The Spatial Disaggregation of GB and European Agricultural Land Use Statistics

Colin J. McClean

Environment Department, University of York, Heslington, York, YO10 5DD, UK.
Tel: +44 (0)1904 432995
Fax: +44 (0)1904 432998
E-mail: cjm8@york.ac.uk

1. Introduction

Spatially and temporally detailed land use data are an increasingly sought, yet generally lacking, environmental data source. Among the many drivers of demand for such data is the requirement to understand how climate change is influencing both global and local patterns of biodiversity. It is understood that human interactions with the distribution of biodiversity, in the form of land use change, need to be taken into account when both modelling the current distribution of biodiversity and future impacts on it (Parmesan & Yohe, 2003; Pearson & Dawson, 2003; Thuiller *et al.*, 2004). Recent snapshots of land cover derived from remotely-sensed data are proving highly valuable for teasing out the importance of climate versus land use change (Termansen *et al.*, 2006). However, longer run time series are required to fully understand changes in biodiversity's spatial distribution over recent climate history. Such data are also required in order to determine how farmers adapt spatially to policy signals. In the past, these signals have included CAP reform and other external factors such as world commodity prices, but an understanding of how farmers adapt may prove invaluable in assessing how land use may change as a response to climate change.

Very good records of major agricultural land use in GB have been collected since the 1860s, including areas of crops grown and livestock reared. For many years this was an agricultural census, in that all farmers returned records. Recently, the June Agricultural Census has become a very large survey. For confidentiality reasons, individual records for farm businesses cannot be released. Results of the survey have been released using a number of different geographies over the years. At the highest spatial resolution data have been released over time at parish level, groups of parishes level, ward level and now super output area level. All of these geographies have been developed with human activities in mind. As a result, these data cannot be integrated easily with data on biodiversity. Furthermore, changes in the units over time do not allow straightforward comparison of agricultural activity across years. Other similar data sets exist for other areas of the world, and suffer from similar limitations. For example, Eurostat release statistics for agricultural land use at a number of different nomenclature of territorial units for statistics (NUTS) levels, in the case of the UK regions reported, these data will be from the agricultural census.

Therefore, there is a need to spatially disaggregate data if they are to be compared through time using some common geography compatible with biodiversity-type data (Howitt and Reynaud, 2003). This need to change geographies has been appreciated by geographers for some time, particularly when dealing with data reported for differing administrative regions (Tobler, 1979; Openshaw, 1984). The work reported here attempts to disaggregate two different scales of such data, both are easily accessible via the internet and other published sources, the GB agricultural census data from reports at the county-level and Eurostat regional agricultural statistics data. The former are disaggregated via a 1 km grid resolution to a 10 x 10 km grid resolution, while the latter are disaggregated via a 5 km grid resolution to a 25 x 25 km grid resolution. Previous work has disaggregated such data (Moxey *et al.* 1995; Howitt and Reynaud, 2003; Huby *et al.*, 2006; You and Wood, 2006), but the approach here demonstrates that a relatively simple and quick method can give results of sufficient quality for many purposes.

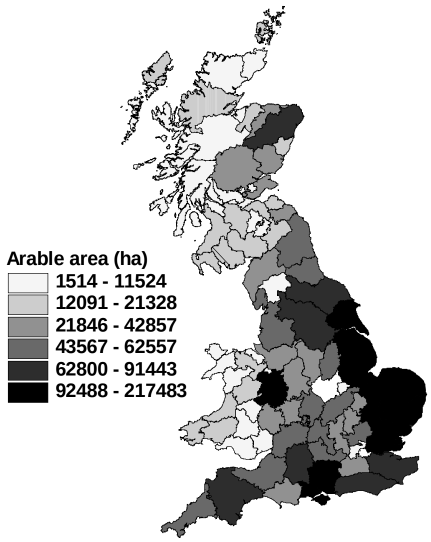
2. Methods

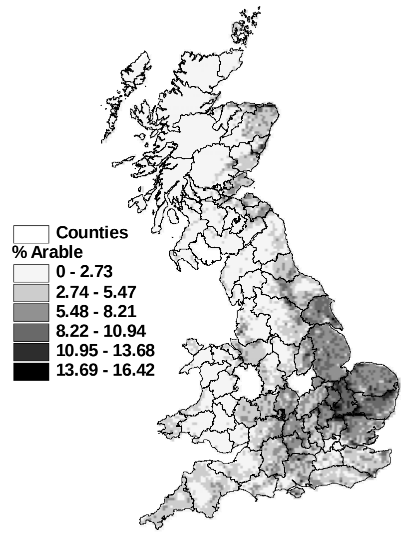
The method used develops from work by Moxey *et al* (1995) and simplifies a more recent approach developed in Huby *et al.* (2006). In the former, a 1 km grid resolution land capability map was used as a key to redistribute parish-level census data using an econometric approach. In the latter, as in the GB part of this work, a 1 km land cover map is used in the place of a specific land capability measure. However, they serve the same purpose in that both are used to guide the allocation of the recorded areas of agricultural land use at the coarser geography to grid cells at a higher resolution. For the GB disaggregation the Centre for Ecology and Hydrology's Land Cover Map of Great Britain 1990 is used. 1 km^2 grid cell records of the percentage cover of the nine most relevant cover types are used: coniferous woodland; deciduous woodland; tilled land; marsh/rough grass; grass shrub heath; dwarf shrub heath; managed grass; heath/moor grass; urban and suburban. For the European data, the equivalent Corine land cover data are used.

The algorithm used to disaggregate the census data has two stages. For each individual land use recorded in the census/Eurostat data, for each year, a simple linear regression model of suitability for that land use is constructed using proportions of land cover types found in the set of administrative boundaries for which the land use data were originally reported. These models are then applied to a finer, 1 km^2 (GB) or 5 km^2 (Europe) representation of the land cover map to produce surfaces of the likely suitability for particular land uses in each grid cell. Given that these are national/regional models based on county/NUTS-level data, it is not necessary that the estimates for each grid cell are particularly accurate, these surfaces are only used as a guide to the second stage of the dissagreation process. The second stage is a simulation process whereby randomly selected units of land use from the census/Eurostat data for a county/NUTS are allocated to 1 km/5 km grid cells on the basis of the suitability scores estimated in the models from the first stage. This simulation stage is run many times to estimate a mean pattern of land use with 10 km/25 km grid cells (summing results for the 1 km/ 5km cell simulations).

3. Results

An example disaggreagation of the GB county-level arable data are illustrated in figure 1. Initial results are visually promising, showing changes in agricultural land use practices across GB that appear sensible for the periods considered. Tests of the stability of estimates and the spatial pattern are made using different simulation





a) b)

Figure 1. a) Areas of arable land reported for counties in 1950 b) disaggregation of data to 5 km grid cells (percentage of 5 km grid cell that is predicted to be arable).

parameters, including the number of iterations of the simulation stage. Tests of the accuracy of the spatial disaggregations are also performed. GB diaggregations are compared to agricultural census data reported at ward level, while European disaggregations are compared to GB county-level data. The approach does not explicitly incorporate spatial dependencies within the data, however spatial patterns in residuals will be interpreted to ascertain whether improvements could be made in the first stage of the method. These will be reported in the full-text of this paper, along with observations about how well disaggregation works at these two varying scales. Potential extensions of the approach to other applications will be considered.

References

Howitt, R and Reynaud, A. (2003) Spatial disaagregation of agriculural production data usng maximum entropy, *European Review of Agricultural Economics*, 30, 359-387.

Huby M., Cinderby S., Crowe A.M., Gillings S., McClean C.J., Moran D., Owen A., White P.C.L. (2006) The association of natural, social and economic factors with bird species richness in rural England, *Journal of Agricultural Economics*, 57, 295-312.

Moxey, A., McClean, C. and Allanson, P. (1995) Distributing agricultural census cover data by areal interpolation, *Soil Use and Management*, 11, 21-25.

Openshaw, S. (1984) The modifiable areal unit problem, *Concepts and Techniques in Modern Geography* 38: 41.

Parmesan, C. and Yohe, G. (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature,* 421, 37-42.

Pearson, R. G. and Dawson, T. P. (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology & Biogeography*, 12, 361–371.

Termansen, M., McClean, C. J. and Preston, C. D. (2006) The use of genetic algorithms and Bayesian classification to model species distributions. *Ecological Modelling,* 192, 410–424.

Thuiller, W., Araújo, M. and Lavorel, S. (2004) Do we need land-cover data to model species distributions in Europe? *Journal of Biogeography,* 31, 353–361.

Tobler, W.R. (1979) Smooth pycnophylactic interpolation for geographical regions. *Journal of the American Statistical Association*, 74, 519-530.

You, L. and Wood, S. (2006) An entropy approach to spatial disaggregation of agricultural production, *Agricultural Systems,* 90, 329-347.

Biography

Colin McClean is a senior lecturer. He has 20 years experience in applying spatial analysis to research fields including ecology, environmental economics, geomorphology and hydrology. As well as the work presented here, recent research has assessed the potential impacts of climate and land-use changes on the British and African flora.