Environment and Rural Affairs Monitoring & Modelling Programme (ERAMMP)

ERAMMP Technical Annex-2 Report-60TA2: Quality Assurance of IMP Land Use Scenario Runs

Harrison, P.A.[1], Dunford, R.[1], Whittaker, F.[1], Mondain-Monval, T.[1], Beauchamp, K.[2], Cooper, J.[3], Dickie, I.[4], Fitch, A.[1], Gooday, R.[5], Hollaway, M.[1], Holman, I.P.[6], Jones, L.[1], Matthews, R.[2], Sandars, D.[6], Seaton, F.[1], Siriwardena, G.M.[3], Smart, S.M.[1], Thomas, A.R.C.[1], Trembath, P.[1], Vieno, M.[1], West, B.[1], Williams, A.G.[6]

[1] UK Centre for Ecology & Hydrology, [2] Forest Research, [3] British Trust for Ornithology, [4] eftec, [5] ADAS, [6] Cranfield University

Client Ref: Welsh Government / Contract C210/2016/2017
Version 1.0.0
Date: 05-August-2022

Funded by:

Llywodraeth Cymru
Welsh Government

Canolfan Ecoleg a Hydroleg y DU
UK Centre for Ecology & Hydrology

Version History

Version	Updated By	Date	Changes
1.0.0	ERAMMP Programme Team	05/08/2022	Published

Mae’r adroddiad hwn ar gael yn electronig yma / This report is available electronically at: **www.erammp.wales/60TA2**

Neu trwy sganio’r cod QR a ddangosir / Or by scanning the QR code shown.

Mae’r ddogfen yma hefyd ar gael yn Gymraeg / This document is also available in Welsh

Series	Environment and Rural Affairs Monitoring & Modelling Programme (ERAMMP)
Title	ERAMMP Technical Annex Report-60TA2: ERAMMP Integrated Modelling Platform Quality Assurance
Client	Welsh Government
Client reference	C210/2016/2017
Confidentiality, copyright and reproduction	© Crown Copyright 2022 This report is licensed under the Open Government Licence 3.0.
UKCEH contact details	Bronwen Williams UK Centre for Ecology & Hydrology (UKCEH) Environment Centre Wales, Deiniol Road, Bangor, Gwynedd, LL57 2UW 01248 374500 erammp@ceh.ac.uk
Corresponding author	Paula Harrison, UKCEH paulaharrison@ceh.ac.uk
Authors	Paula Harrison[1], Rob Dunford[1], Freya Whittaker[1], Thomas Mondain-Monval[1], Kate Beauchamp[2], Joe Cooper[3], Ian Dickie[4], Alice Fitch[1], Richard Gooday[5], Mike Hollaway[1], Ian Holman[6], Laurence Jones[1], Robert Matthews[2], Daniel Sandars[6], Fiona Seaton[1], Gavin Siriwardena[3], Simon Smart[1], Amy Thomas[1], Phil Trembath[1], Massimo Vieno[1], Adrian Williams[6] [1] UK Centre for Ecology & Hydrology, [2] Forest Research, [3] British Trust for Ornithology, [4] eftec, [5] ADAS, [6] Cranfield University
Contributing authors & reviewers	James Skates (Welsh Government), Ken Stebbings (Welsh Government)
How to cite (long)	Harrison, P.A., Dunford R., Whittaker, F., Mondain-Monval, T., Beauchamp, K., Cooper, J., Dickie, I., Fitch, A., Gooday, R., Hollaway, M., Holman, I.P., Jones, L., Matthews, R., Sandars, D., Seaton, F., Siriwardena, G.M., Smart, S.M., Thomas, A.R.C., Trembath, P., Vieno, M., West, B. Williams, A.G. (2022). *Environment and Rural Affairs Monitoring & Modelling Programme (ERAMMP).* ERAMMP Technical Annex-2 Report-60TA2: ERAMMP Integrated Modelling Platform (IMP) Land Use Scenarios Quality Assurance. Report to Welsh Government (Contract C210/2016/2017)(UK Centre for Ecology & Hydrology Projects 06297 & 06810)
How to cite (short)	Harrison, P.A. et al. (2022). ERAMMP Technical Annex-2 Report-60TA2: ERAMMP Integrated Modelling Platform (IMP) Land Use Scenarios Quality Assurance. Report to Welsh Government (Contract C210/2016/2017)(UKCEH 06297/06810)
Approved by	Bridget Emmett (UKCEH) James Skates (Welsh Government)

Abbeviations Used in this Report

BBS	Breeding Bird Survey
BPS	Basic Payment Scheme
BTO	British Trust for Ornithology
BSFP	British Survey of Fertiliser Practise
CAP	Common Agricultural Policy
COMEAP	Committee on the Medical Effect of Air Pollutants
DA	Disadvantaged Area
DAMS	Detailed Aspect Method of Scoring
DMU	Decision Making Unit
EFT	ERAMMP Farm Type
ERAMMP	Environment and Rural Affair Monitoring and Modelling Programme
ESC	Ecological Site Classification
EUID	ERAMMP Unique Identifier
FBI	Farm Business Income
FTE	Full time equivalent worker.
GLM	generalised linear modelling
GHG	Greenhouse Gas
GLU	Grazing Livestock Unit
GMEP	Glastir Monitoring and Evaluation Programme
HWP	Harvested Wood Product
HMS	Harmonised Monitoring Scheme
IMP	Integrated Modelling Platform
JAS	June Agricultural Survey
JNCC	Joint Nature Conservation Committee
LAM	Land Allocation Module
LFA	Less Favoured Areas (the sum of DA + SDA).
LISS	Low Impact Silvicultural Systems
LULUCF	Land Use, Land Use Change and Forestry
MAE	Mean Absolute Error
NAEI	National Atmospheric Emissions Inventory
NPV	Net present value
QA	Quality Assurance
QC	Quality Control
RFT	Robust Farm Type
RSPB	Royal Society for the Protection of Birds
SDA	Severely Disadvantaged Area
SRF	Short Rotation Forestry
SRO	Senior Responsible Officer
UKCEH	UK Centre for Ecology and Hydrology
WFD	Water Framework Directive
WG	Welsh Government

Abbreviations and some of the technical terms used in this report are expanded on in the programme glossaries: https://erammp.wales/en/glossary (English) and https://erammp.cymru/geirfa (Welsh)

Contents

1 Introduction

This document reports Quality Assurance (QA) of the Integrated Modelling Platform (IMP) developed within the Environment and Rural Affairs Monitoring and Mapping Project[1] (ERAMMP). It focuses on the version of the IMP used to simulate the Land Use Scenarios delivered to the Welsh Government (WG) between August 2020 and March 2021[2]. This document should be read in conjunction with its partner document, IMP Trade Scenarios assumptions (see Annex-1 of the IMP Land Use Scenarios Final Report; ERAMMP Report-60[2]). Together, these documents act to support those in WG who use the IMP to interpret its outputs. Further information to support the QA of the IMP is available to WG in the slidepacks, data dictionaries and data cubes.

1.1 What is the ERAMMP IMP?

The IMP is a linked-model system that has been co-designed with the WG to support decision-making around the future of Welsh agriculture, and Wales's natural environment. A full description of the IMP is given in Chapter 2 of the IMP Land Use Scenarios Final Report (ERAMMP Report-60). This is summarised in this section to provide context for the QA reporting.

The IMP has been designed to enable rapid exploration of the effects of policy and management interventions on farm viability, land use change and various public goods in Wales. It takes an integrated approach that recognises that policy effects in one sector have indirect effects in other sectors. This allows assessment of potential unintended consequences of policy interventions and appraisal of potential trade-offs and synergies between payments for public goods.

To do this, the IMP has been constructed as a chain of specialised models covering agriculture, forestry, land use allocation decisions, water, air, soils, biodiversity, ecosystem services and their valuation. The models pass data between them representing different biophysical and socio-economic interactions between sectors (Figure 1.1). The modelling works at the farm scale and considers each farm in Wales with a workforce of > 1 Full Time Equivalent (FTE).

The top of the IMP modelling chain (the yellow boxes) focusses on identifying the potential profitability of all possible land use types on a given >1FTE farm. This includes taking into consideration that farm's constraints (e.g. climate, soil type, elevation and special designations). The Land Allocation Model (LAM) compares the profitability of the different alternatives to decide whether the farm retains its farm type, transitions to another farm type or leaves full-time farming as a result of loss of revenue. The LAM then passes the bottom-

---

[1] www.erammp.wales (English) and www.erammp.cymru (Welsh)

[2] Harrison, P.A., Dunford, R., Beauchamp, K., Cooper, J., Cooper, J.M., Dickie, I., Fitch, A., Gooday, R., Hollaway, M., Holman, I.P., Jones, L., Matthews, R., Mondain-Monval, T., Norris, D.A., Sandars, D., Seaton, F., Siriwardena, G.M., Smart, S.M., Thomas, A.R.C., Trembath, P., Vieno, M., West, B., Williams, A.G., Whittaker, F., Bell, C. (2022). *Environment and Rural Affairs Monitoring & Modelling Programme (ERAMMP)*. ERAMMP Report-60: ERAMMP Integrated Modelling Platform (IMP) Land Use Scenarios. Report to Welsh Government (Contract C210/2016/2017)(UK Centre for Ecology & Hydrology Projects 06297 & 06810)
www.erammp.wales/60

of-chain models (green boxes) the spatialised information from the top-of-chain models (yellow boxes); including Decision-Making Unit (DMU) level land use, associated with the final farm outcome selected by the LAM. The bottom-of-chain models use this data to produce outputs associated with public goods, ecosystem services and their value.

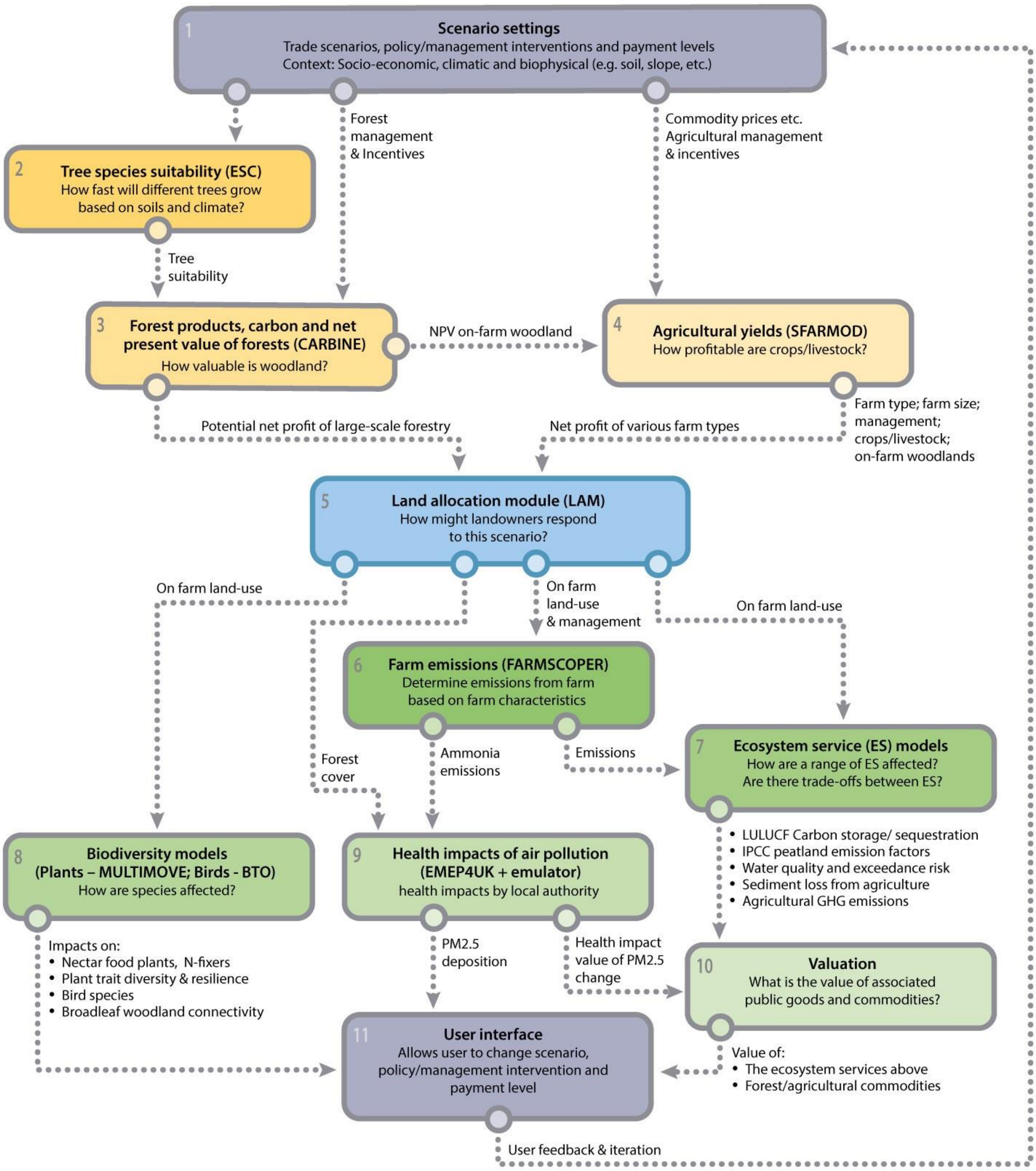


Figure 1.1: An overview of the linked models within the ERAMMP Integrated Modelling Platform and the types of data that are passed between them.

1.2 Quality Assurance – what is it and why does it matter?

Understanding the strengths, weaknesses, opportunities, and limits of any modelling system is vital so users of the model understand what is and is not possible to infer from the outputs.

QA provides the critical reflection needed to understand these limits. The IMP is designated as business critical and is one source of information used to support decision-making in policy, as such, this QA is mandated by the UK Government's Review of quality assurance of government analytical models (HM Treasury: United Kingdom, 2013) and the Aqua Book (HM Treasury: United Kingdom, 2015).

The Aqua book sets out the four principles of analytical QA to support the delivery of fit-for-purpose analysis:

- **Proportionality of response:** The extent of the analytical quality assurance effort should be proportionate in response to the risks associated with the intended use of the analysis.
- **Assurance throughout development:** Quality assurance considerations should be considered throughout the life cycle of the analysis and not just at the end.
- **Analysis with RIGOUR:** Quality analysis needs to be Repeatable, Independent, Grounded in reality, Objective, have understood and managed Uncertainty, and the results should address the initial question Robustly.
- **Verification and validation**: Analytical quality assurance is more than checking that the analysis is error-free and satisfies its specification (verification). It must also include checks that the analysis is fit for the purpose for which it is being used (validation).

As described in Section 1.1, the IMP comprises a series of linked models with data flows representing real-world interdependencies. The range and complexity of the models means there is no single QA activity. Instead, QA has been delivered through a range of activities, with each adding to the overall level of QA. Each component of the IMP has undergone QA led by an expert modelling team, full details of the model QA can be found in Chapters 3 - 14. Briefly, these approaches include:

- **Version control:** the management of different versions of inputs, outputs, and models.
- **Verification:** the process through which the model is reviewed to ensure it is error free and meets specification.
- **Documentation of assumptions:** the presentation of key parameters and assumptions to build understanding.
- **Expert Assessment (Consortium and External**): using expert knowledge within the consortium and externally (including a WG expert group) to assess the data, assumptions, methodology and outputs.
- **Validation:** the process through which the model is reviewed to ensure it is fit-for-purpose including comparison or contextualisation of baseline model runs with independent datasets or alternative modelling approaches.
- **Peer Review:** many of the models have significant history within the academic literature, justifying their application within the IMP. Others follow agreed standard approaches used for government reporting: these are also considered fit-for-purpose.
- **Uncertainty Analysis (Sensitivity Testing):** including sensitivity analysis of key parameters and an assessment of the implications on the results produced. This stage also reviews the relevance of pre-defined assumptions.

- **Building understanding**: presentation of baseline results to aid interpretation of other scenarios. Often including supporting expert interpretation.

1.3 How to use this document?

Chapter 2 provides an overview of the QA approaches that have been applied to the IMP. The remainder of this document is divided into sections focussing on each of the individual components of the IMP (Chapters 3 - 14). Within each section, QA approaches are detailed to allow the WG expert group and the ERAMMP modelling team to have confidence in their understanding of the limits of the modelling results. Each chapter addresses a different part of the model chain and the modelling teams involved have taken approaches most appropriate to their model.

This document should be read in conjunction with the IMP Assumptions Document (Annex 1), which sets out the key assumptions as agreed with WG both across the modelling framework as a whole and within each individual model.

Where possible, each model has been validated against a baseline scenario to ensure the model is fit-for-purpose and grounded-in-reality. For this purpose, a ‘baseline’ scenario was developed to represent something close to current conditions. It is parameterised as a farming system with CAP Pillar 1 Basic Payments and cost-neutral Pillar 2 additional payments. Where possible, 2015 is the year used for the data to match with the Land Cover Map 2015 data used to parameterise the modelling. The full parameterisation of the IMP baseline is detailed in the Annex 1 Assumptions Document (particularly assumptions 5-12).

2 Summary of the IMP QA

2.1 Development of the IMP

Due to its designation as business critical, the complexity of the modelling chain and its use as support within policy decision-making, the IMP demands a comprehensive analytical QA response to satisfy the four Aqua Book principles.

To address these principles the ERAMMP IMP was developed following the principles of co-creation, taking an iterative approach that involved the modelling consortium and Government experts throughout. The principles of RIGOUR were strictly adhered to with all assumptions underlying the modelling approach agreed, transparently documented and signed-off by a Senior Responsible Officer (SRO) within WG following a multi-stage iterative discussion between modellers and end users. This framework, as illustrated by Figure 2.1, addresses both proportionality to response and assurance throughout development.

The co-creation approaches started with an agreement over the set of models and datasets to be included in the modelling platform. For a model to be included in the platform it had to conform as best as possible to the following selection criteria. The model had to be:

1. Well-tested in previous research and policy applications;
2. Appropriate for multi-scale spatially-explicit policy assessment studies;
3. Able to produce a wide range of policy-relevant outputs;
4. Responsive to a wide range of environmental, policy and market drivers;
5. Capable of using readily available public data as inputs;
6. Able to enable quantification of uncertainties for the estimations;
7. Suitable for integration, in that points of contact exist between the models; and
8. Easily adapted allowing implementation within the proposed time frame.
9. Datasets were selected by a joint WG-modelling team working group to ensure that the best available and most recent datasets were utilised.

Early interactions between the modelling team and WG also focused on the types of policy questions that were expected to be asked of the modelling platform. These included the exploration of the impacts of different interventions aligned to WG policy objectives (such as payments to farmers associated with a new sustainable farming scheme) and external drivers (such as changes in commodity prices due to new trading relationships) on Welsh agricultural, land use and ecosystem service outcomes. These were used to co-develop a detailed specification describing the individual system components, linkages (i.e. which outputs from which models will form inputs to other models), how they respond to different drivers (including policy drivers), and the spatial and temporal scale of simulation.

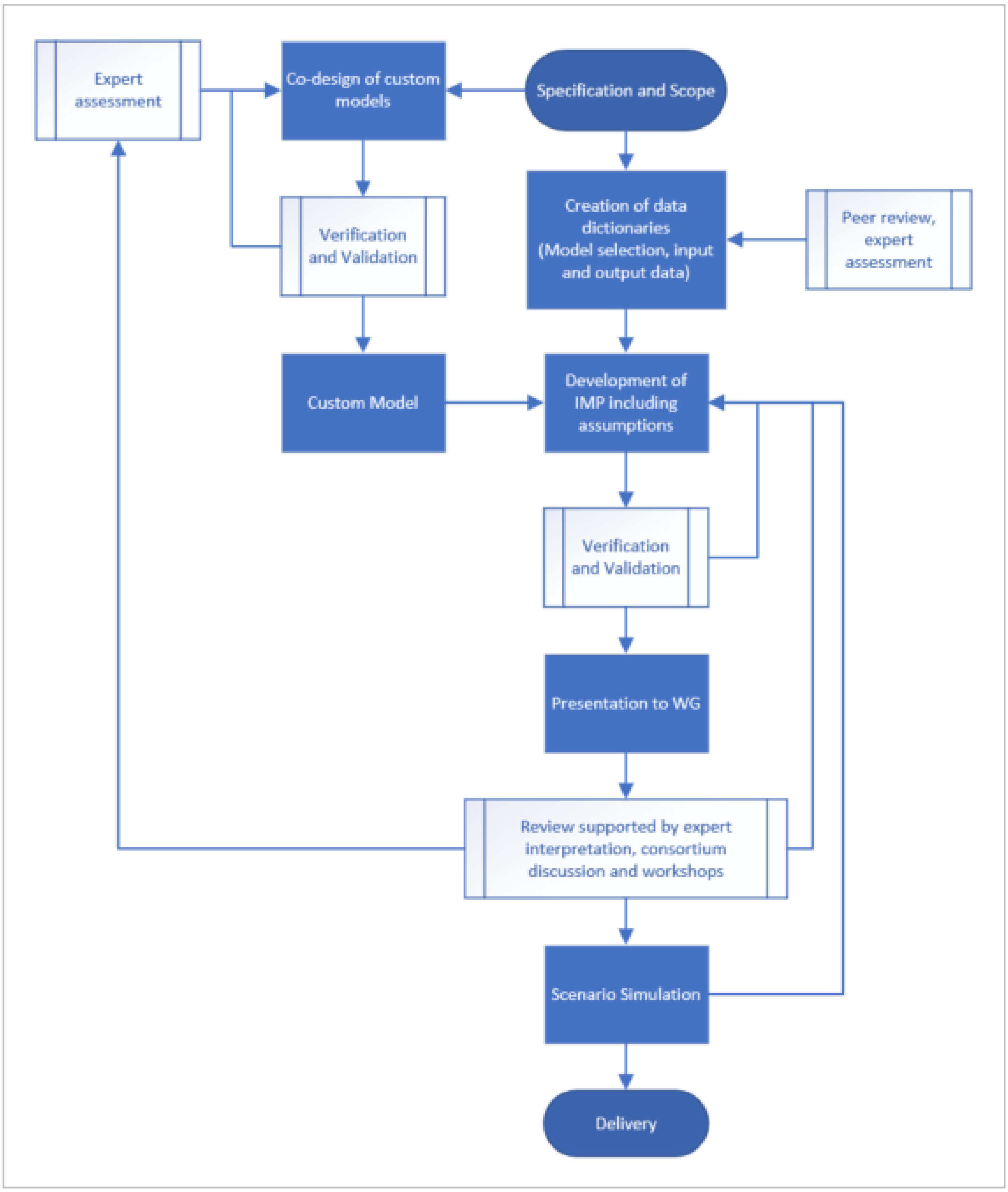


Figure 2.1 Schematic showing the design, build, test and review stages of the IMP development

2.2 Overview of model QA

To assure quality throughout, each individual model has undergone QA led by an expert team. Full details of each component QA are detailed in Chapters **Error! Reference source not found.** - **Error! Reference source not found.**. This Chapter provides an overview of the QA processes undertaken in each model (Table 2.1). Each model was subject to version control, analyst self-check, internal verification, assumptions documentation and internal peer review.

Table 2.1 QA processes by model

Model	Version Control	Verification	Assumption Documentation	Expert Assessment	Validation	Peer Review (PR) and Standard Approaches (SA)	Sensitivity Testing
SFARMOD agricultural model	□	□	□	□	□	PR	□
ESC-CARBINE-NPV forestry models	□	□	□	□	□	PR	□
Land Allocation Module	□		□	□	□	□	□
BTO bird models	□	□	□	□	□	PR	□
MultiMOVE plant model	□	□	□	□	□	PR	□
Habitat Connectivity	□	□	□	□	□	□	□
Farmscoper emissions model	□	□	□	□	□	PR	□
Water Quality	□	□	□	□	□	Partial	□
Air Quality	□	□	□	□	□	PR	□
Carbon	□	□	□	□	□	SA	□
Valuation	□	□	□	□	□	SA	□

Version control:

The IMP uses a soft model coupling approach that moves away from hard-wired integrated models to provide a customisable modelling framework that can adapt to changing WG needs. This soft-coupling approach is key to the flexible integration, as ‘people’ (academics and WG working in partnership) are the enablers of fast model adjustment to evolving WG business critical policy questions. However, this approach requires a strict approach to QA and data management to ensure correct application and consistency across the IMP.

Each data pass in the IMP is representative of a real-world interdependency and as such, any iteration in the ‘upstream’ models must be cascaded correctly through the chain. This was facilitated by the generation of Unique Identifiers, or ERAMMP Unique Identifiers (EUIDs). An EUID was assigned to each model, input and output which facilitates traceable data flow to ensure version control, **verification**, and **repeatability**. A copy of the EUID database is available to WG and can be accessed by contacting the UKCEH IMP team.

Verification:

Verification is the process by which modellers check and understand that their model is functioning as expected. It has been carried out on all models with processes and checks tailored to each model. Examples include checking code for errors, setting checks to catch common errors in code or modelling teams using their own expert judgement to assess their model's performance is within expected parameters.

Assumptions documentation:

For **transparency** and **repetition**, all model assumptions are documented in the Assumptions Document (Annex 1). All assumptions have been reviewed, tested, and signed-off by the WG Senior Responsible Officer (SRO). The assumptions documented reflect the final agreements of a considerable period of iteration between the consortium modelling teams and a range of experts within Welsh Government. This applies to all models within the modelling chain and across the modelling framework (e.g. the choice of 1 FTE cut-off). The iterative process to explicitly define and test model assumptions increases the robustness of analysis by presenting the results in the context of residual uncertainty and limitations to ensure it is used appropriately.

The assumptions document has been made available to WG.

Expert assessment:

Each model underwent expert assessment (consortium and external) to independently check model verification, validation, and any implications on linked models. This addresses the principle of **independence** by involving a range of perspectives across the modelling team and WG. Whilst there are limits to which bias can be constrained, this document and the assumptions document are efforts to be transparent so that any biases can be addressed if and when they are raised.

Throughout the IMP development, results were shared with WG, supported by expert assessment and documentation. This provided opportunities for challenge by the end user and increased the robustness of analysis and subsequent decision-making.

Peer review and standard approaches:

The modelling chain uses both academic peer review and agreed, standard approaches used for government reporting, which addresses to some extent, model choice **uncertainty**.

Academic peer review of models is an important step in the assessment of model's fitness-for-purpose. Most models within the ERAMMP IMP chain have a significant history of application within academic literature for addressing similar questions to those they are used for in ERAMMP. A review of supporting literature for each model is provided in subsequent chapters. Where a model has been specifically developed for use in the ERAMMP IMP (e.g. LAM), addition checks, expert assessment and where possible, validation and sensitivity testing were undertaken.

In other cases, (e.g. Water Quality) the coefficients are derived from a peer-reviewed model (FARMSCOPER) and combined with the outputs of another peer-reviewed model (SFARMOD); to provide extra confidence the combined outputs are also independently evaluated.

The carbon accounting and ecosystem service valuation modelling components of the ERAMMP IMP use standard approaches used for government accounting. The carbon accounting follows LULUCF carbon accounting procedures, whilst the valuation of ecosystem services follows Treasury Green Book guidance on appraisal and evaluation.

Validation:

Due to the complexity of the modelling chain, the IMP was validated by assessing the results of each model element. All models were validated where possible, although the specific approach taken varies depending on the model and the available data. A baseline scenario was generated for this purpose. Full model validation was not always possible, either due to the methods employed or lack of available data. In these cases, thorough sense checks were undertaken. Validation addresses uncertainty and attempts to challenge the perceptions of both the WG and the IMP consortium by making connections between the analysis and its real-world consequences. In doing so, it ensures the context of the problem is properly grasped and the analysis is grounded-in-reality.

Sensitivity testing:

Sensitivity testing is used to **address uncertainty** about key parameters. Where there is significant dependency on an uncertain assumption, effort has been made to control and communicate the implications of that uncertainty. This is particularly the case for the newly developed LAM. The LAM recognises that there are complex human and financial factors that affect changes to farm type. It is not possible to model these complex relationships, which are instead reflected by co-developed rules and Farm Business Income (FBI) thresholds. Downstream models are heavily reliant on the outcome of the LAM and as such, sensitivity testing was carried out on key parameters including, the minimum simulated FBI required to continue full-time farming. This provided opportunities to challenge assumptions and understand their implications.

3 Farmland – SFARMOD

Authors: Daniel Sandars and Ian Holman

3.1 Introduction to the model QA

This section introduces the Silsoe Whole Farm Model (SFARMOD) and quality assurance steps taken to understand it. SFARMOD is the on-farm agricultural model that assesses, for each >1FTE farm in Wales, how profitable a range of farm types would be. The model is briefly introduced below (Section 3.2) and four sections illustrate efforts to improve confidence with, and understanding of, the model and its outputs:

- Peer Review: Documenting previous studies published in the peer review literature which detail how SFARMOD has been applied and validated (Section 3.2.1);
- Validation: Comparison of SFARMOD's broad land use assignments for the Baseline Scenario against Land Cover Map 2015 and June Agricultural Census (Section 3.2.2);
- Validation: Comparison of SFARMOD's cropping and grassland output for the Baseline Scenario with the June Agricultural Survey data (Section 3.2.3);
- Validation: Comparison of SFARMOD's stock numbers for the Baseline Scenario with Welsh June Agricultural Census (Section 3.2.4).

3.2 Introduction to the model

The Silsoe Whole Farm Model (SFARMOD) (Annetts and Audsley, 2002) is a constrained optimising strategic farm planning model based on profit maximisation, solved by linear Programming (LP). It has been extensively applied across a range of farm types and scales (e.g. Hutchings et al., 2018; Holman et al., 2018). SFARMOD finds the optimum stocking, cropping, manure usage, fixed costs, labour and profit for given land quality, climate and a selection of available resources, constraints, costs and revenues. For the livestock farms that dominate in Wales, it provides an economic optimum farm management that ensures that the feed and bedding demand of the optimised livestock numbers through the year can be met by a farm-specific combination of on-farm feed production and bought-in concentrates.

The nutritional demands of livestock are represented by fortnightly demands for metabolisable energy, crude protein and dry matter intake, along with bedding demands to meet welfare needs, which must be met within acceptable tolerances. Within their grazing seasons, suitable stock are all fed grazed grass (based on disaggregated yield using Qi et al., 2018), with supplements, mainly for dairy cows. The model chooses the least cost ration (considering grass silage, a self-fed forage crop (roots), whole crop silage, maize silage, straw and concentrates), so that grass use is normally maximised. Input data are derived from Nix, ABC, Welsh Farm Business Survey and British Survey of Fertiliser Practice.

Due to data constraints, we use representative ERAMMP Robust Farm Types (e.g. general cropping, lowland cattle and sheep, dairy) to define a set of realistic farming systems per Decision-Making Unit (DMU) to solve with SFARMOD. Each full-time Welsh farm is therefore modelled as a set of DMUs based on farm-specific discretised (banded) soil, rainfall, slope, altitude and recent farm type and land cover. Each DMU is optimised independently and then additively combined to obtain the solution for the farm. Within scenario runs, an optimised solution is derived for all feasible farm types for each farm.

3.2.1 Peer Review: SFARMOD

The details of the original SFARMOD model, its application and validation can be found in Annetts and Audsley (2002). Since then SFARMOD has been applied across a wide range of contexts and scales (e.g. regional - Holman et al., 2005, Audsley et al., 2008; national – Holman et al., 2016; Papadimitriou et al., 2019a; continental - Audsley et al., 2014; Harrison et al., 2019; Papadimitriou et al., 2019b; Lee et al., 2019), including inter-model comparison (Hutchings et al., 2018), sensitivity (Fronzek et al., 2019; Kebee et al., 2015) and uncertainty analyses (Brown et al., 2015; Dunford et al., 2015).

3.2.1 Validation: Comparison of broad land use classes between SFARMOD, Land Cover Map 2015 and June Agricultural Census

The June Agricultural Census/Survey (JAS) of Agriculture and Horticulture is a survey of agricultural activity in Wales and the rest of the UK. It is performed annually as a stratified sample of farms across Wales, with higher (100%) sampling rate in “large” and “very large” farms. For reference, the total sample was 11,069 farms in 2019. As a primary source of information about how Welsh farms are managed the JAS is a useful benchmark comparator for SFARMOD modelling results.

Note: to allow a direct comparison with SFARMOD outputs the JAS data was first scaled to remove farms of <1FTE (Full Time Equivalent). As the JAS results suggest that there are 8561 farms that are >1FTE, compared to the 7726 farms > 1FTE in the IMP, the JAS results were further scaled to represent the equivalent number of farms to the IMP (see Appendix A).

To provide spatially-explicit insights into each farm’s farming systems, SFARMOD used DMUs that incorporate simplified land use classes based on amalgamation of Land Cover Map 2015 (LCM2015) target classes. These simplified classes represent arable (based on the LCM2015 class of “arable and horticulture”), semi- or improved grassland (aggregation of “improved grassland”, “neutral grassland” and “calcareous grassland”) and unimproved grassland / rough grazing (aggregation of “acid grassland”, “fen, marsh and swamp”, “heather”, “heather grassland” and “bog”) to inform its selection of appropriate farming systems.

The comparative evaluation of land use between JAS 2015-16, LCM2015 and the SFARMOD simulated baseline scenario is shown in Table 3.1. In terms of total farmed area, the process of rescaling the JAS to remove <1FTE farms and to re-scale to the same farm number is close. However, there is a discrepancy between the grassland types, with the JAS estimating a smaller area of rough grazing and a greater area of semi-improved/improved grassland areas (given by the categories of temporary or permanent grassland) than the LCM2015 and the SFARMOD baseline. Nevertheless, despite these uncertainties arising from different data sources and different grassland classifications, there is a good overall match between these datasets and the broad land use outcomes from the SFARMOD solutions.

Table 3.1 Comparison of broad land use in >1FTE farms (7726) according to Land Cover Map 2015, JAS and SFARMOD baseline

Source	**Arable (ha)**	**Semi- or improved grassland (ha)**	**Unimproved grassland /Rough grazing (ha)**	**Total farmed area (ha)**
LCM2015	63,353	613,120	248,246	924,720
JAS (2015-16) [scaled]	66,460	669,212	176,581	912,254
SFARMOD baseline	65,859	610,427	248,431	924,717

3.2.2 Validation: Comparison of SFARMOD cropping and grassland output with the June Agricultural Survey data

Table 3.2 shows a comparison between the re-scaled JAS estimates and the different crops and grassland types simulated by SFARMOD. The JAS and SFARMOD total crop areas (flagged in green on the table) are very similar (given the inevitable uncertainties in the JAS scalars), although there are differences between individual crops and between the total areas of arable and break crops. Part of the difference is due to a reduced set of crops within SFARMOD to manage computability; this means that some crops in the JAS are unrepresented in SFARMOD (e.g. oats, peas and beans), whilst stubble turnips and whole crop silage are used by SFARMOD to represent the diversity of crops included within the broad JAS category of “stockfeed”. It is also apparent that there is some interplay between crops in the SFARMOD solutions with overestimated areas of spring cropping maize and stock feed crops (+9527 ha) being largely balanced by underestimated areas of spring barley and oats (-8879 ha).

The break crops within a rotation are opportunistic by nature and their inclusion in a rotation is sensitive to a balance of many cultural and financial factors such as rotational benefits, local markets, and timeliness of demands on the machinery and workforce. Consequently, given that arable crops represent only 6.5% of Welsh farmland, SFARMOD has produced a good overall representation of arable farming in Wales.

With the grassland systems, SFARMOD has an apparent bias towards temporary grass compared to permanent grass, which is defined by a point in a distribution of ley lengths/ reseeding intervals. The apparent over-estimation of rough grazing by SFARMOD compared to the JAS refers to the differences between JAS and LCM2015 described in the earlier section. The effect of more temporary grass is to see more forage crops in rotation and higher grass yields. The practical effect is to increase the possible stocking and inputs. The effect of more rough grazing is the opposite.

Table 3.2 Validation of ERAMMP crop and grassland types against June Agricultural Census 2015-16

	Welsh June Agricultural Census Survey 2015-16	Scaling to remove < 1FTE farms	Adjusted Cropping	Modelled BASELINE
Wheat	21,835	77%	16,819	14,702
Winter barley	6,926	77%	5,335	7,387
Spring barley	15,136	77%	11,659	6,795
Oats	5,212	77%	4,015	-
Other cereals	654	77%	504	-
Total Cereals	**49,763**		**38,332**	**28,883**
Potatoes (early)	919	77%	708	1,318
Potatoes (main)	1,927	77%	1,484	63
Maize	10,003	77%	7,705	13,563
Stockfeed	14,761	77%	11,370	15,039
Field peas & beans	849	77%	654	-
Oilseed rape	4,812	77%	3,707	6,993
Other crops	2,570	77%	1,980	-
Bare fallow	675	77%	520	-
Total break crops	**36,516**		**28,128**	**36,976**
Total arable crops and bare fallow	**86,279**		**66,460**	**65,862**
New grass	125,692	76%	95,473	142,724
Permanent grass	876,173	65%	573,738	467,703
Sole rights rough grazing	222,075	80%	176,581	248,431
Total Grass	**1,223,940**		**845,793**	**858,858**
Total Cereals, break crops, and grass	**1,310,219**		**912,254**	**924,720**

3.2.3 Validation: Comparison of SFARMOD baseline stock numbers with Welsh June Agricultural Census

Grazing Livestock Units (GLU) have been used to provide an approximate basis of comparison for stock numbers. This allows the breeding and finishing stock of various ages to be combined to a single figure for dairy, beef, cattle, and sheep to inform feed requirements and stocking rates.

The standard GLU constants from the Nix Farm pocketbook were re-estimated for Welsh sheep and beef to take account of the large numbers of light breeds of sheep that uniquely characterise the traditional breeds in the mountains of Wales. Welsh sheep are distributed 40% hill, 47% upland and 13% lowland and these have standard grazing livestock unit estimates of 0.06, 0.07 and 0.08, respectively (Anthony pers. comm. 2022). Table 3.3 shows the results of these calculations.

Table 3.3 Calculation of Grazing Livestock Units based populations of dairy, beef and sheep from the Welsh JAS

	Welsh June Agricultural Census Survey 2015-16	**Welsh GLU factor**	**Grazing Livestock Units**
Male cattle <1year	137,637	0.34	46,797
Beef females <1	106,393	0.34	36,174
Male 1-2	94,252	0.65	61,264
Beef females1-2	83,166	0.65	54,058
Male 2+years	40,591	0.65	26,384
Beef f 2+ not calved	41,894	0.75	31,421
Beef f 2+ calved	166,692	0.75	125,019
Total Beef GLU			**381,116**
Dairy female <1	77,864	0.34	26,474
Dairy 1-2	70,039	0.80	56,031
Dairy 2+ not calved	54,120	0.80	43,296
Dairy 2+ calved	246,331	1.00	246,331
Total Dairy GLU			**372,132**
Total cattle	1,118,979		
Breeding ewes	3,661,555	0.070	256,309
Ewes for cull	311,372	0.070	21,796
Ewes for first time breeding	769,974	0.053	40,424
Rams	107,167	0.070	7,502
Other sheep (female)	48,354	0.070	3,385
Other sheep (male)	29,500	0.070	2,065
Lambs	4,576,055	0.035	160,162
Total sheep	9,503,977		
Total sheep GLU			**491,642**

Table 3.4 shows that there is a generally good match between the SFARMOD baseline and the re-scaled JAS total GLUs, with a difference of less than 3.5%. There is a slight model under-estimation of beef GLUs and over-estimate of dairy GLUs. The apparent under-estimation of sheep numbers (and to a lesser extent beef GLUs) probably reflects three main issues. Firstly, the metabolic feed requirement from rough grazing is simulated using the grass yield model of Qi et al. (2017), but there is considerable uncertainty in the Qi model outputs due to a lack of yield data from rough grazing systems. Secondly, SFARMOD implicitly assumes good agronomic practice so that over-stocking land is prevented by both feed limitations and soil condition. Thirdly, SFARMOD makes a rationale economic decision regarding the use of supplementary feeds to augment the metabolic feed production of rough grazing, which may reduce stocking rates relative to local practice.

Table 3.4 Comparative evaluation of SFARMOD baseline stocking numbers with Welsh June Agricultural Census

Stocking	Welsh June Agricultural Census Survey 2015-16	Scaling to remove <1FTE farms	Adjusted livestock	Modelled baseline
Total Beef GLU	381,116	71%	270,592	249,031
Total Dairy GLU	372,132	89%	331,197	362,720
Total Sheep GLU	491,642	83%	408,062	362,917
Total GLU			**1,009,851**	**974,668**

4 Forest suitability, carbon balance & economics - ESC-CARBINE-NPV

Authors: Kate Beauchamp, Robert Matthews and Vadim Saraev

4.1 Introduction to the model QA

This section introduces the tree species suitability model (Ecological Site Classification, ESC), the forest sector carbon accounting model (CARBINE), the forest economic calculations of Net Present Value (NPV), the processing model linking woodland data across scales (GLUE) and the quality assurance steps taken to understand them. The models are outlined in more detail in Section 4.2 and section 4.2.1 to 4.2.8 describe the quality assurance steps taken:

- Validation/ Contextualisation: The ESC model has been validated through peer review, survey data, and long-term practitioner testing. ESC outputs have been compared to published values and model runs from published reports.
- Validation/ Contextualisation: CARBINE outputs have been compared to published values, articles and model runs from published reports.
- Presentation/ Interpretation: Baseline results for all models have been outlined in their respective sections; all values have been checked by Forest Research experts.

4.2 Introduction to the modelling

ESC: The Ecological Site Classification model (ESC; Pyatt et al. 2001) assesses tree species suitability (0-1) and forest productivity (Yield Class, m3ha-1yr-1) for a given site based on six climatic and soil variables (accumulated temperature, moisture deficit, continentality, windiness, soil moisture regime and soil nutrient regime). This model is used within the IMP chain to identify the most suitable species to plant given different priorities for forest management. It is also used to guide where climax vegetation is likely to be trees (if suitability is sufficient) or short scrub (if suitability is low) following land abandonment.

ESC was implemented as an R script at a 250m resolution grid across Wales for eleven key tree species, using CHESS climate variables for the baseline period (1981-2000) to derive accumulated temperature (day degrees greater than 5 degrees centigrade) and moisture deficit (the balance between evaporation and precipitation in the growing season (April to September)). Continentality (seasonal variability of the climate, CONRAD Index, follows Birse (1971) and Bendelow and Hartnup (1980)) and windiness (DAMS, detailed aspect method of scoring, Quine and White (1993, 1994)) scores were used. Soil Moisture Regime and Soil Nutrient Regime were derived from the dominant soil properties type as identified in the Cranfield dataset (1:250,000).

For each of three forest types the highest yielding species (conifers and short rotation forestry, SRF) or most ecologically suitable species (broadleaf) was used to represent that forest type for each cell: productive conifers (Sitka spruce (Picea sitchensis), Douglas fir (Pseudotsuga menziesii), Scots pine (Pinus sylvestris)); native broadleaves (oak (Quercus petraea, Quercus robur), beech (Fagus sylvatica), aspen (Populus tremula), birch (Betula pendula, Betula pubescens); short rotation forestry (Sitka spruce (Picea sitchensis), red alder (Alnus rubra), poplar (Populus nigra)).

Using these three forest types, five forest management-types were simulated: productive conifers (thin-fell & Low Impact Silvicultural Systems (LISS); native broadleaves (LISS & no-thin-no-fell); Short Rotation Forestry (SRF – a 25-year rotation).

The IMP chain integrates the CARBINE model with ESC, to calculate carbon storage in woodland and wood-related products. Tree species, management, yield class, climate zone, soil class and previous land use were used to lookup the carbon and greenhouse gas (GHG) balances of the forestry systems as calculated by the CARBINE model.

CARBINE: The CARBINE forest sector carbon accounting model calculates the development of carbon stocks over time in all key woodland carbon pools (trees, deadwood, litter, soil); the production of wood over time, representing key raw product types (sawlogs, small roundwood and bark), GHG emissions from fuels, materials and machinery involved in creating and managing the woodlands (Matthews et al., 2021a,b). Carbon and GHG assessments were made for each of the five forest management types, for each 250m grid location and for three time horizons (2020-2025, 2026-2050 and 2051-2100).

For ERAMMP, this model has been extended by the addition of a module for the calculation of the Net Present Value (NPV) of forestry in a way that is comparable with the NPV values calculated in an agricultural context.

Economic NPV: Forest economic values were calculated using costs for establishment, management, and harvesting, per hectare, based on a central estimate of 2015 data. Harvested timber product volumes from CARBINE and product values from 2015 were used to calculate revenue (Saraev, 2017a,b). NPV were calculated according to Green Book methods using Green Book discount rates (HM Treasury, 2018) annualised over the length of the rotation.

Linking woodland data across scales (GLUE): ESC-CARBINE-NPV data are all run at a 250m resolution. Additional code was developed to average the woodland data to the sub-farm resolution of Decision-Making Units (DMU) used by other models.

4.2.1 Peer Review: ESC & CARBINE

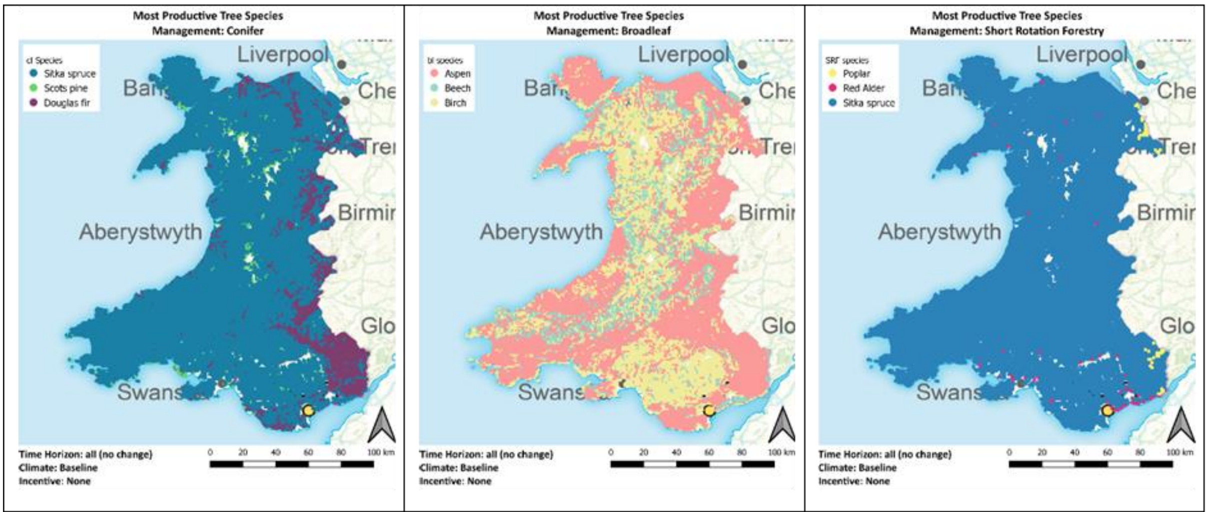
The ESC model was developed by Forest Research to assess tree species suitability for sites across the UK (Pyatt et al. 2001) and has been developed to incorporate climate change projections to explore the impacts of future climate on tree species suitability (Broadmeadow et al., 2009; Ray, 2008a, 2008b; Ray et al., 2010). The ESC model has been used to research tree health impacts on UK woodland composition and species alternatives (Broome et al., 2018; Ray et al., 2021) and to explore how changes in forest management and climate change will impact the provision of ecosystem services, to support regional and national policy development (Beauchamp et al., 2016; Ray et al., 2015; Ray et al., 2019). The ESC decision support tool was developed to support practitioner decision-making and is used widely across the forestry sector by the Forestry Commission and private managers to support planning. Research engagement with practitioners and surveys of established stands allows feedback into model validation and refinement. The ESC model and outputs have therefore been extensively peer-reviewed and validated on the ground.

The CARBINE model has been applied in national greenhouse gas inventories under the United Nations Framework Convention on Climate Change (Brown et al., 2021) and forms the basis of the UK's GHG emissions and removals due to afforestation, deforestation, and forest management under the Kyoto Protocol (Thomson et al., 2020). Values also underpin the UK Woodland Carbon Code (Jenkins et al., 2018; UKWCC, 2020, 2021). Model values have been extensively validated through literature review and national and international assessment, since the model was first produced in 1988 (Thompson & Matthews, 1989). Mensuration values have been extensively validated through field assessments and publication (Matthews et al., 2016).

Economic models of forest management have been validated through publication and peer review (Saraev 2017a, b). Net Present Value was calculated according to Green Book guidelines (HM Treasury 2018). Methods are consistent with published approaches (e.g. Hardaker, 2018). References to further papers are included in the forestry section of the reference list at the end of this report.

4.2.2 Building Understanding: Understanding the baseline distribution of the highest yielding species across Wales from ESC outputs.

Figure 4.1 shows which species is the highest yielding for each forest type, under the baseline climate. This represents the most likely species that would be planted by a landowner for each forest management type, and so is indicative of the forest type.



1 Productive Conifers 2 ■ Sitka Spruce 3 ■ Scots pine 4 ■ Douglas fir	5 Native Broadleaves 6 ■ Aspen 7 ■ Beech 8 ■ Birch 9 * Oak not selected	10 Short Rotation Forestry 11 ■ Poplar 12 ■ Red Alder 13 ■ Sitka Spruce

Figure 4.1 The most productive (highest yielding) tree species, as assessed by ESC for the baseline climate period, is selected at 250m resolution for each of the three forest types.

The take-home-messages are that:

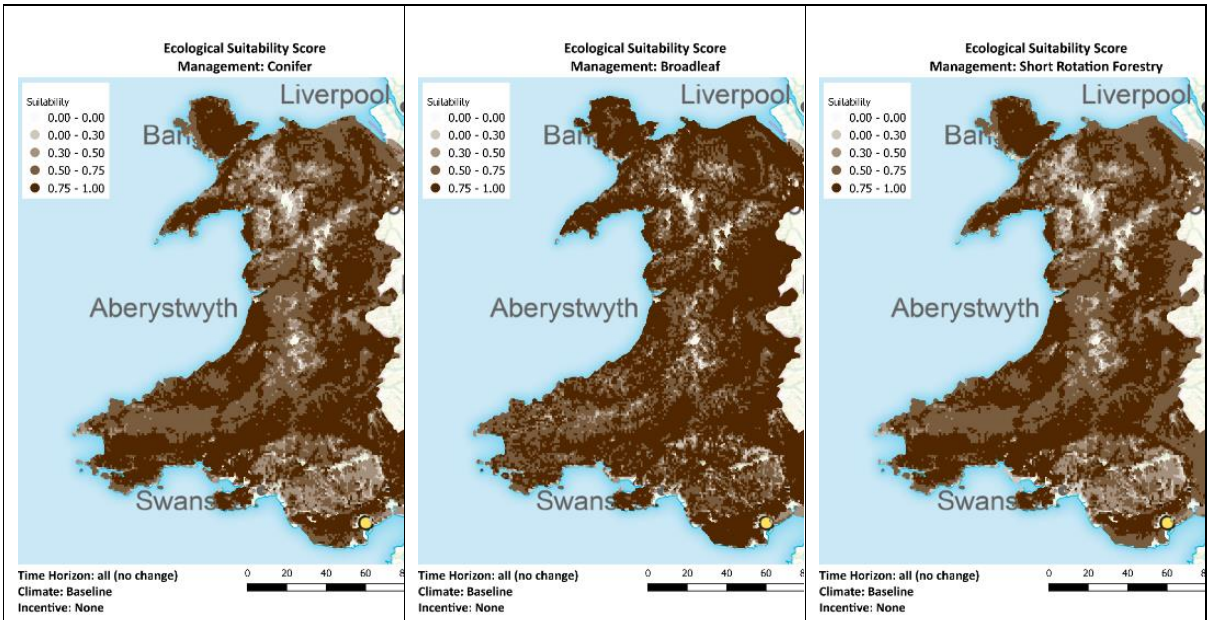
- For **productive conifers**: Sitka spruce is selected as the highest yielding species across most of Wales, with Scots pine selected in drier locations and Douglas fir in eastern locations on rich soils.
- For **native broadleaves**: Aspen is selected in lowland areas and beech at higher elevations with scattered areas of birch. Oak is not selected in this scenario as it has a lower yield class.

- For **short rotation forestry**: Sitka spruce is widely selected as the highest yielding species with only very small patches of poplar or red alder.

4.2.3 Building Understanding: Understanding the distribution of ESC suitability scores for each forest type

The ESC suitability score identifies how suitable a given tree species is for a site, on a scale from 0 to 1, with <0.3 unsuitable, 0.3 – 0.5 marginal, 0.5 – 0.75 suitable, and 0.75 – 1 as very suitable. Sites are only modelled as suitable for broadleaf woodland if the suitability score is greater than 0.3 (marginal, suitable or very suitable), and for productive conifers if greater than 0.5 (suitable or very suitable).

Figure 4.2 shows modelled suitability is highest in lowland areas. Suitability for all forest types decreases as elevation and windiness increase and as temperatures decrease. Soil and rainfall also affect suitability. Average suitability is similar across all forest types (0.70 – 0.83 Table 4.1 ; equivalent to 'suitable' in ESC).



Conifer	Broadleaf	Short Rotation Forestry
■ 0 | ■ 0.0-0.3 | ■ 0.3-0.5 | ■ 0.5-0.75 | ■ 0.75-1		

Figure 4.2 Ecological tree species suitability for the selected tree species (see Figure 4.1) for each of the three forest types, for the baseline climate period, no climate change (ESC score 0 unsuitable to 1 very suitable)

Table 4.1 Statistical overview of ESC suitability value and yield class by forest type

	Suitability		Yield Class	
	Mean	**Median**	**Mean**	**Median**
Broadleaf	0.79	0.83	8.1	8
Conifer	0.72	0.75	20.0	21
Short Rotation Forestry	0.70	0.72	19.6	21

4.2.4 Building Understanding: Understanding levels of baseline productivity (yield class and wood production) from ESC

Forest productivity (growth) is measured as the yield class (m3ha-1yr-1), which in the models is calculated as the product of the maximum potential yield class and the ESC suitability score for each species. As with suitability, modelled yield class is highest in lowland areas and decreases with increasing elevation and windiness and decreasing temperatures for all forest types (Figure 4.3). Modelled yield class is highest for productive conifers and SRF (means of 20 and 19.6, respectively) than for broadleaves (mean yield class 8.1) as conifers have a higher yield class than broadleaves, even where suitability scores are the same or higher; and in particular due to the prevalence of Sitka spruce which is high yielding.

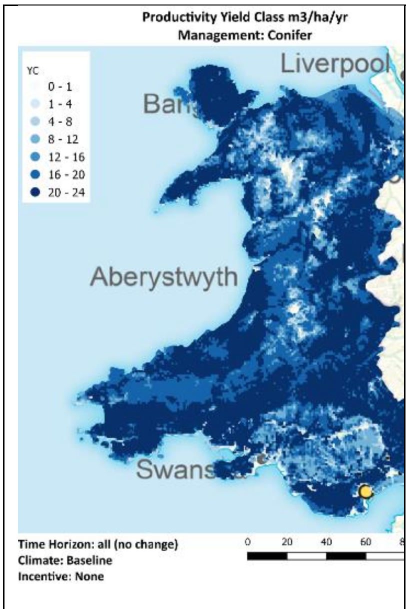
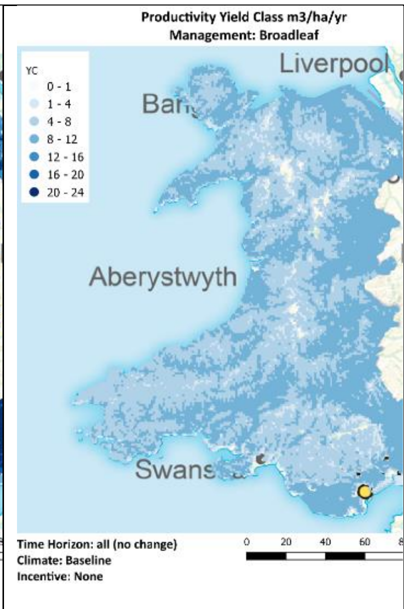
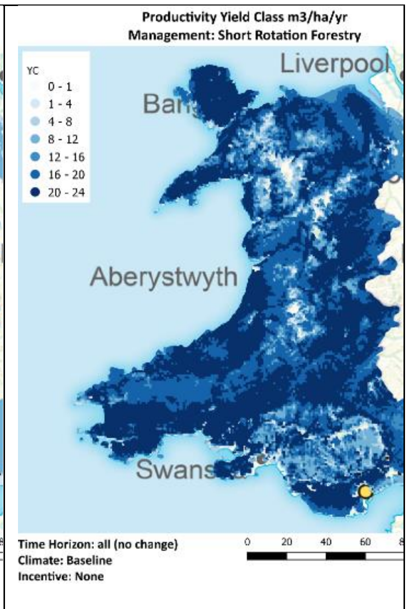
		
Conifer	Broadleaf	Short Rotation Forestry
0-1 | 1-4 | 4-8 | 8-12 | 12-16 | 16-20 | 20-24		

Figure 4.3 Forest productivity; the highest potential yield class for each forest type. Yield class a measure of growth or timber production, measured in m3ha-1yr-1

4.2.5 Building Understanding: Understanding the carbon sequestration potential of new woodland from CARBINE

Potential carbon sequestration of new woodland is calculated as the sum of a) the change in carbon stock in trees, deadwood and litter; b) the change in carbon stock in soil; c) the change in carbon stock in harvested wood products; and d) the GHG emissions due to forest operations. Note these changes do not include any changes in carbon sequestration resulting from changes in climate or atmospheric composition.

As shown in Figure 4.4, SRF management results in net emissions in lower yielding areas due to emissions from forest management and where the carbon in harvested wood products is lost when material is burned. Conifer stands managed under Low Impact Silvicultural Systems (LISS) have lower levels of harvesting and higher levels of carbon on-site per hectare than conifer stands managed under clear-fell. Likewise, unmanaged broadleaf stands (no-thin-no-fell) have lower levels of harvesting than those managed under LISS, therefore results show higher carbon sequestration per hectare. Conifer stands managed under LISS sequester more carbon than broadleaf LISS stands due to faster growth rates.

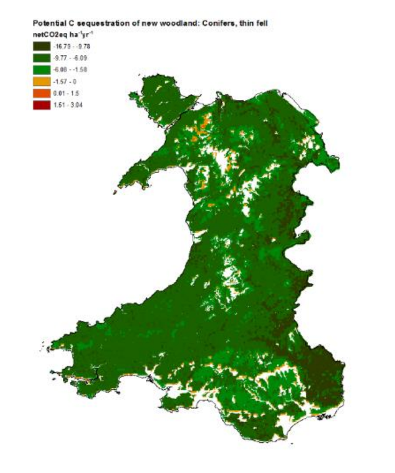
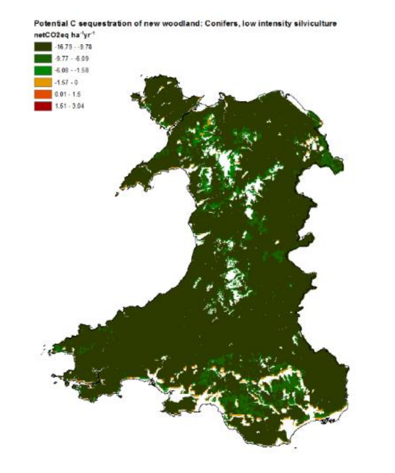
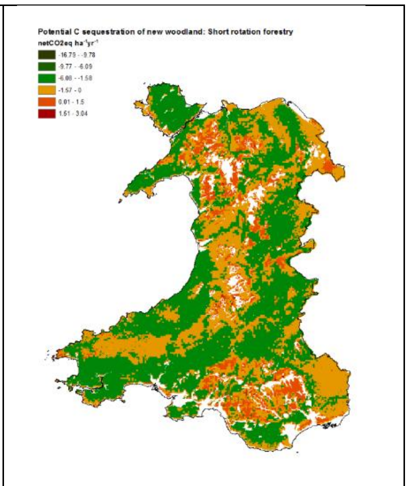
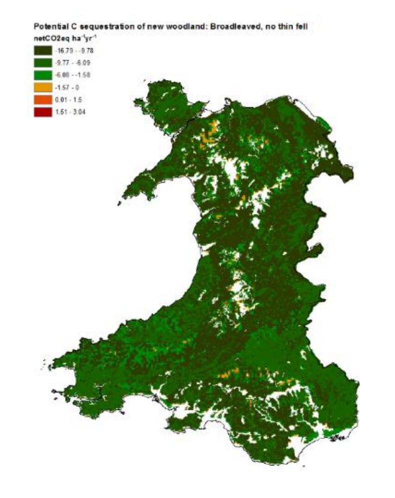
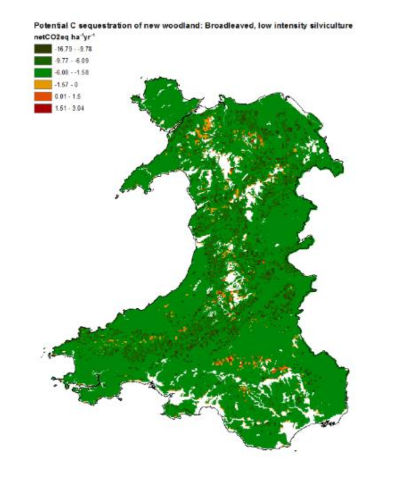
		
Conifer (thin fell)	Conifer (low intensity)	Short Rotation Forestry
		
Broadleaved (no thin fell)	Broadleaved (low intensity)	
■ -16.79 to -9.78 | ■ -9.77 to -6.09 | ■ -6.08 to -1.58 | ■ -1.57 to 0 | ■ 0.01-1.5| ■ 1.61-3.04 Positive numbers are emissions / Negative are sequestration.		

Figure 4.4 Carbon sequestration (tCO2eqv/ha/yr) by forest type and management: a) conifer thin fell, b) conifer LISS, c) broadleaf LISS, d) broadleaf no-thin-no-fell, e) short rotation forestry). Green represents net sequestration, red net emissions

4.2.6 Building Understanding: Wood production by forest type and management

Figure 4.5 shows the average annual wood production (harvested timber) to 2100 for the four forest management scenarios where timber is harvested (i.e. excluding broadleaf no-thin-no-fell). Conifer forests under a standard thinning regime and felled at the end of the rotation have the highest potential for wood production per hectare per year, followed by short rotation forestry (Figure 4.5). This is because conifer species are faster growing (higher yielding) and have shorter rotation lengths. Broadleaved woodlands under LISS management produce the lowest volume in these scenarios. These values influence forest economics and NPV. There is no harvesting and no wood production in the broadleaf no-thin-no-fell scenario (not shown).

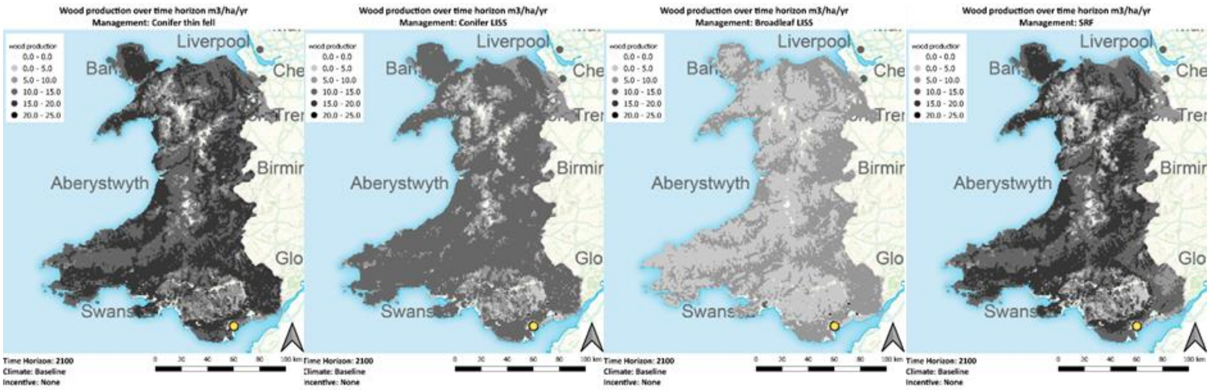


Figure 4.5 Modelled wood (timber) production to 2100 (m3/ha/yr) by forest type and management, for conifer thin fell, conifer LISS, broadleaf LISS, SRF; there is no harvesting in broadleaf no-thin-no-fell

Table 4.2 presents the mean values of timber harvested across Wales, in cubic metres per hectare per year; values to 2100 are averages of those in Figure 4.5. Values to 2050 are lower, as less material is harvested, mainly thinning for conifer stands under LISS and thin-fell, and broadleaf LISS stands, and SRF which goes through a single cycle (25 years). Conifer stands have completed one full rotation to 2100, SRF 3 cycles, and broadleaf stands are yet to complete a full harvesting cycle.

Table 4.2 Mean wood production (harvested timber, m3ha-1yr) across Wales by forest management type, for 2050 and 2100.

Mean Wood production	**2050**	**2100**
Broadleaf LISS	3.1	4.7
Broadleaf no-thin-no-fell	0	0
Conifer LISS	6.1	10.2
Conifer thin fell	6.1	13.8
Short Rotation	11.8	13.2

4.2.7 Building Understanding: Forest Economics (CARBINE-NPV)

Figure 4.6 shows the NPV calculated for discount rates of 0 and 3.5 for different forest and management types. At a discount rate of 0, NPV declines from: conifer thin fell, to conifer LISS, to broadleaf LISS (positive NPV values); to broadleaf no-thin-no-fell and SRF (negative values). As discount rate increases, NPV decreases (Table 4.3). At a discount rate of 5, conifer forests have a positive NPV, but others are negative. NPV varies spatially in the same pattern as wood production, suitability and yield class (Figure 4.5, Table 4.1).

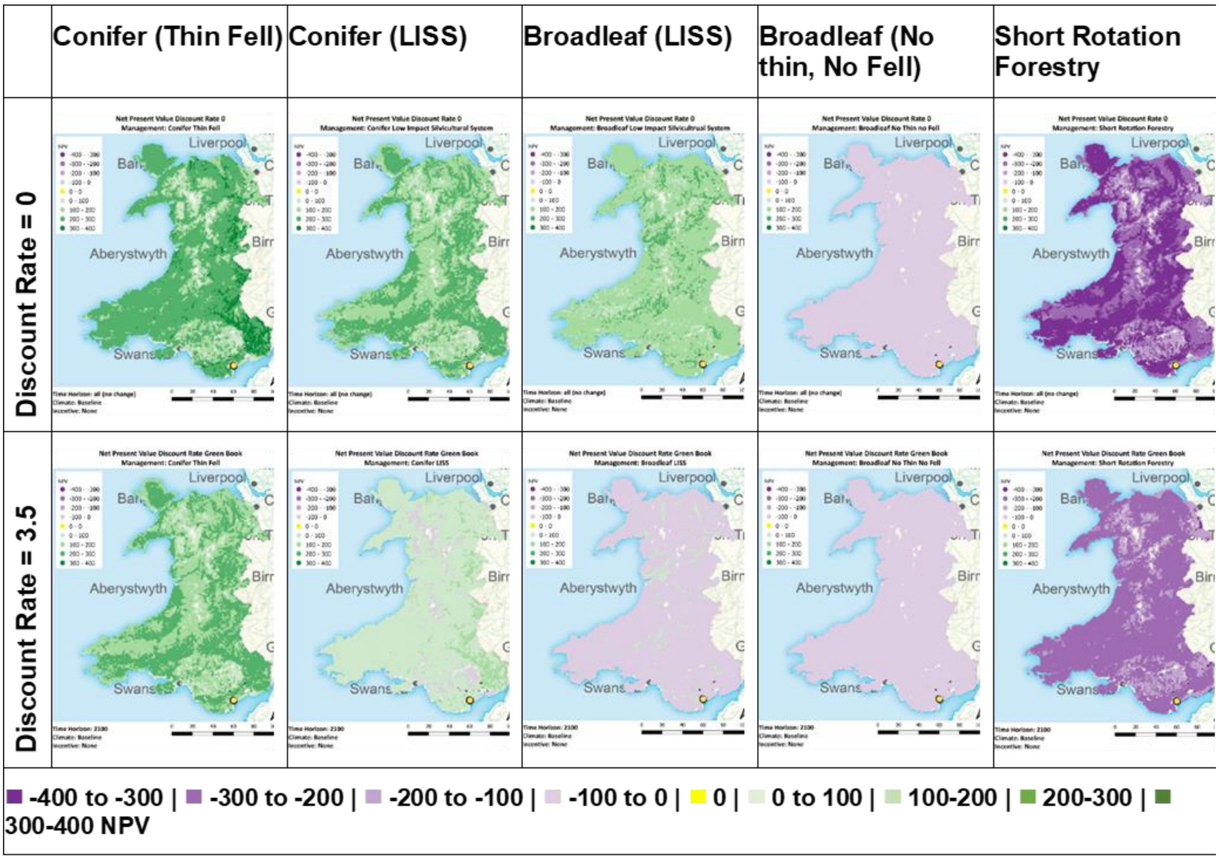


Figure 4.6 Net present values calculated by forest type and management, for conifer thin fell, conifer LISS, broadleaf LISS, SRF; there is no harvesting in broadleaf no-thin-no-fell for a discount rate of zero and the Green book discount rate of 3.5%.

Table 4.3 Mean Net Present Value by forest management type for three different discounting rates (zero, green book (3.5%) and 5%).

Mean NPV	NPV0	NPV3.5	NPV5
Conifer thin fell	217.4	74.3	27.5
Conifer LISS	175.7	52.8	14.6
Broadleaf LISS	170.6	-15.8	-92.8
Broadleaf no-thin-no-fell	-10.1	-66.2	-102.0
Short Rotation Forestry	-283.4	-231.6	-218.8

4.2.8 Error checking: ESC downscaling process

The ESC-CARBINE_NPV outputs were scaled proportionately from the 250m resolution to the resolution of the DMUs, which is used by the downstream models. During the scaling process, areas ecologically unsuitable or unavailable for forestry due to planting restrictions are removed. Figure 4.7 outlines the scaling process.

As a QA step the downscaled data were sense checked for common errors. For example, checks were conducted to ensure plantable area was less than the DMU area and ensure that values are within the range of minimum and maximum values. Then, from the resultant datacube, random DMUs were chosen and overlain in ArcGIS with the original ESC-carbine data for a manual comparison of the results. The sense checking and validation steps show that species choice, suitability scores and yield class at 250m and DMU level are all within expected ranges and follow expected geographic and climatic trends. These values underpin the timber production volumes from the different forest management scenarios, which in turn underpin the values for carbon sequestration and annualised NPV; these values are also within expected limits and follow expected spatial trends and differences between forest and management type. NPV values were compared to farm economic values to confirm NPV value ranges were within the appropriate range.

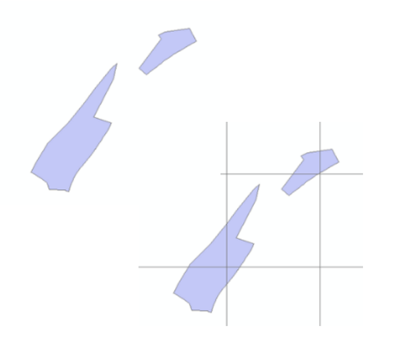
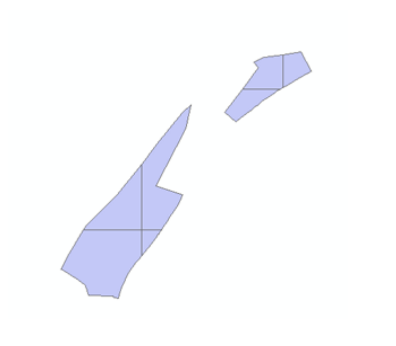
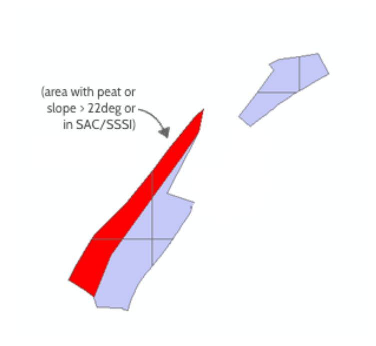
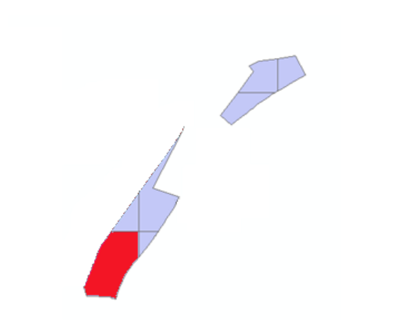
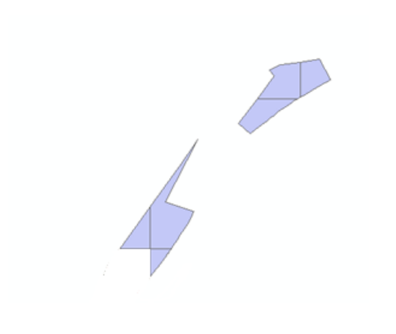
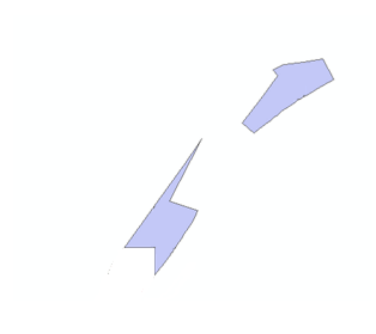
		
The ESC carbine data cube is converted into a spatial grid and intersected with the DMU polygons	This creates a spatial dataset with multiple polygons per DMU. Each new polygon has the attributes of the ESC-carbine-NPV datacube	Areas with restrictions are deleted: peat soils, slope > 22°, SACs (buffered by 500m), SSSIs (buffered by 300m)
		
Areas ecologically unsuitable for forestry are discarded; ESC suitability score thresholds: conifers >=0.5; broadleaf >=0.3	Each remaining polygon has ESC-CARBINE-NPV values calculated for the partial 250m grid squares suitable for forest	The polygons are then aggregated at the DMU level and spatial average values calculated

Figure 4.7 Overview of the scaling process used to convert ESC-CARBINE-NPV data from 250m to DMU resolution.

5 Land Allocation Module (LAM)

Authors: Mike Hollaway and Ian Holman

5.1 Introduction to the model QA

The Land Allocation Model (LAM) uses the outputs of the upstream models to create the final allocation of land use that is passed to the downstream group of models. To improve understanding and confidence in the model, the following steps have been undertaken:

- Validation/ Contextualisation: the LAM allocations for baseline were compared with the Newcastle Brexit Report (Section 5.2.1) and the Welsh Farm Business Survey (Section 5.2.2);
- Expert Assessment (External): the assumptions underlying the LAM were given explicit scrutiny by the Welsh Government expert Group (Section 5.2.3);
- Sensitivity analysis: the influence of changing key LAM parameter values was assessed (Section 5.2.4);
- Expert Assessment (Consortium): expert sense-checking of LAM outputs for the land use scenarios (Section 5.2.5).

5.2 Introduction to the modelling

The LAM is a heuristic-based decision model that selects the expected long-term outcome for each farm (and their associated land use and livestock selections) across Wales. The LAM considers the optimised SFARMOD farm solutions (crop types/areas, grassland types/areas, stocking types/numbers, fertilisation etc) under a given scenario – these include the optimised solution for a holding's current farm type (e.g. dairy), but also for all alternative farm types (e.g. mixed livestock, specialist sheep etc). The LAM estimates the Farm Business Income (FBI) for each farm based on the SFARMOD net farm profit, allowing for miscellaneous non-agricultural income, unpaid labour and those fixed costs that do not change directly with farm plans (e.g. land ownership or tax). The latter three components not simulated by SFARMOD are calculated according to farm-type specific "All sizes" data from the Welsh Farm Business Survey. For those farm types for which there are no data in the Welsh Farm Business Survey (Cereals and General cropping), data from agro-climatically similar regions in England is used.

The LAM recognises that there are complex human and financial factors that affect the likelihood of a change in farm type in response to changing economic circumstances. These are reflected in the rules and FBI thresholds that are used to identify farms under economic pressure; farms that remain in their current farm type; and farms that transition to more profitable alternative farm types. Farms under economic pressure fail to achieve a simulated FBI of less than £6000 p.a. (based on Hubbard, 2019; Hubbard et al., 2018) and either leave full-time agriculture, are sold and converted to an alternative farm type or are afforested, depending on the viability and environmental suitability of alternative farm types and forestry (from ESC-CARBINE). Farms remain in their current farm type if they exceed the minimum FBI threshold, but either fail to achieve an FBI that provides the financial resources needed to change farming system or if there is no sufficiently financially attractive alternative to incentivise transition. Farms transition to a more profitable alternative farm types if the FBI uplift is sufficient to both incentivise change and to meet the cost of additional borrowing required to make the change (based on Welsh Farm Business Survey data and Andrew Moxley pers. comm).

The LAM sits between the upstream forest and agricultural profitability models and the downstream ecosystem service and biodiversity modelling. The model is responsible for the final allocation of land use that is passed downstream to the other models. The LAM passes through the SFARMOD (cropping, grassland, livestock and nutrients) and ESC-CARBINE outputs for each baseline ERAMMP Farm Type directly to downstream models.

5.2.1 Validation: Comparison of Welsh Farm Business Income from the LAM and Newcastle University Brexit Report from 2019

The simulated distribution of FBI from the LAM was compared to that in the Newcastle University 'Brexit' report (Hubbard, 2019) which was based on Farm Business Survey data and a budgetary simulated model (Figure 5.1). The apparent over-estimation of baseline FBI by the LAM through the lower half of the distribution can be explained, at least in part, by:

1. The LAM using SFARMOD optimised outputs that assume farmers are profit maximisers;
2. The Newcastle report including "part-time" farmers who represented almost 9% of their sample. Only farms with >1FTE are included within the IMP;
3. The different, although overlapping, baseline economic periods (2015/16 vs 2013/14-2015/16).

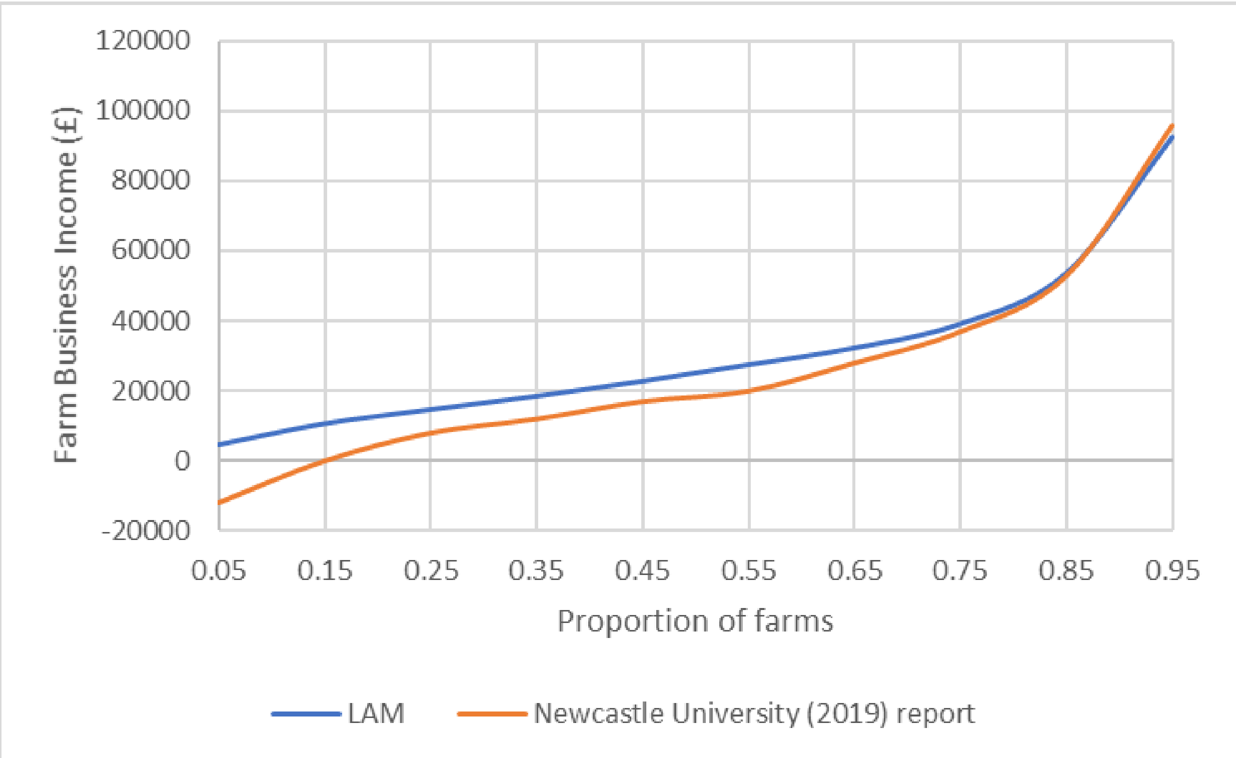


Figure 5.1 Comparison of estimated Welsh Farm Business Income distribution from the LAM and Newcastle University (2019) report.

5.2.2 Validation: Comparison with Welsh Farm Business Survey

Figure 5.2 provides a comparison of the median simulated FBI from the LAM (green bars) against the "All sizes" average FBI from the Welsh Farm Business Survey (2015/16, blue bars), with the uncertainty bars providing the range in FBI across the period 2012/13 to 2017/18. Blue Welsh FBS bars are missing where there is no equivalent farm type reported (e.g. cereals and general cropping). The graphs show the high annual volatility in the actual FBI of both lowland and upland/hill dairy farms compared to all other farm types in the Welsh Farm Business Survey (FBS). Farm planning decisions are about the unknown future, so SFARMOD basis its decisions on price and yield expectations (rather than future actual prices) in a similar manner to the subjective expectations of farmers. Our expected prices are informed by historic prices going back a number of prior years using the John Nix and ABC pocket books. In contrast, the Welsh farm Business survey is based on a statistical sample of historical facts.

Direct comparisons are impacted by several methodological uncertainties related to i) the actual costs incurred, and output prices received by farms throughout the year, ii) farm-specific uncertainties in unpaid labour, miscellaneous non-agricultural income etc used by the LAM in reflecting all full-time farms; and iii) uncertainties within the Welsh FBS data. However, whilst there are inevitable mismatches between the LAM and the Welsh FBS data, given these uncertainties, it is encouraging to see that the LAM outputs are within the reported range of recent years for all farm types (Figure 5.2).

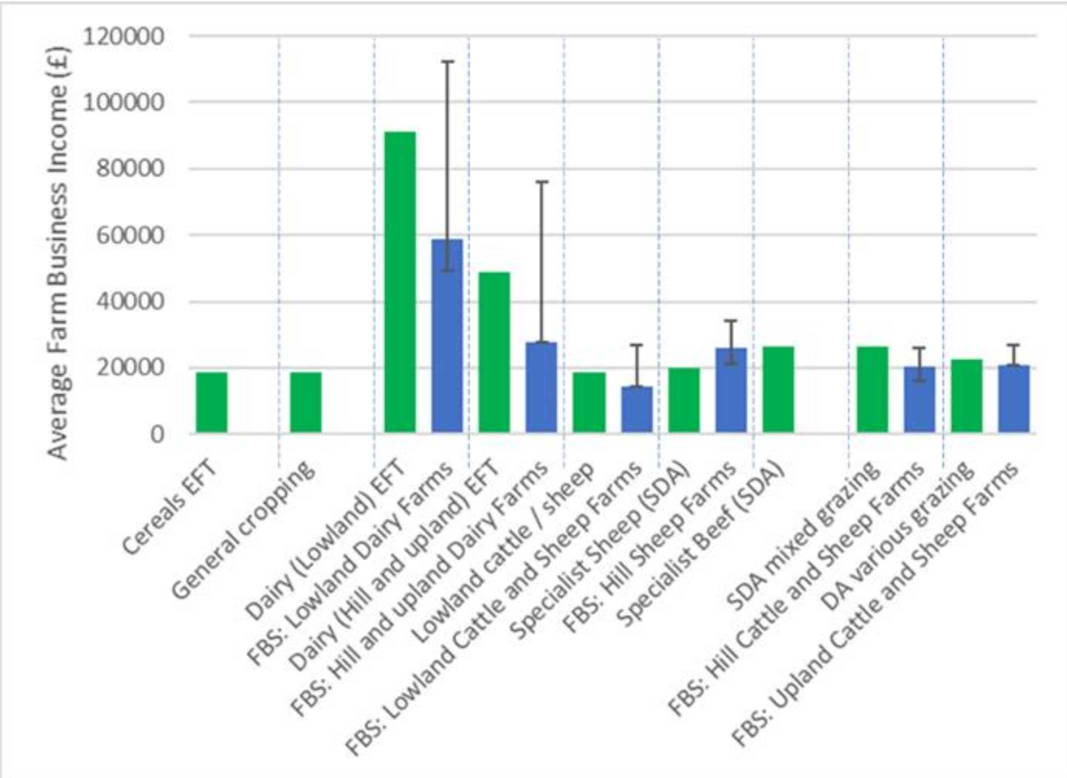


Figure 5.2 Comparison of simulated median Farm Business Income from the LAM (green bars) against the "All sizes" average Farm Business Income from the Welsh Farm Business Survey (2015/16, blue bars) by farm type, with the uncertainty bars indicating the range in Farm Business Income across the period 2012/13 to 2017/18

5.2.3 Expert Assessment (External) Welsh Government Expert Group consultation on LAM decision rules and thresholds

LAM decision rules and key decision thresholds have been reviewed and agreed with a Welsh Government expert group. Given the diverse and farm-specific decision contexts associated with farm transition (e.g. farmer age, indebtedness, values, personal preferences etc), it was agreed that the LAM does not attempt to predict farm type change over a fixed given time horizon; nor is it trying to predict individual farmer behaviour. The implementation of the rules and thresholds by scientists at UKCEH in the IMP were QA-ed by scientists at Cranfield University through examination of LAM output.

5.2.4 Sensitivity analysis: Key LAM parameters assessed through sensitivity analysis

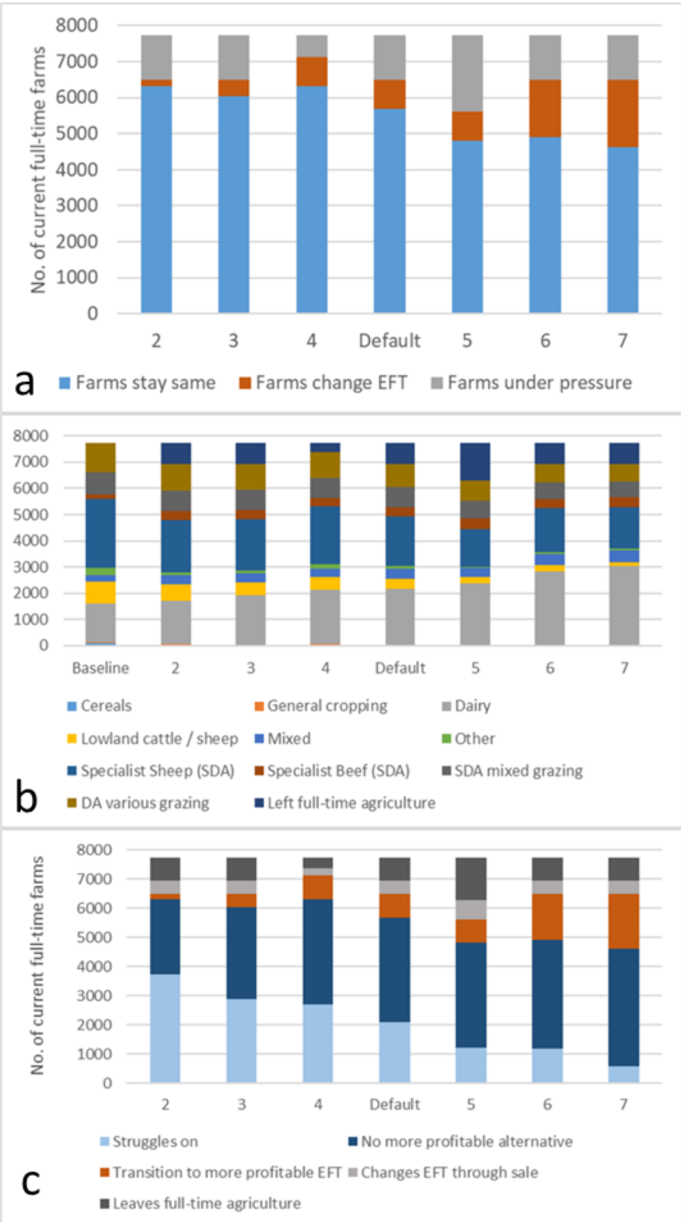
Sensitivity analysis of the LAM was performed on the No Basic Payment Scheme (NoBPS) runs of SFARMOD. Three settings influence potential change in the farm type in the LAM:

1. The minimum annual FBI required to continue full-time farming:
    - This was agreed at a low value of £6,000 p.a. to represent the tenacity of many farmers to keep farming.
2. A minimum annual FBI for a farm to be performing at a level to contemplate transition was agreed as the national minimum wage for 1 FTE (~£13,000 p.a.).
3. The required increase in annual FBI to make farm type transition sufficiently attractive. This has two components:
    - An agreed minimum FBI uplift required to motivate transition of £5,000 p.a. or 25% of the baseline farm type's FBI;
    - An agreed additional annual FBI uplift to finance any increased tenants' capital requirement for farm type transition as 10% of increased tenants' capital requirement.

The LAM sensitivity analysis explored how changing these critical thresholds or percentages affected the outcome of ERAMMP Farm Type (EFT) changes across seven bundles of LAM settings, in which the agreed LAM thresholds were modified by up to +/- 50%. These seven bundles, in which between one and three LAM thresholds were modified in each run, were chosen to make deliberate EFT change more or less difficult and enforced EFT change more or less likely. These analyses (labelled Default and 2-7) are described below:

1. Default values;
2. Making deliberate EFT change more difficult – increased minimum FBI to contemplate transition from £13,000 to £19,500; increase tenants capital multiplier from 0.1 to 0.15 and change minimum FBI uplift to greater of £10,000 or 25%;
3. Making deliberate EFT change slightly more difficult – increased minimum FBI to contemplate transition from £13,000 to £16,000; keep tenants capital multiplier to 0.1 and change minimum FBI uplift to greater of £7,500 or 25%;
4. Making enforced change more unlikely – reduce minimum FBI viability threshold from £6,000 to £3,000;
5. Making enforced change more likely - increase minimum FBI viability threshold from £6,000 to £9,000;
6. Making deliberate EFT change slightly more likely - reduce minimum FBI to contemplate transition from £13,000 to £10,000; decrease tenants capital multiplier from 0.1 to 0.05 and keep minimum FBI uplift to greater of £5,000 or 25%;
7. Making deliberate EFT change more likely - reduce minimum FBI to contemplate transition from £13,000 to £8,000; decrease tenants capital multiplier from 0.1 to 0.05 and change minimum FBI uplift to greater of £5,000 or 20%.

Figure 5.3 shows how the number of farms within LAM transition types and EFTs changes with the different bundles of LAM thresholds. Despite the large changes in LAM settings across the seven runs as outlined above, Figure 5.3a shows that the high-level outcomes do not fundamentally change from the default LAM NoBPS result of most current full-time farms continuing in their current EFT ("farms stay same"), with limited numbers of farms being able to plan a deliberate transition to a more profitable EFT ("farms change EFT"). However, the setting does significantly change the relative balance in Figure 5.3c between farms whose EFT is unchanged due to "struggling on" (making sufficient FBI to keep going but not enough to contemplate transition) and those for which the economic incentive to transition is insufficient ("no more profitable alternative"). The large changes in LAM settings across the runs do not fundamentally change the default LAM NoBPS distribution of EFTs in Figure 5.3b. However, as the settings are changed to make deliberate transitions more likely (runs #6 and 7 compared to runs #3 and 2), the number of dairy farms increases as this is the most profitable alternative for many farms.



"Farms stay same" – EFT unchanged as either (1) "Farm struggles on" as FBI > minimum level but insufficient to enable transition or (2) FBI > minimum level and there is "no more profitable alternative" (taking account of transition requirements)

"Farms change EFT" – FBI of current EFT > minimum level but farm is able to "transition to more profitable EFT"

"Farms under pressure" – farms either (1) "leave full-time agriculture" (as there is no EFT that achieves minimum threshold) or (2) "Changes EFT through sale" as current EFT has FBI < minimum threshold but another EFT can achieve FBI > minimum threshold)

Figure 5.3 Change in distribution of number of farms within (a) simplified LAM transition classes, (b) full LAM transition classes across runs with a range of LAM settings and (c) ERAMMP Farm Types

Figure 5.4 explores the spatial sensitivity of LAM outcomes to changes in settings to assess the spatial coherence and plausibility of simulated EFT changes. The upper panels of Figure 5.4 show the simplified LAM transitions, focusing on the two runs in which deliberate/planned EFT change is made most difficult (#2) or more likely (#7). The opportunities for planned transitions to a more profitable alternative farm type (shown in blue) increase, but mostly in the low elevation regions. The lower panels relate to the simulations in which enforced change (through leaving full-time agriculture) is made more unlikely (lower minimum FBI in #4) or more likely (higher minimum FBI in #5). Consequently, the “farms under pressure” increase, mostly in the Severely Disadvantaged Areas.

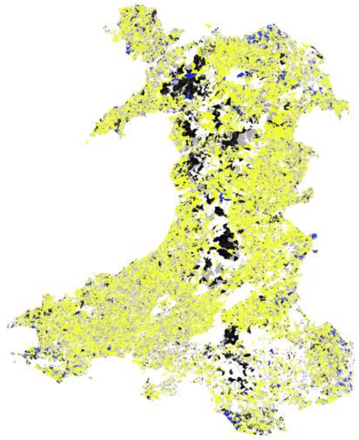
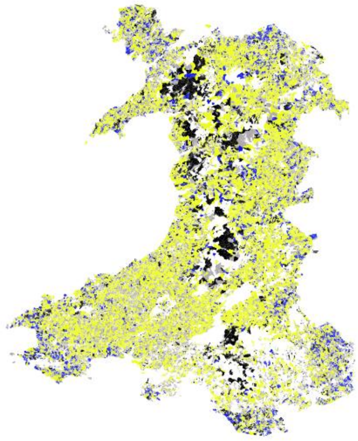
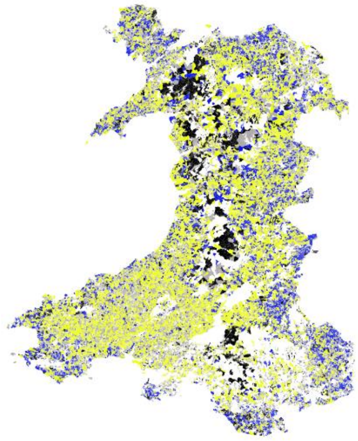
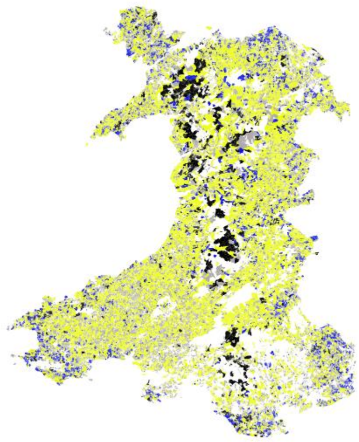
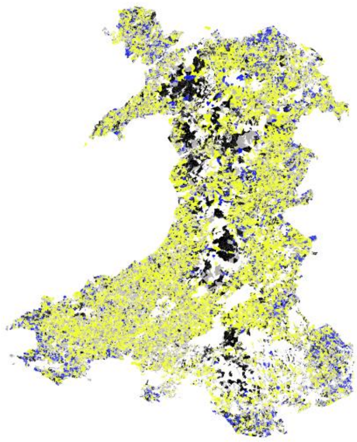
#2 Making deliberate EFT change more difficult	#1 Default	#7 Make deliberate EFT change more likely
		
#4 Making enforced change more unlikely		#5 Making enforced change more likely
	■ Yellow: Farms that have stayed the same ■ Blue: Farms that have changed EFT ■ Black: Farms under pressure	

Figure 5.4 Sensitivity of spatial transitions to changes in the LAM settings

5.2.5 Expert Assessment (Consortium): Sense-checking of land use scenario outputs

Quality assurance of the LAM outputs for the scenarios was by expert judgement, evaluating the magnitude and direction of change in modelled FBI for each farm type from the baseline against scenario changes to key SFARMOD input and output prices. The spatial distribution of farm type changes was also evaluated by expert judgement for plausible spatial coherence. In particular, that farms under pressure were mostly located within agroclimatically-marginal areas within the current Severely Disadvantaged Areas; and that simulated farms making deliberate transitions to more profitable EFTs were largely located within lowland regions with more productive agroclimates.

5.2.6 Validation: Comparison of baseline land-cover from SFARMOD/LAM and Land Cover Map with field observations at each quadrat location?

GMEP baseline plot locations were intersected with the SFARMOD/LAM baseline depiction of land-use types across Wales. In most cases there was high agreement (Figures 5.5 and 5.6). Lower agreement was seen in the more dynamic land use types likely to turnover at a greater rate spatially and through time. When broken down by the broad habitat to which the polygon containing each plot was assigned during field survey again most plot locations were in agreement indicating that the downstream MultiMOVE plant suitability model could process the upstream land use transitions from SFARMOD/LAM given consistency between the two source of land cover information (Figure 5.5).

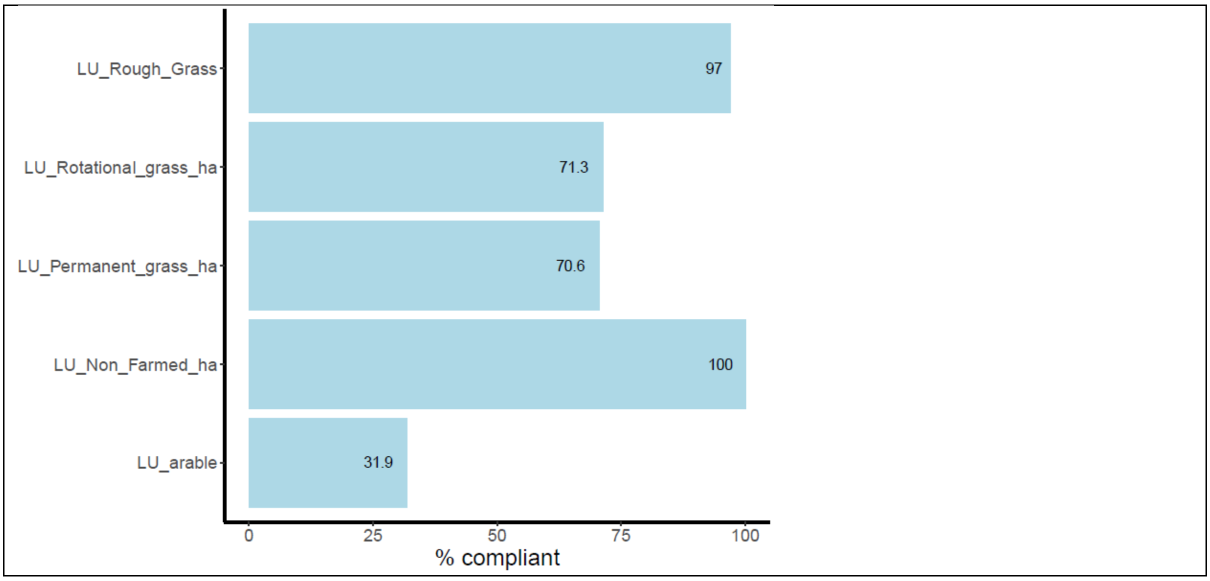


Figure 5.5 Percentage of GMEP baseline plot locations coinciding with each SFARMOD/LAM land use type where this type agreed with the land cover observed during field survey between 2013-16.

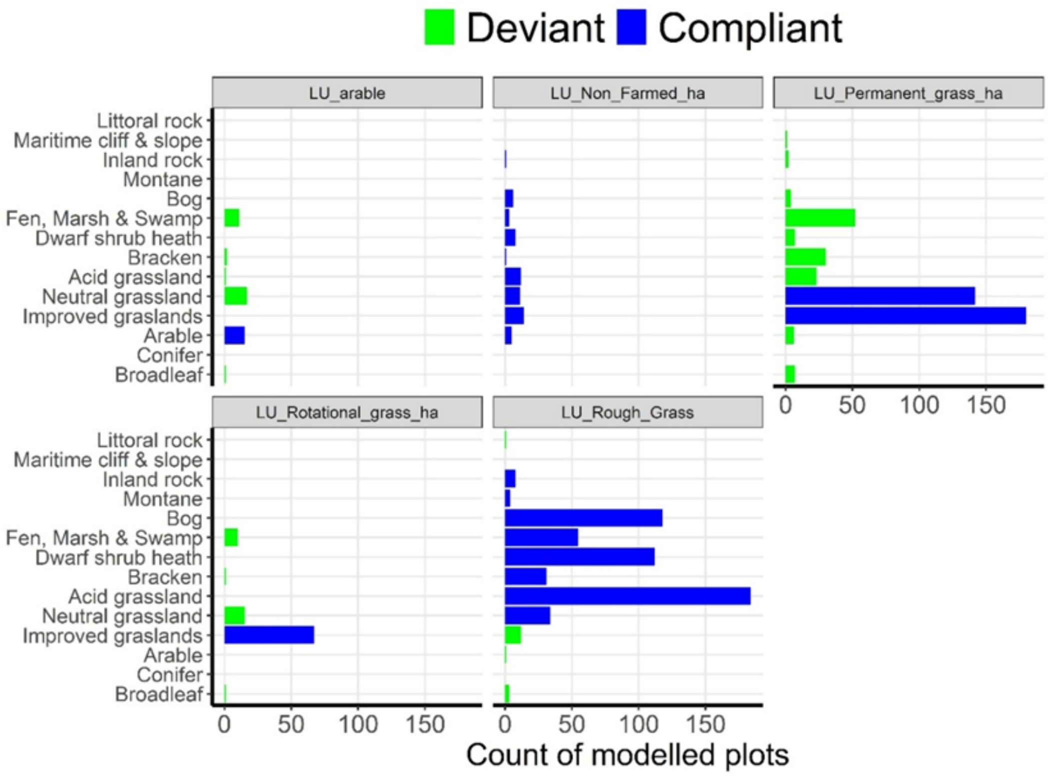


Figure 5.6 Count of GMEP baseline plots by SFARMOD/LAM land use type and by broad habitat as assigned by the field surveyors when visiting each plot in 2013-16. Blue bars indicate agreement between the land cover/land use types and denote plots where land use transitions were modelled under scenarios of land use change. Green bars indicate where SFARMOD/LAM and field survey did not match. These locations remained stable and no modelled change was applied.

6 Biodiversity: Birds – BTO models

Authors: Joe Cooper and Gavin Siriwardena

6.1 Introduction to the model QA

This section presents the models used to investigate the influence of land allocation on bird abundance and distribution, and the quality assurance methods used to assess the outputs. These models use volunteer survey datasets from the BTO/JNCC/RSPB within 1 km squares across Wales. The quality assurance methods used are:

- Peer review: The modelling approach has been applied in other contexts and published in the academic literature.
- Expert Assessment: The baseline results have been presented and discussed in Section 6.2.2.
- Error Checking: Models were checked for the influence of outliers (Section 6.2.3).
- Validation: Population estimates were compared against percentage occurrence from Breeding Bird Surveys and results for habitat specialist species were sense-checked by experts (Section 6.2.4).

6.2 Introduction to the modelling

The BTO modelling framework is designed to assess the impact of land-use change on bird populations. Species-specific models were developed using a combination of landscape composition datasets and volunteer bird survey data from the BTO/JNCC/RSPB Breeding Bird Survey (BBS) (Freeman et al., 2007). Bird surveys were conducted within 315 spatially randomised 1 km squares across Wales, with counts extracted for 68 species from the period 2013-2018 using previously established methods (Plummer et al., 2020).

Landscape composition was first assessed from 1 km square summaries of SFARMOD-LAM outputs. Land not covered by SFARMOD was summarised from the National Forest Inventory, Land Cover Map Plus (Rowland et al., 2017), Woody Cover Product (Schofield, 2016), Detailed River Network and Digital Terrain Map. This resulted in 34 landscape metrics for spatial cover, land characteristics and farming intensity for each 1km square of Wales, for the baseline and each land use scenario. A generalised linear modelling (GLM) framework was used to produce unique models for each species in R. Error structure (Poisson or negative binomial) was selected that best accounted for overdispersion and zero-inflation present in the data, and model covariates were those which had a significant influence on species abundance (as detailed in Plummer et al., 2020).

Models are applied to the baseline dataset (consisting of all 1 km squares across Wales), using the R function, predict.glm. This provided bird counts for each square, which were summed to provide an overall population estimate for each scenario, with confidence intervals constructed through methods developed by Krinsky & Robb (1986). Population sizes were compared between scenarios to assess for significant changes in size for individual species and between species groups (defined in Bladwell et al., 2018).

6.2.1 Peer review: BTO model in the academic literature

Species-specific, spatial models of bird abundance using BBS data have been produced previously within a wide range of studies. Relevant examples include an evaluation of the relative importance of landscape, cropping and field boundary factors as influences on bird abundance (Siriwardena et al., 2011), measurement of range expansions in response to climate warming (Massimino et al., 2015), modelling of bird responses to urban form to inform planning (Plummer et al., 2020) and land use based prediction modelling that was included in the UK National Ecosystem Assessment (Bateman et al., 2013ab). Specific predictive modelling for Wales was integrated within the quick start modelling programme under ERAMMP (Thomas et al., 2021). These and other studies demonstrate the utility of the data source and the reliability of inference from established modelling techniques. Furthermore, the latter four studies incorporate the extension of inferential models to deliver predictions.

6.2.2 Expert Assessment: Sense-checking the inputs and outputs of the BTO model

To ensure that the BTO model within the IMP was as robust and repeatable as possible, the following steps were taken to automate the process and minimise common sources of error:

- Each scenario was run within an R project, with data saved such that it could be reproduced at any time.
- In order to avoid situations where there was a strong linear relationship between any two model covariates, all were assessed for collinearity via the cor() function in R (Table 6.1). No pair of values scored more than 0.9, a threshold previously utilised in BBS-habitat modelling (Plummer et al., 2020). The six covariate pairs which scored between 0.7-0.9 were all products of the SFARMOD-LAM model (crop types) and were retained, as a substantial proportion of the differences between each scenario run related to crop types.
- We set a value of 20 BBS squares in which each covariate required a non-zero count. This was designated to retain ecologically distinctive habitat types which feature rarely in Wales, e.g. saltmarshes. Covariates which appeared in less than 20 squares were amalgamated with the most biophysically similar covariate (e.g. lowland acid grassland, lowland calcareous grassland & lowland neutral grassland were combined into semi-natural grassland). The amalgamations are detailed in the assumptions document.
- Coastal squares (defined as those of land area < 900 m2) were removed from modelling as these regions are poorly covered by BBS surveys.
- For each species, model covariates were only assessed for significant correlation with abundance, if they were non-zero in > 20 BBS squares with a non-zero count. This was performed to decrease the likelihood of false associations between species and rarer or ecologically unrealistic habitats.
- All model covariates had a significant impact on at least one species present in the modelling.

Table 6.1 In order to avoid situations where there was a strong linear relationship between any two model covariates, all were assessed for collinearity via the cor() function in R

	hab_diversity	mean_altitude	cover_maize	cover_wheat	cover_barley	cover_rape	cover_other_farmed	cover_broadleaf_wood	cover_conifer_wood	cover_shrub	cover_rotational_wood	area_hedgerow	area_other_scrub	cover_rotational_gr	cover_perm_gr	cover_rough_gr	cover_improved_gr	cover_semi_nat_gr	cover_acid_gr	cover_heather_gr	cover_heather	cover_saltmarsh	cover_littoral_zone	cover_suburban	cover_urban	cover_freshwater	cover_inland_rock	cover_boggy	no_cows	no_sheep	cover_winter_bare	cover_winter_crop	yield_grass	yield_crop
hab_diversity		-0.42	0.31	0.38	0.38	0.37	0.08	0.18	0.01	-0.10	0.03	0.20	0.17	0.32	0.43	-0.20	0.13	0.10	-0.19	-0.16	-0.16	0.09	0.06	-0.17	-0.17	0.08	-0.03	0.01	0.39	0.03	0.45	0.32	0.41	0.42
mean_altitude	-0.42		-0.27	-0.23	-0.21	-0.22	0.08	-0.12	0.12	0.18	0.02	-0.22	-0.02	-0.26	-0.26	0.57	-0.27	-0.12	0.48	0.20	0.20	-0.19	-0.23	-0.32	-0.20	-0.05	0.01	-0.10	-0.30	0.16	-0.33	0.13	-0.25	-0.21
cover_maize	0.31	-0.27		0.10	0.16	0.22	-0.04	-0.04	-0.08	-0.10	-0.03	0.07	-0.01	0.60	-0.02	-0.15	-0.02	0.04	-0.13	-0.04	-0.05	0.01	-0.02	-0.04	-0.04	-0.01	0.24	-0.01	0.53	-0.10	0.56	0.11	0.36	0.05
cover_wheat	0.38	-0.23	0.10		0.83	0.77	0.04	-0.06	-0.07	-0.09	0.05	-0.07	-0.05	0.06	0.08	-0.11	-0.09	-0.03	-0.12	-0.04	-0.04	-0.03	-0.01	-0.07	-0.03	-0.05	-0.04	-0.04	0.08	-0.09	0.47	0.19	0.07	0.63
cover_barley	0.38	-0.21	0.16	0.83		0.84	0.03	-0.06	-0.08	-0.09	0.03	-0.05	-0.03	0.07	0.09	-0.11	-0.08	-0.04	-0.13	-0.04	-0.04	-0.05	-0.02	-0.08	-0.04	-0.03	-0.05	-0.04	0.08	-0.06	0.52	0.19	0.05	0.68
cover_rape	0.37	-0.22	0.22	0.77	0.84		0.03	-0.06	-0.07	-0.08	0.02	-0.04	-0.04	0.08	0.06	-0.09	-0.09	0.05	-0.12	-0.04	-0.04	-0.04	-0.01	-0.07	-0.04	-0.01	-0.04	-0.03	0.09	-0.08	0.53	0.18	0.11	0.69
cover_other_farmed	0.08	0.08	-0.04	0.04	0.03	0.03		-0.04	0.39	0.38	0.08	-0.12	-0.11	0.13	-0.13	-0.08	-0.11	-0.05	-0.05	-0.05	-0.05	-0.04	-0.06	-0.09	-0.06	-0.06	0.00	-0.03	0.06	-0.10	0.13	-0.07	-0.03	0.03
cover_broadleaf_wood	0.18	-0.12	-0.04	-0.06	-0.06	-0.06	-0.04		0.09	0.04	0.10	-0.13	0.19	-0.06	0.06	-0.16	0.12	-0.03	-0.08	-0.06	-0.07	0.03	-0.11	-0.05	-0.06	0.13	-0.06	-0.03	-0.07	-0.01	-0.07	-0.10	0.04	-0.02
cover_conifer_wood	0.01	0.12	-0.08	-0.07	-0.08	-0.07	0.39	0.09		0.32	0.04	-0.11	-0.11	-0.07	-0.16	0.01	-0.08	0.00	0.00	-0.02	-0.02	-0.03	0.07	-0.07	-0.04	0.13	-0.03	0.00	-0.12	-0.05	-0.09	-0.14	-0.11	-0.05
cover_shrub	-0.10	0.18	-0.10	-0.09	-0.09	-0.08	0.38	0.04	0.32		0.06	-0.19	-0.18	-0.11	-0.20	-0.02	-0.15	-0.05	0.01	-0.04	-0.03	-0.03	-0.07	-0.11	-0.06	0.01	-0.04	-0.04	-0.14	-0.08	-0.13	-0.19	-0.16	-0.09
cover_rotational_wood	0.03	0.02	-0.03	0.05	0.03	0.02	0.08	0.10	0.04	0.06		-0.13	-0.02	-0.05	-0.07	-0.09	-0.05	0.01	-0.03	-0.03	-0.02	-0.02	-0.06	-0.05	-0.03	-0.02	0.04	-0.03	-0.08	-0.08	-0.03	-0.12	-0.03	0.07
area_hedgerow	0.20	-0.22	0.07	-0.07	-0.05	-0.04	-0.12	-0.13	-0.11	-0.19	-0.13		-0.18	0.15	0.27	-0.07	0.37	0.07	-0.12	-0.02	-0.06	0.01	0.09	-0.09	-0.13	-0.07	0.04	0.02	0.25	-0.02	0.10	0.16	0.23	0.03
area_other_scrub	0.17	-0.02	-0.01	-0.05	-0.03	-0.04	-0.11	0.19	-0.11	-0.18	-0.02	-0.18		0.00	0.28	-0.06	0.18	0.04	0.05	-0.06	-0.03	0.02	-0.14	0.05	-0.10	0.09	-0.09	-0.06	0.03	0.15	-0.03	0.09	0.15	-0.02
cover_rotational_gr	0.32	-0.26	0.60	0.06	0.07	0.08	0.13	-0.06	-0.07	-0.11	-0.05	0.15	0.00		-0.02	-0.20	0.02	-0.04	-0.19	-0.06	-0.06	-0.02	-0.04	-0.07	-0.06	-0.06	-0.04	0.00	0.87	-0.14	0.86	0.23	0.52	0.03
cover_perm_gr	0.43	-0.26	-0.02	0.08	0.09	0.06	-0.13	0.06	-0.16	-0.20	-0.07	0.27	0.28	-0.02		-0.27	0.15	-0.02	-0.25	-0.13	-0.12	-0.06	-0.09	-0.23	-0.18	-0.01	-0.11	-0.04	0.26	0.41	0.00	0.50	0.38	0.16
cover_rough_gr	-0.20	0.57	-0.15	-0.11	-0.11	-0.09	-0.08	-0.16	0.01	-0.02	-0.09	-0.07	-0.06	-0.20	-0.27		-0.28	-0.06	0.17	-0.05	-0.02	-0.07	-0.09	-0.17	-0.08	0.00	0.03	-0.03	-0.23	0.24	-0.21	0.49	-0.15	-0.09
cover_improved_gr	0.13	-0.27	-0.02	-0.09	-0.08	-0.09	-0.11	0.12	-0.08	-0.15	-0.05	0.37	0.18	0.02	0.15	-0.28		0.07	-0.18	-0.01	-0.08	-0.08	-0.02	0.01	-0.09	-0.06	0.03	-0.01	0.02	-0.15	-0.03	-0.22	0.08	-0.05
cover_semi_nat_gr	0.10	-0.12	0.04	-0.03	-0.04	0.05	-0.05	-0.03	0.00	-0.05	0.01	0.07	0.04	-0.04	-0.02	-0.06	0.07		0.01	-0.01	0.01	0.00	0.03	0.00	0.13	0.03	-0.02	0.36	-0.05	-0.06	-0.05	-0.10	-0.03	-0.04
cover_acid_gr	-0.19	0.48	-0.13	-0.12	-0.13	-0.12	-0.05	-0.08	0.00	0.01	-0.03	-0.12	0.05	-0.19	-0.25	0.17	-0.18	0.01		0.01	0.00	-0.06	-0.09	-0.15	-0.08	-0.01	-0.04	-0.04	-0.24	-0.04	-0.21	-0.14	-0.20	-0.14
cover_heather_gr	-0.16	0.20	-0.04	-0.04	-0.04	-0.04	-0.05	-0.06	-0.02	-0.04	-0.03	-0.02	-0.06	-0.06	-0.13	-0.05	-0.01	-0.01	0.01		0.02	-0.02	0.02	-0.02	0.01	0.01	0.05	-0.02	-0.08	-0.07	-0.07	-0.15	-0.09	-0.04
cover_heather	-0.16	0.20	-0.05	-0.04	-0.04	-0.04	-0.05	-0.07	-0.02	-0.03	-0.02	-0.06	-0.03	-0.06	-0.12	-0.02	-0.08	0.01	0.00	0.02		-0.02	0.00	-0.06	-0.03	-0.04	0.03	-0.02	-0.09	-0.07	-0.07	-0.13	-0.06	-0.05
cover_saltmarsh	0.09	-0.19	0.01	-0.03	-0.05	-0.04	-0.04	0.03	-0.03	-0.03	-0.02	0.01	0.02	-0.02	-0.06	-0.07	-0.08	0.00	-0.06	-0.02	-0.02		0.27	0.00	0.04	0.06	-0.01	-0.02	-0.04	-0.06	-0.04	-0.12	-0.06	-0.04
cover_littoral_zone	0.06	-0.23	-0.02	-0.01	-0.02	-0.01	-0.06	-0.11	0.07	-0.07	-0.06	0.09	-0.14	-0.04	-0.09	-0.09	-0.02	0.03	-0.09	0.02	0.00	0.27		-0.03	0.01	-0.08	0.04	0.00	-0.04	-0.11	-0.04	-0.14	-0.11	-0.03
cover_suburban	-0.17	-0.32	-0.04	-0.07	-0.08	-0.07	-0.09	-0.05	-0.07	-0.11	-0.05	-0.09	0.05	-0.07	-0.23	-0.17	0.01	0.00	-0.15	-0.02	-0.06	0.00	-0.03		0.27	0.06	-0.02	-0.02	-0.13	-0.16	-0.10	-0.34	-0.17	-0.06
cover_urban	-0.17	-0.20	-0.04	-0.03	-0.04	-0.04	-0.06	-0.06	-0.04	-0.06	-0.03	-0.13	-0.10	-0.06	-0.18	-0.08	-0.09	0.13	-0.08	0.01	-0.03	0.04	0.01	0.27		0.02	-0.02	0.02	-0.10	-0.11	-0.07	-0.22	-0.12	-0.06
cover_freshwater	0.08	-0.05	-0.01	-0.05	-0.03	-0.01	-0.06	0.13	0.13	0.01	-0.02	-0.07	0.09	-0.06	-0.01	0.00	-0.06	0.03	-0.01	0.01	-0.04	0.06	-0.08	0.06	0.02		0.00	0.02	-0.08	-0.07	-0.05	-0.10	-0.02	0.00
cover_inland_rock	-0.03	0.01	0.24	-0.04	-0.05	-0.04	0.00	-0.06	-0.03	-0.04	0.04	0.04	-0.09	-0.04	-0.11	0.03	0.03	-0.02	-0.04	0.05	0.03	-0.01	0.04	-0.02	-0.02	0.00		-0.01	-0.06	-0.06	-0.05	-0.08	-0.03	-0.05
cover_boggy	0.01	-0.10	-0.01	-0.04	-0.04	-0.03	-0.03	-0.03	0.00	-0.04	-0.03	0.02	-0.06	0.00	-0.04	-0.03	-0.01	0.36	-0.04	-0.02	-0.02	-0.02	0.00	-0.02	0.02	0.02	-0.01		-0.02	-0.01	-0.02	-0.05	-0.02	-0.04
no_cows	0.39	-0.30	0.53	0.08	0.08	0.09	0.06	-0.07	-0.12	-0.14	-0.08	0.25	0.03	0.87	0.26	-0.23	0.02	-0.05	-0.24	-0.08	-0.09	-0.04	-0.04	-0.13	-0.10	-0.08	-0.06	-0.02		-0.16	0.76	0.42	0.62	0.11
no_sheep	0.03	0.16	-0.10	-0.09	-0.06	-0.08	-0.10	-0.01	-0.05	-0.08	-0.08	-0.02	0.15	-0.14	0.41	0.24	-0.15	-0.06	-0.04	-0.07	-0.07	-0.06	-0.11	-0.16	-0.11	-0.07	-0.06	-0.01	-0.16		-0.14	0.51	0.10	-0.06
cover_winter_bare	0.45	-0.33	0.56	0.47	0.52	0.53	0.13	-0.07	-0.09	-0.13	-0.03	0.10	-0.03	0.86	0.00	-0.21	-0.03	-0.05	-0.21	-0.07	-0.07	-0.04	-0.04	-0.10	-0.07	-0.05	-0.05	-0.02	0.76	-0.14		0.29	0.48	0.40
cover_winter_crop	0.32	0.13	0.11	0.19	0.19	0.18	-0.07	-0.10	-0.14	-0.19	-0.12	0.16	0.09	0.23	0.50	0.49	-0.22	-0.10	-0.14	-0.15	-0.13	-0.12	-0.14	-0.34	-0.22	-0.10	-0.08	-0.05	0.42	0.51	0.29		0.41	0.23
yield_grass	0.41	-0.25	0.36	0.07	0.05	0.11	-0.03	0.04	-0.11	-0.16	-0.03	0.23	0.15	0.52	0.38	-0.15	0.08	-0.03	-0.20	-0.09	-0.06	-0.06	-0.11	-0.17	-0.12	-0.02	-0.03	-0.02	0.62	0.10	0.48	0.41		0.13
yield_crop	0.42	-0.21	0.05	0.63	0.68	0.69	0.03	-0.02	-0.05	-0.09	0.07	0.03	-0.02	0.03	0.16	-0.09	-0.05	-0.04	-0.14	-0.04	-0.05	-0.04	-0.03	-0.06	-0.06	0.00	-0.05	-0.04	0.11	-0.06	0.40	0.23	0.13	

6.2.3 Error checking: Checks applied to each predictor dataset (baseline and scenarios) to make sure these were valid

The following steps were taken to ensure that each predictor dataset for every run (baseline and land use scenario) to trap any common errors that might occur:

- All coverage estimates in each square were positive.
- All coverage estimates in each square totalled to within 50 m2 of the overall land cover.
- Each scenario was formed of the same 1 km grid squares as the baseline.
- In each scenario, all coverage estimates which were not involved in the LAM remained identical in cover.
- The total cover area of a scenario was within 25,000 m2 of the baseline (less than 1.5 m2 difference in area per 1 km grid square).

6.2.4 Validation: Assessment of BTO model predictive ability using cross-validation

Cross-validation is a statistical approach designed to evaluate the performance of predictive models, through scoring the accuracy of model relationships on test datasets. The procedure involves randomly allocating each species-habitat model input dataset into 10 subsets (folds). A model is trained using 9 folds (training data) and used to predict the output in the remaining fold (test data). This process is iterated with each fold utilised as test data once (Roberts et al., 2017). The deviation between the training data predictions and the real data are evaluated through two measures. Spearman's rank correlation coefficient (rs) is used to assess the degree of agreement between predicted and true values, whilst mean absolute error (MAE), is used to assess the mean difference in counts between true and predicted values (Willmott & Matsuura, 2005).

Spearman's rank correlation coefficient rs values for the bird-habitat models averaged at 0.41 (range 0.66 – 0.12) (Figure 6.1), indicating on average each species model accounted for at least a moderate level of the variation in counts. Such values are comparable with previously conducted species-habitat models in the wider literature (e.g. Evans et al., 2009; Plummer et al., 2020). Low values (<0.3) tended to be a consequence of low sample size in the original BBS surveys, with seven species excluded from the modelling through having a rs value of less than 0.1. Further evidence for model efficacy was provided by the MAE results. The mean predicted count differed to the observed values by an average of 4.1 (1.02 SE), and less than 3 in 44 of the 68 species. Selection of the overdispersion threshold (6), was also a product of sensitivity testing, based upon minimising MAE and increasing rs.

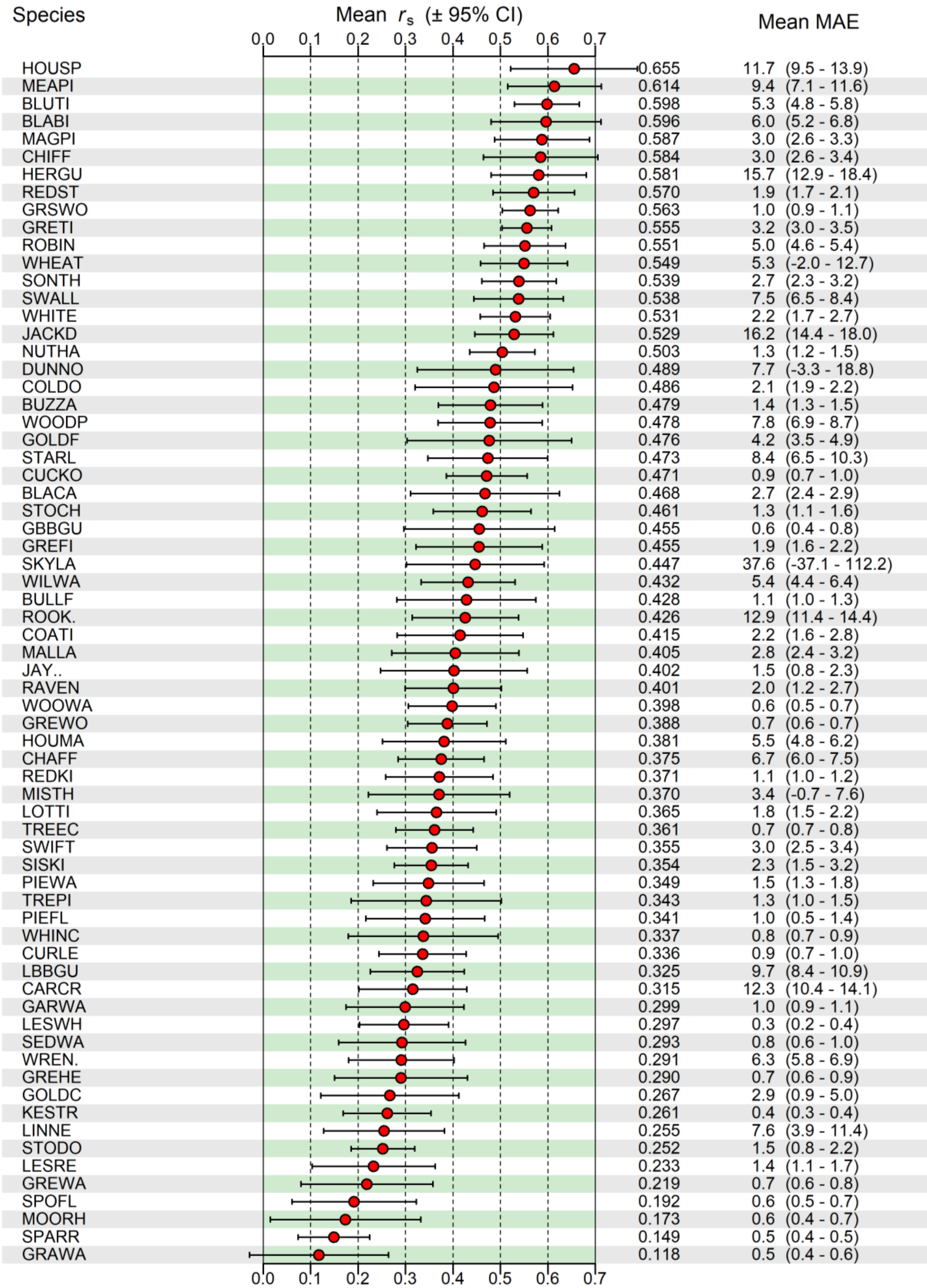


Figure 6.1 The predictive ability of bird-habitat models for 68 bird species, as assessed through the mean spearman's rank correlation coefficient and mean absolute error (both with 95% confidence intervals), observed from block 10-fold cross validation. Species are denoted by their BBS 5-letter code and are ordered from highest spearman's rank correlation coefficient to least.

6.2.5 Expert assessment: Comparison of BTO model predictions with other data sources.

Baseline predictions were assessed for suitability against data from wider sources, including wider literature, expert assessment, and other models. This was primarily designed to make sure predictions were realistic and the relationship between individuals and model metrics behaved as expected.

- Scenario changes and predictions were discussed with the IMP plant modelling team, in particular focussing on whether habitat indicators behaved similarly in both models.
- Indicator species were selected for different model metrics and assessed for whether their relationship between these metrics was positive. In all specialist species checked (n=12), relationships between species counts and covariates were positive. As examples, pied flycatcher had a strong association with broadleaf woodland, whilst collared dove was associated with suburban cover.
- Population predictions were compared against percentage occurrence from the BBS surveys, with predicted population sizes matching expectations based upon how common a species was in surveys.
- For land use scenarios T1 and T6, the substantial increases in coniferous woodland meant our model only had a limited number of training data to inform predictions (as coniferous-dominated squares were rare in BBS surveys). In test runs, this resulted in unrealistically high predicted counts for species associated with coniferous woodland. As a result, we devised a method to reduce impact of rare habitats undergoing substantial increases in cover within a particular scenario and retain predictions which were realistic whilst also reflecting the substantial nature of scenario changes.
- A threshold value of 1000 individuals for a 1km square was set, based upon a high-bound for an observable maximum from BBS survey data. Any model which resulted in predictions which exceeded this value were flagged. An iterative method for outlier removal was devised based upon Cook's distances (Cook, 1977), a measure of the influence on the observed data when performing a regression-based analysis. We applied this to the model input dataset when predictions were flagged. This resulted in a lowering of the covariate-count estimate, applying a brake to keep predictions realistic within a particular scenario, whilst retaining the baseline model when covariate changes are less substantial. For more information, please see the model technical guide.

7 Biodiversity: Plants - MultiMOVE

Authors: Simon Smart and Bede West

7.1 Introduction to the model QA

This section presents the MultiMOVE model ensemble and the relevant quality assurance steps that have been undertaken. This set of models is used to predict plant species composition at baseline and for the same locations in scenarios given changes in soil conditions and vegetation height expected to arise as a consequence of the change in land use predicted by the LAM. The quality assurance steps are:

- Peer review: MultiMOVE has been discussed in several peer reviewed papers (Sections 7.2 and 7.2.1).
- Error Checking: Transfer of SFARMOD/LAM outputs into MultiMOVE (Section 7.2.2).
- Validation and Expert Assessment: Soil trajectories (Section 7.2.3).
- Validation: Reproduction of plant species composition of baseline quadrats (Section 7.2.4).
- Error Checking: Verification of the MultiMOVE workflow (Section 7.2.5).

7.2 Introduction to the modelling

MultiMOVE comprises a small ensemble of Species Niche Models for British plants (see Smart et al., 2010, DeVries et al., 2010, Rowe et al., 2013, Henrys et al., 2015 and Smart et al., 2019 for full details of model building, testing and application). Five statistical modelling techniques are used to model the probability of occurrence of 1188 higher and lower plants in terms of seven environmental variables: i) substrate pH, ii) fertility, iii) soil moisture, iv) canopy height, v) annual rainfall, vi) max July and vii) min January temperature. These variables were measured or estimated at fine resolution in quadrat samples ranging from 14.14 x 14.14m (200m2) to 2x2m (4m2) in which full plant species lists were recorded. The species' input data for model building were therefore presence/absence records. A large GB-wide database of 32,727 quadrats was used to build the models resulting in coverage of all habitat dominants and numerous rare and subordinate species. Importantly this means that MultiMOVE includes those species that deliver the most ecosystem function and service by virtue of their frequency and abundance across ecosystems. Some species groups are omitted: alien casuals (alien species that flourish occasionally, but which do not form self-replacing populations (Richardson et al., 2000)) are excluded as are coastal halophytes and several floating and submerged aquatics because of lack of data.

Because species dispersal is not modelled, the probability values that are output from MultiMOVE should be interpreted as habitat suitability indices (Smart et al., 2019). Therefore, an output value close to the maximum possible for the species (pmax) suggests that abiotic conditions are estimated to be appropriate for establishment and persistence should the species be able to reach the location. In order to increase realism and constrain the species pool modelled in any one location, we restrict this pool to the number of species observed in each sample plot at baseline plus the extra species recorded in the wider 10x10km square where these can be accessed from the online database maintained by the Botanical Society of the British Isles (https://database.bsbi.org/). Lastly, the output values from MultiMOVE for each of the five techniques is transformed into a single weighted model average. This is achieved by applying weights to each output based on a prior cross-validation test of the ability of each technique for each species to predict hold-out samples of the training data (see Smart et al., 2019 for details). Because MultiMOVE models each species separately, we aggregate results afterwards (a predict first, assemble later approach) to generate measures of functional group richness as necessary or present results at the individual species level (Ferrier & Guisan, 2006).

When used in predictive mode, new values for each of the seven input variables are supplied to MultiMOVE and each of the five statistical modelling techniques are used to solve each of the species' models. The input values can be adjusted using any number of flexible approaches to yield a new configuration of input variables representing a scenario of environmental change. As the MultiMOVE model is trained on high quality, fine spatial resolution data for known sites and a large number of species it provides very high local realism and very flexible model application based on minimal data demands. Unlike many other models within the IMP, MultiMOVE models are applied at the small vegetation patch scale for specific sites rather than across a consistent raster grid. This means that predictions are kept at the patch and field scale at which land use decisions need to be made without any averaging that uncouples predictions from the field data. It should be noted that the cost of this field-based realism is that the models contain no inherent dynamism and are therefore constrained in their ability to produce novel outcomes beyond the range of the conditions used to train the model. Any model projecting outside of the domain on which it was trained must be viewed as highly uncertain. MultiMOVE issues a warning that input data represent novel ecological space, hence, the resulting predictions can be isolated and carefully inspected, while the warnings themselves act as a useful check on the plausibility of the input data and the potential for increasingly unoccupied niche space into which species must either adapt or disperse.

7.2.1 Peer Review: MultiMOVE papers and niche model validation

MultiMOVE has been tested on several occasions in other settings (Rowe et al., 2015; DeVries et al., 2010). Most recently, an independent assessment of each modelled niche axis for each species was carried out by eliciting judgements from two expert botanists with experience of the British flora (Smart et al., 2019). This was done in parallel with a quantitative assessment of model fit whereby models were retrained and used to predict repeated random hold-out samples of the training set (Smart et al., 2019). The results showed excellent ability of the models to predict presence/absence patterns in cross-validation samples (Figure 7.1).

Given the value of independent assessment but the large number of species in the models, a website now exists to crowd-source expert opinion (https://shiny-apps.ceh.ac.uk/find_your_niche/).

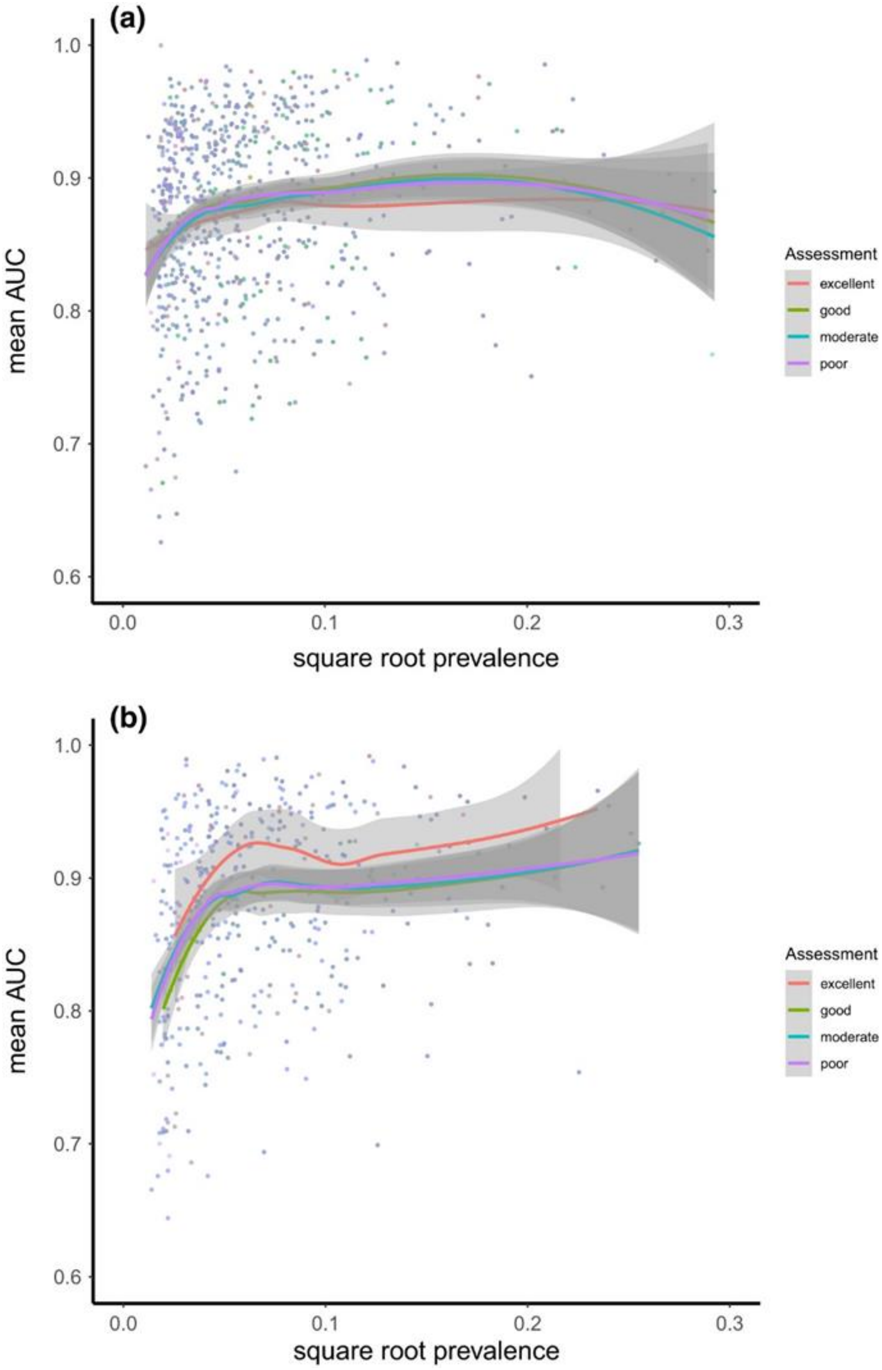


*Figure 7.1 Comparison of expert assessments - (a) Expert 1; (b) Expert 2 - for each species niche axis combination versus AUC statistics for the associated model and the prevalence of each species in the training data used to build each model. Loess smoothers are fitted to each species * niche axis combination grouped by the assessment category awarded by the expert. Thus, each point is a species * niche axis combination whose position is defined by its prevalence on the x-axis and the mean AUC for the species model on the y-axis. Note that prevalence (the proportion of presences/ total number of quadrats) was square-root-transformed to spread the data more evenly across the x-axis (from Smart et al. 2019).*

7.2.2 Error Checking: Deciding whether a GMEP plot location can change land use

To model the impacts of predicted land use change on plant habitat suitability, the land use transitions output from SFARMOD/LAM needed to be processed through MultiMOVE. This requires additional steps that translate the land use transitions predicted at baseline locations into changes in the soil and vegetation height inputs that MultiMOVE uses to filter the species composition of the assemblage observed at baseline. Two key rule-checking steps are required to manage the transfer of SFARMOD/LAM outputs into MultiMOVE. The validation and management of these transfers is described below.

Two major rules govern whether the vegetation and soil in each baseline GMEP plot location can be allowed to change (Figure 7.2).

Firstly, we determine the match between the predicted land use/land cover type from SFARMOD/LAM and the habitat observed by the field surveyor when they recorded each plot. Where the two agree then change is allowable; these are termed 'compliant' plot locations. Where they disagree, the predicted land use transition is not processed through MultiMOVE and the location remains stable; these are termed 'deviant' plot locations. This is a conservative approach that means modelled change only occurs in locations where the upstream land use models and field observations at baseline agree. This leads to another test comparing the two datasets at baseline (see LAM section: Figures 5.5 and 5.6). For projected land use change to fully impact plant species via MultiMOVE, then all plot locations should be compliant indicating that SFARMOD/LAM and the field classification agree. In reality, because SFARMOD/LAM use LCM2015 as a source dataset for land use, the matching described above amounts to a test of whether the field survey agrees with the satellite classification. In most cases, it does (Figures 5.5 and 5.6). Hence, Figure 5.5 indicates that for the majority of the quadrat locations modelled by MultiMOVE, the LAM land use type agreed with the observed habitat present at baseline. Greatest disagreement was where the LAM expected permanent grassland, but the observed habitat was different.

Secondly, having established a set of matching locations, the next decision criteria focus on allowing or disallowing land use change. In broad terms this means that semi-natural habitat types such as Fen, Marsh & Swamp, Bog, Bracken, Dwarf Shrub Heath and Broadleaved woodland are not permitted to transition to more intensive land use types within MultiMOVE. This is because it would mean contravening cross-compliance regulations under the Whole Farm Code and the Environmental Impact Assessment Regulations (https://gov.wales/sites/default/files/publications/2021-02/environmental-impact-assessment-guidance.pdf). In these cases, the QA step prevents transitions, which is a conservative approach that minimises incorrect prediction because reality on the ground trumps the habitat type estimated from the satellite-derived LCM2015.

Each scenario involved varying patterns of predicted land use change at each plot location. The same set of rules was applied on a habitat-specific basis and the resulting decision matrices saved for each scenario to aid auditability and repeatability.

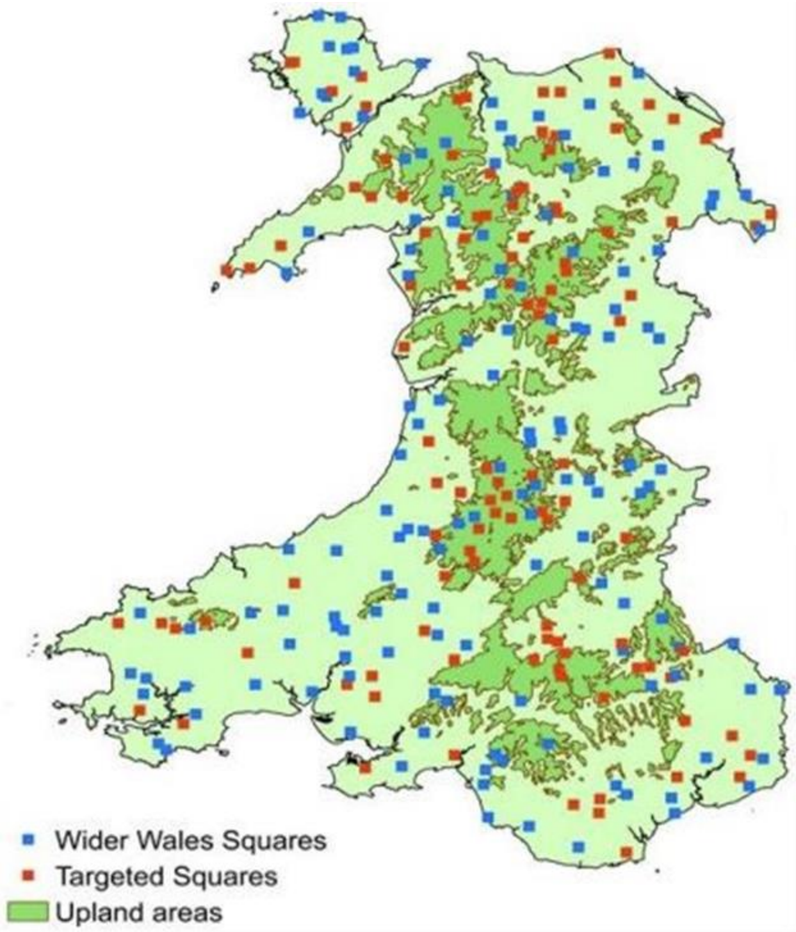


Figure 7.2 Locations of the GMEP 1 km squares across Wales. Up to 5 2x2m plots are located in each square. Soil, habitat type and plant species composition were recorded in each and used as the baseline dataset for MultiMOVE modelling in the IMP.

7.2.3 Validation & Expert Assessment: Soil trajectories

The soil trajectories within MultiMOVE were derived by assembling observed and experimentally induced changes in soil conditions from published sources. These changes were summarised as mean changes per year with uncertainty. The plausibility of the models was subject to expert validation as part of the PhD supervision process for Bede West. In summary the direction and magnitudes of change in soil variables given varying broad habitat starting points was checked against the combined 60 years' experience of land use change research by Prof Davey L. Jones and Dr Simon Smart. No outliers or implausible trajectories were observed in the literature-derived transitions.

Care was taken to closely match the methods of each study found during literature review to the GMEP methods or ensure data values could be converted to match. Non-UK studies were omitted. This search resulted in datasets of varying size as requests were made to study authors to provide full datasets but not always yielded, and also including relevant open access data; these are:

- Project supplementary material from grassland restoration from arable (Pywell et al., 2007).
- UK Department for Environment, Food and Rural Affairs (Defra) report (Wagner et al., 2014).

- Rothamsted Research, Park Grass long-term experiment data (Rothamsted Research, 2016).
- Project on restoration from farmland (Pywell et al., 1994).
- Full data set provided by author from Marrs et al. (2018).
- Summary data published in McGovern et al. (2014).
- Defra project on managing grassland diversity (Defra, 2015).
- Elan Valley grasslands report (Hayes & Lowther, 2014).
- A 12-year fertilizer and lime experiment, supplementary material from (Kirkham et al., 2011).

The benefit of modelling at each of the GMEP plot locations is high realism reflecting the availability of fine resolution soil input data and observed species composition. The disadvantage is that these locations cannot be dynamically modelled using a soil biogeochemical model because prohibitive amounts of additional data would be required at each of the quadrat locations modelled. An empirical and data-driven approach was taken to assembling plausible trajectories of change in the soil variables used as input to MultiMOVE using a wide range of published papers. Our goal was to build replicated time series of change in pH, %C and %N from experiments, modelling and chronosequence studies where soil change was driven by a management intervention corresponding to the land use transition of interest. Average rates of change per year were then extracted and applied to each set of starting values in each plot predicted by SFARMOD/LAM to undergo the respective land use transition. The assembly of these data was carried out by Bede West as part of his ENVISION-funded PhD. Construction and plausibility of the soil models have been reviewed by Dr S. Smart and Prof D.L. Jones as part of the PhD supervision process and are due for publication soon.

7.2.4 Validation: Comparison of MultiMOVE predictions with the plant species composition of baseline quadrats

Further testing was carried out by investigating whether, given soil, climate and vegetation height at each plot location, MultiMOVE was able to successfully predict the observed species composition in the baseline quadrats (Figure 7.3). This was achieved by comparing the model's performance (predicted habitat suitability) for all observed species in plots with a random draw of the predicted suitabilities from all the modelled species. Performance was good with most observed species (statistic) having higher probabilities than a random draw. Where models were less reliable and results were closer to random (left of the line in Figure 7.4) this could largely be explained by low numbers of observations in the dataset (Figure 7.4). Hence the vast majority of species were more likely to be correctly predicted as being present when they were in fact present in baseline quadrats.

Further support for good predictive ability across the models comes from a cross validation exercise (Figure 7.5). This involved building all models repeatedly, but leaving out a single random quadrat sample from the training data. The predictions from the models are compared with observed species in the hold-out sample and the number of correctly and incorrectly predicted species are noted. Values were accumulated across 10,000 cross-validated model checks for each modelling technique and species. The statistic takes into account the need to correctly predict when a species is present without at the same time predicting presence when in fact the species is absent. So ideally the true positive rate should be high and the false positive rate low. This ratio increases as the AUC statistic increases (Figure 7.5).

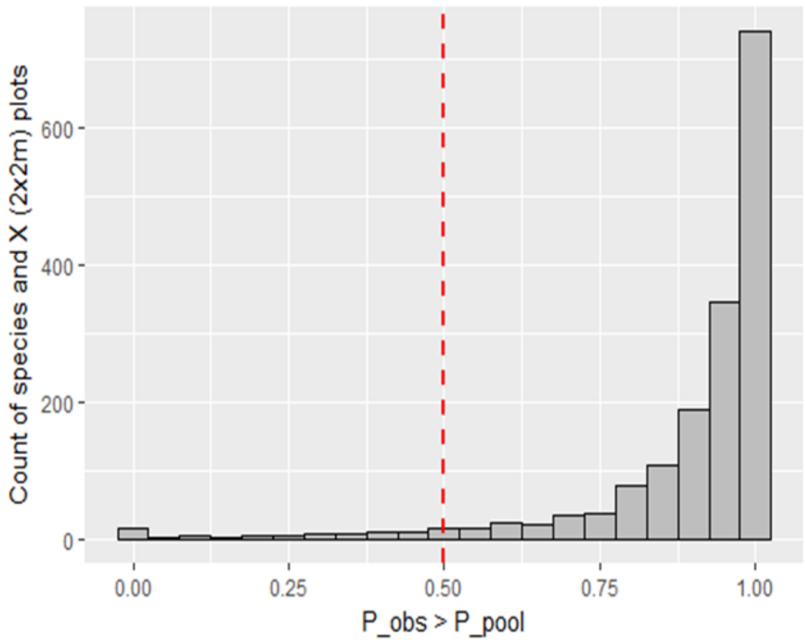


Figure 7.3 Results of randomisation testing of the ability of MultiMOVE to reproduce observed species composition in baseline quadrats. The bar on the extreme right shows that in c.750 observed occurrences of species in different baseline quadrats the predicted habitat suitability (P_obs) was always greater than a random draw (P_pool) of predicted habitat suitabilities for the other plant species in the local pool (present in 2x2m quadrats + present in wider 10km square).

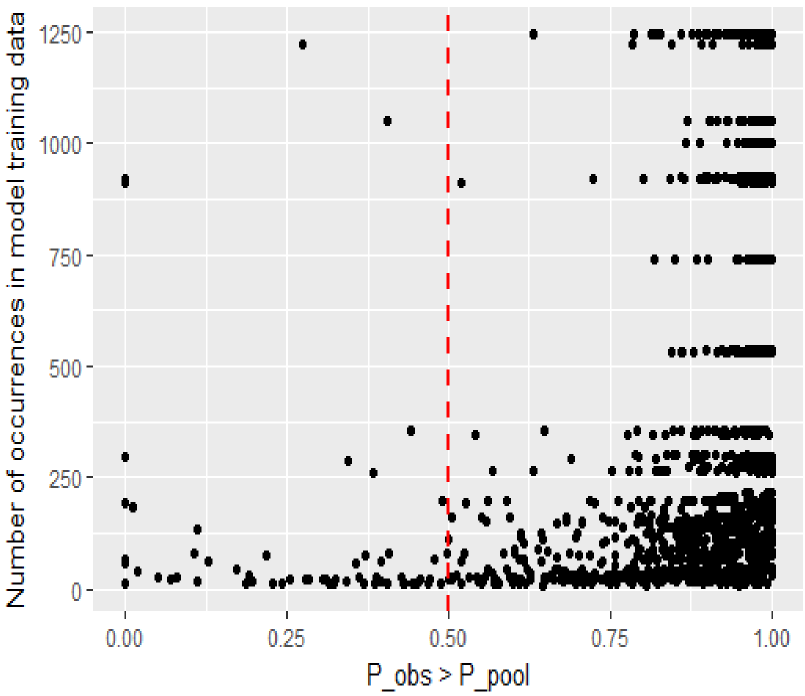


Figure 7.4 Relationship between predicted suitability for observed versus randomly drawn suitability values (P_obs > P_pool) and number of presences in the training data used to build MultiMOVE. As expected rare species end up with less reliable niche models in MultiMOVE.

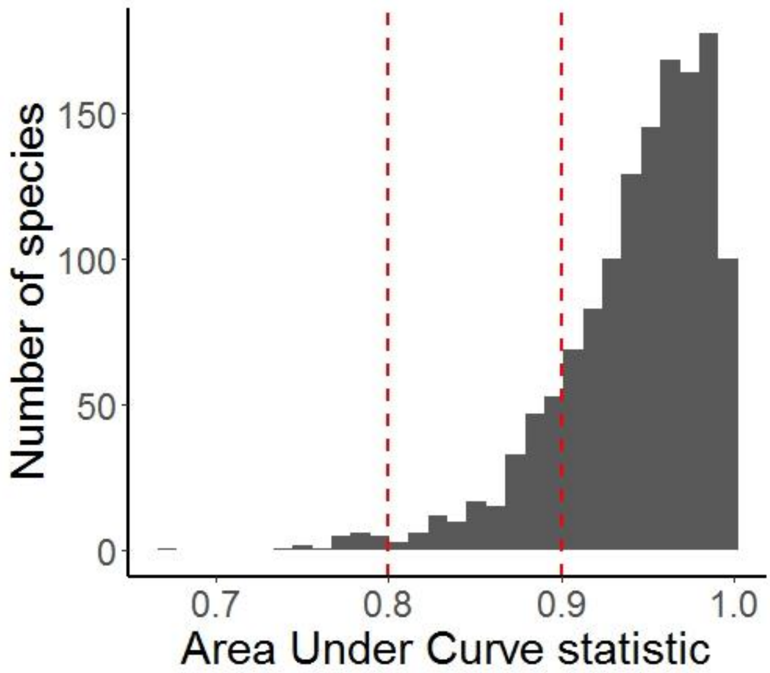


Figure 7.5 AUC statistics for MultiMOVE models. A value of one indicates prefect prediction with no false positives. Values above 0.8 are generally considered to indicate very good predictive ability.

7.2.5 Error Checking: Checks in the MultiMOVE workflow

Each modelled scenario run is produced using an R project dedicated to that scenario and that can be rerun in its entirety. Each project consists of three R code files comprising code-chunks that generate saved output that can also be reproduced at any time. Several automated checks are built into the code within the MultiMOVE R package to flag any common errors.

Firstly, the data are checked to ensure input variables are within the range of the training data with warnings issued when this is not the case. All data were within the range of the training data within the Land Use Scenario modelling.

Secondly, the LAM-MultiMOVE linking tables are checked and compared across scenarios to ensure consistency and trap errors. These critical tables (one per scenario) join the land use transitions predicted in SFARMOD/LAM with the associated shift in vegetation height and soil conditions needed by MultiMOVE. Each table is structured such that every unique land use transition from SFARMOD/LAM is assigned a binary indicator that allows the transition to be modelled by MultiMOVE or not (due to cross-compliance preventing the change in land use: see Figure 5.5 and Figure 5.6 in Section 5.2.6). Those allowable transitions are then classified by the type of transition involved. Then each location undergoing a particular transition type is attached to a sequence of numbers representing the amount of change in vegetation height and soil input variables to be modelled over time in response to the particular land use transition and given the particular vegetation at the start.

The key QA checks undertaken here aim to ensure that the linking has worked correctly. The checks include assuring that i) all LAM DMU have a spatial link to a GMEP baseline quadrat location, and ii) that all LAM-modelled transitions are represented by a profile of expected changes in soil conditions and canopy height. Both these checks were passed in the Land Use Scenario modelling.

8 Habitat connectivity

Authors: Amy Thomas and Eleanor Warren Thomas

8.1 Introduction to the model QA

The habitat connectivity models identify areas where new habitat, as generated by the LAM, would connect two patches of unconnected habitat types. The quality assurance steps are:

- Expert Assessment: The model code has been checked and the outputs have been visually checked to ensure they are identifying land that would create connectivity given the model parameters.

8.2 Introducing the modelling

The model identifies existing connectivity and opportunity for new connectivity between patches of habitat. It does this by identifying the area or "zone" around a habitat patch that would be accessible to biodiversity within the patch, based on a user input dispersal distance and patch size. Where the "accessible zone" contains another patch of the same habitat type, the habitats may be considered connected. If "accessible zones" overlap, new woodland in the "overlapping zone" would create a connection between the existing habitat patches. Different species types will have differing habitat size requirements and dispersal capabilities, so model parameterisation, in terms of patch size and dispersal distance, is crucial to the interpretation of results (Table 8.1).

The method was developed, and an ArcGIS toolbox created by A. Thomas. The code was then reviewed and edited by E. Warren and A. Thomas, and applied across Wales using a range of parameters derived from the literature, as shown in Table 8.1. A. Thomas then used the outputs from these model runs to identify which DMUs could create connectivity for existing broadleaf woodland (under any of these parameterisations) if they were planted as woodland or regenerated natural woodland.

Table.8 1 Parameterisation applied for the habitat connectivity model implementation. Dispersal distances and minimum patch sizes from: Watts et al., 2010; Natural England Nature Recovery Networks Evidence Review. Note: 10 ha for woodland specialist birds (Dolman et al., 2007).

Dispersal distance/ patch size	**100m: snails**	**200m: woodland specialist plants**	**500m: invertebrates**	**1km: max. for snakes; amphibians; moths**	**2km: max. for woodland flora/fauna**
1 ha: low area requirements	not modelled	modelled	modelled	not modelled	not modelled
10 ha: high area requirements	not modelled	modelled	modelled	not modelled	not modelled
40 ha: NE recommended minimum size for wildlife site	not modelled	modelled	modelled	not modelled	modelled

8.2.1 Expert Assessment (Consortium): Interpreting model outputs in the baseline scenario

There are no appropriate datasets for comparison. Whilst the model outputs could be compared to the outputs of alternative connectivity modelling tools, this would be a comparison of approaches with differing assumptions, modelling different aspects of connectivity. Hence, such a comparison would not explore uncertainty in the model output as it should be interpreted, rather it would explore the variation between different types of connectivity (e.g. areas which are already linked to one patch of woodland vs areas which have potential to join two patches of woodland).

Based on review of the code, and visual assessment of outputs, we are confident that the code does what it should and is identifying land which would create connectivity as per the parameters. However, these output data should be used appropriately, and may be unsuitable for some uses. For example, data should not be used for targeting of new woodland planting because whilst theoretically planting trees in the identified locations would create connectivity for the relevant types of biodiversity, it cannot be said that new woodland in those locations would create a discernible benefit for the relevant types of biodiversity. Actual benefits in practice would be strongly dependant on other factors such as the existing levels of the relevant types of biodiversity in the habitats being connected, as well as other factors such as barriers in the landscape, and potential negative impacts of connectivity, e.g. for disease transmission, movement of invasive species or impacts on predator prey relationships.

9 Agricultural Pollutant Coefficients - Farmscoper

Authors: Richard Gooday, Amy Thomas, Daniel Sanders and Christopher Feeney

9.1 Introduction to the model QA

Farmscoper provides farm pollutant emission coefficients for nitrate, phosphorus, sediment, methane, nitrous oxide and ammonia. These coefficients are combined with agricultural scalar data (e.g. area of improved grassland or quantity of beef slurry applied to arable land) from SFARMOD to predict total agricultural pollutant loads for the different scenarios.

This section describes validation of the Farmscoper coefficients independently from the SFARMOD data by using an alternative input agricultural scalar dataset (referred to as the "ADAS 1 km2" data). It also compares the ADAS 1 km2 input dataset with SFARMOD output so as to understand the influence of any differences. Note that actual QA of the combination of Farmscoper coefficients with SFARMOD input data is described separately in Section 10 (for GHGs) and Section 11 (for water pollutants).

The contents of Section 9 are thus:

- Validation: Comparison of modelled agricultural loads using the Farmscoper coefficients and the non-SFARMOD input agricultural scalar data ("ADAS 1 km2") with previously published results; including a comparison of the spatial distribution of predicted pollutant distributions (Section 9.2.1).
- Validation: Comparison of the agricultural scalar data from SFARMOD with that derived from the ADAS 1 km2 data (Section 9.2.2). This section is important to help explain any differences in Sections 10 and 11.

9.2 Introduction to the modelling

The Farmscoper model (Gooday et al., 2014) was used to calculate annual average losses of sediment, nitrate, phosphorus to water; and of ammonia, nitrous oxide and methane to air. Farmscoper was populated with cropping and livestock data for three representative farm systems (extensive grazed livestock; intensive grazed livestock and arable systems). This was done using data derived from the June Agricultural Survey for Wales, alongside other information of farm management taken from national surveys. Where possible data for Wales were used (e.g. the Welsh Farm Practice Survey (Anthony et al., 2012)), or for England and Wales (e.g. the Defra Farm Practice survey).

This farm management information included, for example, the proportion of manure managed as slurry and the uptake of various mitigation measures. The three farm systems were then modelled on each of the three soil types and six climate zones available within Farmscoper. Pollutant loss coefficients per unit input (e.g. pollutant loss attributable to dairy slurry per kg of N in dairy slurry) were calculated from the Farmscoper source-apportioned pollutant loss results.

In the IMP, the agricultural pollutant loads are then calculated by multiplying SFARMOD data (e.g. kg of dairy slurry N) at the DMU level by the appropriate coefficients for the Farmscoper soil type, climate zone and farm type of that DMU. Pollutant loads at other spatial scales are then determined by summing the results for the DMUs.

This coefficient approach allowed the losses predicted by Farmscoper to scale with the input data from SFARMOD, which would vary between DMUs and would change as a result of different scenarios. Coefficients are based on average climate data (1961-1990) (as that is the data within Farmscoper v4) and do not reflect variations in weather between years.

9.2.1 Validation: Assessing the IMP Farmscoper coefficients using ADAS 1 km2 June Census Data

To independently QA the Farmscoper coefficients developed for use in the IMP before they are integrated with the SFARMOD agricultural scalar data, the coefficients were combined with an alternative scalar dataset created using the ADAS 1 km2 June Agricultural Survey (JAS) dataset for 2014 (Lee et al., 2015). The total pollutant loads calculated using this approach have been compared against three alternative datasets:

1. Anthony et al. (2012), which used 2004 JAS data to calculate pollutant loads as part of an assessment of the impacts of Tir Cynnal and Tir Gofal;
2. The SEPARATE database (Zhang et al., 2014), which contains modelled pollutant loads for all WFD waterbodies in England, taken from the PSYCHC (Davison et al., 2008) and NEAP-N (Lord and Anthony, 2000) models using JAS data for 2010;
3. GHG and ammonia emissions from the National Atmospheric Emissions Inventory (data for 2018).

The ADAS 1 km2 agricultural census dataset contains land use and livestock data (hectares of cropping and count of stock per 1 km2) derived from holding level June Agricultural Survey data. In order to use this dataset with the Farmscoper coefficients created for use in the IMP, appropriate scalars were derived for each livestock and crop category (e.g. kg N in dairy excreta per dairy cattle, kg N in beef slurry applied to grass per beef cattle). These scalars could then be multiplied by the 1 km2 census data to provide the total scalar data to use with the Farmscoper coefficients in order to calculate pollutant losses (whereas in the IMP, the total scalar data for each DMU is determined by SFARMOD). This allowed for an independent assessment of the Farmscoper coefficients outside of the IMP, and an additional dataset against which to compare aspects of the IMP calculations. For each 1 km2, it was also necessary to determine the appropriate Farmscoper soil category, which is based upon HOST class, and Farmscoper climate zone, based upon annual average rainfall.

Table 9 1 shows that the predictions as part of the QA are close to those of Anthony et al. (2012) for nitrate and close to SEPARATE for phosphorus and are between the values for sediment. Nitrous oxide emissions are close to those of Anthony et al. (2012), but both are greater than the NAEI figures. Ammonia emissions in the QA are also higher than in the inventory (they were not calculated as part of Anthony et al. (2012)). Both Farmscoper and Anthony et al. (2012) used the IPCC (1996) methodology as a basis for their calculations, using default coefficients derived for Western Europe (see Baggott et al., 2006). The methodology used in the NAEI has evolved since, with changes to many coefficients and more sophisticated tier 3 approaches adopted for some calculations (see Brown et al., 2021), including Welsh-specific calculations of N excreta from livestock. These changes explain some of the differences between the emissions from Farmscoper and the NAEI. The calculated emissions can also be very sensitive to management assumptions (e.g. the ammonia loss from a kg of excreta at grazing is a fraction of that if the excreta are deposited in the yard). Small differences in where cattle are assumed to spend their time, or how manure is managed can thus have a big impact.

Table 9 1 National annual average agricultural pollutant loads (kt yr-1)

Area	Sediment	Nitrate	Phosphorus	Ammonia NH3-N	Nitrous Oxide	Methane
ADAS QA	393	37.6	1.04	25.4	8.0	150
Anthony et al. (2012)	220	36.5	0.65	N/A	8.5	135
SEPARATE	479	26.3	0.97	N/A		
NAEI	N/A			18	5.2	139

The data in Anthony et al. (2012) allow for a comparison of the area and source apportionment of the pollutant loads with those predicted as part of the QA process (Table 9.2 and Table 9.3 respectively). Grassland is unsurprisingly the dominant area for all pollutants, with the most notable difference being the greater contribution to the sediment load from rough grazing (19% in the QA, 29% in Anthony et al. 2011). This is related to the large difference in total sediment load predicted for the two approaches, as shown in Table 9.1 (i.e. the absolute load from rough grazing is approximately equal for the two approaches). There is slightly more pollution predicted from non-field areas (steadings, tracks etc.) in the QA approach, but there is arguably greater uncertainty around assumptions made for these losses (e.g. time spent in yards, cleaning efficacy, etc.) and thus there is no reason to say one of the modelled outputs is correct and the differences are small enough to have limited impact on any results or conclusions drawn from the modelling. The apportionment by source is also largely comparable for the two approaches, although the QA approach assumes a greater proportion is from dairy animals, and slightly less from beef. Contributions from pigs and poultry are small in both approaches, with a maximum contribution to the load of 3%.

Table 9 2 Comparison of area apportionment for national annual average agricultural pollutant loads a) derived using the Farmscoper export coefficients created for the IMP and scalar data derived from ADAS 1 km2 agricultural census data for 2014 and b) from Anthony et al. (2012) calculated using 2004 agricultural census data.

	Area	Sediment	Nitrate	Phosphorus	Ammonia	Nitrous Oxide	Methane
ADAS QA	Arable	16.3	12.4	8.5	5.1	9.2	0.1
	Grass	64.0	74.4	74.2	46.8	76.1	60.8
	Rough	18.6	6.5	8.3	0.0	4.2	2.4
	Woodland	1.1	4.4	1.1	-	0.8	-
	Non-field	-	2.3	8.0	48.0	9.7	36.7
Anthony et al. (2012)	Arable	12.6	10.9	7.8		3.7	0.3
	Grass	58.6	82.7	73.2	No Data	80.9	49.4
	Rough	28.8	5.2	14.4		4.9	8.5
	Woodland						No Data
	Non-field	0.0	1.0	4.6		10.4	41.8

Table 9 3 Comparison of source apportionment for national annual average agricultural pollutant loads a) derived using the Farmscoper export coefficients created for the IMP and scalar data derived from ADAS 1 km2 agricultural census data for 2014 and b) from Anthony et al. (2012) derived, calculated using 2004 agricultural census data

	Source	Sediment	Nitrate	Phosphorus	Ammonia	Nitrous Oxide	Methane
ADAS QA	Dairy	-	20.7	10.0	38.1	21.4	29.5
	Beef	-	15.5	8.9	31.3	15.5	29.8
	Sheep	-	17.9	12.7	17.4	31.3	40.6
	Pigs	-	0.1	0.0	0.2	0.1	0.0
	Poultry	-	0.7	0.5	1.1	0.6	0.0
	Fertiliser	-	17.2	10.9	12.0	20.8	-
	Soil / Residue	100.0	27.9	57.0	-	10.4	-
Anthony et al. (2012)	Dairy	-	18.5	8.4	No Data	8.3	29.5
	Beef	-	21.2	11.8		15.9	33.4
	Sheep	-	16.0	8.8		19.4	36.5
	Pigs	-	0.2	0.0		2.0	3.0
	Poultry	-	3.0	0.8		1.0	0.5
	Fertiliser	-	19.8	8.3		12.8	-
	Soil / Residue	100.0	21.3	61.9		42.5	-

Table 9 4 part 1 (Sediment and Nitrate-N): Comparison of spatial distribution of pollutant losses (as kg ha-1 of agricultural land) for national annual average agricultural pollutant loads a) derived using the Farmscoper export coefficients created for the IMP and scalar data derived from ADAS 1 km2 agricultural census data for 2014 and b) from Anthony et al. (2012) calculated using 2004 agricultural census data summarised at WFD waterbody scale.

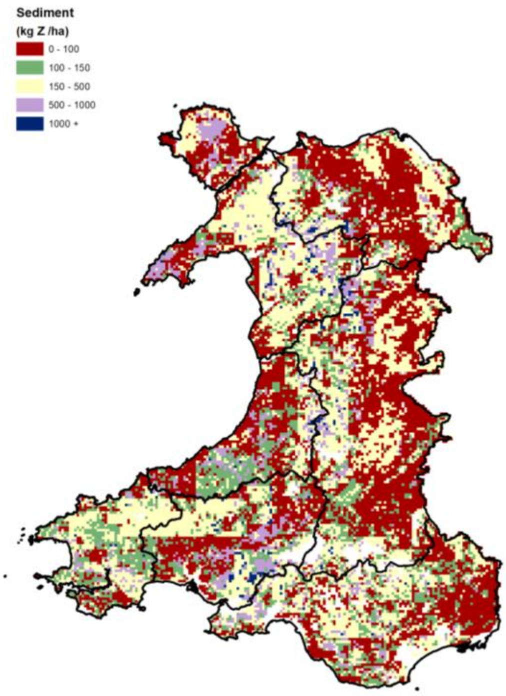
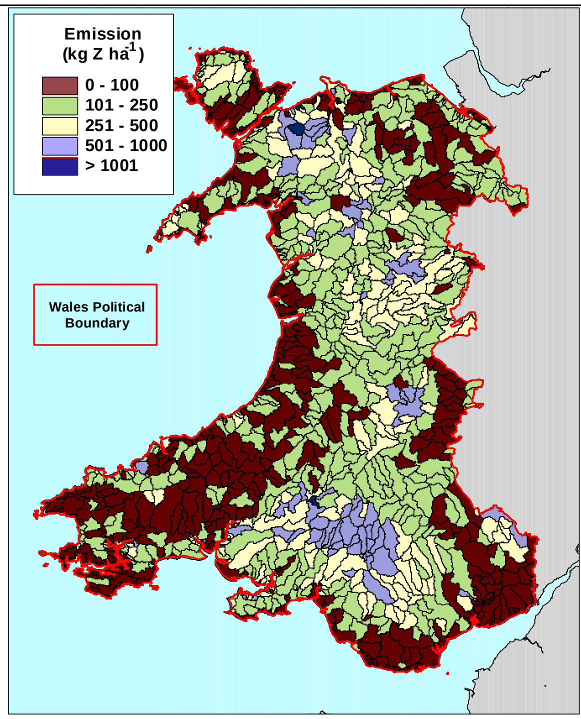
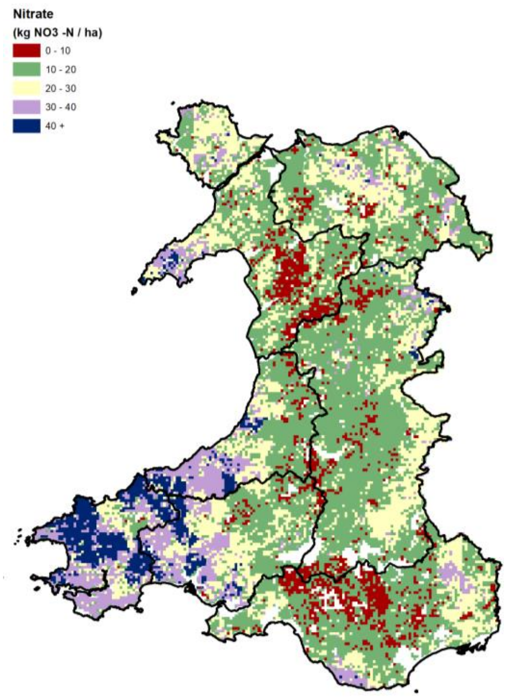
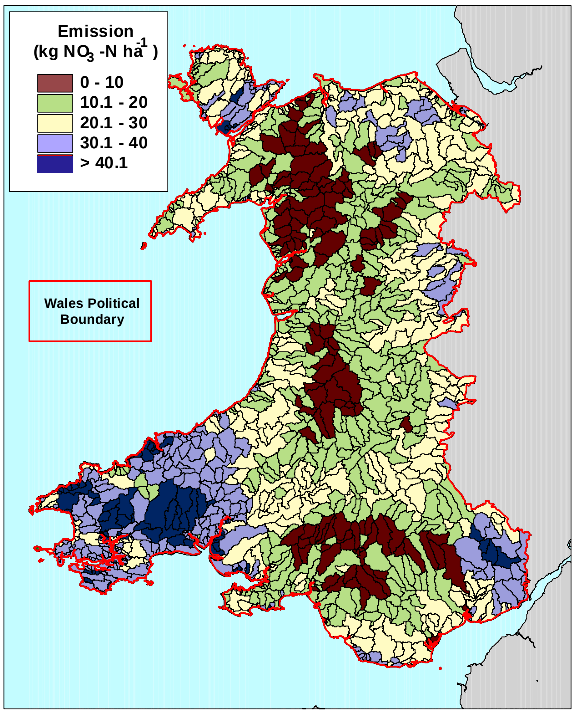
	ADAS QA of Farmscoper coefficieints for IMP, using 2014 Agricultural Census	**Anthony *et al.* (2012) using 2004 Agricultural Census**
Sediment		
Nitrate		

Table 9 4 part 2 (N20 and Phosphorous): Comparison of spatial distribution of pollutant losses (as kg ha-1 of agricultural land) for national annual average agricultural pollutant loads a) derived using the Farmscoper export coefficients created for the IMP and scalar data derived from ADAS 1 km2 agricultural census data for 2014 and b) from Anthony et al. (2012) calculated using 2004 agricultural census data summarised at WFD waterbody scale.

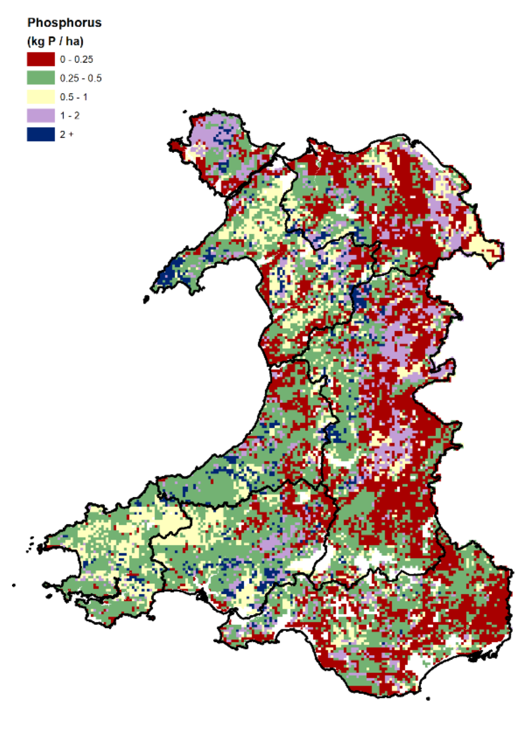
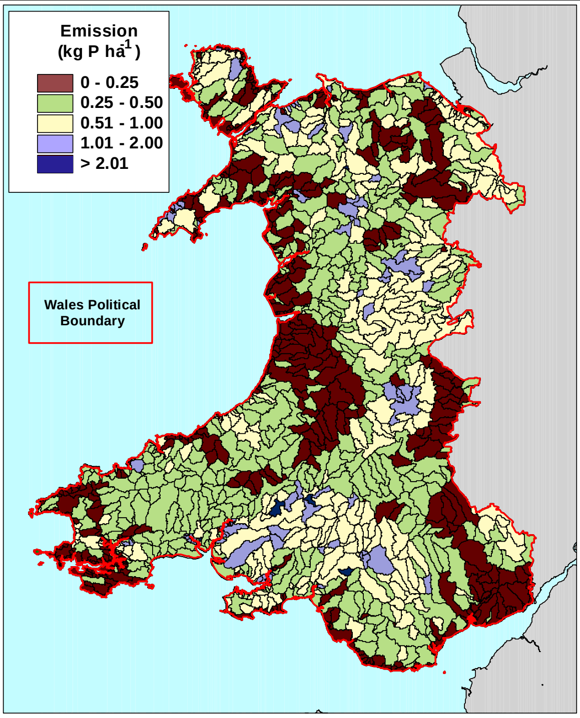
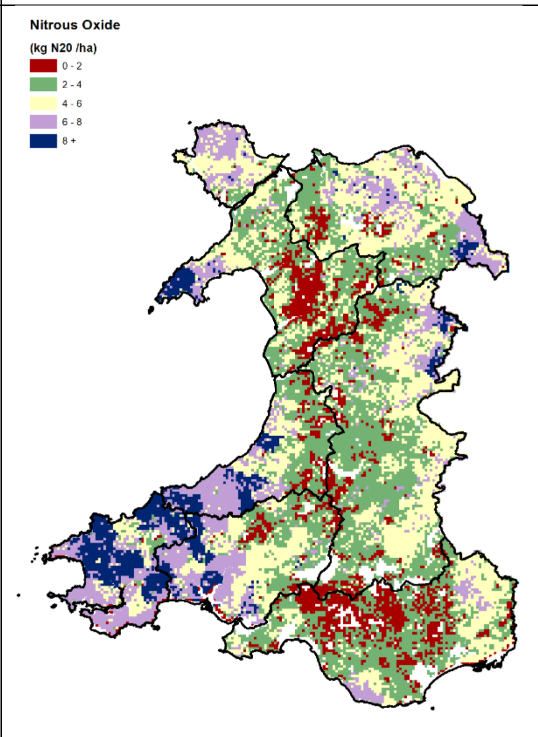
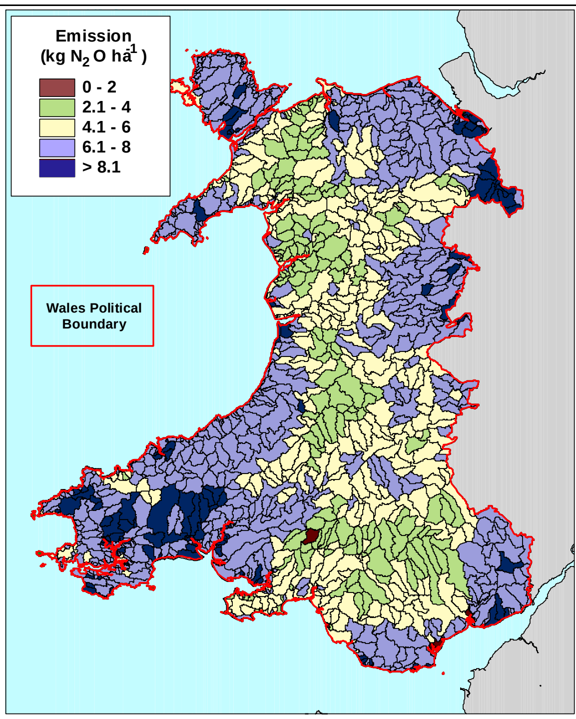
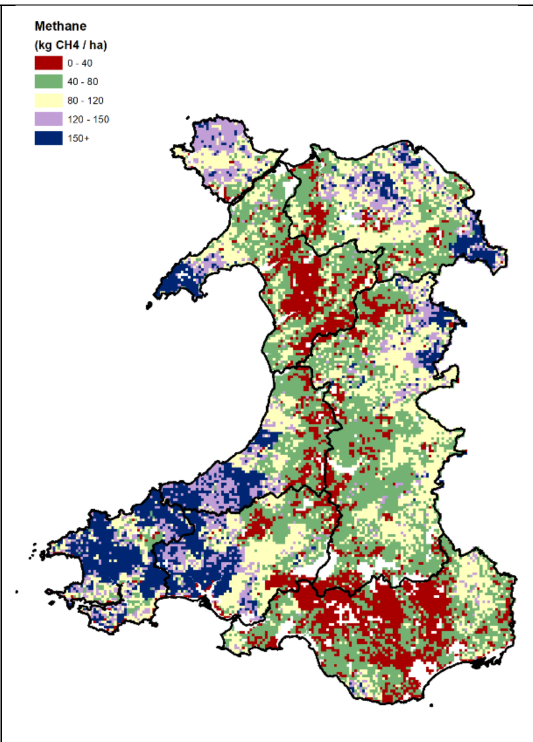
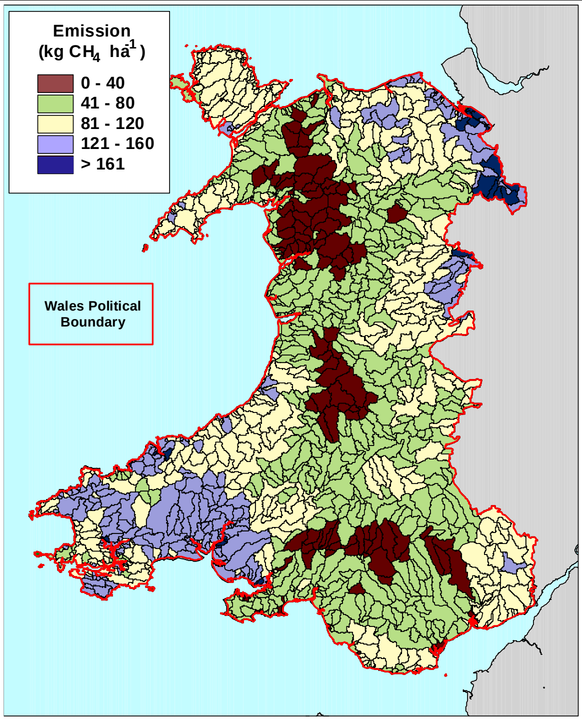
Phopsohrus		
Nitrous Oxide		

Table 9 4 part 3 (Methane): Comparison of spatial distribution of pollutant losses (as kg ha-1 of agricultural land) for national annual average agricultural pollutant loads a) derived using the Farmscoper export coefficients created for the IMP and scalar data derived from ADAS 1 km2 agricultural census data for 2014 and b) from Anthony et al. (2012) calculated using 2004 agricultural census data summarised at WFD waterbody scale.

Methane





The validation shows that the Farmscoper datacube can recreate the spatial patterns and emission totals from previous modelling work, when combined with an independent dataset from the IMP, particularly given the uncertainties in environmental and farm management data. There are some discrepancies in GHG and ammonia emissions, related to both i) changes in the official calculation approach since Farmscoper was developed and this project was started and ii) the sensitivity of the calculations to assumptions on certain farm management practices.

9.2.2 Validation: Comparison of the agricultural scalar data from SFARMOD with that derived from the ADAS 1 km2 data

To QA the SFARMOD agricultural scalar data that is used with the Farmscoper coefficients in the IMP, a comparison was performed between the SFARMOD scalar with those derived from ADAS 1 km2 agricultural census data for 2014 (used to construct Table 9 1). As SFARMOD only models farms greater than one full time equivalent (FTE), assumptions were made about the land, stocking and management on farms < 1 FTE (from data for such farms in the JAS). The ADAS scalar data, that from SFARMOD and for the small farms are shown in Table 9 5.

Table 9 5 Comparison of IMP scalar totals (from SFARMOD and for farm < FTE) with the scalars derived from ADAS 1km2 agricultural census data for 2014 (used to construct table 0-1). Values are in '000s

		ADAS QA	IMP		Difference (IMP – ADAS)
			SFARMOD output	Small Farms (<1 FTE)	
Area (ha)	Arable	90	66	14	-10
	Grassland	1,210	610	333	-267
	Rough grazing	257	248	30	22
	Woodland	296	69	32	-195
Fertiliser N applied (kg)	Arable	9,861	7,395	1,122	-1,344
	Grassland	65,429	41,907	12,221	-11,302
Fertiliser P applied (kg)	Arable	2,749	1,492	406	-850
	Grassland	9,982	11,430	5,666	7,114
Dairy (kg N)	FYM to arable	365	73	4	-287
	FYM to grassland	3,285	3,497	40	252
	Slurry to arable	1,230	164	19	-1,047
	Slurry to grassland	11,073	31,850	174	20,951
	Excreta at grazing	13,269	15,985	267	2,983
	Excreta total	29,222	51,569	601	16,448
Beef (kg N)	FYM to arable	1,160	11	256	-893
	FYM to grassland	8,083	8,819	1,805	2,542
	Slurry to arable	164	0	36	-127
	Slurry to grassland	1,472	17,405	324	16,258
	Excreta at grazing	20,426	17,007	4,546	1,127
	Excreta total	31,305	43,242	8,778	12,491
Sheep (kg N)	FYM to arable	131	0	13	-118
	FYM to grassland	1,176	2,734	97	1,654
	Excreta at grazing	58,929	28,005	4,793	-26,132
	Excreta total	60,236	30,739	5,050	-25,480
Pig (kg N)	FYM to arable	4		3	
	FYM to grassland	38		29	
	Slurry to arable	5		3	
	Slurry to grassland	47		27	
	Excreta at grazing	0		0	
	Excreta total	94		101	
Poultry (kg N)	Manure to arable	98		28	
	Manure to grassland	885		256	
	Excreta at grazing	0		6	
	Excreta total	983		444	

Table 9-5 shows there is a difference of 267,000 ha in the area of grassland, with a much lower figure used in the IMP. The 2019 report on the JAS data for Wales[3] shows an

[3] June 2019 Survey of Agriculture and Horticulture: results for Wales (gov.wales)

increase of approximately 200,000 ha since 2010, with the following note “The increase in the total land on farm holdings in recent years is believed to be linked to issues with registration of land. Principally by the re-registering of existing land and also continued registration of land no longer in agricultural use”. Therefore, it is likely that the IMP data is more correct, with the ADAS QA dataset affected by this issue. Nitrogen fertiliser rates expressed per hectare are broadly comparable between the ADAS QA data (based of British Survey of Fertiliser Practice (BSFP) information) and SFARMOD (which are initially based on recommended rates from RB209), but phosphorus fertiliser rates in SFARMOD based on an independent analysis of BSFP data for 2017) are higher than ADAS QA figures.

Section 3.2.4 shows that stocking numbers in SFARMOD are comparable to national totals. However, in Table 9 5, total nitrogen in excreta for beef and dairy cattle are higher in the IMP data, whilst the sheep total nitrogen in excreta are lower. SFARMOD calculates a volume of material deposited (plus any dilution) and then converts this to a nitrogen load using a concentration, rather than directly assigning a nitrogen load, and considers a cow and its followers together rather than individual stock categories, so it is difficult to directly compare with the ADAS QA and identify causes of the differences. SFARMOD also assumes that much more beef excreta is handled as slurry than in the ADAS QA. The difference in sheep excreta is largely due to the ADAS QA using a lowland ewe value for sheep excreta, whereas SFARMOD has used a lower value that takes into account the large numbers of light breeds of sheep that uniquely characterise the traditional breeds in the mountains of Wales.

Pigs and poultry were not included in the IMP, but Table 9 5 shows that the amount of excreta produced by these livestock is less than 1% of the total from all livestock, thus the consequences of this are not important at a large scale.

10 Greenhouse Gas and Ammonia Emissions

Authors: Amy Thomas and Richard Gooday

10.1 Introduction to the model QA

As described in Section 9, pollutant loss coefficients for methane, nitrous oxide and ammonia were derived from Farmscoper and combined with data from the SFARMOD to predict total agricultural pollutant loads for the different scenarios. The QA of the pollutant coefficients is described in Section 9. This section presents the QA of agricultural greenhouse gas (GHG) and ammonia outputs derived from the combination of SFARMOD outputs and Farmscoper coefficients. Section 11 focuses on water quality outputs.

The QA of GHG and ammonia emissions comprises 3 sections:

- Building Understanding: Presenting the breakdown between the IMP modelled farms > 1 FTE and small farms under this threshold (Section 10.2.1).
- Validation: Comparison of IMP total emissions for ammonia and GHG with totals from the ADAS QA (Section 9) and the National Atmospheric Emissions Inventory (NAEI) (Section 10.2.2).
- Validation: Comparison of the IMP emissions at 10 x 10 km2 with NAEI data (10.2.3).

10.2 Introduction to the modelling

Ammonia and GHG emissions are calculated at the DMU level by aggregating the SFARMOD scalar data for that DMU with the relevant Farmscoper coefficients, accounting for the DMU climate zone, soil type and farm type (see Section 9.2 for more detail). The same process is applied for farms not modelled by SFARMOD (those < 1 FTE), where typical stocking rates and management are used to calculate the scalar data to combine with the Farmscoper coefficients.

10.2.1 Building Understanding: Presenting the breakdown between the SFARMOD farms (> 1 FTE) and small farms under this threshold

Table 10.1 shows the IMP results from linking SFARMOD outputs with the Farmscoper coefficients and the results for farms less than 1 FTE which are not modelled by SFARMOD.

The results presented here were requested by WG to aid with the interpretation of model outputs by determining the contribution of the SFARMOD modelled farms relative to smaller farms with < 1 FTE. SFARMOD farms contribute 92% for ammonia and 86% for nitrous oxide, despite only occupying 70% of the agricultural land (see Table 9 5), reflecting the greater stocking density and fertiliser use on these farms.

Table 10.1 Total modelled annual average agricultural pollutants for CH4 and N2O, NH3 for the IMP baseline, disaggregated for the farms modelled by SFARMOD and the smaller farms.

	SFARMOD	**Farms < 1 FTE**	**IMP total**
Methane (kt CH_4)	120	13	133
Nitrous Oxide (kt N_2O)	5.6	0.8	6.4
Ammonia (kt NH_3-N)	31	3	34

10.2.2 Validation: Comparison of IMP total emissions for ammonia and GHG with totals from the ADAS QA and the NAEI

Total outputs of agricultural pollutants for air quality and GHG emissions are shown in Table 10.2. These values shown are the modelled outputs from the IMP plus the projections for small farms (i.e. the “Total” values in Table 10.1 where national values are calculated using the IMP-Farmscoper coefficients linked to ADAS 1 km2 JAS data for 2014).

IMP totals are close to the ADAS QA totals in Table 10.2 and to the NAEI data for both methane and nitrous oxide. However, for ammonia, the IMP is higher than both the ADAS and NAEI values. This is related to differences in management assumptions and the large variations in ammonia losses depending upon where excretion occurs. Table 10.2 shows that there is much more dairy excreta in the IMP than in the ADAS QA. This will cause much more ammonia pollution due to the amount excreted on hardstanding (whilst the cows are waiting to be milked). For other pollutants, this is not so important and so differences in dairy excreta between IMP and ADAS QA are negated by comparable differences in sheep excreta.

The final column in Table 10.2 shows nitrous oxide levels adjusted for peat N2O emissions. Peatland N2O emissions are modelled separately (see Section 13.2.1) using a different method based on wetland coefficients. The value presented here is the total N2O value from Table 10.1 minus the N2O stored in peatland. This is highlighted to avoid double counting of peatland N2O.

Table 10.2 Total modelled annual average agricultural air quality and GHG emission for the IMP baseline compared to totals from Table 9-1 (calculated using the Farmscoper coefficients and data derived from ADAS 1 km2 agricultural census data for 2014) and National Atmospheric Emissions Inventory totals for 2018.

	Methane (kt CH_4)	**Nitrous oxide* (kt N_2O)**	**Ammonia (kt NH_3 N)**	**Peat-corrected nitrous oxide* (kt N_2O)**
Total (IMP + <1 FTE farms)	133	6.4	34	6.3
ADAS 2014 + FS-IMP (Table 8.1)	150	8.0	25	N/A
National Atmospheric Emissions Inventory (2018)	139	5.2	18	N/A

10.2.3 Validation: Comparison of distribution of emissions for Air Quality and GHG from SFARMOD-Farmscoper with NAEI

Figure 10.1 shows model outputs compared to data from the National Atmospheric Emissions Inventory (NAEI) 2018. These are gridded (1 km resolution) air quality data (t) for methane (CH4), ammonia (NH3) & nitrous oxide (N2O) for the year 2018. Individual raster/ascii grids for each type of emissions sector were used for the QA[4] . Data were aggregated to 10 km due to uncertainty in the agricultural data used in NAEI at 1 km (NAEI aggregate data to coarser scales and then distribute averages on a 1 km grid due to data sensitivity).

The over-prediction in ammonia at the national scale is systematic, and so the IMP modelled data is nearly always greater than the NAEI data. For all three pollutants, there is a correlation between the IMP output and the NAEI, but the spatial agreement here is worse for ammonia (r2 = 0.51) than for CH4 and N2O (r2=0.95 and 0.76 respectively). This reflects the greater difference in national predictions. Some of this scatter may be associated with pig and poultry farms, which are locally significant but only contribute a few percent of the total at the national scale and were not included in the IMP. Note that the spatial pattern is not important for nitrous oxide and methane as GHG benefits are quantified at the national scale. For ammonia, the disagreement identified here will feed into the health modelling in Section 12.

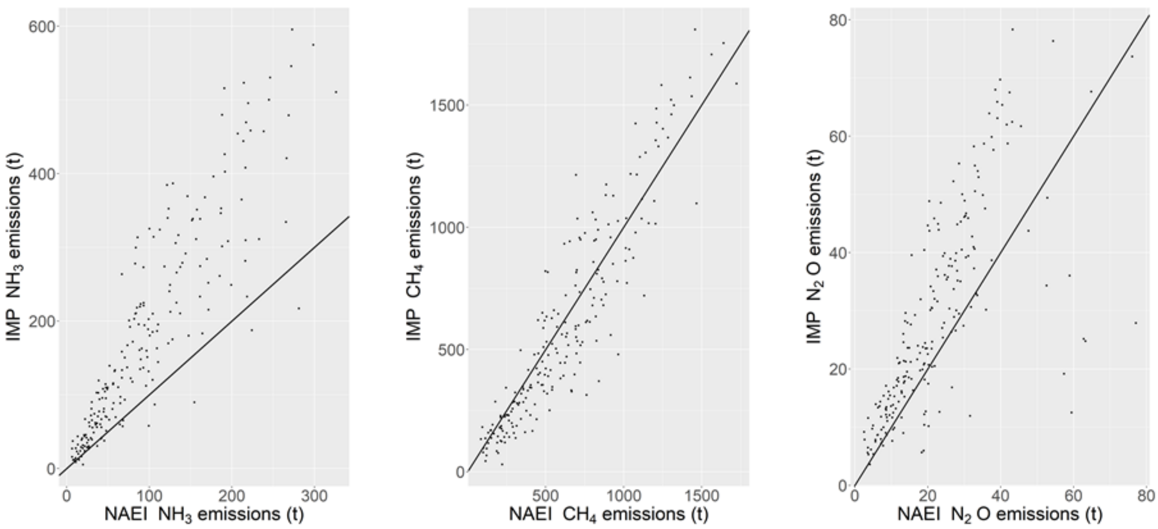


Figure 10.1 Comparison of IMP outputs to 10 X 10 km gridded NAEI data for NH3, CH4 and N2O.

[4] For each pollutant, the NAEI provide a breakdown by sector - the relevant layer for the agricultural sector is labelled "agric- SNAP 10 (Agriculture, Forestry and Landuse Change)" on the NAEI website.

11 Water quality

Authors: Amy Thomas

11.1 Introduction to the model QA

As described in Section 9, pollutant loss coefficients for nitrate, phosphorus and sediment were derived from Farmscoper and combined with data from the SFARMOD to predict total agricultural pollutant loads for the different scenarios. The modelling then uses statistical relationships derived from loads and observed concentrations to predict changes in concentrations under the different modelled scenarios. The QA in this section includes comparison of the results against some of the data used in Section 9, plus additional validation against river water quality measurements. This is split into four sections:

- Validation: Comparison of IMP total emissions for nitrate, phosphorus and sediment with totals from the ADAS QA (Section 9) and the SEPARATE dataset (Section 11.2.1).
- Validation: Comparison of IMP catchment scale emissions for nitrate, phosphorus and sediment with modelled values from the SEPARATE dataset (Section 11.2.2)*.
- Validation: Comparison of IMP derived concentrations for nitrate and phosphorus with observed water quality data (NRW Water Framework Directive (WFD) data and Harmonised Monitoring Scheme) (Section 11.2.3).
- Validation: Comparison of IMP WFD status categorisations with NRW allocated WFD status[5] (Section 11.2.4).

*Note, for sediment it is only possible to compare against modelled data (section 11.2.2) as large uncertainties in predicting and measuring sediment concentrations meant there was no comparison against observed monitoring data for sediment (section 11.2.3).

11.2 Introduction to the modelling

Nitrate, phosphorus and sediment loads are calculated at the DMU level by linking the SFARMOD scalar data for that DMU with the relevant Farmscoper coefficients, accounting for the DMU climate zone, soil type and farm type (see Section 9.2 for more detail). The same process is applied for farms not modelled by SFARMOD (those < 1 FTE), where typical stocking rates and management are used to calculate the scalar data to combine with the Farmscoper coefficients. The DMU results can then be aggregated to produce results for coarser spatial scales (catchment or national).

The IMP totals for WFD catchments were calculated by aggregating the DMU outputs within each catchment. To calculate total loads for N and P, non-agricultural sources of pollutants were accounted for using SEPARATE outputs. The SEPARATE dataset (Zhang et al., 2014;

[5] This is important as changes in WFD status affect the environmental valuation of scenarios assigned in the IMP (see Section **Error! Reference source not found.**).

data available online[6]) provides estimates of discharged loads of nitrogen, phosphorous and fine-grained sediments to rivers in England and Wales from multiple sector sources (agricultural and non-agricultural), reported at WFD catchment scale).

As part of the project which created SEPARATE (Defra project WQ0223), algorithms were developed to convert from loads to concentration for N (as 95th percentile) and Orthophosphate (OP) (as annual average) in order to validate the SEPARATE outputs. These algorithms accounted for dilution, retention and any contributions from groundwater and the proportion of the total P load that was soluble (i.e. OP). We used these to calculate concentrations from the IMP accumulated total loads. Data for N and P are processed to units reflecting the relevant monitoring data: annual average concentration for P and 95th percentile for N.

River sediment concentrations are controlled by event driven inputs and in-river processes occurring over a range of timescales, so it is hard to measure average concentrations using infrequent grab samples and difficult to predict these from annual average inputs to watercourses as predicted by the IMP. Therefore, sediment outputs are calculated only as annual average loads, rather than concentrations.

To assign WFD status for Phosphorus, site specific thresholds provided by NRW were used, which are based upon altitude and alkalinity. Phosphorus status can be assessed at multiple locations throughout a catchment, whereas the modelling predicts P concentration at the most downstream point, thus the most downstream threshold has been used to assign status. Nitrate status is based upon EU Nitrate Directive target of 50 mg l-1.

When considering the scenario impacts predicted by IMP-Farmscoper, the following points should be considered:

- Farms < 1 FTE do not respond to the scenarios, but these contribute between 12% (for nitrate) and 25% (for sediment) of the total agricultural load.
- Data outputs relate to a new long-term average reflecting land use and management for the scenario - some measures might change soil P status or soil organic N supply, which would respond over a period of 10+ years.
- There is no accounting for time lags in groundwater catchments, which could be particularly important for nitrate.
- Predicted loads are based on average climate data (1961-1990) (as that is the data within Farmscoper v4) and do not reflect variations in weather between years.
- Changes in water quality are not modelled for lakes, but these may be important for recreation and associated businesses in Wales.
- We should also note that SEPARATE data used here represent pollution for 2010, which may introduce some error into calculations of total loads and concentrations.

[6] https://data.gov.uk/dataset/3e698568-8492-4dfd-aa11-3439d77cd71a/source-apportionment-of-annual-nutrient-and-sediment-loads-to-rivers-in-england-and-wales-from-the-separate-framework

11.2.1 Validation: Comparison of IMP total emissions for nitrate, phosphorus and sediment with totals from the ADAS QA

Table 11.1 compares Nitrate-N, Phosphorous and Sediment outputs from the IMP (from both SFARMOD and the < 1 FTE farms) with the outputs from ADAS QA of the Farmscoper coefficients (as shown in Section 9) and SEPARATE data. IMP totals are lower than, but comparatively close to, the ADAS QA values for nitrate and phosphorus (90% and 85% of the total respectively), whilst the IMP sediment value is only 66% of the ADAS QA value. The IMP sediment totals are also much lower than the SEPARATE values. Sediment predictions are very sensitive to soil type, and there is some disagreement between the values from the two modelling assessments, with no way to determine which is 'correct'. There will also be differences in the climate data used, which is more important for sediment than the other pollutants. It should be noted that this is a comparison between modelled outputs, and Table 11.1 is thus indicative of minimal disagreement being introduced at the national scale by using agricultural data from SFARMOD compared to that used in the ADAS QA (derived from 1km2 June Agricultural Survey).

Table 11.1 Total modelled annual average agricultural pollutants for of N, P and sediment, for the IMP baseline compared to totals from Table 9 1 from the ADAS QA (calculated using the Farmscoper coefficients and data derived from ADAS 1km2 agricultural census data for 2014)

	Nitrate (kt NO_3-N)	**Phosphorus (kt P)**	**Sediment (kt)**
IMP: SFARMOD	30.1	0.72	194
IMP: Farms <1FTE	4.1	0.18	68
IMP Total	34.2	0.90	262
ADAS QA	37.6	1.04	393
SEPARATE	26.3	0.97	479

11.2.2 Validation: Comparison of IMP catchment scale emissions for nitrate, phosphorus and sediment with modelled values from the SEPARATE dataset

Figures 11.1 to 11.4 use agricultural pollutant data from the IMP and non-agricultural data from SEPARATE, compared with the total loads from SEPARATE, as ultimately it is the total load from all sectors that determines the impact of any change in the agricultural load and any associated change in concentration or WFD status (which are discussed in the following sections).

Figure 11.1 shows there is a reasonable correlation for sediment, but a systematic under prediction (although this is to be expected given the differences in Table 11.1). Comparison is better for cumulative load ($r2= 0.487$) than local ($r2= 0.004$), suggesting that disagreement averages out at larger spatial scales. Figure 11.2 suggests there is no obvious spatial pattern to the disagreement, although IMP values are slightly higher in the west and lower in the east.

Correlations are much better for nitrate ($r2 = 0.71$) and phosphorus ($r2 = 0.66$) as shown in Figure 11.3. There are some catchments where IMP loads are much smaller than SEPARATE, but even the SEPARATE ones are low so the contribution to errors in the national loads or any cost-benefit analysis would be small, and WFD status is unlikely to be an issue as concentrations will also be small. As before, performance for accumulated load is improved compared to local load predictions since disagreement averages out at larger spatial scales (r2 values for cumulative N and P are 0.95 and 0.82 respectively, plots not shown). Figure 11.4 shows the phosphorus loads are often higher in the west of Wales, the same as for sediment (and so are a combination of higher particulate P and greater mobilisation due to environmental conditions). Differences in nitrate do not always correlate with spatial differences in phosphorous and sediment since nitrate losses are more dependent upon agricultural inputs than environmental data than is the case for phosphorus and sediment.

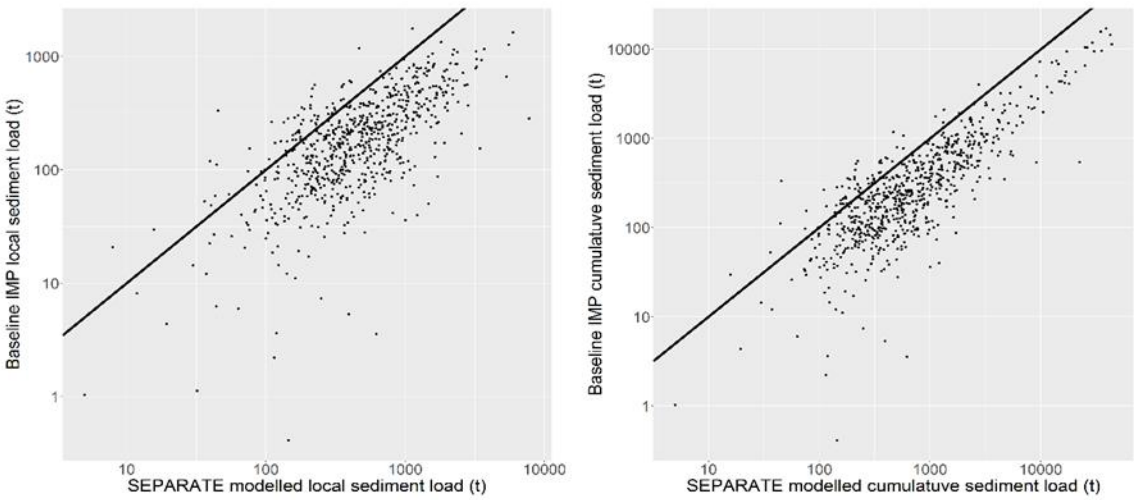


Figure 11.1 Graphs comparing IMP modelled local (left) and cumulative (right) sediment loads to an alternative set of modelled values from SEPARATE. Black lines indicate the 1:1 relationship. Data are shown at the WFD catchment scale. Plots are filtered to catchments where IMP has >=70% coverage: total number of catchments covered by both IMP & SEPARATE = 422.

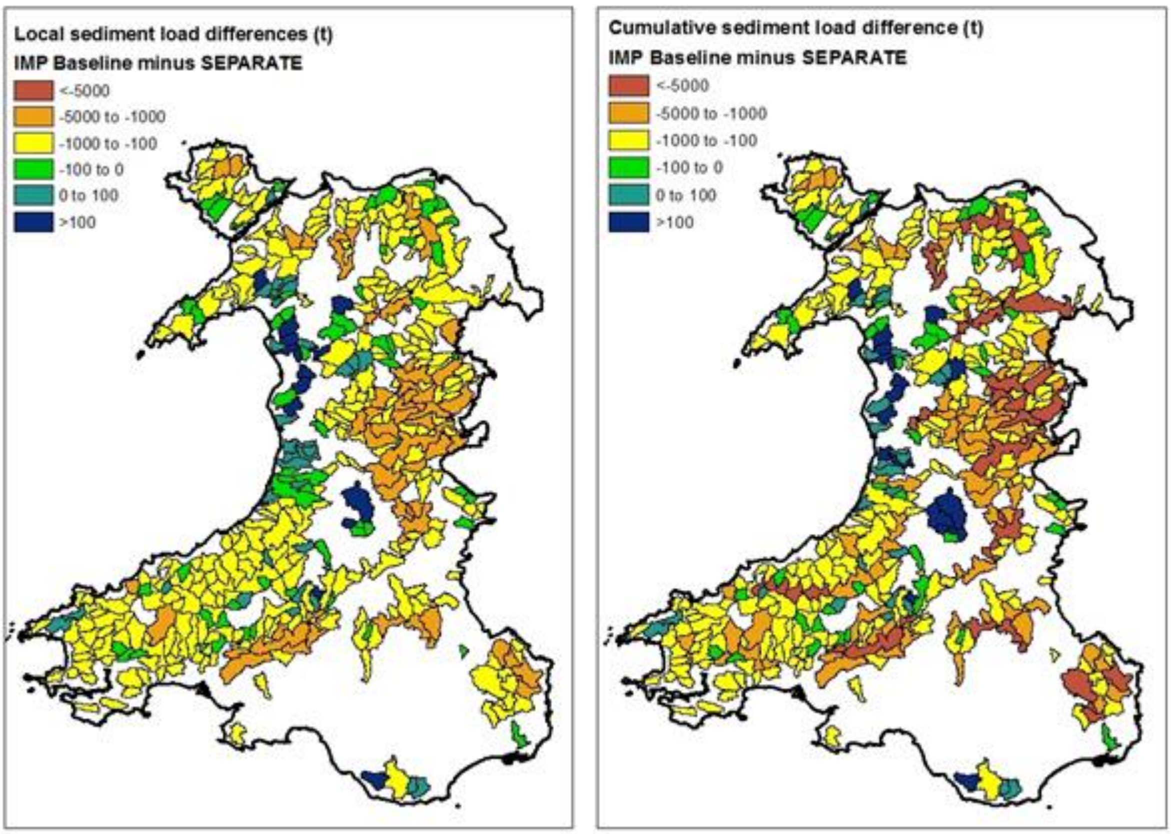


Figure 11.2. Spatial comaprison of IMP local and cumulative WFD waterbody sediment loads from all sectors with values from SEPARATE, to show where disagreements are greatest.

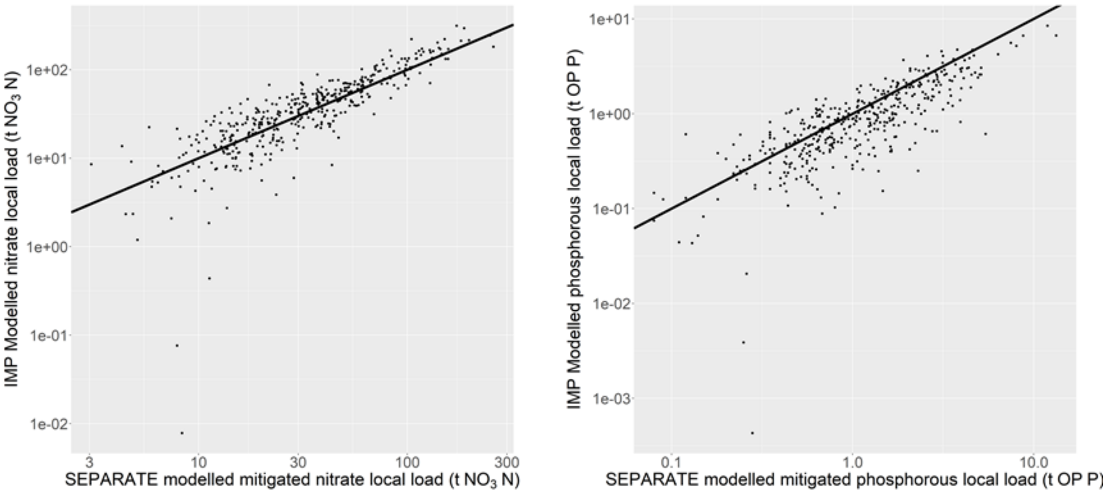


Figure 11.3 Graphs comparing IMP modelled local N (left) and P (right) loads (tons) to modelled mitigated loads (tons) from SEPARATE. Graphs display the relationship (with the 1:1 line in black) between IMP and SEPARATE. Data are shown at the WFD catchment scale. Plots are filtered to catchments where IMP has >=70% coverage: total number of catchments covered by both IMP & SEPARATE = 422.

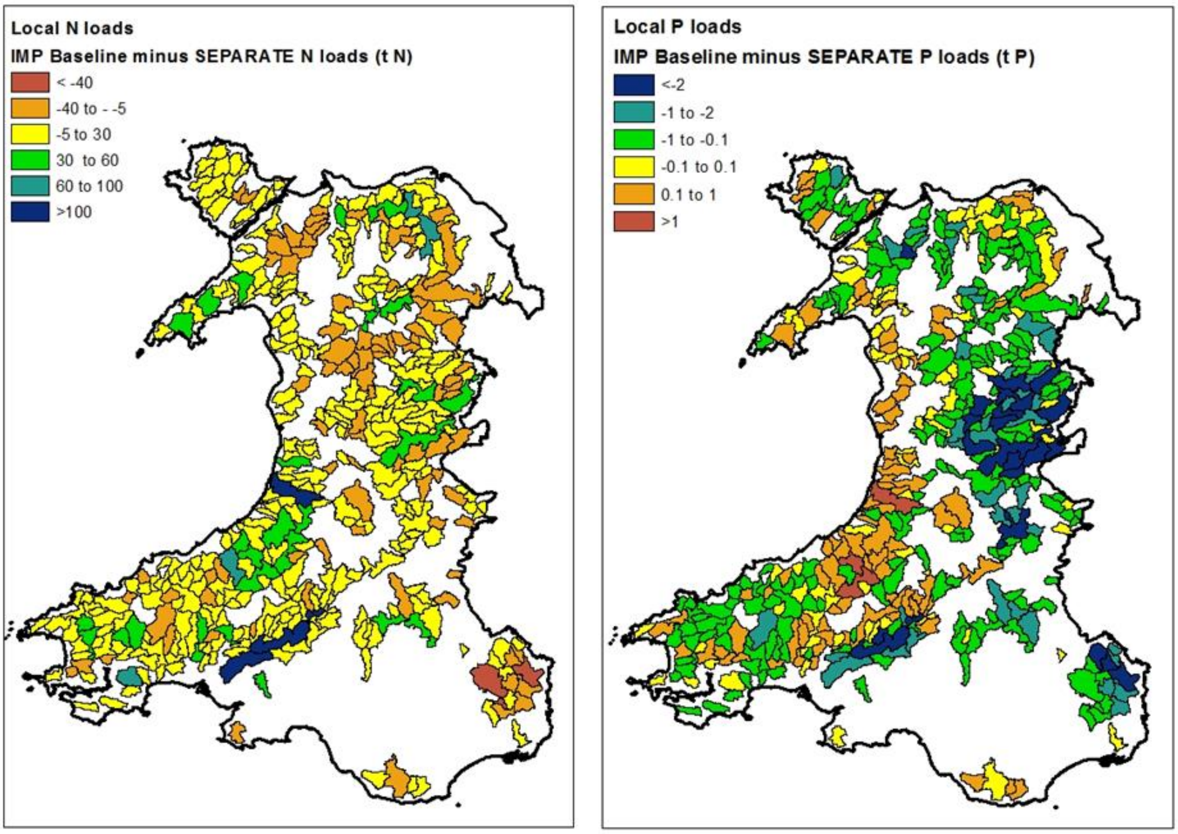


Figure 11.1 Maps comparing IMP modelled local N (left) and P (right) loads (tons) to modelled mitigated loads (tons) from SEPARATE. Maps display the differences (IMP minus SEPARATE) in N and P loads. Data are shown at the WFD catchment scale.

11.2.3 Validation: Comparison of IMP derived concentrations for nitrate and phosphorus with observed water quality data

Two separate measured datasets of in-stream N and P concentrations were used for assessment of predicted concentration at the WFD catchment scale: the Harmonised Monitoring Scheme (HMS) and NRW data collected for regulatory purposes.

For both the HMS and the NRW WFD data, we selected the most downstream monitoring point on the main river. The modelled data effectively represents the point where the main waterbody flows into the next sub-catchment; some of the disagreement between modelled and measured values will reflect the spatial discrepancy between the measurement point and the location represented by the modelling (and the corresponding catchment areas).

NRW WFD data

These data were taken for the WFD classification interim cycle 2 2017, (data 2014-2016). NRW collected the raw data, modified and checked the data for suitability for WFD classification. Frequency of sampling over this time period varied, from site to site and between N and P - eight samples is the minimum requirement to have confidence to use the data for classification. We filtered the available data to sites on the main river with over eight samples, within 500m of a river outflow of the catchment, downstream of any tributaries. These were deemed appropriate to our analysis since they should be representative of the water quality at the outflow of the WFD catchment.

Data on OP, and NO3 concentrations (for monitoring sites appropriate to our analysis) were available for 54 and 282 catchments respectively. From the NRW dataset, a time-weighted average OP concentration was determined for each waterbody by taking the mean of all OP measurements for that location. The 95th percentile NO3 concentration was calculated taking the value at quantile 0.95 of all NO3 measurements. Note that any measurement value indicated by < in the dataset would be halved, which is a standard procedure.

Note of caution on TON and OP data (NRW): There are potential issues with some of the analysis results for TON and orthophosphate in this time frame. NRW considers all Total Oxidised Nitrogen (TON) and Orthophosphate (OP) data determined by the "low range" method for all freshwater sampling points between the dates of 01/07/2014 and 31/07/2016, to be anomalous. (That is to say the reliability of the above data cannot be guaranteed, due to the perception that environmental levels were not themselves increased during that period).

HMS (Harmonised Monitoring Scheme) data

The HMS dataset was established to provide an archive of water quality data UK-wide. This is used to provide information for international obligations, including analyses of long-term trends and estimating river borne inputs of certain water quality determinants to the sea. The data are jointly owned by Natural Resources Wales, Environment Agency, Scottish Environmental Protection Agency, and Northern Ireland Agri-Food and Biosciences Institute. Data are available for 230 sites across the UK, starting from 1975.The data include long-term data from a small number of catchments in Wales (35 of those modelled by IMP). Very few (4) points were suitable for use in QA once filtered to meet our criteria as per the NRW monitoring data.

Averages were calculated over the most recent available 10 years (2003-2013) of OP & NO3 concentrations. Note that the frequency of sampling is variable - from 19 to 55 for OP & NO3 over the 10-year period compared here.

Results

Figure 11.5 shows that the IMP modelling does not routinely over or underestimate in-stream N concentrations when compared to measured data. The exception to this is WFD catchments which were assigned the minimum N concentration of 0.1 mg l-1 in the SEPARATE calculations but have much higher values in the monitoring data. Some of this under-estimation of concentrations will reflect agricultural pollution from England, which was not modelled here.

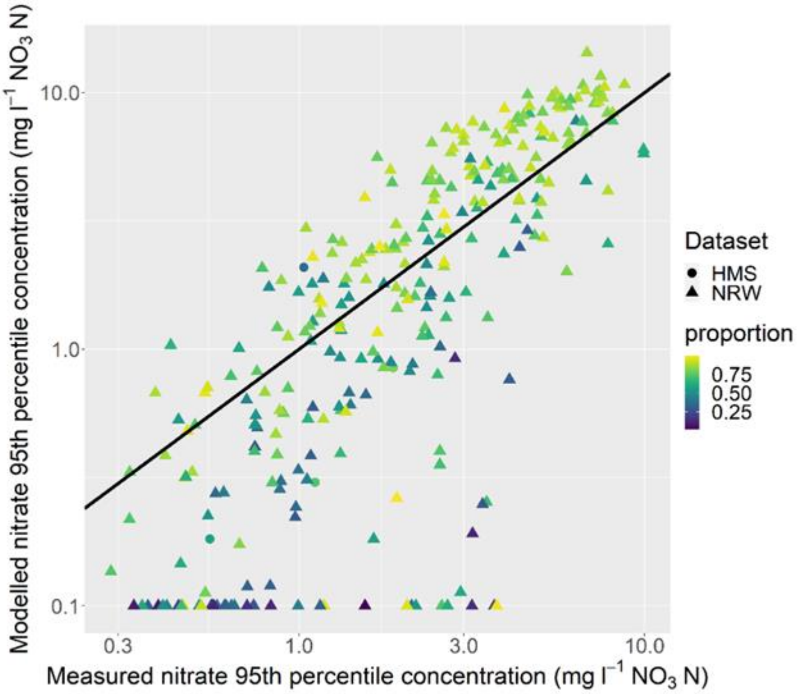


Figure 11.5 Graph comparing modelled cumulative N concentration from IMP outputs with the measured data from NRW (n=282) and HMS (n=4). Contains NRW information © Natural Resources Wales and database right. All rights reserved. NRW points were filtered to remove erroneous outliers (N>110), catchments with very low coverage in IMP (<30%), and catchments where estimates were based on < 8 samples.

The correlation is low for the NRW dataset (r2 = 0.18). This indicates that for any individual data point (representing a WFD sub-catchment) there is potentially a large difference, reflecting in part the uncertainty in model input data and environmental data at local scales, i.e. either SFARMOD is over-predicting manure and pollution inputs in these areas, or assuming more intensive land use, or the soils and climate categorisation within Farmscoper is not representative of that catchment. However, better performance for local loading when compared to SEPARATE data (Figure 11.5) suggests that converting to concentrations is responsible for more of the apparent error in the modelled data - the relationships presented are not too dissimilar to those achieved in the validation of the SEPARATE outputs, reflecting the difficulty in predicting concentrations at the national scale, using nationally available datasets and without resorting to local calibration. It is also worth noting that measured data will also contain sources of error, and are not a perfect representation of average concentrations, since they are based on intermittent sampling and may miss important peaks in pollution associated with episodic nutrient inputs or precipitation events.

As per nitrate, the correlation for OP (Figure 11.6) is low for the NRW dataset (r2 = 0.31). Unlike for nitrate, where the model struggled with low concentrations, for OP the modelling tends to over-predict at low concentrations. However, errors at low concentrations are not important when considering the valuation of any scenario impacts as such catchments will more than likely already be at Good or High status (see next section).

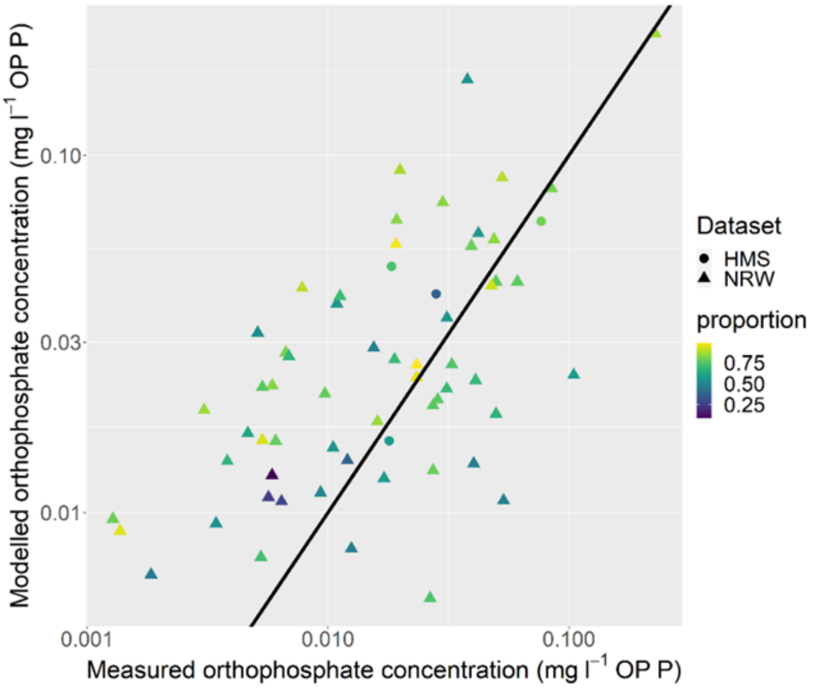


Figure 11.6 Graph comparing modelled cumulative P concentration from IMP outputs with the measured data from NRW (n=54) and HMS (n=4) Contains Natural Resources Wales information © Natural Resources Wales and database right. All rights reserved. NRW points were filtered to remove catchments with very low coverage in IMP (<30%), and catchments where estimates were based on < 8 samples.

Comparison of the spatial distribution of N concentrations is shown in Figure 11.7. This shows that most of the IMP over-prediction occurs in the southwest and so may be associated with the high dairy stocking values shown in Table 9.5. A lot of the highest NRW data (which will include some of the data along the x-axis in Figure 11.5) are isolated catchments. This suggests that there may be other factors associated with the monitoring data that may influence interpretation such as:

i) Sampling being targeted at known pollution events;

ii) There being significant point sources targeted by the monitoring (e.g. a sewage treatment works) but not reflected in the SEPARATE data;

iii) There being significant local agricultural issues (such as pig or poultry farms) not captured in the modelling data; or,

iv) Genuine errors.

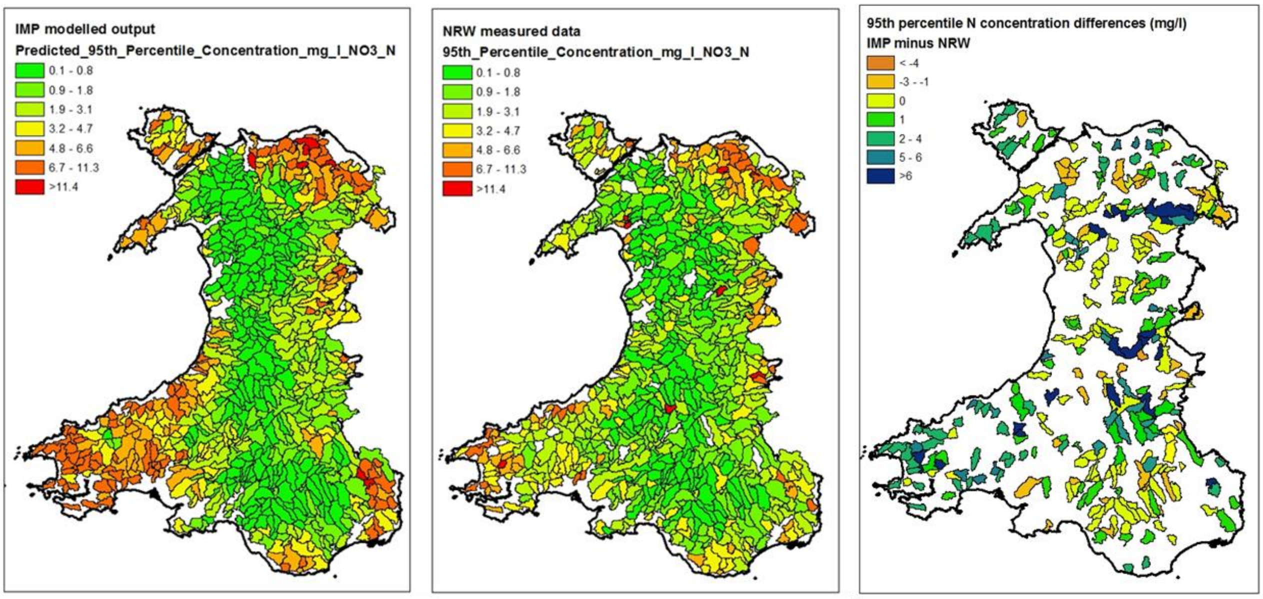


Figure 11.7 Maps of modelled N concentration and the measured data from NRW by WFD catchment. Contains Natural Resources Wales information © Natural Resources Wales and database right. All rights reserved.

The spatial data for OP (Figure 11.8) shows the much smaller monitoring dataset available, and how it is more reflective of central upland catchments where agricultural pressures and concentrations are low (plus the Cleddau catchment, where this is not the case). The observed data are thus not such a good dataset for the validating the modelled concentrations predicted on the northern and coastal parts of Wales that are generally higher than the range of the observed data. The major difference in Figure 11.8 is cluster of high measurements in the Southeast not reflected in the modelling, which may reflect similar issues as per nitrate.

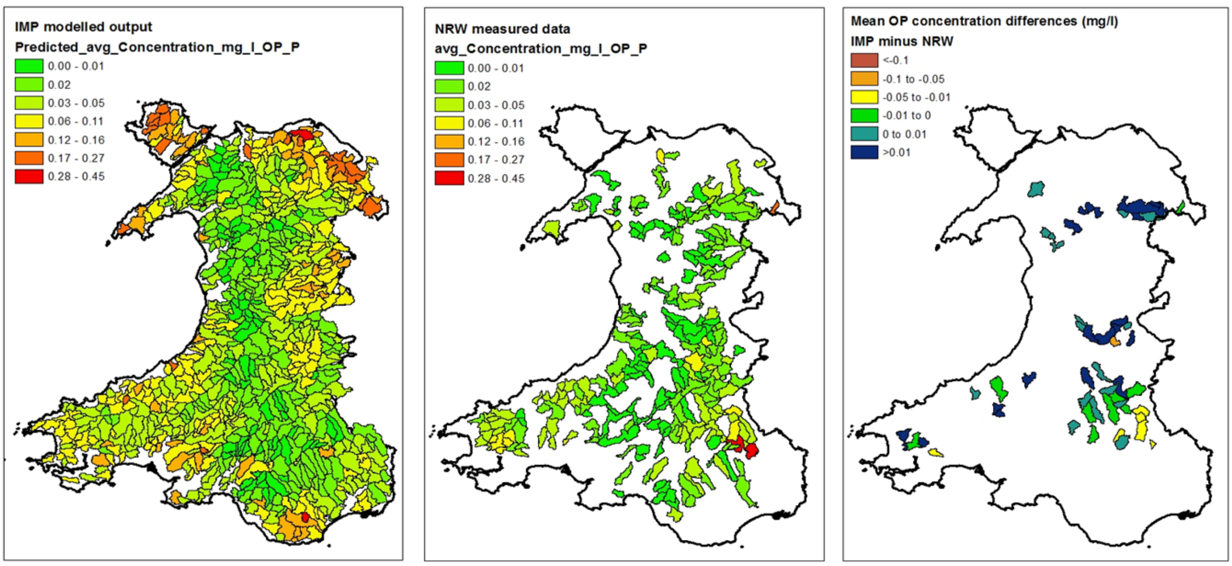


Figure 11.8 Maps of modelled OP concentration and the measured data from NRW by WFD catchment. Contains Natural Resources Wales information © Natural Resources Wales and database right. All rights reserved.

11.2.4 Validation: Comparison of IMP WFD status categorisations with NRW allocated WFD status

WFD status was also assessed for P, as this is subject to the relationship between predicted values and thresholds, and any error in WFD status and change in status will affect valuations. It is therefore important to note that small discrepancies in modelled N and P may be important in terms of status, depending on how close a catchment is to the relevant threshold. P status can be characterised as either 'High', 'Good', 'Moderate', 'Poor' or ‘Bad’. Unique thresholds exist for each of these categories and for each waterbody. The thresholds vary between waterbodies according to altitude and alkalinity.

The data used in this assessment are as follows:

P thresholds:

These were provided by NRW (sent by Dean Rhoden and Rhian Thomas on 21/5/2020) for individual sites; the most downstream of these sites within each waterbody was used here. WFD P status within the IMP are calculated using these thresholds, to be compared with the NRW statuses. Not all catchments had available data to construct these thresholds. It should be noted that WFD status is generally calculated by NRW from assessment at monitored points, whereas for IMP outputs we have calculated based on the WFD catchment accumulated concentration.

WFD P Status:

These data are the cycle 2 2018 interim classification, which are subject to similar issues to the NRW P data for the 2017 interim classification, namely: Phosphorus data quality problems in 2015-16, which meant some P data from this period were excluded for the 2018 classification. This led to instances where the previous classification result is retained in the 2018 dataset. Consequently, much of this data reflects the 2015 classification, which would have used data collected between 2011-2014. This problem applies to about three-quarters of river water bodies in the 2018 classification. It does not affect water bodies where we used the more sensitive P method (mainly water bodies with lower nutrient concentrations). Not all catchments had an NRW assessed status available, due to insufficient monitoring points. The data are available online from NRW.[7]

Many catchments were incorrectly predicted by >1 class (98 were >1 class worse, and 18 were >1 class better), and this may reflect the difference in methodology for assigning status, as much as disagreements in P concentration. The discrepancy will affect valuation of change in WFD status; since we tend to predict worse status. We may, for example, model change from poor to moderate, which would be measured as a change from moderate to good. This matters as changes between different statuses are valued differently within the valuation model (see Metcalf et al., 2012 “NWEBS”, also Section 14.1.1): in general an uplift from moderate to good is the highest value, and bad to poor is the lowest value.

[7] https://waterwatchwales.naturalresourceswales.gov.uk/en/

Additionally, a modelled improvement in status for an IMP scenario, may represent only change in concentration in-stream (i.e. improvement without crossing status thresholds). The spatial pattern of disagreement shown in the maps below broadly follows that seen for P concentration. We should note that worse status is predicted by IMP in upland areas whilst more WFD catchments in the northeast of Wales are predicted to have better status than suggested by the WFD interim classification.

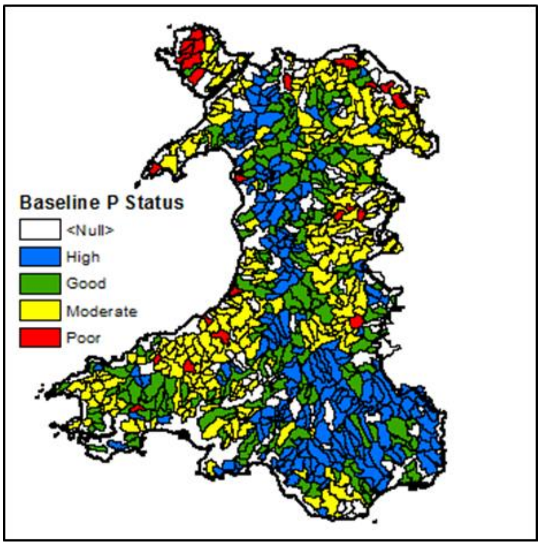


Figure 11.9 Map of baseline WFD P status produced by IMP, using thresholds from data provided by NRW to produce status from OP concentration (note: none of the Baseline model results breached the Poor/Bad threshold, hence, no waterbodies are displayed as 'Bad' here). Contains Natural Resources Wales information © Natural Resources Wales and database right. All rights reserved

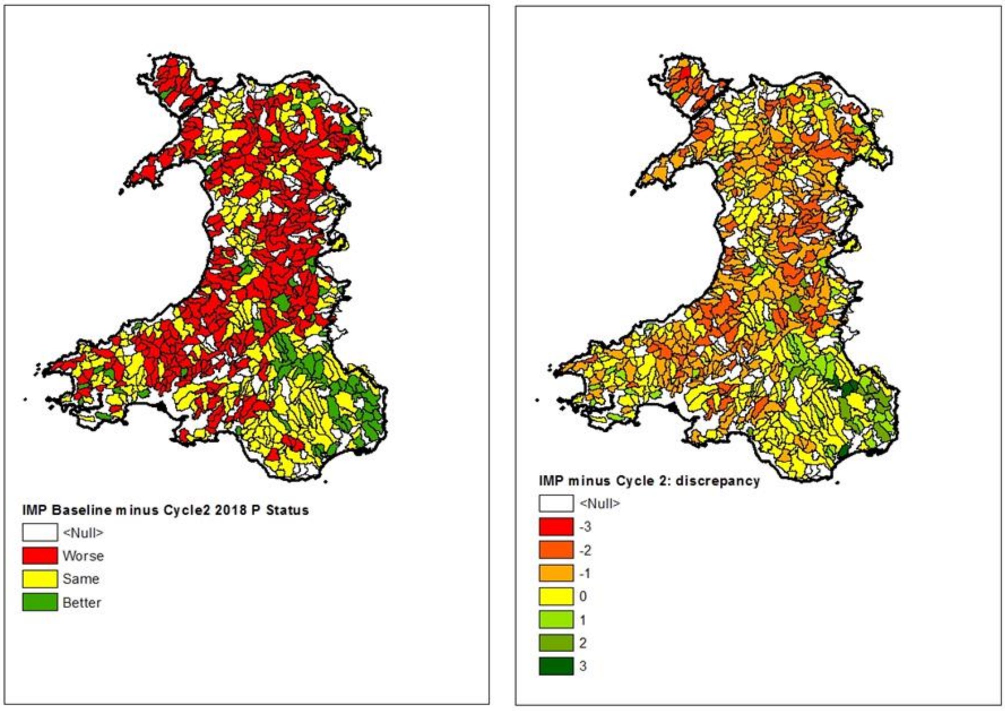


Figure 11.10 Maps of discrepancy in baseline WFD P status produced by IMP, in comparison to NRW assigned P statuses for the 2018 Cycle 2 WFD classification cycle (Note: P statuses are assigned by NRW for 628 of the catchments, hence, some catchments are displayed as "<Null>"). Contains Natural Resources Wales information © Natural Resources Wales and database right. All rights reserved

12 Air Quality & Health - MetaEMEP4UK

Authors: Alice Fitch, Laurence Jones, Janice Scheffler, Edward Carnell and Massimo Vieno

12.1 Introduction to the model QA

The Meta-EMEP4UK calculations assess the changes in air quality as a result of changes in land use affecting woodland area and agricultural ammonia (NH3) emissions higher up the modelling chain. This section introduces these models and carries out the following quality assurance:

- Peer Review: EMEP4UK has been extensively published.
- Validation: Comparison of modelled concentrations from the main EMEP4UK model runs compared with AURN daily air quality monitoring data.
- Validation: QA of Meta-EMEP4UK and comparison of results to previous modelling of air pollution removal by trees for Wales.

12.2 Introduction to the modelling

Changes in air quality as a result of land use management and land use change were calculated using the meta-model Meta-EMEP4UK. This predicts the change in PM2.5 concentration at a grid cell level (approx. 5 x 5 km), resulting from changes in land use in a given scenario compared with the baseline. Inputs required for this calculation are the change in NH3 emissions, current PM2.5 levels, and the area of new (or felled) woodland within a 40 x 40 km grid (here interpreted as a 9x9 cell grid) calculated as a proportion of the total 9x9 cell grid.

To create the meta-model a series of model runs were made using the atmospheric chemistry transport model EMEP4UK (Vieno et al., 2016). This involved running two scenarios: a bespoke land use change and emissions change scenario (SCENARIO) was constructed which incorporated different combinations of the full range of variation in ammonia emissions and woodland planting likely to occur under any of the land use and management scenarios envisaged under policy change. The scenario was constructed using the entire UK to provide greater opportunity to incorporate variation in background PM2.5 concentrations and other atmospheric chemistry and meteorological variables required for EMEP4UK, as well as random and independent variation in ammonia emissions and woodland area within pre-specified ranges at the required scale of 40 x 40 km. A baseline scenario (BASELINE) with the current pattern of ammonia emissions, woodland and other land covers was also run. Both scenarios were run with 2015 emissions and 2015 meteorology. EMEP4UK uses WRF version 3.7.1 as its meteorological input, using hourly 3D meteorological data.

The parameters for the meta-model were calculated by comparing SCENARIO and BASELINE runs in EMEP4UK. The statistical meta-model calculates the change in PM2.5 concentrations as a function of change in ammonia emissions, change in woodland cover and background PM2.5 concentrations. The model structure was constrained by adjusting the intercept in order to ensure that modelled PM2.5 concentrations did not change if there was no change in woodland and no change in ammonia concentrations.

The resulting meta-model equation (adjusted R2 = 40.7%) is:

Meta-EMEP4UK = [-0.20409]+(-0.18950*(Change in frac of woodland within 9x9 cell window * baseline PM2.5)) + (0.000003*Change in NH3 emissions)

Meta-EMEP4UK directly calculates a change in pollutant (PM2.5) concentrations as a result of changes in the input parameters. The change in PM2.5 concentration is population-weighted to give an estimate of the change in exposure of the population to air pollutants known to be damaging to human health. The change in exposure is converted to health impact metrics using response functions derived from COMEAP (Committee on the Medical Effect of Air Pollutants) 2010 and Atkinson et al. (2014); compiled by independent experts for governmental use on the impact of PM2.5 on respiratory hospital admissions, cardiovascular hospital admissions, Loss of Life Years, and the health costs associated with these. Population data used within these calculations are from the UK Office of National Statistics 2011 census. Health impacts are calculated as a proportional change in health outcome based on existing mortality and morbidity data by local authority. From these health impacts, an economic value is estimated (see Jones et al., 2019 for a full description).

To run Meta-EMEP4UK, the aggregated change in ammonia emissions and woodland proportion were calculated at a 40 x 40 km grid as inputs. All calculations on changes in PM2.5 concentration were output at this grid resolution, and then combined with population data for estimation of population-weighted change in exposure at Local Authority level. Subsequent calculation of health outcomes and economic value were conducted for each Local Authority.

12.2.1 Peer review: EMEP4UK

EMEP4UK is an established atmospheric chemistry transport model used for a range of air quality modelling applications. The original EMEP model is described in Simpson et al. (2012). The enhanced UK parameterisation in EMEP4UK is described in Vieno et al. (2016).

12.2.2 Validation: Comparison of modelled concentrations from the main EMEP4UK model runs compared with AURN daily air quality monitoring data

In this section we describe the validation of the main EMEP4UK model runs which were used to create the meta-model for this project (Meta-EMEP4UK). Validation is reported for the baseline run of EMEP4UK which used 2015 UK vegetation and emissions.

Daily air pollution concentrations for O3, NO2, PM10 and PM2.5 from the BASELINE run were evaluated against the 2015 AURN daily air quality monitoring network data (full report available on request). An example for Cardiff Centre is shown in Figure 12.1, which shows good agreement of the EMEP4UK outputs compared with measured data at the Cardiff monitoring site.

For each EMEP4UK model run, QA checks of emissions, surface concentration budgets, surface concentration plots, and wet and dry deposition were undertaken and compiled into a QAQC document. Also stated are the difference in configuration files of the BASELINE and SCENARIO runs, and whether any warnings were raised. The concentration plots (e.g. Figure 12.2) allow sense-checking of the concentration levels and the spatial pattern of changes under each scenario which, combined with the validation against individual air quality monitoring locations described above, allows a more detailed QA assessment of the outputs. Selected EMEP4UK model outputs and input data for the meta-model are shown in

Figure 12.2. BASELINE ammonia and PM2.5 concentrations closely match other modelled interpolated concentration fields for these pollutants for 2015.

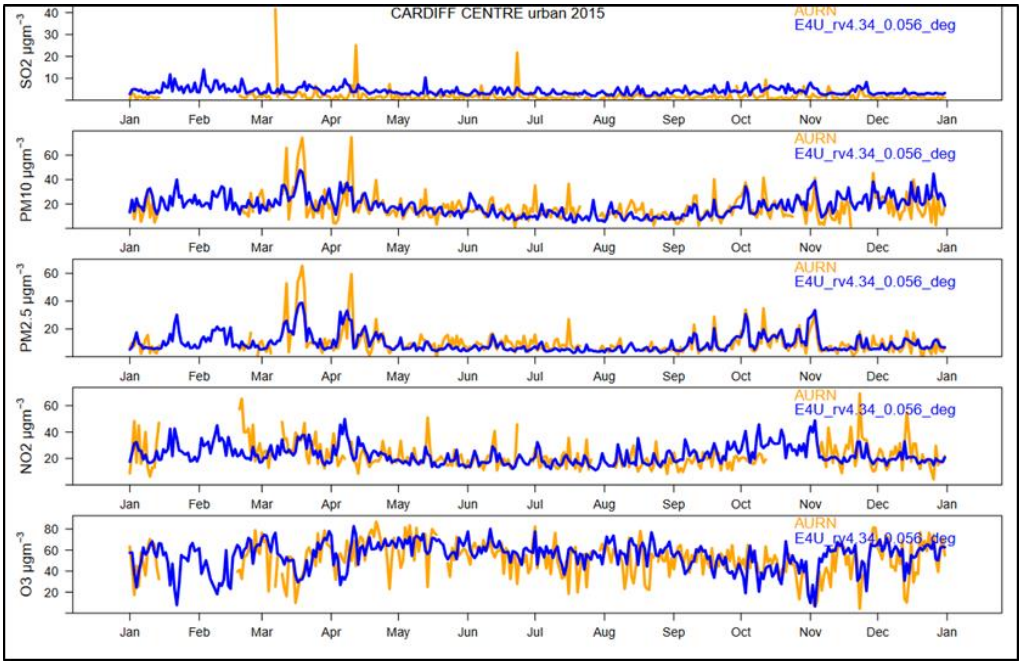


Figure 12.1 Modelled concentrations from the BASELINE run of EMEP4UK (blue lines EMEP4UK) compared with AURN daily monitoring data for Cardiff Centre (orange lines).

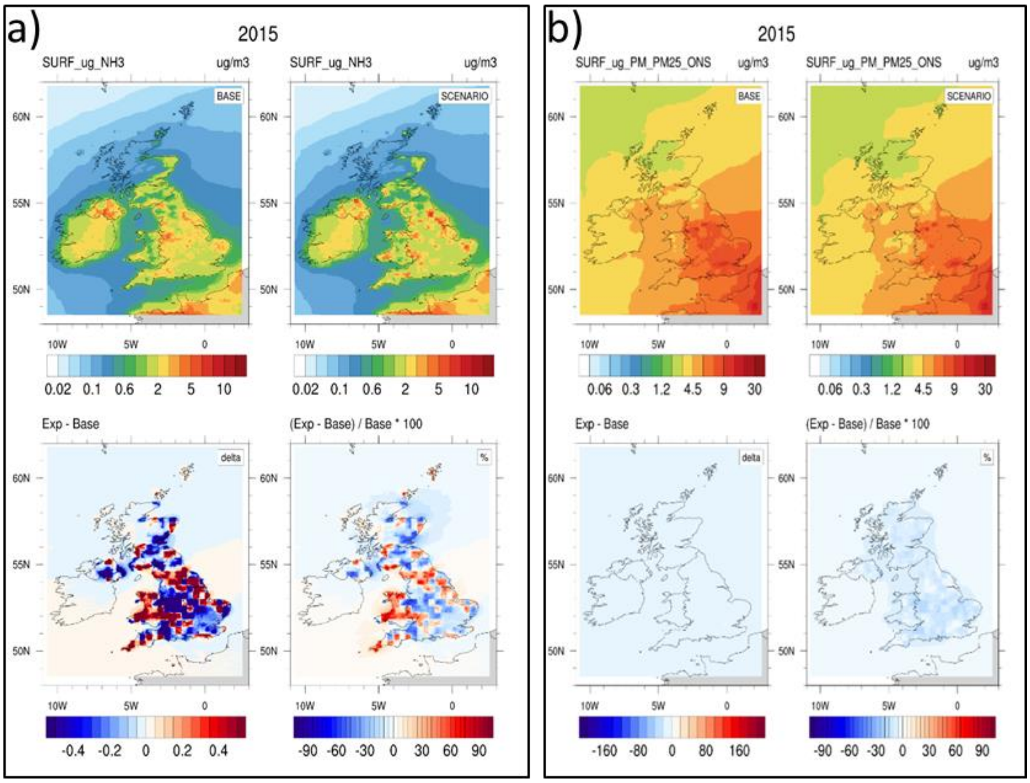


Figure 12.2 Composite panels show a) ammonia concentrations in the EMEP4UK BASELINE and SCENARIO runs, and the absolute and percentage difference between the two, b) PM2.5 concentrations in the EMEP4UK BASELINE and SCENARIO runs, and the absolute and percentage difference between the two.

Separately, the model outputs can be compared against expected findings from the literature, for example in Figure 12.3. This shows that the decrease in PM2.5 concentrations is greatest at high initial PM2.5 concentrations and where there is the largest increase in new woodland. Changes in PM2.5 due to increases or decreases in ammonia emissions are proportional to the change in ammonia and are largely independent of change in woodland area.

All model outputs are compiled into netCDF with model version, and emission year included in the file name. Comprehensive metadata on model run and outputs is included within each file as well as in the file naming conventions[8]. This allows easy tracking of which model versions and data sources were used in individual runs, facilitating rapid Quality Assurance if results need to be checked.

12.2.3 Validation: QA of Meta-EMEP4UK and comparison of results to previous modelling of air pollution removal by trees for Wales

In this section we describe the QA processes for the meta-model Meta-EMEP4UK, and describe comparison of outputs with previous modelling runs for Wales Natural Capital Accounts.

When running Meta-EMEP4UK, flags are in place to raise warnings if input data are in a different format or the wrong units, and intermediate data produced during the process is outputted and checked against provided data, e.g. change in NH3. The code to process data into the input format needed for Meta-EMEP4UK and the code to calculate population-weighted pollution concentrations and aggregate Meta-EMEP4UK output have all been independently checked for script errors.

Health outcome and economic estimates from Meta-EMEP4UK were compared with previous assessments using different model versions of EMEP4UK to calculate health benefits of pollution removal by natural vegetation for the UK Natural Capital Account (Jones et al., 2017) and for the Wales Natural Capital Accounts (Engledew et al., 2019) . The estimated pollution removal per unit area of woodland was in very close agreement with these previous calculations.

After QA checks conducted on each analysis, the outputs from Meta-EMEP4UK are passed onto environmental economists at eftec for subsequent economic analysis. eftec have worked with air pollution health associated costs previously, so this ensures an additional quality check on output.

Example results are shown below for different scenarios. These show: calculated change in PM2.5 concentration (Figure 12.4), change in health outcomes at local authority level (Figure 12.5), and the same results in table format showing the numerical estimates (Table 12.1).

---

[8] An example file name is below, the sections in bold provide run information for version tracking: EMEP4UK_emep-ctm-**rv4.34_wrf3.7.1**_ERAMMP_{runtype}_trend**2015_**emiss**2015**_UK_**2015_**fullrun.nc .

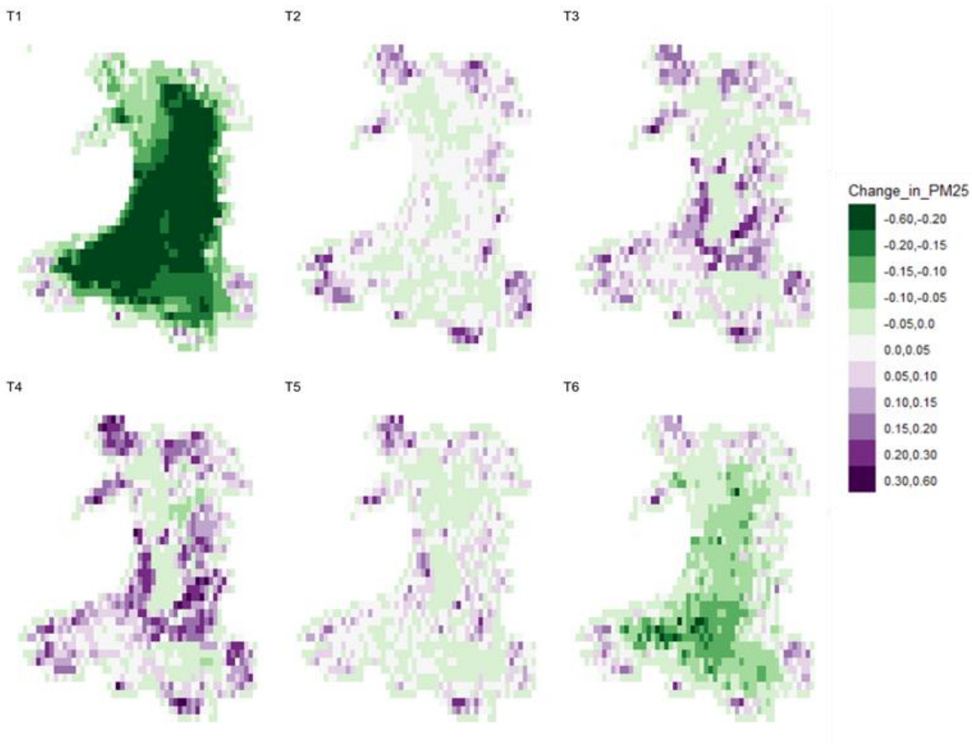


Figure 12.4 Change in PM2.5 concentrations (µg m-3) at 5x5 km grid cell resolution. Purple shows increase and green shows a reduction in PM2.5 concentrations.

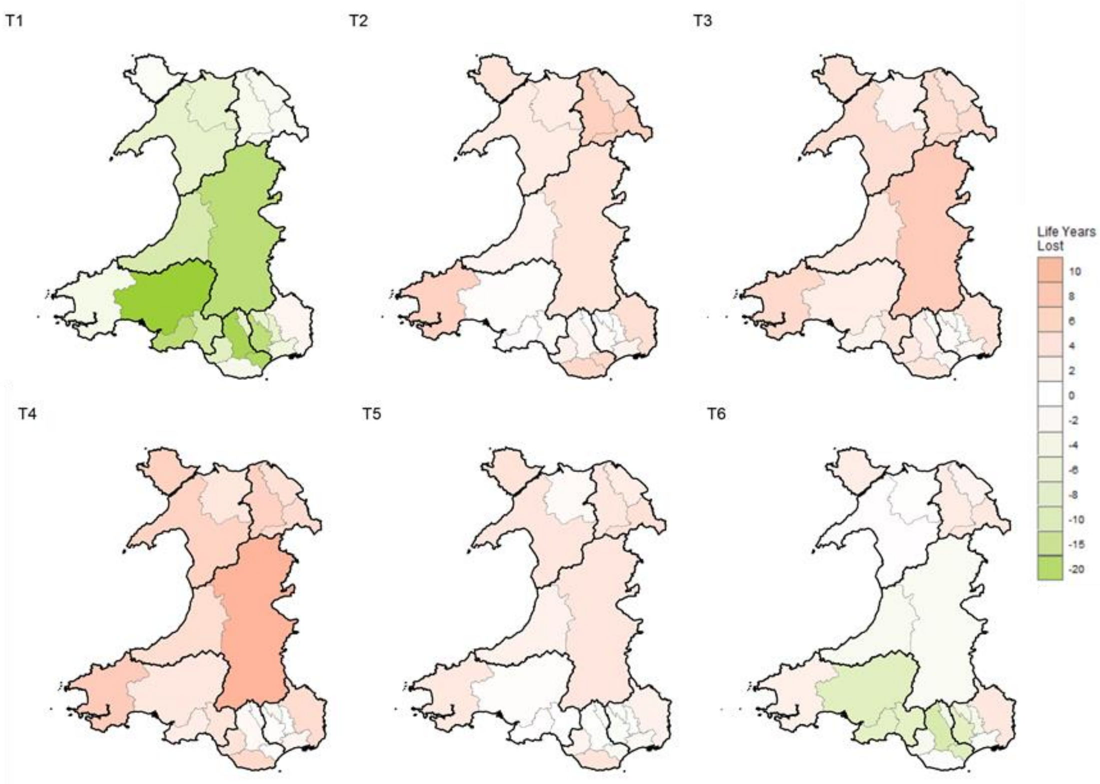


Figure 12.5 Change in Life Years Lost, at Local Authority level. Green shows a reduction, pink shows an increase.

Table 12 1 Health costs (in £2012 prices) associated with change in PM2.5 concentrations.

Code	Unitary authority	T1	T2	T3	T4	T5	T6
W06000001	Isle of Anglesey	£43,384	-£135,757	-£145,424	-£224,371	-£135,672	-£89,270
W06000002	Gwynedd	£242,157	-£96,829	-£157,690	-£216,809	-£117,836	-£8,886
W06000003	Conwy	£230,839	-£95,630	-£54,895	-£122,241	-£23,917	£16,039
W06000004	Denbighshire	£81,580	-£231,088	-£165,332	-£222,225	-£114,499	-£95,591
W06000005	Flintshire	£70,504	-£156,023	-£128,803	-£149,544	-£90,563	-£55,635
W06000006	Wrexham	£54,444	-£225,809	-£172,428	-£175,671	-£139,647	-£108,973
W06000008	Ceredigion	£393,621	-£53,092	-£112,023	-£165,235	-£64,873	£67,859
W06000009	Pembrokeshire	£105,141	-£222,964	-£183,551	-£252,311	-£110,601	-£74,844
W06000010	Carmarthenshire	£935,352	-£12,053	-£89,425	-£114,678	-£9,344	£305,201
W06000011	Swansea	£660,768	-£3,509	-£62,456	-£87,904	-£2,786	£278,125
W06000012	Neath Port Talbot	£450,405	-£17,126	-£120,290	-£132,009	-£2,108	£270,464
W06000013	Bridgend	£303,746	-£68,778	-£103,082	-£136,778	-£58,940	£79,035
W06000014	Vale of Glamorgan	£102,737	-£191,345	-£121,655	-£182,567	-£129,443	£17,935
W06000015	Cardiff	£753,951	-£182,918	-£25,396	-£12,975	£59,217	£316,159
W06000016	Rhondda Cynon Taf	£817,432	-£38,856	-£65,361	-£35,941	£8,602	£389,855
W06000018	Caerphilly	£639,883	-£13,209	-£28,454	-£16,510	£39,983	£287,934
W06000019	Blaenau Gwent	£233,156	£8,287	£58,951	£70,819	£44,858	£115,702
W06000020	Torfaen	£236,405	-£9,292	£9,610	£6,060	-£3,490	£96,004
W06000021	Monmouthshire	-£49,523	-£153,293	-£137,691	-£171,530	-£64,452	-£113,917
W06000022	Newport	£159,979	-£61,422	-£29,339	-£51,218	-£32,152	£45,077
W06000023	Powys	£656,544	-£136,706	-£252,906	-£370,589	-£117,890	£55,263
W06000024	Merthyr Tydfil	£246,054	-£53	£20,167	£19,934	£29,719	£117,887

13 Carbon modelling - LULUCF methodologies

Authors: Robert Matthews, Kate Beauchamp and Amy Thomas

13.1 Introduction to the model QA

The carbon ecosystem services models provide outputs on LULUCF carbon stock and change, using the relationships between land use and soil types as well as outputs from ESC and CARBINE (Section 4). Emissions of GHG from peat (or wetlands) are also calculated, using predicted land cover. This section introduces the modelling and the quality assurance steps taken, which are, in brief:

- Building Understanding: Presenting the LULUCF stocks at baseline.
- Validation: Comparison of modelled agricultural carbon stock with published data (section 13.2.2).
- Validation: Comparison of baseline wetland emissions outputs with published data (section 13.2.3).
- Expert Assessment (Consortium): Understanding woodland carbon change in an example land use scenario (13.2.4).
- Expert Assessment (Consortium): Understanding wetland GHG emissions in an example land use scenario (13.2.5).
- Expert Assessment (Consortium): Understanding agricultural land carbon change in an example land use scenario (13.2.6).

13.2 Introduction to the model

For agricultural land, carbon stocks in soils and biomass are calculated using LULUCF coefficients for Wales. For soils, these represent soil carbon in the top 1m according to relationships between land use and soil types (organic, organomineral, mineral, other). For biomass, coefficients vary with land use but not soil type.

For scenarios of land use change, time series data are required for valuation purposes since carbon price varies over time. Annual changes in carbon stock in agricultural systems were accounted for using LULUCF methods, assuming a non-linear rate of change and that some transitions occur more slowly than others. For example, for conversion of grassland to arable, losses of carbon stock are initially high and decrease exponentially over time.

The equation for annual change is ft = k(Cf - C0)e-kt

where t = time; k = time constant of change; Cf = assumed equilibrium carbon density for new land use; C0 = assumed equilibrium carbon density for baseline land use.

Example rates of change in soil carbon stock are shown in Figure 13.1.

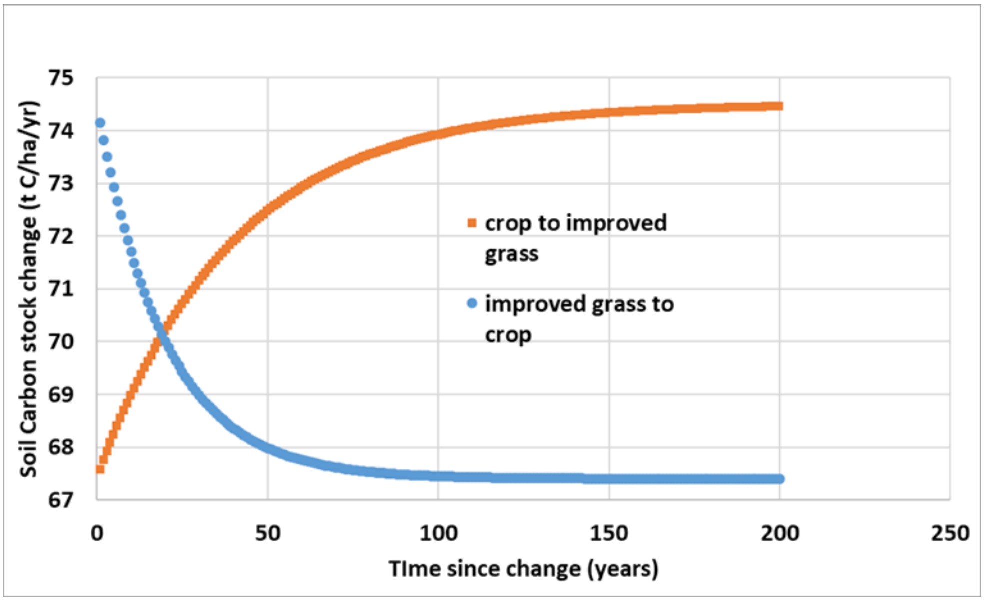


Figure 13.1 Example rates of soil carbon stock change over time for different land use transitions.

Carbon stock and change are calculated at the spatial resolution of the DMU. Note that because each DMU is a composite of land use, the method represents change between composites.

Vegetation biomass change is assumed to occur in year one. The rate of woodland carbon change is valued using data from the ESC and CARBINE models described in Section 4, applying rates averaged over the three time periods for simplicity. Rotational grassland/arable is assigned the same soil carbon stock as arable, due to assumed frequent soil disturbance. Note that the magnitude of change is small relative to total stock in vegetation and the top 1m of soil.

For peat, GHG emissions are calculated using an approach aligned with planned future inventory methods. Coefficients are derived from the draft wetland supplement (Evans et al., 2017) to align with LULUCF inventory methods and are therefore not directly based on modelled nutrient inputs (which will affect N2O emissions from peat soils). These coefficients are used to model baseline and scenario emissions based on the simulated land use from SFARMOD and the LAM (see section 10). For the land use scenarios, we assume emissions from land reach an equilibrium immediately following instantaneous land use change in 2020, whereas in reality it may take around 30 years for vegetation assemblages to converge with reference states. Conversely, the water table may recover relatively quickly, and a large proportion of emissions from improved land reflect N2O from nutrient inputs, which should respond much more rapidly. For scenarios of woodland creation, it is assumed that new woodland cannot be planted on peat, and the peat portion of any field that is replanted to woodland will revert to short vegetation.

Agricultural GHG emissions are calculated at the DMU level by combining each of the SFARMOD loading outputs (for fertiliser input, livestock excreta and land use areas) with the relevant Farmscoper coefficient, accounting for the climate zone, soil type and farm type. Changes are assumed to take place immediately.

13.2.1 Building Understanding: Presenting the LULUCF stocks at baseline

Baseline agricultural carbon stocks in soils and vegetation are predicted at the DMU resolution, using LULUCF coefficients as outlined in Section 13.2. The data are aggregated to NRW regions and small agricultural area for mapping as tonnes per hectare of modelled land (see Figure 13.2). Totals for Wales are shown in Table 13.1; these are compared to other available data in section 13.2.2.

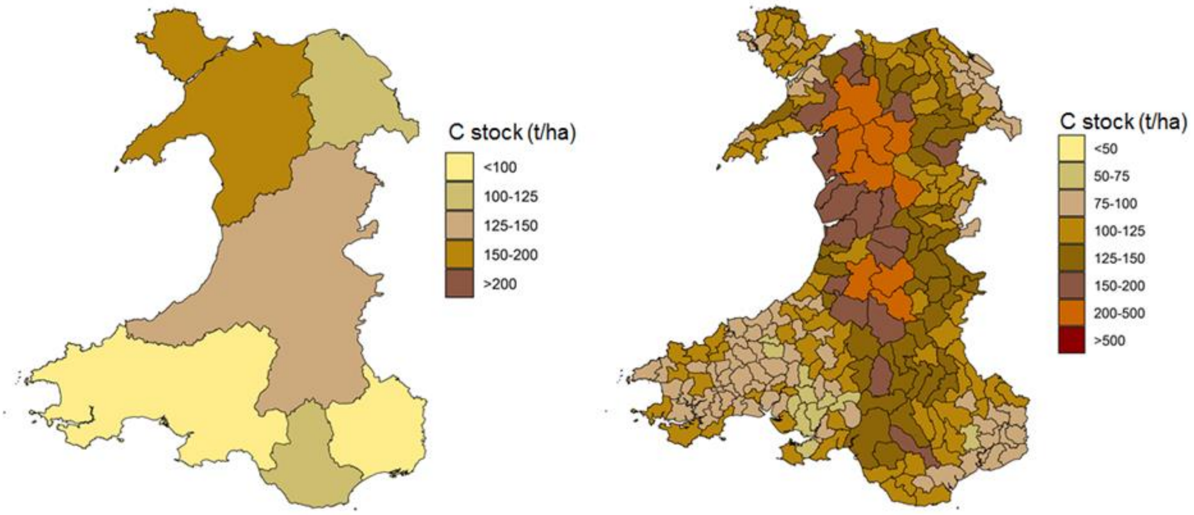


Figure 13.2 Baseline LULUCF carbon stock, mapped by NRW region (left) and small agricultural area (right).

Table 13.1 IMP modelled totals of agricultural stocks of carbon in soils and biomass.

Indicator	Total for Wales
Total soil C (kt)	170,537
Total biomass C (kt)	2,862
Total area (ha)	937,522
Average soil C density (t/ha)	182

13.2.2 Validation: Comparison of modelled agricultural carbon stock with published, inventory aligned data

LULUCF soil stock data and soils and carbon from the LULUCF inventory have been used to QA the modelled carbon stock outputs. These provide national level totals with habitat type breakdown and are derived from measured data published in Bradley and Milne (2005). The LULUCF inventory coefficients derived from this work are the same as those applied in the IMP modelling, therefore this step tests the implementation of the coefficients. Table 13.2 and Table 13.3 show discrepancy for totals and per hectare values.

Table 13.2 Difference in Wales soil carbon stock 0-100 (Kt C) (shown as LULUCF minus IMP).

	Arable	Pasture	Natural	Not modelled	Total	Total modelled
Mineral	17,525.8	-56,678.0	-7,612.4	-22,484.4	-69,249.1	-46,764.6
Organic	484.3	-3,690.6	-25,365.4	-11,719.7	-40,291.5	-28,571.8
Organo-mineral	1,268.2	-8,672.6	-11,355.5	-10,490.0	-29,249.9	-18,759.9
Other	-838.2	-16,081.7	-9,421.8	-3,976.3	-30,318.1	-26,341.7
ZERO	0.0	0.0	0.0	0.0	0.0	0.0
ALL	18,440.1	-85,122.9	-53,755.2	-48,670.5	-169,108.5	-120,438.0

Table 13.3 Difference in land use area values for Wales (km2) (shown as LULUCF minus IMP).

	Arable	Pasture	Natural	Not modelled	Total	Total modelled
Mineral	1,439.7	-3,496.7	-430.6	-1,724.1	-4,211.7	-2,487.6
Organic	5.1	-45.9	-283.6	-105.8	-430.3	-324.5
Organo-mineral	79.4	-543.4	-606.9	-576.3	-1,647.2	-1,070.9
Other	-167.1	-2,816.4	-1,115.7	-694.9	-4,794.1	-4,099.2
ZERO	-16.5	-104.1	-138.9	-685.7	-945.1	-259.4
ALL	1,340.7	-7,006.5	-2,575.7	-3,786.9	-12,028.4	-8,241.5

The main discrepancies between the datasets can be summarised as follows:

- There is an over-estimation of arable stocks and area, and under-estimation for pasture. This may be partly due to arable and rotational grass being combined in the IMP modelling, with possible contributions from increases in arable area since 2005. There will also be a significant area of grassland in Wales that is not part of a farm modelled by IMP (as also seen in Table 9-5).
- There is an under-estimation of the area of natural land (and associated carbon), which largely reflects the fact that much of the land in the “natural” category is not on farms modelled by the IMP.

The tables below indicate that the soil carbon discrepancies (Table 13.4) are in line with the coefficients used (Table 13.5), with minor (<0.5 kt/km2) disagreements reflecting rounding errors. Therefore, the disagreement between IMP outputs and the inventory data reflects differences in area assigned to each land use, and there are no identified issues with the implementation of coefficients.

Table 13.4 IMP coefficients (as kt/km2) from LULUCF.

	Arable	Natural	Pasture
Mineral	12.2	17.6	16.2
Organic	95.1	89.6	80.7
Organo-mineral	15.8	18.7	16.1
Other	4.9	8.4	5.7

Table 13.1 Density of discrepancy as (kt/km^2).

	Arable	Natural	Pasture
Mineral	12.2	17.7	16.2
Organic	95.4	89.4	80.3
Organo-mineral	16.0	18.7	16.0
Other	5.0	8.4	5.7

13.2.3 Validation: Comparison of baseline Wetland GHG emissions with wetland coefficients from published, inventory aligned data

To QA the wetland emissions outputs, the baseline outputs were checked against the wetland coefficients being applied in the IMP modelling, thereby testing the implementation of the coefficients. Wetland coefficients were taken from the emissions inventory wetland supplement (Evans et al., 2017, Table 4.1). Figure 13.3 indicates a consistent relationship between area and modelled emissions, and Table 13.6 shows that the mean of this relationship matches the coefficient for the relevant land use type, indicating that the coefficients have been implemented correctly.

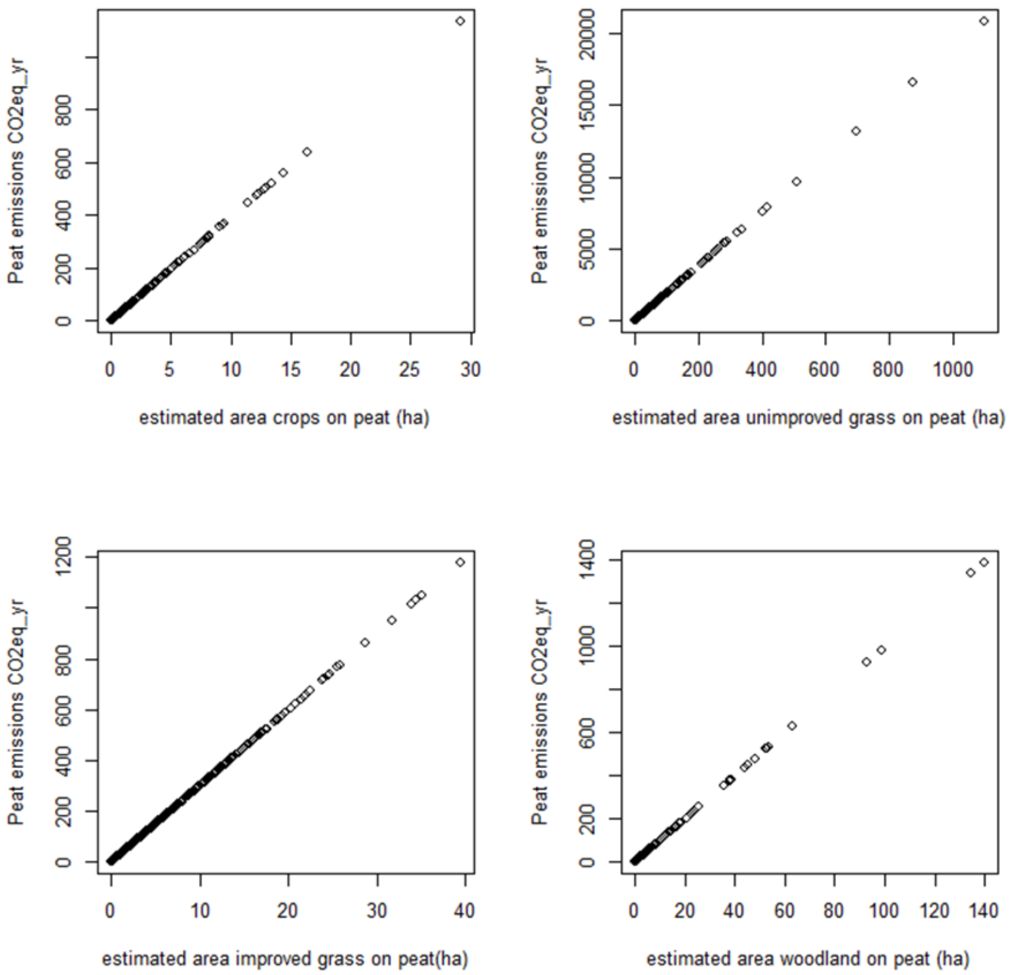


Figure 13.3 Baseline modelled wetland GHG emissions as tCO2 equivalents, plotted at the DMU level, as 4 scatterplots split by land use type, plotted against the area (ha) of peat modelled on that land use.

Table 13.6 Comparison of baseline modelled wetland GHG emissions per hectare with the wetland coefficients implemented within the IMP to verify the coefficients were correctly used.

Land use	Mean of modelled baseline emissions per estimated ha of land use on peat	Coefficients implemented in the IMP
Cropland	38.98	38.98
Unimproved grass	19.02	19.02
Improved grass	29.89	29.89
Woodland	9.91	9.91

13.2.4 Expert Assessment (Consortium): Understanding woodland carbon change in an example land use scenario

Modelled carbon sequestration in woodland was assessed for an example scenario to sense-check the data. The CARBINE and ESC data have been re-aggregated to the DMU level as described in section 4.2.8, then adjusted in post-processing to better represent the time periods. These were then incorporated within the LULUCF carbon modelling. It was therefore important to check that the data outputs being used in the scenarios were sensible, in case of errors in processing and implementation. Woodland carbon sequestration rates vary with existing land cover, soils, climate and other site factors; hence, the figures displayed are specific to the spatial pattern of woodland creation/regeneration simulated for the example scenario (T1).

The rates of carbon change over time in the data shown in Table 13.7 and Figure 13.4 follow the expected patterns. For the first 5 years, managed systems (native broadleaf and productive conifer) may be expected to be net emitters (hence positive values for 2020-2025 in Table 13.7), due to disturbance of soils, loss of baseline vegetation and slow vegetation growth. From 2025-2050, net sequestration is modelled as the system becomes established. This sequestration rate slows from 2050-2100.

Natural revegetation to mixed forest was predicted to sequester carbon in soils during 2020-2025 and 2025-2050, where it takes place on arable land (but not on grassland). From 2050-2100, sequestration is modelled as the system becomes established; it is expected that there is a delay in the timing of this sequestration relative to managed woodland.

Table 13.7 Change in carbon for each woodland type represented in the T1 scenario, provided as average netLULUCF_CO2eq/ha/yr, accounting for change in soils and vegetation, and in the case of managed woodland, harvested wood products and GHG emissions from management.

	2020-2025	2025-2050	2050-2100
Natural revegetation to mixed forest	-0.05	-0.05	-5.44
Native broadleaf	3.82	-7.24	-4.40
Productive conifer	3.45	-17.39	-3.48

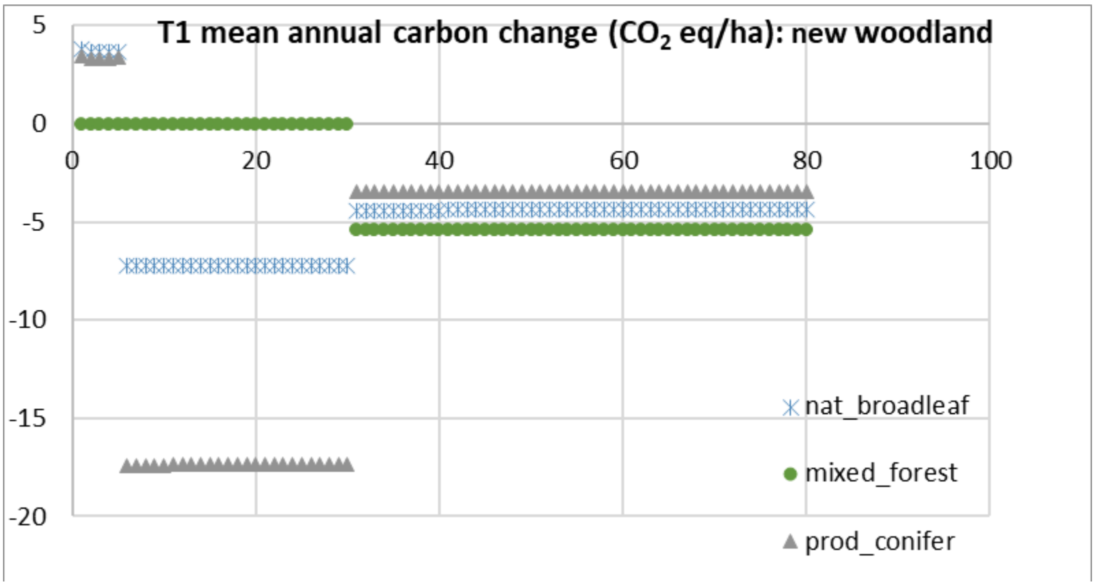
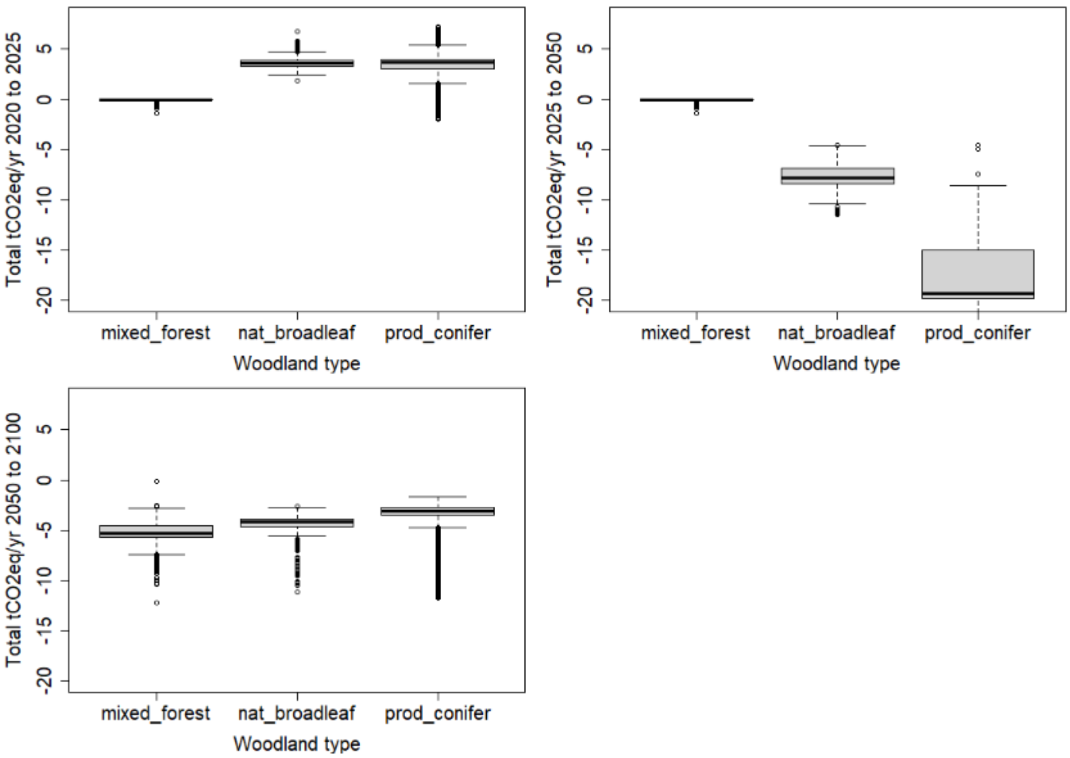


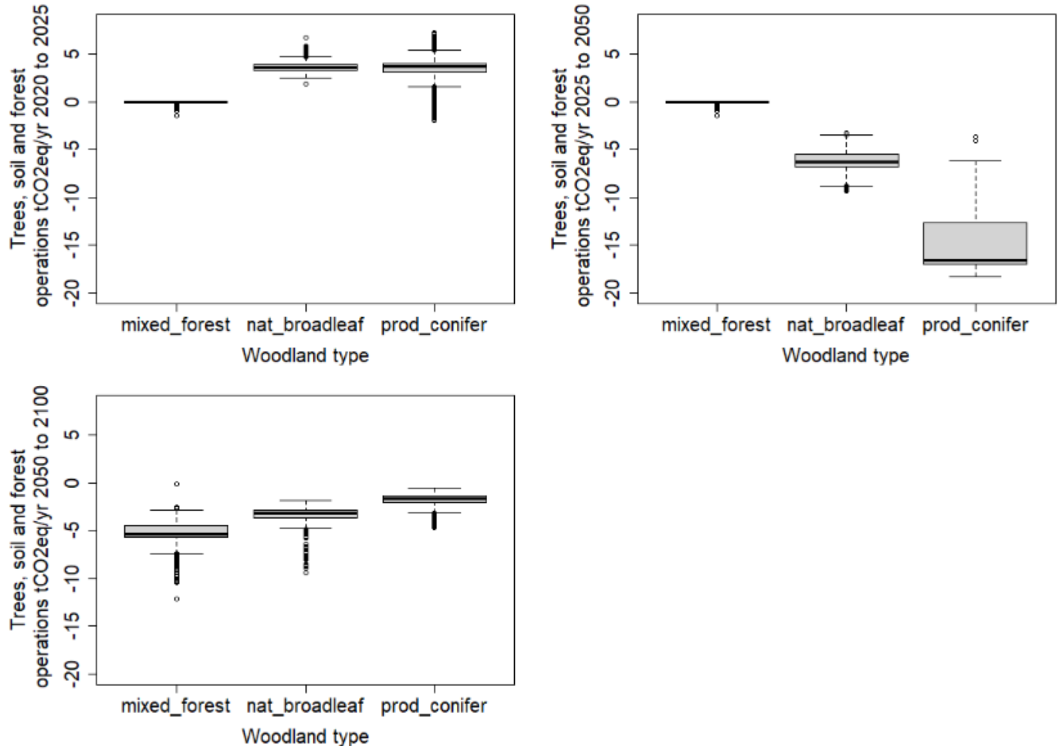
Figure 13.4 The modelled carbon stock change as tCO2 equivalents per hectare, plotted as the annual mean for each modelled woodland type (data from table 13.7).

Figure 13.5 shows the range of modelled values within each woodland type and highlights the contribution from harvested wood products (HWP). The range reflects differences in soil type, baseline land cover, climate and other parameters, which affect the CARBINE and ESC predictions of carbon sequestration rates. Overall, the greatest magnitude and range in sequestration rates was modelled from 2025 to 2050, and these are generally greatest for areas planted to conifer, although the range overlaps with the range for broadleaf. The breakdown to separate HWP indicates that they are only responsible for a relatively small proportion of modelled sequestration, with the exception of outliers for conifers in the 2050-2100 time period.

(i)



(ii)



(iii)

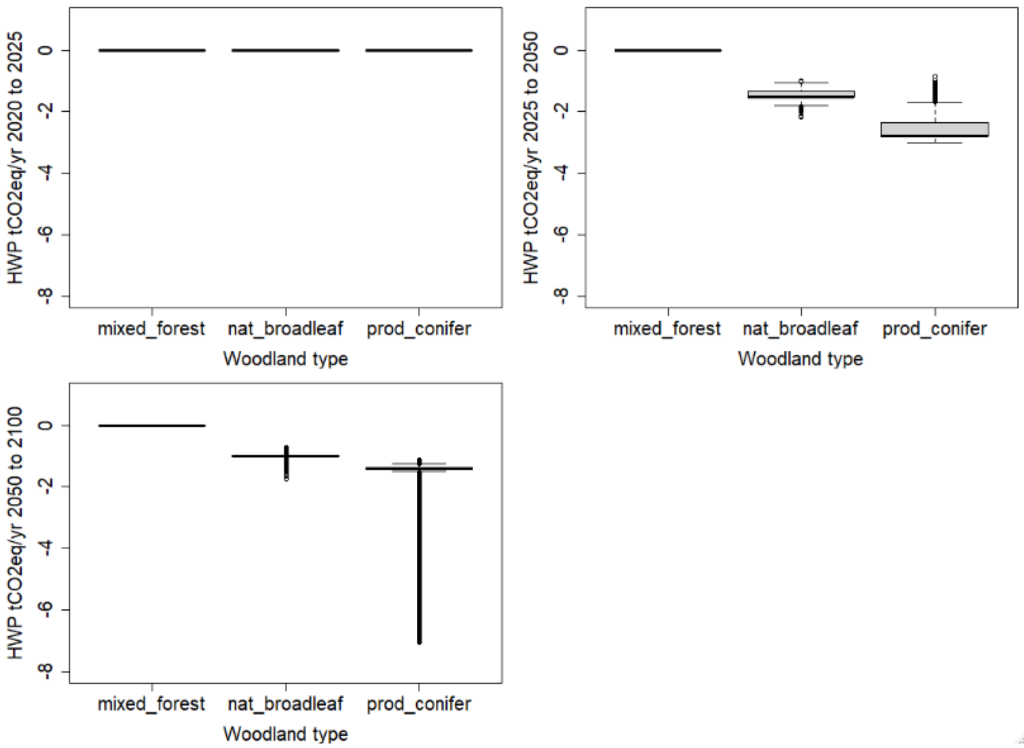


Figure 13.5 Box and whisker plots of T1 modelled carbon stock change as tCO2 equivalents per hectare per year, for each modelled woodland type. Plots are shown for each of our three modelled time periods, and display data for (i) total net change (to explore the range of data shown in Table 12.7 and Figure 12.3) and then a breakdown to show contribution from (ii) trees, soil and forest operations and (iii) harvested wood products (HWP). Boxplots show the median value (dark horizontal line) relative to the interquartile range (between first and third quartiles, shown as shaded area), range of data (dashed vertical line with horizontal lines for highest and lowest values) with outliers also shown as dots.

13.2.5 Expert Assessment (Consortium): Understanding wetland GHG emissions in an example land use scenario

The modelled change in wetland GHG emissions were assessed for an example scenario (T1) to sense-check the implementation of the coefficients. Table 13.8 shows that the emissions per hectare of peat match the coefficient for the relevant land use type, indicating that the coefficients have been implemented correctly.

Table 13.8 Comparison of T1 modelled wetland GHG emissions per hectare with the wetland coefficients implemented within the IMP.

Land use	Mean of modelled T1 emissions per estimated ha of land use on peat	Coefficients implemented within the IMP
Cropland	38.98	38.98
Unimproved grass	19.02	19.02
Improved grass	29.89	29.89
Woodland	9.91	9.91

13.2.6 Expert Assessment (Consortium): Understanding agricultural land carbon change in an example land use scenario

The modelled carbon sequestration in agricultural land was assessed for an example scenario (T1) to check the implementation of the LULUCF methodology for change in carbon stocks over time. Annual average sequestration rates are shown in Figure 13-6, with negative values indicating sequestration (as per LULUCF convention). The plot on the left shows that for the creation of shrub grassland, large sequestration is simulated in the first year. The rate of sequestration is much lower from year 2, and further decreases over the modelled 80 years. This is in line with our expectations: the relatively large changes in biomass carbon are modelled as taking place in the first year (as per LULUCF standard approach for non-woodland), and the average changes in soils were smaller, decreasing over time as expected. The plot on the right shows net emissions for land staying in agriculture in the T1 scenario, which are greatest in year 1, exponentially decreasing over the modelled time period. This is in line with our expectations, since the trend for land staying in agriculture is dominated by transitions from permanent to rotational grassland, which would be expected to result in an exponential rate of soil carbon loss.

Figure 13.7 shows the difference in carbon stock between 2020 and 2100. The data conform to our expectations; in general, more carbon is sequestered for new woodland and is emitted for land remaining in agricultural use. Land becoming shrub grass might be expected to always sequester carbon; however, the data show that sometimes less carbon is sequestered. This reflects the coefficients applied from LULUCF, which indicate the highest carbon stocks for organic soil managed as arable. This is because of spatial patterns in the baseline data that may not always be representative for all transitions, which is a limitation of a space for time approach. The overall pattern is likely to be similar across scenarios, with variation reflecting the different areas undergoing different transitions.

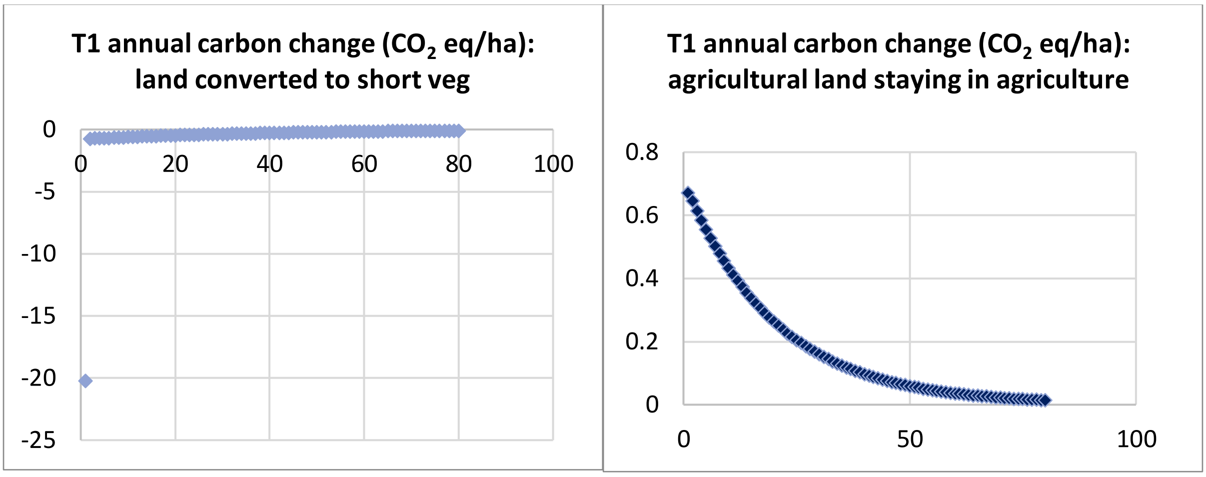


Figure 13.6 Modelled carbon stock change as tCO2 equivalents per hectare, for land converting from agricultural use to short vegetation (left), and land staying in agriculture (right) for an example scenario (T1). Both plots are the mean of all sites in that category.

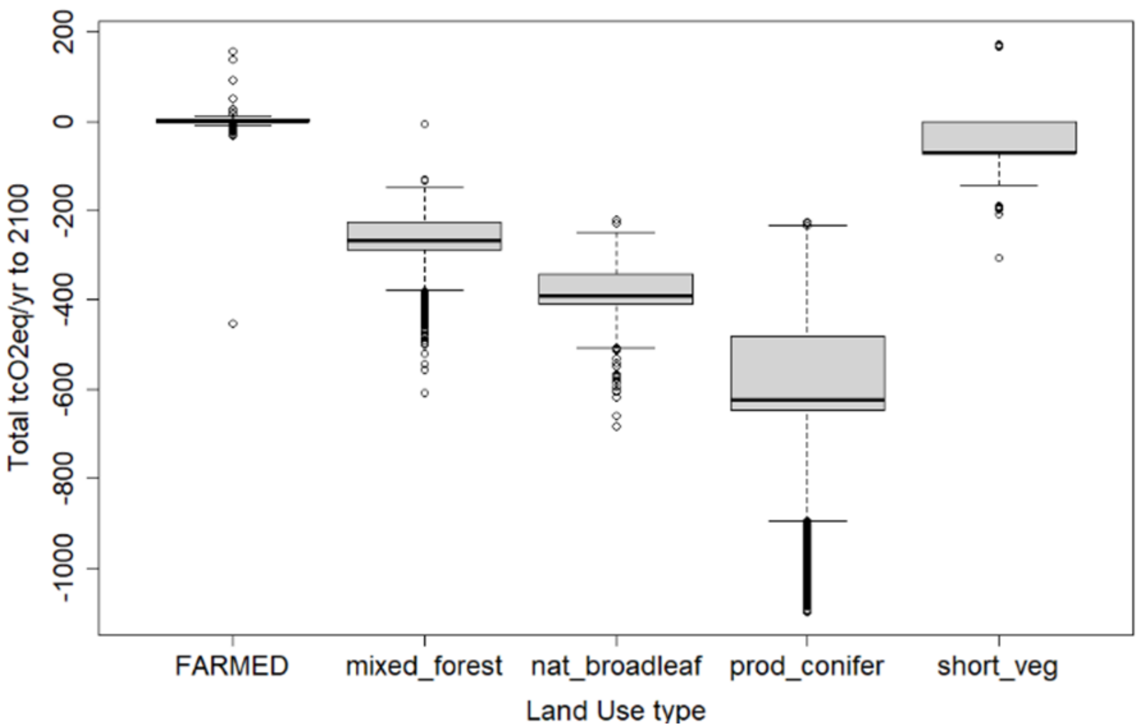


Figure 13.7 Modelled carbon stock change between 2020 and 2100 under an example scenario (T1). Negative values represent net sequestration. The values for mixed_forest, nat_broadleaf and prod_conifer are the same as in Figure 13-5, but are included here for comparison.

Table 13.9 and Figure 13.8 allow comparison of T1 modelled outputs with the LULUCF coefficients. Model outputs can only be split into "farmed" and "short vegetation" (i.e. farmed DMUs cannot be split into the land use types represented by the coefficients). This is because each DMU has a mix of land cover types, which transitions to a new mix of land cover types. Carbon stock for the scenario land cover mix is modelled based on change from the baseline land cover mix at the DMU level. In order to model change without knowing the spatial pattern within the DMU it is necessary to model for the aggregated DMU, across the relevant mix of land use types. Therefore, scenario carbon stock is only known for the aggregated DMU. Table 13.9 shows that the average model outputs are within the expected range for the coefficients. The T1 outputs for "short veg" match the coefficients, with minor (<0.1 kt C /km2) disagreements which may reflect rounding errors or potentially land having not quite reached equilibrium. The T1 outputs for farmed DMUs fall within the range covered by arable, rough grass and pasture. Figure 13-8 shows that the range of output values for the farmed DMUs in the T1 scenario also fall within the coefficient range in Table 13.9 (short vegetation DMUs have consistent carbon stock for each soil type), this indicates appropriate implementation of the coefficients for the scenario.

Table 13.9 LULUCF coefficients for carbon stock in soils and vegetation, compared to the modelled mean stocks for 2100 (these are calculated using LULUCF methods to model change from baseline). Coefficients are split by soil type and land use, whereas the T1 modelled outputs are split by soil type and then into farmed and short vegetation, since each DMU has a mix of land use.

	Coefficients kt C /km^2				**T1 output mean kt C /km^2 2100**	
	Arable	**Rough grass**	**Pasture**	**Short veg**	**Farmed**	**Short veg**
Mineral	12.70	17.91	16.48	18.63	14.33	18.47
Organic	95.60	89.91	80.96	90.63	89.72	90.53
Organo-Mineral	16.30	19.00	16.39	19.72	18.51	19.67
Other	5.40	8.73	5.98	9.45		9.45

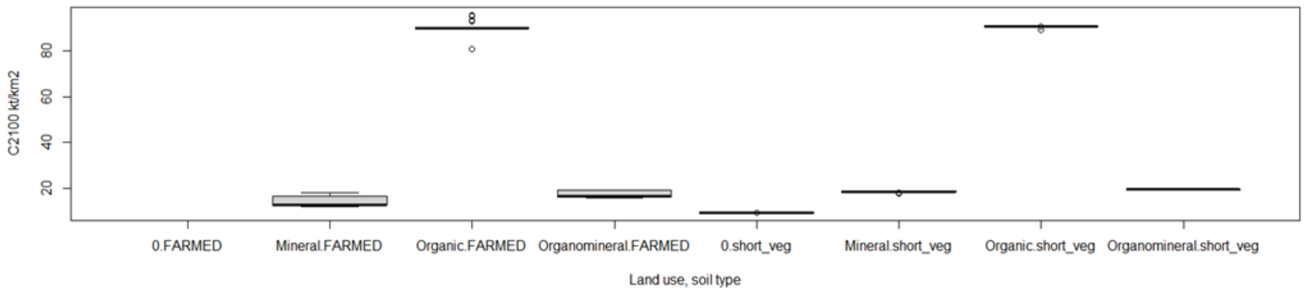


Figure 13.8 The ranges of modelled carbon stock for 2100 by land use (Farmed and Short vegetation) and soil type (other (O), mineral, organic and organomineral) combinations.

14 Valuation

Authors: Ian Dicke, Amy Thomas and Sophie Neupauer

14.1 Introduction to the model QA

Valuation takes the outputs from the ecosystem service models and uses best available information to provide monetary values to go alongside the physical indicators modelled. The QA steps taken are detailed in the section below.

14.1.1 Expert Assessment/ Peer Review: Setting out the approach to Quality Assurance within the valuation modelling using best available valuation evidence

This step of the analysis provides valuation results as part of the IMP outputs. The approach taken is summarised using the best available valuation evidence from UK public sector sources. The quality assurance steps that have been applied are described in the HM Treasury Aqua Book (HM Treasury, 2015), which provides guidance on producing quality analysis for government.

Three ecosystem services, greenhouse gas sequestration, water quality and air quality, are valued based on the outputs from preceding steps in the IMP. A description of the models used to determine the value of each of the ecosystem services, and the assumptions made during each, are outlined below.

Greenhouse gas sequestration:

- The ecosystem services modelling provides peat GHG tCO2e/yr, net LULUCF tCO2e/yr and baseline/scenario non-peat GHG tCO2e/yr for six Welsh regions.
- Each tCO2e is valued based on relevant annual values from the BEIS (2019) valuation of energy use and greenhouse gas emissions for appraisal. The document containing these values (data tables 1 to 19: supporting the toolkit and the guidance) are available at https://www.gov.uk/government/publications/valuation-of-energy-use-and-greenhouse-gas-emissions-for-appraisal.
- The models assume that the rate of tCO2e change for agriculture and wetlands remains constant over the time period. This means the timing of impacts cannot be determined from scenario definitions that are input to the modelling process.

Water quality:

- The ecosystem services modelling provides baseline and scenario status/concentration for N and P in Welsh waterbodies, disaggregated at the waterbody level. The IMP also provides the proportion of each waterbody for six Welsh regions.
- Changes in waterbody status are valued using Metcalfe (2012) and NERA Economic Consulting (2007) “The benefits of Water Framework Directive Programmes of Measures in England and Wales” (“NWEBS”).
- Estimates are based on a sixth of the central NWEBS values, which are the best estimates of willingness-to-pay (WTP) for changes in waterbody status. The analysis

of water quality only looks at one of the six components of waterbody status, therefore taking 1/6th of the value.

- In cases where the waterbody status changes by more than one step in status (e.g. from high to bad), NWEBS values for each individual change are added (e.g. for waterbodies changing from moderate to bad, we add the value of moderate to poor, and poor to bad).
- Changes in water quality status are assumed to be immediate.
- The valuation assumes that the willingness-to-pay for a deterioration in water quality status (i.e. from good to moderate) is the same as the willingness-to-pay for an improvement in water quality status (i.e. moderate to good).

Air quality model:

- The ecosystem services modelling provides the annual volume and monetary value of air pollutant emissions and removals by local authorities in Wales.
- In terms of the timings, a third of the PM2.5 value is due to changes in NH3 emissions - this change is assumed to occur in year 0 and remain constant over the time period. Two thirds of the value is due to tree planting - the full value of tree planting is achieved after 40 years and, based on expert judgement, is assumed to linearly increase from year 0.

All models:

- All monetary valuations are aligned to methods in HM Treasury (2018) The Green Book, which is available at https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-governent. Values were converted to current prices using HM Treasury (2020) GDP deflators at market prices, and money GDP; available at https://www.gov.uk/government/collections/gdp-deflators-at-market-prices-and-money-gdp.

Table 14.1 Record of Quality assurance compliance.

Quality assurance	**Record of compliance**
Ensure version control	Version numbers are included in cover tab. Document recording date/time raw data is received for data used in the final model.
Analysist self-checks	Confirmation that the analyst conducted self-checks during model development (including checking sum-totals, spot checking results, checking patterns in physical and monetary results).
Quality assurance guidelines in place	Confirmation that the quality assurance guidelines have been followed, including documenting key assumptions, logging comments by reviewers and following the general guidance.
Periodic reviews occurred	Confirmation that periodic reviews occurred throughout the modelling process. Results are also reviewed in conjunction with other modelling results in the IMP for expected changed (e.g. reductions in livestock lead to decreases in GHG emission and water pollution).
Internal peer review	An internal peer review has been completed.

15 Concluding remarks

Within the land use scenario analysis the ERAMMP IMP simulates the potential effects of changes in farm gate prices on agriculture, land use and the natural environment in Wales. The nature of decision-making around these areas is inherently complex due to the range of interdependencies between different drivers, sectors, and the varied actors within them. The IMP's integrated approach recognises that drivers or policies in one sector may have consequences or effects in others. Despite efforts to represent these complex relationships and interdependencies, all models, by necessity, are a simplification of reality, but can still provide very useful insights if applied for a specific purpose and with caution. In this case, the IMP has been designed to support policy decision-making that will have real-world impacts. It is therefore essential that the model outputs are critically evaluated if they are to be used with genuine understanding and confidence. QA provides the critical reflection needed. **This document has reported the key QA issues for each model within the IMP chain as well as for key areas of intersection between models. Newly developed models (e.g. the LAM) have undergone additional scrutiny and sensitivity analysis.**

As the IMP supports the development of core elements of government policy, it is designated as business critical and as such is mandated by the UK *Government's Review of quality assurance of government analytical models*[9] and *Aqua Book*[10] . For successful QA, there must be both a modelling environment that creates conditions in which QA processes can operate effectively and clear process for every stage of the model life-cycle. An environment and the processes that foster effective QA are delivered through compliance with the four Aqua Book principles of analytical QA:

- **Proportionality of response**: The extent of the analytical quality assurance effort should be proportionate in response to the risks associated with the intended use of the analysis.
- **Assurance throughout development:** Quality assurance considerations should be considered throughout the life cycle of the analysis and not just at the end.
- **Analysis with RIGOUR**: Quality analysis needs to be Repeatable, Independent, Grounded in reality, Objective, have understood and managed Uncertainty, and the results should address the initial question Robustly.
- **Verification and validation:** Analytical quality assurance is more than checking that the analysis is error-free and satisfies its specification (verification). It must also include checks that the analysis is fit for the purpose for which it is being used (validation).

This document has sets out how the ERAMMP IMP team have addressed the requirements of the Review of quality assurance of government analytical models and Aqua Book.

---

[9] Review of quality assurance of government models, https://www.gov.uk/government/publications/review-of-quality-assurance-of-government-models
[10] The Aqua Book: Guidance on producing quality analysis for government https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/416478/aqua_book_final_web.pdf

- The ERAMMP IMP was developed following the principles of co-creation, taking an iterative approach that involved the modelling consortium and government experts throughout.
- The principles of RIGOUR were strictly adhered to with all assumptions underlying the modelling approach agreed, transparently documented and signed-off by an SRO within Welsh Government following a multi-stage iterative discussion (see Annex 1 for the assumptions document).
- In addition, modelling teams employed a range of appropriate methods for quality assurance, including validation, sensitivity analysis, contextualisation, and interpretation, and detailing historical peer review (summarised in this document as technical annex 2).

15.1 Addressing RIGOUR

Core to the Aqua Book are the RIGOUR principles of the work being Repeatable, Independent, Grounded in reality, Objective, have understood and managed Uncertainty, and the results should address the initial question Robustly. The following section summarises how the QA of the IMP has addressed these principles throughout the project.

Repeatability:

For an analytical process to be considered valid, it is reasonable to expect that when using the same inputs and constraints, the analysis will produce the same output.

- All assumptions and settings for key parameters, have been transparently documented.
- An auditable record of data passes between models for each scenario was kept using a unique identifier for each model, input and output.
- Each model was independently verified by an expert modelling team to ensure correct implementation and application.

Independent:

Analysis should be free of prejudice or bias and in doing so, care should be taken to balance the views across stakeholders and experts.

- The IMP was developed using an iterative approach considering a range of different perspectives from within the modelling team and across Welsh Government.
- Consultations and workshops were held with a range of Welsh Government representatives and sectoral experts.

Grounded in reality:

Connections between the analysis and its real-world consequences must be made by challenging the views and perspectives of all stakeholders. In doing so, this encourages the context of the problem to be fully understood.

Where possible, each model component was validated by comparing a baseline scenario against appropriate datasets.

- The assumptions document was widely circulated within WG to ensure those with sectoral expertise could sense-check the assumptions made.
- Where no real-world value was known, sensitivity analysis was undertaken to quantify the implications of different parameterisations.

Objective:

Objective analysis should reduce potential bias and enables the end user to genuinely understand the results and therefore enables effective interpretation and application.

- An iterative co-learning approach between the modelling team and WG was taken.
- Effective engagement through regular meetings and workshops provided opportunities for suitable challenge.
- The IMP framework was developed with a focus on transparency, with the model assumptions, results, and QA widely shared.

Uncertainty-managed:

This principle demands the identification, management, and communication of uncertainty throughout the analytical process.

- Parameter uncertainty was addressed by sensitivity analysis.
- Prediction uncertainty was addressed, where possible, through validation against a baseline.
- Peer reviewed models, methods and the application of standard techniques was used wherever possible to minimise uncertainty of model choice.
- Where an uncertainty was identified, this has been communicated within the consortium and WG.

Robust:

This principle argues the need to provide the analytical result in the context of uncertainty and limitations.

- Expert interpretation of results was provided to support WG in the use of the IMP outputs in decision-making.
- All assumptions and their implications were communicated transparently.

This document, in combination with the assumptions document and supporting information details the framework within which the IMP was developed and the key QA processes undertaken for each model. Care must be taken in the interpretation of the results and must consider the assumptions made at each step of the modelling process.

15.2 Conclusion

An integrated modelling platform is complex by its very nature, as is the socio-ecological environment around which it is trying to support decisions. There are challenges with data availability, modelling capability and limits to what it is possible to know and understand. By explicitly addressing these challenges and being open and transparent about methodologies, assumptions, limitations and the QA steps taken to understand them it is possible to gain insights into the behaviour of the socio-ecological system. Furthermore, by building the system iteratively through co-creation with government end users it is possible to better support decision-making through shared learning – about not only the system modelled, but where questions can and can't be answered. In this way, an integrated modelling system can be understood as something other than a black box. Instead, becoming a system that is able to challenge presumptions, identify opportunities and avoid unintended consequences whilst helping governments and academics learn from each other. QA is critical to building understanding of the system, providing confidence in the limits of knowledge, avoiding modelling hubris and ultimately, providing better joined-up, cross-sectoral information to help decision-makers plan for the future.

16 Appendix A: Matching the Welsh Agricultural Survey to >1 FTE ERAMMP farms

In the ERAMMP Integrated Modelling Platform only Welsh farm holdings that are above 1 full time equivalent (FTE) are included. Validation data held in the June Agricultural Survey (JAS) includes both these farms and those < 1 FTE. To compare SFARMOD outputs with JAS data the < 1 FTE farms need to be removed from the JAS data.

This is done by using a scalar estimated from analysis of Standard Output and Standard Labour Requirements thresholds in the Welsh Farm Business Survey provided by the Welsh Government. The scalars used are listed:

- Arable land 0.85
- New grass 0.84
- Permanent grass 0.71
- Sole rights rough grazing 0.9
- Dairy cattle numbers 0.99
- Beef cattle numbers 0.78
- Sheep numbers 0.9

Within the JAC, there are 7,850 holding meeting both Standard Output and Standard labour Requirement Thresholds. This does not precisely align with the cut off used in with ERAMMP. Within the ERAMMP IMP there are 7,726 agricultural holdings with > 1 FTE modelled.

Table A-1: Key variables by farms being above thresholds for Standard Output and Standard Labour Requirement (from the Welsh Agricultural Survey, June 2019)

	Both	SO only	SLR only	Neither	Total	Both	SO only	SLR only	Neither	Total
Farms	7,850	2,299	711	13,947	24,807	32%	9%	3%	56%	100%
Economic proxies										
Output	1,684,393	120,229	11,211	89,901	1,905,735	88%	6%	1%	5%	100%
Labour requirement	28,997	1,439	1,062	3,413	34,911	83%	4%	3%	10%	100%
Land on farms										
Perm Grass	800,339	133,733	16,806	175,440	1,126,318	71%	12%	1%	16%	100%
Rough grazing	217,631	6,623	6,204	23,621	254,080	86%	3%	2%	9%	100%
New grass	134,702	15,340	1,038	10,214	161,294	84%	10%	1%	6%	100%
Crops & horticulture	80,538	10,895	193	2,969	94,595	85%	12%	0%	3%	100%
Woods & others	86,688	11,679	2,351	26,943	127,661	68%	9%	2%	21%	100%
All land	**1,319,898**	**178,271**	**26,591**	**239,188**	**1,763,948**	**75%**	**10%**	**2%**	**14%**	**100%**
Cattle										
Dairy cows	248,994	2,112	6	480	251,592	99%	1%	0%	0%	100%
Beef cows	127,755	24,510	226	11,326	163,817	78%	15%	0%	7%	100%
Calves	269,583	37,703	203	13,528	321,017	84%	12%	0%	4%	100%
Other cattle	317,394	50,939	223	14,862	383,418	83%	13%	0%	4%	100%
All cattle	963,726	115,264	658	40,196	1,119,844	86%	10%	0%	4%	100%
Other livestock										
Sheep	8,594,792	166,077	138,353	634,394	9,533,616	90%	2%	1%	7%	100%
Poultry	8,043,504	363,486	2,928	79,881	8,489,799	95%	4%	0%	1%	100%
Pigs	16,160	3,557	186	4,528	24,431	66%	15%	1%	19%	100%
Goats	7,328	403	327	4,964	13,022	56%	3%	3%	38%	100%
Horses	15,422	874	9,705	19,219	45,220	34%	2%	21%	43%	100%
Each farm is checked against thresholds for Standard Output and Standard Labour Requirement										
Thresholds										
Output (€)	25,000									
Labour (FTE)	1									
Summary categories										
Both	Farms above both thresholds									
SO only	Farms above SO threshold but below SLR threshold									
SLR only	Farms below SO threshold but above SLR threshold									
Neither	Farms below both thresholds									

17 References

Introduction

HM Treasury (2013) Review of quality assurance of government models, https://www.gov.uk/government/publications/review-of-quality-assurance-of-government-models

HM Treasury (2015) The Aqua Book: Guidance on producing quality analysis for government https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/416478/aqua_book_final_web.pdf

SFARMOD

Annetts JE and Audsley E (2002) Multiple objective linear programming for environmental farm planning Journal of the Operational Research Society (2002) 53, 933–943. doi:10.1057/palgrave.jors.2601404

Audsley, E., Trnka, M., Sabaté, S., Maspons, J., Sanchez, A., Sandars, D., Balek, J., Pearn, K. Interactively modelling land profitability to estimate European agricultural and forest land use under future scenarios of climate, socio-economics and adaptation (2014) Climatic Change, 128 (3-4), pp. 215-227. DOI: 10.1007/s10584-014-1164-6

Audsley, E., Pearn, K.R., Harrison, P.A., Berry, P.M. The impact of future socio-economic and climate changes on agricultural land use and the wider environment in East Anglia and North West England using a metamodel system (2008) Climatic Change, 90 (1-2), pp. 57-88. DOI: 10.1007/s10584-008-9450-9

Brown C, Brown E, Murray-Rust D, Cojocaru G, Savin C, Rounsevell M (2015). Analysing uncertainties in climate change impact assessment across sectors and scenarios. Climatic Change 128(3-4): 293-306

Dunford R, Harrison PA, Rounsevell M (2015). Exploring scenario and model uncertainty in cross-sectoral integrated assessment approaches to climate change impacts. Climatic Change 132(3): 417-432

Fronzek, S., Carter, T.R., Pirttioja, N., Alkemade, R., Audsley, E., Bugmann, H., Flörke, M., Holman, I., Honda, Y., Ito, A., Janes-Bassett, V., Lafond, V., Leemans, R., Mokrech, M., Nunez, S., Sandars, D., Snell, R., Takahashi, K., Tanaka, A., Wimmer, F., Yoshikawa, M. Determining sectoral and regional sensitivity to climate and socio-economic change in Europe using impact response surfaces (2019) Regional Environmental Change, 19 (3), pp. 679-693 DOI: 10.1007/s10113-018-1421-8

Harrison PA, Dunford RW, Holman IP, Cojocaru G, Madsen MS, Chen PY (2019). Differences between low-end and high-end climate change impacts in Europe across multiple sectors. Reg Environ Change 19(3): 695-709. doi:10.1007/s10113-018-1352-4

Holman IP, Harrison PA, Metzger MJ (2016). Cross-sectoral impacts of climate and socio-economic change in Scotland - implications for adaptation policy. Regional Environmental Change 16(1), 97-109 DOI: 10.1007/s10113-014-0679-8

Holman, I.P. Nicholls, R.J. Berry, P.M. Harrison, P.A. Audsley, E. Shackley, S. and Rounsevell, M.D.A. (2005). A regional, multi-sectoral and integrated assessment of the impacts of climate and socio-economic change in the UK: II Results. Climatic Change, 71, 43-73.

Hutchings NJ, Gulzari SO, de Hann M, Sandars D (2018) How do farm models compare when estimating greenhouse gas emissions from dairy cattle production? Animal 12(10): 2171-2180

Kebede, A.S., Dunford, R., Mokrech, M., Audsley, E., Harrison, P.A., Holman, I.P., Nicholls, R.J., Rickebusch, S., Rounsevell, M.D.A., Sabaté, S., Sallaba, F., Sanchez, A., Savin, C., Trnka, M., Wimmer, F. Direct and indirect impacts of climate and socio-economic change in Europe: a sensitivity analysis for key land- and water-based sectors (2015) Climatic Change, 128 (3-4), pp. 261-277. DOI: 10.1007/s10584-014-1313-y

Lee H; Brown C, Seo B, Holman I, Audsley E, Cojocaru G, Rounsevell M (2019). Implementing land-based mitigation to achieve the Paris Agreement in Europe requires food system transformation. Environ. Research Letters 14: 104009

Papadimitriou L, Trnka M, Harrison P, Holman I (2019a). Cross-sectoral and trans-national interactions in national-scale climate change impacts assessment—the case of the Czech Republic. Reg. Environ. Change 19: 2453–2464

Papadimitriou L, Holman IP, Dunford R, Harrison PA, (2019b). Trade-offs are unavoidable in multi-objective adaptation even in a post-Paris Agreement world. Sci Total Environ. 696: 134027

Forestry and Carbon

Beauchamp, K., & Bathgate, S., Ray, D., Nicol, B. (2016) Forest ecosystem service delivery under future climate scenarios and adaptation management options: a case study in central Scotland. Scottish Forestry. 70. 30-41. www.researchgate.net/publication/311562293_Forest_ecosystem_service_delivery_under_future_climate_scenarios_and_adaptation_management_optionsa_case_study_in_central_Scotland

Bendelow, V. C. and Hartnup, R. (1980). Clim atic classification of England and Wales. Soil Survey Technical Monograph No. 15. Rothamsted Experimental Station, Harpenden. 27 pp.

Birse, E.L.. (1971). Assessment of climatic conditions in Scotland, 3: The bioclimatic sub575 regions. Macauley Land Use Research Institute, Aberdeen.

Broadmeadow, MSJ., Webber, J., Ray, D ., Berry, P. (2009). An assessment of likely future impacts of climate change on UK forests. www.researchgate.net/publication/237049052_An_assessment_of_likely_future_impacts_of_climate_change_on_UK_forests

Broome, A., & Ray, D., Mitchell, R., Harmer, R. (2018). Responding to ash dieback (Hymenoscyphus fraxineus) in the UK: woodland composition and replacement tree species. Forestry: An International Journal of Forest Research. www.researchgate.net/publication/328803210_Responding_to_ash_dieback_Hymenoscyphus_fraxineus_in_the_UK_woodland_composition_and_replacement_tree_species

Brown, P., Cardenas, L., Choudrie, S., Del Vento, S., Karagianni, E., MacCarthy, J., Mullen, P., Passant, N., Richmond, B., Smith, H., Thistlethwaite, G., Thomson, A., Turtle, L. & Wakeling, D. (2021). UK Greenhouse Gas Inventory, 1990 to 2019. Annual Report for Submission under the Framework Convention on Climate Change. BEIS: London, https://naei.beis.gov.uk/reports/reports?section_id=3

Evans, C., Morrison, R., Burden, A., Williamson, J., Baird, A., Brown, E., Callaghan, N., Chapman, P., Cumming, A., Dean, H., Dixon, S., Dooling, G., Evans, J., Gauci, V., Grayson, R., Haddaway, N., He, Y., Heppell, K., Holden, J., Hughes, S., Kaduk, J., Jones, D., Matthews, R., Menichino, N., Misselbrook, T., Page, S., Pan, G., Peacock, M., Rayment, M., Ridley, L., Robinson, I., Scowen, M., Stanley, K., and Worrall, F. (2017) Interim report on project SP1210: Lowland peatland systems in England and Wales – evaluating greenhouse gas fluxes and carbon balances. Report to Defra, Centre for Ecology and Hydrology, Bangor.

Hardaker, A. (2018). Is forestry really more profitable than upland farming? A historic and present day farm level economic comparison of upland sheep farming and forestry in the UK, Land Use Policy, 71, 98-120

HM Treasury (2018) The Green Book CENTRAL GOVERNMENT GUIDANCE ON APPRAISAL AND EVALUATION https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/938046/The_Green_Book_2020.pdf

Jenkins, T.A.R., Mackie, E.D., Matthews, R.W., Miller, G., Randle, T.J. and White, M.E. (2018) FC Woodland Carbon Code: Carbon Assessment Protocol (v2.0). Forestry Commission: Edinburgh, https://www.woodlandcarboncode.org.uk/images/PDFs/WCC_CarbonAssessmentProtocol_V2.0_March2018.pdf

Matthews, R.W., Jenkins, T.A.R., Mackie, E.D. & Dick, E.C. (2016). Forest Yield: a Handbook on Forest Growth and Yield Tables for British Forestry. Forestry Commission, Edinburgh.

Matthews, R., Ražauskaitė, R., Hogan, G., Mackie, E., Sayce, M. & Randle, T. (2021a). The CARBINE model: a technical description. Forest Research technical report. In preparation.

Matthews, R. (2021b) Quantifying the sustainable forestry carbon cycle (QFORC) Forest Research Report to Scottish Forestry

Pyatt, D. G., Ray, D., & Fletcher, J. (2001). An Ecological Site Classification for Forestry in Great Britain Forestry Commission Bulletin 124

Quine, C. P. and White, I. M. S. (1993). Revised windiness scores for the windthrow hazard classification: the revised scoring method. Forestry Commission Research Information Note 230. Forestry Commission, Edinburgh, and

Quine, C. P. and White, I. M. S. (1994). Using the relationship between rate of tatter and topographic variables to predict site windiness in upland britain. Forestry 67 (3), 245-256.

Ray, D. (2008a) Impacts of climate change on forestry in Wales Forestry Commission Wales Research Note FCRN301 www.researchgate.net/publication/237049039_Impacts_of_climate_change_on_forestry_in_Wales

Ray, D. (2008b) Impacts of climate change on forestry in Scotland – a synopsis of spatial modelling research Forestry Commission Scotland Research Note FCRN101 www.researchgate.net/publication/237049035_Impacts_of_climate_change_on_forests_in_Scotland_-_a_preliminary_synopsis_of_spatial_modelling_research

Ray, D., Bathgate, S., Moseley, D., Taylor, P., Nicoll, B., Pizzirani, S. et al. (2015) Comparing the provision of ecosystem services in plantation forests under alternative climate change adaptation management options in Wales. Regional Environmental Change 15, pages 1501–1513 https://link.springer.com/article/10.1007/s10113-014-0644-6

Ray, D., Marchi, M., & Rattey, A., & Broome, A. (2021) A multi-data ensemble approach for predicting woodland type distribution: Oak woodland in Britain. Ecology and Evolution. 11 (14) 9423-9434 www.researchgate.net/publication/352568021_A_multi-data_ensemble_approach_for_predicting_woodland_type_distribution_Oak_woodland_in_Britain

Ray, D., Morison, J., Broadmeadow, M. (2010) Climate change: impacts and adaptation in England's woodlands Forestry Commission England Research Note FCRN201 www.researchgate.net/publication/269275863_Climate_change_impacts_and_adaptation_in_England's_woodlands

Ray, D., & Petr, M., Mullett, M., Bathgate, S., & Marchi, M., & Beauchamp, K. (2019) A simulation-based approach to assess forest policy options under biotic and abiotic climate change impacts: A case study on Scotland's National Forest Estate. Forest Policy and Economics. 103. 17-27. https://www.sciencedirect.com/science/article/pii/S1389934116304026

Saraev, V., Edwards, D., Valatin, G. (2017a) Timber, carbon and wind risk: towards an integrated model of optimal rotation length Forestry Commission Research Report FCRP028 www.forestresearch.gov.uk/documents/813/FCRP028.pdf

Saraev, V., MacCallum, S., Moseley, D., Valatin, G. (2017b) Valuation of Welsh Forest Resources Forest Research Report to Welsh Government www.forestresearch.gov.uk/documents/1680/FR_Valuation_Welsh_Forest_Valatin.pdf

Thompson, D.A. & Matthews, R.W. (1989). The storage of carbon in trees and timber. Forestry Commission Research Information Note 160. Forestry Commission: Edinburgh.

Thomson, A., Buys, G., Clilverd, H., Malcolm, H., Henshall, P. and Matthews, R. (2020). Projections of Emissions and Removals from the LULUCF Sector to 2050/2100 (consistent with the 1990-2017 inventory). Prepared by the UK Centre for Ecology & Hydrology (UKCEH) and Forest Research for the Department for Business, Energy and Industrial Strategy (BEIS) as part of the UK National Atmospheric Emissions Inventory. BEIS: London. https://naei.beis.gov.uk/reports/reports?section_id=7

UKWCC (2020) UK Woodland Carbon Code Carbon Calculation Spreadsheet Version 2.2, 9th January 2020

UKWCC (2021) UK Woodland Carbon Code: Standards and Guidance. https://woodlandcarboncode.org.uk/standard-and-guidance

Birds

Bladwell, S., Noble, D.G., Taylor, R., Cryer, J., Galliford, H., Hayhow, D.B., Kirby, W., Smith, D., Vanstine, A. & Wotton, S.R. (2018). The state of birds in Wales 2018. RSPB, BTO, NRW and WOS. RSPB Cymru, Cardiff

Cook, R. D. (1977). Detection of influential observation in linear regression. Technometrics, 19 (1), 15–18

Eaton, M., Aebischer, N., Brown, A., Hearn, R., Lock, L., Musgrove, A., Noble, D., Stroud, D. and Gregory, R. (2015). Birds of Conservation Concern 4: the population status of birds in the UK, Channel Islands and Isle of Man. British Birds, 108(12), 708-746.

Evans, K.L., Newson, S.E. and Gaston, K.J. (2009), Habitat influences on urban avian assemblages. Ibis, 151: 19-39. https://doi.org/10.1111/j.1474-919X.2008.00898.x V

Freeman, S. N., Noble, D. G., Newson, S. E. & Baillie, S. R. (2007). Modelling population changes using data from different surveys: the Common Birds Census and the Breeding Bird Survey, Bird Study, 54:1, 61-72, DOI: 10.1080/00063650709461457

Krinsky, I., & Robb, A. L. (1986). On Approximating the Statistical Properties of Elasticities. The Review of Economics and Statistics, 68(4), 715–719. https://doi.org/10.2307/1924536

Massimino, D, Johnston, A. & Pearce-Higgins, J.W. (2015). The geographical range of British birds expands during 15 years of warming, Bird Study, 62:4, 523-534, DOI: 10.1080/00063657.2015.1089835

Plummer, K. E., Gillings, S., & Siriwardena, G. M. (2020). Evaluating the potential for bird-habitat models to support biodiversity-friendly urban planning. Journal of Applied Ecology V

Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J.J., Schröder, B., Thuiller, W. and Warton, D.I. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography, 40(8),913–929.

Rowland, C.S., Morton, R.D., Carrasco, L., McShane, G., O'Neil, A.W., & Wood, C.M. (2017) Land Cover Map 2015 (1km percentage target class, GB). NERC Environmental Information Data Centre. https://doi.org/10.5285/505d1e0c-ab60-4a60-b448-68c5bbae403e

Scholefield, P. A., Morton, R. D., Rowland, C. S., Henrys, P. A., Howard, D. C., & Norton, L. R. (2016). Woody linear features framework, Great Britain v.1.0. NERC Environmental Information Data Centre. https://doi.org/10.5285/d7da6cb9-104b-4dbc-b709-c1f7ba94fb16

Siriwardena, G.M., Cooke, I.R. and Sutherland, W.J. (2012). Landscape, cropping and field boundary influences on bird abundance. Ecography, 35: 162 - 173. https://doi.org/10.1111/j.1600-0587.2011.06839.x

Thomas, A., Cosby, B.J., Gooday, R., Lyons, H., Siriwardena, G., Kettel, E., Matthews, R., Beauchamp, K., Petr, M. and Emmett, B.A. (2021). Rapid adaptive modelling for policy support towards achieving Sustainable Development Goals: Brexit and the livestock sector in Wales. Environmental Science & Policy, 125, 21-31.

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate research, 30(1), 79-82.

Plants

de Vries, W, et al. (2010). Use of dynamic soil-vegetation models to assess impacts of nitrogen deposition on plant species composition and to estimate critical loads: an overview. Ecol. Applications. 20, 60-79.

Defra. (2015). Managing grassland diversity for multiple ecosystem services. Retrieved from http://sciencesearch.defra.gov.uk/Default.aspx?Menu=Menu&Module=More&Location=None&Completed=0&ProjectID=17251

Ferrier, S. & Guisan, A. (2006). Spatial modelling of biodiversity at the community level. Journal of Applied Ecology, 43: 393-404. https://doi.org/10.1111/j.1365-2664.2006.01149.xHayes, M. J., & Lowther, R. A. (2014). Conservation management of species-rich grasslands in the Elan Valley, Radnorshire. Retrieved from https://www.elanvalley.org.uk/sites/default/files/fileman/Hayes__Lowther_2014_-_Conservation_management_of_species-rich_grasslands_in_the_Elan_Valley_Radnorshire_-_Evidence_report_8.pdf

Henrys, PA et al. (2015) Niche models for British plants and lichens obtained using an ensemble approach. New Journal of Botany 5, 89-100.

Kirkham, F., Dunn, R., Tallowin, J., Bhogal, A., & Chambers, B. (2011). BD1468 Appendix 2, Are inputs of fertilizer and lime compatible with the restoration of species diversity to hay meadows? Evidence from a 12 year fertilizer and lime experiment. Retrieved from Department for Environment, Food and Rural Affairs (Defra), Nobel House, 17 Smith Square, London SW1P 3JR helpline@defra.gsi.gov.uk website: http://sciencesearch.defra.gov.uk/Default.aspx?Module=More&Location=None&ProjectID=16130

Marrs, R., Sánchez, R., Connor, L., Blackbird, S., Rasal, J., & Rose, R. (2018). Effects of removing sheep grazing on soil chemistry, plant nutrition and forage digestibility: Lessons for rewilding the British uplands. Annals of Applied Biology, 173. doi: 10.1111/aab.12462

McGovern, S., Evans, C. D., Dennis, P., Walmsley, C., Turner, A., & McDonald, M. A. (2013). Resilience of upland soils to long term environmental changes. Geoderma, s 197–198, 36–42. doi: 10.1016/j.geoderma.2012.12.018

Pywell, R. F., Bullock, J. M., Tallowin, J. B., Walker, K. J., Warman, E. A., & Masters, G. (2007). Enhancing diversity of species-poor grasslands: An experimental assessment of multiple constraints. Journal of Applied Ecology, 44, 81–94. doi: 10.1111/j.1365-2664.2006.01260.xPywell, R. F., Webb, N. R., & Putwain, P. D. (1994). Soil fertility and its implications for the restoration of heathland on farmland in southern Britain. Biological Conservation, 70, 169–181.

Richardson, D. M., Pysek, P., Rejmanek, M., Barbour, M. G., Panetta, F. D., & West, C. J. (2000). Naturalization and Invasion of Alien Plants: Concepts and Definitions. Diversity and Distributions, 6(2), 93–107. http://www.jstor.org/stable/2673320

Rothamsted Research. (2016). Park Grass soil pH. Retrieved from https://doi.org/10.23637/KeyRefOAPGsoilpH

Rowe, EC et al. (2015) Field survey based models for exploring nitrogen and acidity effects on plant species diversity and assessing long-term critical loads. In: W. de Vries et al. (eds.), Critical Loads and Dynamic Risk Assessments, Environmental Pollution 25, DOI 10.1007/978-94-017-9508-1_11

Smart, SM, et al. (2019) Assessment of a large number of empirical plant species niche models by elicitation of knowledge from two national experts. Ecology & Evolution. DOI: 10.1002/ece3.5766

Smart, SM, et al. (2017) Changes in the frequency of common plant species across linear features in Wales from 1990 to 2016: implications for potential delivery of ecosystem services. New Journal of Botany 7, 112-124. doi.org/10.1080/20423489.2017.1408190

Smart, SM et al. (2015) Common plants as indicators of habitat suitability for rare plants; quantifying the strength of the association between threatened plants and their neighbours. New Journal of Botany 5, 72-88.

Smart, SM, et al. (2010) Empirical realized niche models for British higher and lower plants – development and preliminary testing. Journal of Vegetation Science 21, 643-656.

Wagner, M., Bullock, J. M., Hulmes, S., Hulmes, L., Peyton, J., Amy, S., … Pywell, R. F. (2014). BD1459: Techniques to enhance the establishment and persistence of poor-performing species in grassland restoration - RP00199. Retrieved from http://publications.naturalengland.org.uk/publication/5700071847886848?category=513173

Air Quality

Atkinson, R. W., S. Kang, H. R. Anderson, I. C. Mills, and H. A. Walton. 2014. “Epidemiological Time Series Studies of PM2.5 and Daily Mortality and Hospital Admissions: A Systematic Review and Meta-Analysis.” Thorax 69: 660–665. doi:10.1136/thoraxjnl-2013-204492.

Engledew, M., Maclean, K., Thomas, T., Fitch, A., Robinson, D.A. & Jones, L. (2019) Environment and Rural Affairs Monitoring & Modelling Programme – ERAMMP Year 1 Report 24: Welsh National Natural Capital Accounts - Ecosystem Service Accounts for Woodland, Farmland and Freshwater Habitats. Report to Welsh Government (Contract C210/2016/2017). Centre for Ecology & Hydrology Project NEC06297

Jones, L., Vieno, M., Fitch, A., Carnell, E., Steadman, C., Cryle, P., Holland, M., Nemitz, E., Morton, D., Hall, J. and Mills, G., 2019. Urban natural capital accounts: developing a novel approach to quantify air pollution removal by vegetation. Journal of Environmental Economics and Policy, 8(4), pp.413-428.

Simpson, D., A. Benedictow, H. Berge, R. Bergström, L. D. Emberson, H. Fagerli, C. R. Flechard, et al. 2012. “The EMEP MSC-W Chemical Transport Model – Technical Description.” Atmospheric Chemistry and Physics 12: 7825–7865. doi:10.5194/acp-12-7825-2012.

Vieno, M., M. R. Heal, M. L. Williams, E. J. Carnell, E. Nemitz, J. R. Stedman, and S. Reis. 2016. “The Sensitivities of Emissions Reductions for the Mitigation of UK PM2.5.” Atmospheric Chemistry and Physics 16: 265–276. 10.5194/acp-16-265-2016.

Valuation

Bradley RI, Milne R, Bell J, Lilly A, Jordan C, Higgins A "A soil carbon and land use database for the United Kingdom." Soil Use and Management 21.4 (2005): 363-369.

Evans, C., Artz, R., Moxley, J., Smyth, M-A., Taylor, E., Archer, N., Burden, A., Williamson, J., Donnelly, D., Thomson, A., Buys, G., Malcolm, H., Wilson, D., Renou-Wilson, F., Potts J. (2017). Implementation of an emission inventory for UK peatlands. Report to the Department for Business, Energy and Industrial Strategy, Centre for Ecology and Hydrology, Bangor.88pp.

HM Treasury (2018) The Green Book, which is available at https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-governent.

HM Treasury (2020) GDP deflators at market prices, and money GDP; available at https://www.gov.uk/government/collections/gdp-deflators-at-market-prices-and-money-gdp

This page blank.

ERAMMP Programme Office
UKCEH Bangor
Environment Centre Wales
Deiniol Road
Bangor, Gwynedd
LL57 2UW
+ 44 (0)1248 374500
erammp@ceh.ac.uk

www.erammp.wales
www.erammp.cymru