Science of the Total Environment xxx (xxxx) 153824

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com

Exploring the effects of land management change on productivity, carbon and nutrient balance: Application of an Ensemble Modelling Approach to the upper River Taw observatory, UK

Kirsty L. Hassall[a], Kevin Coleman[b,1], Prakash N. Dixit[b,1], Steve J. Granger[c], Yusheng Zhang[c], Ryan T. Sharp[b], Lianhai Wu[c], Andrew P. Whitmore[b], Goetz M. Richter[b], Adrian L. Collins[c,1], Alice E. Milne[b,1,*]

[a] Computational and Analytical Sciences Department, Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK
[b] Sustainable Agriculture Sciences Department, Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK
[c] Sustainable Agriculture Sciences department, Rothamsted Research, North Wyke, Oakhampton EX20 2SB, UK

ARTICLE INFO

Editor: Fernando A.L. Pacheco

Keywords:
Agroecosystems modelling
Trade-offs
Ensemble modelling
Nutrients flows
Arable and livestock systems

ABSTRACT

Agriculture is challenged to produce healthy food and to contribute to cleaner energy whilst mitigating climate change and protecting ecosystems. To achieve this, policy-driven scenarios need to be evaluated with available data and models to explore trade-offs with robust accounting for the uncertainty in predictions. We developed a novel model ensemble using four complementary state-of-the-art agroecosystems models to explore the impacts of land management change. The ensemble was used to simulate key agricultural and environmental outputs under various scenarios for the upper River Taw observatory, UK. Scenarios assumed (i) reducing livestock production whilst simultaneously increasing the area of arable where it is feasible to cultivate (PG2A), (ii) reducing livestock production whilst simultaneously increasing bioenergy production in areas of the catchment that are amenable to growing bioenergy crops (PG2BE) and (iii) increasing both arable and bioenergy production (PG2A + BE). Our ensemble approach combined model uncertainty using the tower property of expectation and the law of total variance. Results show considerable uncertainty for predicted nutrient losses with different models partitioning the uncertainty into different pathways. Bioenergy crops were predicted to produce greatest yields from Miscanthus in lowland and from SRC-willow (cv. Endurance) in uplands. Each choice of management is associated with trade-offs; e.g. PG2A results in a significant increase of edible calories (6736 Mcal ha^{-1}) but reduced soil C (−4.32 t C ha^{-1}). Model ensembles in the agroecosystem context are difficult to implement due to challenges of model availability and input and output alignment. Despite these challenges, we show that ensemble modelling is a powerful approach for applications such as ours, offering benefits such as capturing structural as well as data uncertainty and allowing greater combinations of variables to be explored. Furthermore, the ensemble provides a robust means for combining uncertainty at different scales and enables us to identify weaknesses in system understanding.

1. Introduction

Over the last century, agriculture has done much to keep up with the demands of a rapidly growing population; not, however, without detrimental impacts on the environment (Foresight, 2011). The challenges associated with increasing the production of healthy and accessible diets sustainably are captured neatly within the Sustainable Development Goals (SDGs) proposed by the United Nations (UN General Assembly, 2015). Arguably, all of the goals have links to agricultural production but of particular relevance are those related to food security, health and

* Corresponding author.
E-mail addresses: kirsty.hassall@rothamsted.ac.uk (K.L. Hassall), kevin.coleman@rothamsted.ac.uk (K. Coleman), prakash.dixit@rothamsted.ac.uk (P.N. Dixit), steve.granger@rothamsted.ac.uk (S.J. Granger), Yusheng.zhang@rothamsted.ac.uk (Y. Zhang), ryan.sharp@rothamsted.ac.uk (R.T. Sharp), lianhai.wu@rothamsted.ac.uk (L. Wu), andy.whitmore@rothamsted.ac.uk (A.P. Whitmore), goetz.richter@rothamsted.ac.uk (G.M. Richter), adrian.collins@rothamsted.ac.uk (A.L. Collins), alice.milne@rothamsted.ac.uk (A.E. Milne).
[1] These authors have contributed equally.

https://doi.org/10.1016/j.scitotenv.2022.153824
Received 10 December 2021; Received in revised form 31 January 2022; Accepted 8 February 2022
0048-9697/

wellbeing, sustainable energy production, sustainable food production and consumption, mitigation of climate change and the protection of terrestrial ecosystems (Goals 2, 3, 7, 12, 13, 15, respectively). The polarity of different SDGs, e.g. mitigating climate change and biodiversity loss versus the defeat of hunger (Nilsson et al., 2016; Scharlemann et al., 2020), raises a series of fundamental questions about how best to capitalize on synergies and resolve potential conflicts between multiple societal needs and policy objectives.

In 2018, we held a series of combined expert and stakeholder workshops to "vision" the future of agriculture within the UK with the aim of making it more sustainable in the face of the needs of a growing population (Aarhus University, 2020). One group considered livestock systems in the UK and came up with the somewhat extreme plan that livestock should only be farmed where it was not possible to cultivate arable crops sustainably. Our experts perceived that this would reduce the emissions associated with agriculture whilst still maintaining or increasing the production of food in terms of calories per ha. To achieve goals associated with sustainable energy production, further action is needed, and undoubtedly one of the largest contributions agricultural land can make to sustainable energy production is the introduction of perennial bioenergy crops (Agostini et al., 2015; Gregory et al., 2018). We are thus faced with complex and challenging decisions on how agricultural land should be used to support the ambitions of the SDGs.

Decision making can be informed by simulation models. However, to be useful and explicit, models need to capture and acknowledge uncertainty in the whole system to avoid unintended biases or misleading results (Willcock et al., 2020). This is particularly true in the case of natural systems, where the complex processes that underpin dynamics and responses are not fully understood and whereby contrasting model assumptions can lead to divergent results. In climate modelling, it is commonplace to use ensembles of models to address the challenges related to predicting complex systems (Suarez-Gutierrez et al., 2021). These ensembles are produced by running simulations with more than one model, model class and/or set of model parameters and then combining the outputs. The advantage being that the conclusions are then not dependent on a single set of assumptions and parameters, hence capturing better the uncertainty within the system (Willcock et al., 2020). Combining outputs from several models has improved climate predictions by reducing noise or unforced variability that may be present in each of the input models (Bradley et al., 2017; Taylor et al., 2012; Wallach et al., 2016). It is accepted that biases exist in individual models and that some may be more appropriate for use in different geoclimatic settings (where they were developed for example) or different circumstances depending on their specific strengths. Nonetheless, there is considerable value in deriving an ensemble of results to help estimate the uncertainty in predictions of future responses and this variability in outcomes has been used as a kind of fingerprinting technique to help attribute cause to effect (e.g. Gedney et al., 2006). Wallach et al. (2016) describe the application of the ensemble approach to crop modelling, highlighting both improved predictions (see also Yin et al., 2017) and closer collaboration within the community as key advantages.

Although climate and climate change scientists have made much use of these techniques (e.g. Aryal and Zhu, 2020; Virkkala et al., 2021; Zhu et al., 2013), and increasingly the method has gained attention with yield prediction (Kafatos et al., 2017; Martre et al., 2015; Ruane et al., 2016; Wallach et al., 2016) much less work has been reported with ensembles in other areas of environmental science such as the environmental footprints of modern intensive farming. Willcock et al. (2020) provide one notable example which investigated an ensemble approach in relation to ecosystem services provision in sub-Saharan Africa. They conclude that their ensemble was more accurate than individual model predictions, and that, importantly, the variation with the ensemble not only provides a measure of precision but is a proxy for accuracy. Other examples relevant to agroecosystems tend to focus on a single outcome. For example, Gaillard et al. (2018) used an ensemble of three process-based models to study nitrous oxide emissions from agriculture, and Yin et al. (2017) used ensembles to predict grain nitrogen. Ehrhardt et al. (2018) go further with ensembles by considering joint predictions of productivity and emissions. Apart from these studies, ensemble modelling within the context of agroecosystems remains limited, occasional comparisons have been made between models (McVoy et al., 1995; Riggers et al., 2019; Smith et al., 1997) but, these have stopped short of combining outputs of agroecosystem models to develop a robust prediction system. There are several reasons why ensemble modelling is particularly challenging with respect to agroecosystems. These include the availability of sufficient data and models to form an ensemble, the need to align model input and output variables to build a model ensemble effectively, and the challenges associated with coercing models to a common temporal or spatial scale (Gneiting and Raftery, 2005).

Against the above background and ongoing scientific gaps, we developed a multi-model ensemble approach by combining the outputs of four agroecosystems models within a hierarchical statistical framework and used this to explore the likely impacts of land management change on multiple productivity and environmental objectives. Each model is different from the others both in terms of the processes considered and how these are captured. Each model has its own area of specialisation and no one model simulated all the productivity and environmental objectives we considered. The benefits of the ensemble approach in this context are therefore two-fold: (1) we can capture a greater range of model outputs than if we were to use just a single model, and (2) we can quantify both within and between model uncertainty.

To ensure that our results were realistic we based our simulations on an experimental observatory catchment (41.3 km^2) for which we have a large amount of data on the current state of soil, water and nutrient flows and survey information on management practices across farming sectors. The observatory is located in the south-west of England in an area where farming is largely characterised as a mixture of livestock and arable. The specific scenarios modelled were: (a) "business as usual" (BAU), representing current land cover and associated farm management; (b) producing meat in areas only where it is not possible to grow arable crops (meat-livestock restricted to moorland areas and permanent grass converted to arable where possible) (PG2A), and (c) reducing meat production whilst simultaneously increasing bioenergy production in areas of the catchment that are amenable to growing *Miscanthus* or short rotation coppice (PG2BE). Further, these scenarios, were viewed on a continuum, whereby: (d) increasing, both arable and *bioenergy* production, whilst restricting meat production to upland areas (PG2A + BE). Through this scenario analysis we aim to demonstrate the utility of the ensemble as an approach for agroecosystems modelling to capture uncertainty, to highlight weakness in systems understanding, and to explore the trade-offs incurred from different land management options.

2. Methodology

2.1. Study catchment

Our simulated landscape was based on the upper River Taw observatory (Granger et al., 2021). The River Taw catchment is located in Devon, South West England, and in its entirety, drains an area of 914 km^2 (Environment Agency, 2020; Granger et al., 2017). The headwaters rise in the south on the Dartmoor granite plateau ca. 550 m above sea level. The river then flows northwards 72 km to the Taw/Torridge estuary, and the Bristol Channel. Our study focusses on part of the upper River Taw catchment (Fig. 1). The area is approximately 15 km in length stretching from the source of the river to just south of the town of North Tawton. Aside from the unimproved semi-natural grass in the uplands (>200 to 300 m asl) below the heathlands of Dartmoor, the land use is predominantly one of improved agricultural grassland in the lowlands (<200 m asl). This supports beef, dairy and sheep produc-

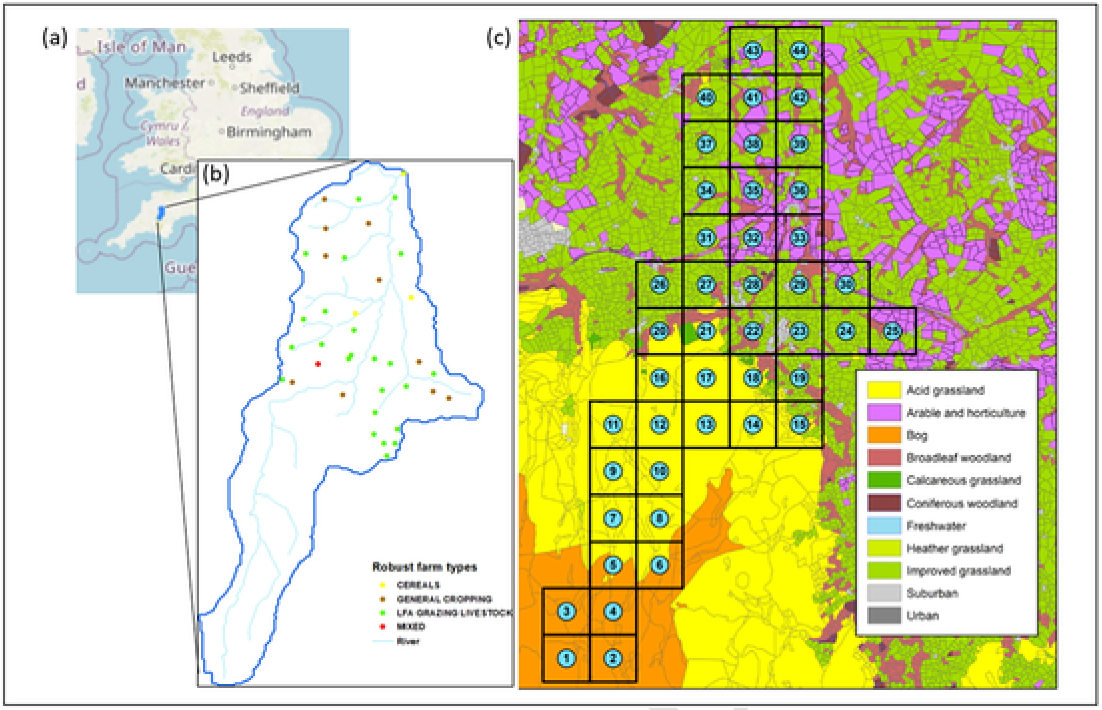

Fig. 1. (a) The location of the upper River Taw catchment in Devon, UK; (b) the location of farms within the upper River Taw catchment, and (c) the land cover allocation of each 1 km × 1 km cell. Acid grassland, heather grassland and calcareous grassland are allocated to "rough grazing".

tion. Cereals and fodder maize are also produced, particularly on the sandier, free-draining soils. Human settlement consists of scattered farmsteads and small rural towns and villages.

For our modelling, we partitioned the study catchment into 44 cells of size 1 km × 1 km (Fig. 1). Grid cells 1–4 were excluded from the modelling framework as they are organic soils and assumed to be bog. We determined the dominant soil series for each cell using the NATMAP Vector data product from the National Soil Resources Institute© Cranfield University (NSRI) (see Table S1). Key soil properties (e.g. pH, organic C) needed for the models were extracted from NATMAP. A total of seven different soil types are present across the modelled catchment (Table S1). The topography of each cell (elevation and slope) was derived from the Integrated Hydrological Digital Terrain Model (CEH, 2017). Each cell in the catchment area was allocated to one of three zones: high rainfall moorland (grid cells 1–8), low rainfall moorland (rough grassland in the uplands, grid cells 9–18, 20) or mixed farming in the lowlands known as a rural land register area, (RLR, grid cells 19, 21–44) (see Table S1), which is agricultural land on which grants and subsidies can be claimed. For the time period simulated, average daily temperatures ranged between 9 and 10.3 °C for RLR, 7.7 and 9.4 °C for low rainfall moorland and 7.6 to 8.2 °C for high rainfall moorland. The mean number of ground frost days per year ranged 87.1–90.5 for RLR, 87.5–97.1 for low rainfall moorland and 92.5–96.3 for high rainfall moorland (see Table S2). To determine the BAU land cover in each 1 km × 1 km grid cell (i.e., the proportion of land allocated to non-agriculture, arable, improved grassland, rough grazing and woodland) we used the Land Cover Map 2015 (Rowland et al., 2017) (see Table S3). Each individual model used land cover information in a slightly different way (see details below).

Arable cropping was derived from the Crop Map for England (CROME, 2019) (S.I. Tables S4 and S5). Our models run at various spatial and temporal scales and this is reflected in the weather variables that are used by each (see descriptions below).

2.2. Models

The ensemble comprised: (i) the AGRicultural Environment MOdelling and Systems Analysis (AGREMOSA) modelling framework, incorporating arable, grassland and biomass crops (Cerasuolo et al., 2016; Qi et al., 2017; Richter et al., 2006); (ii) the Rothamsted Landscape Model (RLM) (Coleman et al., 2017, 2021); (iii) the Soil-Plant-Atmosphere Continuum SYStem Model (SPACSYS) (Wu et al., 2007, 2015; Wu et al., 2019), and (iv) the Catchment Systems Model (CSM) (Zhang et al., 2022). These models were selected because they had been validated for a wide range of conditions, including those similar to the study catchment, and we had full access to the code and their process description. Additionally and importantly, each model is different from the others, both, in terms of the processes considered and how these are modelled (including variation in the temporal and spatial scale of outputs). For example, the AGREMOSA framework focusses more on water-limited production and associated biophysical limitations, whilst RLM, SPACSYS and CSM include nutrient dynamics. SPACSYS and AGREMOSA were developed for field-scale prediction whilst RLM is intended to describe the dynamics of multiple linked fields and CSM predicts process outputs at farm to catchment scale. Each model is described in brief below.

2.2.1. AGREMOSA (AGRicultural Environment MOdelling and Systems Analysis)

AGREMOSA is a modelling and optimization framework of process-based models simulating water-limited production of arable and perennial crops, which include grassland and biomass crops for bioenergy. It is based on the STAMINA modelling framework (Richter et al., 2006), which covers a wide range of arable crops, following the principles of SUCROS (van Laar et al., 1992) and was calibrated for wheat (Richter et al., 2010). It was successively expanded implementing the sink-source interaction approach of LINGRA (Hoglind et al., 2001; Schapendonk et al., 1998) to simulate growth of grasslands and perennial biomass crops like tall grasses (Ni et al., 2019; Triana et al., 2011) and short-rotation coppice (SRC)-willow (Cerasuolo et al., 2016). All models in the AGRE-

MOSA modelling framework assume that nutrients are well managed and therefore are not limiting. AGREMOSA simulates the water and energy balance at an hourly and plant phenology and growth at a daily time-step.

The LUCASS (Light Use and Carbon Assimilation in Salix Species); (Cerasuolo et al., 2016) is a process-based growth model for SRC-willow included in the AGREMOSA framework. It simulates development and growth of *Salix* spp. at the stand scale, considering phenological and morphological plant development (sink formation), light interception, photosynthesis, and respiration (source formation). The organs of the above ground biomass (leaves, branches, and stems) and below ground biomass (stool and all roots) are considered as sinks, and the carbon allocation to these sinks is phenologically controlled and balanced following the principles of the sink-source interaction model proposed for grass (Schapendonk et al., 1998). The LUCASS model was calibrated in two locations in the UK with and without water stress using carbon partitioning data of a 2-year rotation following the year of establishment. It was validated for two successive 2-year rotations for stem, leaves and stool development and using final harvest of 3-year rotations at Rothamsted Research and Long Ashton (South West England) for 'Endurance' and 'Tora' cultivars by Richard et al. (2019). The LINGRA model was calibrated and validated using generalized phenology, photosynthesis and carbon allocation rates to account for the effects of senescence in extensive and semi-natural grasslands in comparison to permanent and temporary grassland using data from across Great Britain (Qi et al., 2017).

The weather data used to run the scenarios were derived from the met station located at the North Wyke Farm Platform. The arable crops simulated for the scenarios described in this paper were winter wheat, spring and winter barley (see Table S6a for management dates). Three types of grassland management systems were simulated: temporary grassland within the arable rotation (Table S6a), permanent grassland and rough grazing. Temporary grassland consists of frequently resown single productive species e.g. perennial ryegrasses (*Lolium perenne*), assuming an annual N application of 300 kg N ha^{-1}. Permanent grasslands consist of a mixture of sown and indigenous grasses and legumes of intermediate productivity receiving an annual N application of 150 kg N ha^{-1}. Rough grazing is low productivity semi-natural grassland containing various herbaceous species receiving no synthetic N inputs (Defra, 2010b; Qi et al., 2017). AGREMOSA implicitly accounts for nutrient effects on sink and source parameters (Qi et al., 2017).

As AGREMOSA does not explicitly simulate livestock to estimate biomass we assume the grass is cut. In the temporary grassland system, grass cuts are assumed to be three times a year (30th May, 20th July and 30th September) whilst in permanent grassland and rough grazing systems, cutting occurs twice a year (21st June and 30th October). Overall, combining seven soil types with three weather zones defined nine distinct soil × weather combinations for AGREMOSA. These were: (i) grid cell 5, (ii) grid cells 6–8, (iii) grid cells 9–18, (iv) grid cell 19, (v) grid cell 20, (vi) grid cell 21, (vii) grid cells 22–27, 29–30, 33, (viii) grid cells 28, 32, 34–44 and (ix) grid cell 31. We assumed that the high rainfall moorland cells (grid cells 5–8) are unsuitable for bioenergy crops and only the lowlands (RLR zone; grid cells 19, 21–44) were suitable for arable and temporary grassland production (see also Table S3).

2.2.2. The Catchment Systems Model

The CSM (Zhang et al., 2022) simulates the environmental footprint of farm systems within a defined catchment based on the Robust Farm Types (Defra, 2010a) present, and in so doing produces farm and landscape scale outputs. The framework combines modules for the interaction between key environmental factors and farm systems, scaling from farm to landscape, non-agricultural pollutant sources, factors generating temporal mismatch for pollution mitigation impacts at farm and landscape scale and life cycle assessment mid-point impacts. CSM integrates a number of well-established process-based models and data layers used extensively for policy support in the UK. Losses of phosphorus and sediment to water are generated using PSYCHIC (Phosphorus and Sediment Yield CHaracterisation In Catchments; (Collins and Anthony, 2008; Collins et al., 2007, 2008, 2009, 2021; Davison et al., 2008; Stroemqvist et al., 2008). Nitrate losses to water are simulated using NEAP-N (Lord and Anthony, 2000; Wang et al., 2016). Soil carbon is generated as a stock (to 1-m depth) assuming soils are in equilibrium using an IPCC Tier 1 methodology (Eggleston et al., 2006) augmented with information on the impact of major farm management practices on the stocks assigned to different land cover types. Nitrous oxide emissions use the IPCC methodology for losses from fertilizers, excreta and managed manure. Ammonia emissions use the tools reported by (Chadwick et al., 2005; Webb and Misselbrook, 2004). Methane losses are generated using (Eggleston et al., 2006). Farm production is not modelled but instead uses publicly available information on regional yields supplemented by bespoke farm surveys in the study catchment. This information is used in the estimation of gross margins on the basis of its monetization.

The CSM uses annual mean weather from HadUK at the 1 km × 1 km spatial scale (Met Office, 2018). For the study catchment, farm types comprised cereal, general cropping, less favoured area (LFA) grazing and mixed. Farm type is based on the Robust Farm Type typology (Defra, 2010a) which is estimated using the dominant contribution to standard outputs. Fertilizer application rates for cereal, dairy, general cropping, mixed and other livestock (lowland grazing livestock and LFA grazing) farm systems were extracted from the British Survey of Fertiliser Practice (Defra, 2019) and supplemented by some commercial farm business surveys in the study catchment. Manure dressing data available in the BSFP suggested no significant change in manure spreading practices for spring crops, autumn crops or grassland between 2013 and 2017. Accordingly, the manure allocation scheme of (Zhang et al., 2017) was applied. Livestock types, counts and ages for the RLR area were sourced at holding level using the June Agriculture Survey (Defra, 2016) supplemented by more recent farm surveys (see Table S7). Livestock information for the moorland areas used information in (Comber et al., 2008). Monthly livestock activity is apportioned between percentage time grazing, housed and in farmyards for each type and age category of animal. Farm system economics include the annual capital (amortised as needed), fixed and operational costs of farm management and operations, including best management practices taken up as a result of regulation, incentivization and advice, as well as the monetized value of farm production (e.g. grain, milk, wool, eggs, carcasses).

2.2.3. The Rothamsted Landscape Model

The Rothamsted Landscape Model (RLM); (Coleman et al., 2017) is a process-based model that simulates soil processes (including soil organic matter, soil nutrient and water dynamics), livestock production, crop growth and yields (wheat, barley, and oats, oilseed rape, beans, sugar beet, forage maize, potato, onions and grass). The model components are based on well-established models such as RothC (Coleman and Jenkinson, 2014), LINTUL (Wolf, 2012), and Century (Parton et al., 1994) as well as many new routines such as an updated root model, and an improved water model, as described in Coleman et al. (2017). The model was calibrated and validated using data from UK arable and grassland systems as described in Coleman et al. (2017).

The RLM runs on a daily time step. Separate weather data sets were used for each of the three zones (high rainfall moorland, low rainfall moorland and RLR). These data were based on the weather for the zone centroid. The model used the soil and topographic variables defined above. As slope is accounted for in the RLM each grid cell is assumed unique. Livestock numbers for each cell were derived from the farm level data used by CSM and were simulated as a homogenous distribution across each of the RLR, high and low rainfall moorland zones, giving an estimated stocking density of each livestock type per cell. Lambs were assumed to be finished at 40 and 38 kg in the RLR and moorland respectively and beef were assumed to have a finishing weight of 517.65 kg (Nix, 2020). The RLM contains a novel crop rotations generator which, based on the crops grown in a given area (in this case the Taw catchment between 2016 and 2019, see Tables S4 and S5) and certain agronomic rules, stochastically generates realistic crop sequences for that area (Sharp et al., 2021). In this way realistic rotations of crop production for a particular area can be simulated. The timing and type of fertilizer is also important in RLM and N was assumed to be applied as ammonium nitrate and P as Triple Superphosphate with the split across time based on RB209 (Defra, 2010b). Table S6b summarises mean fertilizer N and P amounts and timings, plus sowing and harvest dates for the RLM.

2.2.4. SPACSYS (Soil-Plant-Atmosphere Continuum SYStem)

The SPACSYS model (Liu et al., 2013; Wu et al., 2007, 2015; Wu et al., 2019 see also S.I.) is a field scale, weather-driven dynamic simulation model. It includes components for plant growth and development, nitrogen cycling, carbon cycling, phosphorus cycling, soil water redistribution and heat-energy transformation. The soil water component includes representation of water flow to and through field drains as well as downwards through the soil layers. A microbial-based module on nitrification-denitrification was included to estimate nitrogenous gas emissions from soils (Wu et al., 2015). Recently, a component to describe livestock growth was added (Wu et al., 2022). More details about the model are presented in S.I. The SPACSYS model can run on a range of time steps but uses a daily time-step in this study (Wu et al., 2021). Parameters have been calibrated and validated near the study site as well as other sites with different climate and soil conditions for grazing livestock (e.g. Carswell et al., 2019; Wu et al., 2016), silage (e.g. Sándor et al., 2020) and cereal and legume crops (e.g. Bingham and Wu, 2011; Liang et al., 2019; Liu et al., 2013). A sensitivity analysis for the model was conducted with respect to changes in 61 input parameters and their influence on 27 output variables (Shan et al., 2021).

Similar to RLM, separate weather data sets for each of the three zones were used by SPACSYS. SPACSYS assumed a homogeneous distribution of livestock across each of the RLR, low and high rainfall moorland zones. For the livestock categories, the ages at the start of a simulation were set from the data provided for use in CSM and initial liveweights were estimated based on their age (HowMonk, 2022) (see S.I Table S8). It was assumed that ewes give birth in late March each year. The average lamb birth percentages are 118, 114 and 137 for RLR, high and low rainfall moorland zones, respectively, based on the provided ewe and lamb numbers. Lambs are assumed to wean at 105 days after birth and finish when they reach 45 kg/head. When beef cattle and heifers reach 550 kg/head, they are assumed to be finished. Inputs and outputs of all components are organised as a database in Microsoft® SQL Server. Arable cropping was derived from the CROME (see Fig.1 and Tables S4 and S5). Table S6c summarises fertilizer N and P amounts and timings, plus sowing and harvest dates for SPACSYS.

2.3. Ensemble framework

Given the varying temporal and spatial resolutions of each of the component models, our ensemble framework *Combines Hierarchical Information in a Probabilistic manner*. Table 1 details the different outputs produced by each model in this study (see S.I. Fig. S1 for details on how each model feeds into the different productivity and environmental metrics). Each model may define outputs slightly differently due to the contrasting model frameworks. To achieve a level of consistency, all outputs were converted to the same units and in the case of crop yield, converted to edible kcal ha^{-1} year^{-1} accounting for expected field-gate and processing losses. Similarly, both livestock and bioenergy systems were also converted into calorific value (kcal ha^{-1} year^{-1}) (S.I. Tables S9 and S10). Specifically the ensembled variables are: edible calories including cereals, dairy and livestock outputs (kcal ha^{-1}), energy crop biomass (kg DM ha^{-1}), soil carbon (t C ha^{-1}), carbon emissions (t C ha^{-1}) and nutrient losses as N (kg N ha^{-1}) and P (kg P ha^{-1}). We note that although grass biomass is simulated, we do not explicitly consider this in the ensemble but rather the impact on livestock production. Similarly, predicted differences in gross margin between BAU and PG2A are also reported but because this estimate is from only one model (CSM) it is not considered in the ensemble.

Table 1
Outputs produced by each of the four agroecosystems models in this study.

Output variables	Units	Models			
		RLM	SPACSYS	AGREMOSA	CSM
Production					
Crop yield	Edible kcal ha^{-1}	X	X	X	
Milk yield	Edible kcal ha^{-1}	X			
Livestock (beef and lamb) for meat	Edible kcal ha^{-1}	X	X		
Grass biomass	kg ha^{-1}			X	
Energy crop biomass	kg ha^{-1}		X	X	
Gross margin	£ ha^{-1}				X
Carbon					
Soil carbon (0–1 m depth)	t C ha^{-1}	X	X		X
Soil CO_2 release	t C ha^{-1}	X	X		
Animal CO_2 release	t C ha^{-1}		X		
Soil CH_4 emissions	t C ha^{-1}		X		X
Animal CH_4 emissions	t C ha^{-1}	X	X		
Nutrient losses					
NO_3 – leaching, runoff and drains[a]	kg N ha^{-1}	X	X		X
NH_4 – leaching, runoff	kg N ha^{-1}	X	X		
N_2O – emissions	kg N ha^{-1}	X	X		X
N_2 – emissions	kg N ha^{-1}	X	X		
NO – emissions	kg N ha^{-1}		X		
NH_3 – emissions	kg N ha^{-1}		X		X
P[b] – leaching, runoff and drains[a]	kg P ha^{-1}	X			X[b]

[a] RLM and SPACSYS partition pathways into leaching and runoff. CSM partitions pathways into leaching, runoff and preferential flow via field drains.
[b] CSM partitions P separately into dissolved and particulate P fractions.

There are several alternative means of combining individual model outputs into an ensemble (Wallach et al., 2016; Willcock et al., 2020). Methods of ranking and weighting are described in Knutti et al. (2010). Bayesian model averaging is a commonly used statistical method (Aryal and Zhu, 2020; Wallach et al., 2016; Zhu et al., 2013). In this latter approach one starts by assigning weights a priori for each model (often equal weights), and then update the weights based on model agreement with observations. Here, because we have the additional complexity of multiple outputs equal weights were used throughout (committee averaging).

Individual model outputs are of interest at the scale at which they are produced. However, in order to combine sources of model uncertainty, outputs are required at a common scale whilst properly accounting for the different sources of variability. This is done using the tower property of expectation and the law of total variance (Eqs. (1) and (2)):

$$E(Y) = E(E(Y|X)) \tag{1}$$

$$\mathrm{Var}(Y) = E(\mathrm{Var}(Y|X)) + \mathrm{Var}(E(Y|X)) \tag{2}$$

Let X_{rtl}^{ψ} be the annual total for a particular output, from model x for landuse ψ, at location l, in year t, for simulation r. Annual totals are calculated over the hydrological year (1st October–30th September).

Then, $X_{tl}^{\psi} = E(X_{rtl}^{\psi} | R = r)$ is the average output at location l, in year t under land use ψ and $\sigma_{tl,\psi}^{2} = \mathrm{Var}(X_{rtl}^{\psi} | R = r)$ is the associated variance due to independent simulations.

Thus, $X_{l}^{\psi} = E(X_{tl}^{\psi} | T = t)$ is the average output at location l, over the 14-year simulation period (2004–2018) under each land use ψ. The associated variance is given by,

$$\begin{aligned}\sigma_{l,\psi}^{2} &= \mathrm{Var}\left(X_{tl}^{\psi} | T = t\right) \\ &= \mathrm{Var}\left(E\left(X_{rtl}^{\psi} | R = r\right) | T = t\right) + E\left(\mathrm{Var}\left(X_{rtl}^{\psi} | R = r\right) | T = t\right),\end{aligned} \tag{3}$$

and captures the temporal uncertainty whilst propagating through the uncertainty due to independent simulations.

Each grid cell is assumed to be partitioned into different land uses that will vary under the scenarios defined below. Consequently, the expected output per grid cell is given by a weighted average:

$$X_l = \sum_{\psi} \omega_{\psi} X_l^{\psi} \tag{4}$$

where ω_{ψ} is the proportion of land in location l that is in land use ψ. The associated variance is given by:

$$\sigma_l^2 = \mathrm{Var}\left(X_{tl} | T = t\right) = \sum_{\psi} {\omega_{\psi}}^{2} \mathrm{Var}\left(X_{tl}^{\psi} | T = t\right) \tag{5}$$

Finally, to get the expected total output for the entire study catchment for model x, we have $X = \sum_l X_l$ with associated variance $\mathrm{Var}(X) = \sum_l \mathrm{Var}(X_l)$ assuming independence of locations.

To investigate trade-offs, an ensembled output was calculated by taking the catchment level outputs for each model; via the process described above for RLM, SPACSYS and AGREMOSA and by summing the farm scale outputs predicted by CSM and combining these via the tower property of expectation and the law of total variance (Eqs. (1) and (2)). This ensembled output was used to explore the BAU and PG2A scenarios (Section 3.2).

2.4. Scenarios

AGREMOSA, RLM and SPACSYS were run for each of the 40 non-bog 1 km × 1 km grid cells shown in Fig. 1. For each cell, we simulated all feasible agricultural land classes (arable, improved grassland, rough grazing and bioenergy) and combined the outputs from these in a weighted sum to generate the estimate from each cell under each scenario.

Simulating a change in land cover and associated management from permanent grassland to arable (PG2A), was achieved by altering the weighted averages used in BAU. Explicitly, for all grid cells in the moorland, permanent grassland proportions were kept as per BAU acknowledging that this would be an unsuitable environment for arable cultivation. The grid cells within the RLR were gradually reweighted in intervals of 10% of the original permanent grassland proportion in favour of arable cultivation and disfavouring livestock systems. The overall area dedicated to agricultural land use was kept constant within each grid cell (Fig. 2).

In addition to expanding arable systems, we also considered scenarios to include cultivation of perennial crops for the bio-economy (PG2BE). These include growing short rotation coppice willow and growing *Miscanthus*. This approach followed that of PG2A where land use weights were altered to reflect an increasing contribution of bioenergy crops. Explicitly, for all grid cells in the high rainfall moorland, permanent grassland proportions were kept as per BAU acknowledging that this would be an unsuitable environment for bioenergy cultivation. The grid cells within the low rainfall moorland and RLR were gradually reweighted in favour of bioenergy cultivation and disfavouring livestock systems. Preferentially, rough grazing areas were converted first before conversion of the permanent grassland. The overall area dedicated to agricultural land use was kept constant within each grid cell (Fig. 2).

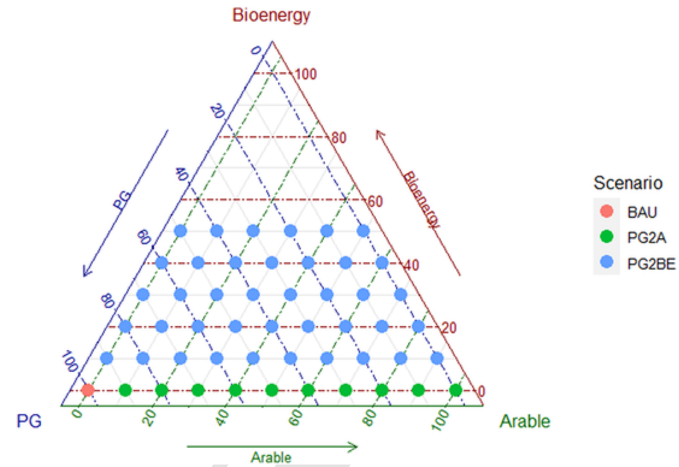

Fig. 2. Illustration of how the land use changes under different scenarios for a single grid cell assumed to have 100% of productive land in permanent grassland under BAU.

The final scenario (PG2A + BE) explored conversion of grassland systems to both arable and bioenergy production. This conversion assumed improved grazing in the lowlands was preferentially converted to arable and rough grazing in the uplands was preferentially converted to bioenergy. Bioenergy crops were also allowed to grow in the low rainfall moorland replacing rough grazing. As per the previous two scenarios, the overall area dedicated to agricultural land use was kept constant within each grid cell (Fig. 2). As described above, each grid cell was converted proportionally according to the amount of productive grassland in each cell under BAU.

For CSM, simulating a change in management from permanent grassland to arable (PG2A), was achieved by converting grazing farm types to arable. Only LFA grazing model farms in the RLR area were considered for conversion to cereals. For converted farms, the same proportion of grassland (permanent grass and improved grass) as found on existing cereal farms was extrapolated to the farms undergoing conversion with similar numbers of sheep and lambs per grassland area included. The insignificant numbers of poultry birds on existing cereal farms were not included in the scenario. For converted farms, a similar proportion of crop types and areas to those on existing cereal farms were introduced to those farms being converted, with fertilizer application rates and manure spreading practices changed to represent those on current cereal farms in the upper River Taw study area. No bioenergy crops are currently included in CSM.

Scenarios were run on comparable time scales using weather data from 2004 to 2018 and were set to represent the “near present”. For RLM, SPACSYS and AGREMOSA outputs were aggregated to an annual time scale for each cell and for CSM, which runs on longer time frames, a single annual average across the catchment was produced for each output.

3. Results

3.1. Scenario components

3.1.1. Arable production

The simulated yields for winter wheat, winter and spring barley, field beans, oilseed rape and maize were similar across the three process-based models that simulated these outputs and across grid cells (Fig. 3). All cells were in the RLR and so ran over the same weather scenarios, but soil type varied between grid cells (Table S1). Although there was no substantial difference across the grid cells, grid cell 31 did stand out as somewhat different from the others, producing slightly larger yields in the AGREMOSA model and smaller within crop variation in SPACSYS. The cell is associated with the Neath soil series which is characterised as being slightly less heavy in texture compared with the other RLR grid cells (Table S1).

In terms of edible calories, the RLM consistently predicts a greater average level of production across the grid cells compared with AGREMOSA and SPACSYS (Fig. 4); however, the variation associated with this model is also far greater. This is because the RLM generates crop sequences stochastically with the constraint that the steady state accords with the observed crop proportions. Thus, production varies substantially compared with the other two models which use fixed proportions. The greater average production is driven by a greater diversity in cropping with higher yielding crops.

3.1.2. Grassland and livestock production

The temporary grassland system produced the highest amount of annual biomass compared to permanent grassland or rough grazing systems. The average annual biomass production for rough grazing was 2.8 t ha^{-1} whilst the average biomass production for permanent grassland and temporary grassland were 9.8 t ha^{-1} and 14.7 t ha^{-1}, an in-

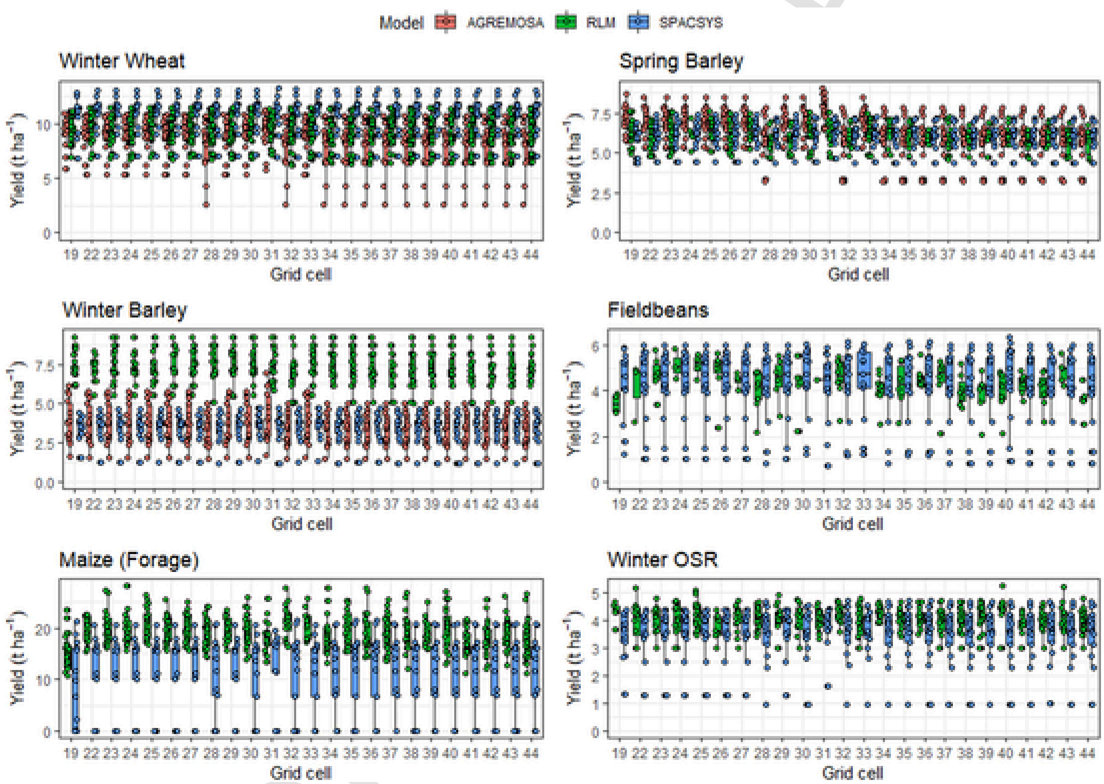

Fig. 3. Yields of common crops grown in the upper River Taw observatory simulated for harvest years 2006–2018 by AGREMOSA, RLM and SPACSYS for each grid cell. Note that the scale of ordinates varies between crops.

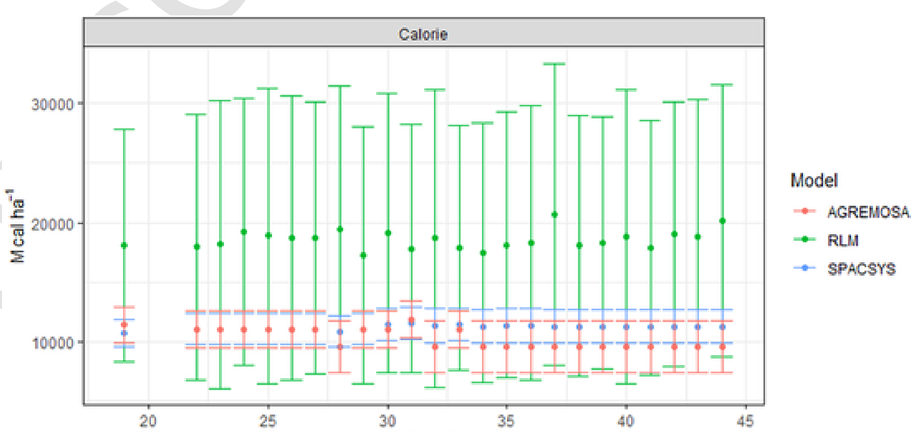

Fig. 4. Average annual production of edible calories from arable simulations from AGREMOSA, RLM and SPACSYS over all grid cells.

crease of 2.5 and 4.3 times, respectively, compared to rough grazing. Grid 31 produced the highest biomass because it had the highest plant available water of 174 mm in a soil depth of 1.25 m. The variation in the biomass production across the 40 grids depends on the weather and the soil C along with the plant available water in the soil profile. Livestock production is a linear function of stocking rate (Table S7), and therefore simulations of production generated using SPACSYS and RLM accord.

3.1.3. Bioenergy production

The average annual biomass production for *Miscanthus* across all 36 grid cells that were deemed suitable for bioenergy production, was 10.8 t ha^{-1} (Fig. 5). In the uplands, grid cells 9–18 and 20 produce the least biomass but more than rough grassland, with yields close to 8.0 t ha^{-1}. The productivity follows the temperature pattern across the study catchment, with the lowest average temperature (in grid cells 9–18 and 20 at 7.05 °C) contributing the least to production. The grid cells with highest yields fell in the RLR zone which has an average temperature of 8.26 °C (grid cells 19 and 21–44). The grid cells with the highest temperature were associated with the lowest rainfall and vice versa. Thus, it is evident that, in general, the water stress was not a factor in the study area and that the production was more limited by lower temperature translating to lower solar radiation which regulates the daily biomass accumulation by the crop.

The two cultivars of willow, Endurance, a broad leaf (BL) cultivar and Tora, a narrow leaf (NL) cultivar, are widely grown across the UK. A three-year rotation cycle was assumed for coppicing. In the uplands, in the uplands, the BL cultivar Endurance produced an average annual biomass of 11.3 t ha^{-1} (ranging from 10.8 t ha^{-1} in grid cell 9 to 14.4 t ha^{-1} in grid cell 19) which was about 18% higher than the average annual biomass of the NL cultivar Tora, which produced 9.8 t ha^{-1} (ranging from 9.1 t ha^{-1} in grid cell 9 to 12.2 t ha^{-1} in grid cell 19). Like *Miscanthus*, in the lowlands (for example grid cells 19 and 31) produced the highest biomass for both the cultivars (Fig. 5). The overall trend for willow production follows that of *Miscanthus* production in the study catchment.

The biomass production for the BL cultivar Endurance was higher than the economical yield in the region, which is taken as 9.0 t ha^{-1}. The old NL cultivar Tora is expected to produce economically viable yields (> 9.0 t ha^{-1}) (Lovett et al., 2009) in only 25 of the 36 grid cells. Comparing the two bioenergy crops, it appears that the *Miscanthus* is more productive overall in the lowlands (RLR zones) of the upper River Taw observatory where the temperature is slightly higher (19, 21–44) (see Fig. 5).

Miscanthus yields from SPACSYS were on average 2.55 t ha^{-1} lower than those simulated by AGREMOSA, but the between and within cell variation was similar for both models (see Fig. S2).

3.1.4. Carbon

The modelled soil carbon and carbon emissions (Fig. 6) largely accord between SPACSYS and RLM. Note the outliers from SPACSYS which were caused by initial high soil C and N content in G19, 20 and 21. We note that RLM does not account for CH_4 emissions from soil, whilst SPACSYS only accounts for CH_4 emissions from soil in non-arable land uses. Further, RLM does not account for CO_2 emissions from animals. Soil carbon is predicted to be largest under rough grazing and least under arable. Carbon emissions as CO_2 are substantially greater than those as methane, with largest total emissions associated with grazing.

CSM predicts catchment scale averages for methane and soil C under BAU to be 0.05 t ha^{-1} and 144.1 t ha^{-1}, respectively, with predictions under the PG2A scenario of 0.01 t ha^{-1} and 142.4 t ha^{-1}. Similar to SPACSYS and RLM, predictions show a reduction in soil C as more land is converted to arable, and a reduction in CH_4 emissions with arable conversion.

3.1.5. Nutrient losses

The predicted nutrient losses (N and P annual average t ha^{-1}) have substantial associated uncertainty. Figs. 7 and 8 show how the uncertainty in particular pathways partitions into temporal variability, spatial variability, various nutrient pathways, and model uncertainty. It is clear that assumptions about process pathways for soluble mineral N

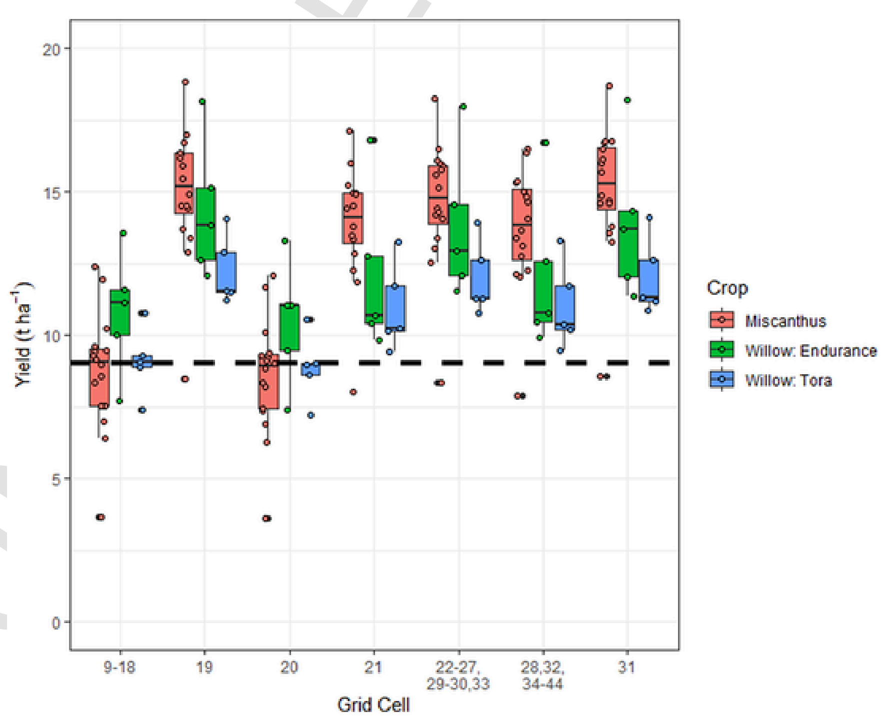

Fig. 5. Annual yields for bioenergy crops simulated using AGREMOSA for grid cells not allocated to high rainfall. These are: (i) grid cells 9–18, (ii) grid cell 19, (iii) grid cell 20, (iv) grid cell 21, (v) grid cells 22–27, 29–30, 33, (vi) grid cells 28, 32, 34–44) and (vii) grid cell 31. Annual willow yields were taken to be one third of the harvest yield at the end of the 3-year coppicing. The black dotted line shows the current economically viable yield for bioenergy crops (Lovett et al., 2009).

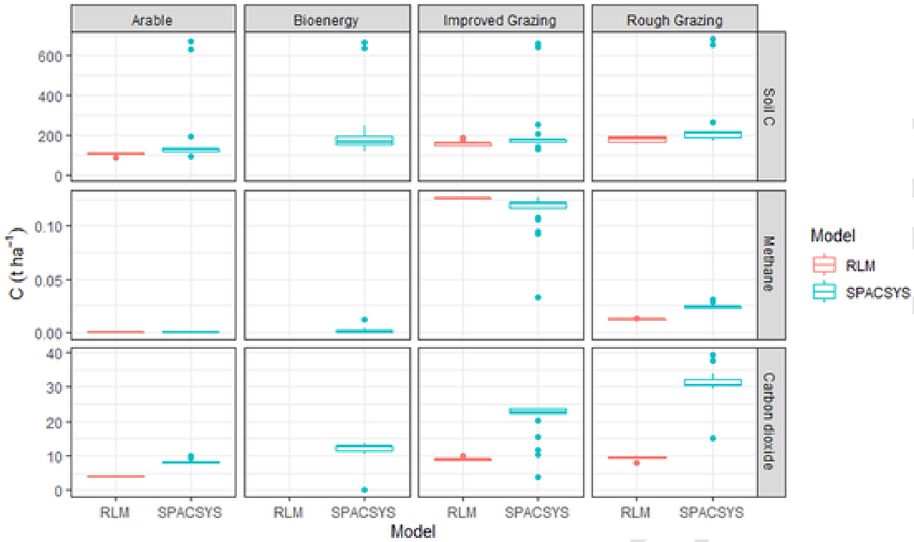

Fig. 6. Boxplots of the annual average of each carbon component listed (soil C, methane and carbon dioxide) for each grid cell × land use combination. Estimated soil carbon relates to the top 1 m of the soil.

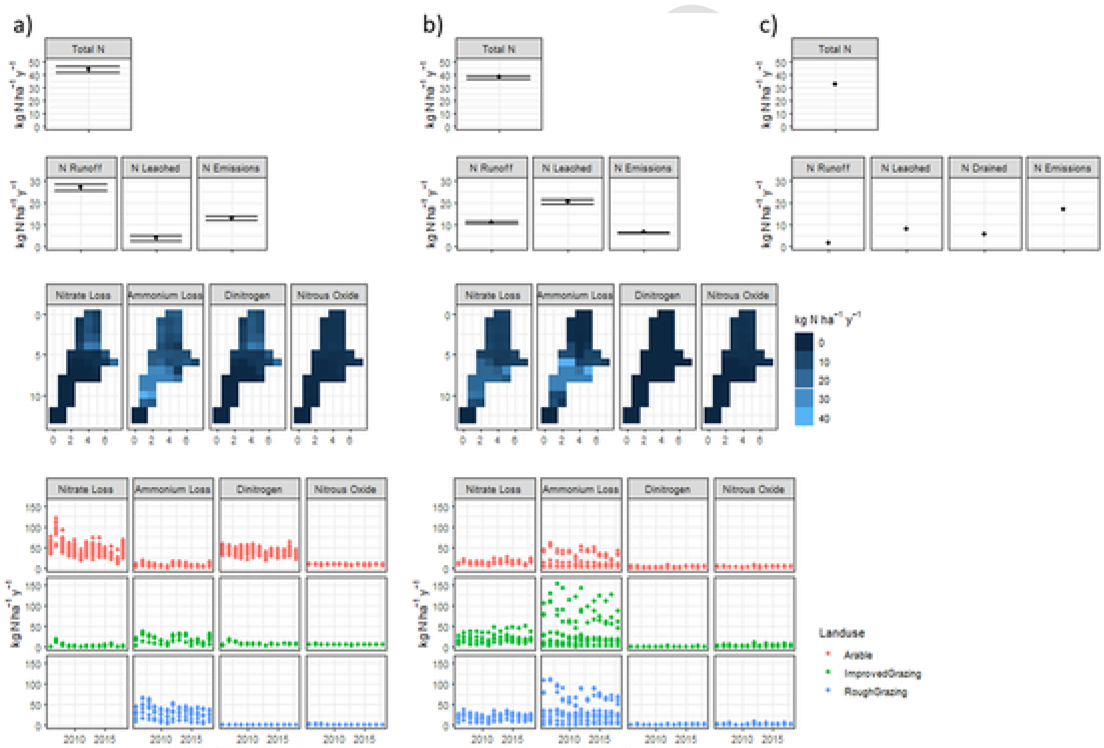

Fig. 7. Partitioning of uncertainty illustrated for the N loss pathways in the a) RLM, b) SPACSYS and c) CSM under the BAU scenario. The top row shows the total N lost per year and the second row shows the breakdown of this according to the N pathways. The third row shows how losses vary spatially and the fourth row shows temporal variation N losses according to land use.

are quite different between SPACSYS and RLM whereby SPACSYS allocates losses preferentially to leaching whereas RLM allocates losses preferentially to runoff (Fig. 7). This is partially due to the RLM accounting explicitly for topology in its water and nutrient flow processes and reflects the different nature of each agroecosystems model and how the partitioning differs in the different models. The spatial patterning of ammonium losses between the models is somewhat similar with larger losses associated with grazing areas; however, the models do not agree on nitrate losses, with RLM predicting greater losses associated with arable compared to grassland and SPACSYS vice versa. This adds greater uncertainty into the absolute predictions under the BAU scenario. Despite the differences between the predicted pathways for nutrient flows the total N losses agree well (Fig. 7). Similarly, P losses estimated by RLM and CSM accord (Fig. 8).

3.2. Ensemble

The ensemble model predictions of production, nutrient losses, C emissions and soil C, along with associated uncertainty (reported as standard deviation) are shown in Fig. 9. Production in terms of edible calories varies substantially from 6.44 T Cal to 27.34 T Cal across the entire study catchment. The range in outputs is largely driven by con-

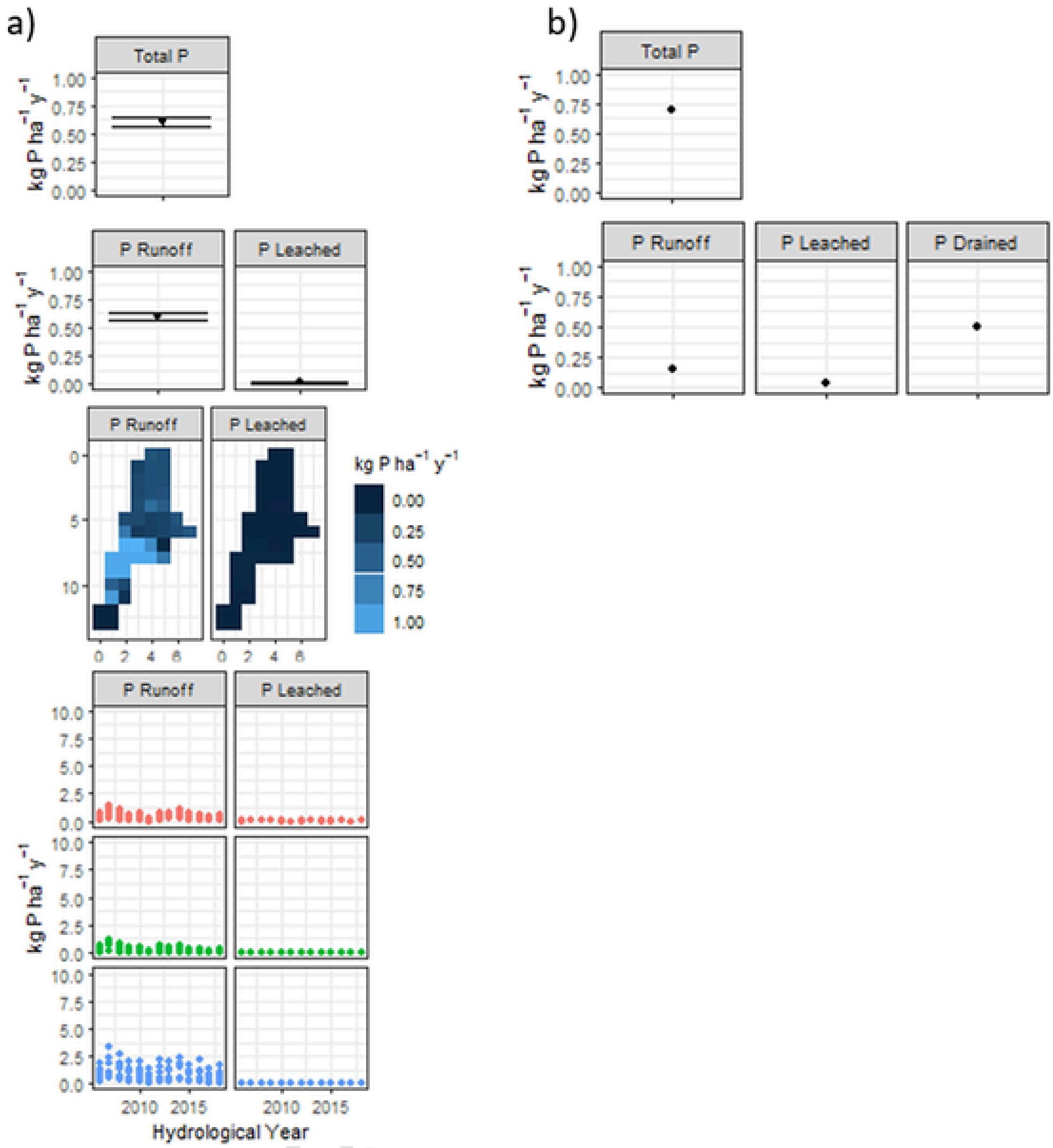

Fig. 8. Partitioning of uncertainty illustrated for the P loss pathways in the a) RLM, and b) CSM under the BAU scenario. The top row shows the total P lost per year and the second row shows the breakdown of this according to the P pathways. The third row shows how losses vary spatially and the fourth row shows temporal variation P losses according to land use.

version of grassland to arable (Fig. 9a). Associated uncertainty was greatest under PG2A (where the possible maximum conversion to arable has been implemented) due to the increased uncertainty in productivity within crop rotations over time. Bioenergy productivity increases as more land is converted to *Miscanthus*, which increases the associated uncertainty in production (Fig. 9).

Variation in predictions of soil C were largely related to conversion to arable with the lowest soil carbon observed in a complete arable conversion (0.558 Mt ± 0.036) and the highest in BAU (0.651 Mt ± 0.038). Conversely, the largest C emissions are associated with BAU and this is because emissions are associated with livestock land management in the form of CO_2 and CH_4. There are approximately equal reductions in C emissions due to both PG2A and PG2BE on an area basis. This, again, is directly related to the associated reduction in livestock. However, the reduction in uncertainty is markedly higher under bioenergy conversions (standard deviation ranges by approximately 0.042 Mt C) compared to arable conversions (standard deviation ranges by approximately 0.014 Mt C).

The losses of N are driven by a conversion to bioenergy production. Along the PG2A axis, average N (across the whole study catchment) decreases from 240.7 t N to 177.28 t N, whilst under complete bioenergy conversion, losses increase to 1150.36 t N. This is representative of the modelled SPACSYS outputs which predict much greater N losses under the bioenergy conversion particularly in the low rainfall moorland (Fig. S3). The increase in N losses as bioenergy production increases is associated with a marked increase in uncertainty, both temporally and spatially (Fig. S3).

Some caution is needed when interpreting the uncertainty (Fig. 9b) as the ensemble amalgamates different sources of variability across the scenario space. For example, C emissions decrease most substantially in the standard deviation along the increasing bioenergy axis. Thus, as more land is converted to bioenergy, predictions of C emissions become more certain. However, this is confounded with the fact that as more land is converted to bioenergy, contributions from multiple models decrease, with SPACSYS having the overriding influence. Thus, for these outputs, the bioenergy axis also represents fewer model contributions.

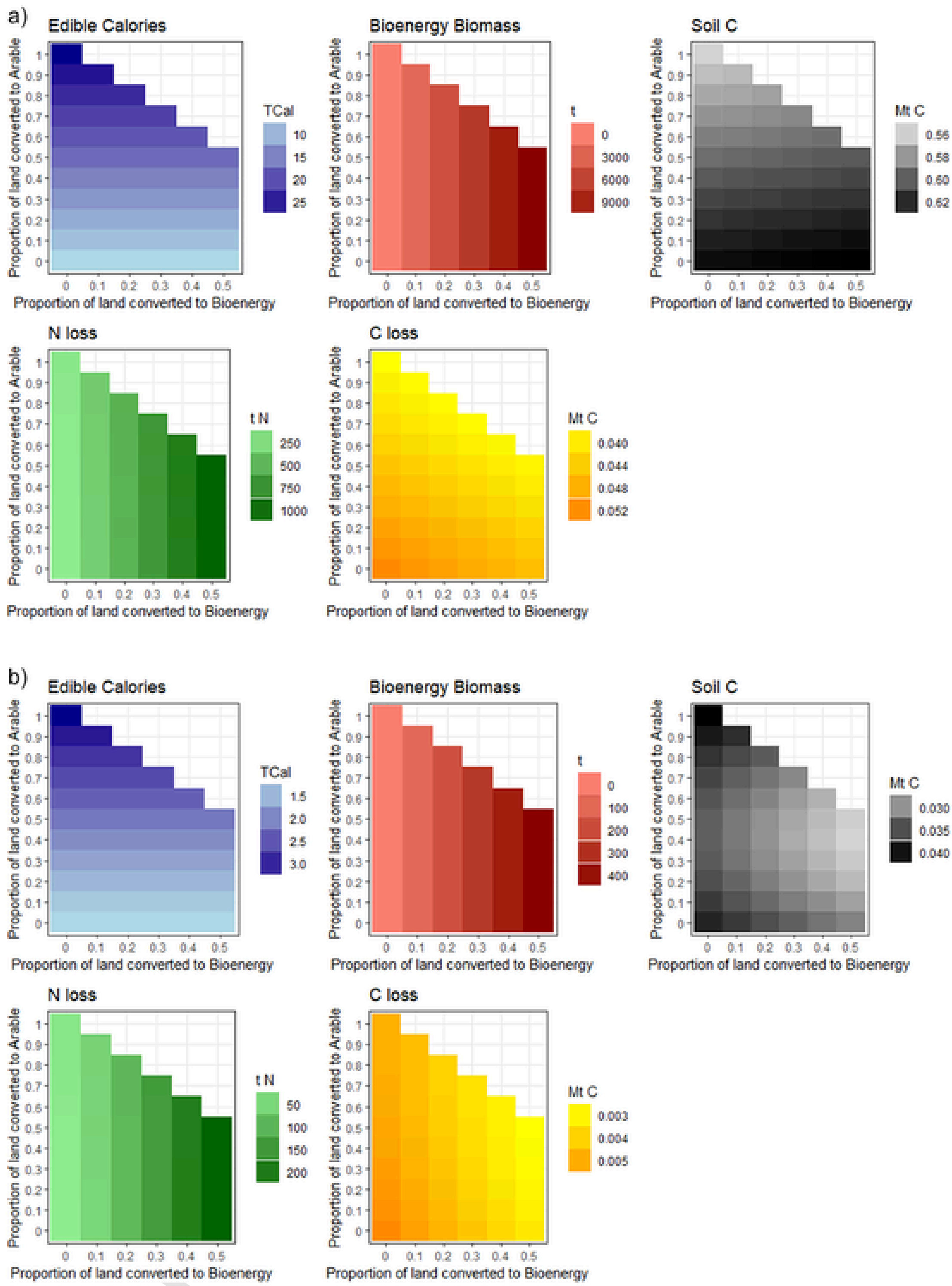

Fig. 9. Predicted nutrient losses, C emissions, soil C and productivity metrics under increasing arable and bioenergy conversion as simulated using the model ensemble. Measurements are the combined total over the 3427.5 ha of suitable land in the study catchment. Top panel (a) shows the ensembled average, Bottom panel (b) shows the ensembled standard deviation. Ensembled values under the changing scenarios are obtained from scaling the grid level ensemble of SPACSYS, RLM and AGREMOSA.

The fact that this coincides with a decrease in the standard deviation implies the inter-model uncertainty is greater than intra-model uncertainty. In comparison, the uncertainty associated with N increases as bioenergy production increases. This implies the intra-model uncertainty in SPACSYS increases along this axis and is indeed the case as the associated uncertainties with these predictions are largest under bioenergy production and so as more land is converted, more uncertainty is apparent. It is interesting to note that the uncertainty in soil C follows a different pattern. As with C emissions, a general trend of decreasing standard deviation can be observed along the x-axis and is associated with less inter-model uncertainty compared to intra-model. However, there is simultaneously, a changing trend in uncertainty as arable productivity increases. RLM predicts slightly less uncertainty in soil C under arable conversion from permanent grassland, whilst SPACSYS predicts a greater level of uncertainty. Given the overall uncertainty is given by the variance of the mean soil C (under each model) and the mean of the variance of soil C (under each model), the non-monotonic trend in uncertainty is explained by this partitioning.

The simulations produced by CSM are associated with the BAU state (where conversion to arable and bioenergy crops is equal to zero) and the state whereby the maximum feasible land has been converted to arable (proportion converted to arable equals one and proportion converted to bioenergy equals zero). The results from the ensemble of all our models are shown in Fig. 10. Predicted differences between BAU and P2G are given in Table 2 along with the standard error of the difference. Our predictions show a substantial increase in calories under the PG2A scenario but other changes in outputs for the study catchment are less distinct (with expected changes being within standard error bounds). Soil C is expected to be lower under PG2A, yet C emissions are also smaller. This is due to smaller CO_2 emissions being associated from the arable system (Fig. 6). Although not significant, nutrient losses are not substantially different between the two scenarios.

4. Discussion

4.1. Production

Across our models, production of arable crops (Fig. 3) agrees with that published by Defra (June Agriculture Surveys) and in the farming literature (BSPB, 2020; Cammarano et al., 2020; Nix, 2020). For wheat, spring barley, winter barley and winter OSR regional statistics for the period simulated report yields between 5.6–8.3 t ha^{-1}, 4.6–6.0 t ha^{-1}, 5.6–7.9 t ha^{-1} and 2.6–4.4 t ha^{-1}, respectively (see Table S11). For maize and fields beans (where no regional statistics are available) we found our simulations compared well with national averages of dry matter of 17 t ha^{-1} and 4.2 t ha^{-1}, respectively. Similarly, the simulated grasslands biomass production in this study accords with the biomass productions reported for temporary and permanent grasslands in Scotland by Jones (2013) who found that a young grazing ley, similar to temporary grassland, is capable of producing 12–14 t ha^{-1} dry matter per year, whilst permanent pasture can produce 9–10 t ha^{-1} dry matter per year. However, these levels of production can only be sustained in soils that have adequate water, nutrient inputs and reserves and appropriate grazing management.

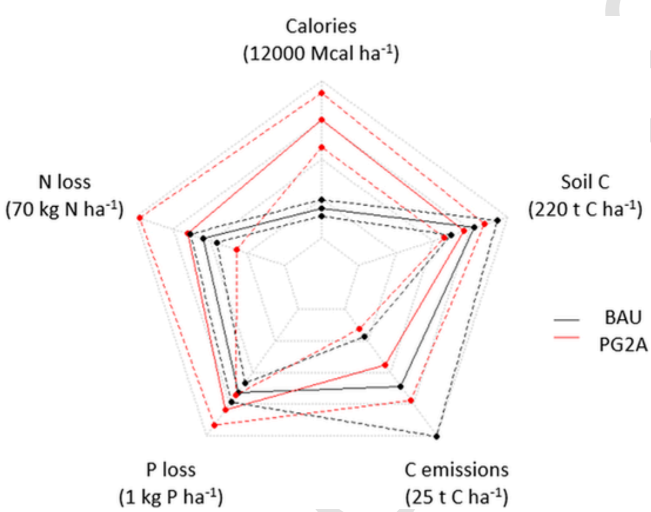

Fig. 10. Predicted change of the catchment level outputs from BAU to PG2A using the full model ensemble. Solid lines are the mean and dashed show ±standard deviation. The upper bounds of each axis are given in brackets followed by units. The lower bound is zero in all cases.

Table 2
Predicted changes from BAU to PG2A across the study catchment.

	Production/ Mcal ha^{-1}	Gross margin/£ ha^{-1}	Soil C/t C ha^{-1}	C emissions/ t C ha^{-1}	N loss/ kg N ha^{-1}	P loss/kg N ha^{-1}
Expected change	6735.9	76.16	−14.48	−4.32	7.21	0.14
Standard deviation	2152.8	66.24	45.11	12.06	23.53	0.13

Although the biomass production of permanent grassland is less than temporary grassland, it should be emphasized that the cost of production is also less. It also has a denser sward and can carry more stock. In this study, whilst simulated livestock removed biomass and contributed to nutrient flows and emissions, the simulated biomass did not dictate stocking rates; rather, we used stocking rates derived directly from commercial farm data (Nix, 2020). These data reflect the stocking capacity of various types of grass land. However, looking at the impact of the biomass production on livestock, every 1 t ha^{-1} dry matter increase in grass equates to a potential increase in stocking rate of 1.4 ewes per hectare or 100 kg of beef live weight gain ha^{-1} year^{-1} (Jones, 2013). This means that compared to rough grazing, there could be an increase of 9.8 ewes ha^{-1} in stocking rate or 700 kg of beef live weight gain ha^{-1} year^{-1} in permanent improved grassland and an increase to 16.8 ewes ha^{-1} in stocking rate and 1200 kg of beef live weight ha^{-1} year^{-1} in temporary grassland.

The simulated average annual *Miscanthus* biomass related very well with reported data in the literature. Production estimates of 10.83 t ha^{-1} obtained from AGREMOSA and 8.30 t ha^{-1} from SPACSYS compared well to the average annual biomass of 8.94 t ha^{-1} from on-farm observations across the UK (Richter et al., 2016). Site-specific biological yield potentials estimated using an empirical model ranged from 5 to 15 t ha^{-1} with an average of 11 t ha^{-1} (Richter et al., 2016). Overall, our simulations ranged from 7.87 to 14.49 t ha^{-1} with the lower yield from the moorland grid cells with lower average temperatures. *Miscanthus* yield from 14 experimental arable sites across the UK ranged from 5 to 18 t ha^{-1}, averaging 12.8 (±2.9) t ha^{-1}. Our predictions for the study catchment are well within the range of those results and mostly well above the economic threshold of 9 t ha^{-1} (Lovett et al., 2009).

The average annual willow biomass production in south west England was reported by Richard et al. (2019) to be 11 and 10.3 t ha^{-1} for the Endurance and Tora cultivars, respectively. Whilst our results matched the observed annual biomass production of Endurance they slightly underestimated the observed annual biomass production of Tora (9.4 vs 10.3 t ha^{-1}). We found that the grid cells with lower temperatures (low rainfall moorland) produced lower annual biomass. This is also reported by Richard et al. (2019) whose simulations suggested that warmer climate increased the average productivity of willow by 0.5–2.5 t ha^{-1} depending on the cultivar.

Looking at the overall production it appears that the temporary grasslands and *Miscanthus* produce the highest biomass followed by willow-SRC and arable crops (see Fig. S4). The grass biomass also has potential for biogas production and 20 million tonnes of biomass could provide up to 12.5% of the total gas output or 25% of the gas imports to the UK in 2017 assuming standard conversion rates from grass biomass to biogas (Qi et al., 2018). The Miscanthus and willow could provide the energy of 19.6 MJ kg^{-1} and 19.2 MJ kg^{-1} translating to 4.68 Mcal kg^{-1} and 4.58 Mcal kg^{-1}, respectively, in terms of their calorific value of combustion (Daraban et al., 2015; Piskier, 2017). It can be observed that the biomass production for *Miscanthus* and willow-SRC is lower in the uplands grid cells with lower temperatures and higher rainfall grid cells 9–18 and 20 compared to potentially improved grassland production. This suggests that these areas should not be converted into bioenergy crops as that would mean overall loss in production. On the other hand, the lowlands (RLR grid cells) with comparatively higher temperatures (even with lower rainfall) grid 19 and grid cells 21–44, are good candidates for converting into arable (based on edible calories produced; Fig. 10) or bioenergy crops, preferably Miscanthus, with slight biomass losses compared to temporary grassland. The economic competitiveness for BE crops comes from low cultivation and fertilizer inputs after establishment (McCalmont et al., 2017).

It must be emphasized that our models of production are not without limitations. For example, they do not consider waterlogging which could have an impact on production in the areas of higher rainfall (low and high rainfall moorland zones) especially under changing climate (Harkness et al., 2020; Ploschuk et al., 2018). Even the lowest rainfall in our simulations is 1200 mm which could be sufficient to cause waterlogging and in fact, most soil series across the study catchment are highly prone to seasonal waterlogging given their clay content (Pulley and Collins, 2020). Our models also consider that threats to crop production, such as weeds and diseases, are well managed, and that either crops are non-nutrient limited (in the case of AGREMOSA) or have fertilizer managed according to best practice recommendations (in the case of the other models). In practice, farmers may choose to under fertilize resulting in nutrient limitations or over fertilize resulting in greater nutrient losses. For example, Qi et al. (2017) reported that when the annual N usage on grassland dropped to about 99 and 52 kg N ha^{-1} on temporary and permanent grassland, respectively, much below the recommended respective economic optimums of 300 kg and 150 N ha^{-1} (Hopkins et al., 1990; Morrison et al., 1980), it resulted in concomitant yields of about 45 and 39% below the respective attainable dry matter yields.

4.2. Carbon

Defra (2010c) report soil C stocks in the top 15 cm of soil to be in the region of 58–120 t ha^{-1} for UK arable systems, 100–150 t ha^{-1} for permanent pasture and for rough grazing estimates are in the region of 210–230 t ha^{-1}. Estimates from our models are associated with approximately the top 1 m of the soil but accord well with these numbers, with few outliers (Fig. 6). Predicted differences in soil C as land is converted to *Miscanthus* are small (Figs. 6 and 9) and this accords with the findings of Zatta et al. (2014) who found an insignificant change in soil C when UK grassland was converted to *Miscanthus*.

Predictions of methane from the models are largely driven by livestock numbers. SPACSYS estimates animal emissions based on feed intake and RLM uses IPCC Tier 2 models to estimate livestock emissions (Milne et al., 2015) and so observed differences in emissions result from varying interpretation of stocking rates.

Oertel et al. (2016) measured an average CO_2 emission from cropland of 6.7 t C ha^{-1} $year^{-1}$ from 41 studies (range 3.9–69.5 t C ha^{-1} $year^{-1}$). These are similar to the modelled values for arable of 5.0 and 8.0 t C ha^{-1} $year^{-1}$, for the RLM and SPACSYS, respectively. For grassland, an average of 8.5 t C ha^{-1} $year^{-1}$ was observed from 47 studies (range 1.8–30.0 t C ha^{-1} $year^{-1}$) and, again, these were similar to the modelled results from the RLM (15 t C ha^{-1} $year^{-1}$ in improved grassland, and 10 t C ha^{-1} $year^{-1}$ in rough grazing) and from SPACSYS (17.5 t C ha^{-1} $year^{-1}$ in improved grazing and 30 t C ha^{-1} $year^{-1}$ in rough grazing).

4.3. Nutrient losses

Predicting nutrient losses and emissions is notoriously complex and uncertain given the partitioning between dissolved and particulate forms and the complexities of the source-mobilisation-delivery transfer continuum (Lloyd et al., 2019). Evidence is emerging that processes other than nitrification and denitrification are far more important than previously assumed for gaseous N production from soils (van Groenigen et al., 2015). Processes such as nitrifier denitrification (Wrage et al., 2001), in situ N_2O reduction (Schlesinger, 2013), anammox (Mulder et al., 1995), Feammox (Sawayama, 2006), dissimilatory nitrate reduction to ammonium (DNRA) (Tiedje, 1988), and co-denitrification (Spott et al., 2011) are highlighted in current literature, but information on process rates and their dynamics in response to environmental factors is scant.

All of our models have been shown to give good predictions of certain aspects of the P and N cycles. For example SPACSYS has been validated against measurements of N_2O and crop N offtake (Wu et al., 2015) and RLM against measurements of grain N and P, and N leached (Coleman et al., 2017). National scale nitrous oxide and methane emissions from agriculture predicted using the calculations embedded in CSM have been found to agree strongly with the UK GHG inventory dataset (Zhang et al., 2017).

In this study, our modelled losses of P were consistent between models. In the case of phase partitioning, P losses are typically strongly controlled by erosion and thereby the source-mobilisation-delivery cascade for fine-grained sediment, meaning that models must include erosion and sediment dynamics to simulate total P losses representatively. This accounts for the inclusion of PSYCHIC predictions for both sediment and P in the CSM framework which have been evaluated at field scale both in the study catchment on the North Wyke Farm Platform (Collins et al., 2021) and nationally (Collins et al., 2008) and more strategically across England and Wales at catchment scale (Collins and Anthony, 2008; Collins et al., 2009; Zhang et al., 2017). Erosion processes are extremely variable both in time and space, raising ongoing challenges for modelling this aspect of agroecosystems (Evans et al., 2016, 2017).

Predicted total N losses were similar across the three models; however, the predicted dominant pathway for N losses varied between models. For example, RLM predicted the largest losses to be in the form of nitrate whereas SPACSYS attributed most losses to be in the form of ammonium. For RLM the dominant pathway was runoff, for SPACSYS leaching and CSM emissions. These apparent contradictions reflect the complexity of the processes considered, the fact that models make different assumptions about the processes and delivery pathways involved, and that the models were originally parameterised using data from other locations in the UK (Coleman et al., 2017; Stromqvist et al., 2008; Wu et al., 2016). Nitrate losses to water using the process-based representation in CSM have been shown to agree well with national monitoring data (1980–2010) collected at 33 monitoring stations as part of the Harmonized Monitoring Scheme, yielding a Kling-Gupta efficiency of 0.65 (Zhang et al., 2017). Nutrient losses depend strongly on soil physical properties. For example, estimated losses of nitrate and ammonium from grid cell19 from SPACSYS showed larger values than adjacent grids because soil texture for the cell 19 is sandy loam that has lower bulk density and higher hydraulic conductivity, which cause nutrients to move out quickly. Under BAU, average losses were predicted to be 31.64 kg N ha^{-1} $year^{-1}$, 30.98 kg N ha^{-1} $year^{-1}$ and 15.19 kg N ha^{-1} $year^{-1}$ from SPACSYS, RLM and CSM, respectively. These predictions fall within the range observed from the 15 experimental catchment flumes at the North Wyke Farm Platform, part of which is in the catchment, but are high compared with the observed average (6.7 kg N ha^{-1} $year^{-1}$ with a range of the measured data 0.0–81.9 kg N ha^{-1} $year^{-1}$).

In relation to bioenergy, N losses were predicted to be large in comparison to arable and livestock management (Fig. 9a). This was surprising as nutrient losses are usually reported to be low from low-input biomass crops like Miscanthus (Davis et al., 2015; Ferchaud et al., 2020) and SRC-willow (Aronsson and Bergström, 2001; Dimitriou and Aronsson, 2004). Indeed, due to their high nutrient uptake bioenergy crops (e.g. willow) are used for purifying sewage water (Dimitriou et al., 2012). Wastewater or sewage sludge applications to willow SRC in newly established fields is practiced to achieve a more balanced fertilizer and to recirculate nutrients contained in sewage sludge (phosphorus and nitrogen) to agricultural soils (Dimitriou and Rosenqvist, 2011). Exploring our simulation results further revealed that the particularly large losses are associated with the high rainfall moorland grid cells where the growth of *Miscanthus* is predicted to be poor (Fig. 7). This would limit the potential to take up the soil nutrients and this in combination with the higher rainfall than the RLR portion of the study catchment are the primary reasons that nutrient losses are predicted to increase substantially. This has been reported in relation to extraction of forest biomass (de Oliveira Garcia et al., 2018; Paré and Thiffault, 2016). In practice it is likely this crop would not be grown in these areas. In addition, our predicted increases in nutrient loss with conversion to bioenergy are associated with large increases in uncertainty and highlight a need for experimental evidence to parameterise the nutrient flows associated with bioenergy crops grown on moorland.

4.4. Trade-offs between objectives

The results of the ensemble modelling approach reveal well-known and expected trade-offs. For example, by increasing the area associated with arable production we increase the production of edible calories, yet the soil C under arable conversion is predicted to be lower than under BAU. This is to be expected because permanent grassland tends to have a larger amounts of soil C than arable soils. The lower carbon in arable soils are due to lower C inputs from arable systems, disturbance of soil structure and aggregates, removal of the grain/seed/tubers, straw and other residues (Johnston et al., 2009; Smith et al., 2005). As expected, reducing livestock areas, either through conversion to arable or bioenergy, reduces C emissions. Conversion to arable is predicted to slightly increase expected nutrient losses, although uncertainties associated with these predictions are substantial (Fig. 10). We note a more surprising trade-off associated with our predictions related to conversion to bioenergy. This is that increasing bioenergy results in increasing nutrient losses as discussed above. This is associated with a marked increase in uncertainty (Figs. 9b and S4) highlighting a need for more data relating to N losses associated with bioenergy crops in moorland areas.

4.5. The ensemble approach

Model ensembles are now commonplace in atmospheric sciences (Gneiting and Raftery, 2005) and climate/global change impact assessment using crop growth models (Rodríguez et al., 2019; Wallach et al., 2018). The widely accepted advantage of this approach is that structural uncertainty, as well as input uncertainty, are accounted for explicitly. In our study we combined four models in an ensemble to explore questions around land management change. Input uncertainty was restricted to stochastic combinations of weather and crop, and the various interpretations of the available input data for the study area (see below). The ensemble approach was not without a combination of advantages and challenges, some of which are common to atmospheric sciences.

In our case, apart from the clear advantage of capturing uncertainty in the understanding of complex processes, the model ensemble also allowed us to compare a broader range of objectives than any one of the models would permit individually. Here, for example, AGREMOSA includes bioenergy crops whereas the crop libraries in CSM and RLM do not. For all but one of the outputs (gross margin; CSM only), at least two models contributed to the predicted outcome. A second clear advantage is that the model ensemble allows us to identify information weak spots. This accords with the findings of Willcock et al. (2020) who, similar to us, concluded that ensemble variation is often a proxy for lack of accuracy and not simply a measure of precision. Here, the largest relative uncertainties and indeed model conflicts were related to the understanding of nitrogen losses. This is perhaps not surprising given the complexity of the cycling, mobilisation and delivery processes involved but certainly an advantage of our approach is that it openly challenges scientists' model-based assumptions identifying areas where further data should be sought. We defer to the quote by George E.P. Box "All models are wrong, but some are useful". During our analysis at least two model assumptions were found to lead to counterintuitive outcomes that required further investigation and new development. In most cases, however, there was insufficient empirical evidence to support changes and so differences in model predictions were attributed to structural uncertainty.

Although ensemble modelling has been widely adopted in atmospheric science (Suarez-Gutierrez et al., 2021), and (to some extent) in crop yield prediction (Yin et al., 2017) and soil organic carbon modelling (Farina et al., 2021; Riggers et al., 2019; Smith et al., 1997) to the best of our knowledge, this trend has not manifested in relation to agroecosystems, and those that do exist, tend to use ensemble predictions as inputs to the system through predictions of climate or rainfall (Dale et al., 2017; Georgakakos and Carpenter, 2006). Crucially, none have attempted to make an ensemble of predictions of scenarios for change in agroecosystems. An obvious challenge here is model availability and complexity. We had access to models from four modelling groups, but the challenges associated with aligning data inputs were notable. Aligning model outputs to accord on units, temporal and spatial scale, and surprisingly even meaning, was not straightforward and required detailed discussions. This is demonstrated in Fig. S1, whereby models feed into the different components of the agroecosystems at different levels of derivation. For example, under P loss, CSM contributes discretely through both dissolved and particulate P. For approaches such as the one described here to be implemented more widely, model and data availability must be improved and standardisation of terminology would also greatly facilitate the ensemble process. Terminology is an increasing issue in the data-driven sciences. As elsewhere, (Arnaud et al., 2020) nutrient cycling would benefit from a universally accepted ontology and dictionary of terms to ensure that all scientists mean the same thing by the terms they use and describe results in an identical manner or at least one where the rules for interconversion are objective and clear.

It is the view of the authors that model variability should not be restricted to the structural components or quantitative outputs from simulations but, rather, should also include data integration. All four modelling groups aligned simulations based on available input and output data from the upper River Taw observatory. However, the ways in which input data were implemented varied greatly. Such variation in the use of input data has been reported previously (Silgram et al., 2009). For example, crop rotations were accounted for explicitly in RLM, albeit derived from data at a larger spatial resolution. SPACSYS and AGREMOSA accounted for rotations in the aggregation of model outputs, although each model included different crop types and CSM generated temporally-averaged outputs for the mix of crops and associated best management practices on the key farm types in the study area. Similarly, livestock stocking rates were implemented differently, with SPACSYS assigning a set number of livestock to each grid cell, RLM fixing an overall stocking rate, each with differing livestock categories compared to those available in the data and CSM using census and additional farm survey data to establish the animal population details and stocking densities per commercial farm in the study area. Consequently, each modelling group needed to make their own assumptions on how best to align the study catchment data with their required inputs. Perhaps more fundamental, are the differences in underlying environmental inputs, such as weather data and soil type as described in Section 2.2 and associated supplementary tables. Such differences in data interpretation inhibit large simulation studies whereby a fixed set of input variables are varied to then observe the change in output variables. It is our belief that input alignment comes best from observed data allowing this flexibility in interpretation. Thus, to proceed at scale, data need to be readily available for a range of diverse study catchments.

5. Conclusion

We developed an ensemble approach for agroecosystem models of different complexity and used this to explore scenarios for a high rainfall catchment in SW-England. Considering outputs for edible calories, biomass, and carbon and nutrient losses we discussed sustainable options of grassland conversion to arable and bioenergy. In terms of productivity and impacts on carbon and nutrient flows we conclude that:

- In upland areas it was viable to convert rough grassland to SRC-willow but not to *Miscanthus*. In lowlands both types of bioenergy crop were viable, with *Miscanthus* producing greater biomass. Converting grassland in lowland areas to arable resulted in greater production of edible calories.
- Across the catchment, differences in soil organic C are likely to be small when land is converted from grassland to *Miscanthus* (or SRC-willow); losses associated with conversion to arable were large, although uncertainties associated with these predictions were also large.
- Nutrient losses associated with conversion from improved grass to arable are large, but generating trade-offs in reduced GHG emissions

Regarding output uncertainty, model complexity and data and process limitation, we argue that:

- Model ensembles that account explicitly for inter- and intra-model variability are particularly valuable for assessing and interpreting land use and management scenarios. In general, agroecosystems models are deterministic and residual background variability is often ignored.
- By combining our models, we obtained independent replicate simulations that, in part, capture such residual error. In addition, the multiple models also capture uncertainties in the modelling frameworks.
- The model ensemble framework provides a robust mechanism for propagating uncertainty at different scales and enables us to identify weak spots in the different agroecosystem models.

Overall, model ensembles in the agroecosystem context are challenging to implement in practice. With only a relatively small number of models available it is an open question as to what to do when models disagree. Here, emerging questions include: Should a more focussed attempt be made to align model outputs? Should the individual processes be adapted? Should the differences be accepted and interpreted as uncertainty? Here, we have taken all three avenues at various points but note that limited empirical data have been available for comprehensive validation. Future work will have to address challenges, to extend and adapt our approach to different agroecosystem geographies.

Uncited reference

Kaipainen et al., 2004

CRediT authorship contribution statement

Kirsty L. Hassall: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision. **Kevin Coleman:** Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization. **Prakash N. Dixit:** Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Steve J. Granger:** Investigation, Data curation. **Yusheng Zhang:** Methodology, Software, Validation, Investigation, Data curation. **Ryan T. Sharp:** Methodology, Software, Visualization. **Lianhai Wu:** Methodology, Software, Validation, Investigation, Data curation, Writing – original draft, Writing – review & editing, Supervision. **Andrew P. Whitmore:** Conceptualization, Methodology, Validation, Resources, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Goetz M. Richter:** Methodology, Software, Validation, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review & editing, Supervision. **Adrian L. Collins:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Alice E. Milne:** Conceptualization, Methodology, Software, Validation, Formal analysis, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Rothamsted Research receives strategic support from the UKRI-BBSRC (UK Research and Innovation-Biotechnology and Biological Sciences Research Council. This research was supported by the Institute Strategic Programme (ISP) grant "Soils to Nutrition" (S2N) grant number BBS/E/C/000I0330. Met data for North Wyke was supplied by the e-RA database, part of the Rothamsted Long-term Experiments National Capability (LTE-NC), supported by the UK-BBSRC (BBS/E/C/000J0300) and the Lawes Agriculture Trust. The North Wyke Farm Platform Is a UK National Capability funded by grant award BBS/E/C/000J0100. Local farmers are thanked for completing survey returns.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2022.153824.

References

Aarhus University, 2020. TSARA (Targets for Sustainable And Resilient Agriculture) Final Report.

Agostini, F., Gregory, A.S., Richter, G.M., 2015. Carbon sequestration by perennial energy crops: is the jury still out? Bioenergy Res. 8, 1057–1080. https://doi.org/10.1007/s12155-014-9571-0.

Arnaud, E., Laporte, M.A., Kim, S., Aubert, C., Leonelli, S., Miro, B., et al., 2020. The ontologies community of practice: a CGIAR initiative for big data in agrifood systems. Patterns 1. https://doi.org/10.1016/j.patter.2020.100105.

Aronsson, P.G., Bergström, L.F., 2001. Nitrate leaching from lysimeter-grown short-rotation willow coppice in relation to N-application, irrigation and soil type. Biomass Bioenergy 21, 155–164. https://doi.org/10.1016/S0961-9534(01)00022-8.

Aryal, Y., Zhu, J., 2020. Multimodel ensemble projection of meteorological drought scenarios and connection with climate based on spectral analysis. Int. J. Climatol. 40, 3360–3379. https://doi.org/10.1002/joc.6402.

Bingham, I.J., Wu, L., 2011. Simulation of wheat growth using the 3D root architecture model SPACSYS: validation and sensitivity analysis. Eur. J. Agron. 34, 181–189. https://doi.org/10.1016/j.eja.2011.01.003.

Bradley, A.V., Rosa, I.M.D., Brandão, A., Crema, S., Dobler, C., Moulds, S., et al., 2017. An ensemble of spatially explicit land-cover model projections: prospects and challenges to retrospectively evaluate deforestation policy. Model.Earth Syst. Environ. 3, 1215–1228. https://doi.org/10.1007/s40808-017-0376-y.

BSPB, 2020. British Society of Plant Breeder's 2021 Forage Maize Descriptive Lists. http://www.plantbreedingmatters.com/sg_userfiles/BSPB_2021_Forage_Maize_Descriptive_Lists.pdf (accessed 11th Jan 2022).

Cammarano, D., Holland, J., Ronga, D., 2020. Spatial and temporal variability of spring barley yield and quality quantified by crop simulation model. Agronomy-Basel 10. https://doi.org/10.3390/agronomy10030393.

Carswell, A.M., Gongadze, K., Misselbrook, T.H., Wu, L., 2019. Impact of transition from permanent pasture to new swards on the nitrogen use efficiency, nitrogen and carbon budgets of beef and sheep production. Agric. Ecosyst. Environ. 283, 106572. https://doi.org/10.1016/j.agee.2019.106572.

CEH, 2017. Integrated Hydrological Digital Terrain Model [IHDTM]. https://catalogue.ceh.ac.uk/documents/242384d6-ce65-4360-bf4e-3f6b4ed53034 (accessed 11th Jan 2022).
Cerasuolo, M., Richter, G.M., Richard, B., Cunniff, J., Girbau, S., Shield, I., et al., 2016. Development of a sink-source interaction model for the growth of short-rotation coppice willow and in silico exploration of genotypexenvironment effects. J. Exp. Bot. 67, 961–977. https://doi.org/10.1093/jxb/erv507.
Chadwick, D., Misselbrook, T., Gilhespy, S., Williams, J., Bhogal, A., Sagoo, L., et al., 2005. Ammonia emissions and crop nitrogen use efficiency: ammonia emissions from nitrogen fertiliser applications to grassland and tillage land; Factors affecting ammonia emissions from urea-based fertilisers; and Ammonia emissions model. In: Final Report for Defra Project NT2605. p. 71.
Coleman, K., Jenkinson, D.S., 2014. RothC - a model for the turnover of carbon in soil: model description and users guide (updated June 2014). Lawes Agricultural Trust, Harpenden, UK.
Coleman, K., Muhammed, S.E., Milne, A.E., Todman, L.C., Dailey, A.G., Glendining, M.J., et al., 2017. The landscape model: a model for exploring trade-offs between agricultural production and the environment. Sci. Total Environ. 609, 1483–1499. https://doi.org/10.1016/j.scitotenv.2017.07.193.
Coleman, K., Whitmore, A.P., Hassall, K.L., Shield, I., Semenov, M.A., Dobermann, A., et al., 2021. The potential for soybean to diversify the production of plant-based protein in the UK. Sci. Total Environ. 767, 144903. https://doi.org/10.1016/j.scitotenv.2020.144903.
Collins, A.L., Anthony, S.G., 2008. Assessing the likelihood of catchments across England and Wales meeting 'good ecological status' due to sediment contributions from agricultural sources. Environ. Sci. Pol. 11, 163–170. https://doi.org/10.1016/j.envsci.2007.07.008.
Collins, A.L., Stromqvist, J., Davison, P.S., Lord, E.I., 2007. Appraisal of phosphorus and sediment transfer in three pilot areas identified for the catchment sensitive farming initiative in England: application of the prototype PSYCHIC model. Soil Use Manag. 23, 117–132. https://doi.org/10.1111/j.1475-2743.2007.00119.x.
Collins, A.L., Duethmann, D., Taylor, P., 2008. National-scale sediment delivery to watercourses across England and Wales under recent (1970-2004) land use change. In: Schmidt, J., Cochrane, T., Phillips, C., Elliott, S., Davies, T., Basher, L. (Eds.), Sediment Dynamics in Changing Environments. International Association of Hydrological Sciences Publication No. 325, Wallingford, UK, pp. 448–452.
Collins, A.L., Anthony, S.G., Hawley, J., Turner, T., 2009. The potential impact of projected change in farming by 2015 on the importance of the agricultural sector as a sediment source in England and Wales. Catena 79, 243–250. https://doi.org/10.1016/j.catena.2009.05.007.
Collins, A.L., Zhang, Y., Upadhayay, H.R., Pulley, S., Granger, S.J., Harris, P., et al., 2021. Current advisory interventions for grazing ruminant farming cannot close exceedance of modern background sediment loss - assessment using an instrumented farm platform and modelled scaling out. Environ. Sci. Pol. 116, 114–127. https://doi.org/10.1016/j.envsci.2020.11.004.
Comber, A., Proctor, C., Anthony, S., 2008. The creation of a National Agricultural Land use Dataset: combining pycnophylactic interpolation with dasymetric mapping techniques. Trans. GIS 12, 775–791. https://doi.org/10.1111/j.1467-9671.2008.01130.x.
Crome, 2019. Crop map for England. https://data.gov.uk/dataset/8c5b635f-9b23-4f32-b12a-c080e3f455d0/crop-map-of-england-crome-2019 accessed 11th Jan 2022.
Dale, A., Fant, C., Strzepek, K., Lickley, M., Solomon, S., 2017. Climate model uncertainty in impact assessments for agriculture: a multi-ensemble case study on maize in Sub-Saharan Africa. Earth's Future 5, 337–353. https://doi.org/10.1002/2017EF000539.
Daraban, A., Jurcoane, S., I, V., 2015. Miscanthus giganteus - an overview about sustainable energy resource for household and small farms heating systems. 20, 10369–10380 (accessed 11th Jan 2022).
Davis, M.P., David, M.B., Voigt, T.B., Mitchell, C.A., 2015. Effect of nitrogen addition on miscanthus × giganteus yield, nitrogen losses, and soil organic matter across five sites. GCB Bioenergy 7, 1222–1231. https://doi.org/10.1111/gcbb.12217.
Davison, P.S., Withers, P.J.A., Lord, E.I., Betson, M.J., Stroemqvist, J., 2008. PSYCHIC - a process-based model of phosphorus and sediment mobilisation and delivery within agricultural catchments. Part 1: model description and parameterisation. J. Hydrol. 350, 290–302. https://doi.org/10.1016/j.jhydrol.2007.10.036.
de Oliveira Garcia, W., Amann, T., Hartmann, J., 2018. Increasing biomass demand enlarges negative forest nutrient budget areas in wood export regions. Sci. Rep. 8, 5280. https://doi.org/10.1038/s41598-018-22728-5.
Defra, 2010. In: Definitions of Terms Used in Farm Business Management, 3rd ed., DEFRA, London, p. 48.
Defra, 2010. Fertiliser Manual, 8th edition. DEFRA, London.
Defra, 2010. Project SP0567: Assembling UK wide data on soil carbon (and greenhouse gas fluxes) in the context of land management: final report to DEFRA from WCA environment limited. In: http://sciencesearch.defra.gov.uk/Document.aspx?Document=SP0567_9953_FRP.pdf.
Defra, 2016. Farming statistics – final crop areas and cattle, sheep and pig populations as at 1 June 2016, England. https://www.gov.uk/government/statistics/farming-statistics-final-crop-areas-and-cattle-sheep-and-pig-populations-as-at-1-june-2016-england (accessed 11th Jan 2022).
Defra, 2019. In: British Survey of Fertiliser Practice 2018 - Fertiliser Use on Farm Crops for Crop Year 2018. p. 117.
Dimitriou, L., Aronsson, P., 2004. Nitrogen leaching from short-rotation willow coppice after intensive irrigation with wastewater. Biomass Bioenergy. 26, 433–441. https://doi.org/10.1016/j.biombioe.2003.08.009.
Dimitriou, I., Rosenqvist, H., 2011. Sewage sludge and wastewater fertilisation of short rotation coppice (SRC) for increased bioenergy production—biological and economic potential. Biomass Bioenergy 35, 835–842. https://doi.org/10.1016/j.biombioe.2010.11.010.
Dimitriou, I., Mola-Yudego, B., Aronsson, P., 2012. Impact of willow short rotation coppice on water quality. Bioenergy Res. 5, 537–545. https://doi.org/10.1007/s12155-012-9211-5.
Eggleston, H.S., Buendia, L., Miwa, K., Ngara, T., Tanabe, K., 2006. 2006 IPCC Guidelines for National Greenhouse Gas Inventories, Japan.
Ehrhardt, F., Soussana, J.-F., Bellocchi, G., Grace, P., McAuliffe, R., Recous, S., et al., 2018. Assessing uncertainties in crop and pasture ensemble model simulations of productivity and N2O emissions. Glob. Chang. Biol. 24, e603–e616. https://doi.org/10.1111/gcb.13965.
Environment Agency, 2020. WFD River Waterbody Catchments Cycle 2. https://data.gov.uk/dataset/298258ee-c4a0-4505-a3b5-0e6585ecfdb2/wfd-river-waterbody-catchments-cycle-2 (accessed 11th Jan 2022).
Evans, R., Collins, A.L., Foster, I.D.L., Rickson, R.J., Anthony, S.G., Brewer, T., et al., 2016. Extent, frequency and rate of water erosion of arable land in Britain - benefits and challenges for modelling. Soil Use Manag. 32, 149–161. https://doi.org/10.1111/sum.12210.
Evans, R., Collins, A.L., Zhang, Y., Foster, I.D.L., Boardman, J., Sint, H., et al., 2017. A comparison of conventional and Cs-137-based estimates of soil erosion rates on arable and grassland across lowland England and Wales. Earth Sci. Rev. 173, 49–64. https://doi.org/10.1016/j.earscirev.2017.08.005.
Farina, R., Sandor, R., Abdalla, M., Alvaro-Fuentes, J., Bechini, L., Bolinder, M.A., et al., 2021. Ensemble modelling, uncertainty and robust predictions of organic carbon in long-term bare-fallow soils. Glob. Chang. Biol. 27, 904–928. https://doi.org/10.1111/gcb.15441.
Ferchaud, F., Peyrard, C., Léonard, J., Gréhan, E., Mary, B., 2020. Large variations in N2O fluxes from bioenergy crops according to management practices and crop type. Atmosphere 11, 675.
Foresight, 2011. The Future of Food and Farming, Final Project Report. The Government Office for Science, London (accessed 11th Jan 2022).
Gaillard, R.K., Jones, C.D., Ingraham, P., Collier, S., Izaurralde, R.C., Jokela, W., et al., 2018. Underestimation of N2O emissions in a comparison of the DayCent, DNDC, and EPIC models. Ecol. Appl. 28, 694–708. https://doi.org/10.1002/eap.1674.
Gedney, N., Cox, P.M., Betts, R.A., Boucher, O., Huntingford, C., Stott, P.A., 2006. Detection of a direct carbon dioxide effect in continental river runoff records. Nature 439, 835–838. https://doi.org/10.1038/nature04504.
Georgakakos, K.P., Carpenter, T.M., 2006. Potential value of operationally available and spatially distributed ensemble soil water estimates for agriculture. J. Hydrol. 328, 177–191. https://doi.org/10.1016/j.jhydrol.2005.12.018.
Gneiting, T., Raftery, A.E., 2005. Atmospheric science - weather forecasting with ensemble methods. Science 310, 248–249. https://doi.org/10.1126/science.1115255.
Granger, S.J., Heaton, T.H.E., Pfahler, V., Blackwell, M.S.A., Yuan, H.M., Collins, A.L., 2017. The oxygen isotopic composition of phosphate in river water and its potential sources in the Upper River Taw catchment, UK. Science of the Total Environment 574, 680–690. https://doi.org/10.1016/j.scitotenv.2016.09.007.
Granger, S.J., Harris, P., Upadhayay, H.R., Sint, H., Pulley, S., Stone, M., et al., 2021. Novel approaches to investigating spatial variability in channel bank total phosphorus at the catchment scale. Catena 202. https://doi.org/10.1016/j.catena.2021.105223.
Gregory, A.S., Dungait, J.A.J., Shield, I.F., Macalpine, W.J., Cunniff, J., Durenkamp, M., et al., 2018. Species and genotype effects of bioenergy crops on root production, carbon and nitrogen in temperate agricultural soil. Bioenergy Res. 11, 382–397. https://doi.org/10.1007/s12155-018-9903-6.
Harkness, C., Semenov, M.A., Areal, F., Senapati, N., Trnka, M., Balek, J., et al., 2020. Adverse weather conditions for UK wheat production under climate change. Agric. For. Meteorol. 282. https://doi.org/10.1016/j.agrformet.2019.107862.
Hoglind, M., Schapendonk, A., Van Oijen, M., 2001. Timothy growth in Scandinavia: combining quantitative information and simulation modelling. New Phytol. 151, 355–367. https://doi.org/10.1046/j.0028-646x.2001.00195.x.
Hopkins, A., Gilbey, J., Dibb, C., Bowling, P.J., Murray, P.J., 1990. Response of permanent and reseeded grassland to fertilizer nitrogen. 1. Herbage production and herbage quality. Grass Forage Sci. 45, 43–55. https://doi.org/10.1111/j.1365-2494.1990.tb02181.x.
HowMonk, 2022. How Much Does a Cow Weigh? (Beef, Dairy, Calf – with charts). https://howmonk.com/how-much-does-a-cow-weigh/ (accessed 26 January 2022).
Johnston, A.E., Poulton, P.R., Coleman, K., 2009. Soil organic matter: its importance in sustainable agriculture and carbon dioxide fluxes. In: Advances in Agronomy, 101. Elsevier Academic Press Inc, San Diego, pp. 1–57.
Jones, R., 2013. Better soil and grassland management for Scottish beef and lamb producers. In: https://www.qmscotland.co.uk/sites/default/files/QM2564%2BSoils%2BBooklet%2BAW%2B1213%2Blr.pdf.
Kafatos, M.C., Kim, S.H., Lim, C.-H., Kim, J., Lee, W.-K., 2017. Responses of agroecosystems to climate change: specifics of resilience in the mid-latitude region. Sustainability 9, 1361.
Kaipainen, T., Liski, J., Pussinen, A., Karjalainen, T., 2004. Managing carbon sinks by changing rotation length in European forests. Environ. Sci. Pol. 7, 205–219. https://doi.org/10.1016/j.envsci.2004.03.001.
Knutti, R., Abramowitz, G., Collins, M., Eyring, V., Gleckler, P., Hewitson, B., et al., 2010. Good Practice Guidance Paper on Assessing and Combining Multi Model Climate Projections.
Liang, S., Zhang, X.B., Sun, N., Li, Y.F., Xu, M.G., Wu, L.H., 2019. Modeling crop yield and nitrogen use efficiency in wheat and maize production systems under future climate change. Nutr. Cycl. Agroecosyst. 115, 117–136. https://doi.org/10.1007/s10705-019-10013-4.
Liu, Y.Y., Wu, L., Watson, C.A., Baddeley, J.A., Pan, X.B., Zhang, L.Z., 2013. Modeling

biological dinitrogen fixation of field pea with a process-based simulation model. Agron. J. 105, 670–678. https://doi.org/10.2134/agronj2012.0412.
Lloyd, C.E.M., Johnes, P.J., Freer, J.E., Carswell, A.M., Jones, J.I., Stirling, M.W., et al., 2019. Determining the sources of nutrient flux to water in headwater catchments: examining the speciation balance to inform the targeting of mitigation measures. Sci. Total Environ. 648, 1179–1200. https://doi.org/10.1016/j.scitotenv.2018.08.190.
Lord, E.I., Anthony, S.G., 2000. MAGPIE: a modelling framework for evaluating nitrate losses at national and catchment scales. Soil Use Manag. 16, 167–174.
Lovett, A.A., Sunnenberg, G.M., Richter, G.M., Dailey, A.G., Riche, A.B., Karp, A., 2009. Land use implications of increased biomass production identified by GIS-based suitability and yield mapping for miscanthus in England. Bioenergy Res. 2, 17–28. https://doi.org/10.1007/s12155-008-9030-x.
Martre, P., Wallach, D., Asseng, S., Ewert, F., Jones, J.W., Rötter, R.P., et al., 2015. Multimodel ensembles of wheat growth: many models are better than one. Glob. Chang. Biol. 21, 911–925. https://doi.org/10.1111/gcb.12768.
McCalmont, J.P., Hastings, A., McNamara, N.P., Richter, G.M., Robson, P., Donnison, I.S., et al., 2017. Environmental costs and benefits of growing miscanthus for bioenergy in the UK. Glob. Chang. Biol. Bioenergy 9, 489–507. https://doi.org/10.1111/gcbb.12294.
McVoy, C.W., Kersebaum, K.C., Arning, M., Kleeberg, P., Othmer, H., Schröder, U., 1995. A data set from North Germany for the validation of agroecosystem models: documentation and evaluation. Ecol. Model. 81, 265–300. https://doi.org/10.1016/0304-3800(94)00197-p.
Met Office, 2018. HadUK-Grid Gridded Climate Observations on a 1km grid over the UK for 1862-2017. Centre for Environmental Data Analysis date of citation. http://catalogue.ceda.ac.uk/uuid/2a62652a4fe6412693123dd6328f6dc8 (accessed 11th Jan 2022).
Milne, A.E., Glendining, M.J., Lark, R.M., Perryman, S.A.M., Gordon, T., Whitmore, A.P., 2015. Communicating the uncertainty in estimated greenhouse gas emissions from agriculture. J. Environ. Manag. 160, 139–153. https://doi.org/10.1016/j.jenvman.2015.05.034.
Morrison, J., Jackson, M.V., Sparrow, P.E., 1980. The response of perennial ryegrass to fertilizer nitrogen in relation to climate and soil. In: Technical Report No 27. Grassland Research Institute, Hurley, U.K. (accessed 11th Jan 2022).
Mulder, A., Vandegraaf, A.A., Robertson, L.A., Kuenen, J.G., 1995. Anaerobic ammonium oxidation discovered in a denitrifying fluidized-bed reactor. FEMS Microbiol. Ecol. 16, 177–183. https://doi.org/10.1016/0168-6496(94)00081-7.
Ni, Y.Z., Mwabonje, O.N., Richter, G.M., Qi, A.M., Yeung, K., Patel, M., et al., 2019. Assessing availability and greenhouse gas emissions of lignocellulosic biomass feedstock supply - case study for a catchment in England. Biofuels Bioprod. Biorefining 13, 568–581. https://doi.org/10.1002/bbb.1966.
Nilsson, M., Griggs, D., Visbeck, M., 2016. Policy: map the interactions between sustainable development goals. Nature 534, 320–322. https://doi.org/10.1038/534320a.
Nix, 2020. The John Nix Pocketbook for Farm Management.
Oertel, C., Matschullat, J., Zurba, K., Zimmermann, F., Erasmi, S., 2016. Greenhouse gas emissions from soils a review. Chemie Der Erde-Geochem. 76, 327–352. https://doi.org/10.1016/j.chemer.2016.04.002.
Paré, D., Thiffault, E., 2016. Nutrient budgets in forests under increased biomass harvesting scenarios. Curr. Forest. Rep. 2, 81–91. https://doi.org/10.1007/s40725-016-0030-3.
Parton, W.J., Schime, D.S., Ojima, D.S., Cole, C.V., 1994. A general model for soil organic matter dynamics: sensitivity to litter chemistry, texture and management. In: Bryant, R.B., Arnold, R.W. (Eds.), Quantitative Modeling of Soil Forming Processes. SSSA Special Publication 39, Madison, WI, pp. 147–167.
Piskier, T., 2017. A method of estimation of the caloric value of the biomass. Part I – biomass energy potential. J. Mech. Energy Eng. 1, 189–194.
Ploschuk, R.A., Miralles, D.J., Colmer, T.D., Ploschuk, E.L., Striker, G.G., 2018. Waterlogging of winter crops at early and late stages: impacts on leaf physiology, growth and yield. Front. Plant Sci. 9. https://doi.org/10.3389/fpls.2018.01863.
Pulley, S., Collins, A.L., 2020. Sediment loss in response to scheduled pasture ploughing and reseeding: the importance of soil moisture content in controlling risk. Soil Tillage Res. 204. https://doi.org/10.1016/j.still.2020.104746.
Qi, A.M., Murray, P.J., Richter, G.M., 2017. Modelling productivity and resource use efficiency for grassland ecosystems in the UK. Eur. J. Agron. 89, 148–158. https://doi.org/10.1016/j.eja.2017.05.002.
Qi, A.M., Holland, R.A., Taylor, G., Richter, G.M., 2018. Grassland futures in Great Britain - productivity assessment and scenarios for land use change opportunities. Sci. Total Environ. 634, 1108–1118. https://doi.org/10.1016/j.scitotenv.2018.03.395.
Richard, B., Richter, G.M., Cerasuolo, M., Shield, I., 2019. Optimizing the bioenergy water footprint by selecting SRC willow canopy phenotypes: regional scenario simulations. Ann. Bot. 124, 531–542. https://doi.org/10.1093/aob/mcz006.
Richter, G.M., Rana, G., Ferrara, R.M., Ventrella, D., Acutis, M., Trevisiol, P., et al., 2006. Stability and Mitigation of Arable Systems in Hilly Landscapes. Final Report. European Commission, Brussels.
Richter, G.M., Acutis, M., Trevisiol, P., Latiri, K., Confalonieri, R., 2010. Sensitivity analysis for a complex crop model applied to durum wheat in the Mediterranean. Eur. J. Agron. 32, 127–136. https://doi.org/10.1016/j.eja.2009.09.002.
Richter, G.M., Agostini, F., Barker, A., Costomiris, D., Qi, A.M., 2016. Assessing on-farm productivity of miscanthus crops by combining soil mapping, yield modelling and remote sensing. Biomass Bioenergy 85, 252–261. https://doi.org/10.1016/j.biombioe.2015.12.024.
Riggers, C., Poeplau, C., Don, A., Bamminger, C., Hoper, H., Dechow, R., 2019. Multi-model ensemble improved the prediction of trends in soil organic carbon stocks in German croplands. Geoderma 345, 17–30. https://doi.org/10.1016/j.geoderma.2019.03.014.
Rodríguez, A., Ruiz-Ramos, M., Palosuo, T., Carter, T.R., Fronzek, S., Lorite, I.J., et al., 2019. Implications of crop model ensemble size and composition for estimates of adaptation effects and agreement of recommendations. Agric. For. Meteorol. 264, 351–362. https://doi.org/10.1016/j.agrformet.2018.09.018.
Rowland, C.S., Morton, R.D., Carrasco, L., McShane, G., Neil, A.W., Wood, C.M., 2017. Land Cover Map 2015 (vector, GB). https://doi.org/10.5285/6c6c9203-7333-4d96-88ab-78925e7a4e73.
Ruane, A.C., Hudson, N.I., Asseng, S., Cammarano, D., Ewert, F., Martre, P., et al., 2016. Multi-wheat-model ensemble responses to interannual climate variability. Environ. Model Softw. 81, 86–101. https://doi.org/10.1016/j.envsoft.2016.03.008.
Sándor, R., Ehrhardt, F., Grace, P., Recous, S., Smith, P., Snow, V., et al., 2020. Ensemble modelling of carbon fluxes in grasslands and croplands. Field Crop Res. 252, 107791. https://doi.org/10.1016/j.fcr.2020.107791.
Sawayama, S., 2006. Possibility of anoxic ferric ammonium oxidation. J. Biosci. Bioeng. 101, 70–72. https://doi.org/10.1263/jbb.101.70.
Schapendonk, A., Stol, W., van Kraalingen, D.W.G., Bouman, B.A.M., 1998. LINGRA, a sink/source model to simulate grassland productivity in Europe. Eur. J. Agron. 9, 87–100. https://doi.org/10.1016/s1161-0301(98)00027-6.
Scharlemann, J.P.W., Brock, R.C., Balfour, N., Brown, C., Burgess, N.D., Guth, M.K., et al., 2020. Towards understanding interactions between sustainable development goals: the role of environment–human linkages. Sustain. Sci. 15, 1573–1584. https://doi.org/10.1007/s11625-020-00799-6.
Schlesinger, W.H., 2013. An estimate of the global sink for nitrous oxide in soils. Glob. Chang. Biol. 19, 2929–2931. https://doi.org/10.1111/gcb.12239.
Shan, Y., Huang, M.B., Harris, P., Wu, L.H., 2021. A sensitivity analysis of the SPACSYS model. Agriculture-Basel 11. https://doi.org/10.3390/agriculture11070624.
Sharp, R.T., Henrys, P.A., Jarvis, S.G., Whitmore, A.P., Milne, A.E., Coleman, K., et al., 2021. Simulating cropping sequences using earth observation data. Comput. Electron. Agric. 188, 106330. https://doi.org/10.1016/j.compag.2021.106330.
Silgram, M., Anthony, S.G., Collins, A.L., Stromqvist, J., Bouraoui, F., Schoumans, O., et al., 2009. Evaluation of diffuse pollution model applications in EUROHARP catchments with limited data. J. Environ. Monit. 11, 554–571. https://doi.org/10.1039/b901181a.
Smith, P., Smith, J.U., Powlson, D.S., McGill, W.B., Arah, J.R.M., Chertov, O.G., et al., 1997. A comparison of the performance of nine soil organic matter models using datasets from seven long-term experiments. Geoderma 81, 153–225.
Smith, J., Smith, P., Wattenbach, M., Zaehle, S., Hiederer, R., Jones, R.J.A., et al., 2005. Projected changes in mineral soil carbon of European croplands and grasslands, 1990–2080. Glob. Chang. Biol. 11, 2141–2152. https://doi.org/10.1111/j.1365-2486.2005.001075.x.
Spott, O., Russow, R., Stange, C.F., 2011. Formation of hybrid N2O and hybrid N-2 due to codenitrification: first review of a barely considered process of microbially mediated N-nitrosation. Soil Biol. Biochem. 43, 1995–2011. https://doi.org/10.1016/j.soilbio.2011.06.014.
Stroemqvist, J., Collins, A.L., Davison, P.S., Lord, E.I., 2008. PSYCHIC - A process-based model of phosphorus and sediment transfers within agricultural catchments. Part 2. A preliminary evaluation. Journal of Hydrology 350, 303–316. https://doi.org/10.1016/j.jhydrol.2007.10.044.
Stromqvist, J., Collins, A.L., Davison, P.S., Lord, E.I., 2008. PSYCHIC - a process-based model of phosphorus and sediment transfers within agricultural catchments. Part 2. A preliminary evaluation. Journal of Hydrology. 350, 303–316. https://doi.org/10.1016/j.jhydrol.2007.10.044.
Suarez-Gutierrez, L., Milinski, S., Maher, N., 2021. Exploiting large ensembles for a better yet simpler climate model evaluation. Clim. Dyn. 57, 2557–2580. https://doi.org/10.1007/s00382-021-05821-w.
Taylor, K.E., Stouffer, R.J., Meehl, G.A., 2012. An overview of CMIP5 and the experiment design. Bull. Am. Meteorol. Soc. 93, 485–498. https://doi.org/10.1175/bams-d-11-00094.1.
Tiedje, J.M., 1988. Ecology of denitrification and dissimilatory nitrate reduction to ammonium. In: Zehnder, A.J.B. (Ed.), Biology of Anaerobic Microorganisms. Wiley, New York, pp. 179–244.
Triana, F., Ragaglini, G., Bonari, E., Cerasuolo, M., Richter, G.M., 2011. Modelling the water balance of different grass species used for bioenergy. Biomass and energy crops. Asp. Appl. Biol. 112, 163–170.
UN General Assembly, 2015. Transforming our world: the 2030 agenda for sustainable development.
van Groenigen, J.W., Huygens, D., Boeckx, P., Kuyper, T.W., Lubbers, I.M., Rutting, T., et al., 2015. The soil N cycle: new insights and key challenges. Soil 1, 235–256. https://doi.org/10.5194/soil-1-235-2015.
van Laar, H., Goudriaan, J., van Keulen, H., 1992. Simulation of Crop Growth for Potential and Water-limited Production Situations, (as Applied to Spring Wheat), Cabo-dlo. (accessed 11th Jan 2022).
Virkkala, A.M., Aalto, J., Rogers, B.M., Tagesson, T., Treat, C.C., Natali, S.M., et al., 2021. Statistical upscaling of ecosystem CO2 fluxes across the terrestrial tundra and boreal domain: regional patterns and uncertainties. Glob. Chang. Biol. https://doi.org/10.1111/gcb.15659.
Wallach, D., Mearns, L.O., Ruane, A.C., Rötter, R.P., Asseng, S., 2016. Lessons from climate modeling on the design and use of ensembles for crop modeling. Clim. Chang. 139, 551–564. https://doi.org/10.1007/s10584-016-1803-1.
Wallach, D., Martre, P., Liu, B., Asseng, S., Ewert, F., Thorburn, P.J., et al., 2018. Multimodel ensembles improve predictions of crop–environment–management interactions. Glob. Chang. Biol. 24, 5072–5083. https://doi.org/10.1111/gcb.14411.
Wang, L., Stuart, M.E., Lewis, M.A., Ward, R.S., Skirvin, D., Naden, P.S., et al., 2016. The changing trend in nitrate concentrations in major aquifers due to historical nitrate loading from agricultural land across England and Wales from 1925 to 2150. Sci. Total Environ. 542, 694–705. https://doi.org/10.1016/j.scitotenv.2015.10.127.

Webb, J., Misselbrook, T.H., 2004. A mass-flow model of ammonia emissions from UK livestock production. Atmos. Environ. 38, 2163–2176. https://doi.org/10.1016/j.atmosenv.2004.01.023.
Willcock, S., Hooftman, D.A.P., Blanchard, R., Dawson, T.P., Hickler, T., Lindeskog, M., et al., 2020. Ensembles of ecosystem service models can improve accuracy and indicate uncertainty. Sci. Total Environ. 747, 141006. https://doi.org/10.1016/j.scitotenv.2020.141006.
Wolf, J., 2012. In: User Guide for LINTUL4 and LINTUL4V: Simple Generic Model for Simulation of Crop Growth Under Potential, Water Limited and Nitrogen Limited Conditions. Wageningen UR, Wageningen, p. 58.
Wrage, N., Velthof, G.L., van Beusichem, M.L., Oenema, O., 2001. Role of nitrifier denitrification in the production of nitrous oxide. Soil Biol. Biochem. 33, 1723–1732.
Wu, L., McGechan, M.B., McRoberts, N., Baddeley, J.A., Watson, C.A., 2007. SPACSYS: integration of a 3D root architecture component to carbon, nitrogen and water cycling-model description. Ecol. Model. 200, 343–359. https://doi.org/10.1016/j.ecolmodel.2006.08.010.
Wu, L., Rees, R.M., Tarsitano, D., Zhang, X.B., Jones, S.K., Whitmore, A.P., 2015. Simulation of nitrous oxide emissions at field scale using the SPACSYS model. Sci. Total Environ. 530, 76–86. https://doi.org/10.1016/j.scitotenv.2015.05.064.
Wu, L., Zhang, X., Griffith, B.A., Misselbrook, T.H., 2016. Sustainable grassland systems: a modelling perspective based on the north wyke farm platform. Eur. J. Soil Sci. 67, 397–408. https://doi.org/10.1111/ejss.12304.
Wu, L.H., Blackwell, M., Dunham, S., Hernandez-Allica, J., McGrath, S.P., 2019. Simulation of phosphorus chemistry, uptake and utilisation by winter wheat. Plants-Basel 8. https://doi.org/10.3390/plants8100404.
Wu, L., Curceac, S., Atkinson, P.M., Milne, A., Harris, P., 2021. A case study on the effects of data temporal resolution on the simulation of water flux extremes using a process-based model at the grassland field scale. Agric. Water Manag. 255, 107049. https://doi.org/10.1016/j.agwat.2021.107049.
Wu, L., Harris, P., Misselbrook, T.H., Lee, M.R.F., 2022. Simulating grazing beef and sheep systems. Agric. Syst. 195, 103307. https://doi.org/10.1016/j.agsy.2021.103307.
Yin, X., Kersebaum, K.C., Kollas, C., Baby, S., Beaudoin, N., Manevski, K., et al., 2017. Multi-model uncertainty analysis in predicting grain N for crop rotations in Europe. Eur. J. Agron. 84, 152–165. https://doi.org/10.1016/j.eja.2016.12.009.
Zatta, A., Clifton-Brown, J., Robson, P., Hastings, A., Monti, A., 2014. Land use change from C3 grassland to C4 miscanthus: effects on soil carbon content and estimated mitigation benefit after six years. GCB Bioenergy 6, 360–370. https://doi.org/10.1111/gcbb.12054.
Zhang, Y., Collins, A.L., Jones, J.I., Johnes, P.J., Inman, A., Freer, J.E., 2017. The potential benefits of on-farm mitigation scenarios for reducing multiple pollutant loadings in prioritised Agri-environment areas across England. Environ. Sci. Pol. 73, 100–114. https://doi.org/10.1016/j.envsci.2017.04.004.
Zhang, Y., Griffith, B., Granger, S., Sint, H., Collins, A.L., 2022. Tackling unintended consequences of grazing livestock farming: multi-scale assessment of co-benefits and trade-offs for water pollution mitigation scenarios. J. Clean. Prod. 336, 130449. https://doi.org/10.1016/j.jclepro.2022.130449.
Zhu, J., Forsee, W., Schumer, R., Gautam, M., 2013. Future projections and uncertainty assessment of extreme rainfall intensity in the United States from an ensemble of climate models. Clim. Chang. 118, 469–485. https://doi.org/10.1007/s10584-012-0639-6.