Recent Land Cover Change

Report for the Office for Environmental Protection

Clare Rowland, Dan Morton, Chris Marston and Lindsay Maskell

Issue number 1
08/01/2024

UK Centre for Ecology & Hydrology

Contents

1. Introduction

The Office for Environmental Protection (OEP) assesses government's progress towards improving the natural environment, including legally binding targets set through the Environment Act 2021 and outlined in the 2023 Environmental Improvement Plan (EIP) for England [1]. The 2023 EIP for England identifies 'Improving Nature', as a key goal.

The purpose of this commissioned report is to provide the OEP with monitoring data to inform their assessment of progress towards improving nature, in particular the Environment Act 2021 long-term wildlife-rich habitat restoration or creation target and to provide a critical appraisal of the applicability of the data for informing progress and policy.

This report summarises the work conducted by the UK Centre for Ecology and Hydrology (UKCEH) for the OEP. The work is split across three work packages (WPs):

WPA: Changes in land cover over time. Here, the OEP requested summary statistics from the UKCEH Land Cover Map (LCM) data for 1990 to 2021, along with commentary, plus comparison of LCM and the Living England habitat map (Kilcoyne *et al.,* 2022).
WPB: Application of the land cover and land use data. For this section of the report, the OEP requested an assessment of the capability of existing data sets for monitoring wildlife rich habitat (see WPB for specific questions).
WPC: Using land cover and land use data to inform the development of, and delivery of, the land use framework. For the final section of the report, the OEP requested commentary on the role of land use and landcover data for supporting and developing the Land Use Framework for England.

[1] Environmental Improvement Plan (publishing.service.gov.uk)

2. WPA: Changes in land cover over time

2.1 Land Cover Map

Land Cover Map production timeline

The UKCEH Land Cover Map series began in the early 1990's with the first Great Britain (GB) wide land cover map (Fuller *et al.,* 1994). However, the work here focusses on the land cover maps produced and released since 2016, which have been more consistent, in terms of land cover class and spatial structure, than the earlier land cover data sets. The land cover maps used in this analysis are listed in Table 1, including a note on the method used to create them, with the year of production shown in green.

The key factors to note from Table 1 are that the methods underlying the LCM data have continued to develop, and that the LCM data have been produced over a number of years, including production of historical LCM's for the 1990's-2010. These factors are important for understanding the Land Cover Map statistics (this is discussed further in section 2.4).

Table 1: Summary of the production timeline and processing method used for different Land Cover Map data sets (publication/completion year in green). **[a]BEIS variant refers to LCM's produced for the then department of Business Energy and Industrial Strategy (BEIS) in a project in 2021.**

	LCM method	2017	2018	2019	2020	2021	2022	2023
LCM1990	Traditional							
LCM1994	Cloud-based BEIS[a] variant							
LCM1998	Cloud-based BEIS variant							
LCM2002	Cloud-based BEIS variant							
LCM2006	Cloud-based BEIS variant							
LCM2010	Cloud-based BEIS variant							
LCM2015	Traditional							
LCM2017	Cloud-based v1							
LCM2018	Cloud-based v1							
LCM2019	Cloud-based v1							
LCM2020	Cloud-based v1							
LCM2021	Cloud-based v2							
LCM2022	In production							

Explaining the LCM methodologies is beyond the scope of the current report, but briefly, the traditional LCM method is based largely on the method used to produce

LCM2007 (Morton *et al.,* 2011). The traditional method used for LCM2015 and LCM1990 had two major advances on the LCM2007 method, specifically the use of the Random Forest classification algorithm and the use of a core set of training areas; both factors that reduced the time taken to produce an LCM in comparison to LCM2007. Despite this, the traditional method used for LCM2015 and LCM1990 was still time-consuming, and in 2020 the first cloud-production-based LCM data sets were released with LCMs 2017-2019 (Morton *et al.,* 2020a-c). This much faster production process enabled annual LCM's from 2017 onwards, and processing of new historical LCM data sets, as shown in Table 1. However, the cloud-based production method has also undergone some developments (Marston *et al.,* 2023; Rowland *et al.,* 2021), which in turn have an impact on the resulting LCM data.

Land Cover Map data sets

The UKCEH Land Cover Maps are delivered in a range of different formats, including vector and raster formats, with the raster data sets varying between 10m (LCM2019 onwards) and 1km resolution. When the LCM data sets are produced, per-pixel classifications are created first and these are then ingested into the LCM spatial framework – this is a UK-wide set of polygons created by generalising Ordnance Survey MasterMap data (Morton *et al.,* 2011). This vector data set is then used to create a 25m raster data set, known as the 25m rasterised polygons data. The 25m rasterised polygon data set is the legacy product, which retains spatial consistency throughout the available LCM time-series, so it was used here. Access details and references for the LCM data are summarised in Table 2.

Table 2: Reference and digital object identifiers (where available) for the Land cover Map (LCM) data sets used.

LCM	Digital object identifier	Reference
LCM1990	Land Cover Map 1990 (25m raster, GB) v2	Rowland *et al.,* (2020)
LCM1994	Not currently published	Rowland *et al.,* (2021)
LCM1998	Not currently published	Rowland *et al.,* (2021)
LCM2002	Not currently published	Rowland *et al.,* (2021)
LCM2006	Not currently published	Rowland *et al.,* (2021)
LCM2010	Not currently published	Rowland *et al.,* (2021)
LCM2015	Land Cover Map 2015 (25m raster, GB)	Rowland *et al.,* (2017)
LCM2017	Land Cover Map 2017 (25m raster, GB)	Morton *et al.,* (2020)
LCM2018	Land Cover Map 2018 (25m raster, GB)	Morton *et al.,* (2020)
LCM2019	Land Cover Map 2019 (25m raster, GB)	Morton *et al.,* (2020)
LCM2020	Land Cover Map 2020 (25m raster, GB)	Morton *et al.,* (2021)
LCM2021	Land Cover Map 2021 (25m raster, GB)	Marston *et al.,* (2022)

Calculating Land Cover Map Statistics

Summary statistics were calculated for England using an R script (Appendix 1). R statistics is a widely used statistical package (R Core Team, 2023) and facilitates the production of reproducible workflows. The summary statistics for England were calculated with the Ordnance Survey (OS) high and low water tidal extents. The data provided by the OS directly are in polyline format. However, for the analysis here polygon boundary data is more appropriate. The Office for National Statistics (ONS) also use high and low tide extents to constrain their geographical statistics. Once a year the ONS get the latest high and low water mark data from the OS and convert it from polyline data to polygon data (see ONS, 2011 for more details). ONS publish the polygon version of the OS's high and low tide data sets on their website, in full and generalised resolution versions; the full resolution data were used here. The most recent mean high and low mark data sets were downloaded (Table 3) and used to extract all the statistics presented in this report. ONS do have boundary data for some earlier years, with their earliest data for 2011, however none of the earlier years were downloaded or explored, so it is not clear how much change they would have on the extracted statistics. The assumptions (for this report) are that:

a) It is better to use a single consistent set of mean high and low tide boundaries, rather than to change the boundaries for the relatively few, recent years, where using different boundaries may have been possible.
b) The change in area of England between 2011 and 2022 is relatively small, so would have minimal impact on the statistics presented here.

A final critical factor is that the Land Cover Map spatial framework is currently fixed, so it covers the same extent for 1990 through to 2021, and consequently determines which areas are mapped.

Table 3: Vector data sets used to extract the statistics using high and low tide marks.

Data set	Source (filename in brackets)
Mean high water mark (clipped to the coastline)	https://geoportal.statistics.gov.uk/datasets/ons::countries-december-2022-boundaries-gb-bfc/about (CTRY_DEC_2022_GB_BFC.shp)
Mean low water mark (extent of the realm)	https://geoportal.statistics.gov.uk/datasets/ons::countries-december-2022-boundaries-gb-bfe/about (CTRY_DEC_2022_GB_BFE.shp)

Land Cover Map classes

The UKCEH LCM data sets are typically released in 21-class and 10-class versions (Table 4). However, the land cover change data set, released in 2020, used a 6-class version of the LCM data set to map land cover change between 1990 and 2015. For the purposes of change detection, aggregating classes is useful as it allows classes that are less accurately classified to be combined to increase accuracy. For example, for LCM2021, the classification at the 21-class level has an overall accuracy of 82.6% and at the 10-class level the accuracy increases to 86.5% (see Tables 3 and 4 in Marston *et al*., 2023). This is because some of the classes, where there was confusion, have been combined (aggregated). LCM2021 is shown in Figure 1.

Table 4: Thematic classes for the 21- and 10-class versions of Land Cover Map data sets used here, showing relationship with broad habitats. Note – class numbers are included for the 21-class data (class numbers are used in the raster data sets).

LCM target class name	Target class	Associated Broad Habitat	Aggregate class
Broadleaved woodland	1	*'Broadleaved, mixed and yew woodland'*	Broadleaf woodland
'Coniferous woodland'	2	*'Coniferous woodland'*	Coniferous woodland
'Arable and horticulture'	3	*'Arable and horticulture'*	Arable
'Improved grassland'	4	*'Improved grassland'*	Improved grassland
'Neutral grassland'	5	*'Neutral grassland'*	Semi-natural grassland
'Calcareous grassland'	6	*'Calcareous grassland'*	
Acid grassland	7	*'Acid grassland'*	
'Fen, marsh and swamp'	8	*'Fen, marsh and swamp'*	
Heather	9	*'Dwarf shrub heath'*	Mountain, heath, bog
Heather grassland	10		
'Bog'	11	*'Bog'*	
'Inland rock'	12	*'Inland rock'*	
Saltwater	13	Saltwater	Saltwater
Freshwater	14	Freshwater	Freshwater
'Supra-littoral rock'	15	*'Supra-littoral rock'*	Coastal
'Supra-littoral sediment'	16	*'Supra-littoral sediment'*	
'Littoral rock'	17	*'Littoral rock'*	
Littoral sediment	18	*'Littoral sediment'*	
Saltmarsh	19		
Urban	20	*'Built-up areas and gardens'*	Built-up areas and gardens
Suburban	21		

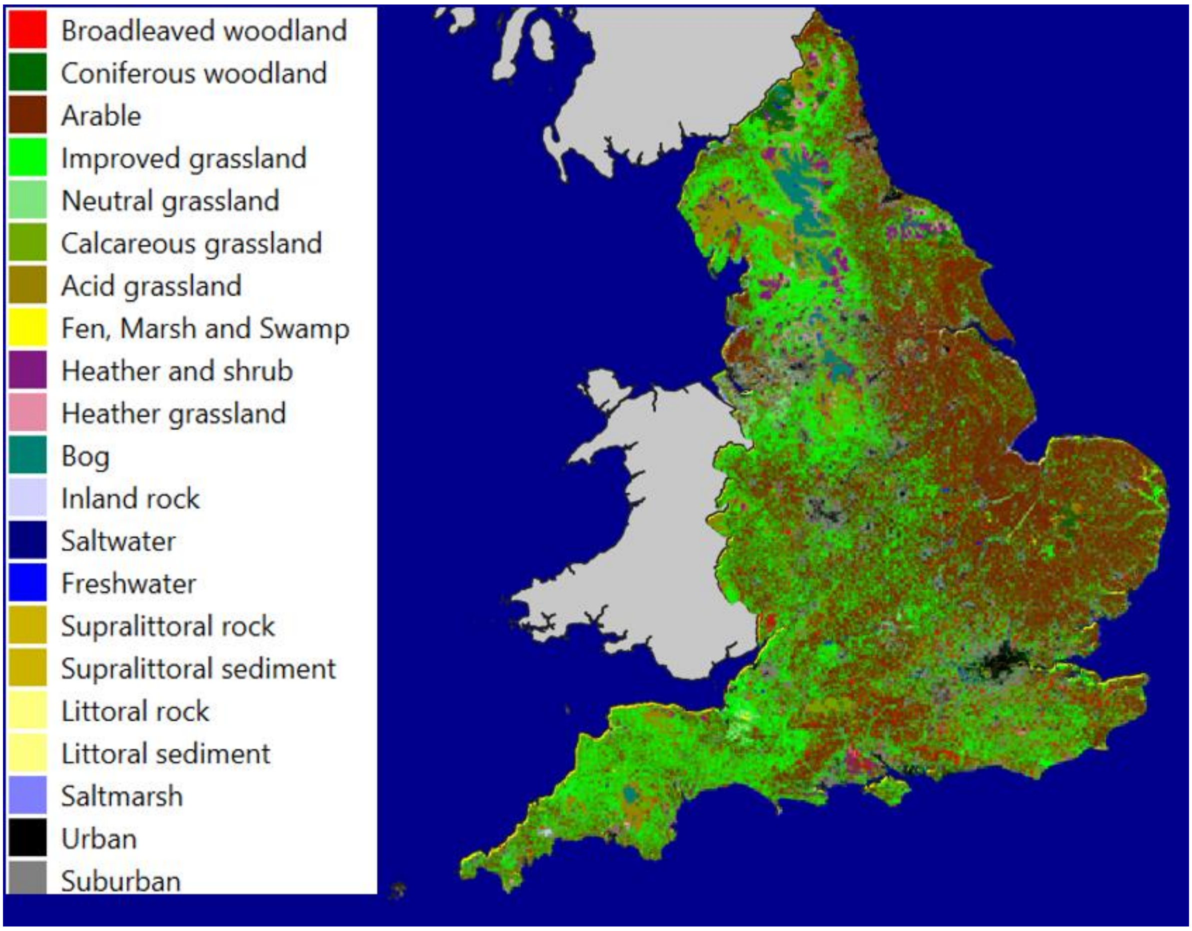


Figure 1: LCM2021 for England.

Accuracy of the Land Cover Map data

For any land cover data set it is important to determine the accuracy of the data. The accuracy of land cover data sets is assessed through validation, where the land cover data is compared to a set of reference points of known land cover type. LCM2021 was validated on 35182 reference points collected from a variety of sources (see Marston *et al.,* 2023 for details). Table 5 is reproduced from Marston *et al.,* (2023) and shows how the accuracy varies across the different classes in LCM2021. A perfectly classified land cover map, with perfect reference data, would have all the values down the main diagonal (the main diagonal is highlighted in Table 5). Numbers greater than zero, that are not on the main diagonal, illustrate areas where different land cover classes are being confused. Note, for illustrative purposes the 10-class version of the confusion matrix is shown, the 21-class version is in Marston *et al.,* (2023).

Table 5 includes the User's and Producer's accuracy. The Producer's accuracy shows how well the reference data are classified by LCM. So, in the case of Arable, the producer's accuracy shows that 88.8% of the arable reference points are correctly classified. Whereas the user's accuracy gives the probability that a pixel classified as a particular class is actually that class. The user's probability for Arable is 96.1%, so there's a 96.1% chance that an arable pixel in the classification is Arable. Together the two numbers suggest that when LCM2021 maps arable in the classification, it is generally arable (high user's accuracy of 96.1%), but there is a tendency for LCM2021 to under-estimate the amount of arable when compared to the reference data (producer's accuracy of 88.8%). Taken together the two numbers illustrate areas of uncertainty, which can be used to drive improvements. The information can also be used to statistically improve the estimates of land cover (see for example Henrys & Jarvis, 2019).

The confusion in land cover classifications arises mainly because either the classes are not spectrally distinct, so they are difficult for the classification algorithm to separate, or because the training data has mislabelled training samples, for example, an arable field labelled as improved grassland. The core set of training areas used for training the land cover map is derived from existing land cover map data, but increasingly this is being filtered with additional data sets, such as the National Forest Inventory (NFI) data set for woodland classes, or satellite derived indices to separate, for example, improved grassland and acid grassland.

The confusion between arable and improved grassland is thought to be due to several issues, including some out-dated arable/improved grassland validation data, and a slight mismatch between agricultural seasons and the calendar year used by the land cover map, which means that a field maybe both arable and improved grassland within a calendar (i.e., an LCM) year, as agricultural land rotates between arable and grassland.

Coniferous woodland being classified as broadleaf woodland, seems to be partly due to newly planted coniferous stands being mis-classified. The confusion around semi-natural grassland is due to confusion of some semi-natural grasslands with more improved grassland, this particularly affects Neutral grassland (Marston *et al.,* 2023). Plus, there is a more general tendency for confusion between the semi-natural grassland classes and the upland classes, as semi-natural habitats often occur together in mosaics, or show transitions through several classes across an environmental gradient like a hillside.

Classification accuracy and improvements

In terms of accuracy, comparisons suggest that the Land Cover Map data is better than the CORINE (Coordination of Information on the Environment) land cover data set for England (Burke *et al.,* 2020). When a final version of Living England for 2021/22 is published, rather than the current interim product, then the Land Cover Map and Living England can be compared; doing this will require validation data that is suitable for both classifications (see also section 2.2 Comparison of land cover estimates).

Comparing the accuracy of LCM to the accuracy of national data sets produced in other countries is more complicated, because the challenges involved in land cover mapping vary from country to country, as do the classes required in the resulting classifications and the quality and quantity of the validation data. The National Land Cover Database (NLCD) is the US equivalent to LCM. The NLCD accuracy assessments are conducted at two-levels, Level I, which has 8-classes and Level II, which has 16-classes. For NLCD2019, the Level II overall accuracy was 77.5% and the Level I accuracy was 83.1% (Wickham *et al.,* 2023), which are lower than the 82.6% (21-class) and 86.5% (10-class) of LCM2021 (Marston *et al.,* 2023).

For the UKCEH Land Cover Map we do not currently have a target accuracy, although other land cover map activities have proposed target accuracies, for example, an overall accuracy of 85% and individual class accuracies > 70% (Thomlinson *et al.,* 1999). The Department of Energy Security and Net Zero (DESNZ) has set a target accuracy of an overall accuracy of greater than 95%, with no individual class user and producer accuracies less than 90%, for the UK's Greenhouse Gas Inventory work. However, this is for a simplified set of six land cover classes (forestland, cropland, grassland, wetlands, settlements (urban), other).

The current UKCEH land cover map strategy for improving accuracy is via incremental improvements in individual classes, as funding or improved training data allow. Increased funding would enable more substantial improvements. Potential improvements could be targeted at the input data sets and improving the separability of different classes, or on improving the quality and spatial distribution of the training areas. Improvements to the quality, timeliness and distribution of the validation data set may also help, as it would provide a better evidence-base for targeting improvements.

Land Cover Map Statistics

The land cover statistics were extracted for England at the high and low tide mark for each of the LCM years. Statistics, using the high tide mark as a boundary constraint, are shown in Table 6 and Figures 2-4.

Table 5: Confusion (or correspondence) matrix for the 10-class version of LCM2021 (main diagonal highlighted) for the UK. The values are the number of field reference points, except for final row and column, which are percentages. Producer's accuracy = percentage of ground reference points classified correctly. User's accuracy = probability (expressed as a percentage) of a pixel of a particular class being correctly classified.

		Reference data											
		Broadleaf woodland	Coniferous woodland	Arable	Improved grassland	Semi-natural grassland	Mountain, heath & bog	Saltwater	Freshwater	Coastal	Built-up areas & gardens	Total points	User's Accuracy (%)
Classification data	**Broadleaf woodland**	1704	218	19	73	34	37	0	9	3	15	2112	**80.7**
	Coniferous woodland	55	649	3	1	6	24	0	0	0	1	739	**87.8**
	Arable	22	3	10102	306	31	14	0	6	1	26	10511	**96.1**
	Improved grassland	100	4	1027	4835	434	116	0	8	31	33	6588	**73.4**
	Semi-natural grassland	83	16	158	468	3390	325	0	12	1	7	4460	**76.0**
	Mountain, heath & bog	24	7	28	18	249	2641	0	4	2	9	2982	**88.6**
	Saltwater	0	0	0	0	0	0	73	0	13	0	86	**84.9**
	Freshwater	13	0	1	4	4	0	0	548	0	2	572	**95.8**
	Coastal	3	0	3	12	26	3	9	2	796	1	855	**93.1**
	Built-up areas & gardens	170	1	29	252	63	40	0	6	14	5702	6277	**90.8**
	Total points	2174	898	11370	5969	4237	3200	82	595	861	5796	35182	
	Producer's Accuracy (%)	**78.4**	**72.3**	**88.8**	**81.0**	**80.0**	**82.5**	**89.0**	**92.1**	**92.5**	**98.4**	**Overall accuracy: 87%**	

Table 6: Area of each Land Cover Class, according to the current set of LCM data, for England in km^2, with the high tide mark as a boundary constraint, from the 25m rasterised version of the LCM data sets.

	LCM1990	LCM1994	LCM1998	LCM2002	LCM2006	LCM2010	LCM2015	LCM2017	LCM2018	LCM2019	LCM2020	LCM2021
Broadleaved woodland	8775	8633	9075	9182	9310	9764	9767	10825	10489	10748	10792	10454
Coniferous woodland	2625	2742	2788	2879	2951	2962	2979	2853	2804	2803	2635	2722
Arable	47788	49653	49547	49097	48062	47339	47757	46634	46380	46068	45582	44599
Improved grassland	44803	41509	40575	40570	41683	41418	42914	41601	42624	42523	41871	41505
Neutral grassland	1304	898	1076	1245	1073	1271	608	797	731	775	1027	1446
Calcareous grassland	2213	2738	2713	2707	2715	2692	816	1079	985	1052	1328	2120
Acid grassland	4337	5002	5137	5172	4879	4952	4712	4833	4734	4694	4721	4465
Fen, Marsh, Swamp	281	311	367	308	318	287	114	106	100	98	134	298
Heather	2502	2275	2225	2265	2258	2236	1742	1671	1612	1581	1737	1898
Heather grassland	1044	806	823	850	825	787	957	1064	1031	1048	1112	1158
Bog	1344	1729	1691	1730	1746	1775	1963	1956	2036	2015	1938	2034
Inland Rock	407	275	263	267	277	268	235	165	179	182	166	190
Saltwater	30	19	23	21	21	23	11	10	11	11	12	10
Freshwater	664	733	761	787	779	801	898	948	949	964	886	918
Supra-littoral rock	63	52	52	51	53	49	32	37	36	37	57	58
Supra-littoral sediment	180	154	161	158	161	156	196	213	238	213	260	155
Littoral rock	5	5	3	3	4	4	15	22	22	22	20	23
Littoral sediment	114	147	144	139	135	133	118	107	108	106	112	95
Saltmarsh	384	269	273	283	291	290	400	637	620	615	683	389
Urban	2707	3042	3102	3200	3279	3331	3356	3454	3628	3644	3424	3499
Suburban	8890	9468	9662	9545	9639	9918	10870	11431	11126	11245	11962	12424
Total	130459	130459	130459	130459	130459	130459	130459	130443	130443	130443	130459	130459

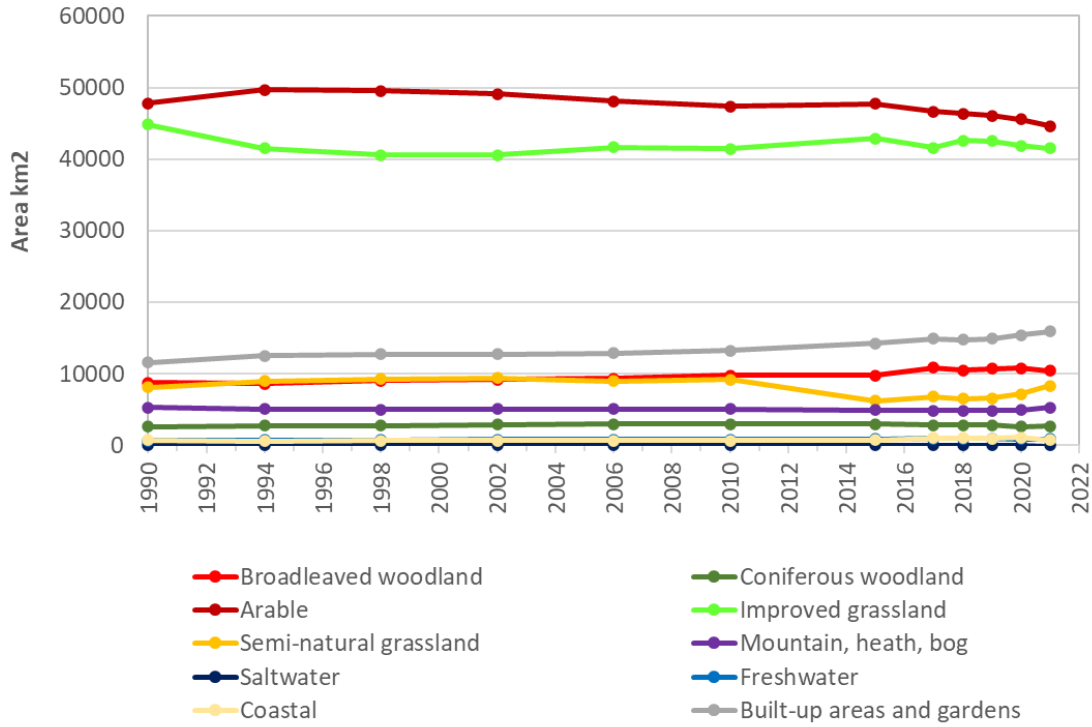


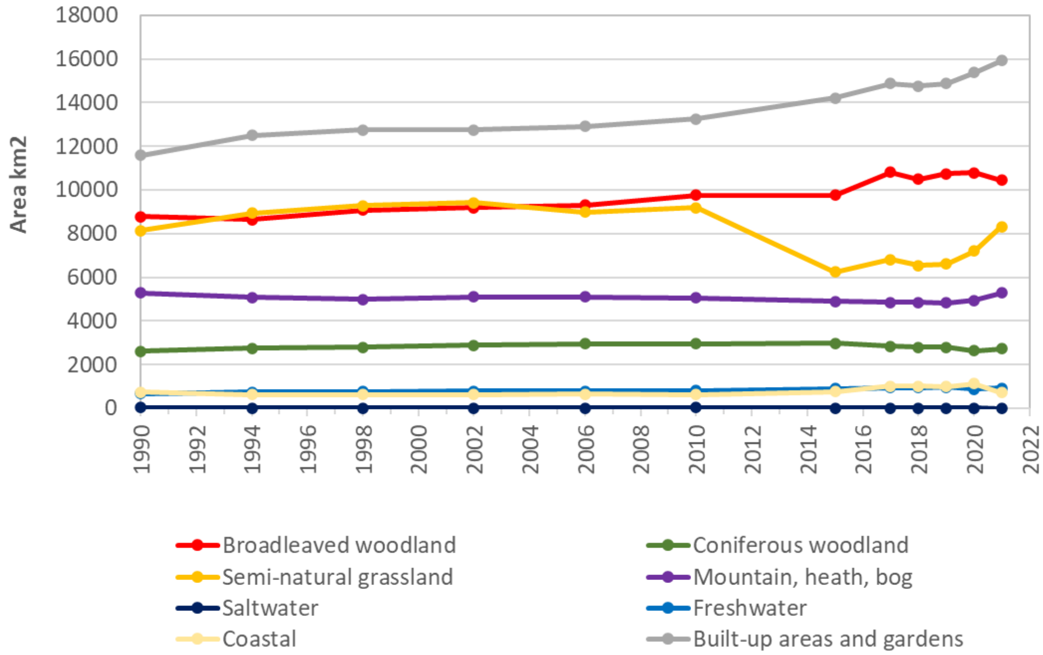
Figure 2: Land Cover Change for England from the aggregate-class versions of the LCMs for 1990 to 2021 (using mean high water as boundary constraint).

Figure 2 shows the change in Land Cover in England from the 10-class version of the Land Cover Maps. It is dominated by Arable and Improved grassland (the main land cover classes in England), so to show the trends in the smaller classes, Figure 3a) and b) duplicate the information in Figure 2, but with a rescaled y-axis. The main trends in land cover change are:

- Decreases in the amount of Arable land in England since the mid-1990's (Figure 2)
- Gradual, on-going increases in the extent of the *Built-up areas and gardens* and *Deciduous woodland* classes (Figure 3).

Semi-natural grassland (Figure 3a) shows an abrupt change in 2015, this is due to changes in method in LCM2015. LCM2015 is the earliest of the recent land cover maps (see Table 1 for details of the production timeline and method) and there were some issues that led to the under-estimation of *Semi-natural grassland* in and a slight over-estimate in the amount of *Improved grassland*. This became a focus in the following land cover maps, so has improved, but this does affect the accuracy of the *Semi-natural grassland* and the *Improved grassland* classes over time. Without further research it's not clear whether the increase in arable between 1990-1994, and the converse decrease in *Improved grassland* is real or methodological.

a) Classes < 18000km² in extent.



b) Classes < 6000 km² in extent.

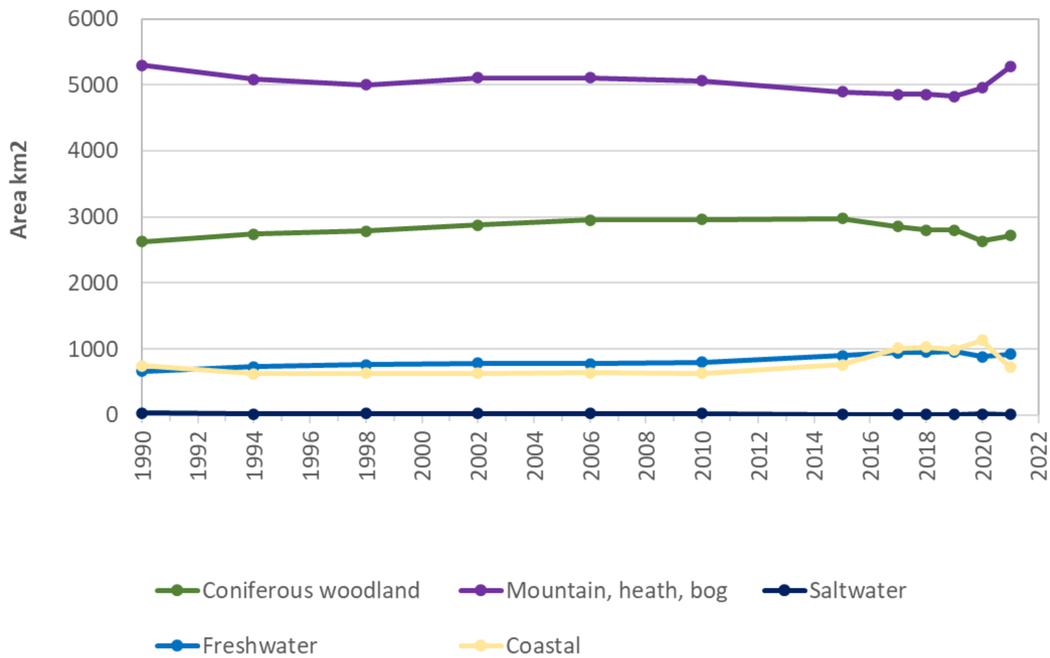


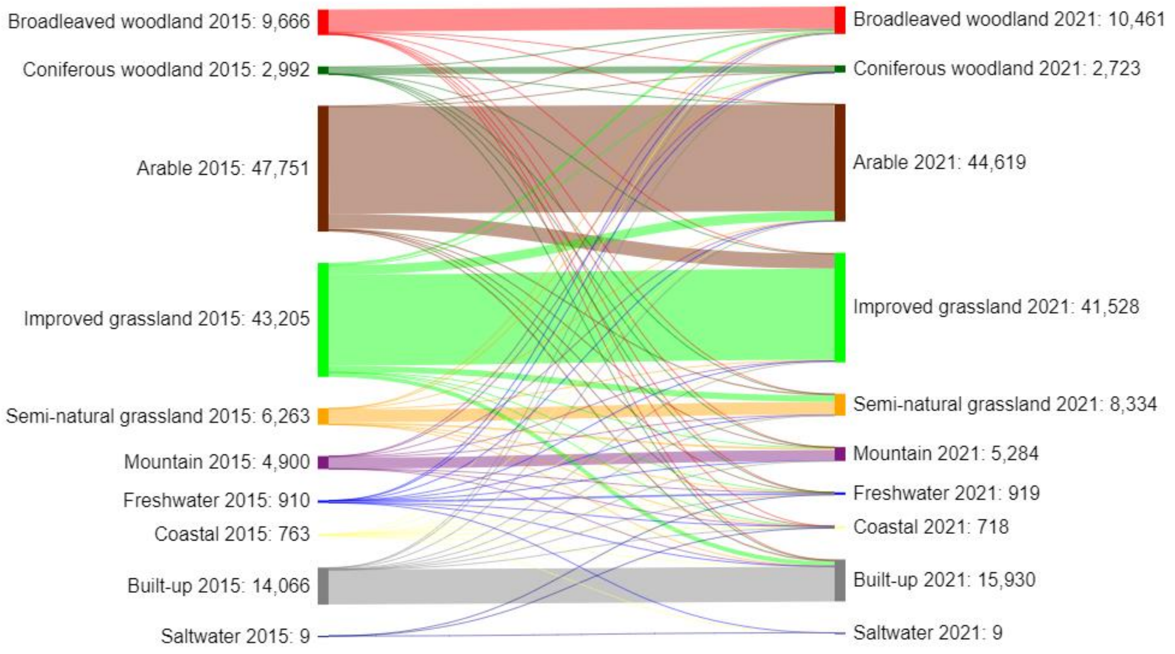
Figure 3: As figure 2, but re-scaled to show rarer classes better, for a) for classes less than 18000 km² in extent for England, and b) for classes less than 6000 km² in extent.

Figure 2 and 3 show the changes in the overall amounts of land cover (net change), but they do not show the flows between the different land cover types (gross change). Figure 4 shows two Sankey plots, which show the flows of land cover change, with the width of the bars representing the scale of the different flows. Figure 4a) shows all the flows, for the 10 aggregated classes, whilst Figure 4b) shows a simplified version, showing the top 5 flows, which along with existing land cover account for 95% of the land in England; the 3 smallest classes (Freshwater, Saltwater and Coastal) are merged into a coastal & water class).

Most land does not change class between 2015 and 2021, so the width of the stable areas e.g. Arable in 2015 to Arable in 2021 is wide, whilst the flow between classes is narrower e.g. from *Improved grassland* in 2015 to *Urban* in 2021.

Figure 4 shows that the increases in *Built-up areas and Gardens* and *Deciduous Woodland*, between 2015 and 2021, are mainly due to the loss of *Improved grassland*. Sankey plots can also capture the ‘churn’ caused by arable-grassland rotations.

a) All flows and 10-classes



b) Top 5 flows and 8-classes (coastal, freshwater and saltwater classes merged)

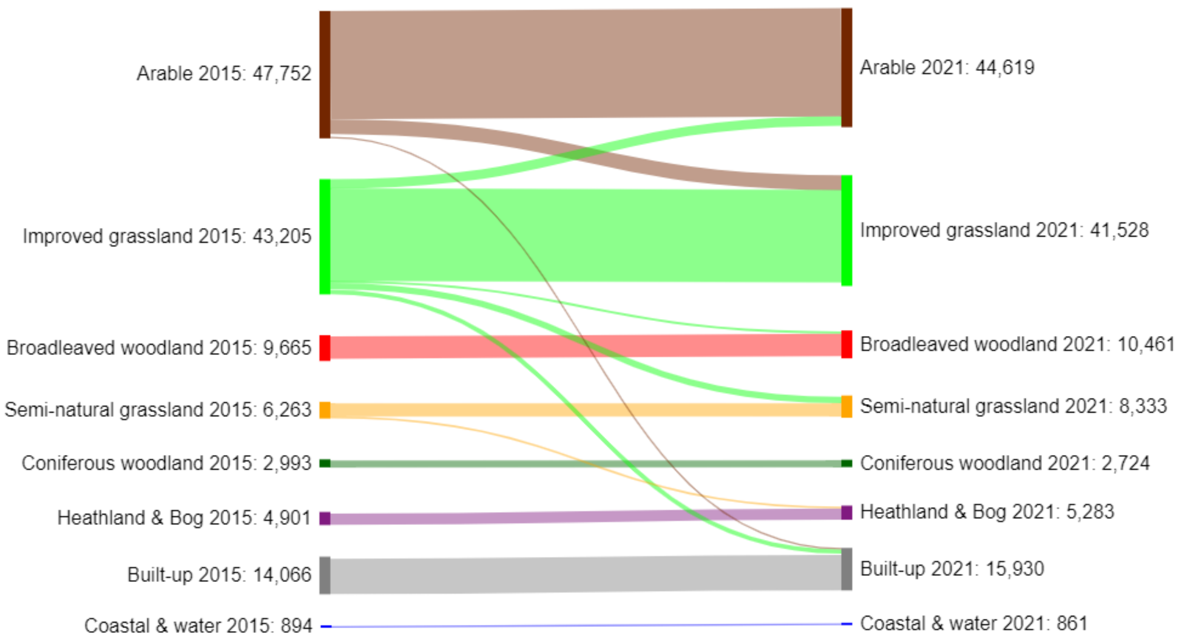


Figure 4: Sankey plots showing the flow of land cover in km^2 between LCM2015 and LCM2021, for a) all flows and b) the top 5 flows, which along with existing land cover constitute 95% of the total area of England. Plot produced with Sankeymatic.

2.2 Comparison of land cover estimates

This section compares the UKCEH land cover map area estimates with those from two other data sets, specifically Living England and Countryside Survey 2007. The Living England data set is currently under-development by Natural England and is intended, on maturity, to provide the data for the D1 Habitat Quantity and Connectivity components of the Department for Environment Food & Rural Affairs (Defra) Outcome Indicator Framework; the Quality component of D1 will be provided from other sources. Countryside Survey is included as the other established source of habitat extent data for England.

Living England (LE) Habitat Map

The Living England Habitat Map is a national habitat layer, derived from satellite data, to support the Environmental Land Management (ELM) System and the Natural Capital and Ecosystem Assessment (NCEA) Pilot. The Living England map is intended to provide land cover for the habitat extent and connectivity elements of Defra's D1: Quantity, Quality and connectivity of habitats indicator[2]. An interim version of Living England (the phase 4 data) is currently available, although it will be replaced by an updated version in 2024. Currently the D1 results just provide a single set of habitat extent estimates based on the Phase 4 data (see below).

The Phase 4 Living England data set was downloaded from: https://naturalengland-defra.opendata.arcgis.com/datasets/Defra::living-england-habitat-map-phase-4/about to extract the statistics for comparison with LCM. The data were also explored to understand the dates of the imagery going into the classifications. Living England is based on spring and autumn composite images with the earliest images from September 2019 and the latest images from April 2021 (Figure 5). It has 16 detailed classes that are (like LCM) based largely on the Biodiversity Action Plan Broad Habitat classes (as defined in Jackson, 2002). The 16 detailed classes can also be aggregated into a simpler set of LE Broad Habitats that have some similarities with the 10 aggregated classes used by the Land Cover Map.

[2] D1: Quantity, quality and connectivity of habitats (defra.gov.uk)

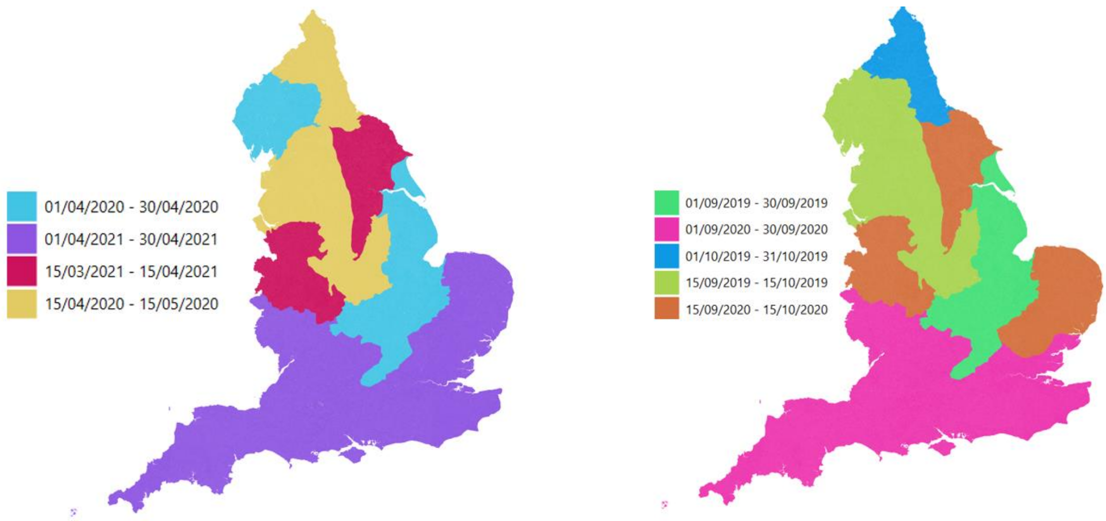


Figure 5: Dates of a) spring and b) autumn imagery for the Living England phase 4 map.

Table 7: Living England habitat classes (based on Table 1 of Kilcoyne *et al.*, 2022).

Living England - Detailed Habitat	Living England - Broad Habitat
Acid, Calcareous, Neutral Grassland	Grassland
Arable and Horticultural	Cropland
Bare Ground	Bare Ground
Bare Sand	Bare Ground
Bog	Wetland
Bracken	Grassland
Broadleaved, Mixed and Yew Woodland	Woodland
Built-up Areas and Gardens	Urban
Coastal Saltmarsh	Coastal
Coastal Sand Dunes	Coastal
Coniferous Woodland	Woodland
Dwarf Shrub Heath	Heath
Fen, Marsh and Swamp	Wetland
Improved Grassland	Grassland
Scrub	Woodland
Water	Freshwater
Unclassified	Unclassified

Countryside Survey

Monitoring of Biodiversity Action Plan (BAP) Broad Habitats in the UK is achieved, in part, through the Countryside Survey programme. The aim of the Countryside Survey is to provide information on the state of the countryside as a whole, rather than just protected areas or key habitats. The Countryside Survey programme began in 1978, with further surveys conducted in 1984, 1990, 1998 and 2007. Originally, Countryside Survey surveyed all squares in a single survey year, with reporting following soon after completion of the survey year. Since 2019 a 5-year rolling programme has been running, with a subset of squares surveyed each year, because it now takes 5-years to cover all the squares this has yet to report.

The field survey is underpinned by a stratification of the UK called the Land Classification; it assigns each 1km square in the UK to a specific land class based on a multivariate classification of environmental characteristics, including climate and geology (Bunce *et al.,* 1996). Intensive field surveys are conducted for a number of randomly located 1km squares within each Land Classification class, with 591 squares surveyed in 2007 (Carey *et al.,* 2008). The stratified random sampling of the field survey is necessary to extrapolate the field survey results to larger geographic areas and is used here to provide an independent estimate of land cover extent for England. Note the values used here are from Countryside Survey 2007, which is the most recent published year.

Harmonising classes between Land Cover Map and Living England

Comparing different land cover data sets is complicated, because products may be acquired at different times, with different classes, and also different interpretations of classes that are apparently the same i.e. Broadleaf woodland might have different definitions between products, particularly at the boundaries of the class around issues such as scrub, minimum tree height, minimum tree cover and treatment of mixed woodland.

In terms of timing, Living England and LCM differ slightly. The current version of Living England uses images from September 2019 to April 2021, whilst LCM2021 uses images from the 1st Dec 2020 to 31st Jan 2022 (Marston *et al.,* 2023) and LCM2020 uses images from January 1st 2020-December 31st 2020 (Morton *et al.,* 2022). This means that temporally, Living England currently overlaps with LCM's 2019, 2020 and 2021.

In terms of classes, Table 8, shows how the Living England and Land Cover Map classes have been compared. Generally, the main terrestrial habitats/land covers are relatively similar between the two land cover data sets, as they are both based on the BAP Broad Habitats (Jackson, 2002), so the correspondence is straightforward, but for other classes the correspondence maybe poorer and in some cases there is not an appropriate class. For example, Living England has a Bracken class and a Bare Ground class, for which LCM has no direct equivalent. In the results here, Bracken, has been kept as a separate class, without an LCM equivalent, because Countryside Survey does have a Bracken class. However, Bracken could justifiably have been included in the Acid, Calcareous, Neutral Grassland class, as in LCM Bracken will generally be incorporated into the LCM Acid Grassland class. The coastal classes and the Bare Ground are more difficult to harmonise, and there maybe differences between the classes that are not accounted for here.

Table 8: Comparison of classes between Land Cover Map and the Living England map.

Harmonised class	LCM classes	Living England classes
Broadleaved woodland and scrub	Broadleaved woodland	Broadleaved, Mixed and Yew Woodland
		Scrub
Coniferous woodland	Coniferous woodland	Coniferous woodland
Arable and Horticulture	Arable	Arable and Horticultural
Improved grassland	Improved grassland	Improved grassland
Acid, Calcareous, Neutral Grassland	Neutral grassland	Acid, Calcareous, Neutral Grassland
	Calcareous grassland	
	Acid grassland	
Fen, Marsh, Swamp	Fen, Marsh, Swamp	Fen, Marsh, Swamp
Dwarf Shrub Heath	Heather	Dwarf Shrub Heath
	Heather grassland	
Bog	Bog	Bog
Bare ground	Inland Rock	Bare ground
	Littoral rock	
	Supra-littoral rock	
Water	Saltwater	Water
	Freshwater	
Coastal sand dunes	Supra-littoral sediment	Coastal sand dunes
Bare sand	Littoral sediment	Bare sand
Saltmarsh	Saltmarsh	Coastal Saltmarsh
Built-up Areas and Gardens	Urban	Built-up Areas and Gardens
	Suburban	

2.3 Comparison of land cover statistics

Figure 6 and Figure 7 show the land cover from the most recent Land Cover Map data set (2021), plus Countryside Survey results from 2007 and Living England phase 4 data (2019-2021). Figure 7 shows that for woodland (coniferous and broadleaf), and Arable, the data sets have relatively similar values. However, grassland is much more variable, with Land Cover Map having most Improved grassland and least Acid, Neutral and Calcareous (ANC) grassland, whilst Living England has least Improved and most ANC grassland, with Countryside Survey falling between the two in both cases. For Built-up Areas and Gardens, LCM has the highest values, Living England the lowest, and CS is slightly higher than Living England.

The most noticeable difference in Figure 7b is between the different Fen, Marsh, Swamp (FMS) values. This is primarily due to a known over-estimation of FMS in the current version of Living England (Kilcoyne *et al.,* 2022). However, previous work has shown that it is difficult to detect small patches of habitats, like FMS from remote sensing (Morton *et al.,* 2011), as many of the patches detected by field survey are below the level that the Land Cover Map has traditionally been able to detect. Living England and LCM are similar for Bog and Water but are more different for Dwarf Shrub Heath. Understanding the differences between the mapping of the different classes can be time-consuming to untangle and typically requires detailed comparison of the spatial data. Comparison of the spatial data can also reveal undocumented differences in the boundaries of seemingly similar classes.

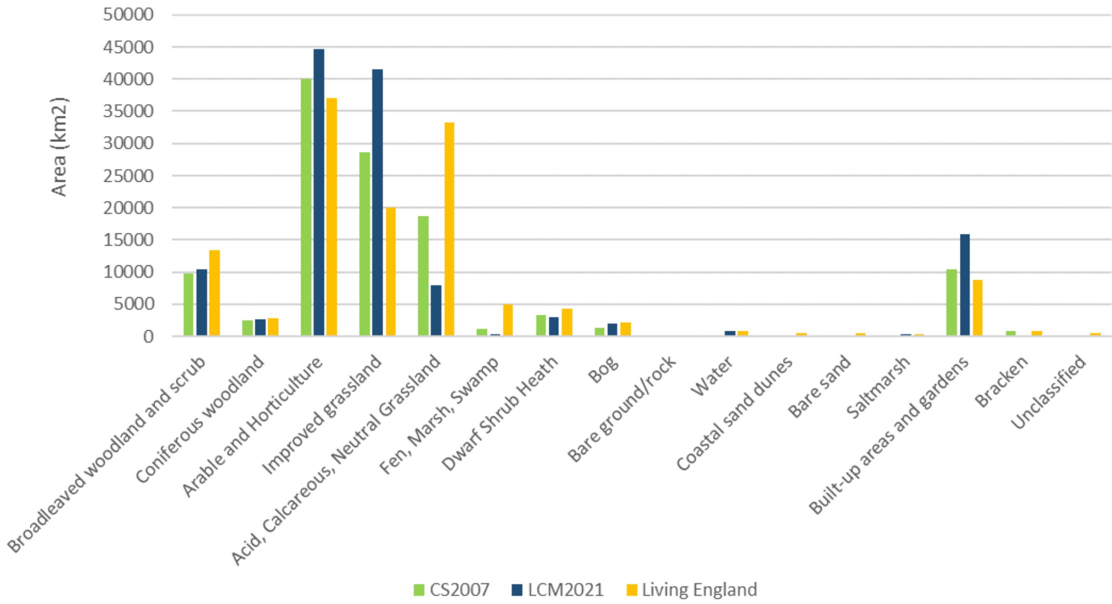
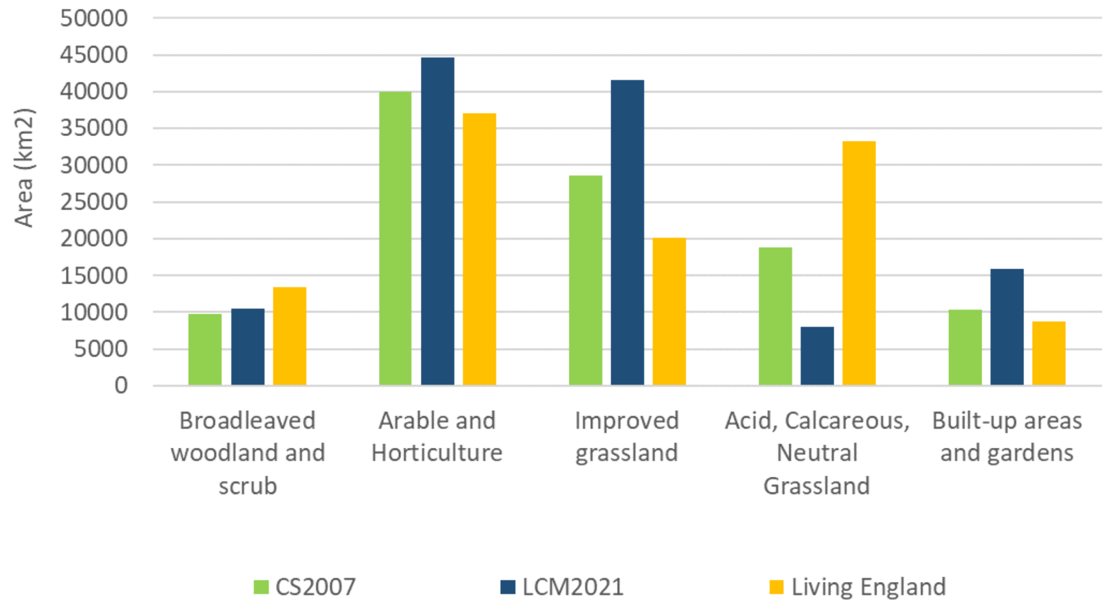


Figure 6: Comparison of land cover area (km^2) from Living England, Land Cover Map 2021 and Countryside Survey for *all classes*.

a) vegetated classes exceeding 5000km^2



b) vegetated classes below 5000km^2

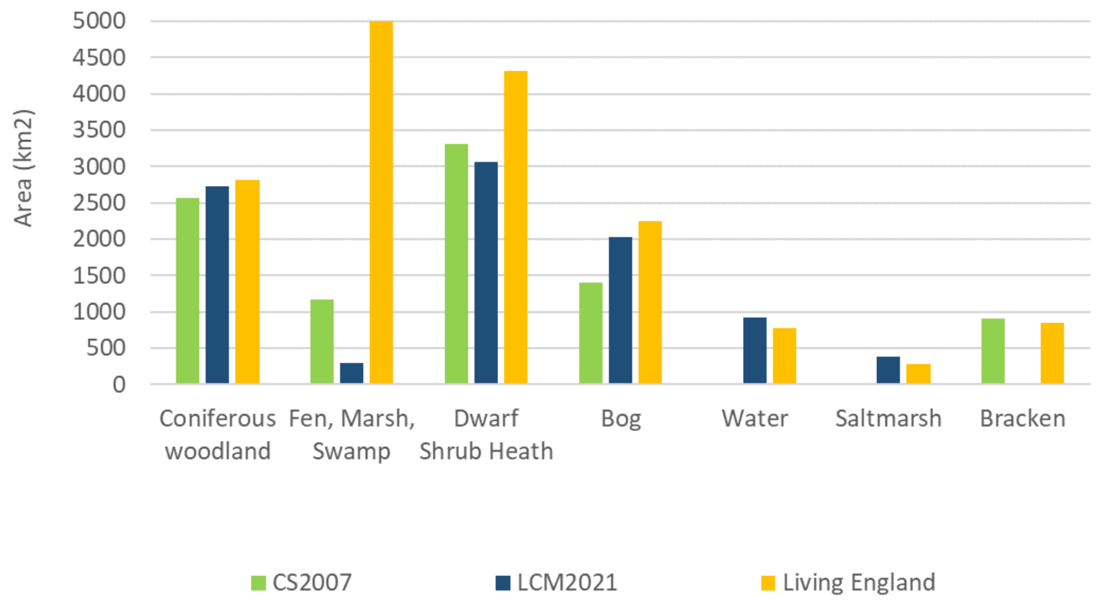


Figure 7: As Figure 6, but for a) vegetated classes exceeding 5000km^2, and b) vegetated classes below 5000km^2 (excluding coastal classes).

2.4 Impact of high and low tide extents on land cover estimates

The OEP requested that statistics were calculated for both the mean high and low water marks, so that the difference in habitat extent between the two boundary conditions could be calculated, whilst also ensuring that all coastal habitats were captured.

LCM statistics for England were calculated using both mean high and low tide water marks as boundary constraints. Figure 8 shows the difference spatially between the two different extents. Note, that the mean low water mark data set extends in some cases, such as in the Bristol channel (inset in Figure 8), to include the area between the mainland and islands (ONS, 2011), and inland on the tidal stretches of rivers like the Thames.

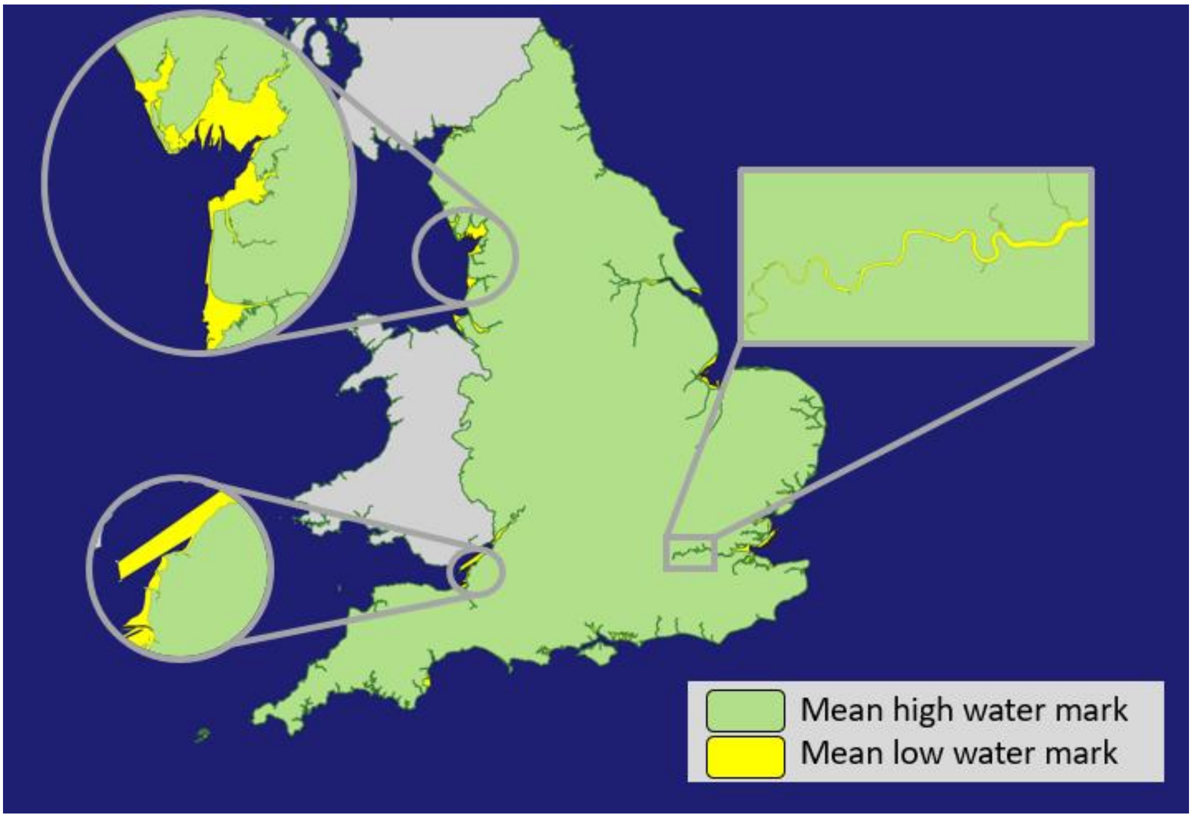


Figure 8: Spatial difference between the mean high-water mark (clipped to the coastline) and mean low water mark (extent of the realm) for England.

Exploring the difference between land cover extent clipped to the high and low water marks respectively, shows that the differences are primarily in the extent of the Saltwater and Littoral sediment classes, which are the main inter-tidal classes (Figure 9). Some of the terrestrial classes (Table 9) do have none-zero values. This is probably due to terrestrial land cover polygons intersecting with the low tide mark line on narrow stretches of coastline and on narrow, inland stretches of tidal rivers, but this has not been explored in detail.

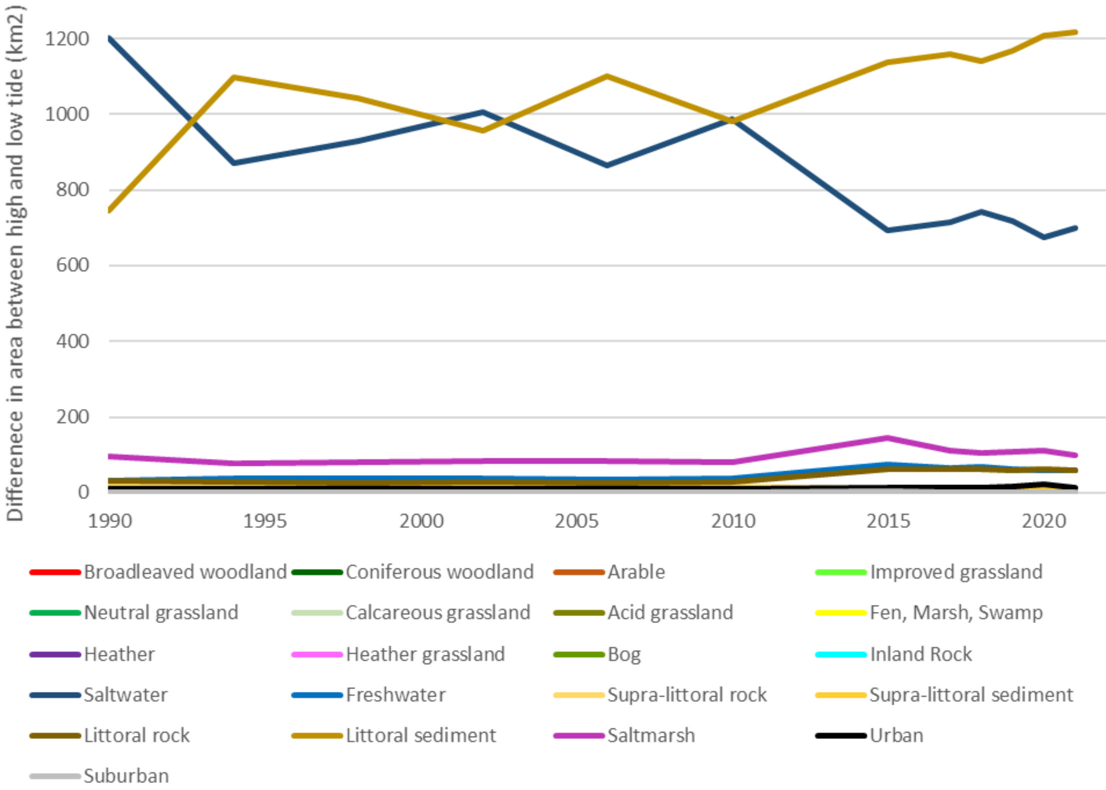


Figure 9: The difference in area (km^2) between the extent of different land cover types when clipped to the mean high or low water mark.

Table 9: The difference in area (km²) between the extent of different land cover types when clipped to mean high and low water mark. Ranked from largest to smallest using 2021 data.

	1990	1994	1998	2002	2006	2010	2015	2017	2018	2019	2020	2021
Littoral sediment	744	1098	1043	955	1102	983	1138	1160	1140	1167	1209	1216
Saltwater	1201	871	929	1006	866	986	692	714	741	717	673	699
Saltmarsh	95	79	80	84	84	82	144	111	106	109	112	97
Littoral rock	31	29	25	29	26	28	62	63	63	58	61	58
Freshwater	33	37	37	39	36	37	74	67	70	62	59	58
Urban	9	9	9	9	9	9	14	14	13	16	21	12
Fen, Marsh, Swamp	6	7	7	7	8	7	6	6	5	5	5	6
Improved grassland	5	4	5	5	4	5	6	4	4	4	3	6
Supra-littoral rock	24	16	17	17	17	17	11	6	6	7	4	4
Arable	6	10	10	8	8	7	6	7	3	4	3	4
Supra-littoral sediment	8	4	4	5	4	4	11	9	11	12	14	4
Broadleaved woodland	4	2	2	3	3	3	4	3	3	3	3	2
Suburban	2	2	2	2	2	2	3	2	2	2	2	2
Neutral grassland	1	1	1	1	1	1	1	1	0	0	1	1
Calcareous grassland	1	1	1	1	1	1	0	0	0	0	0	0
Bog	0	0	0	0	0	0	0	0	0	0	0	0
Coniferous woodland	0	0	0	0	0	0	0	0	0	0	0	0
Inland Rock	0	0	0	0	0	0	0	0	0	0	0	0
Heather	0	0	0	0	0	0	0	0	0	0	0	0
Heather grassland	0	0	0	0	0	0	0	0	0	0	0	0
Acid grassland	0	0	0	0	0	0	0	0	0	0	0	0
Total	2172	2172	2172	2172	2172	2172	2171	2165	2165	2165	2172	2172

2.5 Caveats / uncertainties in the application of the outputs in understanding changes in land cover.

- Accurately detecting Land Cover Change is difficult in the UK, because of the relatively small area of many UK land cover parcels and the relatively small scale of much of the change that occurs.
- The LCM time-series from 2017 onwards is much more consistent than pre-2017 in its methodological approach, however, there is still an ongoing evolution of methods which affects the ability to detect change. This is partly because land cover mapping is undergoing a paradigm shift from one-off production of land cover data sets to something more akin to model development, where the model is improved over time and the model output is reprocessed and re-released at appropriate intervals; this version-based approach has been common in some global coarse-scale land cover products, but not in national-scale land cover mapping.
- In terms of the OEP's assessment, some of the grassland classes of most concern, and interest, are difficult to differentiate in the field and from EO data, especially for areas of grassland at the boundaries of different grassland types. This is noticeable in the figures above, as *Semi-natural grassland* particularly shows abrupt changes, which are due to changes in production methodology, rather than a real change in land cover. Currently, this introduces a high level of variability in some of the classes most likely to contain wildlife rich classes, specifically Neutral grassland. Consequently, whilst the current LCM data are useful in providing an overview of the land cover trends and the drivers and pressures on land cover and land cover change in England, like increasing urbanisation and afforestation, they are not currently sensitive enough to detect the relatively small levels of change in individual semi-natural classes (this is discussed more in section 3.3).
- Validation is important to understand the uncertainties in land cover data sets, however updating and refreshing validation data is difficult as there are limited sources of these data sets, especially for some of the key wildlife rich habitats. Validation data for land cover change is also particularly sparse, as is validation data for the earlier land cover maps due to less open-access digital data being available, particularly for the 1990's.
- Comparing data from different sources is complex, especially when there are differences in timing, differences in thematic definitions of classes, or abilities of products to map different classes. The issues around comparing Land Cover Map data to Countryside Survey data are discussed in detail in chapter 4 of Morton *et al.,* (2011) and also in Fuller *et al.,* 2005, but the same issues also affect the comparison of different land cover products. For example, a woodland class from the National Forest Inventory differs from a woodland class in the UKCEH LCM data. This is because the requirements

for the NFI data are different from the requirements and methods that underpin the UKCEH Land Cover Map. The NFI data is designed to inform operational forestry decisions, whereas the LCM data are designed for wall-to-wall mapping, so everywhere in the UK is assigned a class. These requirements affect the data flowing into the two mapping programmes: NFI is informed by analysis of aerial photography and also administrative records about new planting[3], whereas LCM is based on image classification of satellite data. In practice, this means that the NFI is likely to know about a forest before it is planted, or soon after planting, whereas LCM will only identify a woodland as woodland, when enough trees have reached a level of maturity that means that woodland is the predominant spectral signal for the given location in the satellite data. This depends on woodland type, planting density and the proficiency of woodland management and may take many years from initial planting.

- Timing of data sets – LCM uses the year that the satellite data were acquired in as part of the name, so LCM2021, uses satellite data from 2021 and was published in 2022. This means that there is always a lag between the satellite data being acquired, the land cover data set being produced and published. Other data sets, or portals displaying summary statistics, may take different approaches. For example, the Defra D1 indicator has statistics for Living England, which are described as Land Cover Statistics, 2022. The 2022 in this case appears to refer to the year the data set was published, rather than the underlying satellite data, which range from April 2019-July 2021. This underlines the fact that it is important to understand the specification of a data set before use. This becomes especially important if data sets are used to produce summary statistics, which may be used widely by users not familiar with the specification and constraints of the underlying data sets.

2.6 Future developments

Land cover mapping in the UK, particularly for the UKCEH land cover map series, is undergoing a transition from irregular one-off land cover maps (i.e. mapping) to annual production of land cover maps (and possibly even more frequent) (i.e. monitoring). However, additional research is required to determine how to create the best estimates of land cover change from the land cover map series, given the uncertainty in the land cover mapping. There are methods of reducing the uncertainty by post-classification filtering of the time-series (Rowland *et al.,* 2021), there are methods of incorporating error bars around the annual estimates (Fuller *et al.,* 2005) and also methods, used for other environmental monitoring, of filtering annual time-series to infer trends (and to reduce the inter-annual variability caused

[3] https://www.forestresearch.gov.uk/tools-and-resources/national-forest-inventory/about-the-nfi/

by uncertainty in the data). Some, or all, of these methods could be incorporated as the UK moves to a more mature land cover mapping capability. This all points to an increasing role for statistical analysis as an important part of operational land cover change monitoring.

Increasing the volume and quality of land cover change data is only valuable, if the data can be used effectively. Development of data visualization methods to support understanding of land cover change data and trends in land cover change is also important. The Sankey plot (Figure 4) is a useful tool for displaying flows of land cover between years, but is not easy to calculate, so it is also important for future development of land cover data sets to include data portals that enable analysis and extraction of key statistics for users without specialist skills.

3. WPB: Application of land cover data

3.1 Land cover data to monitor wildlife richness

England's long-term biodiversity target is for the creation or restoration of 500,000 hectares (5000km^2) of wildlife-rich habitat outside of the protected site network by 2042[4]. This includes an interim target of restoring, or creating, 140,000 ha of a range of wildlife-rich habitats outside of protected sites by 2028, compared to 2022 levels. At least 80% of the 500,000 hectares wildlife-rich target is expected to come from paying farmers and land-owners to improve the countryside, including through peatland restoration and biodiverse woodland. There is also an intention to increase woodland in England by planting 180,000 hectares (1800km^2) of woodland, also by 2042.

In response to these targets, the OEP posed the following questions:

- How can land-use and landcover data be used to assess progress against the Government's "wildlife-rich habitat restoration or creation" target? What are the limitations of applying this data?
- To what extent can wildlife rich habitat be inferred from current LCM data outputs?
- To what extent can wildlife rich habitat be inferred from current Defra's land-use statistics and other products being developed through NCEA such as the Living England Map and priority habitat inventories?

Definition of wildlife rich habitats

In the biodiversity target regulation, wildlife rich is defined as: "*a habitat that is one of the following types of habitat and which is of sufficient quality that it is, or will be, capable of supporting flora and fauna which are typically found in the habitat in question—*
(a) a habitat type of principal importance for the conservation of biodiversity listed by the Secretary of State under section 41 of the Natural Environment and Rural Communities Act 2006(7) (biodiversity lists and action (England));
(b) another habitat type listed in Schedule 1".[5]

For England there are 46 habitats of principal importance (previously known as Priority Habitats).[6] The 46 habitats of principal importance for England are more

---

[4] Environmental Improvement Plan (publishing.service.gov.uk)

[5] The Environmental Targets (Biodiversity) (England) Regulations 2023 (legislation.gov.uk)

[6] Habitats and species of principal importance in England - GOV.UK (www.gov.uk)

finely resolved than the Broad Habitat-based categories that LCM and Living England use. For further consideration of the key ecological issues, see Annex 1 of the Defra Biodiversity and Freshwater Targets: Detailed Evidence report[7] , which considers a range of additional factors that are important to the creation of wildlife rich habitat.

3.2 How can land-use and land cover data be used to assess progress against the Government's "wildlife-rich habitat restoration or creation" target?

In terms of methods, there are four *potential roles* for land cover data in terms of monitoring progress towards the Government's "wildlife-rich habitat restoration or creation" target:

(1) To directly monitor the restoration and creation of new wildlife rich habitats, with *land cover data as the primary source of data.*

(2) To directly monitor the restoration and creation of new wildlife rich habitats *to corroborate other sources of data.*

(3) To monitor the wider countryside to assess whether the gains from targeted creation and restoration are undermined by losses elsewhere.

(4) To provide an interim data source, whilst other monitoring mechanisms are developed.

Below we discuss these roles with respect to LCM data, before exploring the trends that LCM shows for wildlife rich habitats. Following that we review the Living England and Priority Habitat Inventory data sets, before considering some more general issues that could impact on the ability to monitor progress towards the wildlife rich habitat biodiversity target.

3.3 To what extent can wildlife rich habitat be inferred from current LCM data?

Here we briefly assess how LCM would perform against the four potential roles for land cover data outlined above. Additional discussion on the constraints of using land cover, particularly with respect to assessing wildlife rich habitat restoration, are also included in the section: What are the limitations of applying land cover data?

[7] Biodiversity terrestrial and freshwater targets Detailed evidence report.pdf (defra.gov.uk)

Role 1 - **Directly monitoring the restoration and creation of new wildlife rich habitats, with *land cover data as the primary source of data.***

The Land Cover Map cannot directly map wildlife richness, and LCM classes are not thematically resolved enough to directly map newly created wildlife rich areas as habitats of principal importance. Consequently, LCM data is not suitable as the sole or primary source of data for monitoring the creation of all wildlife rich habitats. The capability of LCM to map habitat restoration specifically is addressed in the section on: What are the limitations of applying land cover data?

Role 2 - **Directly monitoring the restoration and creation of new wildlife rich habitats *to corroborate other sources of data*.**

The Land Cover Map data should be able to detect new areas of wildlife rich habitat once they become sufficiently established, although it will detect new habitat areas at the Broad Habitat level, rather than at the level of the more finely resolved habitats of principal importance. So, LCM data should be able to corroborate other sources of data, such as the priority habitat inventory.

Role 3 - To monitor the wider countryside to assess whether the gains from targeted creation and restoration are undermined by losses elsewhere.

LCM will be able to monitor the wider countryside to identify areas of potentially wildlife rich habitat that are lost to other land cover types.

Role 4 - **To provide an interim data source, whilst other monitoring mechanisms are developed.**

LCM data have an immediate role in the provision of land cover and land cover change data, particularly whilst Living England is in development. The LCM series can provide information on land cover extent, the spatial configuration of land cover, and land cover change. For example, Defra used LCM data to understand the extent of semi-natural land in England (in the Biodiversity targets detailed evidence report[8]) and to capture the spatial configuration of land cover, during the developmental phase of the connectivity component of the D1 indicator[9] (see also section 3.5 Recommendation on which data to use).

---

[8] Biodiversity terrestrial and freshwater targets Detailed evidence report.pdf (defra.gov.uk)

[9] D1: Quantity, quality and connectivity of habitats (defra.gov.uk)

Inferring wildlife rich habitats from current LCM data

Figure 10 shows the trends in wildlife rich habitat from LCM data. For Figure 10 we define wildlife rich habitat as: Broadleaf woodland, Neutral, Calcareous and Acid grassland, Bog, Fen, marsh and swamp, Heather, Heather grassland, Bog, Inland Rock, Supra-littoral rock, Supra-littoral sediment, Littoral rock, Littoral sediment and Saltmarsh. These classes are included on the basis that they contain the main, large-scale habitats of principal importance, and as such are the classes most likely to include large-scale nature friendly habitats.

The classes excluded from the wildlife-rich habitat category here, are Arable, Coniferous woodland, Improved grassland, Freshwater, Saltwater, Urban and Suburban classes. In England, these classes may include small areas of wildlife-rich habitat, such as arable field margins, but not typically the extensive, large-scale areas of wildlife-rich habitat.

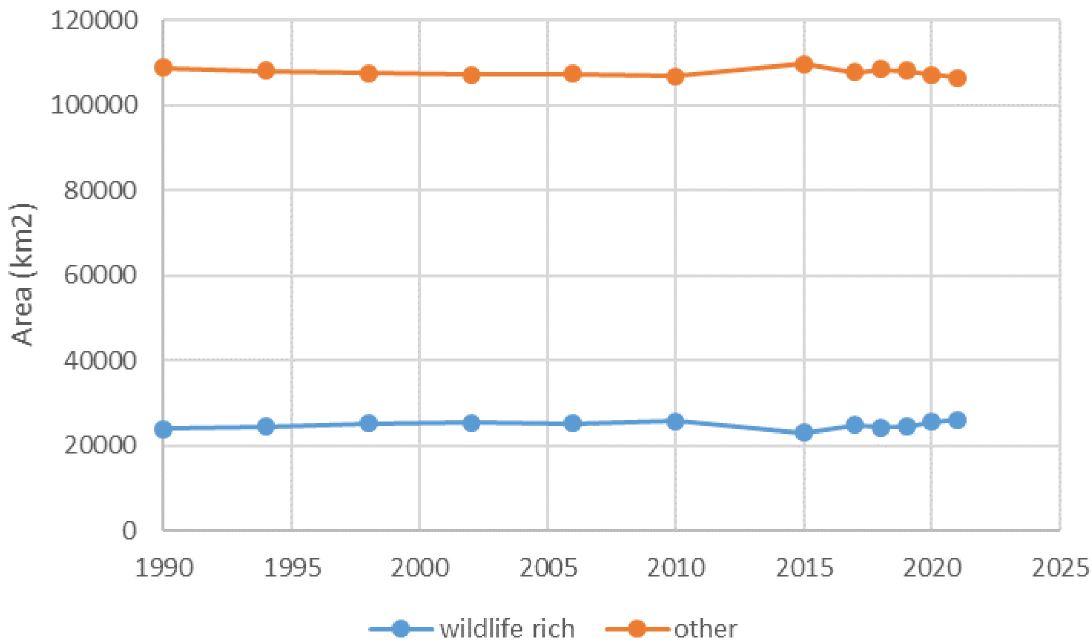


Figure 10: Extent of wildlife rich habitat in England based on the LCM series. Here, wildlife rich is defined as all habitats, except conifer, arable, improved grassland, water or urban classes. Data clipped to the low-tide mark. Note – here water is included in the 'other' class, in Table 10, water is assessed separately.

Table 10 shows the LCM2021 confusion matrix aggregated into three categories: wildlife rich, other and water. When aggregated into these classes, the wildlife rich class is slightly less accurate at ~89% than the other classes, which are generally closer to 92-95%. This is primarily due to some confusion between the 'wildlife rich' category and 'other' category. Improving the accuracy of the classes aggregated into the 'wildlife rich' class would improve the ability of the LCM series to assess the change in 'wildlife rich' habitat.

Future work will resolve this issue by re-processing the whole LCM time-series to further improve the separation of the semi-natural classes, particularly the grassland classes.

Table 10: LCM2021 confusion matrix for 'wildlife rich' habitats versus other land cover, and water, based on 35,182 points. Wildlife rich is defined here as all habitats, except conifer, arable, improved grassland, water or urban classes. All values are points, except percentage accuracies.

		Reference data				
		Wildlife rich	Other	Water	Total points	User's accuracy (%)
Classification data	**Wildlife rich**	***9321***	1052	36	10409	**89.5**
	Other	1121	***22974***	20	24115	**95.3**
	Water	30	7	***621***	658	**94.4**
	Total points	10472	24033	677	**Overall accuracy: 93.6**	
	Producer's accuracy (%)	**89.0**	**95.6**	**91.7**		

Consistency of data over OEP's assessment period, 2015-2021

The LCM data assessed here are the best time-series of data currently available for England. However, as Table 1 shows the underlying methods are still evolving, which leads to some issues with consistency across the time-series. The main issue, in the time period 2015 - 2021, is that Neutral and Calcareous grassland are under-estimated in LCM2015, with an associated over-estimation of improved grassland. This under-estimation reduces through LCMs 2017-2019. Overall, this leads to an over-estimate of the flow (or apparent creation) of landcover classes between 2015 and 2021, as the underestimation is corrected. Note this will affect Figure 4 by overestimating the flow of improved grassland to semi-natural grassland.

To corroborate the increase in wildlife rich habitat, between 2015-2021, we can assess change over a longer time-period. Between LCM2010 and LCM2021 there is an increase in wildlife rich habitat of 38,134ha (381km^2) or 3467ha/year.

This can be compared to Defra's annual average of 16,595 ha/year for both *habitat creation and restoration* between 2011 and 2019 (Table 11). Focussing on the habitats that LCM can detect gives 8,598 ha annually of *habitat restoration and creation*. The LCM estimate is about 40% of the Defra figure, which (ignoring time-lag between habitat creation and habitats being detectable) would be consistent with a mix of approximately 40% habitat creation and 60% habitat restoration.

Both the Defra and LCM figures support a recent increase in wildlife rich habitat.

Table 11: Habitat creation and restoration rates between 2011 – 2019 for all mechanisms (based on table in Annex 2 of Defra's detailed evidence report [10]) and assessment of land cover map type data to detect. **Highlights shows habitats LCM-type data: should detect (green), uncertainty about detection (yellow) and undetectable (orange).**

Habitats	Annual restoration / creation (ha)	Will LCM-type data detect creation of new habitat
Grassland	3591	Should detect creation
Lowland heath	505	Should detect creation
Uplands	1182	Should detect creation
Coastal floodplain grazing marsh	859	Should detect creation
Wetland	422	Should detect creation
Coastal	218	Maybe, if not inter-tidal
Wood-pasture & parkland	261	No, no suitable class
Traditional Orchards	18	No, no suitable class
Woodland creation	2039	Should detect creation
Plantations on Ancient Woodland Sites (PAWS)	2717	Maybe, depends on species
Arable field margins	3674	No, too narrow
Mix of Priority Habitats (undefined)	984	Unclear
Ponds	70	No, too small
Hedges	55	No, too narrow
Habitats not detectable by LCM	**4078**	
Habitats where there is uncertainty about LCM's ability to detect	**3919**	
Maximum area detectable by LCM type data (assuming all habitat creation, in practice it will include habitat restoration)	**8598**	
Annual average extent of restored and created habitats between 2011 and 2019	**16595**	

10 https://consult.defra.gov.uk/natural-environment-policy/consultation-on-environmental-targets/supporting_documents/Biodiversity%20terrestrial%20and%20freshwater%20targets%20%20Detailed%20evidence%20report.pdf

3.4 Defra data sets

To what extent can wildlife rich habitat can be inferred from current Defra's land-use statistics and other products being developed through NCEA such as the Living England Map and the priority habitat inventories?

The Living England data set, the Priority Habitat data set, and the Nature Recovery Network, all potentially contribute to monitoring wildlife rich habitats and are discussed here.

Living England map – Calculating wildlife rich land from the Living England map produces a figure of 60,536 km^2. This is significantly more than the LCM figure of 26,173 km^2 and the Countryside Survey 2007 figure of 36,013 km^2. The Countryside Survey figure rises to 39,547km^2 when including boundary and linear features, which are narrow features (such as hedges and ditches) below the resolution of the current satellite-based land cover data sets.

The high wildlife rich figure from Living England appears to be mainly due to the known over-estimates in the current version of the Living England habitat map, specifically the over-estimation of semi-natural grassland and Fen, Marsh & Swamp, and consequent under-estimation of Improved grassland.

The difference between the Countryside Survey and Land Cover Map estimates for wildlife rich habitat appears to be largely down to differences in the estimates of Neutral Grassland, with Countryside Survey 2007 having an estimate of 14,426km^2 compared to 1,446km^2 in LCM.

Living England have published separate confusion matrices for each of the Biogeographical Zones used in their classification, so it is possible to create an aggregated confusion matrix for wildlife rich versus other classes, as done for LCM (Table 10). However, a similar confusion matrix is not presented here, as the Living England validation does not reflect the known issues in the current version of the Living England data set.

The Living England habitat map is currently under development, with a new release planned for 2024, at which point it will be timely to revisit the question of: to what extent can wildlife rich habitat be inferred from Living England data?

Priority Habitat Inventories – The Priority Habitats Inventory is a spatial data set showing the extent of 25 priority habitats in England (NE, 2022). It collates a range of field-survey data collected by Natural England and their partners. A new significantly upgraded version has recently been released and further developments are planned (NE, 2022). Because it relies on targeted field-surveys it is unclear what its role would be in monitoring progress towards the Biodiversity

Target. However, it may have a role in tracking larger areas of habitat creation and restoration, particularly those that Natural England and their partners are involved in. It may also have a role in providing training and validation data for satellite-based data sets.

Nature Recovery Network - The Environment Act 2021 has also established county-based local nature recovery strategies (LNRS), which will cover the whole of England and are intended to enable mapping of Nature Recovery Networks locally and nationally[11].

We have reviewed the developments under the NCEA requested by the OEP and that we are aware of, however, there may be other developments that would contribute to the monitoring of wildlife rich habitats of which we are not aware.

3.5 Recommendation on which data to use

To assess changes in land cover and to understand habitat contributions towards the wildlife rich habitat target, we think that the UKCEH Land Cover Map series is currently the best source of data for England, because:

- it provides a time-series of habitat data. Whereas Living England is currently only available as a single, interim product, until the revised product becomes available in 2024.
- the interim Living England data set has major issues, with some classes, such as the large overestimation of ‘Fen, Marsh and Swamp’ in the Lake District.

When Living England maps are available for two different time points, enabling the calculation of land cover change, then it will be appropriate to compare the performance of both Living England and LCM for mapping land cover and change in land cover. The results of the validation will then provide the evidence-base for deciding which is most appropriate to use. The comparison of the two data sets may need repeating at intervals, as the products develop (see also section 3.2(4)).

3.6 What are the limitations of applying land cover data?

Satellite-based land cover data sets can potentially be used to map newly created habitats if the land cover/habitat classes are appropriate. However, there are caveats principally that:

- The accuracy with which classes are mapped varies in EO-based data sets. Some classes are mapped more accurately than others, making it possible to accurately map smaller levels of change in some classes than others.

[11] Nature Recovery Network - GOV.UK (www.gov.uk)

- Different methodologies are likely to be required to monitor restoration versus creation from satellite data, although this maybe complicated by differences in terminology between ecological and technical perspectives.

These issues, and others, are explored below.

Accuracy – the key issue affecting whether a land cover data set can be used to map land cover change is the accuracy of the data set and the accuracy of key classes. Aggregating classes may increase the accuracy with which change can be mapped, if the aggregation reduces uncertainty. The accuracy of individual land cover data sets is typically summarised with a confusion matrix (e.g. Table 5). It is also important to acknowledge that satellite-derived data sets tend to be accompanied by relatively comprehensive assessments of their accuracy, so that their uncertainties are relatively well known. Such accuracy assessments are rarely conducted for data sets from other sources. This can potentially make satellite-derived land cover data appear less accurate than other data sets with unquantified uncertainties.

Spatial scale - the accuracy of land cover data sets is affected by the size of the patches of habitat being classified, with larger areas typically easier to classify. England is very heterogeneous, and patches of semi-natural habitat often occur at small scales. Ecologically, the size of successful wildlife sites depends on the habitats in question, their quality, and their connectivity (Crick *et al.,* 2020). In their handbook for the Nature Recovery Network, Crick *et al.,* (2020) suggest larger sites may be more beneficial, but these could still be complicated for existing land cover type data sets to map, if the sites comprise of a mosaic of small habitat patches, below the ability of the EO data sets to accurately detect. Crick *et al.,* (2020) also suggest that it would be beneficial for some smaller sites to expand; this expansion could be difficult to detect in land cover mapping data sets, especially those using fixed spatial units, if land cover methods do not adapt.

Time-scales – timescales are important from both the ecological and technological perspectives**.** The time taken to achieve a wildlife rich habitat will vary depending on the habitat, the spatial context and the approach taken to habitat creation or restoration (Maskell *et al.,*2014; POST, 2022). Similar factors will also influence how rapidly EO-based methods are able to detect and correctly identify the new habitat. For example, as noted earlier, newly planted woodland needs to reach a sufficient coverage to be identified as woodland from land cover mapping methods. It is important to note that the land cover mapping methods are sensitive to the spectral distinctiveness of the habitat and detecting a specific habitat would not mean that the new habitat was wildlife rich. Potentially the land cover mapping could detect a newly created habitat many years before it became wildlife rich. The wildlife richness (and associated time lags) of newly created habitats is an ecological issue and beyond the scope of this report.

Novel habitats - In terms of habitat creation some cases maybe relatively easy to detect from EO, such as the conversion of quarries to heathland, or the removal of coniferous woodland from peatland. It is not clear whether the habitat creation will include novel habitats, but these are likely to be difficult to identify correctly in land cover mapping via image classification, as such methods are trained on labelled examples of each class. More generally, rare classes can also be more difficult to classify, because of the lack of suitable training examples. This can be reduced, but not entirely resolved, by expanding classification tiles, so that classifications are run over a larger area, so rare habitats are more likely to be included. Alternatively, the training areas of rarer class can be over-sample training areas, so that the training sample is balanced across different classes (all classes have the same number of training pixels regardless of their rarity). For LCM a mix of over-sampling rarer classes and expanding classification tiles is used.

Role of remote sensing in monitoring restoration versus creation – the biodiversity target covers both habitat restoration (change within a habitat) and creation (change between habitats). Land cover mapping deals in categories/classes, so change is only registered when a parcel of land changes from one land cover class to another. Consequently, wildlife rich habitat creation may be detectable over time using land cover data sets, but restoration (i.e. within habitat change) will not generally be. However, not all remote sensing focusses on labelling parcels of land with a categorical label, so other forms of remote sensing may be more appropriate for tracking the trajectory of parcels of land over time as they are restored, and habitat condition improves (McKenna *et al.,* 2022).

However, the split between habitat restoration and creation will not always determine the methodology. As such, care may be needed around terminology, for example, if peatland restoration includes the removal of coniferous woodland on peatland, this would generally be detected as a land cover change, rather than a change in the quality of the existing habitat. This suggests some consideration of the specific habitat restoration and creation trajectories maybe required to ensure that monitoring methods detect the relevant changes.

Appropriateness of data for monitoring – In the early years of land cover mapping land cover data sets were typically produced as one-off products. For monitoring, data sets should be produced regularly, to a known schedule, with documentation that explains the method used to create the data set, the validation data and results, and the land cover or habitat classes.

Consistency over time is important and needs to be balanced with the general trend for EO-based methods to improve over time. In some applications it is acceptable to reprocess data sets across the full time-series, however, this is more complicated if the first time point is being used as a baseline for regulatory purposes.

This will require thought in the future to balance the benefits of increasing spatial resolution and potentially thematic detail (i.e. the number of land cover/habitat classes) that can be mapped, against consistency with the baseline data. The LCM results presented here have used the 'legacy' style 30m land cover data, which is consistent since 1990. However, this reduces the potential benefits of newer, higher resolution data.

Spatial units – remote sensing data sets can use a variety of spatial units, including pixels, polygons capturing real-world objects (e.g. fields), and segments (areas of an image with similar spectral responses). The choice of spatial unit becomes important when identifying change over time, as if the spatial units poorly represent the shape and extent of the change, then the accuracy and speed of detection are likely to be affected.

LCM currently uses both pixels, which limit assumptions about spatial structure, and generalised polygons, which aim to capture real-world objects. The Living England phase 4 map currently uses image segments. Image segments are clusters of pixels that have similar spectral properties at one point in time. They may be useful in some circumstances, but good segments are difficult to create.

Image segmentation algorithms perform best in highly managed, simplified landscapes, where there are sharp boundaries between land cover/habitat types. In the UK this will typically occur in lowland agricultural landscapes. In very busy landscapes, such as urban and suburban areas, or in landscapes where habitat transitions are gradual, such as semi-natural and upland areas, segmentation algorithms often struggle (Smith and Morton, 2010). In these locations segments are usually impure i.e. they contain more than one land cover class.

Segments are a function of image, time and space; segments from images at the same location but at a different time will often give significantly different segmentation results. The segment framework should therefore be recomputed each time new images are used, so that the segmentation and subsequent classification are coincident in time and space. However, a complication of this is that the segments will change with each classification phase. Image segments can vary substantially between different images and are not well-suited for change detection, so this may be a challenge for the Living England product, depending on how it develops.

Image segments were also used in LCM2000, but were dropped when LCM2007 was produced, in favour of generalised cartographic objects, which represent real-world objects. The LCM spatial framework was created for LCM2007, so some areas will now be outdated i.e. areas where urban areas have expanded into what were arable fields in 2007. However, this issue does not affect the LCM per-pixel products (produced since LCM2017), as the per-pixel products make no assumptions about landscape structure.

Inter-tidal habitats – country-wide land cover mapping programmes will typically capture inter-tidal areas in a mix of tidal states (as shown by the churn between saltwater and littoral sediment in Figure 9) unless specifically designed to map inter-tidal areas well.

Thematic and spatial developments – the regulatory requirement is to monitor change to 2042. EO is currently developing rapidly, in part due to cloud-computing and advances in machine learning, so there are likely to be advances in what can be measured thematically (i.e. in terms of land cover data sets). This means that it will be useful to re-produce the entire series of LCM data (and Living England) as methods continue to develop, which will reduce inconsistencies across the time-series and improve the capacity to detect change (see also section 2.5). This may be complemented with increases in spatial resolution, as targeted work with 3m satellite data now possible. 3m resolution data provides the opportunity to potentially monitor narrower, linear features that maybe important in increasing wildlife-rich habitats. However, any such methods would require sufficient data to back-date them to the point at which the regulations come into force.

Availability of appropriate field-data – satellite-based land cover mapping benefits from field observations for training and validation. Specifically, it requires spatially and thematically accurate field data for areas of sufficient size, with a single land cover type and using the classes used in the land cover data.

3.7 OEP's use of land cover change data

The OEP presented an analysis of existing Government data on land use change in England in their 2023 report (OEP, 2023; pages 34-36). Assessing the OEP's analysis raises two questions:

1. Was the analysis appropriate?
2. Was the data appropriate?

The data available to OEP were the official land-use statistics published by the government and supplied by the Ordnance Survey[12,13]. The OEP correctly identified changes in the way the statistics were reported, specifically the use of the mean high tide mark in 2022, but not in 2018, that limited their analysis of trends in the data. It was also not possible to recalculate the summary statistics from the original OS data. Therefore, the OEP analysed the data appropriately and included appropriate caveats.

The land-use data published annually by the government focusses on land-use and are intended to track planning and flooding issues, but include classes, such as

[12] https://www.gov.uk/government/statistics/land-use-in-england-2018

[13] https://www.gov.uk/government/statistics/land-use-in-england-2022

'Natural land', 'Forestry and woodland' and 'Rough grassland'; in some cases these classes will cover wildlife rich areas. For monitoring the biodiversity targets, the Land-Use change data sets published by the government [14] may have some role as they show information about the original land-use and the new land-use it's converted to. However, the land-use data are not designed to track habitats and it is not clear how the definitions of habitats relate to habitats used for biodiversity purposes and hence wildlife richness. The land-use data potentially enables newly created habitats to be tracked, although the classes are not well-aligned to the biodiversity targets. Other issues which would require consideration, if these data were to continue being used, are:

- The timing of the start and end dates is not currently ideal for monitoring the biodiversity target, although it would presumably be possible to process over different time periods.
- The update cycle of the underlying OS data sets will affect the currency of data, particularly in rural areas.
- The current format of data would not allow the separation of newly created habitat outside of protected areas, compared to that within; the biodiversity target of interest is the creation, or restoration, of habitat outside of protected areas.
- These data would not provide any information on restored habitat areas.
- Better understanding of the definition of the different land-use classes and how they relate to habitats and specifically wildlife-rich habitats.

In summary, the land-use data are not designed to track changes in habitat extent, so they are not ideal as a primary source of data, but they may be useful for corroborating trends identified in other data sets. The LCM data analysed in this report are more appropriate for monitoring wildlife rich habitat, but will be better when re-processed for additional consistency, especially if improved separation of grassland types is developed and implemented.

[14] https://www.gov.uk/government/statistics/land-use-change-hectarage-2019-to-2022/land-use-change-statistics-hectarage-2019-20-to-2021-22-statistical-release

4. WPC – The role of land cover data in the land use framework

Government targets for net-zero and biodiversity, alongside the need to provide food and homes, are placing increasing pressure on land-use. To enable evidence-based decisions, the Land Use Framework has been proposed. A key element of the Land Use Framework is the requirement for accessible, spatially explicit evidence [15] and a shared evidence-base to support decision-making [16].

Land cover data has a key role in informing a Land Use Framework. It can help the development and delivery of better-informed land use decisions in several ways including:

- **Providing an inventory of existing land cover**. This could be providing statistics about land cover, or recent land cover change, at county-level, or catchment-level.
- **Comparison, validation and corroboration of land cover and land-use data.** LCM data can be compared against other data sets to provide corroboration or validation.
- **Providing maps and information on the spatial distribution of land cover**. Understanding the spatial distribution of different land cover types is important for a wide range of environmental issues. For example, trees provide a wide range of services beyond carbon sequestration, but the role of trees in flood mitigation, recreation and other ecosystem services is strongly dependent on location (Burke *et al.,* 2023), so spatially explicit data and modelling is required to inform decisions. Maps are also important visual aids for enhancing engagement with stakeholders when consulting on land-use decisions.
- **Providing data on the spatial configuration of land cover**. The spatial configuration of habitats is important for assessing habitat connectivity, which is crucial for understanding how species can move across the landscape. Habitat connectivity is recognised as a key component of habitat quality in Defra's D1 indicator. LCM data has often been used for mapping connectivity (Mancini *et al.,* 2022) and for informing decisions about where to create new habitat to improve habitat connectivity and thus biodiversity, such as the work by Highways England to improve wildlife corridors across their landholdings[17].
- **Development of tools to aid uptake of data** - LCM has an important role in providing information, in conjunction with other data sets and domain knowledge, to help inform decisions about land-management and land-use. However, users need data packaged in appropriate ways for them to use. This requires the development, and ongoing maintenance, of appropriate tools. For example, LCM data is used by

---

[15] https://ffcc.co.uk/land-use-framework

[16] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1163819/2023-05-23_FINAL_NLDP_report_compressed.pdf

[17] https://www.ceh.ac.uk/sites/default/files/2021-06/UKCEH-Case-study-Informing-decision-making-by-mapping-land-cover.pdf

the e-planner tool, which is aimed at land-managers and farmers and provides them with information and evidence to inform their land-use decisions[18].

- **Underpinning environmental modelling.** Many of the users of Land Cover Map data use land cover data alongside other environmental data sets to understand the interactions between land cover data and specific environmental phenomenon, such as disease spread, biodiversity change and water quality. Land cover data is therefore a key data set required to underpin the modelling required to support evidence-based decisions, and to explore the impact of different scenarios of future land-use in the Land Use Framework. Modelling work typically involves data sets from other monitoring schemes, such as the inclusion of UK National Biodiversity Network species records to explore biodiversity (Chetcuti *et al.,* 2019), or the inclusion of river monitoring observations to understand the impact of land cover on water quality (Ritson *et al.,* 2019). This can also include the use of LCM data to explore scenarios of future land-use or species distribution under climate change.
- **Underpinning policy and regulation.** Land Cover data is used for policy and regulation. For example, Natural England use the Land Cover Map data, alongside newt surveys, to create Great Crested Newt risk zones[19], which are used to assess and manage the impact of new development on newt populations, helping developers meet their legal requirement to protect the Great Crested Newt. LCM data is also used to calculate the UK's natural-capital accounts, which are published by the Office for National Statistics[20].
- **Trans-boundary data** – Environmental issues cross-boundaries, so LCM data provides an important source of consistent UK-wide data, which is critical for many environmental projects and for providing a UK-wide shared evidence base.
- **Understanding the impact of land cover change** – the LCM time-series is a valuable resource for exploring the impact of land cover change on environmental phenomenon, like carbon stocks (Fryer & Williams, 2021), bird species (Broughton *et al.,* 2020) and flood plain condition (Entwistle *et al.,* 2019).

The UKCEH land cover map data sets have been widely used for over 30 years, so there is a wide body of research and knowledge showing how they (or other land cover data) can be used to understand a wide-range of environmental issues, across a broad range of disciplines, including health, ecology, hydrology, air pollution and net-zero.

In terms of limitations, the existing land cover data sets are better suited to application at broad scales. More spatially resolved data sets, such as the Environment Agencies airborne Lidar data [21], may also be required for local-scale decisions and monitoring, for example monitoring of individual trees.

---

[18] https://e-planner.ceh.ac.uk/About

[19] https://www.gov.uk/government/publications/great-crested-newts-district-level-licensing-schemes-for-developers/developers-how-to-join-the-great-crested-newt-district-level-licensing-scheme#:~:text=a%20mitigation%20licence.-,How%20district%20level%20licensing%20works,applying%20for%20a%20mitigation%20licence.

[20] https://www.ons.gov.uk/economy/environmentalaccounts/bulletins/seminaturalhabitatnaturalcapitalaccountsuk/2021

[21] https://www.data.gov.uk/dataset/f0db0249-f17b-4036-9e65-309148c97ce4/national-lidar-programme

5. Acknowledgements

The high and low water mark data were sourced from the Office for National Statistics and licensed under the Open Government Licence v.3.0, they contain OS data © Crown copyright and database right [2023].
The Sankey plot was produced with the Sankyematic application at: https://sankeymatic.com/build/

6. References

Broughton, R. K., Parry, W., & Maziarz, M. (2020). Wilding of a post-industrial site provides a habitat refuge for an endangered woodland songbird, the British Willow Tit *Poecile montanus kleinschmidti*. Bird Study, 67(3), 269-278.

Bunce, R.G.H., Barr, C.J., Clarke, R.T., Howard, D.C. and Lane, A.M., 1996. Land classification for strategic ecological survey. Journal of environmental management, 47(1), pp.37-60.

Burke, T., Whyatt, J. D., Rowland, C., Blackburn, G. A., & Abbatt, J. (2020). The influence of land cover data on farm-scale valuations of natural capital. *Ecosystem Services*, *42*, 101065.

Carey, P.D., Wallis, S., Chamberlain, P.M., Cooper, A., Emmett, B.A., Maskell, L.C., McCann, T., Murphy, J., Norton, L.R., Reynolds, B., Scott, W.A., Simpson, I.C., Smart, S.M. and Ullyett, J.M., (2008) Countryside Survey: UK Results from 2007. NERC/Centre for Ecology & Hydrology, 105pp. (CEH Project Number: C03259).

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.

Chetcuti, J., Kunin, W. E., & Bullock, J. M. (2019). A weighting method to improve habitat association analysis: tested on British carabids. *Ecography*, 42(8), 1395-1404.

Crick, H. Q. P., Crosher, I. E., Mainstone, C. P., Taylor S. D., Wharton, A., Langford, P., Larwood, J., Lusardi, J., Appleton, D., Brotherton, P. N. M., Duffield, S. J. & Macgregor N. A. (2020) Nature Networks Evidence Handbook. Natural England Research Report NERR081. Natural England, York.

Entwistle, N. S., Heritage, G. L., Schofield, L. A., & Williamson, R. J. (2019). Recent changes to floodplain character and functionality in England. *Catena*, 174, 490-498.
Fryer, J., & Williams, I. D. (2021). Regional carbon stock assessment and the potential effects of land cover change. *Science of The Total Environment*, 775, 145815.

Fuller, R.M., Groom, G.B. and Jones, A.R., 1994. The land-cover map of Great Britain: an automated classification of Landsat thematic mapper data. *Photogrammetric Engineering and Remote Sensing*, 60(5), pp.553-562.

Fuller, R.M., Cox, R., Clarke, R.T., Rothery, P., Hill, R.A., Smith, G.M., Thomson, A.G., Brown, N.J., Howard, D.C. and Stott, A.P., 2005. The UK land cover map 2000: Planning, construction and calibration of a remotely sensed, user-oriented map of broad habitats. International Journal of Applied Earth Observation and Geoinformation, 7(3), pp.202-216.

Henrys, P.A. and Jarvis, S.G., 2019. Integration of ground survey and remote sensing derived data: Producing robust indicators of habitat extent and condition. Ecology and evolution, 9(14), pp.8104-8112.

Jackson, D. Guidance on the Interpretation of the Biodiversity Broad Habitat Classification (Terrestrial and Freshwater Types): Definitions and the Relationship with Other Habitat Classifications; Joint Nature Conservation Committee: London, UK, 2000.

Kilcoyne E, A.M., Clement, M., Moore, C., Picton Phillipps, G.P., Keane, R., Woodget, A., Potter, S., Stefaniak, A. & Trippier, B. (2022). Living England: Technical User Guide. NERR108. Natural England. Online at: https://publications.naturalengland.org.uk/publication/4918342350798848 last accessed January 2024.

Mancini, F., Hodgson, J. A., & Isaac, N. J. (2022). Co-designing an Indicator of Habitat Connectivity for England. *Frontiers in Ecology and Evolution*, 10.

Marston, C. G., O'Neil, A. W., Morton, R. D., Wood, C. M., and Rowland, C. S.: LCM2021 – The UK Land Cover Map 2021, *Earth Syst. Sci. Data Discuss*. [preprint], https://doi.org/10.5194/essd-2023-78, in review, 2023.

Marston, C.; Rowland, C.S.; O’Neil, A.W.; Morton, R.D. (2022). Land Cover Map 2021 (25m rasterised land parcels, GB). NERC EDS Environmental Information Data Centre. https://doi.org/10.5285/a1f85307-cad7-4e32-a445-84410efdfa70

Maskell, L., Jarvis, S., Jones, L., Garbutt, A. and Dickie, I., 2014. Restoration of natural capital: Review of evidence. Online at:

https://www.theoep.org.uk/sites/default/files/reports-files/ncc-research-restoration-natural-capital-review.pdf, last accessed: 18/10/2023.

McKenna, P. B., Lechner, A. M., Hernandez Santin, L., Phinn, S., & Erskine, P. D. (2023). Measuring and monitoring restored ecosystems: can remote sensing be applied to the ecological recovery wheel to inform restoration success?. *Restoration Ecology*, 31(1), e13724.

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.

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 Map 2018 (25m rasterised land parcels, GB). NERC Environmental Information Data Centre. https://doi.org/10.5285/25c6451b-5c88-40da-9a63-c3ec473e4874

Morton, R. D.; Marston, C.G.; O'Neil, A. W.; Rowland, C. S. (2020c). Land Cover Map 2019 (25m rasterised land parcels, GB). NERC Environmental Information Data Centre. https://doi.org/10.5285/f15289da-6424-4a5e-bd92-48c4d9c830cc

Morton, R.D.; Marston, C.G.; O'Neil, A.W.; Rowland, C.S. (2021). Land Cover Map 2020 (25m rasterised land parcels, GB). NERC EDS Environmental Information Data Centre. https://doi.org/10.5285/6c22cf6e-b224-414e-aa85-900325baedbd

Natural England (2022) User Guide for Natural England's Priority Habitats' Inventory Version 3.0 November 2022, online at: https://www.data.gov.uk/dataset/4b6ddab7-6c0f-4407-946e-d6499f19fcde/priority-habitats-inventory-england, Last accessed: 18/10/2023

Office for Environmental Protection (2023) Progress in improving the natural environment, 2021/2022. Online at: Progress in improving the natural environment in England, 2021/2022 | Office for Environmental Protection (theoep.org.uk), Last accessed: 17/10/2023

Office for National Statistics (ONS) (2011) Boundary Dataset Guidance – Information Note, online at: https://geoportal.statistics.gov.uk/documents/boundary-dataset-guidance-2011-to-2021/explore accessed: 26/09/2023

Parliamentary Office of Science and Technology (POST). (2022) POSTbrief 48: Restoration and creation of semi-natural habitats. UK Parliament. https://doi.org/10.58248/PB48

R Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Ritson, J. P., Croft, J. K., Clark, J. M., Brazier, R. E., Templeton, M. R., Smith, D., & Graham, N. J. D. (2019). Sources of dissolved organic carbon (DOC) in a mixed land use catchment (Exe, UK). Science of The Total Environment, 666, 165-175.

Rowland, C., Buys, G., Clark, L., Correia, V., Ditchburn, B., Evangelides, C., Higgins, A., Marston, C., Morton, D., Olave, R., O'Neil, A., Tomlinson, S., Walker, A., Watterson, J., Whitton, E., Williamson, J. and Thomson. A. (2021) BEIS Land-Use Tracking Project – WPB Final Report. Unpublished UKCEH led report for BEIS.

Rowland, C.S.; Marston, C.G.; Morton, R.D.; O'Neil, A.W. (2020). Land Cover Map 1990 (25m raster, GB) v2. NERC Environmental Information Data Centre. https://doi.org/10.5285/1be1912a-916e-42c0-98cc-16460fac00e8

Rowland, C.S.; Morton, R.D.; Carrasco, L.; McShane, G.; O'Neil, A.W.; Wood, C.M. (2017). Land Cover Map 2015 (25m raster, GB). NERC Environmental Information Data Centre. https://doi.org/10.5285/bb15e200-9349-403c-bda9-b430093807c7

Smith, G.M.; Morton, R.D. (2010) Real world objects in GEOBIA through the exploitation of existing digital cartography and image segmentation. *Photogrammetric Engineering & Remote Sensing*, *76*(2), pp.163-171.

Wickham, J., Stehman, S. V., Sorenson, D. G., Gass, L., & Dewitz, J. A. (2021). Thematic accuracy assessment of the NLCD 2016 land cover for the conterminous United States. *Remote Sensing of Environment*, **257**, 112357.

Appendix 1: R script used to calculate summary values

```
# script to calculate LCM statistics for England for the OEP.
#
# Input data sets:
# - LCM data for 1990, 1994, 1998, 2002, 2006, 2010, 2015, 2017-2021(2?)
#
# - OS/ONS polygons for high and low water
#
# Processing steps:
#
# 1 - create masks for England high and low water extents
# 2 - loop through LCM's and:
#       i) clip to low and high water marks
#       ii) calculate statistics (calc_stats function)
#       iii) save as individual csv files
# 3 - merge csv files for high and low water(merge_csv function)
#
# Clare Rowland, 18/09/2023

#packages required
library(raster)
library(sf)
library(terra)


# 1 - create mask for England extent
print ("Create masks for low and high water extent for England...")

# location of key files:
# - LCM data set to provide the template for the mask
lcmgrid <- rast('D:/LCM_data/3_raster_21class/25m/gb2020lcm25m_b1.tif')

# - ONS data for high and low water marks
high_water <- 'D:/oep/ons/BFC_high/CTRY_DEC_2022_GB_BFC.shp'
low_water <- 'D:/oep/ons/BFE_low/CTRY_DEC_2022_GB_BFE.shp'

# read in vector data
vec_data_high <- st_read(high_water,  quiet = TRUE)
vec_data_low <- st_read(low_water, quiet = TRUE)

field_id <- ('CTRY22NM')
field_value <- ('England')

# extract england only for high water
england_high_vec <- subset(vec_data_high, vec_data_high$CTRY22NM ==
field_value)
st_write(england_high_vec,  append=FALSE,
'D:/oep/checks/BFE_high_water_england.shp')

# extract england only for low water
england_low_vec <- subset(vec_data_low, vec_data_low$CTRY22NM ==
field_value)
st_write(england_low_vec,  append=FALSE,
'D:/oep/checks/BFE_low_water_england.shp')
```

```
# create mask for England
high_water_mask <- rasterize(england_high_vec, lcmgrid,
field=1,filename='D:/oep/high_water.tif', overwrite=FALSE)
low_water_mask <- rasterize(england_low_vec, lcmgrid,
field=1,filename='D:/oep/low_water.tif', overwrite=FALSE)


# 2 - loop through LCM's and clip to low extent and high extent


calc_stats <- function(mask_data_set, year, tidal_state){

  lcm_file <-
paste0('D:/LCM_data/3_raster_21class/25m/gb',year,'lcm25m_b1.tif')
  print (paste('input filename:', lcm_file))

  #open lcm data
  lcmdata <- rast(lcm_file)

  #output filenames
  raster_name <-
paste0('D:/oep/lcm_raster/LCM_eng_',year,'_',tidal_state,'.tif')
  csv_name <-
paste0('D:/oep/lcm_stats/LCM_eng_',year,'_',tidal_state,'.csv')

  print (paste("Calculating the histogram...", year))

  # calculate values
  vals <- mask_data_set * lcmdata
  vals <- replace(vals, vals == 0, NA) # required for 2015, which has 0
values
  v <- values(vals, na.rm = TRUE)
  h <- table(v)
  write.csv(h, csv_name)
  writeRaster(vals,raster_name, overwrite=TRUE)

}

merge_csv_files <- function(tidal_state){
  #this function merges all the low, or all the high, tide csv files
  years <- c('LCM1990', 'LCM1994',
'LCM1998','LCM2002','LCM2006','LCM2010','LCM2015',
             'LCM2017', 'LCM2018', 'LCM2019', 'LCM2020', 'LCM2021')
  pattern_string <- paste0("*\\",tidal_state,'.csv$')
  csv_list <- list.files(('D:/oep/lcm_stats/'), pattern=pattern_string,
full.names = TRUE)
  # assigned rowname is Freq so skip first row
  df_list <- lapply(csv_list, read.table, sep = ",",  na.strings = c("",
"NA"), skip=1)
  names(df_list) <- years
  op_list <- lapply(df_list, `[[`, 3)
  merge_csv_name <-
paste0('D:/oep/lcm_stats/LCM_england_',tidal_state,'_tide_mark.csv')
  #save output
  write.csv(op_list,merge_csv_name)
```

```
}
#re-open raster mask files if necessary
high_water_mask <- rast('D:/oep/high_water.tif')
low_water_mask <- rast('D:/oep/low_water.tif')

for (year in c(1990, 1994, 1998,2002,2006,2010, 2015, 2017, 2018, 2019,
2020, 2021)){
  #for (year in c(2015)) {
  calc_stats(low_water_mask, year, 'low')
  calc_stats(high_water_mask, year, 'high')
}

# 3 - merge csv files for high and low water
merge_csv_files('low')
merge_csv_files('high')
```

Contact

enquiries@ceh.ac.uk

@UK_CEH

ceh.ac.uk

Bangor

UK Centre for Ecology & Hydrology
Environment Centre Wales
Deiniol Road
Bangor
Gwynedd
LL57 2UW

+44 (0)1248 374500

Edinburgh

UK Centre for Ecology & Hydrology
Bush Estate
Penicuik
Midlothian
EH26 0QB

+44 (0)131 4454343

Lancaster

UK Centre for Ecology & Hydrology
Lancaster Environment Centre
Library Avenue
Bailrigg
Lancaster
LA1 4AP

+44 (0)1524 595800

Wallingford (Headquarters)

UK Centre for Ecology & Hydrology
Maclean Building
Benson Lane
Crowmarsh Gifford
Wallingford
Oxfordshire
OX10 8BB

+44 (0)1491 838800

UK Centre for Ecology & Hydrology