



UK Centre for
Ecology & Hydrology

Improving Land Use Change Tracking in the UK Greenhouse Gas Inventory: Final Outputs Report

Client Ref: Department of Energy Security and Net Zero (formerly Department of Business, Energy and Industrial Strategy)

Issue Number 1

UKCEH, Forest Research, Ricardo, Agri-Food & Biosciences Institute.

Date 10/11/2021



afbi AGRI-FOOD
& BIOSCIENCES
INSTITUTE



Improving Land Use Change Tracking in the UK Greenhouse Gas Inventory: Final Outputs Report

Title Improving Land Use Change Tracking in the UK Greenhouse Gas Inventory: Final Outputs Report

Client Department for Business, Energy and Industrial Strategy (now Department of Energy Security and Net Zero)

Client reference Contract Ref 415000038732

UKCEH reference UKCEH project 07643 Tracking Land Use/Combined Final Outputs Report

UKCEH contact details Amanda Thomson
t: 0131 4458575
e: amath@ceh.ac.uk

Author Peter Levy (1), Clare Rowland (1), Gwen Buys (1), Liz Clark (2), Vera Correia (2), Ben Ditchburn (2), Christopher Evangelides (3), Alex Higgins (4), Christopher Marston (1), Daniel Morton (1), Rodrigo Olave (4), Aneurin O'Neil (1), Beth Raine (1), Sam Tomlinson (1), Anthony Walker (2), John Watterson (3), Esther Whitton (2), Jenny Williamson (1) and Amanda Thomson (1)

(1) UK Centre for Ecology & Hydrology, (2) Forest Research, (3) Ricardo, (4) Agri-Food & Biosciences Institute.

Approved by Amanda Thomson

Signed



Date 10/11/2021

Table of Contents

| | |
|---|----|
| WORK PACKAGE-A APPROACH AND DATASETS REPORT | 6 |
| Work Package-A Approach and Datasets Executive Summary | 7 |
| 1 Introduction | 8 |
| 1.1 Tracking land-use change | 8 |
| 1.2 Approach | 8 |
| 2 Mathematical representation of land use and land-use change | 10 |
| 2.1 Notation | 10 |
| 2.1.1 U | 10 |
| 2.1.2 A | 10 |
| 2.1.3 B | 11 |
| 2.1.4 G and L | 11 |
| 2.2 Data Assimilation | 11 |
| 2.3 Rationale for two-stage procedure | 12 |
| 3 QA/QC Using Chess Data | 12 |
| 3.1 Introduction | 12 |
| 3.2 Examples | 13 |
| 4 Data Sources | 20 |
| 4.1 Countryside Survey | 22 |
| 4.2 Monitoring Landscape Change | 29 |
| 4.3 National Countryside Monitoring Scheme | 29 |
| 4.4 Agricultural Census | 30 |
| 4.4.1 Results | 31 |
| 4.5 Agricultural Holdings Data | 36 |
| 4.6 Forestry Commission Data | 37 |
| 4.7 IACS | 39 |
| 4.8 Land Cover Map | 40 |
| 4.9 Land Cover Plus Crops | 42 |
| 4.10 CORINE Land Use | 44 |
| 4.10.1 Background | 44 |
| 4.11 CROME | 47 |
| 4.12 Ordnance Survey Land Use Change Statistics | 47 |
| 4.13 Land Use and Cover Area Frame Survey | 50 |
| 5 Comparison of Data Sources | 51 |
| 5.1 Results | 51 |
| 5.2 Discussion | 56 |
| 6 Estimating B by Least-Squares Optimisation | 57 |
| UKCEH report ... version 1.0 | 1 |

- 6.1 Introduction..... 57
- 6.2 Methods 58
- 6.3 Results 58
- 6.4 Discussion 62
- 7 Estimating B by MCMC 63
 - 7.1 Introduction..... 63
 - 7.2 Methods 63
 - 7.2.1 Observational data 63
 - 7.2.2 The model..... 63
 - 7.2.3 Prior distributions..... 63
 - 7.2.4 Likelihood function..... 64
 - 7.2.5 MCMC algorithm 65
 - 7.3 Results 65
 - 7.4 Discussion 71
- 8 Estimating Land-Use Change Spatially..... 72
 - 8.1 Introduction..... 72
 - 8.2 Methods 73
 - 8.3 Results 73
 - 8.4 Discussion 88
- 9 Discussion 88
- 10 References 91
- WP-A QA REPORT..... 92**
- 1 Model QA Log for LUC Track..... 93
 - 1.1 Documentation..... 93
 - 1.2 Structure & Clarity..... 93
 - 1.3 Verification 93
 - 1.4 Validation 94
 - 1.5 Data & Assumptions..... 94
- 2 Comparison to Approaches Used by Other Countries 94
 - 2.1 Approach 95
 - 2.2 Responses from Key People 95
 - 2.2.1 Marcelo Rocha..... 95
 - 2.2.2 Paulo Canaveira, Portugal 96
 - 2.2.3 Eric Arts, Netherlands..... 97
 - 2.3 Review Relevant Recent Literature and Conferences..... 97
 - 2.3.1 Analysis of LULUCF actions in EU Member States as reported under Art. 10 of the LULUCF Decision 97
 - 2.3.2 Towards Spatially Explicit Land Representation in LULUCF Inventories 98

| | |
|--|------------|
| WORK PACKAGE A FINAL RESULTS REPORT | 100 |
| Work Package A Final Results: Executive Summary..... | 101 |
| 1 WP-A Introduction..... | 103 |
| 1.1 Tracking land-use change..... | 103 |
| 1.2 Approach | 105 |
| 2 WP-A Quantifying Random Uncertainty in Land-Use Data Sources..... | 106 |
| 2.1 Introduction..... | 106 |
| 2.1.1 Representation of random uncertainty..... | 106 |
| 2.2 Methods | 107 |
| 2.2.1 Estimating σ^{obs} from reference data..... | 107 |
| 2.2.2 Effects of survey frequency and survey start/stop dates..... | 108 |
| 2.3 Results | 108 |
| 2.4 Discussion | 109 |
| 3 WP-A Quantifying Systematic Uncertainty in Land-Use Data Sources..... | 110 |
| 3.1 Introduction..... | 110 |
| 3.1.1 Representation of systematic uncertainty | 110 |
| 3.2 Methods | 111 |
| 3.2.1 Creating reference datasets: | 111 |
| 3.2.2 False positive rates | 113 |
| 3.2.3 False negative rates..... | 113 |
| 3.3 Results | 113 |
| 3.3.1 Data-source-specific false positive and false negative rates..... | 116 |
| 3.4 Updating B, D, G and L..... | 117 |
| 3.5 Conclusions..... | 118 |
| 4 WP-A Assessing systematic errors in estimates of land-use change: sensitivity to survey-interval length..... | 118 |
| 4.1 Introduction..... | 118 |
| 4.2 Methods | 119 |
| 4.3 Results | 120 |
| 4.3.1 IACS:..... | 121 |
| 4.3.2 LCM..... | 122 |
| 4.3.3 LCC: | 123 |
| 4.3.4 CROME:..... | 124 |
| 4.3.5 CORINE: | 125 |
| 4.3.6 Summary Table..... | 125 |
| 4.4 Discussion and Conclusions..... | 127 |
| 5 WP-A Estimating false-positive rates in detection of land-use change based on classification accuracy..... | 128 |

| | | |
|-------|--|------------|
| 5.1 | Introduction..... | 128 |
| 5.2 | Methods | 128 |
| 5.3 | Results | 129 |
| 5.4 | Conclusions..... | 131 |
| 6 | WP-A Using life tables in modelling land-use change | 131 |
| 6.1 | Introduction..... | 131 |
| 6.2 | Methods | 132 |
| 6.3 | Results | 132 |
| 6.4 | Discussion | 134 |
| 7 | WP-A Initial Results: England..... | 135 |
| 8 | WP-A Initial Results: Scotland..... | 143 |
| 9 | WP-A Initial Results: Wales..... | 152 |
| 10 | WP-A Initial Results: Northern Ireland | 161 |
| 11 | WP-A Discussion | 170 |
| 12 | References | 172 |
| | WORK PACKAGE B FINAL REPORT | 173 |
| 1 | WP-B Introduction..... | 174 |
| 1.1 | LULUCF land categories..... | 174 |
| 1.1.1 | Forest land..... | 174 |
| 1.1.2 | Cropland | 174 |
| 1.1.3 | Grassland | 174 |
| 1.1.4 | Wetlands | 175 |
| 1.1.5 | Settlements | 175 |
| 1.1.6 | Other Land..... | 175 |
| 1.2 | Summary of Work package A | 175 |
| 2 | WP-B1: Produce Land Cover Data/Land-Use Data | 176 |
| 2.1 | Production of new Land Cover Map data (UKCEH) | 176 |
| 2.1.1 | Satellite data..... | 176 |
| 2.1.2 | Ancillary data..... | 178 |
| 2.1.3 | Training data..... | 179 |
| 2.1.4 | Classification method | 182 |
| 2.2 | Generation of new Forestry data for Great Britain (Forest Research)..... | 183 |
| 2.2.1 | Production of the 1990 woodland map | 183 |
| 2.2.2 | Production of 1995, 2000 and 2005 woodland maps | 185 |
| 2.2.3 | Production of 2010, 2015 and 2020 Interim woodland maps | 188 |
| 2.2.4 | Woodland creation data..... | 192 |
| 2.3 | Generation of new forestry data for Northern Ireland (AFBI) | 193 |

| | | |
|-------|---|-----|
| 2.3.1 | Northern Ireland 2020 Woodland Basemap | 193 |
| 3 | WP-B.2 – Optimise Land Cover Change Data | 196 |
| 3.1 | LCM Post-classification filtering to improve accuracy | 196 |
| 4 | WP-B.3 – Validation of datasets | 199 |
| 4.1 | Evaluation of biases in LCM data sets (validation) | 200 |
| 4.2 | Validation of BEIS LCM’s: individual LCMs | 201 |
| 4.3 | Comparison with IACS data | 206 |
| 4.4 | Manually derived change data sets: Change to woodland | 213 |
| 4.5 | Manually derived change data sets: Change to urban | 213 |
| 4.6 | Manually derived change data sets: Change to water | 214 |
| 4.7 | Manually derived change data sets: Change to other land..... | 214 |
| 4.8 | Summary of LCM data products..... | 214 |
| 4.9 | Discussion of LCM results..... | 215 |
| 5 | WP-B.4 – Method inter-comparison | 217 |
| 6 | Quality Assurance and Data Management..... | 217 |
| 6.1 | Data management..... | 217 |
| 6.2 | Model QA | 217 |
| 6.3 | Independent validation of satellite data classification | 218 |
| 6.3.1 | Validating through NDVI..... | 220 |
| 6.3.2 | Limitations and future recommendations | 221 |
| 6.4 | Comparison to methods used by other countries..... | 222 |
| 6.4.1 | Canada..... | 223 |
| 6.4.2 | Australia..... | 236 |
| 6.4.3 | France | 257 |
| | Acknowledgements | 260 |
| 7 | References | 260 |
| | Appendix 1: Development of a 1990 Woodland Map for the BEIS Land Use Project..... | 262 |
| | Appendix 2: Ancillary information for intercountry comparison..... | 287 |

WORK PACKAGE-A APPROACH AND DATASETS REPORT

WP-A Approach and Datasets Report

Authors: Peter Levy, Sam Tomlinson, Gwen Buys and Amanda Thomson

UK Centre for Ecology & Hydrology

Date: 14/11/2020

Work Package-A Approach and Datasets

Executive Summary

This report describes work on the project “Improving Land Use Change Tracking in the UK Greenhouse Gas Inventory” for the Department for Business, Energy & Industrial Strategy (reference TRN 2384/05/2020). The aim of the project was to make improved estimates of land-use change in the UK, using multiple sources of data. We applied a method for estimating land-use change using a Bayesian data assimilation approach. This allows us to constrain estimates of gross land-use change with national-scale census data, whilst retaining the detailed information available from several other sources. We produced a time series of maps describing our best estimate of land-use change given the available data, as well as the full posterior distribution of this space-time data cube. This quantifies the joint probability distribution of the parameters, and properly propagates the uncertainty from input data to final output. The output data has been summarised in the form of land-use vectors. The results show that we can provide improved estimates of past land-use change using this method. The main advantage of the approach is that it provides a coherent, generalised framework for combining multiple disparate sources of data, and adding further sources of data in future would be straightforward. Future work could focus on more detailed analysis of existing data sets, introducing independent constraints where possible, and obtaining further relevant data sets. The code is available via GitHub.

1 Introduction

1.1 Tracking land-use change

The tracking of land use and land-use change is fundamental to producing accurate and consistent greenhouse gas inventories (GHGI) for the Land Use, Land-Use Change and Forestry (LULUCF) sector. This is necessary to meet the international requirements of the Kyoto Protocol to the UN Framework Convention on Climate Change (UNFCCC) and the Paris Agreement and the national requirements of the UK's Climate Change Act and related legislation within the UK's Devolved Administrations.

The estimation of land-use change in the current UK GHGI is based on a combination of infrequent CEH Countryside Surveys and afforestation/deforestation statistics from the GB Forestry Commission. It uses Approach 2 (non-spatial land-use change matrices) as described in the KP Guidance. However, this matrix-based approach, and its implementation in the UK, have some important limitations. Firstly, the non-spatial matrix-based approach is insufficient for tracking annual land-use change: the matrices have no time dimension and are defined independently each year. There is therefore no possibility of representing a sequence of land-use on the same parcel of land (such as afforestation followed by deforestation, or crop-pasture rotations). Secondly, the data used to estimate these matrices in the UK are rather limited. The CEH Countryside Surveys were only carried out approximately once per decade, and whilst the geographical extent was very wide, the actual ground area surveyed was small as a fraction of the total UK area. The afforestation/deforestation statistics from the Forestry Commission have good national coverage (excluding Northern Ireland) but do not contain any information on the spatial location or land use prior to afforestation or following deforestation.

In October 2019, the UNFCCC Expert Review of the UK 1990-2017 GHG inventory raised concerns in relation to the reporting requirements of the second commitment period of the Kyoto protocol. They questioned whether the current approach is appropriate for the identification and tracking of lands where the elected Article 3.4 activities occur (i.e. Cropland Management, Grazing Land Management and Wetland Drainage and Rewetting). They recommended that the UK explore how to make the best possible use of available data and prepare and implement a work-plan to enable the use of these data. The UK has already explored several approaches to land use tracking, including a data assimilation approach to integrate available land-use data into land-use vectors, which was successfully piloted in Scotland (Levy et al. 2018). This project builds on that approach to assess gross land-use change, and land-use history for the whole of 7 8 SECTION 1. INTRODUCTION the UK from 1990 to 2019.

As well as improving accuracy of the GHGI, a time series of spatially explicit land-use change would enable better tracking of mitigation activities and improve baseline data for scenario modelling. These baseline data are needed for understanding the potential of land-based mitigation and adaptation options. The government's ambitions for Net Zero by 2050 or sooner means that the LULUCF sector will have an increasingly critical role in the UK's overall GHG balance. This kind of scenario modelling will become very important to inform the setting of future carbon budgets and monitor progress towards the UK's legal obligations to GHG emissions reductions. An accurate spatio-temporal land-use change data set would be useful to other stakeholders and UK government departments. For example, from the perspectives of biodiversity conservation, air quality, or ecosystem services, there are clear applications of these data for understanding and tracking the effects of land use.

1.2 Approach

If we had reliable maps of land use each year, we could infer land-use change by difference. However, even with advances in satellite sensors, GIS and spatial data handling, the accuracy of change detection from EO-based products is generally too poor to do this; the different EO products are inconsistent (with each other, and with themselves over time), irregular, and become more infrequent as we go

back in time. Change is more reliably detected by repeat ground-based surveys, but these have other short-comings. For example, the annual June Agricultural Census gives a long record of areas in different land uses, but does not provide spatial data, or any information on gross change (i.e. what land uses have changed to which other land uses). The CEH Countryside Survey did provide spatial data with gross change, but without complete coverage, and only at infrequent intervals.

In light of the above, some data assimilation method, which combines the spatiotemporal data with non-spatial repeat survey data, would appear to provide a solution. To this end, we previously developed a methodology using a Bayesian data assimilation approach, and this has been applied successfully to Scotland (Levy et al. 2018). This method allowed for the use of a wider range of data types, including high-resolution spatial data, and combined them in a mathematically coherent way. Importantly, the method produced the appropriate data structure needed for modelling the effects of land-use change on GHG emissions - the set of unique land-use vectors (i.e. unique sequences of land use, or land-use histories) and their associated areas. An important feature is that the uncertainty in land-use change can be easily propagated to provide the uncertainty in GHG emissions, because the procedure explicitly handles the distribution of plausible vectors of land-use change. The approach provides a general framework for combining multiple disparate data sources with a simple model which describes how these data sources inter-relate. This allows us to constrain estimates of gross land-use change with reliable national-scale census data, whilst retaining the detailed spatial information available from several other sources. Here, we apply this methodology to improve and update the tracking of land-use change for the UK. Our aim was to apply a Bayesian approach to make spatially- and temporally-explicit estimates of land-use change in the UK, using multiple sources of data.

The tasks required to achieve this aim were:

- A.1 updating and obtaining new data, and the necessary processing;
- A.2 running the data assimilation algorithm to produce a time series of maps (deliverable A.1);
- A.3 summarising the output in vector format and characterising the uncertainty with respect to random, systematic, and correlated error terms (deliverables A.1 & A.2);
- A.4 documenting the code and data processing workflow (deliverable A.3);

The remainder of this report describes these four tasks. The Methods section describes the basic methodology for the data assimilation. The Data Sources section and subsequent sections describes the data sources assessed for inclusion. The Data Assimilation section describes the data assimilation procedure in some more detail, with results in subsequent sections. Finally, we discuss the results compared with the previous method, and consider areas where further work is needed.

All the code is written in R using the “literate programming” paradigm implemented with Rmarkdown, which combines the source code, text/graphical output, documentation, and report text within the same document. This ensures integrity of documentation, code and corresponding outputs. All the Rmarkdown files are held in a GitHub repository, for version control and wider accessibility. The documentation is rendered using bookdown and made publicly available as a web site via GitHub Pages. This documentation describes the data processing workflow so as to make it reproducible.

1. Methods

2 Mathematical representation of land use and land-use change

We begin by describing the data structures which are used to represent land use and its change over time, **U**, **A**, **B**, **G**, and **L**, as in Levy et al. (2018).

2.1 Notation

We use the mathematical convention of representing vectors, matrices and arrays as uppercase bold (e.g. **U**), and individual elements thereof as uppercase italic (e.g. U_{xyt}). Objects from R code are shown in sans-serif typeface, e.g. `s_U[[t]][x,y]`. The R code follows the mathematical notation defined in the journal paper, and uses a consistent naming convention, based on a few rules. Details of the notation are given in the [GitHub documentation](#). In short, the maths are written in LaTeX, and the R code tries to mirror this.

2.1.1 U

The AFOLU Guidance (IPCC 2006) recommends use of six types of land for broad descriptive purposes: forest, grassland, cropland, settlements, wetlands and other land. However, the area of grassland in the UK is very large, and heterogeneous in terms of soil carbon. For the purposes of modelling soil carbon in the current GHGI, grassland is subdivided into improved grassland and semi-natural/rough grazing land, on the basis of aggregating classes in the CORINE classification. The soil carbon parameters associated with these two grassland types are given by Bradley et al. (2006). Wetlands in the UK are very small, consisting only of those areas undergoing active commercial peat extraction, areas of inland water and flooded land. GHG emissions associated with peat extraction are calculated separately, and emissions associated with transitions to/from wetlands are considered negligible.

Here, we represent land use u as a number of discrete states from the set {woods,crops,grass,rough,urban,other}{woods,crops,grass,rough,urban,other}, encoded as integers 1-6. “Woods” means all forested land, “grass” is short for improved grassland, and “rough” is short for rough grazing and semi-natural land.

At a single location (x,y) , land use can change between these states over time, represented by the vector \mathbf{U}_{xy} . Spatially, we represent land use on a grid, where each grid cell contains a vector of land use. Combining the spatial and temporal dimensions, we have the 3-D space-time array $\mathbf{U}=\{U_{xyt}\}$.

2.1.2 A

We denote the area occupied by each land use u at time t as A_{ut} , obtained by counting the frequency of land uses in \mathbf{U}_t :

$$1. \quad A_{ut} = \sum_{x=1}^{n_x} \sum_{y=1}^{n_y} [U_{xyt} = u] A_{\text{gridcell}}$$

where the square brackets are Iverson notation, evaluating to 1 where true and zero otherwise, and A_{gridcell} is the area of a single grid cell. We denote the set of all these areas as $\mathbf{A}=\{A_{ut}\}$. By difference, we obtain the areas of net land-use change:

$$2. \quad \Delta A_{ut} = A_{ut} - A_{ut-1}$$

2.1.3 B

At each time step, we have a square transition matrix, which we collect in the 3D array **B**:

$$\mathbf{B} = \begin{bmatrix} 0 & \beta_{12} & \beta_{13} & \dots & \beta_{1n_u} \\ \beta_{21} & 0 & \beta_{23} & \dots & \beta_{2n_u} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \beta_{n_u1} & \beta_{n_u2} & \beta_{n_u3} & \dots & 0 \end{bmatrix}_{t=1} \begin{bmatrix} 0 & \beta_{12} & \beta_{13} & \dots & \beta_{1n_u} \\ \beta_{21} & 0 & \beta_{23} & \dots & \beta_{2n_u} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \beta_{n_u1} & \beta_{n_u2} & \beta_{n_u3} & \dots & 0 \end{bmatrix}_{t=2} \dots \begin{bmatrix} 0 & \beta_{12} & \beta_{13} \\ \beta_{21} & 0 & \beta_{23} \\ \vdots & \vdots & \vdots \\ \beta_{n_u1} & \beta_{n_u2} & \beta_{n_u3} \end{bmatrix}$$

where the elements represent the gross area changing from one land use to another that year, and there are n_u possible land uses. For example, β_{23} in $\mathbf{B}_{t=2}$ is the area changing from land-use type 2 to land-use type 3 between years 1 and 2. The transition matrix at time t can be derived from the array layer \mathbf{U}_t by comparison with the previous layer at $t-1$. Each element is given by

$$3. \quad \beta_{ij} = \sum_{x=1}^{n_x} \sum_{y=1}^{n_y} [U_{xyt-1} = i \wedge U_{xyt} = j] A_{\text{gridcell}}$$

2.1.4 G and L

For a given time step, the net change in the area occupied by each land use is given by the gross gains (the vector of column sums, G_{ut}) minus the gross losses (the vector of row sums, L_{ut}):

$$G_{ut} = \sum_{i=1}^{n_u} \beta_{iu} \quad \text{for time } t$$

$$L_{ut} = \sum_{j=1}^{n_u} \beta_{uj} \quad \text{for time } t$$

where i and j are the row and column indices, respectively, and

$$4. \quad \Delta A_{ut} = G_{ut} - L_{ut}$$

2.2 Data Assimilation

Almost all of the information on the history of land use can be expressed in the form of **U**, **B**, **A**, **G**, and **L**, and the equations above tell us how these are inter-related. **U** contains complete information about the system. **B** contains partial information about the system, which can be summarised in the form of **A**, **G**, and **L**, but does not directly specify **U**. In themselves, **A**, **G**, and **L** do not directly specify either **U** or **B** but can be used as constraints in their estimation.

Observations from the data sources described in subsequent sections provide information in the form of one or more of the data structures described above. Our approach here is to use the above equations as a simple model to relate the different observations via data assimilation in a two-stage process. Firstly, we use a Bayesian approach to estimate the parameters in **B**, given all the available information. Secondly, we use the posterior distribution of **B** and likelihood for the location of land use and land-use change to estimate the posterior distribution of **U**. The maximum *a posteriori* probability (MAP, the mode of the posterior distribution) realisations represent our best estimate of land use and land-use change, given the available data.

We refer to this as a “Bayesian data assimilation” procedure. The term “data assimilation” is usually applied to methods which use observed data to update the *state variables* of a model; terms such as “calibration” and “inversion” are usually applied to methods which use observed data to update the *parameters* of a model. The distinction is not very important, as the underlying mathematics is essentially the same, and particularly here where the distinction between parameters and states is itself sometimes blurred. We use the term “data assimilation” because it better conveys the idea that observations are brought together to form something greater than the sum of their parts. The procedure is Bayesian because it updates the parameters according to Bayes Theorem, i.e. using prior knowledge and an explicit calculation of the probability of the data given the parameters (the likelihood). This has the advantages that it provides a robust means of estimating uncertainty on the estimates, and it allows us to combine data from different sources, as well as informative priors.

2.3 Rationale for two-stage procedure

The ultimate problem is to estimate the state of the 3-D space-time array $\mathbf{U}=\{U_{xyt}\}$ using the available observations from different data sources. If we are only interested in the best estimate of *land use* at a given location and time, the assimilation of data is relatively simple. We want to estimate $p[\hat{u}|\mathbf{u}]$, that is, the probability of an estimated land-use state \hat{u} being correct, given a vector of observations \mathbf{u} . For example, with a vector of observations $\mathbf{u}=\text{"crop", "crop", "grass"}$ from three data sources (e.g. LCM, IACS and CORINE), we might estimate the probability that the true land-use state is “crop” to be 2/3. If we know the precision of these different observations, we can adjust this accordingly.

However, if we are interested in *land-use change at national scale*, the problem is more complex. The crux of this is that calculating the probability of correctly estimating a change to a new state u^{new} at location xy is no longer independent at each location. For example, the prediction of whether a particular forest changes to grassland at time t depends on how many other forests we predict to change to grassland elsewhere. Put mathematically, the probability of correctly estimating a change to the new state $p[\hat{u}_{xy}^{new}|\mathbf{u}_{xy}]$ becomes dependent on the estimates of the B matrix, so

$$p[\hat{u}_{xy}^{new}|\mathbf{u}_{xy}, \mathbf{B}]$$

But similarly, \mathbf{B} is dependent on the areas changing state. This inter-dependence of the two key unknowns makes it difficult to solve in a single step. For this reason, we split the problem into two parts: firstly, we estimate the \mathbf{B} matrix using all available data; secondly, given the constraint of this matrix, we estimate the locations of land-use change. These two problems are different in nature, and accordingly, we solve them in different ways: the former we solve using MCMC; the latter we solve using a form of importance sampling.

3 QA/QC Using Chess Data

3.1 Introduction

The game of chess provides a convenient analogue for land-use change. Because the system is small enough to count and visualise easily, and because we know the true state at all times, this is an ideal way to test our algorithms, providing quality assurance and quality control (QA/QC) for our code.

Chess is played on a 8 x 8 board, so a suitably small domain of 64 squares. We can consider the chess board to have three states analogous to land uses: occupied by either white pieces or black pieces, or empty. Mathematically, $u=\text{white, empty, black}$. At the start, there are two rows occupied by

white, two rows occupied by black, and four empty rows ($A_{\text{white}}=16$, $A_{\text{black}}=16$, $A_{\text{empty}}=32$). At each move by white, one piece is moved, so its origin square changes from white to empty, and its new square changes either from empty to white, or from black to white (if an opposing piece is taken). So for example, after the first move by each side, we have a **B** matrix:

```
##          white empty black
## white    15     1     0
## empty     1    30     1
## black     0     1    15
```

As the game proceeds and pieces are captured, the numbers of white and black cells decrease (A_{white} and A_{black}) and the number of empty cells (A_{empty}) increases. The starting position is always the same, so there are predictable spatial patterns in where each state is likely to be found. Thousands of chess games are freely available online and can be used as pseudo-data for showing how the data assimilation procedure works, and testing it with simple cases.

3.2 Examples

The plots below show some example analyses which test the functions for (i) the calculation of the **B** matrix from the time series of spatial maps (or board states) **U** and (ii) the calculation of the time series of area **A**, and the gross gains and losses **G** and **L**, from the **B** matrix.

Here we use a single game of chess to provide the true course of land-use change over 83 years (moves). To simplify, we restrict this to the states at $t = (1, 11, 15, 80)$ years. If we designate this to be the true states, we can simulate two sets of observations with noise by selecting different time slices, in this case $t = (1, 3, 13, 60)$ and $t = (1, 12, 45, 83)$. This simulates data sets which are related to the true state but differ because of measurement error. These **U** states are plotted below.

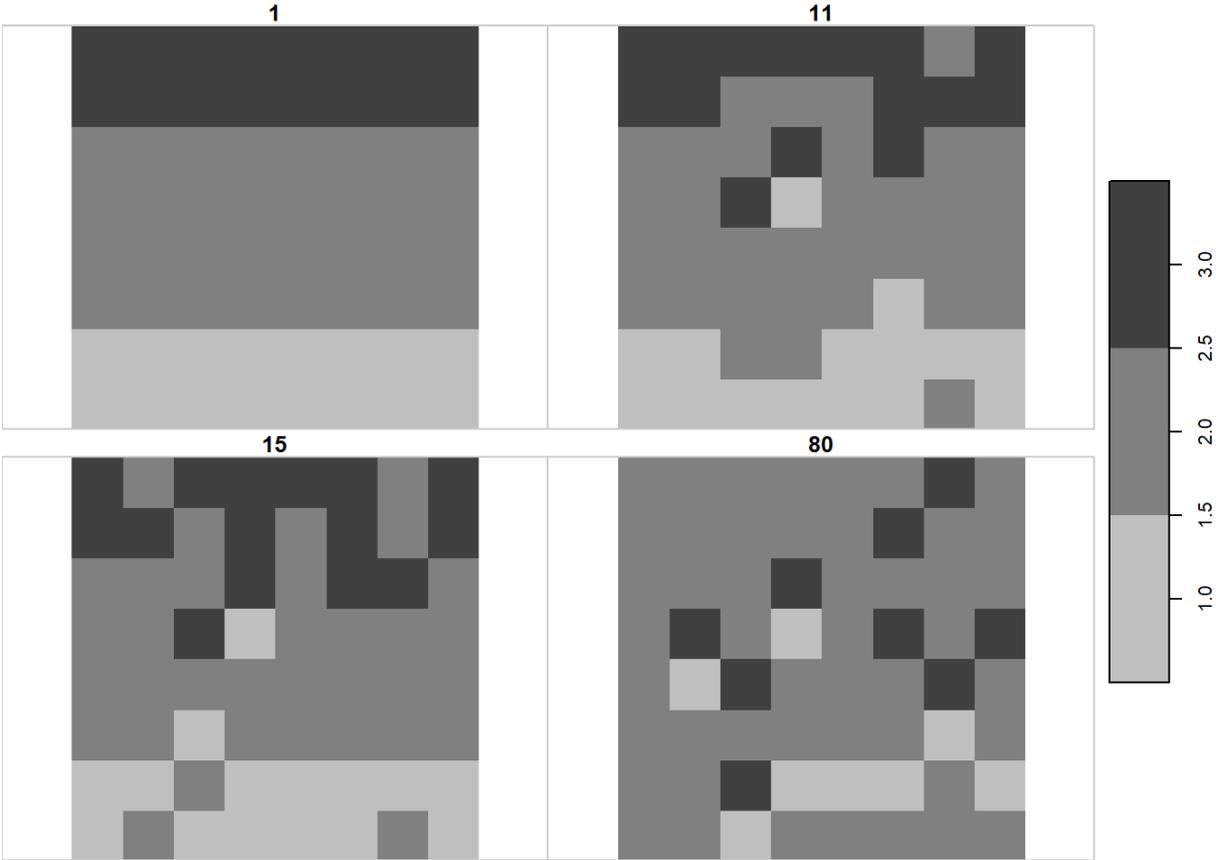


Figure 3.1: True state of land use U at four time points, based on chess data

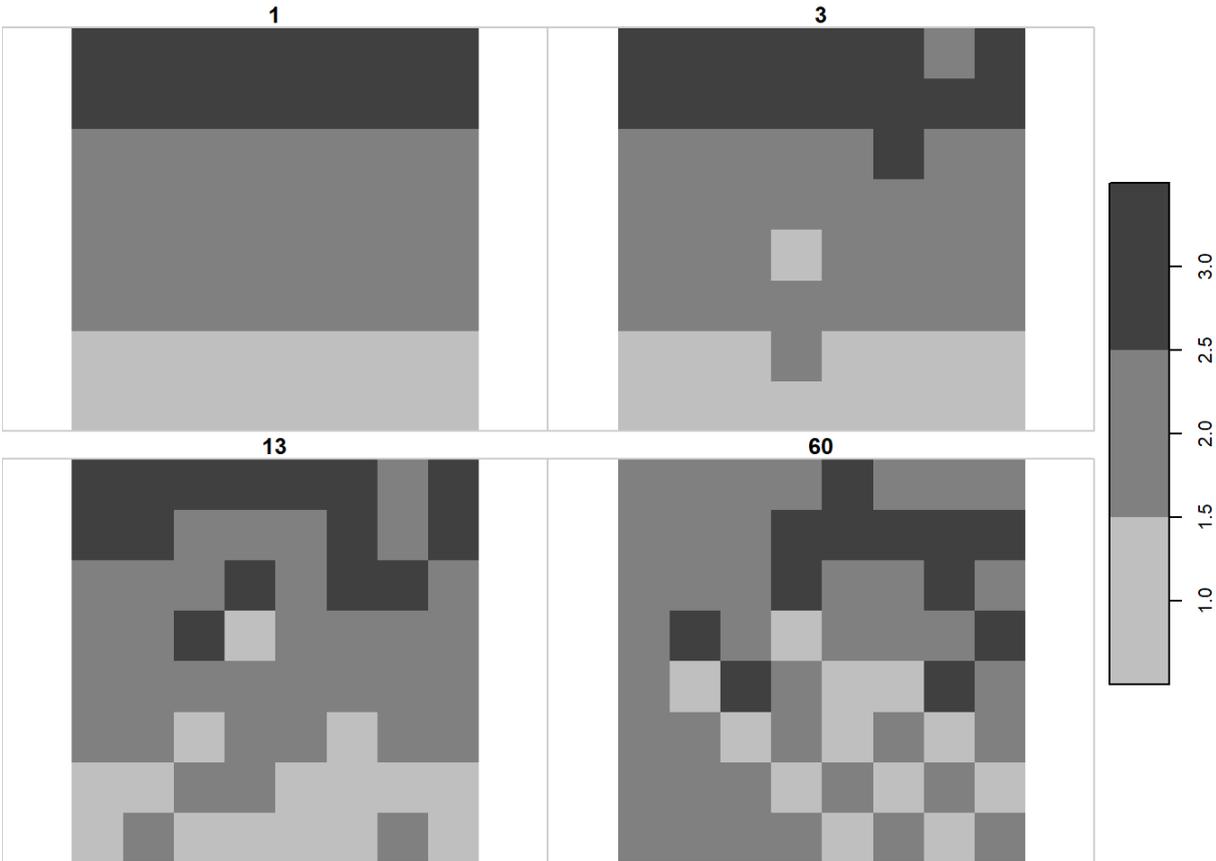


Figure 3.2: Simulated observations of land use U at four time points, based on chess data, using time slices at different times from those designated as the true state. Denoted 'Obs1' in following figures.

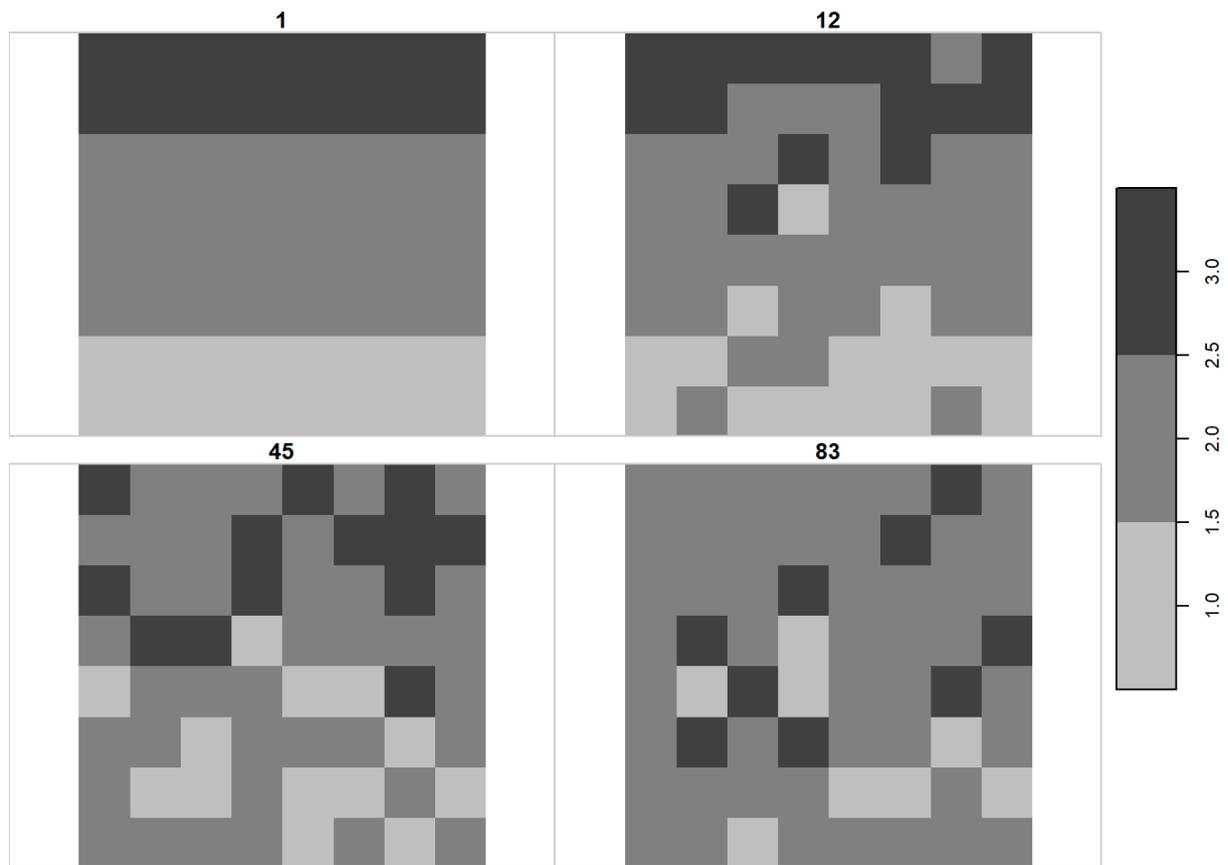


Figure 3.3: Simulated observations of land use \mathbf{U} at four time points, based on chess data, using time slices at different times from those designated as the true state. Denoted 'Obs2' in following figures.

Subsequently, we have used these to test the performance of the data assimilation algorithm when confronted with imperfect data (results not shown). Here, we use these data as a basic QA/QC test, verifying that the functions correctly calculate the time series of \mathbf{B} , \mathbf{A} , \mathbf{G} , and \mathbf{L} . Verification is by simple cross-checking with the known values, which are easily calculated. The area of each cell is assumed to be 10 m^2 to provide a proper check on the consistency of units. The plots below show the time series of \mathbf{B} , \mathbf{A} , \mathbf{G} , and \mathbf{L} .

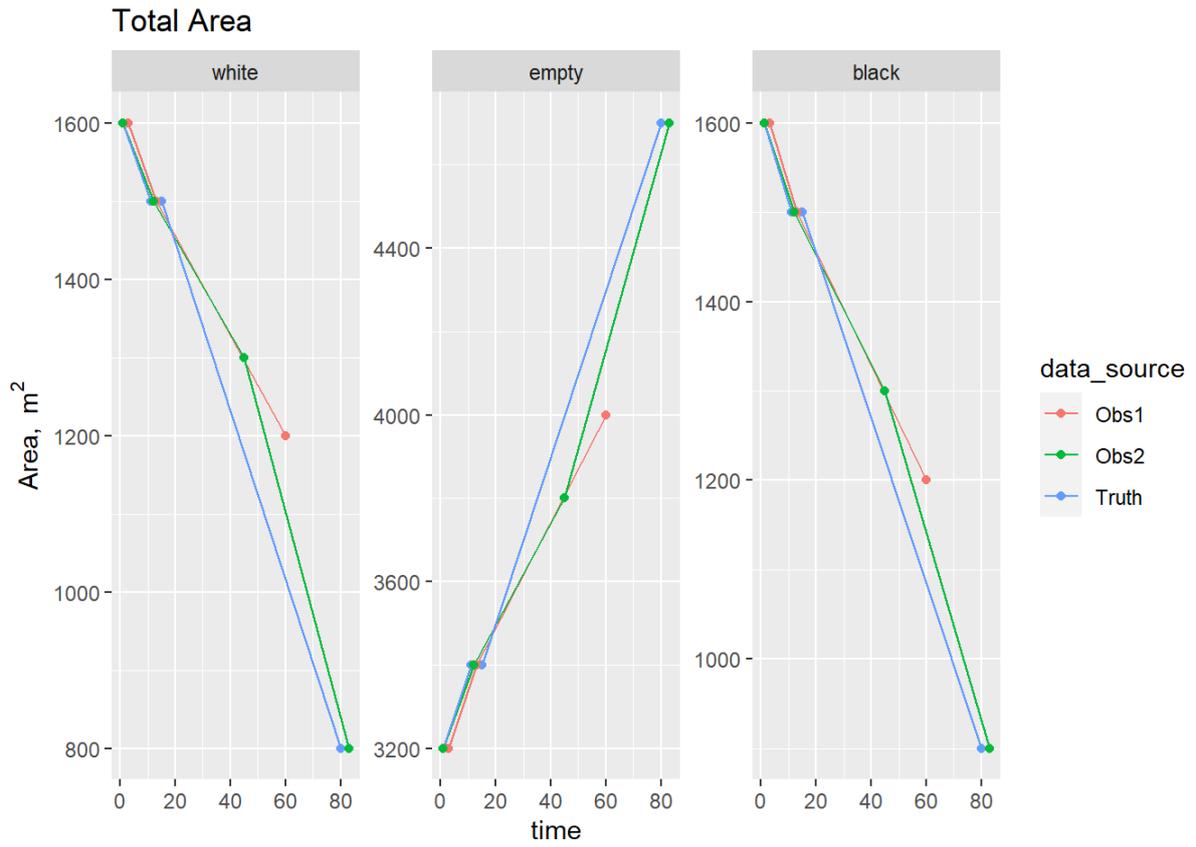


Figure 3.4: Time series of area **A** from simulated data sources based on chess data.

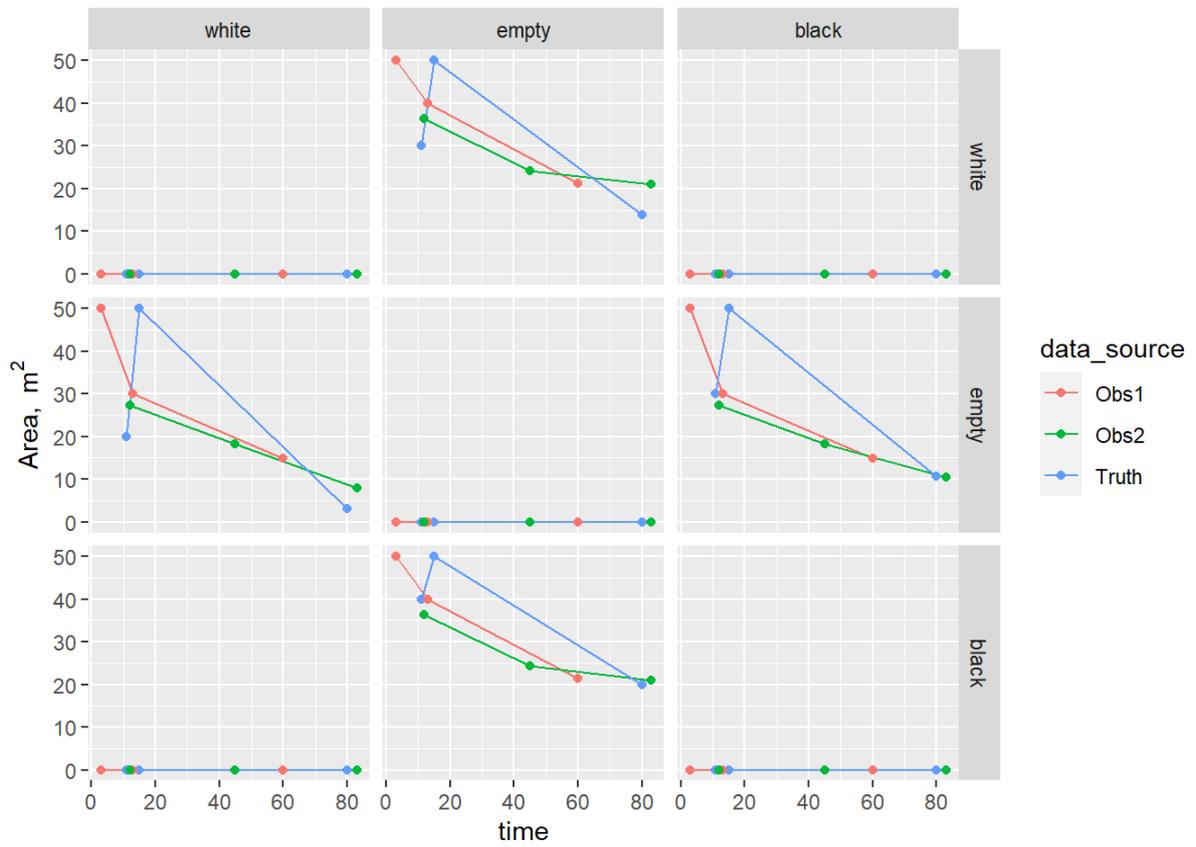


Figure 3.5: Time series of area transitions **B** from simulated data sources based on chess data.

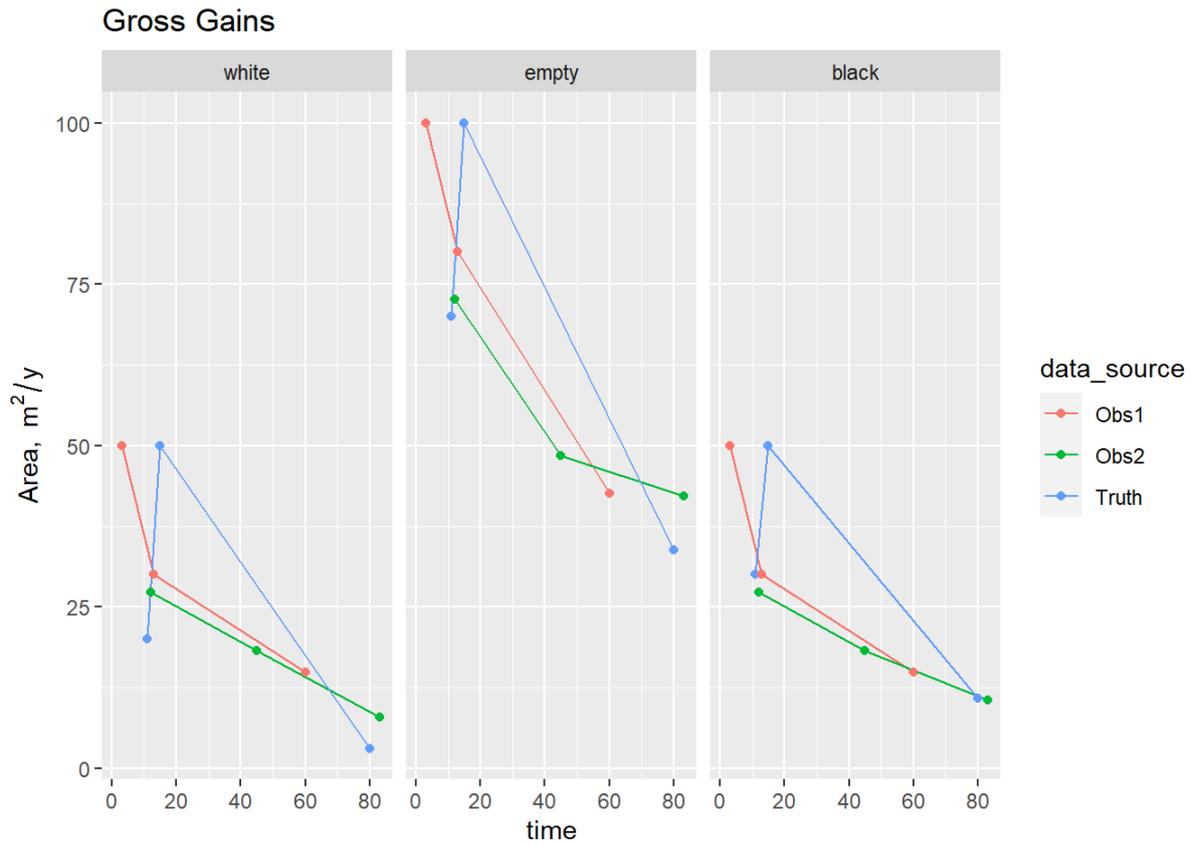


Figure 3.6: Time series of area gains **G** from simulated data sources based on chess data.

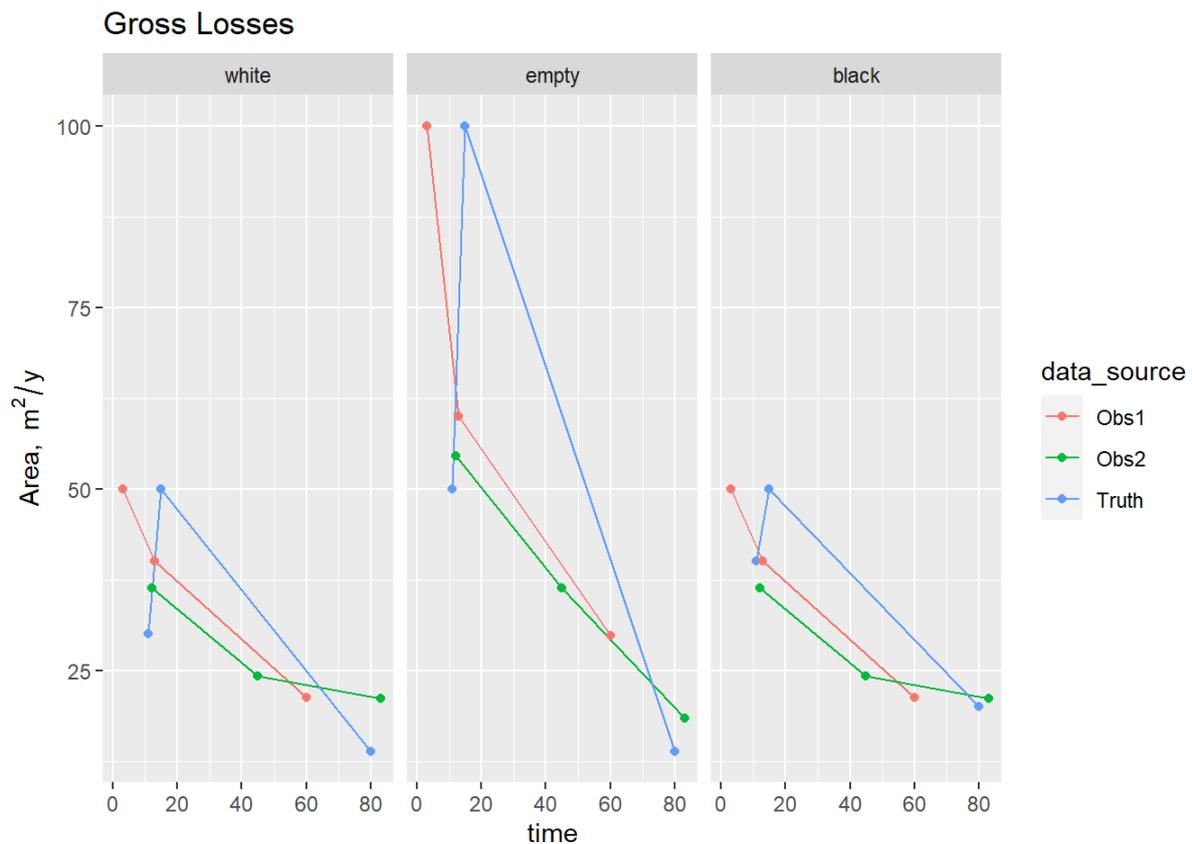


Figure 3.7: Time series of area losses L from simulated data sources based on chess data.

4 Data Sources

The UK is relatively rich in data sets that provide information on different categories of land-cover and land management over time. However, these data sets are collected for different purposes (e.g. agricultural statistics, forest inventory or urban development statistics), may have different classification definitions and spatial resolution and rarely cover the entire time span (1990-2020) or region (all countries of the UK) of interest. In the Bayesian methodology, all observations have an associated uncertainty, and it is straightforward to include multiple data sources of differing quality (i.e. observations with different uncertainties). In the previous work, eight different data sources with these varied characteristics, with three different data structures, were combined in our posterior estimate of land use and land-use change. It follows from Bayes Theorem that the impact of new data on the posterior predictions of land use will depend on their associated uncertainty, and how much they differ from other observations and prior expectations. We prioritise publicly-available data sets, because they provide transparency, ease-of-access and a higher likelihood of continuing revision and availability.

Given our starting point of having a working method for Scotland, we need to extend this to the rest of the UK. The first task here is assessing the available data and the addition of further data sources into the DA procedure. This requires updating existing data sets, and adding sources which give additional data for England, Wales and Northern Ireland. The main data sets of relevance are shown in Table 4.1.

The first stage is to review these data sets for their potential, in terms of information content, consistency, and ease of use, and prioritise work to include them on this basis. The next stage is to add these new data sets, which require a variable degree of data processing - translating other land-use classifications into the LULUCF classification, transforming raw data into one or more of the data structures we require as necessary, placing the data in a common spatial framework (the OSGB grid), and calculating time series and/or rasterising as required. Each of the data sets come with a native classification for land use or cover which require integration. To do this, the information in these data sets was assigned to the appropriate LULUCF categories using a systematic approach. The available descriptions and metadata for the dataset classes were compared to the UK LULUCF definitions and the rationale for assigning each data class will be recorded. This process was carried out with discussion across the project team, so that ambiguous cross-classifications were considered as a group and agreed. The most awkward cases are where classes are added / removed or definitions change over the time span of the dataset, which introduces discontinuities in the data, and methods for handling this are documented.

Table 4.1: Data sources considered for use in the project.

| Dataset name | Spatial coverage | Time period | LULUCF category |
|---|--|--------------------------------|--|
| Land Cover Map | UK | 1990, 2015, 2017-2019 (annual) | Forest Land, Cropland, Grassland, Wetlands, Settlement, Other Land |
| Countryside Survey | UK | 1990, 1998, 2007 | Cropland, Grassland, Settlement, Other Land |
| Northern Ireland Countryside Survey | NI | 2000, 2007 | Cropland, Grassland, Forest Land, Settlement, Other Land |
| National Forest Estate Sub-compartments | GB | Historic-2019 | Forest Land |
| National Forest Inventory | GB | Historic-2019 | Forest Land, Forest land converted to other land use categories |
| Forestry Statistics | UK | Historic-2019 | Forest Land |
| UK June Agricultural Census | England, Scotland, Wales, Northern Ireland | Historic-2019 | Cropland, Grassland |
| EDINA | GB | | |
| Monitoring Landscape Change | England, Wales | 1947, 1969, 1980 | Forest Land, Cropland, Grassland, Wetlands, Settlement, Other Land |

Table 4.1: Data sources considered for use in the project.

| Dataset name | Spatial coverage | Time period | LULUCF category |
|--|--|------------------------------|--|
| National Countryside Monitoring Scheme | Scotland | 1947, 1969, 1980 | Forest Land, Cropland, Grassland, Wetlands, Settlement, Other Land |
| CORINE Land Cover | UK | 1990, 2000, 2006, 2012, 2018 | Forest Land, Cropland, Grassland, Wetlands, Settlement, Other Land |
| UKCEH Land Cover plus crops | GB | Annual 2015-2019 | Cropland, Grassland |
| MHCLG land-use change statistics | England | Annual 2013-2018 | Settlement |
| Integrated Administration and Control System (IACS) | England, Scotland, Wales | 2014 - onwards | Cropland, Grassland |
| Holdings-level agricultural data | England, Scotland, Wales, Northern Ireland | 2000 - onwards | Cropland, Grassland |
| Northern Ireland Forest Service Sub-compartment Boundaries | NI | 2020 | Forest Land |
| Northern Ireland Woodland Base Map | NI | 2020 | Forest Land |
| Northern Ireland Agricultural Census - 5km Grid | NI | 1997-2016 | Cropland, Grassland, Rough Grazing, Farm woodland |

The specific processing and issues with each data source are described in the following sections.

4.1 Countryside Survey

The Countryside Survey (UKCEH 2020a) is a series of national (GB) field surveys carried out in 1984, 1990, 1998/99 and 2007 by the Centre for Ecology and Hydrology and the former the Institute of Terrestrial Ecology. The surveys consist of detailed field observations, collected in 1 km sample grid squares across Great Britain (384 squares in 1984, 508 in 1990 and 569 in 1998). Each successive survey resamples squares included in the previous survey. Around 0.24% of Great Britain's land surface was surveyed in the

1998 survey (Barr et al. 2003). The sampling is stratified using a system of classifying geographical locations depending on many mapped attributes ('ITE Landclass').

The Northern Ireland Countryside Survey (DAERA 2016) was initiated in the late 1980s, with a baseline land cover survey in Northern Ireland. This estimated the area of different types of Primary Habitats from a random sample set of quarter kilometre grid squares. A monitoring resurvey was carried out in 1998 to determine the extent of change using the same sample grid squares and methods as in the baseline. A third time series of the NICS was carried out in 2007 to survey 287 grid squares.

The Countryside Survey (CS) classifies land use into so-called "Broad Habitats" (Jackson, 2000). The table below shows the correspondence between CS Broad Habitats and LULUCF classes.

Table 4.2: Correspondence between Broad Habitat classes used in the Countryside Survey and the LULUCF classes used in the present work.

| | Broad.Habitat | LULUCF_ID | LULUCF_name |
|----|-------------------------|------------------|--------------------|
| 1 | Broadleaved and Mixed | 1 | forest |
| 2 | Coniferous Woodland | 1 | forest |
| 3 | Arable and Horticulture | 2 | crop |
| 4 | Improved Grassland | 3 | grass |
| 5 | Neutral Grassland | 4 | rough |
| 6 | Calcareous Grassland | 4 | rough |
| 7 | Acid Grassland | 4 | rough |
| 8 | Fen Marsh and Swamp | 4 | rough |
| 9 | Dwarf Shrub Heath | 4 | rough |
| 11 | Bog | 4 | rough |
| 12 | Inland Rock | 4 | rough |
| 13 | Saltwater | 0 | NA |
| 14 | Freshwater | 0 | NA |
| 15 | Supra-littoral Rock | 6 | other |

Table 4.2: Correspondence between Broad Habitat classes used in the Countryside Survey and the LULUCF classes used in the present work.

| | Broad.Habitat | LULUCF_ID | LULUCF_name |
|----|----------------------------|------------------|--------------------|
| 16 | Supra-littoral Sediment | 6 | other |
| 17 | Littoral Rock | 6 | other |
| 18 | Littoral Sediment | 6 | other |
| 19 | Littoral Sediment | 4 | rough |
| 20 | Built-up Areas and Gardens | 5 | urban |

For each CS 1-km survey square that coincided between survey years (544 in 1998-2007), the area that changed from one land-use class to another was calculated, using ArcGIS software. The **B** matrix values were calculated by summing these for each land-use type.

To extrapolate the CS squares to national scale, the “ITE Land Class” was used, which represents a broad classification of environments across the UK. Northern Ireland is treated as a single uniform unit as its smaller area means that it does not display the climatic or elevation variations evident across the rest of the UK. A mean transition matrix was calculated for each ITE Land Class. The national-scale matrices were calculated as the mean over all ITE Land Classes, weighted by their area coverage. To interpolate between survey dates, we assumed that the rates of change were constant during the period between surveys.

The plots below show the estimates of land-use change produced by the Countryside Survey observations. The data are shown over the whole period available, 1970 to 2020.

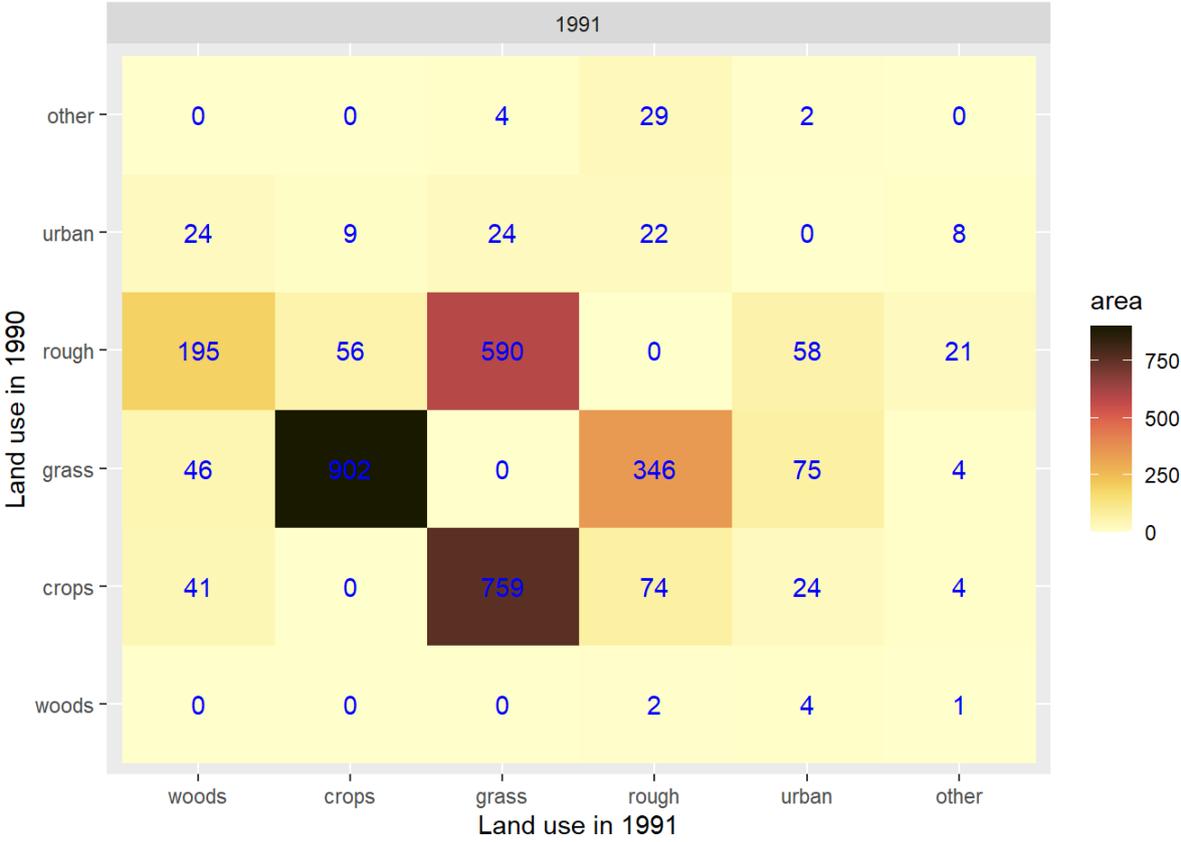


Figure 4.1: Example **B** matrix showing the areas changing land use between 1990 and 1991 in km².

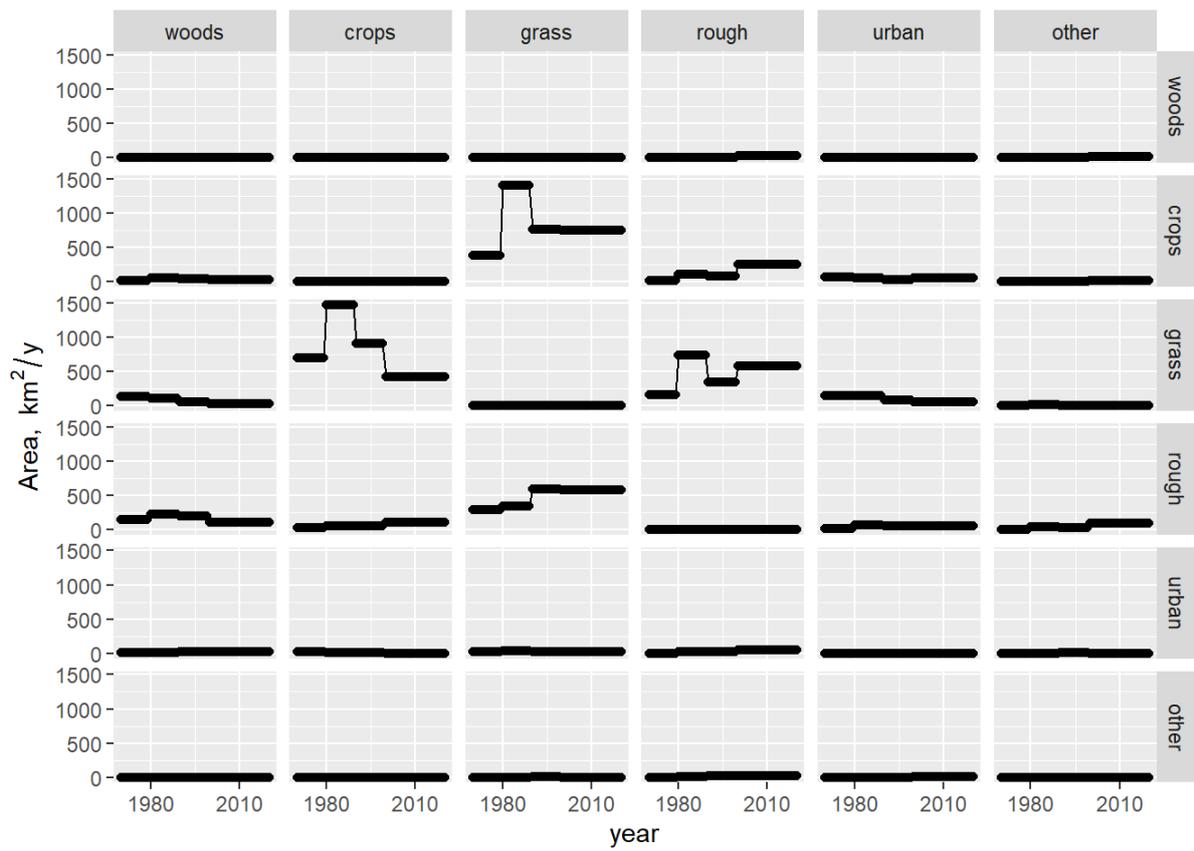


Figure 4.2: Time series of **B** matrix, showing the areas changing land use over time. The layout of panels follows the matrix itself, so rows represent the starting land use, columns represent the end land use. We assumed that the rates of change were constant during the period between surveys.

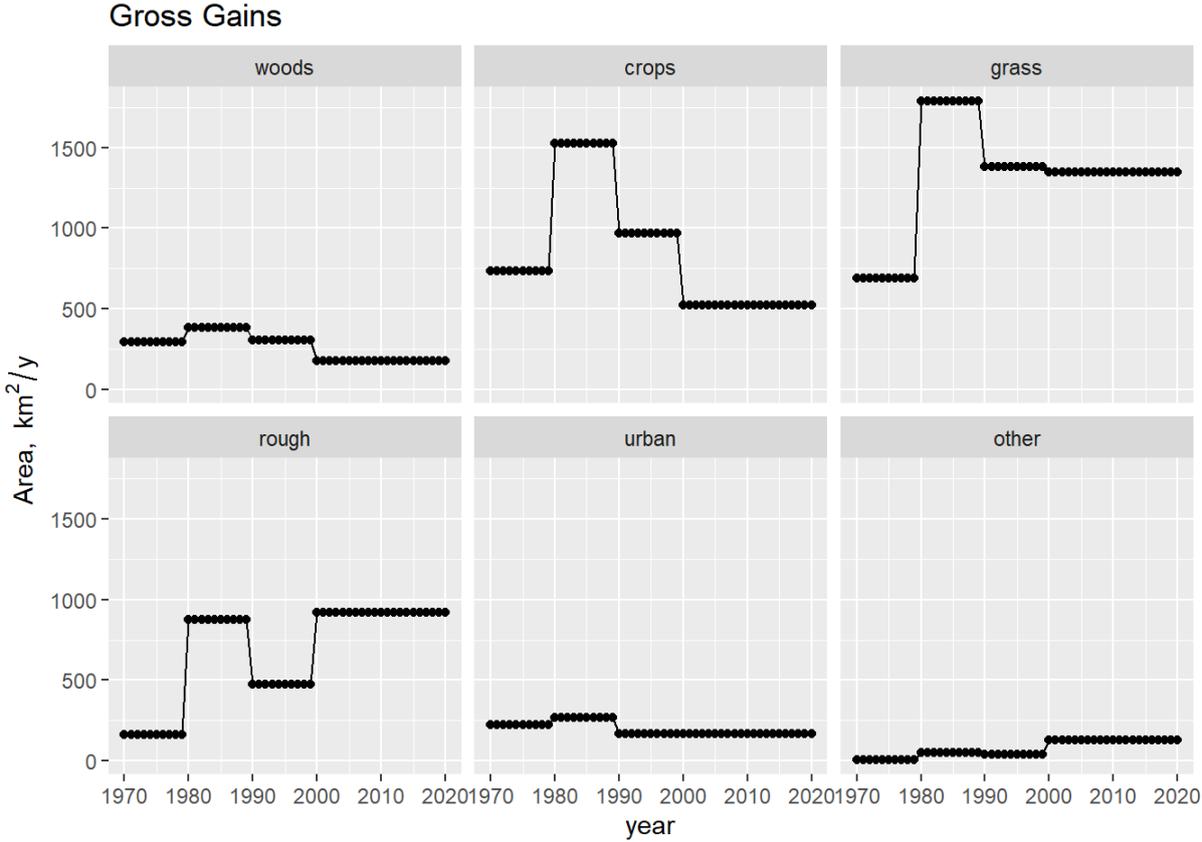


Figure 4.3: Time series of implied area gains **G** to each land use, from CS data. We assumed that the rates of change were constant during the period between surveys.

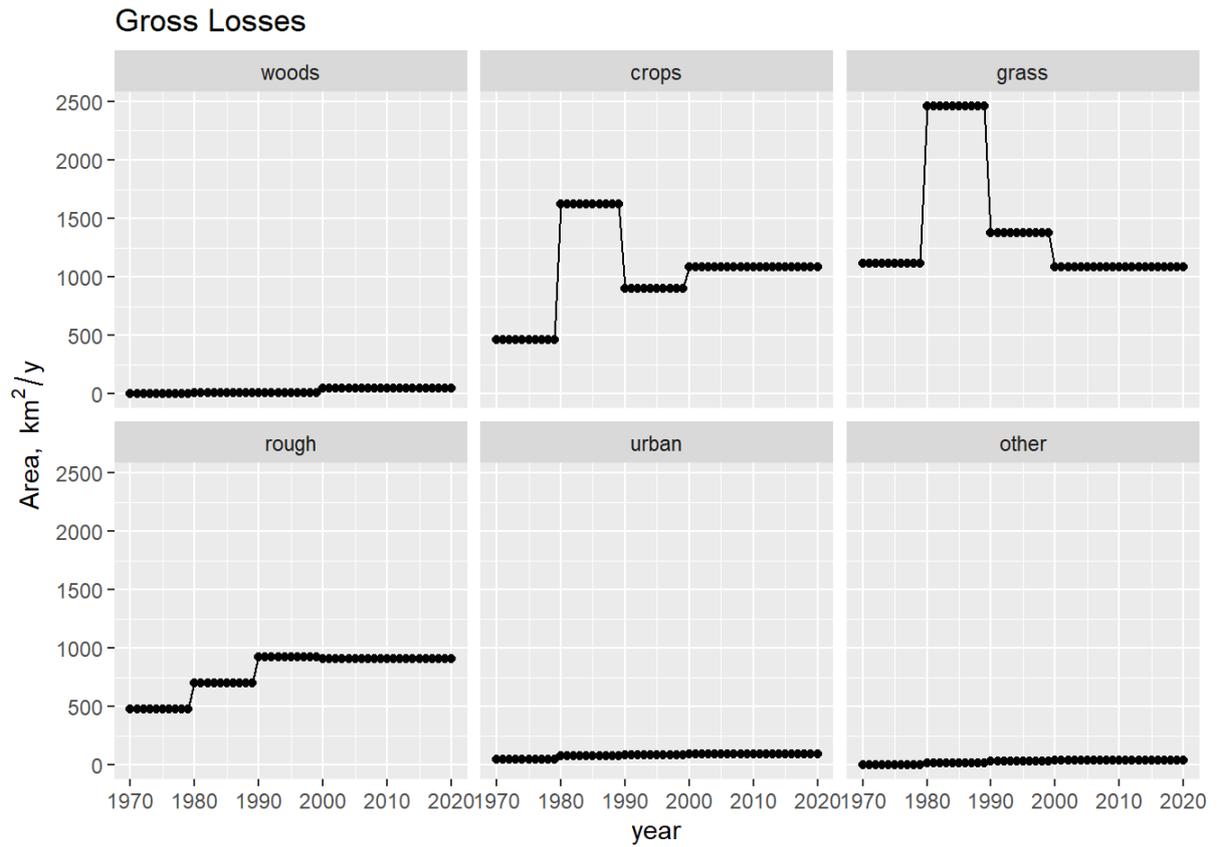


Figure 4.4: Time series of implied losses of area L from each land use, from CS data. We assumed that the rates of change were constant during the period between surveys.

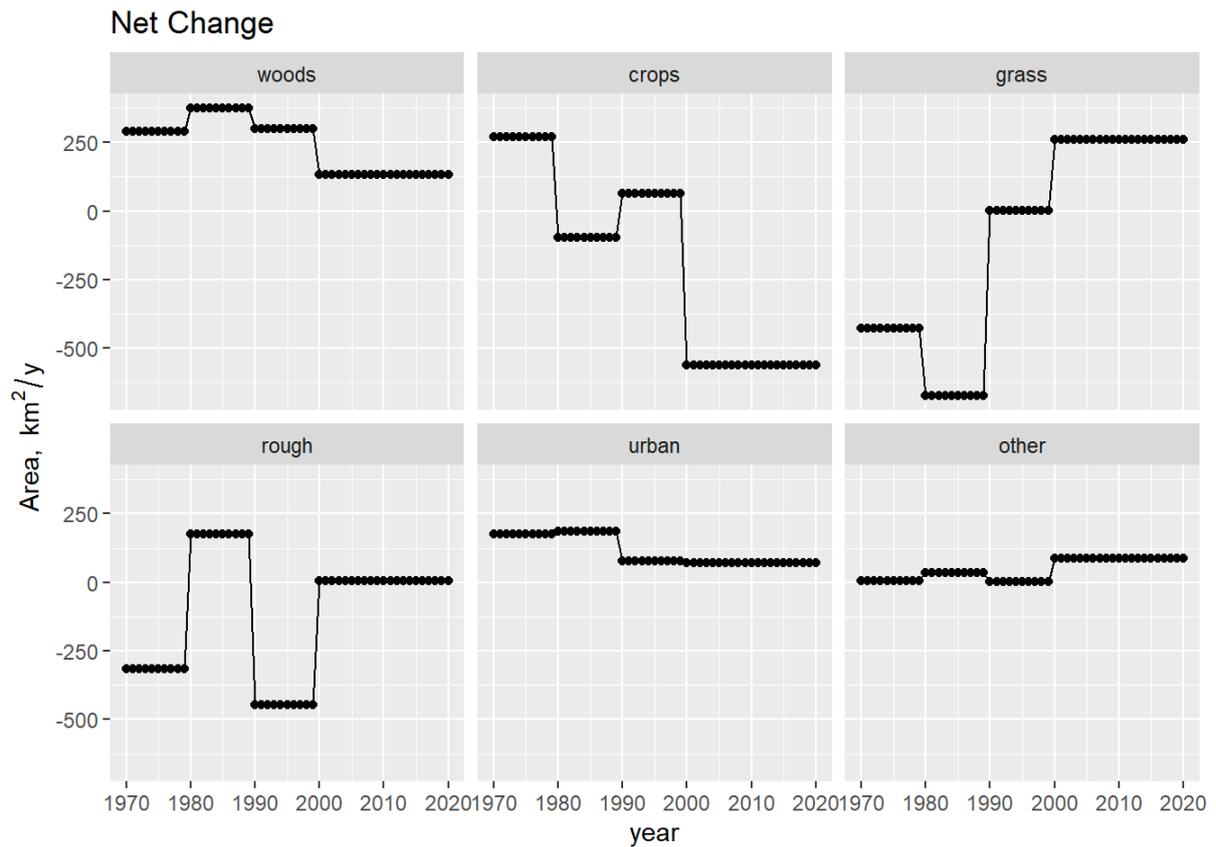


Figure 4.5: Time series of implied net change in area ΔA of each land use, from CS data. We assumed that the rates of change were constant during the period between surveys.

4.2 Monitoring Landscape Change

The Monitoring Landscape Change project (Hunting Technical Services Ltd 1986) assessed land use change in England and Wales between 1947, 1969 and 1980. The results were derived from aerial photo interpretation of stratified random samples within soil strata in each county (based on the county boundaries established in 1974). The sample covered 2.4% of England and Wales in total. Areas of different land use types (in 1947, 1969 and 1980) at the county level, and matrices of change at the regional level, are published in the MLC reports (volumes 1, 3-6 and 10) (Hunting Technical Services Ltd 1986). Because these date from 1947 & 1980, they were not used in the current project which focusses on 1990-2019, but they are available to be used in future work.

4.3 National Countryside Monitoring Scheme

This project (Mackey, Shewry et al. 1998) assessed land cover change in Scotland between 1947, 1973 and 1988 using aerial photography. A stratified random sample of 467 sample squares (7.5% of Scotland's land area) was used, the size of which was determined by the need to detect national changes of 10% or more in major features of interest, with 95% confidence. Stratification was done by the regional classification of Landsat multi-spectral scanner images. Again, because the data pre-date the 1990s, they were not used in the current project, but they are available to be used in future work.

4.4 Agricultural Census

Agricultural Census data provide annual records of the total area in the main agricultural land uses. The Agricultural Census is conducted in June each year by the government agriculture department in each of the DAs. Farmers declare the agricultural activity on their land in the form of ca. 150 items of data via a postal questionnaire. The results are collated at national scale. This is a long-running data set with near-complete coverage of agricultural land, relatively consistent over time. The Agricultural Census data are reported as national statistics and to the FAO. Hence it is desirable for our estimates of land-use change to be consistent with these data as far as possible. The classifications used differ by DA, and have varied over time, but the table below summarises the correspondence between the most consistently used Agricultural Census classes and LULUCF classes.

Table 4.3: Correspondence between the land-use classes reported in the Agricultural Census and the LULUCF classes used in the present work.

| LULUCF_ID | LULUCF_name | AgCensus_name |
|-----------|-------------|--|
| 1 | woods | Woodland |
| 2 | crops | Total.crops |
| 2 | crops | Uncropped.arable.land..No.set.aside.1983.1989...f. |
| 3 | grass | Temporary.grass..sown.in.the.last.5.years. |
| 3 | grass | Land.used.for.outdoor.pigs..g. |
| 3 | grass | Grass.over.5.years.old |
| 4 | rough | Common.rough.grazing..e. |
| 4 | rough | Sole.right.rough.grazing |

Some step changes occur in the data because of methodology changes, and the data may be reported twice for that year, from the two different methods. To remove the effect of these step changes from the data, the difference each step change makes is calculated and the difference applied to all older data. Where a step change appears in the data without any reported methodological change, the time series is smoothed to remove the disjunct, which is assumed to be an artefact of unknown methodological change. A GAM model is fitted to the data set prior the step change, and separately fitted to the data set after the step change. The difference between the two smoothed predictions at the time of the step change is used to estimate the magnitude of the artefact. This is removed from the data set prior to the step change. Clearly there is some subjectivity in this procedure, and further analysis time could be spent on this. However, some smoothing of the data is required, because several changes cannot be taken at face value as true land-use change. Results for each DA are shown below.

4.4.1 Results

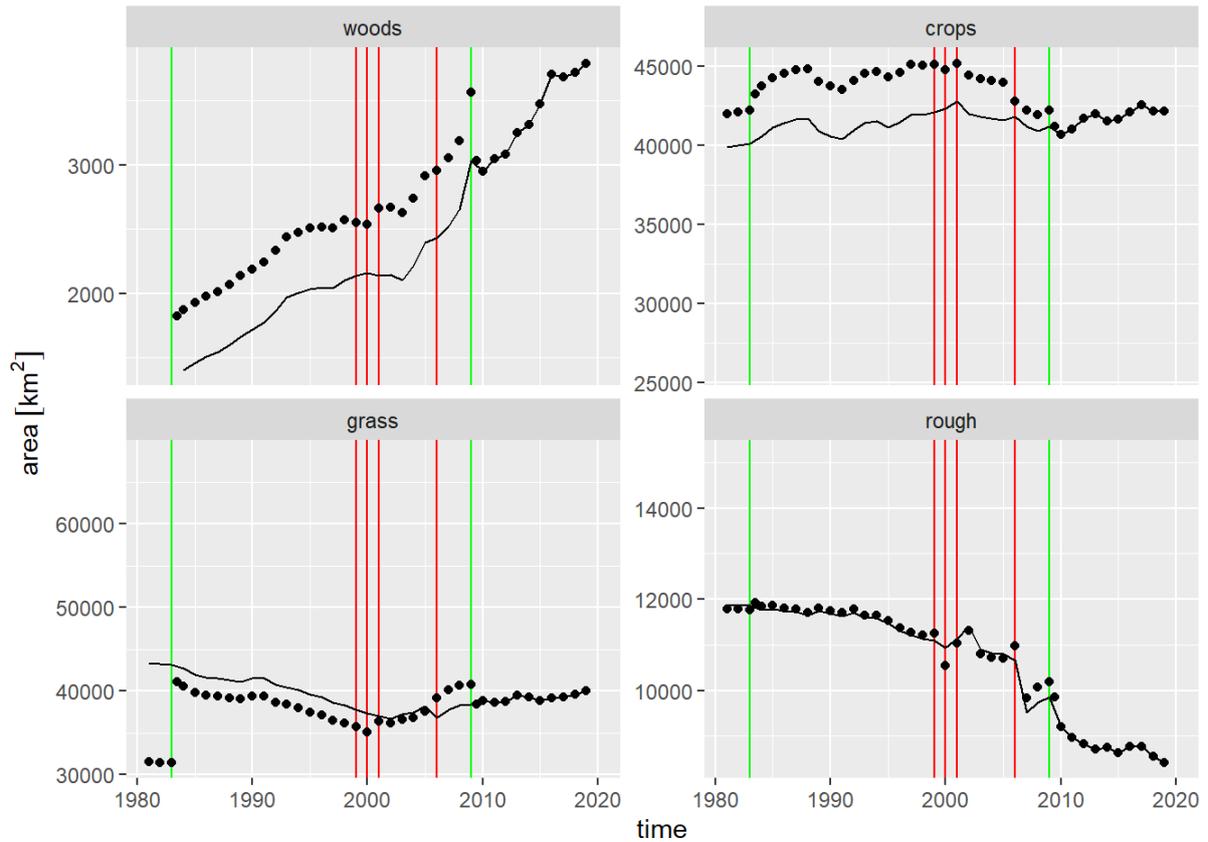


Figure 4.6: Time series of land use from the June Agricultural Census in England, 1980-2019. Points show uncorrected data; lines show corrected data; Green vertical lines show known step changes; Red vertical lines show suspected step changes which have been adjusted for.

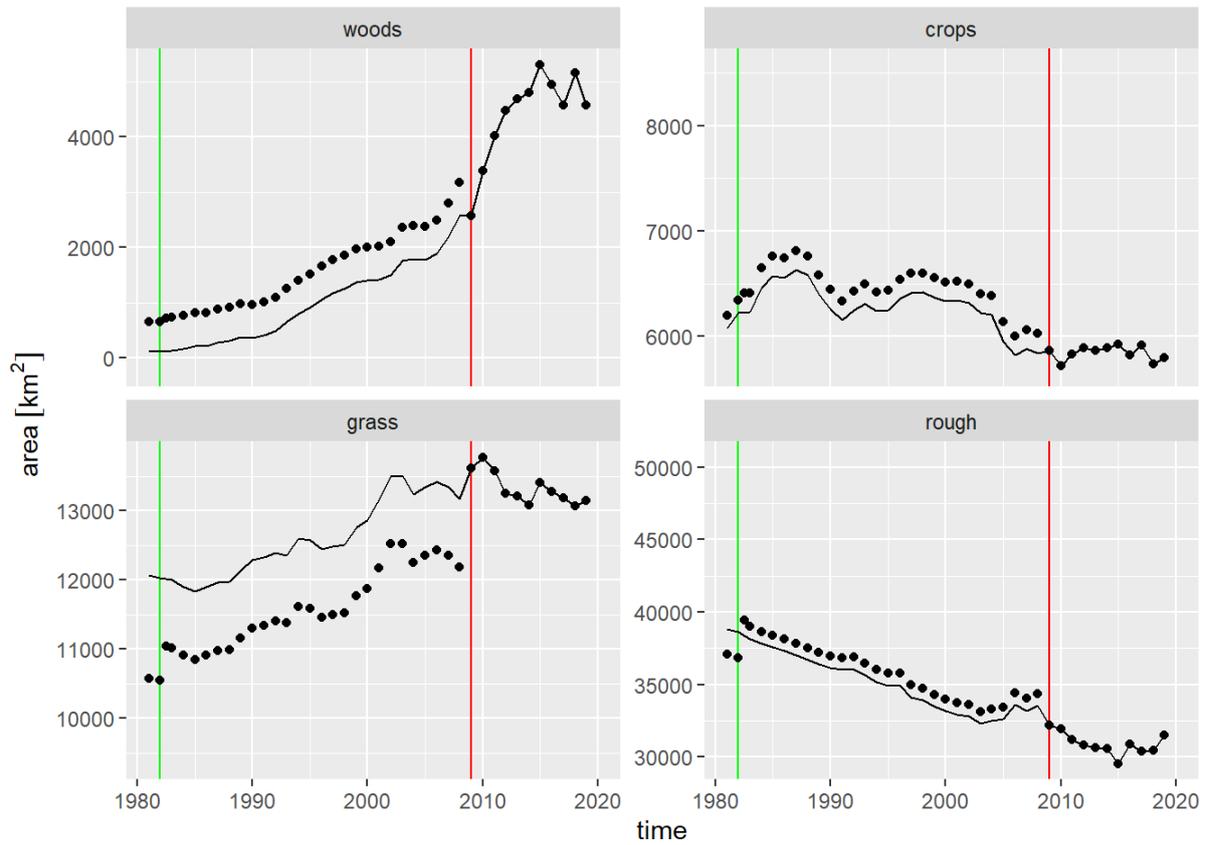


Figure 4.7: Time series of land use from the June Agricultural Census in Scotland, 1980-2019. Points show uncorrected data; lines show corrected data; Green vertical lines show known step changes; Red vertical lines show suspected step changes which have been adjusted for.

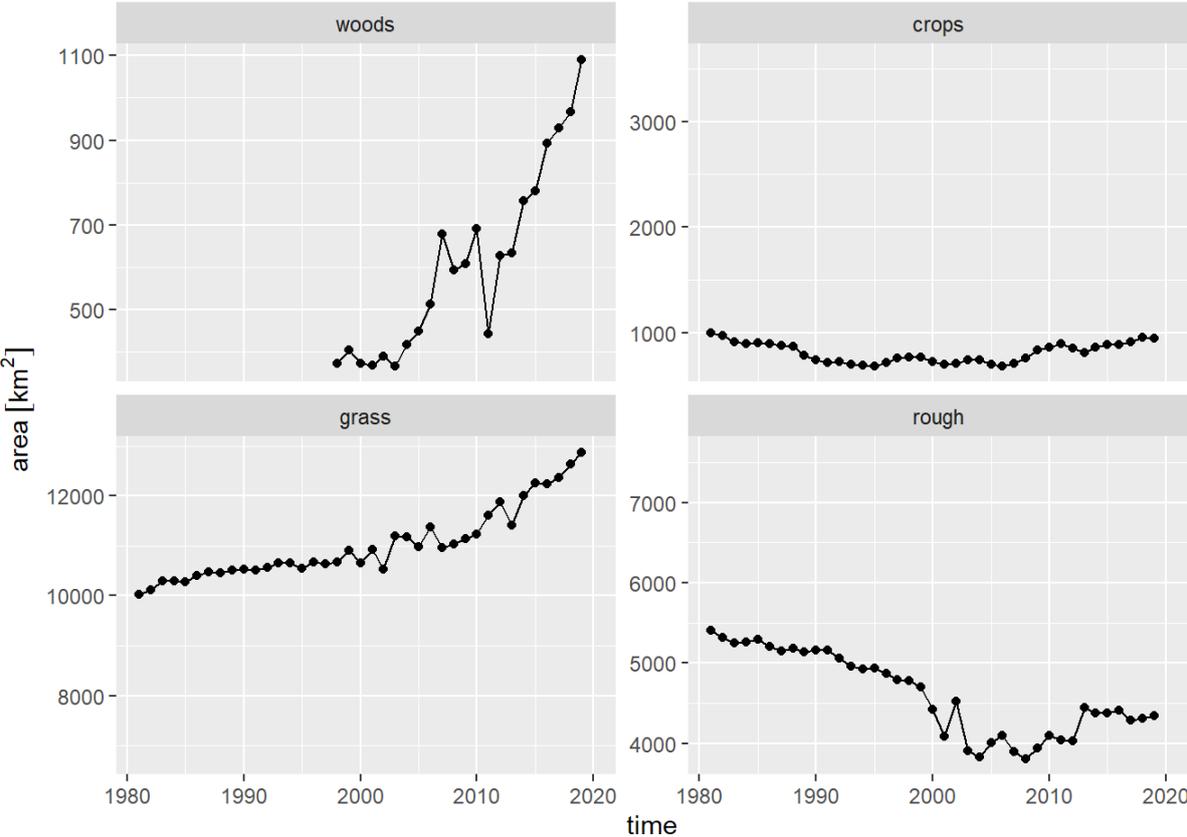


Figure 4.8: Time series of land use from the June Agricultural Census in Wales, 1980-2019.



Figure 4.9: Time series of land use from the June Agricultural Census in Northern Ireland, 1980-2019.

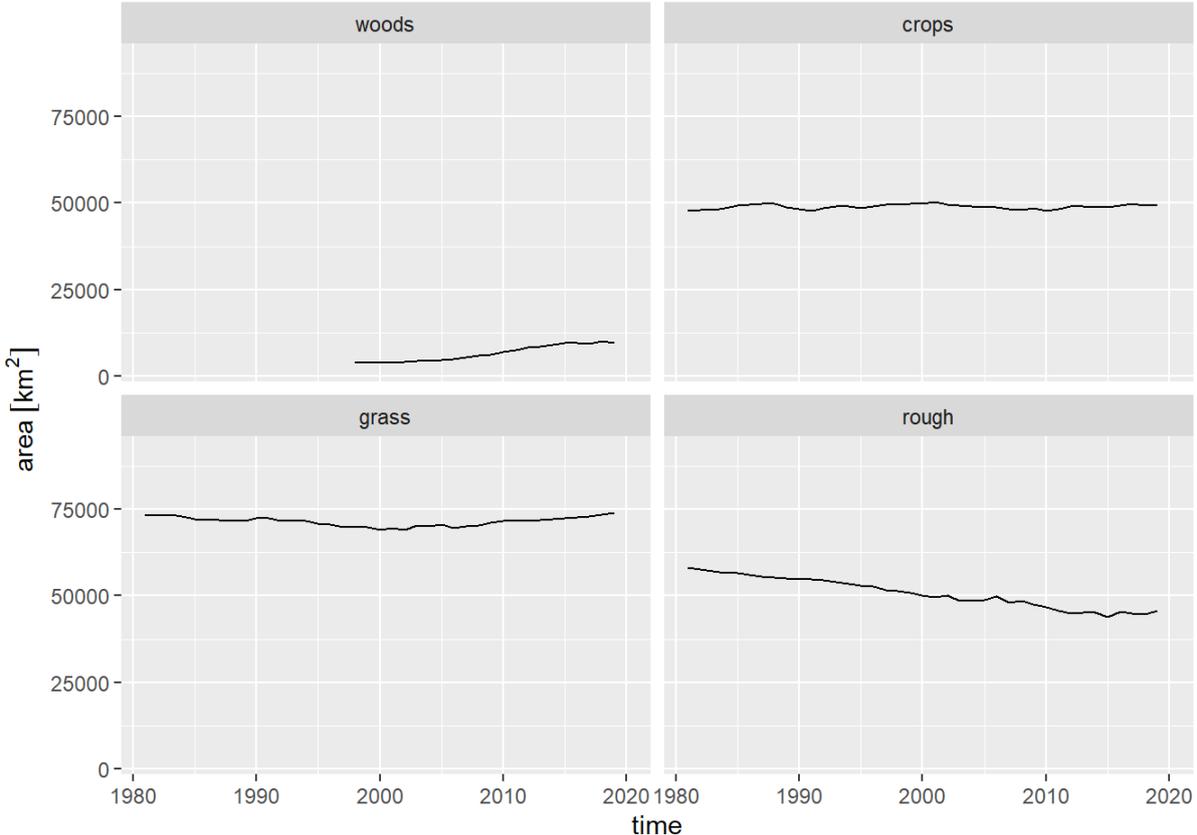


Figure 4.10: Time series of land use from the June Agricultural Census in the UK, 1980-2019.



Figure 4.11: Time series of net land-use change from the June Agricultural Census in the UK, 1980-2019.

4.5 Agricultural Holdings Data

Survey data for agricultural holdings were supplied by Scottish Government for Scotland for the years 1990 to 2018. These data, which are the finest level of detail in the agricultural census, provide information on livestock numbers, crop and grass areas (and more) at the individual farm level (ca. 51,000 unique holdings). The data are compiled from an annual questionnaire to a sample population (ca. 20 – 50%) that undergoes periodic methodological changes (such as the questions asked or the categories used). From a spatial perspective, this dataset does not provide information beyond the registered location of the holding; the actual location of crops and grass is not provided. From 2008, the location of the holding is provided using the postcode (c. 80% of holdings – the other c. 20% are parish located) and prior to this, only the parish is provided.

The processing task had several steps. We firstly had to remove information not pertaining to crops and grass and standardise the remaining data. Given the diversity of information requested by the census over the years, it is challenging to ensure that changing questions and sub-types of various crop and grass types were standardised across the time series. Secondly, the locations of the holdings were extracted from the spatial information provided, dependent on the year. Holdings were located in the centroid of the parcel to which they were associated with. Where the parcel is a parish or postcode, this can represent a large spatial extent where we have very low-precision spatial information (particularly in NW Scotland where parishes and even postcodes are large areas). The data were re-categorised to the LULUCF classification and summed by specific XY location. The net change in area per LULUCF category from one year to the next was calculated for each unique XY location. Absolute

values of gains and losses were separated and then interpolated via inverse distance weighting (Figure 4.12).

The data are the finest detail of agricultural ownership in Scotland and, when totalled, agree well with the national Agricultural Census. This is to be expected and the spatial information gained from the holding location (particularly postcodes), along with the detailed categories, is advantageous but limited. In Scotland this issue is magnified due to the large parish/postcode sizes in remote parts of the country. One of the principal problems with the data is that it is a survey, not a comprehensive census, sampling around half of farms in a normal survey year. Estimates are made for those holdings that are not surveyed based on trends. Census data are always exposed to error via misreporting from the respondent, but this is not thought to be a significant source of error.

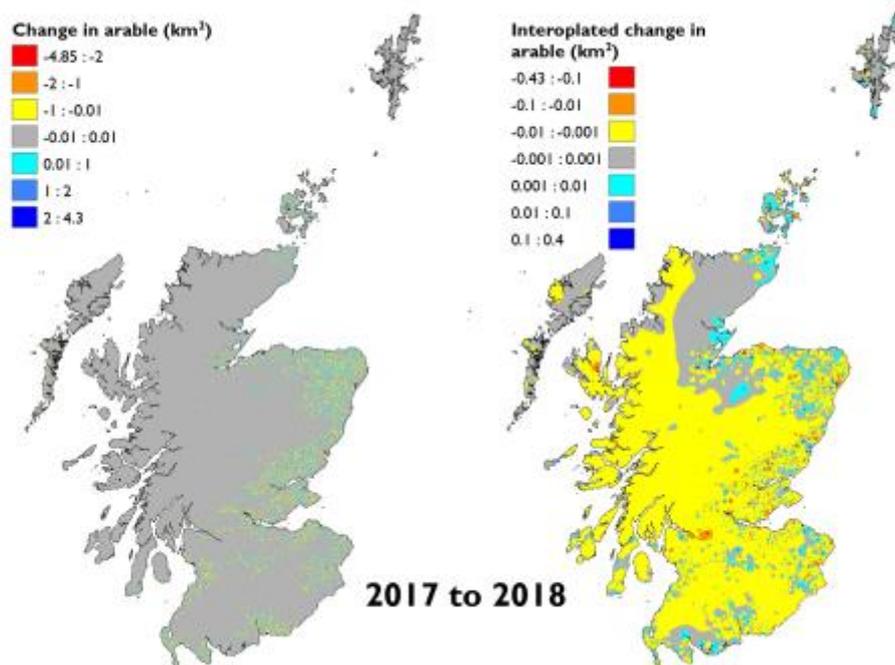


Figure 4.12: Spatial distribution of the gains and losses in arable land from Agricultural Census data in Scotland for 2017-2018. The map on the left shows the raw data for each agricultural holding. The map on the right shows an interpolation of this using inverse distance weighting.

The same analysis can be applied to the equivalent data for England, Wales and Northern Ireland, but permission to use the data for this project had still not been obtained at the time of writing. For the current analysis, the holding-level data for Scotland are used to estimate the spatial maps of the likelihood for each land use in a later section.

4.6 Forestry Commission Data

Forestry Service Northern Ireland Data

Several publicly-available data sets from the GB Forestry Commission and Forestry Service Northern Ireland were obtained. [Forestry Statistics](#) is a compilation of statistics on woodland, forestry and primary wood processing in the UK, published annually for the UK and broken down to individual countries where possible. This is the basis for the afforestation and deforestation time series that is used in the current GHGI. For this project, we did not retrieve raw data, but used the data previously provided for the current GHGI (by Paul Henshall, Forest Research).

The National Forest Inventory and the National Forest Estate Sub-compartments (SCDB) data sets were downloaded from the FC Open Data website (Forestry Commission 2020). The National Forest Inventory (NFI) is a rolling programme, starting in 2009, designed to provide accurate information about the size, distribution, composition and condition of forests and woodlands and about the changes taking place in the woodlands through time. The NFI covers any forest or woodland in Great Britain of at least 0.5 hectares in area with a minimum width of 20 m, and that have at least 20% tree canopy cover (or the potential to achieve this). The NFI includes a digital woodland map which is updated annually using more recent aerial photography, interpretation of satellite imagery and administrative records of newly planted areas covered by government grant schemes. There is also a field survey of a representative sample of randomly selected 15,000 one hectare (100 m x 100 m) plots across Britain. The first assessment of the sample sites took place between 2009-2015 with the second 'cycle' of assessments taking place between 2015-2020. About two thirds of plot locations are permanent and will be revisited by the survey team during the second and subsequent 'cycles' of ground surveying. Data for 'small woods' (0.1-0.5 hectares) are based on sample field survey and/or newly emerging high resolution remote sensing data.

The National Forest Estate Sub-compartment Database (SCDB) is the operational GIS data, which has polygonal spatial data for forest stands in GB. For the forest stands managed by FC (and its devolved bodies), there is considerable descriptive data (planting year, species, yield class), though this is largely absent from the privately-managed stands. To impute the planting year where this was missing, we sampled from the distribution of planting years produced for the current GHGI methodology by Forest Research. This means that for private forestry, the planting year is not accurate on an individual stand basis, but the overall distribution matches that of the rest of the forest estate.

The equivalent of the Sub-compartment Database from the Forestry Service Northern Ireland was provided by AFBI. These data were reprojected to OSGB, and merged with the GB data, and rasterised to 25-m resolution to produce a consistent forest map for the UK (Figures 4.13). This includes the small woodland areas within the limits of this resolution, and a time dimension in the form of the planting year. The latter is imperfect information, since we do not know if this is the date of new planting or replanting but is the best available at present.

PlantingYear



Figure 4.13: Map of forest area and date of planting across the UK from FC NFI, SCDB, and FSNI data.

4.7 IACS

The Integrated Administration and Control System (IACS) is a European-wide spatially explicit dataset at the field level that serves as a register of agricultural subsidy claims under the EU Common Agricultural Policy. IACS records field-level land use (crop type, grassland age, forest coverage), field geometry and its association to a farm holding. This has large, but not complete spatial coverage in the UK. The administration of the system differs across the DAs and has changed over time. In England, this is now run by the Rural Payments Agency (RPA) but we refer to this as IACS for back-consistency. IACS uses a very large number of classes, which has changed over time, and a table showing the correspondence between these and LULUCF classes is too large to display here but available on GitHub. Although UKCEH already holds the data for use in other projects, gaining explicit permission to use the data for this project proved very protracted. To date, explicit permission has only been granted for the English IACS(RPA) data. No other IACS data are shown here for this reason. However, the Scottish data have already been used in this framework previously and adding these back in together with Welsh and Northern Irish data when available should be straightforward.

English data were available for 2004 to 2019, as a vector product. Processing involved re-classifying the IACS classes to the six LULUCF classes. Data from 2015 onwards had a different classification, and a further sub-set of data were obtained to try to make the time series consistent. However, the issues are not all resolved at the time of writing, and the post-2015 data are not used here.



Figure 4.14: Spatial distribution of LULUCF land-use classes in the UK according to IACS.

4.8 Land Cover Map

The UKCEH Land Cover Maps (LCMs) are parcel-based thematic classifications of satellite image data covering the United Kingdom. The most recent Land Cover Maps release includes LCM 2017, LCM 2018, LCM 2019 (UKCEH [2020b](#)). It also includes a revised LCM 1990 covering the whole of the UK to support comparisons with LCM 2015 and the newer products. Using the revised LCM 1990, quarter-century change datasets have been created by comparison with LCM 2015. We note that these products are based on land cover, rather than land use. Land cover relates to the physical properties of the surface only, not how it is used. For example, clear-felled forest may not appear as forest cover to a satellite sensor, but its land use may remain so if it is to be replanted. Land use is therefore less directly detectable from satellites.

Maps were available for 1990, 2015, 2017, 2018 and 2019, as a 25-m raster product. Processing involved the reprojecting and merging of data for GB (in OSGB36 coordinates) and NI (in Irish Grid coordinates). We re-classified the 21-class UKCEH Land Cover Map data to the six LULUCF aggregated classes. The table below shows the correspondence between LCM classes and LULUCF classes.

| LCM_ID | LCM_name | LULUCF_ID | LULUCF_name |
|--------|----------------------|-----------|-------------|
| 1 | Broadleaved Woodland | 1 | forest |
| 2 | Coniferous Woodland | 1 | forest |

| LCM_ID | LCM_name | LULUCF_ID | LULUCF_name |
|--------|-------------------------|-----------|-------------|
| 3 | Arable and Horticulture | 2 | crop |
| 4 | Improved Grassland | 3 | grass |
| 5 | Neutral Grassland | 4 | rough |
| 6 | Calcareous Grassland | 4 | rough |
| 7 | Acid grassland | 4 | rough |
| 8 | Fen, Marsh and Swamp | 4 | rough |
| 9 | Heather | 4 | rough |
| 10 | Heather grassland | 4 | rough |
| 11 | Bog | 4 | rough |
| 12 | Inland Rock | 4 | rough |
| 13 | Saltwater | 0 | NA |
| 14 | Freshwater | 0 | NA |
| 15 | Supra-littoral Rock | 6 | other |
| 16 | Supra-littoral Sediment | 6 | other |
| 17 | Littoral Rock | 6 | other |
| 18 | Littoral sediment | 6 | other |
| 19 | Saltmarsh | 4 | rough |
| 20 | Urban | 5 | urban |
| 21 | Suburban | 5 | urban |

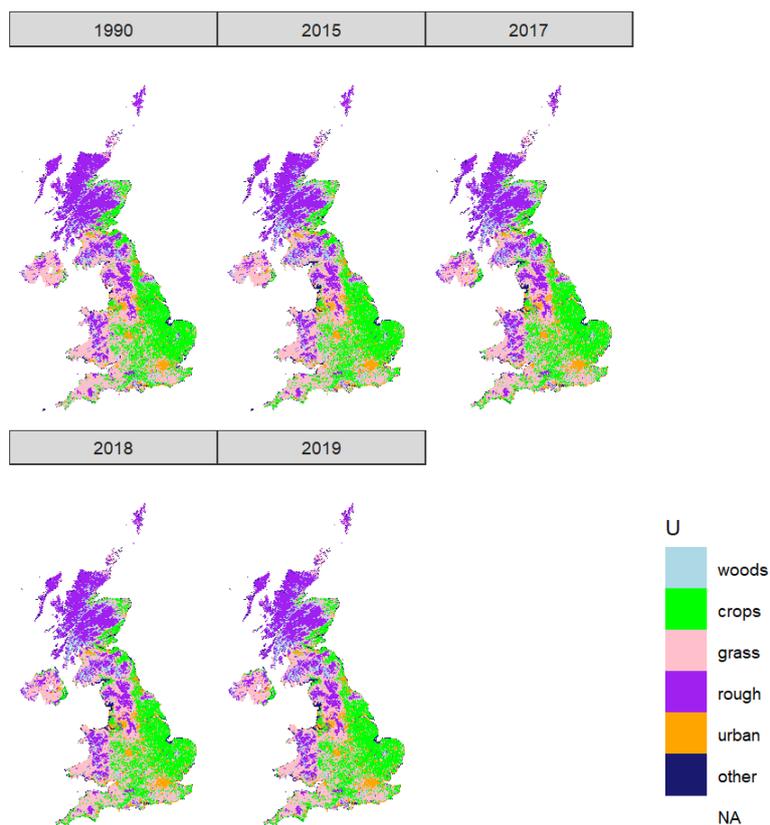


Figure 4.15: Spatial distribution of LULUCF land-use classes in the UK according to LCM.

4.9 Land Cover Plus Crops

UKCEH Land Cover® plus: Crops (LCC) (UKCEH 2016) is based on the Land Cover Map parcel framework. Every parcel which is larger than 2 ha and categorised as agricultural land is coded with crop type information from satellite data - Copernicus Sentinel-1 C-band SAR (Synthetic Aperture Radar) and, from 2016 onwards, Sentinel-2 optical data. Data are available for 2015 (partial GB coverage only), 2016, 2017, 2018 and 2019 cropping years.

Maps were available for 2015 (partial GB coverage only), 2016, 2017, 2018 and 2019 cropping years, as vector product. Processing involved re-classifying the 11-class UKCEH Land Cover Map data to the six LULUCF classes. The table below shows the correspondence between LC: Crop classes and LULUCF classes.

| LCCROP.ID | LCCROP.DESC | LULUCF_class_ID | LULUCF_class_name |
|-----------|-------------|-----------------|-------------------|
| 1 | Beet | 2 | crop |
| 2 | Field Beans | 2 | crop |
| 3 | Grass | 3 | grass |
| 4 | Maize | 2 | crop |

| LCCROP.ID | LCCROP.DESC | LULUCF_class_ID | LULUCF_class_name |
|-----------|---------------|-----------------|-------------------|
| 5 | Oilseed Rape | 2 | crop |
| 6 | Other crop | 2 | crop |
| 7 | Peas | 2 | crop |
| 8 | Potatoes | 2 | crop |
| 9 | Spring Barley | 2 | crop |
| 10 | Spring Wheat | 2 | crop |
| 11 | Winter Barley | 2 | crop |
| 12 | Winter Oats | 2 | crop |
| 13 | Winter Wheat | 2 | crop |

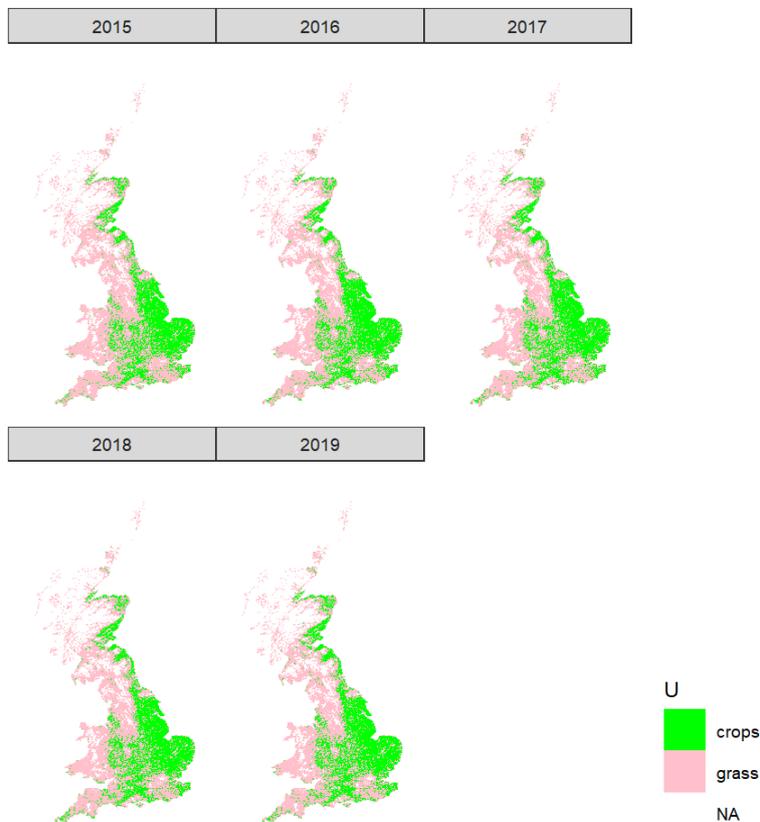


Figure 4.16: Spatial distribution of crop and grass land-use classes in the UK according to LCC.

4.10 CORINE Land Use

4.10.1 Background

The CORINE Land Cover product (“CORINE Land Cover — Copernicus Land Monitoring Service”) consists of an inventory of land cover in 44 classes. CORINE is produced within the framework of the Initial Operations of the Copernicus programme (the European Earth monitoring programme previously known as GMES) on land monitoring. It attempts to provide consistent information on land cover and land cover changes across Europe. CORINE is produced by semi-automatic interpretation of high-resolution satellite imagery, following a standard methodology and nomenclature with the following base parameters: - 44 classes in the hierarchical three level Corine nomenclature; - Minimum mapping unit (MMU) for status layers is 25 hectares; - Minimum width of linear elements is 100 metres.

Maps were available for 2000, 2006, 2012, and 2018 as a 100-m raster product (version updated 2020). Processing involved the reprojecting from Lambert Equal Area projection to OSGB36 coordinates and re-classifying the 44-classes to the six LULUCF classes. The table below shows the correspondence between CORINE classes and LULUCF classes.

| CLC_CODE | Corine_name | LULUCF_ID | LULUCF_name |
|----------|--|-----------|-------------|
| 1 | Continuous urban fabric | 5 | urban |
| 2 | Discontinuous urban fabric | 5 | urban |
| 3 | Industrial or commercial units | 5 | urban |
| 4 | Road and rail networks and associated land | 5 | urban |
| 5 | Port areas | 5 | urban |
| 6 | Airports | 5 | urban |
| 7 | Mineral extraction sites | 5 | urban |
| 8 | Dump sites | 5 | urban |
| 9 | Construction sites | 5 | urban |
| 10 | Green urban areas | 3 | grass |
| 11 | Sport and leisure facilities | 3 | grass |
| 12 | Non-irrigated arable land | 2 | crop |

| CLC_CODE | Corine_name | LULUCF_ID | LULUCF_name |
|----------|---|-----------|-------------|
| 13 | Permanently irrigated land | 2 | crop |
| 14 | Rice fields | 2 | crop |
| 15 | Vineyards | 2 | crop |
| 16 | Fruit trees and berry plantations | 2 | crop |
| 17 | Olive groves | 2 | crop |
| 18 | Pastures | 3 | grass |
| 19 | Annual crops associated with permanent crops | 2 | crop |
| 20 | Complex cultivation patterns | 2 | crop |
| 21 | Land principally occupied by agriculture with significant areas of natural vegetation | 3 | grass |
| 22 | Agro-forestry areas | 1 | woods |
| 23 | Broad-leaved woods | 1 | woods |
| 24 | Coniferous woods | 1 | woods |
| 25 | Mixed woods | 1 | woods |
| 26 | Natural grasslands | 4 | rough |
| 27 | Moors and heathland | 4 | rough |
| 28 | Sclerophyllous vegetation | 4 | rough |
| 29 | Transitional woodland-shrub | 4 | rough |
| 30 | Beaches dunes sands | 6 | other |
| 31 | Bare rocks | 6 | other |

| CLC_CODE | Corine_name | LULUCF_ID | LULUCF_name |
|----------|-----------------------------|-----------|-------------|
| 32 | Sparsely vegetated areas | 6 | other |
| 33 | Burnt areas | 6 | other |
| 34 | Glaciers and perpetual snow | 6 | other |
| 35 | Inland marshes | 4 | rough |
| 36 | Peat bogs | 4 | rough |
| 37 | Salt marshes | 4 | rough |
| 38 | Salines | 6 | other |
| 39 | Intertidal flats | 0 | NA |
| 40 | Water courses | 0 | NA |
| 41 | Water bodies | 0 | NA |
| 42 | Coastal lagoons | 0 | NA |
| 43 | Estuaries | 0 | NA |
| 44 | Sea and ocean | 0 | NA |
| 45 | NODATA | 0 | NA |

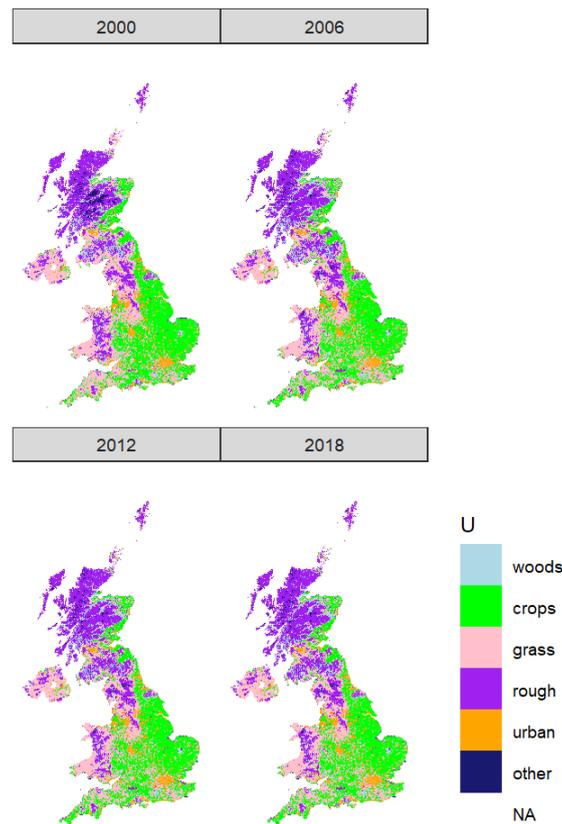


Figure 4.17: Spatial distribution of LULUCF land-use classes in the UK according to CORINE.

4.11 CROME

The Crop Map of England (CROME) dataset is a tessellated (hexagonal) vector product derived via remote sensing techniques. It concentrates on classifying areas of crops in England using the Sentinel satellite constellation, a random forest classification and ground-truth data. Ground-truth data were recorded by RPA field inspectors during the spring/summer of each year. Each hexagonal cell was ~0.4 hectares and represents one of 80 classes (55 cereal crops, 15 leguminous crops, 1 energy crop, 3 grassland types, 3 tree classes, water, non-agricultural land and mixed).

Unfortunately, the data are only made available on a county-by-county basis for each year. Requests to the RPA for a bulk download have not yet provided the full dataset. Retrieving and processing all 46 files for each year is feasible but time-consuming, and not deemed efficient use of time when we should be able to get the complete data in bulk, albeit at a later date. A further issue is that the data are trained and evaluated via RPA field inspections which are the same inspections used to validate the IACS data. Since these data are already used in the project, it is not clear whether we would gain any new information from this data set. We have therefore not used CROME data in the current project, but anticipate it being available in the near-future for examination.

4.12 Ordnance Survey Land Use Change Statistics

The Land Use Change Statistics (LUCS) dataset, created by the Ordnance Survey (OS), is a local authority-level dataset which concentrates on changes in development activities, usually associated with housing and development of land. Polygon data is drawn from the OS database (e.g. MasterMap)

that has undergone change and also changing address data with regards residential development. The data produced is derived from inferred intelligence that has not been directly observed by field or other survey method and may be in error with regards real world use. However, this method does allow for large scale analysis at a very high level of detail and spatial resolution. Data parcels that specifically underwent change for 2017 to 2018 were provided by the OS as vector data (c. 1.1 million parcels). The parcels are classified into 28 categories – predominantly building and development based, with some agricultural land and forestry – which were reclassified to LULUCF categories. The table below shows the correspondence between OS classes and LULUCF classes.

| LUCS_ID | LUCS_name | LUCS_DESCRIPTION | LULUCF_ID | LULUCF_name |
|---------|-----------|---|-----------|-------------|
| 1 | ~B | Unidentified building | 5 | urban |
| 2 | ~M | Unidentified general manmade surface | 5 | urban |
| 3 | ~S | Unidentified structure | 5 | urban |
| 4 | ~U | Unknown surface type with no classification | 6 | other |
| 5 | A | Agricultural Land | 2 | crop |
| 6 | B | Agricultural Buildings | 5 | urban |
| 7 | C | Community Buildings | 5 | urban |
| 8 | D | Defence | 5 | urban |
| 9 | F | Forestry/Woodland | 1 | forest |
| 10 | G | Rough grassland | 4 | rough |
| 11 | H | Highways and roads | 5 | urban |
| 12 | I | Industry | 5 | urban |
| 13 | J | Offices | 5 | urban |
| 14 | K | Retail | 5 | urban |
| 15 | L | Leisure (indoor) | 5 | urban |
| 16 | M | Minerals and Mining | 5 | urban |

| LUCS_ID | LUCS_name | LUCS_DESCRIPTION | LULUCF_ID | LULUCF_name |
|---------|-----------|-----------------------------|-----------|-------------|
| 17 | N | Natural Land | 3 | grass |
| 18 | O | Outdoor Recreation | 3 | grass |
| 19 | Q | Communal Accommodation | 5 | urban |
| 20 | R | Residential | 5 | urban |
| 21 | RG | Residential Gardens | 3 | grass |
| 22 | S | Storage and Warehousing | 5 | urban |
| 23 | T | Transport | 5 | urban |
| 24 | U | Utilities | 5 | urban |
| 25 | V | Vacant Land | 6 | other |
| 26 | W | Water | 6 | other |
| 27 | X | Undeveloped Land | 3 | grass |
| 28 | Y | Landfill and Waste Disposal | 5 | urban |

The land-use change matrix derived from these data is shown below. The values in most elements of the matrix are extremely small, although the transitions to urban appear reasonable. Without further investigation, it is not clear which elements we can validly use, so these data are not used further in the current analysis.

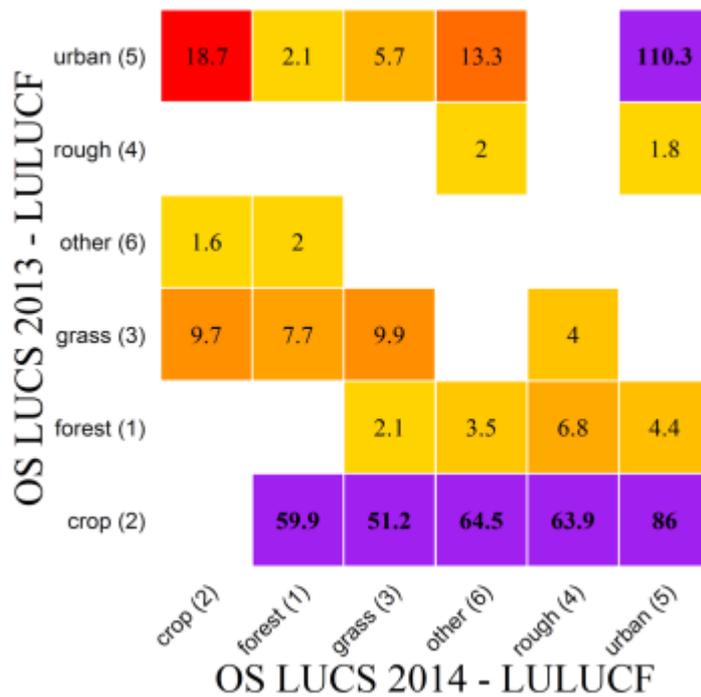


Figure 4.18: Land-use change matrix derived from LUCS data for 2017-2018. Units are km².

4.13 Land Use and Cover Area Frame Survey

The Land Use and Cover Area Frame Survey (LUCAS) is point database of direct observations throughout the EU every three years from 2006 and classifies observations into a Land Use (LU) and Land Cover (LC). Across the EU, around 350,000 points were surveyed in 2018, either by person (~70 %) or by interpretation of photos (~30 %). The points are selected via stratification of roughly 1 million locations in Europe that lie on a 2 x 2-km grid. The data also includes soil data at roughly 10 % of surveyed locations. The classification is derived for not only the point location but also for the land in the surrounding 20 metres. The classification of the data is hierarchical; this incorporates differing levels of detail starting at a set of broad classes. For LC, there are 8 main categories that are further divided to 83 sub-classes. For LU, there are 14 main categories that are further divided to 33 sub-classes. In total, there are roughly 17,000 points across the UK in the LUCAS dataset. The classification system is comparable to that of FAO and CORINE.

The data have been obtained but need some further time to process and analyse, as their use requires an understanding of the data structure and sampling design for proper interpretation. Despite the large number of points, a very small area appears to be represented (only 20 metres around each point), and which of these is actually from a ground survey remains unclear, so quite how this can be interpreted is still open to discussion. We anticipate the data will be of some use in the future, but are not used in the current analysis.

5 Comparison of Data Sources

5.1 Results

Having assembled all the available data sources, we can now plot them on a common series of axes to compare their absolute magnitudes and relative trends in time.

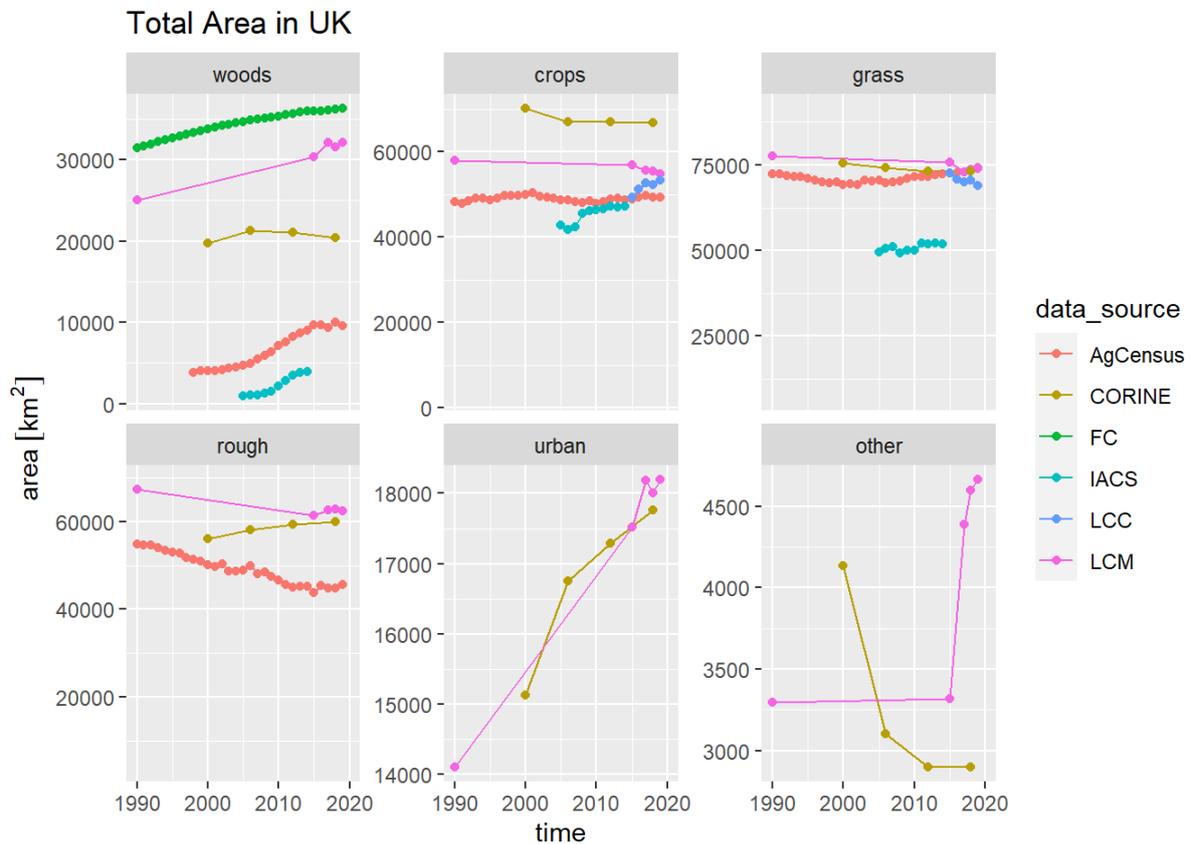


Figure 5.1: Time series of land-use areas **A** observed by different data sources. Note that IACS data only cover England, so are included for comparison of trends only.

Figure 5.1 shows the total areas of each land-use class estimated by the available data sources. For the total area of woodland, Agricultural Census and IACS only record the woodland area occurring on agricultural land, so these should not be expected to match the actual totals; they are only included to compare the rate of increase. However, between the remaining sources, there is a wide range in estimates, varying by 16000 km², with FC > LCM > CORINE. CORINE shows an apparent decline in woodland area, which has to be considered suspect. LCM and FC show similar trends, but LCM estimates the woodland area to be ~ 6000 km² lower. Because the FC data is largely based on ground-based survey and aerial photography rather than satellite, we assume this to be more reliable.

The Agricultural Census shows the total crop area to be relatively stable at ~50000 km². The latter part of the IACS data appear in close agreement with this, but they account only for England. Adding in the crop area of Scotland, Wales and NI suggests that the IACS estimates exceed the Agricultural Census by around 10000 km². The IACS crop data show a step change after the first three year, which is probably an artefact of changes in the methodology of IACS reporting. The LCC data start in close agreement with Agricultural Census, but show a strongly rising trend of ~5000 km² not seen in other

data sources. Both CORINE and LCM show higher crop areas, with a declining trend not seen in the ground-based data.

For grassland, the outlying line of IACS data can be discounted as it covers only England, and is included only for comparison of trends. Otherwise, estimates seem in closer agreement in absolute terms, although trends are not very consistent across data sources. The total area of rough grazing and semi-natural land shows a similar decline in both Agricultural Census and LCM, although their initial starting points differ by 10000 km². CORINE shows the reverse, which appears implausible. For urban and built-up land however, LCM and CORINE show very close agreement on the absolute area and trend. Estimates of the total area of other land uses are only provided by LCM and CORINE, and neither appears plausible as genuine land-use change. This other land is mostly coastal zone, and it is more likely that this apparent change is due to differences in satellite imagery and the algorithms used to classify them.

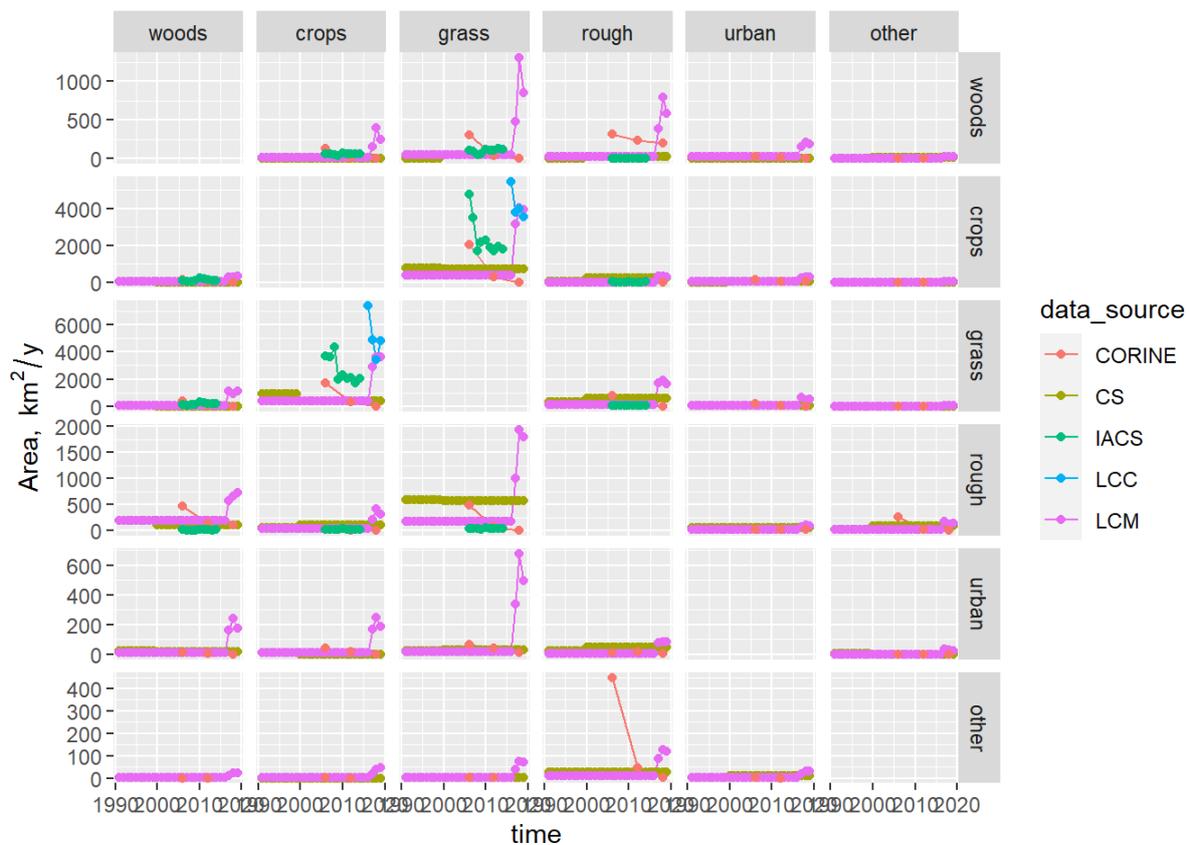


Figure 5.2: Time series of area changing from each land use to every other land use (the matrix **B**) observed by different data sources. LCM and CS values between surveys were interpolated values as constant annual rates.

Figure 5.2 shows the **B** matrix over time. The y axis scales are somewhat skewed by a number of outlying high values. The plots on the top row, representing deforestation, are dominated by the high values from post-2015 LCM. Given the abrupt change with the previous LCM estimate, these would seem implausible. For the conversions from crop to grass and vice versa, the LCC and post-2015 LCM data give very high values (4000 to 6000 km²) compared with CS and pre-2015 LCM (< 500 km²). IACS also shows similarly high values when we account for the fact that this only cover England and also with a step-change in the time series that appears implausible. Similar patterns are seen in other panels e.g. other to rough, where a single CORINE value dominates. Because of these order-of-magnitude

differences, it is hard to discern subtler features here. As an alternative presentation, we can show the matrix numerically, with the colour scale indicating the magnitude of the area (Figure 5.3).

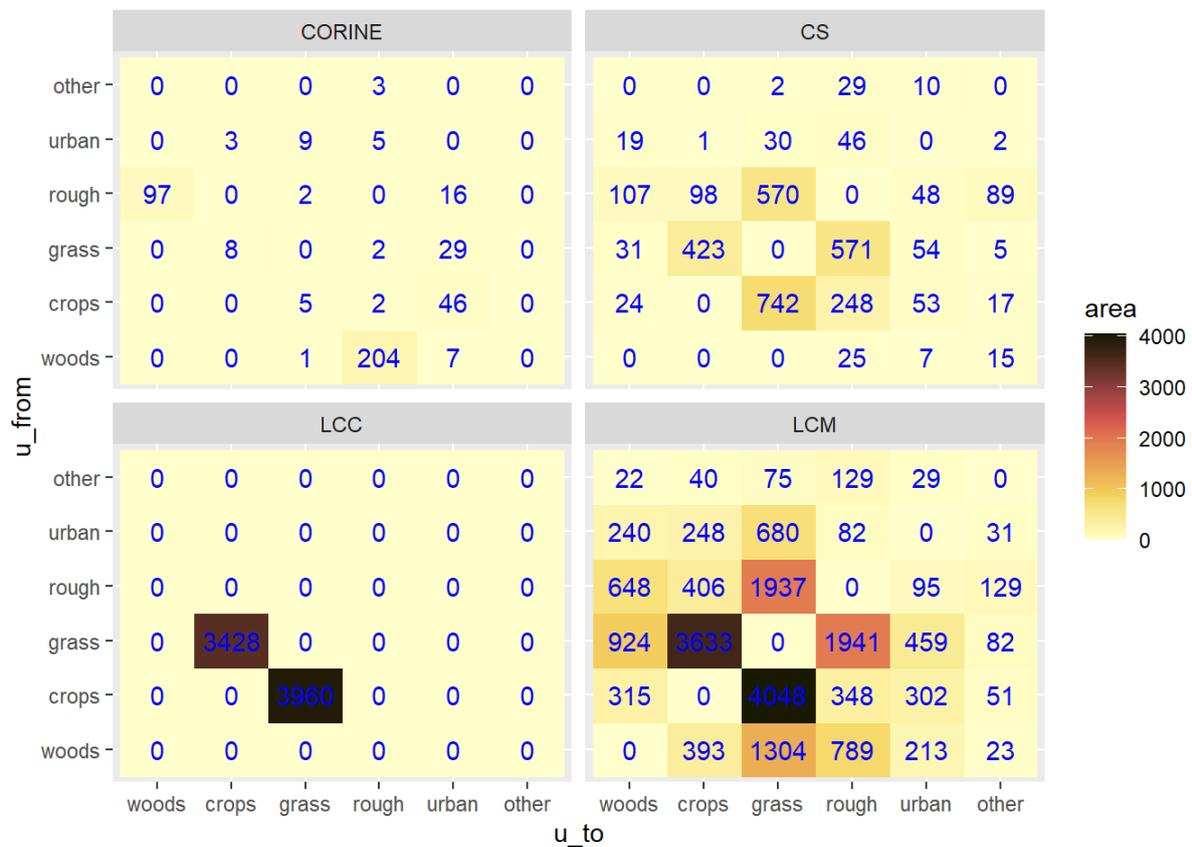


Figure 5.3: Area changing from each land use to every other land use (the matrix **B**) observed by different data sources between 2017 and 2018.

Figure 5.3 indicates that LCC and LCM give B values that are approximately an order of magnitude higher than CS and CORINE in 2017-2018. The former give crop-grass transitions close to 4000 km²; CS averages 583 km². For grass-rough transitions, LCM gives values close to 2000 km²; CS estimates 570 km². To examine the general patterns over time more clearly, we can plot the gross changes of each land-use class i.e. the sums of the rows and columns in Figures 5.4 and 5.5.

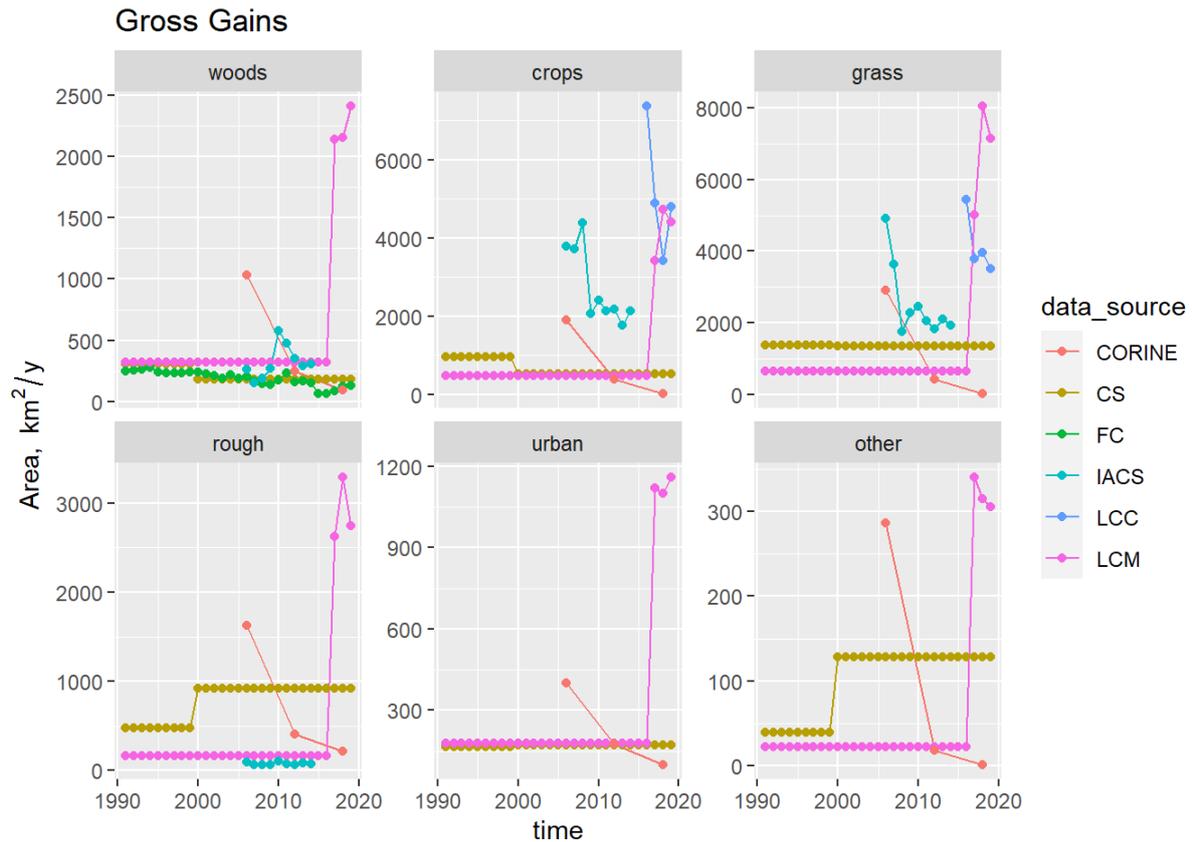


Figure 5.4: Time series of gross gain of area G to each land use observed by different data sources. LCM and CS values between surveys were interpolated values as constant annual rates.

Figure 5.4 shows the gross gain in area of each land use. For woodlands, we would place highest confidence in the FC data (green symbols), which shows the gross gain to have declined from 250 km² to near zero in 2016. The 1990-2015 LCM data gives a slightly higher mean compared to this, and the CS data are not truly independent of the FC data as they share some data in common. The post-2015 LCM data show much higher rates, over 2000 km², which are not believable. CORINE also shows high afforestation in 2006, but sharply declining; again this is not believable, assuming the FC data are reasonably reliable.

A similar lack of agreement is generally apparent in the data for the other five land uses. The 1990-2015 LCM and CS data show reasonable agreement. The post-2015 LCM, LCC and IACS data show much higher gross gains. In the case of crops, if we factor in that the IACS data only account for England, then these three data sets are in some agreement that the gross gain in crop is in the range 3000-7000 km². However, their patterns over time make us suspicious of their validity. The large year-to-year variability seems unlikely, and the patterns in LCC and LCM are not consistent.



Figure 5.5: Time series of gross loss of area **L** from each land use observed by different data sources. LCM and CS values between surveys were interpolated values as constant annual rates.

Examining the gross losses shows a similar picture (Figure 5.5). According to FC data (green symbols), deforestation rates have remained very low, less than 50 km². The 1990-2015 LCM data gives a close but slightly higher mean compared to this. The post-2015 LCM data show much higher rates, over 2000 km², which again seem implausible. CORINE also shows high deforestation in 2006, which then decreases. Across all land uses there is essentially the same overall pattern as in the gross gains, with the 1990-2015 LCM and CS data showing small losses, whilst the post-2015 LCM, LCC and IACS data show much larger losses.

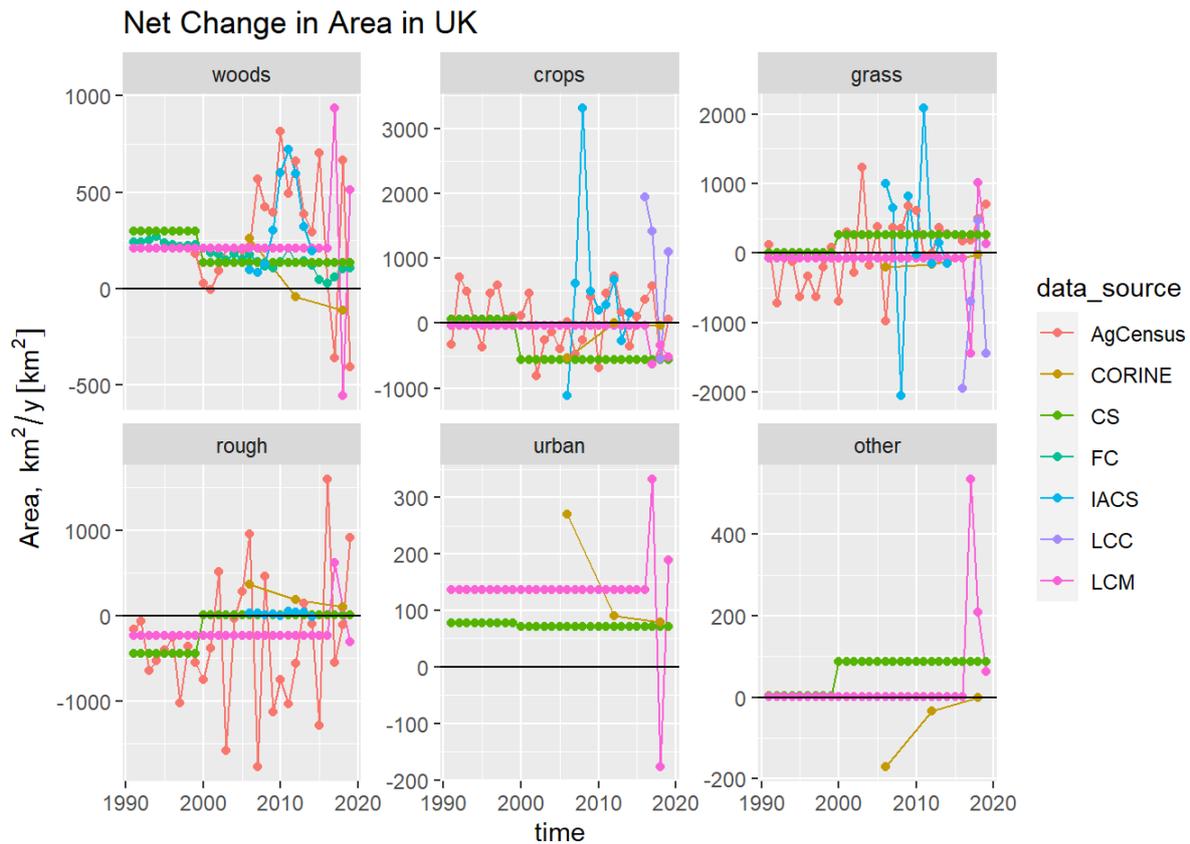


Figure 5.6: Time series of net change in area of each land use ΔA observed by different data sources.

The net effect of these gains and losses is shown in Figure 5.6. The same broad pattern is apparent. The 1990-2015 LCM and CS data give lower estimates, with little temporal resolution because of the frequency of the surveys. The Agricultural Census shows annual variability, oscillating from positive to negative, but at a lower level. LCC, post-2015 LCM data, and IACS data show higher net rates of change, with rates of net change 3-4 times higher and a more noisy pattern.

5.2 Discussion

From the above results, it appears that more work remains to be done on the EO-based data sets before they are suitable for use as direct measures of land-use change. Whilst the larger gross changes that they observe may be correct, their patterns over time make them appear less plausible. Particularly, the temporal patterns are not consistent (i.e. they do not increase and decrease at the same times), and the large year-to-year variability seems unlikely, given the way land management works. It appears that their sensitivity for detecting change alters over time. In the CORINE time series, both the gains and losses are initially large, and both decline at a similar rate over time. not useful. With LCM, there is a clear difference between the magnitude of gross change detected in recent years compared with pre-2015. The IACS data are also not consistent in time, and the magnitude and variability of changes appears questionable.

None of the data sources here represent absolute truth. We cannot rule out the possibility that the much higher gross change rates are indeed correct, and that the Agricultural Census and CS underestimate these. However, we should consider other factors as well as the patterns in the data themselves. First is the parsimony principle: if small gross changes can reproduce the observed net change, this is fundamentally more probable than the occurrence of large gross changes, all else being

equal. In the Bayesian methodology, we can specify this as a prior probability distribution, and the extent to whether “all else” really is “equal” is quantified by the observed data. We use this approach later. Secondly, the errors produced by EO-based and ground-based estimates are quite different. If we estimate land-use change by differencing maps, errors in the maps appear as spurious land-use change. So this method will necessarily over-estimate land-use change, to an extent that can be predicted from the predictive accuracy of the individual maps. The error variance in the land-use change estimate will be the sum of the variance in the two maps used. Future work may provide a quantification of this. The systematic errors present in ground-based estimates are less clear-cut, as it depends on the methodology used. However, because it requires sampling effort to detect change, it is commonly assumed that it is easier to under-estimate change. Again, future work may clarify this.

Given the uncertainties in the EO data that may be tackled at a later stage, we focus on using ground-based observations for the data assimilation in WP-A, as originally proposed. We elect to use CS, FC and Agricultural Census data to estimate the B matrix. We use the spatial pattern in the other data sources (LCM, LCC, IACS, CORINE) to estimate the likelihood of where change occurs, but not in the estimation of the B matrix itself. This is a conservative approach, as the emphasis is on the same input data sets as used in the current inventory method but constrains them with the net change measured by the national-scale Agricultural Census data, and the spatial pattern contained in LCM, LCC, IACS, and CORINE data. In this way we can try to minimise confounding changes to the methodology for estimating land-use change, and the data sources used in its estimation.

The issues with the time trends and the plausibility of much higher gross rates of change in these data sets are issues which may be resolved in future work, which will focus on EO data. There are various approaches to further examining this. For example, if the IACS data are reliable, they should match the spatial pattern in the Agricultural Holdings data. If the two data sets are at odds, we cannot have great confidence in both. The Agricultural Census contains some data which would act as a constraint on our estimates of the gross changes to/from grass. The area of grassland less than five years old is specified. Assuming a uniform distribution, one fifth of this would be converted each year on average, which would give a ballpark estimate of the gross gain and loss. The complications arise in the age distribution of grassland (what area is > 5 years old?) and whether the distribution is really uniform. Lastly, some other data sets exist for comparison, but have not become available within the time frame of this project (further IACS data, CROME, and LUCAS).

6 Estimating B by Least-Squares Optimisation

6.1 Introduction

As a first step in the data assimilation (DA) procedure, we show that it can be done in a relatively simple way. In this section, we use a least-squares (LS) optimisation algorithm to estimate the B matrix parameters. This is done for a number of reasons.

Firstly, for the purposes of elucidation, we show that the DA procedure can be separated from the Bayesian aspect.

Secondly, there is also a practical reason: when estimating the B parameters by MCMC in the following section, it is helpful to have an idea of sensible starting values. MCMC chains can be initialised with entirely random values but starting at least one chain in a region of high posterior probability speeds up the process considerably.

Lastly, by doing the DA in a non-Bayesian way, we can illustrate the advantages of the Bayesian approach.

6.2 Methods

Firstly, we define a function to calculate the root-mean-square error (RMSE) for a given B matrix. RMSE is commonly used as a measure of the differences between observed values and those predicted by a model. Here, we have observations of land-use change in the form of ΔA , G , L and B observed by several different data sources. For comparison, we have the given B matrix and the resulting predictions of ΔA , G and L that this produces via the equations in Methods. The difference between the observations and predictions gives the residual; for each term ΔA , G , L , B this is squared, averaged, and the square root taken to give the RMSE. Because all terms have the same dimensions ($\text{km}^2 / \text{year}$), we can simply add them. There are various permutations on the exact way in which the RMSE could be calculated here, but this is sufficient for our purposes. There are numerous optimisation algorithms, which will vary parameters iteratively to minimise a function. Here we use the algorithm of Byrd et al. (1995). For each year between 1990 and 2019, we run this algorithm to find the B matrix parameter set which has the smallest RMSE value, i.e. parameter set which best fits to the observed data.

6.3 Results

The figures below (6.1, 6.2, 6.3, and 6.4) show the same observed data as in the previous section, with the addition of the estimates from least-squares optimisation (black solid line). The least-squares fit is largely as expected, fitting through the centre of the observations, and following the same trends. Each of the terms ΔA , G , L and B are given equal weighting, and within these, all of the observations are given equal weighting.

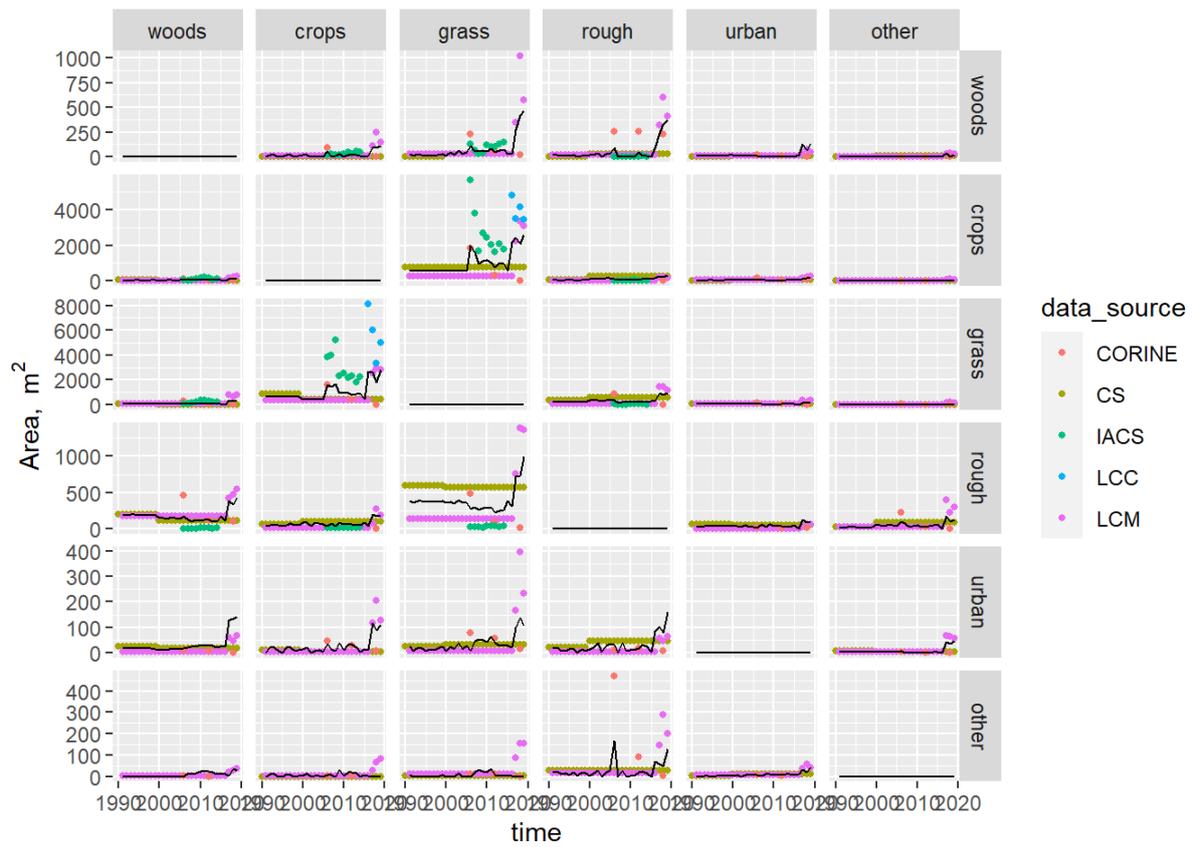


Figure 6.1: Time series of area changing from each land use to every other land use (the matrix **B**) observed by different data sources. Coloured symbols show the observations; the black line shows the best-fit values estimated by least-square optimisation. LCM and CS values between surveys were interpolated values as constant annual rates.

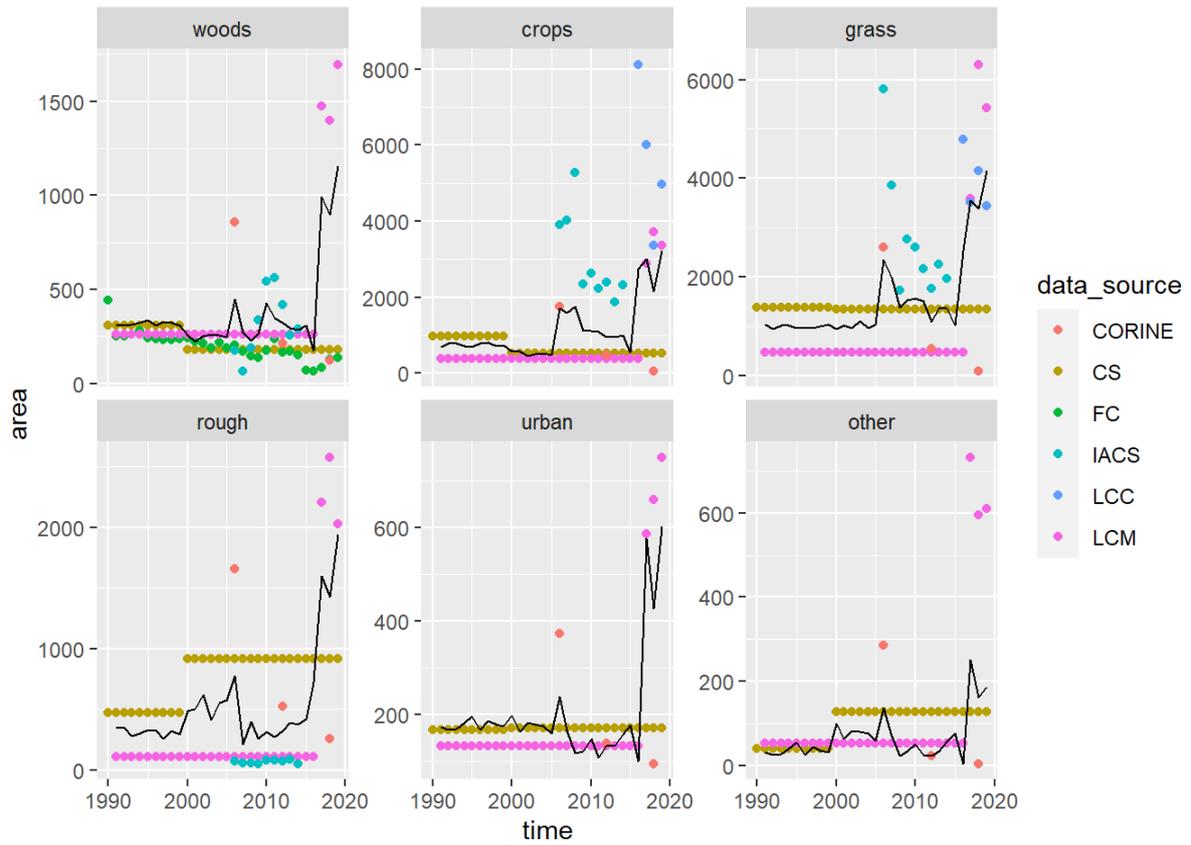


Figure 6.2: Time series of gross gain in area G of each land use observed by different data sources. Coloured symbols show the observations; the black line shows the best-fit values estimated by least-square optimisation. LCM and CS values between surveys were interpolated values as constant annual rates.

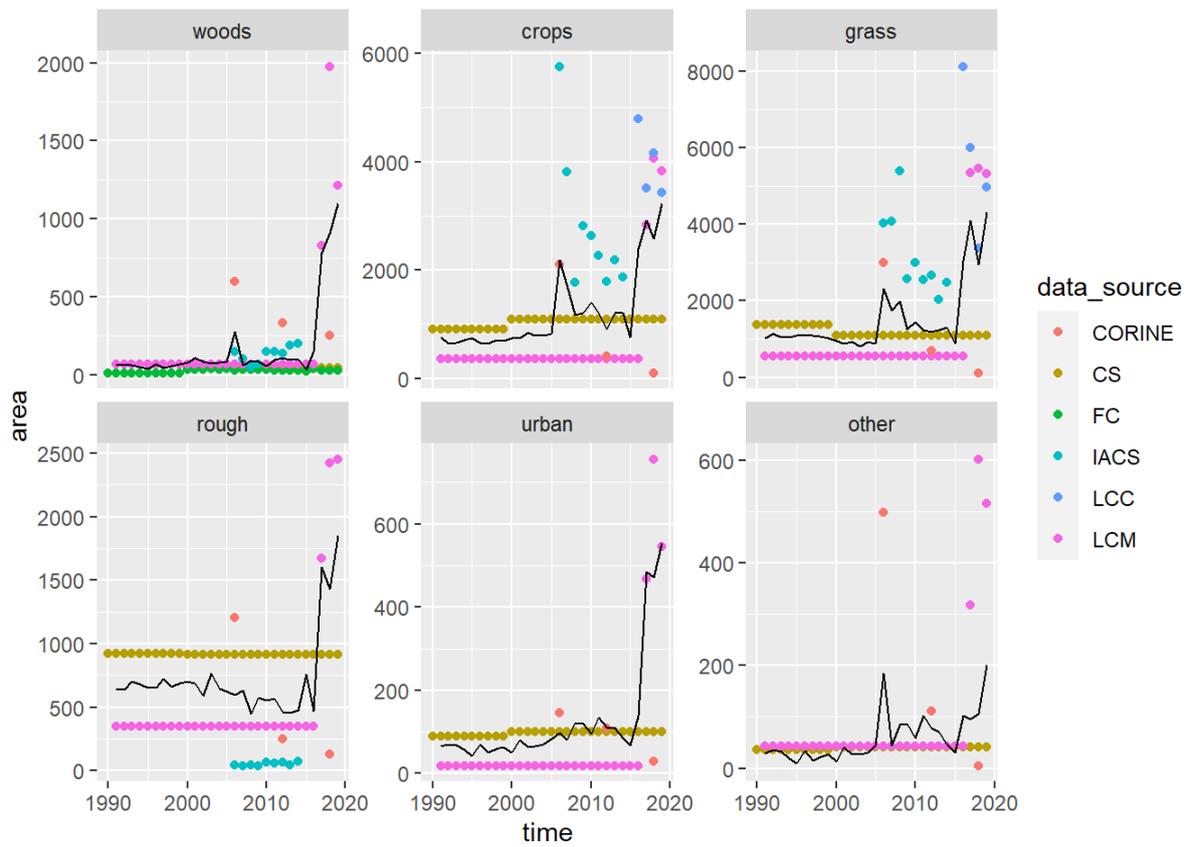


Figure 6.3: Time series of gross losses of area L from each land use observed by different data sources. Coloured symbols show the observations; the black line shows the best-fit values estimated by least-square optimisation. LCM and CS values between surveys were interpolated values as constant annual rates.

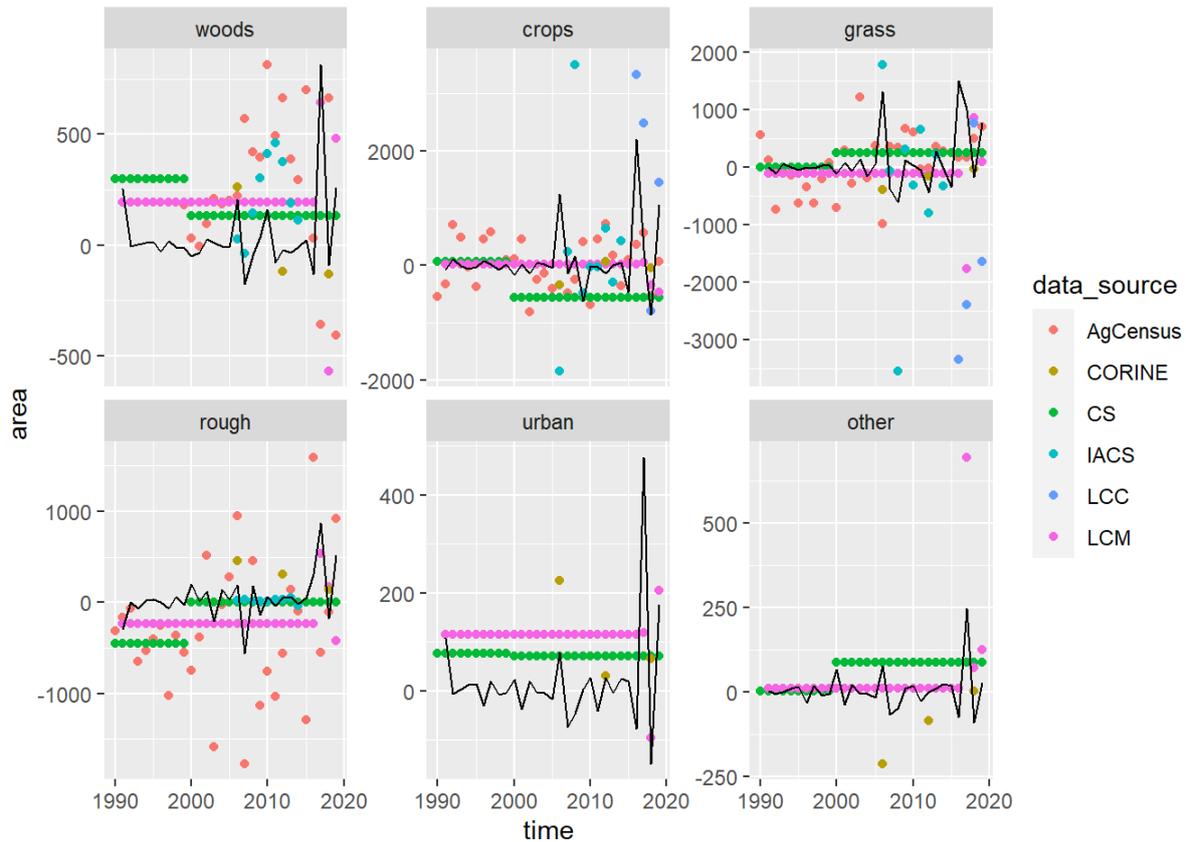


Figure 6.4: Time series of net change in area ΔA of each land use observed by different data sources. Coloured symbols show the observations; the black line shows the best-fit values estimated by least-square optimisation. LCM and CS values between surveys were interpolated values as constant annual rates.

6.4 Discussion

One purpose of the exercise is to show that the DA procedure can be done in a simpler, non-Bayesian way, and the figures above show that this is possible. The practical purpose, to provide a quick method to give initial values for MCMC chains is demonstrated by the results above.

More informatively, two main weaknesses of the non-Bayesian approach are apparent. Firstly, no quantification of uncertainty is provided. The LS algorithm identifies the single parameter set which has the best fit with the data, given our definition of RMSE. However, it gives us no estimate of the probability of any other parameter set, nor does it tell us anything about the confidence we should have in predictions.

Secondly, our definition of RMSE is somewhat arbitrary. We implicitly assume that small RMSE values relate to high likelihoods, and that our best-fit parameter set might approximate to that with the maximum likelihood. However, we have not explicitly calculated the likelihood. Where this really matters is if we want to make the likelihood function more complex, to allow us to combine data sources with different dimensions and degrees of uncertainty. We implicitly give each observation the same uncertainty, and we can only add up the different terms in our RMSE because they happen to have the same units. If either of these were not the case, we could attempt to add in some arbitrary weightings to address this. However, the Bayesian approach explicitly calculates the likelihood, which is the probabilistically-correct formulation of the problem. In the present context, it is the answer to the question “what is the probability of observing the data if we know B ?” (Oijen 2020). For each

observation, no matter what its source, measurement uncertainty or units, we obtain a probability, not an arbitrary residual. Because we are working in the common currency of probability, we can multiply the values for all observations to give the likelihood for the whole observation set - parameter set combination. This is the approach we apply in the next section.

7 Estimating B by MCMC

7.1 Introduction

As described in the **Methods** section, we perform the data assimilation (DA) as a two-stage procedure. The previous section showed how the first stage can be done in a simple way with least-squares optimisation. In this section, we re-do the data assimilation using a Bayesian approach. The aim is to find not only the B matrix parameters with the maximum likelihood, but also the full posterior distribution for the parameters and predictions. This quantifies the uncertainty in our estimates and predictions of land-use change. To do this, we use a Markov chain Monte Carlo (MCMC) algorithm. MCMC is the standard method for Bayesian parameter estimation where the problem is too complex for an analytical solution to be found. It provides a means of sampling from an unknown posterior distribution, so long as the likelihood of the data can be evaluated.

7.2 Methods

The method used here closely follows that described by Levy et al. (2018), so our description here is brief, and focuses on the relatively minor differences.

7.2.1 Observational data

Based on the earlier discussion of data sources, we elected to use CS, FC and Agricultural Census data to estimate the B matrix. Other data sources can be still used to estimate the spatial pattern in land-use change, and can be added to the estimation of the B matrix in future. At present, the other data sources examined were either inconsistent in quality and the focus of ongoing and future work, incomplete, or unequivocal permission for use had not been obtained.

7.2.2 The model

The model used here is the simple arithmetic of matrices described in the **Methods** section. Given a matrix B , this calculates the gross and net changes in area G , L and ΔA . The only additional element, which is simple but potentially confusing for terminology, is that as well as the B values acting as parameters, we also have observations of these, and this comparison is included in the likelihood. In mathematical terms, we can think of this part of the model as being the identity function, i.e. a function which returns the input value, so $y=f(B)=B$.

7.2.3 Prior distributions

Levy et al. (2018) used the CS data as the prior distribution for the B matrix. This made sense in that context, where there were more data sources available, and CS had been used as the basis for previous estimates. In the present case, in order to treat all the data sources consistently, it is easier to consider

this as an observational data set like the others. We then need to specify a prior distribution for B independently of the data. We chose two options for comparison (Figure 7.1):

- a uniform distribution, where all B parameters could vary between 0 and 10000 km², and
- a half-normal distribution, with mean zero and standard deviation $\sigma=3000$ km². “Half-normal” means that only positive values are possible. This provides a relatively weak version of the parsimony principle, that lower rates of gross change are more probable (all else being equal). The strength of this assumption can be altered by varying σ .

Large differences were not seen between these two, and all results shown here use only the uniform prior assumption for simplicity, though there is scope for further exploration here.

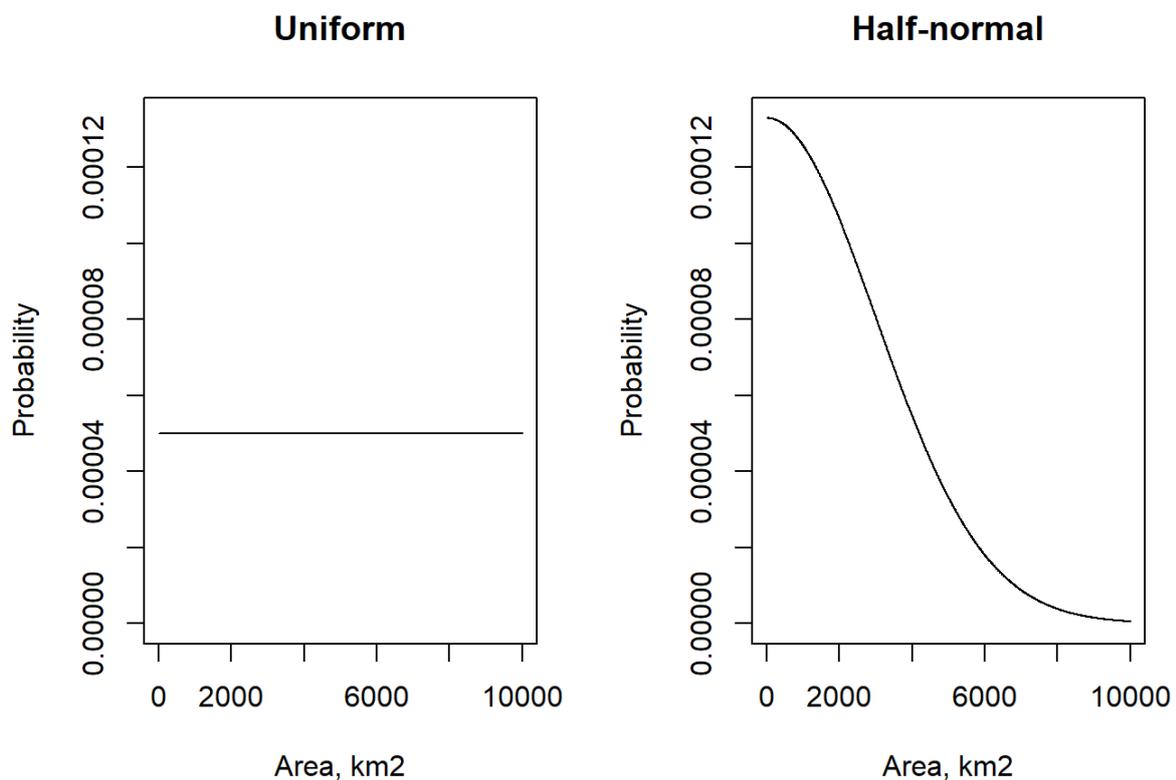


Figure 7.1: Two alternative prior distributions for the B parameters.

7.2.4 Likelihood function

The likelihood function is very similar to the function used to calculate RMSE in the previous section. The key difference is that, for each observation, a likelihood is calculated, assuming that measurement errors show a Gaussian distribution (hence the use of the `dnorm` function) and are independent of each other. To put this in mathematical notation, the likelihood of observing the net change ΔA^{obs} is

$$\mathcal{L} = \prod_{u=1}^{n_u} \frac{1}{\sigma_u^{\text{obs}} \sqrt{2\pi}} \exp\left(-(\Delta A_u^{\text{obs}} - \Delta A_u^{\text{pred}})^2 / 2\sigma_u^{\text{obs}2}\right) \quad (19.1)$$

where ΔA_u^{pred} is the prediction for the change in land use u , given the parameter matrix \mathbf{B} , and σ_u^{obs} is the measurement uncertainty. There are analogous terms for G , L and B which can all be multiplied.

Rather than assume that measurement uncertainty is the same for all observations, we specify that it is proportional to the y value. That is, observations of large areas come with larger absolute uncertainty. Potentially, we can specify unique uncertainties for each data source and observation. At present, we do not have a basis for specifying these uncertainties based on a proper quantitative analysis, but this could be an area of future work. For now, we assume that all data sources are equal, and assign them the same relative uncertainty.

7.2.5 MCMC algorithm

Many MCMC algorithms are available. We chose the parallelised version of the Differential Evolution Adaptive Metropolis or “DREAMz” algorithm (Vrugt et al. 2008), a development of an earlier Differential Evolution algorithm (ter Braak and Vrugt 2008). These, and other MCMC algorithms are implemented in the R package `BayesianTools` (Hartig, Minunno, and Paul 2017), which allowed us to compare the efficiency of different approaches. Differential Evolution MCMC uses an adaptive algorithm, in which multiple chains are run in parallel. It uses information from the past states of multiple chains, by generating jumps from differences of pairs of past states. It increases efficiency, particularly for high-dimensional problems, producing convergence with fewer iterations and fewer chains, estimated to be about 5–26 times more efficient than the typical Metropolis sampler (ter Braak and Vrugt 2008).

The problem was parallelised at two levels. Firstly, we can parallelise in the time dimension. The transition matrix between each pair of years can be treated independently, so we can run all years on separate processors at the same time. In other words, we treat the time series between 1990 and 2019 as 30 separate tasks, rather than one. Secondly, we can parallelise across MCMC chains. The DREAMz algorithm uses multiple chains, each with a number of internal chains which inter-communicate. So long as we allow for this inter-communication, we can run each (internal) chain on a separate processor. For each pair of years, we ran 120000 iterations with nine chains in total (three chains, each with three internal chains). Running a single chain of this type takes around 10 mins. If we ran all nine chains for 30 years in series, this would take $30 \times 9 \times 10/60 = 45$ hours. Using 270 processors on the NERC/STFC [JASMIN super-computer](#), all chains can be run simultaneously in ten minutes, subject to the job queueing system having capacity.

7.3 Results

Figure 7.2 shows the basic diagnostic check for MCMC output. The chains should start over-dispersed, so that a wide range of parameter space is explored. After a “burn-in” phase, the chains should begin to converge on an area of high probability. To account for auto-correlation between sequential samples, it is standard practice to “thin” the chains by a factor of ten or more, i.e. leaving only every tenth value. Here, the iteration count refers to the thinned samples, so the actual iteration count is a factor of ten higher. Depending on the nature of the model and the data, we might expect something like a normal distribution around the most likely value, though this is not necessarily the case. The diagnostic test for convergence is whether the variation among chains is greater than the variation within chains, and we reach this point after about 4000 thinned iterations.

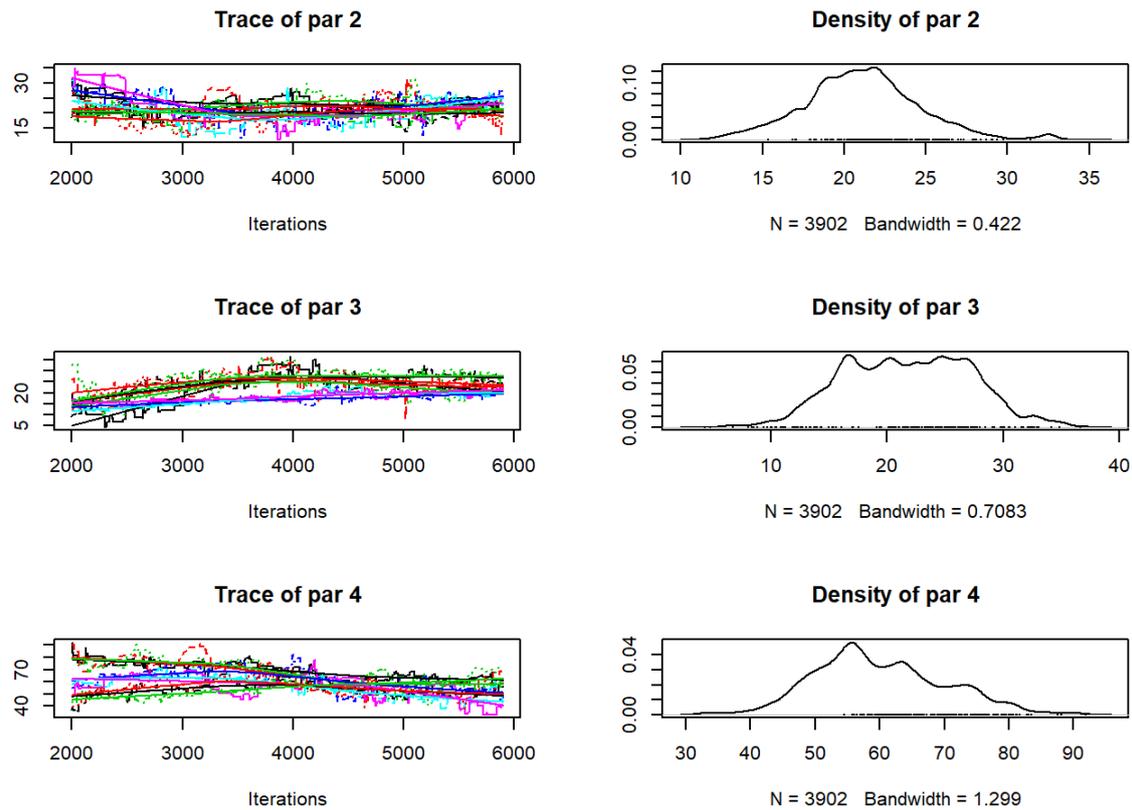


Figure 7.2: Trace plots for three parameters from the MCMC run for 2018-2019. Parameters represent the area changing from woods to crop, grass and rough grazing, respectively. The left-hand plot shows the values sampled from the posterior as the iterations progress in each chain. Different colours show the different MCMC chains. The iteration count refers to the thinned samples after burn-in; the actual iteration count is a factor of ten higher. The right-hand plot shows the density distribution of parameter values, which represents our sampled approximation to the posterior probability distribution.

An important feature of the MCMC approach is that it properly estimates the joint probability distribution for the parameters, and these correlations are automatically incorporated. In the example shown in Figure 7.3, the areas changing from crop to grass and vice versa are clearly positively correlated. This follows because the net change is constrained by the observations; if the crop-to-grass area is large, the grass-to-crop area also has to be large to fit with the observed net change.

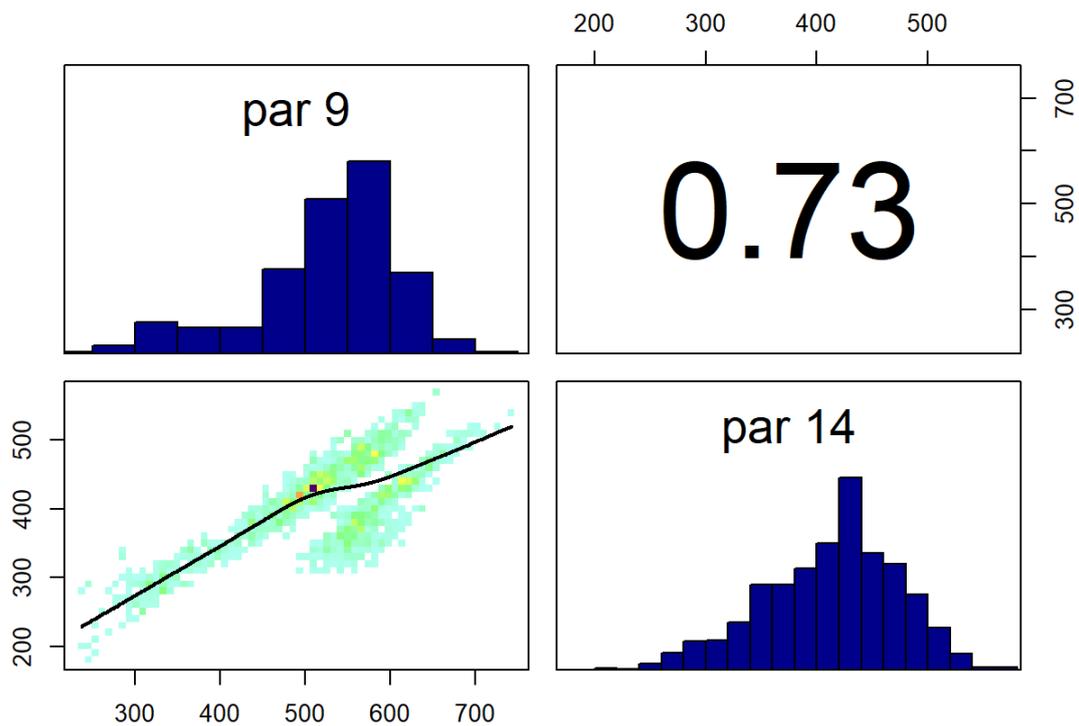


Figure 7.3: Correlation plot for two parameters from the MCMC run for 2018-2019. Parameters represent the area changing from crop to grass and vice versa. The plot shows the correlation between parameters values obtained over the course of the MCMC sampling. This represents our sampled approximation to the correlations in the joint posterior probability distribution for the parameters.

In Figures 7.4, 7.5, 7.6, and 7.7, the solid red line shows the maximum *a posteriori* estimate, which represents the B parameter set with the highest posterior probability, the closest equivalent to a “best-fit” estimate. The shaded area denotes the 95 % credibility interval. In these plots, the MCMC estimates do not fit any single set of observations or the B , G , L or ΔA variables exactly, as they are constrained by all of these simultaneously.

Figure 7.4 shows the time course of the posterior distribution of the B matrix estimated by MCMC. The estimates are strongly influenced by the CS data, as these provide the only observations of B used here. However, annual variability is introduced, as the parameters are varied to simultaneously fit with the Agricultural Census and FC data. In most cases, the posterior estimates fit closely with the CS observations but this is not always the case. To understand the deviations and cause of year-to-year variability is not easy, because this is a very high-dimensional problem. For example, the B parameter for the transition woods-to-rough is lower and more variable than the CS observation after 2000. However, the likelihood of this parameter is influenced by the agreement with all the other observations and variables. The MCMC algorithm attempts to find parameter sets which simultaneously satisfy all the conditions described by the observations e.g. that the area changing from woods-to-rough is a constant 25 km² (CS); that the total gains and losses to/from rough are close to 1000 km² (CS); and that the net change in rough grazing varies between +1500 km² and -1500 km² over the same period (Agricultural Census).

In the case of forestry, where we have the observations of the gross gains and losses, the problem is much better constrained (Figures 7.5 and 7.6). Here, the posterior estimates fit the FC and CS estimates well, but cannot simultaneously reproduce the pattern in the Agricultural Census data (Figure 7.7, top left). The Agricultural Census data, clearly have an influence on the year-to-year pattern in the estimates, but parameter combinations which would reproduce the amplitude of variation seen in the Agricultural Census data clearly have lower likelihood.

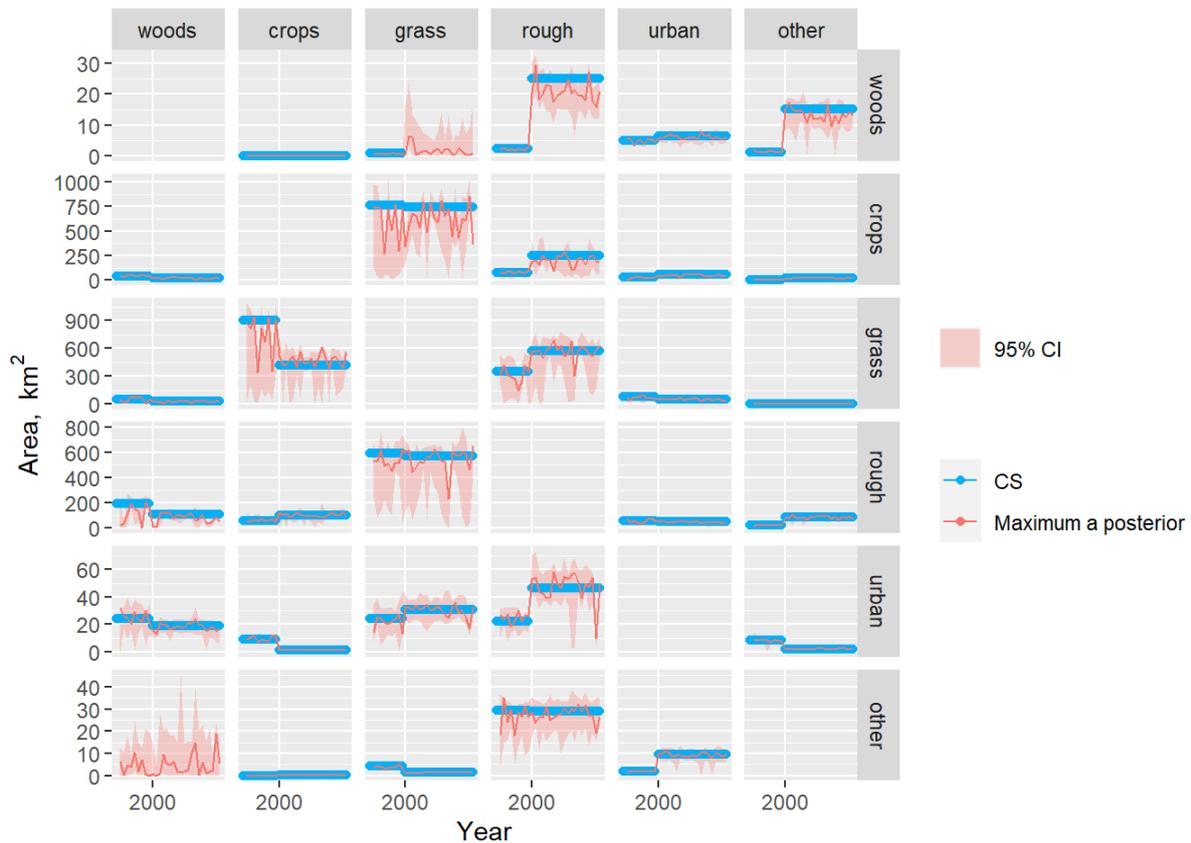


Figure 7.4: Observations and posterior distribution of the transition matrix \mathbf{B} , representing the gross area changing from the land use in each row to the land use in each column each year from 1990 to 2019. The shaded band shows the 2.5 and 97.5 percentiles of the posterior distribution. The maximum *a posteriori* estimate is shown as the solid red line within this. Note the y scale is different for each row.

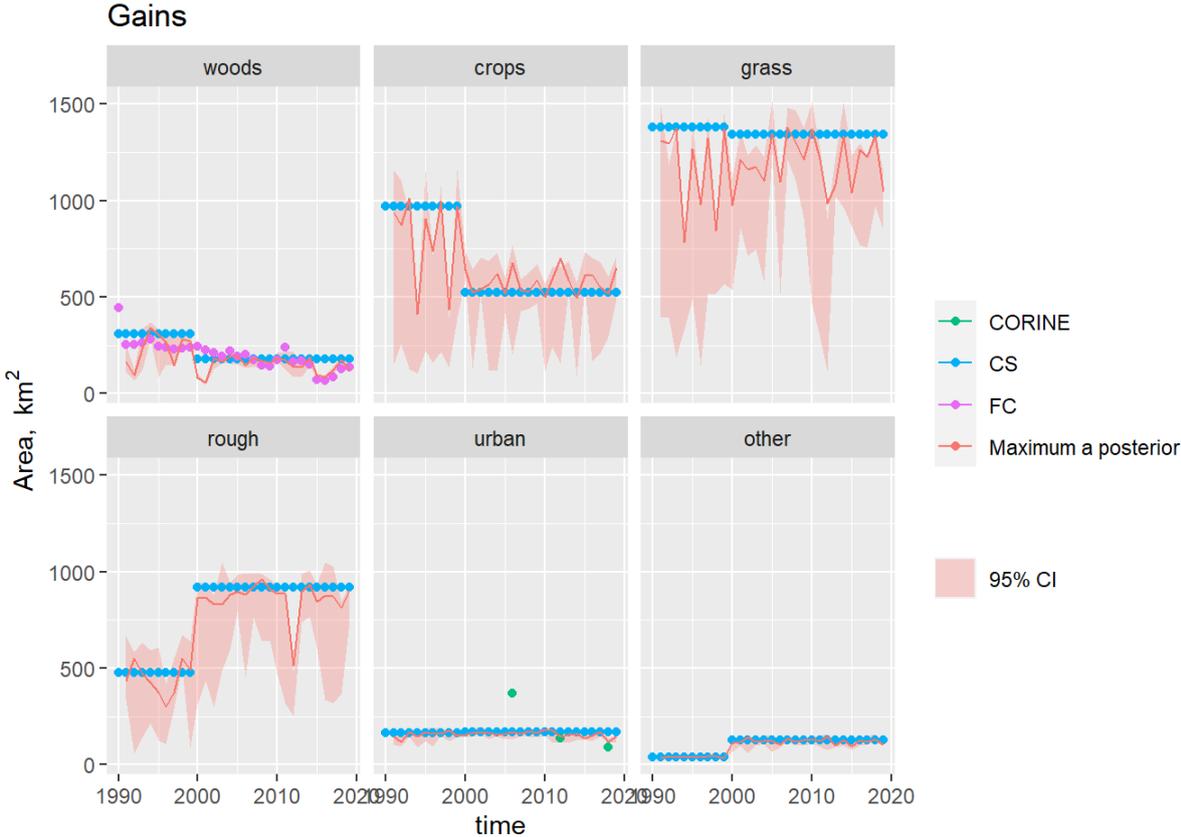


Figure 7.5: Observations and posterior distribution estimated by MCMC of the gross gain in area of each land use G from 1990 to 2019. The shaded band shows the 2.5 and 97.5 % percentiles of the posterior distribution. The maximum *a posteriori* estimate is shown as the solid red line within this.

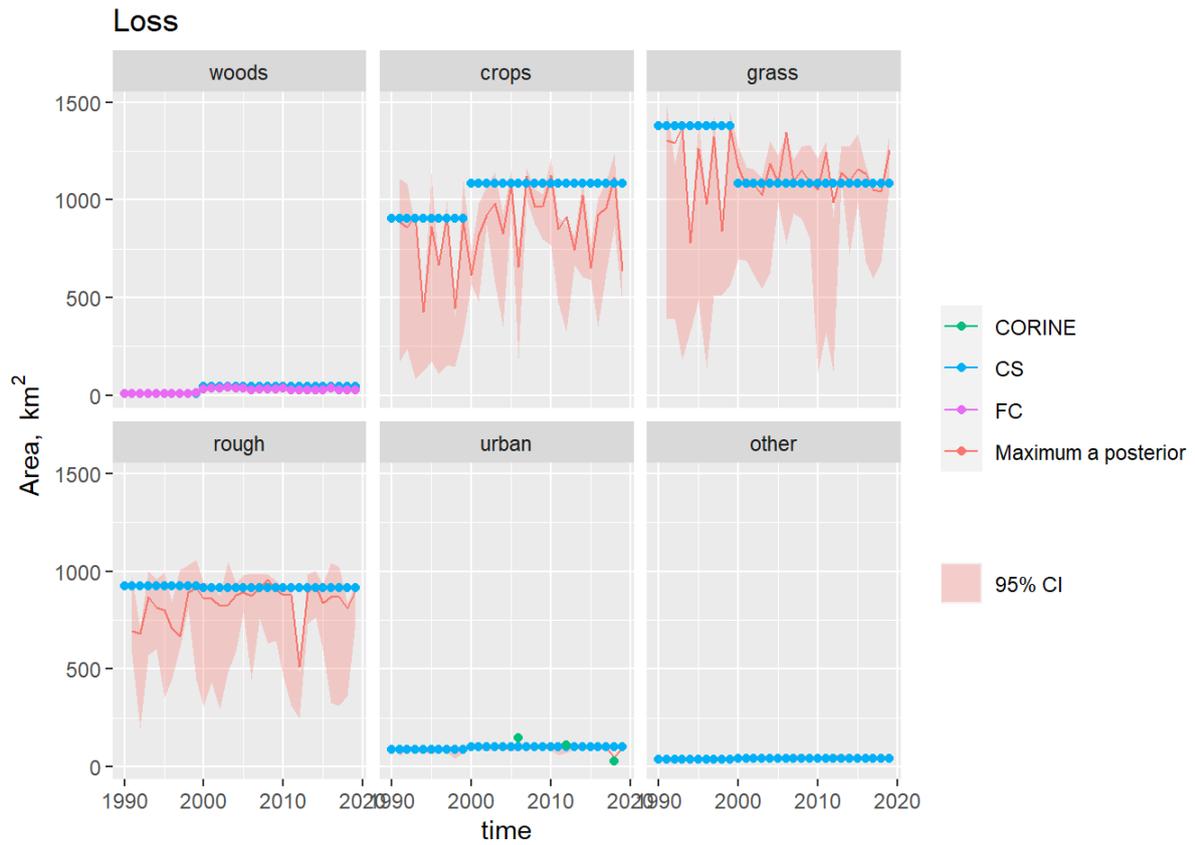


Figure 7.6: Observations and posterior distribution estimated by MCMC of the gross loss of area from each land use L from 1990 to 2019. The shaded band shows the 2.5 and 97.5 % percentiles of the posterior distribution. The maximum *a posteriori* estimate is shown as the solid red line within this.



Figure 7.7: Time series of the net change in area occupied by each land use (ΔA) from 1990 to 2019, showing the observations and posterior distribution of estimates. The shaded band shows the 2.5 and 97.5 % percentiles of the posterior distribution of the net change in area of each land use. The solid red line shows the maximum a posteriori prediction.

7.4 Discussion

In the current GHGI method, the CS and FC data are taken at face value as the estimate of land-use change. The results shown above represent an advance beyond this in several ways. Most obviously, the constraint of national-scale survey is introduced. This means that the B estimates are less subject to the sampling error associated with using only 544 sites across the UK. The Agricultural Census data is an imperfect constraint, as it specifies only net change, but it is based on a very comprehensive and long-standing census programme, and it far more likely to pick up the broad trends in national-scale land use over long periods of time. Secondly, our method quantifies uncertainty in a rigorous way, consistent with probability theory. This aspect could be improved by quantifying the uncertainty on the input observational data in a systematic way. Lastly, although we introduce only one additional data source beyond those used in the current GHGI to estimate B , the framework is in place to introduce any other data sets we choose. For simplicity, we elected not to include the LCM, LCC, CORINE or IACS data in the results presented here. However, adding these in is a simple matter, if we choose to believe the observations or can improve their consistency.

8 Estimating Land-Use Change Spatially

8.1 Introduction

In the previous sections, we have estimated the posterior distribution of the B matrix for each year, using MCMC. Next, we need to use the B matrix to predict land-use change spatially. We do this by starting with the relatively well-known state of land use U at the present-day and move backwards in time. At each time step, the B matrix for that year tells us how many grid cells need to change from each class to every other class. For a given element B_{ij} in the matrix, the candidate cells for change are all those where $u=i$. We do not know exactly which cells actually changed in a given year, but there are several spatio-temporal data sets which gives us information about this. For example, LCC and IACS both have spatio-temporal observations of the time and location of agricultural land use. The Agricultural Holdings data tell us where and when the changes in crop, grass and rough grazing occurred, based on farmer census returns. The FC NFI/SCDB data tell us the location and planting dates of forestry. None of these data sets are perfect, and we may not believe the absolute magnitudes of change implied, but they tell us something about the probability of land being used for a given purpose at a given location and time. Our approach attempts to make the best use of the available data to estimate the location of each land-use change probabilistically. Again we use a Bayesian approach.

Each year, our aim is to sample from the posterior distribution of the land-use map U_{t-1} . Because we have already established the B matrix, and we can consider each change to be independent, the problem is simply to choose the location of each change in accordance with Bayes Theorem. To do this, we can use a method referred to as “importance sampling” (Hartig et al. 2011). Importance sampling is a very efficient form of rejection sampling (aka the “accept-reject” algorithm, or Sequential Monte-Carlo (SMC)). In the previous section we used MCMC to sample from the posterior distribution. This is usually necessary, because we have to evaluate the likelihood function across a wide range of high-dimensional parameter space by iterative numerical sampling. However, in the case of estimating the land-use map U_{t-1} , the form of the model and the likelihood is so simple that we can calculate the likelihood in advance. The model is simply the identity function, and there are only six possible states {woods, crops, grass, rough, urban, other}. For each of these possible states, we can evaluate the likelihood L , the probability that the proposed land-use state \hat{u} is correct, given a vector of observations \mathbf{U} . For example, with a vector of observations $\mathbf{U} = \{\text{crop, crop, grass}\}$ from three data sources (e.g. LCM, IACS and CORINE), we might estimate the probability that the true land-use state is “crop” to be $2/3$. If we know the precision of these data sources, we can adjust this accordingly. Similarly, if we have different observations, such as the probability of an increase in crop area in the local region, we can incorporate these also.

Having established these six likelihoods at each location for each year, we can then use these in the sampling process when choosing from the candidate cells for land-use change. The basis of all sampling algorithms is that the probability of a parameter set belonging to the posterior distribution is proportional to the likelihood. The simplest such algorithm, rejection sampling, works by choosing a huge number of possible parameter sets at random, often from a uniform distribution; the chance of their being accepted is proportional to their likelihood. Importance sampling is similar, but the candidates are sampled with some importance weighting, such that they are somehow sampled approximately in proportion to their likelihood. The inefficient rejection step can be minimised because improbable parameters sets are not being selected in the first place. Here, we have the

extreme case where we exactly know the likelihood in advance, so can use this in weighting the sampling process such that all samples can be ascribed to the posterior distribution. We use the sampling algorithm in the R package `wrswor` to implement this.

8.2 Methods

The implementation of importance sampling described above breaks down into a number of tasks.

Firstly, we evaluate the likelihood function for each of the six possible states {woods, crops, grass, rough, urban, other} for all years and at all locations. Computing this for 30 years at a resolution of 100 m requires ~two hours and a large amount of memory, but is feasible on JASMIN, and considerable stream-lining is possible in future.

Then, we begin with the present-day land-use map U_t , as defined by LCM in 2019, with additional forest where this is reported in the current FC NFI/SCDB or FSNI data. This is a 100-m resolution raster grid, with 91 million cells, of which ~24 million of which are on land. For each year from the present-day going backwards:

- select a B matrix from the 2000 posterior samples produced in the [previous section][Data Assimilation: Estimating B by MCMC];
- for each element B_{ij} , define the candidate cells as those where $u=i$;
- select the number of cells specified by B_{ij} from the set of candidate cells, with the probability of inclusion being proportional to $L_{u=j}$;
- for the selected cells, set the new land use to equal j .
- remove the selected cells from the candidate set so they cannot be resampled that year;
- when this has been completed for all elements B_{ij} , move to the next year back in time and repeat until the starting year is reached (1990 in this case).

The above procedure produces a single sample from the posterior distribution of \mathbf{U} , and it can be repeated to produce as many samples as desired. How many many samples are required will depend on the purpose at hand.

The basics of this procedure are very simple, and it can be repeated there is no good model for predicting which cells change in any given year. Having B_{ij} matrix, the candidate cells for change are all those where $u=i$. the state of land use U at the next time step. Although there are predictable patterns in which cells tend to be used for forest, crops or grassland etc., exactly when a given cell will change land use is not predictable; the process is effectively stochastic for our purposes.

8.3 Results

All the maps here show 1-km resolution data because higher resolution is not easily discernible for large areas on typical graphical devices. However, all these data are originally produced at 100-m, and aggregated to lower resolution for display and summary purposes. The plots below show maps of the likelihood L for each of the six land uses (Figures 8.1-8.6). These show little that is unexpected. The likelihood is expressing the number of data sources which predict that land-use type, as a fraction of the number of data sources available. It thereby effectively averaging predictions of that land-use type across the data sources. The likelihood of the presence of woods is dominated by the FC NFI/SCDB data, so the maps closely resemble this, but with the addition of other data sources which include the location of woods (LCM, CORINE, and IACS). The likelihood of the presence of crops is an amalgamation of all the data sources which include the location of crops (LCM, LCC, CORINE, IACS, Agricultural Holdings). The same principles apply to the other land-use types.

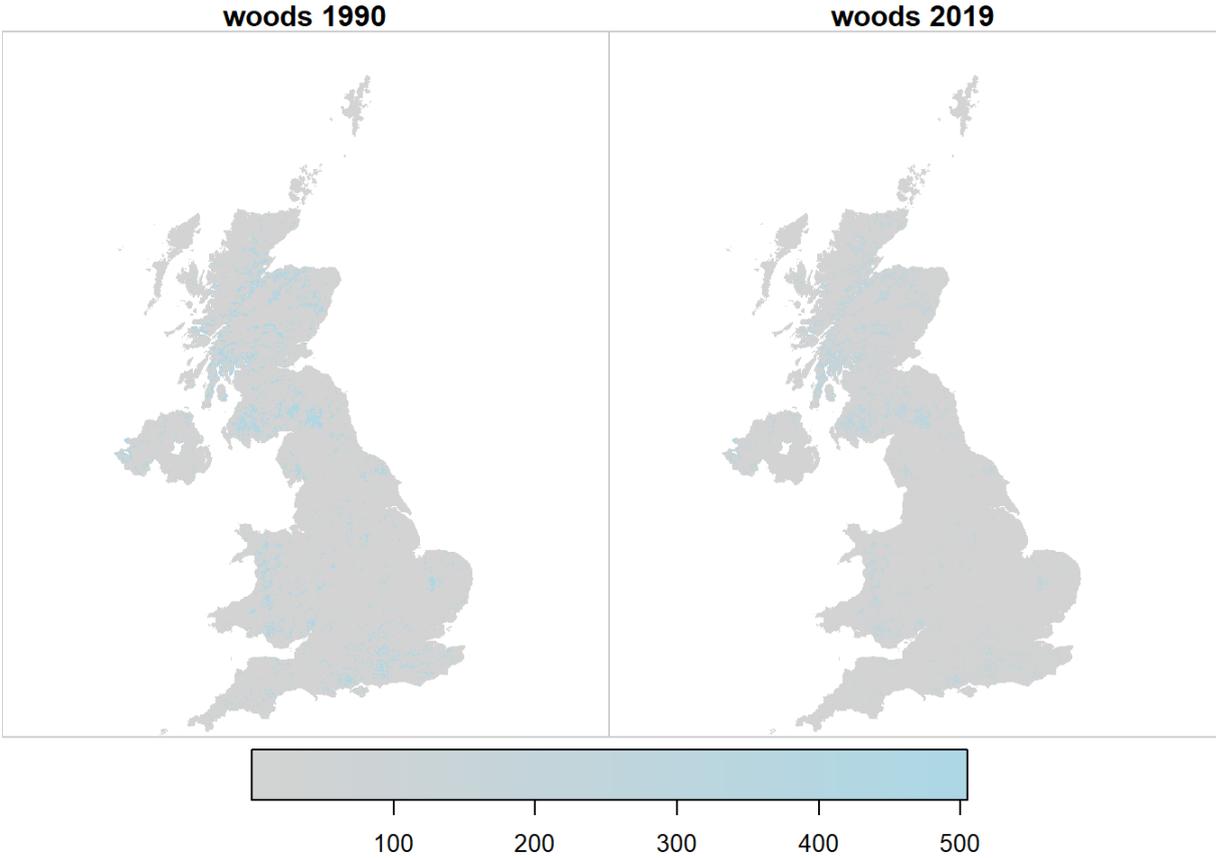


Figure 8.1: Spatial variation in the likelihood L of observing land use u at two example points in time, arbitrarily rescaled for plotting.

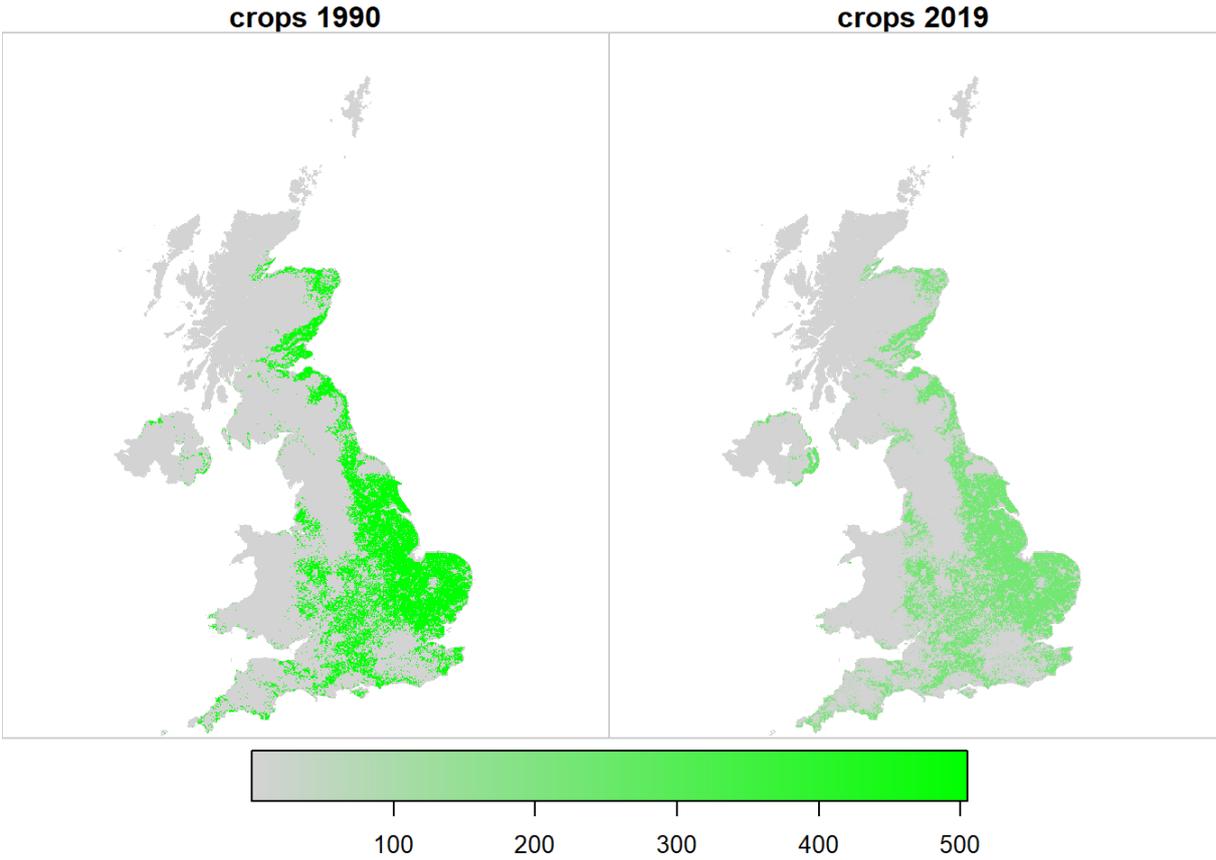


Figure 8.2: Spatial variation in the likelihood L of observing land use u at two example points in time, arbitrarily rescaled for plotting.

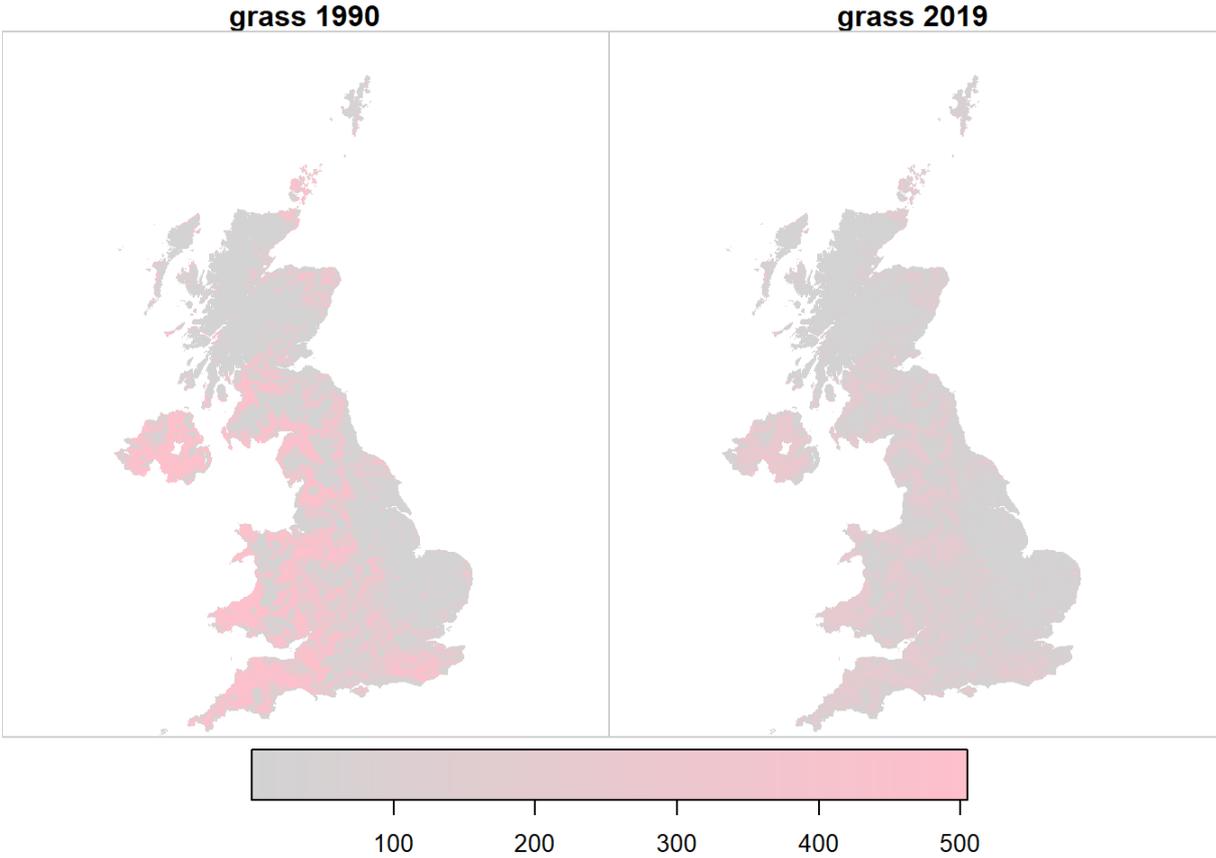


Figure 8.3: Spatial variation in the likelihood L of observing land use u at two example points in time, arbitrarily rescaled for plotting.

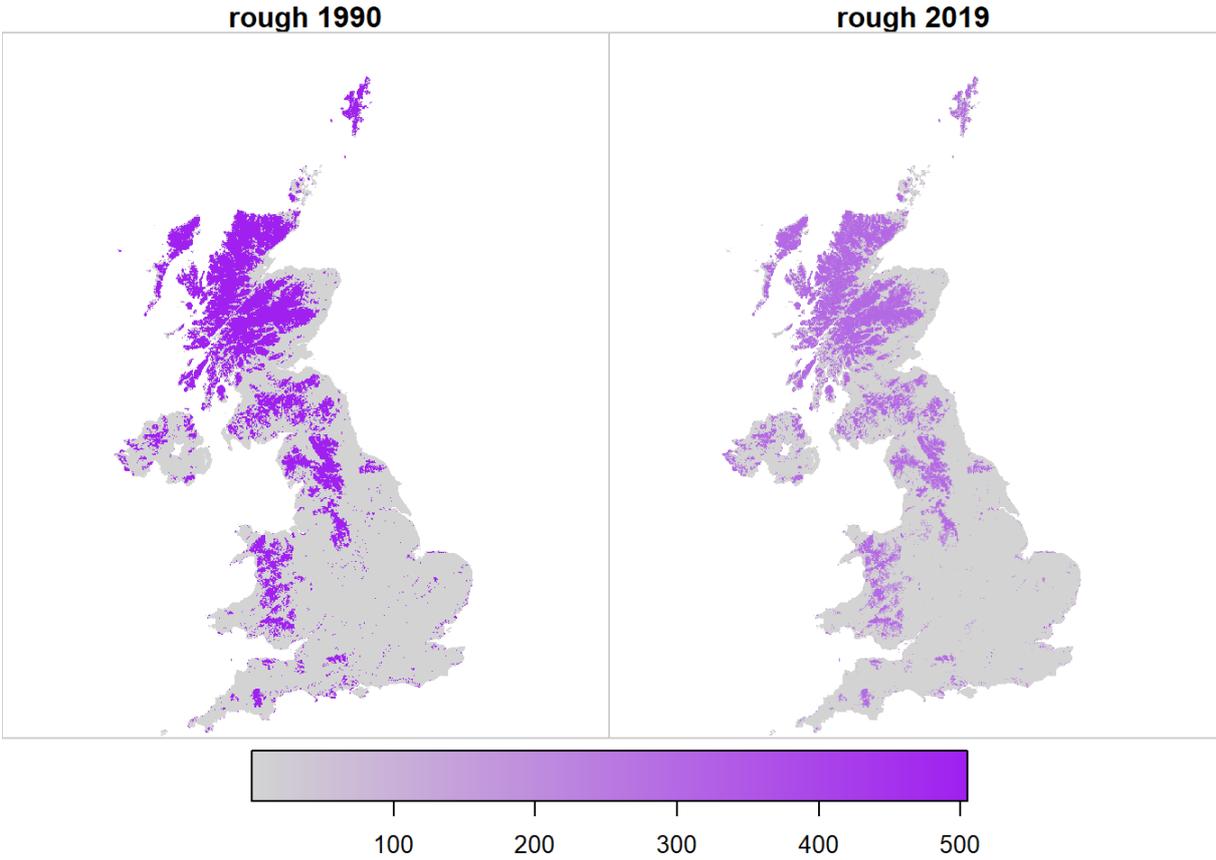


Figure 8.4: Spatial variation in the likelihood L of observing land use u at two example points in time, arbitrarily rescaled for plotting.

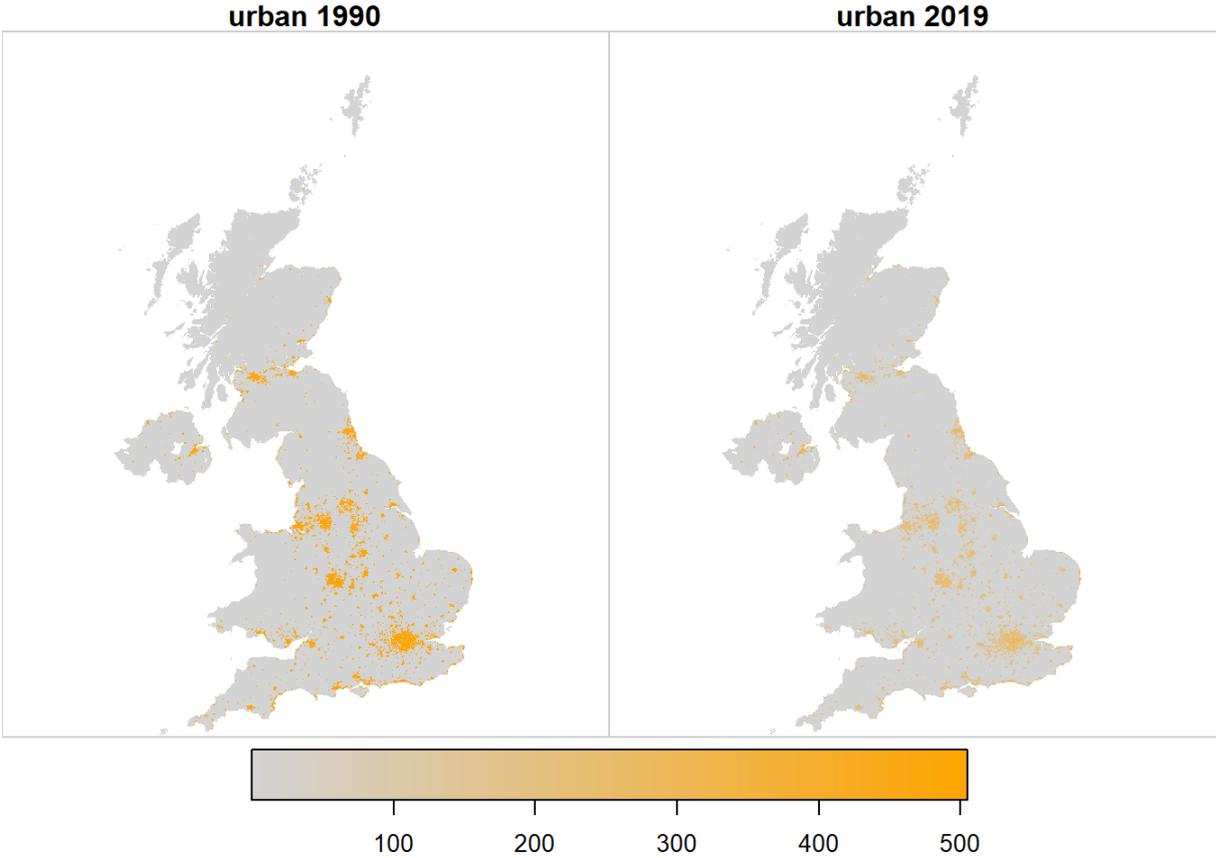


Figure 8.5: Spatial variation in the likelihood L of observing land use u at two example points in time, arbitrarily rescaled for plotting.

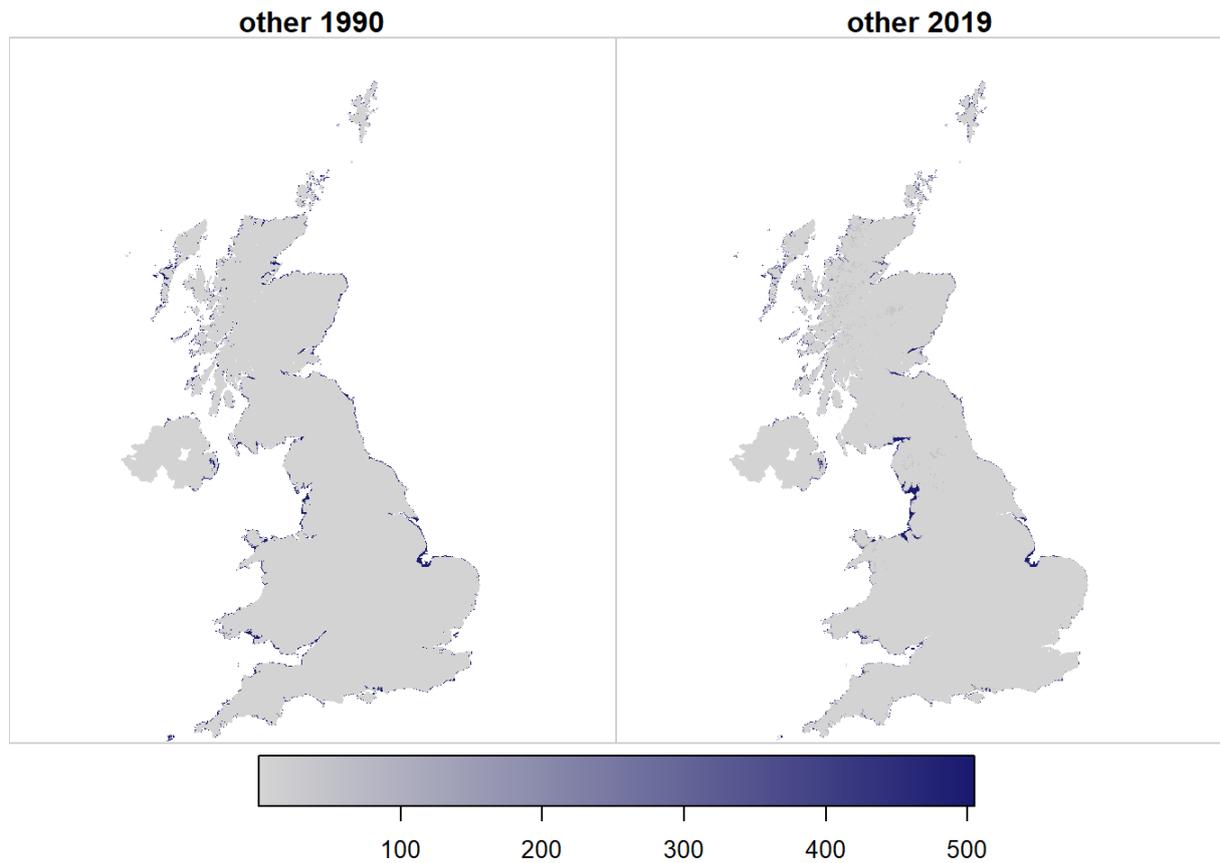


Figure 8.6: Spatial variation in the likelihood L of observing land use u at two example points in time, arbitrarily rescaled for plotting.

Figure 8.7 shows the outcome from the importance sampling algorithm described in the Methods. The 2019 map is defined solely by observations (LCM, FC NHI/SCDB, and FSNI data). The 1990 map is the result of applying the set of land-use changes prescribed in the series of annual B matrices from the posterior distribution. This map is thus based on the magnitude and nature of land-use change from the data sources in which we have highest confidence in their consistency over time (CS, FC and AgCensus), and applies this spatially, using the data sources which provide information on the spatial pattern of land use change. Because the absolute extent of land changing use is small, and it tends to occur with relatively small granularity. Differences are not easily apparent when viewing the whole UK, or even at a Devolved Administration level (Figures 8.8, 8.9, 8.10), but we show these here for completeness.

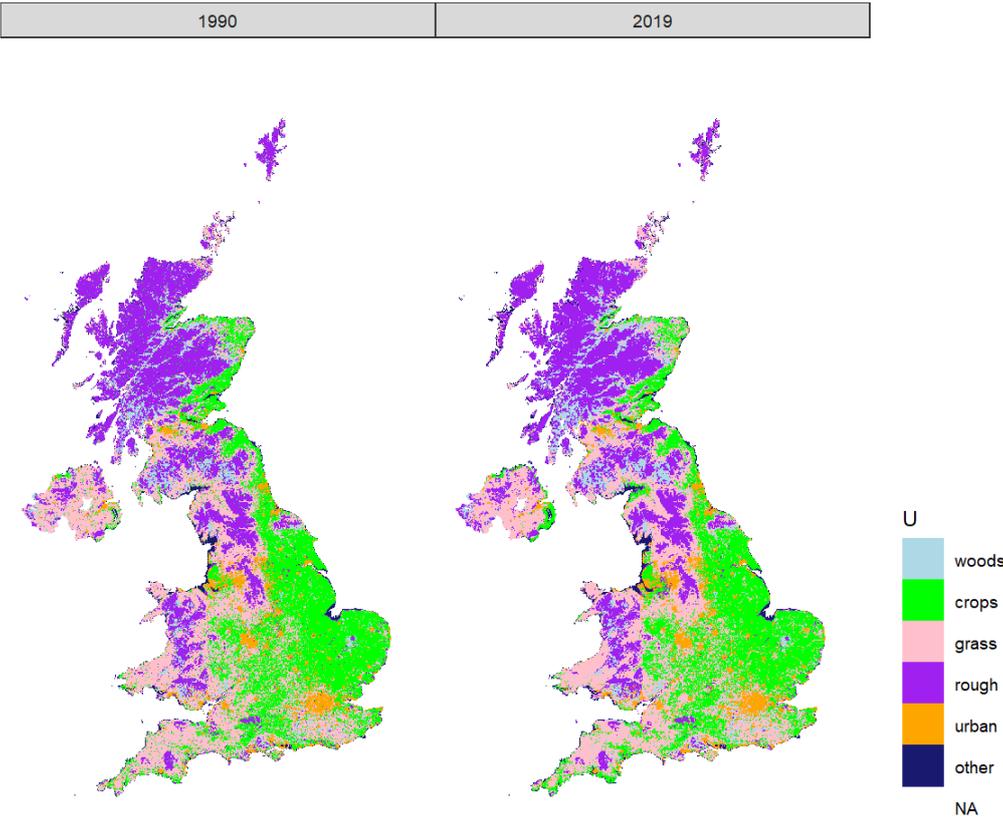


Figure 8.7: Estimated state of land-use U in 1990 and 2019 in the UK.

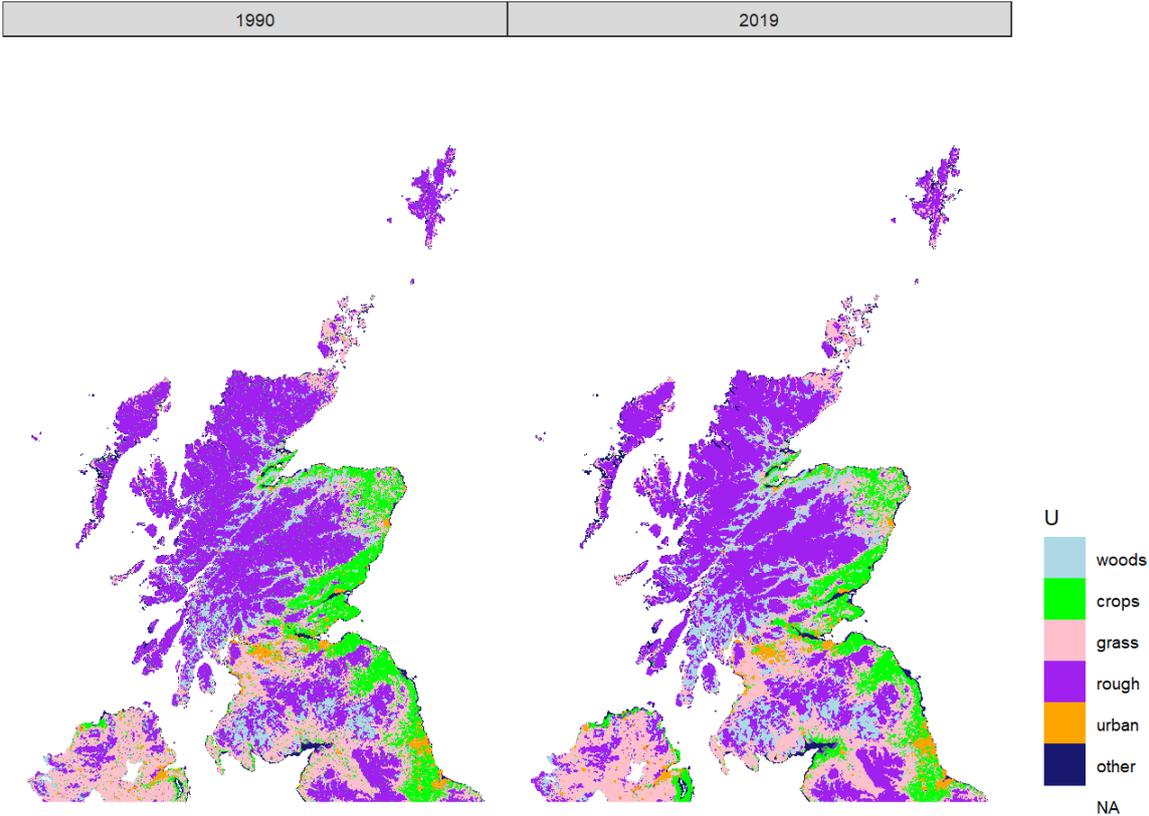


Figure 8.8: Estimated state of land-use U in 1990 and 2019 in Scotland and Northern Ireland.

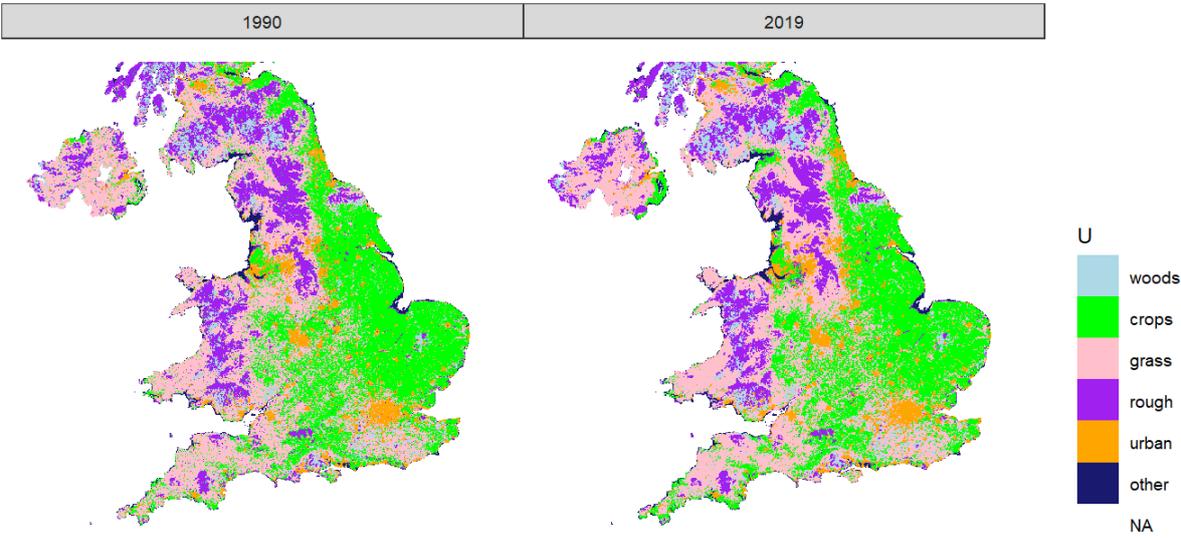


Figure 8.9: Estimated state of land-use U in 1990 and 2019 in England.

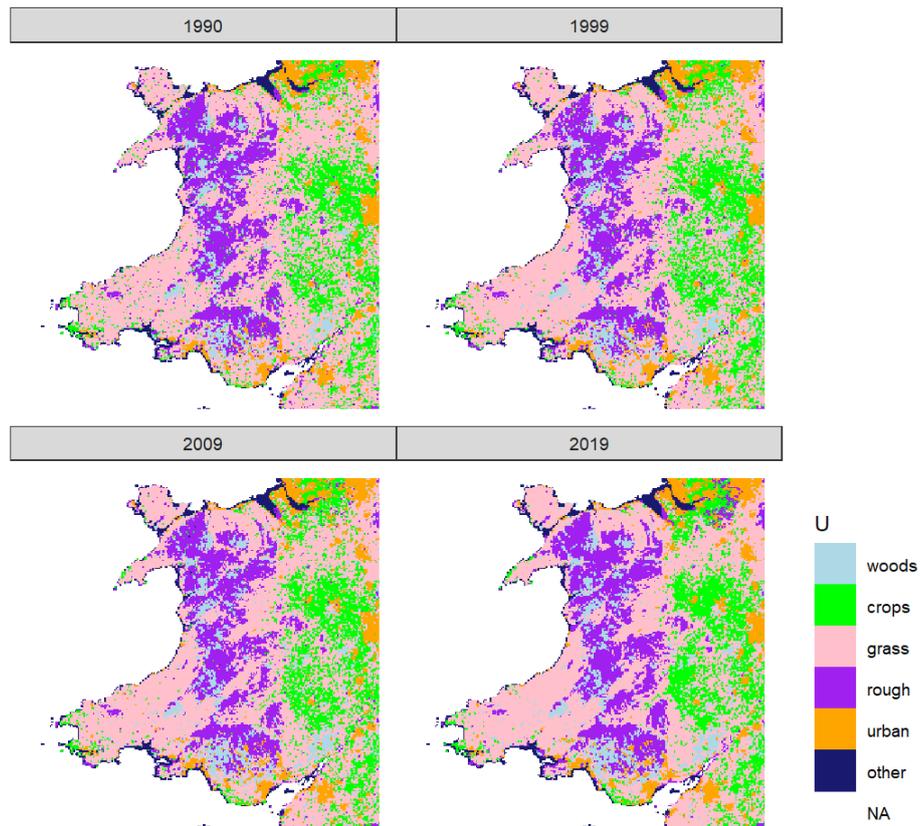


Figure 8.10: Estimated state of land-use U in 1990, 1999, 2009 and 2019 in Wales.

Slightly more useful is to zoom in on smaller regions. Figures 8.11 and 8.12 show two 100 x 100 km regions in Scotland, to illustrate the potential for analysing these results in by comparison with other mapped data. The broad expected patterns, such as some areas with known urbanisation over this period can be identified. However, beyond basic sense-checking, a detailed spatio-temporal analysis of this output has not yet been done.

For the purposes of the UK GHGI, the actual spatial patterns are not important in themselves (although obviously desirable). As currently implemented, there are no explicitly spatial terms in the soil carbon modelling. Instead, we can summarise the spatio-temporal data cube U as the set of unique vectors of land use. To greatly reduce the volume of input and output data, the soil carbon model can be applied to this much smaller set of inputs. We analyse the output of the data assimilation procedure in terms of these vectors in the next section.

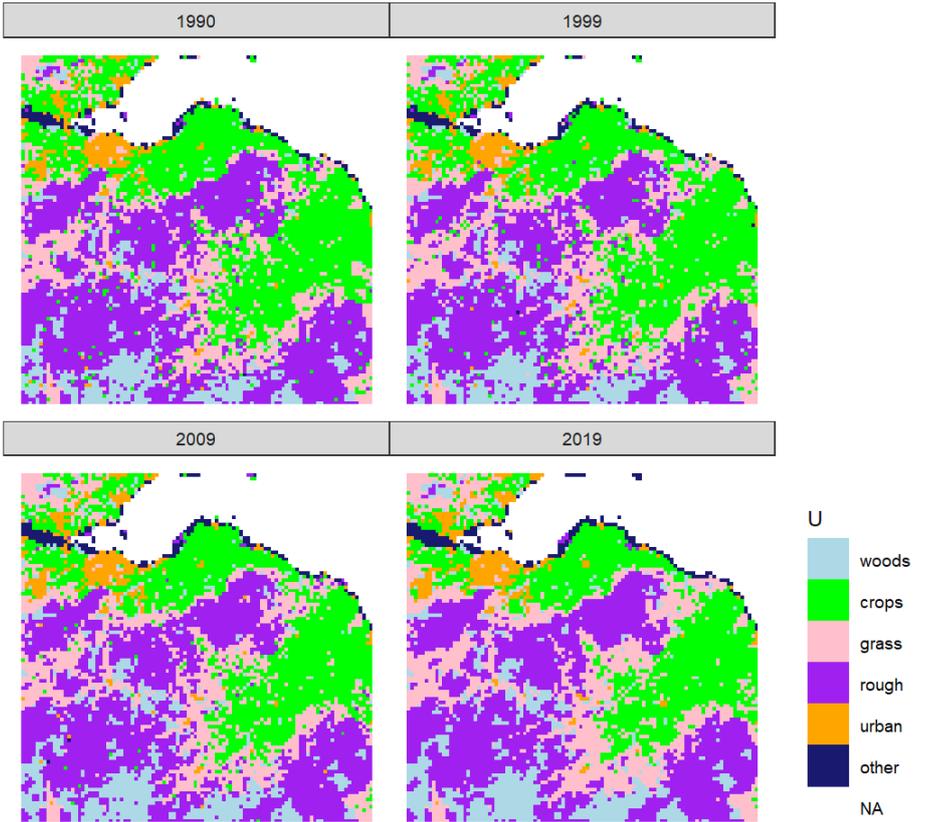


Figure 8.11: Estimated state of land-use U in 1990, 1999, 2009 and 2019 in the 100 x 100 km square containing Edinburgh.

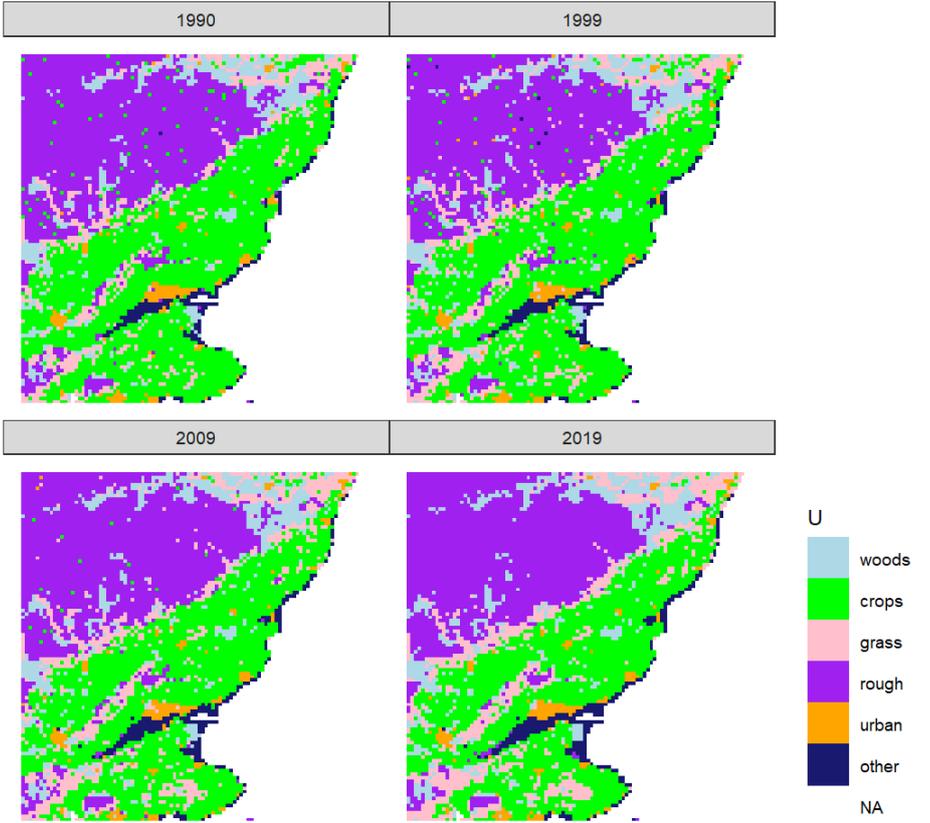


Figure 8.12: Estimated state of land-use U in 1990, 1999, 2009 and 2019 in the 100 x 100 km square centred on Dundee. Figures 8.13, 8.14 and 8.15 show the results in the form of the vectors with greatest area (most frequent).

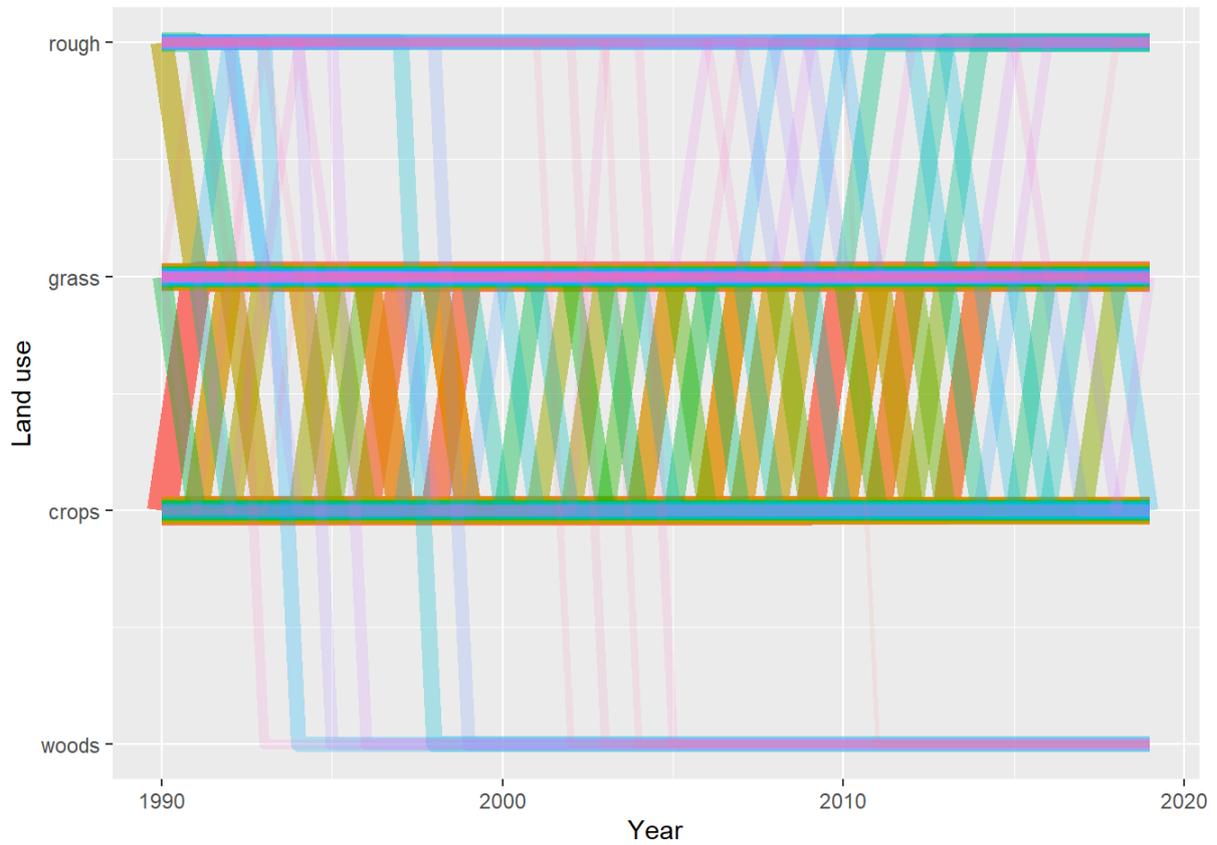


Figure 8.13: Trajectories of the 100 land-use vectors in the posterior U with the largest areas (excluding the six vectors which show no change). Each vector of land use is shown in a different colour, varied arbitrarily to differentiate different vectors. Line thickness and opacity are proportional to the frequency of (or total area occupied by) each vector, so that the dominant vectors are the most visually obvious.

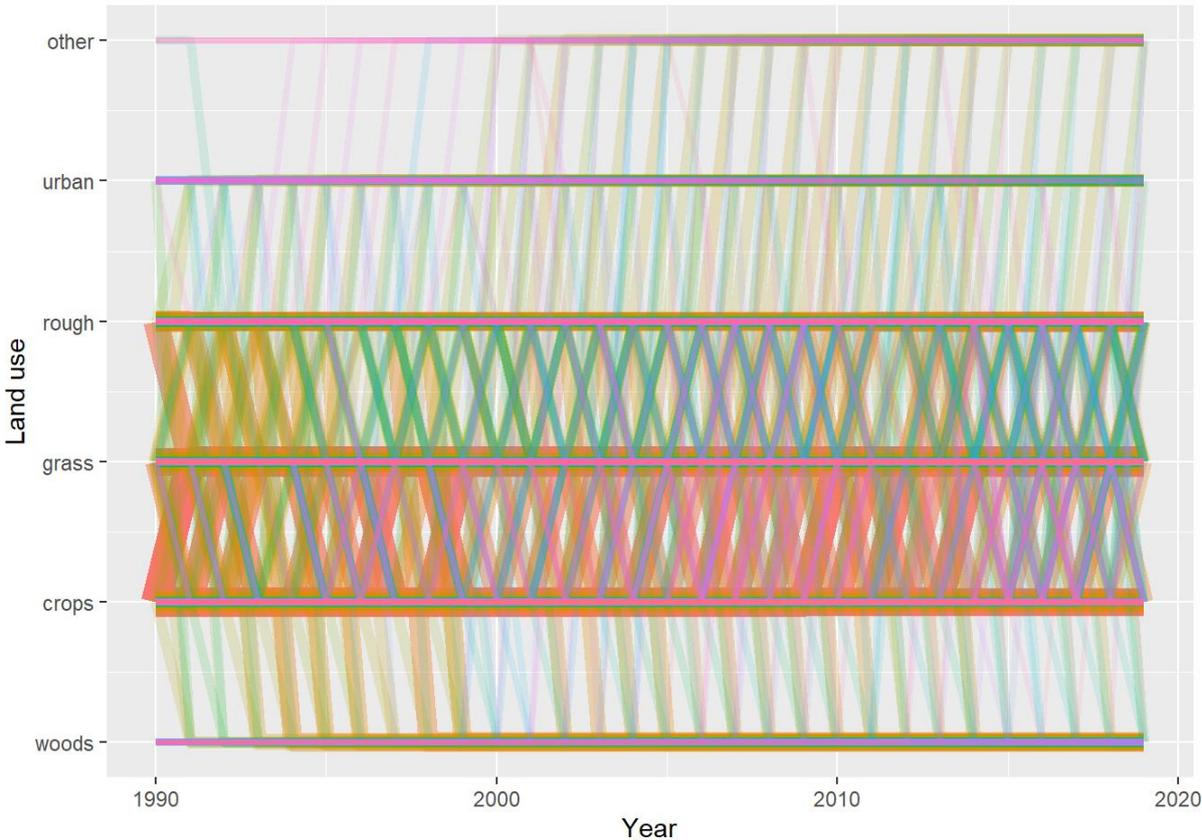


Figure 8.14: Trajectories of the 1000 land-use vectors in the posterior U with the largest areas (excluding the six vectors which show no change). Each vector of land use is shown in a different colour, varied arbitrarily to differentiate different vectors. Line thickness and opacity are proportional to the frequency of (or total area occupied by) each vector, so that the dominant vectors are the most visually obvious.

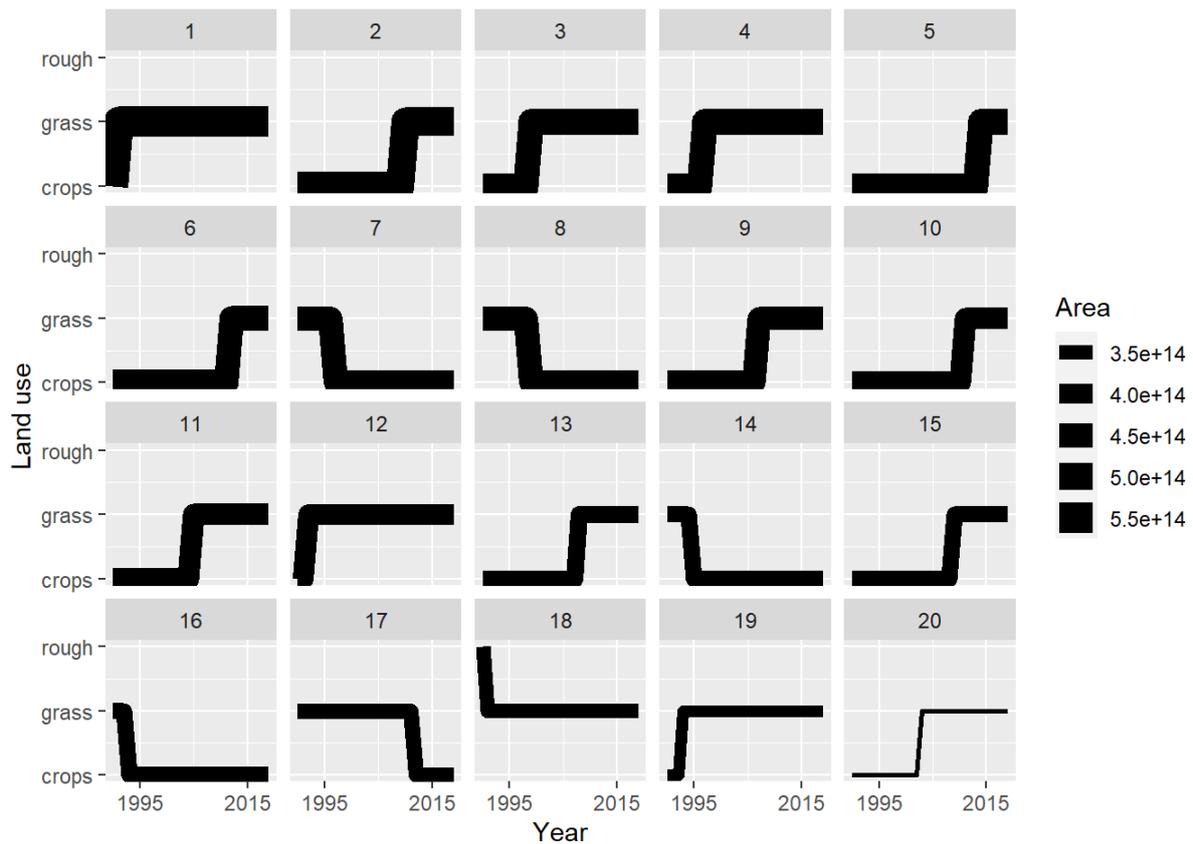


Figure 8.15: Trajectories of the 20 land-use vectors in the posterior U with the largest areas (excluding the six vectors which show no change). Line thickness is proportional to the frequency of (or total area occupied by) the vector

8.4 Discussion

Despite reducing the dimensions, the land-use vectors are still a high-dimensional structure to visualise. The plots in Figures 8.13, 8.14 and 8.15 show two ways to visualise the data. Attempts to make visualisation easier generally comes at the cost of presenting only a restricted amount of the total information. Some time spent exploring other methods for visualisation would be worthwhile. The figures show that the dominant vectors contain the transitions between crop & grass and vice versa. There are no rotational changes in the top 20 vectors (Figure 8.15) i.e. involving changes back and forth at the same location. The frequency of such changes, and whether this matches expectations, remains to be examined. Figures 8.13, 8.14 and 8.15 show only a single set of vectors, based on the MAP B matrix. In fact, we have 2000 samples from the whole distribution, but showing the uncertainty in data like this, with categorical and continuous features, is particularly difficult.

9 Discussion

The results show that we can provide improved estimates of past land-use change using multiple data sources in the Bayesian framework. The main advantage of the approach is that it provides a coherent, generalised framework for combining multiple disparate sources of data. Recalling our original aim of making spatially- and temporally-explicit estimates of land-use change in the UK, using multiple sources of data, we conclude that this aim has largely been met. We have added new data beyond that

which is used in the present inventory method: B is now jointly determined by the annual national-scale Agricultural Census data, rather than just decadal Countryside Survey data; the spatial pattern and resulting vectors are constrained by the high-resolution information contained in IACS, the holdings-level Agricultural Census data, LCM, LCC and CORINE (deliverables A.1). The data have been assimilated so as to use the information from sources which we believe to reliably represent the absolute magnitude of change, together with data sets which provide spatial information. This results in a time series of maps describing our best estimate of land-use change given the available data (deliverable A.1). Furthermore, rather than producing a single time series of maps, we have a set of these which represents the posterior distribution of the data cube U . This distribution quantifies the joint probability distribution of the parameters, and properly propagates the uncertainty from input data to final output (deliverable A.2). The resulting data cube U has been summarised in vector format (deliverable A.1, e.g. Figure 8.13). The code is documented, version-controlled and made available via [GitHub](#) (deliverable A.3).

We suggest the work described here represents an advance on the existing GHGI methodology, but there remain a number of improvements and issues to resolve before this might be made operational. We outline some of the more important ones here.

Possibly most importantly, we should attempt to characterise uncertainty in the different data sources. The Bayesian approach is quite capable of incorporating different uncertainties for different data sources, and indeed, at the level of the individual observation; or in more complex ways, such as accounting for changes in observation uncertainty over time (as we suspect happens in IACS). However, to do this in a rigorous and quantitative way can be quite difficult, and at present we represent all observations to be equal. Some advance on this state at least should be possible. This could involve some more detailed analysis of the data sets we have already. For example, a thorough comparison of the IACS data with the holdings-level Agricultural Census data would be informative. Good agreement would give us some confidence that these are reflecting reality rather than artefacts of the reporting methodology; poor agreement would necessarily raise doubts. The same comparisons could be made between other data sets, such as ground-based and EO-derived data.

Most usefully, we should attempt to introduce independent constraints. For example, as discussed earlier, using data on the area and age structure of temporary grassland would constrain our estimates of the area changing between crop and grass each year. Other forestry data sets, such as annual timber harvest volumes, felling license data, and Woodland Grant Scheme planting data could act as constraints in a similar way.

The CS data was used in the form that it is used in the current GHGI - as a constant annual value over the period between surveys. However, this value is not actually constant; we know the sample mean value over the inter-survey period, but not the annual variability. We should change the way this is represented accordingly. This would have the effect of increasing the uncertainty in the values for individual years, and permitting greater year-to-year variation, as seen in AgCensus.

As discussed in the section on the Agricultural Census, there are some apparent step changes in these data, with no known cause. Our approach of smoothing these out where necessary is effective, but subjective. As the Agricultural Census data play an important part, as the longest running data set with the widest coverage, some effort spent in understanding these would be worthwhile. Whether we apply more or less smoothing to the data is open to debate, and we should understand the sensitivity of the final output to this.

The present project was carried out within a rather short time frame, and the process of obtaining data, and the permission to use it, can be slow. As a result, at the time of completion, we are still lacking permission to use some important data sets that we would want to include (notably IACS and holdings-level data), even though we are already in possession of the data files. The process of requiring the permissions requires significant staff time to establish and maintain communication with

personnel in the numerous different government departments who hold the data and needs some understanding of the legalities of data sharing. This would be an area to prioritise in any future work.

So far, we have pooled all the data for the UK, and allowed the spatial information in the second part of the procedure to differentiate between the four DAs which make up the UK. This approach is certainly more efficient and simpler in terms of programming and data processing. However, it is possible that more accurate results could be achieved by applying both parts of the DA procedure to each of the four DAs separately. This is not essential and would mean that we have to replicate the whole process four times, but merits consideration. Alternatively, if we apply the GHGI inventory model spatially, then this might actually simplify the process, at the expense of larger data volumes.

We have used only two different assumptions about the prior distribution of the B parameters. Preliminary analysis suggested that there was not much sensitivity to this. However, we suggest this is worth further investigation, and a wider range of priors examined.

A potential issue is the assumption of independence of errors. The consequence of assuming non-independence of errors would be to produce unreasonably small uncertainties in the posterior parameters, and that does not seem to be the case here. However, we need to be careful when selecting data sources that these are truly independent estimates and are not double-counting the same data. A much larger source of uncertainty is probably associated with classification. Depending on definitions used to delimit land-use classes, and the consistency of applying these definitions, quite different areas may be calculated for the same nominal classes. There is a real potential problem in combining data from different sources in that we may not be comparing like with like. Here, we reduce this problem by using a relatively coarse land-use classification, with only six classes. This would become more problematic if attempting to distinguish more refined classes.

Whilst the code has good internal documentation (in comment lines, and text within the Rmarkdown files), a document to act specifically as a user manual is still lacking. We consider most parts of the code to be reasonably robustly tested. However, it is likely that there is an artefact of the order of iteration in the importance sampling loop described in section 20. The land-use classes are assigned an arbitrary integer 1-6, and this is determining the order in which land-use changes are computed. We suspect this has only a minor effect, if any. A simple solution would be to randomise the order of iteration every time, so as to remove any bias. This might be an unnecessary complication if not needed, but this should be confirmed.

It is reasonably straightforward to display the uncertainty associated with the final output, the carbon flux from land-use change. This is a time series of a continuous variable, and can be plotted as a line with a shaded band representing 95 % CI limits (Levy et al. 2018). Summarising and visualising the uncertainty in categorical data like land use itself, and its change over time, is much harder. The plots in Figures 8.13, 8.14 and 8.15 can summarise the data, but they do not capture the uncertainty, and some thought needs to be given to this.

Our method here introduces much larger data volumes than the existing method, and a dependence on some level of high-performance computing. The ramifications of this, the potential risks, and the benefits, need to be weighed up. However, the trends in availability and ease-of-use of HPC means that this limitation will diminish over time; low-cost commercial cloud computing services are becoming widely available, such as Amazon Web Services and Google Earth Engine. While it is freely available, the NERC JASMIN facility is ideal for our purposes at present, but there is no particular dependence on this.

The remit of this project was to provide land-use maps and vectors for the GHGI. Because these data are not yet used in the current LULUCF methodology, no testing of the suitability of our output data has taken place. Before taking this work further, it would be wise to consider how the data will actually be used in a new LULUCF methodology. At a minimum, it should be checked that the vector data meets the needs of the soil carbon, land management and CARBINE models. We have previously

demonstrated using the first of these (Levy et al. 2018), but this used a model of minimal complexity, without some of the complexities that are introduced in the operational inventory model. The CARBINE model run by Forest Research uses closely related input data and we should consider how these two activities should relate to one another. Some planning ahead for this would help focus what is needed in future.

10 References

- Braak, Cajo J. F. ter, and Jasper A. Vrugt. 2008. "Differential Evolution Markov Chain with Snooker Updater and Fewer Chains." *Statistics and Computing* 18 (4): 435–46. <https://doi.org/10.1007/s11222-008-9104-9>.
- Byrd, Richard H., Peihuang Lu, Jorge Nocedal, and Ciyou Zhu. 1995. "A Limited Memory Algorithm for Bound Constrained Optimization." *SIAM Journal on Scientific Computing* 16 (5): 1190–1208. <https://doi.org/10.1137/0916069>.
- "CORINE Land Cover — Copernicus Land Monitoring Service." n.d. Land Section. Accessed November 13, 2020. <https://land.copernicus.eu/pan-european/corine-land-cover>.
- DAERA. 2016. "Northern Ireland Countryside Survey | Department of Agriculture, Environment and Rural Affairs." DAERA. November 17, 2016. <https://www.daera-ni.gov.uk/articles/northern-ireland-countryside-survey>.
- Forestry Commission. 2020. "Forestry Commission Open Data." 2020. <https://data-forestry.opendata.arcgis.com/>.
- Hartig, Florian, Justin M. Calabrese, Björn Reineking, Thorsten Wiegand, and Andreas Huth. 2011. "Statistical Inference for Stochastic Simulation Models – Theory and Application." *Ecology Letters* 14 (8): 816–27. <https://doi.org/10.1111/j.1461-0248.2011.01640.x>.
- Hartig, Florian, Francesco Minunno, and Stefan Paul. 2017. "BayesianTools: General-Purpose MCMC and SMC Samplers and Tools for Bayesian Statistics."
- Levy, P., M. van Oijen, G. Buys, and S. Tomlinson. 2018. "Estimation of Gross Land-Use Change and Its Uncertainty Using a Bayesian Data Assimilation Approach." *Biogeosciences* 15 (5): 1497–1513. <https://doi.org/10.5194/bg-15-1497-2018>.
- Oijen, Marcel van. 2020. *Bayesian Compendium*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-55897-0>.
- UKCEH. 2016. "UKCEH Land Cover® Plus: Crops." UK Centre for Ecology & Hydrology. January 25, 2016. <https://www.ceh.ac.uk/crops2015>.
- . 2020a. "Countryside Survey." EIDC. 2020. <https://catalogue.ceh.ac.uk/id/2069de82-619d-4751-9904-aec8500d07e6>.
- . 2020b. "Land Cover Maps." EIDC. 2020. <https://catalogue.ceh.ac.uk/id/c0078881-7d5a-4641-91e2-c271426bc8a1>.
- Vrugt, Jasper A., James M. Hyman, Bruce A. Robinson, Dave Higdon, Cajo J. F. Ter Braak, and Cees G. H. Diks. 2008. "Accelerating Markov Chain Monte Carlo Simulation by Differential Evolution with Self-Adaptive Randomized Subspace Sampling." *International Journal of Nonlinear Sciences and Numerical Simulation*, nos. LA-UR-08-07126; LA-UR-08-7126 (January). <https://www.osti.gov/biblio/960766-accelerating-markov-chain-monte-carlo-simulation-differential-evolution-self-adaptive-randomized-subspace-sampling>.

WP-A QA REPORT

WP-A QA Report: Model Quality Assurance and Comparison to Approaches used by Other Countries

Authors: Gwen Buys ¹, Christopher Evangelides ² and John Watterson ²

¹ UK Centre for Ecology & Hydrology, ² Ricardo

Date: 13/11/2020

1 Model QA Log for LUC Track

The main QA activity carried out in WPA was an assessment of the data processing and model code¹ (known as the `luc_track` model) using the BEIS model QA log for non-Excel models². The first assessment was made shortly after project initiation, hence allowing for areas for improvement to be identified and worked on during the remainder of WPA. The log was then updated in the final week of the project and the following synopsis describes the assessment of the model at that point.

The model log has been completed by the project QA manager Gwen Buys using the guidance provided alongside the log. The log is split into five main sections and requires the user to assess the model against tasks within each section, rating the model between 1 = Excellent and 5 = Significant Issues for each task. The individual scores are then weighted and summed to provide a final model score as a percentage. When assessing the `luc_track` model it was kept in mind that it is a UKCEH model (as opposed to an internal BEIS model) hence some of the task requirements and guidance needed interpretation in terms of relevance to third party models.

1.1 Documentation

The model QA log has ten documentation tasks spanning both model and project management documentation and the current rating for eight of these is 1 or 2. The `luc_track` model is coded in RMarkdown and stored in a GitHub repository resulting in the model scoring well in terms of version control, transparency and guidance. Governance, responsibilities and continuity planning are covered by UKCEH project management systems and the relevant risk assessment and data management planning (DMP) documents have been completed for this project. The tasks for scope and specification documents have currently been rated as 5 as the BEIS templates for these two documents have not been completed for the `luc_track` model. However, in practice much of the required content for these documents already exists within the RMarkdown, the project tender and the Levy et al 2018³ paper. Creating official scope and specification documents for `luc_track` could form part of the QA work in WPB if this is a priority for the BEIS project manager.

1.2 Structure & Clarity

There are four tasks relating to model structure and clarity and the `luc_track` model scores 2 for all of these. The model has been clearly structured and has very consistent (and well described) naming convention throughout the code. Further improvements to clarity, simplicity, comments and the creation of a model map could be made during WPB if this is needed.

1.3 Verification

The model QA log has four tasks relating to model verification, however the guidance for one of these, Inputs and outputs formula correctness, seems to relate only to Excel models so the `luc_track` model has been assessed against the remaining three.

Usability testing and code correctness & testing have been carried out by project partners Ricardo for a subset of the `luc_track` code. They found no errors in cloning the GitHub repository / reading the files and the process was efficient. They commented that the code is well documented and a new

¹ https://github.com/NERC-CEH/luc_track

² <https://www.gov.uk/government/publications/model-quality-assurance-full-log-template-non-excel-models>

³ <https://bg.copernicus.org/articles/15/1497/2018/>

user can easily understand library dependencies and error handling. Hence, these tasks were scored at 2 and 3 respectively and the main recommendations are to check compatibility of the spCEH package (an R package developed by UKCEH) with more recent versions of R and to expand testing to more of the luc_track code. These tasks can form part of the QA work during WPB.

Autochecks & regression testing have been scored as 2. There are checks within the model (e.g. consistent / compatible units) which stop the model running if incompatible choices are made. Additionally there is a system set up to use an alternative dataset (chess games) to test functionality of key aspects of the model.

1.4 Validation

Model validation has five tasks in the BEIS QA log covering methodology correctness and extreme value, uncertainty and re-performance testing. There is also a task for comparison with historical data but this was assessed as not appropriate for the luc_track model as the intention of the model is to utilise historical data to create a timeseries of land use change.

Methodology correctness has been scored as 1 and the three remaining tasks have been scored as 2. The method used in the luc_track code is expanded on from the peer reviewed Levy et al 2018 paper which demonstrates that the model is fit for purpose and produces logical outputs. Previous work (in another project) has been carried out to investigate sensitivity of parameter weighting in the model and further sensitivity is built into the model as each run produces 1000+ iterations of output. In terms of extreme value testing there are some restrictions coded into the model to effectively remove extreme values (including negative land use change areas) and to reduce the influence of unlikely land use change areas.

Re-performance testing has been carried out by Ricardo for a subset of the luc_track code. The work focused on independently testing the mapping from land cover classes to LULUCF classes and was carried out in ArcGIS Pro for three test areas. The tests were successful and found that the mapping of classes was correct for all three test areas. The main recommendation for improving model validation in WPB would be to expand re-performance testing to further sections of the luc_track model.

1.5 Data & Assumptions

There are three QA tasks relating data and assumptions and data transformation. The luc_track model scores 1 for data transformation as the RMarkdown code and the project DMP thoroughly describe the data sources used in the model and how they are formatted and re-classified for use in the model. The tasks for both data and assumptions have been scored 3 as they both reference the use of the BEIS assumptions log template. This has not yet been completed for the luc_track model but the majority of the information needed for the template is already described in the RMarkdown and DMP. If it is a priority for BEIS then an assumptions log can be created for luc_track during WPB.

2 Comparison to Approaches Used by Other Countries

John Watterson, Ricardo

2.1 Approach

BEIS require a “*Validation of research methods as fit for purpose by review of the land-use change mapping approaches used by other countries*”.

In our proposal, we suggested we would conduct two main pieces of work to help meet this requirement:

1) Contact key people, to canvas their opinions.

This including contacting one of the Lead Authors, **Marcelo Rocha**, of the remote sensing chapter in the IPCC 2019 Refinement for specific views on “best practice” of using RS/EO data in LULUCF inventories.

We also suggested having bilateral discussions with: Germany, Sweden, Finland, Canada and Australia. Canada is particularly interesting as they have extensive use of satellite data to generate a detailed mosaic of land use categories.

Pierre Brender (BEIS LULUCF team) recommended contacting:

Paulo Canaveira, Portugal. Paulo presented at a land use meeting in 2019, showing how land use is represented in Portugal and how consistent full time series are compiled while avoiding false land-use changes. Paulo also acts as the Lead Negotiator on LULUCF for the European Union.

Eric Arts, Netherlands. Eric presented at the JRC workshop in 2019 to explain how the inventory team avoided some potential pitfalls in their spatially explicit representation of land-use change⁴.

2) Review relevant recent literature and conferences.

We suggested a review of relevant recent literature to increase our understanding of “best practice” and barriers and solutions to using RS/EO data. Two examples of material that we suggested should be examined are: **a)** The “*Analysis of LULUCF actions in EU Member States as reported under Art. 10 of the LULUCF Decision*” – containing an assessment of EO/RS data by Member States (Paquel, K.; Bowyer, C.; Allen, B.; Nesbit, M.; Martineau, H.; Lesschen, JP.; Arets, E. (2017), Analysis of LULUCF actions in EU Member States as reported under Art. 10 of the LULUCF Decision, a report for DG CLIMA of the EC, and, **b)** The findings of the JRC conference “*Towards spatially explicit land representation in LULUCF inventories Experiences from capacity building activities*”, EC and JRC.

This short report summarises the work done to date. The page limit for the proposal was very tight, and in the proposal, the team were not able to elaborate how task to conduct the “validation of research methods” would be handled in detail. Work Package A has a very modest budget for this task, and therefore only certain elements of the work has been done to date. If Work Package B is commissioned, we will be able to complete this task in much more detail.

2.2 Responses from Key People

This section summarises the responses received as of the 11th November, 2020.

2.2.1 Marcelo Rocha

The IPCC 2019 Refinement, Lead Author

⁴ https://forest.jrc.ec.europa.eu/media/filer_public/88/3d/883d9efe-f824-4515-a9af-a21770fd4651/16_e_aret_spatial_explicit_approach_in_nl.pdf

Volume 4, Chapter 3 “*Consistent representation of lands*” of the IPCC 2019 Refinement <https://www.ipcc-nggip.iges.or.jp/public/2019rf/vol4.html> has a range of useful material about the use of remote sensing and satellite data. Some information is presented in the main body of the chapter, and in Annexes.

In the main body of the Refinement there is useful guidance on good quality approaches to:

3.3.4 Combining Multiple Data Sources: *Typically, countries will combine a variety of different data sources and approaches to estimate areas of land-use*

3.3.5 Derivation of IPCC Land-Use Categories from Land Cover Information: *Addressing gaps in remote sensing data*

3.5 Uncertainties Associated with the Approaches: *Evaluation of land-use and land-use change information generated from remote sensing techniques and estimation of uncertainties; Collection of validation data; Evaluation of sample-based method; Evaluation of wall-to-wall methods*

In the Annex of the Refinement, **Annex 3A.2.4 Tools for data collection Remote sensing (RS) techniques**:

This Annex summarises the range of RS techniques, including recent developments. There is some material setting out what we might judge to be “best practice” approaches to using EO/RS data. This material is presented in sections on:

Time-series consistency (cross references also to Chapter 5 of Volume 1: Time-Series Consistency of the 2006 GLs)

Ground reference data

Integration of remote sensing and geographical information systems

Land-Use classification using remote sensing data (discusses the application image classification algorithms)

Detection of land-use conversion using remote sensing

Time-series classification

Evaluation of mapping accuracy

Ground-Based Surveys (discusses how these surveys may be used to gather and record information on land-use, and for use as independent ground-truth data for remote sensing classification).

Assessment: There is no “single solution” presented in the 2019 Refinement to ensure that the research EO/RS methods the UK uses are fit for purpose. The approach needed will be to use a range of the techniques described in the Refinement, and then document their use transparently in the UK National Inventory report.

2.2.2 Paulo Canaveira, Portugal

No response was received from Paulo after two e-mail requests.

2.2.3 Eric Arts, Netherlands

The maps that the Netherlands use as a basis of their spatially explicit monitoring and reporting for LULUCF are based on a series of base-maps that were originally created for monitoring nature developments. These maps are based on the national topographic maps, which in turn, are based on aerial photographs. An overall description on how land-use is represented and how the land-use change matrix of the Netherlands created is included in chapter 3 of the methodological background report for LULUCF: <https://edepot.wur.nl/517340>.

In terms of QA/QC the initial map interpretation and first quality control is covered by the topographic service. Then in a next step the categories on the topographic maps are further aggregated to a number of (LULUCF) relevant land-use classes and the whole map is gridded (see for instance Kramer et al 2013: <http://edepot.wur.nl/356218>). An English description is only available for the 2009 map: <https://edepot.wur.nl/14207>. Subsequently the land-use change matrix resulting from comparing this map with a previous map is used to assess any strange and counterintuitive changes. These then are further assessed and verified by checking a large number of polygons representing those changes with the underlying aerial photographs. This is what was done after the observation on the 2017 map that net forest area had decreased compared to the 2013 map, which Eric addressed in the JRC presentation; see p.24 in the methodological background report.

Additionally a point-validation is carried out on at least 600 points by comparing the classification on the map with aerial photographs.

Assessment: The Netherlands take a multi-step approach to generating the maps of land use and land use change. Several organisations are involved. Quality control and verification procedures are completed by both the topographic service and GHG LULUCF sector experts. The work and effort is considerable, and involves a lot of manual analysis of maps and photographs.

2.3 Review Relevant Recent Literature and Conferences

This section of the report presents a simple assessment of a recent report and a conference.

2.3.1 Analysis of LULUCF actions in EU Member States as reported under Art. 10 of the LULUCF Decision

Aim of the Study

The aim of this study⁵ was to support the European Commission in the implementation of the LULUCF Decision (no 529/2013) with a focus on its Article 10 provisions. Under Article 10 Member States must submit information on their most relevant current and future LULUCF actions in land use activities such as afforestation, forest management, cropland and grassland management, and wetlands management. The study analyses the initial and progress LULUCF action reports submitted to the European Commission between 2014 and 2017.

Methods

⁵ “Analysis of LULUCF actions in EU Member States as reported under Art. 10 of the LULUCF Decision” – containing an assessment of EO/RS data by Member States (Paquel, K.; Bowyer, C.; Allen, B.; Nesbit, M.; Martineau, H.; Lesschen, JP.; Arets, E. (2017), Analysis of LULUCF actions in EU Member States as reported under Art. 10 of the LULUCF Decision, a report for DG CLIMA of the EC <https://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupMeetingDoc&docid=10585> UKCEH report ... version 1.0

Between May and September 2017, the study team analysed 51 Article 10 reports. The information provided in the reports was first synthesised based on a set of indicators such as scope, goal, planning period, link to national priorities, type of policy instrument, sources of funding, expected impact, and data sources. The synthesis results were then analysed in order to identify the most often reported actions, the policy instruments used to support the actions, the LULUCF priorities shared by the Member States, and the most cost-effective measures. The analysis also tried to find out how widespread the use of spatially explicit data in LULUCF accounting is among the Member States, and how the methodologies for estimating GHG emission could be improved to ensure more accurate results. One of the study tasks was also to try to provide an estimate of the aggregated impact of the activities on GHG emissions, and compare it with relevant findings from other studies in this field. Finally, the analysis was also oriented to those LULUCF actions that could be enhanced to maximize the pursuit of their mitigation potential. The analysis built not only on the Article 10 reports, but also on literature review and expert judgment.

Synthesis findings - Methods used to determine land use and GHG impacts

Assessment: The main relevant observation for our work is that there is no specific mention of quality control or verification measures in the sections of the report looking at the methods used to determine land use and GHG impacts. This is presumably because the authors of the study were both not specifically looking for information about these QC or verification measures, and, the Article 10 reports do not present this level of methodological detail.

Although the Article 10 reports do not include much information on the methodologies used to determine land use or GHG emission and removals relevant to the reported LULUCF actions, analysis based on other recent studies shows a high potential for improvement in both areas. It seems that an improvement of land use data availability and accuracy is possible and could lead to a better quality monitoring of LULUCF activities also in terms of their GHG impacts. The improvement could be done by exploiting the potential to complement the various existing data sets such as Land Parcel Identification System (LPIS) deployed widely under CAP, Eurostat's Land Use and Land Cover Survey (LUCAS) and Copernicus, the EU Earth Observation programme, offering state of the art quality of land use data.

2.3.2 Towards Spatially Explicit Land Representation in LULUCF Inventories

The aim of this conference⁶ was to share experience from capacity building activities.

The relevant presentation⁷ to this study was given by Eric Arets, of the Netherlands LULUCF inventory team, based at Wageningen University. The presentation considered the challenges of the use of land use maps in the Netherlands. It uses an example the changes in deforestation estimated between 2013 and 2017, and examines the reasons for the decrease in net forest area based on the 2013 map. The approach used was to re-examine the analysis based on the aerial photos – by visual inspection – and to check the magnitude of the changes seen.

The analysis found that some of the changes in deforestation between 2013 and 2017 were correct, and were due to: temporary forest; conversion to other nature areas; real deforestation (but before 2013).

⁶ The findings of the JRC conference “Towards spatially explicit land representation in LULUCF inventories Experiences from capacity building activities”, EC and JRC. 29th May, 2019

<https://forest.jrc.ec.europa.eu/en/activities/lulucf/workshops/workshop-2019/>

⁷ https://forest.jrc.ec.europa.eu/media/filer_public/88/3d/883d9efe-f824-4515-a9af-a21770fd4651/16_e_aret_spatial_explicit_approach_in_nl.pdf

Some of the changes had been ascribed incorrectly, and were due to artefacts including: border effects; border along road/"bike road"; was misclassified as forest before; wood harvesting occurring; high voltage lines.

The reasons for some differences between 2013 and 2017 could not be determined.

In some cases, part of a polygon was correct, but part was incorrect.

The conclusions of the presentation were: mapping land use is not very difficult; comparing changes in total areas of different land uses is also doable; spatially explicit monitoring of land-use changes over time according approach 3 is very challenging.

Assessment: It is likely that in LULUCF GHG inventory land use maps there will be both misclassification of land use, and the size of land use change between two time periods. It is possible to improve the quality of the land use maps, but at least in part, it is likely that this verification will need to be done manually and will take time. It is also likely that the existing algorithms used to identify land use and land use change could be improved. In part, these improvements will be based on an understanding of the likely reasons for the errors which will be revealed by the verification studies.

WORK PACKAGE A FINAL RESULTS REPORT

WP-A Final Outputs Report

Authors: Peter Levy, Sam Tomlinson and Beth Raine

UK Centre for Ecology & Hydrology

Date: 10/11/2021

Work Package A Final Results: Executive Summary

This report describes work on the project “Improving Land Use Change Tracking in the UK Greenhouse Gas Inventory” for the Department for Business, Energy & Industrial Strategy (reference TRN 2384/05/2020). The aim of the project was to make improved estimates of land-use change in the UK, using multiple sources of data. We applied a method for estimating land-use change using a Bayesian data assimilation approach. This allows us to constrain estimates of gross land-use change with national-scale census data, whilst retaining the detailed information available from several other sources. Previous reports covered work with existing data sets (WP-A) and on developing new data sets (WP-B). This report describes subsequent work, focussed on quantifying uncertainties and improving the representation of crop-grass rotation.

The random uncertainty term σ for each data source has been estimated, based on agreement with the net change observed in a reference data set, comprising those data sets we believe to be most reliable for net change. The random uncertainty is now represented with both absolute and relative components, as a linear function with intercept and slope terms.

Systematic uncertainties represent biases in the data, and can be characterised as the false positive and false negative rates (i.e. detecting apparent change where none really occurs, and failing to detect actual change, respectively). We compared three different methods for estimating these rates, which gave similar results. These showed very high false positive rates, which mean that large corrections have to be applied to the data. The data assimilation algorithm now includes terms to account for: data-source-specific random uncertainty (σ); data-source-specific systematic uncertainty (false positive and false negative rates); sampling error due to low survey frequency (e.g. decadal cf. annual); and uncertainty in extrapolation outwith (pre/post) survey time period.

Additional data sets, which were not previously available have also been included in the analysis. All the holdings-level agricultural census data available in the UK from 1990 to 2019 have now been included. The CROME data set, covering crops in England, has been added.

We have improved the accuracy of how we represent crop-grass rotations, using the idea of “life tables” or age-specific transition probabilities, from population modelling. The life table probabilities are based on an analysis of the data available from IACS, CROME, LCM, and LCC, and these all show consistent patterns. Of these, IACS has much the longest time span of data, and is the dominant source of information. Using this method, the observed frequencies of transitions from crop to grassland as a function of cropland age (and v.v. for grassland) are now reproduced in the land-use vectors, and thereby approximate the observed frequency of crop-grass rotations.

Since land-use change is now represented spatially, we can separately identify that occurring on mineral soils and organic soils. In summary form, land-use vectors and the matrices of land-use change are now provided separately for mineral and organic soils for each DA.

The land-use change output data are provided as space-time data cubes at 100-m and 1000-m resolution for the period 1950-2020. These are also provided in summary (non-spatial) form as the set of unique land-use vectors, and the matrices of land-use change. In principle, many thousand samples of the posterior distribution are available (currently 20000 samples of the matrices are stored per DA). Many samples of the mapped and vector data can be provided from these, given practical computing and storage constraints, but the optimal way to produce these depends on how they are used in the LULUCF inventory work, and this is open to discussion. Storing land use as a 71-digit character string at 91 million locations (for the time period 1950-2020 at 100-m resolution) requires ~50 MB, or ~5MB in vector summary form, and much less at 1-km resolution.

The workflow for the project uses a “Make”-like paradigm to maintain a reproducible workflow, implemented in the R “targets” package (<https://books.ropensci.org/targets/>). This means that the dependencies among functions and input data are analysed dynamically and stored in a hash table. Any changes to source code functions or input data are detected automatically, and only the invalidated components re-run as necessary. This has several advantages: forcing the workflow to be declared at a higher level of abstraction; only running the necessary computation, so saving run-time for tasks that are already up to date; and most importantly, providing tangible evidence that the results match the underlying code and data, and confirm the computation is reproducible. All the code is under version control on GitHub.

MCMC algorithms typically have to run for many iterations to produce an acceptable sample of the posterior, so that parameter estimates (the Beta matrices) are robust. For each year and each DA, the algorithm was run with nine MCMC chains for 200,000 iterations. Convergence was assessed by a standard method - comparing within-chain variance with among-chain variance via the Gelman-Rubin statistic. As chains move towards convergence, the value of this tends towards 1, and the standard test requires that this value is less than 1.1. This criterion is well met with this number of iterations. It is standard to discard “burn-in” samples prior to convergence, and to thin chains by a factor of ten to remove any autocorrelation. This yields around 16000 posterior Beta samples, which gives potentially the same number of posterior samples of the land-use vectors. However, it may be more practical to combine the calculation of fluxes into this code, without storing all the underlying data.

We produced a time series of maps describing our best estimate of land-use change given the available data, as well as the full posterior distribution of this space-time data cube. This quantifies the joint probability distribution of the parameters, and properly propagates the uncertainty from input data to final output. The output data has been summarised in the form of land-use vectors. The results show that we can provide improved estimates of past land-use change using this method. The main advantage of the approach is that it provides a coherent, generalised framework for combining multiple disparate sources of data, and adding additional sources of data is straightforward.

1 WP-A Introduction

This report describes work carried out on the project “Improving Land Use Change Tracking in the UK Greenhouse Gas Inventory” for the Department for Business, Energy & Industrial Strategy (reference TRN 2384/05/2020). The aim of the project was to make improved estimates of land-use change in the UK, using multiple sources of data, using a Bayesian data assimilation approach. Two previous reports describe the background to the project and results from Work Package (WP) A (Levy et al. 2020), using pre-existing data sets, and work developing new data sets based on Earth Observation in WP-B (Rowland et al. 2021). This report describes subsequent work which focussed on assessing the uncertainties in data sets and incorporating these in the data assimilation procedure. Specifically, we aimed to:

- quantify random uncertainty in the different data sources;
- quantify systematic uncertainty, i.e. biases which may make a data source under- or over-estimate land-use change;
- represent the uncertainty associated with the different sampling frequencies of the data sources (e.g. annual versus decadal surveys);
- handling the uncertainty associated with the different temporal coverage of the data sources (avoiding step changes when data coverage starts or stops);
- incorporate some additional data sets which were not previously available because of data access or processing time constraints;
- improve the representation of the rotational change between crop and grassland.

The above were incorporated in the data assimilation procedure, and results produced for each of the Devolved Administrations (DAs) of the UK. Land-use change on mineral soil and organic soil was estimated separately in each DA.

The remainder of this report describes these tasks and the resulting estimates of land-use change produced after their inclusion.

The first section on quantifying uncertainty in land-use data sources describes how uncertainty is represented, and estimates random and systematic uncertainty by comparison with a reference data set.

The next section describes an alternative method for systematic uncertainty based on the length of the interval between surveys. As a third method, we then look at how errors in map classification propagate into errors in estimates of land-use change.

Next, we describe a method for more accurately representing the frequency of rotational land-use change, using the idea of “life tables” borrowed from population modelling.

The previous reports give detailed descriptions of the methods. However, here we reproduce the basic rationale and approach of the project for background.

1.1 Tracking land-use change

The tracking of land use and land-use change is fundamental to producing accurate and consistent greenhouse gas inventories (GHGI) for the Land Use, Land-Use

Change and Forestry (LULUCF) sector. This is necessary to meet the international requirements of the Kyoto Protocol to the UN Framework Convention on Climate Change (UNFCCC) and the Paris Agreement and the national requirements of the UK's Climate Change Act and related legislation within the UK's Devolved Administrations.

The estimation of land-use change in the current UK GHGI is based on a combination of infrequent CEH Countryside Surveys and afforestation/deforestation statistics from the GB Forestry Commission. It uses Approach 2 (non-spatial land-use change matrices) as described in the KP Guidance. However, this matrix-based approach, and its implementation in the UK, have some important limitations. Firstly, the non-spatial matrix-based approach is insufficient for tracking annual land-use change: the matrices have no time dimension and are defined independently each year. There is therefore no possibility of representing a sequence of land-use on the same parcel of land (such as afforestation followed by deforestation, or crop-pasture rotations). Secondly, the data used to estimate these matrices in the UK are rather limited. The CEH Countryside Surveys were only carried out approximately once per decade, and whilst the geographical extent was very wide, the actual ground area surveyed was small as a fraction of the total UK area. The afforestation/deforestation statistics from the Forestry Commission have good national coverage (excluding Northern Ireland) but do not contain any information on the spatial location or land use prior to afforestation or following deforestation.

In October 2019, the UNFCCC Expert Review of the UK 1990-2017 GHG inventory raised concerns in relation to the reporting requirements of the second commitment period of the Kyoto protocol. They questioned whether the current approach is appropriate for the identification and tracking of lands where the elected Article 3.4 activities occur (i.e. Cropland Management, Grazing Land Management and Wetland Drainage and Rewetting). They recommended that the UK explore how to make the best possible use of available data and prepare and implement a work-plan to enable the use of these data. The UK has already explored several approaches to land use tracking, including a data assimilation approach to integrate available land-use data into land-use vectors, which was successfully piloted in Scotland (Levy et al. 2018). This project builds on that approach to assess gross land-use change, and land-use history for the whole of the UK from 1990 to 2019.

As well as improving accuracy of the GHGI, a time series of spatially explicit land-use change would enable better tracking of mitigation activities and improve baseline data for scenario modelling. These baseline data are needed for understanding the potential of land-based mitigation and adaptation options. The government's ambitions for Net Zero by 2050 or sooner means that the LULUCF sector will have an increasingly critical role in the UK's overall GHG balance. This kind of scenario modelling will become very important to inform the setting of future carbon budgets and monitor progress towards the UK's legal obligations to GHG emissions reductions. An accurate spatio-temporal land-use change data set would be useful to other stakeholders and UK government departments. For example, from the perspectives of biodiversity conservation, air quality, or ecosystem services, there are clear applications of these data for understanding and tracking the effects of land use.

1.2 Approach

If we had reliable maps of land use each year, we could infer land-use change by difference. However, even with advances in satellite sensors, GIS and spatial data handling, the accuracy of change detection from EO-based products is generally too poor to do this; the different EO products are inconsistent (with each other, and with themselves over time), irregular, and become more infrequent as we go back in time. Change is more reliably detected by repeat ground-based surveys, but these have other short-comings. For example, the annual June Agricultural Census gives a long record of areas in different land uses, but does not provide spatial data, or any information on gross change (i.e. what land uses have changed to which other land uses). The CEH Countryside Survey did provide spatial data with gross change, but without complete coverage, and only at infrequent intervals.

In light of the above, some data assimilation method, which combines the spatio-temporal data with non-spatial repeat survey data, would appear to provide a solution. To this end, we previously developed a methodology using a Bayesian data assimilation approach, and this has been applied successfully to Scotland (Levy et al. 2018). This method allowed for the use of a wider range of data types, including high-resolution spatial data, and combined them in a mathematically coherent way. Importantly, the method produced the appropriate data structure needed for modelling the effects of land-use change on GHG emissions - the set of unique land-use vectors (i.e. unique sequences of land use, or land-use histories) and their associated areas. An important feature is that the uncertainty in land-use change can be easily propagated to provide the uncertainty in GHG emissions, because the procedure explicitly handles the distribution of plausible vectors of land-use change. The approach provides a general framework for combining multiple disparate data sources with a simple model which describes how these data sources inter-relate. This allows us to constrain estimates of gross land-use change with reliable national-scale census data, whilst retaining the detailed spatial information available from several other sources. Here, we apply this methodology to improve and update the tracking of land-use change for the UK. Our aim was to apply a Bayesian approach to make spatially- and temporally-explicit estimates of land-use change in the UK, using multiple sources of data.

All the code is written in R using the “literate programming” paradigm implemented with Rmarkdown, which combines the source code, text/graphical output, documentation, and report text within the same document. This ensures integrity of documentation, code and corresponding outputs. All the Rmarkdown files are held in a GitHub repository, for version control and wider accessibility. The documentation is rendered using bookdown and made publicly available as a [web site via GitHub Pages](#). This documentation describes the data processing workflow so as to make it reproducible.

2 WP-A Quantifying Random Uncertainty in Land-Use Data Sources

2.1 Introduction

Several different data sources provide observations of the transition matrix B and the net and gross changes in area of each land use (D, G, L). The method used in WP-A treated all these data sources as equally uncertain and assumed the same relative error for all observations. However, in reality, we know that these data sets have different levels of uncertainty: some data sets are closer to direct observations, are more plausible, and we have greater faith in these. We want to reflect this in the methodology by quantitatively associating different uncertainties with each data set. This is straightforward in principle, but there are several considerations when doing this in practice:

1. We can consider increasing levels of detail:
 - variable-specific uncertainties (i.e. different for B, G, L & D)
 - data source type-specific uncertainties (i.e. different for ground-based vs EO data)
 - data set-specific uncertainties (i.e. different for CS, IACS, LCM etc.)
 - land-use type-specific uncertainties (i.e. different for woods, crops, grass etc.)
 - time-specific uncertainties (i.e. different for 1990, 2000 ... 2019)
2. Rather than continuous data with a simple σ error term, the B observations are count data in a 6 x 6-way classification. When considering land-use *change*, we can compare the 36 elements of this classification from one data source with another (or the truth), so we have a 36 x 36 error matrix (or “confusion” matrix). Various metrics can be calculated which summarise the agreement measured by this matrix.
3. We can specify uncertainty with greater or lesser rigour: there are several possibilities for how we represent “uncertainty” in the mathematical model.
4. We can estimate uncertainty subjectively or base it more closely on data. There are also several possibilities for how we translate measures of uncertainty in the data into the mathematical model.

A limitation is that none of the data sources represents absolute truth, and we have no definitive data set against which to calibrate.

2.1.1 Representation of random uncertainty

The data sources are assimilated in the Bayesian method via the likelihood function, which includes a term σ^{obs} , representing the standard deviation in the probability

density function for the observation. The observation is thus not assumed to be the true value, but subject to errors which make it deviate from this. Random uncertainty is represented by the magnitude of σ^{obs} - large values of σ^{obs} represent high uncertainty (systematic uncertainty is considered in later Sections). For each observation, a likelihood is calculated, assuming that measurement errors show a Gaussian distribution and are independent of each other. In mathematical notation, the likelihood of observing the area changing from land use i to land use j , β_{ij}^{obs} , is

$$\mathcal{L} = \frac{1}{\sigma_{ij}^{obs} \sqrt{2\pi}} \exp\left(-(\beta_{ij}^{obs} - \beta_{ij}^{pred})^2 / 2\sigma_{ij}^{obs2}\right)$$

Where β_{ij}^{pred} is the corresponding prediction, and σ_{ij}^{obs} is the uncertainty in the observation. There are analogous terms for G, L and D which can all be multiplied.

For example, the term for the likelihood of observing the net change in land use u , D_u^{obs} , is

$$\mathcal{L} = \frac{1}{\sigma_u^{obs} \sqrt{2\pi}} \exp\left(-\left(D_u^{obs} - D_u^{pred}\right)^2 / 2\sigma_u^{obs2}\right)$$

We previously assumed that *relative* measurement uncertainty was the same for all observations, i.e. a constant proportion of the observed value. Thus, observations of large areas come with larger absolute uncertainty. Here, we estimate more specific uncertainties σ^{obs} for the different data sources, and potentially this can be extended to be specific for each individual observation.

Two additional issues concern the specification of the random uncertainty when accounting for the effects of the frequency of surveys and avoiding step changes when data sets begin and end. For example, Countryside Survey data come from approximately decadal surveys, but are interpolated and used as if they were constant annual values within each decade. We incorporated some simple methods to include these effects appropriately.

2.2 Methods

2.2.1 Estimating σ^{obs} from reference data

A pre-cursor step is to define a reference data set with which we compare each data source. In the absence of ideal ground-truth data, some subjectivity is inevitable here, and we use the data sets which we believe to be the most plausible or closest to the truth, based on judgement and prior knowledge. For agricultural land, we defined the reference data set as the June Agricultural Census data for crops, grass and rough grazing, as this is a very long record and is submitted annually as part of reporting to FAO. For forests, the pre-existing record of afforestation and deforestation based on FC statistics was used, as this is also long-running and has been submitted as official national data.

We applied a simple method, basing all the σ^{obs} terms on the “lowest common denominator” data set, the time series of net area change D . Although they vary in the higher levels of detail, all data sources produce estimates of D , so we can calculate a comparable metric of agreement across all data sets. Suitable metrics are the root-mean-square error (RMSE), mean-absolute error (MAE) and the correlation coefficient. We used the RMSE and the correlation coefficient, as measures of absolute and relative agreement, respectively. We multiplied these (rescaling the correlation coefficient as $1/(r^2+1)$) to give a single scaling factor for σ^{obs} (“CV” in the table). This metric is used as a scaling factor in estimating σ^{obs} , such that data sources with poor agreement receive high σ^{obs} (high uncertainty). Some subjectivity comes into which metrics to use, and the absolute values of σ^{obs} , but the relative uncertainties (and therefore weighting) are based quantitatively on data.

2.2.2 Effects of survey frequency and survey start/stop dates

Each pair of surveys gives an estimate of land-use change over some time interval. When survey observations are not available annually, we are effectively trying to estimate the change in a population (all years) from a sample (the years when surveys were carried out). Using the analogy with conventional sampling theory, the standard error in our estimate should be lower in more frequent surveys, because more samples are included. We can apply the same logic to infer the appropriate correction to apply to σ^{obs} , such that it reflects the uncertainty about the rate of land-use change in any given year within the inter-survey interval:

$$\sigma_{ann}^{obs} = \sigma^{obs} \sqrt{n}$$

where n is the survey interval length in years. This scaling means that the correction evaluates to 1 for annual data (no effect).

All data sources have a limited period over which they were collected, and this would introduce an artefact at the boundaries. When a data set which tends to provide lower estimates begins, this would tend to pull the mean estimate down at this point, and introduce a step change purely as an effect of data availability. To counter this effect, we apply a similar logic to that above - the available data is an imperfect sample of the surrounding time interval, and the uncertainty increases with fewer samples and extended distance in time. We therefore extend each data set to the limits of the time period considered here (1950-2020), assuming a constant rate, but we force the uncertainty to increase with the square of distance in time (Δ_t^2) beyond the boundaries where data was actually available. The random uncertainty in any given year is

$$\sigma_t^{obs} = \sigma^{obs} \Delta_t^2$$

scaled so that there is no effect within the time bounds of the observed data.

2.3 Results

The table below shows the RMSE, the correlation coefficient and the scaling factor for σ^{obs} (expressed as a coefficient of variation (CV), σ^{obs} as a fraction of the observed value. The ranking shows that CS has the lowest uncertainty and the CROME has

the highest (bearing in mind all the imperfections in the reference data). This produces a quantitative means of accounting for the different relative uncertainties in these data sources.

| | RMSE | r^2 | CV |
|--------|-----------|------------|-----------|
| CS | 654.9496 | 0.2850220 | 0.2000000 |
| CORINE | 830.2048 | 0.3533137 | 0.2293023 |
| LCM | 892.9640 | 0.1650214 | 0.3184481 |
| IACS | 1170.0198 | 0.1882285 | 0.4056547 |
| LCC | 1340.3314 | 0.1350960 | 0.4951190 |
| CROME | 7899.3774 | -0.0819340 | 3.6502540 |

2.4 Discussion

Representing data source-specific random uncertainty is relatively straightforward in principle. We need to estimate appropriate σ values for each data source and use these in the likelihood function. The most fundamental problem is accurately estimating σ in the absence of any data which we regard as “true,” particularly for the B matrices which are key. There is no immediate solution to this, and a pragmatic approach is to define a reference data set, with more or less subjectivity/expert judgement, and potentially with some cross-validation. Here we implemented a simple method, whereby σ for each data set is scaled according to metrics measuring its correspondence with reference data. Currently, this is based only on the net change data, as this makes cross-comparison simplest, but this could be extended to include the gross changes G, L and B .

An alternative approach would be to estimate the uncertainties as part of the data assimilation. This avoids the sticking point of subjectively defining a suitable reference data-set, when all of the available data sources, including ground-truthed data, are flawed in some way. The downside of this approach is that it is more complicated, involves estimating more parameters, and will have greater computation time, but merits some exploration.

3 WP-A Quantifying Systematic Uncertainty in Land-Use Data Sources

3.1 Introduction

The spatial datasets used in the data assimilation for Land Use Tracking will contain systematic errors, related to falsely detecting land-use change when it has not occurred, and missing true land-use change when it does occur. To characterise uncertainties in the data, we want to quantify these false positive and false negative detection rates. These can be estimated by comparison with a reference dataset, and thereby judging where the observed changes in a given data set are correctly identified or not. For observations of the B matrix, this is complicated by the fact that they form a 6 x 6-way classification. When considering land-use *change*, we need to compare the 36 elements of this classification from one data source with another (or the truth), so we have a 36 x 36 error matrix (or “confusion” matrix). This matrix has two distinct types of errors that we ideally want to distinguish: false positives and false negatives, or “user”/“commission” and “producer”/“omission” error/accuracy, in the terminology commonly used in remote sensing.

As with estimating random uncertainty, a limitation is that none of the data sources represents absolute truth, and we have no definitive data set against which to make this assessment.

3.1.1 Representation of systematic uncertainty

Our approach here is to explicitly represent the false positive and false negative error terms in the likelihood function. False positives cause observations to over-estimate change, whilst false negatives produce an under-estimate, and the estimated bias in the observation is a simple function of these error rates (F_P and F_N). The likelihood equation becomes:

$$\mathcal{L} = \frac{1}{\sigma_{ij}^{\text{obs}} \sqrt{2\pi}} \exp\left(-\left(\left(1 - F_P\right)\beta_{ij}^{\text{obs}} + A_N F_N - \beta_{ij}^{\text{pred}}\right)^2 / 2\sigma_{ij}^{\text{obs}2}\right)$$

where A_N is the area in which the false negative errors can occur, given by the number of grid cells where the land-use change ij was not detected. This equation calculates the likelihood of the observed change from land use i to land use j , given that the true value is β_{ij}^{pred} , and with given false positive and false negative rates F_P and F_N , and random uncertainty σ_{ij}^{obs} in the observation. This approach can be implemented in increasingly complex ways:

- estimating the false positive and false negative error rates based on some set of confusion matrices, and thereafter assuming them to be fixed and constant for a given data source;

- as above, but calculating false positive and false negative error rates specific to each type of land-use change (i.e. F_{Pij} and F_{Nij}), and potentially varying in time;
- including the false positive and false negative error rates as unknown parameters to be calibrated, along with the B and σ_{obs} values. This is the most sophisticated solution, as it properly represents the fact that these are not truly known, and allows the values to be an emergent property of the data, given prior information, rather than imposing our guesses. The exact number of these parameters to estimate could vary as above, whether specific to each data source, type of land-use change, and point in time.

3.2 Methods

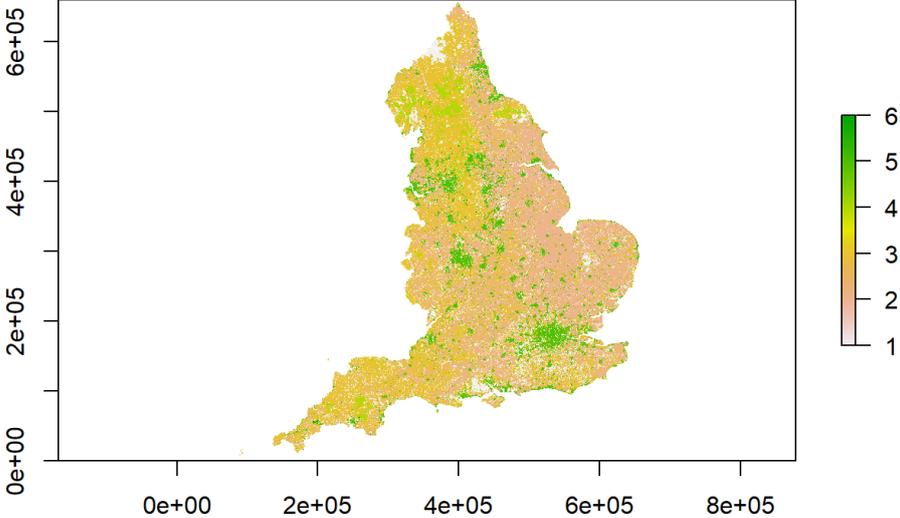
We explored the variability in false positive and false negative rates in the datasets at the land-use, data-source and time levels to identify the most appropriate way to account for uncertainty in the data assimilation. The definition of the reference data, calculation of false positives and negatives, and application of these values into the likelihood function is presented below.

3.2.1 Creating reference datasets:

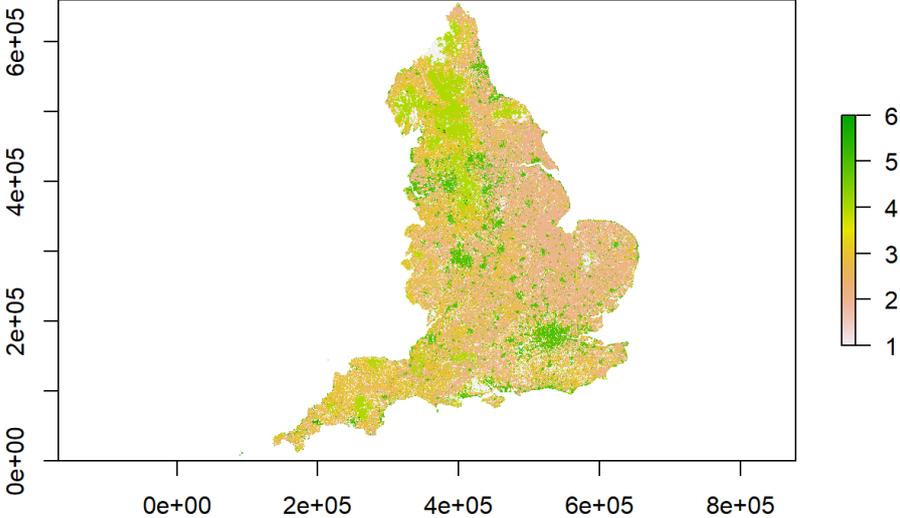
We created the reference data by combining several spatially explicit data sources that are thought to have the highest accuracy and reliability for certain land use types, namely FC, IACS and LCM. These data-sources were added to the reference dataset in order of expected reliability: land use defined in FC data took precedence over IACS, followed by LCM. This resulted in a reference dataset that includes the majority of forest cells from FC, the majority of crop, grass and rough grazing from IACS and the majority of urban and other cells from LCM. This reference data was used to compare the other data-sources against (LCC, CORINE, CROME). In order to test FC, IACS and LCM themselves, we removed the respective dataset from the reference data, and tested land-use change classifications against those present in the remaining reference data (e.g. testing the forest, crop, grassland or rough grazing land-use change in IACS against that defined in FC and LCM).

This method enabled us to build reference rasters for England for 2006, 2010, 2015, 2017, 2018 and 2019. The 2010 reference raster comprising FC, IACS and LCM data is shown below, as well as the 2010 reference raster for testing IACS (comprising FC and LCM data).

Reference data 2010 (FC, IACS and LCM)



Reference data 2010 (FC and LCM only)



The table below shows the reference data used for each data source:

| Dataset | Reference |
|---------|-----------|
| FC | IACS, LCM |
| IACS | FC, LCM |

| Dataset | Reference |
|---------|---------------|
| LCM | FC, IACS |
| LCC | FC, IACS, LCM |
| CORINE | FC, IACS, LCM |
| CROME | FC, IACS, LCM |

Each data source was tested against its appropriate reference data according to the table above. The resulting 36*36 confusion matrix quantifies the correspondence between the data sets, in terms of their agreement over the area of each land use changing to every other land use. The diagonal identifies the area of land-use change that was identified to occur in both the reference and test data sets. The unchanging land on the diagonal can be disregarded as not relevant here.

3.2.2 False positive rates

The false positive rate was calculated as:

$$F_{Pij} = (\beta_{test,ij} - \beta_{ref,ij}) / \beta_{test,ij}$$

where $\beta_{test,ij}$ is the observed area of land changing from use i to j in the test dataset and $\beta_{ref,ij}$ the corresponding value in the reference dataset.

3.2.3 False negative rates

The false negative rate, the rate of failing to observe land-use change from i to j in the test dataset compared to the reference dataset, was calculated as:

$$F_{Nij} = (\beta_{ref,ij} - \beta_{ref,test,ij}) / (A - \beta_{test,ij})$$

where $\beta_{ref,test,ij}$ is the area of land changing from i to j identified in both the reference and test data sets, and A is the total area of England. The denominator is thus the total land area where false negatives could occur.

3.3 Results

Here we use CORINE as an example showing the false positive and negative rates calculated between the data source and the reference dataset between 2006 and 2018 (most recent years available for comparison). The data shows the trend across all of the data-sources: very high false positive rates, with slightly lower false positive rates for the land-use change between crop and grassland. False negative rates were much lower, but are expressed on a very different basis, and not directly

comparable. The matrices below show the false positive/negative rate for each land-use change between the first year (row) and the second year (column).

False positive rates for CORINE between 2006 and 2018:

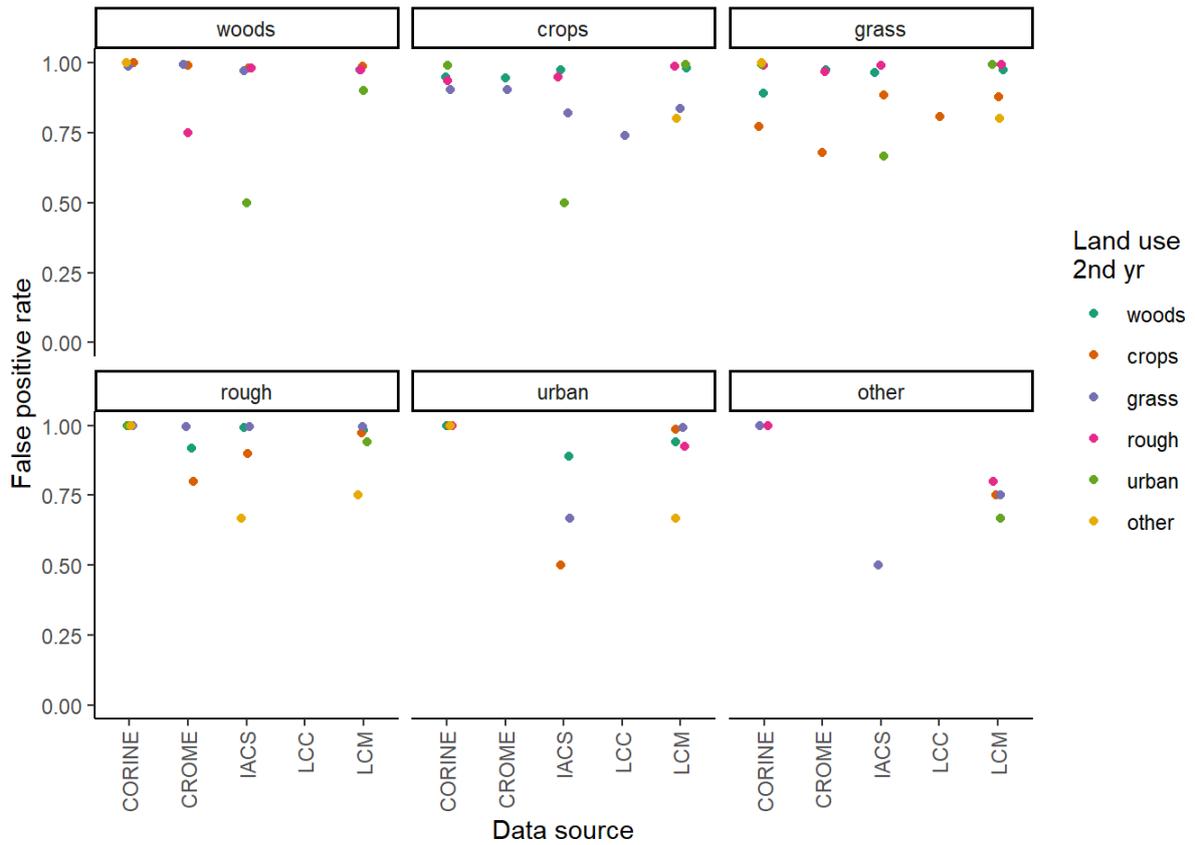
| | woods | crops | grass | rough | urban | other |
|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| woods | | 1 | 0.992 | 1 | 1 | 0 |
| crops | 0.977 | | 0.928 | 0.949 | 0.994 | 1 |
| grass | 0.951 | 0.767 | | 0.996 | 0.991 | 1 |
| rough | 1 | 1 | 1 | | 1 | 1 |
| urban | 1 | 1 | 0.985 | 1 | | 1 |
| other | 0 | 1 | 1 | 1 | 0 | |

False negative rates for CORINE between 2006 and 2018:

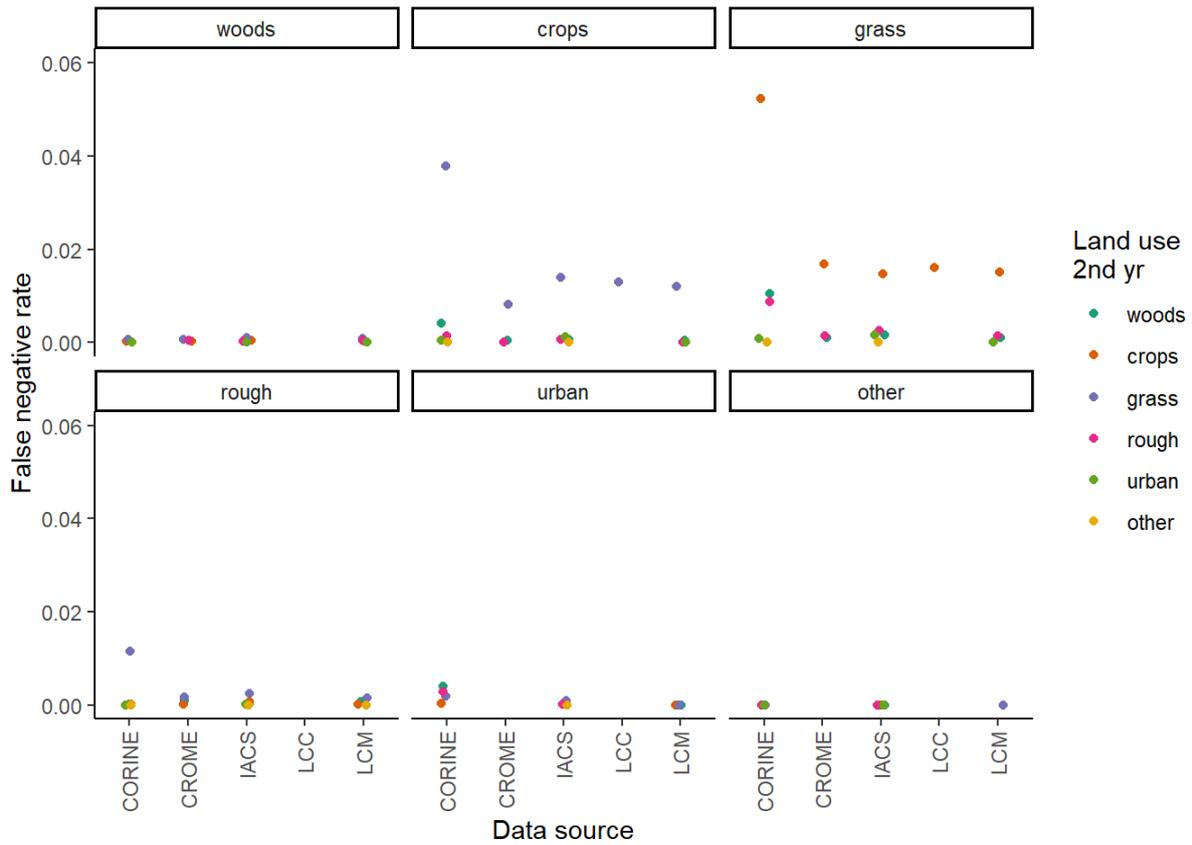
| | woods | crops | grass | rough | urban | other |
|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| woods | | 0 | 0.001 | 0 | 0 | 0 |
| crops | 0.004 | | 0.038 | 0.001 | 0 | 0 |
| grass | 0.01 | 0.053 | | 0.009 | 0.001 | 0 |
| rough | 0 | 0 | 0.011 | | 0 | 0 |
| urban | 0.004 | 0 | 0.002 | 0.003 | | 0 |
| other | 0 | 0 | 0 | 0 | 0 | |

The following graphs show the false positive and false negative rates for each data source, with the grid of graphs showing the land use type in the first year and colour showing the land use in the second year.

False positive rates:



False negative rates:



3.3.1 Data-source-specific false positive and false negative rates

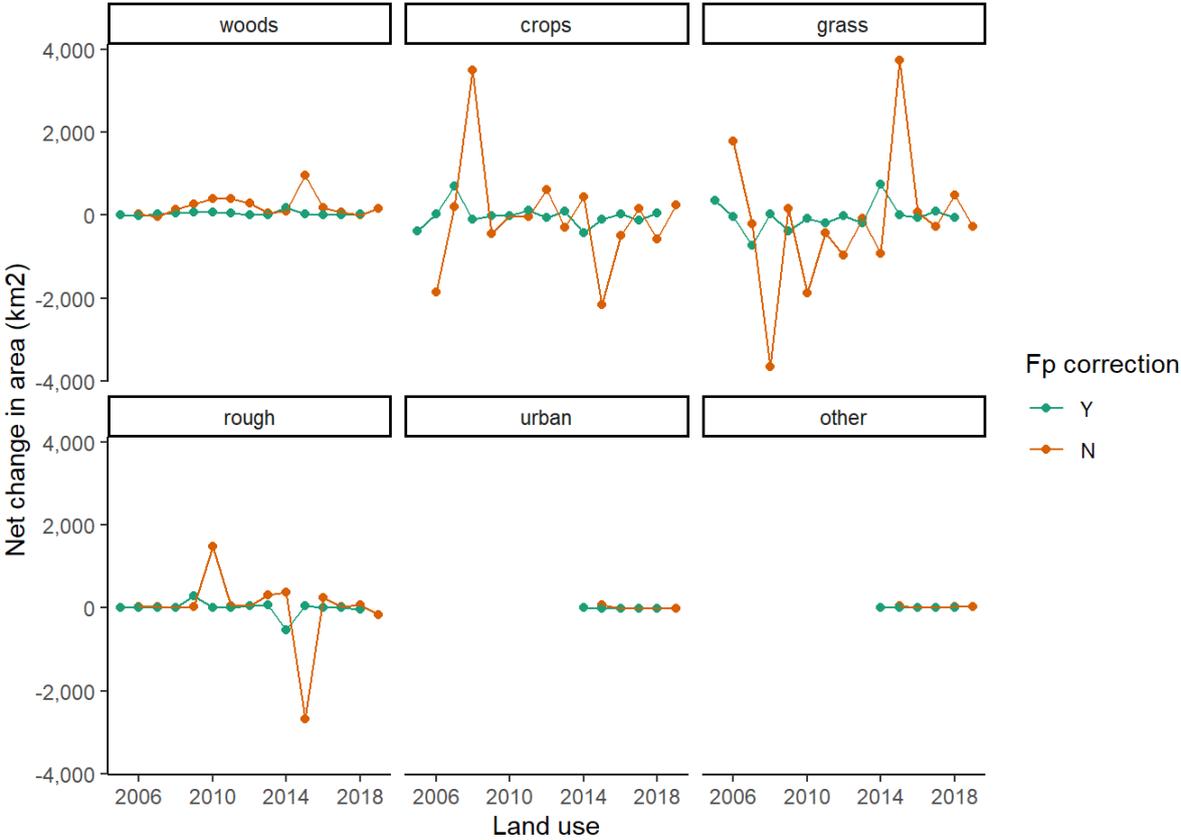
Based on the relatively low variability in false positive rates identified for many of the land-use change classes, we chose to summarise false positive rates at the data-source level rather than land-use level. Data-set level false positive rates were calculated as an average of the land-use level false positive rate, weighted by area of each land-use change.

This gave the following outputs which can be incorporated into the data assimilation method:

| Data source | F_N | F_P |
|-------------|-----------|-----------|
| FC | 0.0020230 | NA |
| LCM | 0.0013481 | 0.8711168 |
| CORINE | 0.0060332 | 0.9029868 |
| LCC | 0.0012670 | 0.6737150 |
| IACS | 0.0020691 | 0.7997226 |
| CROME | 0.0012571 | 0.9329713 |

3.4 Updating B, D, G and L

To examine the effect of these systematic biases on the observations, we can recalculate the B transition matrix. An example is provided below for IACS from 2005 to 2019:



3.5 Conclusions

We can represent data source-specific systematic uncertainty by estimating appropriate false positive and false negative rates for each data source and adding these terms to the likelihood function.

Again, the most fundamental problem is the absence of any data which we regard as “true.” There is no immediate solution to this, and a pragmatic approach is to define a reference data set, with more or less subjectivity/expert judgement, and principles from cross-validation.

False positive rates appear to be very high but are consistently so across most of the land-use categories. With this in mind, we use data-set level F_p and F_n values rather than land-use level values.

4 WP-A Assessing systematic errors in estimates of land-use change: sensitivity to survey-interval length

4.1 Introduction

The spatial datasets used in the data assimilation for Land Use Tracking will contain systematic errors, related to falsely detecting land-use change when it has not occurred, and missing true land-use change when it does occur. To characterise uncertainties in the data, we want to quantify these false positive and false negative detection rates. These have previously been estimated by comparison with a reference dataset, and thereby judging where the observed changes in a given data set are correctly identified or not. However, this depends on the validity of the reference dataset as a standard for comparison, and we know that the reference data set is imperfect. Here we use an alternative method that assesses these error rates by analysing the apparent rate of land-use change as a function of the time interval between surveys. In the absence of systematic errors, no relationship with survey-interval length would be expected, and any apparent sensitivity can be used to infer the error rates.

The observed area changing from one land-use type i to another j between two surveys, β_{ij}^{obs} (in $\text{km}^2 \text{yr}^{-1}$), will be made up of the true rate of change, β_{ij} and systematic and random error terms. Systematic errors comprise false positive (F_p) and false negative rates (F_n). Together with the random error term ϵ_{ij} , we can express our expectation for the observations to be:

$$\beta_{ij}^{\text{obs}} = \beta_{ij} + (F_{ij}^P - F_{ij}^N) + \epsilon_{ij}$$

As a broad approximation, we can assume that the true rate of land-use change, and the error rates, are approximately constant in time. With short intervals between

surveys, there will therefore be proportionately less true change, but the magnitude of the errors will be the same. Conversely, with long intervals between surveys, the magnitude of the errors will still be the same, but there will be proportionately more true change. As the time difference between surveys increases, the observed rate will tend towards an asymptote, equal to the true mean rate of land-use change, $\overline{\beta_{ij}}$, as the random error term ϵ_{ij} has a mean of zero. We can therefore examine the apparent rate of land-use change as a function of the time interval between surveys, and infer the error rates from this relationship. It is not possible to explicitly separate F_p and F_N in this analysis, only their net effect $F^{net} = F_p - F_N$, although the shape of curve indicates which is larger in the data. If F_p and F_N were zero, or exactly balancing each other, the data would show a flat line in the figures below. For the datasets we are testing here, in all cases it appears that false positives are the major source of error.

This method is potentially superior to estimating error rates by comparison with a reference dataset, as it does not require reliable ground-truth data, and only uses intrinsic properties of the observations. As a down-side, the error rates calculated apply to ongoing directional change, and any rotational change is effectively included in the error term. However, we note that this is only really an issue with crop-grass transitions (and v.v.), and we estimate, this affects around ~7 % of the grassland area (assuming that half the area of grassland < 5 years old has a rotation length shorter than the observation period). The method also assumes that the true mean rate of land use change has not changed systematically over the observation period.

4.2 Methods

- Calculate beta matrices for land-use change with all possible permutations of between-survey intervals available with each data source
- Plot the relationship between time-interval length (Δ) and the apparent rate of land-use change
- For each term in the β matrix, we fit an exponential model to the data using nonlinear least squares:

$$\beta_{ij\Delta}^{obs} = \overline{\beta_{ij}} + (A_0 - \overline{\beta_{ij}})\exp(-\exp(k)\Delta)$$

where $\overline{\beta_{ij}}$ is the asymptotic value, equal to the long-term mean rate of land-use change, A_0 is the intercept at $\Delta=0$, and k is the natural logarithm of the rate constant.

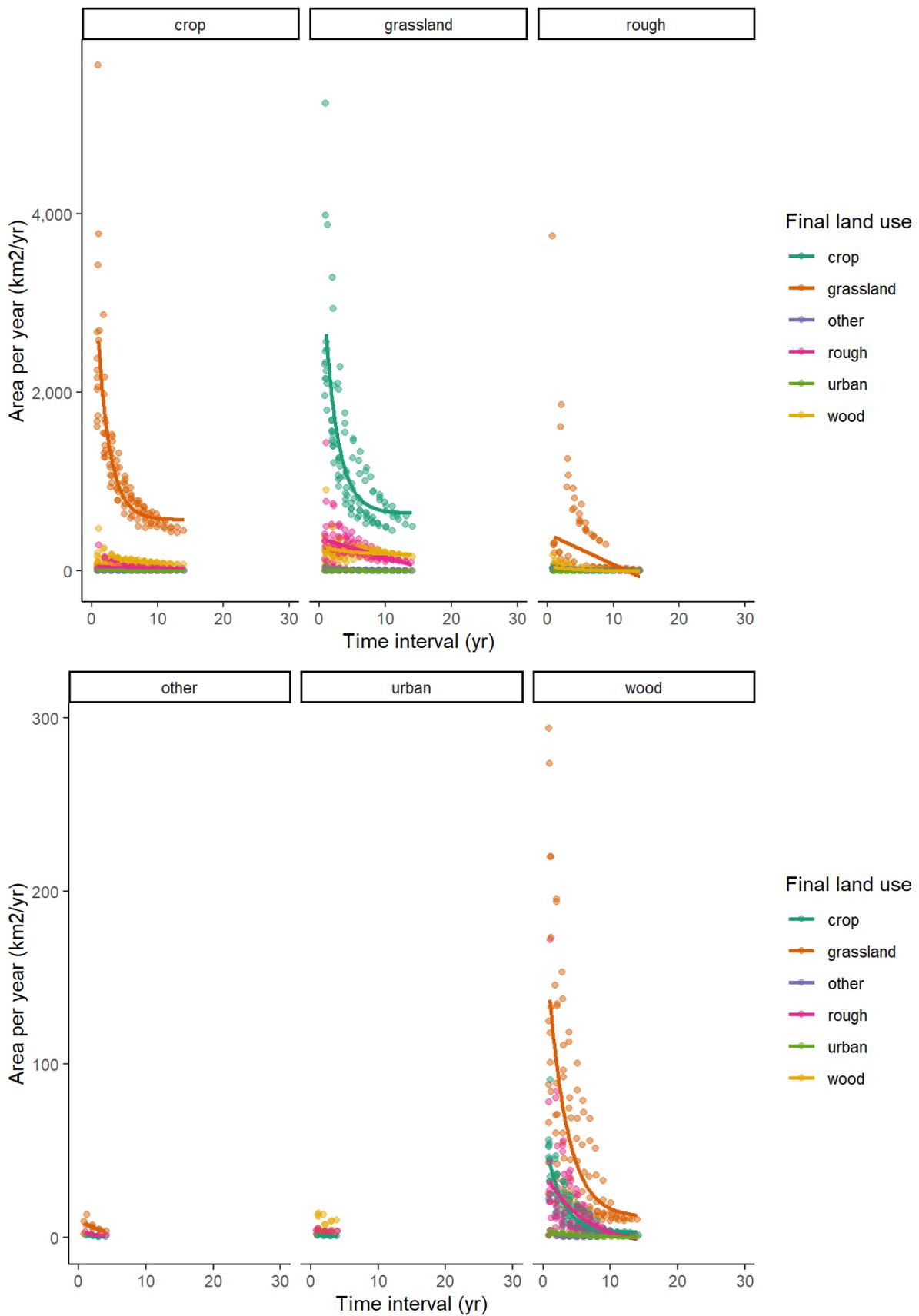
From the fitted model, we obtain estimates of $\overline{\beta_{ij}}$ and the value of $\beta_{ij\Delta}^{obs}$ as a function of the time interval Δ . We can then estimate the mean net error rate for a given time interval from the fitted curve, expressing this as a fraction of the observed rate:

$$F_{ij\Delta}^{net} = (\beta_{ij\Delta}^{obs} - \overline{\beta_{ij}}) / \beta_{ij\Delta}^{obs}$$

4.3 Results

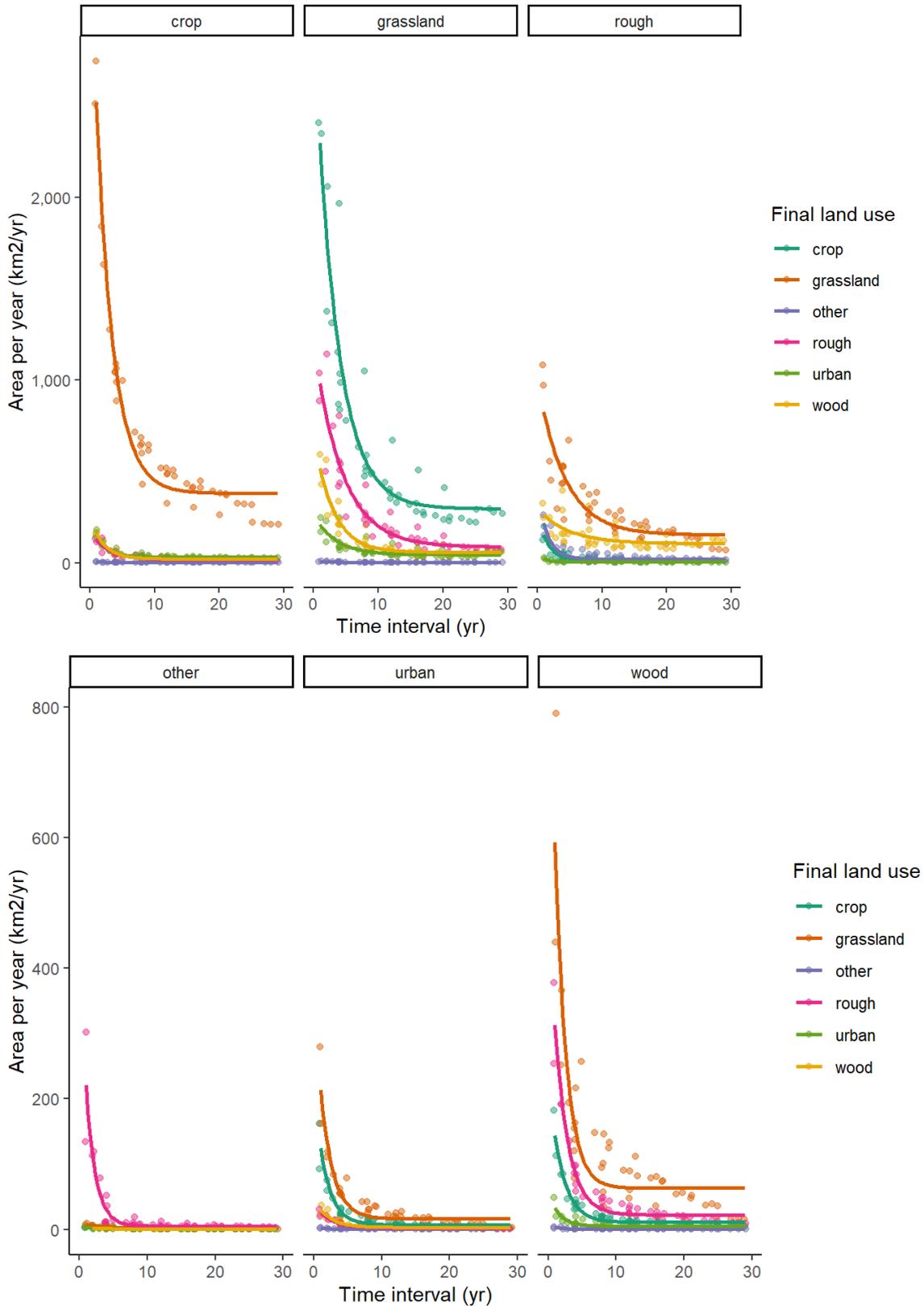
In almost all cases, we see very strong sensitivity to survey-interval length, with much higher apparent rates of change observed at short interval lengths. This implies the observations are dominated by false positives; no relationship would be expected in the absence of such errors. Similar trends are seen in all the data sets examined here.

4.3.1 IACS:



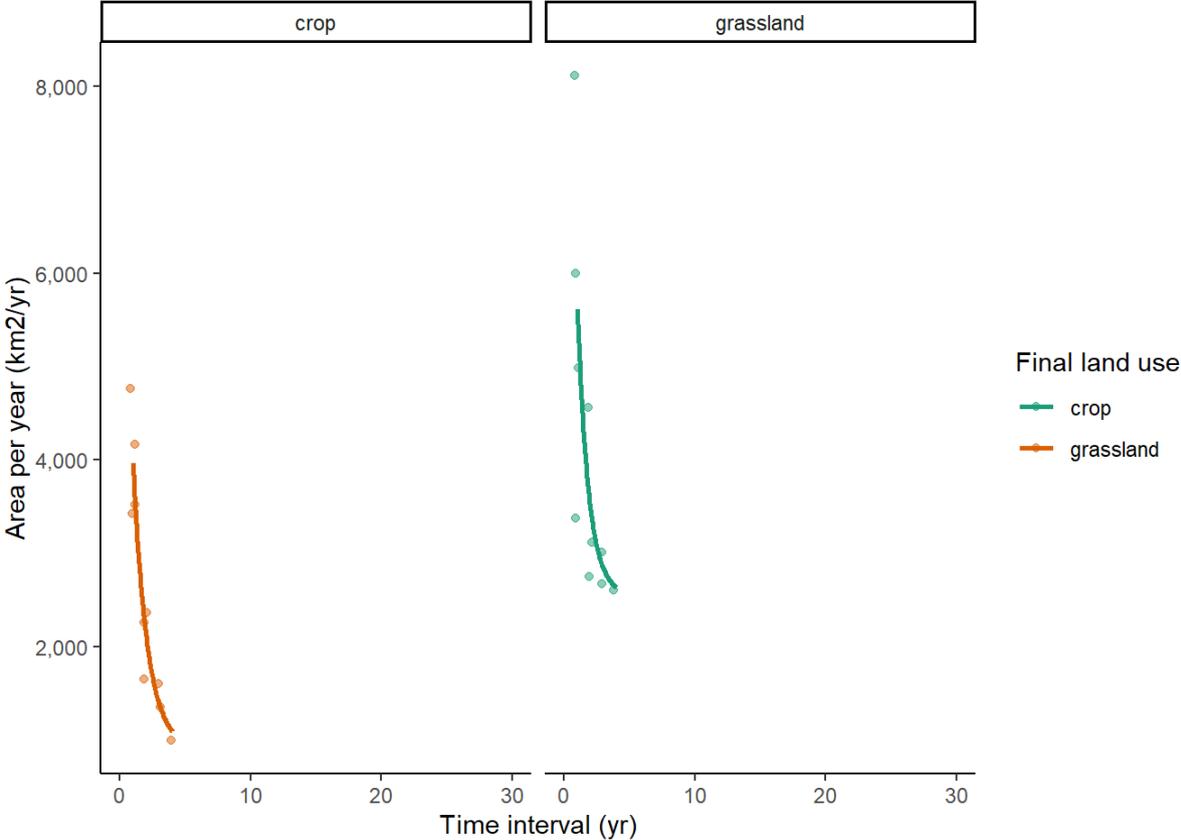
4.3.2 LCM

LCM has the greatest number of datasets to apply this method to with 10 surveys conducted across 29 years. This enables comparison between surveys that give 23 time intervals.



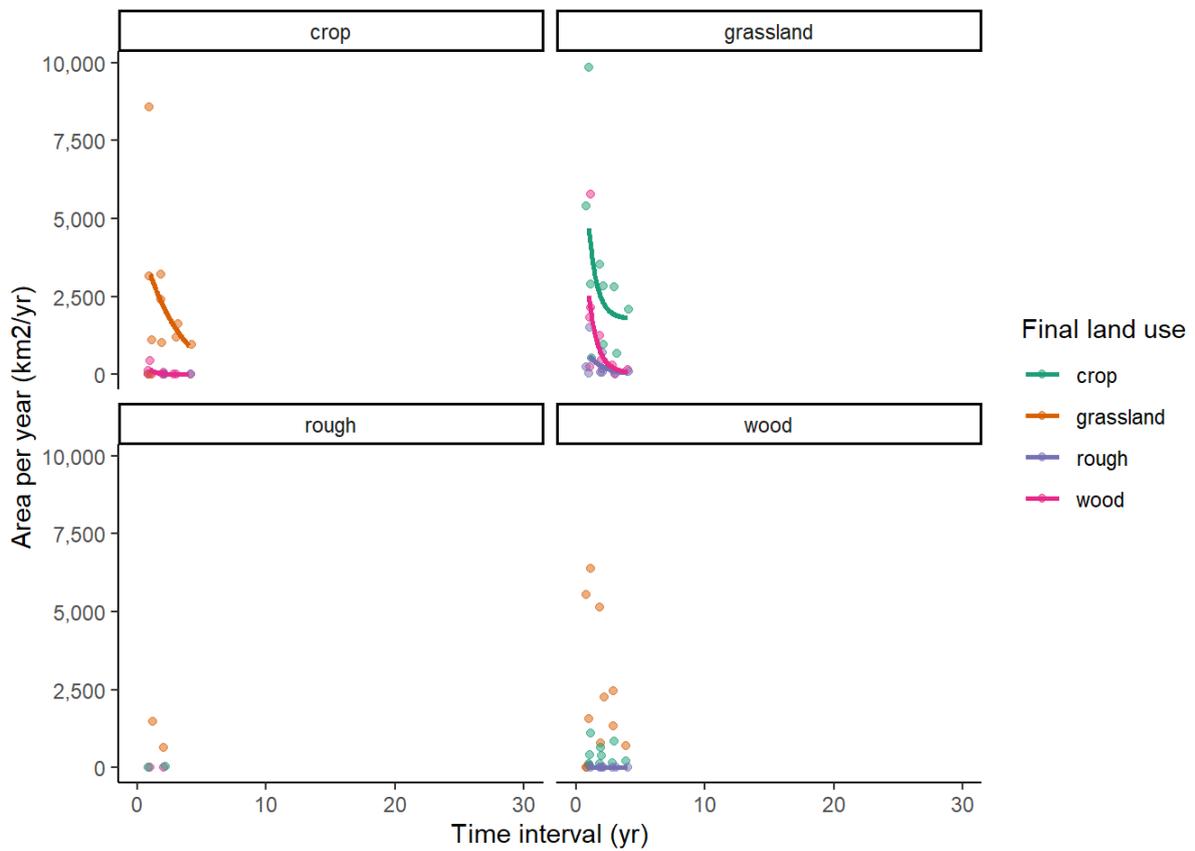
4.3.3 LCC:

LCC only includes crop and grassland land use change so there are less data to test.

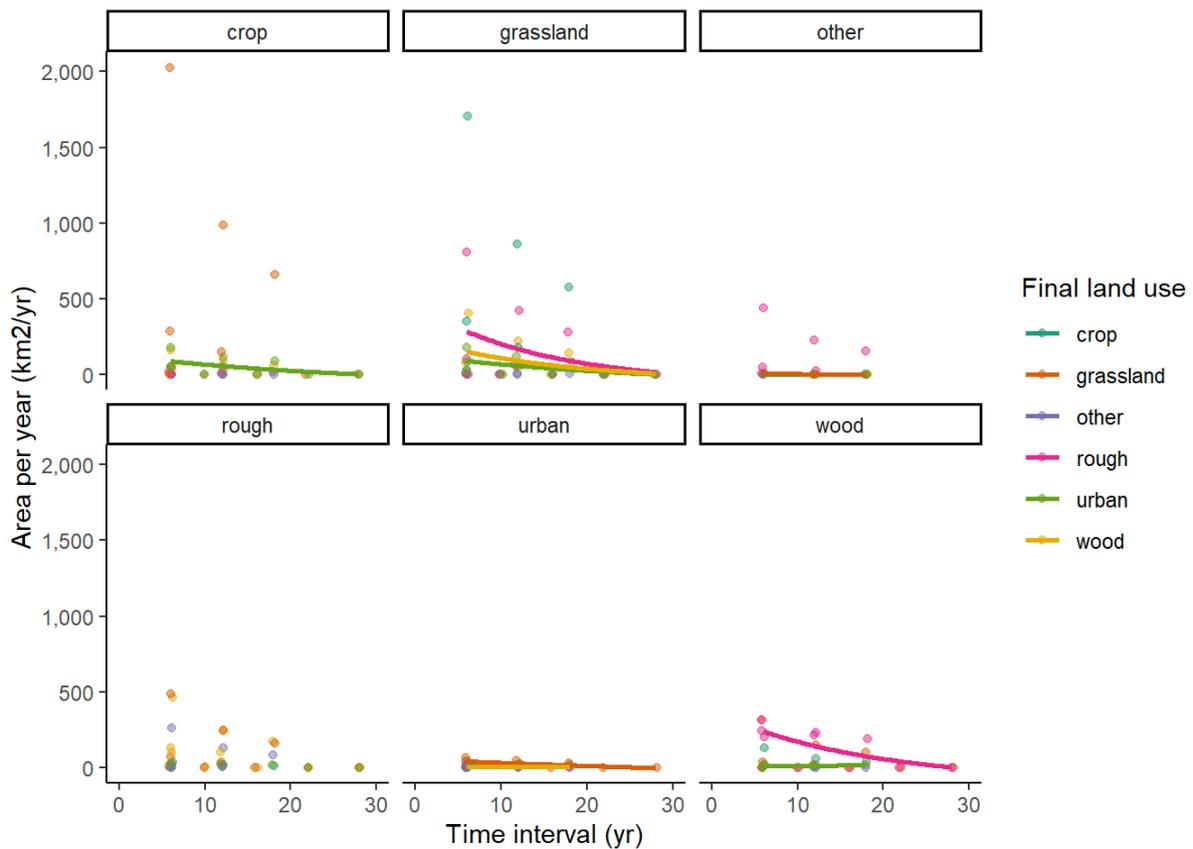


4.3.4 CROME:

Both CROME and CORINE have few surveys meaning estimating a fit to this data is difficult:



4.3.5 CORINE:



4.3.6 Summary Table

The table below shows the error rate for all the different land use change categories for each of the different data sets. These are calculated for the typical time interval in each of the different data sources.

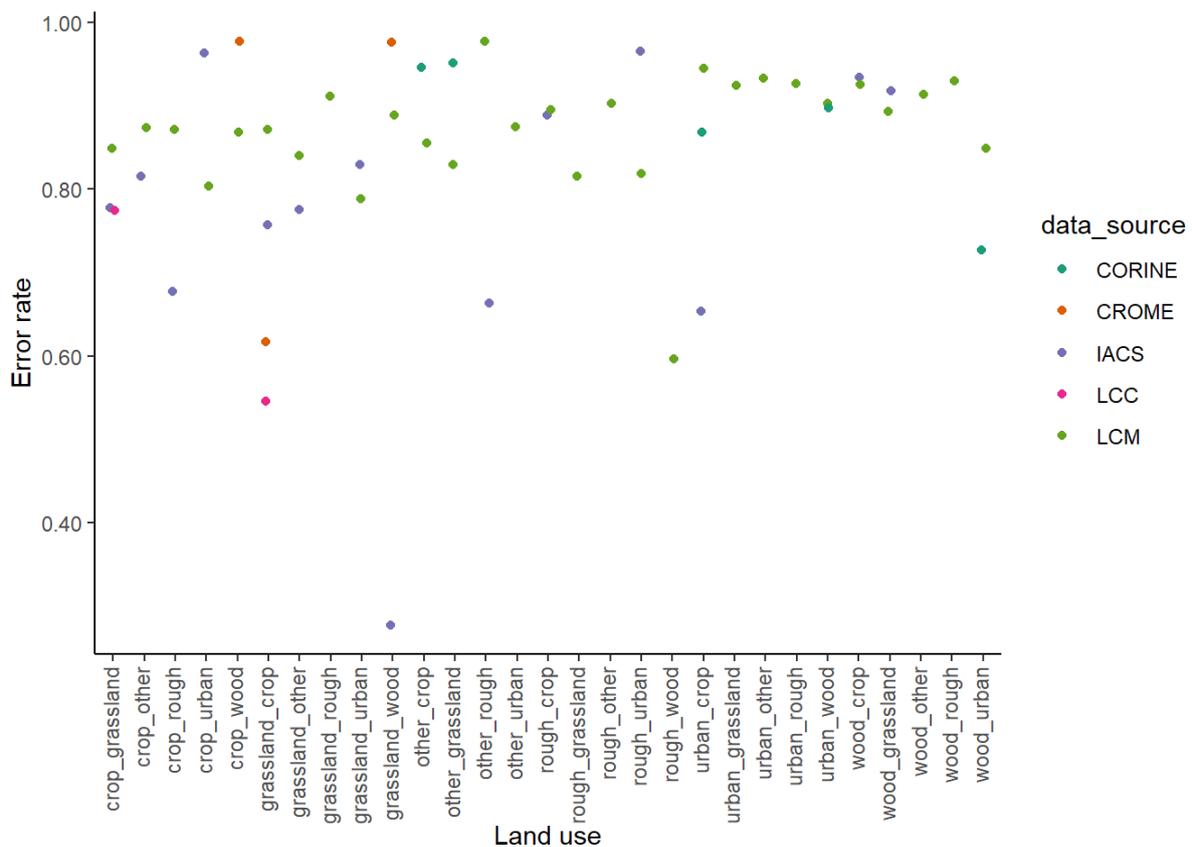
| luc | IACS | LCM | LCC | CORINE | CROME |
|-----------------|------|------|------|--------|-------|
| crop_grassland | 0.78 | 0.85 | 0.77 | NA | NA |
| crop_other | 0.82 | 0.87 | NA | NA | NA |
| crop_rough | 0.68 | 0.87 | NA | NA | NA |
| crop_urban | 0.96 | 0.80 | NA | NA | NA |
| grassland_crop | 0.76 | 0.87 | 0.55 | NA | 0.62 |
| grassland_other | 0.78 | 0.84 | NA | NA | NA |

Improving Land Use Change Tracking in the UK Greenhouse Gas Inventory: Final Outputs Report

| luc | IACS | LCM | LCC | CORINE | CROME |
|-----------------|-------------|------------|------------|---------------|--------------|
| grassland_urban | 0.83 | 0.79 | NA | NA | NA |
| grassland_wood | 0.28 | 0.89 | NA | NA | 0.98 |
| other_rough | 0.66 | 0.98 | NA | NA | NA |
| rough_crop | 0.89 | 0.90 | NA | NA | NA |
| rough_urban | 0.96 | 0.82 | NA | NA | NA |
| urban_crop | 0.65 | 0.94 | NA | 0.87 | NA |
| wood_crop | 0.93 | 0.93 | NA | NA | NA |
| wood_grassland | 0.92 | 0.89 | NA | NA | NA |
| crop_wood | NA | 0.87 | NA | NA | 0.98 |
| grassland_rough | NA | 0.91 | NA | NA | NA |
| other_crop | NA | 0.86 | NA | 0.95 | NA |
| other_grassland | NA | 0.83 | NA | 0.95 | NA |
| other_urban | NA | 0.87 | NA | NA | NA |
| rough_grassland | NA | 0.82 | NA | NA | NA |
| rough_other | NA | 0.90 | NA | NA | NA |
| rough_wood | NA | 0.60 | NA | NA | NA |
| urban_grassland | NA | 0.92 | NA | NA | NA |
| urban_other | NA | 0.93 | NA | NA | NA |
| urban_rough | NA | 0.93 | NA | NA | NA |

| luc | IACS | LCM | LCC | CORINE | CROME |
|------------|------|------|-----|--------|-------|
| urban_wood | NA | 0.90 | NA | 0.90 | NA |
| wood_other | NA | 0.91 | NA | NA | NA |
| wood_rough | NA | 0.93 | NA | NA | NA |
| wood_urban | NA | 0.85 | NA | 0.73 | NA |

The figure below shows the same data plotted for all data sources.



4.4 Discussion and Conclusions

The error rates we obtain from this method are similar to those derived from comparison with the reference data set. The net positive error rates are generally in the range 80-98 %; that is, 80-98 % of the ostensibly observed land-use change did not actually occur. Almost all other values in the range 50-80 % (with only one exception less than this). The conclusion from this is that these observations are extremely over-sensitive, and for whatever reason, differences in imagery (or survey data in the case of IACS) at different times is being recorded as land-use change when none has occurred.

This makes it challenging to extract useful information from these data. However, if we believe the error rates to be consistent, we can specify and correct for these errors in the data assimilation procedure, as described previously. Land-use-change-specific rates can be estimated from this analysis, to capture the variability in errors between different land-use conversions. However, the error rates are broadly similar, and a single value per data source could justifiably be used. To explore this further, confidence intervals can also be calculated for the error rates in the table above, using the standard errors in the parameters from the exponential model fit. If the errors are not consistent in time, and given their magnitude of 80-98 %, an alternative conclusion would be that these observations are not yet reliable enough to include in the inventory procedure for tracking land-use change.

5 WP-A Estimating false-positive rates in detection of land-use change based on classification accuracy

5.1 Introduction

In the “Tracking Land-Use Change” project, several data sources provide a time series of maps of land use. An obvious approach is to estimate land-use change as the difference between these maps over time. However, any error in land classification will also be included in the estimate of land-use change, so it is important to quantify these errors properly and note how they propagate. Here, we show the calculations which propagate the error in classification through to its resulting effect on the estimate of land-use change.

5.2 Methods

The accuracy of land-use maps is usually estimated by comparison with some reference data set. We can then calculate confusion matrices and metrics of overall agreement, of which there are several. Common choices are overall accuracy (α , the fraction of locations where estimated land use agrees with the reference data set) and the κ statistic, which corrects for the probability of chance agreement. κ therefore gives a more robust measure, typically 5-20 % lower than simple percentage agreement. The probability of misclassification can be estimated simply as $1 - \alpha$ or more stringently as $1 - \kappa$. For maps at times t_1 and t_2 , the probabilities of misclassification are denoted p_1 and p_2 . Estimating land-use change involves calculating the difference between maps, and the errors are additive in the result. The probability of estimating erroneous land-use change because of misclassification in a pair of maps can be written as

$$p_{1\cup 2} = p_1 + p_2 - p_{1\cap 2}.$$

That is, the probability of error is the union of two events (misclassification occurring at time 1 or at time 2, minus their intersection $p_{1\cap 2}$, which is the probability of misclassification at both time 1 and time 2, which would otherwise be double-counted. $p_{1\cap 2}$ can be estimated as p_1p_2 assuming that the errors leading to misclassification at times 1 and 2 are independent of each other. In practice, our estimates of the probabilities of misclassification at times 1 and 2 are usually the same ($p_1 \approx p_2$), so this simplifies to:

$$p_{1\cup 2} = 2p_1 - p_1^2.$$

5.3 Results

Estimates of α and κ from some of the data sets used in the LUC Tracking project are shown in the table below.

| Data source | α | κ  |
|-------------|----------|--|
| Corine | 0.80 | 0.64 |
| LCC | 0.91 | 0.82 |
| LCM | 0.88 | NA |

For the purposes of the examples below, we use the value of 0.88, the overall accuracy of the LCM, as a relatively optimistic metric. The value of $p_1 \approx p_2$ is $1 - \alpha$, and therefore = 0.12.

Using this value in Equation yields a probability of estimating erroneous land-use change because of misclassification of 0.226. Because this probability applies at every location on the map, multiplying by the total area yields the expected area of erroneous land-use change. So, when comparing two UK maps which each have a classification accuracy of 88 %, 22.6 % of the area, around 55000 km², will show land-use change where none actually occurs. This provides a huge amount of measurement noise when we are attempting to detect a very small signal: the expected magnitude of actual land use change in the UK is of the order of a few hundred km², and at most a few thousand km², based on Forestry Commission planting rates, Agricultural Census, and urban expansion data. The area of land changing use is thus less than 1 % of the total area, and we would therefore need the probability of misclassification error to be less than this in order to accurately detect true change (meaning the accuracy needs to be > 99 %).

We can extend this to calculate the false positive rates for terms in the β matrix and gross gains and losses, given the appropriate denominators and estimates of the true extent of land-use change. For example, the area of cropland in England is approximately 45000 km², and the area of gross gains and losses are estimated to be in the range 300-800 km² y⁻¹ based on CS. Based on the June Agricultural Census, we might estimate these rates to be higher, perhaps reaching 1000-3000 km² y⁻¹.

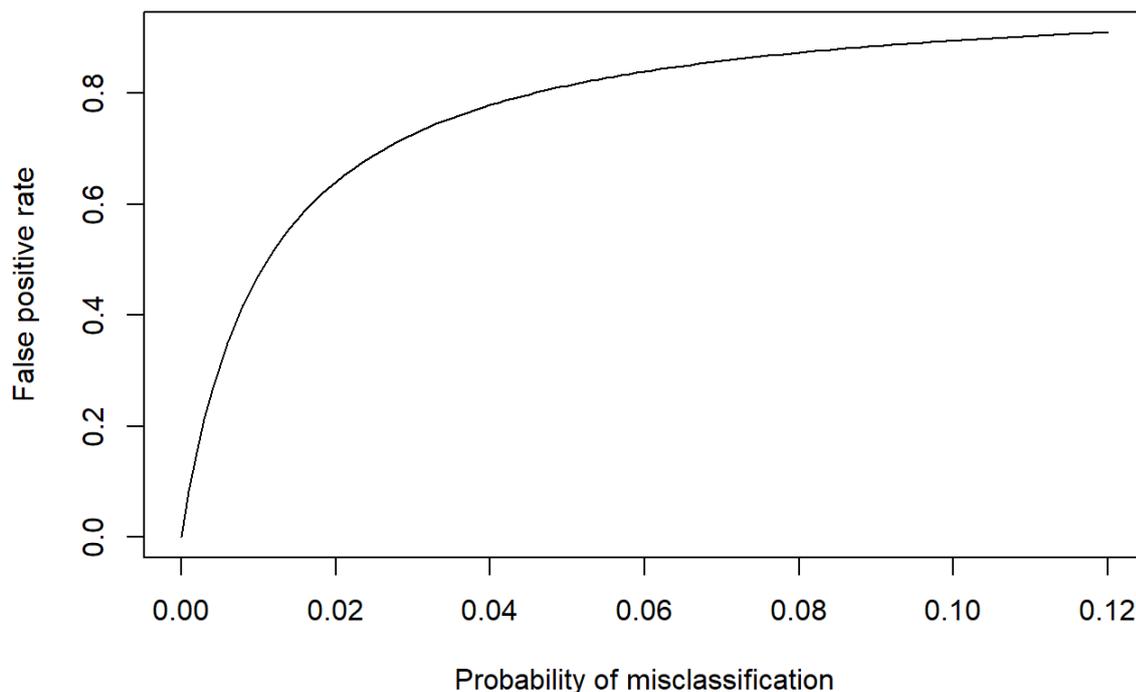
Expressing the estimated true land-use change A_{true} as a fraction of the total recorded land-use change (i.e. true + erroneous), we can calculate the relevant false positive rate, F_P .

$$F_P = 1 - \frac{A_{true}}{A_{true} + A_{false}}$$

If CS rates of land-use change are correct (300-800 km² y⁻¹), the false positive rate is in the range 92.7 to 97.1 %.

If the Agricultural Census rates of land-use change are correct (1000-3000 km² y⁻¹), the false positive rate is in the range 77.2 to 91 %.

Given this relationship between the classification accuracy of individual maps and the resulting false positive rates in detecting land-use change, we can examine the improvement needed to obtain false positive rates below a given level. The figure below shows the change in false positive rate with classification error (1 - α), using the example of cropland gains in England as above, assuming a true rate of change of 1000 km² y⁻¹.



The figure shows asymptotic behaviour because of the form in the previous equation, with a constant true area expressed with an increasing A_{false} term in the denominator. The result is that it takes a substantial decrease in misclassification (or increase in accuracy) from present values to achieve a marked increase in the false positive rate. For example, to reduce the false positive rate to 0.5 requires an accuracy of 0.988. The basic problem is that the true areas of change are very small compared to current error rates, and it would require an order of magnitude improvement in accuracy to reduce the measurement noise to a similar level to the signal we want to detect.

5.4 Conclusions

- Accuracy of land-use classification is in the range 0.8-0.9 in the available data sets. The corresponding probability of misclassification is 10-20 %.
- The errors in individual maps can be propagated to calculate the error in their differences i.e. estimates of land-use change.
- Error rates for land-use change propagating from misclassification will typically be greater than 20 %. This is 1-2 orders of magnitude larger than the expected land-use change.
- The probability of misclassification can be used to calculate the false positive rate of land-use change detection, and values are typically around 90 %.
- Whilst these errors are predictable and can be accounted for, directly detecting the expected land-use change rates <1 % is not currently practicable.

6 WP-A Using life tables in modelling land-use change

6.1 Introduction

This section describes the concept of life tables in modelling land-use change. In the current procedure, we firstly estimate the B matrix each year by MCMC, then estimate where these land-use changes take place in a separate step. This second step uses static maps of likelihood for each land use. That is, for each year, we have a raster containing the likelihood of a given land-use occurring in each cell. This is based on observed data; if several data sets agree that a given cell is used for crops in a given year, there is a high likelihood of any new cropland being placed there by the algorithm (if it is not already cropland). However, these likelihood maps are static: they vary over time according to the data, but they are the same in every simulation. What this misses is the dependence of land-use change on prior history in the grid cell. There are a few cases where this is important. Most importantly, there is rotational grassland, which is used for arable crops for a number of years, before being returned to grassland on a repeating cycle. Thus, the likelihood of grassland changing to cropland is higher for a four-year old grassland than a 50-year old grassland. This phenomenon is not well captured in the current method. For forests, deforestation may be more likely to occur where the trees are at a commercially harvestable age, so the likelihood of transition is not constant, but peaks at around 40-60 years. More generally, land use shows inertia, and change is less likely where no change has happened before.

To capture such “memory” effects (i.e. that the time since past land-use change affects the likelihood of current land-use change), we can use an approach borrowed from population modelling based on “life tables.” In the population modelling context, life tables are a set of age-specific mortality rates. The same idea is referred to as survival analysis, reliability analysis, of time-to-event analysis in various domains. Here, we are modelling the “survival” of land under a given continuous usage. Using the population analogy, a forest is “born” when a grid cell is afforested (from any

other previous land use), and “dies” when it is deforested (converted to any other previous land use). Similarly, the same applies when areas of other land uses are created or destroyed. We can think of this as six populations (woods, cropland, grassland, rough grazing, urban or other land uses), each of which has a specific life table. In this context, rather than mortality rates, the life table is the set of age-specific probabilities of conversion to other uses. So rather than a single dimension, each life table has six columns, for the probabilities to conversion to each of the five other land uses, plus the probability of remaining unchanged.

6.2 Methods

We established the life tables based on observed data, by counting the frequency of the length of all contiguous land uses. The land-use vectors derived from the IACS, CROME, and LCC data sets were used to do this for the six land-use classes considered here. Within these, for each land use, we performed a cross-tabulation of the frequency of transitions to every other land use with age. That is, taking crops as an example, we counted the occurrences of:

- 1-year old crops changing to woods,
- 1-year old crops remaining as crops,
- 1-year old crops changing to grasslands,
- 1-year old crops changing to rough grazing,
- ... etc., and
- 2-year old crops changing to woods,
- 2-year old crops remaining as crops,
- 2-year old crops changing to grasslands,
- 2-year old crops changing to rough grazing,
- ... etc.,

and so forth, up to an age of 10 years, the longest span of continuous data available (in IACS). Normalising by total count, we can convert these frequencies to estimated probabilities. These transition probabilities are usually denoted λ in the context of population modelling (probability of mortality).

6.3 Results

The life tables for cropland and grassland over the first ten years are shown in Figures 6.1 and 6.2.

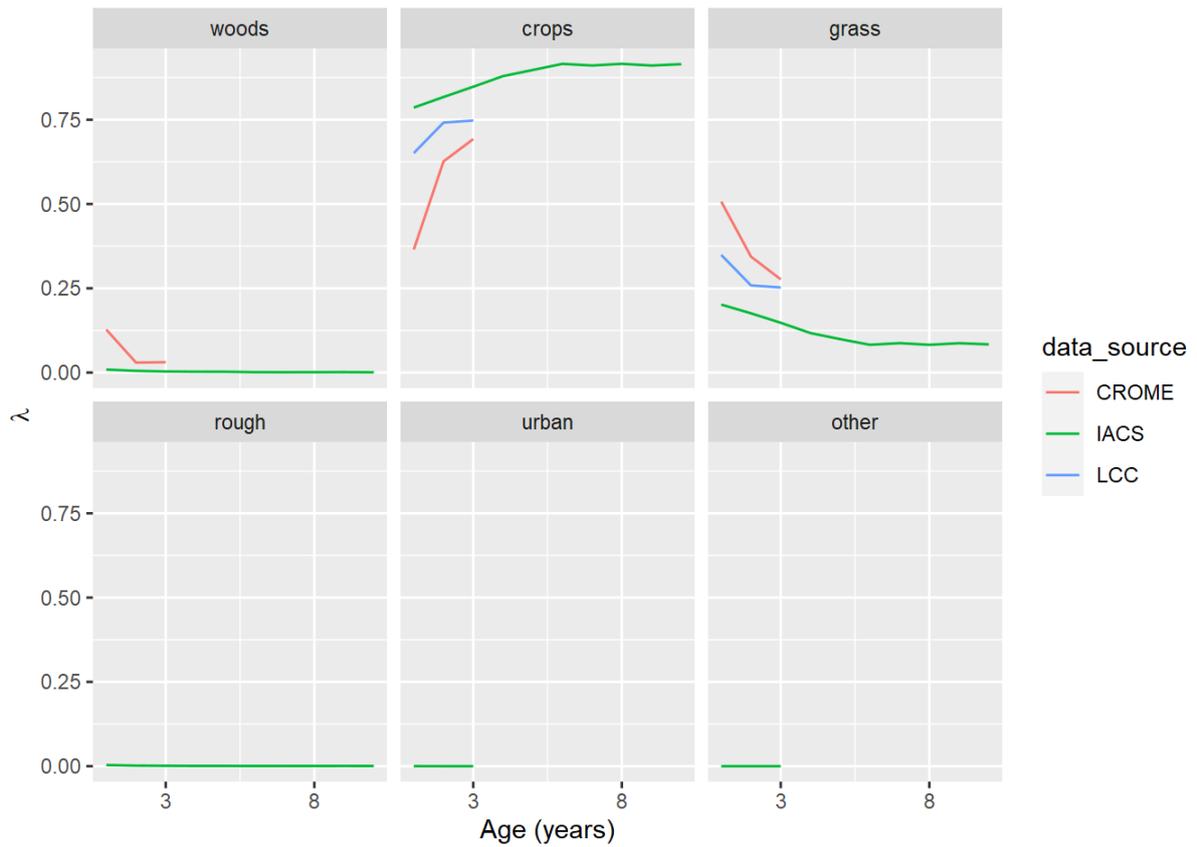


Figure 6.1: The transition probability, λ , for cropland as a function of its age (i.e. time since previous land use). The panel labelled 'crops' shows the probability of cropland remaining cropland.

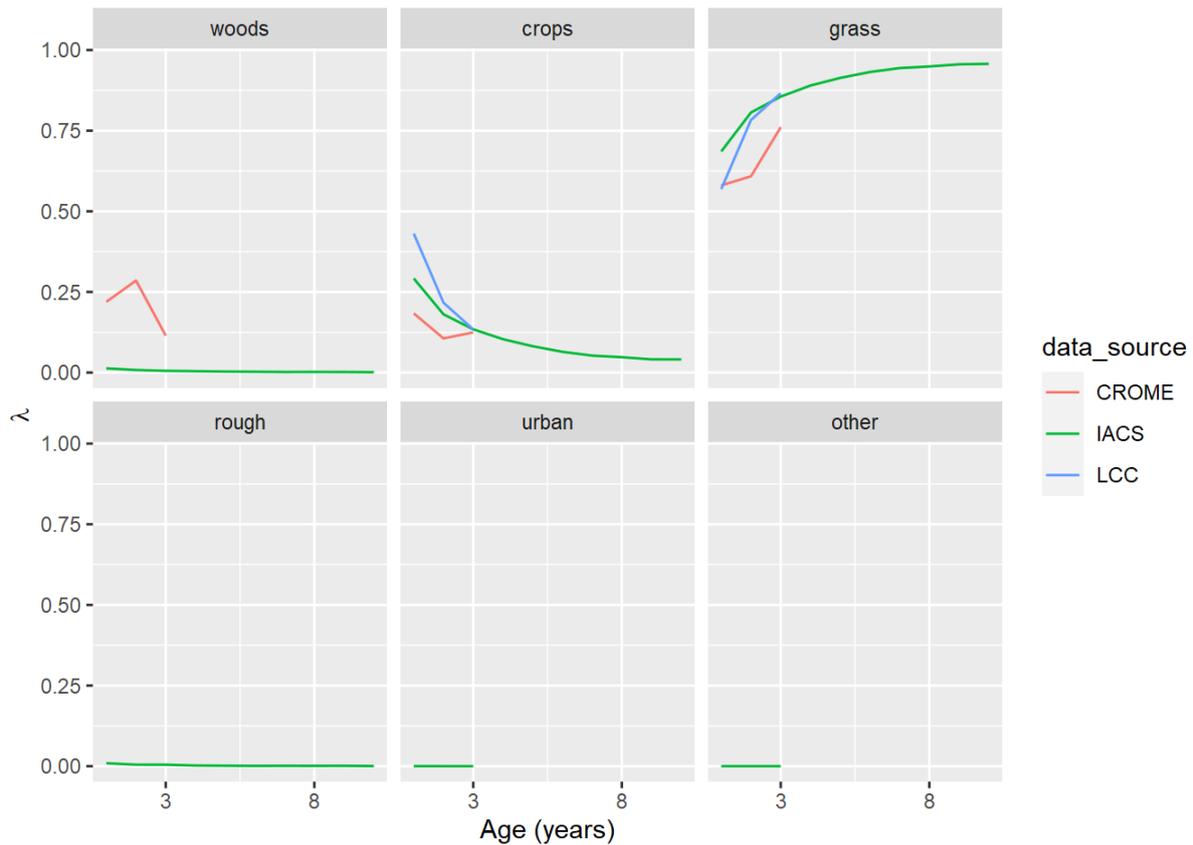


Figure 6.2: The transition probability, λ , for grassland as a function of its age (i.e. time since previous land use). The panel labelled *grass* shows the probability of grassland remaining grassland.

The three data sources show similar patterns, although initial change is generally steeper in CROME and LCC. A 10-year time span is only available in IACS, so definitive comparisons are not possible. In the case of both crops and grasslands, the probability of remaining unchanged increases asymptotically with age; the probability of changing use decreases correspondingly. The most likely transition for cropland is to grassland, and vice versa. The probability of other conversions remains low and roughly constant.

6.4 Discussion

The code for the data assimilation algorithm was adapted to use the life tables dynamically to calculate the likelihoods L in sampling U , going back in time from 2019. Previously, the likelihood calculation was done as a pre-processing step, to calculate a number of static maps, one per year. This is now done dynamically, multiplying a spatial likelihood term L_{static} with the dynamic likelihood term $L_{dynamic}$ (depending on the age of the current land use). Because $L_{dynamic}$ depends only on the age (and not the whole previous history), we simply need to update a raster containing the age of each land use each year. This requires an initial estimate from which to start, and then works backwards. Note that the absolute values in the life tables are not critical; the actual number of cells changing use is determined when we estimate the β matrix values. Indeed, there are reasons not to trust the absolute values as they will include all the false positives discussed in previous sections. It is the shape of these curves

with age that is important. Given we know how many cells are changing from (say) crop to grass each year, the table determines the relative likelihood of this occurring in new croplands versus older more established croplands. Because the possibility of change occurs at every location every year, rotational land use is an emergent property of the simulations, at a frequency approximately the same as in the observed data. In principle this can occur with any land use but is only really significant with changes between crops and grass.

7 WP-A Initial Results: England

The figures below show the results of the data assimilation procedure. All the data sets shown were used in the algorithm, but their relative random uncertainties (σ) determined how much influence they have on the estimates. The spatial data sets were corrected for systematic uncertainties, using the estimated net false positive rate (F_p). Having estimated the posterior distribution of the β matrix, we used this to simulate multiple maps of land use going back in time to 1950. The maps of the likelihood of transition to each land use established in WP-A were updated dynamically, using the life tables described in Section 5.

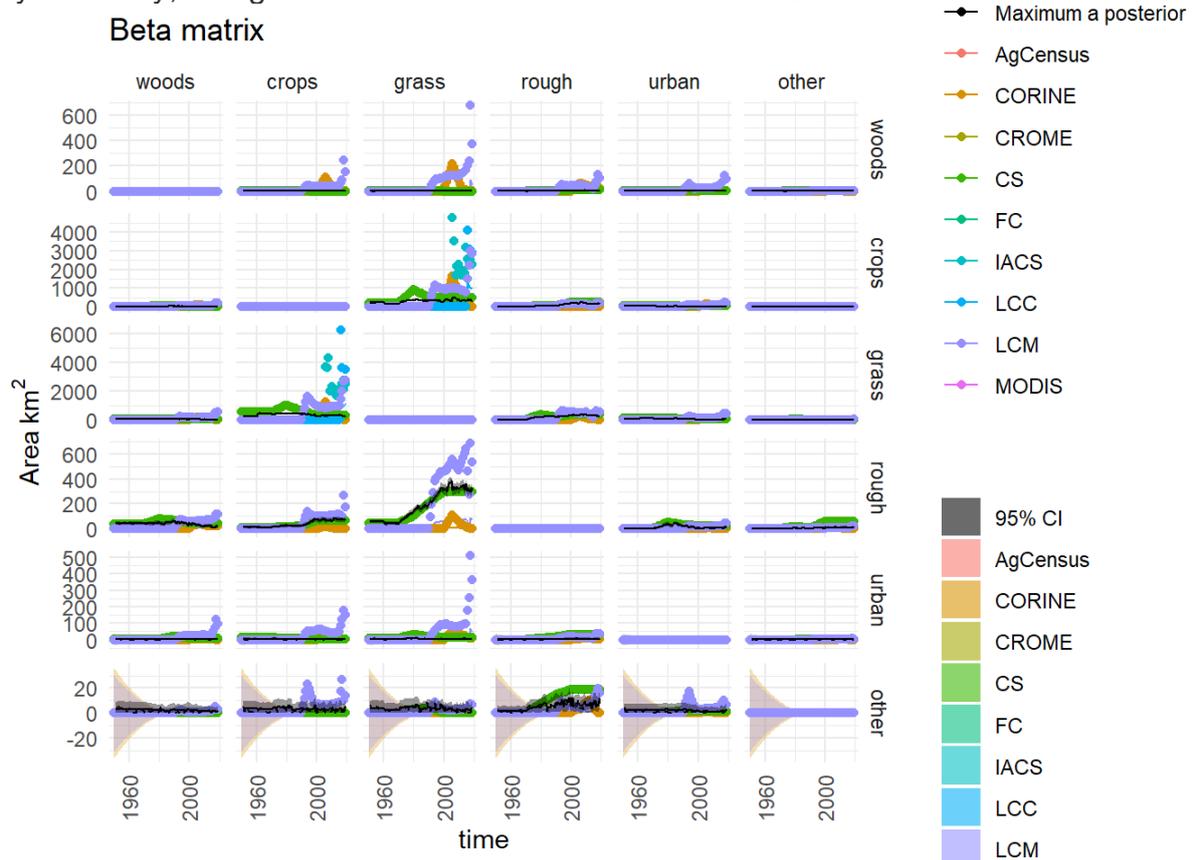


Figure 7.1: Observations and posterior distribution of the transition matrix \mathbf{B} , representing the gross area changing from the land use in each row to the land use in each column each year from 1950 to 2020. The grey shaded band shows the 2.5 and 97.5 percentiles of the posterior distribution. The maximum *a posteriori* estimate is shown as the solid black line within this. Observations from the different data sources are shown as coloured circles. The coloured solid lines show the corrected observations after accounting for systematic uncertainties, and

interpolating. The coloured bands around these lines show the random uncertainty, rescaled as $\sigma/5$ to keep with the axis scale. Because the random uncertainties and the corrections to the observations are generally very large in comparison to the actual change, scaling the axes is difficult. Note that a consistent colour scheme for the data sources is shown, but not all contribute to every figure.

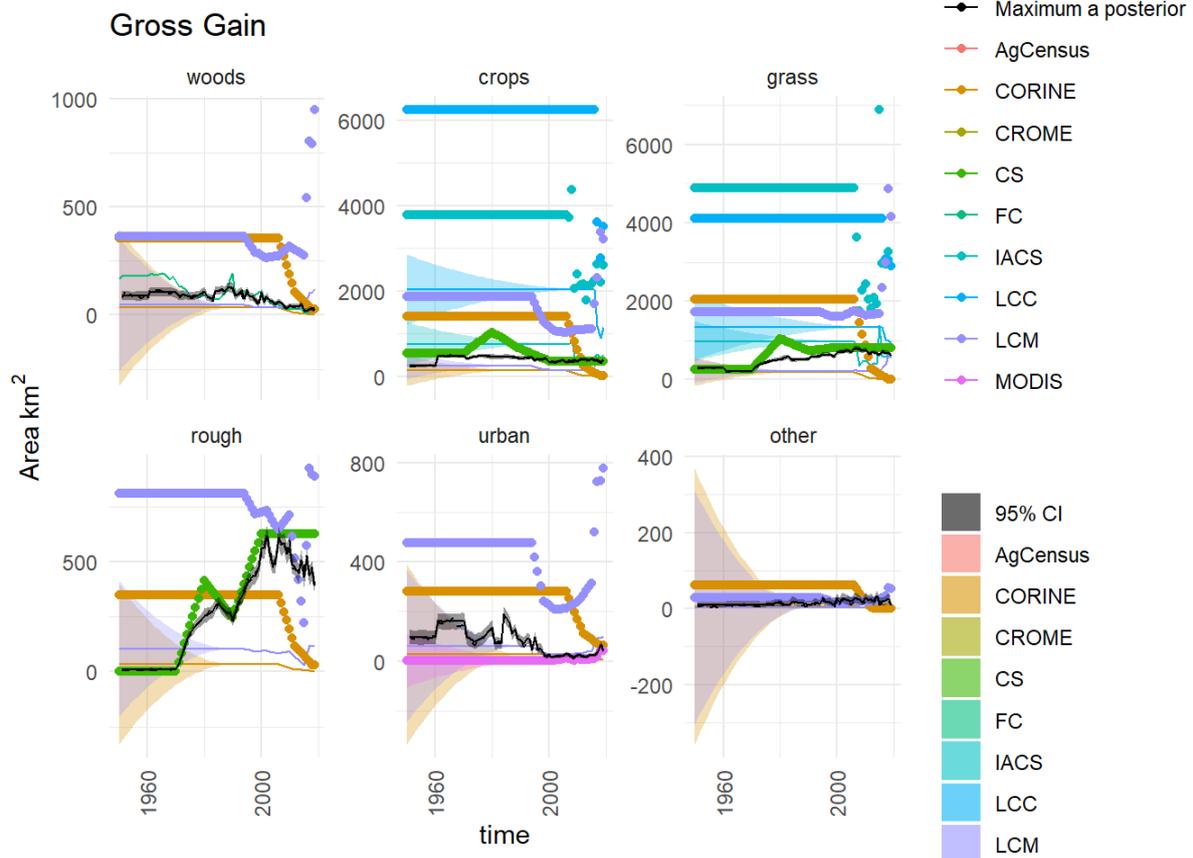


Figure 7.2: Observations and posterior distribution of the gross gain in area of each land use **G** from 1950 to 2020. The grey shaded band shows the 2.5 and 97.5 percentiles of the posterior distribution. The maximum *a posteriori* estimate is shown as the solid black line within this. Observations from the different data sources are shown as coloured circles. The coloured solid lines show the corrected observations after accounting for systematic uncertainties, and interpolating. The coloured bands around these lines show the random uncertainty, rescaled as $\sigma/5$ to keep with the axis scale. Because the random uncertainties and the corrections to the observations are generally very large in comparison to the actual change, scaling the axes is difficult. Note that a consistent colour scheme for the data sources is shown, but not all contribute to every figure.



Figure 7.3: Observations and posterior distribution of the gross loss of area from each land use **L** from 1950 to 2020. The grey shaded band shows the 2.5 and 97.5 percentiles of the posterior distribution. The maximum *a posteriori* estimate is shown as the solid black line within this. Observations from the different data sources are shown as coloured circles. The coloured solid lines show the corrected observations after accounting for systematic uncertainties, and interpolating. The coloured bands around these lines show the random uncertainty, rescaled as $\sigma/5$ to keep with the axis scale. Because the random uncertainties and the corrections to the observations are generally very large in comparison to the actual change, scaling the axes is difficult. Note that a consistent colour scheme for the data sources is shown, but not all contribute to every figure.

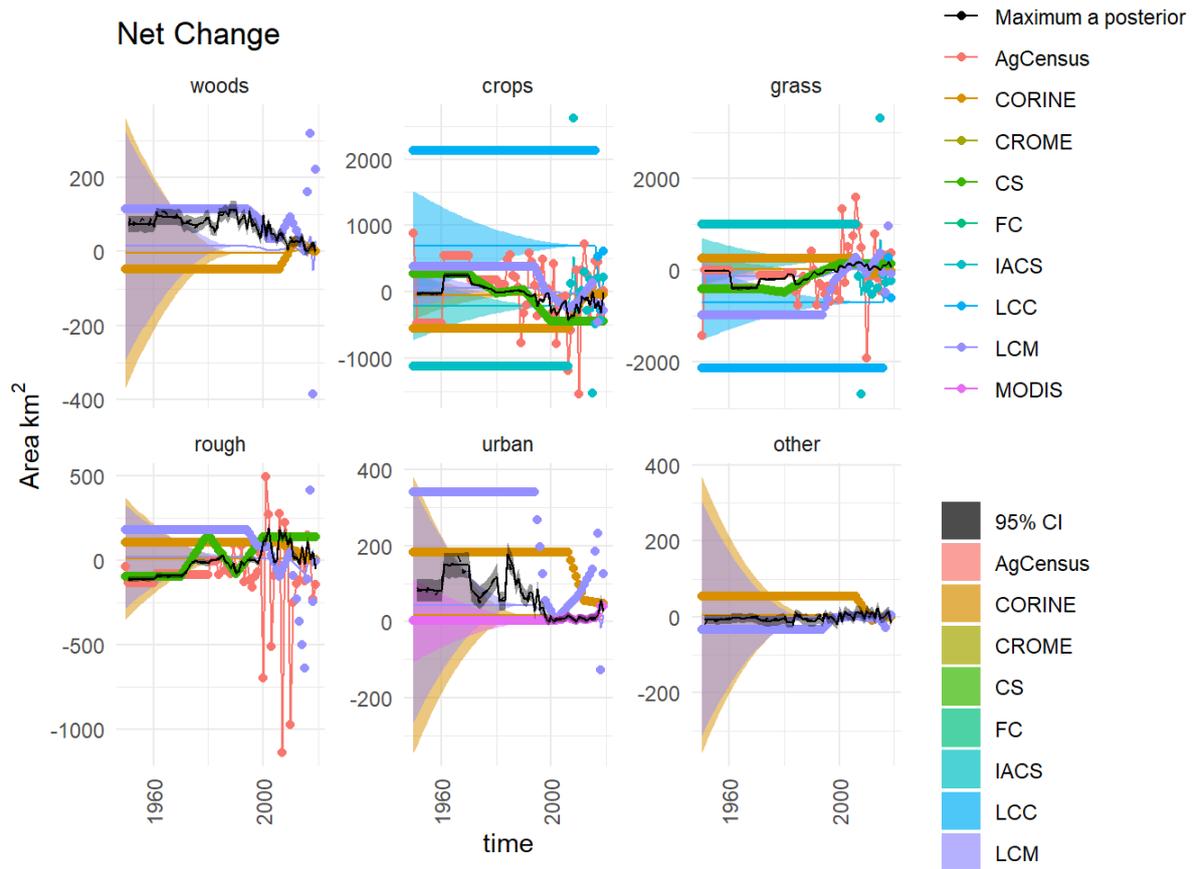


Figure 7.4: Time series of the net change in area occupied by each land use (D) from 1950 to 2020, showing the observations and posterior distribution of estimates. The grey shaded band shows the 2.5 and 97.5 percentiles of the posterior distribution. The maximum *a posteriori* estimate is shown as the solid black line within this. Observations from the different data sources are shown as coloured circles. The coloured solid lines show the corrected observations after accounting for systematic uncertainties, and interpolating. The coloured bands around these lines show the random uncertainty, rescaled as $\sigma/5$ to keep with the axis scale. Because the random uncertainties and the corrections to the observations are generally very large in comparison to the actual change, scaling the axes is difficult. Note that a consistent colour scheme for the data sources is shown, but not all contribute to every figure.

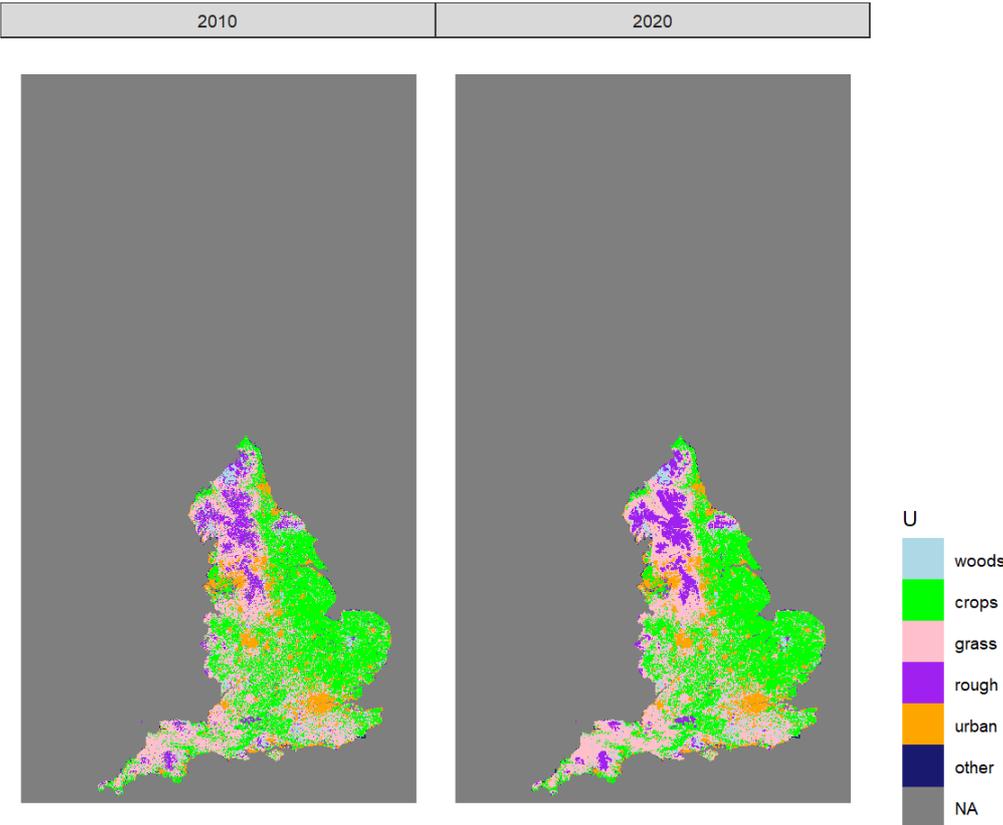


Figure 7.5: Estimated state of land-use in 2010 and 2020 in one realisation of U from the maximum *a posteriori* estimate of B .

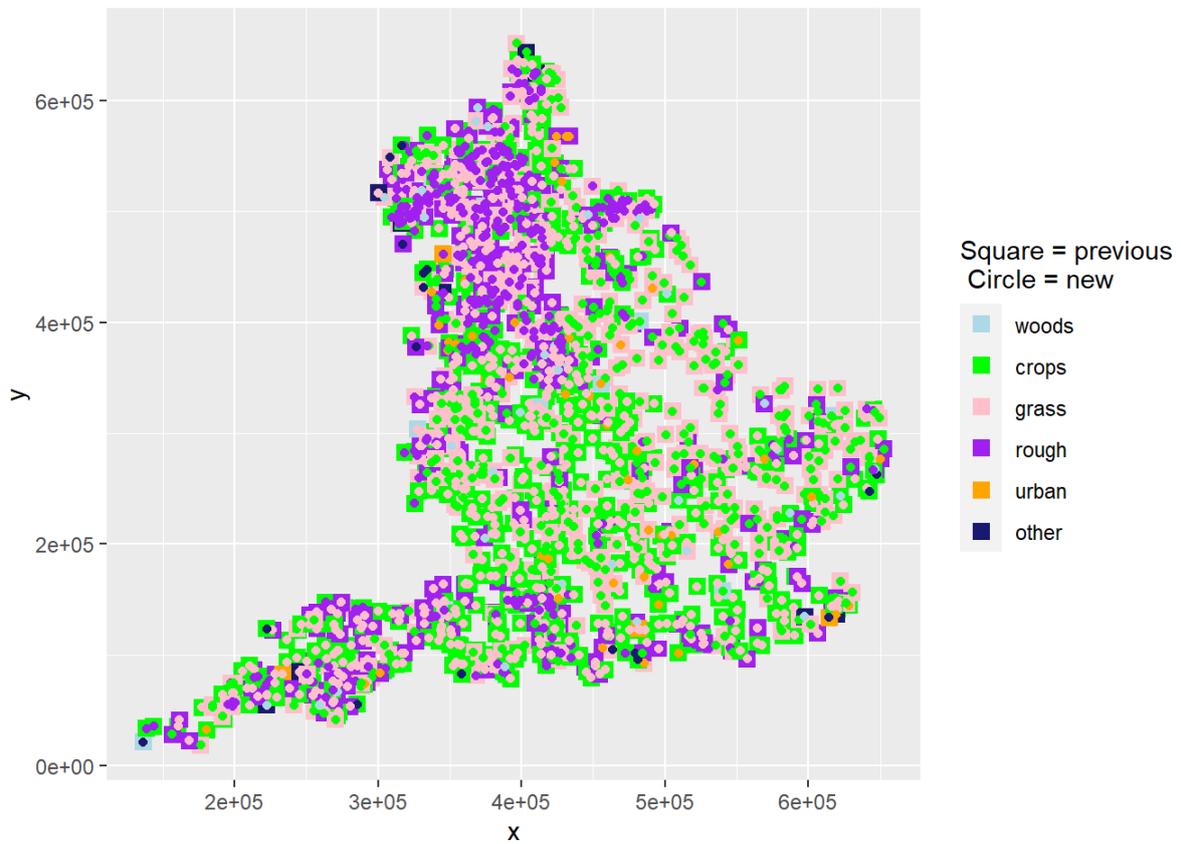


Figure 7.6: The spatial distribution of land-use change between 2010 and 2020 in one realisation of U from the maximum *a posteriori* estimate of B . At each location where land use has changed, the use in 2010 is shown as a coloured square, and the use in 2020 is shown as a coloured circle within this

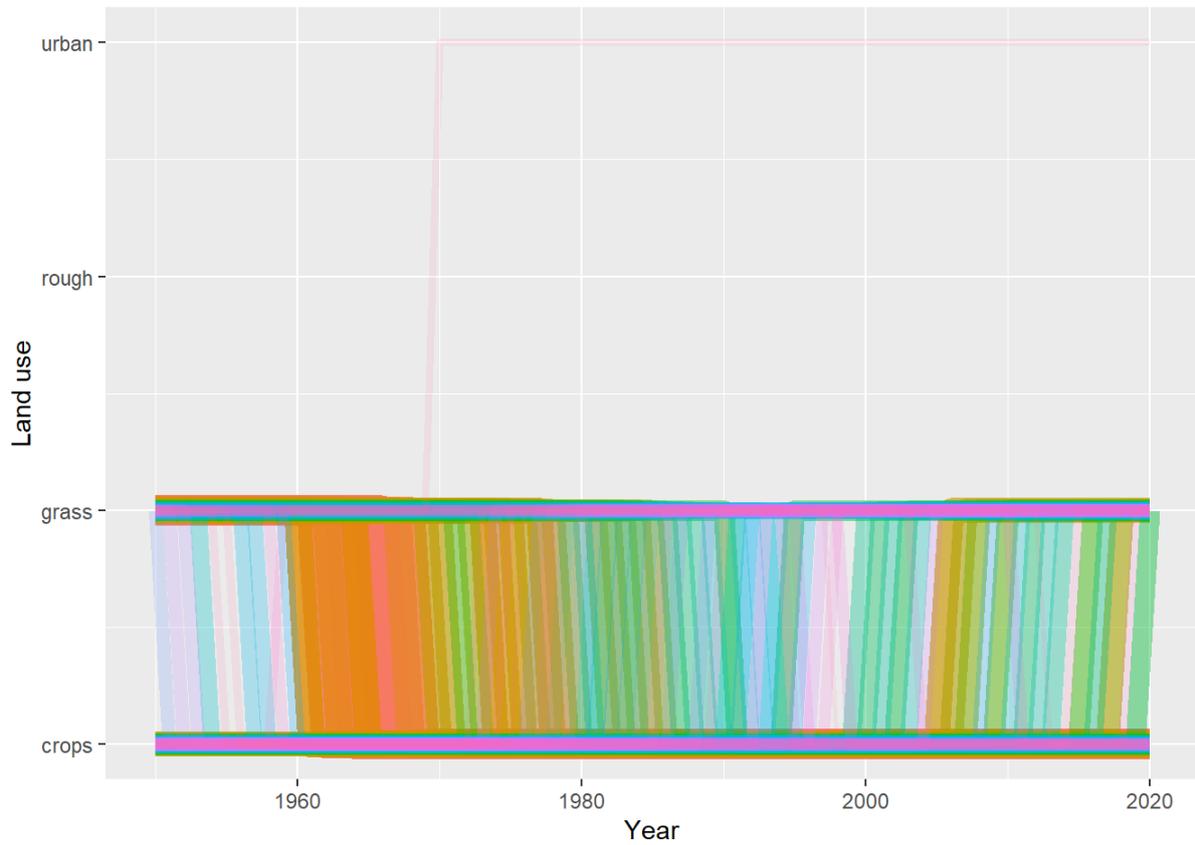


Figure 7.7: Trajectories of the 100 land-use vectors in the posterior U with the largest areas (excluding the six vectors which show no change). Each vector of land use is shown in a different colour, varied arbitrarily to differentiate different vectors. Line thickness and opacity are proportional to the total area occupied by each vector, so that the dominant vectors are the most visually obvious.

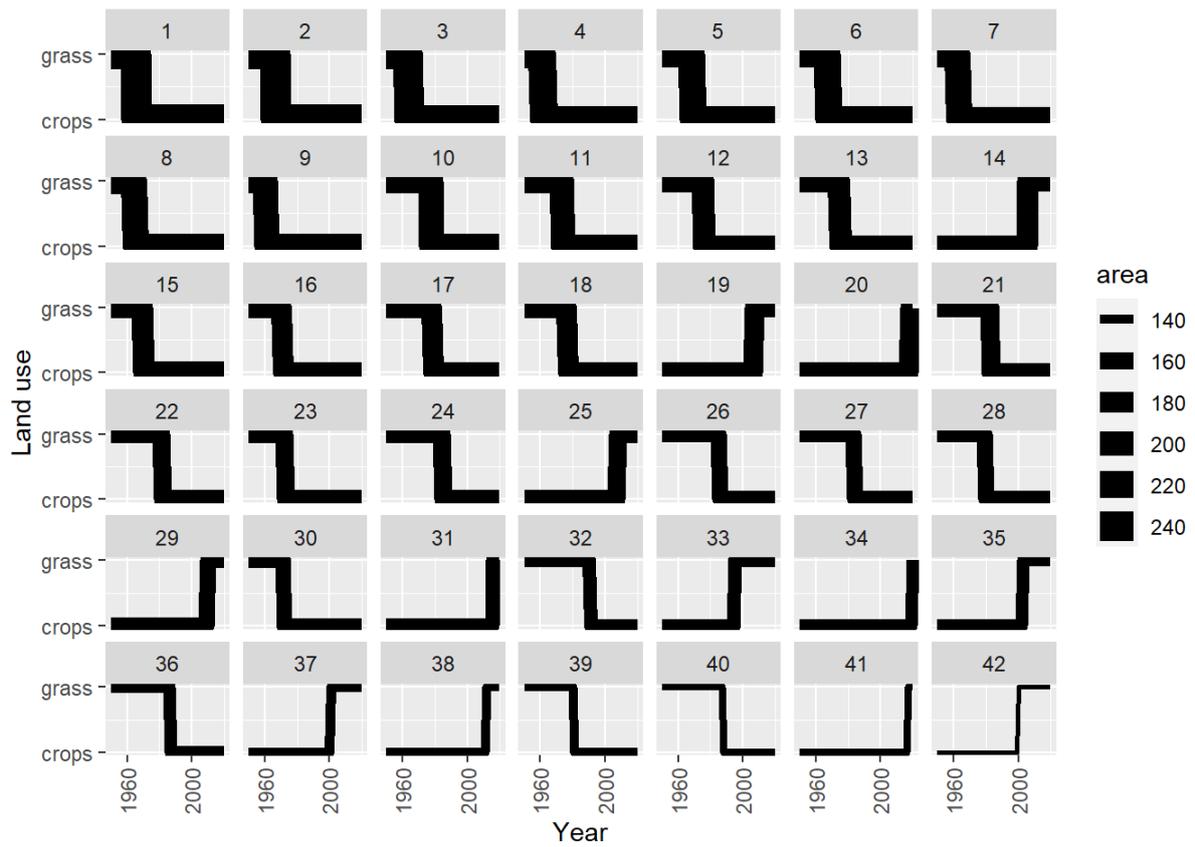


Figure 7.8: Trajectories of the 42 land-use vectors in the posterior U with the largest areas (excluding the six vectors which show no change). Line thickness is proportional to the total area occupied by each vector.

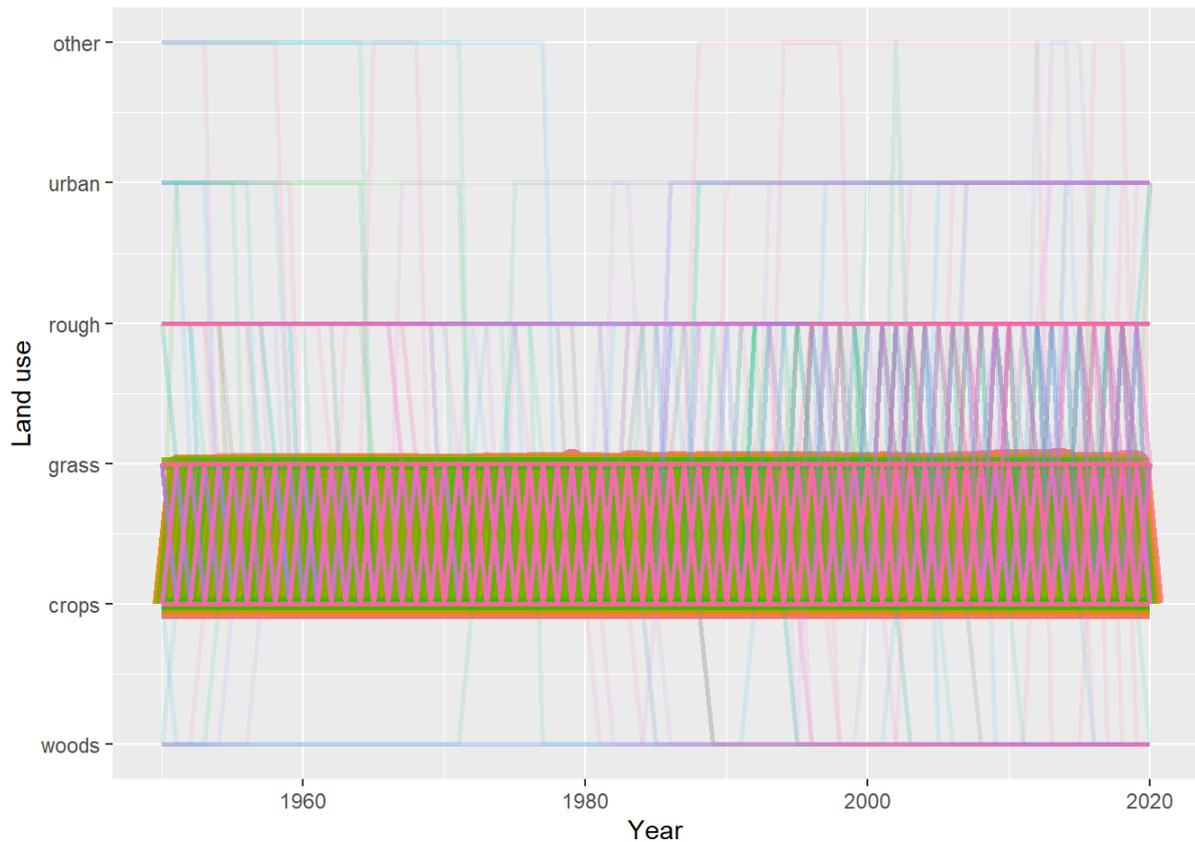


Figure 7.9: Trajectories of the land-use vectors in the posterior U which involve rotational change between crop and grassland (i.e. those which include either (i) transitions from crop to grass and then subsequently from grass to crop, or (ii) transitions from grass to crop and then subsequently from crop to grass). Each vector of land use is shown in a different colour, varied arbitrarily to differentiate different vectors. Line thickness and opacity are proportional to the total area occupied by each vector, so that the dominant vectors are the most visually obvious.

8 WP-A Initial Results: Scotland

The figures below show the results of the data assimilation procedure. All the data sets shown were used in the algorithm, but their relative random uncertainties (σ) determined how much influence they have on the estimates. The spatial data sets were corrected for systematic uncertainties, using the estimated net false positive rate (F_p). Having estimated the posterior distribution of the β matrix, we used this to simulate multiple maps of land use going back in time to 1950. The maps of the likelihood of transition to each land use established in WP-A were updated dynamically, using the life tables described in Section 5.

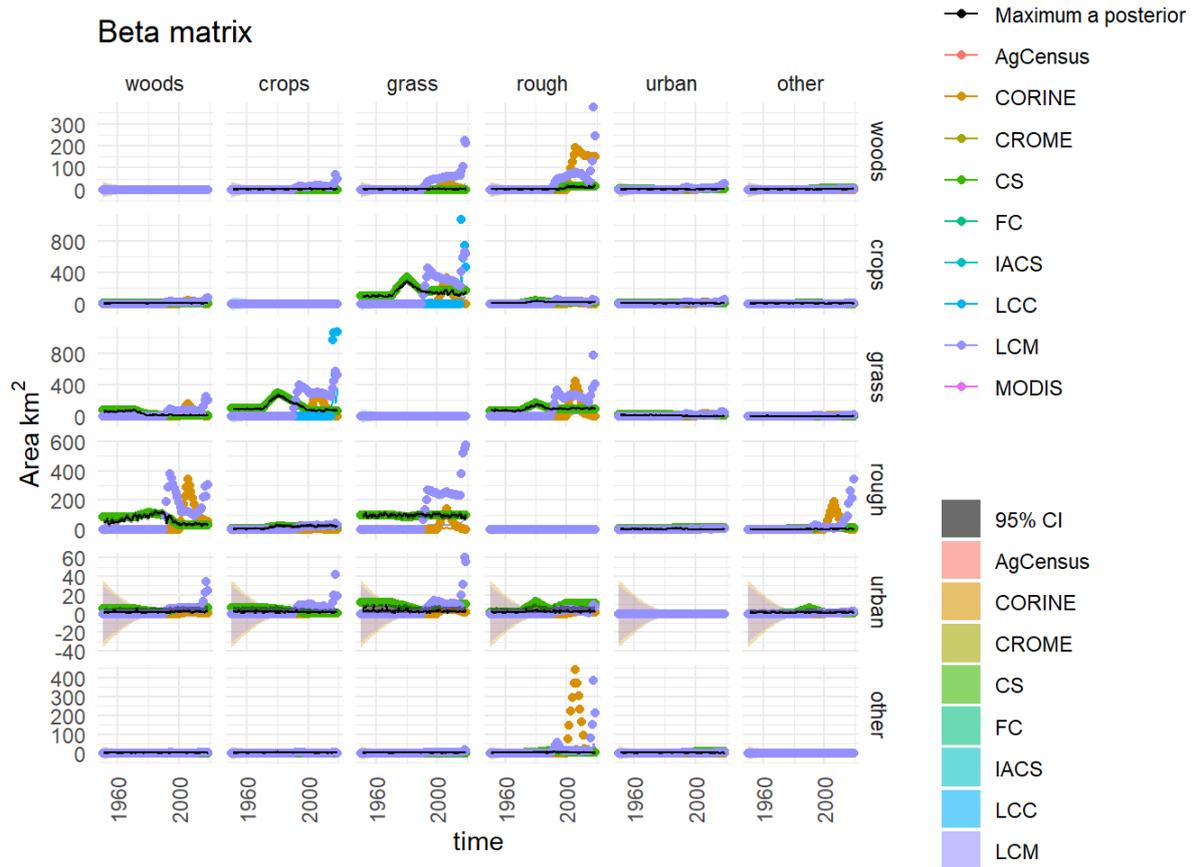


Figure 8.1: Observations and posterior distribution of the transition matrix B , representing the gross area changing from the land use in each row to the land use in each column each year from 1950 to 2020. The grey shaded band shows the 2.5 and 97.5 percentiles of the posterior distribution. The maximum *a posteriori* estimate is shown as the solid black line within this. Observations from the different data sources are shown as coloured circles. The coloured solid lines show the corrected observations after accounting for systematic uncertainties, and interpolating. The coloured bands around these lines show the random uncertainty, rescaled as $\sigma/5$ to keep with the axis scale. Because the random uncertainties and the corrections to the observations are generally very large in comparison to the actual change, scaling the axes is difficult. Note that a consistent colour scheme for the data sources is shown, but not all contribute to every figure.

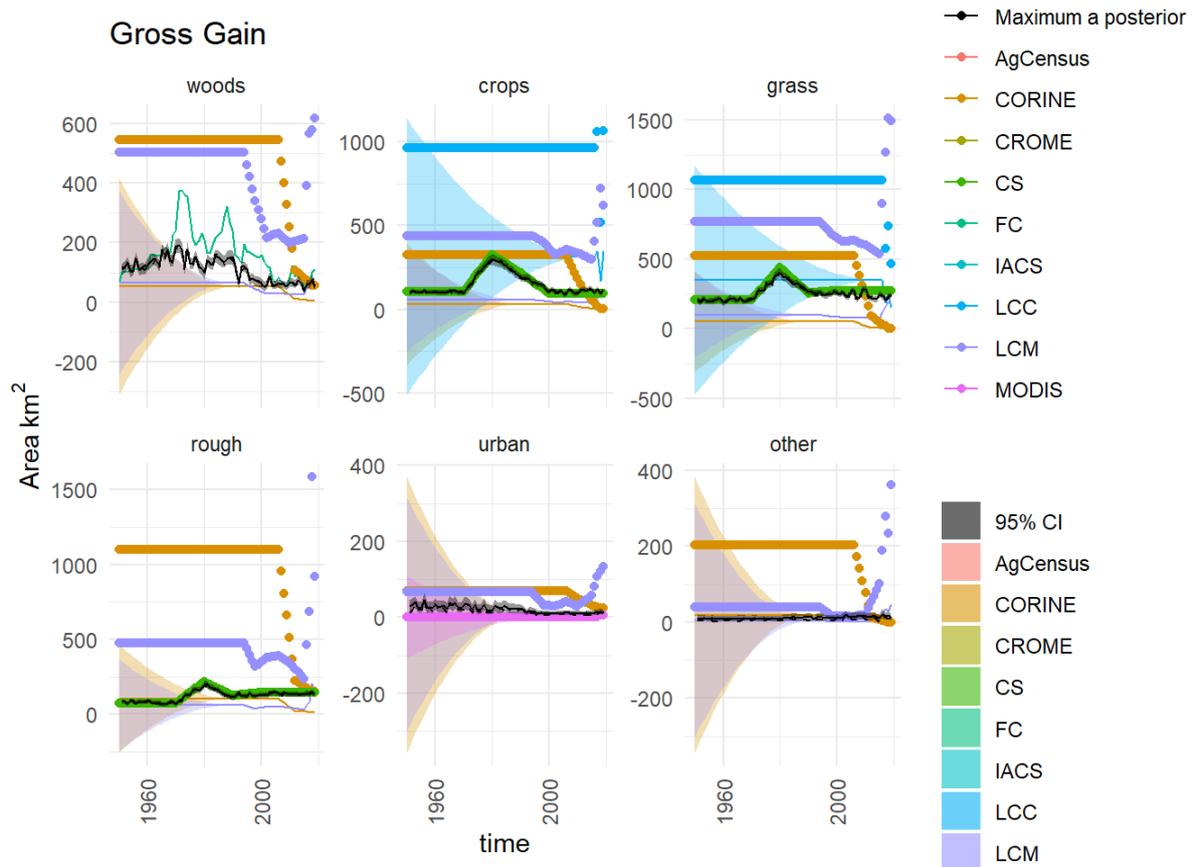


Figure 8.2: Observations and posterior distribution of the gross gain in area of each land use G from 1950 to 2020. The grey shaded band shows the 2.5 and 97.5 percentiles of the posterior distribution. The maximum *a posteriori* estimate is shown as the solid black line within this. Observations from the different data sources are shown as coloured circles. The coloured solid lines show the corrected observations after accounting for systematic uncertainties, and interpolating. The coloured bands around these lines show the random uncertainty, rescaled as $\sigma/5$ to keep with the axis scale. Because the random uncertainties and the corrections to the observations are generally very large in comparison to the actual change, scaling the axes is difficult. Note that a consistent colour scheme for the data sources is shown, but not all contribute to every figure.

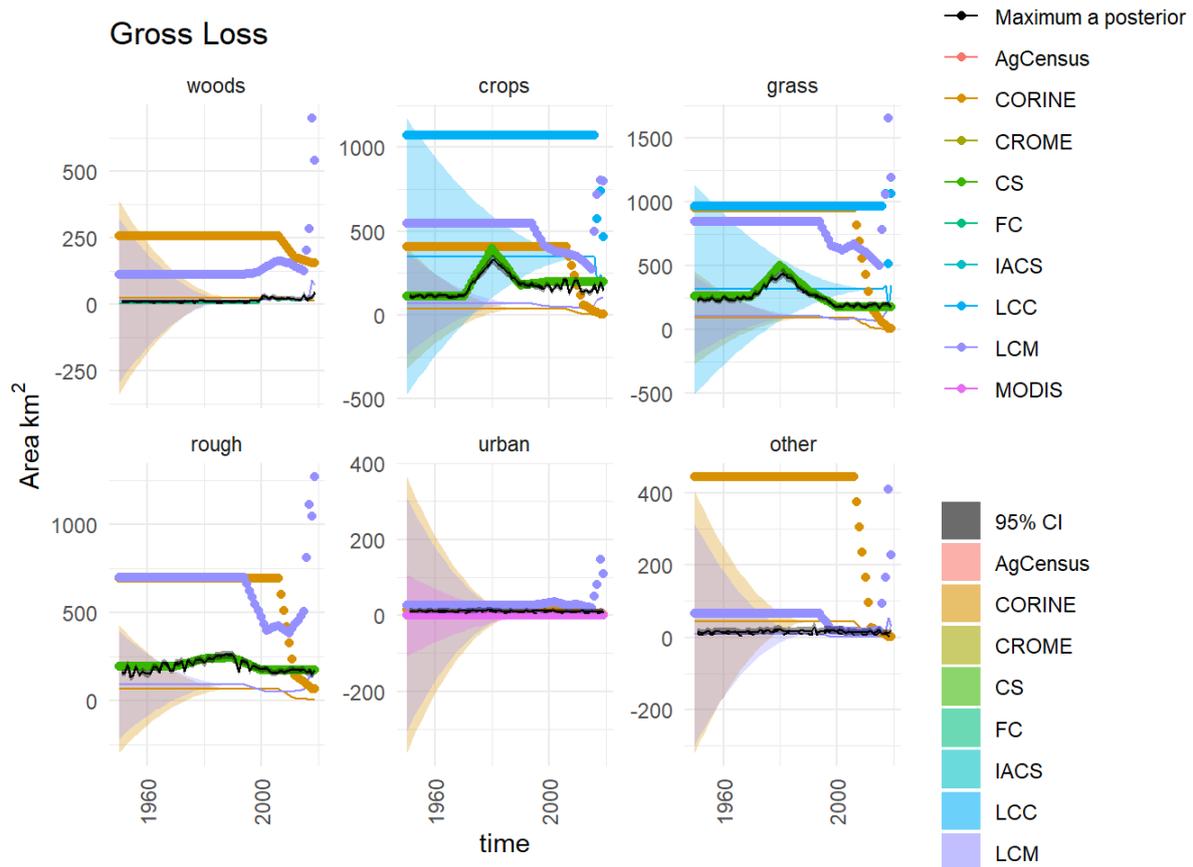


Figure 8.3: Observations and posterior distribution of the gross loss of area from each land use **L** from 1950 to 2020. The grey shaded band shows the 2.5 and 97.5 percentiles of the posterior distribution. The maximum *a posteriori* estimate is shown as the solid black line within this. Observations from the different data sources are shown as coloured circles. The coloured solid lines show the corrected observations after accounting for systematic uncertainties, and interpolating. The coloured bands around these lines show the random uncertainty, rescaled as $\sigma/5$ to keep with the axis scale. Because the random uncertainties and the corrections to the observations are generally very large in comparison to the actual change, scaling the axes is difficult. Note that a consistent colour scheme for the data sources is shown, but not all contribute to every figure.

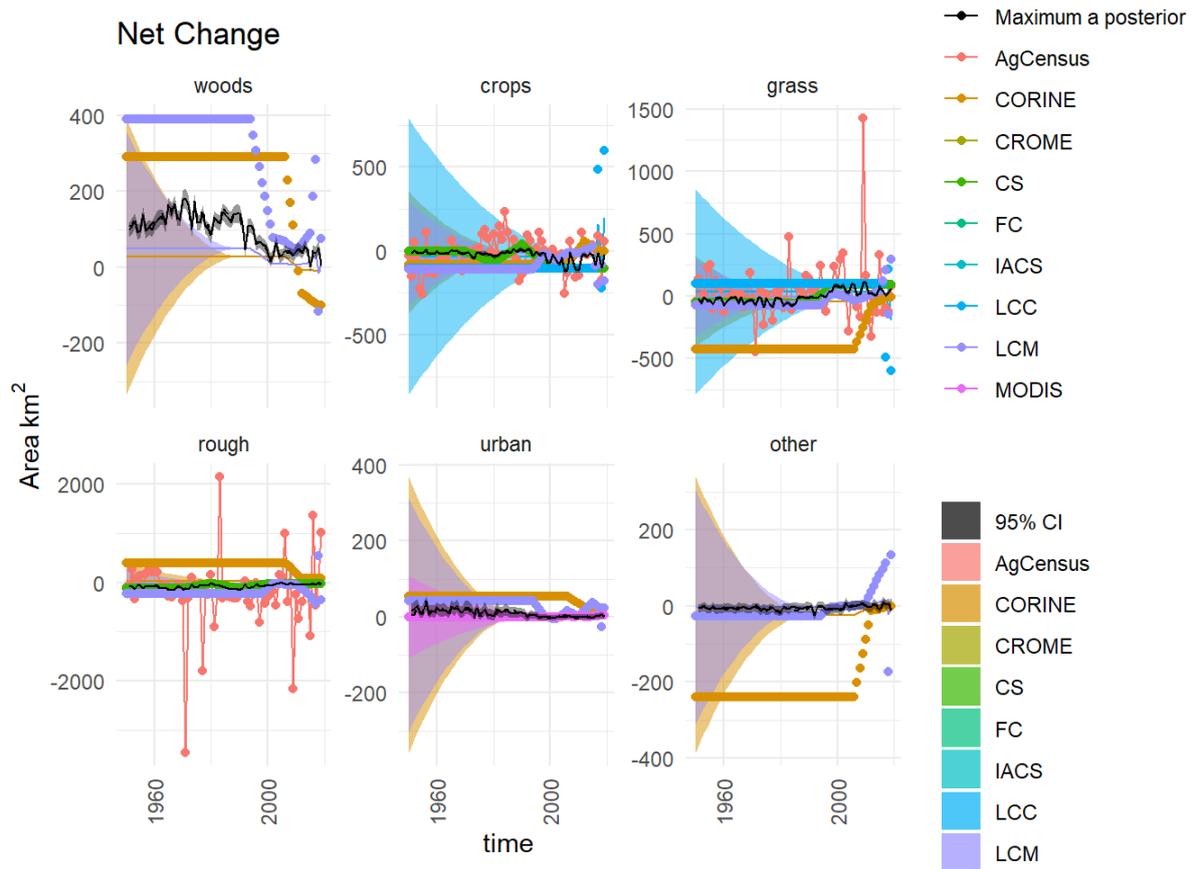


Figure 8.4: Time series of the net change in area occupied by each land use (*D*) from 1950 to 2020, showing the observations and posterior distribution of estimates. The grey shaded band shows the 2.5 and 97.5 percentiles of the posterior distribution. The maximum *a posteriori* estimate is shown as the solid black line within this. Observations from the different data sources are shown as coloured circles. The coloured solid lines show the corrected observations after accounting for systematic uncertainties, and interpolating. The coloured bands around these lines show the random uncertainty, rescaled as $\sigma/5$ to keep with the axis scale. Because the random uncertainties and the corrections to the observations are generally very large in comparison to the actual change, scaling the axes is difficult. Note that a consistent colour scheme for the data sources is shown, but not all contribute to every figure.

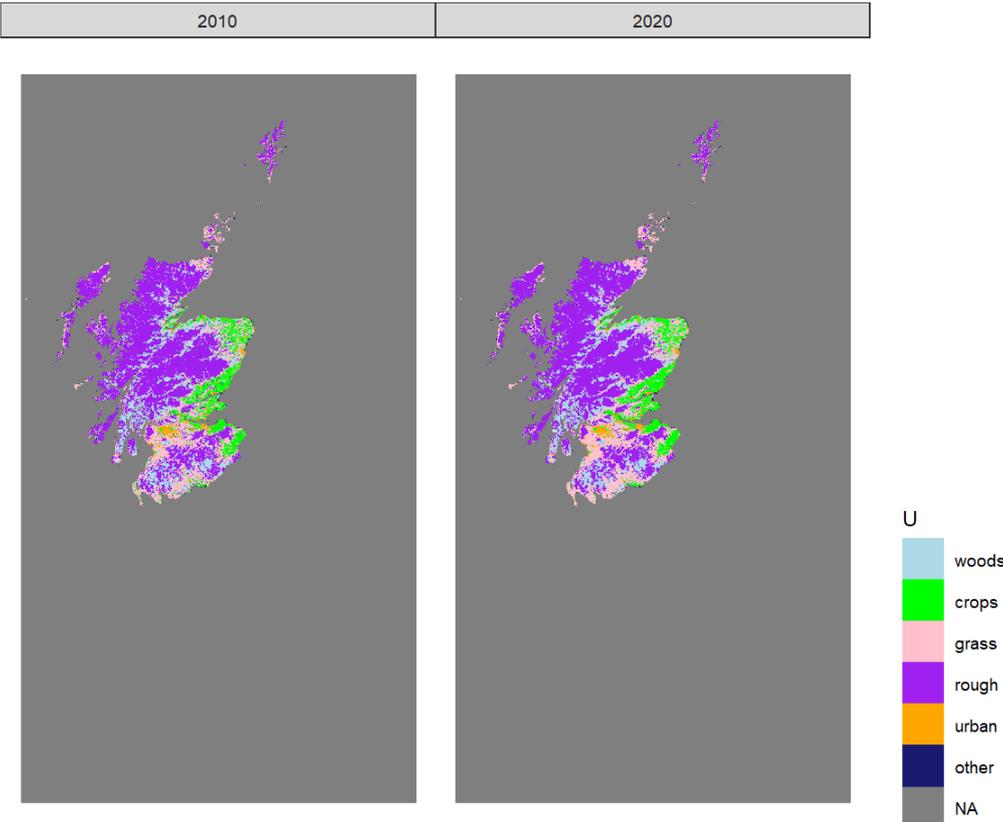


Figure 8.5: Estimated state of land-use in 2010 and 2020 in one realisation of U from the maximum *a posteriori* estimate of B .

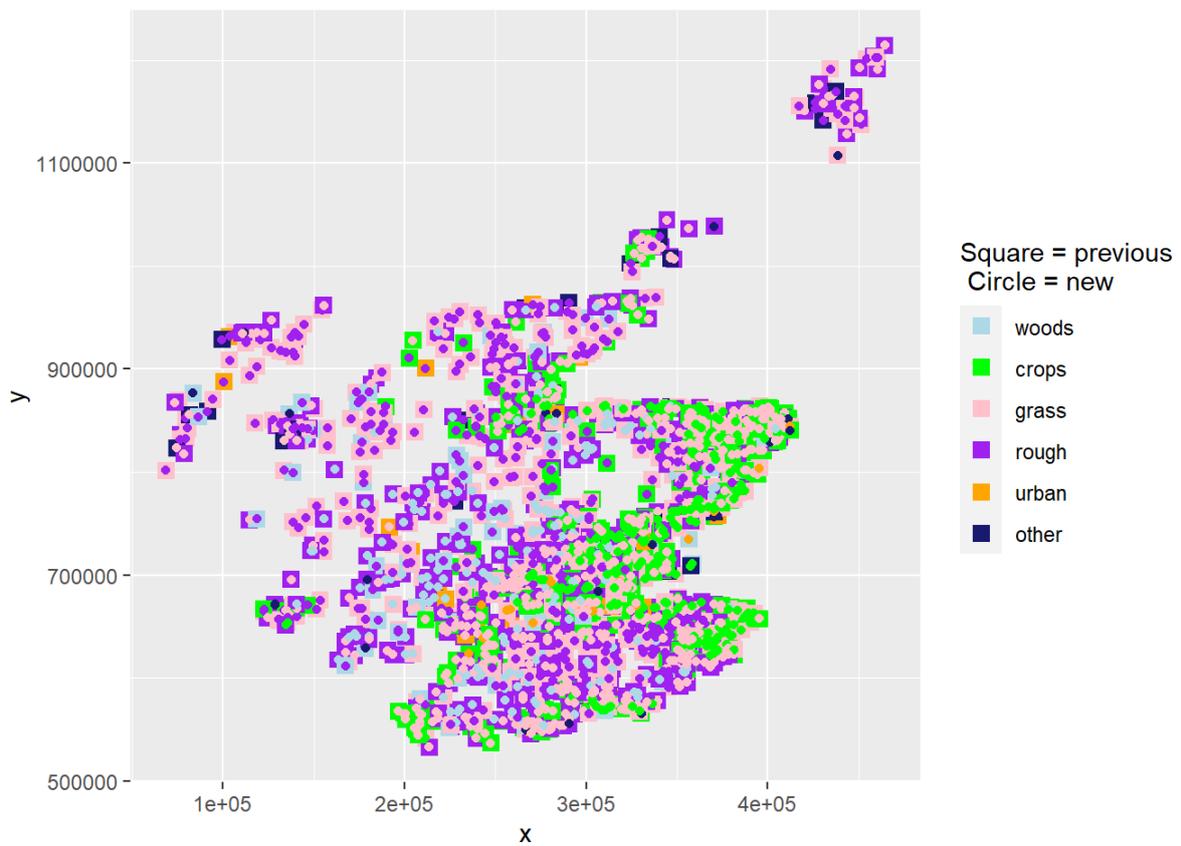


Figure 8.6: The spatial distribution of land-use change between 2010 and 2020 in one realisation of U from the maximum *a posteriori* estimate of B . At each location where land use has changed, the use in 2010 is shown as a coloured square, and the use in 2020 is shown as a coloured circle within this

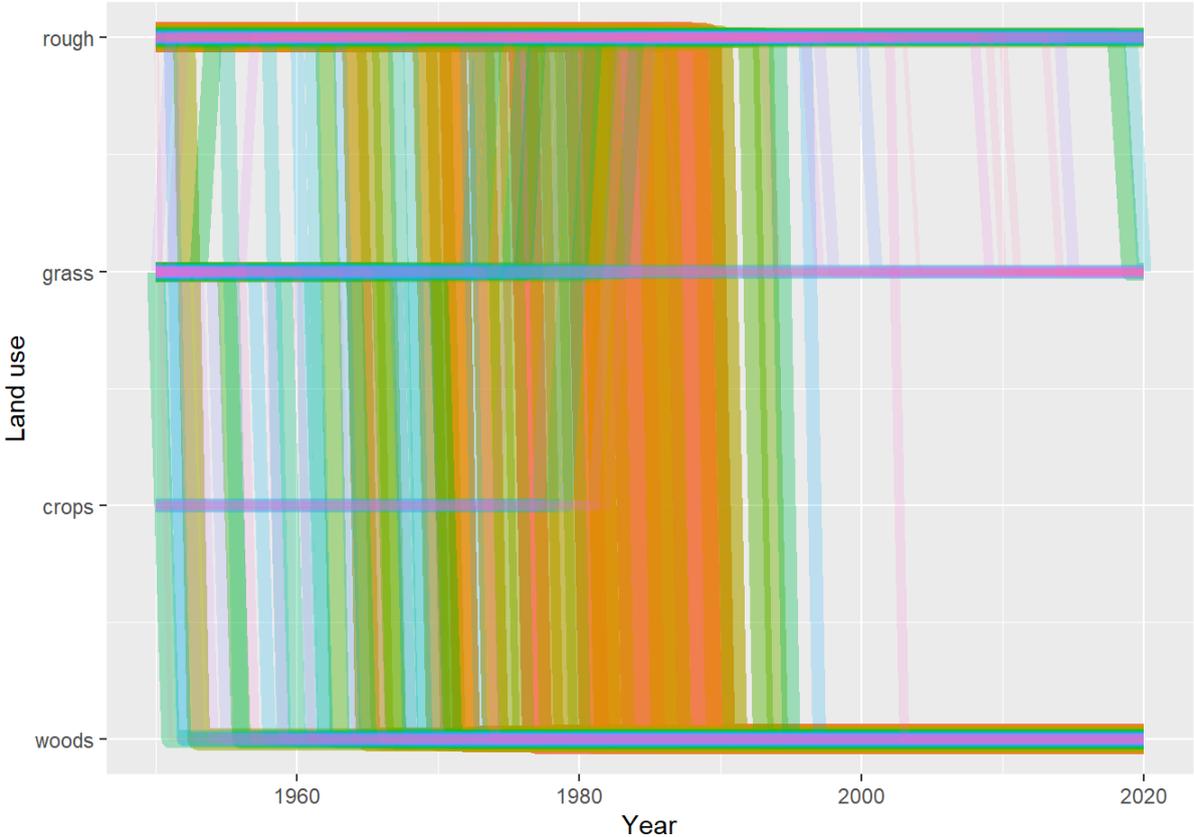


Figure 8.7: Trajectories of the 100 land-use vectors in the posterior U with the largest areas (excluding the six vectors which show no change). Each vector of land use is shown in a different colour, varied arbitrarily to differentiate different vectors. Line thickness and opacity are proportional to the total area occupied by each vector, so that the dominant vectors are the most visually obvious.

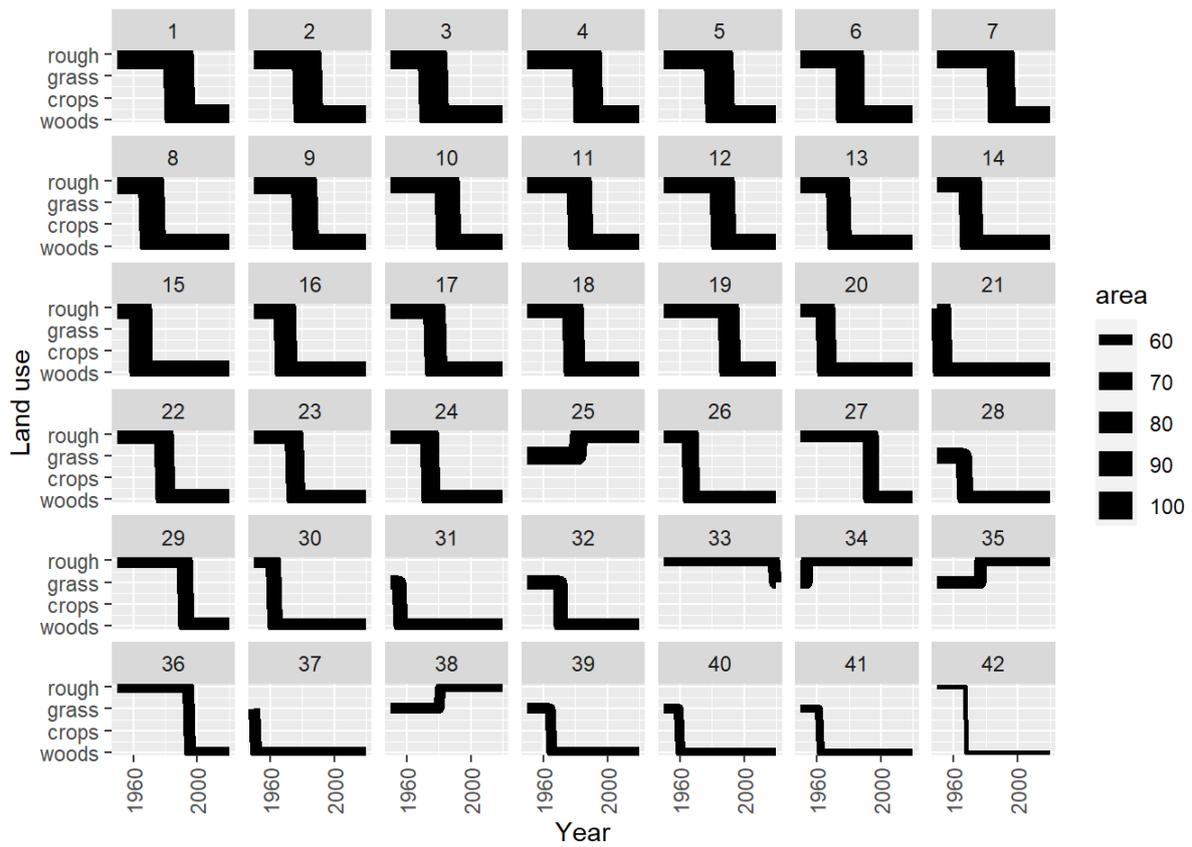


Figure 8.8: Trajectories of the 42 land-use vectors in the posterior U with the largest areas (excluding the six vectors which show no change). Line thickness is proportional to the total area occupied by each vector.

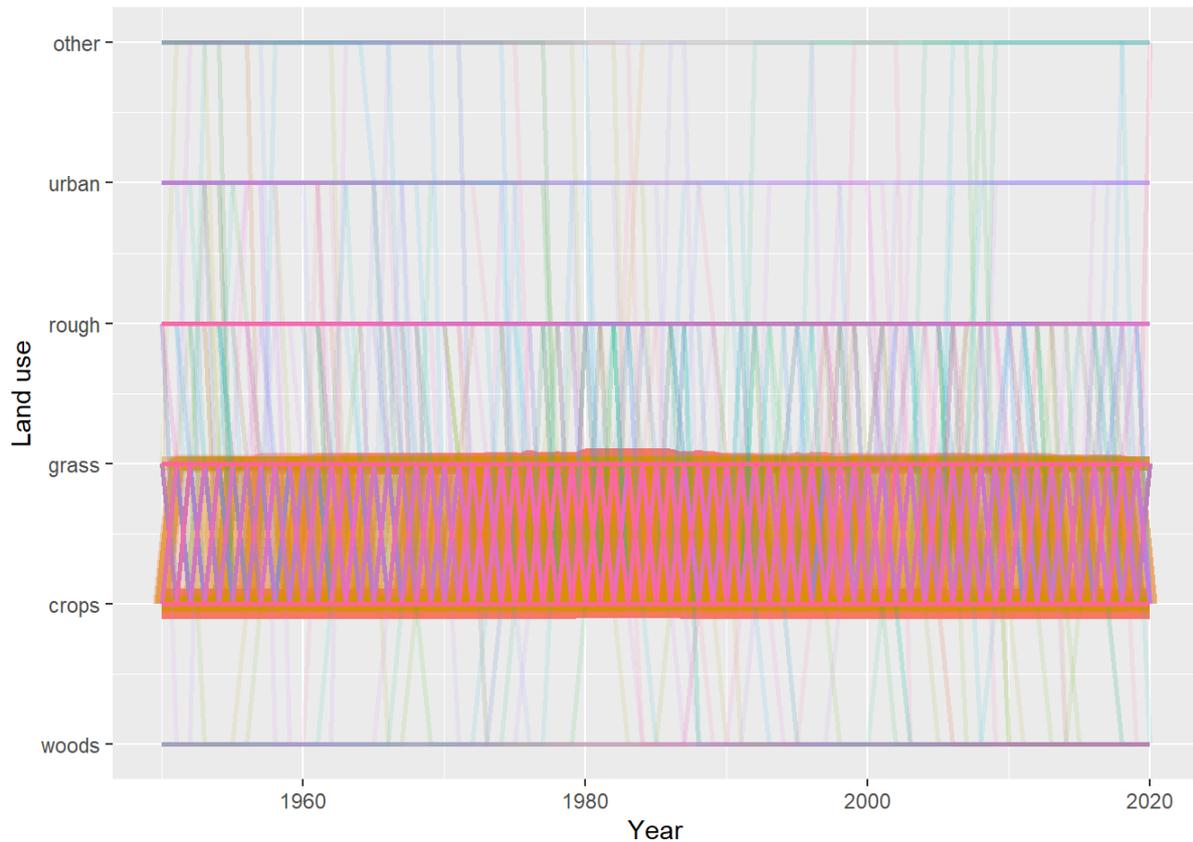


Figure 8.9: Trajectories of the land-use vectors in the posterior U which involve rotational change between crop and grassland (i.e. those which include either (i) transitions from crop to grass and then subsequently from grass to crop, or (ii) transitions from grass to crop and then subsequently from crop to grass). Each vector of land use is shown in a different colour, varied arbitrarily to differentiate different vectors. Line thickness and opacity are proportional to the total area occupied by each vector, so that the dominant vectors are the most visually obvious.

9 WP-A Initial Results: Wales

The figures below show the results of the data assimilation procedure. All the data sets shown were used in the algorithm, but their relative random uncertainties (σ) determined how much influence they have on the estimates. The spatial data sets were corrected for systematic uncertainties, using the estimated net false positive rate (F_p). Having estimated the posterior distribution of the β matrix, we used this to simulate multiple maps of land use going back in time to 1950. The maps of the likelihood of transition to each land use established in WP-A were updated dynamically, using the life tables described in Section 5.

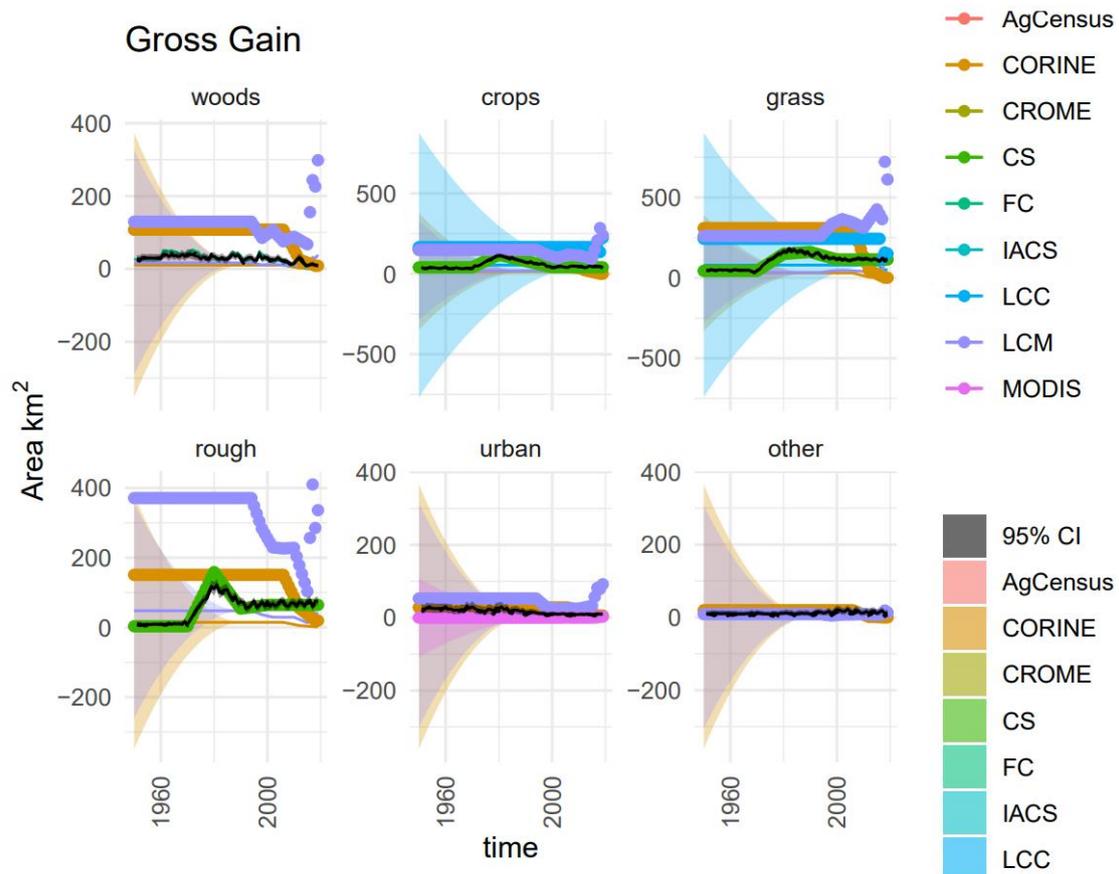


Figure 9.2: Observations and posterior distribution of the gross gain in area of each land use **G** from 1950 to 2020. The grey shaded band shows the 2.5 and 97.5 percentiles of the posterior distribution. The maximum *a posteriori* estimate is shown as the solid black line within this. Observations from the different data sources are shown as coloured circles. The coloured solid lines show the corrected observations after accounting for systematic uncertainties, and interpolating. The coloured bands around these lines show the random uncertainty, rescaled as $\sigma/5$ to keep with the axis scale. Because the random uncertainties and the corrections to the observations are generally very large in comparison to the actual change, scaling the axes is difficult. Note that a consistent colour scheme for the data sources is shown, but not all contribute to every figure.

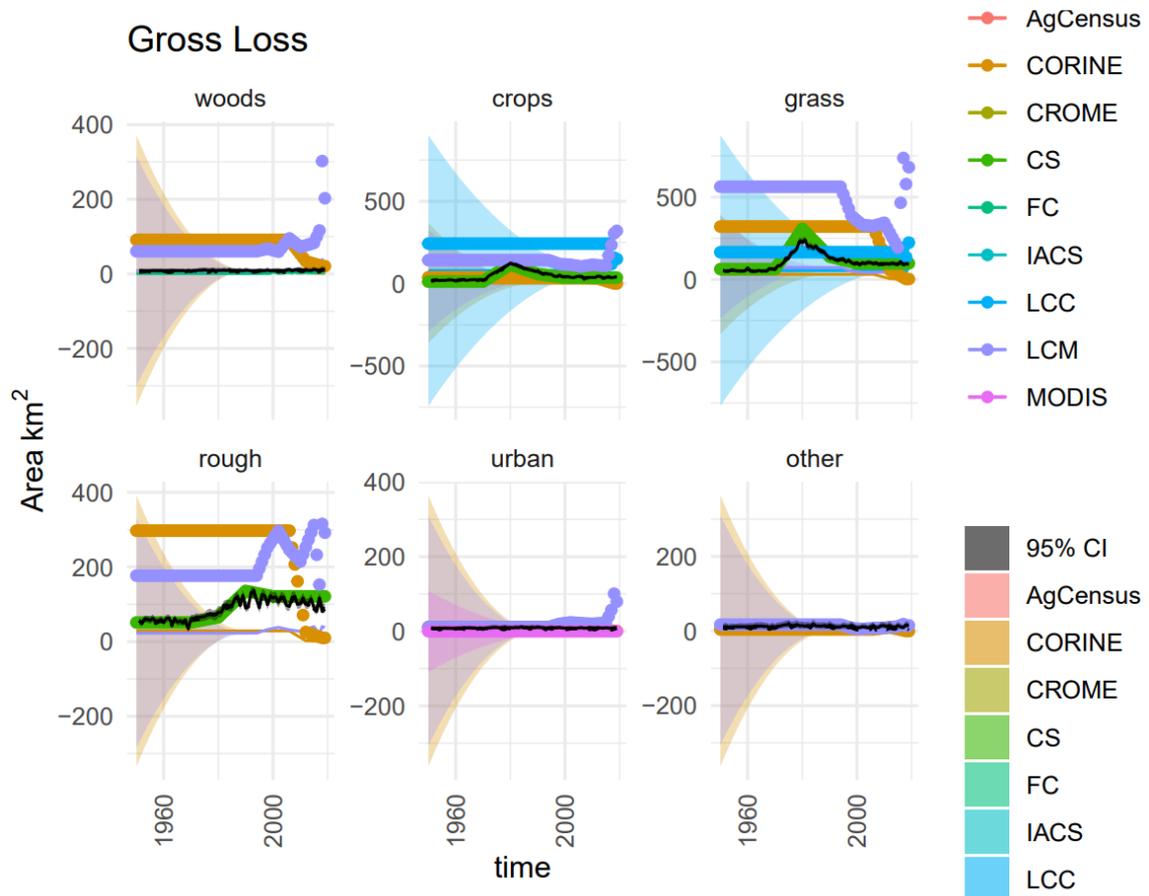


Figure 9.3: Observations and posterior distribution of the gross loss of area from each land use L from 1950 to 2020. The grey shaded band shows the 2.5 and 97.5 percentiles of the posterior distribution. The maximum *a posteriori* estimate is shown as the solid black line within this. Observations from the different data sources are shown as coloured circles. The coloured solid lines show the corrected observations after accounting for systematic uncertainties, and interpolating. The coloured bands around these lines show the random uncertainty, rescaled as $\sigma/5$ to keep with the axis scale. Because the random uncertainties and the corrections to the observations are generally very large in comparison to the actual change, scaling the axes is difficult. Note that a consistent colour scheme for the data sources is shown, but not all contribute to every figure.

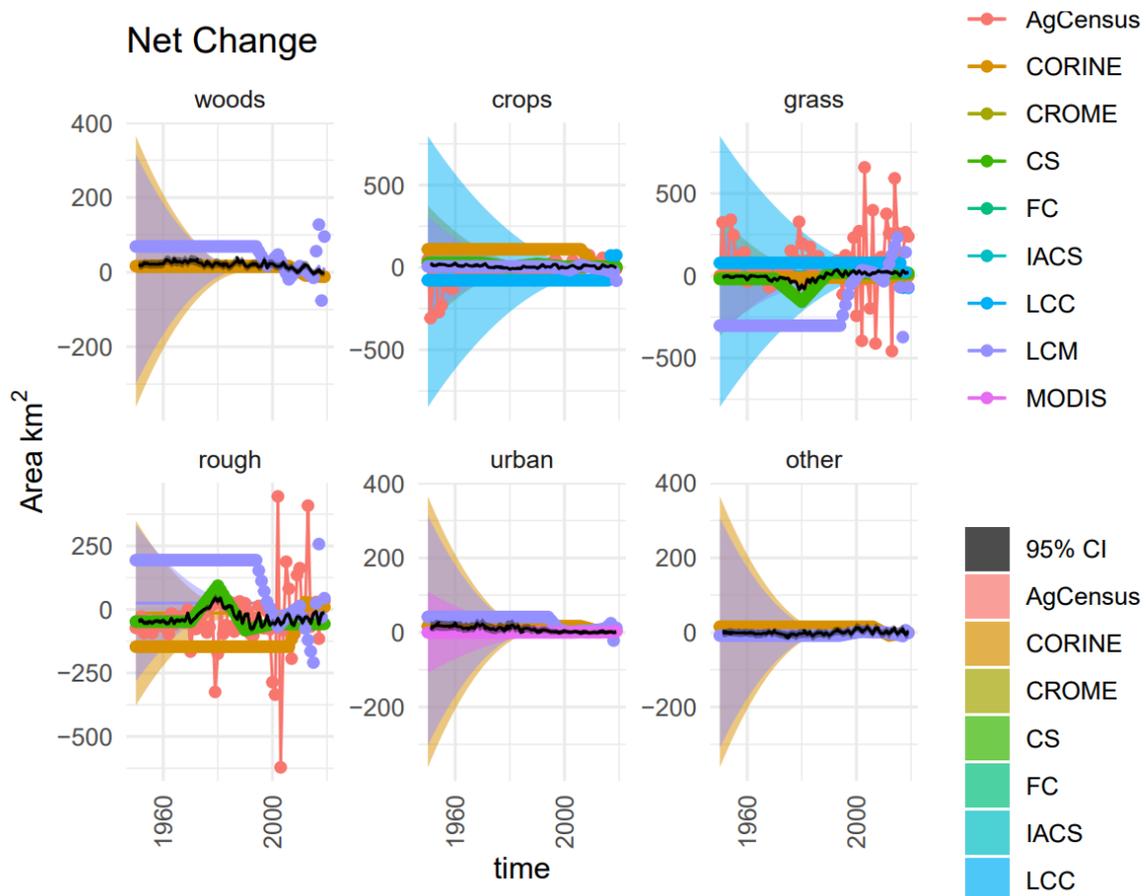


Figure 9.4: Time series of the net change in area occupied by each land use (D) from 1950 to 2020, showing the observations and posterior distribution of estimates. The grey shaded band shows the 2.5 and 97.5 percentiles of the posterior distribution. The maximum *a posteriori* estimate is shown as the solid black line within this. Observations from the different data sources are shown as coloured circles. The coloured solid lines show the corrected observations after accounting for systematic uncertainties, and interpolating. The coloured bands around these lines show the random uncertainty, rescaled as $\sigma/5$ to keep with the axis scale. Because the random uncertainties and the corrections to the observations are generally very large in comparison to the actual change, scaling the axes is difficult. Note that a consistent colour scheme for the data sources is shown, but not all contribute to every figure.

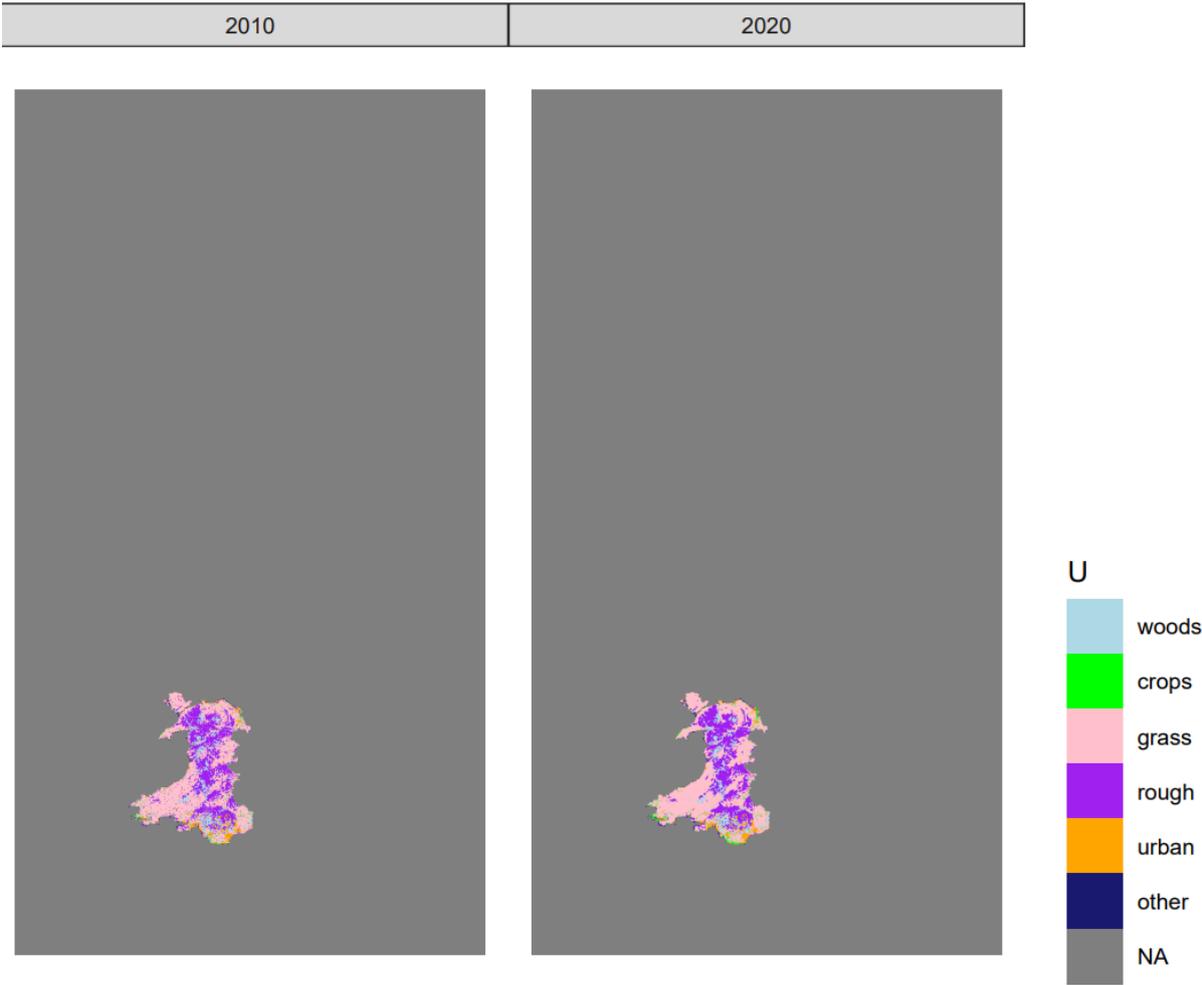


Figure 9.5: Estimated state of land-use in 2010 and 2020 in one realisation of U from the maximum *a posteriori* estimate of B .

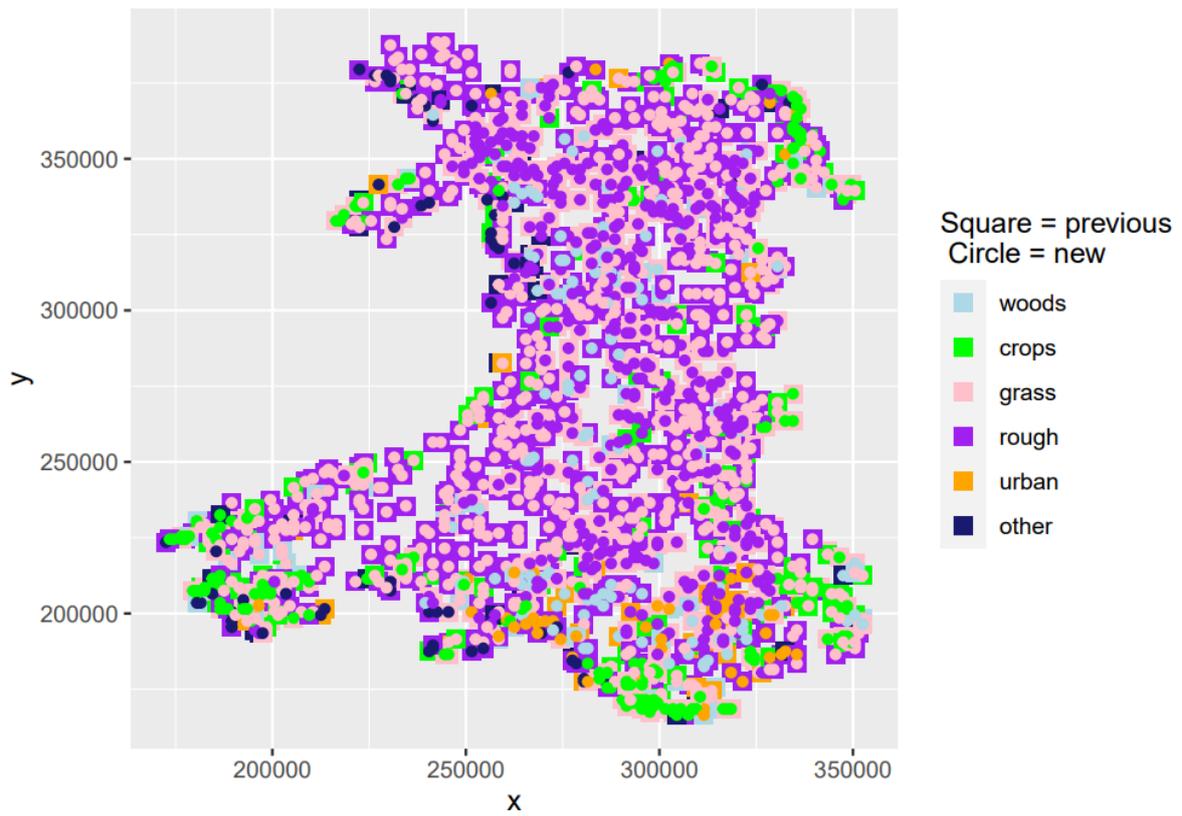


Figure 9.6: The spatial distribution of land-use change between 2010 and 2020 in one realisation of U from the maximum a posteriori estimate of B . At each location where land use has changed, the use in 2010 is shown as a coloured square, and the use in 2020 is shown as a coloured circle within this

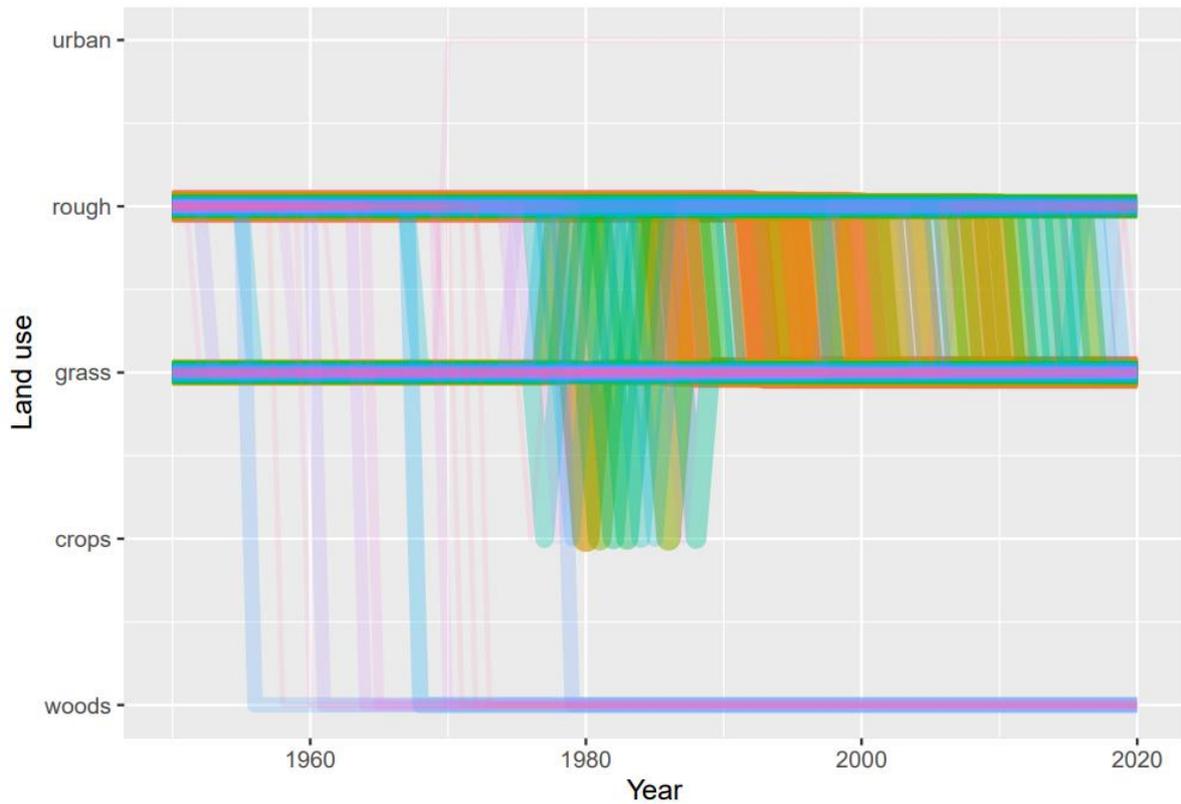


Figure 9.7: Trajectories of the 100 land-use vectors in the posterior U with the largest areas (excluding the six vectors which show no change). Each vector of land use is shown in a different colour, varied arbitrarily to differentiate different vectors. Line thickness and opacity are proportional to the total area occupied by each vector, so that the dominant vectors are the most visually obvious.

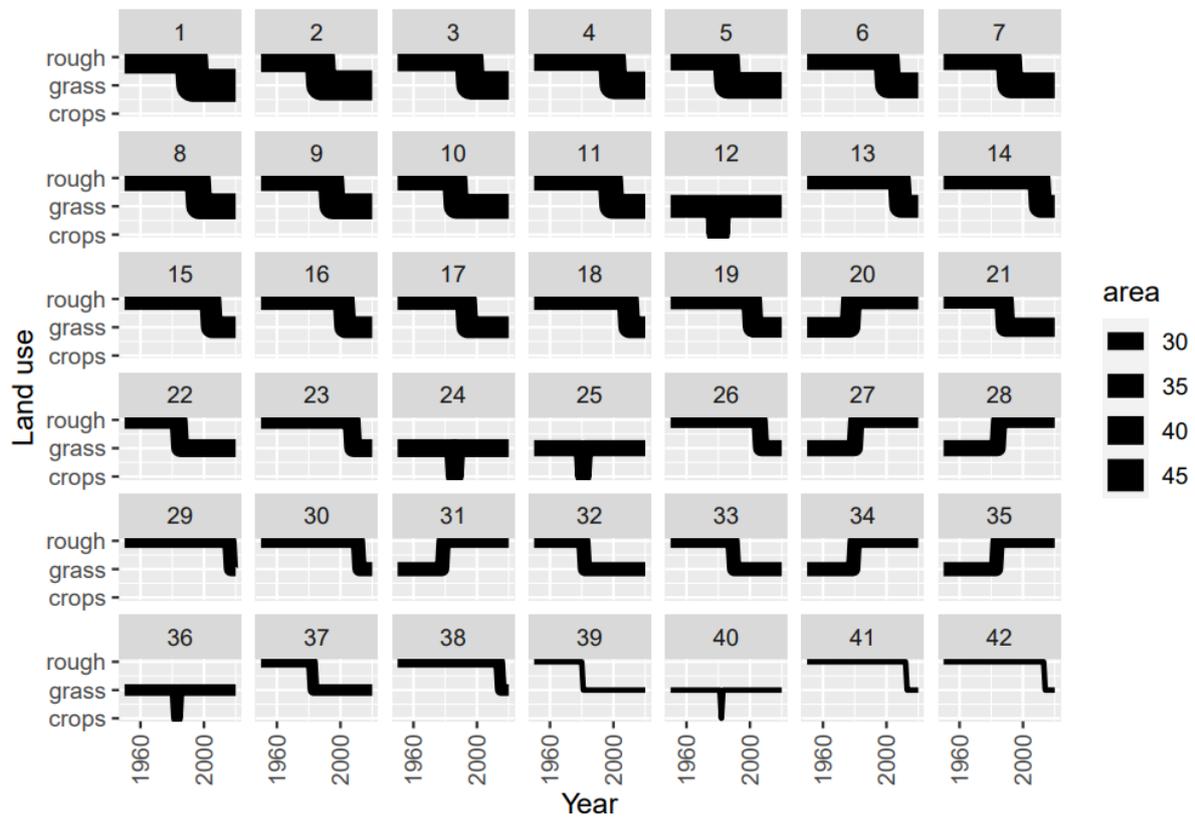


Figure 9.8: Trajectories of the 42 land-use vectors in the posterior U with the largest areas (excluding the six vectors which show no change). Line thickness is proportional to the total area occupied by each vector.

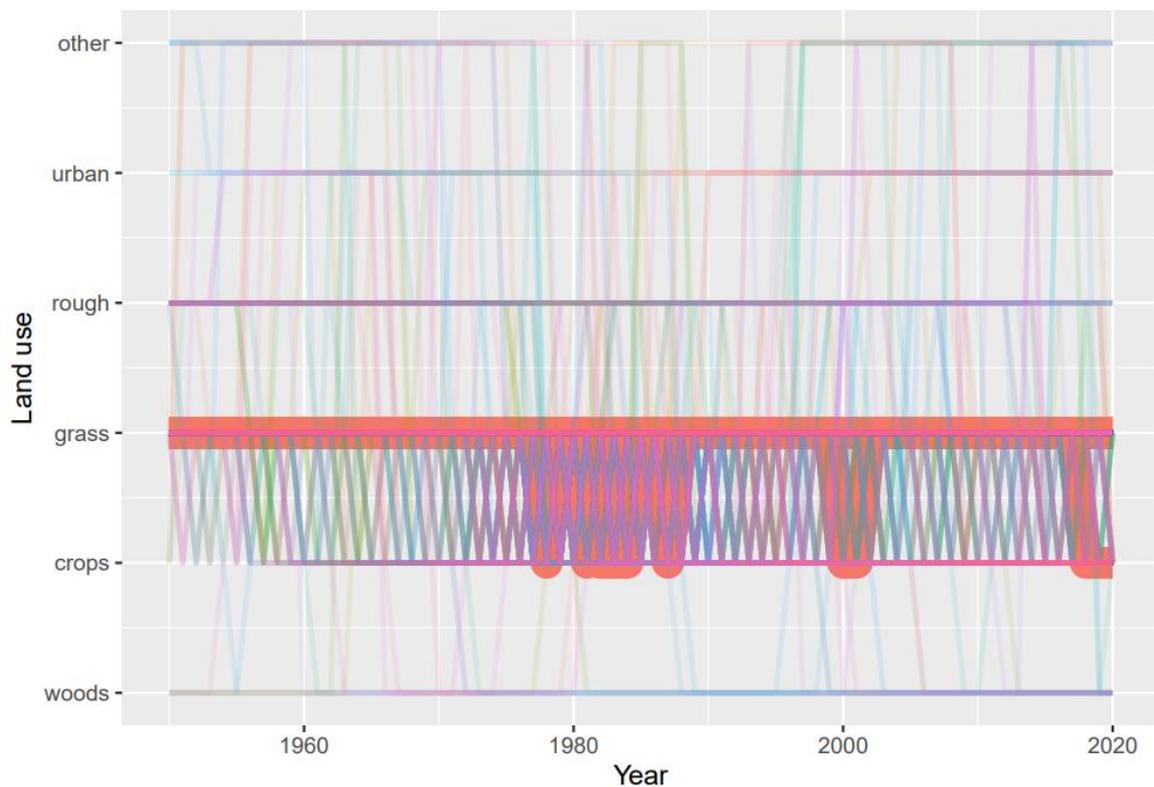


Figure 9.9: Trajectories of the land-use vectors in the posterior U which involve rotational change between crop and grassland (i.e. those which include either (i) transitions from crop to grass and then subsequently from grass to crop, or (ii) transitions from grass to crop and then subsequently from crop to grass). Each vector of land use is shown in a different colour, varied arbitrarily to differentiate different vectors. Line thickness and opacity are proportional to the total area occupied by each vector, so that the dominant vectors are the most visually obvious.

10 WP-A Initial Results: Northern Ireland

The figures below show the results of the data assimilation procedure. All the data sets shown were used in the algorithm, but their relative random uncertainties (σ) determined how much influence they have on the estimates. The spatial data sets were corrected for systematic uncertainties, using the estimated net false positive rate (F_p). Having estimated the posterior distribution of the β matrix, we used this to simulate multiple maps of land use going back in time to 1950. The maps of the likelihood of transition to each land use established in WP-A were updated dynamically, using the life tables described in Section 5.

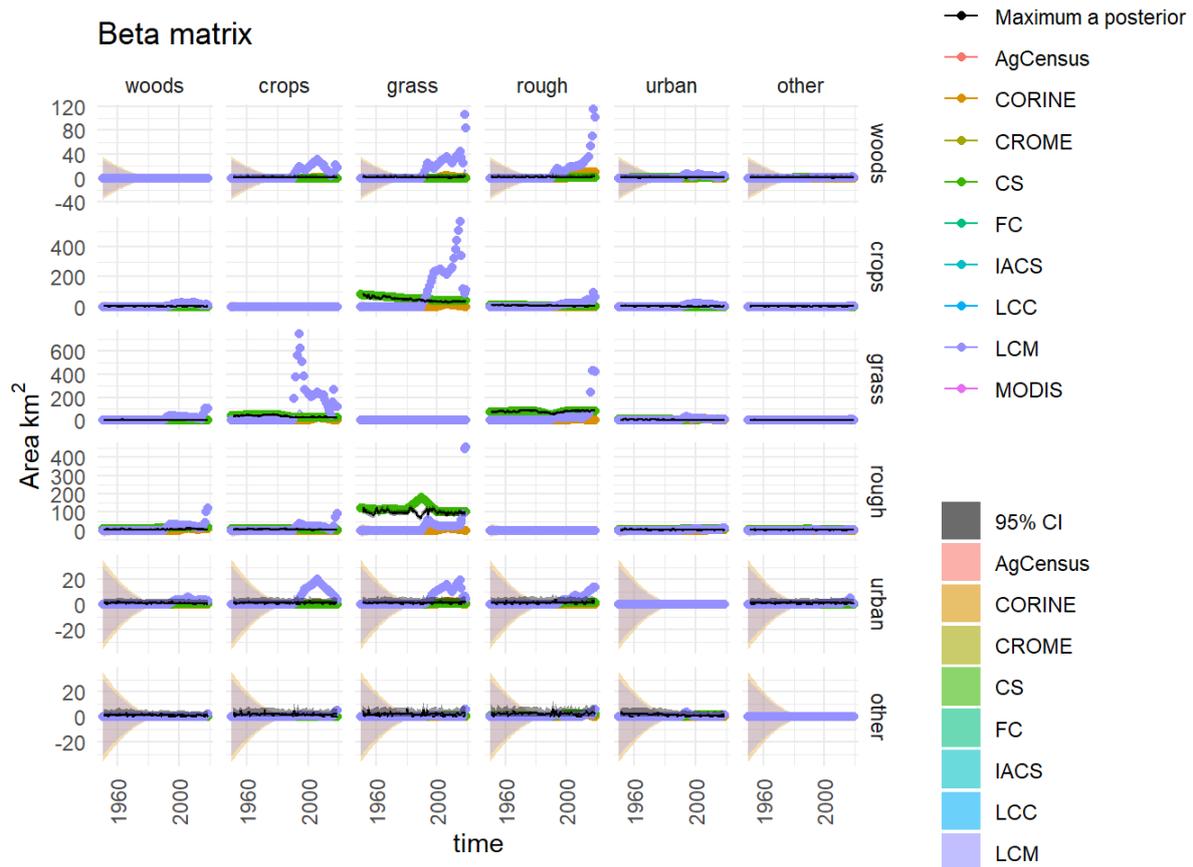


Figure 10.1: Observations and posterior distribution of the transition matrix **B**, representing the gross area changing from the land use in each row to the land use in each column each year from 1950 to 2020. The grey shaded band shows the 2.5 and 97.5 percentiles of the posterior distribution. The maximum *a posteriori* estimate is shown as the solid black line within this. Observations from the different data sources are shown as coloured circles. The coloured solid lines show the corrected observations after accounting for systematic uncertainties, and interpolating. The coloured bands around these lines show the random uncertainty, rescaled as $\sigma/5$ to keep with the axis scale. Because the random uncertainties and the corrections to the observations are generally very large in comparison to the actual change, scaling the axes is difficult. Note that a consistent colour scheme for the data sources is shown, but not all contribute to every figure.

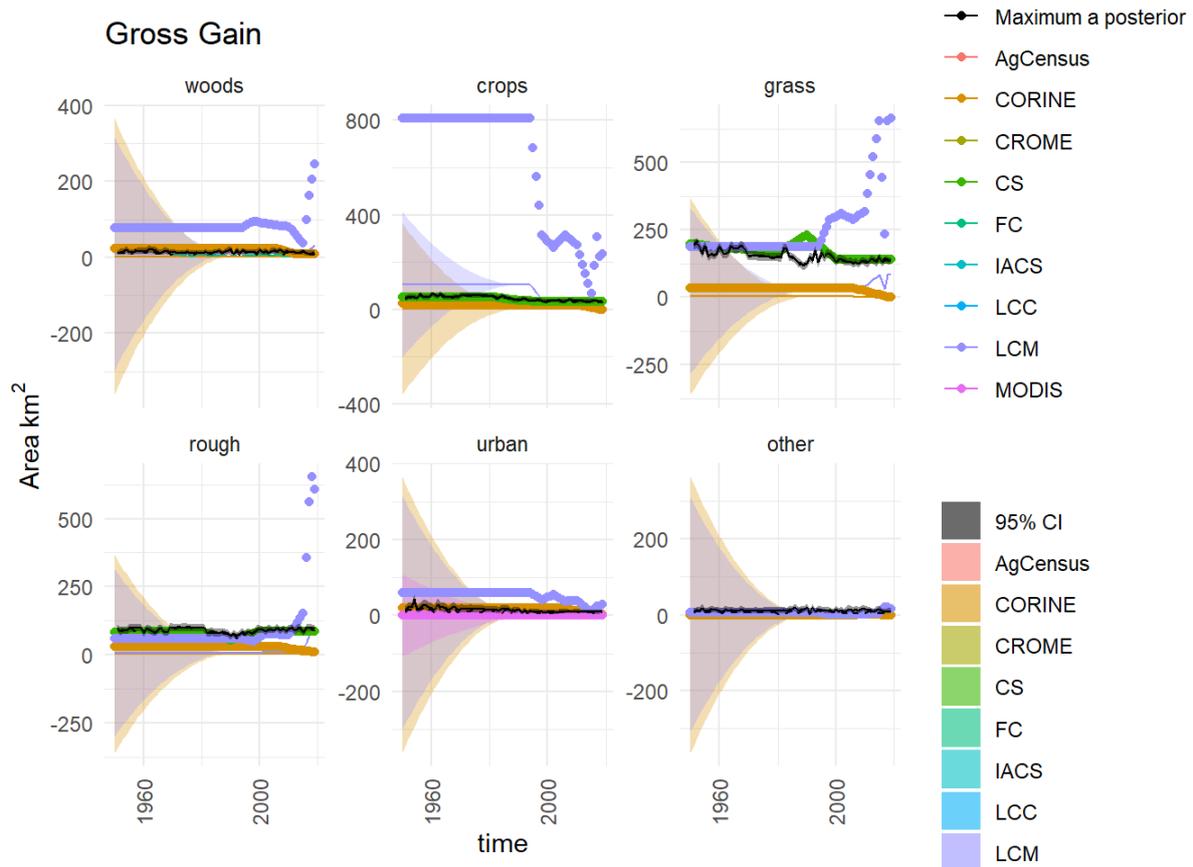


Figure 10.2: Observations and posterior distribution of the gross gain in area of each land use G from 1950 to 2020. The grey shaded band shows the 2.5 and 97.5 percentiles of the posterior distribution. The maximum *a posteriori* estimate is shown as the solid black line within this. Observations from the different data sources are shown as coloured circles. The coloured solid lines show the corrected observations after accounting for systematic uncertainties, and interpolating. The coloured bands around these lines show the random uncertainty, rescaled as $\sigma/5$ to keep with the axis scale. Because the random uncertainties and the corrections to the observations are generally very large in comparison to the actual change, scaling the axes is difficult. Note that a consistent colour scheme for the data sources is shown, but not all contribute to every figure.



Figure 10.3: Observations and posterior distribution of the gross loss of area from each land use **L** from 1950 to 2020. The grey shaded band shows the 2.5 and 97.5 percentiles of the posterior distribution. The maximum *a posteriori* estimate is shown as the solid black line within this. Observations from the different data sources are shown as coloured circles. The coloured solid lines show the corrected observations after accounting for systematic uncertainties, and interpolating. The coloured bands around these lines show the random uncertainty, rescaled as $\sigma/5$ to keep with the axis scale. Because the random uncertainties and the corrections to the observations are generally very large in comparison to the actual change, scaling the axes is difficult. Note that a consistent colour scheme for the data sources is shown, but not all contribute to every figure.

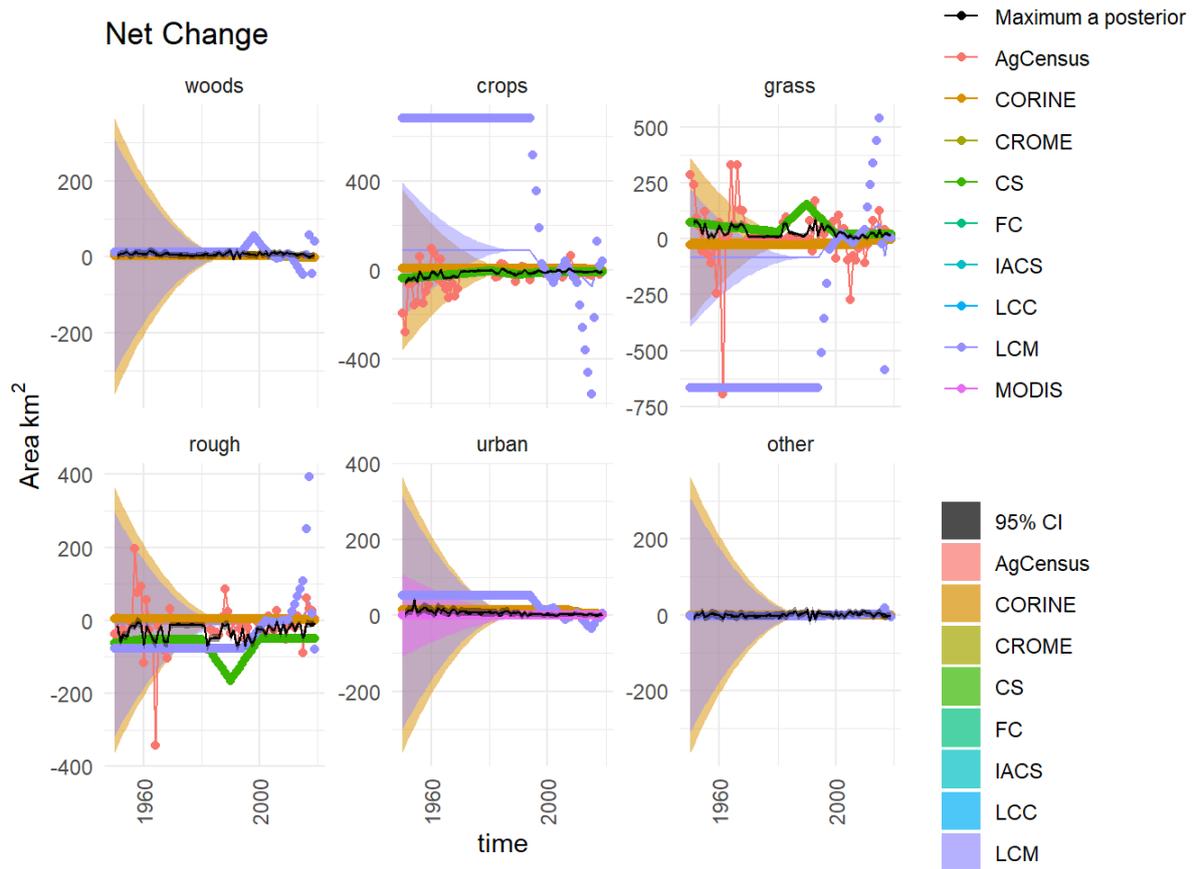


Figure 10.4: Time series of the net change in area occupied by each land use (D) from 1950 to 2020, showing the observations and posterior distribution of estimates. The grey shaded band shows the 2.5 and 97.5 percentiles of the posterior distribution. The maximum *a posteriori* estimate is shown as the solid black line within this. Observations from the different data sources are shown as coloured circles. The coloured solid lines show the corrected observations after accounting for systematic uncertainties, and interpolating. The coloured bands around these lines show the random uncertainty, rescaled as $\sigma/5$ to keep with the axis scale. Because the random uncertainties and the corrections to the observations are generally very large in comparison to the actual change, scaling the axes is difficult. Note that a consistent colour scheme for the data sources is shown, but not all contribute to every figure.

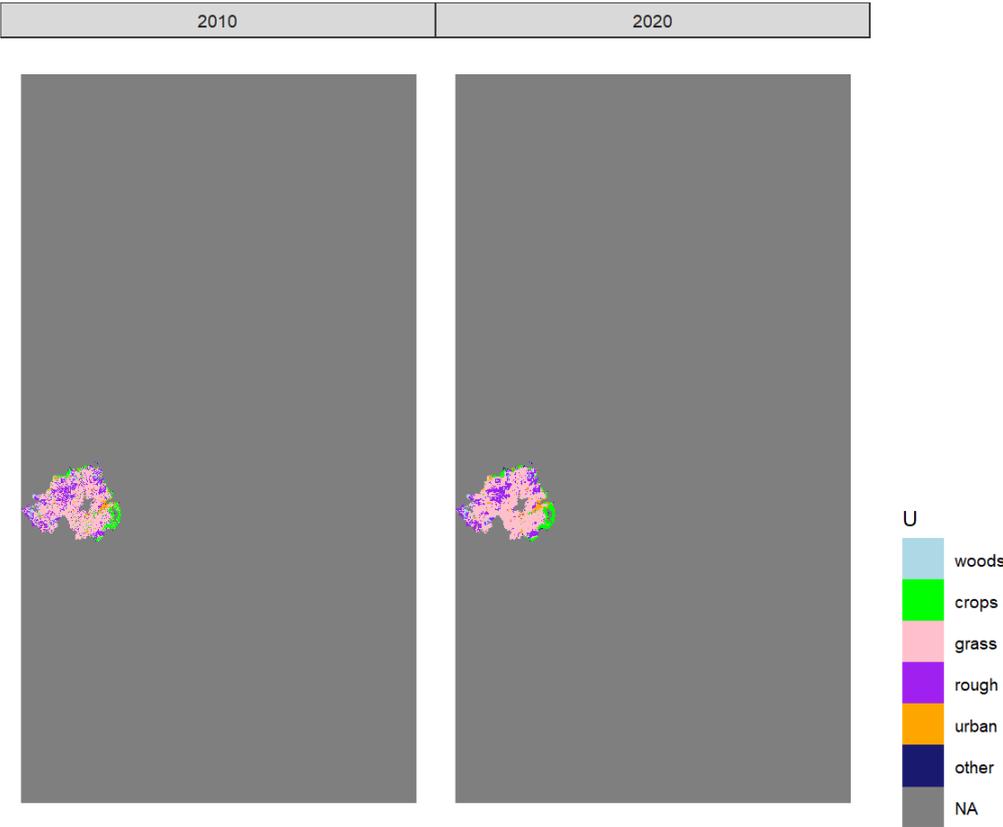


Figure 10.5: Estimated state of land-use in 2010 and 2020 in one realisation of U from the maximum *a posteriori* estimate of B .

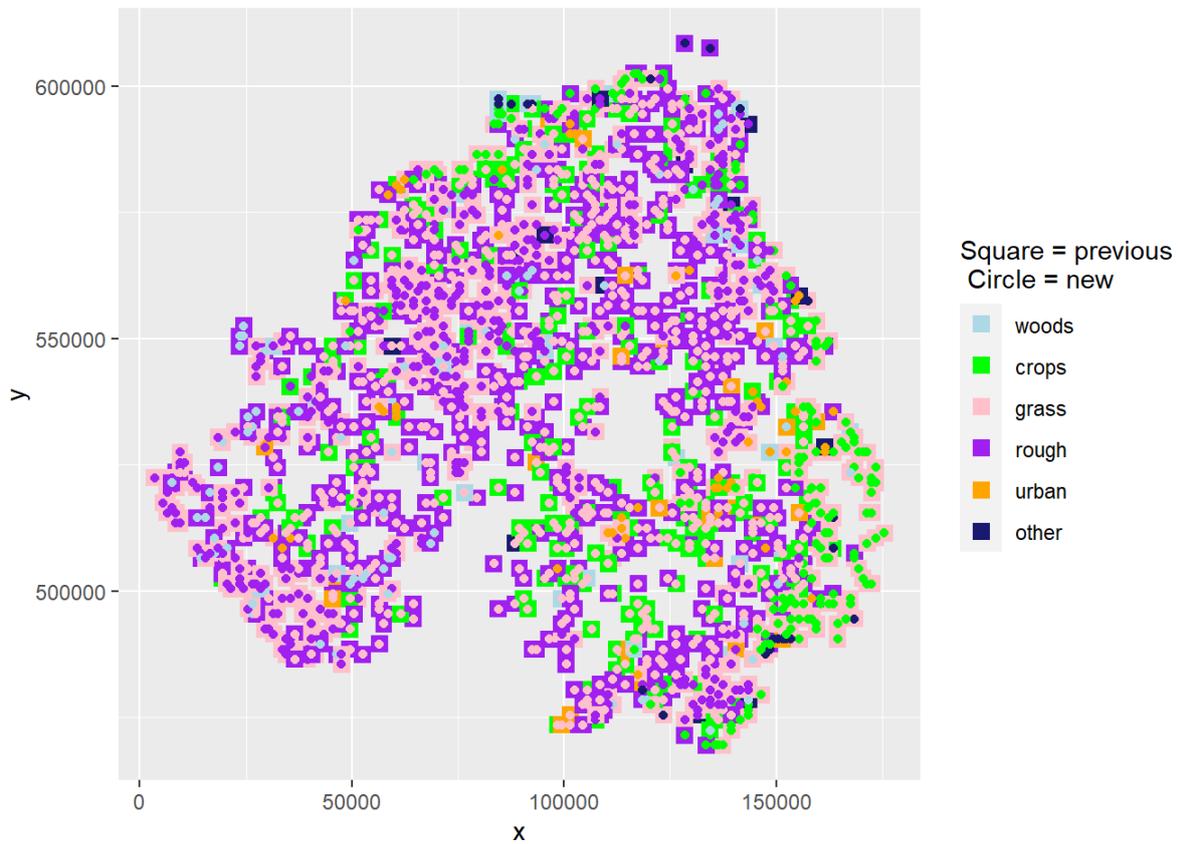


Figure 10.6: The spatial distribution of land-use change between 2010 and 2020 in one realisation of U from the maximum *a posteriori* estimate of B . At each location where land use has changed, the use in 2010 is shown as a coloured square, and the use in 2020 is shown as a coloured circle within this

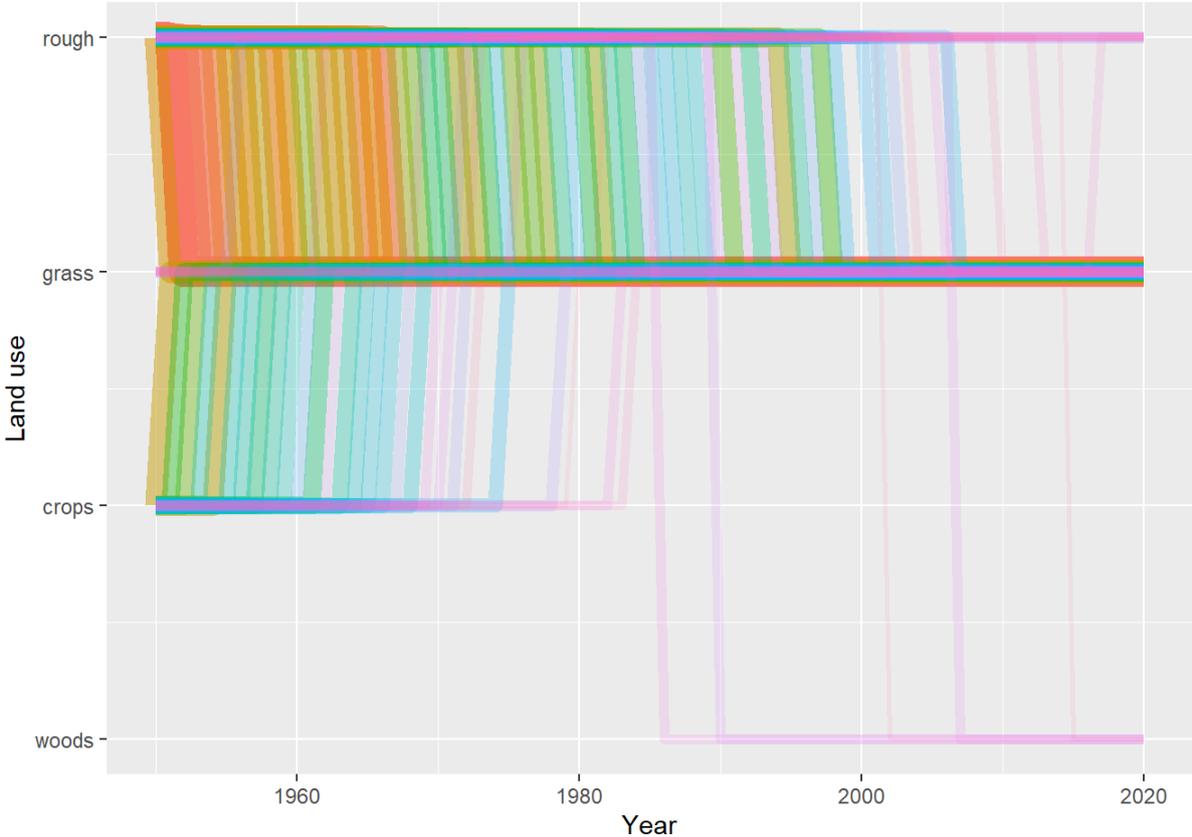


Figure 10.7: Trajectories of the 100 land-use vectors in the posterior U with the largest areas (excluding the six vectors which show no change). Each vector of land use is shown in a different colour, varied arbitrarily to differentiate different vectors. Line thickness and opacity are proportional to the total area occupied by each vector, so that the dominant vectors are the most visually obvious.

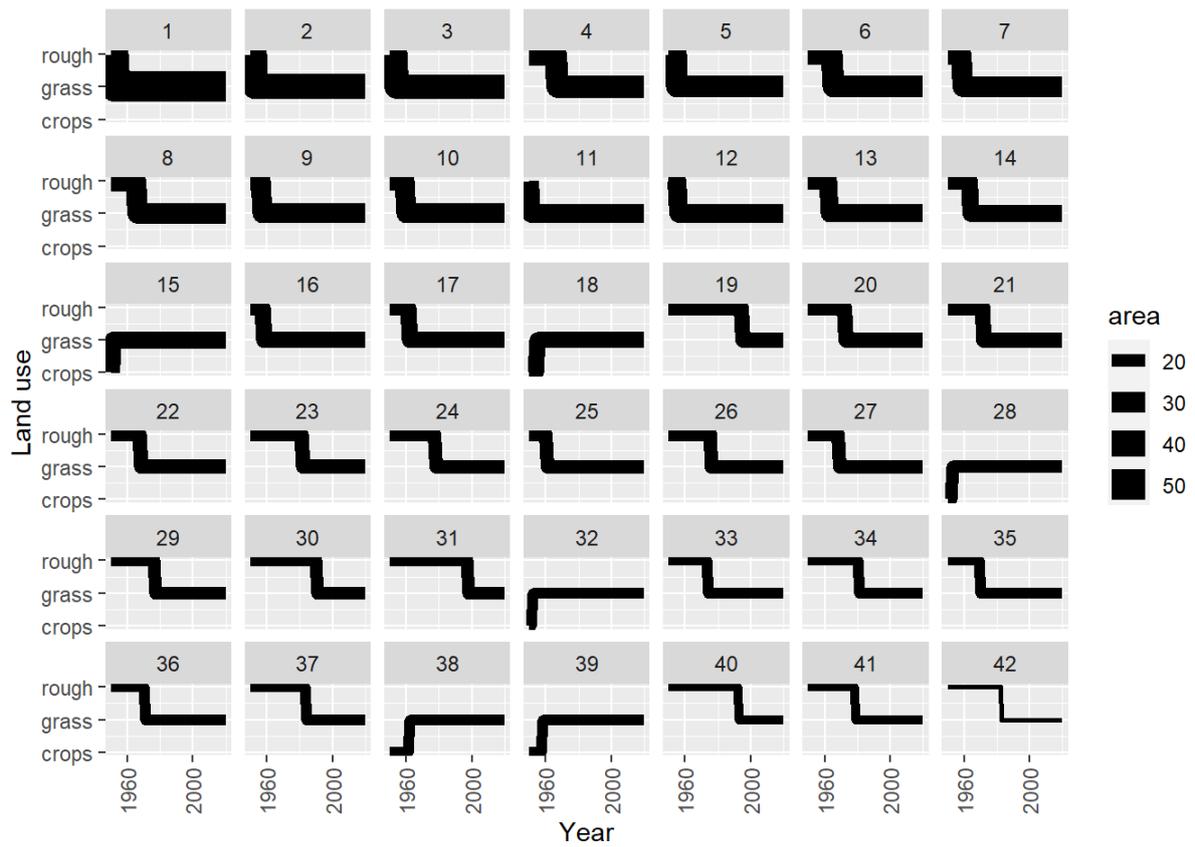


Figure 10.8: Trajectories of the 42 land-use vectors in the posterior U with the largest areas (excluding the six vectors which show no change). Line thickness is proportional to the total area occupied by each vector.

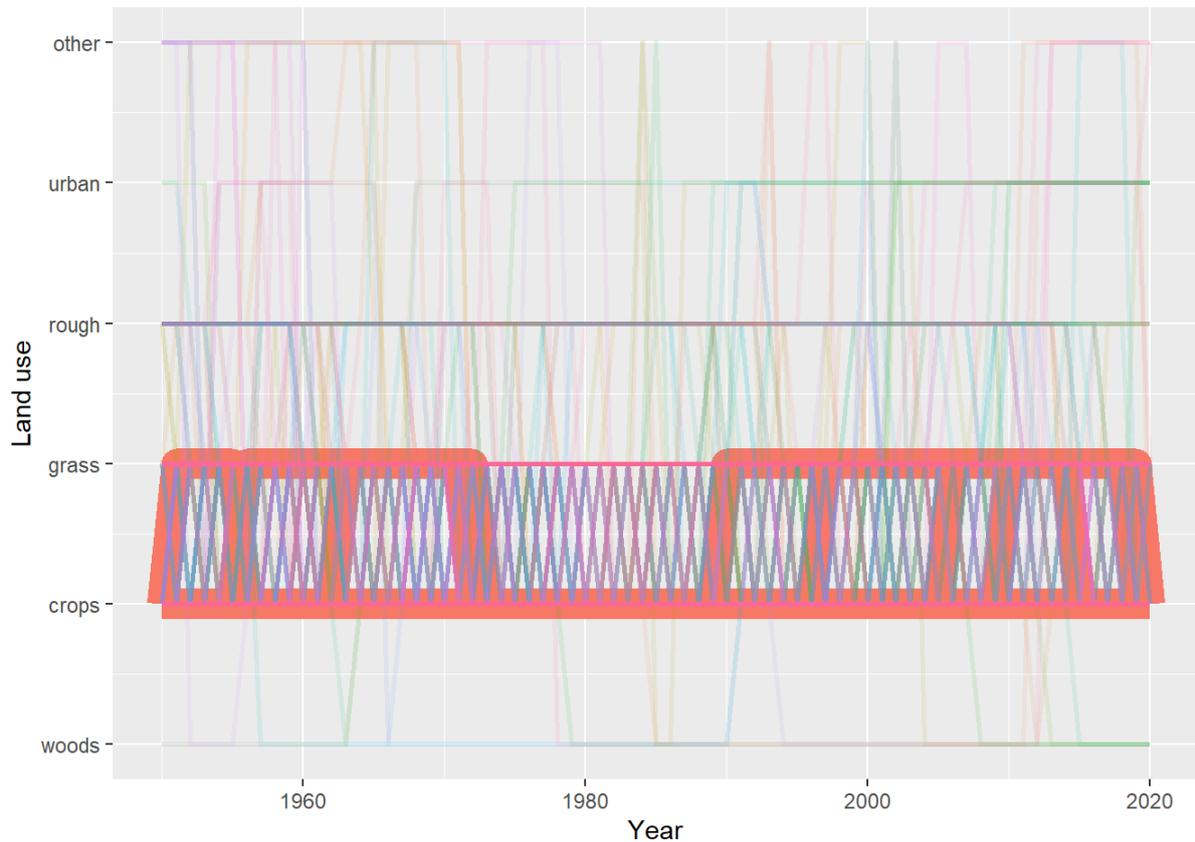


Figure 10.9: Trajectories of the land-use vectors in the posterior U which involve rotational change between crop and grassland (i.e. those which include either (i) transitions from crop to grass and then subsequently from grass to crop, or (ii) transitions from grass to crop and then subsequently from crop to grass). Each vector of land use is shown in a different colour, varied arbitrarily to differentiate different vectors. Line thickness and opacity are proportional to the total area occupied by each vector, so that the dominant vectors are the most visually obvious.

11 WP-A Discussion

In the original proposal we envisaged that WPA would be based on ground-based data that were already available, whilst WPB would provide new data based on remote sensing. We proposed to compare land-use change estimates based on these individually and combined. We anticipated WP-B would produce at least two data sets which provide reliable, high-resolution data on land-use change, with the detail to provide the Beta matrices at least every few years back to 1990. However, based on the analysis of the false positive rate in previous sections, land-use change from the remotely sensed data appears to be wrong 80-90 % of the time. Furthermore, the new woodland maps gave no information on the Beta matrices and did not provide a plausible time series for new planting/deforestation, compared with the existing FC statistics. Basing an analysis solely on the WPB data taken at face value therefore seems not very useful, and we restrict the comparison to WPA versus the combination of WPA and WPB, where the latter are corrected for estimated false positive rates using ground-based data (denoted WPAB).

In the results of [WP-A](#), the data assimilation estimates remain close to the Countryside Survey data because these are the only data that fully specify the beta matrix. The decadal-scale trends in these data remain clear in the assimilated estimates. The Agricultural Census data only specify the net change, and there is considerable year-to-year noise in the time series. The assimilation algorithm effectively smooths out a lot of this noise but follows the general trend. The Forestry Commission data for afforestation and deforestation are followed quite closely because these are specified with high precision.

Contrasting these with results from [WP-AB](#), the main effect is further smoothing of the time series. This is because it is now a weighted average of several more data sets, and these additional datasets do not show a strong coherent pattern. Particularly, these datasets do not show the sharp decadal trends seen in the CS data, so this is smoothed out. All these data sets have uncertainties considerably larger than the Agricultural Census data, so their weighting is relatively low. The overall effect of this is that the [WP-AB](#) estimates are smoother, but the effect on the general pattern and the absolute magnitude of change is relatively small.

In terms of assessing which of these combinations should be used in the inventory, the answer is not entirely clear. Using only [WPA](#) data is obviously simpler, as it requires fewer data sources. Given the analysis of the false positive rates in the remotely sensed data, it is clear these are dominated by spurious differences, and most of the apparent change is simply noise. It is therefore questionable whether there is any real value in including these. The argument for including these data is that it is plausible that they do include some true signal as well as noise – their accuracy in detecting change is low but not zero. Given that we can account for the false positive rate and down-weight those datasets with highest random uncertainty, it makes some sense to use all the information that we have available. The practical effect is to smooth out features in the [WPA](#) data which we know to be artefacts (the sharp decadal trends in CS data), so given that the analysis has been done already, there is no immediate problem with using the additional data.

One assumption that is implicit is that the biases in the [WPB](#) data are fixed, known constants, so we can reliably correct the data with the calculated false positive rates. The false positive rates are almost certainly not fixed, but quite how they vary, and whether this is systematic is hard to know. Further work would be needed to estimate this. A more elegant approach would be to include the false positive rates as parameters to be estimated in the data assimilation algorithm, at the cost of increasing the complexity and computation time.

Whilst we have improved the representation of the CS data by using linear interpolation rather than decadal step changes, this still leaves an artefact in the data (a different linear trend each decade). This could be improved with a smoothing routine (e.g. LOESS or GAM) which would give a better approximation to what we expect is the true pattern of change over time. The same point can be applied to the Agricultural Census data, where much of the year-to-year noise seems implausible and smoothing this out to some degree would seem more realistic.

The data assimilation routine still gives relatively high weight to the CS data, based on the unique way it monitors change. However, exactly how much uncertainty is introduced when extrapolating from the sample 1-km squares to national scale is

unclear, and the methodology for doing this is now rather obscure. Revisiting this, with an appropriate method for quantifying the uncertainty in national-scale estimates, would be worthwhile if these data remain central to estimates of change.

One effect of introducing the WPB datasets is that the spatial attribution of land-use change becomes a more confused picture: we have a number of data sources which give conflicting (and largely erroneous) information on where land use has changed. The effect of this on the time series of maps produced by the data assimilation algorithm is to fragment the distribution of land use types. For example, if we used only LCM as the basis of past change spatially, the algorithm would reproduce the spatial pattern contained in LCM data. If we mix several inconsistent data sets, we get a change in the spatial pattern which is largely incoherent. More sophisticated techniques for combining the spatial data are required to address this.

12 References

Levy, P., M. van Oijen, G. Buys, and S. Tomlinson. 2018. "Estimation of Gross Land-Use Change and Its Uncertainty Using a Bayesian Data Assimilation Approach." *Biogeosciences* 15 (5): 1497–1513. <https://doi.org/10.5194/bg-15-1497-2018>.

Levy, P.M., Tomlinson, S., Buys, G. and Thomson, A. (2020). Improving land use change tracking in the UK Greenhouse Gas Inventory: WP-A Approach and Data Sources Report.

Rowland, C., Buys, G., Clark, L., Correira, V., Ditchburn, B., Evangelides, C., Higgins, A., Marston, C., Morton, D., Olave, R., O'Neil, A., Tomlinson, S., Walker, A., Watterson, J., Whitton, E., Williamson, J., and Thomson, A. (2021). Improving land use change tracking in the UK Greenhouse Gas Inventory: WP B Final Report.

WORK PACKAGE B FINAL REPORT

WP-B Final Report

Authors: Clare Rowland¹, Gwen Buys¹, Liz Clark², Vera Correia², Ben Ditchburn², Christopher Evangelides³, Alex Higgins⁴, Christopher Marston¹, Daniel Morton¹, Rodrigo Olave⁴, Aneurin O’Neil¹, Sam Tomlinson¹, Anthony Walker², John Watterson³, Esther Whitton², Jenny Williamson¹ and Amanda Thomson¹

Date: 30/09/2021

¹ UK Centre for Ecology & Hydrology; ² Forest Research; ³ Ricardo; ⁴ Agri-Food & Biosciences Institute.

1 WP-B Introduction

The tracking of land use and land-use change is fundamental to producing accurate and consistent greenhouse gas inventories (GHGI) for the Land Use, Land-Use Change and Forestry (LULUCF) sector. The aim of this project is to improve and update the tracking of land-use change for the UK's GHGI in order to meet the international requirements of the Kyoto Protocol to the UN Framework Convention on Climate Change (UNFCCC) and the Paris Agreement and the national requirements of the UK's Climate Change Act and related legislation within the UK's Devolved Administrations.

Work Package B aims to improve the maps of land-use and land-use change generated by Work Package A using data from remote sensing. This will require mapping land cover changes for all six Land Use, Land Use Change and Forestry classifications between 1990-2020. This report describes the work undertaken in WPB, including the five sub-tasks and the Quality Assurance (QA) and Data Management components.

1.1 LULUCF land categories

The six LULUCF land categories are Forest Land, Cropland, Grassland, Wetlands, Settlements and Other Land. The UK has specific definitions for each category for GHGI reporting, published in the annual National Inventory Report (Brown et al. 2021). All land within the UK is assigned to one of these six land categories.

1.1.1 Forest land

The UK uses the following definition of forest for reporting under the UNFCCC: land with woody vegetation that meets the following thresholds:

- Minimum area of 0.1 hectares;
- Minimum width of 20 metres;
- Tree crown cover of at least 20 per cent, or the potential to achieve it;
- Minimum height of 2 metres, or the potential to achieve it.

This definition includes felled areas awaiting restocking and integral open spaces (open areas up to 1 hectare).

1.1.2 Cropland

Cropland is defined in accordance with the Agriculture, Forestry and Other Land Use Guidance (IPCC 2006) and includes cropped land, including systems where the vegetation structure falls below the thresholds used for the Forest Land category (e.g. orchards).

1.1.3 Grassland

Grassland that has undergone land-use change and direct management is defined in accordance with the Agriculture, Forestry and Other Land Uses guidance (IPCC 2006)⁸. There are also large areas of extensively grazed semi-natural grassland on mineral soils, which are assigned to the 4.C.1 "undisturbed grassland" sub-category and calculated as the area remaining after all other land use

⁸ "This category includes rangelands and pasture land that are not considered Cropland. It also includes systems with woody vegetation and other non-grass vegetation such as herbs and brushes that fall below the threshold values used in the Forest Land category. The category also includes all grassland from wild lands to recreational areas as well as agricultural and silvi-pastoral systems, consistent with national definitions." IPCC (2006) Vol. 4, Chapter3.

areas are subtracted from the total UK land area. This is the buffer land use category for the UK, so may contain small areas of other land uses that are not directly managed.

Grazing is the main land use on semi-natural peatland habitats that would otherwise fall within in the Wetland category, so areas of peatland habitat not used for peat extraction or that have been rewetted from forest or cropland, are also included in the Grassland category. Areas of grassland that have undergone rewetting (or other activities that restore peatland habitats and reduce emissions) remain within the Grassland category.

1.1.4 Wetlands

The IPCC(2006) definition of wetlands includes areas of peat extraction and land that is covered or saturated by water for all or part of the year (e.g. peatlands) and that does not fall into the Forest Land, Cropland, Grassland or Settlements categories. It includes both reservoirs and natural rivers and lakes. In the UK this category includes areas of both ‘near-natural’ and rewetted peatlands that have undergone restoration activities to restore normal peatland biogeochemical functioning. Areas of former peatland that have been modified for other land use are reported under the appropriate LULUCF category.

1.1.5 Settlements

Settlement is defined in accordance with the Agriculture, Forestry and Other Land Use Guidance (IPCC 2006) and includes all developed land, including transportation infrastructure and human settlements of any size, unless they are already included under other categories. In the UK this category includes domestic gardens and allotments, waste and derelict ground and urban parkland.

1.1.6 Other Land

The IPCC (2006) definition of Other Land are areas that do not fall into the other land use categories. In the UK this category comprises inland and coastal bare rock, and areas of shingle/gravel/mud/sand in the inter-tidal zone.

1.2 Summary of Work package A

The overall aim of the project is to make improved estimates of land-use change in the UK, using multiple sources of data. In WP-A, we applied a method for estimating land-use change using a Bayesian data assimilation approach. This allows us to constrain estimates of gross land-use change with national-scale census data, whilst retaining the detailed information available from several other sources. We produced a time series of maps describing our best estimate of land-use change given the available data, as well as the full posterior distribution of this space-time data cube. This quantifies the joint probability distribution of the parameters, and properly propagates the uncertainty from input data to final output. The output data has been summarised in the form of land-use vectors. The results show that we can provide improved estimates of past land-use change using this method. The main advantage of the approach is that it provides a coherent, generalised framework for combining multiple disparate sources of data, and adding further sources of data in future would be straightforward.

2 WP-B1: Produce Land Cover Data/Land-Use Data

Task B.1 is split into three components, led by UKCEH, FR and AFBI respectively, with each component tasked with producing new land cover or land-use data for the period 1990-2020.

2.1 Production of new Land Cover Map data (UKCEH)

The purpose of the UKCEH component of Task B.1 was to produce new Land Cover Maps (LCM) between 1994 and 2010, to complement the existing LCM data for 1990, 2015, 2017, 2018 and 2019; LCM2020 production ran in parallel with this project. Producing new LCM data sets required the selection, and processing, of satellite data to produce maps of land-use change in accordance with the six LULUCF categories and in line with the IPCC guidelines.

This was achieved by adapting the method developed in Carrasco et al. (2019), which has since been applied to produce UK-wide LCMs for 2017, 2018 and 2019, and developing new post-classification filtering methods (Figure 1).

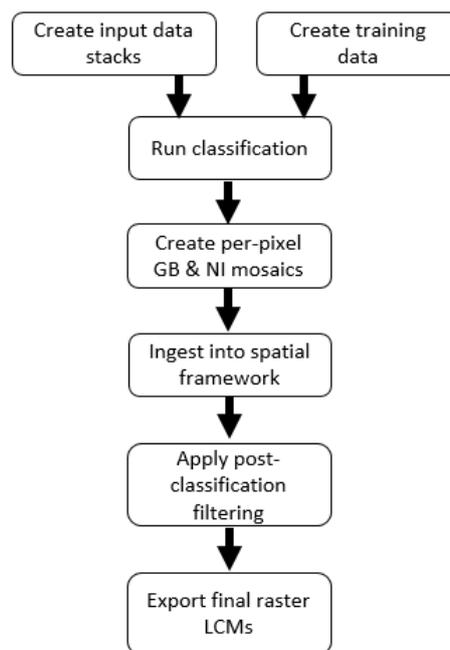


Figure 1: Overview of the LCM production chain.

2.1.1 Satellite data

Google Earth Engine was used to create seasonal composite mosaics for 1990, 1994, 1998, 2002 and 2010, with a pixel size of 25 m. Google Earth Engine is a cloud-processing environment, linked to cloud-storage, which is designed to enable large-scale geospatial analysis, particularly of satellite data (Gorelick *et al.*, 2017). Atmospherically corrected Landsat collection 1 tier 1 data was used (see USGS, 2018 for full data specification; see GEE, 2021 for details of data once ingested into GEE; Wulder *et al.*, 2019), cloud-masking was then applied and the data were then aggregated using the median seasonal

value for spring, summer, autumn and winter (Figure 2). The compositing methods try to identify the best available pixel in a particular time period. The aim is to produce a complete coverage of imagery, free of clouds, cloud-shadow and snow. This significantly reduces the time required to produce Land Cover Maps, because the methods can be readily automated and the use of cloud-free composites for large square areas (circa 100km x 100km) removes the need for time-consuming hole-filing exercises that were crucial parts of ensuring complete UK coverage for previous LCM's (Carrasco *et al.*, 2019).

To aid quality assurance the number of images used for different areas of the UK was also mapped (Figure 3). The seasonal mosaics were created for six Landsat bands: blue, green, red, near-infrared (NIR), short-wave infra-red (SWIR) 1 and SWIR2 and for the years required by the project (Table 1).

Table 1: Summary of Landsat data.

| Year | Data source | Based on images covering |
|------|---------------------------------------|-------------------------------|
| 1990 | Landsat-5 | October 1989 – September 1991 |
| 1994 | Landsat-5 | October 1993 – September 1995 |
| 1998 | Landsat-5 | October 1997 – September 1999 |
| 2002 | Landsat-5 & 7 | October 2001 – September 2003 |
| 2006 | Landsat-5, with gap-fill by Landsat-7 | October 2005 – September 2007 |
| 2010 | Landsat-5, with gap-fill by Landsat-7 | October 2009 – September 2011 |

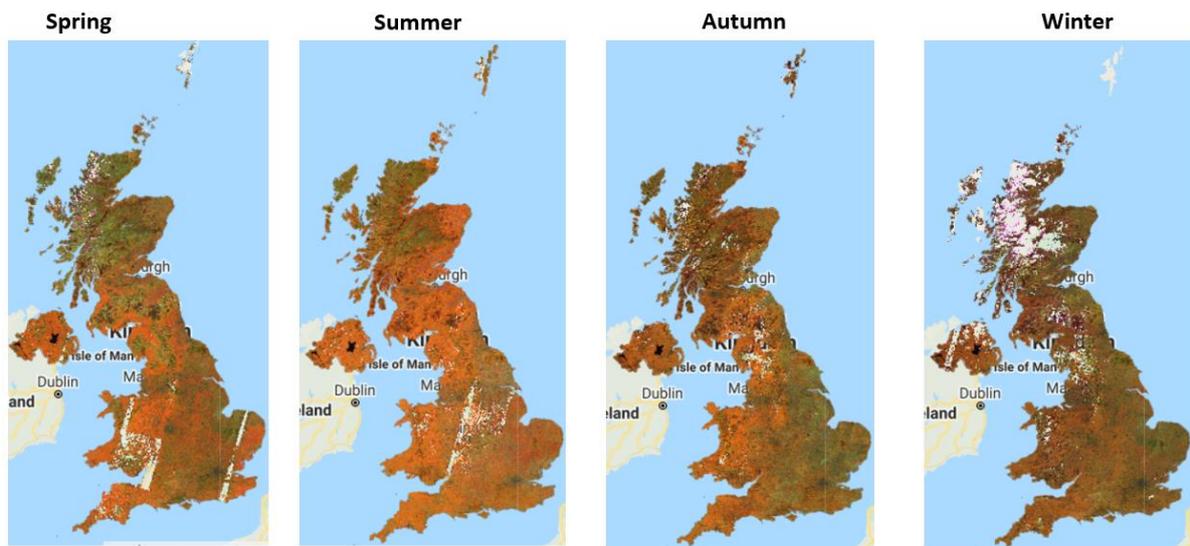


Figure 2: Landsat seasonal mosaics for 1994, shown as false colour composites (NIR, SWIR and red bands displayed as R,G,B respectively).

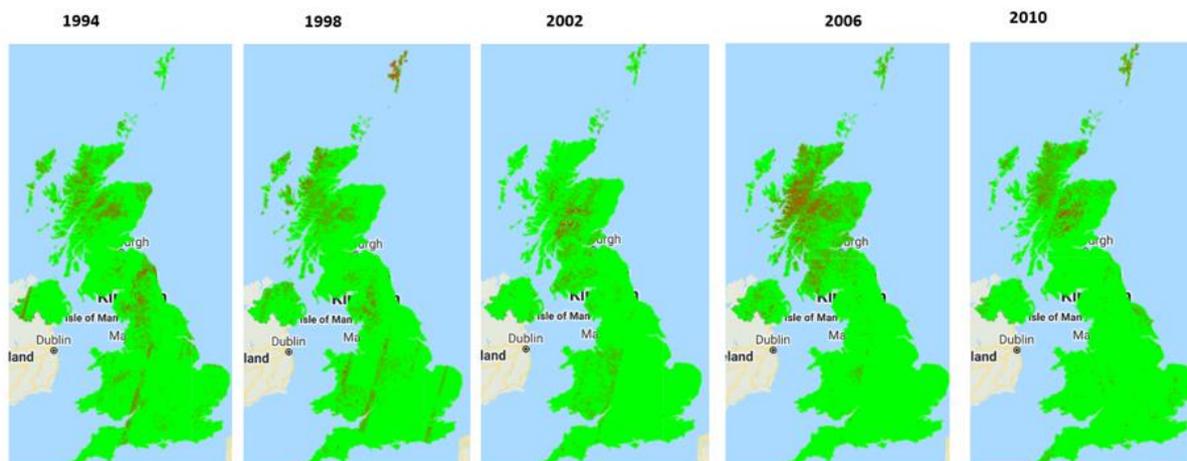


Figure 3: Summary of the number of seasons with satellite data for the new Landsat-based LCM data sets. Red represents 1 season, Bright green represents 4 seasons.

2.1.2 Ancillary data

In addition to the spectral data, the ancillary data sets listed in Table 2 were also used for the image classification.

Table 2: Summary of ancillary data sets.

| Data type | Rationale | Data set name | Coverage | Produced by: |
|---------------------|---|-------------------------------|----------|--|
| Altitude | Constrain land cover classes to appropriate slopes and altitudes | Digital elevation data | GB | Nextmap |
| | | | NI | Land & Property Services |
| Urban extent | Distance from urban, used to limit spectral confusion, especially between arable and urban | Buildings layer | GB | Ordnance Survey (OS) |
| | | Road network | | |
| | | Settlement development limits | NI | NI Statistics and Research Agency |
| Coast line | Constrain coastal classes so they do not occur inland | Mean high/low water mark | GB | OS |
| | | Coastal water | NI | Department of Agriculture, Environment and Rural Affairs (DAERA) |
| Water | Distance from water used to improve classification of habitats often associated with proximity to rivers (e.g. Fen, Marsh and Swamp, and Neutral Grassland) | River network | GB | OS |
| | | Surface water | | |
| | | Freshwater | NI | DAERA |
| Forest | Improve extent of forest, especially for recently | Forest layer | GB | Ordnance Survey |

| | | | | |
|--|---|--|--|--|
| | harvested forest and newly planted forest | | | |
|--|---|--|--|--|

2.1.3 Training data

Training data have been developed from the vector versions of LCM1990 (Rowland *et al.*, 2020a, b) and LCM2015 (Rowland *et al.*, 2017a, b) by using polygons that remained stable (i.e. the same class) between LCM1990 and LCM2015. The training data used the 21 LCM class specification (Table 3), with the classifications re-categorised to the LULUCF classes post-classification. The training data were derived from areas that stayed classified as the same class between LCM1990 (Rowland *et al.*, 2020a, b) and LCM2015 (Rowland *et al.*, 2017a, b), so if a field was classified as crop in LCM1990 and LCM2015 then it could become a crop training area. The LCM vector data set records a number of attributes including the modal coverage, which is the proportion of the polygon covered by the dominant class, as well as an attribute that captures the level of uncertainty from the Random Forest classification. These attributes can be used to refine the selection of training areas and to exclude some of the training polygons that might be less appropriate. Ideally a training polygon should be covered by a single land cover class and should be characteristic of that land cover class, so the modal_proportion attribute will be high, as will the classification certainty value.

Table 3: Mapping from LCM classes to the LULUCF classes. Note LULUCF class 7 was added for water to ensure complete coverage of the UK. Note classes highlighted in red show deviations from the mapping used in WP-A. The deviations were discussed within the team and are based on the carbon content (or lack of) of the classes. Supra-littoral sediment includes vegetated sand-dunes.

| LCM_ID | LCM_name | LULUCF_ID | LULUCF_name |
|--------|-------------------------|-----------|-------------|
| 1 | Broadleaved Woodland | 1 | forest |
| 2 | Coniferous Woodland | 1 | forest |
| 3 | Arable and Horticulture | 2 | crop |
| 4 | Improved Grassland | 3 | grass |
| 5 | Neutral Grassland | 4 | rough |
| 6 | Calcareous Grassland | 4 | rough |
| 7 | Acid grassland | 4 | rough |
| 8 | Fen, Marsh and Swamp | 4 | rough |
| 9 | Heather | 4 | rough |
| 10 | Heather grassland | 4 | rough |
| 11 | Bog | 4 | rough |
| 12 | Inland Rock | 6 | other |
| 13 | Saltwater | 7 | water |
| 14 | Freshwater | 7 | water |
| 15 | Supra-littoral Rock | 6 | other |
| 16 | Supra-littoral Sediment | 4 | rough |
| 17 | Littoral Rock | 6 | other |
| 18 | Littoral sediment | 6 | other |
| 19 | Saltmarsh | 4 | rough |
| 20 | Urban | 5 | urban |
| 21 | Suburban | 5 | urban |

A range of training areas were tested on 100km x 100km tiles for a range of classification years, to inform development of the training data set. The classified results were manually reviewed to identify issues and to identify which performed best. The aim was to identify a set of training areas that could be used for each of the classifications from 1994-2010, and resulted in a final set of training areas for GB defined as:

1. **Stable areas** - Polygons that were the same class in LCM1990 and LCM2015
2. **Modal coverage** - Coverage by the modal class was greater than 80% in LCM1990.
3. **Pixel area** - Polygons containing more than 4 pixels (the purpose of this step is to remove small polygons, although most of these polygons are likely to be removed in the classification phase anyway, as part of that process removes pixels at the edge of polygons).
4. **Additional sea polygons** added - the spatial framework is based on Ordnance Survey data, which covers land, so there are few sea polygons in the LCM data. Consequently sea is classified poorly unless additional polygons are added. In some cases this may primarily be an aesthetic issue for the intermediate classifications, as few sea polygons will remain in the final product, but it is highly visible and is important for the inter-tidal areas.
5. **Additional grassland polygons** added – a post LCM2015 review identified that some grassland types were poorly represented in LCM2015. To improve this situation, additional grassland training areas were produced based on areas identified as SSSI's or phase 1 habitat areas from data published by NRW, NE and SNH. These areas had previously been used in the production of LCM1990.

Coastal classes occurring inland are a long-standing issue with the Land Cover Map and were partly resolved at this stage by targeted edits to the training data, specifically:

6. **Saltmarsh correction** – excess saltmarsh inland was addressed by deleting saltmarsh polygons with a certainty value of < 80%.
7. **Littoral sediment correction** – littoral sediment, primarily the vegetated sand dune polygons that fall in this class, tend to cause false detection of littoral sediment near low-lying rivers in coastal areas. Filtering based on certainty value (as applied for saltmarsh) was assessed but found to be ineffective due to some persistent mis-classification in both LCM1990 and LCM2015 that had resulted in inland areas classified as littoral sediment inland. Because none of the polygon attributes were distinctive enough to enable these training areas to be filtered out a spatial filter was applied. The spatial filter excluded littoral sediment training polygons that fell more than 100m from the mean high water mark, as defined by the OS boundary line data set.

The corrections in steps 7 and 8 are important for classes that are aggregated into the 'Rough grassland' class.

8. **Inland rock** – inland rock was represented by training areas based on LCM1990 only, where the polygon area was greater than 4 pixels, the modal coverage was greater than 85% and the certainty value was higher than 80%.

For GB this resulted in over 3 million polygons (see table 4 for full breakdown).

Table 4: Summary of the number of polygons for each LCM-class in Great Britain.

| LCM-class | Number of polygons |
|-------------------------|--------------------|
| Broadleaved woodland | 281223 |
| Coniferous Woodland | 109929 |
| Arable and Horticulture | 453037 |
| Improved Grassland | 1444715 |
| Neutral Grassland | 5347 |
| Calcareous Grassland | 11070 |
| Acid grassland | 126877 |
| Fen, Marsh and Swamp | 3280 |
| Heather | 27457 |
| Heather grassland | 50127 |
| Bog | 13633 |
| Inland Rock | 2059 |
| Saltwater | 2802 |
| Freshwater | 8460 |
| Supra-littoral Rock | 2938 |
| Supra-littoral Sediment | 3770 |
| Littoral Rock | 1486 |
| Littoral sediment | 4402 |
| Saltmarsh | 5673 |
| Urban | 100127 |
| Suburban | 446943 |
| Total | 3105355 |

The NI data was created separately (Table 5), and followed steps 1-3 of the GB methodology, but then diverged to resolve NI-specific issues, especially around the quality of some of the grassland and deciduous woodland training areas:

1. **Certainty threshold** –all training polygons had to have a certainty > 80% in LCM1990.
2. **Additional grassland filtering** – to remove some poor grassland training polygons, the certainty for grassland training polygons in LMC1990 was increased to over 90%, with polygons less than 90% deleted from the training data set.
3. **Additional woodland filtering** – to remove some poor deciduous woodland training polygons, the woodland class required a certainty of over 80% in LCM2015 (as well as the certainty threshold in step 1).
4. **Manual edits to woodland data** – after the edits above, a persistent issue remained, that was most apparent in Fermanagh, and was caused by long-standing issues with the accurate classification of small fields surrounded by mature trees. The image classification, across multiple LCM's, has struggled to classify these fields as grassland, rather than forest. Consequently, some of these fields were being wrongly selected as woodland training areas. All the LCM vector attributes were reviewed to find a systematic solution to filter out these areas. However, some persisted, so a targeted manual review of the training areas was undertaken. The fields that were manually identified as being unsuitable deciduous woodland training polygons were flagged, deleted from the training data and saved as a separate data set, so they can be re-used in the future if required.

5. **Additional grassland polygons** – as with GB additional grassland areas were added (see GB text for further details).

Table 5: Summary of the number of polygons for each LCM-class in Northern Ireland.

| LCM-class | Number of polygons |
|-------------------------|--------------------|
| Broadleaved woodland | 1377 |
| Coniferous Woodland | 9945 |
| Arable and Horticulture | 2863 |
| Improved Grassland | 42821 |
| Neutral Grassland | 95 |
| Calcareous Grassland | 22 |
| Acid grassland | 1807 |
| Fen, Marsh and Swamp | 120 |
| Heather | 589 |
| Heather grassland | 624 |
| Bog | 877 |
| Inland Rock | 66 |
| Saltwater | 233 |
| Freshwater | 1042 |
| Supra-littoral Rock | 24 |
| Supra-littoral Sediment | 191 |
| Littoral Rock | 51 |
| Littoral sediment | 452 |
| Saltmarsh | 6 |
| Urban | 2089 |
| Suburban | 15443 |
| Total | 80737 |

2.1.4 Classification method

The image data stacks were classified using the Random Forest algorithm (Breiman, 2001). The classifications used up to 10,000 training pixels selected from the specified training areas for each class. For each pixel this produced a land cover class, based on the majority vote of the 200 trees, as well as the probability of the majority class (based on the number of votes for the majority class). For these historic classifications a two-stage classification process was employed. Firstly, the classification was run with an input data stack comprised of satellite imagery and ancillary data layers, then the classification was re-run, using the satellite data, without the ancillary data. The results of the two-classifications are then merged. This two-stage method was developed in Rowland et al., (2017c) to minimise issues caused by the ancillary data being current ~2015. For example, a distance to buildings layer is used to prevent bare arable fields occasionally being classified as urban areas, however, the buildings data is circa 2015, so includes housing estates that did not exist in 1990. The ancillary data has a strong influence on Random Forest, so Random Forest will tend to classify urban land cover where the ancillary layer suggests it, even if the satellite data does not show any signs of it. Hence the need for the two-stage process.

The final part of the classification process included three corrections:

- **Arable in urban areas** – arable land cover is sometimes detected in urban areas, primarily because of the spectral heterogeneity of urban parks. To resolve this the urban greenspace data sets

were used to identify arable in urban areas and convert it to grassland. For GB, this correction used the Ordnance Survey Greenspace layer (OS ref), whilst for NI it used an equivalent layer developed specifically for this project from Open Street Map data.

- **Airport runways** – airport runways are not in the urban ancillary layer, so were being detected by the classification without ancillary data, but not by the classification with ancillary data. This led to airport runways being excluded when the two classifications were merged. To avoid this a correction was implemented that used urban pixels from the classification without ancillary data, within the boundaries of airports. To implement this airport boundaries were taken from the CORINE land cover data set.
- **Coastal correction** – coastal classes, especially saltmarsh, have a tendency to ‘creep’ inland, creating false positives in low-lying areas, next to rivers. To minimise this a coastal correction was developed, based on the coastal extent from LCM1990 and LCM2015. It was based on a coastal buffer zone, with the aim of restricting coastal classes to within this extent. The coastal buffer was created by merging the coastal extent of LCM1990 and LCM2015 and then some manual edits were applied to delete some false positives, after which a 100m buffer was applied. The coastal zone was then used to minimise the encroachment inland of coastal classes, by using non-coastal classes (where mapped by either classification) beyond the coastal zone.

2.2 Generation of new Forestry data for Great Britain (Forest Research)

National forest inventories are carried out by the Forestry Commission to provide accurate, up-to-date information about the size, distribution, composition and condition of forests and woodlands in Great Britain. This information is essential for developing and monitoring policies and guidance which support sustainable forest management.

2.2.1 Production of the 1990 woodland map

Production of the interim 1990 woodland map

The Interim 1990 woodland map (Figure 4) is a combination and manipulation of four data elements that were pre-prepared by Forest Research (FR).

To achieve this NFI have conducted four main areas of analytical work:

1. Production of 1995 woodland map;
2. Elements of the 2006 NFI GIS woodland map;
3. Elements of the NFI map and NIWT map comparison work;
4. Manipulation and reformulation of elements of Forestry Commission new planting scheme GIS data 1990 to 2010.

To produce an interim 1990 map, a logical ruleset that determined what existing GIS map data representing woodlands would be pertinent to the production of this map.

This related to selecting map data that had evidence within it verifying the existence of woodland at 1990, such as tree age at 1998 and 2006 (step 1 and 2). In addition, age of new planting grants was used. Through taking such existing GIS map data and analysing it for its actual and implied age at 1990, woodland polygons could be selected or rejected for existence at 1990, then if appropriate added to the new 1990 map.

The internal logic for probability of existence at 1990 was primarily built from a knowledge of when the data sources were created, the age of woodland when observed and the age of any imagery used when observed. For new planting administrative data on when woodland officers confirmed a successful planting was additionally used.

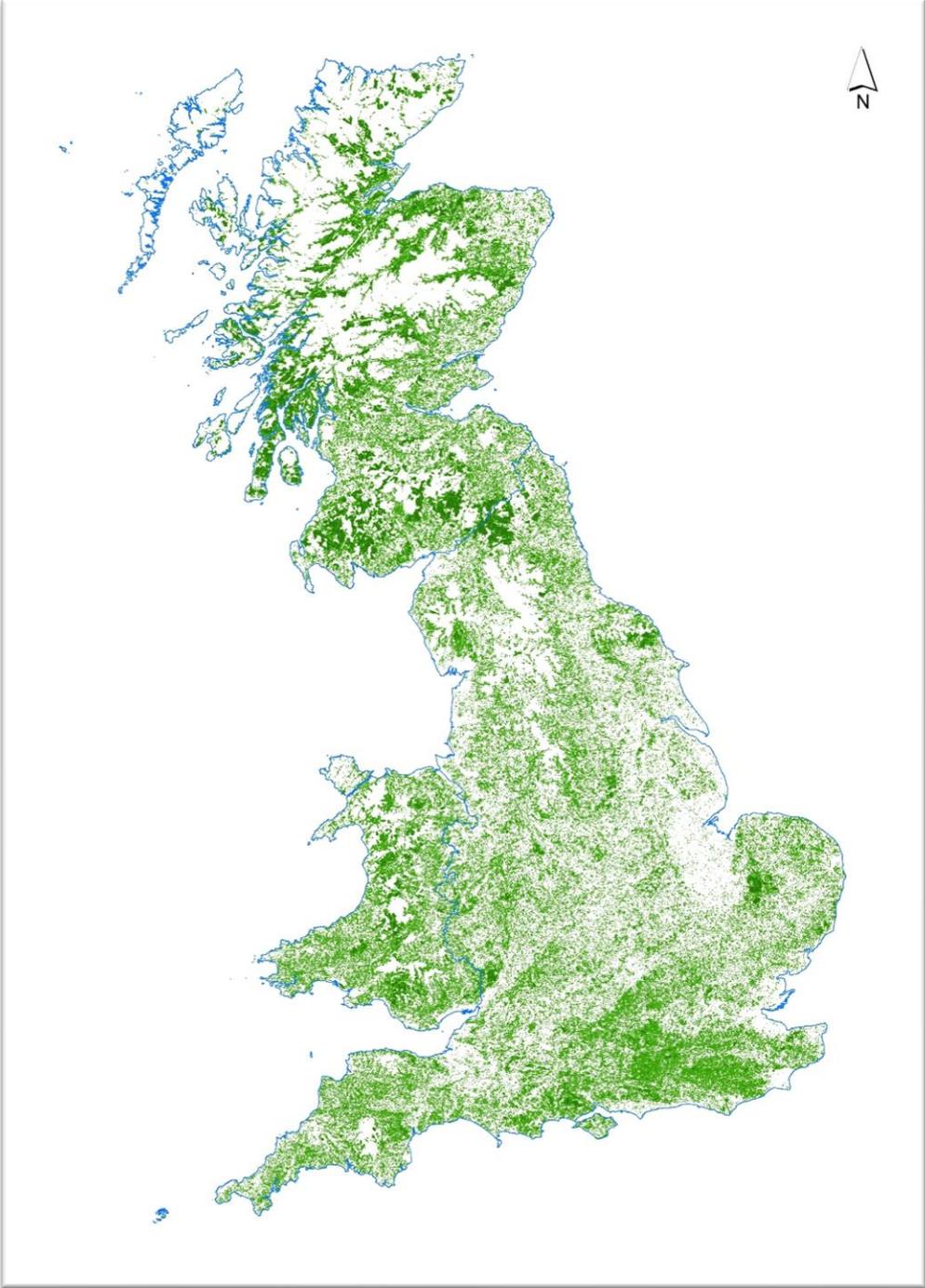


Figure 4: 1990 Interim Woodland map

Production of the final 1990 woodland map

In order to identify 'large' errors of omission within the interim 1990 woodland map, an automated classification of a composite 1990 Landsat image mosaic and subsequent manual verification was carried out. For further detail of the model used for this work, please see Appendix 1: *Development of a 1990 Woodland Map for the BEIS Land Use project* by Antony Walker.

The model found between 4700 and 6351 hectares of woodland not previously included on the woodland map, although this only represents approximately 2% of the woodland area (i.e. the interim woodland map had already covered approximately 98% of woodland). The model performed well over lowland areas, but less well over vegetated coastland and upland, particularly across Scotland and highland areas of England and Wales.

Quality assurance of the final 1990 woodland map

To ensure that all NFI woodland maps keep in line with the same accuracy levels, the areas generated by the model were then manually verified by an expert interpreter.

During this process, areas greater than 1.5 hectares, inclusive (15% of the total outputs produced by the model), were assessed individually against other sources, e.g. 1990 Landsat imagery, older aerial photographs and grant scheme data. Where there was enough evidence of woodland presence, the original output was modified to reflect the relevant woodland area. This work involved the expansion of the original area to include adjacent woodland area or its reduction to exclude open area.

Due to limited support data availability during this stage of the process, as part of the quality assurance process, 100% of the areas that have been identified as woodland were then reviewed by a senior interpreter.

Where discrepancies were found, further investigation took place prior to the inclusion or exclusion of these areas in the final product.

For areas that didn't show enough evidence of woodland, 50% of these were randomly selected and the same quality assurance process was applied.

From this work a total area of 7000 ha were added to the interim 1990 woodland map.

Due to time constraints, areas less than 1.5 hectares weren't manually verified by an interpreter. As the inclusion of unverified data would bring along errors to the total areas, such as the incorporation of open space and other features in the final product, these areas were excluded from the final product.

2.2.2 Production of 1995, 2000 and 2005 woodland maps

The 1995, 2000 and 2005 woodland map are a combination and manipulation of different data elements that were pre-prepared by FR.

To achieve the 1995 woodland map analytical work in four main areas was conducted.

1. 1995-1998 NIWT GIS woodland map, for further information regarding this work, please see section *National Inventory of Woodlands and trees 1995 – 1998 (NIWT)*
2. Elements of the 2006 NFI GIS woodland map;
3. Elements of the NFI map and NIWT map comparison work, for further information regarding this work, please see section *Comparison between NIWT and NFI estimates of woodland area work*;
4. Manipulation and reformulation of elements of Forestry Commission new planting scheme GIS data 1990 to 2010.

To achieve the 2000 woodland map analytical work in four main areas was conducted.

1. Production of the 1995 woodland map;
2. Elements of the 2006 NFI GIS woodland map;
3. Elements of the NFI map and NIWT map comparison work, for further information regarding this work, please see section *Comparison between NIWT and NFI estimates of woodland area work*;
4. Manipulation and reformulation of elements of Forestry Commission new planting scheme GIS data 1990 to 2010.

To achieve the 2005 woodland map analytical work in four main areas was conducted.

1. Elements of the 2006 NFI GIS woodland map;
2. Elements of the NFI map and NIWT map comparison work, for further information regarding this work, please see section *Comparison between NIWT and NFI estimates of woodland area work*;
3. Manipulation and reformulation of elements of Forestry Commission new planting scheme GIS data 1990 to 2010.

National Inventory of Woodlands and trees 1995 – 1998 (NIWT)

The National Inventory of Woodland and Trees 1995–1998 (NIWT) was presented as a series of inventory reports at a national and regional level.

In England and Wales a digital map of all woodland showing Interpreted Forest Types was derived from 1:25000-scale stereo colour aerial photography (1997 for Wales and 1998 for England). In Scotland, the main survey was based on the Land Cover of Scotland (LCS) 1988 project, which used 1:24000-scale aerial photography to create a land cover map. For further information regarding this work, please see LCS88Final report available on https://www.hutton.ac.uk/sites/default/files/soils/lcs88_full_report.pdf. The woodland components of this dataset were extracted to provide the basis for a digital woodland map showing Interpreted Forest Types. The map was then updated to 1995 for new planting within Woodland Grant Schemes and the Forestry Commission woodlands.

The digital map gave the extent of all woodland of 2 hectares or more, and this was progressively updated in preparation for survey work.

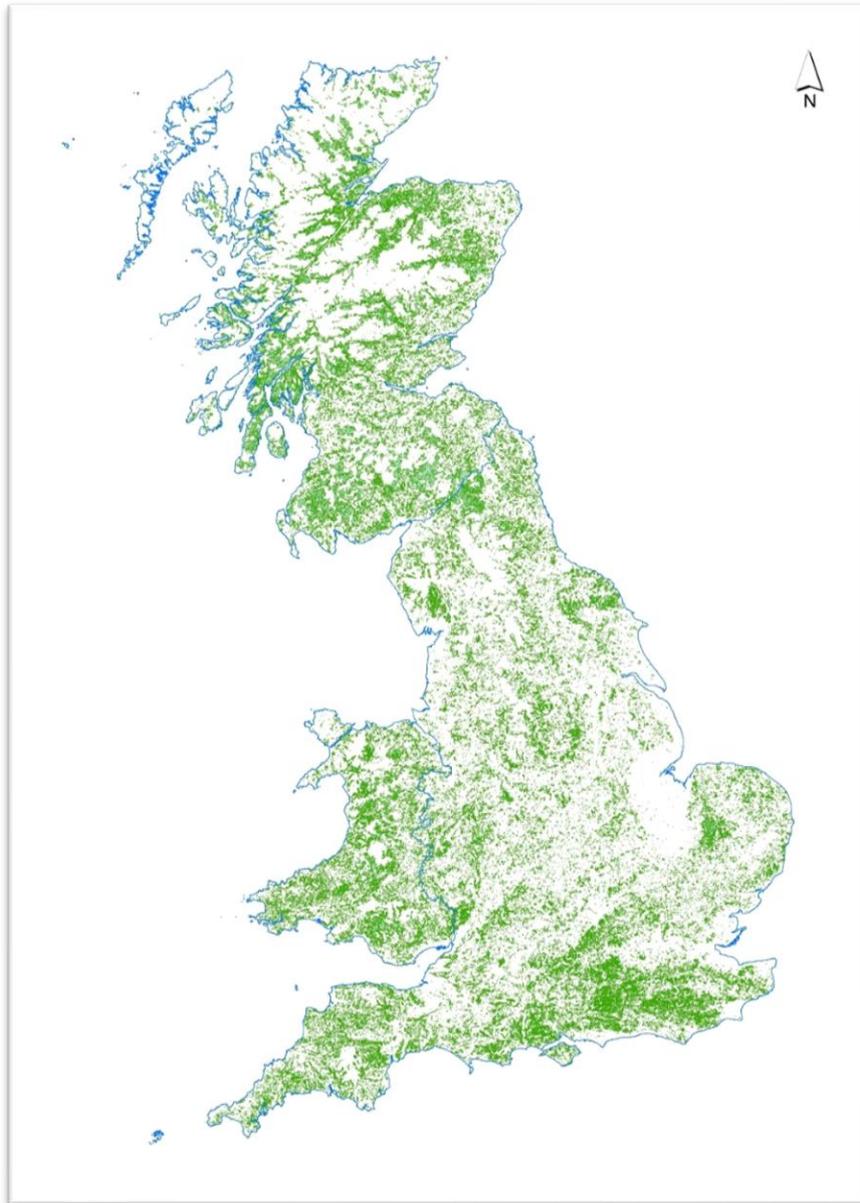


Figure 5 National Inventory of Woodland and Trees 1995 - 1998 woodland map

For further information the NIWT can be found at <https://www.forestresearch.gov.uk/tools-and-resources/national-forest-inventory/national-inventory-of-woodland-and-trees/>.

Comparison between NIWT and NFI estimates of woodland area work

The NIWT used different base years for the maps and initial woodland area estimates for each of the individual countries (1998 for England, 1995 for Scotland and 1997 for Wales). The analysis compared the NFI-based estimated woodland area for Great Britain at March 2010 with the sum of the estimated woodland areas reported by NIWT for each of the countries at each of their respective base years, taking into account the differing periods involved for each of the countries.

NIWT estimated a total woodland area of 2,665 thousand hectares in Great Britain compared to the current NFI-based estimate of total woodland area in 2010 of 2,982 thousand hectares. This represents a net difference between the two estimates of 317 thousand hectares. The causes of this net difference are important to understand, analysis has been undertaken to identify and quantify the various individual sources of this difference, as far as can be determined.

The work to date has concluded that most of the differences are a result of the limitations in previous technology that led to NIWT under-reporting woodland area in some areas and over reporting woodland area in other areas. Anomalies identified in this study as having been due to errors found in the NFI map have been rectified in the latest maps.

These records have been analysed and where pertinent selected for inclusion into the 1990, 1995, 2000, 2005 and 2010 woodland maps.

For further information regarding this work, please see the [National Forest Inventory Woodland Area Statistics for Great Britain](#) report.

2.2.3 Production of 2010, 2015 and 2020 Interim woodland maps

The National Forest Inventory (NFI) is to be a continuous inventory of Britain's woodlands conducted on a five year cycle. The elements of the inventory are a digital map of woodland in Britain constructed from aerial photography, complemented by other sources of information, and a programme of ground surveying of woodland using a representative sample drawn from the woodland and forested areas of Great Britain.

The NFI succeeds a series of single inventories produced by the Forestry Commission, the most recent of which was the National Inventory of Woodland and Trees (NIWT).

The NFI woodland maps covers all forest and woodland area over 0.5 hectare with a minimum of 20% canopy cover (or the potential to achieve it) and a minimum width of 20 metres, including areas of new planting (see *Woodland creation data* section), clearfell and restocked areas. These criteria conform closely to international definitions of woodland and are referred to here as areas of 'NFI woodland'. Wooded areas and individual trees that do not conform to these criteria are referred to by NFI as 'small woods', which are reported on separately⁹.

All forest types are assessed in aerial photography using GIS techniques to establish areas and percentages of tree cover. The woodland is further differentiated into interpreted forest types (IFTs), distinguishing primarily between conifer stands and broadleaved stands, and internal open spaces. All boundaries, woodland and open space are based upon 25-cm-resolution colour aerial photography for England and Scotland and 40-cm-resolution colour aerial photography for Wales. In addition, any open areas of greater than 0.5 hectare that are completely surrounded by NFI woodland are mapped as interpreted open areas (IOAs). Further details regarding the full list of categories in the NFI woodland map can be found in the *NFI description of attributes* document at <https://www.forestresearch.gov.uk/tools-and-resources/national-forest-inventory/how-our-woodlands-might-change-over-time-8211-nfi-forecast-reports/supporting-documents-for-the-nfi/>.

NFI uses an observational approach to estimating woodland area, augmented by reported activity in respect of new planting, and specific tests and conventions in defining the existence and timing of woodland loss in line with international conventions. In the estimation and timing of woodland loss, NFI does not use information on intent or declared policy.

⁹ These smaller wooded features are referred to internationally as 'Other Wooded Land' (OWL).
UKCEH report ... version 1.0

The first NFI woodland map and accompanying report was published in May 2011 (reference date March 2010). Further details can be found in the *National Forest Inventory woodland area statistics for Great Britain* report at www.forestry.gsi.gov.uk/inventory.

The 2006 map comprised the woodland and forest areas identifiable in the aerial photography available at the time. The majority of the photography used was taken around 2006, but smaller areas were taken at other dates ranging from 1999 to 2009. For more detailed information about the date ranges of the imagery used for each country, see Table 6.

Table 6 Aerial photography dates by country used for the 2006 baseline woodland map

| Year flown | % of coverage | | |
|------------|---------------|----------|-------|
| | England | Scotland | Wales |
| 1999 | 2 | | |
| 2000 | 1 | | |
| 2001 | 2 | | |
| 2002 | 4 | 2 | |
| 2003 | 9 | 4 | |
| 2004 | 7 | 19 | |
| 2005 | 27 | 36 | |
| 2006 | 31 | 16 | 100 |
| 2007 | 13 | 8 | |
| 2008 | 4 | 7 | |
| 2009 | | 8 | |

Notes for Table 6:

1. The full 2006 aerial photograph coverage of Wales was provided to the Forestry Commission by the Welsh Assembly.

In addition to this main mapping exercise, areas identified under new planting grant aid schemes between 1990 and 2010 were also added to the map, since when trees are small (i.e. newly planted woodland) it is difficult to clearly identify woodland using aerial photography. To account for this, areas that were identified under the new planting grant aid scheme but were not clearly evident as woodland were added to the map and classified as 'assumed woodland' until the trees became visible. The NFI map was also compared to the 1995-1998 NIWT map and any valid and significant discrepancies identified by the comparison exercise that were greater than 5 hectares in size were added to the 2011 published NFI woodland map. For more information regarding NIWT, please see www.forestry.gov.uk/forestry/inf-d-86xc6c.

Prior to the publication of the first NFI map, additional independent cross-checks of woodland present were made by comparing the NFI map to satellite imagery and alternative assessments of land cover such as those arising from the Countryside Survey. After 2011, satellite imagery was used to identify areas of recent canopy change¹⁰. This included clearfell and windblow areas and areas where a change in land use had occurred since the date of the previous satellite image. Satellite imagery data for 2006, 2009, 2012, 2014-2019 were analysed to produce these updates. The 2006 and 2009 datasets were applied retrospectively to the first NFI map. The 2012, 2014- 2019 imagery sets were used to update the subsequent NFI maps as part of the annual update process.

¹⁰ The areas of canopy loss identified were restricted to areas of over 0.5 hectare in extent.
UKCEH report ... version 1.0

Further improvements to the map were made between 2012, 2015 and 2019 as the following additional datasets became available:

1. Bluesky's National Tree Map (NTM) product for England and Wales.
2. The Native Woodland Survey of Scotland (NWSS).
3. Further NFI/NIWT comparison work (for areas less than 2 hectares).
4. Woodland Carbon Code product.

In addition, if new photography was available, a new aerial photography assessment was undertaken to confirm canopy cover in areas of the map where the photography used in the first assessment was more than 7 years old.

Quality assurance of the NFI woodland maps

The quality and accuracy of the NFI woodland maps produced by FR are of crucial importance as the woodland maps are the base of various statistical reports indicating an extensive and unique record of key information about our forests and woodlands across Great Britain.

The quality assurance process followed to produce the NFI published maps ensures that all data published have been manually validated and confirmed against different data sources to support the changes between the woodland maps.

The NFI uses an observational approach to estimating woodland area and monitor woodland loss. This approach comprises manual interpretation of aerial photography and optical remote sensing imagery available at the time of the assessment and reported activity in respect of new planting.

To ensure the outputs' interpretation is consistent between interpreters, a standardised set of mapping rules was developed ensuring the coherency and uniformity of the woodland maps over the years and across Great Britain. For further information on the NFI mapping rules, please see NFI Map Protocol document available on <https://www.forestresearch.gov.uk/tools-and-resources/national-forest-inventory/>

After an initial training period, junior interpreters are then supervised and mentored by a senior interpreter. During this phase, junior interpreters have an opportunity to deepen their knowledge in the different mapping rules exceptions and on the impact of an inconsistent and bias final output.

Alongside the mentoring phase, at least 20% of the areas proposed for change are reviewed by senior interpreters to ensure that these comply with the NFI mapping rules and it reflects the changes identified within the current monitoring period.

Prior to publication, all geospatial data goes under extensive data integrity checks such as, attributes and topological checks. Additionally, the final outputs are compared with previous versions of the published data. This step ensures that the revised estimates of woodland area are line with the proposed changes and any discrepancies are accounted for.

Due to the volume of data to be analysed and the time constraints of this project, the data used for the reformulation of elements of new planting scheme GIS data 1990 to 2010 in this work, have not been through the same quality control and quality assurance process as the NFI published maps.

The new planting data used to produce the 1990, 1995, 2000 and 2005 woodland maps was mostly taken from its original source. Very limited manual validation and only integrity data checks, e.g. topological checks, were carried out prior to its inclusion in the products above.

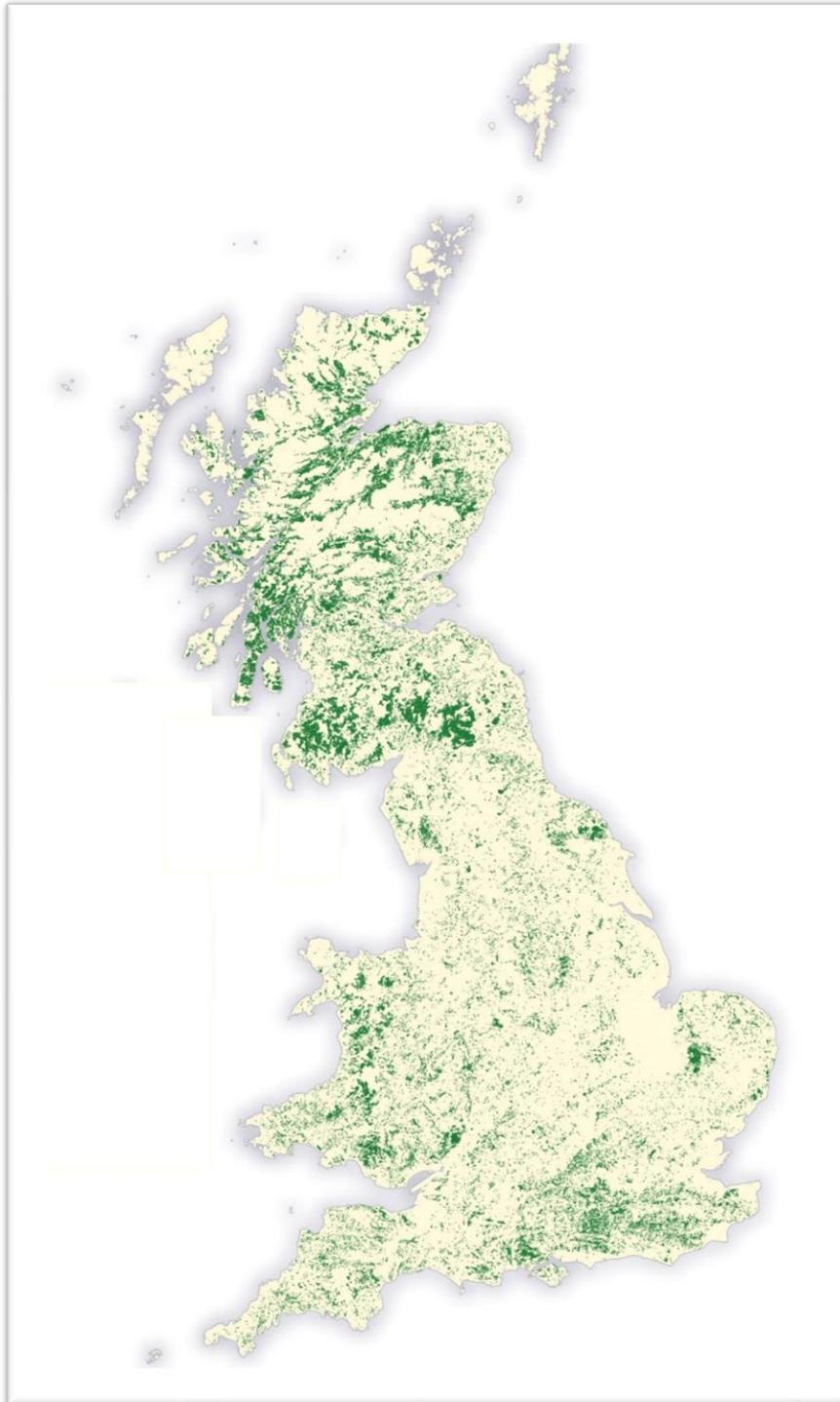


Figure 6 National Forest Inventory 2019 woodland map

2.2.4 Woodland creation data

The afforestation programme across Great Britain began with the creation of the Forestry Commission (FC) in 1919. Since its establishment, the FC, was given the power to acquire and plant land and promote timber supply.

- 1) Public ownership

Since its inception, FC kept detailed paper records of the woodlands under its management. The advances in technology in the 1990s allowed for the inventory of the public forest land to become available geo-spatially. In 2000, the woodland stock maps were brought together with the woodland descriptions into a spatial database. They have been managed and maintained as a single unit since then.

2) Private ownership

To encourage woodland creation and management of existing woodlands, in 1971, the government approved financial incentives to private owners. Although this policy goes back to 1947, geo-spatial records are only available from June 1988 onwards.

Forestry Commission throughout this period kept spatial records of where and when all new woodlands were planted, both public and private. This was initially in paper form and later in GIS.

For the 2010, 2015 and 2020 NFI woodland map, areas of new planting areas, either in public or private land, which haven't already been included in the woodland maps were identified by a geoprocessing function resulting in an output of the balance of the unmapped features.

These records were then assessed by the interpreters against the latest aerial photograph available and, where pertinent, the relevant data was updated to comply with the NFI mapping rules. The relevant data was then selected for inclusion into the woodland maps produced.

During the assessment phase, the preferred classification would be to allocate an attribute, such as young trees. However, due to the imagery availability, there will be areas where no evidence of trees could be found at this stage. In these cases, such areas were included into the woodland areas as Assumed woodland. These areas will feature in various woodland maps until new imagery is available and evidence of trees could be seen.

2.3 Generation of new forestry data for Northern Ireland (AFBI)

2.3.1 Northern Ireland 2020 Woodland Basemap

The Northern Ireland Forest Service Woodland Basemap 2020 (Figure 7) was used as the initial dataset for the identification of woodland polygons in Northern Ireland. Although this indicates some 117,600 ha of woodland, a proportion of the polygons were subsequently found to be scrub or hedgerows (when manually checked with reference to the 2020 Land and Property Service Orthoimage Layer). Scrub dominated polygons within the dataset were identified using the Department of Agriculture, Environment and Rural Affairs (DAERA) Land Parcel Information System (LPIS) Ineligible layer, with some 11,700 ha removed from the Basemap. Detection of the included hedgerows was more

problematic in that a number of the hedgerows are appended onto actual woodland polygons. However, the majority of hedgerows (as single rows of trees) were identified using a polygon thinness ratio formula ($4 * \pi * \text{polygon area} / (\text{polygon perimeter}^2)$), allowing the removal of approximately 1,300 ha from the Basemap.

Although secondary datasets were used to look for additional woodland polygons (Land and Property Service Vegetation Layer, Woodland Trust Sites NI, Woodland Trust Ancient Woodland Inventory), no updates were made. Additional woodland polygons identified from satellite imagery classification (detailed in the section below) were added to the Modified Basemap, resulting in a final 2020 woodland area of approximately 104,700 ha.

The Modified Basemap was merged with the Northern Ireland Forest Service Sub-compartments dataset to allow the inclusion of planting dates for public woodlands. Areas of new woodland plantation were identified, dated and previous land use recorded, evidenced by use of the Land and Property Service Orthoimage Layers (which run in an approximately 3-year update cycle from 2003). No freely available datasets were found prior to 2003, documenting new woodland plantations.

Satellite data

Google Earth Engine was used to create seasonal composite raster mosaics for 1990, 1998, 2002, 2004, 2010, 2014, 2018 and 2020. For all years, atmospherically corrected image collections were used to generate the composite raster; Landsat 5 for maps 1990 – 2010, Landsat 7 & 8 for 2014 and Sentinel 2 for 2018 & 2020 (GEE 2021b, c, d. e)). In order to reduce the impact of cloud cover to a minimum, the composite raster was generated over 3 years, 1 year either side of the target year, except for the 2020 map where 2019 and 2020 source images were used. Cloud-masking was then applied and the data were sampled using the median reducer function to give a representative single pixel value for that time-series. Each mosaic was composed of the bands; blue, green, red, near-infrared (NIR), short-wave infra-red (SWIR) 1 and SWIR2. To include aspects of seasonality in the vegetation cover, composite images for spring, summer and the autumn/winter period were generated for all years apart from 2006, where only spring and summer were used due to significant cloud cover in the rest of the image collection.

Woodland backcast and classification

The Modified Basemap was backcast through the time series with the assumption that the woodlands identified would be generally persistent over that period. In this process polygons flagged as new woodland plantations were sequentially removed based on their planting date and polygons flagged as lost woodland areas were reinstated. The estimated total area of woodland resulting is shown in Table 7.

Image classification was carried out in ESRI ArcPro 2.8 software using the Random Trees classifier (an implementation of the Random Forest classification algorithm (Breiman, 2001)). For each year a single stacked image was created from the seasonal mosaics with up to 18 bands being included. The classification procedure was based on up to 30,000 training pixels per land cover class and using a 200 decision tree limit. With the primary focus of this work being the classification of the woodland polygons identified in the Modified Basemap, the land cover classes in Table 8 were used. A training dataset of 15,000 polygons was sourced from the Forest Service Woodland Basemap 2020 and DAERA LPIS 2020 dataset and used with a 70/30 training/validation split. The area of woodland thus classified for the map sequence and the associated overall accuracy assessment value are illustrated in Table 7.

Table 7 Estimated total woodland areas in Northern Ireland

| Year | Modified Basemap (ha) | Modified Basemap; classification (ha) | Classification accuracy (Kappa) |
|------|-----------------------|---------------------------------------|---------------------------------|
| 1990 | 103,043 | 94,342 | 0.836 |
| 1998 | 103,009 | 95,276 | 0.864 |
| 2002 | 103,183 | 96,485 | 0.883 |
| 2006 | 103,530 | 96,874 | 0.871 |
| 2010 | 103,879 | 98,597 | 0.827 |
| 2014 | 104,203 | 99,544 | 0.854 |
| 2018 | 104,621 | 101,659 | 0.885 |
| 2020 | 104,725 | 101,683 | 0.871 |

Table 8. Land cover classes used in the classification of Northern Ireland woodland polygons.

| Map ID | Land Cover Class |
|--------|-----------------------|
| 1 | Arable |
| 2 | Bog |
| 3 | Broadleaved woodland |
| 4 | Coniferous Woodland |
| 5 | Urban/Suburban |
| 6 | Heather |
| 7 | Improved Grassland |
| 8 | Regenerating Woodland |
| 9 | Scrubland |
| 10 | Heather grassland |
| 11 | Water |

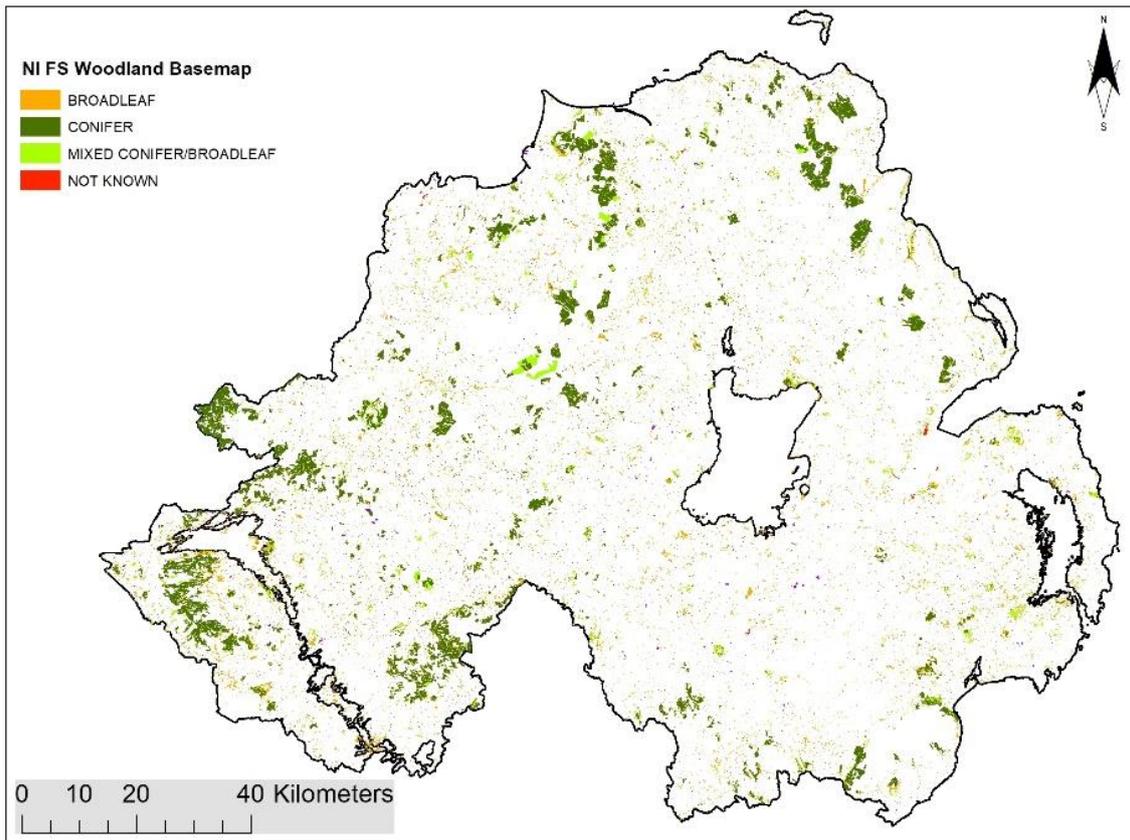


Figure 7. Northern Ireland Forest Service Woodland Basemap 2020.

3 WP-B.2 – Optimise Land Cover Change Data

3.1 LCM Post-classification filtering to improve accuracy

Having the data for eleven Land Cover Maps in a single sqlite database provides the opportunity to applying filtering to remove unlikely land cover changes caused by classification error. This project is the first time that such a database has been available for the UK, so the methods for this kind of processing are currently in their infancy. We conducted a brief literature review which showed that a range of methods have been applied. The most common, and easiest of which, is a running modal filter applied along the time-series, so if a classification result is recorded as class: 1, 2, 1, over 3 dates, then it can be changed to 1, 1, 1, if this is thought to be more appropriate.

To aid the development of post-filtering methods, the database processing was created as follows:

1. The modal class was calculated for each of the LCM polygons for each of the years – for GB this resulted in 11 classifications, for over 6 million polygons, so over 65 million land cover data records. The modal class for LCM's 1994-2020 was calculated using the exactextractr package in R. The 1990 modal class was calculated from the vector version of LCM1990 using an sql query in QGIS:

```
CASE
  WHEN _mode IN (1,2) THEN 1
  WHEN _mode = 3 THEN 2
  WHEN _mode = 4 THEN 3
  WHEN _mode IN (5,6,7,8,9,10,11,16,19) THEN 4
  WHEN _mode IN (13,14) THEN 7
  WHEN _mode IN (20,21) THEN 5
  WHEN _mode IN (12,15,17,18) THEN 6
  ELSE 0
END
```

The attributes for these are named: bc_1990, bc_1994 etc.

2. A new set of modal class columns providing an updated value, if any filtering has been applied to change the values, and that echoes the original value for the polygon if no changes have been applied. These are the values exported to the raster LCM data.
3. Add frequency columns for the seven classes. These attributes count the number of occurrences of the class for each polygon over the time-series. These seven columns are named: freq_c1 etc, where c1 is class 1.

Two related attributes were also calculated to show polygons undergoing rotation between crop/grassland (named: crop_grass_rot) and rotation between crop/grass and rough grassland (named: crop_grass_rg_rot).

4. From the frequency columns, the replacement class is calculated, this is the dominant/modal class across the time-series and it is used as a replacement value, for some of the kbe's. Attribute named: rep_val
5. Then a series of knowledge-enhancement (kbe) rules were created (see Table 9 for details) and applied. The kbe's were designed to tackle two specific types of issues:
 - a. To smooth time-series over time, where the time-series appeared to be unchanged and just exhibited one or two changes in land cover in the middle of the time-series. So, if a forest polygon was forest in 1990 and 2020, but briefly appeared to be rough grassland in 1996. The 1996 value would be change to forest in new_2006 (the appropriate column for the updated value, the original bc_2006 would remain unchanged).
 - b. To resolve specific systematic issues, where there appeared to be a 'drift' in the classification over time. Four corrections of this type were applied (kbe's 6-9, table 9).
6. River correction – one spatial correction is currently applied. The LCM spatial framework was designed not to have narrow, thin polygons, however, some narrow thin features were retained. These features cause issues as the spatial resolution and the spatial precision of the satellite data increases over the time-series, which results in sections of river being classified in later years, but not in earlier years. This gives the appearance of a change to water. To remove these segments a spatial correction was devised, which uses a polyline river data set (Moore *et al.*, 2000), buffers it and converts it to raster. The raster river data is then ingested into the database, as an attribute named: river_prop. In addition, a perimeter area ratio (named: pa_ratio) is calculated for each polygon to help identify long, thin polygons. The

river_prop and pa_ratio are then used in conjunction, to identify polygons containing river segments, which should not be treated as real change (Table 9).

7. Then a series of change columns were added e.g. chng_90_94 for change between 1990 and 1994. The change attributes were populated with 0 for no change over this time period and 1 to indicate a change. These columns were created for each of the time-periods between 1990-2020 and then the number of Land Cover Changes detected for the polygon were recorded in an attribute named: LCC_count.
8. Class-specific columns were then added for each class, to show the year of change to enable review (these can be used to identify single/multiple changes over time) e.g. named attribute: crop populated with years of change, which can be 1994-2019. These columns were primarily created to export subsets of data to explore how particularly classes were faring in the review process. But are also used for the validation sample.

The database, and associated kbe's, were created through an iterative process, with change maps produced, reviewed, issues identified and corrections derived and applied. The filtering was applied cautiously, as there is a risk of over-filtering and removing real change. The process is also complicated by the size of the data-base and the lack of existing data to compare this data with. There are likely to be additional corrections that would improve the quality of this database and future work should explore this area further.

Table 9: Rules applied to filter the Land Cover Change database.

| Rule number | Purpose | Query | kbe code in database |
|-------------|--|---|----------------------|
| 1 | Smooths forestry time-series, requires start and end of time-series to be forest & forest time-series to be forest for more than half-the time-series | bc_1990 == 1 & bc_2020 == 1 & rep_val == 1 & freq_c1 > 6 | 10 |
| 2 | Smooths crop, grass and rough grass time-series, requires start and end as crop/grass and crop/grass for more than half-time-series | (bc_1990 %in% c(2,3,4)) & (bc_2020 %in% c(2,3,4)) & (rep_val %in% c(2,3,4)) & crop_grass_rg_rot > 6 | 20 |
| 3 | Smooths urban time-series, requires start and end as urban & urban for more than half-time-series | bc_1990 == 5 & bc_2020 == 5 & rep_val == 5 & freq_c5 > 6 | 50 |
| 4 | Smooths other time-series, requires start and end as other & other for more than half-time-series | bc_1990 == 6 & bc_2020 == 6 & rep_val == 6 & freq_c6 > 6 | 60 |
| 5 | Smooths water time-series, requires start and end as water & water for more than half-time-series | bc_1990 == 7 & bc_2020 == 7 & rep_val == 7 & freq_c7 > 6 | 70 |
| 6 | Remove change due to inter-tidal polygons that alternate between class 6 and 7 | rep_val >= 6 & freq_c6 > 1 & freq_c7 > 1 | 80 |
| 7 | Remove false change due to different sensitivity to rivers over time (see river correction method above). | freq_c7 > 0 & pa_ratio > 350 & river_prop > 0 | 90 |
| 8 | Removes false urban to other in the Peak District, where there's spectral confusion between urban regeneration & quarries. | bc_1990 == 5 & bc_2020 == 6 & freq_c5 > 3 & freq_c6 > 3 | 100 |
| 9 | Removes false rough grass to inland rock in highlands of Scotland, seems to be due to slight difference in classification. These areas may be more susceptible to errors, because the areas are highly cloudy, suffer from high spectral variability due to topography and also experience high levels of snow cover. This means that the satellite data available to classify them is often limited and of lower quality. | bc_1990 == 4 & bc_2020 == 6 & freq_c4 > 3 & freq_c6 > 3 | 110 |

4 WP-B.3 – Validation of datasets

B.3 validates the accuracy of the Land Cover Data sets by quantifying the biases and random errors. The validation will be applied for the individual Land Cover Maps, the woodland maps and the derived land cover change data. This will require new confusion matrices, with user and producer errors, to be

produced for each of the new LCMs, as well as for each of the land cover change time-periods, thus providing category-by-category basis in line with the AFOLU guidelines

4.1 Evaluation of biases in LCM data sets (validation)

The LCM validation work comprised of four components:

1. **Review of classifications and of temporal filtering** – this has been an ongoing strand of work through the whole project. The evaluating began with qualitative review of the quality of the initial classifications and then the revised classifications, before developing into a review of the preliminary change data as the project progressed. Essentially, the image classification is an iterative process, where initial classifications are run, reviewed, issues assessed and changes made to resolve, as far as possible, the issues identified. In this project, the first challenge to resolve was how best to create the training data set. To this end, a range of training sets were created and their resultant classifications reviewed to identify what worked best. The same process was followed to develop the temporal filtering, with initial sets of change data reviewed with issues identified and resolved, as far as practicable within the constraints of the project.
2. **Confusion matrices for each LCM** – this was conducted using a composite validation data set that was created initially for validation of LCM's 2017-2019 and was then manually reviewed to remove points that were not appropriate for LCM1990, so that it could be used for LCM1990 as well. Consequently, the points should be valid for the period between 1990 and 2020.
3. **Classifications over time – validation of stable/no change fields using IACS data** – change and stability over change for crop and grassland areas was also assessed by comparison against IACS data. Given uncertainties in the IACS data this maybe more of a corroboration, than a validation. IACS data were used to validate change and lack of change of land cover data. Table 10 demonstrates how the EO-date underpinning each of the LCMs was related to the IACS data.
4. **Manually derived change data sets** - Limited data is available on land cover change, so manual interpretation of satellite and aerial photography was used to validate change in: woodland, urban, water and other land. A script was written in Google Earth Engine to display image at the start and the end of the change period, using imagery for the year in question +/- 6 months either side. So to assess change between 2006 and 2010 the filter date ranges would be `.filterDate('2005-07-01', '2007-06-30');` and `.filterDate('2009-07-01', '2011-06-30');`.

The R script used to undertake the analysis of the spatial database, included a function to create random validation data sets with 10 polygons for each change period for each type of change i.e. 1994-1998. This data was exported as sqlite files that were then reprojected to WGS84 and imported into Google Earth as kml files, or into Google Earth Engine as shapefiles. An additional column was added to enable the manual interpreter to record the polygon as: y (for change), n (for no change) and u for uncertain. Where possible, the manual assessment was conducted using the timeline of aerial photography in Google Earth Pro, as this provided better discrimination and assessment of the time of change. Because it's easier to see the establishment/planting of woodland in Google Earth than in the GEE imagery, whilst there is an additional uncertainty when checking the GEE imagery for change. So it is likely that the timing of changes in the earlier time period contain a lag, as they require the woodland signal to become dominant enough for the new woodland to be clear in the Landsat imagery. This could explain some of the poorer correspondence between the change polygons and visual validation using Google Earth imagery for the later time periods. These

factors led to the manual checks against the satellite data being more time-consuming and less certain than the checks against aerial photography.

Table 10: Summary of IACS data and harmonisation with timing of EO image acquisition.

| LCM | IACS temporal equivalence |
|---------|---------------------------|
| LCM2006 | 2006 & 2007 |
| LCM2010 | 2010 & 2011 |
| LCM2015 | 2014 & 2015 |
| LCM2017 | 2017 |
| LCM2018 | 2018 |
| LCM2019 | 2019 |

Combined to look at:

- Change between 2010 and 2006
- Change between 2015 and 2010
- Change between 2017 and 2015
- Change between 2018 and 2017
- Change between 2019 and 2018

4.2 Validation of BEIS LCM’s: individual LCMs

Table 11 and Figure 8 gives the overall accuracy and confidence limits for the reference data set when applied to each of the Land Cover Maps. The confusion matrix statistics were calculated using the caret package in R.

The results show that the overall accuracy is between 86%-89%. Lower overall accuracies are seen for LCM2017 and onwards. This may be because the LCM’s produced in this project benefitted from additional improvements to the method and to the training data set or it may be that the difference in accuracy is due to the satellite data; for LCM2017 onwards the satellite is sourced from one year, compared to 2 years of satellite data for the Landsat-based LCMs (1990-2015). Tables 12-22 give the confusion matrices for LCM’s 1990-2020 and show that the user’s and producer’s accuracies are typically over 90% for forest, urban, other and water. The crop, grass and rough grassland classes tend to exhibit the lowest classification values, however this maybe in part due to issues with the reference data and the assumption that the crop and grassland reference points are not subject to crop/grass rotation; this is discussed further in the IACS validation section. Most of the issues with rough grassland are due to confusion between grassland and rough grassland and reflect confusion around the boundaries of these two classes.

Table 11: Summary of overall accuracy and confidence limits for LCM data.¹ Due to project timing, a preliminary version of LCM2020 was used.

| | Overall Accuracy | 95% confidence limits |
|----------------------------|------------------|-----------------------|
| LCM1990 | 0.8871 | 0.882, 0.8921 |
| LCM1994 | 0.8848 | 0.8796, 0.8899 |
| LCM1998 | 0.8848 | 0.8796, 0.8899 |
| LCM2002 | 0.8894 | 0.8843, 0.8943 |
| LCM2006 | 0.8875 | 0.8824, 0.8925 |
| LCM2010 | 0.8919 | 0.8868, 0.8968 |
| LCM2015 | 0.8815 | 0.8762, 0.8866 |
| LCM2017 | 0.8696 | 0.8641, 0.8749 |
| LCM2018 | 0.8763 | 0.8709, 0.8815 |
| LCM2019 | 0.8711 | 0.8657, 0.8764 |
| LCM2020¹ | 0.8606 | 0.8550, 0.8661 |

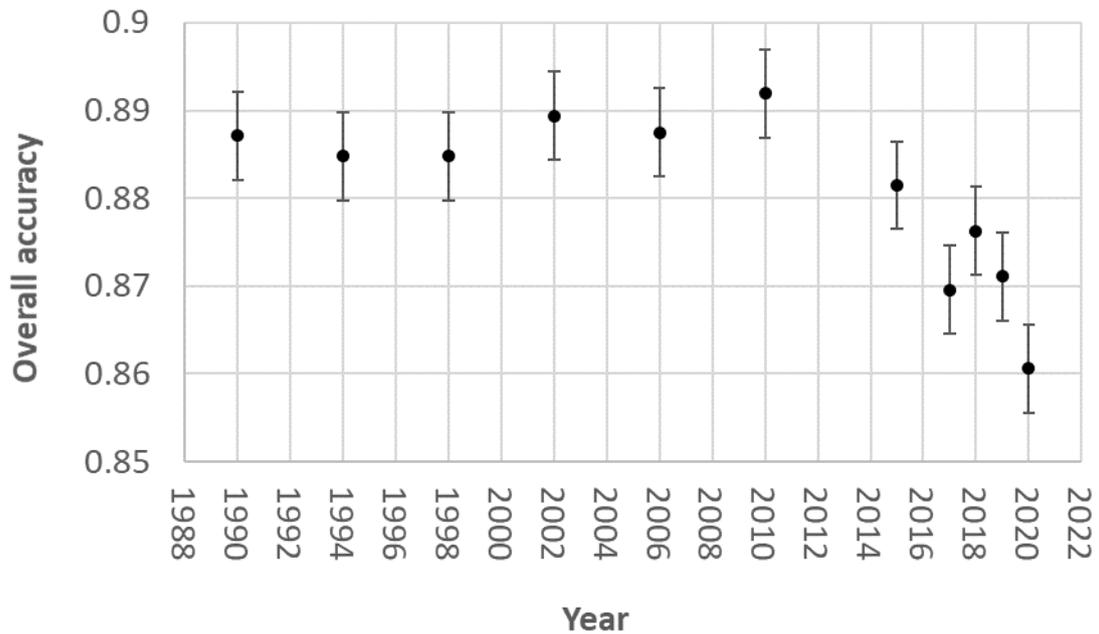


Figure 8: Overall accuracy of the LCM data sets, with 95% confidence limits (due to project timing a preliminary version of LCM2020 was used).

Table 12: Validation of LCM1990, against point reference data set.

| BEIS-LCM1990 | | Reference data | | | | | | | User's Accuracy (%) |
|-------------------------|-----------------|----------------|------|-------|-----------------|-------|-------|-------|---------------------|
| | | Forest | Crop | Grass | Rough grassland | Urban | Other | Water | |
| LCM data | Forest | 1102 | 1 | 13 | 24 | 27 | 1 | 1 | 94 |
| | Crop | 27 | 1491 | 279 | 33 | 42 | 0 | 0 | 80 |
| | Grass | 61 | 166 | 2547 | 580 | 75 | 3 | 0 | 74 |
| | Rough grassland | 67 | 13 | 170 | 4997 | 23 | 11 | 3 | 95 |
| | Urban | 10 | 2 | 3 | 0 | 2751 | 1 | 1 | 99 |
| | Other | 2 | 2 | 1 | 25 | 28 | 252 | 2 | 81 |
| | Water | 0 | 0 | 0 | 2 | 2 | 13 | 332 | 95 |
| Producer's Accuracy (%) | | 87 | 89 | 85 | 88 | 93 | 90 | 98 | |

Table 13: Validation of LCM1994.

| BEIS-LCM1994 | | Reference data | | | | | | | |
|--------------|---------------------|----------------|------|-------|-----------------|-------|-------|-------|-----------------|
| | | Forest | Crop | Grass | Rough grassland | Urban | Other | Water | User's Accuracy |
| LCM data | Forest | 1149 | 4 | 20 | 33 | 8 | 1 | 1 | 94 |
| | Crop | 13 | 1433 | 346 | 36 | 17 | 0 | 0 | 78 |
| | Grass | 37 | 214 | 2419 | 569 | 34 | 1 | 0 | 74 |
| | Rough grassland | 58 | 21 | 221 | 4977 | 12 | 8 | 1 | 94 |
| | Urban | 12 | 3 | 7 | 3 | 2860 | 2 | 1 | 99 |
| | Other | 0 | 0 | 0 | 41 | 17 | 266 | 3 | 81 |
| | Water | 0 | 0 | 0 | 2 | 0 | 3 | 333 | 99 |
| | Producer's Accuracy | 91 | 86 | 80 | 88 | 97 | 95 | 98 | |

Table 14: Validation of LCM1998.

| BEIS-LCM1998 | | Reference data | | | | | | | |
|--------------|---------------------|----------------|------|-------|-----------------|-------|-------|-------|-----------------|
| | | Forest | Crop | Grass | Rough grassland | Urban | Other | Water | User's Accuracy |
| LCM data | Forest | 1177 | 2 | 22 | 46 | 6 | 0 | 1 | 94 |
| | Crop | 13 | 1422 | 332 | 37 | 13 | 0 | 1 | 78 |
| | Grass | 34 | 221 | 2414 | 565 | 31 | 2 | 0 | 74 |
| | Rough grassland | 35 | 28 | 236 | 4959 | 10 | 13 | 3 | 94 |
| | Urban | 10 | 2 | 8 | 4 | 2875 | 5 | 0 | 99 |
| | Other | 0 | 0 | 1 | 48 | 13 | 258 | 2 | 80 |
| | Water | 0 | 0 | 0 | 2 | 0 | 3 | 332 | 99 |
| | Producer's Accuracy | 93 | 85 | 80 | 88 | 98 | 92 | 98 | |

Table 15: Validation of LCM2002.

| BEIS-LCM2002 | | Reference data | | | | | | | |
|--------------|---------------------|----------------|------|-------|-----------------|-------|-------|-------|-----------------|
| | | Forest | Crop | Grass | Rough grassland | Urban | Other | Water | User's Accuracy |
| LCM data | Forest | 1177 | 5 | 35 | 42 | 7 | 0 | 1 | 93 |
| | Crop | 6 | 1419 | 331 | 34 | 12 | 1 | 1 | 79 |
| | Grass | 34 | 229 | 2442 | 533 | 26 | 1 | 0 | 75 |
| | Rough grassland | 40 | 20 | 197 | 5001 | 19 | 11 | 2 | 95 |
| | Urban | 12 | 2 | 7 | 3 | 2871 | 2 | 0 | 99 |
| | Other | 0 | 0 | 1 | 43 | 13 | 263 | 2 | 82 |
| | Water | 0 | 0 | 0 | 5 | 0 | 3 | 333 | 98 |
| | Producer's Accuracy | 93 | 85 | 81 | 88 | 97 | 94 | 98 | |

Table 16: Validation of LCM2006.

| BEIS-LCM2006 | | Reference data | | | | | | | User's Accuracy |
|---------------------|-----------------|----------------|------|-------|-----------------|-------|-------|-------|-----------------|
| | | Forest | Crop | Grass | Rough grassland | Urban | Other | Water | |
| LCM data | Forest | 1183 | 6 | 30 | 47 | 5 | 0 | 1 | 93 |
| | Crop | 11 | 1355 | 330 | 45 | 9 | 0 | 1 | 77 |
| | Grass | 27 | 292 | 2466 | 540 | 22 | 1 | 0 | 74 |
| | Rough grassland | 39 | 20 | 180 | 4989 | 11 | 11 | 1 | 95 |
| | Urban | 9 | 2 | 6 | 3 | 2888 | 4 | 0 | 99 |
| | Other | 0 | 0 | 1 | 35 | 13 | 263 | 2 | 84 |
| | Water | 0 | 0 | 0 | 2 | 0 | 2 | 334 | 99 |
| Producer's Accuracy | | 93 | 81 | 82 | 88 | 98 | 94 | 99 | |

Table 17: Validation of LCM2010.

| BEIS-LCM2010 | | Reference data | | | | | | | User's Accuracy |
|---------------------|-----------------|----------------|------|-------|-----------------|-------|-------|-------|-----------------|
| | | Forest | Crop | Grass | Rough grassland | Urban | Other | Water | |
| LCM data | Forest | 1194 | 7 | 36 | 57 | 7 | 0 | 1 | 92 |
| | Crop | 7 | 1400 | 342 | 26 | 9 | 0 | 1 | 78 |
| | Grass | 29 | 250 | 2457 | 523 | 16 | 1 | 0 | 75 |
| | Rough grassland | 29 | 15 | 168 | 5009 | 13 | 13 | 1 | 95 |
| | Urban | 10 | 2 | 10 | 3 | 2887 | 1 | 0 | 99 |
| | Other | 0 | 1 | 0 | 41 | 16 | 263 | 2 | 81 |
| | Water | 0 | 0 | 0 | 2 | 0 | 3 | 334 | 99 |
| Producer's Accuracy | | 94 | 84 | 82 | 88 | 98 | 94 | 99 | |

Table 18: Validation of LCM2015.

| BEIS-LCM2015 | | Reference data | | | | | | | User's Accuracy |
|---------------------|-----------------|----------------|------|-------|-----------------|-------|-------|-------|-----------------|
| | | Forest | Crop | Grass | Rough grassland | Urban | Other | Water | |
| LCM data | Forest | 1180 | 8 | 24 | 87 | 7 | 1 | 1 | 90 |
| | Crop | 9 | 1407 | 421 | 65 | 6 | 1 | 0 | 74 |
| | Grass | 37 | 245 | 2466 | 669 | 21 | 2 | 1 | 72 |
| | Rough grassland | 25 | 7 | 72 | 4821 | 0 | 5 | 1 | 98 |
| | Urban | 18 | 8 | 29 | 2 | 2913 | 4 | 1 | 98 |
| | Other | 0 | 0 | 1 | 15 | 1 | 266 | 2 | 93 |
| | Water | 0 | 0 | 0 | 2 | 0 | 2 | 333 | 99 |
| Producer's Accuracy | | 93 | 84 | 82 | 85 | 99 | 95 | 98 | |

Table 19: Validation of LCM2017.

| BEIS-LCM2017 | | Reference data | | | | | | | User's Accuracy |
|---------------------|-----------------|----------------|------|-------|-----------------|-------|-------|-------|-----------------|
| | | Forest | Crop | Grass | Rough grassland | Urban | Other | Water | |
| LCM data | Forest | 1202 | 13 | 49 | 127 | 9 | 1 | 1 | 86 |
| | Crop | 6 | 1364 | 411 | 49 | 3 | 2 | 1 | 74 |
| | Grass | 20 | 270 | 2381 | 702 | 25 | 3 | 0 | 70 |
| | Rough grassland | 27 | 18 | 151 | 4755 | 5 | 7 | 2 | 96 |
| | Urban | 14 | 9 | 21 | 7 | 2906 | 2 | 0 | 98 |
| | Other | 0 | 1 | 0 | 18 | 0 | 264 | 2 | 93 |
| | Water | 0 | 0 | 0 | 3 | 0 | 2 | 333 | 99 |
| Producer's Accuracy | | 95 | 81 | 79 | 84 | 99 | 94 | 98 | |

Table 20: Validation of LCM2018.

| BEIS-LCM2018 | | Reference data | | | | | | | User's Accuracy |
|---------------------|-----------------|----------------|------|-------|-----------------|-------|-------|-------|-----------------|
| | | Forest | Crop | Grass | Rough grassland | Urban | Other | Water | |
| LCM data | Forest | 1199 | 15 | 28 | 114 | 9 | 1 | 1 | 88 |
| | Crop | 4 | 1371 | 392 | 28 | 2 | 2 | 1 | 76 |
| | Grass | 22 | 275 | 2437 | 704 | 26 | 2 | 0 | 70 |
| | Rough grassland | 29 | 6 | 137 | 4795 | 1 | 10 | 1 | 96 |
| | Urban | 15 | 8 | 19 | 4 | 2910 | 3 | 0 | 98 |
| | Other | 0 | 0 | 0 | 14 | 0 | 261 | 2 | 94 |
| | Water | 0 | 0 | 0 | 2 | 0 | 2 | 334 | 99 |
| Producer's Accuracy | | 94 | 82 | 81 | 85 | 99 | 93 | 99 | |

Table 21: Validation of LCM2019.

| BEIS-LCM2019 | | Reference data | | | | | | | User's Accuracy |
|---------------------|-----------------|----------------|------|-------|-----------------|-------|-------|-------|-----------------|
| | | Forest | Crop | Grass | Rough grassland | Urban | Other | Water | |
| LCM data | Forest | 1202 | 11 | 25 | 143 | 8 | 1 | 1 | 86 |
| | Crop | 3 | 1369 | 412 | 33 | 1 | 0 | 0 | 75 |
| | Grass | 23 | 273 | 2407 | 725 | 26 | 3 | 0 | 70 |
| | Rough grassland | 27 | 11 | 148 | 4744 | 2 | 8 | 1 | 96 |
| | Urban | 14 | 9 | 20 | 2 | 2911 | 5 | 1 | 98 |
| | Other | 0 | 2 | 1 | 12 | 0 | 262 | 2 | 94 |
| | Water | 0 | 0 | 0 | 2 | 0 | 2 | 334 | 99 |
| Producer's Accuracy | | 95 | 82 | 80 | 84 | 99 | 93 | 99 | |

Table 22: Validation of LCM2020.

| BEIS-LCM2020 | | Reference data | | | | | | | User's Accuracy |
|---------------------|-----------------|----------------|------|-------|-----------------|-------|-------|-------|-----------------|
| | | Forest | Crop | Grass | Rough grassland | Urban | Other | Water | |
| LCM data | Forest | 1184 | 15 | 26 | 150 | 6 | 5 | 1 | 85 |
| | Crop | 3 | 1348 | 407 | 40 | 0 | 1 | 1 | 75 |
| | Grass | 24 | 282 | 2374 | 705 | 23 | 9 | 0 | 69 |
| | Rough grassland | 43 | 15 | 176 | 4735 | 3 | 67 | 4 | 94 |
| | Urban | 15 | 14 | 24 | 9 | 2916 | 13 | 1 | 97 |
| | Other | 0 | 1 | 6 | 17 | 0 | 173 | 4 | 86 |
| | Water | 0 | 0 | 0 | 2 | 0 | 2 | 324 | 99 |
| Producer's Accuracy | | 93 | 80 | 79 | 84 | 99 | 64 | 97 | |

4.3 Comparison with IACS data

To conduct the comparison against the IACS data, the IACS raster layers created by WP-A were used. 20,000 points were randomly distributed across the area for which IACS data were available (England) and then reference values were extracted from the IACS data and from the LCM data for each of the available time periods. The same set of points has been used here for each of the time periods. The IACS and LCM data were rescaled to two-digit numbers, with the first digit representing the land cover for the first point in time and the second digit representing the second point in time. So, a woodland pixel that does not change would be class 1 and then class 1 i.e. 11. A conversion from cropland (class 2) to grassland (class 3) would be 23.

The results of the comparison with the IACS data are summarised in Tables 23 and 24, with full results given in (Tables 25-29). The overall agreement between them varies between 90.29% (change between LCM2010 to LCM2015) to 56% (change between 2018 and 2019).

Table 23 shows the degree of correspondence for the change categories and shows that the maximum values vary from lows of around 25% to high values of around 60%. The values are noticeably lower for the most recent LCM's which have moved onto an annual production cycle. The cause of this requires further analysis. Table 24 suggests that the issue may be due to classification accuracy, however, the decrease observed in producer's and user's accuracy is primarily due to confusion between the Rough grassland and Grassland split. So further exploration of the LCM and IACS data is required to understand the underlying cause of this drop in correspondence. It maybe that more targeted processing of the IACS data would produce better correspondence, as in the current validation data set, no effort was made to avoid points falling near field boundaries. A more serious issue though, may be the mis-match between the timing of the satellite data for the LCM's for 2017 onwards, which use calendar years (Jan. 1st – Dec. 31st), and the cropping cycle, which is affecting accuracy. This could be assessed by targeted reprocessing of the LCM2017-2019 data for an area with substantial arable-grassland churn, to see whether sourcing the satellite data over years that are more tuned to the cropping cycle (i.e. November-October) would produce better correspondence with the IACS change data. This reprocessed data would provide a better basis for understanding the accuracy of annual change data.

The IACS data and the individual ILCM data sets were not compared, but the results of the IACS comparison suggests that this would be useful. For the stable cropland class (denoted as class 22), the

percentage accuracies are between 93%-96%. This is much higher than the accuracies for crop reported in the validation of the stable LCM data and highlights the issues associated with assumption that underlie some of that data that crop and grassland fields remain stable over time.

Table 23: Summary of LCM and IACS correspondence for change categories.

| Focus on change categories | User's accuracy | | Producer's accuracy | |
|----------------------------|-----------------|-----|---------------------|-----|
| | Average | Max | Average | Max |
| 2006-2010 | 25% | 62% | 46% | 51% |
| 2010-2015 | 47% | 68% | 54% | 60% |
| 2015-2017 | 30% | 37% | 31% | 44% |
| 2017-2018 | 12% | 23% | 16% | 26% |
| 2018-2019 | 16% | 29% | 21% | 35% |

Table 24: Summary of LCM and IACS correspondence for stable categories.

| Focus on no change categories | User's accuracy | | Producer's accuracy | |
|-------------------------------|-----------------|-----|---------------------|-----|
| | Average | Max | Average | Max |
| 2006-2010 | 83% | 96% | 79% | 96% |
| 2010-2015 | 95% | 97% | 93% | 95% |
| 2015-2017 | 91% | 95% | 91% | 93% |
| 2017-2018 | 59% | 98% | 58% | 93% |
| 2018-2019 | 59% | 98% | 58% | 94% |

Table 25: Results of comparing IACS data for **change between 2006 and 2010** (to interpret numbers, first digit represents class number at first time point, second digit represents class number at second time point). Pale blue highlights the main diagonal, where the two classifications correspond.

| | | IACS data | | | | | | | | | | | | | | |
|----------|----|-------------------------|----|----|----|----|------|----|----|----|----|------|----|--------------|----|----|
| LCM data | | 11 | 12 | 13 | 14 | 21 | 22 | 23 | 24 | 31 | 32 | 33 | 34 | 41 | 42 | 44 |
| | 11 | 48 | 0 | 1 | 0 | 1 | 8 | 1 | 1 | 3 | 0 | 61 | 0 | 1 | 0 | 3 |
| | 13 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 15 | 0 | 0 | 0 | 0 |
| | 14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 0 | 0 | 0 | 0 |
| | 21 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| | 22 | 6 | 0 | 0 | 0 | 1 | 2189 | 23 | 0 | 0 | 63 | 53 | 1 | 0 | 0 | 1 |
| | 23 | 0 | 0 | 0 | 0 | 0 | 18 | 57 | 0 | 0 | 1 | 59 | 0 | 0 | 0 | 1 |
| | 24 | 1 | 0 | 0 | 1 | 0 | 8 | 4 | 0 | 0 | 1 | 13 | 0 | 0 | 0 | 0 |
| | 25 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| | 31 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 13 | 0 | 0 | 0 | 0 |
| | 32 | 2 | 0 | 0 | 0 | 0 | 13 | 0 | 0 | 1 | 88 | 37 | 0 | 0 | 0 | 1 |
| | 33 | 5 | 1 | 0 | 0 | 0 | 34 | 22 | 0 | 1 | 5 | 1688 | 1 | 0 | 1 | 2 |
| | 34 | 1 | 0 | 0 | 0 | 0 | 1 | 2 | 0 | 1 | 3 | 90 | 0 | 0 | 0 | 0 |
| | 35 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 |
| | 37 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| | 41 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 0 | 0 | 0 |
| | 42 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 9 | 9 | 0 | 0 | 0 | 0 |
| | 43 | 1 | 0 | 0 | 0 | 0 | 3 | 3 | 0 | 0 | 0 | 86 | 0 | 0 | 0 | 0 |
| | 44 | 3 | 0 | 0 | 0 | 0 | 5 | 6 | 0 | 1 | 1 | 451 | 1 | 0 | 0 | 0 |
| | 45 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 51 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 54 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | |
| 55 | 0 | 0 | 0 | 0 | 0 | 3 | 1 | 0 | 0 | 2 | 5 | 0 | 0 | 0 | 1 | |
| 66 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | |
| 77 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 1 | |
| | | Overall accuracy | | | | | | | | | | | | 76.91 | | |

Table 26: Results of comparing IACS data for **change between 2010 and 2015** (to interpret numbers, first digit represents class number at first time point, second digit represents class number at second time point). Pale blue highlights the main diagonal, where the two classifications correspond.

| | | IACS data | | | | | | | |
|----------|----|-------------------------|-------------|-----------|------------|------------|----------|--------------|----------|
| | | 12 | 22 | 23 | 32 | 33 | 34 | 42 | 44 |
| LCM data | 11 | 0 | 21 | 2 | 1 | 2 | 0 | 0 | 1 |
| | 12 | 0 | 2 | 0 | 1 | 0 | 0 | 0 | 0 |
| | 13 | 0 | 2 | 1 | 1 | 0 | 0 | 0 | 0 |
| | 21 | 0 | 10 | 1 | 0 | 0 | 0 | 0 | 0 |
| | 22 | 1 | 4879 | 72 | 43 | 18 | 0 | 0 | 2 |
| | 23 | 0 | 85 | 93 | 2 | 13 | 1 | 0 | 0 |
| | 24 | 0 | 8 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 25 | 0 | 8 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 26 | 0 | 2 | 0 | 0 | 1 | 0 | 0 | 0 |
| | 32 | 0 | 42 | 3 | 150 | 27 | 0 | 0 | 0 |
| | 33 | 0 | 51 | 15 | 29 | 105 | 3 | 0 | 2 |
| | 34 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 |
| | 35 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| | 41 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 42 | 0 | 7 | 0 | 15 | 4 | 0 | 0 | 0 |
| | 43 | 0 | 9 | 3 | 3 | 11 | 0 | 0 | 2 |
| | 44 | 0 | 2 | 0 | 2 | 0 | 0 | 0 | 0 |
| | 45 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| | 47 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| | 52 | 0 | 4 | 0 | 0 | 1 | 0 | 0 | 0 |
| 53 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | |
| 55 | 0 | 12 | 1 | 0 | 1 | 0 | 0 | 1 | |
| 77 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | |
| | | Overall accuracy | | | | | | 90.29 | |

Table 27: Results of comparing IACS data for **change between 2015 and 2017** (to interpret numbers, first digit represents class number at first time point, second digit represents class number at second time point). Pale blue highlights the main diagonal, where the two classifications correspond.

| | | IACS data | | | | | | |
|----------|----|-------------------------|-----|----|----|-----|--------------|----|
| | | 22 | 23 | 24 | 32 | 33 | 34 | 44 |
| LCM data | 11 | 26 | 1 | 0 | 1 | 5 | 0 | 1 |
| | 12 | 5 | 0 | 0 | 0 | 1 | 0 | 0 |
| | 13 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| | 14 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 21 | 4 | 1 | 0 | 0 | 0 | 0 | 0 |
| | 22 | 4789 | 154 | 4 | 59 | 39 | 7 | 2 |
| | 23 | 96 | 96 | 5 | 7 | 41 | 13 | 0 |
| | 24 | 21 | 2 | 1 | 1 | 3 | 0 | 0 |
| | 25 | 3 | 0 | 0 | 0 | 1 | 0 | 0 |
| | 26 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 27 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 31 | 2 | 0 | 0 | 1 | 0 | 1 | 1 |
| | 32 | 87 | 3 | 0 | 65 | 30 | 8 | 0 |
| | 33 | 73 | 22 | 1 | 11 | 172 | 52 | 6 |
| | 34 | 3 | 1 | 0 | 0 | 4 | 1 | 0 |
| | 35 | 3 | 1 | 0 | 0 | 1 | 0 | 1 |
| | 42 | 6 | 0 | 0 | 1 | 0 | 0 | 0 |
| | 43 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| | 44 | 6 | 0 | 0 | 0 | 0 | 1 | 0 |
| | 52 | 7 | 0 | 0 | 0 | 0 | 0 | 0 |
| 53 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | |
| 55 | 14 | 0 | 0 | 2 | 1 | 0 | 2 | |
| 62 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 63 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | |
| | | Overall accuracy | | | | | 85.53 | |

Table 28: Results of comparing IACS data for **change between 2017 and 2018** (to interpret numbers, first digit represents class number at first time point, second digit represents class number at second time point). Pale blue highlights the main diagonal, where the two classifications correspond.

| | | IACS data | | | | | | | | | |
|----------|----|-------------------------|------|-----|----|-----|-----|----|----|--------------|------|
| | | 11 | 22 | 23 | 24 | 32 | 33 | 34 | 42 | 43 | 44 |
| LCM data | 11 | 2 | 27 | 4 | 0 | 3 | 6 | 2 | 0 | 0 | 221 |
| | 12 | 1 | 4 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 2 |
| | 13 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 33 |
| | 14 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 |
| | 21 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 |
| | 22 | 0 | 5026 | 154 | 11 | 129 | 124 | 14 | 14 | 4 | 146 |
| | 23 | 0 | 51 | 72 | 5 | 1 | 60 | 7 | 1 | 1 | 118 |
| | 24 | 0 | 13 | 3 | 0 | 1 | 2 | 0 | 1 | 0 | 10 |
| | 25 | 0 | 4 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 2 |
| | 26 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 31 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 16 |
| | 32 | 0 | 98 | 8 | 1 | 40 | 44 | 8 | 29 | 2 | 72 |
| | 33 | 1 | 103 | 25 | 6 | 17 | 320 | 91 | 18 | 18 | 3482 |
| | 34 | 0 | 3 | 1 | 0 | 2 | 3 | 2 | 0 | 0 | 49 |
| | 35 | 0 | 1 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 9 |
| | 36 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| | 37 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| | 41 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 |
| | 42 | 0 | 18 | 1 | 0 | 2 | 2 | 0 | 0 | 0 | 10 |
| | 43 | 0 | 5 | 0 | 0 | 0 | 3 | 2 | 0 | 0 | 67 |
| 44 | 0 | 14 | 2 | 3 | 0 | 6 | 2 | 2 | 0 | 1514 | |
| 46 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | |
| 51 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | |
| 52 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | |
| 53 | 0 | 2 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 10 | |
| 54 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 55 | 0 | 13 | 1 | 0 | 1 | 4 | 0 | 0 | 0 | 29 | |
| 64 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 66 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 14 | |
| 77 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 11 | |
| | | Overall accuracy | | | | | | | | 55.65 | |

Table 29: Results of comparing IACS data for **change between 2018 and 2019** (to interpret numbers, first digit represents class number at first time point, second digit represents class number at second time point). Pale blue highlights the main diagonal, where the two classifications correspond.

| | | IACS data | | | | | | | | | | |
|----------|----|-------------------------|----|------|-----|----|-----|-----|----|----|--------------|------|
| | | 11 | 12 | 22 | 23 | 24 | 32 | 33 | 34 | 42 | 43 | 44 |
| LCM data | 11 | 0 | 0 | 26 | 3 | 1 | 3 | 5 | 0 | 2 | 0 | 243 |
| | 12 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 2 |
| | 13 | 0 | 0 | 0 | 2 | 0 | 0 | 1 | 0 | 1 | 0 | 11 |
| | 14 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 |
| | 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| | 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| | 21 | 0 | 1 | 7 | 2 | 0 | 0 | 2 | 0 | 0 | 0 | 1 |
| | 22 | 0 | 0 | 5053 | 119 | 7 | 160 | 110 | 2 | 23 | 6 | 157 |
| | 23 | 0 | 0 | 86 | 91 | 3 | 8 | 53 | 2 | 2 | 1 | 69 |
| | 24 | 0 | 0 | 6 | 2 | 0 | 1 | 4 | 0 | 0 | 0 | 8 |
| | 25 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 26 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 27 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| | 31 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 22 |
| | 32 | 0 | 0 | 37 | 7 | 1 | 46 | 63 | 4 | 32 | 2 | 102 |
| | 33 | 1 | 0 | 109 | 27 | 10 | 15 | 330 | 32 | 19 | 12 | 3568 |
| | 34 | 0 | 0 | 1 | 0 | 0 | 1 | 3 | 0 | 1 | 1 | 54 |
| | 35 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 |
| | 41 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 |
| | 42 | 0 | 0 | 14 | 1 | 0 | 2 | 4 | 0 | 1 | 0 | 6 |
| | 43 | 1 | 0 | 2 | 2 | 0 | 1 | 3 | 0 | 1 | 0 | 50 |
| | 44 | 1 | 0 | 17 | 0 | 1 | 0 | 8 | 0 | 0 | 1 | 1524 |
| | 46 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| | 47 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| | 51 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| | 52 | 0 | 0 | 2 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| | 53 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 |
| | 54 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 55 | 0 | 0 | 15 | 1 | 0 | 1 | 6 | 0 | 1 | 0 | 30 | |
| 61 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | |
| 62 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | |
| 64 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | |
| 66 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 16 | |
| 72 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 73 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | |
| 77 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 12 | |
| | | Overall accuracy | | | | | | | | | 56.06 | |

4.4 Manually derived change data sets: Change to woodland

From the 90 points in the woodland change validation data set, 44 were correctly identified as change, 44 were incorrectly identified, and two were classified as uncertain, giving an accuracy of 50% when excluding the uncertain polygons. The polygons that were incorrectly identified as change were due to two factors:

- Mixed polygons – where multiple land cover types occur within a single polygon happen when the spatial framework poorly represents the landscape (Figure 8). These cause issues for the change, because they are changing from apparently being land cover type A to land cover type B at some point along the time-series, with little real change. For this project, the modal proportions (the fraction of the polygon covered by the dominant class) and frequency distributions showing the coverage of each class within the polygon were not calculated. However, it maybe that these metrics could be explored to see if they can provide a systematic correction for such polygons. A similar issue was also found with some urban polygons.
- Classification error in one or both of the classifications.



Figure 9: Example of mixed polygons from the woodland change validation data set.

4.5 Manually derived change data sets: Change to urban

From the 90 points in the urban change validation data set, 28 were correctly identified as change, 38 were incorrect identified, and 24 were classified as uncertain, giving an accuracy of 42%. The large number of polygons that were recorded as uncertain were primarily for the earlier time periods, which relied on satellite data where it was not always clear whether a polygon had changed, this is especially the case for urban areas, as the polygons are typically small. The change to urban is subject to a

systematic level of noise, caused primarily by mixed polygons (Figure 10), but also some misclassification, which reduces the change accuracy achieved. Further work is needed to explore the potential for a systemic correction to rectify this issue.



Figure 10: Example of mixed land cover polygons from the urban change validation data set, which appear to change from non-urban to urban at some point in the time-series.

4.6 Manually derived change data sets: Change to water

From the 90 points in the urban change validation data set, 60 were correctly identified as change, 27 were incorrectly identified, and 3 were classified as uncertain, giving an accuracy of 69%.

4.7 Manually derived change data sets: Change to other land

From the 90 points in the change to other validation data set, 61 were correctly identified as change, with 29 incorrectly identified, giving an accuracy of 68%. Some of the incorrectly identified change to other polygons were coastal polygons. The coastal polygons are inter-tidal areas and can legitimately be classified as other or as water, depending on tidal state. A filtering rule (Table 9, rule 6) was implemented to resolve this, but there still appears to be an issue with some polygons. The coastal correction focussed on polygons alternating between classes 6 (other) and 7 (water), but does not currently resolve issues with polygons that may have been misclassified at some point across the time-series. Future work should review the incorrect coastal polygons so that the implementation of this inter-tidal correction can be revised.

4.8 Summary of LCM data products

The final data sets produced for this project are:

- LCM's in the LULUCF classes for: 1990, 1994, 1998, 2002, 2006, 2010, 2015, 2017, 2018, 2019, 2020 that have undergone some temporal filtering.
- Geospatial databases for GB and NI containing the above LCM data, before and after post-classification filtering.

- Maps showing the application of knowledge-based enhancements, the number of land cover changes and whether a particular area appears to be subject to crop/grassland rotations (binary data set).

4.9 Discussion of LCM results

At the start of this project there were five LCM data sets for the UK covering 1990, 2015, 2017, 2018 and 2019. Following the work in this project there are an additional 5 LCM data sets for 1994, 1998, 2002, 2006 and 2010. This enabled the creation of a spatial database containing all 11 LCM data sets; for GB this dataset contains in excess of 60 million records on land cover at different points in time. This new data set enabled the first large-scale exploration of the potential of temporal filtering to improve the ability to map change for the UK. The results so far show some promise, for example the accuracy achieved with mapping change in water over time, but for other types of change there are still challenges that need to be resolved. This discussion below focusses on some of the key areas outstanding:

Comprehensive review of results – the first stage of follow-on work should be a comprehensive review of what is working well and what is working poorly, building on the validation work conducted here to identify how best to resolve remaining issues. The review should identify the classes and types of change that are currently being mapped with large uncertainties, with a focus on understanding why these issues are occurring, so that they can be explored to identify systematic solutions. This project has developed a range of strategies for resolving specific issues, and these will be useful for resolving the outstanding issues, as some issues may require edits to the training data and re-running of classifications, whilst other may just require some form of post-classification correction. The change to water has already highlighted that a systematic correction can reduce uncertainty and remove false positives.

Further development of post-classification filtering – the results here are based on a cautious application of post-classification filtering, so in most cases up to 2 classification results could be changed across the time-series of 11 values. Future work would benefit from exploring such filtering in greater depth, initially for a small set of test areas that can be checked in great detail to avoid any undesirable outcomes. This would also allow the comparison of the methods applied here, with a running modal filter based method. Further exploration would also help determine to what extent it is possible to develop robust methods to remove false positives/changes.

Detection of woodland change - There is a lag between trees being planted and new woodland being detectable in the satellite data. How long the lag is depends on the scale and method of tree planting, as well as how long the trees take to establish a detectable level of canopy cover. Wooded areas were recorded as change, even if the timing of planting was unclear and in some cases may have been up to 10 years prior to detection in the change data. If LCM is needed for LULUCF activity data on tree planting then more research will be required to understand the typical lag-time for detection of new coniferous and deciduous woodland, both within small woodland patches and large/scale plantations. Note, issues such as rewilding and natural regeneration of woodland will occur on slower time-scales and may have different assumptions, although these processes are operating at a very low level in the UK at the moment. This also links into the scrub issue, which has been highlighted particularly for Northern Ireland.

Revisit satellite processing - Since the satellite data processing was conducted for this project, the results of a significant, large-scale reprocessing of the Landsat data archive have been published, as Landsat Collection 2. This re-processing of the Landsat archive is intended to improve the quality and consistency of the satellite data over time and to improve on Collection 1, which was used in this work (Masek *et al.*, 2020). A particular focus of the Collection 2 improvements has been on the spatial

accuracy of the data (Storey *et al.*, 2019). This improved spatial accuracy should help improve the quality of the resultant classifications, particularly in the 1990's, which in turn may then help reduce some of the false positives currently affecting the woodland and urban change detection.

Iterative construction of the core training data set – Further work should review the core training areas and their classification over time. This might aid identification of poor training areas i.e. where a training area is set as a particular class based on LCM1990 and LCM2015, but is not classified as that class for one or more LCM's between 1994-2010. This check may be particularly relevant for arable polygons, where the assumption is that a polygon that is arable in LCM1990 and LCM2015 is arable for all years in between. This is unlikely to be true for all the arable polygons in the core training data set and their identification and removal is likely to improve the quality of the final classifications.

Integration with WPA – In addition, it would be beneficial to identify which uncertainties should be resolved in this database and which are best tackled in the Bayesian methods developed in WPA, especially as the Bayesian implementation already includes some methods for dealing with improbable land cover dynamics.

Northern Ireland – some additional work is required to review classifications and to develop an accurate core set of training data for Northern Ireland. The accuracy of the Northern Ireland data sets is currently still affected by the high levels of cloud and the impact that these have had on the quality of historic classifications for Northern Ireland, as well as some specific issues caused by the complexity and size of fields in Northern Ireland.

IACS validation - The IACS data set is a complex data set. For this analysis, the same random set of points was used for each of the time periods. Future work should explore in more detail the differences between the two data sets to get a better understanding of what change is being well detected, well change is being poorly detected and where the difference is due to differences in the description of land cover classes and/or the aggregation of classes. Part of this may involve a stratified random sample of points that better samples the change in the IACS data set over time.

LCM spatial framework – some of the false positives in the change detection are due in part to the mixed polygons, where multiple land cover types are found within a single polygon, especially for the urban and woodland change. It may be calculating additional attributes for each polygons will provide a systematic solution for identifying these polygons. However, it may also be that the easiest solution is to explore the potential for remapping these polygon boundaries or recreating the spatial framework.

LCM1984/85 – the time-series filtering was largely conducted with an assumption that the first and last classifications were correct, which places a greater pressure on these classifications. Given this, there would be potential benefits of producing a new LCM for 1984/85 which should be considered.

Initially, the assumption was that having multiple classifications would be beneficial, as it potentially allows poor classification results to be filtered out of the time-series. But having many classifications can be problematic, as small changes in classification sensitivity/errors will wrongly appear to cause change. This is a particular problem with polygons that have mixed land cover. Further work will determine to what extent it is possible to improve the methods for filtering and correcting these systematic errors. It is not yet clear how data from different satellite sensors affects the change data.

Accuracy of change and length of time between LCMs - The LCM's based on a single year of satellite data (LCM2017 onwards) have slightly lower accuracies than the LCM data sets produced in this project. This will affect the ability to detect change between recent years, as demonstrated by the comparison to IACS data. Further work will be required to assess whether the lower accuracy of LCMs 2017 onwards is due to methodological improvements developed and applied in this project, or whether it is due to the different input data, specifically using 2 years of Landsat data compared to one year of Sentinel-2 data. Or whether it is due to using a calendar year (as used in LCM2017 onwards) or a year that is more closely related to the cropping cycle (November-October).

A second issue is about the ideal frequency of LCM production for change mapping. For most classes the rate of change is slow and so for short time periods e.g. a year, the rate of change maybe less than the rate of classification error. The post-classification temporal filtering potentially offers a way for strengthening the change signal, but relies heavily on the first and last LCM's in the time-series, and the method also needs further careful development. Generally, longer time-periods are likely to offer higher levels of change, so that the classification error is less of an issue.

5 WP-B.4 – Method inter-comparison

This task will compare, and combine, the methods and results from WP-A and WP-B to identify the best source of activity data for the inventory. This activity will take place during August and September 2021 and is reported separately (See WP-A Final Report in this document).

6 Quality Assurance and Data Management

6.1 Data management

The UKCEH Data Management Plan (DMP), which was set up during WPA has been reviewed and updated for WPB. Data documentation and storage requirements were discussed with project partners in FR & AFBI, including a reminder of the importance of clear and comprehensive documentation of datasets produced for the project. Datasets have been received by UKCEH and will be reviewed by the project Data Manager in advance of task 4.

6.2 Model QA

Forest Research

Forest Research have completed a BEIS model QA log for their 1990 Woodland Map Code. Guidance and examples were provided by the QA Manager Gwen Buys and completion of the log was coordinated by Vera Correia at FR with contributions from Liz Clark, Eve McAleer, Anthony Walker and Esther Whitton. An initial draft of the log was completed in July 2021, feedback was provided by Gwen Buys and the log was revised and finalised in September 2021. The log is accompanied by supplementary information including a model report and methodology documents. The model achieves a final score of 85% with scores of 95% for structure & clarity and verification.

The methodology used by FR to produce the 2010, 2015 & 2020 woodland maps does not require a model QA log but is documented in a methodology report including a QA section. The process described is standardised and requires extensive user training and cross-checking by senior staff. Limitations are also described allowing for variation if confidence between different aspects of the methodology and results.

AFBI

AFBI have completed a BEIS model QA log for a short piece of code (~125 lines including comments) used to generate cloudfree composite images for Northern Ireland from Landsat 7 & 8. The model scores 72% overall, with scores of 84-90% for documentation, structure and clarity.

Additionally, AFBI have prepared a short document describing manual tasks carried out to complete the 2020 Woodland Basemap for Northern Ireland. The majority of these tasks are to remove spurious data which is outside of the forest definition.

UKCEH

The UKCEH land use mapping team are familiar with the BEIS model QA log, having completed one for a previous project undertaken in 2017 (Applying Earth Observation to assess UK Land use change: Lot 2 Medium Resolution Optical data). Clare Rowland has updated the model QA log to reflect the status of the model as used in this project. The model is a revised version of an existing freely available model 'Random Forest for Remote Sensing'. The original model is available on github and is described in a peer reviewed paper. The version used for LCM classification has a BEIS model QA log with associated scope/specification and assumption documents. The model scores 93.2% which is comfortably over the 90% threshold for BEIS business critical models.

6.3 Independent validation of satellite data classification

The independent validation of satellite data classification was carried out by the team at Ricardo. The main aim was to provide independent assurance of the quality of the land use classification (LUC) methodology that UK CEH use. As a result, the aim was split in the following objectives.

1. Obtain LCM 2017 raster layer (<https://www.ceh.ac.uk/services/lcm2019-lcm2018-and-lcm2017>)
2. Develop a methodology that utilises the Sentinel-2 multispectral bands to create land cover maps
3. Perform visual checks at selected study areas to validate the LCM 2017 by UK CEH (NB. The agreed validation involved the classification change, between 2017 and 2020, however the LCM 2020 is still under development and Ricardo were not able to undertake this analysis)

The methodology (for objective 2) was based on the development of a python script (mainly [ArcPy](#)) that also makes use of 3rd party libraries such (SentinelSat) to query the Copernicus API and download the Sentinel-2 imagery as well as to adopt utilise certain spatial (GIS) techniques to create the land cover maps. The spectral index adopted was the Normalised Differences Vegetation Index (NDVI) and the study areas involved were Lincolnshire (England), North Wales and North Scotland (see figure 11 for the map).

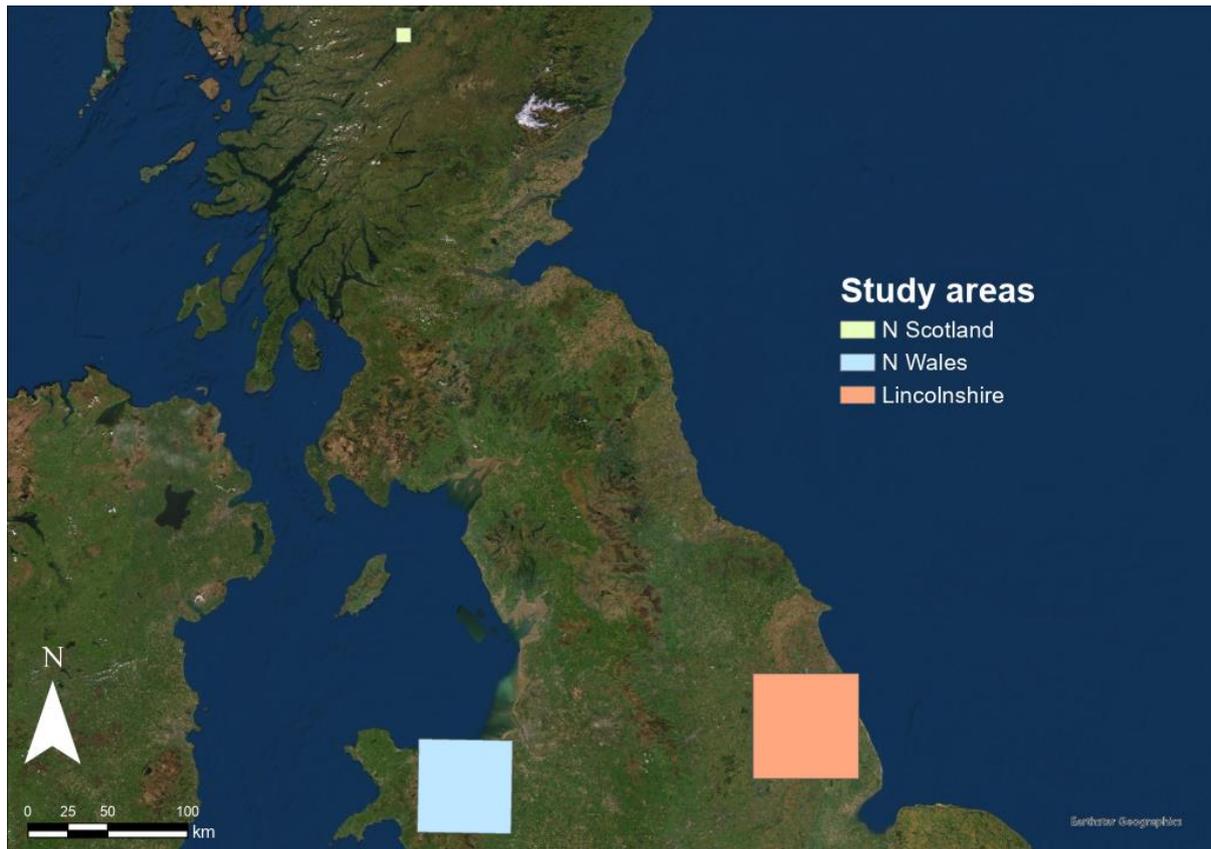


Figure 11. The three study areas used for the validation of LCM 2017

The datasets used have been obtained via [SentinelSat](#) (python library) using the study areas' vector files (converted to geojson) as the footprint for the data to be downloaded. The rules, when querying, were:

- Cloud coverage < 15%
- Sentinel-2 Level 2A Data (Bottom of Atmosphere reflectance images)
- Sensing period (1st March – 10th October)

The reasons for these criteria are based on the fact the clouds (water droplets) interfere with the satellite images - the MSI (Multispectral Instrument) cannot penetrate clouds hence the resultant images are not representative of the ground condition (i.e., classification). Further to that the cloud coverage percentage was set to 15% to allow for satellite images to be used as in the UK it is very unlikely to obtain a Sentinel-2 image that's 100 % cloud free. For North Scotland a higher percentage was used (50 %) to classify the cloud-free pixels as clouds are more common in the mountainous areas of Scotland. It is important to mention that pixels that may have been obstructed by clouds have been omitted from the analysis (via visual checks of the RGB image of that capture). Bottom of Atmosphere images were chosen because the Top of Atmosphere Sentinel-2 products were not available prior to March 2018 and the specific sensing period was chosen to obtain 2 cloud-free images that have been captured 6 months apart ensuring seasonal changes are being represented in the annual composite.

The study areas as previously seen in figure 11 have been chosen to analyse a diverse range of land cover classes, assess and incorporate any seasonal changes (i.e., arable land) and for workflow

efficiency by ensuring that these areas are fully within the boundaries of the corresponding Sentinel-2 tiles (satellite images). Specifically, the area in Lincolnshire is dominated by arable land and grassland as well as smaller regions of buildings and water bodies (river) and the validation exercise aimed to distinguish these via NDVI. The reason for choosing the study area in North Wales was mainly to examine coastal changes and the area in Scotland to identify any changes in the water bodies (i.e., lakes) and rough land (LULUCF ID 4).

6.3.1 Validating through NDVI

Objective 3 of the validation exercise was based on utilising the multispectral bands of Sentinel-2 and the creation of annual NDVI composites using bands 4 and 8 – specifically $NDVI = (NIR - VIS) / (NIR + VIS)$, where NIR is obtained from band 8 and VIS from band 4. NDVI was used as it relies on the fact that spongy and healthy vegetation reflect a lot of light in the near-infrared (NIR) spectrum, as opposed to most non-plant objects. Therefore, NDVI can help to highlight vegetation from other land features, and even help differentiate healthy vegetation from unhealthy vegetation (Evangelides and Nobajas, 2019). The classification was based on the fact low values if NDVI (< 0.1) present rough land (e.g., rock, stone, sand). Moderate NDVI values (0.2-0.3) correspond to bush and meadow vegetation whereas high index values (0.6–0.8) show presence of healthy forests or vegetation with high chlorophyll content.

The example in Lincolnshire is presented below (figure 12) where the NDVI composite of 2017 was overlaid on the LCM 2017 (obtained from CEH).

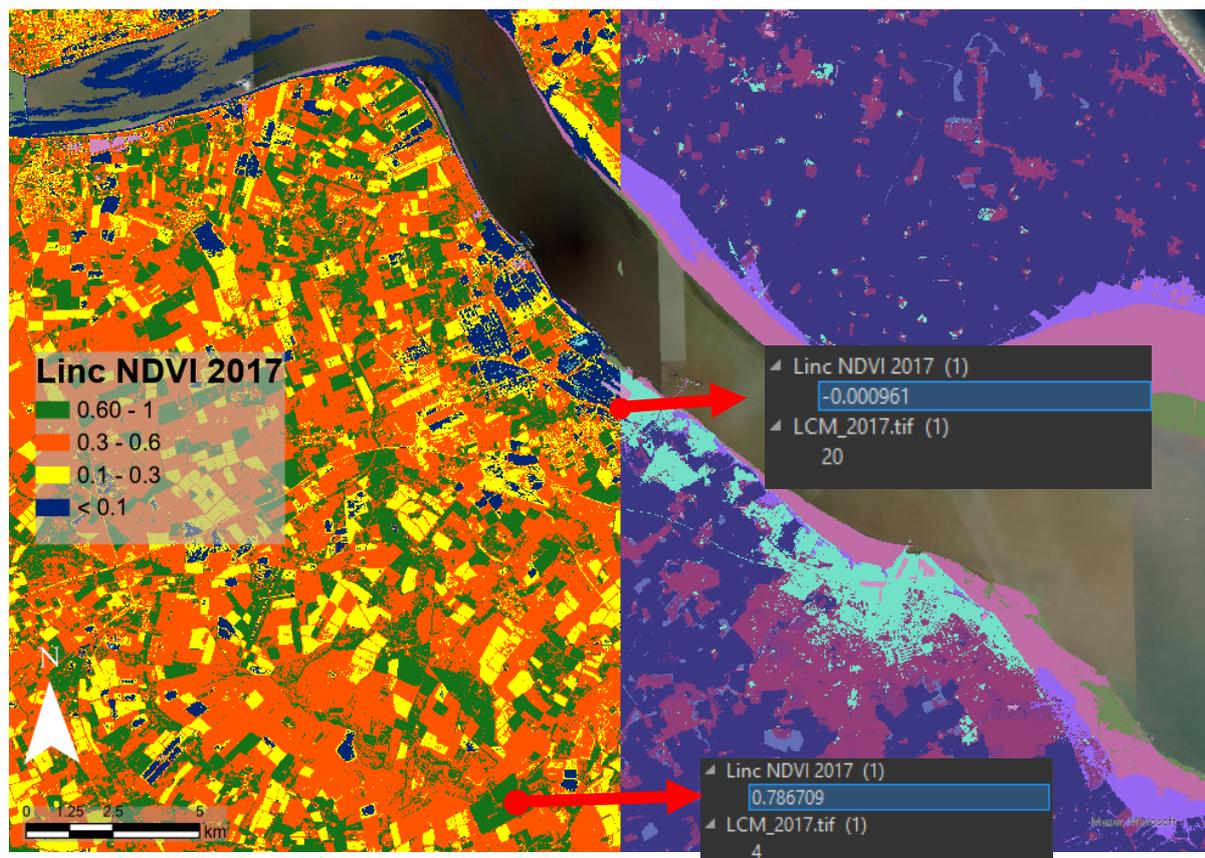


Figure 12: The created NDVI 2017 composite for validation (L) and the LCM 2017 from UK CEH (R)

The spot checks for this area of interest involved two pixel-classification validation checks – one at a suspected urban site and one at a low healthy vegetation region. The one for the urban had an NDVI value below 0.1 and a classified LCM value of 20 corresponding to the ‘Urban’ LCM class. This pixel UKCEH report ... version 1.0

along with the surrounding area in blue (left) and the cyan (right) correspond to the Urban LCM and LULUCF class and as a result validate the classification by UK CEH. The second pixel involved a moderately healthy vegetation site presented in green (left). The pixel for this site had a value between 0.6 and 1 corresponding to the healthier vegetation and an LCM class of 4 (improved grassland) resulting to the 'grass' LULUCF category. The same approach was carried out for the study area in North Scotland (only one satellite image was used for Scotland due to cloud interference). See figure 13 for the study area in North Scotland.

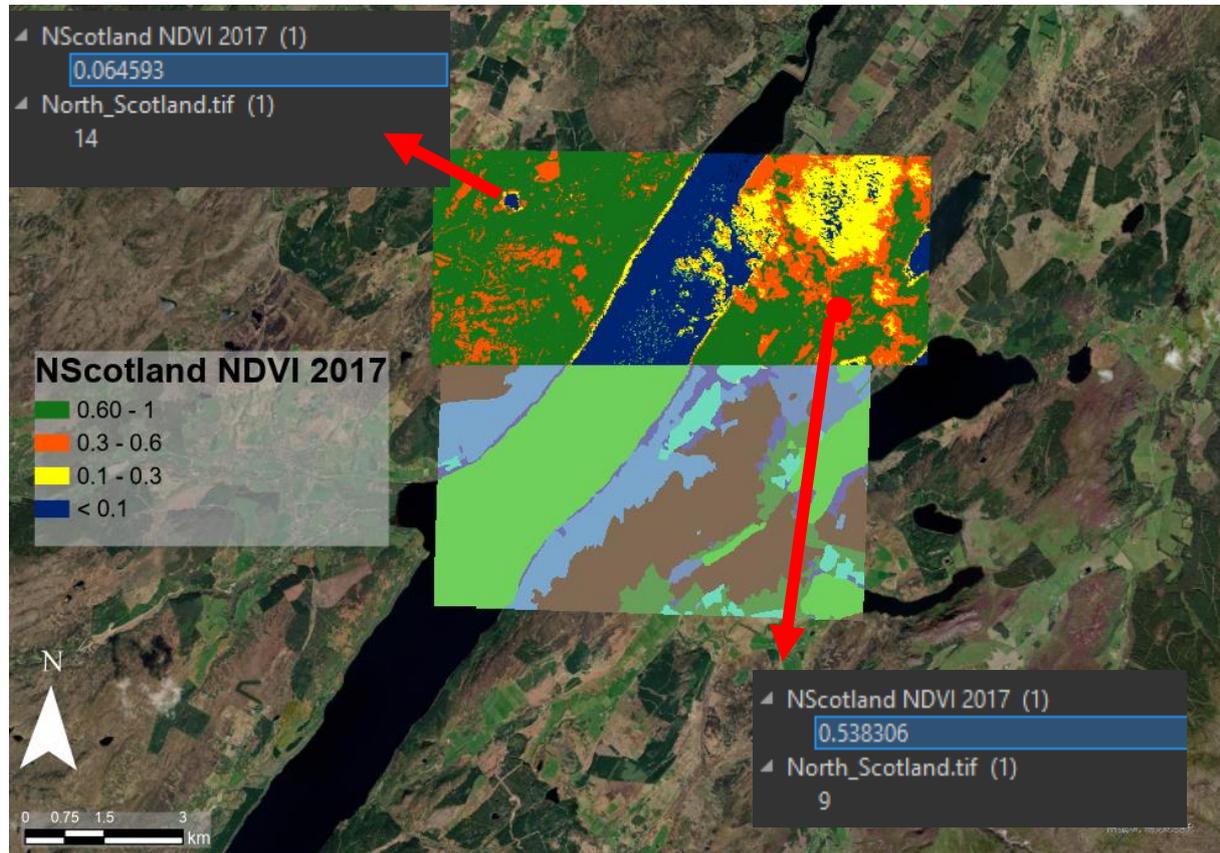


Figure 13: Study area in North Scotland (top map presents the NDVI 2017) and the bottom map the LCM 2017 classification (pixels in yellow present clouds therefore were excluded from the comparison with the LCM product).

The results from this study area also verify the UK CEH LCM classification as the low NDVI (values < 0.1) correspond to water bodies (areas in blue) something that is aligned with the LCM 2017 with the pixels classified as 14 (corresponding to 'Freshwater' and 'Water' in the LULUCF names). The pixel with an NDVI value of 0.538306 that can be interpreted as moderately healthy, corresponds to heather. As a result, these exercises could validate the UK CEH LCM 2017 classification. It is important to mention here that for the study area in North Wales, the SentinelSat library as well as the interactive [Copernicus Open Hub](https://copernicus.openhub.com/) could not retrieve the relevant satellite images for 2017. Although satellite images older than a year are archived and the downloading requires extra time (usually 30' – 1hr), this issue for North Wales' data was not presented for North Scotland and Lincolnshire. The exact reason for this remains unknown despite more recent attempts.

6.3.2 Limitations and future recommendations

From the visual analysis of this exercise the LCM product 2017 could be verified. However, the comparisons of the NDVI and LCM 2017 were not performed at a national scale – the main reason being the time constrain for both the spot checks to be performed nationally and acquiring data for UKCEH report ... version 1.0

the whole UK (to ensure that images used were captured at the same instance). As mentioned before, cloud interference was another limiting factor, especially in the case of North Scotland as only 1 image could be obtained from 2017 that was near cloud-free. This as a result, limited the reliability of this validation as seasonality was not taken into account.

It is important to mention since UK CEH utilised other sources of data and information (in addition to Sentinel-2) as well as other processing techniques, a direct methodological comparison was not made. This validation utilised one source of satellite imagery that is well-documented, in the literature, to be able to provide a free rapid land cover classification based on the reflectance of the spectral bands of Sentinel-2. Future work could incorporate more sources (i.e., the upcoming Landsat-9 mission). Additionally, more satellites images could be used (instead of 2 instances) to improve the reliability and quality when creating the annual composites. More spectral indices could also be used where occasional events such as fires take place. The spectral indices that could improve the mapping component and identify the spatial extent are the Normalised Burnt Ratio $((B8A-B12)/(B8A+B12))$ and Fire Detection Index $(B12/(B8A*B09))$. Future validation steps could also involve checks to determine if the areas are associated with bias.

6.4 Comparison to methods used by other countries

This section provides a selected review of the land use change (LUC) tracking and quality assurance/quality control (QA/QC) procedures that non-UK inventory compliers apply in their LULUCF inventories. We have not attempted a full review of QA/QC procedures in the LULUCF sector of the inventories considered, or, a full review of the procedures used to identify and track all land use activities because to do so would have been beyond the resources available.

Using expert judgement, we have selected two main countries to focus on: Canada, and Australia. In addition, a short commentary has been added about France. These countries have been selected because they have highly developed, mature LULUCF inventories and our expectation before the work started was that the extent and quality of the documentation would be high.

Our analysis from Canada, and Australia is presented first. Following this, there is a short section considering relevant messages from a bi-lateral exchange with France held in London, August 2018. We have extracted key points about LUC tracking and QA/QC from this bilateral exchange, but, have not reviewed the latest French NIR.

The analysis of the methodologies used by Canada and Australia are mainly split into those used in the forest and non-forest sectors. There is some but not an excessive overlap in the analysis between these two sectors.

Several information resources and contacts were identified at the start of this project. Subsequently, we found some of the references to reports and studies were already mentioned in the NIRs, or, were judged not be directly useful for the study. Rather than exclude the information completely from the project report, for completeness, we have included them in Appendix 2. To keep the context of the information in e-mail correspondence, we have included the full e-mails rather than just summarise the information in the e-mails. Perhaps in future studies this information and the contacts could prove useful.

After each major section, we have provided a short section which summarises our assessment of the information reviewed and provides the key points from our analysis. In each of the sections, we have included material from the NIRs and associated reports. This has been done for completeness and to allow our analysis to be directly and quickly cross checked with the source material. Points that we feel are important in the extracts of the text are highlighted in **blue shading**.

6.4.1 Canada

Canada has a methodologically well-developed LULUCF inventory, and a comprehensive National Inventory Report. The NIR is issued in three parts. Information in the NIR relevant to this study includes:

- **Part 1 of the NIR (NIR P1):** Chapters 3 to 7 provide descriptions and additional analysis for each sector, according to UNFCCC reporting requirements. Chapter 8 presents a summary of recalculations and planned improvements.
- **Part 2 of the NIR (NIR P2):** Annexes 1 to 7, provide a key category analysis, an inventory uncertainty assessment, detailed explanations of estimation methodologies.

Part 3 of NIR is not relevant to this study. It presents rounding procedures, summary tables of GHG emissions at the national level and for each provincial and territorial jurisdiction, sector and gas.

This analysis is separated into the sections covering the LUC tracking approaches used in the forest land and non-forest land sectors. As a general point, there is much more methodological detail provided both in the NIR and in supplementary reports regarding the approaches used to estimate the areas of forest land and the transitions between forest land and other land uses.

12.5.1.1 Forest sector

Information presented in the NIR

The method that Canada uses to estimate its GHG emissions and removals in managed forests is presented in Section 6.3.1.2 of Part 1 of the NIR "*Methodological Issues*". Canada applies a Tier 3 methodology for estimating GHG emissions and removals in managed forests. Canada's National Forest Carbon Monitoring, Accounting and Reporting System (NFCMARS) includes a model-based approach (Carbon Budget Model of the Canadian Forest sector, or CBM-CFS3). The origins of this model can be traced back to the late 1980s.

This model integrates forest inventory data and yield curves with spatially referenced activity data on forest management and natural disturbances to estimate forest carbon (C) stocks, stock changes and CO₂ emissions and removals.

The model uses regional ecological and climate parameters to simulate C transfers among pools, to harvested wood products and to the atmosphere. A more detailed description of forest C modelling is presented in Annex 3.5.2.1. "*Carbon Modelling*".

We have looked through the technical details provided in Part 2 of the NIR to identify information that is relevant to land use tracking. Section A3.5.2.5 "Forest Conversion" provides relevant information. The section on "*Quality Assurance / Quality Control of Forest Conversion Data*" indicates that the arrangements for the using the remote sensing analysis in the inventory is complex, and the data is sources from a wide variety of organisations. From the text in the NIR, we conclude that the database itself is used as a mechanism to record decisions about which data are used in the inventory, issues with the data, and solutions to overcoming the issues. There are advantages to this approach, as the data and the documentation about the use of the data are saved together and hence kept directly connected.

The NIR P2 states field validation is conducted on an ongoing basis as resources permit.

The NIR P2 refers to a separate publication which documents the GIS system database of "conversion events". Although not completely clear from the explanation in the NIR, we interpret these events to be land use conversion events.

"The remote sensing interpretation follows defined procedures (Leckie et al., 2010b; Dyk et al., 2015), although it is conducted by a variety of organizations, including provincial government forestry or geomatics groups, remote sensing or mapping companies, research and development organizations and in-house government staff. The basic image analysis quality control (QC) process includes: internal checks within the mapping agency or company by a senior person; real-time quality assurance (QA) by Canadian Forest Service specialists during interpretation, with

feedback provided within days of interpretation of an area; and a final QA and vetting of the interpretation by the Canadian Forest Service. Field validation is conducted on an ongoing basis as resources permit. Each QC point and revision is documented within the Geographic Information System (GIS) database of conversion events (Dyk et al., 2015).

Records of decision as to data used and expert judgement applied, as well as decisions on the resolution of contradictory data, are documented within the overall processing database (Leckie, 2006b) and updated for each new submission (Dyk et al., 2015). Data sources and limitations are recorded, and remote sensing data and interpretations archived.”

The description indicates that both quality control and quality assurance activities are conducted. This procedure of applying both QC and QA measures is consistent with IPCC good practice.

We have examined the publication referenced, Dyk et al., 2015, and this is discussed further below. It is worth noting that this publication was easily located in a web search. Because this publication was easily found, this helps improve transparency of the methodology used.

Later in section A3.5.2.5, details are provided about methodological approaches to tracking one type of land use change: Land Converted to Forest Land.

“Records of land conversion to forest land in Canada were available for 1990–2002 from the Feasibility Assessment of Afforestation for Carbon Sequestration (FAACS) initiative (White and Kurz 2005). Conversion activities for 1970–1989 and 2003–2008 were estimated based on activity rates observed in the FAACS data. Additional information from the Forest 2020 Plantation Demonstration Assessment was included for 2004 and 2005, and an environmental scan was performed to identify additional sources of information on afforestation rates from 2000–2008. Each event, regardless of date, source, type or location, was converted to an inventory record for the purposes of C modelling. All events were compiled in a single data set of afforestation activity in Canada from 1970 to 2008. No new afforestation activity data were identified for 2009 to the current inventory year. Efforts continue to obtain additional data on recent afforestation activities in Canada.

For 1990–2008, the area planted was stratified by ecozone, province and tree species. Total area planted by province and ecozone, in conjunction with the proportion of species planted for each province, was used to calculate area planted by species, resulting in estimates of the area converted to forest, by species, for each RU.”

From this description, it is not explicitly clear if remote sensing has played any role in the assessment of land use change in the forestry sector. However, the report by Dyk et al. (2015) clearly states the role of remote sensing in the tracking of deforestation events.

Assessment:

Assessment: The NIR 2020 from Canada provides a reasonable summary of the work Canada does to track land use change in the LULUCF inventory. It cross references to a much more detailed report that explains the methodology that is used to track one aspect of land use change – deforestation. Methods in supporting documents could be found easily on the internet improving the transparency of the reporting.

Information presented in the description of Canada’s National Deforestation Monitoring System (Dyk et al., 2015)

This report describes the methodology that is used to track the deforestation in Canada’s forest estate. We have examined both the image source data and data handling approaches – but we have not examined the methodology to estimate carbon fluxes in detail.

The important elements to note are:

- Canada’s National Deforestation Monitoring System (NDMS) was designed and implemented to provide information needed by Canada to meet its obligation under the United Nations Framework Convention on Climate Change (UNFCCC) to report the areas affected annually by deforestation.
- The NDMS uses deforestation mapped on a system of sample areas.
- The mapping is based on visual interpretation of satellite imagery supported by available ancillary information, such as high-resolution imagery, forest inventory, and industrial databases, and informed by records-based information and expert knowledge.

- Accurate detection and mapping of deforestation events involves manual interpretation of satellite remote sensing imagery by specialized analysts.
- A key factor in the mapping is to distinguish deforestation from other forest cover losses that occur in Canada.
- The report describes the methods used to monitor and report on deforestation, including stratification and sampling, data sources, mapping and interpretation, estimation procedures, and quality control through all stages of the process.

Data sources

Section 4.2 “Data Sources and Information Collection” states that remote sensing imagery is the most critical input to the process of deforestation monitoring. In the NDMS, image data are used to identify the location, size, and timing of events. These data are also used to determine pre-existing forest types and subsequent post classes. The NDMS employs medium-resolution satellite imagery in combination with high-resolution satellite and air photo data wherever available. Ancillary sources, in the form of spatial data and non-spatial records, are used to support the mapping process and act as supplementary sources for estimation and quality control.

A range of image source data are used to track deforestation:

- **Landsat.** The core image data source is from Landsat. Landsat series satellites have been acquiring data since the early 1970s and provide a consistent data stream for all NDMS mapping periods with a spatial and spectral detail suitable for deforestation mapping. The current Landsat instruments are medium-resolution sensors with 30 × 30 m pixels and a large image footprint (170 × 185 km). Each satellite has a return period of 16 days, with continual image acquisition and scene sidelay that increases at higher latitudes.

Higher-resolution data are also very useful in supporting the interpretation of pre-existing conditions and forest types, and in confirming deforestation and better identifying the post class and post-class modifier. High-resolution imagery is therefore used to supplement core Landsat data wherever possible. Additional image data are sourced from:

- **SPOT 4/5 satellites.** The SPOT 4/5 satellites offer a higher spatial resolution but provide fewer spectral bands than Landsat. A 2.5 m resolution SPOT 5 panchromatic mosaic of the Prairies (circa 2006) was provided by Agriculture and Agri-Foods Canada, and a 10 m SPOT 4/5 national coverage called GeoBase1 was obtained from the Centre for Topographic Information of Natural Resources Canada.
- **Google Earth™ and Bing Maps™** are a valuable resource with an ever-expanding selection of recent high-resolution imagery over population centres and many other regions of the country. This includes terrain display information and crowd-sourced Panoramio™ photographs.

Ancillary data are also incorporated to support the interpretation process. These data may be used to validate or confirm deforestation occurrence or to support the interpretation of prior forest type or post-change land use. Some data types are available nationally, whereas others are present only within limited extents. Some data sets are directly relevant LUC activities in the LULUCF sector, such as: wooded areas from provincial base maps, wetlands, and forest inventories, forest management tenure areas. Other data sets are either indirectly relevant to LUC activities in the LULUCF sector, or are relevant to other inventory sectors: GIS data sets of road networks, hydrology, pit and quarry licence areas, specialized oil and gas pipeline and well pad databases, field oblique photos, and “other data”.

This report focusses on the methodology and data that is needed to track the effects of deforestation. From the text, it is not clear whether these image source data are also used to track other land use changes apart from deforestation in the LULUCF inventory, but we suspect they will be.

Data handling and processing

Section 4.3 describes the approaches that have been used to map the deforestation events, and the data handling procedures, in section 4.3.1.

“Deforestation event mapping is done using core date imagery and multi-band image products. Mapping is broken up into units of mapping work, called “project packages,” to be completed by internal staff, provincial partners, or qualified private sector contractors. All relevant images and supporting data are assembled as a package for mapping in a particular area. Detection, interpretation, digitization, and attribution of individual deforestation events is done manually by trained interpreters according to the established NDMS methods in a standardized ArcGIS geodatabase (Leckie et al. 2012).”

For the mapping approaches, there are two important aspects to note: 1) a range of staff types are used for the mapping work, including internal staff in the Deforestation Monitoring Group, partners, and private sector contractors; 2) the work is done manually. There is no information on the time taken to map the deforestation events, but it is likely to be time consuming.

A figure is provided in the report (labelled as Figure 7 in the source but Figure 14 in this report), to illustrate how individual deforestation events are manually interpreted and identified, and delineated. This shows how core data imagery from the Landsat satellites are used. The caption to the figure explains how the Landsat satellite band combinations and change enhancements that may be used in the mapping process. So called “red triggers” are identified where vegetation loss or change has occurred. This “change enhancement” is shown in the far-right hand column of images, in the figure.

The figure and its caption are reproduced immediately below.

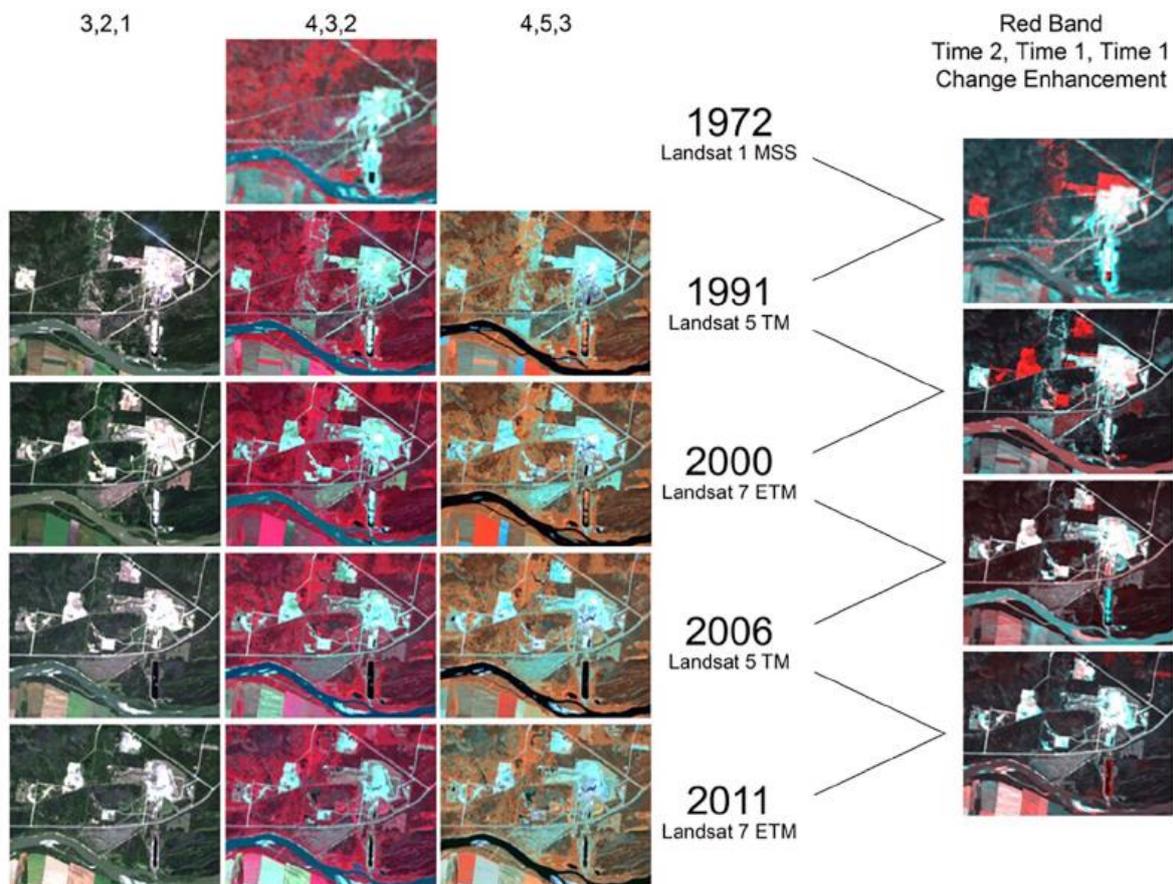


Figure 14. Examples of various Landsat satellite band combinations and change enhancements that may be used in the mapping process. Note the change enhancement at right shows red triggers where vegetation loss or change has occurred. Band

combinations shown are normal colour rendition on left, with colour infrared in the middle. The right-hand column shows Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper (ETM) bands 4,5,3 (i.e., the two near infrared bands and a red band) displayed as red, green, blue.

In addition to the core Landsat data, supplementary data from the SPOT 4/5 satellites and aerial photography are used. Figure 15 illustrates how these supplementary data are used. Of particular note is the use of these data to verify deforestation activities.

This figure and its caption are reproduced immediately below.

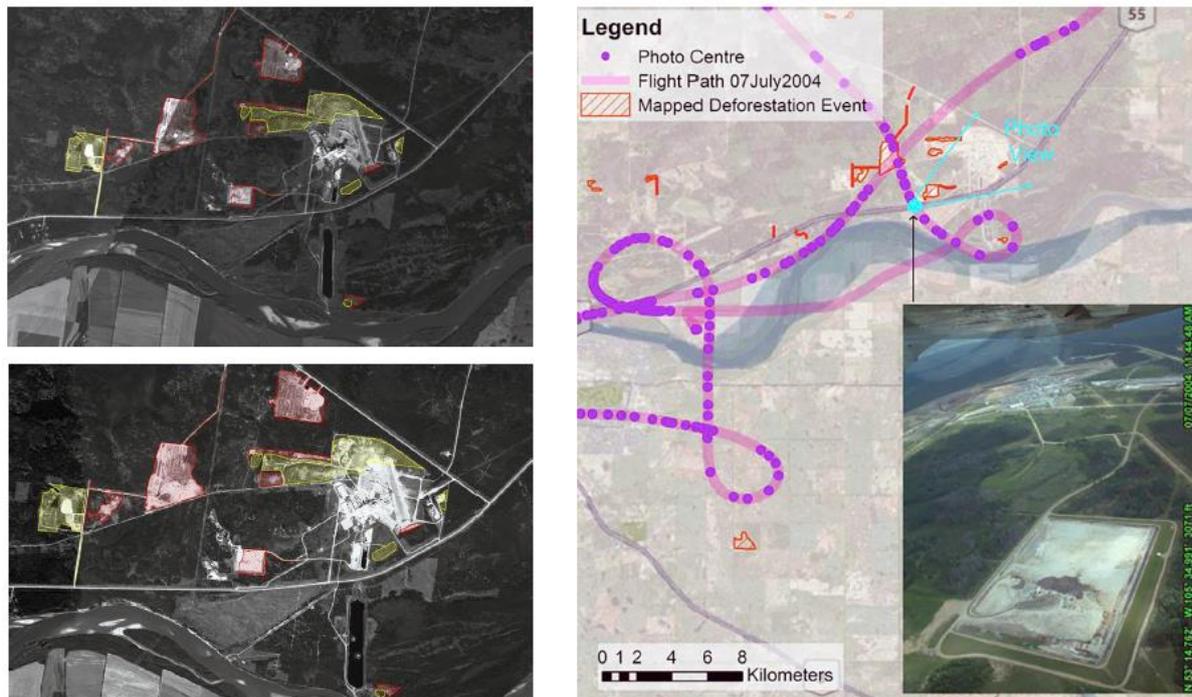


Figure 15. Examples of supplemental data use. Left: Example of confirmed and interpreted events (red and yellow lines) shown as polygons on the SPOT imagery (2.5 m resolution SPOT 5, top left, and 10 m SPOT 4, bottom left). Right: Example of flight path and oblique aerial photograph from a deforestation field verification campaign.

Data interpretation in a GIS system

Typically, the interpretation process is conducted in a GIS environment, where change enhancement and single-date images can be layered and compared. The report refers to two specific manual techniques to interpret the image source data: blinking, and the swipe tool. Both these comparative approaches help visually identify where changes have occurred in image data.

- **Using the Flicker tool** (or “blinking”). If there are 2 layers of data, this tool switches the view between the 2 layers alternately – or “blinks” the layers.
- **Using the Swipe tool.** If there are 2 layers of data, this tool allows the user to swipe horizontally or vertically to reveal the lower layer beneath the top layer.

Georeferenced ancillary imagery can be simultaneously displayed, and a link between ArcGIS and Google Earth allows the synchronization of spatial display extents in both programs.

Section 4.3.3 of the report, “Interpretation”, explains in more detail how the image data are interpreted. Image interpretation is done manually on a cell-by-cell basis. Typically, the change UKCEH report ... version 1.0

enhancement is used first to identify change events. These are then confirmed or refuted as deforestation by the interpreter using the individual core date images supported by other image and ancillary data. The supporting data are used to assess whether an event area was forested before the change, and whether the event meets the deforestation definition. Supporting data are also used to help interpret pretype and post class.

The boundary of a deforestation event is identified in the Landsat images, but higher-resolution imagery is used in a supporting role to improve event boundary delineation.

Maintaining consistency of methodological approach when using a range of image data interpreters

To maintain consistency of interpretation between different people and teams, interpreters follow specific digitizing guidelines provided in the Deforestation Interpretation Guide (Leckie et al. 2012). These guidelines dictate how events shall be mapped, e.g., digitize smooth lines rather than following pixel boundaries, avoid gaps and overlaps between adjacent events, keep the mapping topology clean, etc.

In some cases, there may be some uncertainty about the land use change event. The National Deforestation Monitoring System has a procedure to record this uncertainty. Interpreters record their level of confidence about each deforestation event, i.e., how confident they are that it was forest before the change and that it is indeed a new land use. Comment fields are used extensively to record details specific to each event. These may indicate uncertainty in an event, a request by the interpreter for specific review by the quality control person, or the need for checking the event in a later time period.

Identifying spurious results and outliers

As part of the quality control procedure, there is a step after the mapping and GIS processing steps have been completed to identify spurious results and which we also assume would include the identification of identify outliers.

After the mapping and GIS processing steps have been completed, the initial estimation results are analyzed to identify potential problems. With the nature of deforestation in Canada and intensity of the sample, occasionally spurious estimates may occur.

Section 4.4.3 “Expert Intervention” explains the procedure used. A key observation is that expert knowledge is essential in this part of the methodology. It does not seem that an automated algorithm is used to identify spurious results. Rather, expert judgement is used.

“Canadian Forest Service deforestation monitoring experts have accumulated familiarity with land use change patterns and activities in Canada over many years. These experts scrutinize NDMS estimation process outputs and investigate questionable outputs that may be then modified if a change is justified.

When suspicious estimation outputs are encountered, experts intervene by evaluating available records information and reconnaissance data. For example, when a particular event type is rare and sampling produces estimates that are unexpectedly high or low, expert reviewers may scan available imagery to see what is occurring outside the sample cell boundaries (i.e., in the non-sampled area), and use this to inform expert intervention. Estimation outputs may be deemed suspicious when they contradict previous estimates or other available lines of evidence. All expert interventions are documented for transparency, and in case they need to be reassessed as new evidence emerges.”

Quality control

Section 5 of the report summarises the procedures that are used to control the quality of the work. Some of the quality steps – such as the identification of spurious results – are discussed above. The QC procedure appears thorough, and is a multi-step approach.

The QC steps are split into 3 main steps:

1. initial mapping quality control;

2. revision quality checks; and
3. overall quality assurance or vetting, where a check on the final mapping data is performed.

Strictly, we interpret step 3, the “overall quality assurance or vetting” step, as a quality control step, rather than a quality assurance step. A quality assurance step would involve the use of personnel who are outside the Canada’s National Deforestation Monitoring System.

After the data are calculated in the estimate process, the final numbers and trends are assessed for changes from previous year estimates. Deforestation estimates are also scrutinized by external users before they incorporate the data into their applications.

The report makes a cross reference to an old NIR of Canada, the 2011 NIR, for more details about the quality control approaches in the estimation of the deforestation area estimation.

The report has 3 pages of detail about the QC system used, and we have provided a summary of this here. To note the Deforestation Monitoring Group’s use of documentation of the more complex quality control procedures from the “vetting interpreters”, rather than using the mapping geodatabase, and some automated QC checks at the data processing stage. In the report, the team refer to an external quality control step, which qualifies as a quality assurance step.

- **Quality Control at the Mapping Stage.**

Typical errors caught through quality control at the mapping stage include: incorrect event interpretations or delineations; event omissions or commissions; missing field values; typographical errors; and topology problems. During the quality control phase, an internal Canadian Forest Service reviewer performs detailed checks to ensure consistency with mapping and interpretation rules, and returns a set of points to flag and describe the problem areas to the interpreter or contractor. These quality control points are stored in the mapping geodatabase as a permanent layer. Throughout the entire mapping and revision process, an internal “project story” is maintained by the Canadian Forest Service. This document summarizes activity on a given project, including staff and contractors involved, dates, and actions completed. It also includes detailed explanation of any complex situations or vetting (“vet”) revisions, as well as notes on any issues raised by the quality control or vetting interpreters. To maintain consistency, analysts who are involved in revisions or who vet revisions read this project story before starting work.

- **Quality Control at the Data Processing Stage**

Once revisions are vetted, final data integrity checks are conducted on the project database to find and correct any technical errors, such as missing attributes, mismatched post classes and modifiers, incorrect or reversed event dates, inappropriate line widths, topology errors, overlaps and duplicates, and typos. This stage is partially automated. The final mapping layers can be loaded into the national database only after the integrity checks have been completed. Data quality checks are built into the script that automates the data processing stages. The script both delivers reports to the screen while running, and dumps to log files, which are reviewed by Canadian Forest Service staff to ensure the runs were successful. Transfer of the final GIS tables into the non-spatial deforestation database is also automated. The grid and event tables for each time period are checked before loading, and the number of records loaded is recorded and compared with previous runs to highlight any potential discrepancies. The database queries and tables used to create the annual estimates by industrial category and RU are controlled by Microsoft Access macros to ensure consistency between the runs.

- **Quality Control at the Output Analysis Stage**

Following generation of the estimate tables, the main outputs (annualized deforestation by industrial class and RU, pretype table, and unmanaged forest table) are evaluated closely by Canadian Forest Service experts to ensure that estimates have been calculated properly and that discrepancies from previously calculated values can be explained. To aid in this effort, a comparison spreadsheet is created for each RU showing the deforestation estimates by industrial class for the current and previous year’s national run, together with illustrations and statistics. The experts scrutinize the data

for changes and ensure the new estimates are consistent with the input mapping data. Any specific issues and special adjustments required are documented in a report that becomes part of the record for each year's estimate.

- **External Quality Control**

Users of the deforestation estimates conduct their own reviews of the data prior to use in their systems, and communicate any issues before incorporating the data into their own business processes. When new issues arise, these are dealt with and the NDMS quality control processes are updated to catch similar issues in the future

Assessment:

Assessment: This report describes the methodology that is used to track the deforestation in Canada's forest estate. We have examined both the image source data and data handling approaches. It is a very thorough report, and for the most part, explains very clearly the procedures that are used. It was easy to find on the internet, which ensures transparency of Canada's climate reporting. The mapping is based on visual interpretation of satellite imagery supported by available ancillary information, such as high-resolution imagery, forest inventory, and industrial databases, and informed by records-based information and expert knowledge. The work relies heavily on manual image processing, and the use of experts both in the Deforestation Monitoring Group and other partners. To maintain consistency of interpretation between different people and teams, interpreters follow specific digitizing guidelines provided in a separate guide, the Deforestation Interpretation Guide. There is a comprehensive multi-step quality control procedure, with additional quality assurance. For some part of the QC process, expert knowledge is essential. Records of the QC findings are kept both within the mapping geodatabase, and where more detailed analysis is necessary, separate records are kept with the explanations of any complex situations or vetting ("vet") revisions. Some data quality checks are built into the script that automates the data processing stages. Any specific issues and special adjustments required are documented in a report that becomes part of the record for each year's estimate (of deforestation).

12.5.1.2 Non-forest sector

Information presented in the NIR

Part 1 of the NIR presents summaries of the methodologies used to estimate emissions from the non-forest sectors of the LULUCF sector. Part 2 of the NIR presents more detailed methodologies.

The sections below summarise the information presented in the NIR about estimating and tracking land use change in non-forest sectors. We have cited relevant material from the sections, and provided comments about the approaches used, and the QA/QC procedures applied.

Part 1 of NIR

Section 6.5. Cropland (CRF Category 4.B), pg. 157

Section 6.5.2. **Land Converted to Cropland** (CRF Category 4.B.2), pg. 162

This subcategory includes the conversion of Forest Land and Grassland to Cropland. Emissions from the conversion of Forest Land to Cropland account for more than 90% of the total annual emissions in this category, which decreased from 9.5 Mt in 1990 to 2.8 Mt in 2018. Emissions from the conversion of Grassland are relatively small.

Section 6.5.2.1. Forest Land Converted to Cropland (CRF Category 4.B.2.1), pg. 162

There is no comment in this section about the methods used to track land use conversions. There is a reference to Annex 3.5, which contains more detailed methodologies.

There is a statement about the QA/QC measures applied, with a reference to which organisation does the checks:

“Quality Assurance / Quality Control and Verification: This category has undergone Tier 1 QC checks (see section 1.3, Chapter 1) in a manner consistent with the 2006 IPCC Guidelines. Quality checks were also performed externally by Agriculture and Agri-Food Canada, which derived the estimates of SOC change. The activity data, methodologies and changes to methodologies are documented and archived in both paper and electronic form.”

Section 6.5.2.2. Grassland Converted to Cropland (CRF Category 4.B.2.2), pg. 163

There is no comment in this section about the methods used to track land use conversions.

There is a statement about the QA/QC measures applied; this is a cross-reference to a statement about the general QA/QC procedures that are applied in the inventory:

“Quality Assurance / Quality Control and Verification: This category has undergone Tier 1 QC checks (see section 1.3, Chapter 1) in a manner consistent with the 2006 IPCC Guidelines. The activity data, methodologies and changes to methodologies are documented and archived in both paper and electronic form.”

Section 6.6. Grassland (CRF Category 4.C), pg. 164

Only 6.6.1. **Grassland Remaining Grassland** (CRF Category 4.C.1), pg. 164, is discussed as a category. We suspect this means that there are no use land conversions from grassland to other categories – but there are land use conversions to grassland.

Section 6.7. Wetlands (CRF Category 4.D), pg. 165

Section 6.7.1. Peat Extraction (CRF Categories 4.D.1.1 and 4.D.2.1), pg. 165

Earth Observation (EO) was used to determine the extent of peatland areas converted for peat extraction:

“An EO mapping approach was used to determine the extent of peatland areas converted for peat extraction for 1990, 2007 and 2013 time periods and to identify the proportion of land category types converted (Forest Land and Other Land). Converted areas were allocated into four land management subcategories: active extraction, abandoned, rehabilitated and restored areas based on image interpretation and industry information. National peat production statistics were used to estimate the annual amount of extracted peat (NRCan 2018a). Emissions from peat extraction are reported under Land converted to Wetlands for the first year after conversion and under Wetlands remaining Wetlands thereafter. More information on estimation methodology can be found in Annex 3.5.”

Category specific QA/QC measures are applied, as well as more general QA/QC:

“Quality Assurance / Quality Control and Verification: Section 1.3 in Chapter 1 describes the general QA/QC procedures being implemented for Canada’s GHG inventory. The same procedures apply to this category as well. Industry and academic experts associated with the Canadian Sphagnum Peat Moss Association and Peatland Ecology Research Group provided QC, validation of mapping estimates and a review of domestically derived emission factors.”

Section 6.7.2. Flooded Lands (CRF Categories 4.D.1.2 and 4.D.2.2), pg. 167

There is a statement about the where the land conversion to flooded lands has occurred:

“Since 1970, land conversion to flooded lands occurred mainly in reporting zones 4, 5, 8, 10 and 14 (i.e. Taiga Shield East, Boreal Shield East, Hudson Plains, Boreal Plains and Montane Cordillera).”

“Canada’s approach to estimating emissions from forest flooding is more realistic temporally than the default approach (IPCC 2006), which assumes that all biomass C on flooded forests is immediately emitted. Canada’s method is more refined in that it distinguishes forest clearing and flooding; emissions from the former are estimated as in all forest clearing associated with land-use change. Further, in Canada’s approach, emissions from the surface of reservoirs are derived from measurements, rather than from an assumption (immediate decay of all submerged biomass) that clearly is not verified.”

There is a statement about the QA/QC measures applied; this is a cross-reference to a statement about the general QA/QC procedures that are applied in the inventory:

“Quality Assurance / Quality Control and Verification: Section 1.3 in Chapter 1 describes the general QA/QC procedures being implemented for Canada’s GHG inventory. The same procedures apply to this category as well. For Forest Land converted to Wetlands, also refer to the corresponding subheading in section 6.9, Forest Conversion.”

Section 6.8. Settlements (CRF Category 4.E), pg. 168

Section 6.8.2. Land Converted to Settlements (CRF Category 4.E.2), pg. 169

The NIR notes that there are activity data problem in this category:

“While there are potentially several land categories, including forests that have been converted to Settlements, there are currently insufficient data to quantify areas or associated emissions for all types of land-use change. Significant efforts were invested in quantifying the areas of Forest Land converted to Settlements, as this has been the leading forest conversion type since 2000. A consistent methodology was developed for all forest conversion and is outlined in section 6.9.”

Section 6.8.2.1. Cropland Converted to Settlements (CRF Category 4.E.2.2), pg. 170

Areas of Cropland converted to Settlements were estimated from land-use maps from 1990, 2000 and 2010 by Huffman et al. (2015a) using the methods described in Annex 3.5.

Category specific QA/QC is applied, which relies on manual visual interpretation. To note the use of “verification” of the boundaries of the polygons:

“Quality Assurance / Quality Control and Verification: Polygons from the 2011 census were used to define the boundary of each Census Metropolitan Area and Landsat imagery from the Global Land Surface products from ArcGIS online services was obtained for each area for 1990, 2000 and 2010. Over 200 points were used to verify land cover/land use change for each time period, using visual interpretation. The points were defined using stratified random sampling, 50% on areas of change from Cropland to Settlements and 50% on areas of no change, separated by a minimum distance of 1 km, to avoid statistical bias.”

Section 6.8.2.2. Grassland Converted to Settlements (CRF Category 4.E.2.3), pg. 170

The NIR notes that the large area of grassland means that tracking land use changes are difficult:

“An accurate estimation of this direct human impact in Northern Canada requires that activities be geographically located and that the vegetation present prior to conversion is known—a significant challenge, considering that the area of interest extends over 560 Mha, intersecting with 11 reporting zones. Land-use change areas were estimated based on mapping from image interpretation for the years 1990, 2000 and 2010, as described in Annex 3.5.7.3.”

There is a statement about the QA/QC measures applied; this is a cross-reference to a statement about the general QA/QC procedures that are applied in the inventory:

“Quality Assurance / Quality Control and Verification: Section 1.3 in Chapter 1 describes the general QA/QC procedures being implemented for Canada’s GHG inventory. The same procedures apply to this category as well.”

Assessment:

Assessment: Our assessment of the transparency of this Part 1 of the NIR, with respect to describing land use conversions in Canada, is that the explanations about the methods used to estimate the land use conversions are normally not provided. It would be useful to have even just one or two sentences on the land use tracking methods used in each category, or a clear cross reference to the sections of the NIR where there is a description of the methodologies used. There are brief descriptions of the methods used for tracking land use changes in some sections e.g. 6.7.1. Peat Extraction, and there are some references to Annex 3.5, which contains the detailed methodologies used in the Land Use, Land-Use Change and Forestry Sector (A3.5. “Methodology for the Land Use, Land-Use Change and Forestry Sector”).

Earth Observation data are used: for example, Landsat imagery to determine Cropland converted to Settlements. Visual interpretation is used in this example to verify that the conversions have been identified correctly.

In some cases, the Canadian inventory struggle with lack of activity and are unable to make an assessment of emissions, for example , in the category Land Converted to Settlements.

Often the sections on Quality Assurance / Quality Control and Verification simply refer back to a section on QA/QC in Part 1 of the NIR. This section is informative and shows general compliance with IPCC requirements, but, is quite generic in nature. In our experience, this approach of cross referencing back to a general section on QA/QC to satisfy the UNFCCC reporting requirements on QA/QC is quite common practice in NIRs. In a couple of cases, more detailed approaches to QA/QC are provided, for

example, in section 8.2.1. “Cropland Converted to Settlements”. Providing this extra detail both improves transparency of reporting and shows the country has applied category specific QA/QC, which is likely to improve the quality of the inventory and is IPCC good practice.

Part 2 of NIR –presenting detailed methodologies

Section “A3.5.2.5. **Forest Conversion**” (pg. 129) summarises the approach used for estimating forest areas converted to other land uses.

In summary, the approach for estimating forest areas converted to other land uses is based on three main information sources: systematic or representative sampling of remote sensing imagery, records, and expert judgement:

“The approach for estimating forest areas converted to other land uses is based on three main information sources: systematic or representative sampling of remote sensing imagery, records and expert judgement/opinion.”

“The basic methods have been tested in several pilot projects (Leckie, 2006a), and the methodology has been implemented across the country.”

“The core method involves remote sensing mapping of forest conversion based on samples from Landsat images dated circa 1975, 1990, 2000, 2007 and 2011. Change enhancements between two dates of imagery are produced to highlight areas of forest cover change and identify possible forest conversion events (i.e. “candidate events”). The imagery is then interpreted to determine (1) whether the land cover of the candidate event was forest initially (at Time 1), and (2) the actual land-use change at Time 2 (Leckie et al., 2002, 2010b). This forest conversion interpretation process is strongly supported by additional spatial data, including digitized aerial photographs; snow-covered, leaf-off, winter Landsat imagery; secondary Landsat images from other dates and years; ancillary data, such as maps of road networks, settlements, wetlands, woodland coverage, and mine and gravel pit locations; and specialized databases giving locations of oil and gas pipelines and well pads (Leckie et al., 2006; Dyk et al., 2015). When readily available, detailed forest inventory information is also used.

Change imagery is interpreted and analyzed; each forest conversion event larger than 1 ha is manually delineated. The forest type, maturity and density prior to forest conversion is interpreted,³⁰ and the post-deforestation land use recorded (“post-class”). Confidence ratings on the land use at the initial time and a later time period are used in subsequent quality control and field validation procedures.

Monitoring of forest conversion activity covers all forest areas of Canada and is not limited to the managed forest. The entire forested area of Canada is broadly stratified into regions of expected forest conversion level and dominant cause, which dictate the target sampling intensity. Depending on the expected spatial patterns and rates of forest conversion, sampling approaches range from complete mapping to systematic sampling over the entire analysis unit of interest to a representative selection of sample cells within a systematic grid. For example, in populated areas of southern Quebec and in the Prairie fringe, a 12% sampling rate was generally achieved, with 3.5 × 3.5-km sample cells at the nodes of a 10 × 10-km grid (Figure A3.5–7). A lower sampling rate is used in some of the forest activity zones characterized by low population density, where the main economic activities are forestry and other resource extraction. Special cases of known, localized and large forest conversion activities are also identified, such as hydroelectric reservoirs and oil sands development in Alberta. In such cases, the entire areas are handled as single events (“Hot Spot” in Figure A3.5–6), with spatially complete mapping.

In practice, resource constraints limit the size of the remote sensing sample; wherever possible, a target sampling rate of 12% or 6% was achieved. It is also important to note that different sampling rates may be applied for each time period, in an effort to track differing activity rates between time periods. The total areas, either fully mapped or sampled, cover a large portion of the Canadian land base (Figure A3.5–6), e.g. approximately 346 million hectares (Mha). This total area was mapped over different time periods, of which over 17 Mha were mapped for 1975–1990, 41 Mha were mapped for 1990–2000, 22 Mha were mapped for 2000–2008 and 23 Mha were mapped for 2008–2013.

Mapping is updated on a roughly five-year time cycle and may be integrated progressively by project for the most recent time period.

Records were gathered when available. They consist mostly of information on forest roads, power lines, oil and gas infrastructure, and hydroelectric reservoirs (Leckie et al., 2006). The temporal coverage, availability and applicability of these records are assessed to determine the most appropriate information sources (records or imagery). Records data are sometimes used to aid in the validation of estimates made through image interpretation. In particular for British Columbia, records data are used to provide estimates of conversion activity for power lines and oil and gas activity. A mix of remote sensing image interpretation and records data are used to assess the areas of forest converted as a result of hydroelectric development.

Expert opinion is only called upon when remote sensing sampling is insufficient and records data are unavailable or of poor quality. Expert judgement is also used to reconcile differences between records and remote sensing information and to resolve large discrepancies in each mapped time period (e.g. 1975–1990, 1990–2000, 2000–2008, 2008–2013) area estimate. In such cases, available expert opinion and data sources are brought together, remote sensing and records data are reviewed, and decisions are made (Leckie, 2006b; Leckie et al., 2010a; Dyk et al., 2015). For most estimates and certainly for those with large impact, estimates are derived directly from remote sensing samples.”

The quality control of Quality Assurance / Quality Control of Forest Conversion Data is discussed in the forest land LUC section of this NIR.

Land Converted to Forest Land, pg. 133

This is the statement about the records of land conversion to forest land in Canada; to note that there are issues with completeness of recent afforestation activities:

“Records of land conversion to forest land in Canada were available for 1990–2002 from the Feasibility Assessment of Afforestation for Carbon Sequestration (FAACS) initiative (White and Kurz 2005). Conversion activities for 1970–1989 and 2003–2008 were estimated based on activity rates observed in the FAACS data. Additional information from the Forest 2020 Plantation Demonstration Assessment was included for 2004 and 2005, and an environmental scan was performed to identify additional sources of information on afforestation rates from 2000–2008. Each event, regardless of date, source, type or location, was converted to an inventory record for the purposes of C modelling. All events were compiled in a single data set of afforestation activity in Canada from 1970 to 2008. No new afforestation activity data were identified for 2009 to the current inventory year. Efforts continue to obtain additional data on recent afforestation activities in Canada.”

Section A3.5.4. Cropland, pg. 140

Section A3.5.4.2. Grassland Converted to Cropland, pg. 152

Data sources were a combination of census data and earth observation:

“The area of Grassland remaining Grassland (GLGL) was estimated using a combination of data from the Census of Agriculture and EO data. Area estimates of grassland converted to cropland were based on reconciling changes in land area between GLGL and land in cropland management.”

Section A3.5.4.3. Forest Converted to Cropland, pg. 154

Data sources:

“The approach used to estimate the area of forest land converted to cropland is described in section A3.5.2.3. The annual forest conversion by reconciliation unit, RU, was disaggregated to SLC polygons on the basis of concurrent changes in the area of cropland within SLC polygons. Only polygons that showed an increase in cropland area for the appropriate time period were allocated to forest conversion, and the amount allocated was equivalent to that polygon’s proportion of the total cropland increase within the RU.”

Section A3.5.6. Wetlands, pg. 158

Section A3.5.6.1. Peat Extraction, pg. 158

Data sources included earth observation data. Interpolation and extrapolation were used to create a time series of data:

“An EO mapping approach based on manual delineation and interpretation of aerial photography, satellite imagery and ancillary data was developed to map the extent of peatland areas disturbed by peat extraction for circa 1990, 2007 and 2013 time periods. Through image interpretation, the total disturbed area was allocated into the following four land management subcategories: active extraction areas, abandoned areas, rehabilitated areas and restored areas. Geospatial data developed by the Peatland Ecology Research Group and information provided by industry experts were utilized to aid subcategory allocation. In addition, for a subset of sites, the pre-disturbance land cover class (forest, shrubby or open bog peatland) was determined in order to identify the land category types converted (Forest Land or Other Land).

Annual area estimates were developed using interpolation between mapped time periods and extrapolation after 2013. Annual area estimates for various land management categories were then refined based on secondary data sources. The two main secondary data sources were industry statistics on peatland areas managed for peat extraction in 2015 compiled by the Canadian Sphagnum Peat Moss Association (CSPMA) and a survey of abandoned peat extraction sites in the provinces of Quebec and New Brunswick (Poulin et al., 2005).”

Section A3.5.6.2. Flooded Lands, pg. 160

Three activity data sources were used:

“Data Sources: The three main data sources used to develop area estimates were (1) information on forest conversion due to reservoir impoundment in reporting zones 4 and 5 (see section A3.5.2.3, Forest Conversion); (2) the Canadian Reservoir Database (Duchemin, 2002); and (3) official industry numbers, derived from industry correspondence (Eichel, 2006; Tremblay).”

Section A3.5.7. Settlements, pg. 162

Section A3.5.7.2. Cropland Converted to Settlements, pg. 164

Land-maps were used as primary data sources:

“Urban and industrial expansion has been one of the main drivers of Cropland conversion in Canada. Areas of Cropland conversion to Settlements were estimated based on the land-use maps for 1990, 2000 and 2010 developed in Huffman et al. (2015a). Areas of conversion for the 1990–2000 and 2000–2010 periods were calculated through spatial analysis for each reporting unit and divided by the number of years in order to develop constant annual conversion rates. Areas of conversion were extrapolated after 2010.”

Section A3.5.7.3. Grassland Converted to Settlements, pg. 165

Polygon areas formed the activity data needed, based on land cover maps:

“Land-use change was derived from the difference in polygon areas for each date, providing an area of change between the time periods (i.e. 1990–2000, 2000–2010), that was divided by the total years in the time period to produce a constant annual rate of change. The same annual rate of land-use change was applied for the years prior to 1990 and following 2010. The pre-conversion land-use type for each of the land-use change polygons was based on available land cover maps (Wulder et al., 2008; Hermosilla et al., 2016), visual interpretation and vegetation indices of concurrent imagery to avoid including areas in other land-use categories (e.g. Forest Land, Cropland, Wetlands and Other Land). Furthermore, deforestation events above 60 degrees were also used to confirm that areas determined as forest conversion to settlements were excluded, to avoid double-counting.”

Assessment

Assessment: The primary approach to estimating forest areas converted to other land uses is based on using remote sensing mapping of forest conversion based on samples from Landsat images dated circa 1975, 1990, 2000, 2007 and 2011. This approach is supplemented with other records, and expert judgement. Expert judgement is only used when remote sensing sampling is insufficient, and records data are unavailable or of poor quality. Expert judgement is also used to reconcile differences between records and remote sensing information and to resolve large discrepancies in each mapped time period (e.g. 1975–1990, 1990–2000, 2000–2008, 2008–2013). In practice, resource constraints limit the size of the remote sensing sample; wherever possible, a target sampling rate of 12% or 6% was achieved.

Other land conversions are based on a mix of data types: land cover maps, census data, and earth observation. A mixture of aerial photography, satellite imagery and ancillary data are used, although in some cases the NIR does not state exactly what data sources are used. Some of the data are relatively old – for example the last set of data for Cropland converted to Settlements dates from 2010, and the later time series is based on extrapolated data.

In many cases, a lot of effort is needed to create a consistent time series of activity data. Methods of interpolation and extrapolation are used, which are consistent with IPCC good practice approaches.

In Part 2 of the NIR, which contains details of the methodologies used to assess LUC, there are no specific subsections on QA/QC. There are specific subsections on uncertainties though. It would increase transparency if even brief category specific details on QA/QC were provided.

References

Canada (2020a). *National Inventory Report 1990–2018: Greenhouse gas sources and sinks in Canada. Part 1.* Canada’s submission to the United Nations Framework Convention on Climate Change. Pollutant Inventories and Reporting Division, Environment and Climate Change Canada, Public Inquiries Centre, 12th Floor, Fontaine Building, 200 Sacré-Coeur Boulevard, Gatineau QC K1A 0H3. <https://unfccc.int/documents/224829>

Canada (2020b). *National Inventory Report 1990–2018: Greenhouse gas sources and sinks in Canada. Part 2.* Canada’s submission to the United Nations Framework Convention on Climate Change. Pollutant Inventories and Reporting Division, Environment and Climate Change Canada, Public Inquiries Centre, 12th Floor, Fontaine Building, 200 Sacré-Coeur Boulevard, Gatineau QC K1A 0H3. <https://unfccc.int/documents/224829>

Canada (2020c). *National Inventory Report 1990–2018: Greenhouse gas sources and sinks in Canada. Part 3.* Canada’s submission to the United Nations Framework Convention on Climate Change. Pollutant Inventories and Reporting Division, Environment and Climate Change Canada, Public Inquiries Centre, 12th Floor, Fontaine Building, 200 Sacré-Coeur Boulevard, Gatineau QC K1A 0H3. <https://unfccc.int/documents/224829>

Andrew Dyk, Don Leckie, Sally Tinis, and Stephanie Ortlepp (2015). *Canada's National Deforestation Monitoring System: System Description*. Deforestation Monitoring Group, Digital Remote Sensing. Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre. Information Report BC-X-439. The Pacific Forestry Centre, Victoria, British Columbia. <https://cfs.nrcan.gc.ca/publications?id=36042> (Canada.ca>Natural Resources>Canada Our Natural Resources>Forests and forestry>Canadian Forest>ServicePublications)

Dyk, A.; Tinis, S.; Leckie, D. (2011). *National Inventory Report, 2011: Deforestation area estimation for Canada: Quality control overview*. Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Victoria, B.C. Report DRS-N-031.

6.4.2 Australia

This analysis is separated into the sections covering the LUC tracking approaches used in the forest land and non-forest land sectors. As a general point, there is much more methodological detail provided both in the NIR and in supplementary reports regarding the approaches used to estimate the areas of forest land and the transitions between forest land and other land uses.

12.5.1.3 Forest sector

Australia has a methodologically well-developed LULUCF inventory, and a comprehensive National Inventory Report.

The NIR is issues in three parts. Information in the NIR relevant to this part of the study include:

- **Volume 2 of the NIR:** This volume provides details of the LULUCF methodology. Appendix 6.A.2 “Monitoring change with remote sensing imagery” provides the key source of information for this review.

Volumes 1 and 3 of NIR are not directly relevant to this this section of the study.

Information presented in the main body of the NIR

The method that Australia uses to estimate its GHG emissions and removals in managed forests is presented in 6.2.2 “Methodology”. Predominantly country specific methodologies and Tier 3 models (Table 6.2) are used for LULUCF. The methods used in the estimation of the LULUCF categories of the inventory are described in detail in Appendices 6.A to 6.K.

Carbon modelling

Australia uses a carbon-stock accounting approach. Their 2016 Inventory introduced a new perspective on the data underpinning land sector calculations in the form of carbon-stock accounts compiled under the System of Environmental-Economic Accounting. **These carbon stocks can be spatially mapped using the Full Carbon Accounting Model (FullCAM) architecture underpinning the estimates.** In the NIR 2020, figure 6.7 shows carbon density on the whole of the Australian landscape, and figure 6.8 shows the changes in forest-related carbon stocks with a focus on South-Western Australia. These maps show the higher density of carbon in Australia’s native forests and highlight the mixed stories of land clearing and regeneration over the recent decades.

Australia’s land sector inventory system integrates spatially referenced data with an empirically constrained, mass balance, **carbon cycling ecosystem model (FullCAM)** to estimate carbon stock changes and greenhouse gas emissions (including all carbon pools, gases, lands and land use activities). The system supports Tier 3, Approach 3 spatial enumeration of emissions and removals calculations for the following sub-categories:

- Forest land converted to cropland, wetlands (Flooded Land), grassland, and settlements

- Grassland, cropland and settlements converted to forest land; and
- The agricultural/grazing system components of cropland remaining cropland and grassland remaining grassland.

Spatial enumeration is achieved through the use of a time-series (since 1972) of Landsat satellite data which is used to determine change in forest and sparse woody vegetation extent at a fine spatial disaggregation. The forest cover change information is coupled together with spatially referenced databases of climate and land management practices which allows a comprehensive quantification of emissions (see Appendices 6.A and 6.B in the NIR 2020).

Land sector reporting within Australia's National Inventory System integrates a wide range of spatially referenced data through a process based empirical model (Tier 3) to estimate carbon stock change and greenhouse gas emissions at fine spatial and temporal scales. Analysis and reporting includes all carbon pools (biomass, dead organic matter (DOM) and soil), all principal greenhouse gases (CO₂, CH₄ and N₂O), and covers both forest and non-forest land uses. A Tier 3 method is used to estimate carbon stock changes for agricultural soils, living woody biomass (excluding perennial woody horticulture) and dead organic matter.

FullCAM is a process based ecosystem model that calculates greenhouse gas emissions and removals in both forest and agricultural lands using a mass balance approach to carbon cycling. The FullCAM framework and its development are described in Richards (2001) and Richards and Evans (2004).

Land monitoring systems

The approach Australia uses for land monitoring is described in section 6.3.2 "Land monitoring systems". Remote sensing provides a key source of data. Supplemental data is provided from the Land Use Mapping programme.

"Australia uses Approaches 1 and 3 as described in the 2006 IPCC Guidelines for National Greenhouse Gas Inventories to monitor land use, land use change and forestry.

The principal monitoring system is a remote sensing programme used to identify forest lands and changes in forest cover. Significant improvements to the remote sensing programme were made in 2016 (see Appendix 6.A for details).

The remote sensing programme is implemented by the Department of Industry, Science, Energy and Resources. The system monitors national forest cover on an annual basis using Landsat satellite data (collected by MSS, TM, ETM+ and OLI sensors). The time series of national maps of forest cover extends across 27 time epochs from 1972 to 2018 and has been assembled on an annual basis since 2004. These maps are able to detect fine scale changes in forest cover at a 25 m by 25 m resolution.

Within forest land remaining forest land, data on areas of forest management are drawn from Australia's National Forest and Wood Products Statistics (ABARES 2019a), Australia's State of the Forests Report (ABARES 2008) and Lucas et al. (1997).

Supplementary spatial information from the Land Use Mapping programme of Australia's Bureau of Agricultural Resource Economics and Sciences (ABARES, 2014) is used to identify land areas in the grassland, wetlands and other land categories. Cropland has been updated to the September 2017 revision of these areas (ABARES, 2017). Settlements has also been updated using this revision and supplemented by spatial data from other unpublished sources. The other land categories are expected to be progressively updated over time."

Of note is the use of short time interval monthly satellite data to detect carbon stock changes following fires. This is discussed in section 6.4.1.3 "Other native forests".

"A time-series of monthly satellite data is used to identify the time and location of fires, which are simulated at the 25m x 25m plot size. The AVHRR burnt area product produced by the Western Australian Land Authority (Landgate), is tailored to Australian conditions and based on the visual interpretation of fire areas by experienced operators. The data was assessed by the Royal Melbourne

Institute of Technology (RMIT, 2014) and compared with a range of alternative datasets, and was found to be the most suitable and highest quality time series data available.”

Information presented Appendix 6 A” Land cover change” of the NIR

Appendix 6A contains most of the information relevant to this review. The information provided is **very detailed** and shows considered investment to develop image processing capabilities.

The Annex states that to process satellite data, a detailed protocol of remote sensing specifications for land cover change was developed by Furby (2002) through extensive pilot testing (Furby and Woodgate, 2002) to ensure time series consistency of methods, and the provision of spatially accurate land cover change data through time. These specifications determine the exact way that images are acquired, processed and classified.

In all, three references to technical reports regarding the development and testing of remote sensing image data with Furby as an author are cited. It is possible to access two of these reports, but they are not available on a “live” web site, and are only available on an Australian government web archive site. It took some time to find the second reference. The third report is unpublished, and we could not assess if it contains useful information for this review. These difficulties reduce the transparency of the GHG inventory.

Furby, S., 2002. *Land Cover Change: Specifications for Remote Sensing Analysis*. National Carbon Accounting System Technical Report No. 9 (236pp), Australian Greenhouse Office, Canberra.

Available at:

<https://catalogue.nla.gov.au/Record/1027610>

<https://webarchive.nla.gov.au/awa/20090722031721/http://pandora.nla.gov.au/pan/102841/20090717-1556/www.climatechange.gov.au/ncas/reports/pubs/tr09final.pdf>

Furby, S. and Woodgate, P., 2002. *Pilot Testing of Remote Sensing Methodology for Mapping Land Cover Change*. National Carbon Accounting System Technical Report No. 16 (354pp), Australian Greenhouse Office, Canberra.

Main report available at:

<https://catalogue.nla.gov.au/Record/2550466>

https://webarchive.nla.gov.au/awa/20030708155556/http://pandora.nla.gov.au/pan/23322/20020220/www.greenhouse.gov.au/ncas/files/pdfs/tech16_jun01.pdf

Furby, S., 2016. *General Guidelines for Thresholding Images Using Image Matching, Version 13*, Unpublished Report, CSIRO Data61

Report unpublished

The QA / QC of the activity data for detecting gains and losses of woody vegetation is described in Appendix 6.A.4.

The Australian 2020 NIR provides more than 20 pages of detail about the image processing and land use tracking approaches used. This provides good transparency in the reporting, but in addition to the supporting reports, constitutes a lot of material to review and synthesise.

To provide a “fast assessment” of the information in the Australian 2020 NIR, we have extracted sections below directly from the NIR. In these sections, the key points are highlighted. In this draft of the review, we have not provided much extra commentary, but as before, have included an “Assessment” section after the material extracted from the NIR. The figure numbers used are those in the NIR.

Satellite Data Processing

The summary of the methodology indicates that machine learning algorithms for change detection are used. This is potentially very interesting

“The sequence of data processing stages have been streamlined since the development of the Australian Geoscience Data Cube in 2014. Migration of legacy data processing methods to the Data Cube

environment has been completed including use of machine learning algorithms for change detection. The process to produce the assessment of Australia-wide land cover change consists of:

- image compositing of highest quality cloud free pixels acquired during the summer season for the southern tiles and the winter season for the northern tiles, from the Data Cube;
- mosaicing¹² of multiple images to the individual map tiles for each time sequence;
- perform a single-epoch 3-class classification using the Random Forests classifier;
- conditional probability network (CPN) analysis (Kiiveri *et al.*, 2001), each year over the entire time series; and
- attribution¹³ of change to direct human-induced change. “

Image acquisition and selection

“The time series of available Landsat images extends from 1972 to 2019. The selection of periods for analysis, shown in Table 6.A.1, was designed to give maximum temporal resolution immediately before and after 1990 and for the period from 2004 onwards to maximise accurate detection of trends in land cover change over time.

Since 2004 imagery has been delivered on an annual basis. Figure 6.A.1 shows the 37 map tiles used in the remote sensing programme (red), the north-south seasonal divide used for image capture (blue line) and the paths/rows of Landsat imagery (yellow).”

¹² Mosaicing aggregates images into the map tiles shown in red in Figure 6.A.1, removing overlaps in the original 185 km*185 km images and optimising cloud removal.

¹³ Attribution uses a combination of automation and visual inspection of the image sequence to determine the cause of land cover change and determine subsequent/existing land use.

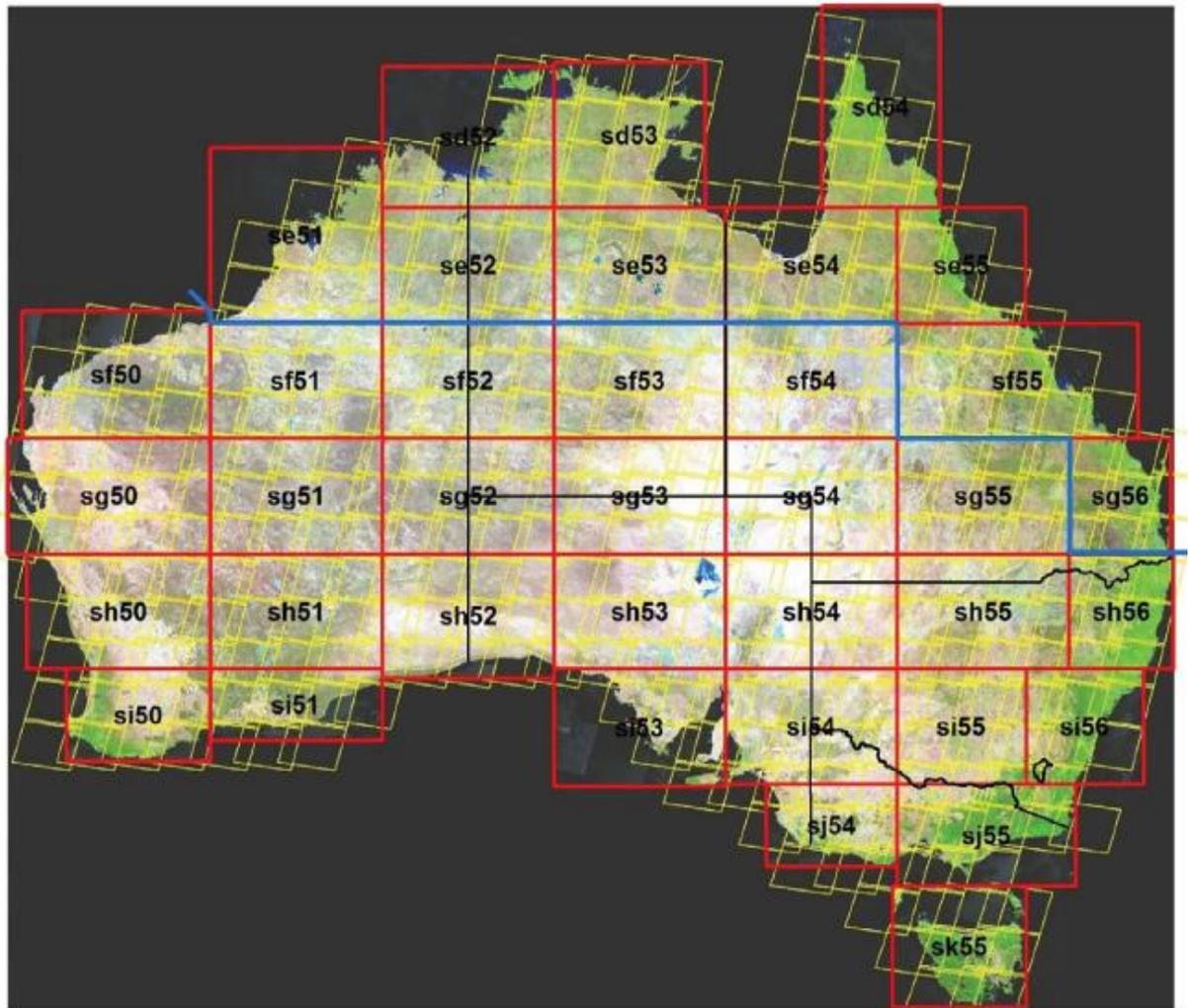


Figure 6.A.1 37 1:1 million scale map tiles used in the remote sensing programme.

“Selection of suitable Landsat scenes from the Data Cube is fully automated. For a given location, the season from which the scene should be selected is identified and the best (cloud-free) image is automatically allocated from the stack within the Data Cube. The image selection criteria (Furby, 2002) require the images to be within three months of the nominated target date. The target dates vary between the north (winter or dry season) and south (summer) of the country and aim to provide the best possible forest discrimination. The precise date allocated to each land cover change (clearing and regrowth) pixel is randomly generated by FullCAM, within the sequence of coverage dates for the relevant map tile. This method provides a random (unbiased over a large sample) distribution of initialisation dates (timing of land cover change event) for the carbon model, within the constraint of the two dates in the overall interval of the image sequence.”

Table 6.A.1 Landsat Image sequence.

| Year | Resolution (m) | Time since previous image (yrs) |
|------------------|----------------|---------------------------------|
| 1972 | 50 | - |
| 1977 | 50 | 5 |
| 1980 | 50 | 3 |
| 1985 | 50 | 5 |
| 1988 (early) | 25/50 | 3 |
| 1989 (end) | 25/50 | 2 |
| 1991 (early) | 25 | 1 |
| 1992 | 25 | 1 |
| 1995, 1998 | 25 | 3 |
| 2000, 2002, 2004 | 25 | 2 |
| 2005–2019 | 25 | 1 |

Mosaicing

“Scene selection and compositing is automated so multiple images can be combined within each path/row to create a cloud free composite (Furby, 2016). Figure 6.A.2 shows how a mosaic is constructed using multiple images within each path and row, resulting in a composite cloud free image. However, in inherently cloudy locations, some gap filling from earlier imagery may be required.”



Figure 6.A.2 Image selection procedure, to create composite cloud free imagery mosaics

Unit of analysis – spatial resolution of the imagery

“The ‘natural’ pixel size of the 1972 to 1985 Landsat MSS (57 m × 79 m) is resampled to a 50 × 50 m pixel. The 30 × 30 m native resolution of the Landsat TM, ETM+ and OLI data available after 1985 is produced as 25 × 25 m pixels. This approach deals with the change in pixel size of the various Landsat sensors over time and supports the need for spatially and temporally consistent integration with other spatial data used in FullCAM.

To apply the pixel-by-pixel analysis over the period where the pixel size changed from 50 m to 25 m, a 50 m MSS equivalent (in both spatial and spectral resolution) is derived from the 1989 TM (25 m) data, and then forest extent is calculated separately from both the 50 and 25 m data sets. Differences in the extents of forest between these two outputs are due to “sensor change”. An overlap technique is used to ensure time-series consistency such that the assessment of land cover change for 1988–89 is then based on a 50 m to 50 m comparison, while the 1989–1991 data is a 25 m to 25 m comparison. As part of continuous improvement, processing of 1988 Landsat TM data at 25m spatial resolution has been completed, replacing the 50 m resolution MSS data for 1988. Consequently the entire land cover time series data has been recalculated making use of best available data while maintaining time series consistency. This approach is consistent with good practice for ensuring time-series consistency where the instruments used to collect activity data change or degrade through time (IPCC, 2003 page 5.58).

All Landsat derived data are used at a consistent 25 m resolution for the full time series analysis by resampling the 50 m pixels (1972–1985 products) into four 25 m pixels. A spatial-temporal model (see the [Conditional Probability Network](#) section below) is used to reduce the effect of “mixed” isolated and edge pixels in the overlap period. The ability to determine, from 1988 onwards, the effects of land use change to 0.2 ha minimum areas is robust, given that this area is greater than the pixel resolution and the approach used removes mixed and other pixels which are temporally and spatially inconsistent.

Resampling Landsat TM, ETM+ and OLI sensor data to 25 m pixels is common practice and provides consistency over the multiple resolutions of Landsat sensors while ensuring uniformity across the time series. Quality assurance and validation processes confirm that accurate results are achieved with this resampled data.”

Use of Landsat 8 Data

“Observations of recent land cover change have been derived from the latest sensor on-board the Landsat 8 satellite, Operational Land Imager (OLI). OLI is an advanced sensor designed to collect improved quality data, ensuring continuity of previous instruments – Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) sensors. Landsat 8 products supplied through the Australian Geoscience Data Cube are in a new format known as the Australian Reflectance Grid (ARG25). ARG25 is a pre-processed product corrected for geometric distortions and calibrated as absolute surface reflectance, hence the specifications of this new product are quite different to the previous Landsat 5 and 7 data products used for the national inventory Land Cover Change Programme (LCCP). To ensure time series consistency and compatibility with the existing LCCP, a detailed technical assessment of the geometric and radiometric consistency and interoperability between these two products was undertaken.

Geometric consistency was assessed by matching about 13,300 ground control points (GCP) drawn from the LCCP scenes held in the national inventory data library and the corresponding ARG25 scenes. Assuming that the correlation matching succeeds in correctly registering each point, the position residuals provide a measure of the accuracy of co-registration of the two datasets. This analysis showed that whilst the temporal geometric accuracy of ARG25 products is highly consistent, several GCPs had residual matching errors ranging from 1, 2 and greater than 2 pixels compared to the LCCP products. The mis-registration, if not accounted for, would result in false change being reported. To resolve this, the mean residual vector for each ground control point (GCP) was calculated and applied to the LCCP scenes to align with the ARG25 product base. The scene specific transformation coefficients ensure that the two products are aligned and consistent to within a pixel for the entire country.

The second step was to assess the radiometric consistency between the ARG25 and LCCP products using 339 image pairs from the 2005 continental coverage. The two products were paired up based on Landsat path and row, and image acquisition date. Null pixels in either image were discarded. Pixels located in very dark or very bright regions in the LCCP images were also excluded from the analysis, since such values may have potentially saturated during the pre-processing. The remaining pixels were linearly regressed against each other, assuming that the relationship will be strongly linear if both products are internally consistent in relation to radiometric characteristics. Correlation values were calculated for each band, gain, and offset combination. The gain and offset values for converting LCCP pixel values into ARG25 pixel values can be expressed as –

$$\text{ARG 25} = \text{gain} \times \text{LCCP pixel value} + \text{offset}$$

The relatively high correlations found in the 2005 coverage confirm that there is a strong linear relationship, across all bands, between the LCCP values and the equivalent ARG25 image values. A scene-specific, linear transformation coefficient for each band was calculated to convert the LCCP calibrated pixel values to be consistent with the ARG25 surface reflectance values (Devereux, et al. 2013). The time series consistency of this method was also assessed for selected sites using eight years of surface reflectance data.

Based on this study, from 2015 the ARG25 Landsat 8 datasets (Figure 6.A.3) have been processed to a consistent quality, LCCP compatible tile based mosaic which are then subjected to image classification to derive forest probability maps.”

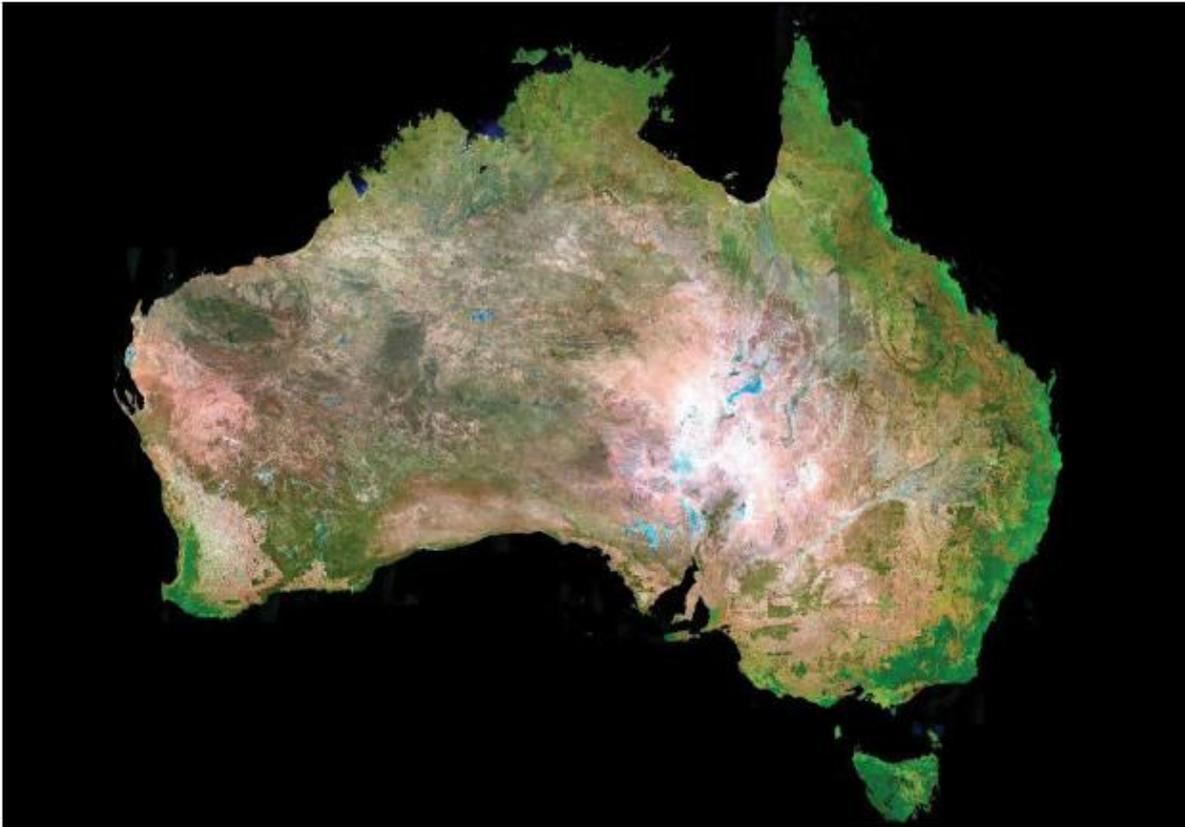


Figure 6.A.3 2019 Landsat 8 surface reflectance image of Australia

3-class Random Forests classifier

“A new method of classifying woody vegetation has been adopted in this National Inventory update. The method has changed from a thresholding approach using simple decision boundaries, to a Random Forests (RF) classifier (Breiman, 2001). The RF classifier uses a sophisticated decision-tree approach, building a large number of trees from samples of training or reference data to create a class prediction. For a given pixel, the average prediction across all the trees is taken. It also allows class membership probabilities to be undertaken concurrently, requires minimal manual intervention and is readily extended to any number of classes of interest.

This method incorporates previous National Inventory innovations such as the move from a 2-class (forest, non-forest) classification to a 3-class classification (forest, sparse, non-woody). Figure 6.A.4 compares the previous 2-class product with the current 3-class outputs. Background image is from UrbanMonitorTM 2014 (Figure 6.A.4 (A)), and a Landsat false colour composite 2014 (B). Forest is highlighted green and Figure 6.A.4 (D) shows sparse vegetation (in orange) that was detected using the 3-class algorithm. As the entire range of woody vegetation needs to be monitored for reporting under the Kyoto Protocol second commitment period and the Paris Agreement, it is essential to create a product that better encompasses all woody vegetation (Figure 6.A.5).

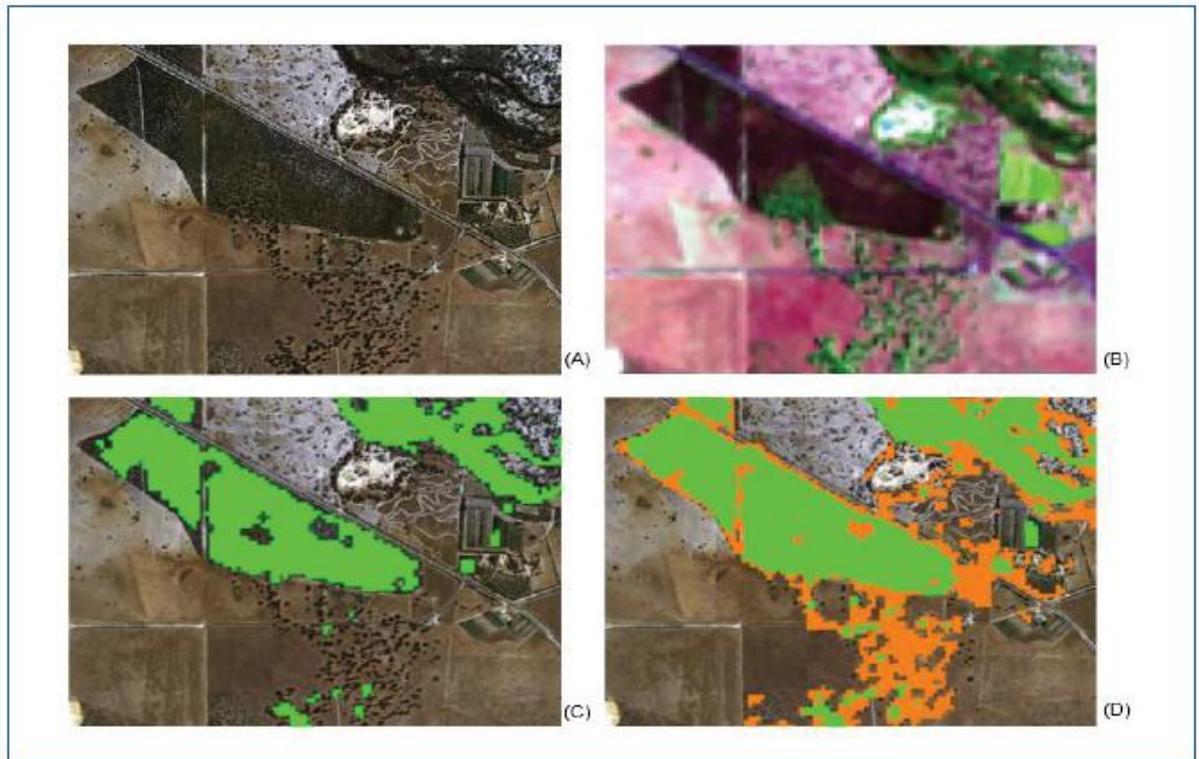


Figure 6.A.4 Comparison of traditional 2-class forest and non-forest product with the new 3-class product

The Random Forests classification was performed on Landsat 8 imagery for the current epoch in a semi-automated manner, to investigate the parameter settings required to optimize the performance of the algorithm. The classifier was fitted independently to each of the stratification zones used in the previous method, which encompass local soil, vegetation and land use types. The relative importance of the individual input variables (ie spectral bands 1–6, spectral indices 7–8, texture bands 9–10, texture index 11) are tracked per stratification zone, and results can be used to modify the variables used in future updates.

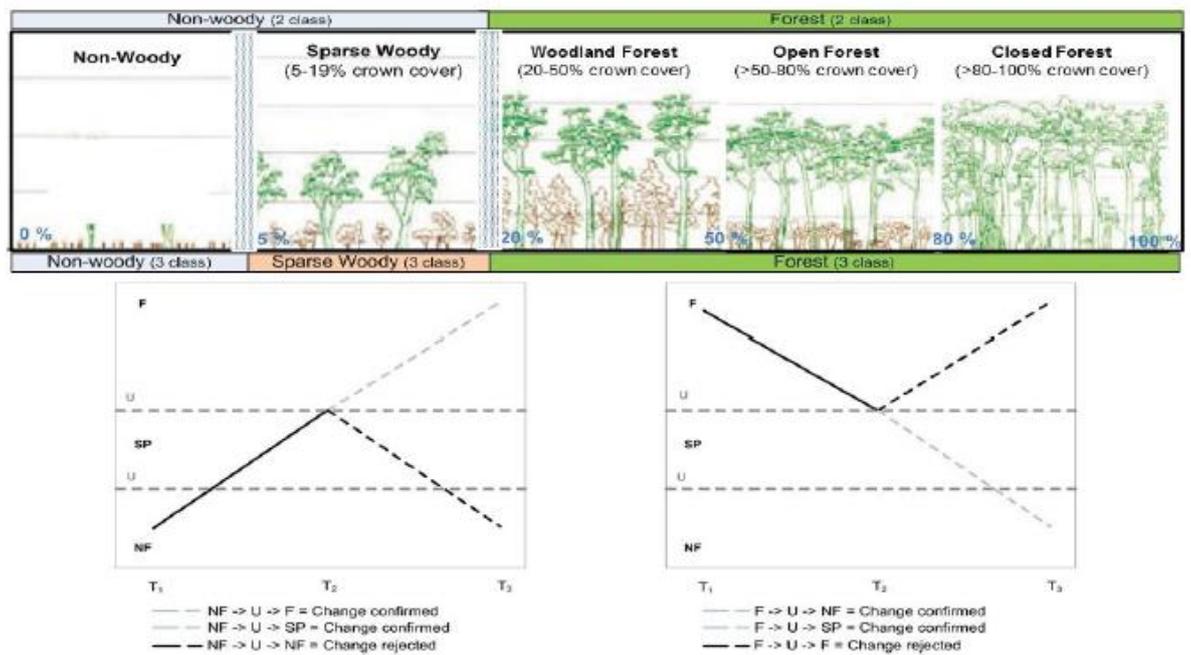
The Conditional Probability Network (CPN) outputs for 2018 were used as the training sample or “base” to train the RF classifier for the new update. Twenty percent of this data is extracted randomly and reserved to calculate an independent accuracy assessment. Early testing indicated that woody extent and change classifications were very sensitive to the choice of training samples, and the RF classifier produced much higher probabilities of class membership than the previous thresholding approach. This is most noticeable in the sparse class, which has historically experienced the greatest uncertainty. As a result, training samples were restricted to more pure examples of each class to enable the classifier to determine the boundary between them.

Early results also showed that the RF classifier could classify an area cleared in the latest epoch as having experienced a reduction in the probability of forest, but not necessarily reduce the probability enough to enable the CPN to correctly identify the change, given multiple years of high forest cover probabilities before the change event. To correct this problem for this update, a change mask was created by comparing the spectral index values between 2018 and 2019. Any pixels that fell under the change mask were excluded from the training sample.

Ultimately a combination of reduced error rates for sparse in 2019, the use of a change mask and temporal rules restricting forest to sparse conversion leading up to 2019 were employed, resulting in products more consistent with earlier versions.

In future, the single-epoch classification will be refined to enable a multi-temporal classification to be performed across all epochs, to ensure consistency across the time series. Once all refinements have

been made and automation is fully implemented, this should assist in moving towards the planned use of Sentinel 1 and 2 imagery.



Source: Adapted from Australia's State of the Forests Report 2013

Figure 6.A.5 3-class algorithm to detect entire range of woody vegetation

Conditional Probability Network analysis

Remote sensing pilot testing demonstrated the need for time-series consistency in image data pre-processing, analysis and subsequent formation of time-series woody/sparse/non-woody labels. The operational standards (Furby, 2002) give explicit emphasis through documented rule sets to each of these areas. For time-series classification, these standards also include the use of a joint spatial-temporal model, in this case a Conditional Probability Network (CPN) (Caccetta, 1997; Kiiveri *et al.* 2001, 2003), for determining a time-series of woody/ sparse/non-woody classes. This process produces superior woody extent and change results compared to a process reliant on pair-wise differencing of image pairs. The use of pair-wise differencing methods can lead to change estimates that are affected by errors due to seasonally changing land management effects (introducing large contiguous areas of false change), or by subtle sampling differences where mixed pixels have varying composition of woody/non-woody from year to year (producing many isolated false change pixels or edge effects at woody boundaries).

The land cover change programme uses Conditional Probability Network (CPN) analysis to strengthen confidence in the 'woody', 'sparse woody' and 'non-woody' classification of a pixel (previously 'forest' or 'non-forest'). This is achieved using a series of spatial and temporal rules to create woody vegetation and land cover conversion datasets. The temporal rules bias against unlikely events such as multiple one year conversions between woody and non-woody, as the CPN empirically assesses the logic of vegetation cover status of a pixel at a point in time, compared to the previous and subsequent images. This helps to eliminate false change from a single image that may be due to anomalies in the data such as unseasonal greenness, wetness or flooding, or missing data. The rules are particularly effective when the time between observations is less than that of a forest growth and harvest cycle.

The spatial rules consider the labelling of a pixel in the context of its spatial surroundings, where labels that are consistent with the neighbouring labels are reinforced as opposed to those that are inconsistent (e.g., isolated pixels). This method evaluates the status of adjoining pixels as well as the pixel of interest,

which has the effect of reducing ‘flickering’ false change in scattered and edge woody pixels. It also ensures that individual and small clusters of forest pixels have a high classification certainty in relation to their neighbouring pixels and through time, minimising false detection of individual woody pixels and minimising false change in woody classification that would otherwise occur as a result of small changes in the crown cover of isolated pixels. The spatial and temporal rules work together to provide spatial and temporal consistency, minimising temporally varying “mixed pixel” effects (due to spatially varying sampling from independent satellite overpass from year to year) and subsequent error in pixel and change labelling.

This comparative analysis of the same land unit over time was made possible by the accurate and consistent geographic registration and spectral calibration of the image sequences, providing the ability to ‘drill’ through time on a pixel-by-pixel basis. Geographic registration ensures that the same pixel is being looked at through the time sequence. It also avoids incorrect change status determination due to substitution of neighbouring pixels that could have different forest cover status, relative to the correct pixel for that location. Spectral inconsistency can also potentially increase the area attributed to clearing and regrowth events by variable status determination due to image calibration difference. This is addressed by consistent (spectral) calibration, thereby preventing the identification of false clearing or regrowth events and results in a more accurate land cover change map. Consistent registration and calibration are both required to ensure robust multi-temporal change analyses.

The CPN allows areas of missing data, such as those due to cloud cover in the Landsat imagery, to be filled in based on the cover status of the earlier and later images (see Figure 6.A.6). With the advent of optimal cloud free image selection from the Data Cube, the amount of missing data is reduced. However gap filling is still necessary in places due to imperfect automated cloud masks and the lack of available data for locations that are inherently cloudy.

There is also potential for sub-pixel shifts to change the forest/non-forest status on the edges of forest systems where a small edge portion of the pixel may have previously been just over the forest area, but a small shift in geographical registration (e.g., 10 m) would be enough to move the pixel out of the forest area. The spatial rules take the status of adjoining pixels into account and so reduce false change in isolated and edge woody pixels.



Figure 6.A.6 Images of forest extent and change, showing how the CPN gap-fills missing data due to cloudy imagery

Forest extent and change analysis

Once the change in forest cover status has been determined for each pixel for a point in time, the spatial relationship of each change pixel to other surrounding or nearby change pixels is assessed to identify isolated pixels with forest cover that do not form part of a forest system. This allows for the identification of pixels that are isolated trees not meeting the minimum canopy criterion defining a forest, as opposed to those pixels that may be part of sparse linear features such as roadsides and riparian zones which do meet the canopy criterion. A minimum mapping unit filter is applied to remove the isolated pixels from the data to be used for attribution.

The area of land cover change is determined as the sum of the changed pixels through time. This approach minimises inclusion of pixels that represent gaps in the forest canopy. An independent study which looked at the implication of the inclusion or exclusion of forest canopy gaps in this way found that the resultant area estimate could vary significantly between approaches (ERIC, 2001). The approach used only includes the area of forest canopy loss and not 'gaps' in the forest canopy. This provides a much lower estimate of area cleared than specified in clearing permits, which usually define the area bounded by the clearing, including gaps in forest canopy cover. Subsequent carbon stock and emissions estimates are computed consistently with the spatial area calculation method. That is, the carbon stock values should reflect the area under canopy, and are not an average that includes 'gaps' between areas of tree canopy.

Using the 3-class product allows us to identify six types of land cover changes in the landscape, namely:

- non-woody to sparse
- non-woody to forest
- sparse to forest

- sparse to non-woody
- forest to non-woody, and
- forest to sparse

Land cover changes related to forest cover gain and loss are reported as *land converted to forest* and conversions of forest land to other land classifications (sections 6.5, 6.7, 6.9, 6.11 and 6.13), whereas changes in sparse woody cover are reported in the *grassland remaining grassland*, *wetlands remaining wetlands* and *settlements remaining settlements* categories (sections 6.8, 6.10 and 6.12) consistent with the 2006 IPCC guidelines.

Attribution of change

The high resolution automated spatial assessment across the continent identifies land cover change resulting from many causes. For unique identification of conversion to another land use, it is necessary to attribute the change event as either direct human-induced and permanent or due to natural temporary effects or methodological artefacts. Land cover change due to temporary tree dieback, natural dynamics of tree mortality and recruitment, drought and both seasonal and inter-annual variability (causing green 'flushes' of growth with similar spectral signals to regrowth) are also identified and excluded by means of an automated, rule based monitoring system. This monitors the temporary loss of forest cover for x number of years to determine if a permanent change in land use or deforestation has occurred. Qualified technical staff use visual image backdrops such as Landsat, Google Earth™, Planet™ and Sentinel Hub™ to differentiate permanent land use change events from those of temporary forest cover loss events such as harvesting or forest fire.

This attribution is achieved by the development of a series of 'masks' to exclude change due to:

- intermittent water features and irrigation areas that may give a false change signal;
- drought and growth flushes; and,
- terrain illumination.

In each national inventory cycle, the method of attribution is continually updated and improved to increase efficiency and reduce the subjectivity of visual attribution of change.

Plantation typing

To allow for more accurate modelling of emissions and removals from newly established forests (under *Grassland converted to Forest Land*), new plantings (reforestation) identified in the remote sensing imagery are mapped into three classes; native forest (environmental plantings), hardwood plantation and softwood plantation. Plantation forests are those that are identified as being due to deliberate human action and are identified by type (e.g., introduction of non-endemic species), evidence of establishment practices (e.g., rip lines) and planting patterns (e.g., rows and stand geometry). The identification of conversion from non-forest to forest follows the same general approach and same remote sensing data as described above. Plantation classes are identified by discrimination against regionally specific ground data. The method uses an automated spectral discrimination and is described in Caccetta and Chia (2004). Currently, only Landsat TM, ETM+ and OLI data is used for plantation classification. The 3-class method has also been applied to plantation typing.

6.A.4 Quality Assurance and Quality Control

Programme implementation

During the initial implementation of the remote sensing programme, pilot tests were used to train and develop industry capacity, refine methods and software and to develop logistical systems to maximise both output and opportunity for quality assurance and quality control (QA/QC). The results of the pilot studies are published in Furby and Woodgate (2002).

The approach to programme administration provides for centralised progress monitoring and QA/QC at each stage in the processing of the Landsat data. Each processing stage is a regionally defined package of work based on 37 1:1,000,000 (1:1 M) map tiles of Australia (Figure 6.A 1).

The QA/QC and data validation procedures for each of these items in Australia's land cover change methods are summarised below – see also Furby (2002, 2016). Some of the resource intensive processes undertaken in previous years are no longer valid as multiple steps have been integrated and automated. As a result, QA/QC procedures have also been streamlined, resulting in significant savings and efficiency.

Mosaicing

All mosaiced images (quadrants and time slices) for a particular map sheet are assessed at the same time. Due to the automated processing of imagery in the Data Cube, QA/QC of the mosaiced imagery has been streamlined to a single step since NIR 2016. Each data set is checked to ensure completeness and consistency of the composite images (Furby, 2016).

3-class Random Forests classifier

As the use of the Random Forests classifier is a new process in this National Inventory, QA/QC was quite a significant part of the overall implementation. The classifier was run in a semi-automated manner as there are a number of variables that can be tuned to optimize the performance of the classification algorithm. In future, the aim is to fully automate the implementation of the classifier.

Semi-automation allowed QA/QC to be undertaken to investigate a number of elements:

- methods of training sample selection, i.e. using default automated settings versus using modified training samples to remove all omission and commission errors
- use of a more 'typical' base year from which to create training samples, for individual stratification zones
- the use of change masks to exclude areas with a change in spectral index values between 2018 and 2019 from the training sample
- setting of suitable probability thresholds of change within indices, per map sheet and stratification zone
- tracking of the relative importance of individual input variables to probabilities for individual map sheets and stratification zones; and
- monitoring of prediction accuracies per stratification zone.

Undertaking all these investigations led to a greater understanding of how the RF classifier performed, and the impact of certain parameters on the probability predictions. As the choice of training sample data was found to greatly influence the results, this was a major focus of the QA process.

After extensive testing, it was determined that the threshold for inclusion in the training sample should be allowed to vary by class, dependent on the dominant vegetation cover of each map sheet.

CPN products for the current epoch were then compared to the cover class probabilities of previous epochs, to identify the impact from the change in classification methodology. This change has generally resulted in a shift in woody extent and change statistics which has implications for the emission calculations derived from this data. To compensate for the different nature of the 2019 RF probabilities, experiments were performed to adjust the CPN parameters to compensate for the observed shifts and produce a result more consistent with previous updates.

When the probability images have passed assessment and are mosaiced, the resultant images and key intermediate products are assessed for mosaicing accuracy, completeness and standardised formatting.

A final assessment report is completed, detailing the results and whether any further data review is required.

CPN products

When the CPN datasets are supplied to the Department's Geospatial team, they undergo a supplementary QA review process. The purpose of this review is to provide an independent logic check to identify any issues which may have impacts on future geospatial processing and modelling, before there is a significant resource allocation.

The review assesses the following components of the CPN products:

- An initial contents check is conducted to ensure the correct number of CPN dataset components have been supplied per tile.
- Check that designated change transitions between neighbouring epoch woody definitions are logical and correct across the time series on a pixel by pixel basis.
- Ensure that for each tile the CPN dataset's individual components for the time series contain pixel values that are within the acceptable range for that component.
- Check that for each tile the CPN dataset's individual components for the time series have correct spatial extents, geographic projection, pixel resolution and no null pixel entries.
- Produce a summary of percentage difference between the previous NIRs CPN run with the updated CPN run, to determine any variations which would be considered extreme and should be investigated further.
- A sample visual review is undertaken of the distribution of pixel values within the CPN dataset's individual components to ensure they are consistent with the previous NIR and with satellite imagery (e.g., forest classification is consistent with forest shown in associated Landsat imagery for the same year).
- For plant type designations, check they occur over the expected spatial extent when related to the associated forest cover datasets for 1990.

If any issues are found from the above assessment the dataset is returned to the remote sensing specialists for investigation. Only when all aspects of the review are satisfactorily resolved are the CPN datasets available for spatial attribution and FullCAM estimates.

Continuous Improvement and Verification Programme

Periodic review of the CPN products, to ensure human-induced vegetation change is not being omitted, is conducted separately to the NIR. This review is undertaken within a continuous improvement and verification programme (CIVP).

The CPN products identify woody vegetation cover and change, and undergo expert geospatial review using high resolution imagery and external datasets to isolate areas of human-induced change. This attribution of human-induced change is a vital part of each NIR. The ongoing verification programme provides an assessment of the CPN products prior to attribution, while attribution by expert operators ensures that errors of omission and commission related to human-induced clearing and regrowth are minimised in the inventory.

Figure 6.A.7 shows the history of the CIVP and the relevant details for each iteration. CIVP-3 was established as an extension of CIVP-2 in response to an ERT recommendation, to determine the commission and omission errors associated with using the CPN algorithm to assess land cover change.

| Program: | CIVP-1 | CIVP-2 | CIVP-3 | CIVP-4 |
|----------------------------------|----------------------------------|-----------------------------------|-----------------------------------|---|
| Year: | 2004 | 2012 | 2014 | 2017 |
| Coverage: | 37 tiles | 19 tiles | 19 tiles | 11 tiles |
| Number of points | 12,564 | 7,680 | 1,214 | 4,520 |
| Time series: | 1972–2000 | 2002–2010 | 2001–2012 | 2011–2014 |
| Products assessed: | Forest & non-forest | Forest & non-forest | Change product only | Forest, sparse & non-woody, change products |
| Resources used for verification: | Aerial photos, satellite imagery | High resolution satellite imagery | High resolution satellite imagery | Very high resolution satellite imagery |

Figure 6.A.7 The series of continuous improvement and verification programmes

For CIVP-4 the new CPN 3-class woody vegetation product (forest, sparse and non-woody) was assessed across 11 tiles that contribute the most emissions to the national inventory, to determine the accuracy of the product and to identify areas for improvement. The method established during CIVP-2 was followed in CIVP-4, where 400 points were created across each tile using a stratified random sample. The vegetation classification at each point was cross-tabulated against the visual assessment of vegetation type undertaken by experienced operators using very high resolution satellite imagery (see table 6.A.2).

At points where the CPN identified change in vegetation cover between 2011–2014, an assessment of the likelihood of change during that period was also undertaken. As the CPN algorithm uses data from earlier and later years to determine vegetation change for each pixel, the time period for assessment of change in CIVP-4 was selected to ensure the change classification had stabilized using data from later years. In the latest assessment, the CPN land cover change product was verified using very high resolution satellite imagery acquired between 2009 and 2014. Imagery earlier than 2011 was consulted in case there was a lag between change being detected by the CPN in 2011 and change occurring prior to that year.

Of the 4520 points assessed across 11 tiles, 88 per cent had experienced no change (NC) across the time period. Based on the CPN classification, these points were identified as forest throughout (FT), sparse throughout (SPT), or non-woody throughout (NWT). The operator determined if these classifications were definitely correct, or probably correct, if imagery was not clear or not available at the right time. Probably non-woody throughout was not assessed as this category was considered to be difficult to distinguish from probably sparse. Table 6.A.2 shows the CPN product identified forest and non-woody areas consistently better than the identification of sparse vegetation. Commission errors indicate where the classification is deemed incorrect, while omission errors are where points should have been given the classification but weren't.

| Verification | CPN classification | | | |
|---------------------|---------------------------|------------------|---------------------------|-------------------------|
| | Number of points | % correct | % Commission error | % Omission error |
| Forest | 1546 | 98 | 2 | 2 |
| Sparse | 685 | 66 | 24 | 13 |
| Non-woody | 1722 | 96 | 6 | 4 |

Table 6.A.2 CIVP-4 verification results for the 3-class woody vegetation product where no change was indicated

As sparse was a new class of woody vegetation and due to the difficulties detecting it remotely using medium resolution data, it was expected that the errors would be moderate. Despite these errors, the

3-class product has improved the prediction of woody and non-woody vegetation when compared to the previous forest and non-forest classes. Forest was predicted as correct for 96 per cent of the points in CIVP-2 compared to 98 per cent in CIVP-4, while non-forest was definitely correct 76 per cent of the time for CIVP-2 compared to 96 per cent for CIVP-4 (Lowell et al. 2012). Point data records from the verification programme could be used as extra sites to train the CPN algorithm and further improve the woody vegetation product.

The results for the points that had experienced change during 2011–2014 are shown in table 6.A.3, with the number of sample points for each classification cross-tabulated against the operators' assessment. Green cells indicate correct detection of change or no change (NC), red cells are erroneously detected change, lavender cells are undetected deforestation and blue cells are undetected regeneration. Of the points where the CPN had identified change ($n = 550$), 26 per cent were classified by the CPN as deforestation (DEF), 63 per cent were regeneration (REG) and 11 per cent indicated cyclic change (CYC). In this report DEF and REG refer to all cleared or regeneration pixels as indicated by imagery and associated processing. This is not to be confused with deforestation as used in the Kyoto Protocol that specifically refers to human-induced land conversion. A small number of points were uncertain (U) due to poor imagery available to confirm the classification. Pixels classified as CYC suggest errors in the classification given that rapid change, such as forest to non-woody and back to forest, is unlikely to occur over such a short time.

It is imperative that errors of omission related to human-induced change are minimised to give confidence that the inventory has captured all true clearing and regeneration within the given year.

Results of the operator assessment in table 6.A.3 take into account transitions such as forest to sparse and vice versa. For the purpose of this exercise such transitions were included as the verification programme was undertaken to assess the implications of introducing a new sparse category into the vegetation classification and its impact on the change product. Therefore the 71 DEF points shown in the table are inclusive of these transitions which do not reflect vegetation clearing.

The 27 DEF points and 11 REG points that were incorrectly classified by the CPN in table 6.A.3 were subject to further evaluation by additional operators. Initial investigation indicated that 73 per cent of these points had no evidence of clearing or regrowth, however they reflected the classification and operator uncertainty between the forest-sparse and sparse-non-woody decision boundaries

Combined errors of omission for DEF and REG were 0.4 per cent of the total 4520 points, while errors of commission were 7 per cent. These results are comparable to those of previous verification programmes (see table 6.A.4), with 0.3 per cent omission errors over 7680 points and 3 per cent commission errors. The higher commission errors in CIVP-4 are related to the addition of the sparse category into the woody vegetation product, as almost all points incorrectly identified as change had been classified by the CPN as sparse at some time in the change period. Errors may also be partly explained by the smaller sample size in CIVP-4.

The commission error of 7 per cent within the CPN change products identified by CIVP-4 justifies the continuation of the attribution process by geospatial experts to ensure that non-human induced change (i.e. false positive change) does not enter the inventory accounts.

Once the Random Forests classifier has been extended back through the time series, further verification of the 3-class CPN products produced using this new methodology will be undertaken.

“Controls” (of omission errors) and “Attribution” (the final QC step, of attribution of changes identified in cover change maps by the CPN) are discussed in the NIR, but are not presented here in any details. Important to note though is a QC step where:

“Pixel level comparisons were undertaken of woody vegetation loss between the national inventory data and the Queensland Government Department of Environment and Science (DES) vegetation monitoring system. An assessment was made of the level of agreement between the two datasets for the period 1988 to 2018 (see Figure 6.A.9). Using the improved 3-class change data, there is a high level

of agreement (within 10 per cent) between the two systems, although at a few places the clearing pattern does not match.”

“Each area of disagreement was reviewed carefully and the national inventory revised accordingly, where appropriate, using the improved 3-class change product.”

“A similar process was also undertaken using vegetation monitoring data for NSW from 1988 to 2014. All areas identified by NSW Department of Planning, Industry and Environment (DPIE) as cleared in the past were checked to determine if they were already part of the national inventory. This analysis showed a high level of agreement, and areas of disagreement were carefully reviewed and the inventory revised if appropriate.”

6.A.5 Refining the CPN algorithm

To address the errors of commission and omission related to the sparse classification identified in the CPN woody vegetation products (see continuous improvement and verification programme section in 6.A.4), it is necessary to refine the CPN algorithm.

Since the publication of the 2016 National Inventory Report, the Department has undertaken fieldwork to collect woody vegetation data using a LiDAR (light detection and ranging) drone and optical sensors over national parks in the Bourke region of NSW. The vegetation in this area is difficult to classify as the landscape is highly modified through clearing and grazing, vegetation responds to climatic cycles such as drought, and high resolution imagery is not always available. There are also numerous ERF projects in the area where human-induced revegetation is occurring and being monitored using the woody vegetation data.

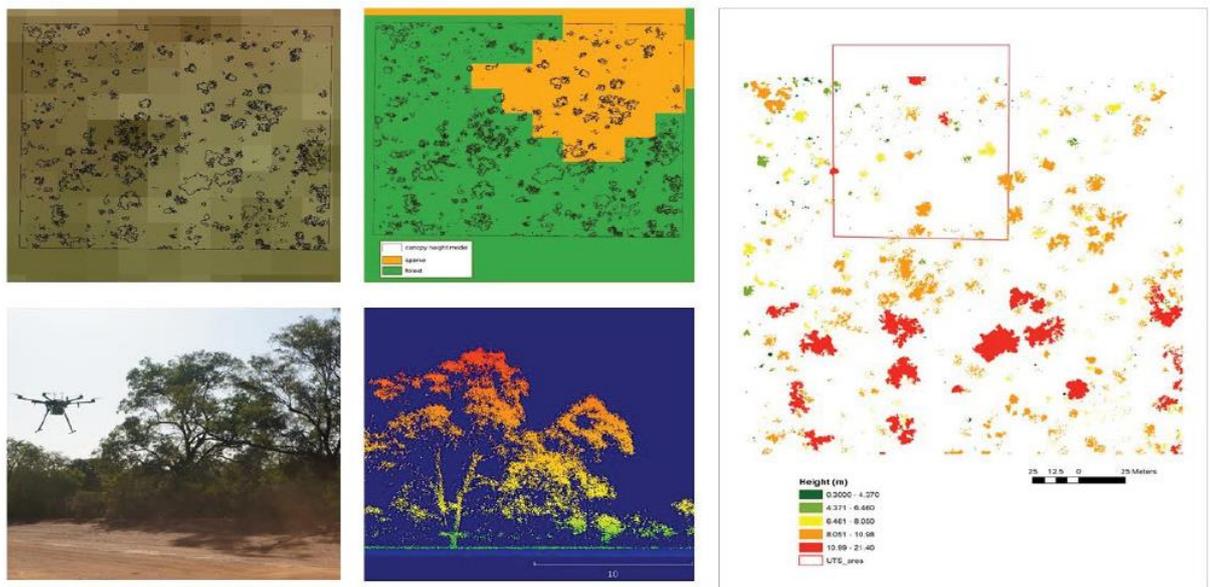


Figure 6.A.12 Examples of outputs from LiDAR drone analysis

Assessment

Assessment:

- Land sector reporting within Australia’s National Inventory System integrates a wide range of spatially referenced data through a process based empirical model (Tier 3) to estimate carbon stock change and greenhouse gas emissions at fine spatial and temporal scales
- Empirically constrained, mass balance, carbon cycling ecosystem model (FullCAM)
- The principal monitoring system is a remote sensing programme used to identify forest lands and changes in forest cover

- These maps are able to detect fine scale changes in forest cover at a 25 m by 25 m resolution
- Supplementary spatial information from the Land Use Mapping programme of Australia's Bureau of Agricultural Resource Economics and Sciences
- Appendix 6A of the 2020 NIR contains most of the information relevant to this review. The information provided is **very detailed** and shows considered investment to develop image processing capabilities
- In all, three references to technical reports regarding the development and testing of remote sensing image data with Furby as an author are cited. Access to these reports is not easy and therefore hinders transparency, but, this limitation is partially offset by high levels of detail in the NIR 2020.
- The Australian 2020 NIR provides more than 20 pages of detail about the image processing and land use tracking approaches used. This provides good transparency in the reporting.
- Landsat image data are used.
- Machine learning algorithms for change detection are used in satellite data image processing.
- Since 2004 imagery has been delivered on an annual basis
- Selection of suitable Landsat scenes from the Data Cube is fully automated. For a given location, the season from which the scene should be selected is identified and the best (cloud-free) image is automatically allocated from the stack within the Data Cube
- Mosaicing: Scene selection and compositing is automated so multiple images can be combined within each path/row to create a cloud free composite
- A spatial-temporal model (Conditional Probability Network, or CPN) is used to reduce the effect of "mixed" isolated and edge pixels in the overlap period. The ability to determine, from 1988 onwards, the effects of land use change to 0.2 ha minimum areas is robust, given that this area is greater than the pixel resolution and the approach used removes mixed and other pixels which are temporally and spatially inconsistent.
- Programme of work conducted to ensure that the latest Landsat 8 image data are time series consistent and compatible with earlier Landsat data. Mis-registration, if not accounted for, would result in false change being reported
- "3-class" random forest (RF) classifier introduced to ensure the entire range of woody vegetation is monitored for reporting under the Kyoto Protocol second commitment period and the Paris Agreement.
- The Conditional Probability Network (CPN) outputs for 2018 were used as the training sample or "base" to train the RF classifier for the new update. Twenty percent of this data is extracted randomly and reserved to calculate an independent accuracy assessment.
- The land cover change programme uses Conditional Probability Network (CPN) analysis to strengthen confidence in the 'woody', 'sparse woody' and 'non-woody' classification of a pixel (previously 'forest' or 'non-forest').
- The CPN allows areas of missing data, such as those due to cloud cover in the Landsat imagery, to be filled in based on the cover status of the earlier and later images (see **Figure 6.A.6**)
- Forest extent and change analysis: Once the change in forest cover status has been determined for each pixel for a point in time, the spatial relationship of each change pixel to other surrounding or nearby change pixels is assessed to identify isolated pixels with forest cover that do not form part of a forest system. This allows for the identification of pixels that are isolated trees not meeting the minimum canopy criterion

- An independent study which looked at the implication of the inclusion or exclusion of forest canopy gaps in this way found that the resultant area estimate could vary significantly between approaches
- Attribution of change: it is necessary to attribute the change event as either direct human-induced and permanent or due to natural temporary effects or methodological artefacts. This attribution is achieved by the development of a series of 'masks' to exclude change due to: intermittent water features, drought and growth flushes, terrain illumination

Specific comments on QA/QC

Specific comments on QA/QC

Extensive QA/QC procedures

- **Programme implementation:** During the initial implementation of the remote sensing programme, pilot tests were used to train and develop industry capacity, refine methods and software and to develop logistical systems to maximise both output and opportunity for quality assurance and quality control (QA/QC). The results of the pilot studies are published in Furby and Woodgate (2002).
- **Mosaicing:** Due to the automated processing of imagery in the Data Cube, QA/QC of the mosaiced imagery has been streamlined to a single step since NIR 2016. Each data set is checked to ensure completeness and consistency of the composite images
- **3-class Random Forests classifier:** A wide range of test conducted to impact of certain parameters on the probability predictions
- **Conditional Probability Network products:** When the CPN datasets are supplied to the Department's Geospatial team, they undergo a supplementary QA review process. The purpose of this review is to provide an independent logic check to identify any issues which may have impacts on future geospatial processing and modelling, before there is a significant resource allocation
- **Continuous Improvement and Verification Programme:** Periodic review of the CPN products, to ensure human-induced vegetation change is not being omitted, is conducted separately to the NIR. This review is undertaken within a continuous improvement and verification programme (CIVP).
- **"Controls" (of omission errors) and "Attribution"** (the final QC step, of attribution of changes identified in cover change maps by the CPN): Pixel level comparisons were undertaken of woody vegetation loss between the national inventory data and the Queensland Government Department of Environment and Science (DES) vegetation monitoring system
- **Refining the CPN algorithm:** fieldwork to collect woody vegetation data using a LiDAR (light detection and ranging) drone and optical sensors over national parks in the Bourke region of NSW; see Figure 6.A.12

12.5.1.4 Non-forest sector

The NIR is issues in three parts. Information in the NIR relevant to this study includes:

- **Volume 2 of the NIR:** This volume provides details of the LULUCF methodology. Appendix 6.A.2 "Monitoring change with remote sensing imagery.
- **Volume 3 of the NIR:** This volume provides methodological detail about the calculations under the KP in Chapter 11. "Kyoto Protocol LULUCF".

Volume 1 of the NIR is not directly relevant to this study.

Appendix 6 A “Land cover change”, in volume 2, is almost exclusively concerned with the approaches for modelling forest carbon. We have only found one clear reference made to non-forest sectors, concerning an update to the method used to identify the location of settlements. The NIR states that the dataset was derived from the 2017 ABARES catchment scale land use data, unpublished sources and visual assessment of high-resolution imagery.

The text on the updates to the settlements dataset states:

“One of the land use categories required by the IPCC 2006 Guidelines is the location of human settlements, and the transitions that occur between settlements and other land use categories. For the National Inventory Report, settlements include areas of residential and industrial infrastructure, including cities, towns, and transport networks (within settlements).

An updated settlements layer was incorporated in the latest NIR to take account of the expansion in settlement areas that have occurred since the last update in 2014 (see figure 6.A.11). The dataset was derived from the 2017 ABARES catchment scale land use data, unpublished sources and visual assessment of high resolution imagery.

The updated settlement dataset was added as a base land use layer for FullCAM spatial simulations. This will allow modelling of emissions and reporting of land conversions such as grasslands or croplands converted to settlements, which is one of the ERT recommendations. Further work is planned to develop a time series of base land use data for all IPCC land use categories.”

Section 6.A.2 “Monitoring change with remote sensing imagery” provides the details about how remote sensing is used to identify land use and track land use change in the inventory. However, we have not found direct references in this section to the approaches used to identify and track non-forest

In Volume 3 of the NIR, information about provided about tracking land use change in cropland management and grazing land management.

Section 11.7.1 “Identification of land subject to **cropland management**” states that Landsat satellite data is used:

“Forest land converted to cropland from 1990 to the inventory year 2018 is identified based on attribution of the Landsat time series and is included under deforestation. Forest land converted to cropland prior to 1990 is identified based on attribution of the Landsat time series and is included under cropland management.

Land converted to forest land, or land that is identified as forest land from the Landsat series, is excluded from croplands.”

The QA/QC arrangements for cropland management are discussed in section 11.7.5 “Quality Assurance – Quality Control”, which refers back to Chapter 6 of Volume 2 (section 6.6.4).

Section 11.8.1 “Identification of land subject to **grazing land management**” states that Landsat satellite data is used:

Grazing land management lands includes grasslands, grasslands with sparse woody cover, and certain specified lands with forest cover – limited to situations in which the presence of grassland has been observed from the Landsat time series and where there has been no change in land use since 1990; or where burning takes place.

The QA/QC arrangements for grazing management are discussed in section 11.8.5 “Quality Assurance – Quality Control”, which refers to Chapter 6.8 of Volume 2.

Assessment

Assessment: It seemed surprising that there was almost no mention of the approaches used to identify land use and track land use change in the non-forest sectors in the main methodological sections of the NIR. The section of the NIR that presents the KP methodologies explicitly states that Landsat data are used for identifying cropland and grassland.

Our view is that it is very likely that Australia uses Landsat satellite data to identify land use, and track land use data in the non-forest sectors. Landsat data are normally not available for every calendar year, and it is also very likely that gap filling is needed – and that considerable work is needed to create time series consistent data.

The Australian NIR is large and it is just possible that there are references to the activity data used to identify and track land use in the non-forest sectors that we did not find in this study. Assuming that this information is not presented in the NIR, or perhaps is very difficult to find, the transparency of the NIR would be enhanced if it explicitly stated how land use identification and tracking in the non-forest sectors was done.

Australia use the same approach to Canada to explain the QA/QC approaches used, and cross references back to earlier sections of the NIR. This saves space in the report, at the expense of reduced transparency.

References

Australia (2020a). *National Inventory Report 1990–2018: Greenhouse gas sources and sinks in Canada. Volume 1.* Australia’s submission to the United Nations Framework Convention on Climate Change. Australian Government Department of Industry, Science, Energy and Resources. <https://unfccc.int/documents/228017>

Australia (2020b). *National Inventory Report 1990–2018: Greenhouse gas sources and sinks in Canada. Volume 2.* Australia’s submission to the United Nations Framework Convention on Climate Change. Australian Government Department of Industry, Science, Energy and Resources. <https://unfccc.int/documents/228017>

Australia (2020c). *National Inventory Report 1990–2018: Greenhouse gas sources and sinks in Canada. Volume 3.* Australia’s submission to the United Nations Framework Convention on Climate Change. Australian Government Department of Industry, Science, Energy and Resources. <https://unfccc.int/documents/228017>

6.4.3 France

The UK had a LULUCF bilateral exchange with France from Wednesday 29th to Friday 31st August, 2018. The meeting was held at BEIS conference centre, 1 Victoria St, Westminster, London SW1E 5ND.

Attending in person: **CITEPA (France):** Etienne Mathias; **UK BEIS:** Peter Coleman; Pierre Brender; **UK Forest Research:** Paul Henshall; **Ricardo:** John Watterson; Glen Thistlethwaite

Dialling in: **CEH (Edinburgh)** Heath Malcolm, Amanda Thomson, Gwen Buys, Hannah Clilverd

The focus of the meeting was a discussion about the latest UNFCCC GHG inventory review reports for the UK and France – and what lessons and best practice could be shared to implement improvements based on the reviewer comments.

Identification and tracking of land use change was considered during one of the presentations from CITEPA. The slides below are reproduced from one of the presentations given by CITEPA.

In the French inventory, a combination of approaches is used to identify land use and track land use change. Remote sensing data have been used since 2006; see slides below.

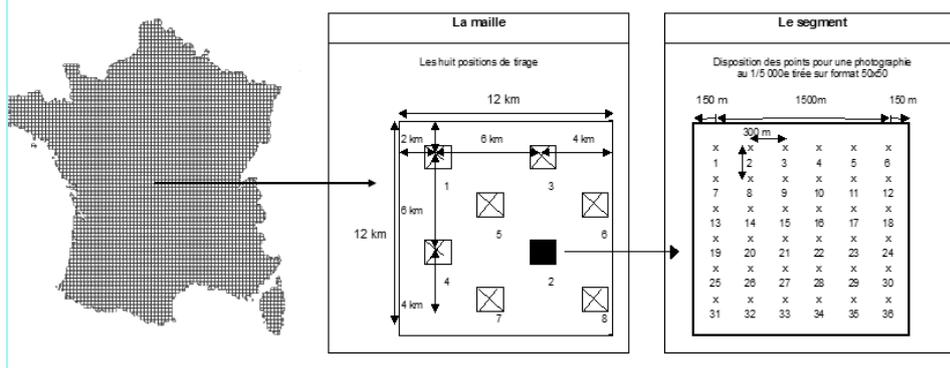
TERUTI surveys

- An annual area frame sampling survey.
- A two-stage systematic sampling design :
 - PSUs (Primary Sample Unit) of 324 ha
 - SSUs (Secondary Sample Unit) of 9 m²
- A double observation (2 nomenclatures) : for each SSU, the surveyor has to determine, on the spot :
 - land cover : wetland, broadleaved forest, wheat, roads ...
 - land use : agriculture, forestry, residential, leisure ...
- Different geographical basis:
 - Remote sensing data (since 2006)
 - Aerial photographs (before 2005).
 - Field documents are used since 1962.
 - Agriculture registry (since 2012)

The French have a comprehensive sampling approach to determine land use and track land use change; see slides below.

TERUTI series

- Different periods :
 - 1981 - 1988
 - 1992 - 2004
 - 2006 - 2010
 - 2012 - 2014



Also, there were some conclusions from the meeting that we believe are relevant for this study, and these are listed below.

1. **Further verification of the forest carbon modelling remains a high priority.** France experience supported the fact that this is not straightforward, and we learned that often they make more

general assumptions (“we consider it in equilibrium and assume no net emissions or sink” – seemed quite common), whereas in UK we seem to aim to model every individual aspect. Does this lead to lowest uncertainties or a risk of bias?

2. **The UK should ensure the NIR is pretty “self contained”**. France have tried separate stand alone reports to supplement the NIR - had lots of ERT findings – and have reverted to putting everything into the NIR.
3. UK approach to carbon stock change in soil in forests SCOTIA modelling is much more advanced than the equilibrium assumption used in France – **but little data to verify. Develop UK verification approach and refine.**

Assessment

Assessment: Verification increases confidence in a GHG inventory, but, it can be hard to select appropriate verification methodologies and they can take time to implement. However, the benefits are worth the effort invested.

The experience of France when it tried to use “stand alone” reports to mainly provide all the detail of the methodologies used in the LULUCF sector was “not positive”; the message for the UK is that it should ensure that the NIR is sufficiently “self-contained”. This will reduce the likely burden of questions from the UNFCCC review team, and, increase transparency of reporting.

Acknowledgements

Landsat data were provided courtesy of the U.S. Geological Survey.

7 References

- Breiman, L. Random forests. *Mach. Learn.* 2001, 45, 5–32
- Brown et al. (2021) UK Greenhouse Gas Inventory, 1990 to 2019. Annual Report for Submission under the Framework Convention on Climate Change. https://naei.beis.gov.uk/reports/reports?report_id=998
- Carrasco, L., O’Neil, A.W., Morton, R.D. and Rowland, C.S., (2019) Evaluating Combinations of Temporally Aggregated Sentinel-1, Sentinel-2, And Landsat 8 For Land Cover Mapping with Google Earth Engine, *Remote Sensing*, 11(3), 288, <https://doi.org/10.3390/rs11030288>
- GEE (2021) https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LT04_C01_T1_SR
- GEE (2021b) https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LT05_C01_T1_SR
- GEE (2021c) https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LE07_C01_T1_SR
- GEE (2021d) https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LE07_C01_T1_SR
- GEE (2021e) https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 202, 18–27. doi:10.1016/j.rse.2017.06.031
- IPCC (2006) Guidelines for National Greenhouse Gas Inventories: Agriculture, Forestry and Other Land Use
- Masek, J., M.A. Wulder, B.L. Markham, J. McCorkel, C.J. Crawford, J. Storey, and D.T. Jenstrom. (2020). Landsat 9: empowering open science and applications through continuity. *Remote Sensing of Environment* 248. doi:10.1016/j.rse.2020.111968
- Moore, M.R.V., Morris, D.G., Flavin, R.W., (2000) CEH digital river network of Great Britain (1:50,000) <https://catalogue.ceh.ac.uk/documents/7d5e42b6-7729-46c8-99e9-f9e4efddde1d>
- Rowland, C.S.; Morton, R.D.; Carrasco, L.; McShane, G.; O’Neil, A.W.; Wood, C.M. (2017a) Land Cover Map 2015 (vector, GB). NERC Environmental Information Data Centre. <https://doi.org/10.5285/6c6c9203-7333-4d96-88ab-78925e7a4e73>
- Rowland, C.S.; Morton, R.D.; Carrasco, L.; McShane, G.; O’Neil, A.W.; Wood, C.M. (2017b). Land Cover Map 2015 (vector, N. Ireland). NERC Environmental Information Data Centre. <https://doi.org/10.5285/60764028-adeb-4316-987a-14b3b21a8f9a>
- Rowland, CS., Morton, RD., Carrasco, L., & O’Neil, A. (2017c) Applying Earth Observation to assess UK Land Use change: Lot 2 Final Report, Final Report for BEIS Ref. TRN 1235/11/2016, Department for Business, Energy and Industrial Strategy, London, 48pp.

Rowland, C.S.; Marston, C.G.; Morton, R.D.; O'Neil, A.W. (2020a). Land Cover Map 1990 (vector, GB). NERC Environmental Information Data Centre. <https://doi.org/10.5285/304a7a40-1388-49f5-b3ac-709129406399>

Rowland, C.S.; Marston, C.G.; Morton, R.D.; O'Neil, A.W. (2020b). Land Cover Map 1990 (vector, N. Ireland). NERC Environmental Information Data Centre. <https://doi.org/10.5285/d6a3588b-23a8-4715-88e9-e21ab0060727>

Storey, J.C., Rengarajan, R., Choate, M.J. Bundle Adjustment Using Space-Based Triangulation Method for Improving the Landsat Global Ground Reference. *Remote Sens.* 2019, 11, 1640. <https://doi.org/10.3390/rs11141640>

USGS/EROS (2018) LSDS-1656 Landsat Collection 1 Level 1 Product Definition, Version 2, April 2018, Online at: <https://www.usgs.gov/media/files/landsat-collection-1-level-1-product-definition>

Wulder, M.A., Loveland, T.R., Roy, D.P., Crawford, C.J., Masek, J.G., Woodcock, C.E., Allen, R.G., Anderson, M.C., Belward, A.S., Cohen, W.B., Dwyer, J., Erb, A., Gao, F., Griffiths, P., Helder, D., Hermosilla, T., Hipple, J.D., Hostert, P., Hughes, M.J., Huntington, J., Johnson, D.M., Kennedy, R., Kilic, A., Li, Z., Lymburner, L., McCorkel, J., Pahlevan, N., Scambos, T.A., Schaaf, C., Schott, J.R., Sheng, Y., Storey, J., Vermote, E., Vogelmann, J., White, J.C., Wynne, R.H., Zhu, Z., Current status of Landsat program, science, and applications, *Remote Sensing of Environment*, 225 (2019), pp. 127-147, 10.1016/j.rse.2019.02.015

Appendix 1: Development of a 1990 Woodland Map for the BEIS Land Use Project

1 Introduction

As part of the BEIS Land Use project, Forest Research (FR) were tasked with developing a model to produce a woodland map for England, Wales and Scotland from 1990, which would help in tracking how land use has changed over time. This document aims to explain the methodology of the model used to develop this 1990 woodland map acting as both a technical and user guide. The scope of the model is to identify 'large' areas of woodland not identified within the existing woodland map and manually verify those areas found ensuring cartographic standards. The model is restricted to areas of omission and not areas of commission (ie will not consider regions included in the 1990 woodland map which were not at that time treed).

2 Definitions

Existing woodland map: from here on will be termed the 'interim woodland map'.

Large: Large areas of woodland as termed in the scope document will include areas greater or equal to 0.5 hectares. However, manual verification will only take place on areas greater or equal to 1.5 hectares.

3 Data sources

The primary data sources for this model include the FR interim woodland map and seasonal Landsat imagery from 1990 (plus or minus 1 year), vegetation indices and a digital elevation model (DEM) as explained below. Each of the datasets were processed to include a coordinate reference system (CRS) aligned to GB OS grid (EPSG: 27700).

3.1 Interim woodland map

The interim woodland map is an ESRI shapefile representing woodland areas across England, Wales and Scotland (see *figure A1.1* and *figure A1.2*). The methodology to derive this map is included in a document entitled *Preliminary1990WoodlandMap_Methodology_vc_bd_18112020_Final.doc*.

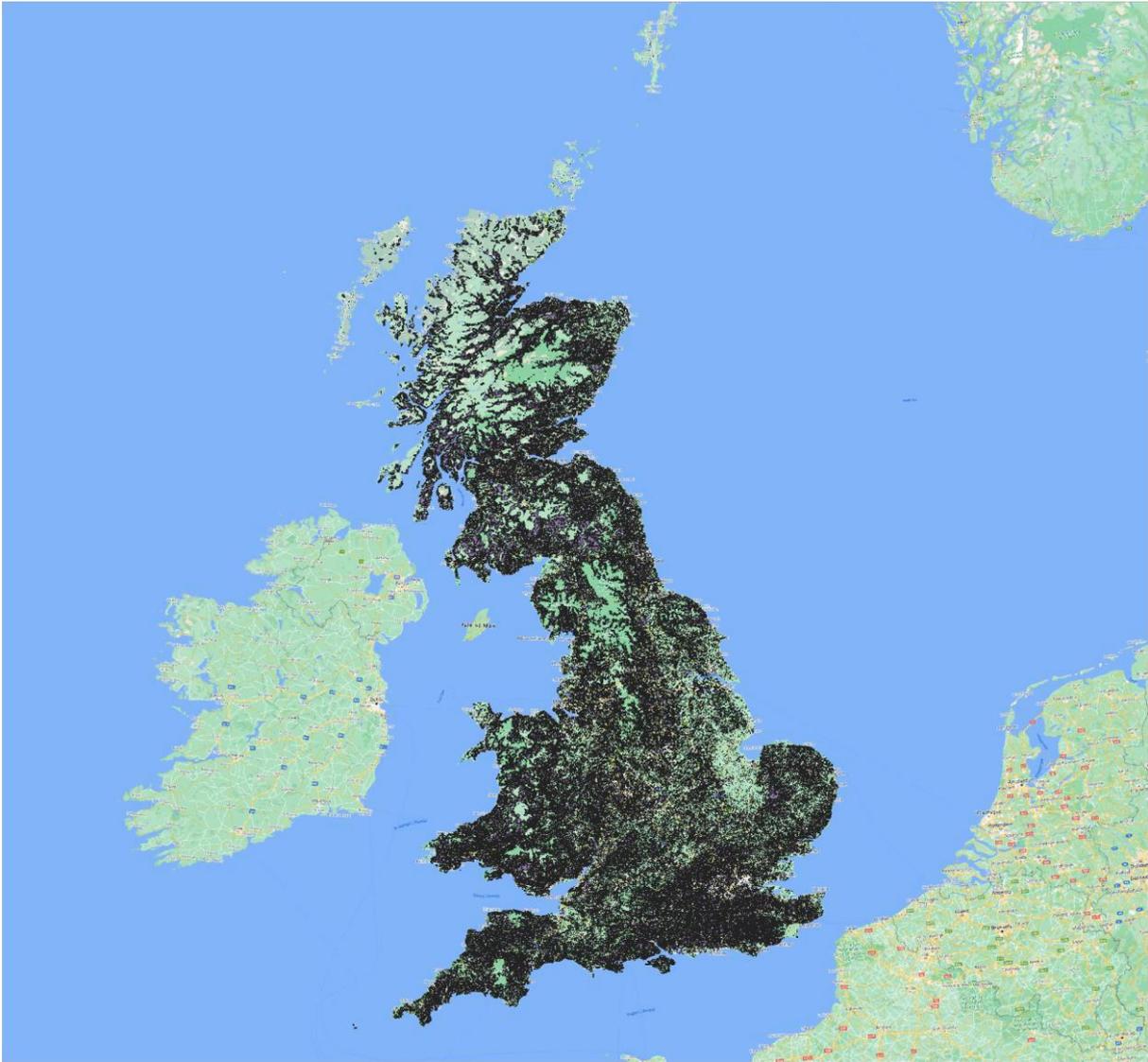


Figure A1.1: ESRI shapefile of interim woodland 1990 map

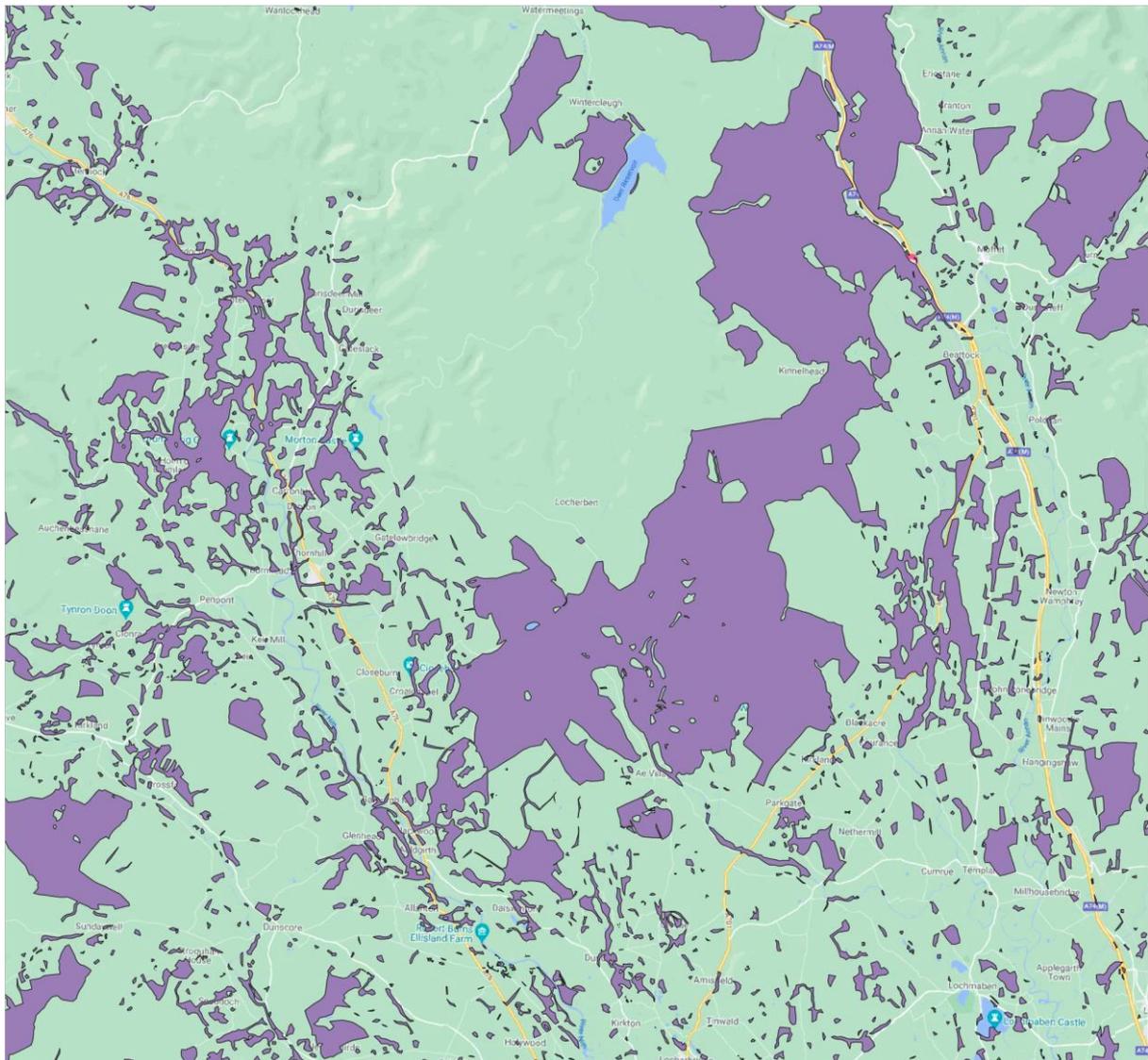


Figure A1.2: Close up of 1990 interim woodland map at 1:100,000

3.2 Landsat imagery

The Landsat imagery was supplied by the Centre for Ecology and Hydrology (CEH) split into England, Wales and Scotland (with some overlap) and three seasonal periods (summer, spring, autumn) as raster .TIF files as detailed in *table A1.1* and the Landsat Thematic Mapper (TM) bands as detailed in *table A1.2*. All images were captured by Landsat 4 or 5 and have a spatial resolution of 30m (see *figures A1.3 to A1.8*).

| Image (File Name) |
|--|
| England_Wales_<autumn/spring/summer>_b1_b3.tif |
| England_Wales_<autumn/spring/summer>_b4_b7.tif |
| Scotland_<autumn/spring/summer>_b1_b3.tif |
| Scotland_<autumn/spring/summer>_b4_b7.tif |
| NI_<autumn/spring/summer>_b1_b3.tif |
| NI_<autumn/spring/summer>_b4_b7.tif |

Table A1.1: Imagery supplied by CEH

| Image Band | Landsat TM Band |
|------------|--|
| Band 1 | B1: Visible - Blue - (0.45 - 0.52 μm) |
| Band 2 | B2: Visible - Green - (0.52 - 0.60 μm) |
| Band 3 | B3: Visible - Red - (0.63 - 0.69 μm) |
| Band 4 | B4: Near Infrared (0.76 - 0.90 μm) |
| Band 5 | B5: Short-wave Infrared (1.55 - 1.75 μm) |
| Band 6 | B7: Mid-Infrared (IR) (2.08 - 2.35 μm) |

Table A1.2: TIFF image bands and corresponding Landsat TM bands

The following details were supplied by CEH regarding image creation:

- Process works by identifying cloud-free pixels at a point and then taking the median value (although there is an option to change the statistics i.e. greenest pixel, mean).
- Cloud masking is applied to each granule separately prior to mosaicking.
- Black areas occur in the image where there is no suitable (cloud-free) data.
- All data are Thematic Mapper images from Landsat 4 and 5.



Figure A1.3: Spring England/Wales Landsat 1990 imagery (RGB)



Figure A1.4: Summer England/Wales Landsat 1990 imagery (RGB)



Figure A1.5: Autumn England/Wales Landsat 1990 imagery (RGB)



Figure A1.6: Spring Scotland Landsat 1990 imagery (RGB)



Figure A1.7: Summer Scotland Landsat 1990 imagery (RGB)



Figure A1.8: Autumn Scotland Landsat 1990 imagery (RGB)

3.3 Unusual light bands within the Landsat imagery

The light band .tif files represent areas of the CEH Landsat imagery where there is a band of visible light that is clearly different (lighter) than areas either side. These are due to the way that the satellite captures the imagery and the mosaicking process. Light bands identified include diagonal bands in the SE, the SW, mid England and SE Scotland (see figures A1.9 and A1.10). These were identified and drawn manually in GIS software and converted to a raster.

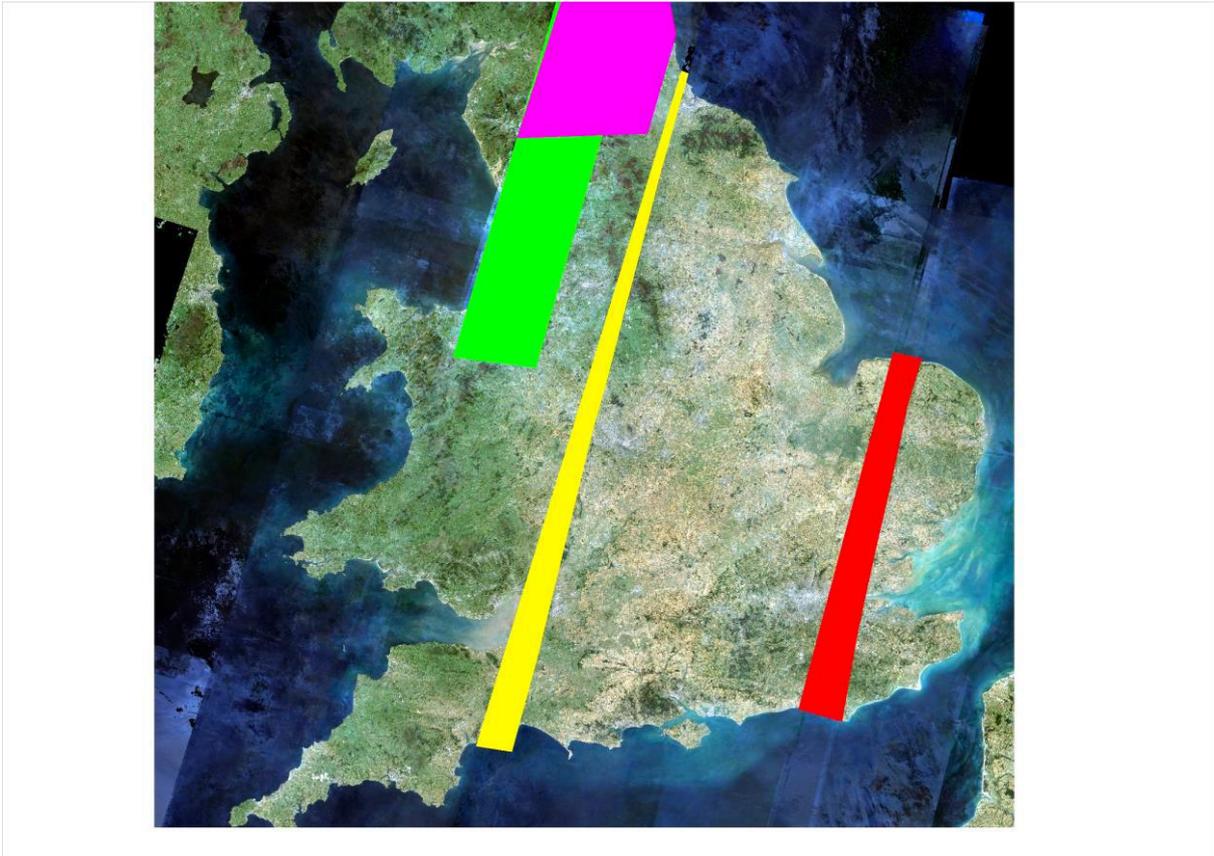


Figure A1.9: Light bands across England_Wales image (Red = SW light band, Yellow = SE light band, Green = Mid light band, Pink = SE Scotland light band)



Figure A1.10: Light bands across Scotland image (Yellow = SE light band, Green = Mid light band, Pink = SE Scotland light band)

3.4 Seasonal vegetation index images

To improve the woodland detection process, the seasonal Landsat images were used to create seasonal vegetation index images. These were developed in separate R code and represent the following indices, where the bands represent spectral bands on Landsat 4 – 7.

- Normalised difference vegetation index (NDVI) = $(B4 - B3) / (B4 + B3)$
- Green normalised difference vegetation index (GNDVI) = $(B4 - B2) / (B4 + B2)$
- Enhanced Vegetation Index (EVI) = $2.5 * ((B4 - B3) / (B4 + 6 * B3 - 7.5 * B1 + 1))$
- Advanced Vegetation Index (AVI) = $[B4 * (1 - B3) * (B4 - B3)]^{1/3}$
- Soil Adjusted Vegetation Index (SAVI) = $((B4 - B3) / (B4 + B3 + 0.5)) * (1.5)$
- Normalized Difference Moisture Index (NDMI) = $(B4 - B5) / (B4 + B5)$
- Moisture Stress Index (MSI) = $B5 / B4$
- Green Coverage Index (GCI) = $(B4 / B2) - 1$
- Bare Soil Index (BSI) = $(B5 + B3) - (B4 + B1) / (B5 + B3) + (B4 + B1)$
- Normalized Difference Water Index (NDWI) = $(B2 - B4) / (B2 + B4)$
- Atmospherically Resistant Vegetation Index (ARVI) = $(B4 - (2 * B3) + B1) / (B4 + (2 * B3) + B1)$
- Structure Insensitive Pigment Index (SIPI) = $(B4 - B1) / (B4 - B3)$

The *Normalized Difference Vegetation Index (NDVI)* is a numerical indicator that uses the red and near-infrared spectral bands. NDVI is highly associated with vegetation content. High NDVI values

correspond to areas that reflect more in the near-infrared spectrum. Higher reflectance in the near-infrared correspond to denser and healthier vegetation (see figures A1.11 and A1.12).

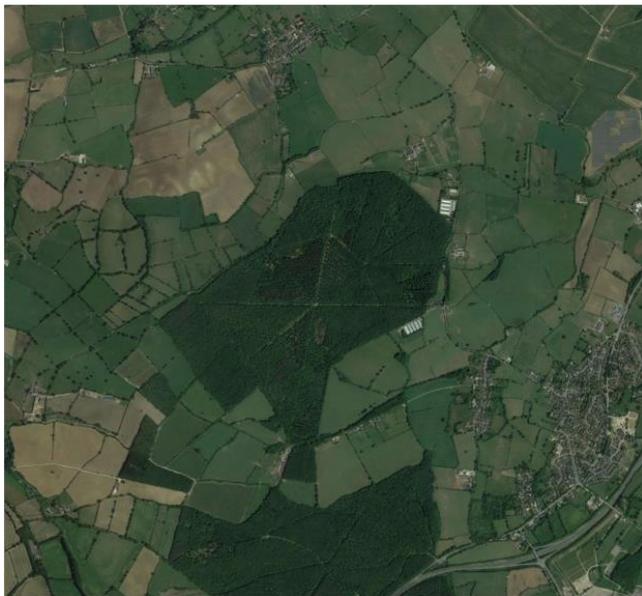
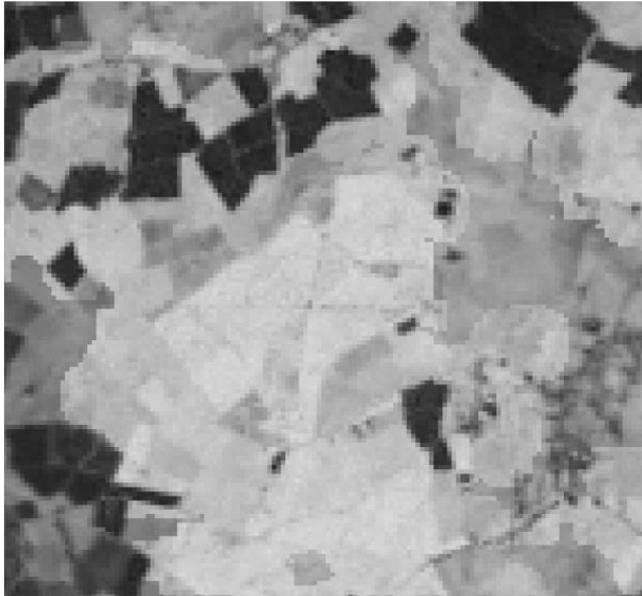


Figure A1.11 and A1.12: Example vegetation index: NDVI where forested areas have higher NDVI values

The *Green Normalized Difference Vegetation Index* (GNDVI) is modified version of NDVI to be more sensitive to the variation of chlorophyll content in the crop.

The *Enhanced Vegetation Index* (EVI) is similar to NDVI and can be used to quantify vegetation greenness. However, EVI corrects for some atmospheric conditions and canopy background noise and is more sensitive in areas with dense vegetation. It incorporates an “L” value to adjust for canopy background, “C” values as coefficients for atmospheric resistance, and values from the blue band. These enhancements allow for index calculation as a ratio between the R and NIR values, while reducing the background noise, atmospheric noise, and saturation in most cases.

Advanced Vegetation Index (AVI) is a numerical indicator, similar to NDVI, that uses the red and near-infrared spectral bands. Like NDVI, AVI is used in vegetation studies to monitor crop and forest variations over time. Through the multi-temporal combination of the AVI and the NDVI, users can discriminate different types of vegetation and extract phenology characteristics/parameters.

Soil Adjusted Vegetation Index (SAVI) is used to correct NDVI for the influence of soil brightness in areas where vegetative cover is low.

Normalized Difference Moisture Index (NDMI) is used to determine vegetation water content.

Moisture Stress Index (MSI) is used for canopy stress analysis, productivity prediction and biophysical modelling. Interpretation of the MSI is inverted relative to other water vegetation indices; thus, higher values of the index indicate greater plant water stress and in inference, less soil moisture content.

In remote sensing, the *Green Chlorophyll Index* (GCI) is used to estimate the content of leaf chlorophyll in various species of plants. The chlorophyll content reflects the physiological state of vegetation; it decreases in stressed plants and can therefore be used as a measurement of plant health.

Bare Soil Index (BSI) is a numerical indicator that combines blue, red, near infrared and short wave infrared spectral bands to capture soil variations. These spectral bands are used in a normalized manner. The short wave infrared and the red spectral bands are used to quantify the soil mineral composition, while the blue and the near infrared spectral bands are used to enhance the presence of vegetation.

Normalize Difference Water Index (NDWI) is use for the water bodies analysis. The index uses Green and Near infra-red bands of remote sensing images. The NDWI can enhance water information efficiently in most cases. It is sensitive to build-up land and result in over-estimated water bodies. The NDWI products can be used in conjunction with NDVI change products to assess context of apparent change areas.

As the name suggests, the *Atmospherically Resistant Vegetation Index* (ARVI) is the first vegetation index, which is relatively prone to atmospheric factors (such as aerosol). The formula of ARVI index invented by Kaufman and Tanré is basically NDVI corrected for atmospheric scattering effects in the red reflectance spectrum by using the measurements in blue wavelengths.

The *Structure Insensitive Pigment Index* (SIPI) is good for analysis of vegetation with the variable canopy structure. It estimates the ratio of carotenoids to chlorophyll: the increased value signals of stressed vegetation.

3.5 Digital Elevation Model

A digital elevation model, a raster representing different heights of the UK, was also included to cap the maximum elevation of woodland for a region to help distinguish areas of woodland and vegetated hilltops that have similar spectral properties. This DEM was obtained from *OpenDEM* and aligned to a GB OS grid reference (see figure A1.13).

3.6 Regional shapefiles

The UK was divided into 50 regional files termed 'checkout areas' (see figure A1.14). Note that the light bands/cross border regions were assigned region numbers 51 to 56.



Figure A1.13: DEM of England, Wales and Scotland

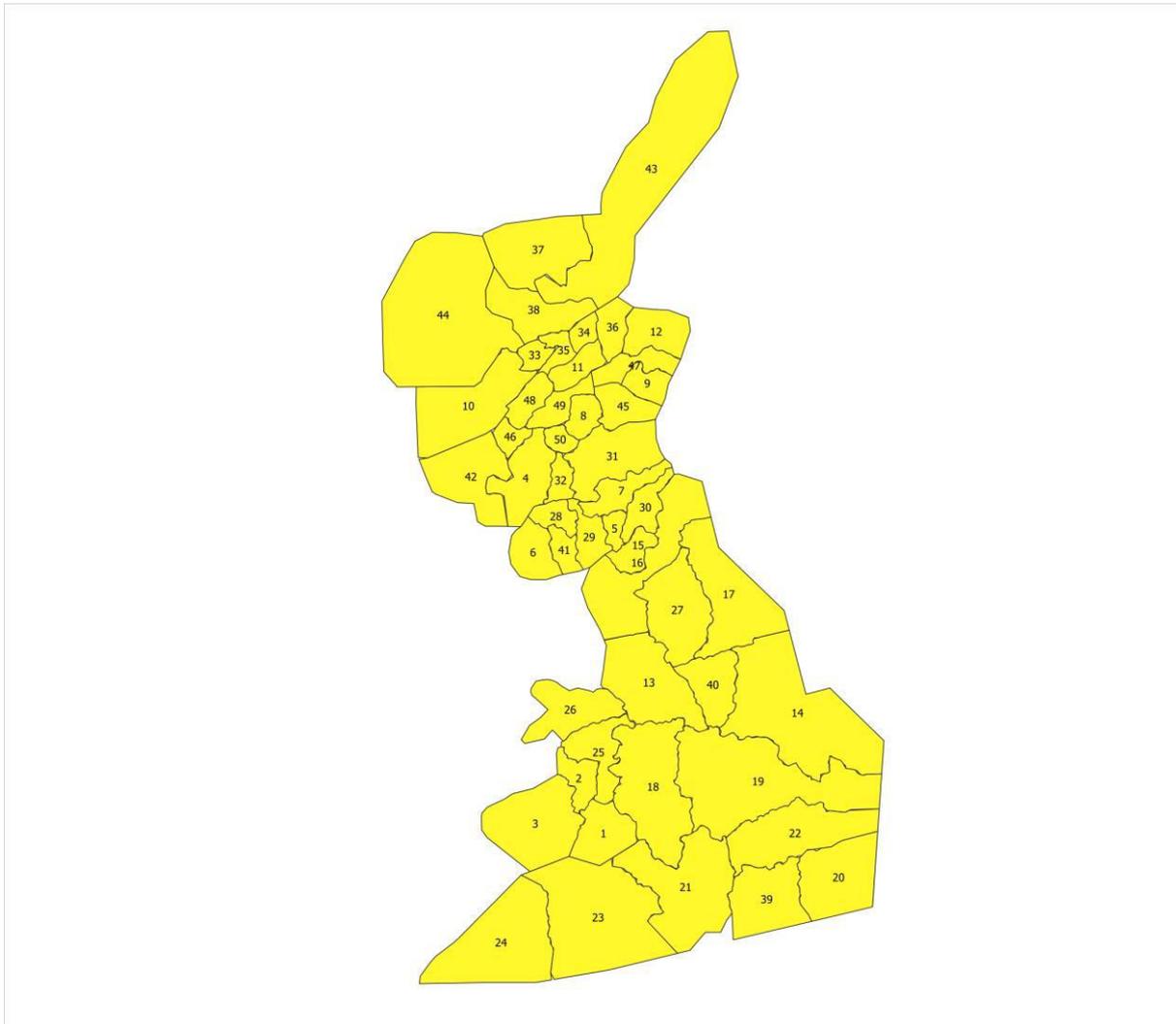


Figure A1.14: Regional shapefile with OBJECT_ID label

4 Algorithm methodology

The data sources previously identified were included in a machine-learning model developed in the R programming language to classify the Landsat images via a per-pixel classification. The code developed represents the following steps:

1) *Load libraries*: The load libraries module loads all pre-requisite R libraries/packages necessary for the model. The model code provides descriptions of each package installed. Each package can fit into one of five broad categories: geospatial data (rgdal, rgeos, gdalUtils, sf, sp and RStoolbox); raster manipulation (raster and terra); parallel computations (parallel, doParallel, foreach); data analysis (data.table, Hmisc); machine learning (caret, nnet, randomForest, ranger, e1071, parallelSVM).

2) *Set model regional and seasonal parameters*: The model was designed to run one region and a combination of two seasons at a time (i.e. 'spring-summer' or 'spring-autumn' or 'summer-autumn'). The advantage of running two seasons at a time it is less likely than three regions to have 'no-data' areas. By including two seasons at a time, it has a higher probability of finding true woodland and reduce the number of false hits.

3) *Run first_phase/combine shapefiles code*: This module simply runs the functions 'first-phase' or 'combine shapefiles'. 'First phase' function requires the region number, country, season 1, season 2,

whether there is a light band that needs to be removed from the region, and the location of that band. Combine shapefiles requires the country and region number.

4) *First phase function*: The first phase function represents the 'first phase' of the algorithm which requires the attributes as detailed in *step 3*. The first step of this function is to load in the relevant raster files necessary for the algorithm as *SpatRaster* objects using the *Terra* package. Using the *Terra* package improves speed and efficiency in raster calculations over the *Raster* package, but not all functions from other libraries recognise *SpatRaster* objects and hence requires conversion from *SpatRaster* to *Raster* objects and reverse throughout the code. Data loading includes light band rasters (if required), the Landsat imagery, the vegetation indices, and the UK DEM. The rasterized interim woodland map was also loaded along with a shapefile (*Spatvector* object) representing all the regions across England, Wales and Scotland (approximately 50). A region number was selected representing the area of interest (AOI).

5) *Crop and mask layers to the AOI*: A list of all the different images is formed (lists contain multiple objects which do not have to be of the same type). The interim woodland map is cropped and masked to the AOI, so only the woodland map within the AOI is left. All the objects within the raster list are then cropped to the extent of the remaining woodland map and converted to a multi-layer *Spatraster* object, with the country wide files removed for memory saving. This *region_rast* file is masked by the AOI leaving a *Spatraster* object with all the different layers with the same shape as the extent of the woodland map.

6) *Remove pixels within the light band*: If a light band is present, the light band areas are removed from the AOI and pixels inside the light band are given a value of NA, and areas outside the light band a value of 1. When the light band is then multiplied by the *Spatraster* object, the pixels within the light band are then given a value of NA and will be dealt with at the end.

7) *Create a multi-layered woodland raster and extract quantile values*: The woodland map is then multiplied by the region *Spatraster* object leaving a *Spatraster* object that only has the pixels that occur within woodland. This is then converted to a *raster brick* object and the 50th percentile of the NDVI (summer or autumn) layer is found, along with the 95th percentile of the DEM layer. All pixels in all layers are set to NA if the NDVI of the summer or autumn layer is less than the 50th percentile mark or the DEM layer is less than the 95th percentile.

8) *Unsupervised clustering and identifying non-woodland*: The *raster brick* object is then entered into an unsupervised clustering machine learning model with 10 clusters and 100,000 samples. This model is then used to predict the cluster for the whole of the *raster brick* object and this cluster layer is converted to a *Spatraster* and written to file. The values of all the layers are converted to a data frame and ordered descending based on the summer or autumn image. Through investigation, it was found that the clusters with the higher summer spectral values (or autumn for the spring/autumn layer) were more likely to be areas of woodland that were not trees (ie gaps in the canopy, tracks, grassland etc). (see *figure A1.15*).



Figure A1.15: Cluster map following unsupervised classification (10 clusters)

9) **Development of a woodland and not_woodland map:** The first four clusters were added to a 'not_woodland_map' and the bottom six clusters added to a 'woodland_map'. The NDVI 50th percentile figure is used to make a layer where NDVI is low which is then merged with the not_woodland layer ie a layer is formed that is either not_woodland as determined by the clustering method or has a low NDVI. This is then multiplied by the regional multi-layered *Spatraster* so as to find the pixel values for all the non_woodland/low NDVI layer. Similarly, a woodland equivalent is made by multiplying the regional *Spatraster* by the woodland_map layer. This enables a binary classifier to be developed. All the values are extracted to a data.table for the woodland map, and a second for the non_woodland areas, and the non_woodland areas filtered further by the 95th percentile of the DEM (ie only areas less than the 95th percentile are kept). The DEM column is then removed from each data table and any rows with NA omitted. A sample of the non_woodland data.table is then taken that matches the same number of rows of the woodland data.table, and a class (as factors) added to each category (0 for non_woodland, 1 for woodland).

10) **Ranger machine learning model:** The data tables are then combined and added to a probability machine learning '*ranger*' model which is a fast C++ implementation of *Random Forest*. This is then used to predict on the *region raster* which are below the DEM 95th percentile and NDVI values above the 50th percentile and then converted back to a *Spatraster* object. The areas within the woodland map

are then removed to leave only the pixels outside woodland. Any areas within a light band are also removed (see figure A1.16).



Figure A1.16: Probability map following ranger machine learning (black = highly probable, white = highly unlikely)

11) **Filtering low probability and extreme values:** Only high probability cells are then kept. The layer values of the remaining pixels are then converted to a data.table and filtered by the 95th percentile DEM. The data.table is then converted to a matrix and the 2.5th percentile and 97.5th percentile values extracted from each layer. The matrix is then filtered to remove values that fall outside the 2.5th to 97.5th values for each layer, and converted back to a data.table. This removes extreme pixels. (See figure A1.17).



Figure A1.17: Probability map outside of the interim map
(red = interim woodland map, black = highly probable, white = highly unlikely)

12) Finding how closely new found woodland matches woodland within the interim woodland map:

A data.table is formed from the first 12 bands of the woodland data.table (ie the Landsat imagery bands), and a similar data.table extracted for each of these bands from the likely woodland areas that are outside woodland. For each row of the outside woodland data.table, the woodland data.table is filtered where band 1 is +/- 10% of the outside woodland data table value and so on for each band. The number of rows left at the end give a weighting to each row within the outside woodland data table. Those with a high weighting are much more probable to be trees than those with a low weighting. Also using the *find.matches* package a number of similar rows within a tolerance are found between the outside woodland and original woodland database. This gives a distance calculation between the outside woodland row and the woodland database. The theory being that those with lots of similar rows are more likely to be trees. This is carried out in parallel to improve speed. The outside woodland data table is filtered so those with a distance ≥ 3.5 are removed (this figure was derived by experimentation).

13) Forming an ESRI shapefile polygons dataset of woodland outside the interim woodland map: A raster layer is formed from the outside woodland data table and assigned the British grid reference CRS. Touching pixels are clumped together (using 8 directions) and filtered so only

those with 5 pixels together or greater are kept (ie \geq half an hectare). This is then formed into a *Spatvector polygons* object and an area in hectares assigned and an ID. A probability figure is then also assigned by finding the mean probability of all the cells in each shape. This is then written to file as an ESRI shapefile (see figure A1.18).



Figure A1.18: ESRI shapefile of woodland found outside of the interim woodland map (red = interim woodland map, yellow = new found woodland \geq 0.5 hectares)

14) **Combining seasonal shapefile:** The combine shapefiles function aggregates the spring-summer, spring-autumn, summer-autumn into one shapefile, with a union between overlapping features. A small buffer is added to each shapefile so they join at corners, and then the buffer is removed. This is then written to file.

15) **Complete for all regions:** The steps above are repeated for each region (where light bands are treated as a separate region).

5 User guide

The software was developed in R with view to be operated/developed by an experienced machine learning/data engineer or scientist and as a one-off product. If it is to be used again, possibly to develop woodland maps in other years other than 1990, it is advisable to operate the code using RStudio. RStudio will recognise missing libraries and offer to install them automatically. All of the files and data sources necessary to operate the code will be installed in the following Forest Research location, but

these data sources should not be referenced directly. Instead, they should be copied to a location suitable and the working directory should be updated in the *setwd* command of '**Set model and regional parameters**'.

When running the code on a region, the country, region_no, season 1 and season 2 variables should be set, and whether there is a light band in the region and the location of that band (note some areas may have multiple light bands).

It is advisable to run the code line-by-line in the first instance to check all data sources are located in the correct place, or at least module by module.

5.1 Location of data sources

All Landsat imagery should be copied into a folder in the working directory as "Images v2" with the format *country_season_bx_by_v2.tif* where x and y refer to the start and end bands of that image e.g. *England_Wales_summer_b1_b3.tif*

All vegetation indices should be copied into a folder in the working directory as "Images v2/NDVI" with the format *country_season_xxxx.tif* where xxx refers to the vegetation index e.g. *England_Wales_summer_NDVI.tif*

The DEM should be copied into a folder in the working directory as "DEM" with the format *DEM_country.tif* e.g. *DEM_England_Wales.tif*

There are four light bands identified in this image and should be contained in the working directory as "Spatial data" given the format *band_location_light_band.tif* e.g. *SE_light_band.tif*. If new data is being used then new light bands may need to be identified and rasterized using GIS software.

The rasterized interim woodland map should be copied into a folder called "Preliminary_1990_WoodlandMap_Nov2020" with the format *country_rasterized_woodland_map.tif* e.g. *England_Wales_rasterized_woodland_map.tif*

The regional shapefile should be copied into a folder called "Spatial data" and given the name *GB_CheckoutAreas.shp*. The code is looking for a field called OBJECTID_1 to distinguish between different areas. Either create that field if it is missing or rename in the R code.

5.2 Further user code refinements

Most of the rest of the code will run without operator interference providing all the data sources are in the right place and the code does not run into memory or processing issues. The user can refine the model in the following areas:

- When the code performs an unsupervised classification, the model picks out the four highest values of the summer green spectral value. If this was increased to five or six then the model would be more discerning removing more false hits (but also possibly some true hits). Similarly, if decreased to three, the model would include more areas as woodland but most likely increase false hits. The user can perform this by adding *as.numeric(ordered_vector[x])*

to the *not_woodland_vector*. It is also important that this is also added to the *woodland_vector*. If one is removed, it should also be removed from the *woodland_vector*.

- Similarly, the code sets a quantile for NDVI as 50th, and DEM as 95th in the *DEM_quantile* and *quantile_season* variables. These could be increased/decreased by modifying

```
quantile_season <- quantile(brick_for_prediction[[nlayer]],probs=c(0.5,0.999));
```

```
DEM_quantile <- quantile(brick_for_prediction[[35]],probs=c(0,0.95))
```

- The user has some flexibility over the probability that is set to filter out *non_woodland* following the *ranger Random Forest* machine learning step. Increasing the probability will remove false hits but also

remove true hits and lowering the probability will capture more forest but also more false hits, by modifying

```
outside_woodland[outside_woodland>0.001] <- NA
```

```
outside_woodland[outside_woodland<=0.001] <- 1
```

- Extreme values can be restricted/increased from 2.5th and 97.5th in the matrix step by modifying `q1 <- quantile(woodland_matrix[,1],c(0.025,0.975))` and for all subsequent qx variables (where x represents 1 to 34).
- The distance figure, currently set to 3.5 can also be increased/decreased to include more/less potential woodland by modifying `outside_dt[outside_dt$distance>=3.5] <- NA`
- If a smaller/larger polygons feature size is required, the line `excludeID <- f$value[which(f$count < 5)]` should be modified by increasing/decreasing 5 (where 5 refers to the number of pixels).

5.3 Verification

To verify the model, 20% of the features found from 37 regions (out of 50) were manually checked against modern day aerial imagery. The 13 areas not checked were due to time limitations on the project and were all north of the central belt in Scotland. A simple assumption was made that if the feature is woodland today (or at the data of the aerial imagery), or felled, or a modern housing development, then it would be assumed to be woodland in 1990. If on the other hand the image was a vegetated fell, or an agricultural field or vegetated coastland etc, it will be assumed to be a false hit although it is possible agricultural fields have replaced some forestry. All of this is based on the fact that woodland is increasing in area, not decreasing and mature trees (particularly broadleaf) have a low probability of removal since 1990.

In total, 8569 features across England, Wales and Southern Scotland were identified as being potentially omitted woodland greater than or equal to 0.5 hectares. Of these, 2783 (32.5%) were \geq 1ha and 5786 (67.5%) were $<$ 1ha. To analyse 20% of these features resulted in a total of 1700 features. 20% of each region were manually identified and given a rating of possible or unlikely.

Possible % varied between regions ranging from 0% and 100%, with a mean value of 59% and a median value of 66%. Across England these values were nearer 75% but much lower across Scotland. This is because vegetated but none treed upland has very similar properties on the course Landsat imagery to treed woodland.

Based on the possible% and number of features, an estimated 4700 hectares were found of woodland outside of the preliminary woodland map. Upscaling this to include all 50 regions could be between 4700 hectares (if no further woodland was found) to 6351 hectares if the same proportion of woodland was found across the northern Scotland regions (although the actual figure is likely to be between the two) (see figure A1.19).



Figure A1.19: Map of newfound woodland from 1990 outside of the interim woodland map

(note: pre validation ie includes false hits)

A decision was made to focus the efforts of the GIS team on areas equal to or above 1.5 hectares to improve the woodland map, and for now store the half hectare to 1.5-hectare regions for a future time.

Examples of true hits are shown in *figure A1.20* and example of false hits shown in *figure A1.21* (often vegetated upland).



Figure A1.20: Example of true hits



Figure A1.21: Example of false hits (often vegetated woodland)

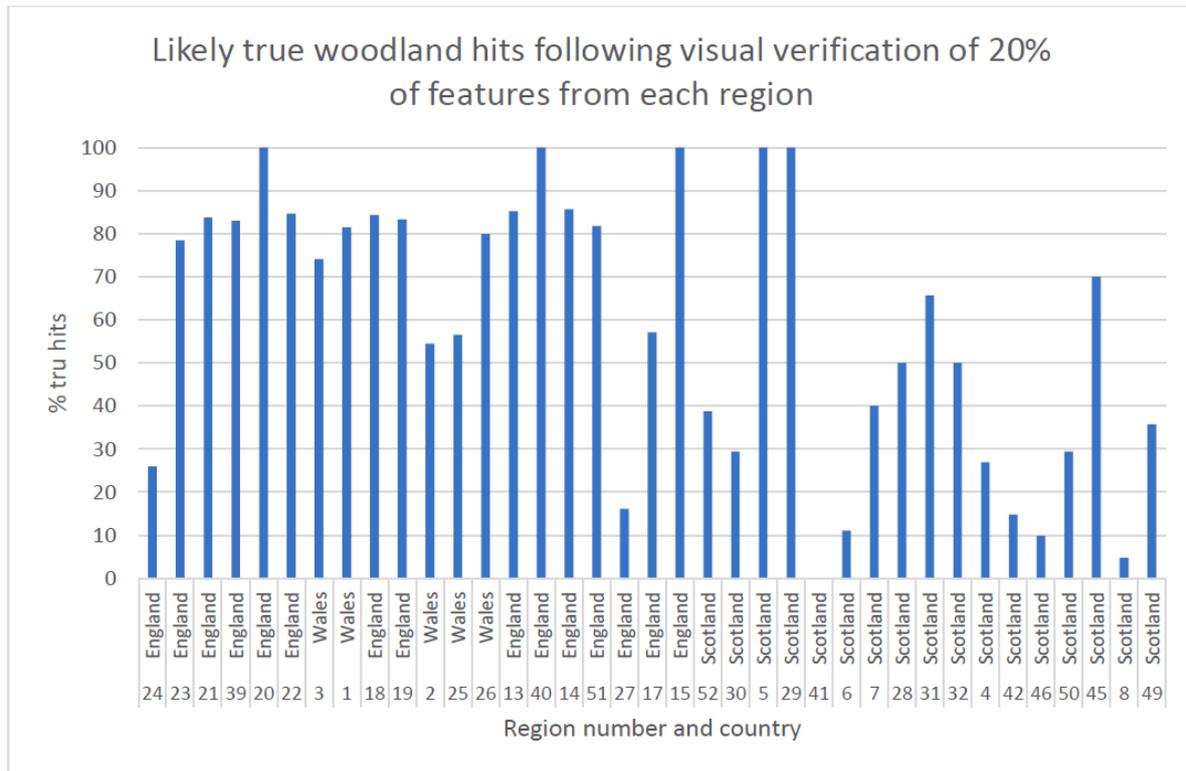


Table A1.3 and *figure* A1.22 shows the % of true hits for each country after visually verifying 20% of all features from each region. Note that this is based on using modern day aerial imagery and if a woodland is present (or clearfell) it is assumed to be a true hit. The estimation of area is based on those areas above 1ha having an average area of 1.5 hectares, and those areas less than 1 hectare having an average area of 0.75 hectares. The estimation is based on the % true hits of this total area.

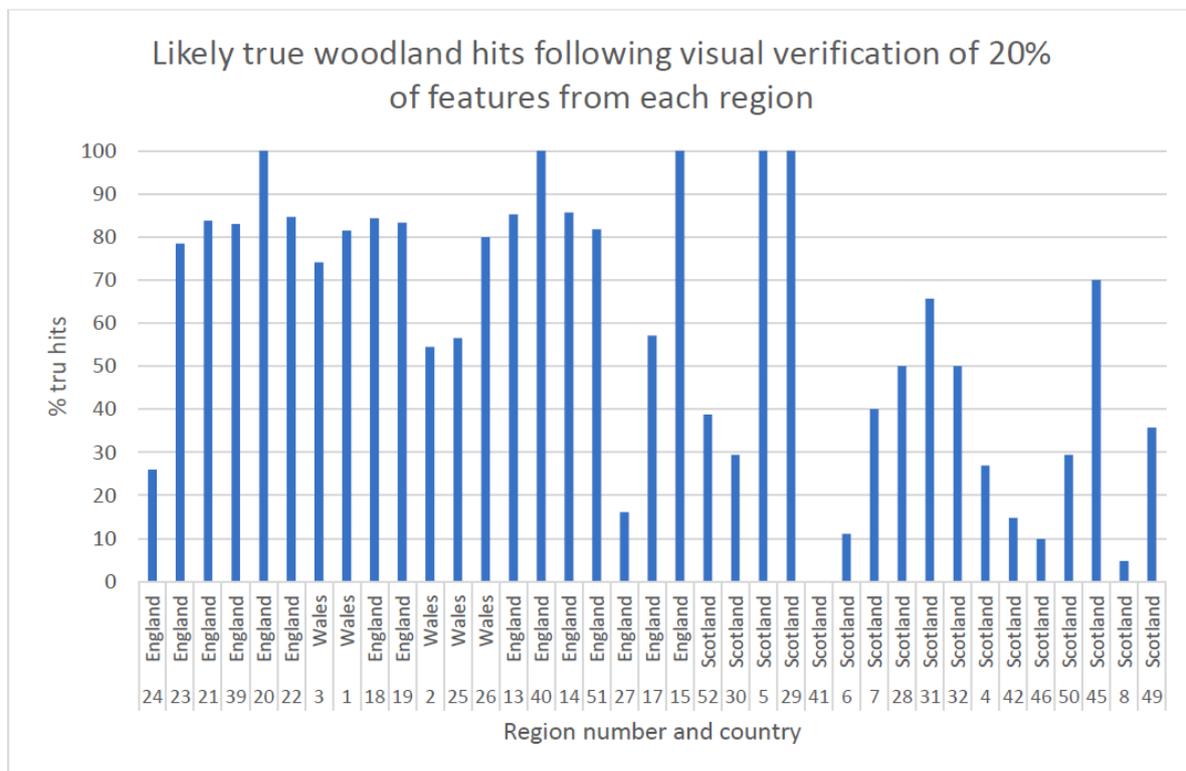


Figure A1.22: Likely true woodland hits following visual verification of 20% of features from each region

Table A1.3: Likely false and true hits of 20% of features for each region by country and an estimated area

| Country | Total features found | Number verified (20%) | Total likely true hits | Total likely false hits | % true hits | Estimation of area of true hits of all features (ha) |
|--------------|----------------------|-----------------------|------------------------|-------------------------|-------------|--|
| Total | 7139 | 1700 | 938 | 765 | 55.2% | 4656 |
| England | 3224 | 637 | 489 | 151 | 76.7 | 2399 |
| Wales | 1588 | 316 | 226 | 90 | 71.5 | 1132 |
| Scotland | 3757 | 747 | 223 | 524 | 29.9% | 1125 |

6 System requirements

This model employed a 64-core AMD Ryzon Threadripper system with 256GB of memory and was not tested on lower specification systems. Although this is a high end professional system, it is likely the model would run on a lower specification system, although memory requirements will be high due to loading in multiple country wide raster files.

7 Conclusions

This report details the background, methodology (a technical guide), a user guide and verification of a model to find areas of omission from an interim woodland map for England, Wales and Scotland. The model found between 4700 hectares and 6351 hectares of woodland not previously included on the woodland map, although this only represents approximately 2% of the woodland area (ie the interim woodland map had already covered approximately 98% of woodland). The model performed well over lowland areas, but less well over vegetated coastland and upland, particularly across Scotland and highland areas of England and Wales.

Appendix 2: Ancillary information for intercountry comparison

This section of the report is an Appendix of “ancillary information”. Several information resources and contacts were identified at the start of this task. Subsequently, we found some of the references to reports and studies were already mentioned in the NIRs, or, were judged not be directly useful for the study.

Rather than exclude the information completely from the project report, for completeness, we have included them in this Appendix. To keep the context of the information in e-mail correspondence, we have included the full e-mail rather than just summarise the information in the e-mails. Perhaps in future studies they may prove useful.

Moja global and FLINT software

Moja global is a collaborative project under the Linux Foundation that supports ambitious climate action by bringing together a community of experts to develop open-source software – including the groundbreaking FLINT software – which allows users to accurately and affordably estimate greenhouse gas emissions and removals from forestry, agriculture and other land uses (AFOLU).

The project’s members aim to support the widest possible use of credible emissions estimation software by countries, institutions and project managers. All members collaborate to achieve this goal, which not only reduces duplication and costs, but also provides more reliable software that responds better to user needs and delivers comparable, transparent and credible estimates.

<https://moja.global/about/>

From: Rowland, Clare S. <clro@ceh.ac.uk>

Sent: 15 March 2021 10:01

To: Buys, Gwen B. <gnew@ceh.ac.uk>; Watterson, John <John.Watterson@ricardo.com>

Cc: Tomlinson, Sam J. <samtom@ceh.ac.uk>; Evangelides, Christopher <Christopher.Evangelides@ricardo.com>

Subject: RE: LUC Tracking WPB Data Management & QA Catch-up

Hi John

I came across this last week: <https://moja.global/tools-of-moja-global/>

I’m not sure how it fits in and whether it’s something that should be in your review – Australia and Canada seem to be very involved with it, so apologies if you’ve already looked at it.

Best wishes

Clare

Australian spatial mapping for LULUCF

From: Rowland, Clare S. <clro@ceh.ac.uk>

Sent: 08 February 2021 12:20

To: Watterson, John <John.Watterson@ricardo.com>

Cc: Buys, Gwen B. <gnew@ceh.ac.uk>; Tomlinson, Sam J. <samtom@ceh.ac.uk>; Evangelides, Christopher <Christopher.Evangelides@ricardo.com>

Subject: FW: Australian spatial mapping for LULUCF

Hi John

Attached is the discussion with Robert Waterworth from 2014 – the questions and answers at the bottom of the thread are all very relevant to our current work.

I think after this email was sent other priorities intervened and I don’t think I ever followed it any further.

Improving Land Use Change Tracking in the UK Greenhouse Gas Inventory: Final Outputs Report

Best wishes

Clare

From: Robert Waterworth <rob.waterworth@anu.edu.au>
Sent: 12 July 2014 02:48
To: Rowland, Clare S. <clro@ceh.ac.uk>
Cc: Penman, Jim <j.penman@ucl.ac.uk>; Buys, Gwen B. <gnew@ceh.ac.uk>
Subject: Re: Australian spatial mapping for LULUCF

Dear Clare,

No problems. The only issue I realised this morning is that I left off the key reference sources! During the NCAS build we published everything we did in a technical report series. It has now been archived but can still be accessed at <http://pandora.nla.gov.au/tep/23322>. Tech reports 9 and 16 are key ones. The methods changed a little as we went along, but this will give you an idea. I can also put you in contact with the CSIRO team who designed and still run the QA/QC on the Australian products if that would help.

Cheers

Rob

PS: this page contains the reports as well as some other documents that may be useful: <http://pandora.nla.gov.au/pan/23322/20080818-0001/www.climatechange.gov.au/ncas/publications/index.html>

Robert Waterworth

Visiting Fellow

Fenner School of Environment and Society

Australian National University

Australia: Ph: +61 407 137173 | Mob: +61 428 343 864

Indonesia: Mob: +62 8111493964

Email: rob.waterworth@anu.edu.au

(Note: j.penman@ucl.ac.uk e-mail address is no longer active)

On 11/07/2014, at 11:28 PM, "Rowland, Clare S." <clro@ceh.ac.uk> wrote:

Dear Rob,

Thank you for your comprehensive reply, it's extremely helpful. I'll have a read through the documents you attached and have a think about your comments, as well as having a look into the Canadian system.

I suspect i'll be back in touch some time next week, when i've had time to read through everything and have a think about it. I may want to take you up on your offer to put me in contact with Werner Kurz, but i'll let you know next week.

Thank you again for all your help,

Best wishes

Improving Land Use Change Tracking in the UK Greenhouse Gas Inventory: Final Outputs Report

Clare

From: Robert Waterworth [rob.waterworth@anu.edu.au]
Sent: 11 July 2014 12:16
To: Rowland, Clare S.
Cc: Penman, Jim; Buys, Gwen B.
Subject: Re: Australian spatial mapping for LULUCF

Dear Clare,

Thanks for the email. There are a few papers out on how the system works. The key ones are on the system concepts and design (in sustainability science), the plantation modelling and the how the databases are used to integrate the remote sensing with the modelling systems (both in Forest Ecology and Management). All are attached.

The situation you describe is certainly not uncommon and one we needed to deal with in Australia. Consistent representation of lands using remote sensing is not something that I think has been dealt with particularly well in the LULUCF community but is becoming increasingly more important as more countries start to try to move towards use of RS. So I must admit I find your email timely and very interesting from a personal perspective.

Some thoughts (rather than answers, sorry) on your questions are below.

Let me know if you would like to discuss this over the phone sometime. It would also help if you could send me some other details, such as how you plan to use the forest inventory data: for example, will it be used to calculate some emissions factors or will it be used to develop systems that better represent finer scale dynamics (fairly important if moving to fine scale activity data updates). This decision will also drive what you want to get from the RS. Once you start to move towards the more complex systems you are describing I think you also need to be careful to avoid the AD x EF trap (i.e., thinking about the sides separately and assuming they will be multiplied together in a spreadsheet) and focus on how they can re-enforce and improve each other. IMHO this is where the real power of the RS can come to bear on the emissions estimates.

Apart from the Australian system approach, the Canadian system is another option. Although for their inventory it runs using polygons of stands, the Canadian model can use remote sensing data as well. Again, it will depend on how you are planning to do the emissions estimates. I have some knowledge of the Canadian system but for greater detail I can put you in contact with Werner Kurz.

Hope this helps somewhat. I am available all next week if you want to catch up. I will be in the USA the week after then in Canada after that (seeing Werner in fact) but can still make time if it helps.

All the best,
Rob

Robert Waterworth
Visiting Fellow
Fenner School of Environment and Society
Australian National University
Australia: Ph: +61 407 137173 | Mob: +61 428 343 864
Email: rob.waterworth@anu.edu.au<<mailto:rob.waterworth@anu.edu.au>>

Our current issues can probably be summarised as:

1) **How we merge data sets from different sources to create our best estimate of LU in LULUCF classes for a particular year. In particular, how we deal with data collected at different spatial resolutions, different spatial structures (some gridded, some vector-based with the vectors representing real-world objects such as fields) and different temporal resolutions/repeat-periods.**

This is an interesting issue. Without knowing all the data I am not sure the best way forward. We did face a similar situation and ended up using nearly all the data, but running a series of decision trees that both

prioritised data but also used other auxiliary data to help assist in decision making (for example, soil types, rainfall etc). The issue was deciding what data was 'best'. For example, planting maps/areas from plantation companies sounded great and were very accurate for the area planted, but the remote sensing never picked them up. In the end it turned out that the remote sensing was more accurate: sure they were planted but the failure rates (upwards of 80% in some cases) meant that if we had used the manual mapping we would have overestimate reforestation considerably. Neither dataset was necessarily better than the other, but the RS was better for what we needed for LULUCF. As per your note, it really involves knowing the data really well and what it is exactly that you are trying to produce. For spatial resolution, we simply ran at the finest scale we had (resampled Landsat, 25m). All other layers were multiples of this. Vector data was simply rasterised (this made things much easier in the late 1990's). This allowed us to drill through each individual pixel for all the data the model required (land use, climate, soil, forest type etc etc, there are about 5000 layers now so efficiency was/is key). I would note though that this is not necessary these days depending on how you are driving through the data layers.

2) **How we deal with error in the input data – one of our concerns is that some classes will be based primarily (or in part) on classified satellite data, and as such will have a level of error associated with it, which is likely in some cases to be greater than the level of change and may give an excessive rate of change and land disturbance. We are working on the basis that the Forest and Crop data represent very high quality data and the image classifications represent lower quality data.**

Again, an interesting issue. There are two key points here from my perspective: maps at points in time (even using the same method) are not time-series consistent maps, especially when hunting for change and the role of attributing cover change to land use change. To address the first issue in Australia we run a conditional probability network through the time series of forest extent. This meant that the state of a pixel at any point in time was informed by the state in the years previous. This helped increase the time series consistency in change (but always means that the latest year is the most uncertain). To address the second issue we ran a process of attribution to focus on the areas of change. This was very important in our case as land cover change (forest>non-forest and non-forest>forest) is very different to deforestation and reforestation. For example, fire removes canopy and in many years canopy loss from fire (and gain as it regrows in subsequent years) is vastly greater than true clearing (close to an order of magnitude). Drought, flood, pests also all caused LC change. This attribution process is described in the NIR in the but has been glossed over in many other publications. Happy to discuss this one further as it is something that I don't think is being well addressed in the RS community.

3) **How we capture known uncertainties and prevent/minimise their propagation**

I think I have at least partially covered this above, but happy to discuss further. We ran a process of continuous improvement and verification focusing on areas of greatest uncertainty. Having said this we tried to focus on the second part of the question. For emissions estimates estimated uncertainty by running monte carlo analysis (no other way of doing this in a fully integrated system).



BANGOR
UK Centre for Ecology & Hydrology
Environment Centre Wales
Deiniol Road
Bangor
Gwynedd
LL57 2UW
United Kingdom
T: +44 (0)1248 374500
F: +44 (0)1248 362133

EDINBURGH
UK Centre for Ecology & Hydrology
Bush Estate
Penicuik
Midlothian
EH26 0QB
United Kingdom
T: +44 (0)131 4454343
F: +44 (0)131 4453943

LANCASTER
UK Centre for Ecology & Hydrology
Lancaster Environment Centre
Library Avenue
Bailrigg
Lancaster
LA1 4AP
United Kingdom
T: +44 (0)1524 595800
F: +44 (0)1524 61536

WALLINGFORD (Headquarters)
UK Centre for Ecology & Hydrology
Maclean Building
Benson Lane
Crowmarsh Gifford
Wallingford
Oxfordshire
OX10 8BB
United Kingdom
T: +44 (0)1491 838800
F: +44 (0)1491 692424

enquiries@ceh.ac.uk

www.ceh.ac.uk