**66.8**
Urban	2		3									15		1		2		2		1449	325	1799	**80.5**
Suburban	21	1	20	43	1				1			1				2				148	2511	2749	**91.3**
Total	1228	883	2325	4236	618	1056	1845	632	1093	838	1363	215	95	548	60	234	104	256	168	1621	2907	22325	
Producer's Accuracy	**85.0**	**75.8**	**80.8**	**85.0**	**69.7**	**80.0**	**70.7**	**84.8**	**74.5**	**44.7**	**68.0**	**71.6**	**84.2**	**95.3**	**53.3**	**73.1**	**72.1**	**73.8**	**97.0**	**89.4**	**86.4**	**Accuracy**	**79.6**

Confusion matrix for LCM2019

	Broadleaved woodland	Coniferous woodland	Arable	Improved grassland	Neutral grassland	Calcareous grassland	Acid grassland	Fen, Marsh, Swamp	Heather	Heather grassland	Bog	Inland Rock	Saltwater	Freshwater	Supralittoral rock	Supralittoral sediment	Littoral rock	Littoral sediment	Saltmarsh	Urban	Suburban	Total	User's Accuracy
Broadleaved woodland	1033	186	11	58	6	1	37	8	29	17	5	3		2	2					1	17	1416	**73.0**
Coniferous woodland	43	675	4				8		14	16	3			1								764	**88.4**
Arable	2	4	1873	261	6	4	1	3				13		2		1			2	5	11	2188	**85.6**
Improved grassland	47	3	396	3552	170	80	250	14	1	70	49	3				22	2			7	42	4708	**75.4**
Neutral grassland	12	3	5	146	426		2	17	2					6						1	6	626	**68.1**
Calcareous grassland	9	1	1	7	1	865	4		1	4		11										904	**95.7**
Acid grassland	14	2	6	140	4	97	1341		50	236	66	5		1						1		1963	**68.3**
Fen, Marsh, Swamp	3			6				532			3											544	**97.8**
Heather	2	3	3				15	3	815	58	111				1							1011	**80.6**
Heather grassland	25	2	1	3		2	124	3	135	353	192	7			1	1				3	1	853	**41.4**
Bog		2				1	47	6	37	81	929			1								1104	**84.1**
Inland Rock		1	1			6	6		5	3	3	154		1						4		184	**83.7**
Saltwater									1				85	2			3	10				101	**84.2**
Freshwater	5			1			2	4				2		520				1			1	536	**97.0**
Supralittoral rock				1			5		1		1			1	26	5	12	4				56	**46.4**
Supralittoral sediment			2	4	3		2		1		1				2	172		6				193	**89.1**
Littoral rock												1		1	28	8	69	32	1	2		142	**48.6**
Littoral sediment			1										10	4		14	18	197	2			246	**80.1**
Saltmarsh	4		2	14				42						3		5		5	163		2	240	**67.9**
Urban	2		1	2								14		3		1		1		1432	324	1780	**80.4**
Suburban	27	1	18	41	2		1		1			2				5				165	2503	2766	**90.5**
Total	1228	883	2325	4236	618	1056	1845	632	1093	838	1363	215	95	548	60	234	104	256	168	1621	2907	22325	
Producer's Accuracy	**84.1**	**76.4**	**80.6**	**83.9**	**68.9**	**81.9**	**72.7**	**84.2**	**74.6**	**42.1**	**68.2**	**71.6**	**89.5**	**94.9**	**43.3**	**73.5**	**66.3**	**77.0**	**97.0**	**88.3**	**86.1**	**Accuracy**	**79.4**

Appendix 5. Full list of datasets for LCM2017, LCM2018 and LCM2019

Product	Dataset	Land mass	Digital holding dataset name
UKCEH LCM2017	20m Classified Pixels	Great Britain	Land Cover Map 2017 (20m classified pixels, GB)
	Land Parcels	Great Britain	Land Cover Map 2017 (land parcels, GB)
	25m Rasterized Land Parcels	Great Britain	Land Cover Map 2017 (25m rasterised land parcels, GB)
	1km Dominant Cover	Great Britain	NOT YET AVAILABLE
	1km Dominant Aggregate Cover	Great Britain	
	1km Percent Cover	Great Britain	
	1km Percent Aggregate Cover	Great Britain	
	20m Classified Raster Product	Northern Ireland	Land Cover Map 2017 (20m classified pixels, N. Ireland)
	Land Parcel Product	Northern Ireland	Land Cover Map 2017 (land parcels, N. Ireland)
	25m Rasterized Land Parcel Product	Northern Ireland	Land Cover Map 2017 (25m rasterised land parcels, N. Ireland)
	1km Dominant Cover	Northern Ireland	NOT YET AVAILABLE
	1km Dominant Aggregate Cover	Northern Ireland	
	1km Percent Cover	Northern Ireland	
	1km Percent Aggregate Cover	Northern Ireland	
UKCEH LCM2018	20m Classified Raster Product	Great Britain	Land Cover Map 2018 (20m classified pixels, GB)
	Land Parcel Product	Great Britain	Land Cover Map 2018 (land parcels, GB)
	25m Rasterized Land Parcel Product	Great Britain	Land Cover Map 2018 (25m rasterised land parcels, GB)
	1km Dominant Cover	Great Britain	NOT YET AVAILABLE
	1km Dominant Aggregate Cover	Great Britain	
	1km Percent Cover	Great Britain	
	1km Percent Aggregate Cover	Great Britain	
	20m Classified Raster Product	Northern Ireland	Land Cover Map 2018 (20m classified pixels, N. Ireland)
	Land Parcel Product	Northern Ireland	Land Cover Map 2018 (land parcels, N. Ireland)
	25m Rasterized Land Parcel Product	Northern Ireland	Land Cover Map 2018 (25m rasterised land parcels, N. Ireland)
	1km Dominant Cover	Northern Ireland	NOT YET AVAILABLE
	1km Dominant Aggregate Cover	Northern Ireland	
	1km Percent Cover	Northern Ireland	
	1km Percent Aggregate Cover	Northern Ireland	
UKCEH LCM2019	20m Classified Raster Product	Great Britain	Land Cover Map 2019 (20m classified pixels, GB)
	Land Parcel Product	Great Britain	Land Cover Map 2019 (land parcels, GB)
	25m Rasterized Land Parcel Product	Great Britain	Land Cover Map 2019 (25m rasterised land parcels, GB)
	1km Dominant Cover	Great Britain	NOT YET AVAILABLE
	1km Dominant Aggregate Cover	Great Britain	
	1km Percent Cover	Great Britain	
	1km Percent Aggregate Cover	Great Britain	
	20m Classified Raster Product	Northern Ireland	Land Cover Map 2019 (20m classified pixels, N. Ireland)
	Land Parcel Product	Northern Ireland	Land Cover Map 2019 (land parcels, N. Ireland)
	25m Rasterized Land Parcel Product	Northern Ireland	Land Cover Map 2019 (25m rasterised land parcels, N. Ireland)
	1km Dominant Cover	Northern Ireland	NOT YET AVAILABLE
	1km Dominant Aggregate Cover	Northern Ireland	
	1km Percent Cover	Northern Ireland	
	1km Percent Aggregate Cover	Northern Ireland	