Methodology for the development of the ADAS land use database 2010.
Hester Lyons
ADAS.

Introduction
The development of the ADAS 1km^2 agricultural land use database has been described in detail in Lord, (1999), ADAS (2001), Lyons *et* al. (2004) and Comber *et al*. (2008). What follows below is a brief description of the methodology, but the reader is directed to the above references for a more detailed description, and discussion about the issues underlying the methodology.

Agricultural census returns are recorded to a parish (pre-2000), or to a holding address (from 2000 onwards), but the exact location of the agricultural land was until recently unknown. In the case of parishes, the land may not even be within the parish that the holding address falls into (see fig. 1a). Although the 1990 Land Cover Map of Great Britain (LCMGB, Fuller *et al.*, 1994) might have seemed a suitable dataset to map the census data to, there were considerable issues with the identification of arable land in this dataset (see Lord, 1999 for further details). Hence, the decision was made to create a new dataset using different techniques, although the LCMGB was used in its creation.
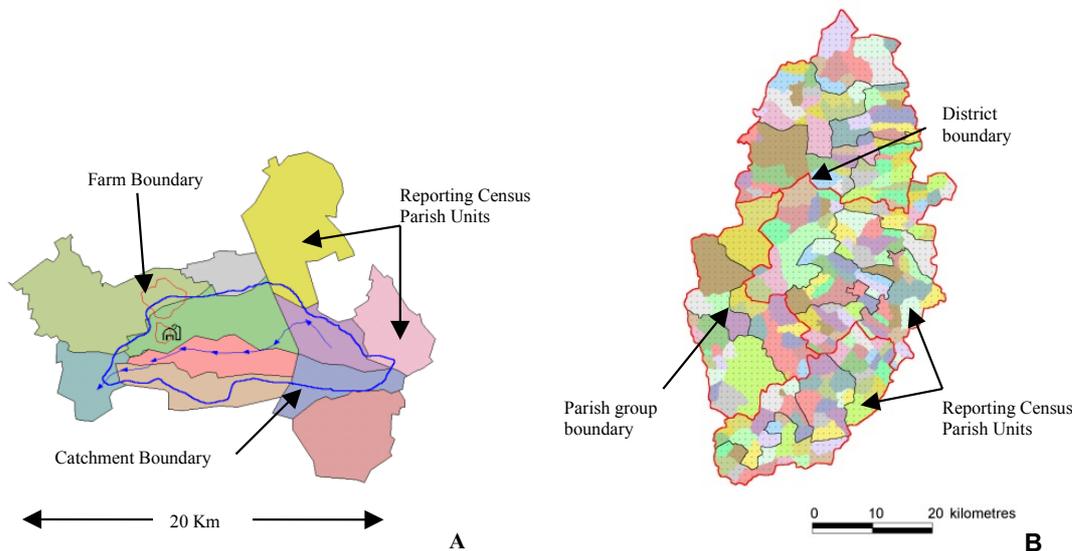
Fig. 1. Spatial units of importance to the project. Fig. 1A shows farm, parish and catchment boundaries. Fig. 1B shows parish, parish group and agricultural district boundaries. The centres of the 1km^2 cells are shown as faint grey dots. From ADAS (2001).

Non-agricultural land use mask
The first step in the production of the ADAS land use database (originally for 1995) was the development of a non-agricultural land use mask. The non-agricultural land use mask maps the areas where we are certain that agriculture does *not* occur. This dataset was derived from OS Strategi® data (using the urban land, water, woodland and infrastructure layers) and the database of land registered as Common Land. Linear features were buffered by their expected widths (see Lord, 1999 for details). This vector dataset was subsequently summarised into 1km^2 cells, giving the area covered by each of the non-agricultural land uses.

Agricultural land dataset
The area within the 1km^2 cells that is not covered by the non-agricultural land use mask, is potentially agricultural land. A technique was therefore developed to allocate land areas from the LCMGB to the 1km^2 squares, based on the area of agricultural land known to exist in a larger spatial unit, and the land cover recognised by LCMGB in each cell. The larger spatial unit in this case was the agricultural district, where the expectation was that the errors of omission and co-mission of agricultural land that are problematic at parish level, would be evened out (see fig. 1b for an illustration of these spatial units). Where the area of 'potential' agricultural land was smaller than the area of agricultural land

reported by the June census, adjacent agricultural districts were merged for the analysis. Where this was the case the adjacent district with the largest area of 'potential' agricultural land was chosen, in the expectation that this was the district where the omission / co-mission error arose.

Land from LCMGB was allocated in order of confidence in the satellite classification, namely Sea, Inland Water, Woodland, Urban, and Rough Grazing. Agricultural land was not allocated explicitly, but was calculated after the allocation algorithm as the difference between the non-agricultural land and the total area of the cell. The algorithm ensured that the total area of agricultural land reported by the cells within the district was identical to that reported by the June census returns. A full description of the above process is found in Lord (1999).

The total areas of arable and grassland were initially inferred from LCMGB by using the proportion of arable land to agricultural grassland in each cell. These proportions were subsequently adjusted by using the proportions reported by the June census at an agricultural parish level, and finally were made to match the district totals by adjustment at the agricultural district level. This final dataset of agricultural land cover matches the total areas of arable land and grassland reported in the June census. A technically detailed description of the process is given in ADAS (2001).

Distribution of cropping and livestock.
Livestock are distributed across the cells using stocking densities derived at a parish group level (see fig. 1b for an illustration of the spatial unit). The stocking densities used the number of animals reported within the parish group by the June census, but the area of agricultural land reported within the parish group by the ADAS land cover database, thereby ensuring that the total number of animals reported by the new dataset matches that of the June census at a parish group level. Grazing animals such as cattle and horses are distributed across intensively managed grassland. Sheep are distributed across grassland and rough grazing, while pigs and poultry are distributed across all intensively managed agricultural land. It is important to bear in mind that the ultimate use of the dataset was for nutrient modelling, and hence the distribution of animals was used as a proxy for the distribution of animal excreta and manures. The exact location of housed animals was, as such, not an issue in the development of this dataset.

The distribution of cropping was more complicated, as there was a requirement to maintain local cropping patterns while maintaining the June census district and national crop areas. Therefore, this distribution was carried out in two stages. Firstly, the crop proportions reported by the June census at parish level were applied to the arable areas in the 1km^2 cells. Secondly, these proportions were then adjusted by using the district total areas until a match was achieved, while maintaining the local pattern as much as possible. A description of the process, which was developed over time, is found in ADAS (2001) and subsequently Lyons et al. (2004). This process finished the development of the ADAS land cover / land use dataset, and is the same for each year the dataset has been produced for (so far, 1970, 1980, 1995, 2000, 2004, 2009 and 2010).

Issues specific to the 2010 version of the dataset.
In the June census of 2010, a number of changes were made to the Defra questionnaire which meant some of the livestock categories were reported in less detail. In addition, the data received from Defra did not have the same detail for poultry as was given on the census form. Therefore, the following additional process were followed, using data at parish group level:
1. Pigs. Previous versions of the June census had a number of detailed weight categories for fattening pigs, important in nutrient modelling. In the 2010 census, these had been reduced to 'pigs < 20 kg' and 'fattening pigs >= 20 kg'. Using the proportions in the detailed categories from the 2009 census, the amalgamated category of 'fattening pigs >= 20 kg' has been disaggregated into the original weight categories.
2. Poultry. Defra supply for layers was for laying hens and pullets together. The data were disaggregated using the 2009 proportions, as above.

Interpretation of the dataset
When using the ADAS land cover / land use database, a number of issues need to be kept in mind.

1. The dataset is highly disclosive (fig. 2 gives some examples). Under the Data Protection Act, no output from the June census may be published that contains information relating to less than five holdings. Tables must not show information relating to less than five holdings in any cell (or from which such information can be deduced), and narrative statements must not

include information about an individual farm holding, whether directly identifiable or not. A data table is available and should be provided with the dataset that shows the number of farms contributing data to each census item, for parish groups (livestock) and agricultural district (cropping). This table can be joined to the 1 km^2 dataset to show areas of disclosivity for each of the census items, and care should be taken during reporting to adhere to the legal requirements of the Data Protection Act.

2. The dataset may only be used for the purposes for which it was supplied, and which is specified in the licence.

3. The dataset is not suitable for use at a very local scale due to the inherent uncertainties in the base data layers (location of land in the June census and classification errors in LCMGB), as well as errors introduced during the iterative methodology in order to match the dataset to the census. Results for individual $1km^2$ cells are expected to be unreliable, and the recommended minimum area of use is $25km^2$.

4. For non-decadal years, the dataset is based on a survey rather than a census. Defra uses an extrapolation methodology to infer changes in holdings that did not report, based on their last return. However, at the end of the decade some holdings have not filed a return for nine years, and may have changed their business structure completely, bought or sold land, or gone out of business altogether. Therefore, the reliability of the dataset decreases as it gets further away from the last full census year.

5. The non-agricultural land use mask (vector) was only developed once, and the process continues to use the LCMGB from 1990. Therefore, large-scale land use changes that have happened since these datasets were developed (e.g. urban expansion, motorway developments or new forestry planting) will not be reflected in the dataset. Likewise, arable or grassland conversion will not be taken account of in the base data (LCMGB), but will be reflected in the final data layer to a certain degree as it is made to match the census at district level.
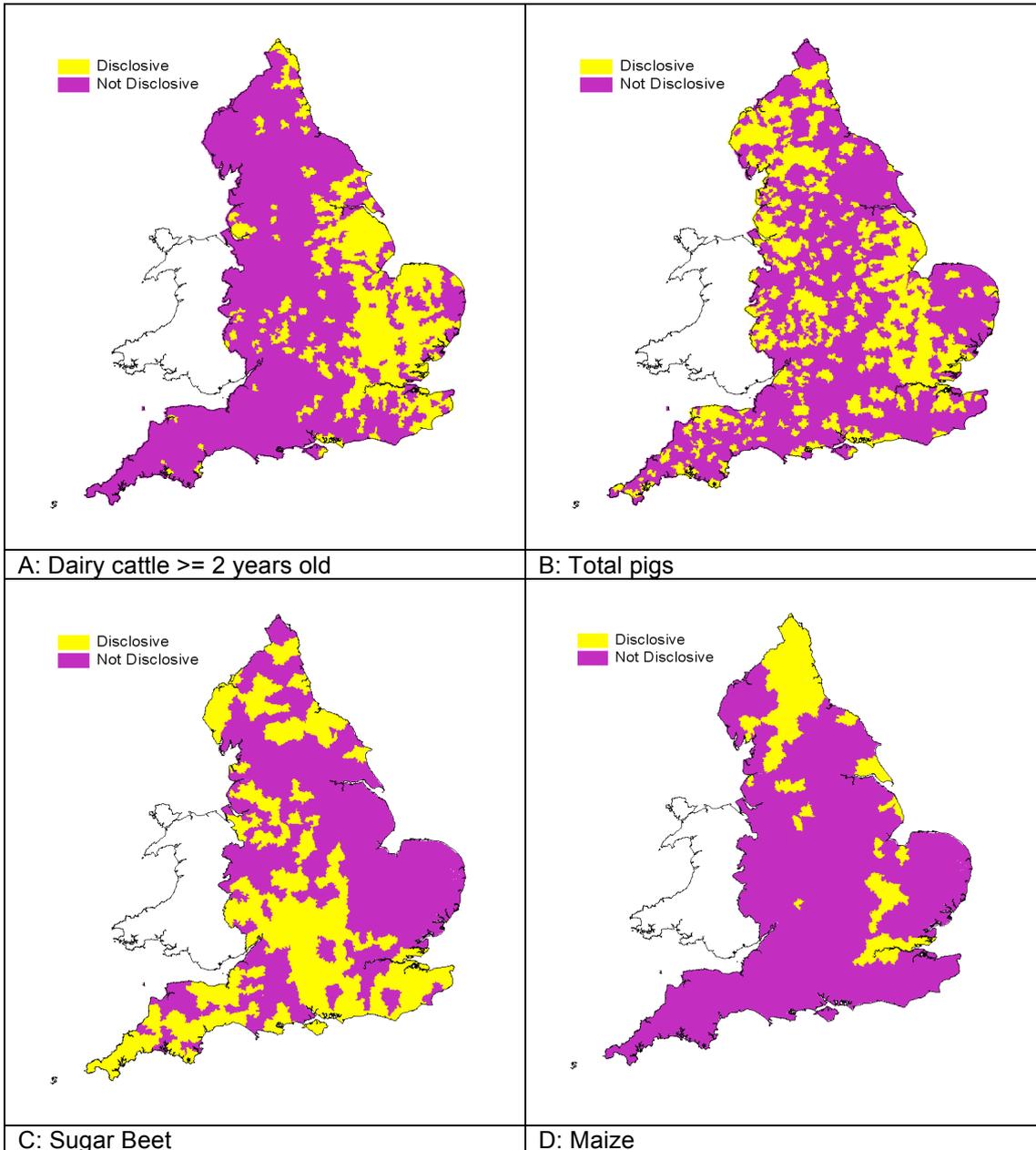
Fig. 2. Areas where the ADAS land use database 2010 is disclosive (i.e. containing data for less than five holdings), at two spatial levels: livestock (A and B) at parish group level, cropping (C and D) at district level.

References

ADAS (2001). *Maintenance of the NT1701 Nitrate Modelling System for MAFF Policy Support.* Defra Project NT2203 Final Report. Defra, UK.
Comber, A., Procter, C., & Anthony, S. (2008). The Creation of a National Agricultural Land Use Dataset: Combining Pycnophylactic Interpolation with Dasymetric Mapping Techniques. *Transactions in GIS* **12 (6)**: 775-791.
Fuller, R., Groom, G. & Jones, A. (1994). The Land Cover Map of Great Britain : An automated classification of LANDSAT Thematic data. *Photogrammetric Engineering and Remote Sensing*. **60**, 553-562.
Lord, E (1999). *Development of an agricultural land use spatial dataset to aid agricultural / environmental spatial modelling and assessment.* Defra Project NT2206 Final Report. Defra, UK.
Lyons, H., Anthony, S., & Henderson, D. (2004). *Changes in the development of the 1 x 1 km ADAS agricultural land use database since the publication of the NT2203 final report.* ADAS Internal Note, ADAS, UK.