16. Development of Bayesian models of future land use change (WP 2.13)

R. Milne
Centre for Ecology & Hydrology, Bush Estate, Penicuik

16.1 Introduction

The guidance (IPCC 2003) for countries required to submit annual estimates of emissions and removals of carbon dioxide to/from the atmosphere under the UNFCCC recommends that land use change should be considered using a matrix of changes of area. A matrix contains data of not only changes in the area in any land category between years but information on the areas moving between each different pair of categories. This detail is required because the rates on emission or removal vary between the different transitions, e.g. carbon is normally lost (as CO_2) more quickly when land is disturbed than is taken up in the reverse process. In most countries annual estimates of land in different categories is usually available but the different transitions for a matrix are seldom produced annually.

In the UK the Forestry Commission, Defra and other bodies produce official land use data annually. The detail is best in England but generally the areas in forestry, agriculture and other land types are published. However land use change (LUC) matrices are only produced intermittently by CEH for Defra as the Countryside Survey. These have been carried out in 1984, 1990, 1998 and another is due in 2007. They allow the land type transition data to be constructed by revisiting the same locations at each survey date and recording the change in land use on a field by field basis.

The land types used for the UK GHG Inventory are Forest Land, Grassland, Cropland, Settlements and Other Land. Grassland is for some estimation purposes split between managed and unmanaged grassland. These types are labelled differently (see caption Equation 16-1) but are directly equivalent.

The question to be addressed in this section of the Land Use Change GHG Inventory contract is whether it is possible to infer annual adjustments to the "medium term" land use change matrices produced from the intermittent surveys using the annually published land areas.

The primary difficulty in answering this is that if there are *n* land categories then the complete the matrix information on *n (n-1)* transitions is required. However the difference in the annual data between two years only provides *n* values. Over longer periods addition data is available from the annual data but these cannot be used directly to assess annual changes to the LUC matrix. It is therefore proposed to use a LUC matrix in which the variation in its elements are described by a simple time series model and the parameters of that model are calibrated against the annual land area data using Bayesian statistical methods.

16.2 Model structure

The approach to the model of using land use change matrices to track changes in stocks of carbon is shown by Equation 16-1.

$$\begin{bmatrix} {}_{A}A_t \\ {}_{G}A_t \\ {}_{W}A_t \\ {}_{D}A_t \\ {}_{O}A_t \end{bmatrix} = \begin{bmatrix} {}_{AA}p & {}_{GA}p & {}_{WA}p & {}_{DA}p & {}_{OA}p \\ {}_{AG}p & {}_{GG}p & {}_{WG}p & {}_{DG}p & {}_{OG}p \\ {}_{AW}p & {}_{GW}p & {}_{WW}p & {}_{DW}p & {}_{OW}p \\ {}_{AD}p & {}_{GD}p & {}_{WD}p & {}_{DD}p & {}_{OD}p \\ {}_{AO}p & {}_{GO}p & {}_{WO}p & {}_{DO}p & {}_{OO}p \end{bmatrix} \begin{bmatrix} {}_{A}A_{t-1} \\ {}_{G}A_{t-1} \\ {}_{W}A_{t-1} \\ {}_{D}A_{t-1} \\ {}_{O}A_{t-1} \end{bmatrix}$$

Equation 16-1: Land use change transition or probability matrix. *p* is probability of transition = fraction of land changing. Each column gives the probability of an area of land e.g. Arable in column 1, changing to a different use. Row 1 gives the probability of land remaining in same use and each other row gives the probability for the transition to a different use e.g. Arable to Grassland The sum of the probabilities in each column is 1 because all land remains in existence. ${}_{X}A_t$ and ${}_{X}A_{t-1}$ area areas of land of type "X" in years *t* and *t-1*. Subscripts: *A* – Arable land (IPCC Cropland), *G* – Grassland (IPCC Grassland), *W* – Woodland (IPCC Forest Land), *D* – Developed land (IPCC Settlements), *O* – Other land (IPCC Other Land

Land use change data is not normally available as the probability or fraction of change for each land use transition between reference dates but as the area of change (or no change) between these dates (Equation 16-2). The probabilities of change are estimated by dividing each entry in a matrix column by the sum of the column.

$$LUC = \begin{bmatrix} {}_{AA}a & {}_{GA}a & {}_{WA}a & {}_{DA}a & {}_{OA}a \\ {}_{AG}a & {}_{GG}a & {}_{WG}a & {}_{DG}a & {}_{OG}a \\ {}_{AW}a & {}_{GW}a & {}_{WW}a & {}_{DW}a & {}_{OW}a \\ {}_{AD}a & {}_{GD}a & {}_{WD}a & {}_{DD}a & {}_{OD}a \\ {}_{AO}a & {}_{GO}a & {}_{WO}a & {}_{DO}a & {}_{OO}a \end{bmatrix}$$

Equation 16-2: *a* is area changing between land types or remaining unchanged. The sum of the matrix columns give the initial areas for Arable (${}_{A}A_{t-1}$), Grassland (${}_{G}A_{t-1}$), Woodland (${}_{W}A_{t-1}$), Developed (${}_{D}A_{t-1}$) and Other Land (A^{O}_{t-1}) *respectively*

Although it is natural to think of the total area in a country that will change from one use to another over a specific period this form of data cannot be readily used by a mathematical model. It is also the case that the total change for the country is made up of decisions by many individual land owners and will involve statistical variability hence an overall probability of change is the most appropriate basis for modelling.

If the area change data has been obtained between two dates more than a year apart (Equation 16-3) then the annual probability of change can be estimated by matrix algebra. This requires calculation of the *n*th root of the measured area matrix but this is easy using a software package that includes matrix algebra.

$$A_t = pA_{t-1}$$
$$A_t = p^n A_{t-n}$$

Equation 16-3: Probability of change of land use over many years using matrix multiplication.

Equation 16-1 describes a LUC matrix that is constant in time, which is the intrinsic assumption from resampling surveys over a specific period. The matrix provides the cumulative change over the period and hence the annual probability matrix is an average for the period. Our purpose however is to construct a matrix whose elements change with time. In principle it would be possible to have every matrix element change every year but the amount of data required to construct this is simply unavailable in the UK at the moment. An alternative approach is to model the matrix element variation using a simple time series model. Each matrix element in such a model would have an initial value that would then change with time but would retain some memory of previous values. A stochastic element is also introduced at each time step. Only LUC transitions that were believed to change significantly with time would require this structure and the probabilities of land not changing can be calculated from the knowledge that the column sum must equal unity. The equation for an example transition, Arable to Grassland, is shown by Equation 16-4. For each LUC matrix transition element where variation is significant two additional parameters are required over the assumption of constant change probability. The two parameters are that controlling memory of past values and the variance of the random process used to introduce stochastic variability at each step.

$${}_{AG}p_t = {}_{AG}r \, {}_{AG}p_{t-1} + {}_{AG}e_t$$

Equation 16-4: Simple model of variation ${}_{AG}p$ the probability of land changing from arable to grassland. ${}_{AG}r$ is a constant that controls the "memory" of past change. ${}_{AG}e$ is a zero mean random process with constant variance that controls external variability introduced to the parameter. *p* is a constant where r = 1.

16.3 Area data

In order to assess the usefulness of the model outlined above and to explore calibration methods recent data for land use and change have been chosen. The Forestry Commission reports annually the area of forest land in each of England, Scotland, Wales and Northern Ireland (FC 2006). The June Agricultural Census is conducted in each of the four UK countries by the appropriate agriculture departments. However comprehensive data on developed areas is only readily available for England. The Ordnance Survey prepares this data on changes in urban land use in England for The Department of Communities and Local Government (DCLG 2006).

The annual data from 1990 to 2005 for the area of Arable Land, Grassland, Woodland, Developed Land and Other Land in England was therefore chosen as the test series for the calibration of the parameters of an annual LUC transition probability matrix model. The time series are shown in Figure 16-1. Longer term annual data from 1950 onwards is available for woodland and agriculture in Great Britain but other land uses and the situation in Northern Ireland are less well documented. Some information is available from the Monitoring Landscape Change reports (MLC 1986) and from surveys in Northern Ireland and this has been used in the GHG Inventory. Further work will be required to construct annual data for each country for each land type but this has been postponed until after initial testing and calibration of the matrix model using the 1990 to 2005 English data.

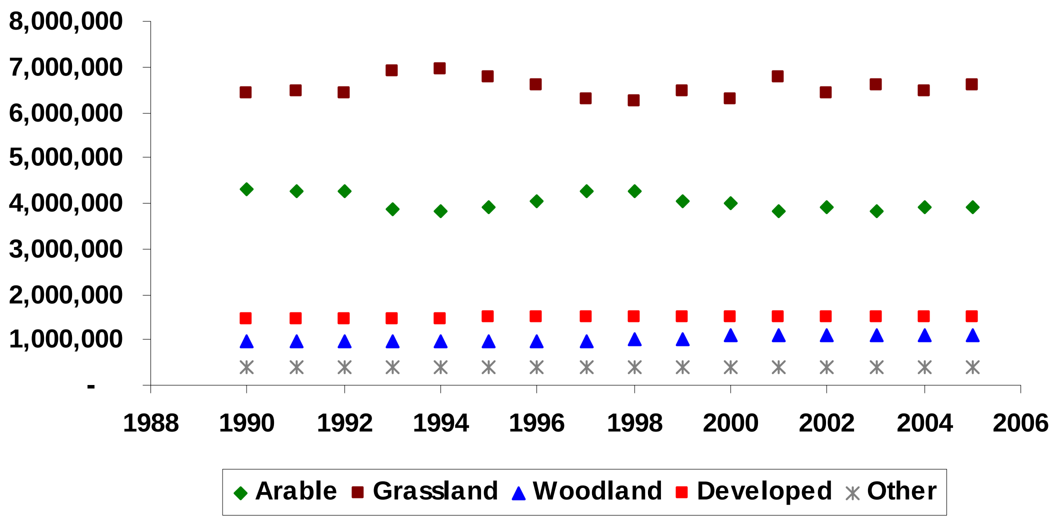

Figure 16-1: Published areas of land use in England. Area of Other land estimated by difference of sum of Arable, Grassland, Woodland and Developed from total area of England.

The data from the Countryside Surveys of 1990 and 1998 (known as Countryside Survey 2000) have been used extensively for UK LUC GHG purposes and the LUC matrix over this period has been selected here to provide preliminary parameter values for the LUC probability matrix (Table 16-1).

Table 16-1: Land use change matrix for land in England for period 1990 to 1998. Units are hectares. Land in the Other category (e.g. rock, water etc) is assumed to remain unchanged

From To	Arable	Grassland	Woodland	Developed	Other	Total 98
Arable	4,053,000	503,030	4,362	5,007	-	4,565,399
Grassland	442,010	5,046,800	69,450	27,180	-	5,585,440
Woodland	27,150	71,350	1,298,000	16,680	-	1,413,180
Developed	17,030	67,690	9,938	1,396,000	-	1,490,658
Other	-	-	-	-	394,700	394,700
Total 90	4,539,190	5,688,870	1,381,750	1,444,867	394,700	13,449,377

Table 16-2: Probability of land use change in England between 1990 and 1998

From To	Arable	Grassland	Woodland	Developed	Other
Arable	0.893	0.088	0.003	0.003	0.000
Grassland	0.097	0.887	0.050	0.019	0.000
Woodland	0.006	0.013	0.939	0.012	0.000
Developed	0.004	0.012	0.007	0.966	0.000
Other	0.000	0.000	0.000	0.000	1.000

Table 16-3: Annual probability of land use change in England on average over period 1990 to 1998. Calculated as 1/8th power of matrix in Table 16-2.

From To	Arable	Grassland	Woodland	Developed	Other
Arable	0.9854	0.0123	0.0001	0.0004	0.0000
Grassland	0.0135	0.9845	0.0068	0.0025	0.0000
Woodland	0.0007	0.0017	0.9922	0.0015	0.0000
Developed	0.0004	0.0016	0.0009	0.9957	0.0000
Other	0.0000	0.0000	0.0000	0.0000	1.0000

16.4 Bayesian calibration and initial testing

The model proposed for assessing annual land use change matrices has many parameters relative to the available data (i.e. published annual area of land use). It is therefore unlikely that the modelled could be fitted using normal statistical techniques. The annual area data is however also subject to uncertainty and from the Countryside Survey data for land use change there is some information on the uncertainty of the matrix elements. It therefore proposed to use Bayesian methods to calibrate the matrix elements (and time series parameters where these are used) to maximise the likelihood of element, i.e. probability of change, values given the uncertainty of the annual area data. Van Oijen *et al* (2005) have described a numerical method of varying the parameters of a model using a Markov Chain Monte Carlo simulation and tracking the likelihood of the output values (in this case annual area of land uses) from the model compared to the measured (in this case published) values until convergence is achieved.

The model described above was implemented within an Excel spreadsheet. Uncertainty ranges for the probability elements of the matrix were set from the Countryside Survey matrix and knowledge of the uncertainty due to the sampled nature of the survey. It has been assumed that the matrix is constant over the period 1990 to 2005 for initial testing and that the annual area data have an uncertainty of +/- 100,000 ha. A full execution of the MCMC procedure to likelihood convergence has not yet been carried out but a test run for 5000 different sets of probability elements starting with those from Table 16-2 is illustrated in Figure 16-2.

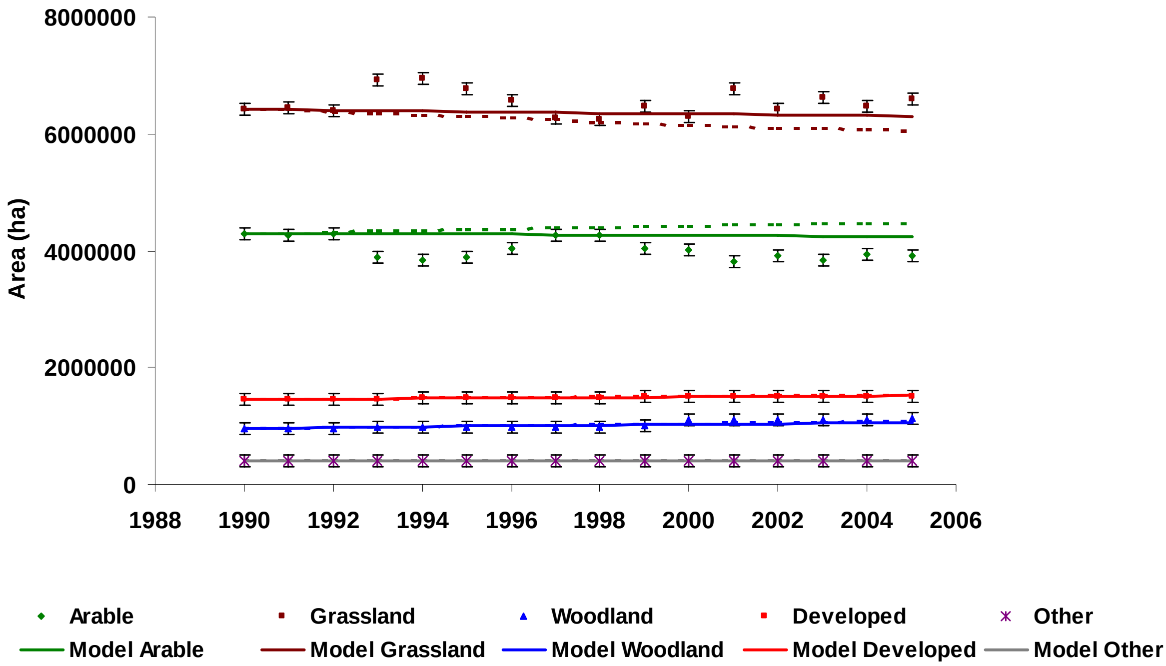

Figure 16-2: Preliminary testing of Bayesian calibration of constant land use transition matrix model for England. Graph shows trend in annual areas of land use estimated from probability matrix with values from Countryside Survey before (dotted line) and after Bayesian calibration (solid line) compared to published annual data (points).

16.5 Next steps

- Include stochastic model for matrix elements into Excel spreadsheet
- Calibrate time variable model against 1990 to 2005 English annual data

16.6 References

DCLG (2006) *Land use Change Statistics* http://www.communities.gov.uk/index.asp?id=1146601

FC (2006) *Forestry Statistics 2006*, http://www.forestry.gov.uk/website/ForestStats2006.nsf/byunique/index_main.html

IPCC (2003). *Good Practice Guidance for Land Use, Land-Use Change and Forestry*. Institute for Global Environmental Strategies (IGES) for the Intergovernmental Panel on Climate Change, Kanagawa, Japan.

MLC (1986). *Monitoring Landscape Change*. Report prepared by Hunting Surveys & Consultants Ltd for Department of the Environment and the Countryside Commission.

van Oijen, M., Rougier, J. and Smith, R. (2005) Bayesian calibration of process-based forest models: bridging the gap between models and data. *Tree Physiology*, **25**, 915 – 927.