

Posen, Paulette E.; Hutchins, Michael G.; Lovett, Andrew A.; Davies, Helen N.

Working Paper

Robust interpolation of agricultural census data to hydrological units and implications for diffuse pollution modelling

CSERGE Working Paper EDM, No. 09-08

Provided in Cooperation with:

The Centre for Social and Economic Research on the Global Environment (CSERGE), University of East Anglia

Suggested Citation: Posen, Paulette E.; Hutchins, Michael G.; Lovett, Andrew A.; Davies, Helen N. (2009) : Robust interpolation of agricultural census data to hydrological units and implications for diffuse pollution modelling, CSERGE Working Paper EDM, No. 09-08, University of East Anglia, The Centre for Social and Economic Research on the Global Environment (CSERGE), Norwich

This Version is available at:

<https://hdl.handle.net/10419/48791>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

**Robust interpolation of agricultural census data to
hydrological units and implications for diffuse
pollution modelling**

By

Paulette E. Posen, Michael G. Hutchins,
Andrew A. Lovett and Helen N. Davies

CSERGE Working Paper EDM 09-08

ROBUST INTERPOLATION OF AGRICULTURAL CENSUS DATA TO HYDROLOGICAL UNITS AND IMPLICATIONS FOR DIFFUSE POLLUTION MODELLING

*Paulette E. Posen¹, Michael G. Hutchins², Andrew A. Lovett¹ and Helen N. Davies²

¹ School of Environmental Sciences, University of East Anglia

² Centre for Ecology and Hydrology, Wallingford

*email: p.posen@uea.ac.uk

ISSN 0967-8875

Abstract

Diffuse pollution from agriculture is often responsible for observed concentrations of agricultural compounds being in excess of the upper limits prescribed by the EU Water Framework Directive (WFD) in some river catchments and reductions in these concentrations will require widespread changes in farm practice. One of the aims of the UK RELU Catchment Hydrology, Resources, Economics and Management (ChREAM) study is to assess likely impacts of WFD implementation on agricultural land use and consequent implications for water quality and farm incomes. An element of this work involves updating an existing diffuse pollution model to reflect present-day land use profiles. Combining agricultural land use data with hydrological spatial units can involve a number of problems arising from the integration of a variety of data formats at a range of spatial and temporal resolutions and the aggregation of source data over different spatial extents. This paper assesses uncertainty arising from areal interpolation of agricultural census data to hydrological units. The work is illustrated through a case study of the River Derwent catchment in north-east England. The study identifies the range of spatial resolutions at which robust estimations of agricultural land use can be made and examines the implications for diffuse pollution modelling.

Keywords: agricultural census, land use, diffuse pollution modelling, Water Framework Directive

Acknowledgements

EDINA at Edinburgh University Data Library and Defra are acknowledged as the sources of the Agricultural Census data. This analysis forms part of ongoing work for the Catchment Hydrology, Resources, Economics and Management (ChREAM) project, funded under the joint ESRC, BBSRC and NERC Rural Economy and Land Use (RELU) programme (award number RES-227-25-0024).

1. INTRODUCTION

1.1 Background

Implementation of the EU Water Framework Directive (WFD) calls for a major restructuring of European water management, based on the natural geographic and hydrological unit of the river basin, with the aim of achieving "good ecological status" in all water bodies by 2015 (CEC 2000). The Environment Agency for England and Wales has identified 11 River Basin Districts, comprising groups of catchments (rivers, streams, lakes and the land draining into them) along with associated groundwater bodies, estuaries, coastal waters and artificial waters such as canals (<http://www.environment-agency.gov.uk>). Each River Basin District will be managed according to the types of water bodies present within them and the

catchment characteristics, which influence their response to pressures placed upon them. Other characterisation criteria relate to factors such as ecologically sensitive areas, important fisheries, or sources of drinking water. The Environment Agency has already completed the first part of the characterisation exercise and identified water bodies at risk of failing to meet WFD targets by 2015. Diffuse pollution from agriculture is known to be a primary cause of excessive levels of compounds such as nitrates, phosphates and pesticides in water bodies, indicating that substantial changes in land management practice will be required to reduce contributions from this source.

One of the aims of the ongoing Catchment Hydrology, Resources, Economics and Management (ChREAM) study (Bateman et al. 2006) is to assess likely impacts of WFD implementation on agricultural land use, and consequent implications for water quality and farm incomes. An element of the study has involved analysis of a number of measures, proposed to the Department for Environment, Food and Rural Affairs (Defra) by Cuttle et al. (2007) that could be used to tackle diffuse pollution from agriculture. An example of one such measure is to convert areas of arable land to un-grazed (extensive) grassland.

Analysis of the likely outcomes (in terms of water quality) of implementing land use change involves combining land use data with hydrological models to determine how agricultural inputs translate into concentrations of agrichemical compounds in water bodies. Combining agricultural land use data with hydrological spatial units can involve a number of problems arising from the integration of a variety of data formats at a range of spatial and temporal resolutions, and the aggregation of source data over different spatial extents (Aalders and Aitkenhead 2006; Defra 2006; Geddes et al. 2003; Huby et al. 2007; Moxey and Allanson 1994; Moxey et al. 1995). For instance, in order to preserve the anonymity of individual farms, agricultural census data for England and Wales are aggregated over the spatial extent of geographical units, previously parishes and currently Super Output Areas (SOAs). These data have been disaggregated by EDINA at Edinburgh University Data Library, to 2 km and 5 km grid resolution 'Agcensus' datasets, based on a 1 km 'Landuse Framework' which defines the spatial extents of seven discrete land use categories (EDINA, <http://edina.ac.uk/agcensus/description.shtml>). In order to provide suitable land use profiles for input into a hydrological model, further manipulation is required to interpolate the Agcensus data to the spatial extent of arable land within each hydrological unit. Due to the way in which the raw agricultural census data were aggregated, there is no true land use information against which to assess the accuracy of this areal interpolation. However, it is possible to evaluate the robustness of the method by testing the sensitivity of interpolation results across a range of spatial scales.

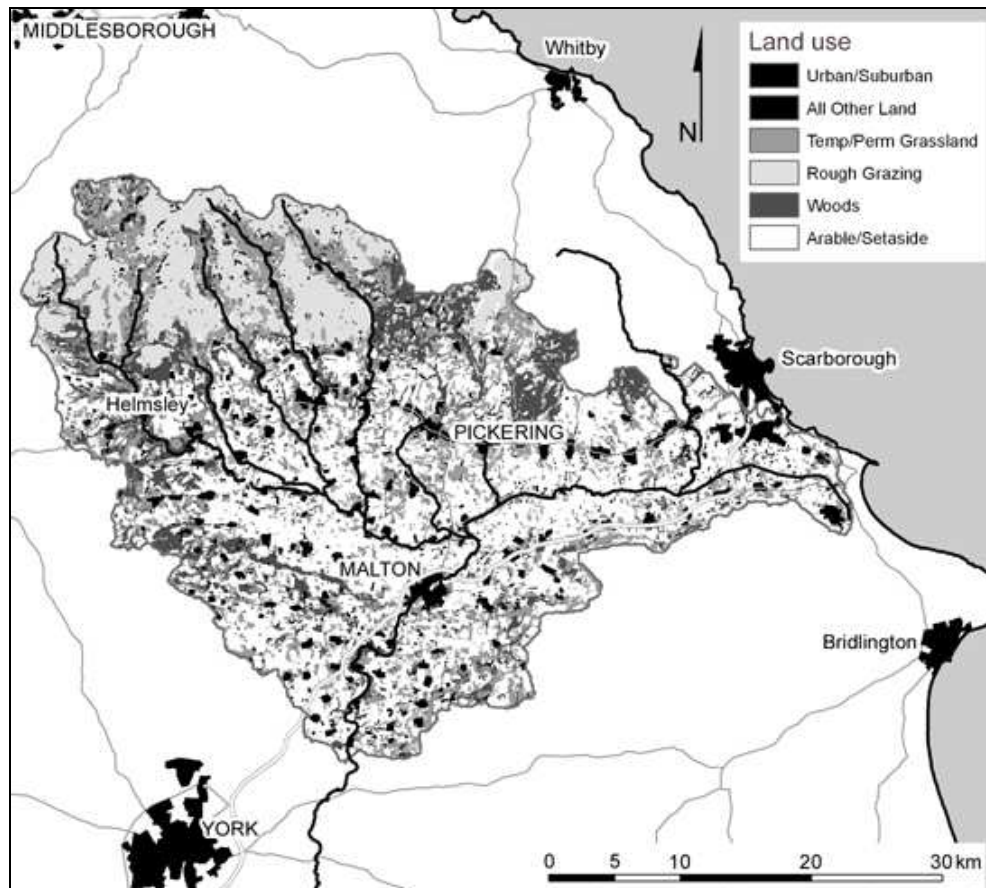
This paper sets out to identify the range of spatial resolutions at which reliable estimations of agricultural land use can be made, and scrutinises the ability to confidently predict the possible outcomes of future, policy-driven, land use change.

1.2 Study Area

The location chosen for detailed analysis was the River Derwent catchment in North Yorkshire, a sub-catchment of the Humber basin, the latter being the main focus of the ChREAM study. The Yorkshire Derwent covers an area of 1600 km², comprising 282 hydrological response units (HRUs) corresponding to areas of land over which surface water drains to discrete river stretches. The catchment includes the rivers Derwent, Rye, Dove, Hertford and their tributaries, encompassing a wide range of topography and land use types, ranging from grazed uplands to lowland arable and small urban areas (Figure 1). Since urban land use occupies less than 8% of the catchment, most of the observed nitrate in rivers can be attributed to diffuse sources. The catchment contains areas of special interest in the context of the Catchment Sensitive Farming (CSF) Programme (Defra 2004), such as the Low Marishes area in the east, which was targeted for more focused analysis in this

study. CSF documentation highlights the Derwent at Low Marishes as being intensively arable with high nitrate and phosphate status.

Figure 1 The Yorkshire Derwent catchment.



1.3 Rationale

The study assessed four different methods of interpolating Agcensus data to hydrological units and compared the results for each method, in terms of areas of agricultural land, at four different spatial scales ranging from a small hydrological unit to the entire Derwent catchment. In this way, it was possible to appraise the suitability of the different methods and to examine whether any method showed particular sensitivity to catchment size. Each of the interpolated land use datasets was then used as input to a diffuse pollution model and comparisons were made between the resulting model outputs in terms of nitrate losses to the river network.

2. THE CASCADE MODEL

Firstly, it was necessary to update land use areas for use in an existing hydrological model employed by the Centre for Ecology and Hydrology (CEH) (Hutchins et al. 2006; Naden et al. 2001). This model comprises two elements, CASCADE and QUESTOR, developed as part of the Land-Ocean Interaction Study (LOIS) (Naden et al. 2001; Neal et al. 2003) to predict fluxes of water, sediment and selected nutrients and contaminants to the North Sea. These models, which use observation data for calibration and adjustment of model parameters, have since been used and tested in numerous studies of surface water quality

(e.g. (Deflandre et al. 2006; Hutchins et al. 2006, 2007; Neal et al. 2006a, b)). This paper focuses on the first model element, CASCADE (CAtchment SCAle DELivery), a spatially distributed water quality model operating at daily time-step, which combines land use data with information on physical parameters such as soil type, geology and rainfall, and predicts the flow of water and pollutants from each HRU to determine the deposition of nutrients into waterways. Sources of pollution considered include nitrogen fertilisers, organic wastes and atmospheric deposition.

The existing CASCADE model was set up by combining the digital ITE Land Cover Map of Great Britain 1990 with parish agricultural census data (MAFF small area statistics) for 1992 to produce series of monthly, crop-specific, loading profiles for nutrients, an example of which is given in Figure 2. The CASCADE analysis is somewhat less aggregated than indicated by Figure 2, and comprises 21 separate land use categories (Table 1).

Figure 2 Loading profiles: Nitrate available for leaching per month for different land uses.

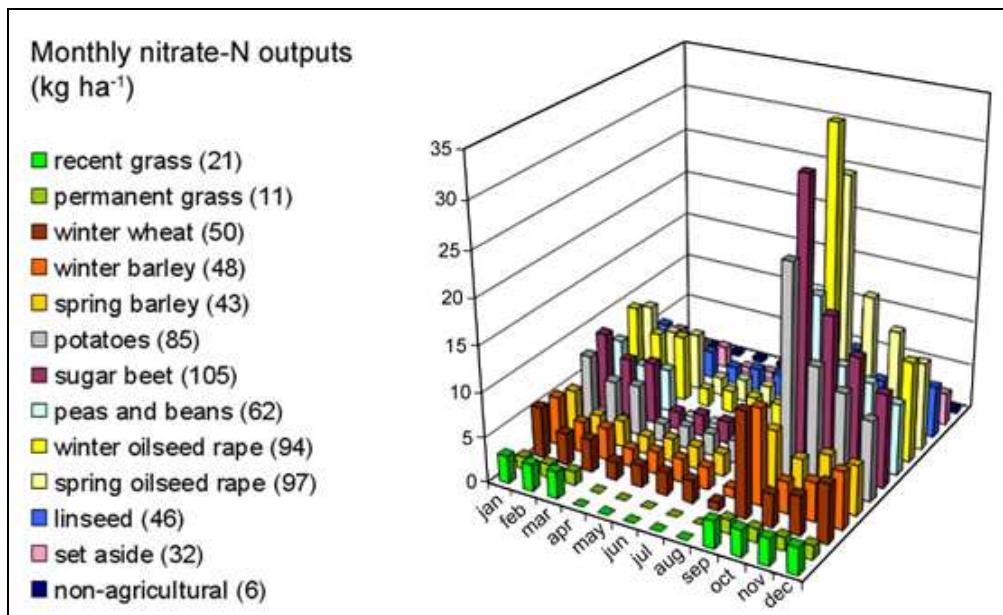


Table 1 Land use classes represented in the CASCADE model.

No.	CASCADE Class	No.	CASCADE Class
1	Recent & temporary grassland	11	Potatoes
2	Permanent grassland	12	Sugar beet
3	Rough grazing	13	Peas & beans
4	Woods	14	Oilseed rape
5	Set aside	15	Linseed
6	All other land	16	Other crops & fallow
7	Winter wheat	17	Vegetables
8	Winter barley	18	Greenhouses
9	Spring barley	19	Fruit
10	Other cereals	20	Suburban
		21	Urban

Nitrate-N (i.e. $\text{NO}_3\text{-N}$: nitrogen as nitrate) inputs from arable land use are calculated based on the concept that nitrate-N available for leaching is supplied through the mineralisation of topsoil organic matter. Two pools of N available for mineralisation are considered: (i) a fraction of readily decomposable material derived from plant residues, and (ii) a larger fraction of more recalcitrant (largely humic) material.

The input values in the current model assume that nitrogen applications are at an economic optimum, but fertiliser inputs change over time due to factors such as changes in land use, cropping patterns and agricultural practice, thereby altering this 'optimum' application rate. As further data on fertiliser use become available (e.g. from the British Survey of Fertiliser Practice (BSFP)) it will be possible to refine the definition of 'optimum' and adjust the monthly input function accordingly. If applications stray from this optimum, the values of Lord (1992) can be applied to define the likely change in crop residue resultant from a unit change in added N. These gradients, which differ above and below the optimum level, are available in the literature (Lord, 1992).

In terms of livestock manure applied to arable land, only cattle manure is considered in the calculation of N inputs. National surveys of fertiliser use reveal that other manure and slurry types (e.g. pig and poultry), although high in N, are only applied to a very small percentage of fields. With respect to poultry manure application there is also high locational uncertainty since there is a significant importing of such material. However, it is possible to build other manures into the model if required. The MANNER model (Chambers et al. 1999) is used to estimate nitrate-N available for leaching from a range of representative application quantities, timings and incorporation routines for both farmyard manure and slurry. For each of the crops considered, MANNER indicated that an additional 2 kg/ha per month available nitrate-N should be included for the months November-January inclusive.

To calculate the nitrate-N contribution from grasslands, agricultural census data are used in conjunction with typical stocking density rates to determine the proportions of grassland used for beef, dairy and cutting systems, respectively. These values are then combined, in the N-CYCLE model (Scholefield et al. 1991), with information on typical fertiliser application rates for grasslands, climate, grassland ages and soil types, and the output is accounted for in the CASCADE loading profiles.

Deposition of atmospheric N has been assumed spatially invariant (at 20 kg/ha) for modelling purposes, and partitioned on a monthly basis using information relating to seasonal variability of atmospheric inputs (Goulding et al. 1998).

2.1 Definition of the 'steady-state' model

CASCADE is currently set up to model 'steady-state' conditions according to inputs of contemporary fertiliser use and land use practice. The dynamics of achieving the steady-state are not yet incorporated in the model, but empirical data are currently used to calculate the time lapse between change in farming practice and the attainment of a new steady-state.

As the existing CASCADE model is based on early 1990s datasets, which may not necessarily reflect current agricultural practice in terms of nutrient inputs and outputs etc, a 'present day' dataset was compiled based on (CEH) Land Cover Map 2000 (LCM 2000) (Fuller et al. 2002) and the Defra Agricultural Census for 2004. A detailed description of the data update process is given in Section 3.

The CASCADE model was run twice, firstly using the original land use profiles from the early 1990s and then again using the updated land use data from 2004. Both model runs used the same meteorological data (from the 1990s) to eliminate climatic variability. The results from these model runs served two purposes:

1. To highlight any significant differences in fertiliser use or land use management between the two 'steady-state' conditions.
2. To evaluate the model by comparing predictions of nitrate-N output for the latest dataset with current observations.

3. UPDATING AND RECLASSIFYING THE LAND USE AREAS

The ITE LCMGB 1990, which was used as a basis for the original CASCADE input data, is classified very differently to LCM 2000, as shown in Table 2. Furthermore, the 21 CASCADE classes (Table 1) had to be matched with their most closely associated respective classes in LCM 2000, following the procedure used in the early 1990s model, in order to facilitate allocation of agricultural land use at the HRU scale. It was necessary, therefore, to group and match both the CASCADE and LCM 2000 land use classes into the six broad categories shown in Table 3.

Table 2 Comparison between ITE LCMGB 1990 and LCM 2000 classifications.

Number	ITE LCMGB 1990 Class	Number	Equivalent LCM 2000 Class
1	sea/estuary	1	sea/estuary
2	inland water	2	inland water
3	beach/coastal bare	3	littoral rock
		4	littoral sediment
4	saltmarsh	5	saltmarsh
		6	supra-littoral rock
		7	supra-littoral sediment
17	upland bog	8	bogs
24	lowland bog		
5	grass heath	9	dwarf shrub heath
8	rough/marsh grass	10	dwarf shrub heath
9	moorland grass	11	montane habitats
10	open shrub moor		
11	dense shrub moor		
12	bracken		
13	dense shrub heath		
25	open shrub heath		
15	deciduous woodland	12	broad-leaved/mixed woodland
16	coniferous woodland	13	coniferous woodland
18	tilled land	14	arable and horticultural
6	mown/grazed turf	15	improved grassland
		16	setaside grass
7	meadow/verge	17	neutral grass
		18	calcareous grass
		19	acid grass
		20	bracken
		21	fen, marsh, swamp
20	suburban	22	suburban and urban
21	continuous urban		
22	inland bare ground	23	inland bare ground
14	scrub/orchard		no direct equivalent
19	ruderal weed		no direct equivalent
23	felled forest		no direct equivalent

Table 3 Land use categories for use in the CASCADE model.

LCM 2000 broad land use categories	CASCADE classes & subdivisions
Temporary/Permanent grassland	Recent, temporary and permanent grassland
Rough grazing	Rough grazing
Woodland	Woods
Arable/Setaside	Arable (subdivided into 13 crop types)/Setaside
Urban/Suburban	Urban/Suburban
All other land	All other land

Two broad categories of land use required special treatment with respect to the matching process, namely the grassland and urban/suburban classes, as described in Sections 3.1 and 3.2.

3.1 Differentiation between agricultural and non-agricultural grassland

Firstly, a combination of visual interpretation and statistical correlation was needed to split the LCM 2000 grassland categories between agricultural and non-agricultural use, based on information from the Agricultural Census. (At the time of the current study, the most recent EDINA Agcensus data available related to the 2004 June Agricultural Survey, so this data set was used to represent ‘present day’ values, in terms of hectares of crops and livestock numbers.) Agricultural grassland areal values from Agcensus data at the spatial extent of the Super Output Area (SOA) were compared with the areal extent of grassland classes identified in LCM 2000, the findings of which are summarised below. The analysis was performed on an area defined by all SOAs intersecting the Yorkshire Derwent catchment (Figure 7).

3.1.1 Spatial correspondence between Agcensus 2004 and LCM 2000 grassland categories for the Yorkshire Derwent catchment and surrounding areas

Good spatial correspondence was observed between the LCM 2000 total grassland area and the Agcensus grasslands, with Pearson Correlation values of +0.968 when Agcensus Class G5 (Rough Grazing) was excluded, and +0.958 when this class was included (Figure 3).

Agcensus Classes G1 and G2 (Temporary and Permanent Grassland, respectively) corresponded most closely with LCM 2000 Classes 5.1, 6.1 and 7.1 (Improved, Neutral and Calcareous Grasslands, respectively), with a Pearson Correlation value of +0.961 (Figure 4).

Weaker, but still significant, correlations were seen between AgCensus G5 (Rough Grazing) and LCM 2000 Classes 6.1 and 8.1 (Neutral and Acid Grasslands) (Figure 5), and a similar level of correlation when AgCensus G5 was plotted against LCM 2000 Classes 8.1 (Acid Grassland) plus Classes 9.1-12.1 (Bracken, Shrub Heath, etc) (Figure 6). The outliers in Figure 6 represent (i) Scarborough 2 SOA, a large part of which falls within the N. York Moors National Park (Figure 7; Figure 8(a)), therefore has a low grazing area compared to LCM grassland/shrub heath area; (ii) Hambleton 1 SOA (Figure 7; Figure 8(b)), where the rough grazing area exceeds the combined LCM acid grassland/shrub heath areas.

Figure 3 Correlation between LCM 2000 total grassland area and Agcensus 2004 grassland classes: (a) excluding Rough Grazing Class G5, and (b) including Rough Grazing Class G5, from the Agcensus data. (N.B. Each point in the correlation plots represents one Super Output Area (SOA) in the Derwent catchment and surrounding area.)

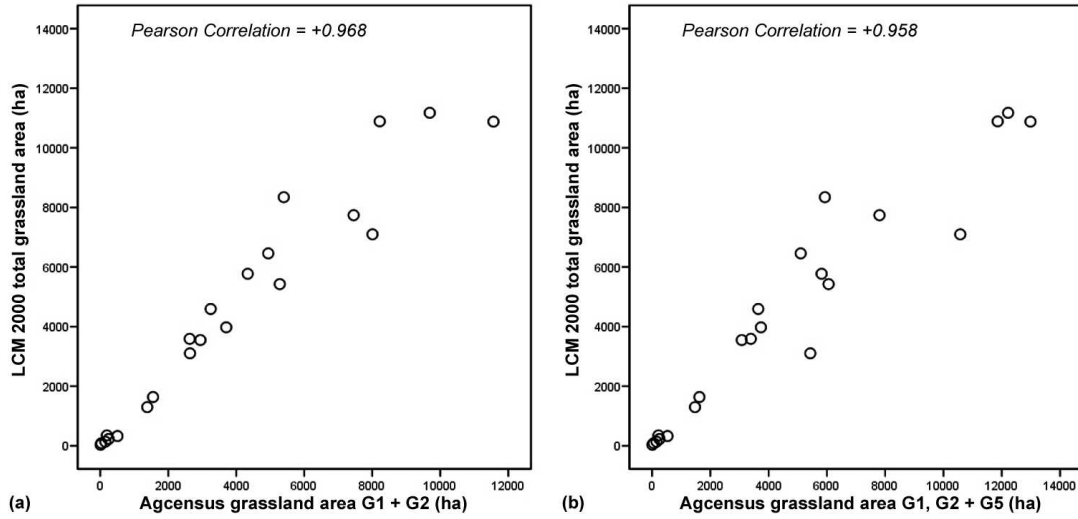


Figure 4 Correlation between LCM 2000 Classes 5.1, 6.1 and 7.1 (Improved, Neutral and Calcareous grasslands, respectively) and Agcensus Classes G1 and G2 (Temporary and Permanent grassland, respectively).

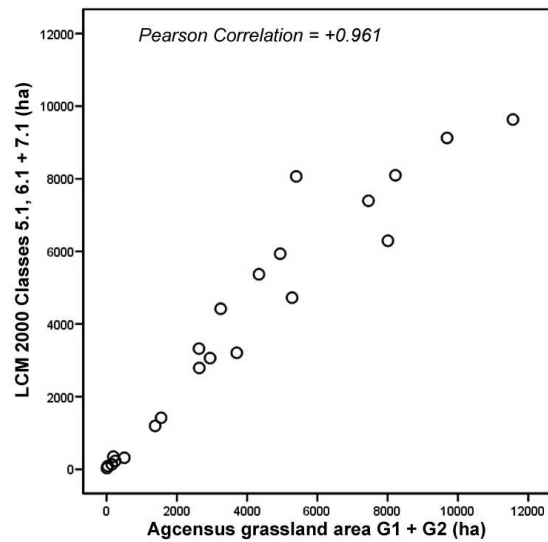


Figure 5 Correlation between LCM 2000 Classes 6.1 and 8.1 (Neutral and Acid grasslands) and AgCensus G5 (Rough Grazing).

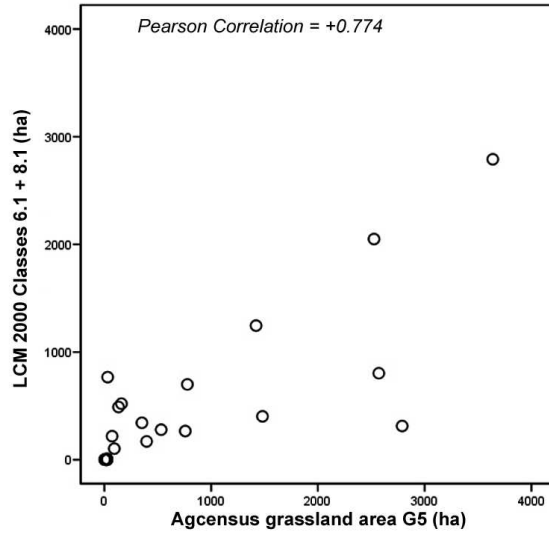


Figure 6 Correlation between LCM 2000 Classes 8.1 (Acid grassland) plus Classes 9.1-12.1 (Bracken, Shrub heath, etc) and AgCensus G5 (Rough Grazing).

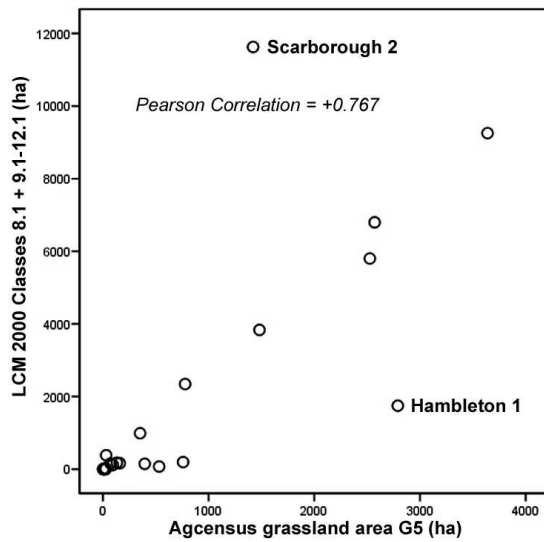


Figure 7 Spatial relationship between the Yorkshire Derwent catchment and associated SOAs, with locations of Scarborough 2 (S2) and Hambleton 1 (H1) SOAs indicated.

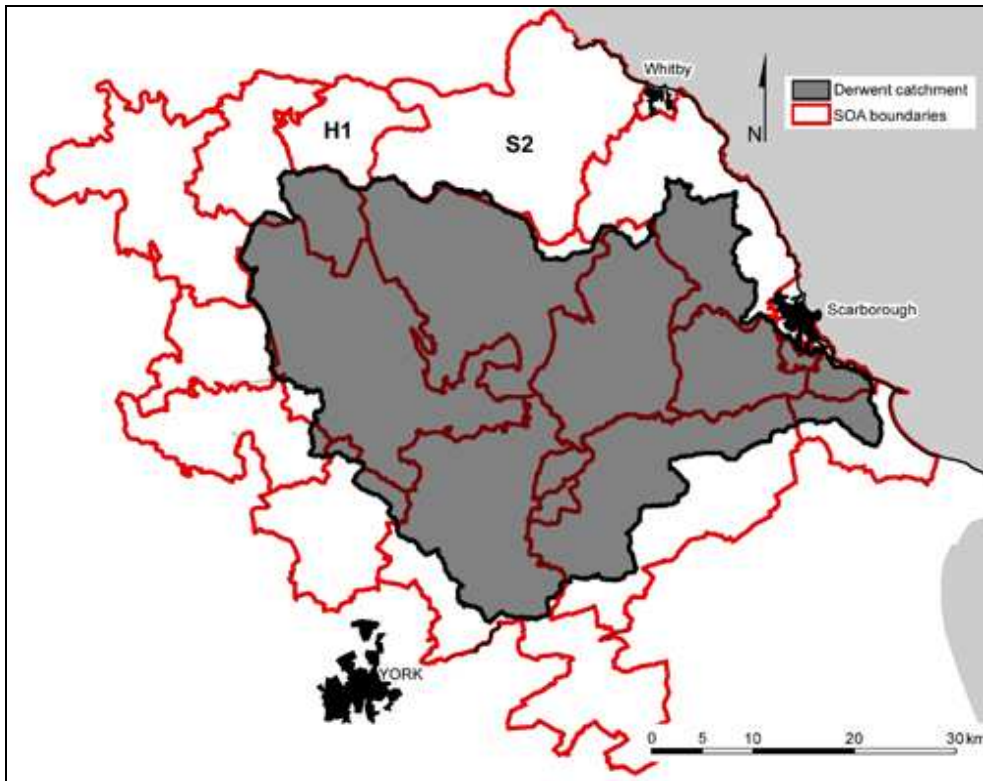
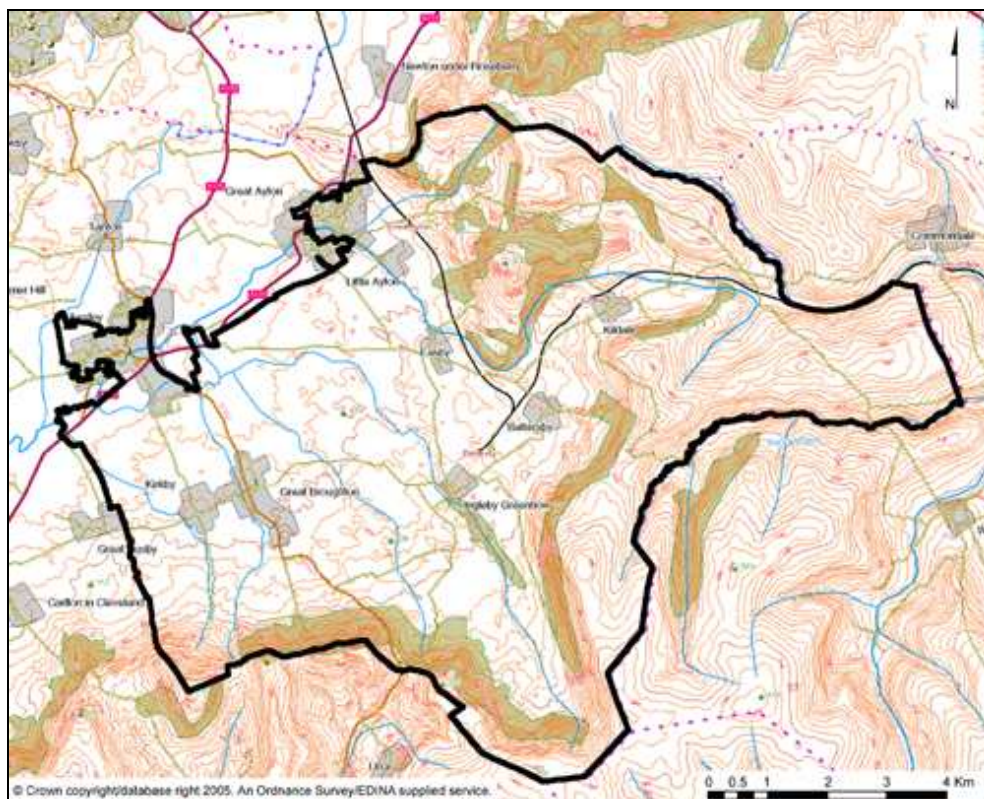


Figure 8 (a) Location of Scarborough 2 SOA (indicated by the solid black boundary), a large part of which falls within the N. York Moors National Park.



Figure 8 (b) Location of Hambleton 1 SOA (indicated by the solid black boundary), where Agcensus 'Rough Grazing' values exceed the 'Acid grass' and 'Heathland' areas identified by LCM2000.



Not surprisingly, the recorded Agcensus grassland areas (G1, G2 and G5) were, in most cases, smaller than the corresponding LCM 2000 areas, since not all grassland is designated farmland.

3.1.2 Distribution of livestock

A visual examination of the spatial distribution of LCM 2000 grassland categories with respect to Agcensus livestock classes suggested a preferential use of (i) Acid and Neutral Grasslands for beef cattle, and (ii) Improved and Calcareous Grasslands for dairy cattle. The latter observation was borne out statistically (Figure 9(a)); however, the best statistical correlation for beef cattle corresponded with Improved Grassland (Figure 9(b)) and Improved Grassland was strongly correlated with the total cattle area (Figure 10). The correlation in Figure 11 indicates that sheep are distributed throughout all combined LCM 2000 grassland categories.

Figure 9 Spatial correlation between (a) LCM 2000 Improved and Calcareous Grasslands and Agcensus dairy cattle numbers, and (b) LCM 2000 Improved Grasslands and Agcensus beef cattle numbers.

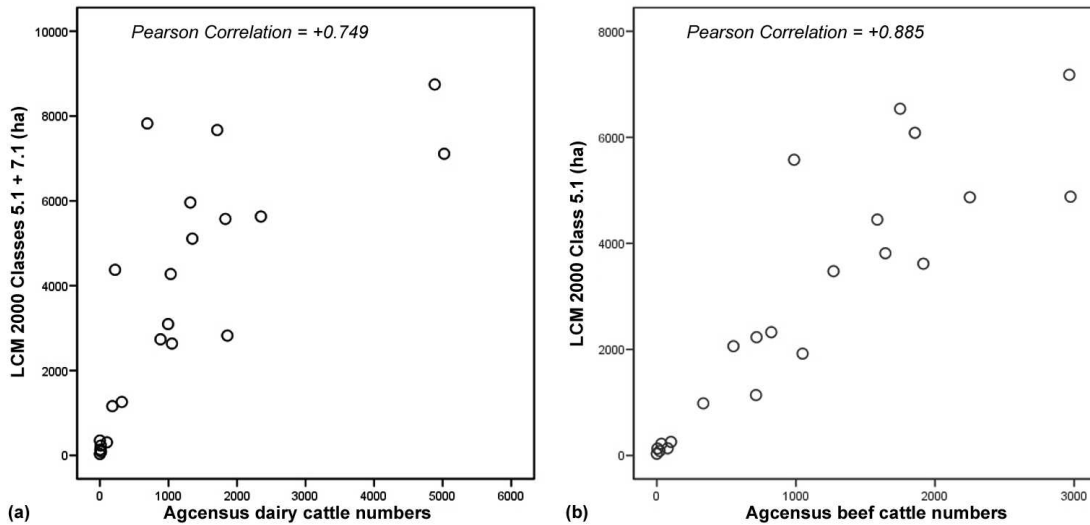


Figure 10 Spatial correlation between LCM 2000 Improved Grassland and Agcensus total cattle numbers.

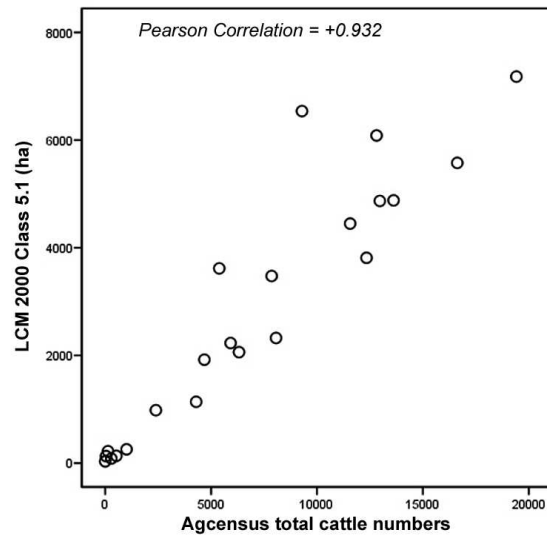
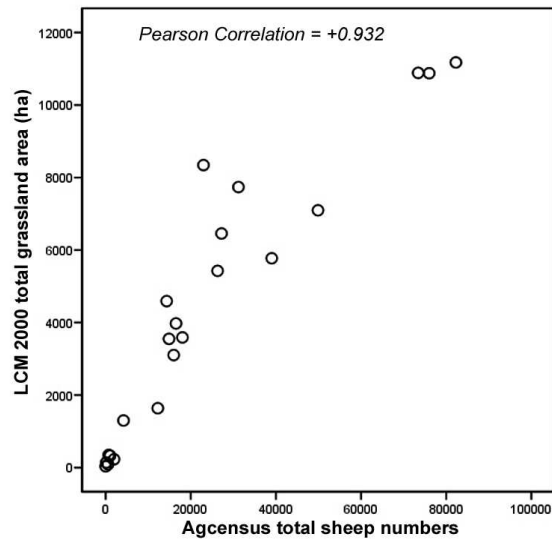


Figure 11 Spatial correlation between total combined LCM 2000 grassland classes and Agcensus total sheep numbers.



3.1.3 Summary of grassland and livestock distribution

- i. The spatial extent of Agcensus grasslands corresponds well with LCM 2000 total grassland areas.
- ii. There is a close spatial association between LCM 2000 Improved and Calcareous Grasslands and, in turn, Neutral Grassland tends to occur close to Calcareous Grassland.

- iii. Agcensus Temporary and Permanent Grasslands occupy almost identical areas, Permanent Grassland being marginally more extensive.
- iv. Agcensus Rough Grazing corresponds most closely (though not exclusively) with LCM 2000 Neutral and Acid Grasslands, and Heathland categories, predominantly on higher ground.
- v. Land in the LCM 2000 'Rough Grazing' category is predominantly used for sheep, although sheep are not confined to Rough Grazing areas.
- vi. The Agcensus 'Total Cattle' category has good spatial correspondence with LCM 2000 Class 5.1 (Improved grassland).
- vii. Agcensus 'Beef Cattle' have good spatial correspondence with LCM 2000 Class 5.1 (Improved Grassland), but they area also found on LCM 2000 Classes 6.1 & 8.1 (Acid and Neutral Grasslands).
- viii. Agcensus 'Dairy Cattle' are found preferentially on LCM 2000 Classes 5.1 & 7.1 (Improved and Calcareous Grasslands).

3.1.4 Other livestock considerations

The spatial distributions of pigs and fowl are less clear with respect to the LCM 2000 classifications, but these important livestock categories are not uniformly distributed in the EDINA 2 km interpolations so, in the absence of further information, it was necessary to use the existing EDINA distribution.

Minor livestock categories in the Agcensus data were included with similar livestock groups in the major categories, as shown in Table 4.

Table 4 Grouping of Agcensus livestock classes for input to CASCADE.

Major livestock class	Subgroups (if any)	Minority classes accommodated in major livestock class	Further subdivision required
Total Cattle	Dairy Herd Beef Herd Other Cattle		
Sheep		Goats	
Pigs			Indoor Pigs Outdoor Pigs
Total fowl	Poultry	Geese Ducks	Indoor Poultry Free Range Poultry

3.1.5 Correspondence of Agcensus and LCM data with CASCADE grassland classes
 CASCADE Classes 1 & 2 (Recent & Temporary Grassland, and Permanent Grassland, respectively) correspond with the respective Agcensus Classes G1 & G2 (Temporary and Permanent Grasslands), and most closely with LCM 2000 Classes 5.1, 6.1 & 7.1 (Improved, Neutral and Calcareous Grasslands).

CASCADE Class 3 (Rough Grazing) corresponds with Agcensus Class G5 (Rough Grazing), and most closely with LCM 2000 Classes 8.1, 9.1, 10.1, 10.2, 11.2 & 12.1 (Acid Grassland, Bracken, Dwarf Shrub Heath, Open Shrub Heath, Fen & Marsh and Bog).

3.2 Redefining the boundaries of urban areas

The next stage in the land use update process involved combining information from LCM 2000 with Ordnance Survey (OS) Meridian 2 digital boundary data (<http://edina.ac.uk/digimap/description/products/meridian.shtml>) in order to redefine the spatial extent of built-up areas. In this way, land formerly identified as agricultural in LCM 2000, which had subsequently been developed (e.g. through housing, road-building or industry) could be reclassified as 'urban', or 'other land' and thus be excluded from the area of land available for agricultural use.

For the Yorkshire Derwent catchment, a detailed visual comparison was made between OS 1:25,000 paper maps and (a) the digital OS Meridian DLUA (Developed Land Use Area) boundaries, representing the spatial extent of built-up areas (namely cities, towns and villages), and (b) the LCM 2000 Urban and Suburban classes.

In general, the OS Meridian data were very accurate in more rural areas, but the paper maps indicated that some peripheral built-up areas of larger conurbations were not included within the DLUA polygons, but *were* identified as urban by LCM 2000. These discrepancies are most likely to be due to the different data collection times for each of the map types.

The LCM 2000 urban/suburban identification was not as successful as the OS Meridian data, particularly in rural areas, where it tended to underestimate or entirely miss some of the smaller settlements (e.g. Figure 12(a)). This can be partly accounted for by the much larger ratio of open space to buildings in small villages and hamlets. However, some large, isolated buildings such as farms or estates were identified as urban/suburban in LCM 2000, as well as features such as motorway interchanges, airfields, quarries, canals and certain other water features (Figure 12(b)).

Figure 12 (a) Example of overlay of LCM 2000 'suburban' and 'continuous urban' classes with OS Meridian 2 DLUA polygons. The boundaries of small villages (defined by DLUA polygons) include areas identified by LCM 2000 as agricultural, woodland, etc.

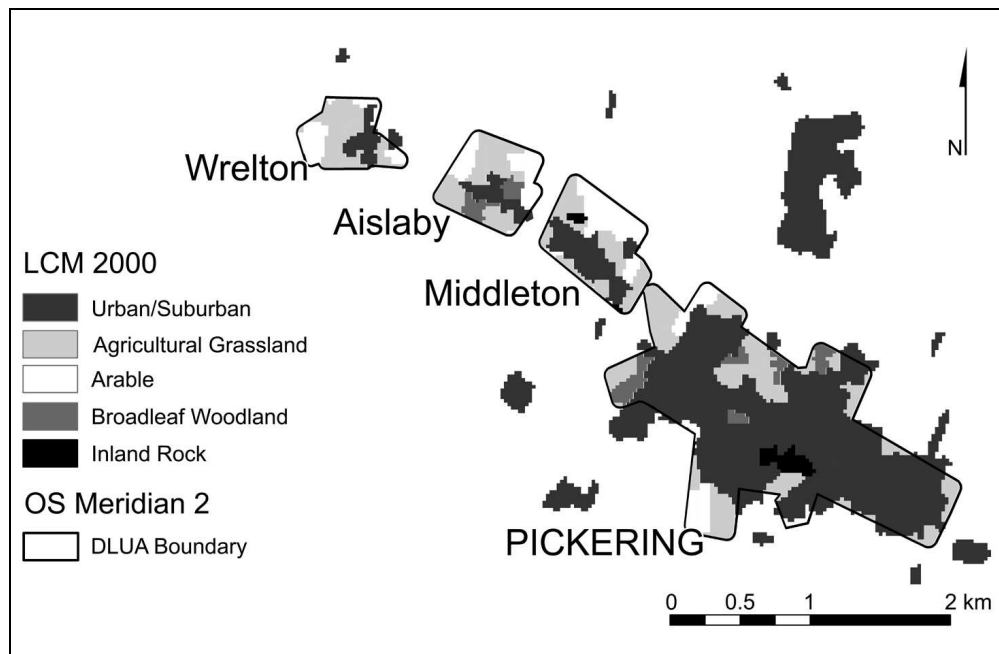
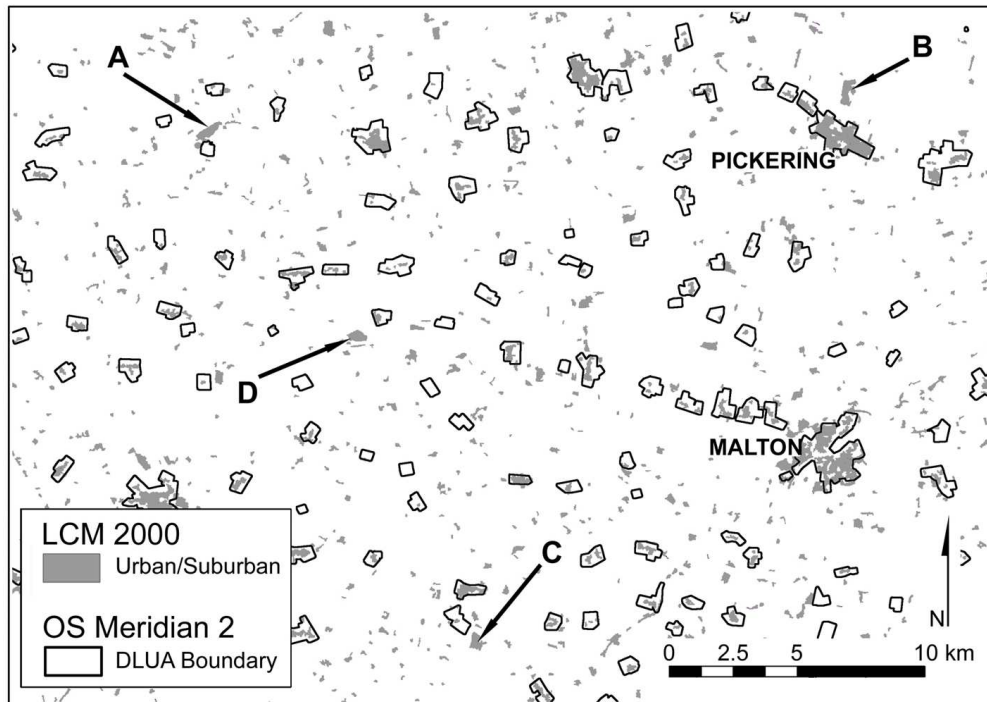


Figure 12 (b) Examples of features in the Pickering and Malton area misrepresented as 'suburban' or 'continuous urban' by LCM 2000. A: Woodland, water; B: Quarry; C: Farm; D: Parkland, castle, quarry.



These findings suggested that LCM 2000 could be useful for redefining the boundaries of urban areas, where they extended beyond the OS DLUA polygons, but that the DLUA polygons were the most reliable indicators of urban/suburban extent. Furthermore, the non-urban areas identified as urban by LCM 2000 generally coincided with non-agricultural sites, so these areas could be included in the CASCADE 'All other land' class.

3.2.1 Method for redefining spatial extent of built-up areas

The following steps, performed using ESRI ArcGIS 8.3 (<http://www.esri.com/>) were used to redefine the spatial extent of built-up areas, in preparation for the land use update for input to the CASCADE model.

- a) LCM 2000 'Continuous Urban' and 'Suburban' classes were extracted from the full data set.
- b) A 'focal operation' was performed to identify the densest clusters of urban and suburban cells.
- c) A subset of LCM 2000 'Continuous Urban' and 'Suburban' clusters was extracted from those identified in (b), representing those clusters that intersected OS Meridian 2 DLUA polygons.
- d) The DLUA polygons were converted to 25 m raster cells (to match the LCM 2000 format) and classified as 'urban'.
- e) The LCM 2000 'urban and suburban' subset extracted in (c) was added to the DLUA raster created in (d).
- f) Using the ArcGIS 'mosaic' function, the result of (e) was reinstated into the full LCM 2000 data set as a revised 'Urban' category, thus redefining the spatial extent of developed areas.
- g) All remaining LCM 2000 'Continuous Urban' and 'Suburban' cells were reclassified as 'All other land'.

This redefinition of 'Urban' and 'All other land' categories facilitated more accurate scaling of the agricultural census data to HRUs.

3.3 Summary of LCM 2000 reclassification to CASCADE broad land use classes

The LCM 2000 categories were matched and reclassified according to CASCADE broad land use classes, as summarised in Table 5.

Table 5 Grouping of LCM 2000 categories in relation to CASCADE broad classes.

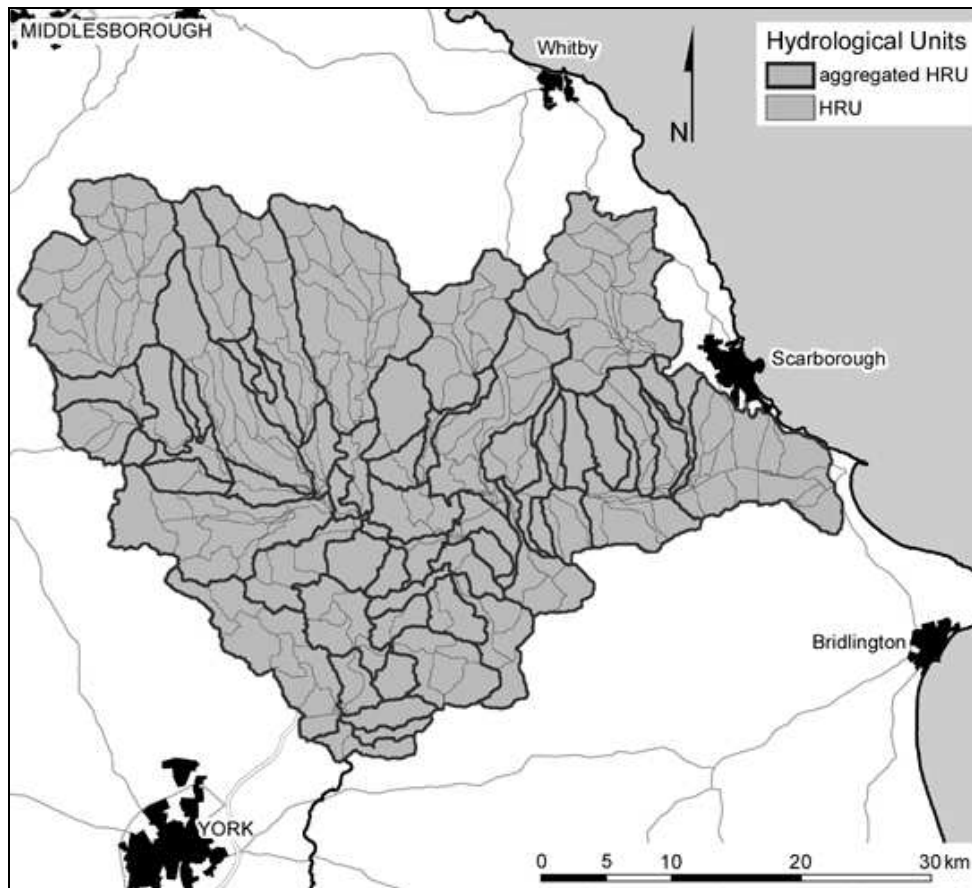
LCM 2000 Class	CASCADE Class
Sea/Estuary Inland Water	Water
Littoral Rock Littoral Sediment Saltmarsh Supra-Littoral Rock Supra-Littoral Sediment Inland Bare Ground	All Other Land
Bogs Dwarf Shrub Heath Montane Habitats Acid Grass Bracken Fen, Marsh, Swamp	Rough Grazing
Broad-Leaved/Mixed Woodland Coniferous Woodland	Woods
Arable and Horticultural Ley Grass Setaside Grass	Arable and Setaside
Improved Grassland Neutral Grass Calcareous Grass	Temporary and Permanent Grassland
Continuous Urban Suburban	Urban

4. AREAL INTERPOLATION OF AGRICULTURAL CENSUS DATA TO HYDROLOGICAL UNITS

Having updated and reclassified LCM 2000 into the broad land use classes used in the CASCADE model, it was necessary to calculate the area occupied by each of these broad classes, within each HRU, before distributing data from the 2004 Agricultural Census within the newly-defined 'Arable and Setaside' and 'Temporary and Permanent Grassland' areas.

The high resolution (25 m x 25 m grid) of LCM 2000 data facilitates the designation of land use areas within HRU boundaries, but the 2 km grid (i.e. 4 km²) resolution of the Agcensus data is coarse with respect to HRUs, which are, typically, 5-8 km² in area. Therefore, as an initial step towards improving the spatial fit of Agcensus data to HRUs, the latter were aggregated into 61 larger spatial units (Figure 13). These aggregated HRUs drain to water quality monitoring points on the river network, and are the spatial units used in the QUESTOR model, which takes outputs from CASCADE and calculates the effects of in-river processes on agricultural pollutants.

Figure 13 Spatial distribution of the 282 hydrological response units (HRUs) (fine boundaries) and 61 aggregated HRUs (bold boundaries) in the Yorkshire Derwent catchment.



The areas, per aggregated HRU, for each of the broad CASCADE classes, are given in Appendix A.1.

Relevant crop types were then extracted from the 2004 Agcensus (2 km resolution) data set, and grouped according to categories used as CASCADE input, depending on their respective nutrient loading profiles (e.g. nitrate available for leaching). These crop groupings are given in Table 6.

Table 6 Agcensus 2004 crop types grouped according to CASCADE crop classes.

Agcensus 2004 Crop Type		CASCADE Crop Category	
G1	Grass sown 2000 or later	1	Temporary grassland
G2	Other grassland (not rough grazing)	2	Permanent grassland
A1	Wheat	7	Winter wheat
A2	Winter barley	8	Winter barley
A3	Spring barley	9	Spring barley
A4	Oats	10	Other cereals
A5	Mixed corn		
A6	Rye		
A7	Triticale		
A10	Potatoes (early)	11	Potatoes
A11	Potatoes (main crop)		
A12	Sugar beet (not for stock feed)	12	Sugar beet
A21	Field beans	13	Peas and beans
A22	Peas for dry harvest		
A24	Oilseed rape (winter)	14	Oilseed rape
A25	Oilseed rape (spring)		
A27	Linseed	15	Linseed
A15 – A18	Crops for stockfeeding	16	Other crops and fallow
A31	Other arable crops		
A32	Bare fallow		
A23	Maize		
A26	Flax		
B99	Total vegetables and salads grown in open	17	Vegetables
F98	Glasshouse/under plastic	18	Greenhouses
C99	Orchards, small fruit, grapes	19	Fruit

The next step was to interpolate these crop categories to the available grassland, arable and setaside areas within each aggregated HRU.

4.1 Interpolation of cropping data

Three different methods were used to interpolate the Agcensus data to the aggregated HRUs, and the resulting areal values were compared and appraised in order to assess what level of detail was required to give the most representative land use profiles for input to the CASCADE model. The three interpolation methods used were:

1. 'Point in polygon' method to match 2 km Agcensus grid squares to aggregated HRUs. This method interpolates data from all 2 km grid centroids falling within respective aggregated HRU boundaries (Figure 14(a)).
2. Division of 2 km Agcensus grid values into four equal 1 km grid squares, followed by a 'point in polygon' match of the 1 km squares to aggregated HRUs. This method includes data from all 'within-boundary' 1 km grid centroids (Figure 14(b)).
3. Areal interpolation of the 1 km squares to obtain a proportional land use profile for each aggregated HRU. This method includes proportional data from all 1 km grid squares intersecting individual aggregated HRUs, regardless of whether the respective 1 km centroids fall within the hydrological boundaries (Figure 14(c)).

Figure 14 Three methods used to interpolate Agcensus data to aggregated HRUs: (a) 2 km 'point in polygon' method; (b) 1 km 'point in polygon' method; (c) proportional interpolation of 1 km grid squares to aggregated HRUs. The points represent respective centroids of the Agcensus 2 km and 1 km grid squares.

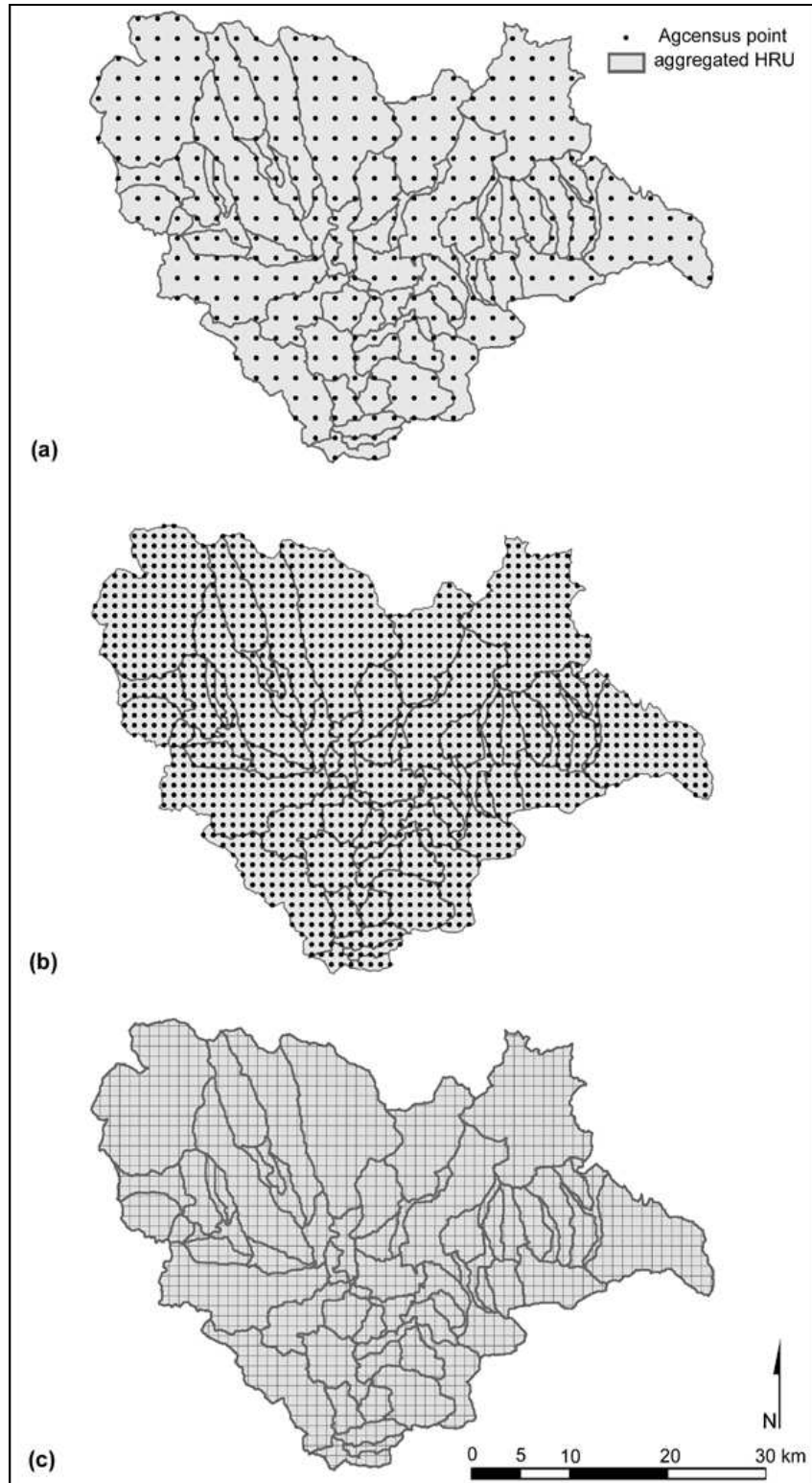
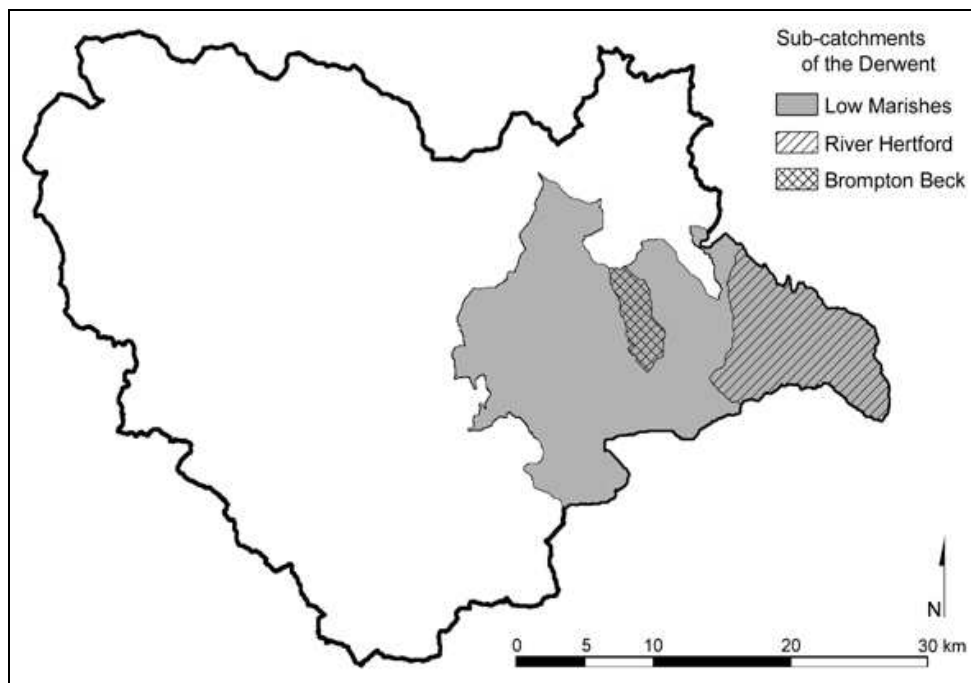


Table 7 Aggregated HRU scaling factors for the CASCADE grassland and arable classes.

aggHRU	grassland scaling factor	arable scaling factor	aggHRU	grassland scaling factor	arable scaling factor
1	1.23	0.93	31	0.88	0.76
2	0.75	0.81	32	0.60	0.99
3	0.94	1.25	33	0.94	1.70
4	0.75	0.90	34	0.94	0.93
5	1.15	0.75	35	0.48	1.48
6	1.15	0.88	36	0.89	1.00
7	2.04	0.79	37	0.51	1.52
8	1.01	1.06	38	0.83	1.35
9	1.35	0.93	39	0.56	1.10
10	2.67	0.97	40	0.48	1.08
11	1.21	1.11	41	1.04	1.20
12	1.22	0.69	42	0.44	1.14
13	1.74	0.96	43	0.57	1.29
14	0.85	1.26	44	0.87	1.21
15	1.03	0.96	45	0.98	1.12
16	0.48	1.00	46	2.17	0.97
17	0.74	1.06	47	0.99	1.02
18	1.21	0.88	48	0.73	1.29
19	0.90	0.98	49	0.73	1.34
20	0.88	0.97	50	1.22	0.65
21	1.28	0.92	51	0.92	0.91
22	1.11	0.81	52	0.86	0.63
23	0.69	0.91	53	0.74	0.62
24	0.86	0.85	54	1.01	1.46
25	1.44	0.71	55	1.09	1.44
26	0.90	1.10	56	0.37	1.44
27	0.54	1.08	57	0.75	1.68
28	0.66	1.27	58	1.44	0.74
29	1.04	1.24	59	0.90	2.53
30	1.03	1.26	60	1.18	0.68
			61	0.99	189.27

The resulting land use datasets (representing absolute areal values for the various Agcensus crop and grassland categories) were evaluated against the available LCM-defined grassland, arable and setaside areas within each aggregated HRU. 'Scaling factors' (Table 7), each representing amounts by which the interpolated data had to be scaled to fit available LCM 2000 areas, were used as a comparative measure between the methods. (The extent of scaling required was likely to have been influenced by many different factors, including changes in land use over time, misclassification of land cover types, and methods used in the compilation of agricultural census data). These results were also compared across a range of spatial scales, defined by sub-catchment area (shown in Figure 15), in an attempt to identify the range of spatial resolutions at which reliable estimations of agricultural land use could be made.

Figure 15 The Yorkshire Derwent catchment and three of its sub-catchments at different spatial scales. The dark outer boundary represents the entire Derwent catchment.



4.1.1 Comments on scaling factors for aggregated HRUs

At the scale of the aggregated HRU, it was decided that all very high or very low scaling factors should be investigated further, particularly in cases where (in the case of high scaling factors) the denominator was >10 ha (i.e. >10 ha of Agcensus arable land/grassland was involved in the interpolation) or, conversely, high numerator values were involved in the case of very low scaling factors. As shown in Table 7, for this particular river catchment, only one scaling factor fell into one of these categories, namely, the arable scaling factor for aggregated HRU 61. In this instance, there was a very large LCM-defined area available for arable cropping (96 ha) compared with the value interpolated from the Agcensus data set (0.51 ha).

Very low scaling factors are often associated with arable/setaside or temporary/permanent grassland in excess of aggregated HRU total land area. This may, in part, be due to the way in which farms make their returns for the June Agricultural Survey, some limitations of which are highlighted by EDINA (<http://edina.ac.uk/agcensus/agcen2.pdf>). It is known, for

example, that landowners with large farms will have their holdings recorded in one Agcensus centroid, even though the farm may well extend beyond the boundary of the corresponding 2 km grid square. In the absence of further information, no action can be taken on this matter and, at the scale of the entire Derwent catchment, the issue tends to even itself out.

4.1.2 *Other discrepancies*

Some other issues that have arisen during interpolating Agcensus data to hydrological units in other sub-catchments, though not in the Yorkshire Derwent, are worthy of mention here.

In some cases, LCM 2000 indicates that there is arable/setaside land or permanent/temporary grassland present where, according to the Agcensus, there is none. Less commonly, the reverse is true. These discrepancies could be due to changes in the amount of setaside land required at different times. In either instance, the following rules are applied:

- a) If LCM 2000 indicates the presence of arable/setaside land and Agcensus shows no arable/setaside
 - the LCM 2000 arable/setaside land value is transferred to the CASCADE 'All Other Land' class
- b) If LCM 2000 indicates the presence of temporary/permanent grassland and Agcensus shows no temporary/permanent grassland
 - the LCM temporary/permanent grassland value is transferred to the CASCADE 'Rough Grazing' class
- c) If LCM 2000 indicates **no** arable/setaside or temporary/permanent grassland and Agcensus indicates that **there is**
 - the CASCADE arable/setaside or temporary/permanent grassland value is taken from (in order of preference)
 1. 'All Other Land'
 2. 'Urban'.

4.2 Interpolation of livestock data

Livestock must also be accounted for in the CASCADE model, and their numbers were interpolated in the same manner as the crop data. However, as livestock numbers are absolute, rather than being related to physical land area, these data were not scaled to any land cover area.

4.2.1 *Further subdivision of certain livestock categories*

Nutrient outputs will differ (i) between indoor and outdoor pigs, and (ii) between indoor and free range poultry. Although the percentages for these categories (at the SOA level) are documented in the Agricultural Census, they do not appear in the Agcensus dataset, so broad generalisations had to be made, based on Defra Farming Statistics averages and limited verbal information from the National Pig Association. Thus, for the purposes of this study, the ratios for the two aforementioned livestock classes were assumed to be (i) 70:30 for indoor vs. outdoor pigs, and (ii) 90:10 for indoor vs. free range poultry.

5. COMPARISON OF THE THREE INTERPOLATION METHODS ACROSS A RANGE OF SPATIAL SCALES

The scaling factors used for comparison of interpolated grassland values for the Yorkshire Derwent and three of its sub-catchments, for each method, are given in Table 8(a), and those for arable land are given in Table 8(b).

Table 8(a) Scaling factors representing amounts by which the interpolated grassland data had to be scaled to fit available LCM 2000 areas.

Sub-catchment	Sub-catchment area (ha)	LCM 2000 area (ha) available for grassland	Scaling factor required to fit Agcensus data to available LCM 2000 space		
			2 km point	1 km point	1 km grid proportional
River Derwent	159480	35501	0.92	0.76	0.95
Low Marishes	36233	6423	0.82	0.63	0.83
River Hertford	8423	1132	0.52	0.51	0.54
Brompton Beck	1584	198	0.57	0.38	0.44

Table 8(b) Scaling factors representing amounts by which the interpolated arable data had to be scaled to fit available LCM 2000 areas.

Sub-catchment	Sub-catchment area (ha)	LCM 2000 area (ha) available for arable land	Scaling factor required to fit Agcensus data to available LCM 2000 space		
			2 km point	1 km point	1 km grid proportional
River Derwent	159480	65453	1.00	0.79	1.06
Low Marishes	36233	21449	0.98	0.72	1.01
River Hertford	8423	5444	1.00	0.95	1.08
Brompton Beck	1584	1115	1.42	0.80	1.14

The most notable results from Tables 8(a) and 8(b) can be summarised as follows:

- a) The scaling factors for the grassland data are all less than 1, indicating larger Agcensus grassland values than available LCM areas, whereas the opposite is true for the majority of arable values, whose scaling factors are close to, or greater than 1.
- b) In the case of the grassland data, scaling factors for all three interpolation methods depart further from the value of 1 as the catchment size decreases. However, the arable scaling factors show no such consistent behaviour.
- c) In almost every case, the 2 km point and proportional results are very similar, whereas the 1 km point values give the poorest fit. The exceptions to this are the River Hertford, which displays quite consistent results across all three methods, and Brompton Beck, where the 2 km method requires a very large scaling factor.

5.1 Interpretation of the scaling results

The scaling factors (Tables 8(a) and 8(b)) indicate that there is, in general, a better fit obtained for arable land than for grassland values. This could, in part, be due to the interpretation of temporary and permanent grassland classes in LCM 2000, in which problems exist in the distinction between some of the unmanaged (e.g. setaside), semi-natural and improved grasslands. This is also reflected (as one would expect) by the poorer fit of grassland values with decreasing catchment size. (The problem may be illustrated in Figure 12(a), where some of the 'agricultural grassland' areas identified within the DLUA boundaries may, in fact, represent domestic gardens, especially in some of the smaller

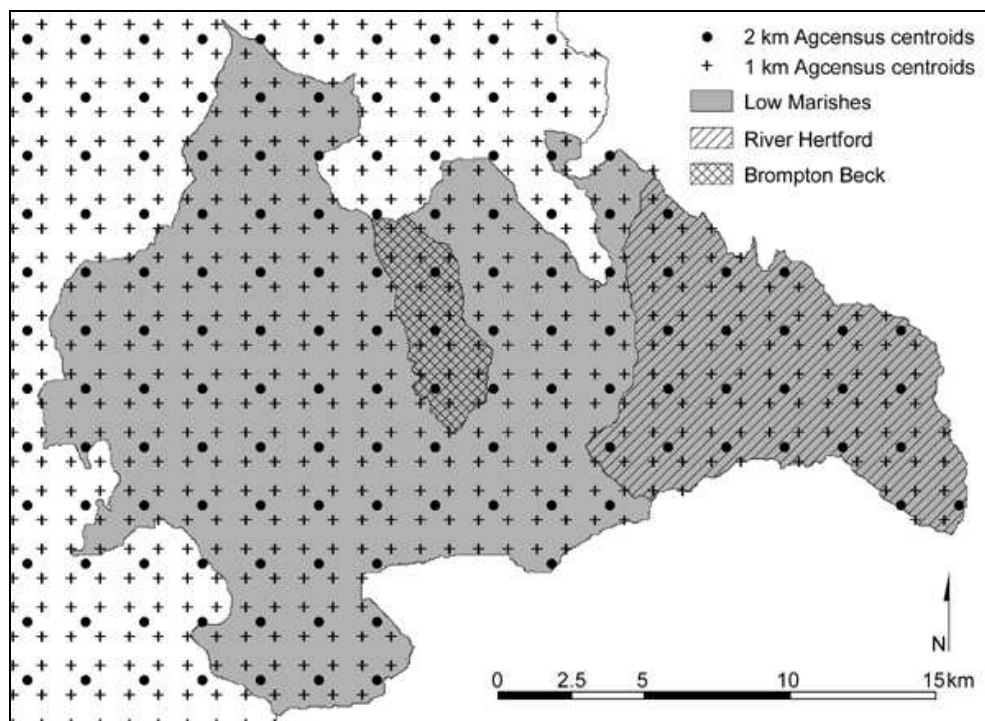
settlements.) The relationship between catchment size and amount of scaling required is not, however, seen in the arable results, and may be influenced by another element of uncertainty, arising from the way in which agricultural census data are compiled, in that land registered to an 'in-catchment' farm may actually be located beyond the catchment boundary, a particular problem with large farm businesses. Additionally, LCM 2000 data were collected in the late 1990s, whereas the current study used Agcensus data from 2004: therefore, land use changes that may have taken place in the intervening period could have led to misrepresentation of catchment profiles.

Of particular interest in Tables 8(a) and 8(b) is the finding that, in almost every case, the 1 km grid interpolation method gives the poorest fit, for both grassland and arable areas, over the entire range of catchment scales. This is thought to be an artefact of the way in which the Agcensus grid squares intersect the catchment boundaries. For example, land use values pertaining to all 1 km grid square centroids falling inside a catchment boundary will be assigned to that catchment. However, depending on the shape and location of the catchment with respect to the 2 km Agcensus grid squares containing the source data, a number of those 1 km centroids may contain values relating to 2 km centroids outside the catchment area. If the land use values are very variable between these 'out-of-catchment' 2 km squares, inclusion of the respective 'in-catchment' 1 km squares will have a marked effect on the interpolated value. The magnitude of the effect will be dependent on how many 'out-of-catchment' grid squares are involved and the variability between their respective land use values. In this study, many of the 1 km scaling factors are consistently less than 1, suggesting that, in these cases, much of the land assigned to the 1 km centroids of squares intersecting catchment boundaries actually lies outside the respective catchments.

This effect can be illustrated by comparing the scaling factors for grassland in the Brompton Beck and River Hertford catchments (Table 8(a)). Scaling factor values for Brompton Beck show a lot of variation between the three methods; however, three 2 km centroids fall inside this catchment and their respective grid squares fill much of the available area, giving the best fit for the 2 km method (Figure 16). The 1 km method performs least well in this case because it includes non-proportional values from five 2 km grid squares that lie outside the catchment boundary. The fit of the proportional method falls in between that of the other two methods, although it should, intuitively, be the best method. However, this method does not perform as well as the 2 km method because of the large number of 'out-of-catchment' grid squares (relative to catchment size) from which it draws its land use values. In contrast, a very similar fit is obtained across all methods for grassland in the Hertford catchment, where the 1 km data performs as well as the rest. Once again, the 2 km grid squares 'fill' the catchment well, but in this case, because of the placement of 2 km centroids in relation to the catchment boundary, most of the 1 km centroids relate to 'in-catchment' 2 km squares, thereby producing very similar land use results for all three methods.

The very high scaling factor for the 2 km method for arable land in Brompton Beck (Table 8(b)) could arise from the fact that only three 2 km centroids fall within this catchment. Consequently, arable land recorded in grid squares intersecting the boundary, but with their centroids just outside the boundary, will not be included in the calculation, thereby leading to an underestimation of arable areas. This is confirmed by the better fit achieved using the proportional method in this catchment.

Figure 16 Placement of 2 km and 1 km grid square centroids in relation to the Brompton Beck and River Hertford catchment boundaries.



Further confirmation is given by the numbers of Agcensus cells contributing to the interpolated values at each catchment size, and the corresponding perimeter/area ratio of each sub-catchment, which influences the spatial extent over which relevant data are amassed (Table 9). The comparatively high perimeter/area ratio of the Brompton Beck sub-catchment, combined with the involvement of a disproportionate number of ‘out-of-catchment’ Agcensus cells in the 1 km method, leads to the greatest variability in results at this catchment size. However, the shape and placement of a sub-catchment such as the River Hertford provides a better fit for the Agcensus data and, consequently, less variability between the outcomes of the different interpolation methods.

Table 9 Perimeter/area ratios for each sub-catchment, and number of Agcensus cells contributing to land use data for each interpolation method.

Sub-catchment	Perimeter/area ratio	Number of contributing Agcensus cells			Comparative ratios between contributing Agcensus cells	
		2 km point	1 km point	1 km grid proportional	proportional/ 2 km point	1 km point/ 2 km point ^a
River Derwent	1.985	411	1599	1732	4.21	3.89
Low Marishes	4.843	92	362	439	4.77	3.93
River Hertford	7.408	22	82	111	5.05	3.73
Brompton Beck	16.225	3	17	26	8.67	5.67

^a One would expect the number of contributing cells in the 1 km interpolation to be roughly four times that of the 2 km method, but in the case of Brompton Beck, this value is elevated due to the spatial relationship between the catchment boundary and the Agcensus grid squares

6. APPLICATION TO NITRATE-N MODELLING USING CASCADE

The three land use data sets were used as input to the CASCADE model (run over the 1998-2004 period) in order to calculate (i) mean nitrate-N values, and (ii) the fraction of total time when nitrate-N concentrations would be above the drinking water limit (11.3 mg l⁻¹ N), for each sub-catchment. The outputs from these modelling runs are presented in Table 10.

The results represent nitrate-N concentrations at the outlet points of each sub-catchment, and consider contributions solely from diffuse inputs derived from the interpolated land use profiles. They do not include the effects of point sources or in-stream processes. Point sources represent less than 10% of the nitrate-N load in all the sub-catchments tested (Hutchins et al. in press). In-river processes serve to reduce the load by 15-40% (Hutchins et al. submitted). Sensitivity of the magnitude of these sources/sinks to choice of land-use data preparation method would be negligible.

Table 10 Nitrate-N results obtained for interpolated land use values, using three different interpolation methods at a range of spatial scales.

Sub-catchment	Sub-catchment area (ha)	Mean nitrate-N level (mg l ⁻¹) modelled from land use values derived from three different interpolation methods		
		2 km point	1 km point	1 km grid proportional
River Derwent	159480	6.365	6.372	6.396
Low Marishes	36233	9.897	10.010	9.943
River Hertford	8423	10.477	10.642	10.498
Brompton Beck	1584	10.990	11.482	11.135

In terms of modelling mean nitrate-N levels (Table 10), the different interpolation methods do not appear to produce marked differences in output within individual catchments, and the only incidence of violation of the drinking water nitrate-N limit (11.3 mg l⁻¹) is seen at the smallest catchment scale, in Brompton Beck. The CASCADE model scales input land use values to the available land area within each catchment and models these values according to relative proportions of crop types. Therefore, large variations between interpolated land use values derived from the three different methods are not reflected in the CASCADE output values, except at the smallest catchment scale. The small variations between CASCADE outputs relate to the number of, and variability (in terms of assemblages of crop and agricultural grassland types) between, 'out-of-catchment' Agcensus grid squares included in the land use calculation for each catchment, as described previously. The mean nitrate-N concentrations shown in Table 10 appear to increase as catchment size decreases, but this is thought to be coincidental, and a reflection of between-catchment land use differences.

A better indication of the variation in modelled nitrate outputs arising from the use of different interpolation methods might be conveyed by the fraction of total time that nitrate-N concentrations would exceed the drinking water limit. As shown in Table 11, this particular parameter of nitrate level (in contrast to mean concentration) seems to be more sensitive to choice of method, at least in the higher nitrate sub-catchments (i.e. Low Marishes, River Hertford and Brompton Beck). This indicator is also more pertinent from a legislative point of view as, specifically, the Directive is violated if the 95th percentile value exceeds 11.3 mg N l⁻¹. Values in excess of the drinking water limit tend to occur after the soils wet up in autumn, when the available leachable N from crop residues is at a maximum, and high

flows flush this N into the river. However, very high flows will not coincide with the highest N concentrations due to a diluting effect.

Table 11 Fraction of total time that modelled nitrate-N levels would be above drinking water limit (11.3 mg l⁻¹), for each interpolation method, across a range of spatial scales.

Sub-catchment	Sub-catchment area (ha)	Fraction of total time that nitrate-N levels would be above drinking water limit (11.3 mg l ⁻¹) ^b		
		2 km point	1 km point	1 km grid proportional
River Derwent	159480	0.000	0.000	0.000
Low Marishes	36233	0.119	0.155	0.127
River Hertford	8423	0.327	0.363	0.334
Brompton Beck	1584	0.405	0.649	0.497

^b Values of > 0.05 would be in violation of the Nitrates Directive

7. FURTHER REFINEMENT OF THE MODEL

Although Agcensus data have already been distributed according to the Landuse Framework used by EDINA, the resulting values are evenly distributed across 2 km grid squares. It was anticipated, therefore, that using the spatial extent of arable and grassland areas defined by LCM 2000 to 'force' those Agcensus land use classes into their appropriate locations *before* performing the interpolation to hydrological units, would offer some improvement of fit within the catchment boundaries, and that this would be reflected by a reduction in scaling factor values. This would not, however, remove the uncertainties regarding potential land use change over time.

A trial of this further refinement to the land use interpolation method resulted in negligible differences to the output crop values which, in turn, would have no discernible effect on the modelled nitrate concentrations in CASCADE. The minimal effect of this new method on the outcomes of interpolated land use values suggests that, at least in the case of the Yorkshire Derwent, agricultural land may be fairly evenly distributed across Agcensus 2 km grid squares or, at the spatial extent of the aggregated HRU, agricultural land is distributed in such a way that the additional refinement to the method has little influence on the total values.

8. CONCLUSIONS AND RECOMMENDATIONS

This work highlights some of the problems associated with assigning agricultural land use data to hydrological units. The indication is that derived land use profiles are more reliable at the scale of a large river catchment such as the Yorkshire Derwent and that reasonable estimations of land use may be made at the scale of sub-catchments such as the River Hertford (i.e. above 8000 ha). However, uncertainty increases in smaller hydrological units, with catchments at the scale of Brompton Beck being most problematic.

It is apparent that there is considerable variability between derived land use estimations for small catchments and that sophisticated interpolation procedures do not necessarily improve the outcomes. Variability relates both to the way in which source data are compiled and manipulated, and to differences in size, shape and location of the spatial units involved in the interpolation process. This analysis highlights the importance of considering these spatial relationships when assessing the reliability of land use estimations, the implications for diffuse pollution modelling at different catchment scales, and the need for caution in interpreting the results, particularly in small-scale catchments.

The study identifies a key need for better land use data and agricultural statistics for catchment management purposes, from projects such as the Defra Observatory Programme (<https://statistics.defra.gov.uk/esg/ace/index.htm>), which was launched in 2005 with a view to monitoring and anticipating farm level changes arising from policy reform and other key drivers.

References

- Aalders, I.H. and Aitkenhead, M.J. (2006) Agricultural census data and land use modeling. *Computers, Environment and Urban Systems*, 30:799-814.
- Bateman, I.J., Brouwer, R., Davies, H., Day, B.H., Deflandre, A., Di Falco, S., Georgiou, S., Hadley, D., Hutchins, M., Jones, A.P., Kay, D., Leeks, G., Lewis, M., Lovett, A.A., Neal, C., Posen, P., Rigby, D. and Turner, R.K. (2006) Analysing the agricultural costs and non-market benefits of implementing the Water Framework Directive. *Journal of Agricultural Economics*, 57:221-37.
- Bateman, I.J., Deflandre-Vlandas, A., Fezzi, C., Hadley, D., Hutchins, M., Lovett, A., Posen, P. and Rigby, D. (2007) *WFD related agricultural nitrate leaching reduction options: Cost estimates derived from farm level survey data & a cost-effectiveness assessment for the Derwent catchment*. CSERGE Working Paper ECM-2007-03, University of East Anglia, pp. 81. Available online at: www.uea.ac.uk/env/cserge/pub/wp/ecm/ecm_2007_03.htm.
- CEC (2000) Council Directive 2000/60/EC of 23 October 2000 establishing a framework for Community action in the field of water policy. Council of the European Communities, Report No 2000/60/EC.
- Chambers, B.J., Lord, E.I., Nicholson, F.A. and Smith, K.A. (1999) Predicting nitrogen availability and losses following application of organic manures to arable land: MANNER. *Soil Use and Management*, 15:137-43.
- Cuttle, S.P., Macleod, C.J.A., Chadwick, D.R., Scholefield, D., Haygarth, P.M., Newell-Price, P., Harris, D., Shepherd, M.A., Chambers, B.J. and Humphrey, R. (2007) *An inventory of measures to control diffuse water pollution from agriculture (DWPA): User manual*. Defra, UK. Available online at: www.defra.gov.uk/farm/environment/water/csf/pdf/UserManual_Jan07.pdf
- Deflandre, A., Williams, R.J., Elorza, F.J., Mira, J. and Boorman, D.B. (2006) Analysis of the QUESTOR water quality model using a Fourier amplitude sensitivity test (FAST) for two UK rivers. *Science of the Total Environment*, 360:290-304.
- Defra (2004) *Developing measures to promote catchment sensitive farming*. Joint Defra-HM Treasury consultation. Available online at: www.defra.gov.uk/farm/environment/water/csf/pdf/Defra-HM-Treasury-consultation.pdf.
- Defra (2006) *Combining Defra June Census and IACS data to construct the first high resolution national 1km² agricultural land cover map*. Defra Agricultural Change and Environment Observatory Research Report No. 04. Available online at: <https://statistics.defra.gov.uk/esg/ace/research/pdf/observatory04.pdf>.
- Fuller, R.M., Smith, G.M., Sanderson, J.M., Hill, R.A. and Thompson, A.G. (2002) The UK Land Cover Map 2000: Construction of a parcel-based vector map from satellite images. *The Cartographic Journal*, 39:15-25.

- Geddes, A., Gimona, A. and Elston, D.A. (2003) Research Article: Estimating local variations in land use statistics. *International Journal of Geographical Information Science*, 17:299-319.
- Goulding, K.W.T., Bailey, N.J. and Bradbury, N.J. (1998) A modelling study of nitrogen deposited to arable land from the atmosphere and its contribution to nitrate leaching. *Soil Use and Management*, 14:70-7.
- Huby, M., Owen, A. and Cinderby, S. (2007) Reconciling socio-economic and environmental data in a GIS context: An example from rural England. *Applied Geography*, 27:1-13.
- Hutchins, M.G., Deflandre-Vlandas, A., Posen, P.E., Davies, H.N. and Neal, C. (submitted) How do river nitrate concentrations respond to changes in land-use? A modelling case-study of headwaters in the River Derwent catchment, North Yorkshire, UK. *Environmental Modelling and Assessment*.
- Hutchins, M.G., Deflandre, A. and Boorman, D.B. (2006) Performance benchmarking linked diffuse pollution and in-stream water quality models. *Arch. Hydrobiol. Suppl. (Large Rivers 17, No. 1-2)*, 161:133-54.
- Hutchins, M.G., Dilks, C., Davies, H.N. and Deflandre, A. (2007) Issues of diffuse pollution model complexity arising from performance benchmarking. *Hydrology and Earth System Sciences*, 11:647-62.
- Hutchins, M.G., Fezzi, C., Bateman, I.J., Posen, P.E. and Deflandre-Vlandas, A. (in press) Cost-effective mitigation of diffuse pollution: Setting criteria for river basin management at multiple locations. Submitted to *Environmental Management*, 23rd July 2008.
- Lord, E.I. (1992) Modelling of nitrate leaching: Nitrate Sensitive Areas. *Aspects of Applied Biology*, 30:19-28.
- Moxey, A. and Allanson, P. (1994) Areal interpolation of spatially extensive variables: A comparison of alternative techniques. *International Journal of Geographical Information Science*, 8:479 - 87.
- Moxey, A., McClean, C. and Allanson, P. (1995) Transforming the spatial basis of agricultural census cover data. *Soil Use and Management*, 11:21-25.
- Naden, P.S., Cooper, D.M. and Boorman, D.B. (2001) Modelling large-scale river basins. In Huntley, D.A., Leeks, G.J.L. and Walling, D.E. (eds.) *Land-ocean interaction : measuring and modelling fluxes from river basins to coastal seas*. IWA Publishing: London.
- Neal, C., Leeks, G.J.L., Millward, G.E., Harris, J.R.W., Huthnance, J.M. and Rees, J.G. (2003) Land Ocean Interaction: processes, functioning and environmental management: a UK perspective, *Science of the Total Environment* 314-316:801-19.
- Neal, C., Neal, M., Hill, L. and Wickham, H. (2006a) River water quality of the River Cherwell: An agricultural clay-dominated catchment in the upper Thames Basin, southeastern England. *Science of the Total Environment*, 360:272-89.
- Neal, C., Neal, M., Hill, L. and Wickham, H. (2006b) The water quality of the River Thames in the Thames Basin of south/south-eastern England. *Science of the Total Environment*, 360:254-71.

Scholefield, D., Lockyer, D.R., Whitehead, D.C. and Tyson, K.C. (1991) A model to predict transformations and losses of nitrogen in UK pastures grazed by beef cattle. *Plant and Soil*, 132:165-77.

Appendix A.1

Areal values (ha), per aggregated HRU (aggHRU) in the Yorkshire Derwent catchment, for each of the broad land use classes used in the CASCADE model.

aggHRU	urban/suburban	all other land	temp/perm grassland	rough grazing	woods	arable/setaside	water	TOTAL
1	39.75	35.06	318.13	0.00	123.81	612.75	0.00	1129.50
2	61.38	36.81	147.56	0.00	27.88	290.63	0.00	564.25
3	348.69	232.81	1244.06	1.00	474.94	4043.25	0.00	6344.75
4	76.38	58.56	290.25	0.00	53.75	644.56	0.00	1123.50
5	124.81	37.75	277.75	0.00	249.06	497.38	0.00	1186.75
6	128.31	158.63	900.94	4.13	355.75	1757.75	0.00	3305.50
7	127.44	91.88	894.38	1.69	312.44	1339.19	0.00	2767.00
8	18.38	45.81	226.25	0.00	53.81	585.75	0.00	930.00
9	132.25	155.63	510.19	0.00	607.00	1494.94	0.00	2900.00
10	130.56	44.00	308.06	0.00	52.56	459.31	0.00	994.50
11	122.13	62.56	322.81	4.56	50.13	894.56	0.00	1456.75
12	235.00	33.88	123.81	0.00	44.13	326.69	0.00	763.50
13	80.56	79.44	620.06	0.00	258.94	1654.50	0.00	2693.50
14	140.13	84.63	402.25	0.00	455.19	1883.81	0.00	2966.00
15	52.69	43.69	163.38	1.19	63.81	840.75	0.00	1165.50
16	115.69	28.88	60.75	0.00	20.00	525.69	0.00	751.00
17	182.25	37.44	191.31	0.00	24.06	1144.94	0.00	1580.00
18	98.63	31.31	281.31	0.00	58.94	1009.56	0.00	1479.75
19	116.50	75.13	194.94	0.00	43.00	680.94	0.00	1110.50
20	442.31	308.44	1782.56	37.31	1507.00	4085.63	0.00	8163.25
21	88.38	8.44	124.13	0.00	34.81	459.75	0.00	715.50
22	51.75	32.75	282.13	7.81	80.69	1104.13	0.00	1559.25
23	50.81	10.31	64.38	0.00	10.31	382.94	0.00	518.75
24	37.63	32.19	185.00	34.06	187.88	622.00	0.00	1098.75
25	178.94	107.13	699.94	23.50	171.13	2040.38	0.00	3221.00
26	98.06	57.31	523.13	0.00	49.69	1664.06	0.00	2392.25
27	1101.25	405.06	1131.69	163.63	177.25	5444.38	0.00	8423.25
28	67.75	1.06	109.81	0.00	5.69	382.69	0.00	567.00
29	208.13	92.38	1489.06	397.38	1919.38	2375.94	0.00	6482.25

aggHRU	urban/suburban	all other land	temp/perm grassland	rough grazing	woods	arable/setaside	water	TOTAL
30	130.38	20.69	387.00	0.00	30.63	775.56	0.00	1344.25
31	65.94	40.13	212.19	0.50	51.06	788.19	0.00	1158.00
32	79.38	76.25	168.88	21.63	97.63	890.00	0.00	1333.75
33	40.06	27.88	323.31	0.69	15.75	916.06	0.00	1323.75
34	72.81	21.44	374.38	57.19	668.88	858.31	0.00	2053.00
35	57.31	43.81	134.69	4.06	314.06	698.06	0.00	1252.00
36	284.56	164.38	1252.06	1550.50	825.06	2771.94	0.00	6848.50
37	119.69	26.44	124.94	0.81	102.38	773.25	0.00	1147.50
38	92.38	39.31	404.31	6.38	22.75	1164.88	0.00	1730.00
39	121.63	35.25	193.19	0.31	30.38	913.25	0.00	1294.00
40	31.06	7.75	59.31	3.00	10.88	405.75	0.00	517.75
41	54.50	37.56	352.88	0.00	21.63	590.19	0.00	1056.75
42	163.50	38.44	197.75	13.81	56.13	1114.63	0.00	1584.25
43	53.25	27.31	146.44	1.94	71.31	695.50	0.00	995.75
44	98.56	33.81	247.94	0.00	105.75	471.94	0.00	958.00
45	8.56	33.31	242.19	0.75	240.81	457.63	0.00	983.25
46	0.00	47.94	174.75	6.31	160.88	117.38	0.00	507.25
47	85.38	119.81	717.19	24.63	422.19	798.31	0.00	2167.50
48	180.06	86.56	658.19	5.75	106.75	1575.94	0.00	2613.25
49	67.75	45.19	349.88	319.63	679.31	618.50	0.00	2080.25
50	4.69	38.00	490.50	36.50	267.50	581.56	0.00	1418.75
51	313.38	240.38	2155.00	2196.69	606.94	865.38	0.00	6377.75
52	28.06	17.56	173.69	9.06	109.25	240.88	0.00	578.50
53	63.00	12.38	109.56	20.75	131.00	164.56	0.00	501.25
54	324.31	370.69	2758.19	4655.00	2223.44	1402.38	0.00	11734.00
55	27.19	68.81	515.06	17.19	443.75	705.25	0.00	1777.25
56	28.25	19.00	70.19	5.38	259.81	428.13	0.00	810.75
57	125.38	105.88	1383.81	1312.75	1565.19	1620.00	0.00	6113.00
58	0.00	10.38	182.56	0.19	163.94	207.44	0.00	564.50
59	193.25	36.56	2163.13	2564.44	4680.44	2023.19	0.00	11661.00
60	154.19	398.56	3991.50	6296.38	1734.44	473.94	0.00	13049.00
61	16.13	103.56	916.81	2069.56	390.25	95.94	0.00	3592.25
Whole catchment	7741.06	4823.94	35501.44	21878.00	24083.13	65452.69	0.00	159480.25

