Countryside Survey

Photograph - iStockphoto

Countryside Survey:
Final Report for LCM2007 – the new UK Land Cover Map

CS Technical Report No 11/07

Final Report for LCM2007 - the new UK Land Cover Map

D. Morton, C. Rowland, C. Wood, L. Meek,
C. Marston, G. Smith, R. Wadsworth, I. C. Simpson

Centre for Ecology & Hydrology
(Natural Environment Research Council)

July 2011

Contents

Acknowledgements

Executive Summary

LCM2007

- Land Cover Map 2007 (LCM2007) is the first UK land cover map with land parcels (the spatial framework) derived from national cartography by a generalisation (simplification) process. This dramatically improves spatial and thematic accuracy and better represents real world objects.
- The GB framework is based on Ordnance Survey Master Map topography layer (hereafter referred to as OSMM) and the NI framework is based on cartographic data from Ordnance Survey Northern Ireland (now part of Land & Property Services).
- The spatial framework has been further refined by supplementing the generalised national cartography with agricultural census data boundaries and image segments.
- LCM2007 is the first land cover map to provide continuous vector coverage of UK Broad Habitats derived from satellite data.
- LCM2007 contains almost 10 million land parcels (8.6 million for GB; 0.9 million for NI).

Heritage

- LCM2007 is the third in the Land Cover Map (LCM) series:
  - o LCM1990 (originally called LCMGB) is the first GB map of land cover produced from satellite data. It is a raster product with land cover given as 25 bespoke target classes.
  - o LCM2000 is a vector land cover map of the UK based on satellite data. Land parcels were derived from image segments. Land cover is based upon UK Biodiversity Action Plan (BAP) Broad Habitats.

Production

- LCM2007 was produced from over seventy satellite images, which were combined into 34 multi-date summer-winter images. The multi-date images were based on one summer and one winter image. The summer-winter images increase the contrast between land cover types and so increase the accuracy of the classification.
  - o 91% of UK was mapped from summer-winter images (compared to 84% for LCM2000 and 87% for LCM1990)
  - o 9% was classified from single-date imagery (compared to 15% for LCM2000 and 11% for LCM1990)
  - o 0.5% was filled manually
- 99.5% of parcels were classified using automated procedures. 0.5% were classified by visual interpretation.
- After initial classification, knowledge-based enhancements were applied to increase the refinement and accuracy of the classification using soil, altitude and urban extent data sets.

Quality Assurance

- Field validation points were collected between 2006 and 2008. 9127 of the field validation points were used to validate LCM2007 giving an overall accuracy of 83% for the LCM2007 classes.
- The accuracy for different land classes varies.

Comparison with Countryside Survey

- LCM2007 has been assessed against a) the 591 Countryside Survey squares, b) against the Countryside Survey estimates of Broad Habitat extent for the UK.
- The comparison of LCM2007 to the 591 1km x 1km squares surveyed by Countryside Survey showed:
  - 62% correspondence at the BH-level, 67% at the aggregate level and 76% at the associated BH-level.
  - For the common Broad Habitats (defined here as those covering more than 10,000km^2 of the UK) the User's accuracy varies from 89% for *'Built-up Areas and Gardens'* to 37% for *'Dwarf, Shrub Heath'*.
- Countryside Survey has an established methodology for calculating National Estimates of the area of each Broad Habitat for the UK from the CS-squares. The CS in 2007 estimates of Broad Habitat were compared to the LCM2007 areas of BH for the UK:
  - The LCM classes which fall within the Countryside Survey upper and lower 95% confidence limits for the UK are: *'Coniferous Woodland'*, **Freshwater**, *'Built-up Areas and Gardens'* and *'Calcareous Grassland',* **Broadleaved Woodland**, *'Acid Grassland'* and 'Inland *Rock'.*

Summary statistics

- LCM2007 shows the UK land cover as being comprised of mainly *'Arable and Horticulture'* and *'Improved Grassland'* (25% each), with the other main land covers being **Semi-natural grassland** (13%) and **Mountain, Heath and Bog** (16%). Urban areas make-up 6% of the UK, as do *'Coniferous Woodland'* (6%) and **Broadleaved woodland** (6%), with coastal classes and **Freshwater** contributing the remaining 2%.

Change mapping

- Using CEH land cover products for change mapping is complex because:
  - The classes used in the 1990 map differ from the two later LCMs.
  - The spatial structure differs across all three LCMs.
  - Classified images typically have an error of around 20%, whereas the level of BH change is likely to be less than this in the period between LCMs.
  - Reliable methods for separating real changes from those due to error and spatial inconsistency have not yet been established.

Product specification

- LCM2007 maps 23 land cover classes, which combine to map 17 terrestrial Broad Habitats.
- LCM2007 maps land cover, rather than land use.
- LCM2007 has a minimum mappable unit of 0.5ha.
- LCM2007 has a rich metadata to enable users to track the processing steps applied to each polygon.

LCM2007 data products

- The main LCM2007 product is the vector data set, which has 10 attributes that document the processing stages applied to the polygon, including the polygon construction (Construct attribute), spectral classification result (ProbList attribute) and knowledge-based enhancements (KBE attribute).
- The main raster data set is a 25m product containing the 23 LCM2007 land cover classes.
- LCM2007 is also summarised as a series of 1km data products, which give percentage cover or dominant land cover for each 1km square. The 1km products are produced for the 23 LCM2007 land cover classes and for the 10 LCM2007 Aggregate classes.

Providing a scientific evidence base

- LCM2007 will be a key component in developing multi-tiered approaches to habitat monitoring and informing evidence-based policy.
- LCM2007 will help to inform environmental policy in many areas including: biodiversity, ecosystem services, landscape planning, habitat connectivity and catchment management.
- Used with other data, LCM2007 will have applications in many other sectors.

Accessing LCM2007 data sets

- The LCM2007 1km raster data sets are available via the CEH Information Gateway [https://gateway.ceh.ac.uk]
- The full vector product and 25m product are available under licence on request from CEH. Please complete the online application on the CEH web site [www.ceh.ac.uk/data] or contact spatialdata@ceh.ac.uk for further details. Please note that licence fees may apply for some users and some applications.

Key figures

- Land Cover Map 2007: Figure 2.1, page 14
- UK land cover percentages: Figure 4.7, page 71
- Examples of LCM2007: Figures 5.2-5.6, pages 73-77

Chapter 1: Introduction

Land cover is the physical material on the surface of the Earth. Examples include artificial materials such as roads and buildings in urban areas, natural and managed vegetated surfaces and inland water. Land cover maps have multiple uses. They are used for the management of natural resources, urban planning, carbon accounting and flood risk modelling, to name only a few. Land cover maps can be obtained by ground-based surveys but for large surface areas this would be impractical and too expensive. To realise cost effective large-scale land cover maps, satellite images are used. This document describes the production of the Land Cover Map for 2007 (LCM2007). LCM2007 is the third digital Land Cover Map produced by CEH from satellite images. It gives land cover in relation to UK Biodiversity Action Plan (BAP) Broad Habitats. The UK BAP is the UK Governments response to the Convention on Biological Diversity (CBD) arising from the Earth Summit in Rio 1992. The complete UK coverage of LCM2007 complements the intensive sampling, but restricted spatial coverage of the Countryside Survey of 2007. CS is a unique study of the natural resources of the UK's countryside which involves ground-based mapping and assessment of associated vegetation, soils and freshwaters.

Three National Land Cover Maps

The first digital land cover map for Great Britain derived from satellite imagery is the Land Cover Map of Great Britain 1990 (LCM1990, formerly LCMGB 1990). This is a 25m x 25m pixel land cover product describing land cover in 1990. LCM1990 pre-dates the CBD so land cover is given as 25 bespoke target classes. It was produced by automated classification techniques where each pixel is assigned to a land cover class based upon its spectral characteristics.

The Land Cover Map for 2000 (LCM2000) is the second map in the series and includes Northern Ireland. The production of LCM2000 used novel object based image analysis (OBIA) techniques that had recently emerged from geographical information science. Instead of representing the land surface as regular sized pixels, OBIA considers the land surface a collection of discrete irregular objects such as forests, lakes, urban areas and fields. Partitioning an image into objects is analogous to the way humans conceptually organise the landscape in order to understand it and the contextual relationships between objects assist with their classification. It is widely accepted that OBIA produces thematic land cover maps with more accuracy than pixel based approaches (Gao and Mas 2008).

Land cover objects for LCM2000 were derived from image segments (see Section 3.3) and were assigned land cover values according to the pixel distributions within. These classifications where then refined using contextual and ancillary information.

LCM2007 builds upon the successes of LCM2000 and employs similar but enhanced classification techniques. The principal difference between LCM2007 and LCM2000 is the source of land cover objects. LCM2007 objects come from generalised digital cartography, refined with image segments. LCM2000 uses only image segments. Deriving objects from digital cartography is advantageous. Cartographic boundaries

very accurately delineate real-world land cover objects (e.g. lakes, fields, settlements, industrial areas, semi-natural areas etc) and this improves the spatial accuracy of LCM2007 over LCM2000. Moreover, the accurate delineation of real-world objects helps to clarify the spectral properties of the land surface and therefore improves thematic accuracy too.

The other major advantage that comes from using national cartography relates to re-usability and change detection. Image segments represent a snapshot of land surface reflectance. Reflectance varies temporally. A segmentation performed in one year will therefore give a different result to the next or previous years. The consequence of this is that successive land cover maps based on image segments will have very different spatial structures. This restricts spatial and temporal comparisons of land cover and therefore the ability to detect change. By using digital cartography LCM2007 overcomes this problem. Many of the boundaries mapped in the UK countryside come from or pre-date the enclosure acts of the 1700s and 1800s. They are relatively static through time when compared to changes in land use and land cover. This relative fixation of land cover objects, the principal units of change, makes the generalised cartography of LCM2007 re-usable. Successive land cover maps that share a common spatial structure will make accurate and spatially explicit change detection a realistic ambition. Therefore in addition to providing up-to-date information on land cover, because of its spatial structure LCM2007 gives us a platform for future UK wide land surface mapping and monitoring exercises.

Please note: *The UK component of the Corine Land Cover 2006 produced on behalf of the European Environment Agency, has been derived from the LCM2007 vector product by a process involving geometric generalisation and thematic transformations.*

Chapter 2: LCM2007 Classification Scheme

Background
The UK Biodiversity Group identified a group of Broad Habitats to cover the complete range of UK habitats, as an initial stage in the development of the UK Biodiversity Action Plan (BAP). The Broad Habitats are described by the Joint Nature Conservancy Committee (JNCC) (Jackson, 2000) and are used by Countryside Survey and LCMs 2000 and 2007. LCM2007 aimed to contribute to the habitat assessment by mapping the UKs Broad Habitats.

> A consistent formatting style will be used when referring to Broad Habitats, LCM2007 classes and LCM2007 aggregate classes. For Broad Habitats the convention is: italics, a capital letter at the start of words and enclose in single speech-marks e.g. '*Improved Grassland*'. LCM2007 classes are given in bold e.g. **Freshwater**, as are LCM2007 Aggregate classes.

LCM2007 classes
LCM2007 classifies the land cover of the UK using classes based on the Broad Habitats, with some minor differences. Table 2.1 summarises the LCM2007 classes and notes their correspondence with the Broad Habitat classification scheme and also with LCM2000. A summary of the Broad Habitats is given in Appendix 1.

Spectral remote sensing can be used to classify land cover, however, the relationship between land cover and habitat is not straightforward. Sometimes land cover and habitat type have a direct and unique match. An example of a unique, direct match would be a patch of coniferous trees, which would map directly to the '*Coniferous Woodland*' Broad Habitat. In some cases the links between land cover and Broad Habitat are more complicated and exhibit a 'one-to-many' relationship. This is particularly true of grassland, which forms the '*Improved Grassland*', '*Neutral Grassland*', '*Acid Grassland*' and '*Calcareous Grassland*' habitats plus some types of '*Bog*' habitats. Grassland is also present, although not as the dominant land cover, in '*Built-up Areas and Gardens*', '*Fen, Marsh and Swamp*' and Saltmarsh, which falls into the '*Littoral sediment*' category. Consequently, assigning grassland to different habitat types is not straightforward.

In the production of LCM2007 grassland is mapped by classifying images into '*Improved Grassland*' and **Rough grassland**, after which knowledge-based enhancement (KBE) rules are applied. The knowledge-based enhancement rules determine whether **Rough grassland** should be reclassified as '*Neutral Grassland*', '*Calcareous Grassland*' or '*Acid Grassland,* or whether it should remain as **Rough grassland** (for details about KBEs see Section 3.7). The requirement for KBEs to separate **Rough grassland** into grassland Broad Habitats is important because grassland Broad Habitats, in the field, may be determined by the presence of a few indicator species rather than the dominant species. The determination between grassland Broad Habitats in LCM2007 is based on soil type, as shown by soil data sets, and consequently does not always reach the same conclusion as field survey. The semi-natural grasslands in LCM2007 therefore differ from the strict Broad Habitat interpretation, as they are partly based on soil data sets, rather than species composition (this is discussed further in Chapter 4). In Table 2.1 Neutral, Acid and

Table 2.1. Summary of LCM2007 classes and Broad Habitat. Green shading highlights Broad Habitats as documented by JNCC (Jackson, 2000).

Broad Habitat	LCM2007 class	Notes on LCM2007 class
'Broadleaved, Mixed and Yew Woodland'	**Broadleaved woodland**	Same as LCM2000 class. Differs from BH due to exclusion of Yew, which is not extensive enough for LCM to map.
'Coniferous Woodland'	*'Coniferous Woodland'*	As BH and LCM2000
'Arable and Horticulture'	*'Arable and Horticulture'*	As BH and LCM2000
'Improved Grassland'	*'Improved Grassland'*	As BH and LCM2000
	Rough grassland	Mix of areas of managed, low productivity grassland, plus some areas of semi-natural grassland, which could not be assigned Neutral, Calcareous or Acid Grassland with confidence by the knowledge-based enhancements.
'Neutral Grassland'	*'Neutral Grassland'*	As BH and LCM2000
'Calcareous Grassland'	*'Calcareous Grassland'*	As BH and LCM2000
'Acid Grassland'	**Acid grassland**	Acid grassland incorporates Bracken. Bracken is a Broad Habitat under certain circumstances (see Appendix 1). Bracken can be mapped using LCM2007 methods, but it depends on image timing, so for consistency it is assigned to *'Acid Grassland'*.
'Fen, Marsh and Swamp'	*'Fen, Marsh and Swamp'*	As BH and LCM2000
'Dwarf Shrub Heath'	**Heather**	As LCM2000; spectral differences between dense heather and heather grassland enable separation spectrally.
	Heather grassland	
'Bog'	*'Bog'*	As BH and LCM2000
'Montane Habitats'	*'Montane Habitats'*	As BH; Altitude cut-off differs from LCM2000
'Inland Rock'	*'Inland Rock'*	As BH and LCM2000
	Saltwater	As LCM2000
'Standing Water and Canals'	**Freshwater**	Merged 2 freshwater BHs, as they cannot be separated from each other using the methods and data used for LCM2007. In many cases small and/or narrow water bodies fall below the MMU.
'Rivers and Streams'		
'Supra-littoral Rock'	*'Supra-littoral Rock'*	As BH and LCM2000
'Supra-littoral Sediment'	*'Supra-littoral Sediment'*	As BH and LCM2000
'Littoral Rock'	*'Littoral Rock'*	As BH and LCM2000
'Littoral Sediment'	**Littoral sediment**	As LCM2000
	Saltmarsh	Priority Habitat and of sufficient extent and spectral distinction to be mapped consistently
'Built-up Areas and Gardens'	**Suburban**	As LCM2000; spectral differences between urban and suburban enable separation spectrally.
	Urban	

Calcareous Grassland are identified as being 'BH and LCM2007' as the aim is to map to the specific grassland Broad Habitat type as far as possible.

The montane distribution for LCM2007 is assigned based on altitude (see Section 3.7), whereas the Broad Habitat definition is based on vegetation type. The Broad Habitat definitions produced by JNCC note (Jackson, 2000) that if other habitats, such as *'Calcareous Grassland'* and '*Bog*', occur within the *'Montane Habitats'* zone they should not be recorded as *'Montane Habitats'*. In the production of LCM2007 it was not possible to determine whether a *'Montane Habitats'* reclassification based on altitude has greater validity than the original spectral classification. Consequently, in LCM2007 above the montane altitude LCM2007 maps three Broad habitats: *'Montane Habitats'*, **Freshwater** and 'Inland Rock'.

Steps such as the application of knowledge-based enhancements to identify the different semi-natural grassland types and the *'Montane Habitats'* areas are documented within the polygon attributes of the vector data set. Consequently, if a user wishes to remove a particular knowledge-based enhancement, either to retrieve a merged class such as Bracken, or to remove the *'Montane Habitats'* altitude-based correction, then they can. This is because the original spectral classification associated with each polygon remains and is documented in the ProbList attribute of the vector data set (see Table 5.1 for more details).

There are a number of habitats where it may not be readily apparent how they are categorised in the JNCC descriptions (Jackson, 2000), including:
- Machair which is a Priority Habitat that goes into the '*Supra-littoral Sediment'* category, rather than the '*Calcareous Grassland'* Broad Habitat (Jackson, 2000, Annex 1: Table 3a http://www.jncc.gov.uk/page-2433-theme=default).
- Urban despoiled land cover types, such as quarries and peat extraction, fall into the '*Inland Rock'*, rather than the '*Built-up Areas and Gardens*' Broad Habitat, in the JNCC Broad Habitat definitions.
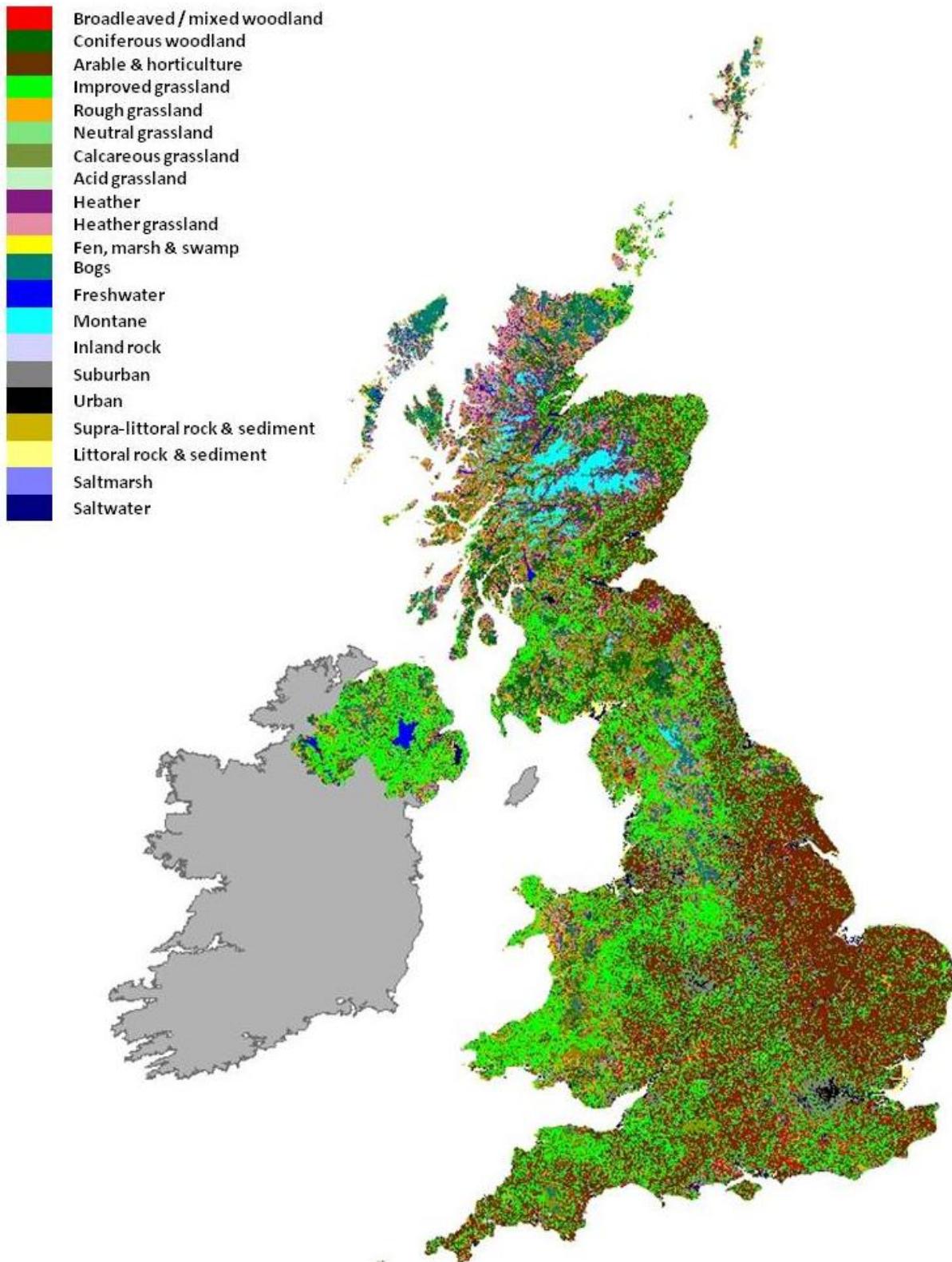
LCM2007 Aggregate classes
Not all applications need the thematic resolution of the complete set of LCM2007 classes, so LCM2007 has a defined set of Aggregate classes, based on merging LCM2007 classes (Table 2.2). The Aggregate classes are used for the 1km raster products and for data analysis and presentation in Chapters 4 and 5.

LCM2007- the new Land Cover Map
The classes described in Table 2.1 were used to classify the UK and produced the map shown in Figure 2.1. A full overview of Land Cover Map 2007 can be seen in Figure 2.1. The distribution of the main Broad Habitats can be clearly identified; for example, the arable land in the eastern part of Great Britain, improved grassland in western England, Wales and Northern Ireland, coniferous plantations in the Scottish Borders and the domination of the north west of Scotland with bog, heather and acid grassland habitats. The larger urban areas are also clearly identifiable, as are the Montane regions in Scotland and Loch Neagh in Northern Ireland.

Table 2.2 Relationship between Aggregate classes, Broad Habitat and LCM2007 classes. Green shading highlights Broad Habitats as documented by JNCC (Jackson, 2000). [1,2] Aggregate class number and LCM2007 class number are used for 1km and 25m raster data sets respectively.

Aggregate class	Aggregate class number[1]	Broad habitat	LCM2007 class	LCM2007 class number[2]
Broadleaf woodland	1	*'Broadleaved, Mixed and Yew Woodland'*	**Broadleaved woodland**	1
Coniferous woodland	2	*'Coniferous Woodland'*	*'Coniferous woodland'*	2
Arable	3	*'Arable and Horticulture'*	*'Arable and horticulture'*	3
Improved grassland	4	*'Improved Grassland'*	*'Improved grassland'*	4
Semi-natural grassland	5	**Rough Grassland**	**Rough grassland**	5
		'Neutral Grassland'	*'Neutral Grassland'*	6
		'Calcareous Grassland'	*'Calcareous Grassland'*	7
		'Acid Grassland'	**Acid grassland**	8
		'Fen, Marsh and Swamp'	*'Fen, Marsh and Swamp'*	9
Mountain, heath, bog	6	*'Dwarf Shrub Heath'*	**Heather**	10
			Heather grassland	11
		'Bog'	*'Bog'*	12
		'Montane Habitats'	*'Montane Habitats'*	13
		'Inland Rock'	*'Inland Rock'*	14
Saltwater	7	**Saltwater**	**Saltwater**	15
Freshwater	8	**Freshwater**	**Freshwater**	16
Coastal	9	*'Supra-littoral Rock'*	*'Supra-littoral Rock'*	17
		'Supra-littoral Sediment'	*'Supra-littoral Sediment'*	18
		'Littoral Rock'	*'Littoral Rock'*	19
		'Littoral Sediment'	**Littoral sediment**	20
			Saltmarsh	21
Built-up areas and gardens	10	*'Built-up Areas and Gardens'*	**Urban**	22
			Suburban	23
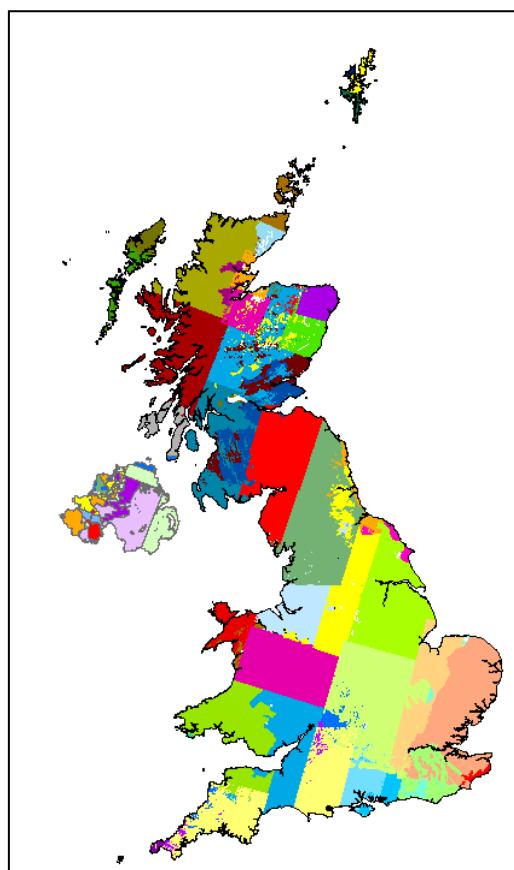
Figure 2.1. Land Cover Map 2007. © NERC (CEH) 2011. © Crown Copyright 2007. Ordnance Survey Licence number 100017572. © Crown Copyright 2011. Licence number 100,427. © third-party licensors.

Chapter 3: Data sources and production methodology

3.1 Satellite imagery

LCM2007 is based primarily on combined summer and winter satellite data acquired by the Landsat-TM5, IRS-LISS3 and SPOT-4 and SPOT-5 sensors, these satellites all have a pixel size of 20-30m. AWIFS, with a pixel size of 60m, was used when other imagery was unavailable. To minimise data volume, but retain spectral information, the red, NIR and MIR bands from the summer and winter images were combined to create 6-band 2-date summer-winter composite images (from here on referred to as composites, see Fuller & Parsell, 1990). Once classified the images produce a patchwork of land cover scenes which cover the UK (Figure 3.1).

Figure 3.1. A mosaic of the classified images used to produce full UK coverage for LCM2007. Different colours show the coverage of the different composite and single-date images (for image details see Appendices 3 and 2). © NERC (CEH) 2011. © Contains Ordnance Survey data © Crown Copyright 2009.
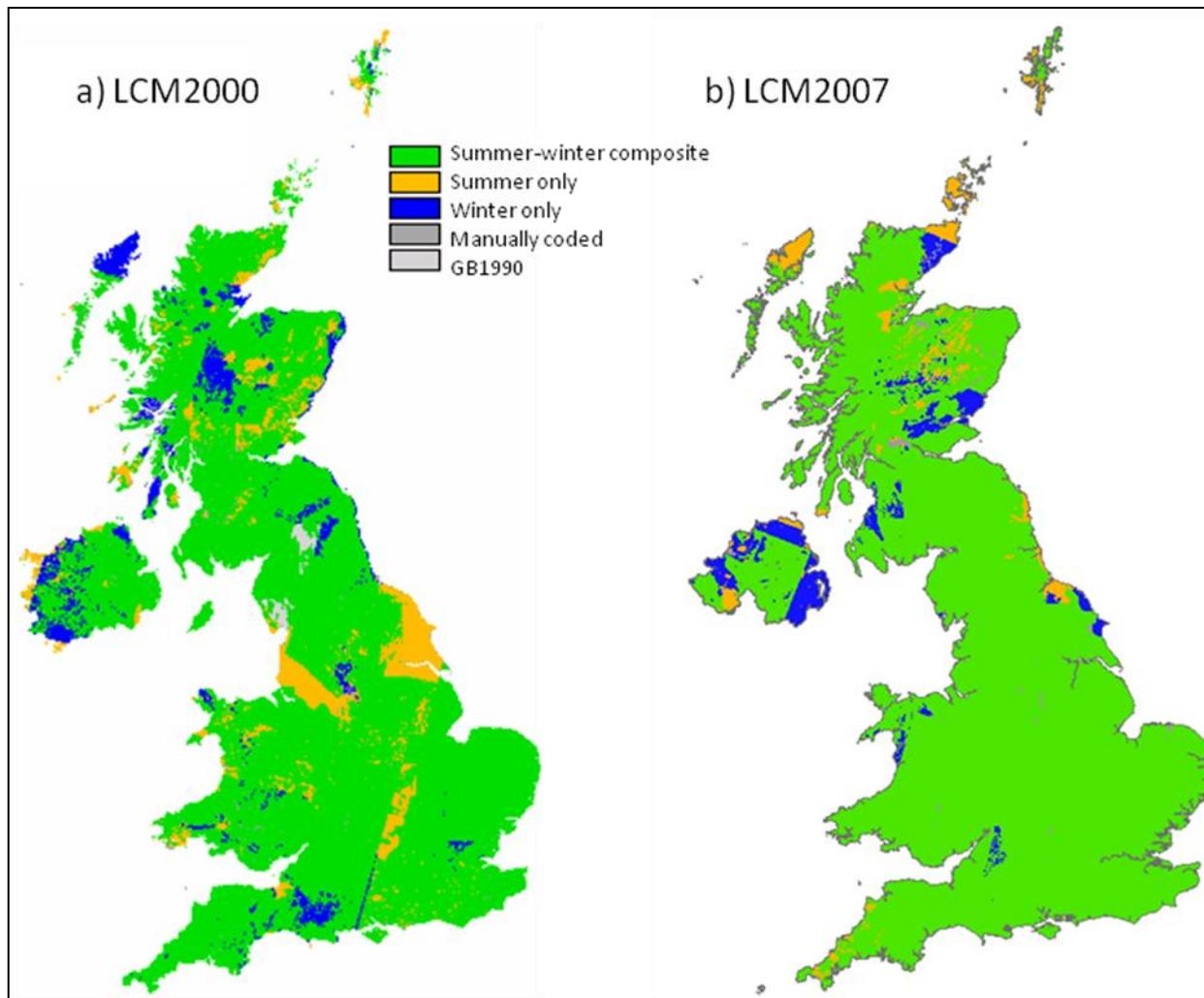
Ideally, the entire UK would be classified using composite images, however cloud-cover and the number of image acquisitions by the relevant satellites limit the amount of data available. For LCM2007 91% of the UK was classified using composite images (Table 3.1), which compares with 84% for LCM2000 and 87% for

GB1990. Where composite images are not available summer-only or winter-only images were used, although they typically have a reduced ability to distinguish between habitats. Figure 3.2 shows the distribution of composite and single-date imagery for LCM2000 (Fig. 3.2a) and LCM2007 (Fig. 3.2b). Single-date imagery was most widely used in Scotland and Northern Ireland for both LCM2000 and LCM2007, because cloud limited the availability of data (Fig. 3.2). The data classified for LCM2007 cover less than three years ranging from 02/09/2005 to 18/07/2008 (see Appendix 1 for further details) compared to over four years for LCM2000 (Fuller *et al.,* 2002) and GB1990 (Fuller *et al.,* 1994).
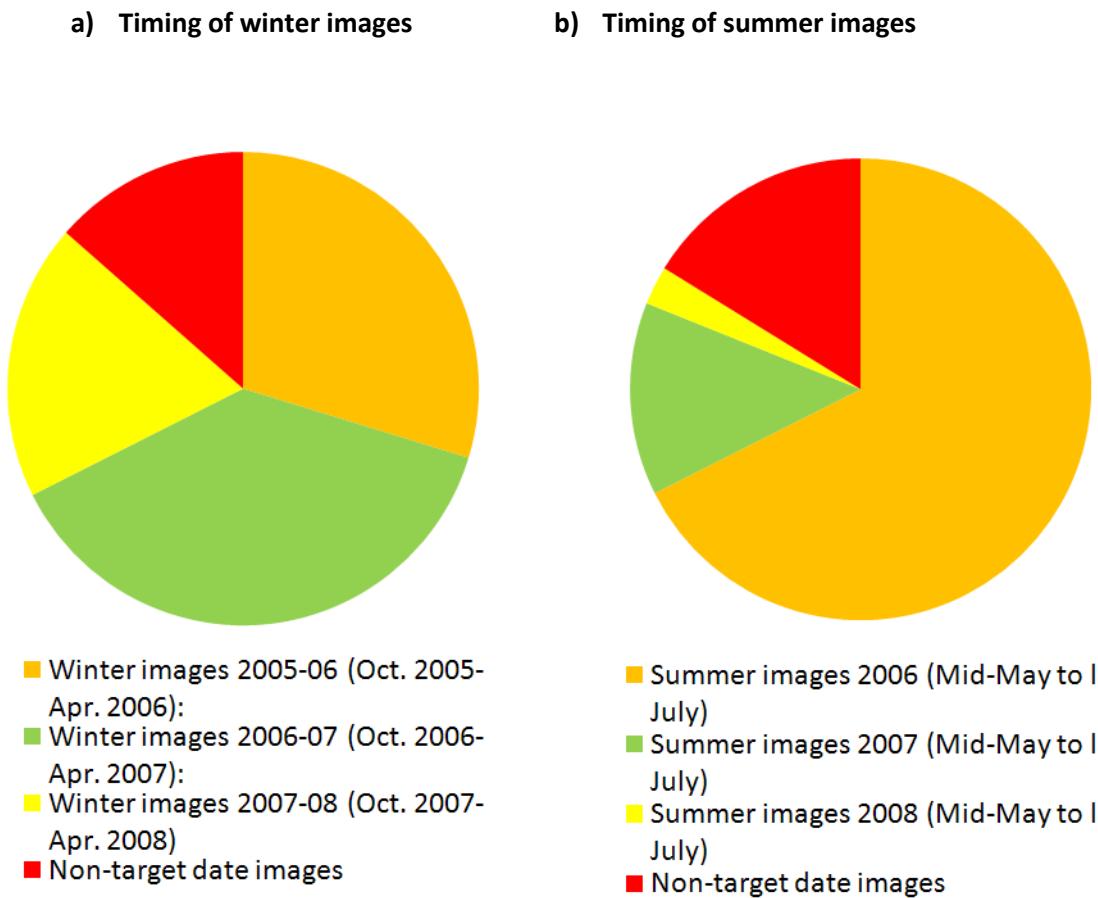
Table 3.1. Percentage of land classified by composite, summer-only or winter-only data for LCM1990, LCM2000 and LCM2007 (figures for LCM1990 and LCM2000 from Fuller *et al.,* 2002).

[1] For details see 'Hole-filling' in Section 3.6 Image Classification

	Total number of images	Number of composites	Area mapped (as a % of UK area)			
			Composite	Summer-only	Winter-only	Other
LCM1990	49	32	87% (56% entirely from target year)	8%	3%	3% unclassified
LCM2000	79	38	84% (23% entirely from target year)	9%	6%	1% filled with GB1990
LCM2007	73	34	91% (6% entirely from target year)	4%	5%	0.5% filled manually[1]
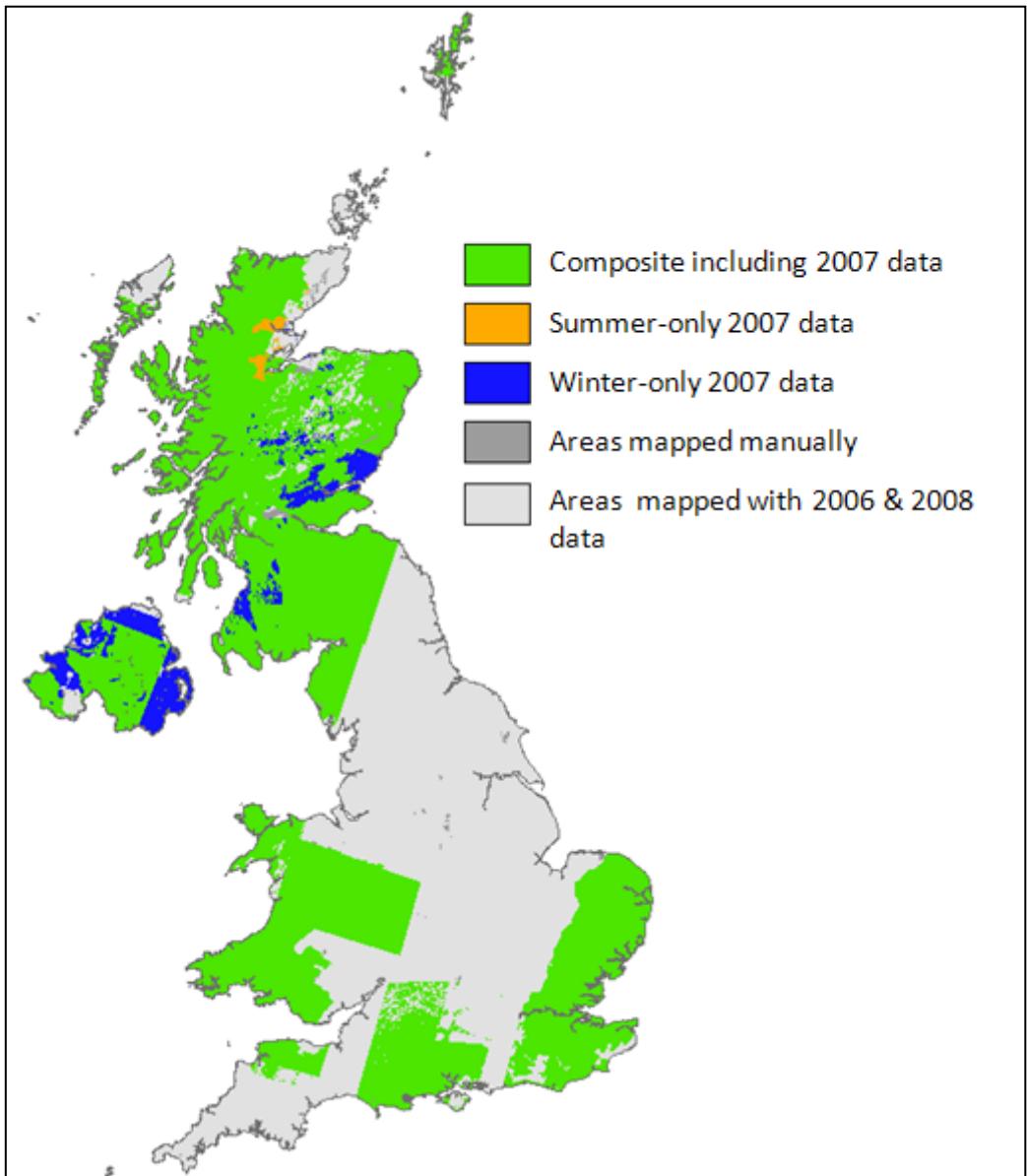
Figure 3.2. Contributions of composite (combined summer and winter imagery) and single-date (summer-only or winter-only) data sets for: a) LCM2000 and b) LCM2007. Note, GB1990 was only used for hole-filling LCM2000; gaps in LCM2007 were filled manually (for details see 'Hole-filling' in Section 3.6 Image Classification).

Winter images 2005-06 (Oct. 2005-
Apr. 2006):
Winter images 2006-07 (Oct. 2006-
Apr. 2007):
Winter images 2007-08 (Oct. 2007-
Apr. 2008)
Non-target date images

Summer images 2006 (Mid-May to late
July)
Summer images 2007 (Mid-May to late
July)
Summer images 2008 (Mid-May to late
July)
Non-target date images

Figure 3.3. Timing of summer and winter images used in LCM2007 composites.

Figure 3.3 shows the timing of the winter and summer images. Winter images are defined as those falling between October and the end of April (through to mid-May for northern Scotland), whilst summer images are considered to be mid-May through to late July (into August for northern Scotland) (Fuller *et al.*, 1994). These seasonal targets aim to capture the extremes of the phenological cycle from winter-senescence to the main growing season before harvest. For accurate classification it is desirable to have a winter image followed by a summer image, so that the images are from the same crop cycle.

LCM2007 is based on a target year of 2007, although other images are required. The target year of 2007 means that the preferred combination would be winter 2006/07 images (green section of Fig. 3.3a) combined with summer 2007 images (green section of Fig. 3.3b), but whilst a high proportion of the winter images are from the winter of 2006/07 only a small proportion of summer images were available for 2007 (Figure 3.3). Therefore, only three composites, covering 6% of the UK, met the preferred winter-summer combination (Table 3.1). Thirteen other composites were from the same crop cycle, but were not in 2007 the preferred year. However, 52% of the UK was mapped by composites including at least one image from 2007, with a further 4% of the UK mapped from single-date images from 2007 (Figure 3.4). Of the remaining composites, in 20 the summer image precedes the winter image, and in the remaining four the winter image preceded the summer image by 18+

Figure 3.4. Areas classified using 2007 images.

months. In a very small number of cases it was necessary to manually classify small areas of Landsat7-ETM+ data (see Hole-filling description in Section 3.6).

The sensors available for land cover mapping at the 20-30m pixel size have changed between the production dates of the three land cover maps. GB1990 was based entirely on Landsat-TM, whereas LCM2000 used Landsat-ETM, Landsat-TM, and LISS. For LCM2007 the available sensors were, in order of preference, Landsat-TM, LISS-3, SPOT-4/5, and AWIFS. Landsat-ETM7 data were available, but due to sensor damage in 2003, the images were not suitable for supervised classification. A breakdown of the composites by sensor shows that the bulk of the imagery used to produce LCM2007 is from Landsat-TM5 and LISS-3 (Table 3.2). The key advantage of using a range of satellites is the increased likelihood of good-quality cloud-free images. The disadvantages include slightly varying processing methods for different

sensors, plus the swathes of the sensors vary, so images from different sensors generally have much less overlap than images from the same sensor.

Table 3.2. Summary of the composites used in LCM2007 by sensor.
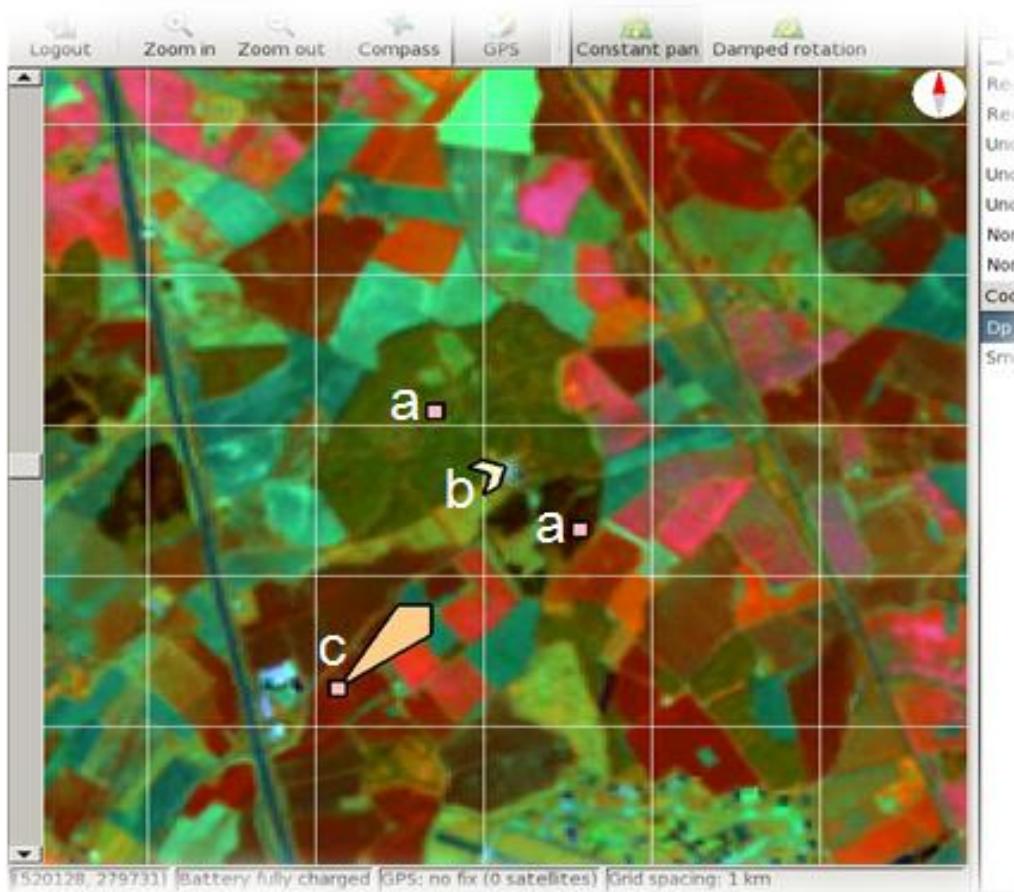
Sensor	AWIFS	SPOT-4/5	Landsat-TM5	LISS-3
LISS-3	5	2	9	5
Landsat-TM5	1	6	4	
SPOT-4/5	1	3		

3.2 Additional spatial data sets
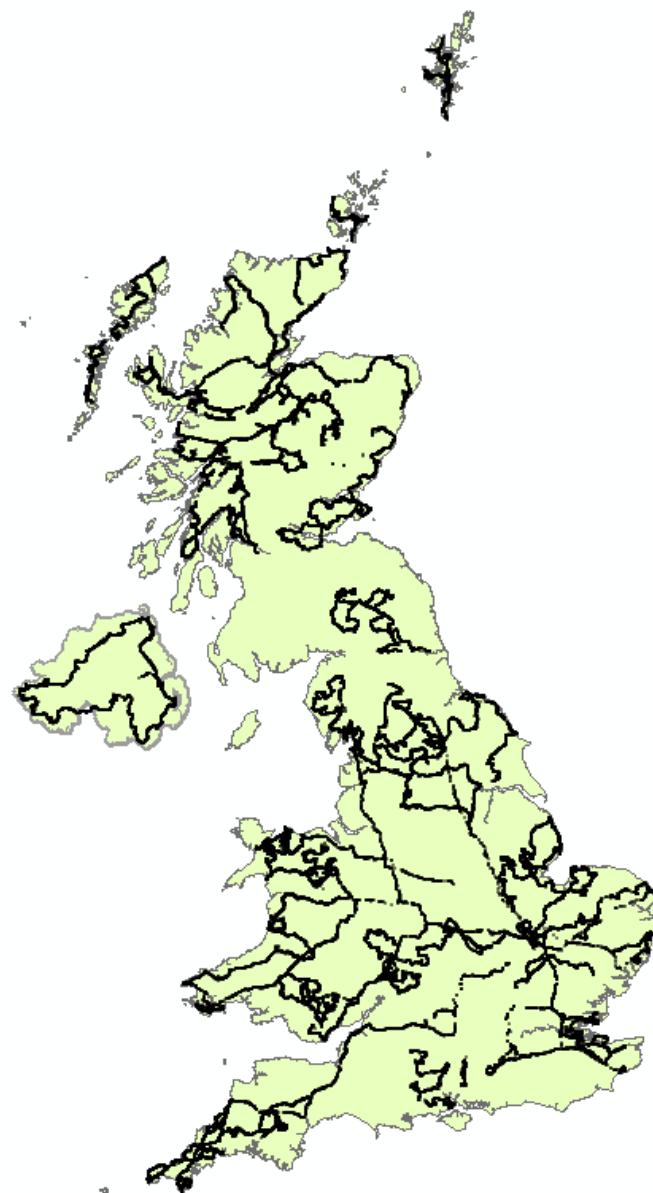
This section provides details about the additional data sets used and covers both internally and externally produced data sets. Two CEH data sets were used, the first being a set of ground reference points collected by CEH between 2006 and 2008 specifically for the purposes of training and validating LCM2007. The second CEH data set was the Countryside Survey Broad Habitat areas data set, surveyed in 2007. The external data sets included national cartographic products for GB and NI, the Digital Elevation Models (DEMs) needed for the altitude-based corrections and various soil data sets.

CEH produced data sets - Ground reference points

A series of field trips were undertaken to collect a set of ground reference points for the purposes of training and validating the land cover classifications. The majority of the field trips took place in 2007, as this was the target summer for LCM2007 image acquisitions. The ground reference points were collected using a tablet PC-based Digital Data collection system developed for CS in 2007. The tablet was connected to a GPS and had a bespoke GIS package which was automatically updated with the tablet/car location, enabling users to easily record land cover at points along the field trip routes (Figure 3.5). Data on LCM2007 land cover classes (as listed in Table 2.1) and location were visually identified and recorded from a vehicle as it travelled along the routes shown in Figure 3.6. Some points were recorded on foot and from view-points when vehicular access was not possible. The ground reference points were distributed widely across the UK and routes were carefully selected to visit rarer land cover types.

Figure 3.5. Example of map area, with satellite image overlaid with gridlines. a) Pink squares show location of field data points; b) yellow chevron shows GPS position, with arrow showing direction of movement; c) large orange pointer shows current field data point. © NERC (CEH) 2011.

Figure 3.6. Distribution of ground reference points collected in the LCM2007 ground survey. Ground reference points are displayed as points, but merge into lines, due to the density of data. © NERC (CEH) 2011. © Contains Ordnance Survey data © Crown Copyright 2009.

Countryside Survey in 2007 data

The Countryside Survey Broad Habitat data collected by CEH in 2007 were used to compare between CS in 2007 and LCM2007 (Chapter 4). Countryside Survey is a unique study of the natural resources of the UK's countryside and has been carried out in 1978, 1984, 1990, 1998 and 2007. Countryside Survey uses a stratified sampling approach which samples one-kilometre squares randomly located within different Institute of Terrestrial Ecology (ITE) Land Classes (Bunce *et al.*, 1996). The stratification by ITE Land Class ensures that survey squares are distributed across all landscape types found in the UK, with each stratum having similar environmental characteristics such as climate and geology. In 2007 the field survey covered a total of 591 1km sample squares spread across England, Scotland and Wales. Areas of habitat were mapped within each square and more detailed samples were made of

vegetation in a series of plots. The full data set collected in Countryside Survey enables national estimates of: the area and the change in area of habitats; the changing condition of vegetation; the changing condition of soils, freshwaters and ponds. However, for the purposes of comparison with LCM2007 the key aspect of the data is the Broad Habitat map of each 1km square.

External datasets (excluding satellite imagery)

Non-CEH data sets were needed to create the spatial framework, inform the knowledge-based enhancements (KBE) (see Section 3.7 for full details) and, in the case of the DEM data set, as input into the raster pre-processing stages. The external data sets used in the production of LCM2007 are summarised in Table 3.3 (the situation was complicated by slightly different products being available for the different countries of the UK).

Table 3.3: Summary of external (non-CEH) data sets used in the production of LCM2007.
[1]*SGRPID is the Scottish Government, Rural Payments and Inspections Directorate*

	England	**Wales**	**Scotland**	**NI**	**Use**
Satellite Imagery	Landsat-TM5 from Eurimage; IRS-LISS3, SPOT-4 and SPOT-5 sensors from European Space Agency; AWIFS from Euromap				*Image classification*
National cartography	Ordnance Survey MasterMap topography layer			OSNI 1:1250 and 1:2500 Large-scale Vector from Land & Property Services	*Spatial framework*
Agricultural census boundaries	Rural Payments Agency, Rural Land Registry.	Welsh Assembly Government, Land Parcel Identification System	SGRPID[1] IACS boundaries from Scottish Government	Not used	*Spatial framework*
DEM	NEXTMap Britain from Intermap Technologies Inc.			Digital Elevation Model as used in LCM2000	*KBE, plus raster pre-processing*
Soils	Soilscapes, from National Soils Resource Institute (NRSI), Cranfield University		SSKIB derived pH for semi-natural soils for upper horizon for dominant soil, and LCS88 Land Cover Map for Scotland from The James Hutton Institute	Soil classification map from Agri-Food and Biosciences Institute	*KBE*
Urban area	Urban area and settlement boundaries from Office for National Statistics		Urban Rural Classification 2003-2004 from Scottish Government	Statistical Classification and Delineation of Settlements of 2005 from Northern Ireland Statistics and Research Agency	*KBE*

National cartography

The Ordnance Survey MasterMap topography layer (OSMM) provides a highly detailed view of Great Britain's landscape, including individual buildings, roads, and areas of land. In total it contains over 400 million individual features. These data were used to produce the spatial framework for GB.

Land & Property Services (LPS) is the mapping agency for Northern Ireland. It was formed by a merger between Ordnance Survey Northern Ireland (OSNI) and Land Registry for Northern Ireland. LPS's Large-scale Vector database is a topographic coverage of Northern Ireland comprised of over 500 individual tiles with map features represented as points, lines and polygons. LPS Large-scale Vector was used to produce the spatial framework for Northern Ireland.

The OSMM and LPS (OSNI) datasets are discussed in more detail in Sections 3.3 and 3.4 respectively.

Agricultural census boundaries

Agricultural agencies in devolved governments maintain up-to-date digital maps of agricultural land, along with woodland and marginal land on which grants or subsidies can be claimed (Table 3.3). These data contain boundaries that are often too new to have been incorporated into UK inventories maintained by the OS and LPS, or where no physical boundary exists for these agencies to map but where fields have been subdivided for distinct agricultural use. Incorporating these boundaries into the framework of LCM2007 improves the delineation of land use and therefore helps to clarify spectral properties of the land surface (see Section 3.3) leading to more accurate land cover classification.

Urban areas

Urban area outlines for larger urban areas in England and Wales were provided by the Office for National Statistics (ONS). The urban area and settlement boundary dataset was created as part of the 2001 Census (ONS, 2004). For Scotland, the Urban-Rural Classification 2003-2004 was provided by the Scottish Government. For Northern Ireland, the Statistical Classification and Delineation of Settlements of 2005 was provided by the Northern Ireland Statistics and Research Agency (NISRA). Additionally, for Northern Ireland some urban areas were manually digitised using satellite imagery data.

Soils

For England and Wales, Soilscapes (at 1:250,000) was provided by Cranfield University (NSRI). The Agri-Food and Bioscience Institute provided a bespoke Soilscape-like product for Northern Ireland (Jordan & Higgins, 2009). The soils information for Scotland came from a bespoke product derived by the Macaulay Land Use Research Institute (now the James Hutton Institute) from 1:250,000 soil maps and the Land Cover of Scotland (LCS88) dataset.

Digital Elevation Model

NEXTMap Britain is a Digital Elevation Model (DEM) for Great Britain created by Intermap Technologies Inc. based on airborne radar data collected during 2002 and 2003.

3.3 Development of spatial frameworks: Great Britain

The spatial framework aims to delineate roughly homogenous groups of pixels (spectral regions) and in doing so optimise the application of the maximum likelihood classifier. The LCM2007 spatial framework for Great Britain is based on the OS MasterMap topography layer (OSMM). The topography layer contains over 450 million individual features representing real-world objects such as fields, roads, buildings (polygons), overhead power lines, overhead phone lines (lines), letter boxes, phone boxes (points) as well as intangible objects such as administrative boundaries. LCM2007 uses only the real-world polygon objects. There are over 100 million real-world polygons and these give complete and continuous coverage of Great Britain.

Minimum Mappable Unit and Minimum Feature Width

OSMM is highly detailed and accurate. Urban areas and rural towns have been surveyed to 1.0m accuracy against the British National Grid. Rural areas have been surveyed to 2.5m accuracy. Consequently, OSMM contains millions of polygons with an extent of just a few square metres. Contrast this with the 20m x 20m spatial resolution of resampled satellite data used for LCM2007 (Figure 3.7). The maximum likelihood object-based spectral classification (Section 3.6) uses the average spectral response within a land parcel to assign a land cover class. A minimum of 4 whole pixels is required to achieve a reliable spectral signature, so clearly there is far too much detail in the OSMM. This is especially evident in urban areas where a single pixel may cover many real-world objects. A land parcel specification with a minimum mappable unit (MMU) of 0.5ha and minimum feature width (MFW) of 20m was specified for LCM2007. This is considered the maximum spatial resolution of land cover that can be achieved from 20m x 20m pixel satellite data.

Generalisation of OS MasterMap topography layer

Spatial generalisation is a process in which unnecessary detail is removed but relevant detail is retained. OSMM was generalised to satisfy the MMU and MFW required for LCM2007. The generalisation involved a series of iterative steps each simplifying the spatial structure by selectively splitting and merging polygons. This process terminated when the rate of simplification relative to the cost of further processing diminished beyond a critical value. Because of the huge volume of data the generalisation of OSMM was computationally intense and required bespoke software written for a 200 node computer-cluster. The density of polygons was reduced to less than 6% of the original.
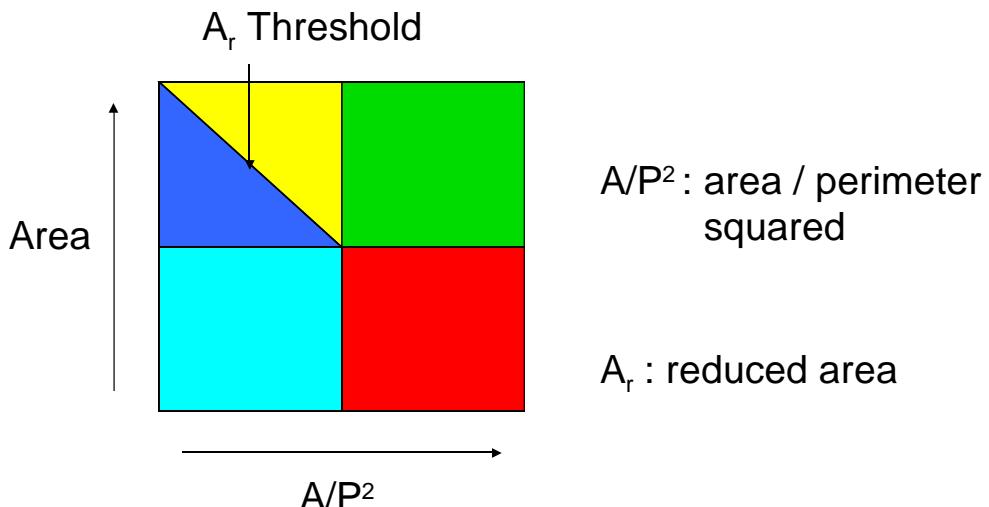
Figure 3.7. OS MasterMap topography layer spatial structure against 20m x 20m satellite data.
© NERC (CEH) 2011. © Crown Copyright 2007, Ordnance Survey Licence Number 100017572.
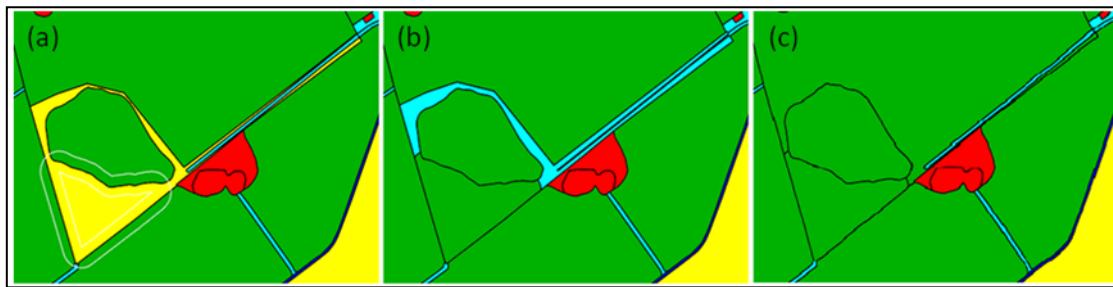
Splitting and Merging

Splitting and merging represents the core of the spatial generalisation. To support splitting and merging polygons were colour coded according to geometric complexity (Figure 3.8). Red polygons are too small (less than the MMU); cyan are too small and too narrow (smaller than the MMU and narrower than the MFW); blue satisfy the parcel criteria for LCM2007 but are elongated; yellow satisfy parcel criteria but have complex boundaries; green polygons represent the ideal, they are broad and have simple boundaries. The objective of the spatial generalisation is to maximise the number of green polygons and eliminate all red and cyan polygons. Blue and yellow polygons are acceptable but are simplified when possible.

The splitting process begins by buffering inwards by 10m. This removes the narrow features of a polygon. Buffering out from this simplified polygon by 20m creates a 'cookie-cutter', which is used to split the original (Figure 3.9a). The resultant pieces are then reclassified as red, cyan, yellow, blue or green (Figure 3.9b) ready for merging. Polygons that are too small or too thin are merged with their most appropriate neighbour (Figure 3.9c).
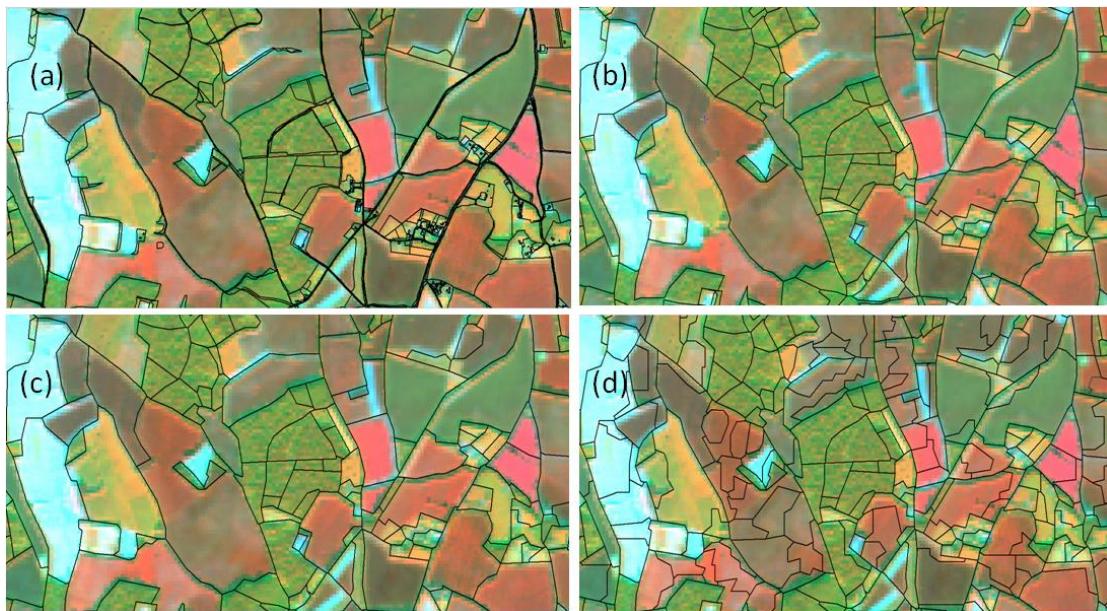
Figure 3.8. Colour classification of polygons based on area and perimeter relationships. *A* is the area, *P* is the perimeter. *Ar* is the reduced area. Reduced area is obtained by buffering inwards by   .

Figure 3.9. Splitting and merging sequence. (a) Shows the original structure and buffering inwards then outwards from the yellow parcel. This creates a 'cookie-cutter' which is used to fragment the yellow parcel giving (b). A merge step (c) merges the fragments and simplifies the spatial structure. © NERC (CEH) 2011. Crown Copyright 2007. Ordnance Survey Licence number 100017572.

An objective when merging is to minimise the spectral heterogeneity in merged parcels. For example, it would be inappropriate to merge an urban parcel with an adjacent lake. The two objects will be spectrally distinct and mixed parcels classify poorly. To guide the merging process therefore attribute data associated to the OSMM land parcels was used. The *descriptiveGroup* attribute assigns a land parcel object to one of 21 groups; for example *Path, Building, Natural Environment*, and so forth. The *descriptiveTerm* refines this; for example *descriptiveGroup = Natural Environment, descriptiveTerm = Rough Grassland*. The *make* attribute specifies whether a feature is man-made or natural. Using this attribute information a concept of spectral similarity was derived. When merging, polygons were merged with those most likely to have similar spectral characteristics.

A typical result of OSMM generalisation is given in Figure 3.10a and 3.10b. Figure 3.10a shows the detail rich OSMM. The generalised result in Figure 3.10b has retained the salient features of the OSMM, but narrow and small features (such as narrow roads, streams, tracks and farmsteads) that violate the MMU and MFW have been removed.

Figure 3.10. Stages in the production of the spatial framework for Great Britain. (a) The detailed line work from OS MasterMap. (b) Generalised OSMM. (c) Generalised OSMM after the inclusion of agricultural boundaries. (d) Generalised OSMM integrated with agricultural boundaries and image segmentations. © NERC (CEH) 2011. © Crown Copyright 2007, Ordnance Survey Licence number 100017572. © third-party licensors.

Integration of agricultural boundaries

Generalised OSMM works very well in urban areas where changes in land cover are invariably separated by some form of boundary mapped by OS. In the agricultural landscape the situation is sometimes less clear cut. Pasture farming requires that boundaries are stock proof and OSMM works well here. It works less well in extensively arable areas, where land cover boundaries are frequently missing. The missing boundaries generally relate to different farming practices being applied to different parts of the same field where no physical boundary exists for the OS to map.

To receive grants from government schemes UK farmers are required to map and submit land parcels outlining the extent of different agricultural land use. These data are maintained by government agencies (Table 3.3) and were used for the second stage of the spatial framework construction. The data resolve many of the missing boundary issues encountered in agricultural landscapes and were merged with the OSMM spatial framework (Figure 3.10c). This increases the number of objects in the spatial framework (Table 3.4). The extra objects account for some of the variability not picked up by the OSMM and in doing so help to improve classification results.

Integration of Image Segments

Within a field there may be spatial differences in fertility due to the effects of shading, aspect, flooding, slope or uneven application of fertilizer leading to spectral variability within a crop. In semi-natural areas, for example in upland areas above land enclosed for arable or pasture the heterogeneity of semi-natural vegetation will lead to spectrally distinct regions than have not been mapped by the OS. In these cases the spatial framework derived from OSMM and agricultural boundary data is insufficient to distinguish spectral regions.
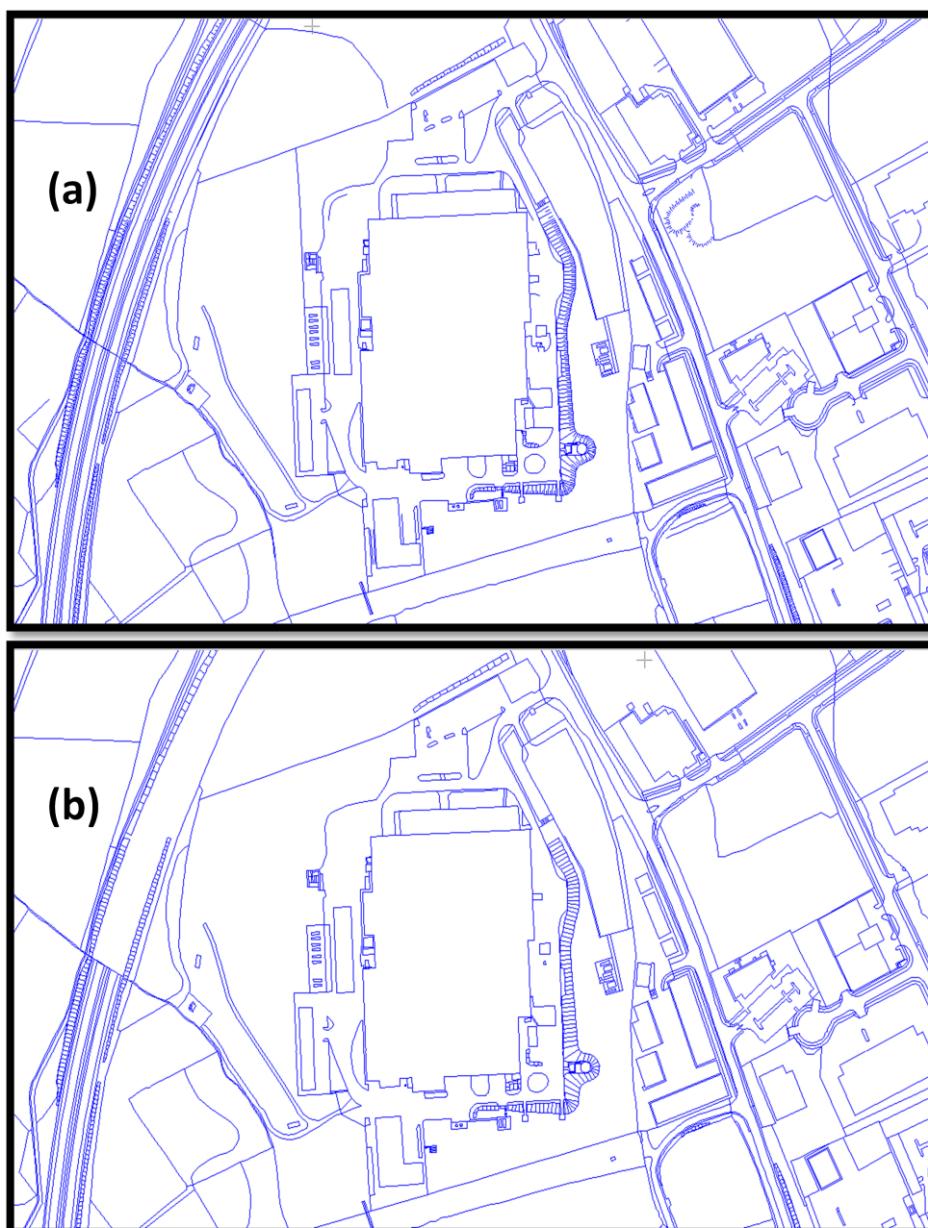
Image segmentation refers to the partitioning of a digital image into multiple segments. A label is assigned to every pixel in an image such that pixels with the same label share certain visual characteristics. Clusters of pixels with the same label identify relatively homogenous areas, image segments. Image segments were derived from satellite images (see Section 3.5). The third and final stage of the spatial framework construction involved the integration of these segments with the land parcels derived from OSMM and agricultural boundary data. The increase in parcel density from this integration is given in Table 3.4. Contrasting Figure 3.10c with 3.10d gives a typical result of this third stage; clearly the segmentation has delineated spectral regions not picked up by OSMM and agricultural boundary data.

Table 3.4. Polygon size frequencies for different stages in the production of the spatial framework for Great Britain.

Polygon size (Sq km)	Total number of OSMM polygons	Number of polygons after generalisation	Number of polygons after inclusion of agricultural boundary data	Number of polygons after image segmentation
< 0.2	105079479	6104041	6624252	8579292
0.2 – 1	108529	116993	98555	54142
1 – 5	10030	10248	9077	3002
5 – 20	866	825	731	131
20 – 100	82	57	34	3
100 – 600	2	0	0	0
Total no. polygons	105198988	6232164	6732649	8636570
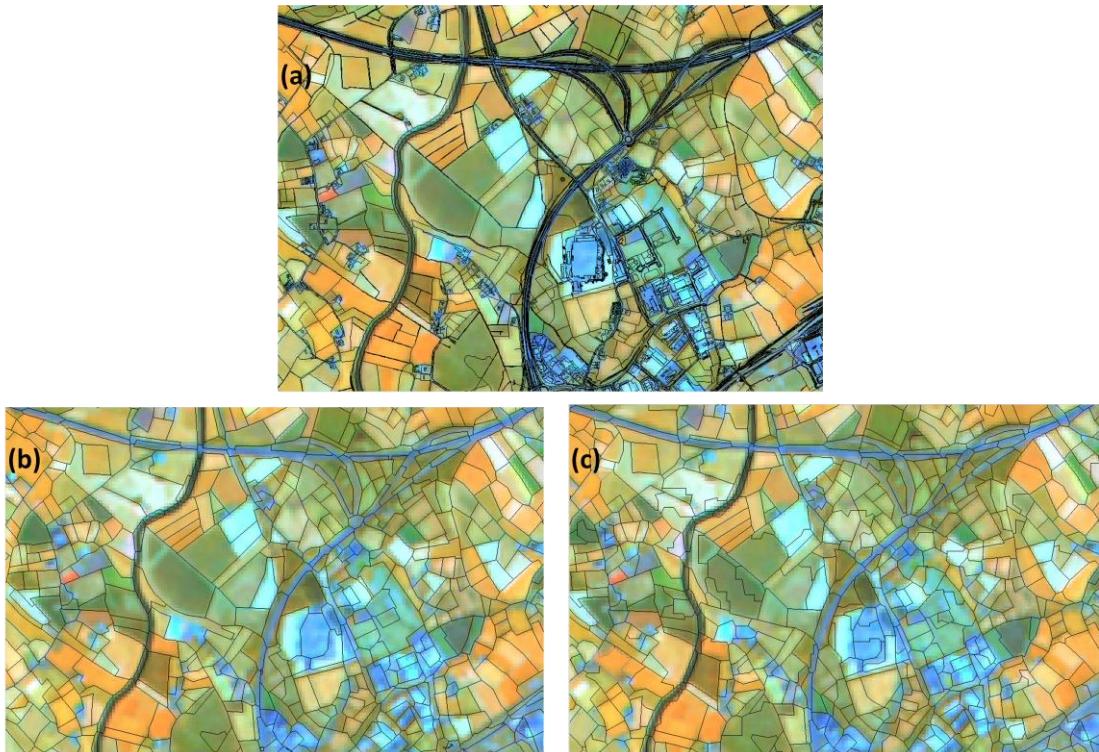
3.4 Development of spatial frameworks: Northern Ireland

The Land & Property Services (LPS) does not maintain an equivalent of OS MasterMap. LPS's Large-scale Vector digital database is a representation of topographic and cartographic map features comprising lines, points and polygons. The polygon coverage is restricted and represents a small percentage of Northern Ireland's land surface. It was therefore necessary to create a polygon coverage from the LPS's Large-scale Vector line work. The polygonisation process retains all lines that enclose an area and dangling lines are discarded. The enclosed areas become the polygons. Figure 3.11a shows a sample of the original LPS line work and Figure 3.11b the resulting polygon coverage.

Figure 3.11. Production of a polygon coverage for Northern Ireland. (a) LPS Large-scale Vector. (b) A polygon coverage created from (a); note that lines not enclosing an area have been removed. © NERC (CEH) 2011. © Crown Copyright 2011, Licence Number 100,427.

Polygons created from the LPS line work have no attribution to support the merging process. The possible consequence is a higher proportion of inappropriate merges which could lead to mixed parcels and therefore poor classifications. In practice however it was found that the subsequent inclusion of lines from image segments resolved the majority of problems of this kind. Figure 3.12 shows the LPS Large-scale Vector (a) and generalised LPS Large-scale Vector before (b) and after (c) the inclusion of image segments.

Figure 3.12. Stages in the production of the spatial framework for Northern Ireland. (a) Detailed line work of LPS Large-scale Vector. (b) Generalised LPS Large-scale Vector coverage. (c) Generalised coverage integrated with image segmentation. © NERC (CEH) 2011. © Crown Copyright 2011, Licence Number 100,427. © third-party licensors.

Agricultural boundary data was not integrated into the NI spatial framework because preliminary work concluded that the effort to do so was disproportionate to the potential spatial refinement benefit. Agricultural datasets are most useful in intensive arable areas where multi-crops are grown within a single large field. Arable production in Northern Ireland is less extensive and intensive than many areas of the UK mainland and field sizes are generally smaller. The exclusion therefore did not have a detrimental effect on classification as spectral mixing within land parcels was well resolved by the inclusion of image segments. In other respects the construction of the spatial framework for Northern Ireland is analogous to that used for Great Britain. Table 3.5 shows the size frequency of polygons at key stages of the spatial framework production.

Generalisation of polygons derived from the LPS Large-scale Vector gave a reduction to approximately 10% of the original polygon density. This is a smaller reduction than was achieved for Great Britain (~6%). The most likely explanation is that Northern Ireland has proportionally less urbanisation than Great Britain and it is in urban areas where the reduction in polygon density from generalisation is

greatest. In dense urban areas generalisation typically achieves reductions to around 1 to 2% of the original density.

Table 3.5. Polygon size frequencies for different stages in the production of the spatial framework for Northern Ireland.
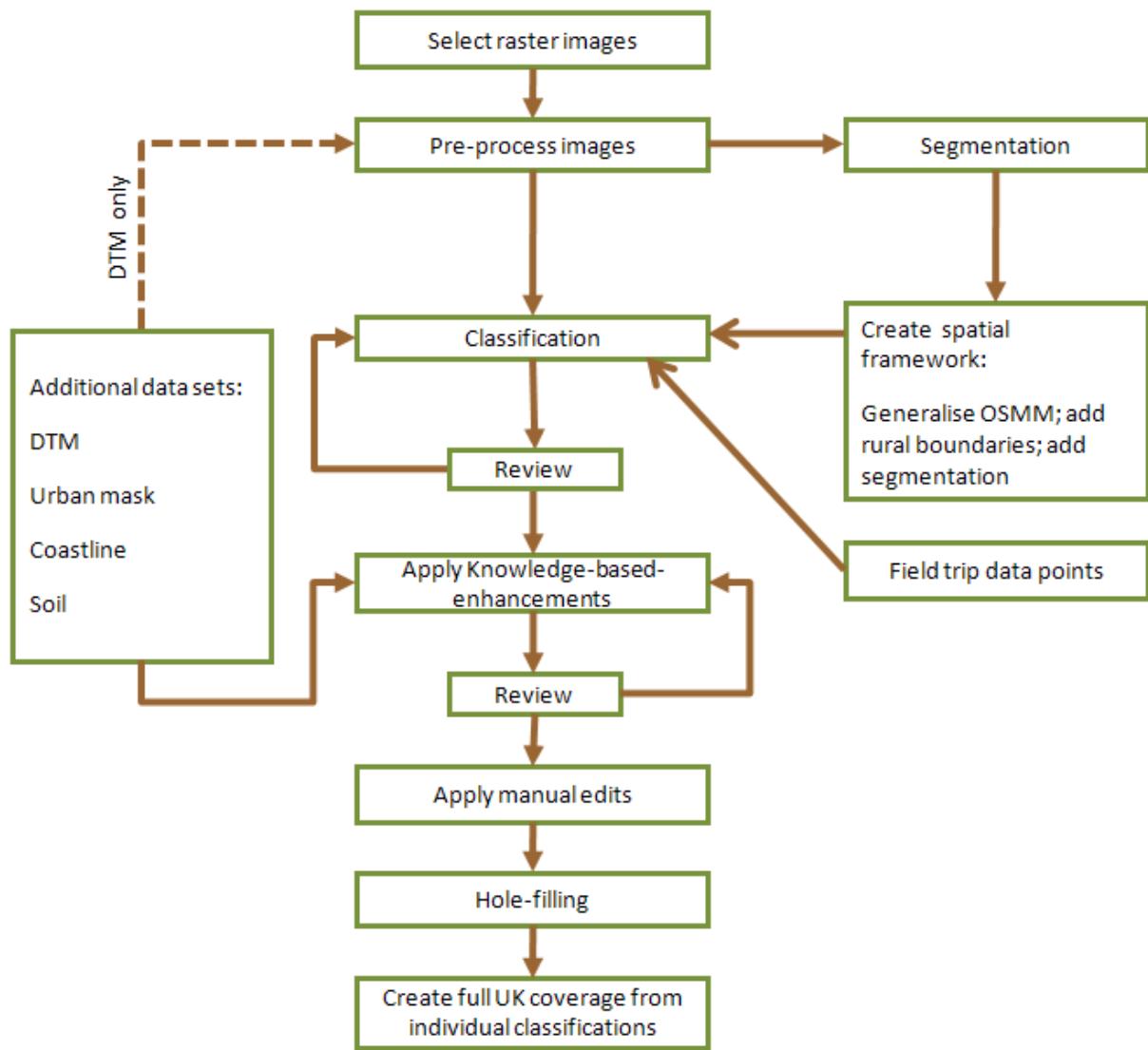
Polygon size (Sq km)	Number of polygons derived from LPS Large-scale Vector	Number of polygons after generalisation	Number of polygons after image segmentation
< 0.2	6796591	547423	737345
0.2 – 1	147588	174685	158916
1 – 5	4724	5297	3589
5 – 20	1989	2072	1244
20 – 100	225	217	43
100 – 600	22	17	15
Total no. polygons	6951139	729710	902299

3.5 Image processing and segmentation

This section of the report describes the raster processing chain for LCM2007 and follows the order in which processes (shown down the centre of Figure 3.13) were conducted.

Selection of raster images

The key criteria for selecting images were the suitability of the sensor (20-30m pixel resolution, plus red, NIR and MIR bands) and appropriate acquisition time. The images used to produce LCM2007 come from Landsat-TM5, LISS-3 and SPOT4/5, with 60m AWIFS data as a backup as described in Section 3.1. 101 images were selected and pre-processed, of these 6 were rejected at some point during processing. Failure was mainly due to contamination by haze and/or cloud, or problems geocorrecting images that had been poorly orthorectified. Not all the images processed were used, as some images were superseded by more appropriate images of the same area. Table 1 of Appendix 2 contains details of all the images processed.

Figure 3.13. Flowchart showing the key stages in the image processing of LCM2007.

Pre-processing of image data

The purpose of image pre-processing is to maximise the spatial and spectral accuracy of the data set. Good pre-processing is a pre-requisite for accurate classification and for multiple image-based, multi-sensor mapping it must be robust and repeatable (Franklin & Wulder, 2002). The pre-processing steps carried out for LCM2007 are similar to those carried out for LCM2000 (Fuller *et al.*, 2002) and to those used in other large-scale land cover mapping (Franklin & Wulder, 2002). The pre-processing flow line for LCM2007 was:

- Import data
- Apply cloud and cloud-shadow masking
- Atmospheric correction
- Geo-registration
- Topographic correction
- Creation of composite images

To ensure consistency a handbook (Bradley *et al.,* 2009) was created to define the pre-processing methods and these were followed by all staff involved in the pre-processing stage. A log was kept for each image documenting the image quality and image specific parameters ensuring traceability and repeatability.

Pre-processing began by importing the data, after which the cloud and cloud-shadow masking was conducted via three stages: firstly, identification of the clouds; secondly, identification of the cloud-shadow; and thirdly by masking the affected areas. Identification of the cloud was conducted by one of four methods (sometimes applied in combination):

- Identifying cloud reflectance on a scatterplot
- Thresholding based on reflectance
- Thresholding based on band ratio
- Manual digitisation

The first three methods require a threshold to be defined, which is then used to create a mask, whilst the fourth method requires the user to digitise round the areas of cloud (and cloud-shadow). The primary method for masking cloud-shadow was by modelling the expected shadow areas, based on the cloud-mask, illumination geometry and shadow-offsets in the x- and y- direction for each image. The cloud-shadow masking worked well when the clouds were of relatively uniform height across the image, but sometimes needed supplementing with manual masking if cloud height varied. A small number of images were deemed to have 'failed' at this stage as the masking showed that the cloud/haze problems were worse than originally suspected.

Atmospheric correction was applied in ENVI using the FLAASH atmospheric correction function, which is based on MODTRAN. The images were converted from digital number to radiance using sensor-specific conversion factors, after which FLAASH was run, producing atmospherically corrected reflectance.

Most of the images were supplied geo-referenced, but due to multiple image suppliers and sensors the geometric reliability varied, so all images were geo-registered to either existing images (typically geo-registered images from LCM2000) or to OS data. The aim was to achieve a root-mean-square-error (RMSE) of < 0.3 pixels. Several images failed at this processing stage, possibly due to poor orthorectification before supply to CEH.

Topographic correction is the process of adjusting the reflectance of an image to take into account variations in slope angle and aspect, so that the same land cover on the north and south slopes of a hill will look similar. The topographic correction code was the same as LCM2000, but ported into an ENVI_IDL compatible format (Bradley *et al.*, 2009). The topographic correction code uses the Minneart model, with the calculation of K means, to correct the radiance values. It used slope, aspect and shaded relief, calculated from NEXTMap Britain and the NI DEM, to correct the reflectance for topography. There were problems with the correction in winter images, with very low sun angles and significant topography, and this resulted in black patches (very low data values) on the north side of hills and mountains. In

many cases this could be avoided by taking alternative images i.e. composite 26 (c26) was taken in preference to composite 4 (c04), which covered the same area of the Lake District, as c26 had a higher winter sun angle producing less topographic shadow. However, in some cases manual editing was required to correct the classification in areas of topographic shadow.

Segmentation

All 60 composite images were segmented, using the method developed for LCM2000. Segmentation results were combined to create a segment-based vector framework for the UK. For full details of the segmentation method see Appendix V of Fuller *et al.,* (2002). The segments were created using 3-band images typically bands 5, 3 and 1 (winter NIR, summer MIR and summer red). The segmentation routine was limited to three input bands, so winter NIR was chosen as the brightest winter band, with red and MIR from the summer for the remaining two bands. In this way, the three wavebands and two dates of the composite are represented in the input to the segmentation. The segmentation, using the same method as LCM2000 (Fuller *et al.*, 2002), involved three stages:

- Application of an edge detector to locate edges in the raster data.
- Segmentation based on seed points located away from edges.
- Post-segmentation generalisation, including dissolving small polygons to produce an acceptable product.

The segmented framework was integrated with the generalised digital cartography from OSMM and LPS Large-scale Vector (Section 3.3 and 3.4). For NI segments from single-date images were used to supplement the segments derived from composite images.

3.6 Image classification

LCM2007 is a parcel-based supervised maximum likelihood classification of the UK predominantly based on medium resolution (20-30m) satellite data. It was produced by classifying 37 composite images and 21 single-date images according to Broad Habitat-based classes (Table 2.1). Supervised classification uses a training data set of observed spectral signatures and known land cover to determine the land cover of unknown parcels. Training areas for the classification were chosen based on the ground reference points collected during the field trips (Section 3.2). Training polygons were selected on the basis, of a) being identified by ground reference points and b) polygon suitability for use as a reference polygon for spectral classification. For areas and/or habitats poorly represented by ground reference points OS maps and web-based mapping services were consulted.

Once the training areas had been selected, they were reviewed and similar spectral signatures corresponding to the same land cover type were grouped into spectral classes after which the classification was run (see Fuller *et al.,* 2002 for additional details). The classification used a maximum likelihood classifier (an automated process) that assigns class probability based on probability contours around the training areas (Allaby & Allaby, 1999). In LCM2007 the probabilities of the top five spectral classes (i.e. those that matched the observed spectral signature for the polygon most closely) were recorded. The process of classification is iterative, with

each classification being reviewed and re-classified until the accuracy reaches the required level, or until the classification can no longer be improved. The review stage would typically identify classes which were not being classified well enough, so additional training areas would be identified and/or the grouping of the spectral classes would be modified, after which the classification would be re-run.

In practice, for one of the more widespread Broad Habitats in a medium sized satellite scene there may be between 5 and 15 spectral training classes, each of which might have up to 15 training polygons. For a heavily arable area more than 30 classes may be needed to capture the spectral signatures of all the fields. These figures vary dramatically between images depending on the Broad Habitats within the scene and the range of spectral variability within the Broad Habitats. In some polygons there is a mix of more than one habitat – this is particularly true for the more semi-natural, less managed areas and in these cases the aim is to identify the most widespread Broad Habitat.

Apply manual edits

Sometimes it was necessary to manually edit the classification. These edits typically occurred when a rare land cover class was located in an image. Trying to classify the rare class sometimes creates uncertainty with other classes and unsatisfactory mapping. An example, of this might be a quarry, which is classed as urban despoiled and sits in the '*Inland Bare'* Broad Habitat. Attempting to classify based purely on spectral signature might lead to confusion with urban land cover, which falls into the '*Built-up Areas and Gardens'* Broad Habitat. Some winter images with problems due to the topographic correction needed manual correction. The decision on class for a manually classified polygon was informed by a variety of sources, including the satellite images (or overlapping images), OS maps and web-based map services.

Hole-filling

Landsat7-ETM+ data were used to manually classify small areas of some ETM+ scenes, when no other imagery was available. This avoided the need to fill holes using LCM2000 data. 12 ETM+ images spanning 22/01/2006 - 30/05/2009 were processed and available for manual classification (see Table 3 of Appendix 2). Manual classification was conducted by overlaying the spatial framework on the raster data and manually selecting all the polygons of each Broad Habitat. It was a pragmatic solution given the problems with ETM+, plus the lack of other suitable data in an appropriate timeframe.

3.7 Knowledge-based enhancements

Knowledge-based enhancements (KBEs) resolve spectral confusion and/or increase the thematic resolution of land cover using contextual and ancillary information. They comprise regionally adaptive rules that reassign land parcels to a more appropriate land cover class and therefore enhance the accuracy of LCM2007. KBE's modified approximately 20% of LCM2007 land parcels.

Spectral confusion

LCM2007 land parcels are classified using red, near and mid infra-red spectral bands. These bands offer the best discriminatory power for vegetated surfaces (Fuller and Parsell, 1990). Spectral confusion occurs when different surface types have similar reflectance properties. For example, an arable field that has been recently ploughed has little or no chlorophyll to absorb red light and will appear bright in this band. Likewise so will a road, a car park, an open-cast quarry, a sandy or shingle beach, and a limestone pavement. If these surfaces have similar reflectance qualities for mid and near infra red light it becomes very difficult to distinguish between them using reflectance alone. Using image composites from different stages in the growing season (Section 3.6) can help to resolve some of this confusion. For example, if an arable land parcel in a winter scene displays a ploughed field and the summer scene displays crops, the composite information will be sufficient to distinguish this parcel from those that are permanently non-vegetated. However, if the summer scene occurs after harvest when the vegetation has been removed the composite information may be of little or no help.

The above gives just one example of a common spectral confusion encountered in the production of LCM2007, but spectral confusion occurs between many land cover types. To detect spectral confusion we examine a land parcel against a knowledge base comprised from contextual and ancillary data. If confusion is suspected knowledge-based enhancement rules (KBE rules) are used to assign a revised land cover class. Taking the arable example further, suppose we detect a land parcel that has been classified as urban, it is shaped like a field, is surrounded by arable fields and exists far from known urban boundaries. It is quite likely that this is an arable field that has been misclassified because at the time of image capture it had a similar reflectance to nearby urban parcels. A KBE rule should therefore pick this up and reassign it to an arable class. Of course there is always the risk of the 'false-positive'. The land parcel may be an out of town industrial area or an open-cast quarry, in which case the reassignment to arable would be inappropriate. In fact it is inevitable that some reassignments are wrong, but the purpose of the KBEs is to 'get it right' most of the time and thereby improve the overall accuracy of the product. To minimise reassignment errors we suppress KBEs in regions where we expect they will be inappropriate. We also include manual inspections and corrections (Section 3.6).

Thematic resolution

KBEs are also used to enhance thematic resolution. LCM2007 aims to provide land cover as target classes related to Biodiversity Action Plan (BAP) Broad Habitats (Table 2.1). However, the information that can be gleaned from optical satellites is frequently insufficient for this. Put simply, some Broad Habitats cannot be determined from optical imagery. UK BAP grassland habitats offer a good example. The UK BAP specifies four grassland Broad Habitats: *'Acid Grassland', 'Calcareous Grassland', 'Improved Grassland'* and *'Neutral Grassland'. Acid, Calcareous and Neutral grasslands* are considered semi-natural and are characterised by their species composition. However, these compositions do not each have a distinct reflectance at the spatial resolutions of satellite data used for LCM2007 and cannot be separated. In fact, using reflectance we can only reliably distinguish improved

from rough grassland (relatively unimproved), and these lie on a continuum with confusion near the middle.

At its extreme improved grassland has typically been modified by extensive fertiliser use and reseeding. It is chlorophyll rich all year round and contains an insignificant amount of dead material. This gives it a very characteristic signal. Conversely at any time of year rough grassland may have a significant amount of dead material from previous seasons; this increases reflectance of the mid and near infra red bands. There is also less chlorophyll, so less photosynthesis and more red reflectance making it spectrally dissimilar from recently improved grassland.

Improved grassland maps directly to a single BAP Broad Habitat. Rough grassland does not. BAP semi-natural grasslands will in general fall within the rough grassland category. However, this represents a many-to-one relationship, so to resolve rough-grassland to a BAP Broad Habitat we need additional information. Suppose that a land parcel has been spectrally classified as rough grassland. After examining its relationship with the knowledge base it is found to be growing on an acid soil and at an altitude typically beyond the limit for improved grassland and arable (the most likely confusion classes). Given this information it becomes highly likely that the Broad Habitat cover is '*Acid Grassland'* and a KBE rule should reassign it.

The Knowledge Base
Ancillary data from a variety of sources was compiled and processed to form the knowledge base. Knowledge was grouped into four types, described below.

Urban context
We have discussed above how urban context can resolve misclassified arable. It can help with many other types of confusion too. For example, a parcel classified as *Inland Bare* in an urban context is more likely to be part of the dense urban fabric and is better represented by an urban class.

Government agencies maintain readily accessible spatial datasets which outline urban and settlement boundaries. Boundaries for England and Wales were obtained from the Office for National Statistics and relate to the 2001 census. Boundaries for Scotland were from the Scottish Executive Urban Rural Classification of 2003-2004. Boundaries for Northern Ireland were obtained from the Northern Ireland Statistics and Research Agency and come from the Statistical Classification and Delineation of Settlements of 2005. Where boundaries were considered insufficient additional urban areas were manually digitised.

Coastal proximity
Coastal proximity helps detect inappropriately labelled land parcels. For example, *Littoral Sediment* encountered 40 miles from coast has certainly been confused with another class; perhaps an urban or an arable land cover.

An outline of the UK Coastline was used to create three zones: the *coastal zone* in which coastal classes are permissible; the *offshore zone* in which terrestrial classes

are not; and the *terrestrial zone* in which coastal habitats cannot occur. Because of the UK's variable coastal topography these zones were defined per satellite scene.

Terrain

Terrain information has many uses. For example coastal classes will not be found at altitude. Parcels with a very steep gradient are unlikely to be '*Arable* or *Improved Grassland*' if the gradient is too steep for agricultural machinery. Urban land cover is very rare in the montane zone.

Terrain for Great Britain was derived from the NEXTMap Britain Digital Elevation Model. For Northern Ireland the terrain model created for LCM2000 was used. For each parcel we recorded altitude, slope and aspect.

Soil Type

Soil type is a key arbiter of land cover type, especially when combined with other contextual information. For example, **Rough grassland** or **Heather grassland** on a bog soil on a shallow slope is quite likely to be *Bog*. Grassland at low altitude in the tidal flooding zone on a saltmarsh soil is quite likely to be **Saltmarsh**.

Soil information for the UK is maintained by distinct bodies. For England and Wales we used the National Soil Resources Institute (NSRI) Soilscapes product. For Scotland we used a bespoke dataset provided the Macaulay Land Use Research Institute (now James Hutton Institute) which combined soil series with land cover information from the Land Cover of Scotland (LCS88) product. The Agri-Food and Bioscience Institute (AFBI) of Northern Ireland provided a bespoke product with soil information given in the NSRI Soilscapes schema.

Soil data from these agencies were generalised into the following soil types: *Acid, Bog, Calcareous, Fen, Neutral, Saltmarsh, Sand dune, Water, and Other.* This level of detail was sufficient for KBEs and casting all the products into the same 'currency' made the development of general rules and algorithms more straightforward.

KBE Algorithms

Seven automated KBE algorithms were developed. The algorithms examine a parcel's context relative to the knowledge base and then apply KBE rules. KBE algorithms were applied in sequence with the results of each stage being the start point for the next. This sequential application optimises the integration of knowledge. A land parcel may go through a sequence of reassignments, ideally each time converging towards the 'truth'. For example, '*Littoral Sediment*' found too far from the coast may change to '*Inland Rock*' but then a subsequent algorithm determines that the parcel exists within an urban boundary so it then becomes '*Built-up Areas and Gardens*'. Records of the application of KBE algorithms are maintained against each parcel and are supplied with the LCM2007 vector product (Section 4.3).

Full details of the algorithms are beyond the scope of this text but an overview of each is given below. The order in which the algorithms are described represents the

order in which they were applied to classification scenes. After the first pass an additional run of the Terrain and Soil algorithms was applied to capture any remaining assignment errors. The general operation sequence of a KBE algorithm is:

- examine a land parcel against context,
- if its classification is considered inappropriate, inspect the five most likely land cover classes from the maximum likelihood classification result.
- If a suitable alternative is found within this list assign this land cover, else apply a specialised sequence of rules to assign a more realistic land cover.

Terrain algorithm

The terrain algorithm is straightforward but quickly picks up a lot of classification irregularities. It takes account of a land parcel's slope and altitude to assign a more appropriate class. For example, it assigns land cover at high altitudes (>800m for southern UK (south of Birmingham) and >600m for Northern UK) to '*Montane Habitats'. 'Arable and Horticulture'* parcels on very steep slopes would also be assigned to a more appropriate land cover, as would coastal classes at high altitude.

Soil algorithm

The soil algorithm is more complex. It integrates soil type with urban context, terrain and coastal proximity to refine or improve the thematic resolution of land cover. This algorithm is essential for separating '*Acid Grassland*', '*Calcareous Grassland*', and '*Neutral Grassland*' from **Rough grassland**. It is also helps with the classification of '*Bog*' and **Saltmarsh**, which are often unreliably, resolved using spectral data alone.

Terrestrial zone algorithm

Coastal habitats cannot occur in the terrestrial zone. Any parcels that have been classified as a coastal habitat that have not been resolved by the terrain algorithm are reassigned by this algorithm to a more appropriate class.

Offshore algorithm

This algorithm picks up classes that are not possible offshore. For example, grassland or arable parcels located offshore are highly likely to be wrong. **Saltmarsh** or '*Littoral Sediment*' are assigned as more likely alternatives.

Within urban boundary algorithm

This algorithm is applied to all parcels that lie within the urban boundary. It resolves common spectral confusions that occur with urban sealed surfaces. For example, urban parcels can have a similar spectral response to areas of naturally exposed bare rock, littoral sediments, and in some cases arable.   In reassigning a new land cover class the algorithm will also take account of the parcel's immediate neighbourhood and geometric properties to obtain the most likely alternative.

Outside urban boundary algorithm

This algorithm is the converse of the above. It checks for parcels assigned to an urban class that lie outside the urban boundary. The urban areas and settlement boundaries frequently do not cover the smaller villages, hamlets and isolated farmsteads that occur throughout the wider countryside. These structures obviously fall within the BAP '*Built-up Areas and Gardens'* Broad Habitat and so to simply reassign all urban parcels that occur outside of the urban context would be wrong.

Distinguishing misclassified urban from true urban is tricky. Local context and geometric properties are key. For example, small urban parcels (<1ha) with irregular boundaries adjacent to objects with similar qualities quite possibly are small settlements or farmsteads, so are ignored. Large urban parcels shaped like fields, surrounded arable fields are highly likely to be misclassified arable.

Water

This algorithm improves the thematic resolution of water. Water absorbs much of the red and infra red spectra and appears dark, except where it is particularly shallow or turbid, and is easily distinguished from the surrounding landscape. However, there is no spectral distinction between coastal and inland water. When a water parcel occurs in the offshore zone it is labelled as Sea water (Ws). When it interacts with the coastal zone or occurs inland the shape of the land parcel is used to categorise it. If it is elongate and coastal then it is labelled as estuarine (We) and if it is elongate then it is labelled as a river (Wr), otherwise it is labelled as a lake (Wl).

3.8 Assembling full UK coverage

LCM2007 is the first UK land cover map to be produced as a continuous vector product. Due to differences in spatial reference systems it is maintained as two spatial databases: one for GB and one for NI. LCM2000 like LCM2007 gives a complete UK vector coverage but it is not continuous. LCM2000 was produced as discrete 100km x 100km tiles. Image segmentation and classifications were performed on a per tile basis. As a consequence of this spatial and thematic inconsistencies occur at boundaries where tiles meet. In order to produce a continuous single layer coverage that exceeds the extent of an individual tile, users have to apply complex merging processes. By providing a continuous coverage LCM2007 overcomes this problem.

Producing a continuous coverage from the many composite and single-date classifications requires a strategy for dealing with overlapping regions and assigning priority to the best classifications. This was achieved by splitting the UK into nine relatively independent chunks (Figure 12, Appendix 3). The chunks were chosen to maximise the natural divisions in the data and the diagonal boundaries reflect the swathes of the satellite data. After the chunks had been defined, the priority of each classification was determined based on the quality of the final classification, with the best classifications having the highest priorities. Table 6 shows the final order of images according to chunk, with arrows showing how images that fall into two or more chunks are catered for. Chunks 6 and 7 are merged in Table 6, because of the high degree of overlap between them. In general, the composites produced the best classifications, but in some cases single-date hole-filling scenes produced better classifications than overlapping composites. This was because the single-date hole-filling images typically covered small areas with a limited range of both spectral and Broad Habitat variation, which minimises the opportunity for spectral confusion between classes resulting in a better classification.

Table 3.6. Order of priority for classifications across the 9 chunks. Note images beginning with 'C' are composite images, whilst those beginning with 'I' are single-date images (see Appendices 2 and 3 for further details). Arrows highlight priority of images appearing in more than one chunk.

Chunk 1	Chunk 2	Chunk 3	Chunk 4	Chunk 5	Chunks 6 and 7	Chunks 6 and 7 (cont.)	Chunk 8	Chunk 9
C36	C12	C19	C55	C27	C54	C24	C51	C47
C17	C01	C18	C22	C26	C52	C25	I90	C48
C16	C13	I23	I76	I21	C58	C50	I91	C49
C01	C38	I51	C20	C08	C40	I45	I93	I26
C37	C56	C20	C09	C04	C41	I39		I85
C29	I29	C09	C21	I76	I93	I36		I86
	I07			C50	C42	I101a		I87
	I21				I102	I79		I88
	C08				C43			I89
	C02				C57			

Figure 3.14. Location of the nine chunks that the UK was divided into to ease production of the continuous vector product (different colours represent different chunks). © NERC (CEH) 2011. © Based upon Ordnance Survey data © Crown Copyright 2009.

3.9 LCM2007 Quality Assurance

Validation and assessment is critical for informed use of any data product, however, for large-scale products it is impossible to do intensively. This section presents a validation of LCM2007 by assessing it against the ground reference data set collected during the field trips. The aim of the comparison is to assess the accuracy of LCM2007. The results are displayed in a correspondence matrix which is a table used to compare two classifications. The correspondence matrix shows where LCM2007 and the ground reference data set correspond and diverge.

Validation of LCM2007 with ground reference data

Selected ground reference points (Section 3.2) were used to validate the classification, both during the classification stage (in conjunction with visual checks) (Section 3.2), and after the final LCM2007 data set was created. This section reports the results of the comparison between the 9127 LCM2007 ground reference polygons and the final LCM2007 product (Table 3.7).

The field trip points were used to identify suitable training areas for the classification (i.e. a training data set), and also to identify suitable points for a ground reference data set against which to validate the product (i.e. a testing data set). The training and testing data set are separate groups of ground reference points. The accuracy of LCM2007 is defined by comparison against the ground reference (testing) data set. In some cases ground reference points and polygons were poorly aligned and in other cases the ground reference points were only appropriate for part of the polygon - this mainly occurred in natural or semi-natural areas, although it did affect some woodlands and fields. Ground reference points were only assigned as validation points if the person conducting the classification was confident that the ground reference point was appropriate for the polygon. This excludes polygons with mixed Broad Habitats (shown by multiple ground reference points within a single-polygon) or where ground reference points fell in adjacent polygons. Once a ground reference point has been accepted as appropriate, for a polygon in the LCM2007 project database, the polygon becomes a ground reference polygon.

Results

The correspondence, between LCM2007 and the 9127 ground reference polygons (Table 3.7), shows that the overall accuracy of LCM2007 is 83%. Information about the classification of individual classes is also given in the correspondence matrix (Table 3.7).

Table 3.7 is a correspondence matrix and is data rich, but due to the number of classes in LCM2007 it is not straightforward to interpret. Using '*Bog*' as an example, the Producer's accuracy (see glossary) quantifies how well areas mapped as '*Bog*' in LCM2007 match the ground reference polygons. Table 3.7 shows '*Bog*' is 93% accurate against the ground reference polygons. The User's accuracy (see glossary) gives the probability of a parcel of a given class being correctly classified. In the case of '*Improved Grassland*', the User's accuracy is 83% and the Producer's accuracy is 89%. Therefore, based on the User's accuracy of 83%, there is a 0.83 probability that a parcel classified as '*Improved Grassland*' will be correctly classified. The Producer's accuracy quantifies how well areas mapped as '*Improved Grassland*' in

LCM2007 match the ground reference polygons. The User's and Producer's accuracy can vary independently of each other, although accurate classes will have high values for both the User's and Producer's accuracies.

The LCM2007 classes are grouped below based on the User's and Producer's accuracy. Note that requiring both accuracy values to be above a threshold means that the classes are grouped by the lowest of their two accuracy values, so the location of '*Bog*' is determined by the User's accuracy (39%), rather than the Producer's accuracy (93%).

> **> 90% for User's accuracy and Producer's accuracy:**
  - *'Coniferous Woodland'*
  - *'Arable and Horticulture'*
  - **'***Littoral Rock'*

> **> 80% for User's accuracy and Producer's accuracy**
  - **Broadleaf woodland**
  - *'Improved Grassland'*
  - **Littoral Sediment**
  - **Urban**
  - **Suburban**
  - **Saltmarsh**
  - **Freshwater**
  - **'***Supra-littoral Rock'*

> **> 70% for User's accuracy and Producer's accuracy**
  - *'Fen, Marsh and Swamp'*
  - **Heather**
  - *'Supra-littoral Sediment'*

> **> 60% for User's accuracy and Producer's accuracy**
  - **Saltwater**

> **> 50% for User's accuracy and Producer's accuracy**
  - *'Calcareous Grassland'*
  - **Heather grassland**

> **< 50% for User's accuracy and Producer's accuracy**
  - **Rough grassland**
  - *'Acid Grassland'*
  - *'Neutral Grassland'*
  - *'Bog'*
  - *'Montane Habitats'*

Accuracy is class-specific, with classes using external data sets for knowledge-based enhancements tending to have the lowest accuracy. The problem with '*Calcareous Grassland*' and '*Neutral Grassland*' is largely due to spectral confusion with '*Improved Grassland*' and **Rough grassland***.* The confusion with **Rough grassland** is partly due to differences between what it is possible to discern in the field and what can be achieved by spectral classification, KBEs and a soil data set. The '*Montane Habitats*' designation is expected to be correct for areas mapped as '*Montane Habitats*', but may underestimate in some areas, especially NW Scotland.

In some cases it may be useful for user's to apply their own bespoke validation, especially where extensive work is planned based on a county or smaller subset of data (see Appendix 4).

Table 3.7. Accuracy of LCM2007 based on 9127 LCM2007 ground reference polygons. Green squares correspond at LCM2007 class-level.
Producer's accuracy = percentage of ground reference polygons classified correctly. *User's accuracy* = probability (expressed as a percentage) of a polygon of a particular class being correctly classified.

LCM2007 \ LCM Ground reference polygons	Broadleaved woodland	Coniferous woodland	Arable and horticulture	Improved grassland	Rough grassland	Neutral grassland	Calcareous grassland	Acid grassland	Fen, marsh and swamp	Heather	Heather grassland	Bog	Montane habitats	Inland rock	Saltwater	Freshwater	Supralittoral rock	Supra-littoral sediment	Littoral rock	Littoral sediment	Saltmarsh	Urban	Suburban	Sum of rows	LCM2007 correspondence: User's accuracy (%)
Broadleaved woodland	593	18	6	3	1	0	0	2	12	10	1	0	0	0	0	1	0	1	0	0	1	0	0	649	91
Coniferous woodland	44	1035	3	6	0	0	0	1	1	1	5	0	0	1	0	1	0	0	1	0	0	0	0	1099	94
Arable and horticulture	5	2	1248	54	15	0	5	0	2	2	4	0	0	10	0	0	3	1	0	0	0	2	0	1353	92
Improved grassland	7	5	64	1047	79	19	13	12	3	3	0	0	0	3	0	0	0	0	0	3	0	1	0	1259	83
Rough grassland	15	10	17	40	177	18	11	26	4	7	19	4	0	4	0	1	0	4	0	2	4	1	0	364	49
Neutral grassland	1	0	0	3	9	22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	35	63
Calcareous grassland	0	1	1	0	4	2	41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	49	84
Acid grassland	2	6	0	8	117	0	6	155	0	2	14	0	7	4	0	0	0	0	0	0	0	0	0	321	48
Fen, marsh and swamp	2	0	0	0	8	0	0	0	98	0	0	0	0	0	0	0	0	0	0	0	1	0	0	109	90
Heather	5	7	0	1	3	0	1	11	1	352	29	3	0	0	0	0	0	0	0	0	0	0	0	413	85
Heather grassland	8	15	1	0	16	0	2	14	1	54	132	0	3	2	0	0	0	1	0	0	0	0	1	250	53
Bog	1	1	2	2	42	0	0	13	1	54	51	114	0	7	0	1	0	0	0	0	0	0	0	289	39
Montane habitats	0	0	0	0	8	0	0	0	0	4	2	0	14	6	0	0	0	0	0	0	0	0	0	34	41
Inland rock	0	0	5	0	2	0	0	0	0	0	0	0	0	177	0	1	0	1	1	1	0	5	0	193	92
Saltwater	0	0	0	0	0	0	0	0	0	0	0	0	1	0	153	61	0	0	0	7	1	0	0	223	69
Freshwater	1	0	0	0	0	0	0	0	1	0	0	0	0	2	3	409	0	0	0	2	1	0	0	419	98
Supralittoral rock	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	23	0	1	0	0	0	0	24	96
Supralittoral sediment	0	0	4	1	10	0	0	0	0	1	0	0	0	0	0	0	1	126	0	18	3	0	0	164	77
Littoral rock	0	0	0	0	0	0	0	1	0	0	0	0	0	0	4	0	0	0	100	3	0	0	0	108	93
Littoral sediment	0	0	0	0	0	0	0	0	1	0	0	0	0	1	14	0	1	9	3	256	11	2	0	298	86
Saltmarsh	0	1	0	0	4	0	0	0	1	0	0	0	0	0	0	0	0	2	0	1	179	0	0	188	95
Urban	1	1	11	6	1	0	0	1	0	1	0	1	0	12	0	1	0	0	0	1	0	559	84	680	82
Suburban	0	1	5	10	3	0	0	0	0	1	0	0	0	1	0	2	0	0	0	0	1	62	520	606	86
Sum of columns	685	1103	1367	1181	499	61	79	236	125	493	257	122	25	230	174	478	28	145	106	294	202	632	605		Correspondence (LCM2007 classes): **83%**
LCM2007 correspondence: Producer's accuracy (%)	87	94	91	89	35	36	52	66	78	71	51	93	56	77	88	86	82	87	94	87	89	88	86		BH correspondence: **85%**

Chapter 4: Comparison with Countryside Survey in 2007

4.1 Introduction

This chapter presents the results of a comparison of LCM2007 with Countryside Survey (CS) in 2007. The aim is to establish confidence in LCM2007 and quantify the correspondence between the LCM2007 and CS by comparing LCM2007 with:

- The 591 Countryside Survey 1x1km squares in Great Britain
- The Countryside Survey estimates of Broad Habitat area for England, Wales, Scotland, Northern Ireland and the UK

The first method produces a correspondence matrix. The second method produces a table showing the LCM2007 estimates of Broad Habitat area with the Broad Habitat estimates of area and confidence limits from the CS in 2007.

4.2 Comparison with Countryside Survey squares

Countryside Survey in 2007 surveyed 591 1x1km squares recording Broad Habitat cover, plus more detailed information (Carey *et al.*, 2008). The LCM2007 and Countryside Survey data sets are very different and to increase compatibility between the two, Countryside Survey polygons below the LCM2007 MMU width and area were excluded, as were the Countryside Survey *Mosaic* and '*Boundary and Linear Features'* classes, as there is no LCM2007 equivalent. Note, the Countryside Survey *Mosaic* class is used when the field surveyors encounter a mix of Broad Habitat types where it is not possible to map discrete areas of a single habitat exceeding the CS MMU. The field surveyors do record proportions of the Broad Habitats comprising the *Mosaic* polygons, but it would not be straightforward to include them in this analysis, so this class was excluded. Other differences, between CS in 2007 and LCM2007, such as LCM2007 mapping '*Standing Open Water and Canals'* and '*Rivers and Streams'* as a composite **Freshwater** class, were accounted for when the correspondence matrix was created. Equivalent data for Northern Ireland were not available to include in this analysis.

An area-based comparison was conducted between the 591 Countryside Survey squares and the corresponding areas of LCM2007. The correspondence was calculated by comparing the polygons in CS with the same area in LCM2007 and recording the area of Broad Habitat. This was conducted for each square in turn, enabling the correspondence for each square to be calculated, and the squares were then aggregated to enable production of correspondence tables for the: UK, England, Scotland and Wales.

Whilst every effort is taken to ensure the accuracy of CS data there are errors, due to habitat definition issues, as well as individual surveyor's interpretations. As a result the comparison between LCM2007 and CS in 2007 is about discerning correspondence and divergence, rather than accuracy. In addition, the timing of field

survey and image acquisition may not coincide, allowing changes to occur in the intervening period. Given that Countryside Survey is about 81% repeatable (Norton *et al.*, 2009) and that LCM2007 aims for an accuracy of > 80% at the Broad Habitat level, then combining these two accuracy levels suggests a likely correspondence of about 65% (80% x 81% = 65%). However, this does not take into account the different spatial structures of the two products, which although much closer than they used to be, still differ. This will have a tendency to reduce areal correspondence values.

The correspondences are reported for the UK, plus England, Scotland and Wales, and are calculated for a range of thematic levels:
- The Broad Habitats common to LCM2007 and CS in 2007.
- Aggregate class level (Aggregate classes are defined in Table 2.2, Chapter 2).
- Broad Habitat Association (BHA) level.

Broad Habitat Association (BHA)

The Broad Habitat Association concept provides an additional measure of thematic correspondence, which has some similarities to the Aggregate class thematic level, but is more targeted. A key factor underlying the Broad Habitat Association (BHA) concept is that the land cover maps provide land cover, which may relate uniquely to one Broad Habitat or may have a one-to-many relationship with several Broad Habitats. For example, deciduous woodland uniquely maps to the '*Broadleaved, Mixed and Yew Woodland*' Broad Habitat, whereas rough grassland has a one-to-many relationship with habitat, and may be '*Acid Grassland*', '*Calcareous Grassland*', '*Neutral Grassland*', poor quality '*Improved Grassland*' or even in some cases '*Fen, Marsh and Swamp*' or '*Bog*'. These distinctions can be difficult to make from remote sensing and even in the field; they are especially difficult when they occur as mosaics where the dominant (most widespread) land cover may vary depending upon the polygon boundaries.

In this type of situation the identification of *'associated'* habitats which form different Broad Habitats, as defined by Jackson (2000), is useful. The cross-habitat links in Table 4.1 formalises some of the uncertainties between mapping land cover and assigning it to a habitat-based classification [see Appendix 5 for further details]. The similarity of land cover for different BH is most pronounced for grassland and semi-natural upland areas, where mosaic landscapes maybe poorly represented spatially by the scale of the parcel-based structure in LCM2007, so these habitats dominate Table 4.1. In practice this means that as well as squares down the main diagonal of the correspondence matrix being accepted as corresponding directly, some additional cells are considered as corresponding directly. For clarity the cells used to calculate the BHA correspondence are shaded orange in Tables 4.3 & 4.5-4.7.

Broad Habitat Association (BHA): BHA identifies prescribed habitat links (Table 4.1) which form allowable correspondence between CS and LCM2007 classes e.g. '*Bog*' is acceptable for '*Dwarf Shrub Heath'* at the BHA-level.

The purpose of using BHAs is to separate correspondences which are slightly different e.g. '*Bog*' and '*Dwarf Shrub Heath*', compared to those that are very different e.g. *'Montane Habitats'* and '*Arable and Horticulture*'.

Table 4.1. Summary of correspondences allowed for Broad Habitat Association (whether due to transitional habitats, mosaic habitats, limitations of KBE's or difference in interpretation between Countryside Survey field survey and LCM).

LCM		CS Field Survey
'Bog'	←→	'Montane Habitats'
	←→	'Dwarf Shrub Heath'
	←→	'Acid Grassland'
'Montane Habitats'	←→	'Dwarf Shrub Heath'
	←→	'Acid Grassland'
	←→	'Bog'
'Dwarf Shrub Heath'	←→	'Acid Grassland'
Rough grassland	→	'Acid Grassland'
	→	'Calcareous Grassland'
	→	'Neutral Grassland'
	→	'Improved Grassland'
	→	'Fen, Marsh and Swamp'
	→	'Bog'
Water	→	'Fen, Marsh and Swamp'
Acid grassland	→	
Broadleaved woodland	→	
Rough grassland	→	
Any grassland	←	'Built-up Areas and Gardens'
Any water	←	'Built-up Areas and Gardens'
Any water	←→	Any water
Saltwater	←→	'Littoral Rock'
	←→	'Littoral Sediment'

Interpretation of correspondence values

Comparing spatial datasets with disparate spatial resolutions and thematic specifications is not straightforward and careful interpretation is required. Figure 4.1 shows three CS squares, with correspondence from 50% to 90%. The figure clearly demonstrates the differences in the level of detail between the Countryside Survey, with a MMU of 0.04ha, and LCM2007, with a MMU of 0.5ha. Spatial differences are particularly evident in Figure 4.1a where small and narrow objects, representing clumps of trees, roads and rivers have been removed by the generalisation of OSMM. There is also displacement and simplification of boundaries where narrow and small features have been 'absorbed' into their surroundings.

Figure 4.1a shows field survey and LCM squares with a direct correspondence of 50%, which sounds low, but it would be wrong to interpret this as a 50% error. Instead, as Figure 4.1a shows the Countryside Survey and LCM2007 squares look quite similar. Both versions of the square are predominantly a mix of '*Improved Grassland*', **Broadleaved woodland** and '*Neutral Grassland*' with the main differences being the extra spatial detail in the field survey square and the distribution of grassland types. The 50% example also shows two '*Fen, Marsh and Swamp*' polygons near the top of the CS square, which in LCM2007 are covered by

a single polygon that extends further to the west and is classified as '*Broadleaf, Mixed and Yew Woodland'.*

The level of spatial complexity (basically the number of polygons) is highest in the lowest correspondence square (Fig. 4.1). The simpler spatial structure of the 70% and 90% correspondence squares produces a better spatial match between the two.

Figure 4.2 shows the same areas as Figure 4.1, but with a single grassland class incorporating Acid, Neutral, Calcareous, Rough and '*Improved grassland'*. The correspondence between the Countryside Survey and LCM2007 data has been updated to reflect the single grassland class. When all grassland is combined into a single class the correspondence for square a) increases from 50% (Fig. 4.1a) to 87% (Fig. 4.2a). Fig. 4.2a shows the mix of grassland and woodland between the two squares is similar, with some of the remaining differences being due to the spatial heterogeneity of the landscape. Therefore, for some uses better results may be achieved by grouping all grassland into a single grassland class.

Table 4.2. Degree of correspondence between CS in 2007 and LCM2007 for Broad Habitat, Broad Habitat Association, and Broad Habitats with a single-grassland class.

Square (in Figs. 5.1 and 5.2)	Correspondence between CS2007 and LCM2007		
	Broad Habitat correspondence (%)	Broad Habitat Association correspondence (%)	BH with single grassland class correspondence (%)
a	50	75	87
b	70	79	85
c	90	92	90

Correspondence matrices

Correspondence matrices can be summarised by giving the Producers accuracy and the Users accuracy. The Producer's accuracy (the percentage values in the bottom two rows of Tables 4.3 & 4.5-4.7) show how well the CS in 2007 Broad Habitats are classified by LCM2007. The User's accuracy (the percentage values in the last two columns of Tables 4.3 & 4.5-4.7) shows the probability that an LCM2007 polygon of a particular class is likely to be that on the ground, based on the comparison with CS2007. The user's and Producer's accuracies are important tools for understanding the correspondence matrix and their respective roles are best illustrated by example. Table 4.3 shows that the Producer's correspondence between CS in 2007 and LCM2007 is 90% for *'Arable and Horticulture'*, whereas the User's correspondence between LCM2007 and CS in 2007 for *'Arable and Horticulture'* is 73%. What this means is that 90% of the CS squares mapped as 'Arable and Horticulture' are also mapped as *'Arable and Horticulture'* in LCM2007; whereas 73% of the land mapped as 'Arable and Horticulture' in LCM2007 corresponds with *'Arable and Horticulture'* in the CS dataset. The differences are subtle but important. It might help to imagine a simplified situation where the whole of the land surface has been mapped as '*Arable and Horticulture'*, when in reality *'Arable and Horticulture'* covers say 10% of the land surface. In this instance Producer accuracy would be 100%, since the whole area known to be *'Arable and Horticulture'* has been mapped as *'Arable and Horticulture'*. This looks encouraging until you consider the User accuracy of only 10%. This low

value implies that 90 % of the land has been incorrectly mapped. Clearly Producer and User accuracy should be considered together, and by doing so map users gain insight into the varying reliabilities of different land cover classes. The 73% User accuracy for LCM2007 *'Arable and Horticulture'* assessed against CS in 2007 implies that LCM2007 estimates for the extent of 'Arable and Horticulture' will be around 27% greater than those coming from CS 2007. And in fact LCM2007 does produce an approximately 30% higher national estimate for *'Arable and Horticulture'* than CS in 2007 (this point is discussed further in Section 4.3).
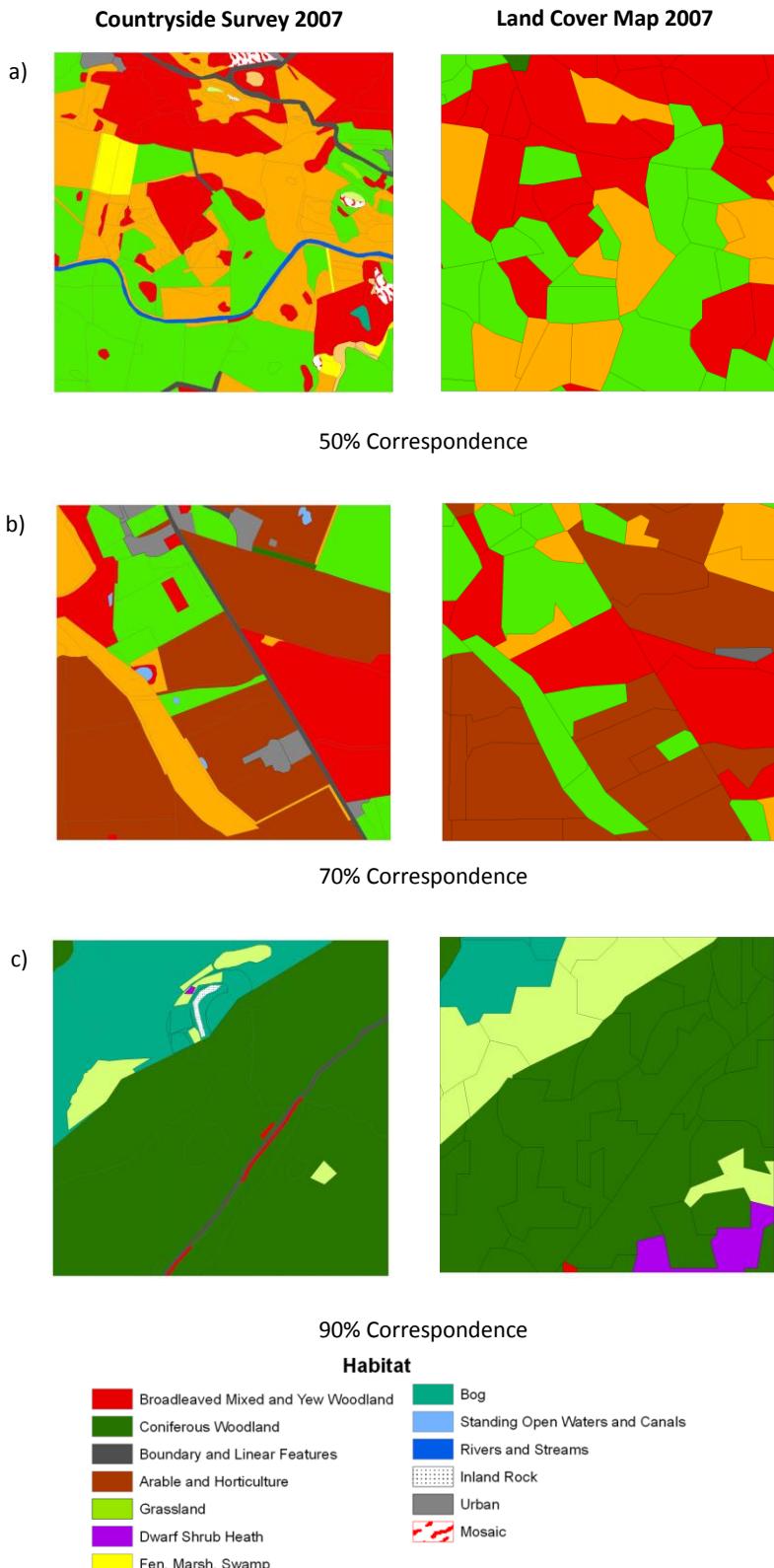
The comparison between the Countryside Survey in 2007 squares for Great Britain and LCM2007 show an overall correspondence of 62% at the Broad Habitat level (Table 4.3), 67% at the aggregate level (Table 4.4) and 76% at the associated BH-level (Table 4.3). For the common Broad Habitats (defined here as those covering more than 10,000km^2 of the UK (based on CS values in Table 4.9)), the Producer's accuracy ranges from 16% (*'Neutral Grassland'*) to 90% (*'Arable and Horticulture'*). The User's accuracy, also for the common Broad Habitats, varies from 37% (*'Dwarf Shrub Heath'*) to 89% (*'Built-up Areas and Gardens'*). The results presented in this section focus on the Producer's correspondence. The issues which would be raised by the User's accuracy are covered in Section 4.3, which focuses on the extent of Broad Habitat from both LCM2007 and CS in 2007.

The BHA are the same as the BH-level for some classes, but for many of the upland classes the correspondence is higher, for example, '*Bog*' goes from 37% at the BH-level to 96% at the BHA-level, as much of the difference between CS in 2007 and LCM2007 was due to the separation of '*Bog*', *'Dwarf Shrub Heath'* and '*Acid Grassland*'.

The correspondence of Countryside Survey '*Littoral Rock*' with LCM2007 is 0% in both the BH-level and BHA-level, as LCM2007 has the areas mapped as '*Littoral Sediment*' and '*Supra-littoral Sediment*'. This confusion is mainly between the coastal classes, so the Aggregate class correspondence for the coastal classes is higher at 66% (Table 4.4). Reducing the thematic resolution to Aggregate class level produces a range of correspondence from 35% for semi-natural grassland, through to 90% for '*Arable and Horticulture'* and 91% for **Freshwater** (Table 4.4)**.**
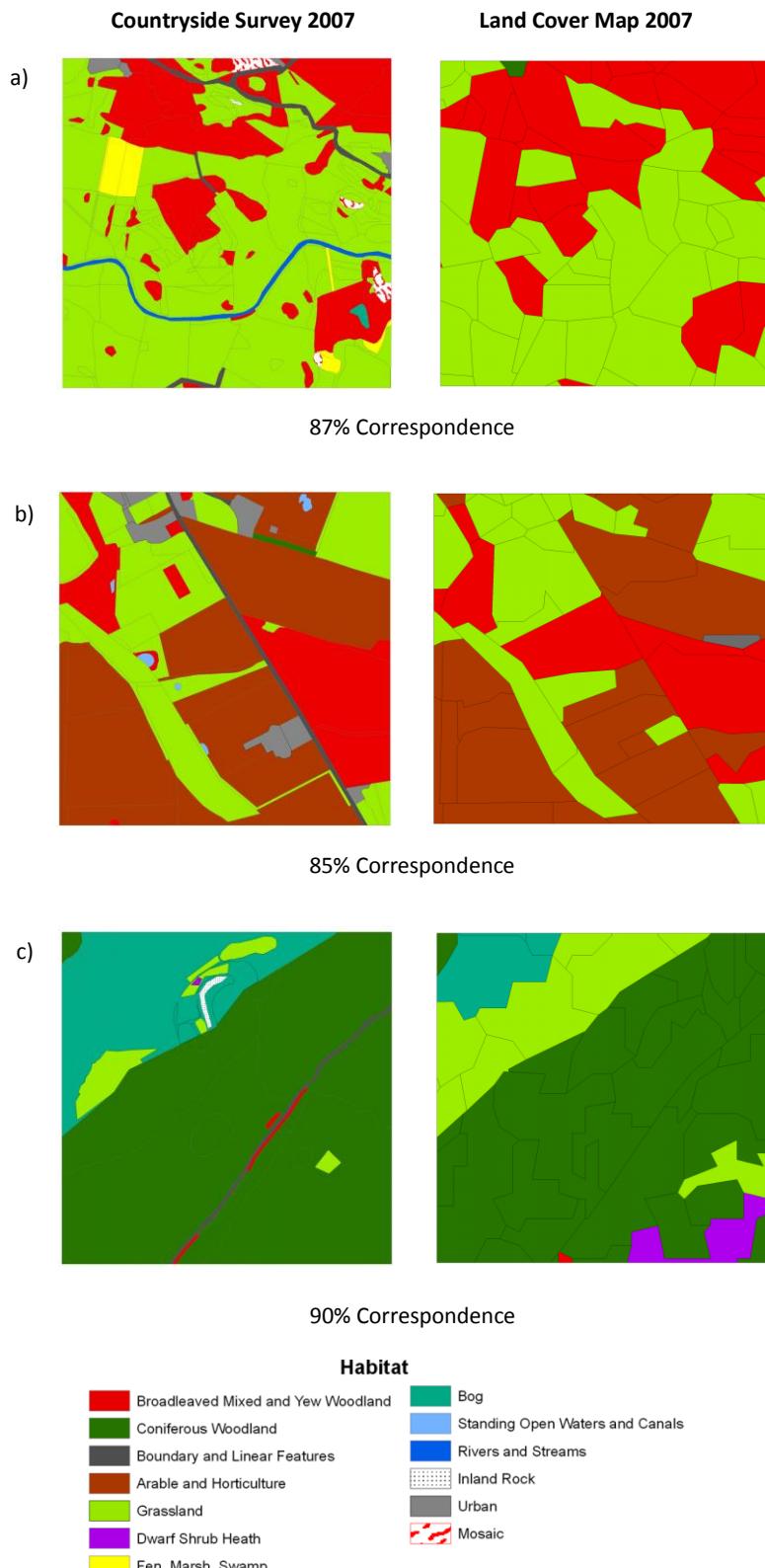
Fig. 4.1 illustrates three of the main differences, between the field survey and LCM2007, as shown by Table 4.3, specifically:
- The higher proportion of '*Fen, Marsh and Swamp'* in CS in 2007 compared to LCM2007 (Fig. 4.1a).
- The mismatch between recording of '*Neutral Grassland'* and '*Improved Grassland'* between the two data sets (Fig. 4.1a and b).
- Differences in the uplands based on the distribution of '*Bog*', '*Acid Grassland'* and *'Dwarf Shrub Heath'* (Fig. 4.1c).

Countryside Survey 2007   **Land Cover Map 2007**

a)

50% Correspondence

b)

70% Correspondence

c)

90% Correspondence

Habitat

■ Broadleaved Mixed and Yew Woodland	■ Bog
■ Coniferous Woodland	■ Standing Open Waters and Canals
■ Boundary and Linear Features	■ Rivers and Streams
■ Arable and Horticulture	▦ Inland Rock
■ Grassland	■ Urban
■ Dwarf Shrub Heath	▨ Mosaic
■ Fen, Marsh, Swamp	

Figure 4.1. Three Countryside Survey 1km squares and their LCM2007 equivalents showing examples of correspondence of a) 50%, b) 70% and c) 90%. © NERC (CEH) 2011. © Crown Copyright 2007. Ordnance Survey Licence number 100017572. © third-party licensors.
Note, the location of CS field survey squares is not disclosed to maintain the scientific integrity and relevance of CS. Consequently, the squares in Figure 4.1 have been spatially transformed to prevent identification of the squares location.

51

a)

87% Correspondence

b)

85% Correspondence

c)

90% Correspondence

Habitat

▮ Broadleaved Mixed and Yew Woodland	▮ Bog
▮ Coniferous Woodland	▮ Standing Open Waters and Canals
▮ Boundary and Linear Features	▮ Rivers and Streams
▮ Arable and Horticulture	▨ Inland Rock
▮ Grassland	▮ Urban
▮ Dwarf Shrub Heath	▨ Mosaic
▮ Fen, Marsh, Swamp	

Figure 4.2. The three CS squares in Fig. 4.1, but with all grassland generalised into a single grassland class. Correspondences are updated to recognise grassland as a single class.
© NERC (CEH) 2011. © Crown Copyright 2007. Ordnance Survey Licence number 100017572. © third-party licensors.
Note, the location of CS field survey squares is not disclosed to maintain the scientific integrity and relevance of CS. Consequently, the squares in Figure 4.2 have been spatially transformed to prevent identification of the squares location.

Table 4.3. Correspondence matrix showing degree of correspondence between LCM2007 and Countryside Survey in 2007 for the 591 field survey squares in Great Britain. (Units are ha). Green squares correspond at BH-level; orange squares correspond at BHA-level. Note, 'Boundary and Linear Features' and 'Mosaic' features were excluded from analysis; the CS in 2007 values for 'Standing Open Water and Canals' and 'Rivers and Stream' were combined to allow comparison with LCM2007, the CS in 2007 values for 'Acid Grassland' and 'Bracken' were also combined, as LCM2007 mapping of 'Bracken' is dependent upon image timing.

LCM2007 / Countryside Survey in 2007	Broadleaved woodland	Coniferous woodland	Arable and horticulture	Improved grassland	Neutral grassland	Calcareous grassland	Acid grassland	Fen, marsh and swamp	Dwarf shrub heath	Bog	Montane habitats	Inland rock	Saltwater	Freshwater	Supralittoral rock	Littoral rock	Littoral sediment	Built up areas and gardens	Sum of rows (ha)	BH: User's correspondence (%)	BH association: User's correspondence (%)
Broadleaved woodland	1212	205	42	131	145	6	59	35	17	6	0	1	2	4	2	0	0	92	1959	62	68
Coniferous woodland	134	2503	9	23	38	0	23	23	21	63	0	1	1	3	1	0	0	26	2870	87	88
Arable and horticulture	103	41	8643	1944	618	0	36	22	18	5	0	7	3	3	3	0	1	424	11871	73	73
Improved grassland	194	38	654	7533	2478	28	177	86	28	21	0	1	4	3	3	0	0	371	11619	65	68
Rough grassland	94	39	137	813	596	29	315	100	43	208	0	6	5	9	15	0	5	135	2549	23	86
Neutral grassland	5	0	24	87	102	1	0	2	0	0	0	0	0	0	0	0	1	15	237	43	49
Calcareous grassland	0	0	0	1	0	1	1	0	0	0	0	1	0	0	0	0	0	0	5	14	14
Acid grassland	36	41	8	166	231	1	1184	94	454	650	5	10	0	2	1	0	0	3	2885	41	83
Fen, marsh and swamp	0	0	0	1	0	0	2	1	0	0	0	0	0	0	0	0	0	0	3	21	21
Dwarf shrub heath	76	148	5	37	96	3	602	77	1082	768	1	16	4	6	3	0	0	10	2933	37	84
Bog	6	4	0	2	7	0	157	31	181	1091	0	1	3	2	0	0	0	1	1488	73	96
Montane habitats	0	0	0	0	0	0	58	1	135	146	19	0	0	0	0	0	0	0	359	5	100
Inland rock	2	3	8	4	11	0	25	0	12	14	0	13	0	0	0	0	0	38	132	9	9
Saltwater	0	0	0	0	0	0	0	0	0	0	0	0	172	0	0	0	2	0	175	98	100
Freshwater	1	2	2	1	2	0	1	0	0	2	0	0	54	346	5	0	0	3	418	83	96
Supralittoral rock	0	0	0	0	0	0	1	0	0	0	0	0	3	0	8	0	0	2	14	56	56
Supralittoral sediment	2	0	26	12	14	28	0	2	0	0	0	0	5	0	0	0	24	4	117	n/a	n/a
Littoral rock	1	0	1	2	3	3	7	1	1	2	0	0	75	0	9	0	1	1	108	0	70
Littoral sediment	0	0	15	6	1	1	0	1	0	0	0	0	251	0	1	1	44	1	322	14	92
Built up areas and gardens	27	1	48	24	30	0	2	4	0	0	0	1	0	2	0	0	2	1109	1251	89	89
Sum of columns (ha)	1895	3026	9622	10786	4371	102	2651	479	1993	2976	25	59	582	381	51	1	80	2234	41314		
BH: Producer's correspondence (%)	64	83	90	70	16	1	45	0	54	37	76	21	30	91	15	0	55	50	BH correspondence	62%	
BH association: Producer's correspondence (%)	64	83	90	77	16	30	87	48	93	96	100	21	95	91	15	0	58	79	BHA correspondence	76%	

Table 4.4. Correspondence matrix showing Aggregate class correspondence between LCM2007 and Countryside Survey in 2007 for the 591 field survey squares in Great Britain. Green squares correspond at Aggregate class level. (Units are ha).

LCM2007 \ Countryside Survey in 2007	Broadleaved woodland	Coniferous woodland	Arable and horticulture	Improved grassland	Semi-natural grassland	Mountain, bog, heath	Saltwater	Freshwater	Coastal	Built up areas and gardens	Sum of rows (ha)	AC correspondence (%)
Broadleaved woodland	1212	205	42	131	245	24	2	4	2	92	1959	62
Coniferous woodland	134	2503	9	23	84	85	1	3	1	26	2870	87
Arable and horticulture	103	41	8643	1944	676	31	3	3	3	424	11871	73
Improved grassland	194	38	654	7533	2769	50	4	3	3	371	11619	65
Semi-natural grassland	135	80	170	1066	2660	1377	5	10	22	153	5678	47
Mountain,heath,bog	84	155	13	43	1068	3479	5	9	5	49	4911	71
Saltwater	0	0	0	0	0	0	172	0	2	0	175	98
Freshwater	1	2	2	1	3	2	54	346	5	3	418	83
Coastal	3	0	42	20	63	3	335	0	87	8	561	16
Built up areas and gardens	27	1	48	24	36	1	0	2	2	1109	1251	89
Sum of columns (ha)	1895	3026	9622	10786	7603	5053	582	381	132	2234	41314	
AC correspondence (%)	64	83	90	70	35	69	30	91	66	50		67

4.2.1 Assessment by class

Broadleaved woodland - the correspondence matrix (Table 4.3) shows that LCM2007 and Countryside Survey in 2007 have a 64% correspondence for broadleaved woodland, with the primary sources of divergence being:

- LCM2007 recording areas as **Broadleaved woodland**, whilst Countryside Survey records them as '*Coniferous Woodland*' (205ha). This may in part be due to mixed woodland (see definition, Appendix 1, as determining the mix between deciduous and coniferous, and hence whether the area falls into the *'Broadleaved Woodland'* or '*Coniferous Woodland'* category, is subject to surveyor interpretation. Deciduous larch forests may also be having an impact, as spectrally they look more similar to deciduous broadleaved woodland than coniferous woodland.

- LCM2007 records some areas as '*Improved Grassland'*, whilst Countryside Survey records them as *'Broadleaved Woodland'* (194ha). This may be due to differences separating the point at which improved grassland with trees becomes woodland, as the transition between the two is based on accurate assessment of percentage tree cover. The differences in polygon structure between the data sets may also affect whether tree cover is great enough to count as woodland.

'Coniferous Woodland' - the correspondence between Countryside Survey in 2007 and LCM2007, at 83%, is very high. However, Countryside Survey '*Coniferous Woodland*' does correspond to LCM2007 **Broadleaved woodland** and '*Dwarf Shrub Heath'* in some cases. The correspondence with '*Dwarf Shrub Heath'* is likely to be

partly spectral confusion, but also recently felled stands, or recently replanted stands, where much of the spectral response is due to heather and grass around the young trees. Rides between woodland stands and open patches are also often '*Acid Grassland*' or '*Dwarf Shrub Heath*' dominated and maybe having an influence here.

'**Arable and Horticulture**' - the correspondence is 90% (8643ha correspond directly). The main divergence between the two data sets for '*Arable and Horticulture*' is with '*Improved Grassland*', although it is minor. Countryside Survey in 2007 recorded 1944 ha of LCM2007 arable as '*Improved Grassland*' and LCM2007 recorded 654ha of Countryside Survey in 2007 arable as '*Improved Grassland*'. This may reflect differences in interpretation between the two data sets, with CS recording agricultural ley as '*Arable*', whereas in the fallow/pasture period LCM would record it as '*Improved Grassland*'.

'**Improved Grassland**' - the correspondence between the two data sets is 70% at the Broad Habitat level, and 77% at the Broad Habitat Association level, which accepts that some low productivity, improved grassland is contained within the LCM2007 **Rough grassland** class. The main mis-match between Countryside Survey in 2007 recorded '*Improved Grassland*' is the 1944 ha which LCM2007 records as '*Arable and Horticulture*'. This difference may relate to:
- Rotation farming where the survey year differed between the two data sets.
- Misclassification in images where recently mown hay may have been indistinguishable from arable.
- Misclassification in images/composites using a spring, rather than summer image, where the spectral separability between arable and improved grassland was at a minimum.
- Heavily grazed land may appear spectrally to resemble arable more than improved grassland.
- Differences in the spatial structure between the two data sets.

'**Neutral Grassland**' - the correspondence between the two data sets is 16%. The low correspondence demonstrates the difficulty of identifying this habitat reliably from satellite data. In the field it is determined on botanical composition and it also includes semi-improved grasslands managed for silage, hay or pasture (Jackson, 2000), which in LCM2007 will often be classified as '*Improved Grassland*'. This is illustrated by the correspondence (Table 4.3) which shows that 56% of the 4371ha of '*Neutral Grassland*' recorded by Countryside Survey is classified by LCM2007 as '*Improved Grassland*' (2478ha), with the rest as a mix of mainly '*Arable and Horticulture*' and '*Acid Grassland*'.

'**Calcareous Grassland**' - Countryside Survey in 2007 recorded 102ha of '*Calcareous Grassland*' of which LCM2007 recorded 1ha, however, LCM2007 recorded 29ha of **Rough grassland** but the KBE (Section 3.7) was not able to convert this with confidence to '*Calcareous Grassland*'. 28ha of '*Calcareous Grassland*' was classified as '*Improved Grassland*' as spectrally they are often very similar. The 28ha of Countryside Survey '*Calcareous Grassland*', which LCM2007 has recorded as '*Supralittoral Sediment*' are probably all Machair grasslands. Machair grassland is a Priority Habitat, which sits under the '*Supralittoral Sediment*'

habitat (Jackson, 2000), though it may have the species composition of a '*Calcareous Grassland*'.

'**Acid Grassland**' - Countryside Survey in 2007 '*Acid Grassland*' shows 45% correspondence with LCM2007 at the Broad Habitat level rising to 87% if adjacent habitats (highlighted orange in Table 4.3), such as '*Dwarf Shrub Heath*', '*Bog*', '*Montane Habitats*' and **Rough grassland** are taken into account. This shows that there is strong agreement between the two data sets over the general type of habitat, in this case upland. The differences are in the assessment of whether grassland or heather dominated, or if the peat is thick enough to classify as 'Bog'. These difficulties and resultant differences are to be expected in upland environments where habitats form complex mosaics.

'**Fen, Marsh and Swamp**' - Countryside Survey records 479ha of '*Fen, Marsh and Swamp*' across the field survey sites, compared to 1ha recorded by LCM2007 (equivalent to 0%). Taking into account BHA increases the correspondence to 48%, with the other main mismatches being with '*Improved Grassland*', '*Dwarf Shrub Heath*' and '*Bog*'. From a remote sensing perspective '*Fen, Marsh and Swamp*' is problematic as it is can be comprised of a wide range of land cover types and many patches of Fen are below the LCM2007 MMU. The small size of '*Fen, Marsh and Swamp*' patches, plus their typically mosaic nature make it difficult to find representative areas, of sufficient size, to conduct a spectral classification. Soil data is of limited use in assisting as it shows the historical land cover, so large swathes of East Anglia have a peaty, fen soil, but subsequent drainage and management have changed them to arable. Rush-pastures are not an issue as in Countryside Survey 2007 a special point was made of not recording them as '*Fen, Marsh and Swamp*' after problems in earlier surveys.

'**Dwarf Shrub Heath**' - Very similar result to '*Acid Grassland*', with correspondence between the two data sets 54% at the Broad Habitat level, increasing to 93% when associated habitats are considered.

'**Bog**' - Similar to '*Acid Grassland*' and '*Dwarf, Shrub Heath*', with Broad Habitat correspondence of 37% and BHA correspondence of 96%.

'**Montane Habitats**' - The areas of '*Montane Habitats*' mapped by the two surveys have a 76% correspondence at the Broad Habitat level, increasing to 100% at the BHA-level.

'**Inland Rock**' - the correspondence between Countryside Survey in 2007 and LCM2007 '*Inland Rock*' is 21% at the Broad Habitat-level. Most of the difference between Countryside Survey and LCM2007 is split between the '*Acid Grassland*' and '*Dwarf Shrub Heath*' classes, which suggest that the mismatch is happening in upland areas. This is probably due to differences in the proportion of rock required before the area is classified as '*Inland Rock*' and differences in the structure of the spatial frameworks.

'**Saltwater**' - the correspondence at the Broad Habitat level for **Saltwater** is 30%, but once littoral coastal classes are included to account for tidal state and

freshwater, then the correspondence reaches 95% at the Broad Habitat Association level.

'Freshwater' - 91% correspondence at the Broad Habitat level.

'Supra-littoral Rock' - 15% correspondence at the Broad Habitat level, although some of the divergence is because of 9ha which LCM2007 has classified as '*Littoral Rock'*. The other main source of divergence is 15ha that LCM2007 has as **Rough grassland** and Countryside Survey has as '*Supra-littoral Rock*', a preliminary review of some of these areas suggests that they are typically sloping areas at the top of cliffs, with a mix of rock and rough grassland.

'Littoral Rock' - Countryside Survey identifies 1ha of *'Littoral Rock'* which LCM2007 maps as *'Littoral Sediment'*.

'Littoral Sediment' – '*Littoral Sediment'* has a correspondence of 55% between the two data sets and primarily gets confused with *'Supra-littoral Sediment'*.

'Built-up Areas and Gardens' - at the Broad Habitat level there is a 50% correspondence, however the way Countryside Survey and LCM2007 map urban areas is different. Countryside Survey counts parcels within the urban boundary as urban and assigns attributes to give more specific land cover information, whereas LCM2007 maps based on land cover rather than context. This is important because woodland, water and grassland within the Countryside Survey urban boundary may be mapped as urban rather than woodland or water or grassland, as LCM2007 would map it.

4.2.2 Country-level results

The country-specific results (Tables 4.5-4.7) are largely a function of the accuracy in the GB correspondence table combined with the area of the Broad Habitats in the different GB countries. Hence, Scotland has the most *'Dwarf Shrub Heath'* and *'Bog'*, both of which show low correspondence at the BH-level, but higher at the BHA-level, so this is reflected in the Scotland country-level results of 55.6% (Table 4.8) and 80.6% at the BH and BHA-level respectively. The more intensively managed nature of England underlies the relatively small percentage change from 67% and 75% at the BH and BHA-levels. The correspondence between the 107 Welsh CS-squares and LCM2007 is 58.2% at the BH-level, so higher than the equivalent Scottish value of 55.6%, but produces the lowest BHA value at 72.3% (Table 4.6). This occurs because the most common Broad Habitats in Wales, based on the extent recorded in the CS-squares (sum of columns in Table 4.7), are '*Arable and Horticulture*' and '*Neutral Grassland'* which are not affected by the BHA. In contrast, the equivalent two classes for Scotland are '*Bog*' and '*Improved Grassland'* (Table 4.6), both of which increase in correspondence when the BHA links are considered.

Table 4.5. Correspondence matrix showing degree of correspondence between LCM2007 and Countryside Survey in 2007 for the 289 field survey squares in England. (Units are ha). **Green squares** correspond at BH-level; **orange squares** correspond at BHA-level.
Note, 'Boundary and Linear Features' and 'Mosaic' features were excluded from analysis; the CS in 2007 values for 'Standing Open Water and Canals' and 'Rivers and Stream' were combined to allow comparison with LCM2007, the CS in 2007 values for 'Acid Grassland' and 'Bracken' were also combined, as LCM2007 mapping of 'Bracken' is dependent upon image timing.

LCM2007 / Countryside Survey in 2007	Broadleaved woodland	Coniferous woodland	Arable and horticulture	Improved grassland	Neutral grassland	Calcareous grassland	Acid grassland	Fen, marsh and swamp	Dwarf shrub heath	Bog	Montane habitats	Inland rock	Saltwater	Freshwater	Supralittoral rock	Littoral rock	Littoral sediment	Built up areas and gardens	Sum of rows (ha)	BH: User's correspondence (%)	BH association: User's correspondence (%)
Broadleaved woodland	855	107	33	75	94	5	23	14	6	1	0	0	0	1	0	0	0	60	1273	67	73
Coniferous woodland	61	400	7	4	3	0	2	15	0	18	0	1	0	1	0	0	0	12	523	76	79
Arable and horticulture	81	29	7356	1221	478	0	17	2	17	1	0	2	0	3	1	0	1	325	9535	77	77
Improved grassland	110	9	445	3803	1513	26	56	11	3	7	0	0	0	2	0	0	0	287	6272	61	65
Rough grassland	27	10	95	320	267	3	58	13	15	5	0	0	1	4	2	0	0	55	875	30	82
Neutral grassland	3	0	24	70	88	1	0	0	0	0	0	0	0	0	0	0	0	13	199	44	50
Calcareous grassland	0	0	0	1	0	1	1	0	0	0	0	1	0	0	0	0	0	0	5	14	14
Acid grassland	12	6	1	38	97	0	318	27	118	22	5	3	0	0	0	0	0	2	649	49	76
Fen, marsh and swamp	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	89	89
Dwarf shrub heath	12	48	3	4	37	0	285	27	351	50	1	1	0	0	0	0	0	3	821	43	84
Bog	0	3	0	0	2	0	89	11	100	142	0	0	0	0	0	0	0	1	350	41	95
Montane habitats	0	0	0	0	0	0	18	0	3	35	0	0	0	0	0	0	0	0	56	0	100
Inland rock	1	0	4	0	3	0	6	0	0	0	0	0	0	0	0	0	0	7	22	2	2
Saltwater	0	0	0	0	0	0	0	0	0	0	0	0	139	0	0	0	2	0	141	99	100
Freshwater	1	0	2	1	1	0	0	0	0	0	0	0	0	138	0	0	0	0	143	96	97
Supralittoral rock	0	0	0	0	0	0	0	0	0	0	0	0	2	0	7	0	0	2	11	64	64
Supralittoral sediment	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	2	1	10	n/a	n/a
Littoral rock	0	0	1	0	1	3	1	0	0	0	0	0	13	0	2	0	0	1	23	0	58
Littoral sediment	0	0	15	2	1	1	0	0	0	0	0	0	97	0	0	1	41	0	158	26	88
Built up areas and gardens	18	1	17	15	16	0	1	0	0	0	0	0	0	1	0	0	0	773	842	92	92
Sum of columns (ha)	1181	613	8003	5554	2607	41	875	122	613	281	6	8	252	149	12	1	46	1543	21907		
BH: Producer's correspondence (%)	72	65	92	68	14	2	36	1	57	51	0	5	55	92	56	0	89	50	BH correspondence	67%	
BH association: Producer's correspondence (%)	72	65	92	74	14	8	88	45	93	91	100	5	99	92	56	0	94	78	BHA correspondence	75%	

58

Table 4.6. Correspondence matrix showing degree of correspondence between LCM2007 and Countryside Survey in 2007 for the 195 field survey squares in Scotland. (Units are ha). Green squares correspond at BH-level; orange squares correspond at BHA-level.
Note, 'Boundary and Linear Features' and 'Mosaic' features were excluded from analysis; the CS in 2007 values for 'Standing Open Water and Canals' and 'Rivers and Stream' were combined to allow comparison with LCM2007, the CS in 2007 values for 'Acid Grassland' and 'Bracken' were also combined, as LCM2007 mapping of 'Bracken' is dependent on image timing.

LCM2007 \ Countryside Survey in 2007	Broadleaved woodland	Coniferous woodland	Arable and horticulture	Improved grassland	Neutral grassland	Calcareous grassland	Acid grassland	Fen, marsh and swamp	Dwarf shrub heath	Bog	Montane habitats	Inland rock	Saltwater	Freshwater	Supralittoral rock	Littoral rock	Littoral sediment	Built up areas and gardens	Sum of rows (ha)	BH: User's correspondence (%)	BH association: User's correspondence (%)
Broadleaved woodland	140	49	5	28	28	0	15	15	7	4	0	0	2	1	1	0	0	20	315	45	56
Coniferous woodland	55	1702	2	18	34	0	19	7	19	45	0	1	1	1	0	0	0	12	1917	89	89
Arable and horticulture	10	7	1085	295	59	0	5	15	1	3	0	3	2	0	2	0	0	46	1534	71	71
Improved grassland	39	15	129	1458	382	2	98	60	12	13	0	1	4	0	2	0	0	43	2259	16	66
Rough grassland	20	17	20	265	163	26	169	61	23	191	0	6	4	4	11	0	5	38	1022	16	89
Neutral grassland	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	82
Calcareous grassland	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Acid grassland	4	15	3	57	89	0	505	54	128	542	0	3	0	1	1	0	0	0	1403	36	88
Fen, marsh and swamp	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Dwarf shrub heath	54	84	0	21	43	2	271	42	552	692	0	13	4	5	3	0	0	6	1792	31	85
Bog	6	1	0	2	4	0	48	18	61	890	0	0	1	2	2	0	0	0	1036	86	96
Montane habitats	0	0	0	0	0	0	39	1	132	111	19	0	0	0	0	0	0	0	302	6	100
Inland rock	0	3	4	4	9	0	12	0	4	14	0	12	0	0	0	0	0	12	74	16	16
Saltwater	0	0	0	0	0	0	0	0	0	0	0	0	26	0	0	0	0	0	27	98	99
Freshwater	0	2	0	0	0	0	1	0	0	2	0	0	43	141	5	0	0	0	194	73	95
Supralittoral rock	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	3	28	28
Supralittoral sediment	2	0	26	12	9	28	0	2	0	0	0	0	0	0	0	0	16	1	95	n/a	n/a
Littoral rock	1	0	0	2	1	0	6	1	1	2	0	0	36	0	7	0	1	1	58	0	61
Littoral sediment	0	0	0	3	0	0	0	0	0	0	0	0	61	0	1	0	0	1	67	1	92
Built up areas and gardens	2	0	27	2	3	0	0	0	0	0	0	1	0	1	0	0	0	124	159	78	78
Sum of columns (ha)	334	1896	1303	2166	825	57	1190	276	940	2509	19	41	186	156	35	0	21	304	12257		
BH: Producer's correspondence (%)	42	90	83	67	20	0	42	0	59	35	100	30	14	90	2	0	2	41	BH correspondence	56%	
Producer's correspondence (%)	42	90	83	80	20	45	87	47	93	97	100	30	89	90	2		2	78	BHA correspondence	81%	

59

Table 4.7. Correspondence matrix showing degree of correspondence between LCM2007 and Countryside Survey in 2007 for the 107 field survey squares in Wales. (Units are ha). Green squares correspond at BH-level; orange squares correspond at BHA-level.
Note, 'Boundary and Linear Features' and 'Mosaic' features were excluded from analysis; the CS in 2007 values for 'Standing Open Water and Canals' and 'Rivers and Stream' were combined to allow comparison with LCM2007, the CS in 2007 values for 'Acid Grassland' and 'Bracken' were also combined, as LCM2007 mapping of 'Bracken' is dependent on image timing.

LCM2007 \ Countryside Survey in 2007	Broadleaved woodland	Coniferous woodland	Arable and horticulture	Improved grassland	Neutral grassland	Calcareous grassland	Acid grassland	Fen, marsh and swamp	Dwarf shrub heath	Bog	Montane habitats	Inland rock	Saltwater	Freshwater	Supralittoral rock	Littoral rock	Littoral sediment	Built up areas and gardens	Sum of rows (ha)	BH: User's correspondence (%)	BH association: User's correspondence (%)
Broadleaved woodland	217	48	5	29	23	0	21	6	5	1	0	1	0	2	1	0	0	12	371	59	64
Coniferous woodland	19	401	0	1	1	0	2	1	2	0	0	0	0	1	0	0	0	3	430	93	94
Arable and horticulture	12	6	201	428	81	0	14	5	0	1	0	2	0	0	0	0	0	52	803	25	25
Improved grassland	45	13	79	2272	583	0	24	14	13	1	0	0	0	1	1	0	0	40	3088	74	75
Rough grassland	46	13	22	228	166	1	88	25	4	12	0	0	1	1	2	0	1	42	652	25	86
Neutral grassland	2	0	0	16	14	0	0	2	0	0	0	0	0	0	0	0	1	2	37	38	43
Calcareous grassland	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Acid grassland	20	20	5	71	44	1	361	13	208	85	0	4	0	0	0	0	0	1	833	43	80
Fen, marsh and swamp	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	2	0	0
Dwarf shrub heath	10	16	2	12	15	1	46	8	180	27	0	2	0	1	0	0	0	1	320	56	79
Bog	0	0	0	0	0	0	20	2	20	58	0	0	0	1	0	0	0	0	101	58	97
Montane habitats	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	100
Inland rock	1	0	0	0	0	0	8	0	8	0	0	0	0	0	0	0	0	18	35	0	0
Saltwater	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	7	94	95
Freshwater	0	0	0	0	0	0	0	0	0	0	0	0	11	67	0	0	0	3	81	83	100
Supralittoral rock	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Supralittoral sediment	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	6	2	13	n/a	
Littoral rock	0	0	0	0	0	0	0	0	0	0	0	0	26	0	0	0	0	0	27	0	99
Littoral sediment	0	0	0	0	0	0	0	1	0	0	0	0	94	0	0	0	2	0	97	2	98
Built up areas and gardens	8	0	3	7	12	0	1	4	0	0	0	0	0	0	0	0	2	212	250	85	85
Sum of columns (ha)	380	517	317	3066	939	4	586	82	440	186	0	11	144	76	4	0	12	388	7149		
BH: Producer's correspondence (%)	57	78	64	74	19	0	62	0	41	31	0	0	5	89	0	0	16	55	BH correspondence	58%	
BH association: Producer's correspondence (%)	57	78	64	82	19	33	88	54	93	98	0	0	96	89	0	0	16	81	BHA correspondence	72%	

Table 4.8 summarises the number of squares compared for GB, England, Scotland and Wales, and the actual land area compared. It also shows the total area compared and the maximum potential area (number of 1km-squares * 100 (number of hectares per square)). The percentage of total area compared (calculated from the maximum potential area and actual area compared) varies as not all 1km-squares are completely recorded by CS, because:

- The square extends beyond the coast - sea is excluded.

- Access was not granted to some areas of the squares (affects 1459 ha).

- Habitat was recorded by CS as 'Mosaic', so was excluded from this analysis (affects 891 ha).

- Habitat was recorded as '*Boundary and Linear Features'*, so was excluded from this analysis (affects 68 ha).
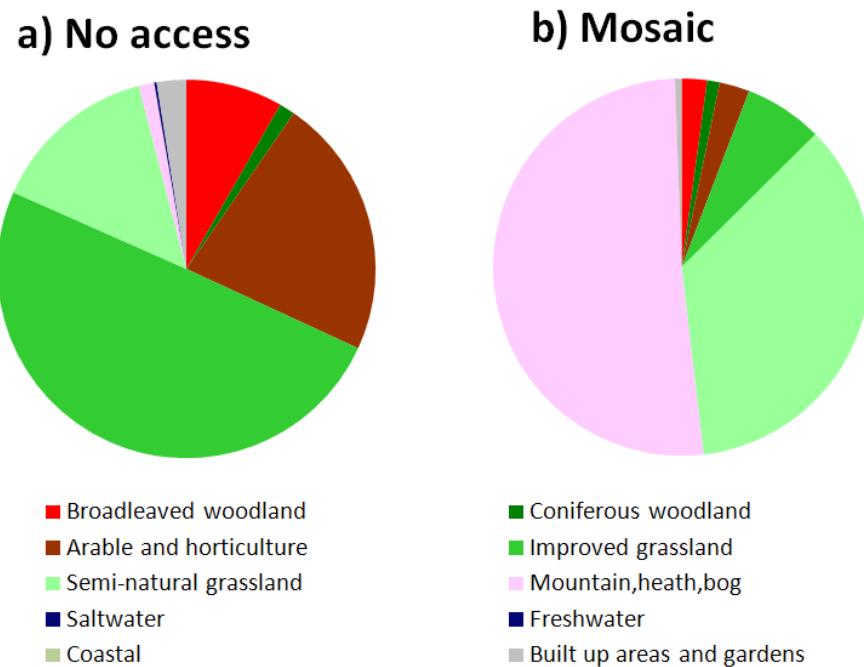
The lowest percentage of CS data used in the comparison is for Scotland, with 63%, probably as the CS 'Mosaic' class is used most often in the uplands (Figure 4.3b). Figure 4.3 also shows that areas classified by LCM2007 as '*Improved Grassland'* dominate the 'No Access' areas. Whether the systematic exclusion of 'No Access' and 'Mosaic' areas produces any systematic biases in correspondence is not clear.

Table 4.8. Summary of the number of squares, potential area and actual area in the comparison between the CS squares and LCM2007, plus the different levels of correspondence.

	GB	England	Scotland	Wales
Number of squares	591	289	195	107
Maximum potential area (ha)	59100	28900	19500	10700
Actual area compared (ha)	41314	21907	12257	7149
Percentage of total area compared (%)	70	75	63	67
Broad Habitats (%)	62	67	56	58
Broad Habitat Association (%)	76	75	81	72

4.2.3 Spatial coverage

The spatial distribution of correspondence at the Broad Habitat (Figure 4.4) and Broad Habitat Association levels (Figure 4.5) is useful for identifying patterns in high and low correspondence. Figure 4.4 shows that the correspondence is lower in semi-natural/upland areas where the spatial structure may vary more between the products and where classes such as '*Dwarf Shrub Heath'* and '*Bog*' may be difficult to separate. Allowing some flexibility in the correspondences using the BHA rules (Table 4.1) creates fewer low correspondences and increases the modal class from 51-60% (Broad Habitat correspondence (Fig. 4.4b)) to 91-100% (Broad Habitat Association correspondence (Fig. 4.5b)). Preliminary analysis of some of the low correspondence areas suggests that some of the problems are due to grassland mismatch, in particular, LCM mapping areas as '*Improved Grassland'* and Countryside Survey mapping areas as other grassland types, however, this will not be the explanation in all cases and would benefit from further analysis.

a) No access

b) Mosaic

- ■ Broadleaved woodland
- ■ Coniferous woodland
- ■ Arable and horticulture
- ■ Improved grassland
- ■ Semi-natural grassland
- ■ Mountain,heath,bog
- ■ Saltwater
- ■ Freshwater
- ■ Coastal
- ■ Built up areas and gardens

Figure 4.3. Piechart showing proportions of LCM2007 Aggregate classes for the polygons categorised by CS as either a) No access or b) Mosaic.
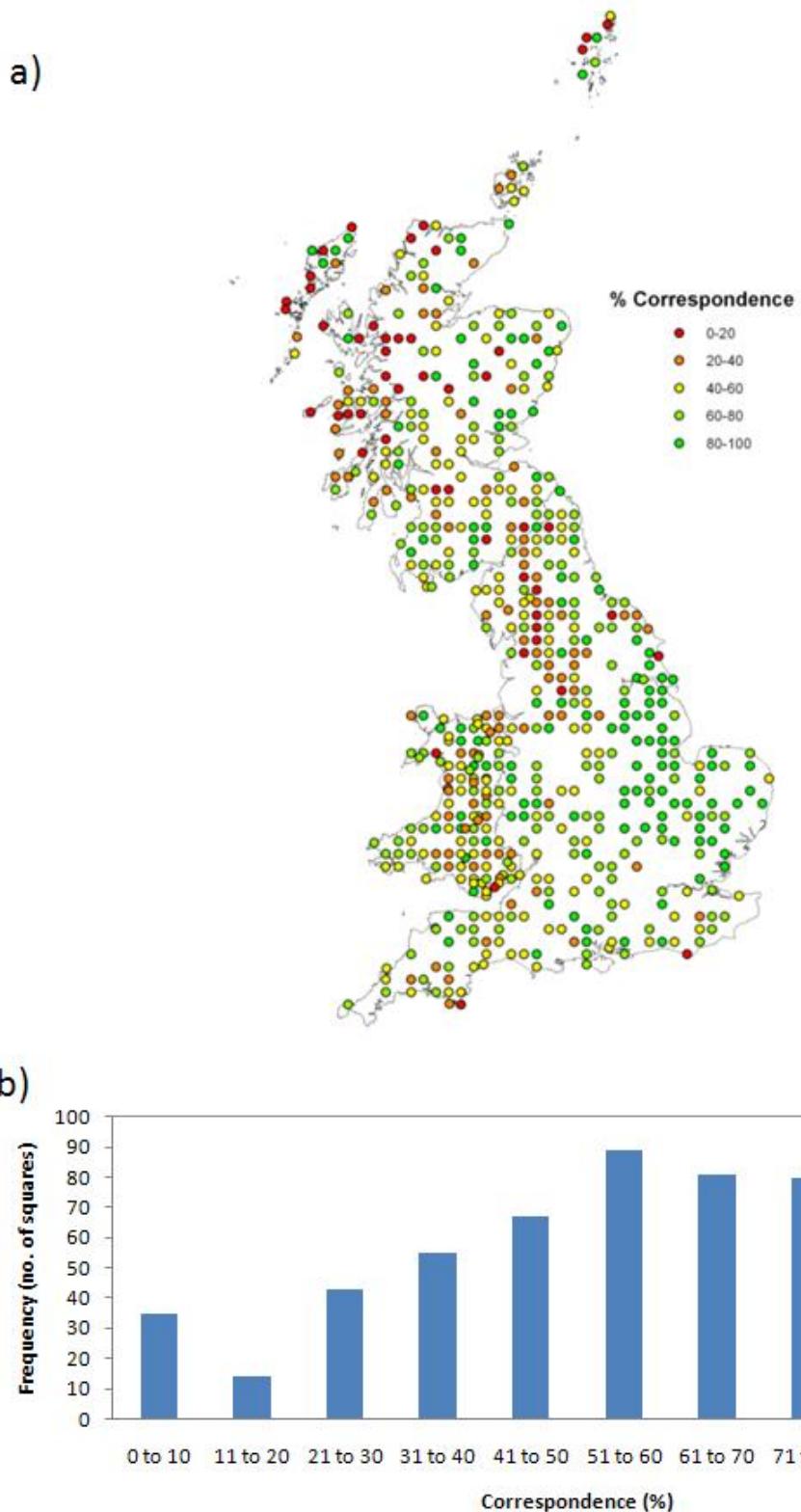
4.3 Extent of Broad Habitats

The Countryside Survey has an established method for calculating National Estimates of the area of each of the Broad Habitats from the field survey data types. LCM2007 provides a very different method of estimating these same Broad Habitat extents, based on extensive mapping of the UK. The comparison in Section 4.1 with the Countryside Survey squares was based on a very specific spatial comparison of the LCM2007 and Countryside Survey products. This comparison is based on area estimates for the UK but does not have a spatial component beyond the reporting of areas at country-level. Tables 4.8 and 4.9 show the Countryside Survey estimates, plus 95% confidence limits, and the LCM2007 estimates for the UK and constituent countries. The results are presented below in three categories depending on the degree to which the LCM2007 estimates fall within the Countryside Survey upper and lower certainty limits.
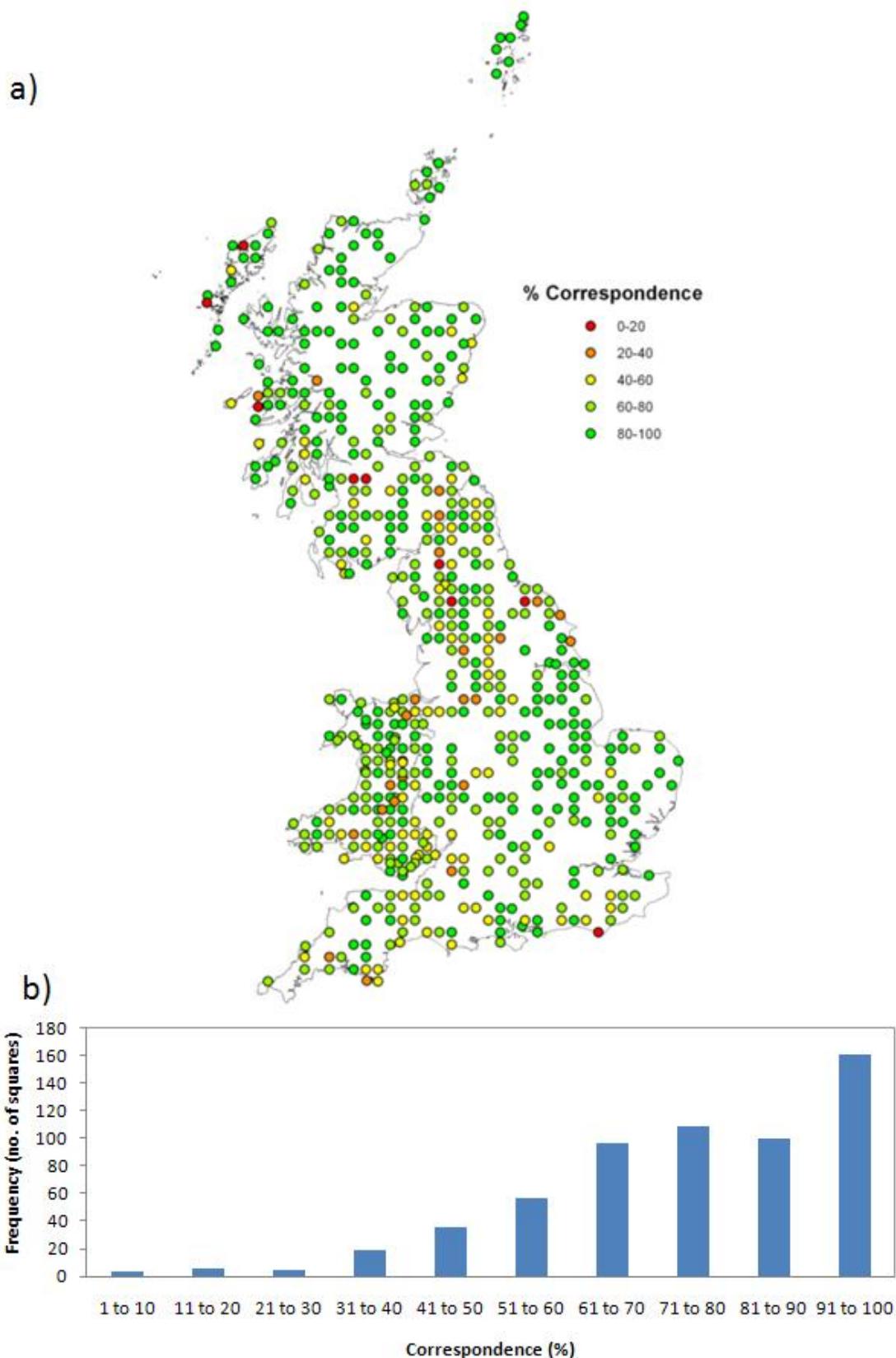
4.3.1 Very similar area estimates

Very similar results are defined as: classes falling within the CS upper and lower 95% confidence limits in **every** country (Table 4.8).

Very similar area estimates occur for: '*Coniferous Woodland'*, **Freshwater**, '*Built-up Areas and Gardens'* and '*Calcareous Grassland'* (Table 4.8*).*

The '*Coniferous Woodland'* result is to be expected as the correspondence analysis showed strong agreement between the Countryside Survey and LCM2007 data sets, however, '*Calcareous Grassland*' is more unexpected, as it was one of the classes

62

Figure 4.4. Correspondence between LCM2007 and the Countryside Survey in 2007 field survey squares for Great Britain plotted a) spatially and b) by frequency for the Broad Habitats. © NERC (CEH) 2011. © Contains Ordnance Survey data © Crown Copyright 2009.

Figure 4.5. Broad Habitat Association correspondence between LCM2007 and the Countryside Survey in 2007 field survey squares for Great Britain plotted a) spatially and b) by frequency for the Broad Habitats. © NERC (CEH) 2011. © Contains Ordnance Survey data © Crown Copyright 2009.

that performed most poorly in the correspondence matrix. This discrepancy highlights the different nature of the correspondence matrix analysis, based on the field survey squares, and this analysis using the estimates of Broad Habitat cover for the UK and constituent countries. In the case of *'Calcareous Grassland'* the answer may be that LCM2007 mapped the largest areas of *'Calcareous Grassland'* across the UK, such as Salisbury Plain and the South Downs well, so the overall estimate coincides with the Countryside Survey estimate.

4.3.2 Similar area estimates

Similar results are defined as: classes falling within the CS upper and lower 95% confidence limits for the UK (Table 4.9).

Similar area estimates occur for: **Broadleaved woodland**, '*Acid Grassland'* and '*Inland Rock'* (Table 4.9).

4.3.2 Dissimilar area estimates

Dissimilar results are defined as: classes falling beyond the CS upper and lower 95% confidence limits for the UK (Table 4.9).

Dissimilar area estimates occur for: '*Arable and Horticulture', 'Improved Grassland'*, '*Neutral Grassland', 'Dwarf Shrub Heath', 'Bog', 'Montane Habitats'* and coastal habitats.

'*Arable and Horticulture'* - the Countryside Survey estimated area at 46,574km^2 is similar to the 46,090km^2 (44,400km^2 of arable plus 1,690km^2 of horticulture) cited in DEFRA's June Survey of Agriculture and Horticulture for the UK in 2007 (DEFRA, 2011). The LCM2007 '*Arable and Horticulture'* estimate at 63,005km^2 exceeds the Countryside Survey upper limit of 51,276km^2. The difference between CS and LCM2007 is likely to be caused by a combination of the following factors:
- Differences in what is classified as *'Arable and Horticulture'* by the two methods.
- Spectral confusion.
- Inclusion of *'Boundary and Linear Features'* in the LCM2007 '*Arable and Horticulture'* area.

Differences in classification of *'Arable and Horticulture'* - The DEFRA June Agricultural Census figures suggest that in 2007, there were:
- 44,400km^2 of arable crops.
- 1,690km^2 of horticultural crops.
- 11,760km^2 of temporary grass.
- 5,990km^2 of uncropped land.

Together this gives a figure of 63,840km^2 for agricultural land, which is close to the LCM2007 figure of 63,005km^2 suggesting that part of the difference between Countryside Survey and LCM estimates may be due to temporary grass and uncropped land being included in the LCM '*Arable and Horticulture'* class. This is likely to be compounded because LCM requires images from several years to

produce full UK coverage. One consequence is that increasing the temporal range of images increases the number of fields changing from temporary grassland to arable and vice-versa. A field in a summer-winter composite which is grass in one image and arable in the other image is likely to be classified as arable, because of the greater spectral variability of this class.

Spectral confusion - The spectral variability of the '*Arable and Horticulture'* habitat is greater than any of the other classes, due to the wide range of crops and because growth stages of the same crop can vary between and across images. This can make it difficult to fully account for all the spectral variability of arable in an image, which is necessary to produce a good classification of the arable areas. It also means that any poorly represented spectral signatures are most likely to be classified as arable, if the image contains substantial arable areas. This may lead to a tendency to overestimate arable extent in some areas.

Differences in mapping of 'Boundary and Linear Features' - Countryside Survey maps boundary and linear features, but many of these are below the MMU of LCM. Consequently, in LCM boundary and linear features are incorporated into field polygons, so an area mapped in Countryside Survey as field-boundary-field, will be mapped in LCM as field-field. This will tend to increase the size of fields, in comparison to Countryside Survey, and hence the area mapped as '*Arable and Horticulture*' or '*Improved Grassland*' by LCM2007. The 5,270km^2 estimated for '*Boundary and Linear Features*' by Countryside Survey is likely to be split in LCM across the '*Arable and Horticulture'* or '*Improved Grassland'* habitats.

'*Improved Grassland'* and '*Neutral Grassland'* - the LCM2007 area for '*Improved Grassland'* is approximately 10,000km^2 more than the Countryside Survey estimate, whilst the LCM2007 estimate of '*Neutral Grassland'* area is approximately 10,000km^2 less than the Countryside Survey estimate. It appears therefore that the difference is due to how the different products deal with neutral grassland. Spectrally it looks like '*Improved Grassland'*, although in the field survey it may be apparent that the species composition is characteristic of '*Neutral Grassland'*. The soil data is of limited use in determining whether the grassland will be have the species composition of '*Neutral Grassland'*, or whether it is '*Improved Grassland'* on a neutral soil.

'*Dwarf Shrub Heath'* and '*Bog*' - Similar to the grassland case, the difference in '*Dwarf Shrub Heath'* and '*Bog*' classes is largely due to allocation between the two habitats. The extent of LCM2007 '*Dwarf Shrub Heath*' and '*Bog*' is 32,090km^2 compared to a value of 31,931km^2 for the combined lower confidence limit of Countryside Survey '*Dwarf Shrub Heath*' and '*Bog*'.
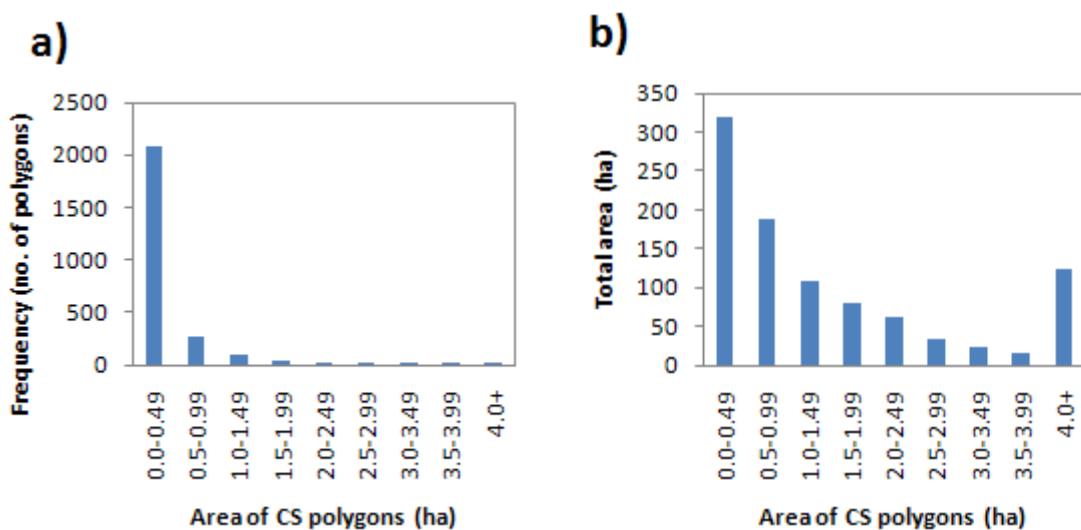
Montane and coastal habitats - are poorly represented in Countryside Survey data (Smart *et al.*, 2010) so it is unsurprising that the area estimates differ for these classes.

'*Fen, Marsh and Swamp'* - The CS in 2007 estimate of '*Fen, Marsh and Swamp'* area for the UK is 4392km^2, with upper and lower confidence limits of 3596km^2 and 5189km^2 respectively. The LCM20007 estimate is 101km^2. '*Fen, Marsh and Swamp'* is so different between the two surveys, because Fen is a mosaic of land cover

types, so areas mapped as '*Fen, Marsh and Swamp*' by CS in 2007, are frequently mapped as **Rough grassland** and '*Acid Grassland*' by LCM2007. The small patch size and different spatial structure of the CS in 2007 and LCM2007 products also contributes to differences between LCM and Countryside Survey recording of Fen (see discussion of Fig. 4.1). The QA of '*Fen, Marsh and Swamp*' (Section 3.9) is higher than the comparison against CS in 2007, because the ground reference polygons are designed to validate LCM2007 and hence the validation polygons match the LCM2007 polygons spatially and the areas of '*Fen, Marsh and Swamp*' are large enough to be detected by LCM2007 methods.

4.3.3 Discussion

This section has focussed on comparing LCM2007 area estimates with those derived from Countryside Survey. It is important to note that the Countryside Survey estimates are created by scaling-up using the ITE Land Classes. It is not clear how much this scaling up affects the estimates; although a sensitivity analysis of the BH area estimates to the different versions of the ITE Land Classes is currently underway (Wood *et al.*, in prep.). When LCM2000 data for the Countryside Survey squares were scaled-up using the Countryside Survey method (to produce UK estimates) the results were noticeably more similar to the Countryside Survey estimates than the results based on the LCM2000 mapped areas (Fuller *et al.*, 2002).

Figure 4.6. Frequency of Fen, Marsh and Swamp polygons recorded by Countryside Survey in 2007 based on a) polygon size and b) total area of polygons for each polygon size class. The first column in a) and b) is 0-0.49ha and as such is below the LCM2007 MMU.

One final factor to take into account is the influence of Countryside Survey parcels below the LCM MMU, which are likely to have a significant effect on the area recorded for some habitats, such as '*Broadleaf, Mixed and Yew Woodland*' and '*Fen, Marsh and Swamp*' which are often characterised by small parcels within a wider landscape dominated by other habitats. Figure 4.6 shows that most '*Fen, Marsh and Swamp*' polygons are below the LCM2007 MMU (Fig. 4.6a) and they account for 33% of the total '*Fen, Marsh and Swamp*' area.

The impact of Countryside Survey patch size on different habitats has not been investigated comprehensively, but is likely to show that other habitats also occur frequently in patch sizes, below the size of the LCM MMU. Understanding the patch size frequency distributions would be useful in further understanding the relationship between the LCM and CS data sets.

4.4 UK Land Cover

Summary estimates for UK Land Cover are given in Figure 4.7 (Broad Habitat values are given in Tables 4.8 and 4.9). LCM2007 shows that for the UK more than 50% is intensive agriculture ('*Arable and Horticulture*' plus '*Improved Grassland*', 51%)) or developed ('*Built-up Areas and Gardens*' (6%)). The remainder is mainly semi-natural, with woodlands covering 12% of the UK - split evenly between **Broadleaved Woodland** and **Coniferous Woodland**. The remaining 30% of the UK is split between **Coastal** (1%), **Semi-natural grassland** and **Mountain, heath and bog**.

The land cover distribution varies dramatically across the four countries of the UK. England has the highest proportion of intensive land use at 76% (40% '*Arable and Horticulture*'; 27% '*Improved Grassland*'; 9% '*Built-up Areas and Gardens*'), followed by Northern Ireland and Wales with 64% of intensive land use ('*Improved Grassland*' and '*Arable and Horticulture*' and '*Built-up Areas and Gardens*'). Scotland has the lowest proportion of intensive land-use with 36%, although it does have the largest proportion of '*Coniferous Woodland*', which could arguably be seen as an intensive land-use. The largest proportion of semi-natural areas are in Scotland, with 36% Mountain, heath and bog and 20% semi-natural grassland. The proportion of **Freshwater** varies widely across the four countries from 1% for Wales and England, 2% for Scotland and 4% for Northern Ireland, mainly due to Lough Neagh.

LCM2000 mapped the UK using the same Broad Habitat classes as LCM20007. Figure 4.8 shows the proportions of land cover based on LCM2000 for the Aggregate classes. The main differences are that LCM2007 records 26% '*Arable*' and 13% semi-natural grassland compared to LCM2000's values of 23% and 17% for '*Arable*' and semi-natural grassland respectively. **Coniferous woodland** changes from 5.4% to 6.1%, whilst **Urban** is slightly less in LCM2007 than in LCM2000 (5.9% compared to 6.7%). How much the figures are affected by the changes to the spatial framework is unclear and how the differences between the two products should be interpreted requires further analysis.

Table 4.8. The coverage (km^2) of Broad Habitats from LCM2007 compared with the Countryside Survey in 2007 estimates and the Countryside Survey upper and lower confidence limits at 95%, for England, Scotland, Wales and Northern Ireland. CS in 2007 and LCM2007 estimates are in km^2. LCM2007 estimates are highlighted: **bold**, where LCM2007 estimates fall within Countryside Survey upper and lower limits; red, where they exceed the Countryside Survey upper limit and blue where they are below the Countryside Survey lower limit.

	England				Scotland				Wales				Northern Ireland			
	Lower CS 95% limit	CS2007	LCM2007	Upper CS 95% limit	Lower CS 95% limit	CS2007	LCM2007	Upper CS 95% limit	Lower CS 95% limit	CS2007	LCM2007	Upper CS 95% limit	Lower CS 95% limit	CS2007	LCM2007	Upper CS 95% limit
Broadleaved, Mixed and Yew Woodland	8208	9814	**9300**	11443	1943	2510	**2642**	3073	1351	1736	1251	2339	651	817	540	983
Coniferous Woodland	1373	2567	**3033**	3494	7311	9560	**9938**	11876	337	1059	**1431**	1808	388	606	**654**	824
Linear Features	3166	3534		4022	807	951		1116	353	476		604	282	310		337
Arable and Horticulture	35356	40016	53329	44198	3748	5341	**7041**	7073	425	728	1769	1109	357	489	867	621
Improved Grassland	25543	28564	35684	31898	7238	9070	11178	11119	6254	7302	**8422**	8753	5308	5730	**7093**	6152
Neutral Grassland[1]	12668	14526	6110	16214	3729	4611	5758	5625	2148	2627	**2271**	3178	2019	2311	1752	2603
Calcareous Grassland	85	297	**359**	599	0	261	**12**	659	0	12	**0**	33	-8	18	**1**	44
Acid Grassland	2903	3956		5106	7895	9827		11855	1363	2106		3300	48	104		159
Bracken	501	911		1383	928	1315		1680	225	375		510	9	26		44
Acid grassland & Bracken	3404	4867	3171	6489	8823	11142	**10240**	13535	1588	2480	**2845**	3810	58	130	217	203
Dwarf Shrub Heath	1922	3315	**3610**	4805	7054	8943	15660	10810	393	1170	**1122**	1740	78	168	732	257
Fen, Marsh, Swamp	686	1175	68	1719	1831	2385	26	2974	235	360	6	591	375	473	2	570
Bog	818	1397	**1965**	2039	16848	20438	7689	23278	48	483	**415**	801	1283	1609	925	1935
Standing Open Waters	-88	331		617	406	885		1574	24	54		91	485	613		742
Rivers and Streams	113	218		269	-30	157		233	28	57		93	40	55		70
Freshwater	25	549	**798**	886	376	1042	**1703**	1807	51	111	**114**	184	524	668	**632**	812
Montane	0	38	366	114	5	383	4520	1191	0	1	17	2	0	2[2]	0	0
Inland Rock	40	95	423	156	350	837	**705**	1507	31	75	**88**	122	-12	54	97	121
Built-up Areas and Gardens	8285	10382	**11690**	12280	1133	1531	**1421**	1983	738	1321	**893**	1722	585	741	**644**	897
Supra-littoral rock	15	59	10	108	187	305	60	454	14	68	**8**	164	1	16	**1**	31
Supra-littoral sediment	0	44	184	106	34	167	**220**	340	3	32	**63**	90	1	20	**16**	38
Littoral rock	0	13	112	29		0	351		0	7	30	13		3[3]		8
Littoral sediment	5	375	1617	947	1	4	604	10	9	54	302	113		3[3]		66

1. LCM2007 Neutral Grassland estimates include LCM2007 **Rough grassland**, although this assignment is not ideal as **Rough grassland** is known to contain a mix of 'Acid Grassland', 'Calcareous Grassland', 'Neutral Grassland' and some 'Improved Grassland'.
2. Northern Ireland 'Montane Habitats' is surveyed in a separate montane stratum covering 735 ha so the actual montane habitat area will always be less than this.
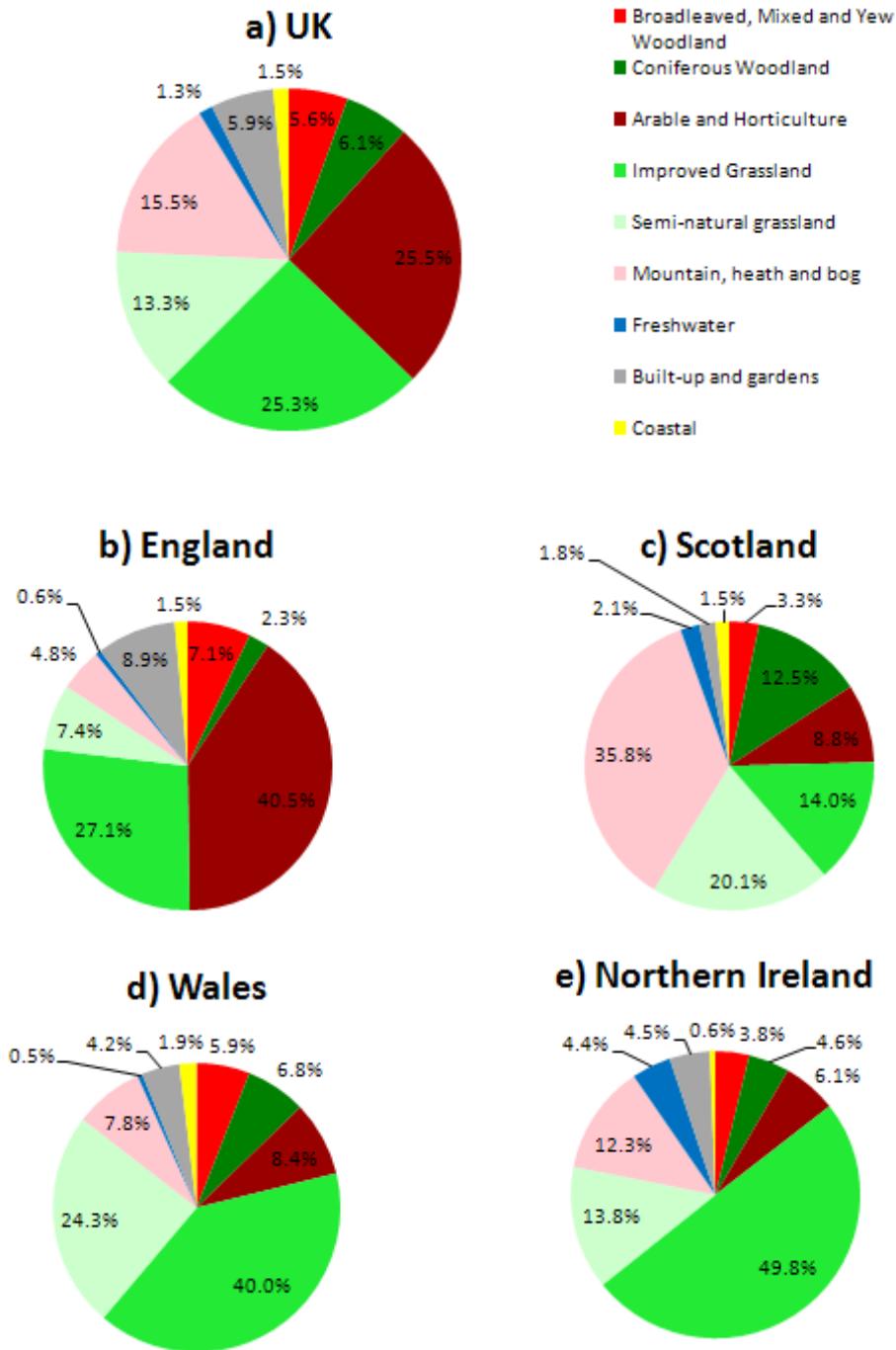**3.** The Northern Ireland Intertidal Broad Habitats are also in a separate stratum which is analysed separately.
4. Acid Grassland & Bracken, plus the Freshwater class both have CS in 2007 areas and 95% confidence limits derived by adding the BH confidence limits, rather than the bootstrapping method used to create the BH confidence limits.

Table 4.9. The coverage (km^2) of Broad Habitats from LCM2007 compared with the Countryside Survey in 2007 estimates and the Countryside Survey 95% upper and lower confidence limits for the UK. CS in 2007 and LCM2007 estimates are in km^2. LCM2007 estimates are highlighted: **bold**, where LCM2007 estimates fall within Countryside Survey upper and lower limits; red, where they exceed the Countryside Survey upper limit and blue where they are below the Countryside Survey lower limit.

	UK			
	Lower CS 95% limit	CS2007	LCM2007	Upper CS 95% limit
Broadleaved, Mixed and Yew Woodland	13107	14894	**13733**	16680
Coniferous Woodland	11145	13909	**15057**	16674
Linear Features	4797	5270		5744
Arable and Horticulture	41872	46574	63005	51276
Improved Grassland	46948	50670	62377	54392
Neutral Grassland[1]	21955	24074	15890	26194
Calcareous Grassland	147	588	**372**	1028
Acid Grassland	13502	15985		18469
Bracken	2033	2627		3222
Acid grassland	15535	18613	**16471**	21691
Dwarf Shrub Heath	11153	13596	21118	16039
Fen, Marsh, Swamp	3596	4392	101	5189
Bog	20778	23929	10972	27080
Standing Open Waters	1079	1884		2689
Rivers and Streams	324	498		672
Freshwater	1403	2382	**3248**	3360
Montane	-235	422	4886	1080
Inland Rock	455	891	**1314**	1328
Built-up Areas and Gardens	11883	13975	**14648**	16067
Other land	606	1140	3633	1679

1. LCM2007 Neutral Grassland estimates include LCM2007 **Rough grassland**, although this assignment is not ideal as **Rough grassland** is known to contain a mix of *'Acid Grassland'*, *'Calcareous Grassland'*, *'Neutral Grassland'* and some *'Improved Grassland'*.

Figure 4.7. Pie charts showing the percentage of Aggregate classes for a) the UK and b)-e) its constituent countries from LCM2007.
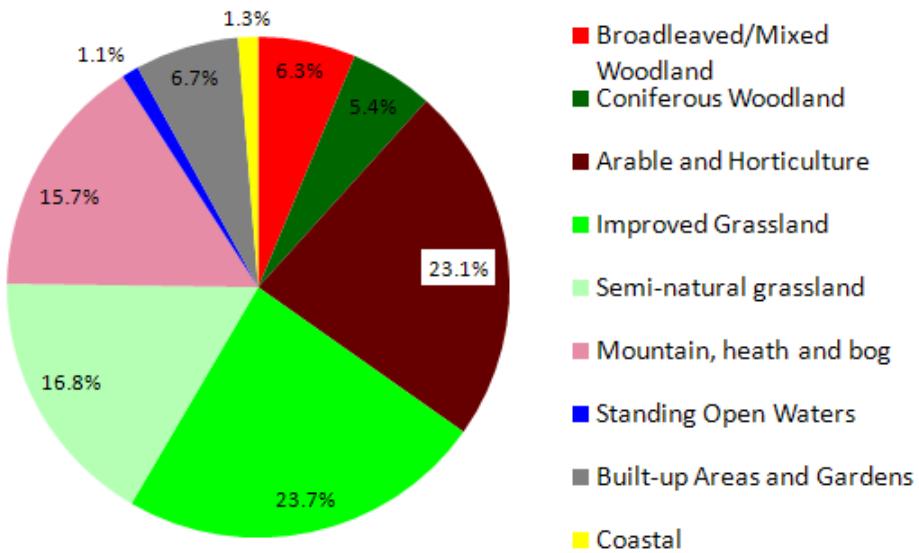
Figure 4.8: Pie charts showing the percentage of Aggregate classes for the UK from LCM2000.

4.5 Summary and discussion

Chapter 4 contains two comparisons between LCM and Countryside Survey. The first comparison was based on the BH mapped in the Countryside Survey 1km squares and LCM2007, whilst the second was the comparison between Countryside Survey National Estimates of Broad Habitat extent and LCM2007 estimates. The two analyses of LCM and Countryside Survey both showed different aspects of the data sets and demonstrated that:

- The level of agreement between LCM2007 and the Countryside Survey data sets varies widely across the Broad Habitats.
- Grassland categories are problematic because of the one-to-many relationship between the observed land cover of grass and the many Broad Habitats where grass is part of the habitat. This led to the development of the Broad Habitat Association rules, which gave a third level of thematic accuracy to assess the correspondences at, in addition, to the Broad Habitat level and the Aggregate class level.
- LCM2007 shows high correspondence with the area of '*Arable and Horticulture*' mapped by CS in 2007 for the 1km squares, however, as the estimates of Broad Habitat extent show LCM2007 estimates a much higher '*Arable and Horticulture*' area. Some possible reasons for this are suggested in Section 4.3.
- '*Fen, Marsh and Swamp*' estimates for the UK vary by an order of magnitude, between CS in 2007 and LCM2007. This is due to the complex mix of land cover types that make up Fen areas, making them difficult to identify with spectral classification, and the small size of many Fen areas. The comparison with the CS 1km squares suggests that a lot of what CS records as Fen and which is large enough to be mappable by LCM2007 is actually recorded as either **Rough grassland** or '*Acid Grassland*'. This raises the possibility that future work could use additional data sets to create a set of KBE's to allocate some of these grassland areas to '*Fen, Marsh and Swamp*'.

Chapter 5: The LCM2007 product range

5.1 Example areas

A series of example areas taken from the LCM2007 vector product are presented in this section. A range of geographical areas have been chosen to illustrate the performance of LCM2007 over a variety of different habitat and landscape types (see Figure 5.1 for the legend), including upland areas (Figures 5.2 & 5.3), calcareous (Figure 5.2), urban (Figure 5.4) and arable and fenland (Figure 5.5).
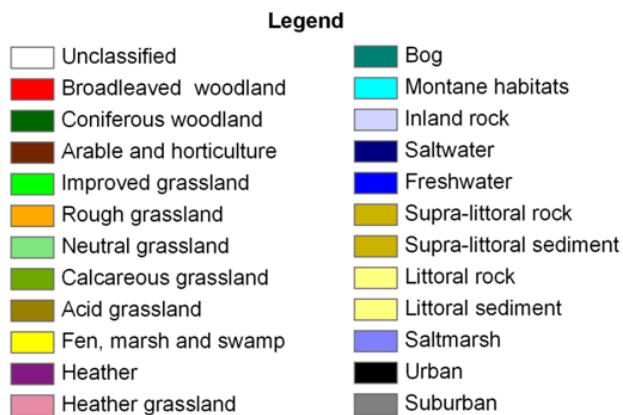

Legend

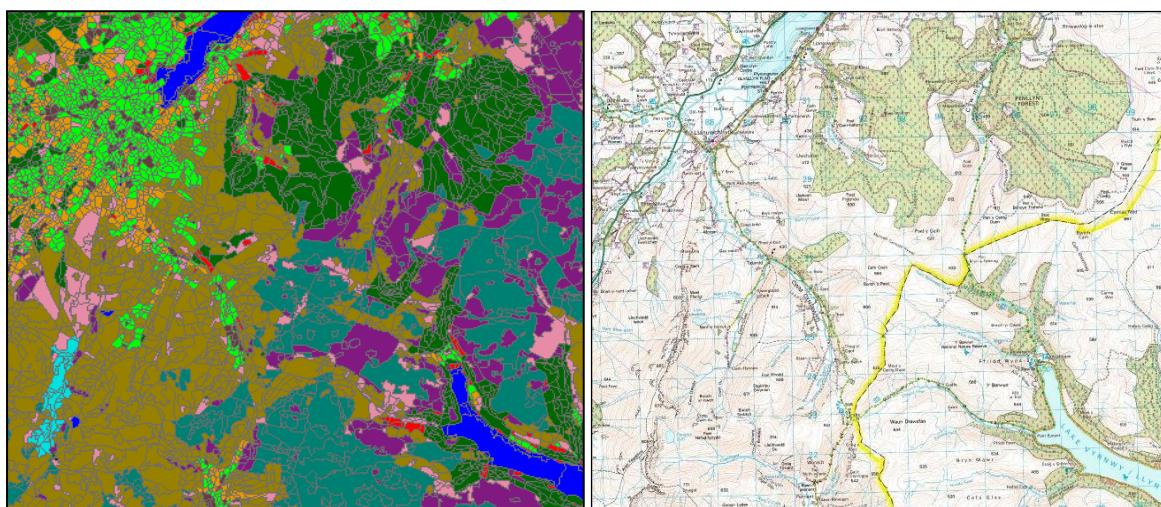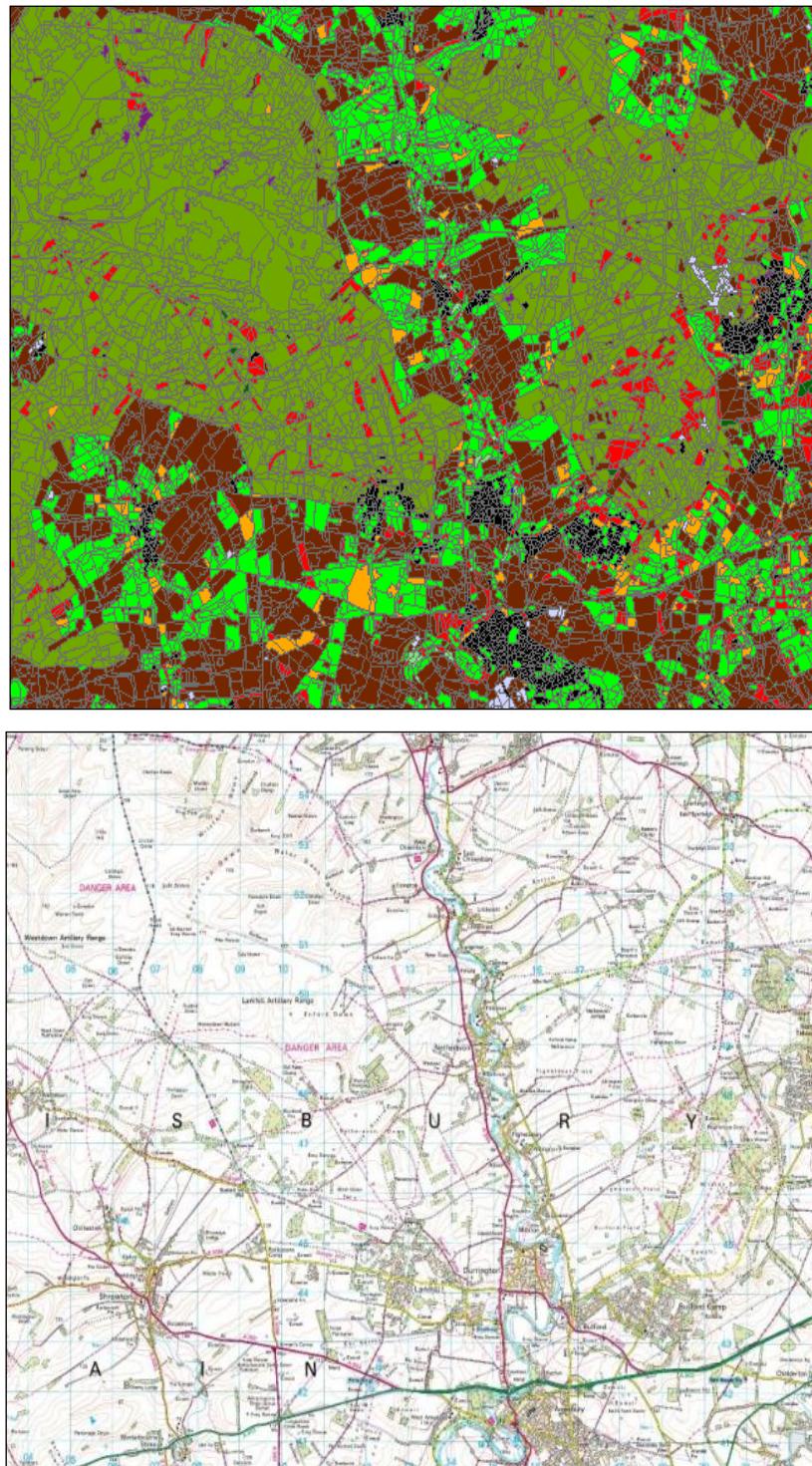|			
☐ Unclassified		Bog	
Broadleaved woodland		Montane habitats	
Coniferous woodland		Inland rock	
Arable and horticulture		Saltwater	
Improved grassland		Freshwater	
Rough grassland		Supra-littoral rock	
Neutral grassland		Supra-littoral sediment	
Calcareous grassland		Littoral rock	
Acid grassland		Littoral sediment	
Fen, marsh and swamp		Saltmarsh	
Heather		Urban	
Heather grassland		Suburban	

Figure 5.1. LCM2007 legend.

Figure 5.2. LCM2007 example and corresponding OS map area of North Wales, approx 15km x 12km. © NERC (CEH) 2011. © Crown Copyright. 2007. Ordnance Survey Licence number 100017572.
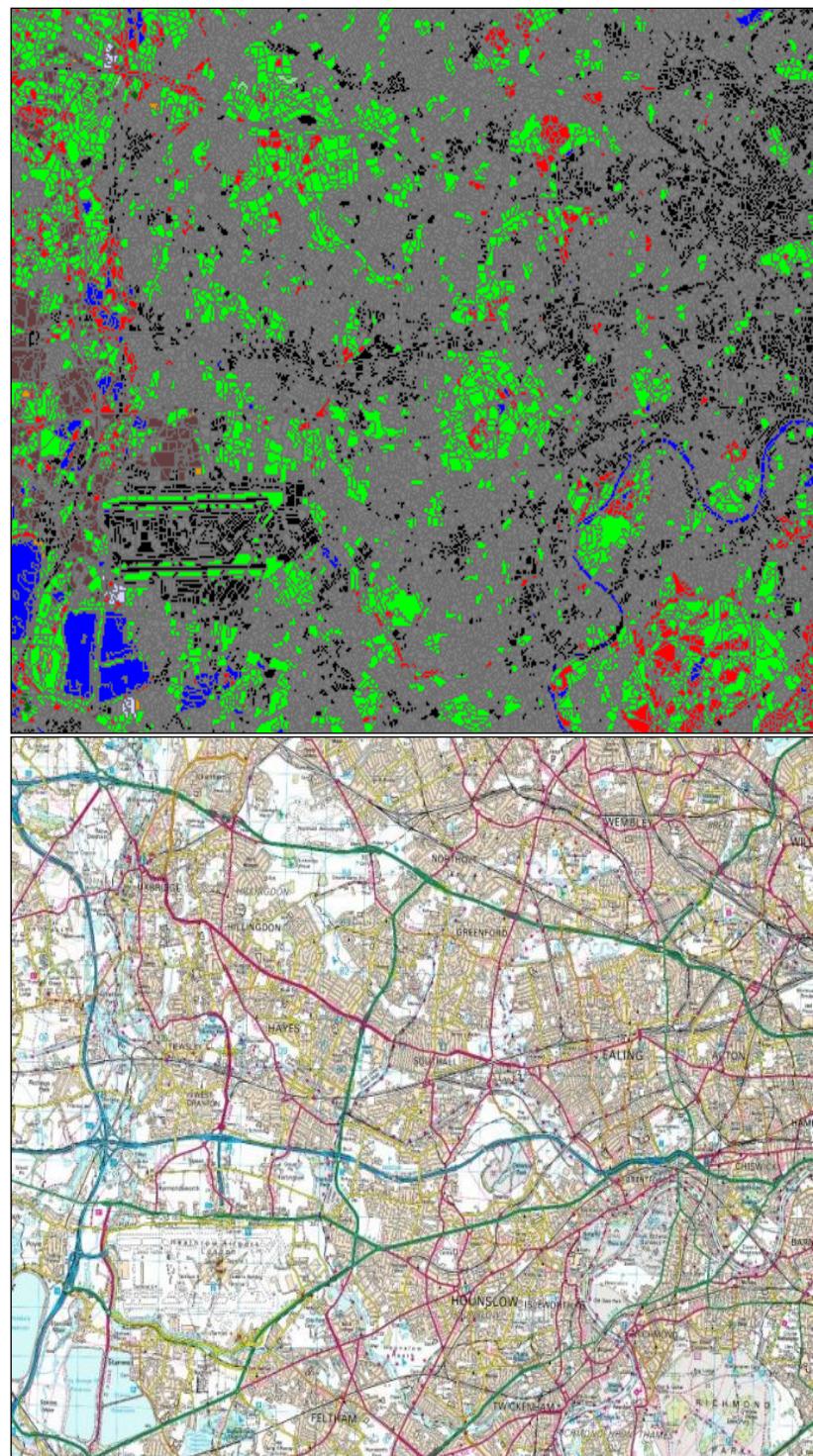
Figure 5.2 shows an area covered by a mixture of upland '*Dwarf Shrub Heath'*, '*Acid Grassland*' and '*Bog*' with the Penllyn Forest to the north and Lake Vyrnwy and Lake Bala to the south-east and north, respectively. To the south-west the montane ridge (cyan colour) of Aran Benllyn is apparent.

73

The second example shows a mixture of the '*Calcareous Grassland'* of Salisbury Plain and its small patches of broadleaved woodland, and the surrounding arable and '*Improved Grassland'* (Figure 5.3). The suburban towns of Amesbury and Tidworth can be seen as well as the industrial areas of Boscombe Down airfield in the south.
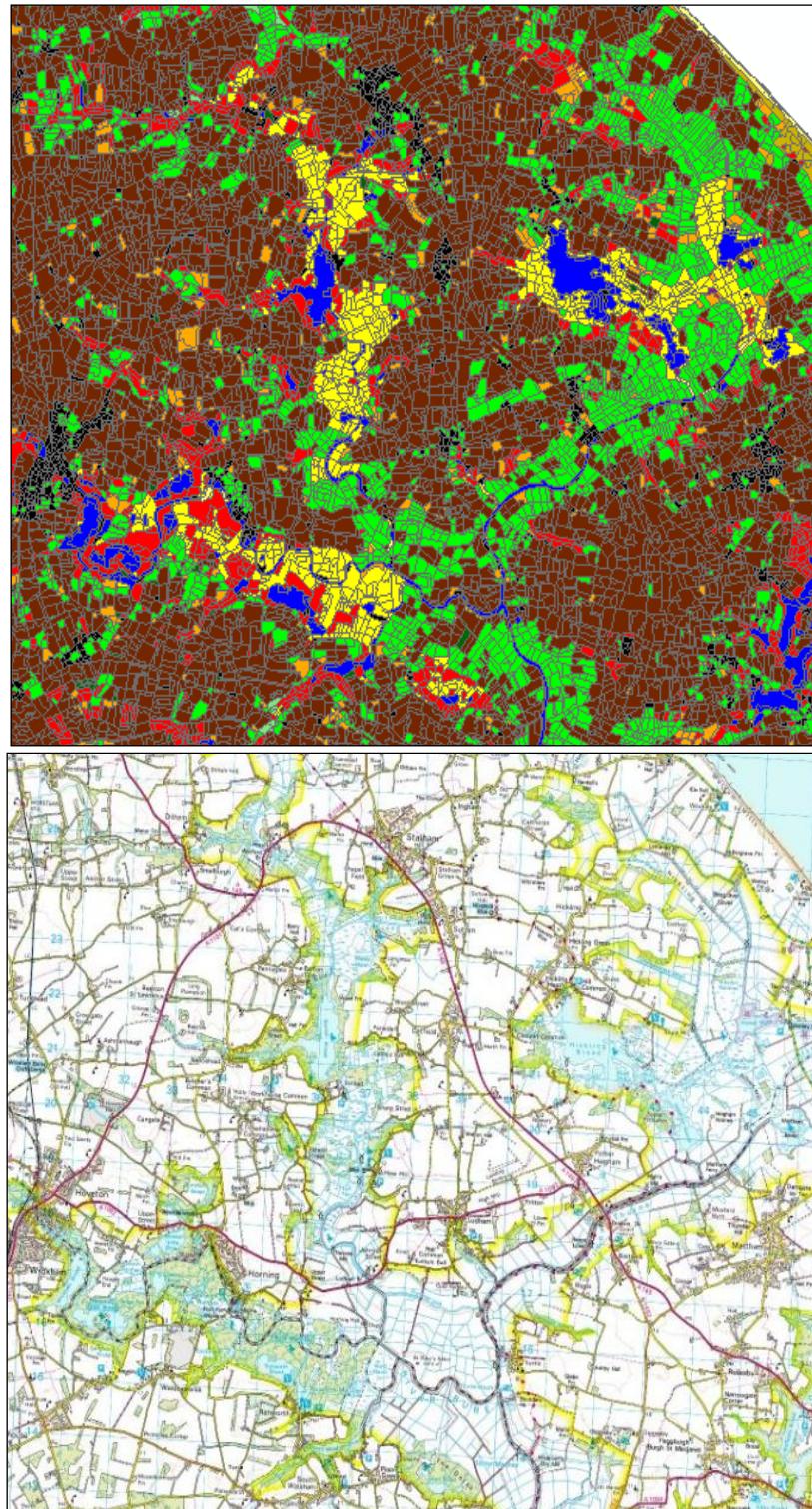

Figure 5.3. LCM2007 map of Salisbury Plain and corresponding OS map area of Wiltshire, approx 19km x 14km. © NERC (CEH) 2011. © Crown Copyright. 2007. Ordnance Survey Licence number 100017572.

Figure 5.4 shows the mainly suburban western London districts. The London parks can be distinguished; Richmond Park, Osterley Park and Kew Botanical Gardens to the south-east, as well as Heathrow airport to the south-west.
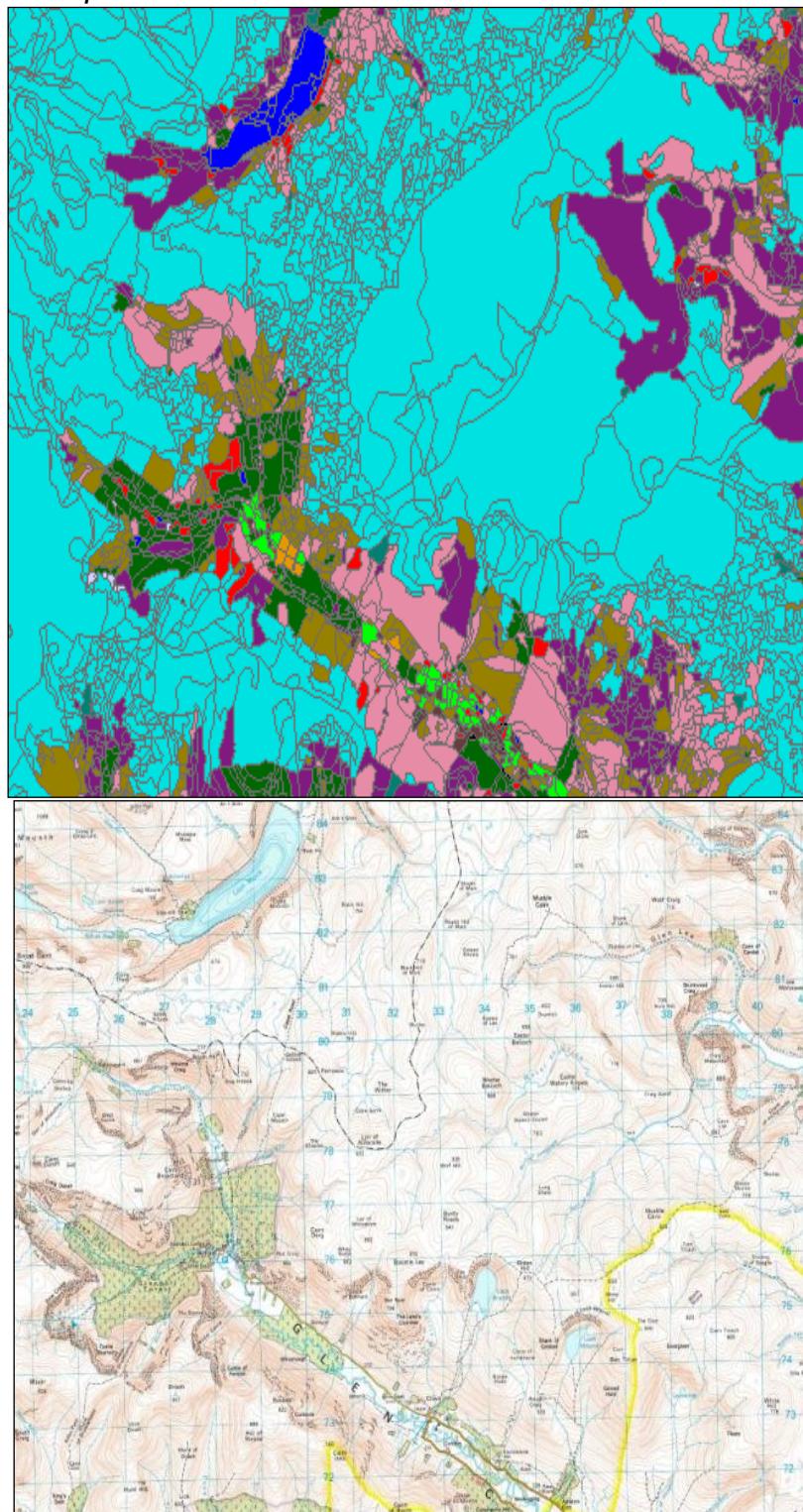

Figure 5.4. LCM2007 map of West London and corresponding OS map area, approx 20km x 14km. © NERC (CEH) 2011. © Crown Copyright. 2007. Ordnance Survey Licence number 100017572.

The majority the area shown in Figure 5.5 is covered by arable land interspersed with fen; namely, Hickling, Barton and Hoverton Great Broad. The land that lies adjacent to the river system is a mixture of '*Improved Grassland*' and **Broadleaved woodland**.

Figure 5.5. LCM2007 example and corresponding OS map area of East Anglia, approx 16km x 13km. © NERC (CEH) 2011. © Crown Copyright. 2007. Ordnance Survey Licence number 100017572.

Figure 5.6 shows the Grampian Mountains and is dominated by '*Montane Habitats'*. The lower elevation valley sides are covered in a combination of '*Acid Grassland'*, '*Dwarf Shrub Heath'* and '*Coniferous Woodland'*. At the bottom of Glen Clova valley small parcels of '*Improved Grassland'* can be identified.

Figure 5.6. LCM2007 example and corresponding OS map area of the Grampian Mountains, approx 16km x 13km. © NERC (CEH) 2011. © Crown Copyright. 2007. Ordnance Survey Licence number 100017572.

5.2 Vector and raster products

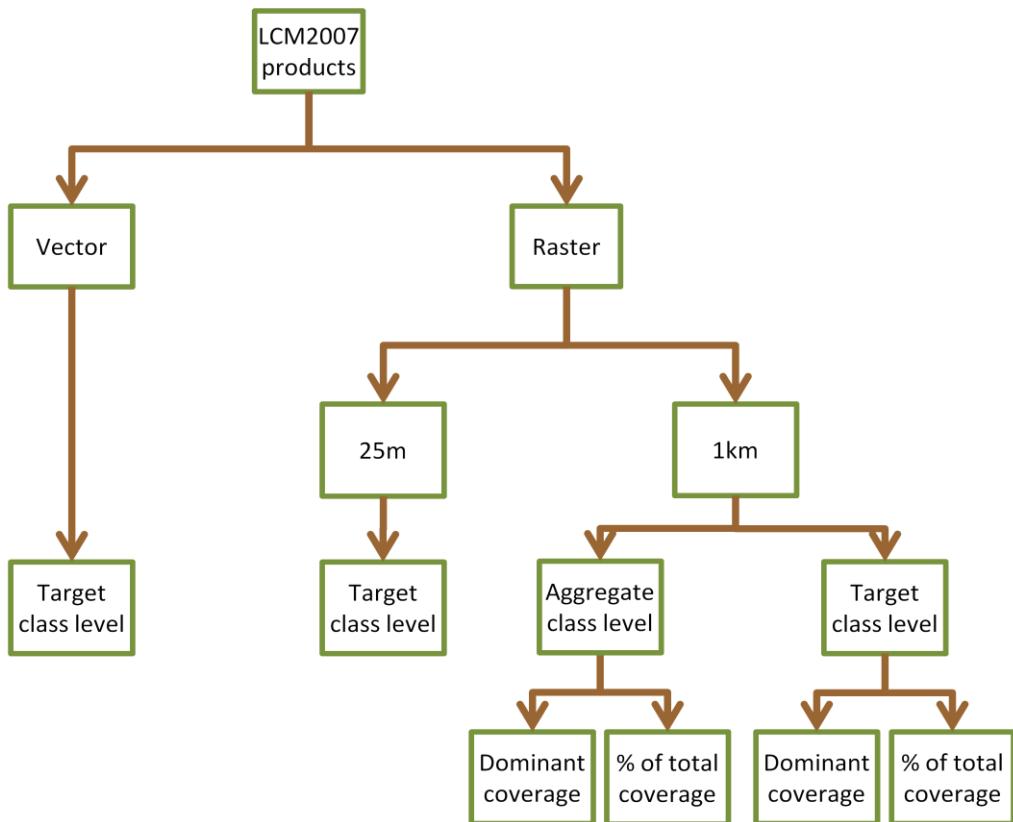A schematic breakdown of the different LCM2007 products and data formats is presented in Figure 5.7.

Figure 5.7. Diagram to show the available LCM2007 data products

Vector data format

The vector data product is provided as polygons (land parcels) with each parcel having a list of attributes attached to it. These include its area, source images, Broad Habitat, and processing details, covering polygon construction, original spectral classification and KBE history (see Table 5.1).

Note: Broad Habitats sub-classes (BHSub attribute in vector data set: Table 5.1) and Field Codes (FieldCode attribute in vector data set: Table 5.1) are included in the vector data set. The BHSub field gives a text description of the FieldCode. Broad Habitat sub-classes are used in the classification process and identify sub-LCM2007 class land cover types. The Broad Habitat sub-classes are aggregated to produce LCM2007 classes (see Table 5.2 for details of groupings). Broad Habitat sub-classes may give additional information, but they are not necessarily recognised with the accuracy or consistency of LCM2007 classes and Broad Habitats. They are included in the data set primarily because they are included in the ProbList attribute (described in Table 5.1), which gives the probability of the top five spectral classes (named at Broad Habitat sub-class level). It is recommended that users apply their own validation (for example the method summarised in Appendix 4) before using data at the Broad Habitat sub-class level if they wish to use it, as it is not covered by the LCM2007 quality assurance.

Raster data format

The raster data have been derived from the vector dataset using the LCM2007 Class number shown in Table 5.2 and are stored as raster datasets at two different resolutions:

1. **25m raster**, consisting of 23 LCM2007 Classes (Table 5.2). See Table 5.3 for metadata.
2. **1km raster** (Table 5.3), data derived from the 25m raster dataset and summarised in two different ways:
   A. **Percentage values,** each 1km pixel provides the percentage cover of a particular land cover at LCM2007 Class level. Multiple sets of data are provided, one for each LCM2007Class.
   B. **Dominant values,** each 1km pixel provides data on the dominant LCM2007 Class at that location.

Table 5.1. Description of the attributes of the LCM2007 vector data set.

Attribute	Description
Parcel_ID	Unique parcel identifier for each parcel, which includes the satellite image the polygon was derived from. Follows the format: 11853977:c20 where 11853977 is the parcel id and c20 means the polygon came from composite image c20. Image numbers may also begin with an 'i', e.g. i76, which signifies a single-date image. All images are documented in Appendix 2.
BH	Dominant land cover at Broad Habitat level e.g. *Coniferous Woodland*
BHSub	Broad Habitat sub-class (Table 5.2). Gives a text description of the FieldCode.
FieldCode	Short text string giving field codes. Note field codes are used in the creation of LCM2007, but the accuracy of the product at this level is not assessed. The LCM2007 team recommend that these codes are used only if users perform their own validation on them.
INTCODE	**RECOMMENDED FOR DISPLAY.** This attribute gives the LCM2007 class as an integer code from 1-23 (see Table 5.2). Note this is often referred to as LCM2007 class number. It is the class recommended for display and is the class validated by the QA. This is the class displayed by the ArcGIS .lyr file.
KBE	Knowledge-based enhancement (KBE). Descriptor detailing the processing history of each segment including the complete list of KBE's applied and the change instigated by the KBE. The KBE notation follows a standard format of 2 or 3 letters followed by a comma followed by original class and then the new class. E.g. • SL,Gr->Bg signifies soil correction (SL) changing Rough Grassland to Bog. This would occur when the polygon was on a bog soil, with no, or only slight, slope. • ALT,Sd->Gr signifies an altitude based correction (ALT) changing sand dune to Rough Grassland. The change to Rough Grassland would be based on the next spectral class is in the probability listing, as long as it was not a coastal class. KBE identifiers: MC - manual correction; SL - soil correction; CM - coastal mask; OM - offshore mask; UR - urban mask; ALT - altitude correction
ProbList	The probability of the polygon belonging to a spectral variant class. The attribute lists the 5 spectral classes that closest match the spectral signature of the polygon e.g. *Gr_d,0.87:Gr_u,0.09:Hga_d,0.04:Gi_n,0.00:Gi_e,0.00* (Table 5.2).
CorePixels	Total number of pixels within the core area of the segment used to perform the maximum likelihood classification.
Construct	History of the construction of the polygon, • OSMM - Polygon derived from generalised OS Master Map data • OSMM:SEG - Generalised OS Master Map polygon(s) as basis with additional segmentation • OSMM:AGC:SEG - Generalised OS Master Map polygon with agricultural payment boundary vectors and segmentation
TotPixels	Total number of pixels in polygon.

Table 5.2. Habitats, LCM2007 classes and Broad Habitat sub-classes for LCM2007.

Broad Habitat	LCM2007 class	LCM2007 class number	Broad Habitat sub-class	Broad Habitat sub-class code (called FieldCode in LCM2007 vector)
'Broadleaved, Mixed and Yew Woodland'	**Broadleaved woodland**	1	Deciduous	D
			Recent (<10yrs)	Dn
			Mixed	M
			Scrub	Sc
'Coniferous Woodland'	*'Coniferous Woodland'*	2	Conifer	C
			Larch	Cl
			Recent (<10yrs)	Cn
			Evergreen	E
			Felled	Fd
'Arable and Horticulture'	*'Arable and Horticulture'*	3	Arable bare	Aba
			Arable Unknown	Aun
			Unknown non-cereal	Aun
			Orchard	O
			Arable barley	Aba
			Arable wheat	Aw
			Arable stubble	Ast
'Improved Grassland'	*'Improved Grassland'*	4	Improved grassland	Gi
			Ley	Gl
			Hay	Gh
	Rough Grassland	5	Rough / unmanaged grassland	Gr
'Neutral Grassland'	*'Neutral Grassland'*	6	Neutral	Gn
'Calcareous Grassland'	*'Calcareous Grassland'*	7	Calcareous	Gc
'Acid Grassland'	**Acid Grassland**	8	Acid	Ga
			Bracken	Br
'Fen, Marsh and Swamp'	*'Fen, Marsh and Swamp'*	9	Fen / swamp	F
'Dwarf Shrub Heath'	**Heather**	10	Heather & dwarf shrub	H
			Burnt heather	Hb
			Gorse	Hg
			Dry heath	Hd
	Heather grassland	11	Heather grass	Hga
'Bog'	*'Bog'*	12	Bog	Bo
			Blanket bog	Bb
			Bog (Grass dom.)	Bg
			Bog (Heather dom.)	Bh
'Montane Habitats'	*'Montane Habitats'*	13	Montane habitats	Z
'Inland Rock'	*'Inland Rock'*	14	Inland rock	Ib
			Despoiled land	Ud
Salt water	**Salt water**	15	Water sea	Ws
			Water estuary	We
Freshwater	**Freshwater**	16	Water flooded	Wf
			Water lake	Wl
			Water River	Wr
'Supra-littoral Rock'	*'Supra-littoral Rock'*	17	Supra littoral rocks	Sr
'Supra-littoral Sediment'	*'Supra-littoral Sediment'*	18	Sand dune	Sd
			Sand dune with shrubs	Sds
			Shingle	Sh
			Shingle vegetated	Shv
'Littoral Rock'	*'Littoral Rock'*	19	Littoral rock	Lr
			Littoral rock / algae	Lra
'Littoral Sediment'	**Littoral sediment**	20	Littoral mud	Lm
			Littoral mud / algae	Lma
			Littoral sand	Ls
	Saltmarsh	21	Saltmarsh	Sm
			Saltmarsh grazing	Smg
'Built-up Areas and Gardens'	**Urban**	22	Bare	Ba
			Urban	U
			Urban industrial	Ui
	Suburban	23	Urban suburban	Us

Table 5.3. Metadata information for the LCM2007 25m and 1km raster data sets.

	Great Britain		Northern Ireland	
	25m	**1km**	**25m**	**1km**
Pixel size	**25m**	**1km**	**25m**	**1km**
Columns / Width (pixels)	28000	700	7800	200
Rows / Height (pixels)	52000	1300	6200	220
Lower left easting (m)	0	0	180000	180000
Lower left northing (m)	0	0	280000	280000
Pixel size (m)	25	1000	25	1000
Data type	Unsigned 8-bit			
Coordinate system	British National Grid		Irish National Grid	
Projection	Transverse Mercator		Transverse Mercator	
Spheroid	Airy		Airy Modified 1849	
Datum	OSGB 1936		Ireland 1965	

Note: Different software packages define coordinates from different parts of the pixel. The values in Table 5.3 refer to the south-west corner of the lower left pixel.

Access to the LCM2007 data sets

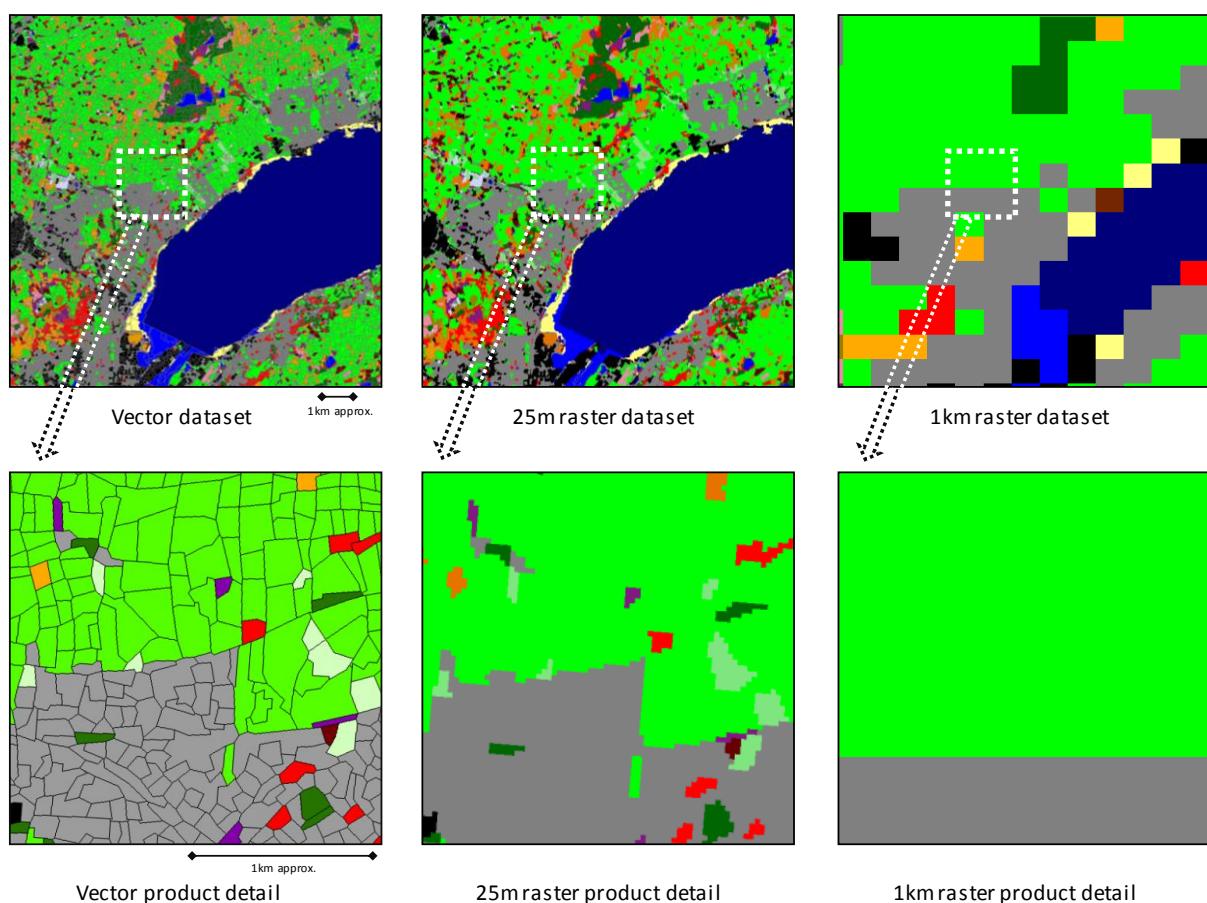The 1km raster data sets are available via the CEH Information Gateway:
https://gateway.ceh.ac.uk

The 25m raster and vector data sets are available under licence from CEH,
please complete the online application on the CEH web site [www.ceh.ac.uk/data]
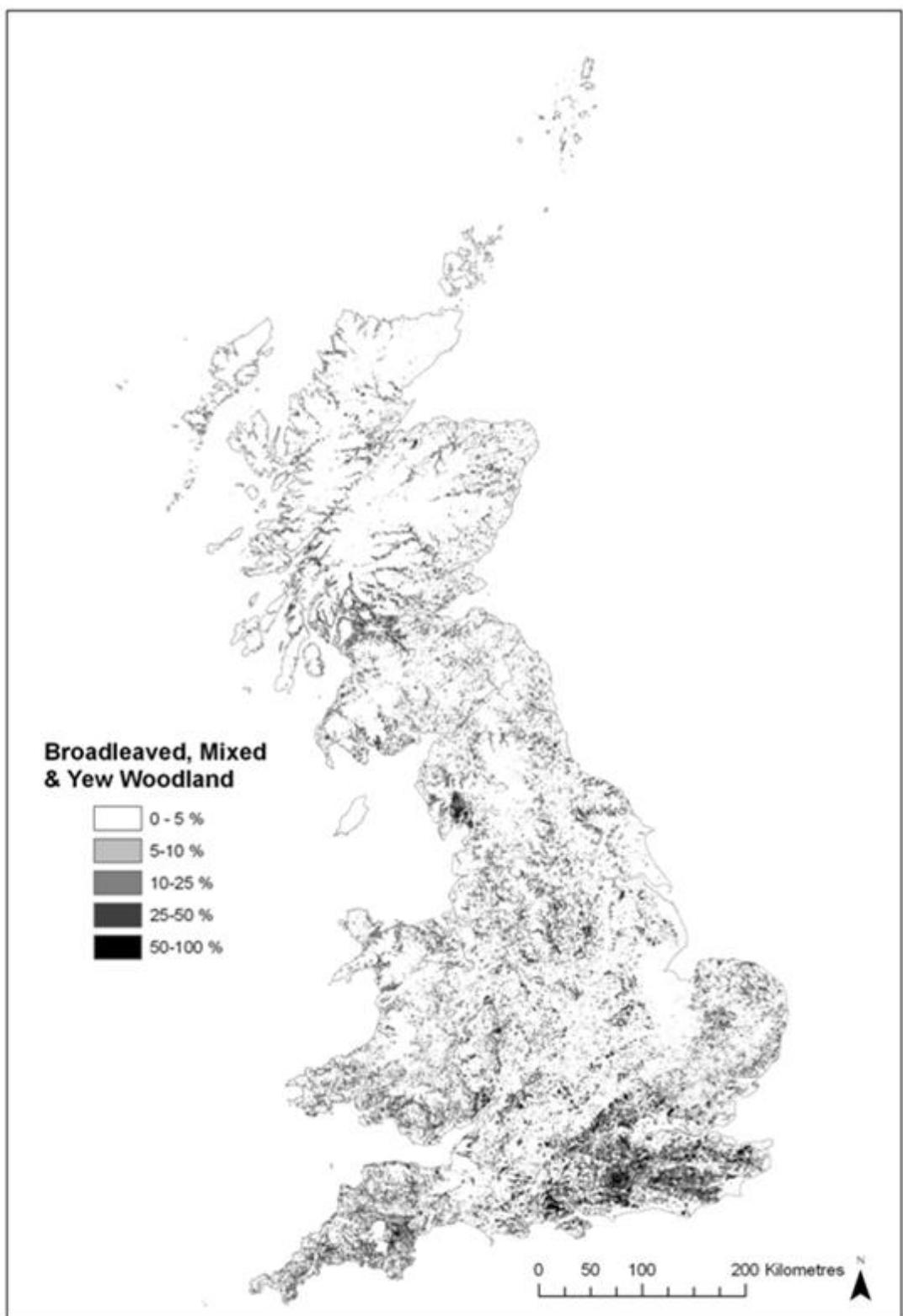or contact: spatialdata@ceh.ac.uk for further details.

Please note that licence fees may apply for some users and some applications.

Example datasets

LCM2007, like LCM1990 and LCM2000, is produced in a range of data formats and at a range of thematic and spatial resolutions. This range of products is required to support a wide range of potential applications. Figure 5.8 shows some examples of the data sets available for LCM2007 and enables comparison of the level of detail associated with each product. Spatially the vector data set and the 25m raster data set are quite similar, however, the advantage of the vector data set is that each polygon has a set of metadata attached. The disadvantage of the vector is that this increases the file size and may make processing unwieldy for some applications/users. The 25m raster data set may be useful in such circumstances, as it provides the same land cover detail but without the additional metadata information and polygon boundaries; this is more appropriate for some applications. The 1km data sets are typically most appropriate for applications modelling the whole of the UK at fairly coarse scale and are often combined with additional data sets, such as meteorological data set or species distribution data. By way of example, the **Broadleaved woodland** percentage land cover 1km raster product is shown in Figure 5.9 - it shows the distribution of broadleaved woodland across the UK.

Vector dataset   1km approx.   25m raster dataset   1km raster dataset

Vector product detail   25m raster product detail   1km raster product detail

Figure 5.8. Comparison of level of detail in the vector dataset, 25m raster dataset and 1km dominant cover raster data set. © NERC (CEH) 2011. © Crown Copyright 2007. Ordnance Survey Licence number 100017572. © third-party licensors.

Figure 5.9. Map showing 1km percentage cover raster product for LCM2007 Broadleaved woodland. © NERC (CEH) 2011. Contains Ordnance Survey data © Crown Copyright 2007, 2009. © third-party licensors.

Chapter 6: Discussion

The Land Cover Map 2007 project has delivered the first continuous parcel-based (polygon) data set for the UK and a suite of derived raster products with 25m and 1km resolution. These products give stock estimates and distributions of Broad Habitat-based classes for the UK, GB, England, Scotland, Wales and Northern Ireland. LCM2007 will support a wide-range of environmental applications and a broad range of users.

LCM2007 has several improvements over the previous CEH land cover maps. A key difference is that the spatial structure has been derived from detailed national cartography (Ordnance Survey MasterMap and LPS Large-scale Vector). This enables LCM2007 to readily integrate with other national products and will therefore increase its uptake and impact. By properly delineating real-world units of land cover, the spatial framework has enabled a more spatially and thematically accurate product than has been previously possible. Another major advantage is that the generalised spatial structure will be re-usable as most of the boundaries in the UK countryside are relatively static compared to changes in land cover. This will increase the efficiency of future national land cover mapping and monitoring exercises, but more importantly, having a common structure for future land cover monitoring activities will facilitate accurate change detection.

Change detection

Identifying, quantifying and understanding land cover change and the ecological effects thereof are essential to meet national and international biodiversity conservation targets. The three CEH land cover maps would seem an ideal resource for this. However, as Chapter 4 demonstrates there are complex issues to be considered when comparing products with different spatial and thematic structures. Comber *et. al.* (2004) have commented on the difficulties of reconciling change between LCM1990 and LCM2000.

LCM1990 is a pixel based product with 25 bespoke habitat classes. LCM2000 is a polygon map based on image segments with habitats related to BAP Broad Habitats. LCM1990 and LCM2000 differ thematically and spatially. LCM2007 is close to LCM1990 thematically but its spatial structure, coming from generalised digital cartography, is very different from the previous two. Clearly comparing CEH land cover products to detect change will not be straightforward; sophisticated methods will be required to delineate real changes from those due to error and methodological differences. It may be possible to make some progress by looking at change in aggregated classes (not necessarily the same ones used in the LCM2007 1km products) or by focussing on the most consistently mapped classes or by using statistical methods that look at the trajectory of change and disregard those that seem implausible. However, it is likely that the most useful way forward will be to re-organise the earlier CEH land cover maps into a common spatial structure based on real-world units of land use. The generalised spatial framework from LCM2007 provides this structure.

LCM2007 and Countryside Survey

LCM2007 has been produced as part of the Countryside Survey (CS) in 2007 project. The field survey of CS and LCM2007 are very different conceptually but complement each other in many ways. The field survey is based on *in situ* observations of broad and fine scale habitat features, for a restricted number of field sites (591 1km x 1km squares in 2007) that are distributed in a stratified random manner across the UK landscape. Statistical procedures are used to extrapolate field observations to produce national estimates for the stock and change of Broad Habitats and it is informative to compare these independent estimates with those produced by LCM2007 (see Chapter 4). Being based on satellite images there are some land cover features that LCM2007 cannot resolve that the CS field survey can. For example, narrow linear features, such as hedgerows and streams, small stands of woodland, narrow roads and small dwellings. The CS field survey can also report on changes in the quality of habitats (e.g. plant species compositions) which is not possible from satellite imagery. However, LCM2007 uniquely provides coast-to-coast coverage of Broad Habitats across the whole of the UK. It would be prohibitively expensive and practically impossible to achieve similar coverage using ground based techniques. Combining *in situ* and satellite derived information is essential to advance our understanding of the UK land surface and to develop optimal strategies for monitoring and managing our land. By considering the evidence-bases of both approaches together and in conjunction with other data sources (see below) our understanding of the UK landscape will be greatly enhanced.

Other sources of information

There are many additional sources of information on UK Land Use/Land Cover. For example, the Forestry Commission maintains detailed maps derived from aerial photography on land used for forestry. There are readily available European land cover and land survey products. For example, Corine Land Cover Maps (1990, 2000 and 2006) derived from satellites, and LUCAS field surveys (2001, 2008 to 2009) respectively. Eurostat and the European Environment Agency are planning to roll out LUCAS, Corine, and five pan-European high resolution land cover layers derived from satellites every three years. To fully understand what is on the land surface, sub-surface (soil and geology) and above surface (climate) information is also useful. Again the UK is very well placed for this with the British Geological Survey maintaining detailed geological inventories; the NSRI, the James Hutton Institute and the AFBI maintaining information on soils; and the Met Office on climate. Future land cover monitoring exercises should aim to develop land information systems that integrate the wealth of information affecting the land surface and present the resulting knowledge in policy relevant ways. This will not be easy, but only by doing so will the UK be able to continue to deliver the socio-economic benefits, supported by a healthy landscape, and fulfil its conservation objectives.

Providing a scientific evidence-base

LCM2007 is a key component in developing a multi-tiered approach to habitat monitoring, with data at a variety of scales (Earth observation, ecological modelling, aerial photography, LiDAR and gound survey) being used to inform evidence-based policy. LCM2007 is the third national-scale digital land cover map produced by CEH and there have been five Countryside Surveys, between 1978 and 2007. The UK is therefore in a very fortunate position. Nowhere else in the world has such rich spatial and temporal data on land cover and habitats. These products provide an invaluable resource to UK scientists, policy makers and managers. CEH's LCMs have been used very widely, for example for the management of natural resources, urban planning, carbon accounting, flood risk modelling and many more. Over 600 data licenses have been granted for LCM1990 and LCM2000. A recent literature search revealed approximately 400 journal publications that have referred to CEH land cover maps. Clearly they have had a huge scientific impact and LCM2007 will continue and enhance this trend.

In conjunction with other data, LCM2007 will be an important evidence-base for policy and other applications in a many sectors including:
• Atmosphere & climate
• Water & catchments
• Ecosystem service assessment
• Marine & coastal
• Ecology & conservation
• Impact assessment
• Health & hazards
• Agriculture
• Landscape planning
• Telecommunications
• Urban studies
• Statistics, information
• Education & publicity
• Carbon accounting

Specific examples of environmental policy areas where LCM2007 may be an important source of information include:
- Providing the most accurate, up-to-date UK land cover map, for use in biodiversity assessment and tracking landscape change.
- Serving as a contextual base map (Broad Habitat level) for higher-resolution habitat mapping in the biodiversity surveillance strategy.
- Enabling habitat network and connectivity analysis to be undertaken, as an input to biodiversity assessment for policy making (e.g. informing woodland expansion), biodiversity reporting (e.g. Natura) and 2020 biodiversity target evaluation.
- Providing a consistent UK frame of reference for informing and evaluating policy, taking account of biogeographical variations and enabling country comparisons.
- Enabling the assessment of the 'naturalness' of river catchments, as part of a standard approach to evaluating rivers for conservation.

European links

The UK component of the Corine Land Cover 2006 (produced on behalf of the European Environment Agency) is being derived from the LCM2007 vector product by a process involving geometric generalisation and thematic transformations. Corine is a key component of pan-European environmental assessment and evidence-based policy making.

LCM2007 benefited from the increase in strategic thinking at a European level, with the European Environment Agency and ESA, amongst others, producing the IMAGE2006 data set for member countries. IMAGE2006 is a pan-European data set of SPOT and IRS data, which provides continuous coverage of Europe at two points in time, to support and build up demand for environmental monitoring based on satellite data sets with approximately 25m resolution. IMAGE2006 data provided the core set of data upon which LCM2007 was based.

Accessing LCM2007 data sets

The LCM2007 1km raster data sets are available via the CEH Information Gateway [https://gateway.ceh.ac.uk]

The full vector product and 25m product are available under licence on request from CEH. Please complete the online application on the CEH web site [www.ceh.ac.uk/data] or contact spatialdata@ceh.ac.uk for further details. Please note that licence fees may apply for some users and some applications.

Glossary of terms and acronyms

Aggregate Classes: There are 10 Aggregate Classes and they are based on combining LCM2007 Classes to a simplified 10-class level (see Table 2.2).

Attribute: Refers here to a data item, held in the geographical information system, recording information about a GIS object; an attribute may be a numerical value (e.g. altitude), an alphabetical code (e.g. Fd = Felled conifer) or a text string labelling or describing the parcel (e.g. the polygon construction attribute).

AWIFS: Advanced Wide Field Sensor, a 56 m spatial resolution, 4 band pushbroom sensor on the IRS platform.

BAP: Biodiversity Action Plan, UK Government's response to the Convention on Biological Diversity (CBD) signed in 1992.

Broad Habitats: A classification by the UK Biodiversity Group to encompass the entire range of UK habitats as an aid to the implementation of, and reporting under, the UK Biodiversity Action Plan. A report providing guidance on the interpretation of the Broad Habitat types is available at: http://jncc.defra.gov.uk/page-2433. For summary of terrestrial Broad Habitats see Chapter 2.

Core pixels: Those pixels of a segment which extracted after shrinking the segment geometry to avoid edge-pixels; they were used in deriving training statistics and/or in deriving a segment's mean reflectance values for use in classification.

CS in 2007: Countryside Survey in 2007.

DTM: Digital Terrain Model.

EO: Earth Observation, field of activity associated with observations of the Earth's surface, usually from satellite sensors.

ETM: Enhanced Thematic Mapper, a sensor on the satellite, Landsat 7, recording visible and infrared reflectance. Damaged in 2003, although still produces some data.

Generalisation: the process of reducing the level of detail in a vector data set.

GB: Great Britain, comprising of England, Scotland and Wales. Also includes the outlying islands of Isle of Wight, Anglesey, Isles of Scilly, the Hebrides, Orkney and Shetland. GB does not include the Channel Islands and the Isle of Man.

GIS: Geographical Information System.

Image classification: Process used to change remote sensing images to classified images. In a classified image each pixel has a class associated with it. Membership of a class is based on the spectral characteristics of the pixel, which is compared to the spectral characteristics of training areas of known class type. The spectral

characteristics of the training areas are used to determine the class membership of each pixel/polygon in the image, using a classification algorithm. The algorithm used for LCM2007 is the Maximum Likelihood Classifier.

Image segmentation: Process of 'breaking up' an image into clusters of pixels with very similar pixel values. It can be used to create polygons from a raster data set.

IRS: Indian Remote sensing Satellite. IRS is a satellite carrying the LISS-III and AWIFS sensors.

Knowledge-based enhancements: Series of rules applied post-classification to aid the determination of Broad Habitat type.

Landsat: US series of satellite platforms which carry the Thematic Mapper (TM) and Enhanced Thematic Mapper (ETM) sensors*.*

LCMGB: Land Cover Map of Great Britain, produced as part of Countryside Survey in 1990. A pixel-based land cover map produced by semi-automated multi-temporal classification and simple knowledge-based enhancement. 25m spatial resolution.

LCM2000: Land Cover Map 2000, produced as part of Countryside Survey in 2000. A segment-based land cover map produced by image segmentation, semi-automated multi-temporal classification and knowledge-based enhancement. 0.5 ha minimum mappable unit.

LCM2007*:* Land Cover Map 2007, produced as part of Countryside Survey in 2007. A parcel-based land cover map, derived from generalised national cartography and classified using multi-temporal image data, followed by post-classification application of a suite of knowledge-based enhancements.

LISS-III*:* Linear Imaging Self Scanner – III, a 23.5 m spatial resolution, four band pushbroom sensor on the IRS platform.

MasterMap: Ordnance Survey MasterMap (OSMM) - a very detailed cartographic product for Great Britain.

MIR: Middle infrared.

MFW: Minimum feature width, the narrowest object, or part of an object that can be mapped within the data set.

MMU: Minimum mappable unit. LCM2007 retains all segments with 9 or more pixels (on the basis that a 3 x 3 pixel segment can contain a 'pure' core pixel); segments with <8 pixels (<0.5 ha in area) are 'dissolved' into surrounding segments, with each pixel individually attached to the neighbouring segment which was most similar in spectral character.

NI: Northern Ireland.

NIR: Near infrared.

Object: A polygon in the LCM2007 GIS database.

OS: Ordnance Survey, the UK national mapping agency, responsible for the production and maintenance of MasterMap.

OSNI: Ordnance Survey of Northern Ireland, national mapping agency for Northern Ireland.

Parcel: A parcel (sometimes called specifically a 'land parcel') is an area on the ground, often a field but perhaps with no boundary such as a woodland, a patch of relatively uniform vegetation (e.g. heath), a built up area, or a water body.

Pre-processing: Early stages of image raster processing, which ensure that atmospheric and topographic effects are minimised within the data and that the data is properly geo-referenced. These stages are conducted before the image is classified.

Producer's accuracy (and Producer's correspondence): Correspondence matrices are often summarised by reporting the User's and Producer's accuracies. The Producer's accuracy is a measure of how well the classification performed against the reference data set (in this report: ground reference polygons in Section 3.9, or CS in 2007 polygons in Section 4.2). In this report the term 'Producer's accuracy' is used in Chapter 3 when LCM2007 is compared to a data set expected to be of higher accuracy than LCM20007. The term 'Producer's correspondence' is used in Chapter 4, when LCM2007 is compared to a data set expected to be of similar accuracy.

Raster: A grid-based data structure used in GIS and image analysis systems.

RPA: Rural Payment Agency, an Executive Agency of the Department for Environment, Food and Rural Affairs (Defra). RPA provides key services for making rural payments, carrying out rural inspections, and livestock tracing.

Segment: The spectrally-defined aggregation of image pixels into a vector polygon, held in the GIS database; the term 'segment' is intended to distinguish the resultant feature from a 'land parcel' which refers to the actual feature on the ground. Often, segments record parcels, but with differences which relate to the underlying 25m structure of the image.

Spatial framework: Is used to describe the network of polygons that make-up the LCM2007 and CS in 2007 vector data sets**.** The spatial frameworks (i.e. the polygons) vary between the two data sets, in part, due to different minimum mappable units.

Spectral classes: Cover types with distinct spectral signatures: for example, 'shaded north-facing' and 'sunlit south-facing' are two distinct spectral classes of 'improved grass'; chalk, clay and peat soil background give different spectral classes of 'wheat'. Spectral classes may differ according to species content, crop variety,

phenology, management practices, atmospheric haze, cloud shadow or any other factor with an impact on the recorded spectral reflectance.

Spot: Satellite pour l'Observation de la Terre - French series of satellites. Some data from SPOT 4 and SPOT 5 were used in LCM2007.

Summer target period: The main growing season for arable crops, from mid-May to late July in southern Britain, or later in Scotland, excluding May but continuing into August; 2007 was the first choice year, then 2008 and 2007.

Target period: See 'Summer target period' and 'Winter target period'.

Training: The procedure by which a sample of known cover types is defined in the image processing system to deduce the spectral characteristics of classes, to form the basis for automatic extrapolation by the system to classify examples of unknown land cover.

TM: Thematic Mapper, a sensor on the Landsat 5 satellite, recording visible and infrared reflectance.

User's accuracy (and Producer's correspondence): Correspondence matrices are often summarised by reporting the User's and Producer's accuracies. The User's accuracy gives the probability that a pixel classified as a class belongs to that class. In this report the term 'User's accuracy' is used in Chapter 3 when LCM2007 is compared to a data set expected to be of higher accuracy than LCM20007. The term 'User's correspondence' is used in Chapter 4, when LCM2007 is compared to a data set expected to be of similar accuracy.

UK: United Kingdom, comprising of England, Scotland, Wales and Northern Ireland. Also includes the outlying islands of Isle of Wight, Anglesey, Isles of Scilly, the Hebrides, Orkney and Shetland. The UK does not include the Channel Islands and the Isle of Man.

Vector: A digital line held as a series of x-, y-coordinates in a geographical information system.

Winter target period: From the time of the first frosts (about October) to late April in southern Britain and well into May in the Scottish Highlands (i.e. until deciduous trees were in full leaf); the target winter was from Autumn 2006 to Spring 2007.

References

Allaby, A., and Allaby, M., _A Dictionary of Earth Sciences_. (1999) Retrieved February 22, 2011 from Encyclopedia.com: http://www.encyclopedia.com/doc/1O13-maximumlikelihoodclssfctn.html

Bradley, A., Thomson, A., Cuevas, M., and Lowe, B., _(_2009) _Raster pre-processing chain for LCM2007 (ENVI)_, Unpublished CEH internal document, pp.62

Bunce, R.G.H., Barr C.J., Clarke R.T., Howard D.C. and Lane A.M.J. (1996) ITE Merlewood Land Classification of Great Britain. _Journal of Biogeography_, **23**, 625-634.

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).

Congalton, R.G., (1991) A review of Assessing the Accuracy of Classifications of Remotely Sensed Data, _Remote Sensing of Environment,_ **37**, pp.35-46

Comber A., Fisher P. and Wadsworth R. (2004). Assessment of a semantic approach to detecting land cover change using inconsistent data sets. Photographic engineering and remote sensing, **70**, pp. 931-938.

DEFRA, (2011) UK timeseries: 1984 - present, http://www.defra.gov.uk/statistics/foodfarm/landuselivestock/junesurvey/junesurveyresults/, [last accessed: April 2011]

Franklin, S.E., and Wulder, M.A., (2002) Remote sensing methods in medium spatial resolution satellite data land cover classifications of large areas, _Progress in Physical Goegraphy, 26, pp. 173-205_

Fuller, R. M., Smith, G. M., Sanderson, J. M., Hill, R. A., Thomson, A. G., Cox, R., Brown, N. J., Clarke, R. T., Rothery, P., and Gerard, F. F., (2002) _Countryside Survey 2000 Module 7. Land Cover Map 2000. Final Report._ NERC/Centre for Ecology and Hydrology. (CEH: Project Report Number C00878) (Unpublished)

Fuller, R.M., Groom, G.B., and Wallis, S.M., (1994) The Availability of Landsat TM images of Great Britain, _International Journal of Remote Sensing_, **15**, pp.1357-1362

Fuller, R.M., and Parsell, R.J., (1990) Classification of TM imagery in the study of land use in lowland Britain: practical considerations for operational use. _International Journal of Remote Sensing_, **11**, pp.1901-1917

Gao Y, and Mas J. F., (2008) A comparison of the performance of pixel-based and object-based classification over images with various spatial resolutions. _Imaging_ (2)8071, pp. 27-35.

Hornung, M., Bull., K.R., Cresser, M., Ullyett, J., Hall., J.R., Langan, S., Loveland, P.J., & Wilson, M.J., (1995) The sensitivity of surface waters of Great Britain to acidification predicted form catchment characteristics, _Environmental Pollution_, **87**, 207-214.

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

Jordan, C. and Higgins, A. (2009). The AFBI Agricultural Land Classification Map of Northern Ireland at 1:50,000 scale. Agri-Food and Biosciences Institute, Belfast. Based on information from Cruickshank, J.G. (ed) 1997, Soil and Environment: Northern Ireland. Agricultural and Environmental Sciences Department, Queen's University, Belfast.

Norton L., Scholefield P., Maskell L., Smart S., Murphy J., Williams P., Scarlett P., Clarke, R., (2009) Countryside Survey Quality Assurance Report: Mapping quality assurance exercise. NERC/Centre for Ecology & Hydrology, 25 pp. (CEH Project Number: C03259).

ONS (2004) Urban area and settlement boundary CD, http://www.statistics.gov.uk/census2001/product_ua_settle_boun_cd.asp [last accessed: 12th May 2011]

Smart, S., Dunbar, M.J., Emmett, B.A., Marks, S., Maskell, L.C., Norton, L.R., Rose, P., and Simpson, I.C., (2010) *An Integrated Assessment of Countryside Survey data to investigate Ecosystem Services in Great Britain*. Technical Report No. 10/07 NERC/Centre for Ecology & Hydrology 230pp. (CEH Project Number: C03259).

Wood, C.M., *et al* (in prep) The influence of devolution on landscape research in Britain: Adjustments to sampling strategies, *journal to be determined*

Appendix 1: Biodiversity Action Plan Broad Habitats

This section provides a brief summary of the JNCC definitions of the Broad Habitats and is based on Jackson (2000).

1. Broadleaved, Mixed and Yew Woodland

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

2. Coniferous Woodland

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

3. Boundaries and Linear Features

This Broad Habitat type covers a range of linearly arranged landscape features such as hedgerows, lines of trees, walls, stone and earth banks, grass strips and dry ditches. These features are not mapped by LCM2007 as they are generally too narrow to be reliably captured from the satellite images. It is included here as its inclusion in Countryside Survey is relevant for the Comparison between LCM2007 and Countryside Survey in Chapter 4.

4. Arable and Horticulture

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

5. Improved Grassland

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

6. Neutral Grassland

This Broad Habitat type is characterised by vegetation dominated by grasses and herbs on a range of neutral soils usually with a pH of between 4.5 and 6.5. It

includes enclosed dry hay meadows and pastures, together with a range of grasslands which are periodically inundated with water or permanently moist.

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

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

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

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

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

12. Bog
This Broad Habitat type covers wetlands that support vegetation that is usually peat-forming and which receive mineral nutrients principally from precipitation rather than ground water. This is referred to as ombrotrophic (rain-fed) mire. The Bog Broad

Habitat includes ericaceous, herbaceous and mossy swards in areas with a peat depth >0.5m.

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

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

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

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

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

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

19. Supralittoral Sediment
'Supralittoral Sediment' occurs above the high water mark, but in areas influenced by wave splash and sea-spray. Includes shingle beaches, sand dunes and machair.

20. Littoral Rock
The geology and wave exposure of the shore influence the form of Littoral Rock habitats, which can be as varied as vertical rock, shore platforms, boulder shores, or

rocky reefs surrounded by areas of sediment. In general, *'Littoral Rock'* tends to be colonised by algae in wave-sheltered conditions, and by limpets, barnacles and mussels as wave-exposure increases.

21. Littoral Sediment

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

22. Inshore Sublittoral Sediment

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

Appendix 2: Satellite Images

Table 1: The complete list of images pre-processed for LCM2007. Note in the Sensors column; L3 is Liss-3, L5 is Landsat-5, Aw is Awifs, S4 is SPOT-4 and S5 is SPOT-5.

Item ID	Sensor	Date	Path	Row	Summer/Winter image	Included in composites:	Composite used in final product	Images used for (single-date) hole filling
i1	LISS-3	10/06/2006	14	31	S	1	Yes	
i2	LISS-3	10/06/2006	14	32	S	1, 38	Yes	
i3	Landsat5	07/11/2005	202	23	W	1, 2	Yes	
i4	Landsat5	07/11/2005	202	24	W	1	Yes	
i5	LISS-3	10/06/2006	14	29	S	2	Yes	
i6	LISS-3	10/06/2006	14	30	S	2	Yes	
i7	Landsat5	07/11/2005	202	22	W	2	Yes	single-date
i8	LISS-3	19/07/2006	17	31	S	3, 10	No	
i9	LISS-3	19/07/2006	17	32	S	3, 10, 36	Yes	
i10	Landsat5	28/01/2006	200	23	W	3	No	
i11	Landsat5	28/01/2006	200	24	W	3	No	
i12	LISS-3	18/07/2006	12	29	S	4, 7, 8	Yes	
i13	LISS-3	18/07/2006	12	30	S	4, 8, 9,55	Yes	
i14	LISS-3	18/07/2006	12	31	S	4, 9	Yes	
i15	Landsat5	09/02/2006	204	22	W	4	Yes	
i16	Landsat5	09/02/2006	204	23	W	4, 6	Yes	
i17	LISS-3	13/07/2006	11	31	S	5	No	
i18	LISS-3	13/07/2006	11	32	S	5	No	
i19	Landsat5	21/11/2005	204	24	W	5, 20	Yes	
i20	LISS-3	13/07/2006	11	30	S	6	No	
i21	LISS-3	18/07/2006	12	28	S	7, 8	Yes	single-date
i22	Landsat5	06/03/2006	203	22	W	7	No	
i23	LISS-3	18/07/2006	12	32	S	9	Yes	single-date
i24	LISS-3	11/05/2006	8	28	S	Failed	No	
i25	LISS-3	11/05/2006	8	29	S	46,48	Yes	
i26	LISS-3	17/07/2006	7	29	S	44, 45, 46, 47,	Yes	single-date
i27	Landsat5	01/11/2006	203	22	W	8	Yes	
i28	Landsat5	01/11/2006	203	23	W	8	Yes	
i29	Landsat5	01/11/2006	203	24	W	9,56	Yes	single-date
i30	Landsat5	01/11/2006	203	25	W	9	Yes	
i31	LISS-3	05/06/2007	14	32	S	12, 13	Yes	
i32	LISS-3	14/07/2005	15	32	S	14, 15	No	
i33	LISS-3	31/08/2005	15	31	S	39	No	
i34	LISS-3	27/04/2006	10	33		Failed	No	
i35	LISS-3	26/05/2005	10	29	S		No	
i36	LISS-3	09/06/2006)	9	27	S	25,	Yes	single-date
i37	LISS-3	09/06/2006	9	28	S	24,	Yes	
i38	LISS-3	11/03/2007	16	32	W	10, 17, 37	Yes	
i39	LISS-3	01/05/2007	7	26	S	42	Yes	single-date
i40	LISS-3	01/05/2007	7	27	S	41	Yes	
i41	LISS-3	27/05/2005	15	31	S	11	No	
i42	LISS-3	11/05/2006	8	28	S	32 , 40	Yes	

i43	LISS-3	11/03/2007	16	31	W	11, 17	Yes	
i44	LISS-3	01/11/2005	13	32	W		No	
i45	LISS-3	09/06/2006	9	26	S	25, 34, 60	Yes	single-date
i46	LISS-3	09/06/2006	9	29	S	24	Yes	
i47	LISS-3	01/05/2007	7	25	S & W	35, 42	Yes	
i48	LISS-3	01/05/2007	7	28	S	40, 41,	Yes	
i49	Landsat5	11/05/2006	201	23	S	16, 17,	Yes	
i50	Landsat5	11/05/2006	201	24	S	16, 17	Yes	
i51	Landsat5	17/06/2006	204	25	S	18	Yes	
i51a						18, 19	Yes	single-date
i52	Landsat5	08/06/2006	205	25	S	18,19	Yes	
i53	LISS-3	16/03/2005	15	32	W	Failed	No	
i54	SPOT-5	06/11/2006	29	244	W	12	Yes	
i55	SPOT-5	19/11/2006	30	245	W	13, 14	Yes	
i56	SPOT-5	9/11/2006	33	243	W	15, 16	Yes	
i57	SPOT-4	15/10/2005	22	242	W	23	No	
i58	SPOT-4	15/10/2005	22	245	W	21	No	
i59	SPOT-4	15/10/2005	22	248	W	18 , 19	Yes	
i60	Landsat5	14/10/2006	205	25	W	22	Yes	
i61	Landsat5	01/10/2007	205	19	W	Failed	No	
i62	Landsat5	01/10/2007	205	20	W	28,50,57	Yes	
i63	Landsat5	01/10/2007	205	21	W	24, 25, 27, 28, 32, 50	Yes	
i64	Landsat5	1/10/2007	205	22	W	24, 27, 32, 50	Yes	
i65	Landsat5	03/05/2007	204	23	S	22, 23	Yes	
i66	Landsat5	03/05/2007	204	24	S	20, 21	Yes	
i67	Landsat5	02/09/2005	204	25	S	18	Yes	
i68	Landsat5	19/07/2006	204	20	S	28, 33	No	
i69	Landsat5	19/07/2006	204	21	S	26, 27, 28, 33	Yes	
i70	Landsat5	19/07/2006	204	22	S	26, 27	Yes	
i71	LISS-3	27/10/2005	12	28	W	26	Yes	
i72	LISS-3	27/10/2005	12	29	W	26	Yes	
i73	AWIFS	10/06/2006	14	28c	S	31, 50	No	
i74	AWIFS	10/06/2006	14	33a	S	30	No	
i75	AWIFS	10/06/2006	14	33b	S	29, 37, 39	Yes	
i76	AWIFS	27/10/2005	12	28c	W	30, 31	No	single-date
i77	AWIFS	27/10/2005	12	28d	W	29, 30, 31	Yes	
i78	AWIFS	27/10/2005	12	33b	W	29 , 30, 36, 38	No	
i79	AWIFS	09/06/2006	9	26c	S	41, 42, 43	Yes	single-date
i80	SPOT-5	03/11/2006	21	233	W	33	No	
i81	SPOT-4	16/02/2006	17	229	W	34, 35, 53, 54	Yes	
i81new	SPOT-4	16/02/2006	17	229	W		No	single-date
i82	SPOT-4	27/09/2007	12	233	S	43	Yes	
i83	LISS-3	01/05/2007	7	29	S	Failed	No	
i84	LISS-3	01/05/2007	7	28	S	Failed	No	
i85	LISS-3	02/04/2007	7	28	S	45, 49	Yes	single-date
i86	LISS-3	2/04/2007	7	29	S	45,49	Yes	single-date
i87	LISS-3	12/04/2007	8	29	S	44	No	single-date
i88	LISS-3	01/05/2007	7	29	S	47, 48, 49	Yes	single-date
i89	AWIFS	09/06/2006	9	26c	S		No	single-date
i90	SPOT-4	23/09/2006	18	227	S		No	single-date
i91	SPOT-4	07/05/2006	18	225	S	51	Yes	single-date

i92	SPOT-4	16/02/2007	18	225	W	51	Yes	
i93	SPOT-5	15/07/2006	18	229	S	53	Yes	single-date
i94	SPOT-4	10/05/2006	18	232	S	54	Yes	
i95	SPOT-4	10/05/2006	21	232	S	52, 57	Yes	
i96	SPOT-4	24/10/2007	21	233	W	52	Yes	
i97	LISS-3	02/05/2007	12	30	W	55	Yes	
i98	SPOT-4	30/04/2006	21	232	W	60	No	
i99	SPOT-4	10/06/2006	17	233	S	58, 59	Yes	
i100	SPOT-4	02/10/2007	18	233	W	59	No	
i101	Landsat5	01/10/2007	205	20	W	25, 58	Yes	single-date

Table 2: Composites created for LCM2007. Note in the Sensors column; L3 is Liss-3, L5 is Landsat-5, Aw is Awifs, S4 is SPOT-4 and S5 is SPOT-5.

Item ID	Item Name	Sensors	Winter image	Summer image
c01	c01l3l5-Midlands	L3, L5	i03(07/11/05); i04(07/11/05);	i01(10/06/06); i02 (10/06/06)
c02	c02l3l5-Lincolnshire	L3, L5	i03(07/11/05); i07(07/11/05);	i05(10/06/06); i06(10/06/06)
c03	c03l3l5-Kent-Anglia	L3, L5	i10(28/01/06); i11(28/01/06)	i08(19/07/06); i09(19/07/06)
c04	c04l3l5-NW-England	L3, L5	i15(09/02/06); i16(09/02/06)	i12(18/07/06); i13(18/07/06); i14(18/07/06)
c05	c05l3l5-S-Wales	L3, L5	i19(21/11/05)	i17(13/07/06); i18(13/07/06)
c06	c06l3l5-N-Wales	L3, L5	i16(09/02/06)	i20(13/07/06)
c07	c07l3l5-Yorkshire	L3, L5	i22(6/03/06)	i12(18/07/06); i21(18/07/06)
c08	c08l3l5-Pennines	L3, L5	i27(1/11/06); i28(1/11/06)	i12(18/07/06); i13(18/07/08); i21(18/07/06)
c09	c09l3l5-Midlands	L3, L5	i29(1/11/06); i30(1/11/06);	i13(18/07/06); i14(18/07/06); i23(18/07/06)
c10	c10l3l3-Kent	L3, L3	i38(11/03/07)	i08(19/07/06); i09(19/07/06)
c11	c11l3l3-EastAngliaNorth	L3, L3	i43(11/03/07)	i41(27/05/05)
c12	c12l3s5-Dorset	L3, S5	i54(6/11/06)	i31(5/06/07)
c13	c13l3s5-Hampshire01	L3, S5	i55(19/11/06)	i31(5/06/07)
c14	c14l3s5-Hampshire02	L3, S5	i55(19/11/06)	i32(14/07/05)
c15	c15l3s5-Sussex	L3, S5	i56(9/11/06)	i32(14/07/05)
c16	c16l5s5-EastAngliaMid	L5, S5	i56(9/11/06)	i49(11/05/06); i50(11/06/06)
c17	c17l5l3-EastAngliaLondon	L5, S5	i38(11/03/07); i43(11/03/07)	i49(11/05/06); i50(11/06/06)
c18	c18l5l5-Cornwall	L5, L5	i67(2/09/05); i59(15/10/05)	i51(17/06/06); i52(08/06/06)
c19	c19l5s4-Cornwall-	L5, S4	i59(15/10/05)	i52(08/06/06)

	west			
c20	c20l5l5-SWWales	L5, L5	i19(21/11/05)	i66(3/05/07)
c21	c21l5s4-SWWales-fill	L5, S4	i58(15/10/05)	i66(3/05/07)
c22	c22l5l5-NWales2	L5, L5	i60(14/10/06)	i65(3/05/07)
c23	c23l5s4-Nwales2-fill	L5, S4	i57(15/10/05)	i65(3/05/07)
c24	c24l3l5-SWScotland	L3, L5	i63(1/10/07); i64(1/10/07)	i37(09/06/06); i46(09/06/06)
c25	c25l3l5-Highlands	L3, L5	i62(1/10/07); i63(1/10/07)	i36(09/06/06); i45(09/06/06)
c26	c26l5l3-NEEngland	L5, L3	i71(27/10/05); i72(27/10/05)	i69(19/07/06); i70(19/07/06)
c27	c27l5l5-ScotBorder	L5, L5	i63(1/10/07); i64(1/10/07)	i69(19/07/06); i70(19/07/06)
c28	c28l5l5-ScotEast	L5, L5	i62(1/10/07); i63(1/10/07)	i68(19/07/06); i69(19/07/06)
c29	c29awaw-SEEngland	Aw, Aw	i77(27/10/05); i78(27/10/05)	i75(10/06/06)
c30	c30awaw-ScCentEngland	Aw, Aw	i76(27/10/05); i77(27/10/05); i78(27/10/05)	i74(10/06/06)
c31	c31awaw-NorthEngland	Aw, Aw	i76(27/10/05) ; i77(27/10/05)	i73(10/06/06)
c32	c32l3l5-SWScotland	L3, L5	i63(1/10/07); i64(1/10/07)	i42(11/05/06)
c33	c33l5s5-SAberdeen	L5, S5	i80(3/11/06)	i68(19/07/06); i69(19/07/06)
c34	c34l3s4-NEScotland1	L3, S4	i81(16/02/06)	i45(09/06/06)
c35	c35l3s4-NEScotland2	L3, S4	i81(16/02/06)	i47(1/05/07)
c36	c36l3aw-Kent	L3, Aw	i78(27/10/05)	i09(19/07/06)
c37	c37awl3-Sussex	Aw, L3	i38(11/03/07)	i75(10/06/06)
c38	c38l3aw-IsleOfWight	L3, Aw	i78(27/10/05)	i02 (10/06/06)
c39	c39awl3-WestAnglia	Aw, L3	i33(31/08/05)	i75(10/06/06)
c40	c40l3l3-MullofKintyre	L3, L3	i48(1/05/07)	i42(11/05/06)
c41	c41awl3-Argyll	Aw, L3	i40(1/5/07); i48(1/05/07)	i79(9/06/06)
c42	c42awl3-RossCromSuth	Aw, L3	i39(1/5/07); i47(01/05/07)	i79(9/06/06)
c43	c43s4aw-SouthOutHeb	S4, Aw	i82(27/09/07)	i79(9/06/06)
c44	c44l3l3-NICentral1	L3, L3	i87(12/04/07)	i26(17/07/06)
c45	c45l3l3-NIWest1	L3, L3	i85(2/04/07); i86(2/04/07)	i26(17/07/06)
c46	46l3l3-NIWest2	L3, L3	i25(11/05/06);	i26(17/07/06)
c47	c47l3l3-NICentral2	L3, L3	i88(1/05/07)	i26(17/07/06)
c48	c48l3l3-NINorth1	L3, L3	i88(1/05/07)	i25(11/05/06)
c49	c49l3l3-NIWest3	L3, L3	i85(2/04/07); i86(2/04/07);	i88(1/05/07)
c50	c50awl5-SW_Scot	Aw, L5		
c51	c51-Shetlands	S4, S4	i92(16/02/2007)	i91(07/05/2006)
c52	c52-Fraserborough	S4, S4	i96(24/10/2007)	i95(10/05/2006)

c53	c53-Orkney	S4, S5	i81(16/02/2006)	i93(15/07/2006)
c54	c54-Moray_Firth	S4, S4	i81(16/02/2006)	i94(10/05/2006)
c55	c55-Nth_Wales	L3, L3	i97(02/05/2007)	i13(18/07/2006)
c56	c56-banbury	L5, L3	i29(01/11/2006)	i1 (10/06/2006) i2(10/06/2006)
c57	c57-montrose	L5, S4	i101(1/10/2007)	i95(10/05/2006)
c58	c58-Scottish Highlands	L5, S4	i101(1/10/2007)	i99(10/06/2006)
c59	c59-Scotland	S4, S4	i100(2/10/2007)	i99(10/06/2006)
c60	C60-Aberdeenshire	S4, L3	i98(30/04/06)	i45(09/06/06)
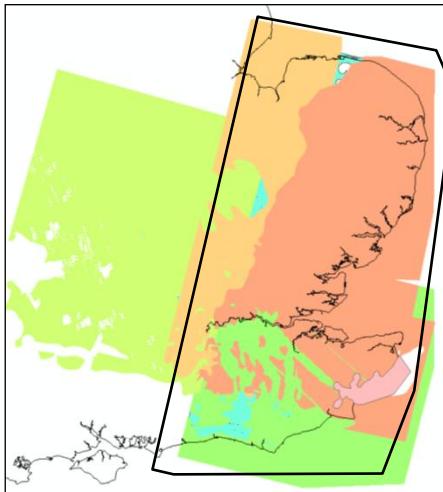
Table 3: Landsat7-ETM+ images available for manual holefilling

Item ID	Sensor	Date	Path	Row
MC1	ETM7	30/05/2009	206	19
MC2	ETM7	30/07/2006	201	23
MC3	ETM7	10/05/2006	202	24
MC4	ETM7	09/06/2006	204	23
MC5	ETM7	22/08/2007	205	21
MC6	ETM7	25/10/2007	205	20
MC7	ETM7	09/06/2006	204	25
MC8	ETM7	22/08/2007	205	20
MC9	ETM7	22/01/2006	205	20
MC10	ETM7	03/11/2008	206	22
MC11	ETM7	30/04/2007	207	22
MC12	ETM7	29/10/2006	206	22

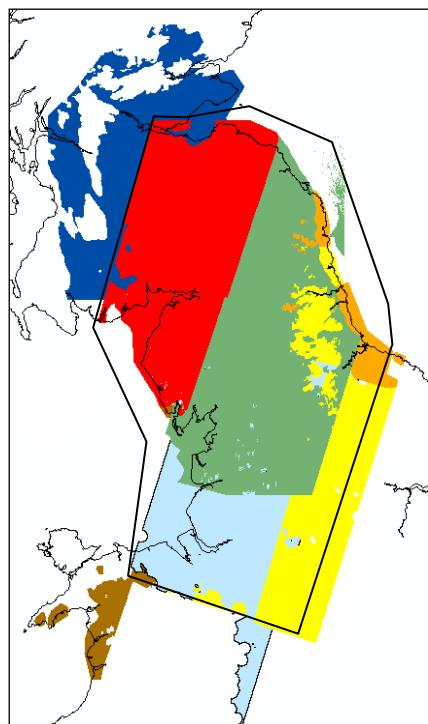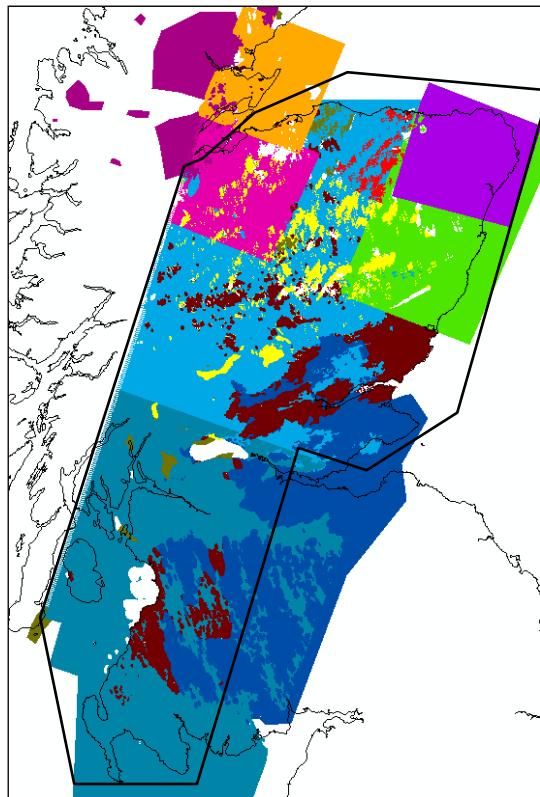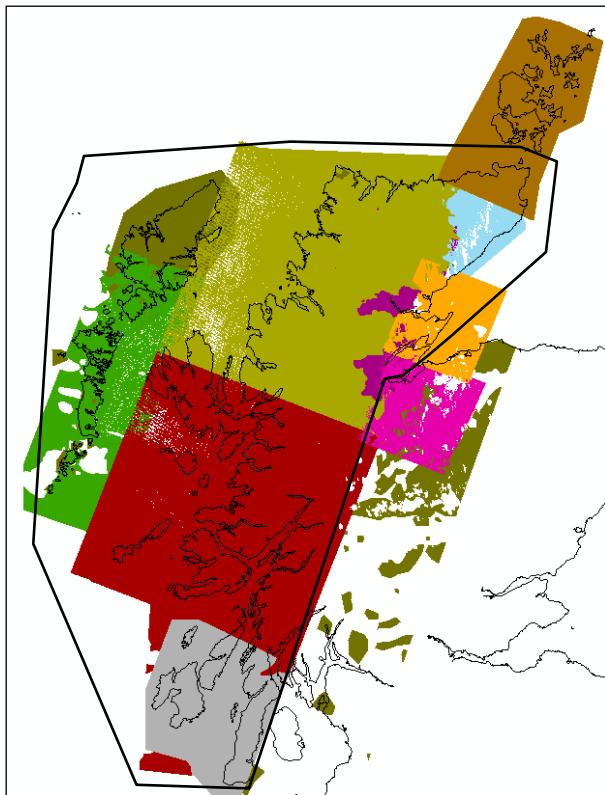# Appendix 3: Creating complete UK coverage

Chunk 1
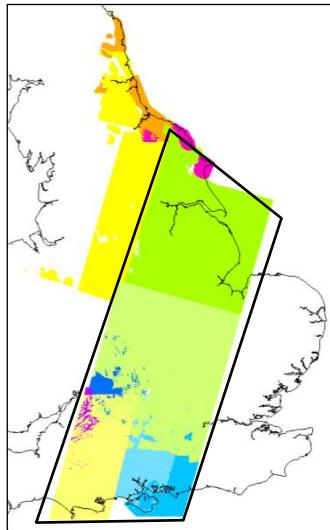
© NERC (CEH) 2011. © Contains Ordnance Survey data © Crown Copyright 2009

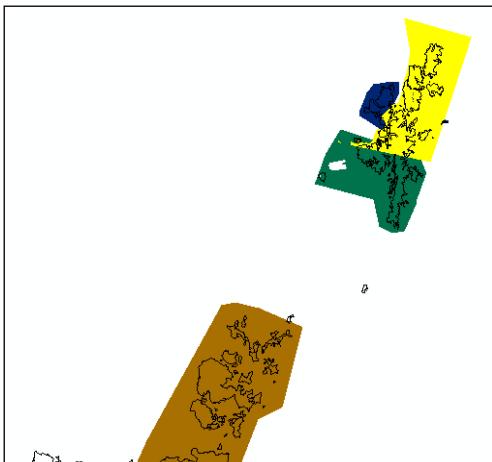| Composite or single-date image	Colour key	Sensor(s)	Summer date	Winter date
C36		L3Aw	19/07/06	27/10/05
C17		L5L3	11/05/06	11/03/07
C16		L5S5	11/05/06	09/11/06
C01		L3L5	10/06/06	07/11/05
C37		L3Aw	10/06/06	11/03/07
C29		AwAw	10/06/06	27/10/05

Chunk 2
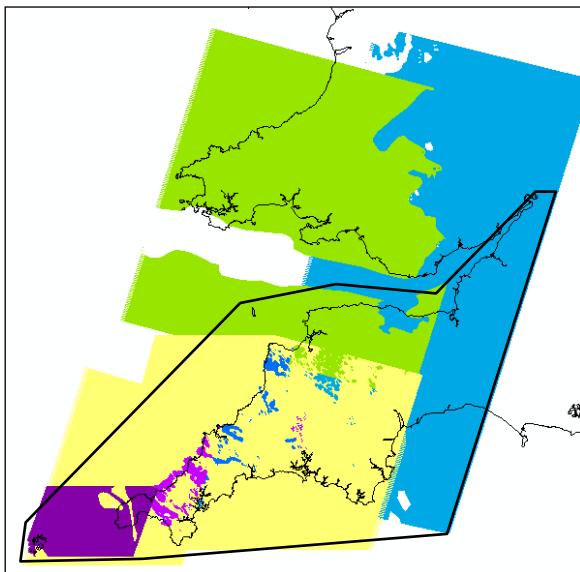
© NERC (CEH) 2011. © Contains Ordnance Survey data © Crown Copyright 2009

Composite or single-date image	Colour key	Sensor(s)	Summer date	Winter date
C12		L3L5	05/06/07	06/11/06
C01		L3L5	10/06/06	07/11/05
C13		L3S5	05/06/07	19/11/06
C56		L3L5	10/06/06	01/11/06
I29		L5	-	01/11/06
I07		L5	-	07/11/05
I21		L3	18/07/06	-
C08		L3L5	18/07/06	09/02/06
C02		L3L5	10/06/06	07/11/05
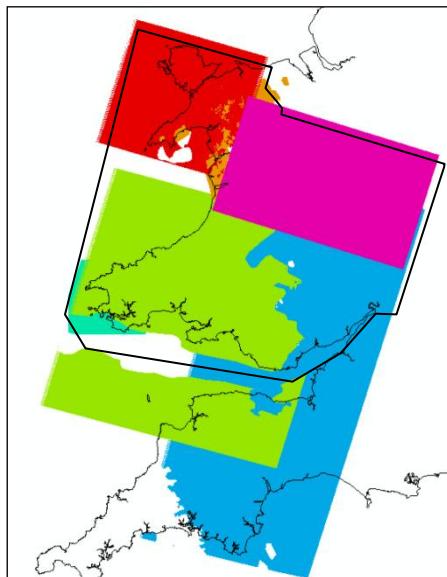
103

Chunk 3

© NERC (CEH) 2011. © Contains Ordnance Survey data © Crown Copyright 2009

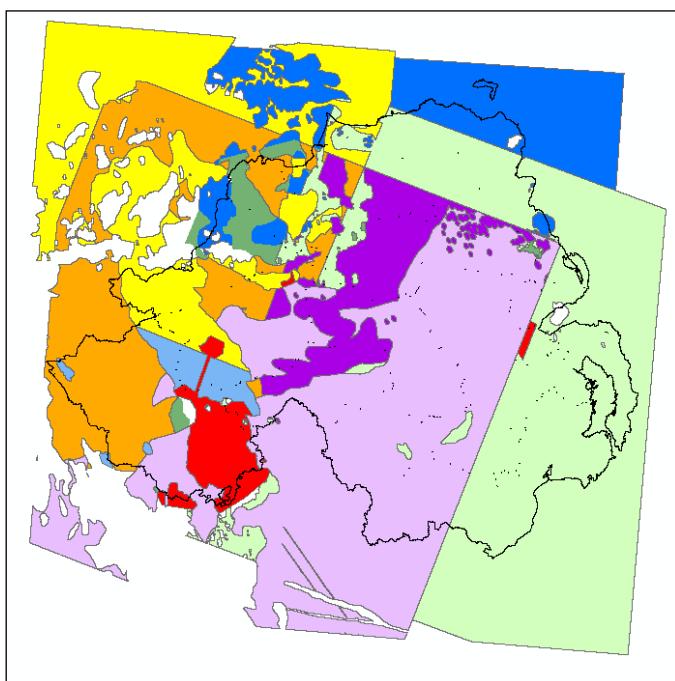| Composite or single-date image	Colour key	Sensor(s)	Summer date	Winter date
C19		L5S4	17/06/06	15/10/05
C18		L5L5	17/06/06	02/09/05
I23		L3	18/07/06	-
I51		L5	17/06/06	-
C20		L5L5	06/05/07	21/11/05
C09		L3L5	18/07/06	01/11/06

Chunk 4

© NERC (CEH) 2011. © Contains Ordnance Survey data © Crown Copyright 2009

Composite or single-date image	Colour key	Sensor(s)	Summer date	Winter date
C55		L3L3	18/07/06	02/05/07
C22		L5L5	03/05/07	14/10/06
I76		Aw	-	27/10/05
C20		L5L5	03/05/07	21/11/05
C09		L3L5	18/07/06	01/11/06
C21		L5S4	03/05/07	15/10/05

Chunk 5

© NERC (CEH)
2011. © Contains
Ordnance Survey
data © Crown
Copyright 2009

Composite or single-date image	Colour key	Sensor(s)	Summer date	Winter date
C27		L5L5	19/07/06	01/10/07
C26		L5L3	19/07/06	27/10/05
I21		L3	18/07/06	-
C08		L3L5	18/07/06	01/11/06
C04		L3L5	18/07/06	09/02/06
I76		Aw	-	27/10/05
C50		AwL5	10/06/06	01/10/07

Chunk 6

© NERC (CEH)
2011. © Contains
Ordnance Survey
data © Crown
Copyright 2009

Composite or single-date image	Colour key	Sensor(s)	Summer date	Winter date
C54		S4S4	10/05/06	16/02/06
C52		S4S4	10/05/06	24/10/07
C58		S3L5	10/06/06	01/10/07
C57		S4L5	10/05/06	01/10/07
C24		L3L5	09/06/06	01/10/07
C25		L3L5	09/06/06	01/10/07
C50		AwL5	10/06/06	01/10/07
I45		L3	09/06/06	-
I39		L3	01/05/07	-
I36		L3	09/06/06	-
I101a		L5	-	01/11/07
I79		Aw	09/06/06	-

Chunk 7

© NERC (CEH)
2011. © Contains
Ordnance Survey
data © Crown
Copyright 2009

Composite or single-date image	Colour key	Sensor(s)	Summer date	Winter date
C54		S4S4	10/05/06	16/02/06
C58		S3L5	10/06/06	01/10/07
C40		L3L3	01/05/07	11/05/06
C41		AwL3	09/06/06	01/05/07
I93		S5	15/07/06	
C42		AwL3	09/06/06	01/05/07
I102		S4	-	16/02/06
C43		S4Aw	09/06/06	27/09/07
I39		L3	01/05/07	-
I79		Aw	09/06/06	-

Chunk 8

© NERC (CEH)
2011. © Contains
Ordnance Survey
data © Crown
Copyright 2009

Composite or single-date image	Colour key	Sensor(s)	Summer date	Winter date
C51		S4S4	07/05/06	16/02/07
I90		S4		23/09/06
I91		S4	07/05/06	
I93		S5	15/07/06	

Chunk 9

© NERC (CEH)
2011.

Composite or single-date image	Colour key	Sensor(s)	Summer date	Winter date
C47			01/05/07	11/05/06
C48			01/05/07	11/05/06
C49			02/04/07	01/05/07
I26			17/07/06	-
I85			-	02/04/07
I86			-	02/04/07
I87			-	12/04/07
I88			01/05/2007	-
189			09/06/2006	-

Appendix 4: Bespoke validation

There is no absolute truth in geography and the "quality" of any geographical "product" (such as a land cover map) depends on the use to which it is put. The purpose of an accuracy assessment ought to be whether the product is "fit for purpose", in this case will using the LCM2007 lead to a better result than if it is not used? Unfortunately, we can not anticipate all the possible uses that the land cover map might be put to; therefore we can only supply a general assessment in terms of comparisons with various field surveys (which are themselves subject to similar levels of inconsistencies).

An unusual and very useful feature of the LCM2007 (and its predecessor the LCM2000) is that it provides parcel level metadata. This means that for all parcels we know what (if any) KBE have been applied and the probability of the 5 most likely land cover variants for that parcel (before any KBE has been applied). This metadata can be mined to allow a user to examine the quality of the categories that are relevant to their study in their location (as the quality of the imagery varies from place to place). In essence the proposed methodology make use of high resolution imagery, (such as that found on GoogleEarth or similar product) to decide whether the LCM2007 attribution is consistent with the imagery. This can be considered as an informal Bayesian approach. It is not always possibly to be definitive as to whether the image is consistent with the LCM2007, so it is suggested that multiple "fuzzy" categories, such as: "plausible", "probably", "possibly", are used. Doing so allows an upper and lower estimate to be made on the accuracy of each target class. The methodology suggests assessing each target classes (relevant to your study) in two parts; parcels with a high probability of belonging to the class and those with a moderate probability. A research tool based on Python scripts has been developed.

Appendix 5: Broad Habitat Association (BHA)

This Appendix summarises the rationale behind the Broad Habitat Association rules (shown in Table 4.1) on a class-by-class basis.

Bog
Issue: *'Bog'* forms part of an ecological continuum covering *'Acid Grassland'*, *'Dwarf Shrub Heath'* and some types of *'Fen, Marsh and Swamp'* and the separation of these habitats can be difficult, as the surface vegetation (i.e. land cover) maybe very similar and the division rests on the depth of peat. The division in the field can account for species presence, plus peat depth, but for LCM2007 the division is based on soil data sets.
Rule*:* LCM classifications of *'Bog'* provide an acceptable correspondence, with CS classifications of *'Dwarf Shrub Heath'*, *'Acid Grassland'*, and *'Montane Habitats'*. This rule is reciprocal, so LCM classifications of *'Dwarf Shrub Heath'*, *'Acid Grassland'* and *'Montane Habitats'* are an acceptable correspondence with CS classifications of *'Bog'*.

Dwarf Shrub Heath
Issue Upland environments containing *'Acid Grassland'* and *'Dwarf Shrub Heath'* may form an ecological continuum from one habitat to another, or a mosaic of both habitats. In such cases it is then difficult to separate the area into discrete polygons of *'Dwarf Shrub Heath'* or *'Acid Grassland'*. The comparison between LCM2007 and CS in 2007 is further compounded by differences in the spatial structure of the two products. The spatial structure affects transitional classes as the size of the polygon affects the proportion of heather and gorse to acid grassland, which determines whether the polygon is classified as *'Dwarf Shrub Heath'* (> 25% of heather or dwarf gorse), or *'Acid Grassland'* (<25% heather or gorse).
Rule*:* LCM2007 *'Dwarf Shrub Heath'* is acceptable correspondence with CS *"Acid grassland"* and reciprocally LCM2007 *"Acid grassland"* is acceptable for CS *'Dwarf Shrub Heath'*.

Montane Habitats
Issue: LCM2007 classifies upland areas based on spectral characteristics, after which a KBE is applied to identify montane areas on the basis of altitude and the presence of vegetation. *'Montane Habitats'* thus replaces polygons which would otherwise be classified as *'Acid Grassland'*, *'Dwarf Shrub Heath'* and *'Bog'*. In most cases this works well, but LCM2007 is unable to separate areas above the montane threshold which are genuinely *'Dwarf Shrub Heath'*, *'Bog'* or *'Acid Grassland'*, rather than *'Montane Habitats'*.
Rule*:* An LCM2007 classification of *'Montane Habitats'* is an acceptable correspondence, at the BH adjacency level, with CS classifications of *'Dwarf Shrub Heath'*, *'Acid Grassland'* and *'Bog'*. This rule is reciprocal, as LCM2007 will not record any montane below the montane height threshold.

Built up Areas and Gardens
Issue*:* LCM2007 allows a variety of land cover types to occur within the urban boundary, whereas CS map polygons as urban, but attach other attributes, which

may identify the land cover type as water, grassland, woodland. In CS areas of grassland > 1ha are mapped as improved grassland, if occurring with an urban area. *Rule*: when CS has recorded a parcel as '*Built-up Areas and Gardens*' it forms an acceptable correspondence if LCM has recognised the same parcel as being grassland, woodland or water. This is not reciprocal.

'*Fen, Marsh and Swamp*'
Issue: '*Fen, Marsh and Swamp*' is characterised by a variety of vegetation types and land cover types, which are often found in a mosaic. This causes 2 problems for a spectral classification: first, parcel-based spectral classifications perform poorly with polygons of mixed land cover. Second, if the '*Fen, Marsh and Swamp*' is large then it may cover several polygons, which may be classified as **Freshwater**, '*Neutral Grassland*', **Rough grassland**, '*Fen, Marsh and Swamp*' and scrub (falls into the **Broadleaved woodland** class).
Rule: if LCM has recorded the polygon as **Freshwater**, **Acid grassland**, **Rough grassland**, and **Broad-leaved woodland** then it forms an acceptable correspondence if CS has recorded '*Fen, Marsh and Swamp*'. This is not reciprocal*.

Rough grassland
Issue: The LCM2007 class **Rough grassland** contains low productivity '*Improved Grassland*' and acid, neutral and calcareous grassland, which could not be assigned with confidence to specific grassland Broad Habitats. For the purposes of Broad Habitat comparison it is nominally assigned to the '**Neutral grassland'** class (as in LCM2000, see Fuller *et al.*, 2002), however for the Broad Habitat Association correspondence it is more appropriate to accept that it contains contributions from a range of grassland BH.
Rule: accept LCM2007 **Rough grassland** if it corresponds with '*Improved Grassland*' and Acid, Neutral or Calcareous grasslands or '*Fen, Marsh and Swamp*'.

Water classes
Issue: LCM2007 maps water, with assignment to **Saltwater** or **Freshwater** based on spatial context.
Rule: LCM2007 **Saltwater** or **Freshwater** are a suitable correspondence with all CS water classes. This rule is not reciprocal.

Tidal areas
Issue: the tidal position at the time of satellite data acquisition or field survey affects the mapping of the saltwater and '*Littoral Rock*' and sediment classes.
Rule: LCM2007 **Saltwater** and CS '*Littoral Rock*' and sediment are acceptable correspondences. This rule is reciprocal.

For further information on Countryside Survey see www.countrysidesurvey.org.uk

Or contact: Countryside Survey Project Office, Centre for Ecology & Hydrology,

Lancaster Environment Centre, Library Avenue, Bailrigg, Lancaster LA1 4AP

Telephone: 01524 595811; *Email:* enquiries@ceh.ac.uk

Countryside Survey

Further information on Countryside Survey can be found at:

www.countrysidesurvey.org.uk