Assessing uncertainties in land cover projections

Abstract

Understanding uncertainties in land cover projections is critical to investigating land-based climate mitigation policies, assessing the potential of climate adaptation strategies, and quantifying the impacts of land cover change on the climate system. Here we identify and quantify uncertainties in global and European land cover projections over a diverse range of model types and scenarios, extending the analysis beyond the agro-economic models included in previous comparisons. The results show a large range in future land cover area projections, with the highest variability occurring in cropland areas. We demonstrate a significant, systematic difference in land cover areas arising from the characteristics of the modelling approach, which is at least as great as the differences between scenarios. This leads us to conclude that a diverse set of models and approaches is required in order to account for model uncertainty when assessing the potential impacts of land cover change on future climate.

1 Introduction

Land use and land cover (LULC) change plays an important role in climate change. LULC change is believed to be responsible for a substantial proportion of total carbon dioxide (CO_2) emissions, 10-20% since 1990[1,2] and approximately a third since pre-industrial times[2], while land-based, climate mitigation measures could contribute substantially to the abatement of future greenhouse gas emissions[3]. Climate change also impacts LULC, both through direct effects on crops and natural vegetation and through land management and land use changes implemented as adaptation responses[4,5]. LULC is not only influenced by climate change, but also by socio-economic factors, such as population dynamics, wealth and urbanisation, which are important for determining demand for agricultural and forestry commodities[6–8].

Modelling at a range of spatial scales has been applied to understand the LULC response to climatic and socio-economic drivers, and to assess the potential for mitigation and adaptation to climate change[9–14]. However, different modelling approaches can produce different outcomes. Uncertainty also arises due to the range of potential socio-economic and climate futures. Attempts have been made to characterise the uncertainty in socio-economic drivers through scenarios, including the IPCC's special report on emissions scenarios (SRES)[15], and more recently, shared socio-economic pathways (SSPs)[16] in combination with representative concentration pathways (RCPs)[17]. Model inter-comparison studies, drawing together the findings of many different modelling approaches, have previously considered aspects of LULC including the agricultural model inter-comparison and improvement project (AgMIP)[18,19], the inter-sectoral impact model inter-comparison project (ISI-MIP)[20], and the coupled model inter-comparison project (CMIP)[21]. CMIP deals primarily with the impact of land use on climate, and AgMIP, which is closely linked to the agricultural sector of ISI-MIP, has a broad focus on various aspects of agricultural models. AgMIP compared the results from 10 global agro-economic models to 2050, demonstrating significant LULC change differences, even within the same scenario, due to differences in model assumptions and parameterisation[19,22]. However, there has been no previous model inter-comparison of LULC projections which examines uncertainty over the breadth of relevant model types. Further knowledge gaps exist in understanding the relative role of model and scenario uncertainty, as well as the influence of model spatial extent, i.e. do global and regional results systemically differ. Understanding uncertainties in LULC projections is critical to investigating the effectiveness of land-based climate mitigation policies, in assessing the potential of climate adaptation strategies, and in quantifying the impacts of land cover change on the climate system.

This study seeks to address these knowledge gaps, and identify and analyse uncertainties in global and European LULC, by comparing projections from a diverse range of models and scenarios. The aim is to quantify the potential range of future LULC and to better understand the associated sources and levels of uncertainty. The study goes beyond existing comparisons in a number of ways. Firstly, it incorporates a wider range of model types, including process or rule-based models in addition to the computable-general equilibrium (CGE) and partial equilibrium (PE) models evaluated in AgMIP. Secondly, it compares models from different spatial extents, including both global and regional-scale models for the European continent. Europe was chosen for this comparison because of the availability of a large number of regional models. Finally, it incorporates a broader range of socio-economic and climate scenarios. Rather than using a small set of common scenarios[18,19], model teams were invited to submit multiple, potentially dissimilar scenarios, which allows the potential extent of scenario space to be more fully covered. This approach also supports the inclusion of a greater diversity of scenarios and models. For example, without the requirement to implement particular scenarios, models that have been developed for different purposes, and thus have implemented different scenarios, can still be included. Consequently a better representation of the range of uncertainty in projected LULC change can be achieved.

Data from 17 models and 70 scenarios were considered (Table 1). Statistical methods were used to augment qualitative insights from comparing differences between the model results. To quantify the relative importance of factors associated with the components of the variability, a multiple linear regression and analysis of variance (ANOVA)[23,24] were used, with variables for the initial condition, model and scenario (climate and socio-economic) factors, and residual or unexplained variability. The robustness of the analysis and completeness of the scenario and model variables were assessed, including through the use of linear mixed effects modelling[25].

The analysis identifies and draws inference from the variability between the LULC projections, and separates the factors driving future LULC uncertainty between the impacts of model-related factors (model type, resolution and extents) and the scenario characteristics. It is not the intention to identify which model or scenario is more plausible, or to indicate which model or approach could be considered more accurate.

2 Variations in modelled land cover areas

Future LULC projections are uncertain, and the results display a wide variation for all assessed land cover types. The global and European land cover over time are shown in Figures 1 and 2, plotted both as absolute areas and scaled to match the FAOSTAT areas at 2010[26]. Global cropland areas follow what might be the anticipated pattern, with relatively small initial differences between scenarios (1290-1650 Mha, 95% interval at 2010), which diverge over time across a range of scenarios (930-2670 Mha at 2100). However, the global pasture and forest areas do not fit this pattern. They demonstrate a relatively large initial variation, which does not change substantially over time. The main reasons for these discrepancies in initial conditions are due to uncertainty in current areas, and differences in the definition of land cover (both in models and in observations). There is a lack agreement particularly over what constitutes pasture and forest, e.g. how to categorise grazed forest land or semiarid grazing[27]. Scaling to a common starting value allows the model trends without these differences to be observed, and this shows the expected pattern of increasing variability over time (Figures 1&2-ii). FAO data[26] was used to display historic values, and is a commonly used source for such data. A small number of scenarios suggest rapid changes in some types of land cover. For example, by 2050 FALAFEL under SSP1 gives a reduction in global cropland of 43%, and LandSHIFT an increase of 76-107% compared to present-day depending on the scenario.

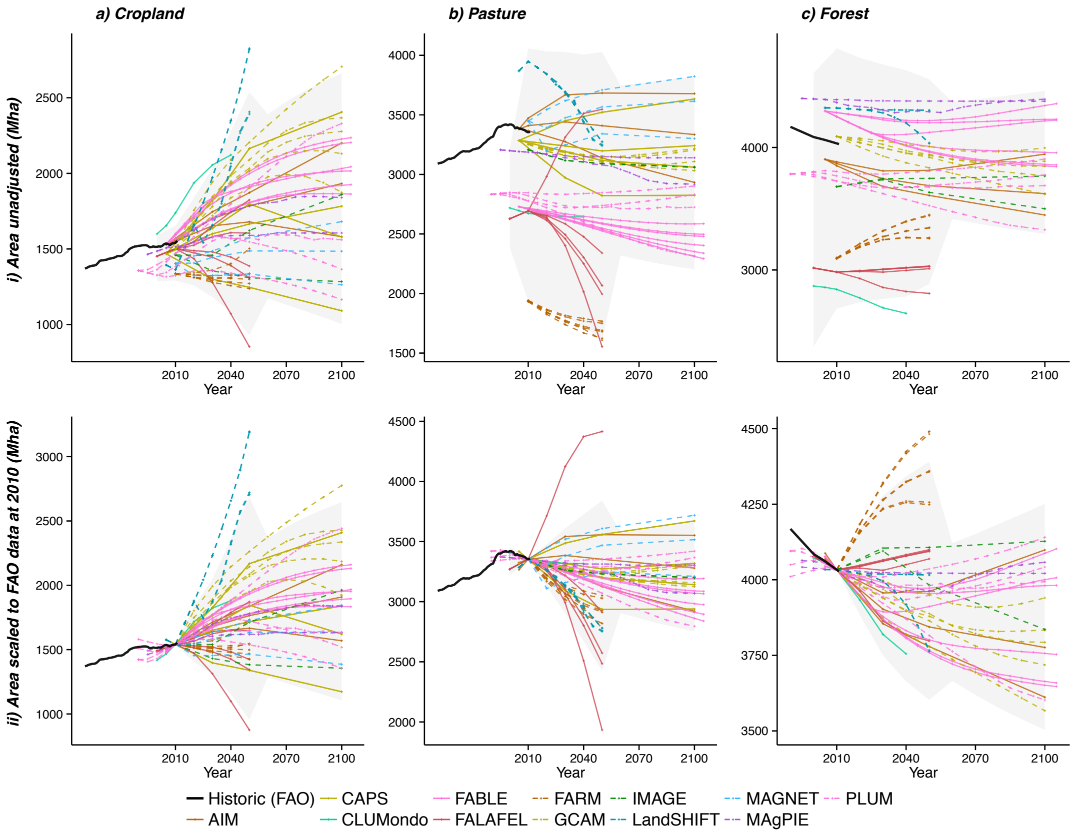

Figure 1. Global modelled land cover areas for cropland (a), pasture (b), and forest (c), from 12 models and a total of 49 scenarios. A historical dataset[26] is shown as a solid black lines, and the 95% interval of model results in as grey shading. The absolute areas are shown in i) and the areas scaled to match the historical data in 2010 are shown in ii). See Table 1 for model and scenario information.

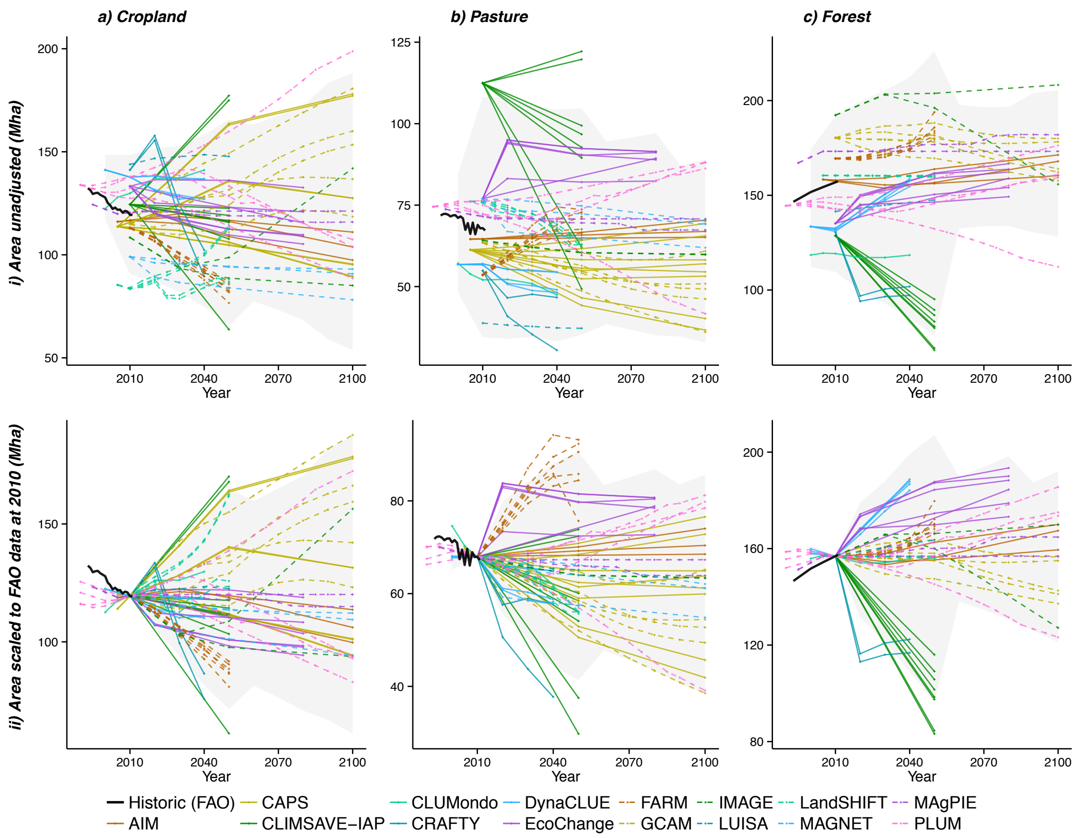

Figure 2. European land cover for 15 models over a total of 59 scenarios based on the EU27 member states. Legend and format consistent with Figure 1.

The European land cover areas (Figure 2) show some of the same patterns of variations as the global areas (Figure 1), e.g. lower initial variation for cropland than for pasture or forest. Some of the European regional models produce many of the more extreme area changes, with CLIMSAVE-IAP, CRAFTY and EcoChange all producing the highest or lowest scaled areas for multiple cover types, although most of the European regional models do not extend past 2050. CLIMSAVE-IAP has a relatively high initial value for pasture, which in the SRES A1 and B1 scenarios decreases rapidly, while forest is lower and decreases substantially in all scenarios, in contrast to the majority of other model results. CRAFTY and EcoChange also have some results in which the direction of change varies through time, while other models tend to show a more consistent direction of change over time for each scenario.

The coefficient of variation, the ratio of the standard deviation to the mean, was used to provide a comparative measure of dispersion across model runs between the global and European areas and the land cover types considered (Figures 3&4-i). This illustrates again that the initial variation is relatively low for cropland, but increases over time. Pasture and forest areas do not exhibit this pattern with global forest area variability decreasing over time, and pasture area variability remaining relatively constant over time; both show a minimum in 2050. The coefficient of variation is generally higher at a European than a global level, particularly for pasture and forest areas.

To analyse the source of these variations, an ANOVA approach was used to show the relative importance of different sources of variance for each land cover type and decadal end year (Figures 3&4-ii). The decomposition was based on 10 variables (Table S3) plus a residual, for the variation not captured by these variables. Higher variance fractions imply that a variable has a higher significance in the regression, and the greater ability to explain the total variance. The initial condition delta has been calculated based on the 2010 baseline area, and therefore 100% of the fraction of variance is associated with it at that point. The significance of the initial condition, in general, decreases over time. For global pasture and forest areas the initial condition remains the most important factor over all time periods.

There is a discontinuity in the results between 2050 and 2060 (Figures 3&4) because a number of models end at 2050. A similar but less substantial effect also occurs between 2080 and 2090 for European data. These effect were removed by rerunning the analysis using only results that extend to 2100 (Figures S5&6), but at the expense of removing approximately half (36 of 70) of the available scenarios. The model results dataset and therefore the analysis do not change for the period 2060-2100 for global areas, and from 2080 in the European data. In the period prior to 2050, cropland and European forest all have more variance associated with scenario variables.

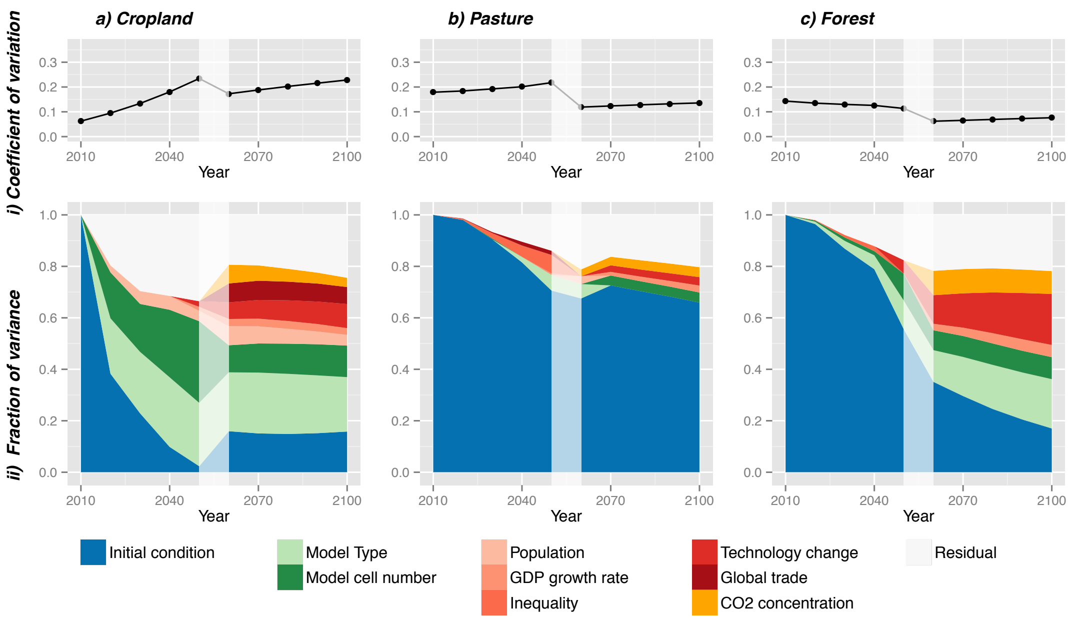

Figure 3. Coefficient of variation (i) and relative importance of different variance components (ii) for global land cover areas between 2010 and 2100. The shaded area between 2050 and 2060 indicates that between these points the set of model results substantially change after 2050. In (ii) variance due to model characteristics is shown in different shades of green and due to scenario characteristics in different shades of red. Figures S5 and S6 show the results from an alternative analysis using only model result that extend to 2100.

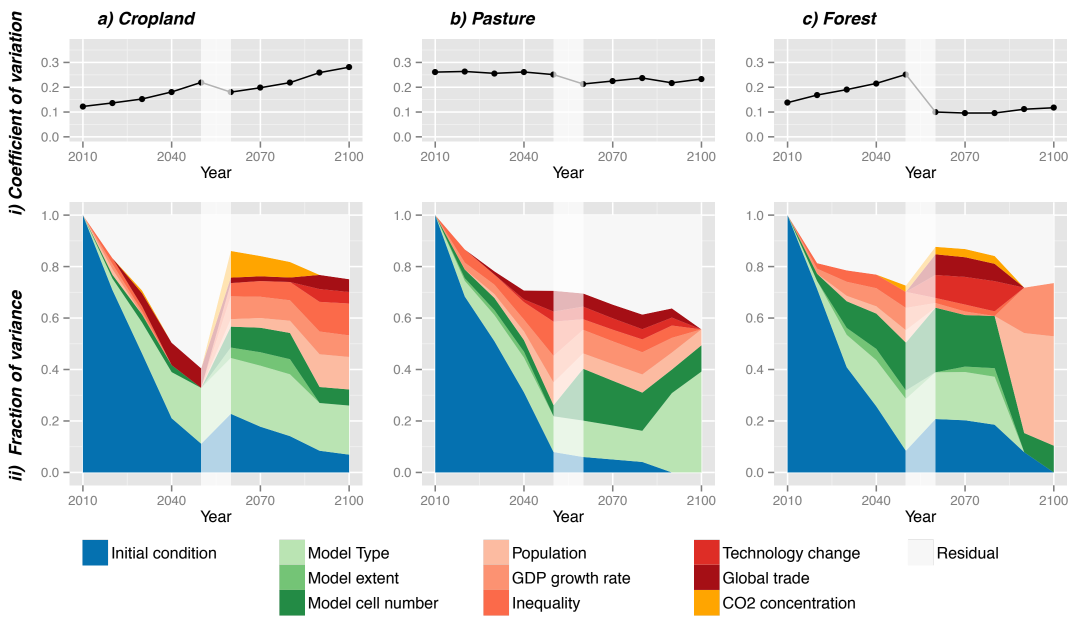

Figure 4. Total coefficient of variation (a) and relative importance of different variance components (b) for European (EU27), format as per Figure 3.

3 Sources of variability

The variables used to characterise the scenarios were shown to have a relatively low fraction of variance for all land cover types, particularly for the global projections (Figures 3&4-ii). The fraction of variance for the model characteristics was similar to, or higher than, that for the variables used to characterise the scenarios in most cases for global areas. This suggests that given only knowledge of the scenario, based on the scenario typologies used, one would only be able to predict a small percentage of the total variation in the results. European data overall has a greater proportion of variance associated with the scenario variables, but still shows a fraction associated with variables used to characterise the models. This indicates that models of a similar type have a level of commonality in behaviour. This may arise because similar model types are more likely to have similar implicit or explicit assumptions, or other commonalities such as the data used to derive model parameter values. Some, albeit lower, association occurred with model resolution, represented as the number of grid cells, which again may be due to model similarities. Model extent for the European data does not have a substantial association, i.e. only limited systemic differences were detected between regional and global model results.

The residual component quantifies the variation that is not associated with any of the regression variables (Table S3), or interactions between them. Thus, if key explanatory variables are not included in the scenario or model typologies then the residual will tend to increase. To check that important variables were not overlooked, a mixed model analysis was conducted[25], a statistical technique which combines fixed effects, based on explanatory variables, and random effects. The mixed model used the regression model as fixed effects, and random effects for the model, and socio-economic and climate scenario (Figures S1&2). This showed that the random effect variances associated with the model were, in some cases, high in comparison to the residual, suggesting that some unknown variables may be missing from the model typology, which if included could improve the fit and reduce the residual, and potentially alter the relative importance of the existing variables. The random effects variances associated with the scenario were however relatively small in most cases for the global and European extents, with global cropland in 2050 and an initial period for European forest data being the exceptions. This suggests that the scenario characterisation was sufficient for the purpose of the analysis. Although alternative sets of variables could be equally valid in describing the scenarios and models, due to correlations in the model inputs and the variables selected, overall the mixed model results provide support for the chosen scenario, as well as some support for the model typologies.

Cropland areas have an initial, relatively low level of variability in the initial condition with a 'cone of uncertainty' increasing with time (Figures 1&2). However, this is not seen in pasture and forest areas, which have high initial variability, increasing little over time, with the variability being lower than for cropland by 2100, both in absolute terms and relative to the land cover area. There are uncertainties and issues around what defines pasture and forest, leading to the potential for differences in the areas used here[27,28]. However, it is hard to justify why the uncertainty would not increase over time. The results could suggest therefore that a larger proportion of future uncertainty associated with cropland has been modelled and quantified. That is to say, more of the potential for future variability in pasture and forest areas remain as epistemic uncertainty[29], perhaps by models and scenario exercises giving greater attention to croplands.

The fraction of variance is also supportive of the view that the uncertainty of global cropland areas is more fully represented, and that this also applies to European areas. This can be seen in European and global cropland and European forest areas showing a higher fraction of variance for the scenario variables, indicating that under alike scenarios the models behave, to some extent, in a similar manner. The variation in Europe is generally higher than that globally for all land cover types, and the fraction of variance explained by the initial conditions within Europe diminishes more quickly in comparison to the global data.

4 Limitations and robustness

The inclusion of 17 models (from the 23 known suitable models), covering a wide range of modelling approaches and research institutions, provides a good representation of the diversity of the LULC modelling community. The inclusion of further models or scenarios could alter the outcome of the analysis if the sample used here is not representative. Higher numbers of scenarios or models would tend to increase the significance of the results and provide greater confidence in the conclusions. The scenarios included are dominated by SRES[15] and SSP[16] based scenarios, as much of the existing land-use modelling effort is based on these scenario frameworks, with the result that more extreme changes may fall outside the range of the land cover projections used here.

Consequently, the true range of outcomes due to scenario uncertainty could be greater than represented here.

The dataset is unbalanced, as models or scenarios may be represented by different numbers of results. However, as each model scenario is given equal weight, models with a larger number of scenarios have a greater impact on the outcome of the analysis. The number of scenarios per model ranges from 1 to 8. To assess the possible impact of this on the results, an analysis was undertaken based on each model having an equal weight, i.e. the weight was the reciprocal of the number of scenarios for that model. This creates a different bias towards the scenarios from models that have fewer scenarios overall. The results were only slightly different from those for which each scenario had an equal weight (Figures S3&4), suggesting that the biases are small in both cases. The equal weighting approach was preferred due to its relative simplicity, and each scenario should be viewed as equally likely. The analysis was also run with the outlying (>1.96 standard deviation from the mean in the last year of the model run) results removed. The outcome showed a greater fraction of variance associated with scenario variables for forest, at a European and global extent, and also for European pasture (Figures S7&8). Although some level of variation in the outcomes was noted in all of the variants (Figure S3-8), the outcomes were sufficiently consistent for the inferences drawn to remain valid and to provide a level of confidence in their robustness. The approach of using unaligned scenarios has the drawback of requiring additional complexity in the analysis, and makes it more difficult to determine differences between models, as the scenarios are not directly comparable. The most suitable approach is therefore dependent on the research question. The design of this study was governed by the objective to explore results from a diversity of models over a wide scenario parameter space.
Implications for land use and land cover uncertaintyThe results suggests that there are systematic differences in future land cover areas based on the modelling approach. To determine which model or model type is 'better', or to obtain a set of modelling assumptions that could be considered definitively accurate is likely to be highly problematic, or even impossible. Such a determination would require choosing between alternative assumptions and the resultant model behaviour, based on some conditions. Although evaluation using historic time series of land cover might appear to offer a potential for such criteria, practical and theoretical issues arise. Firstly, there is a lack of a consistent historic time series of land cover data that can be used as a reference, which is itself an output of other models and therefore subject to a range of uncertainties[31]. Secondly, even the ability to reproduce historic land use change does not ensure that future conditions will be adequately represented. Finally, given a single limited series of historic data this may be implicitly or explicitly used to calibrated and tune the model, therefore greatly diminish any inference that can be drawn from reproducing it. The situation contrasts with the modelling of some other systems (e.g. weather forecasting) where models can be repeatedly confronted with unseen data, to allow a measure of model efficacy to be determined.

A potential interpretation consistent with the results is that cropland and European land cover have received greater research focus, leading to lower variance in initial areas, greater consistency between models, and a higher degree of uncertainty represented in the projections. For example, many LULC models derive forest area change from changes in agricultural area, and do not consider factors such as demand for forest products or non-market ecosystem services[19]. Such an asymmetry in focus would be hard to justify as forests cover 31% of the global land surface, and pasture 26%, but cropland only 11%[26]. The focus on cropland may be due to the importance of food production, as crops provide 90% of the global calories consumed by human[30]. However in the context of climate change other land covers are of importance, reinforced by the changes over the past 50 years with pasture accounting for 60% of the expansion in agricultural land[26]. Furthermore, if other land covers have received less attention in the models, then cropland areas may inadequately account for the interactions between demands for other uses such as timber production or other ecosystem services.

A diverse set of LULC models and approaches is required to explore model uncertainty and to ensure that biases in outcomes from a particular approach do not dominate. This is analogous to the situation regarding model uncertainty in climate projections within the IPCC process, which uses results from multiple earth system models developed at different modelling centres[32]. Further research, including model development and more detailed comparisons, are required in an attempt to identify, understand and if appropriate update models to address the sources of these differences. However, uncertainty in future LULC is likely to remain, and possible even increase, as more processes are represented, and scenario and parameter uncertainty is more fully captured. This has implications for assessing future climate change, and the success of land-based mitigation and adaptation options. Currently some of the models do not consider the impact of climate change, further supporting the view that work remains to better evaluate future LULC uncertainty. Similarly, the impact from the level of future uncertainty in LULC demonstrated here may not be fully explored within the parameterisation of many current earth system models[33]. Uncertainties in the coupled LULC and earth system need to be considered, due to the feedback effects that may dampen or amplify responses. Therefore LULC models, as well as earth system models, need to be studied in a way that allows formal uncertainty analysis of the coupled system.

5 Methods

5.1 Models of land use or land cover

Modelled data were obtained from 17 models able to provide scenario results for land use or land cover areas, with either a global or European geographic extent. Research groups covering a further 6 models were approached, but did not submit data. Table 1 gives details for each of the models included in the analysis. No attempt was made to align the scenarios definitions, initial conditions or other model parameterisation. The land use or cover types from each model were used to provide the areas of cropland, pasture and forest. The definition of these types was based on FAOSTAT[26], e.g. pasture is land used to grow herbaceous forage crops, either cultivated or growing wild, and therefore ranges from intensively managed grassland through to savannahs and prairies. All models were able to provide these three types, in some cases by aggregating more detailed types, except CAPS and MAGNET that provided only cropland and pasture areas. The categorisation was selected to avoid some of the definitional issues, e.g. between managed and unmanaged forest, and to maximise the model coverage. Urban and other natural vegetation or unmanaged areas were not analysed due to the lower numbers of models able to provide these types.

Table 1. Summary of models and scenarios data included in the analysis of land cover results.

Model name and key publication	Spatial resolution data (model, if different)	Spatial extent+	Temporal resolution data (model, if different)	Model type (classification)	Scenario descriptions (number of scenarios)
AIM/CGE [14]	17 regions	Global	2005, 2010, 2030, 2050 and 2100 (annual)	CGE	SSP1, SSP2 and SSP3. (3)
CAPS [12]	0.5 x 0.5 degree grid	Global	2005, 2030, 2050 and 2100	Allocation model using demand from CGE or PE model (Hybrid)	SSP3, SSP5, RCP 4.5 and RCP 8.5, each under estimated model parameters from historical data from Ramankutty et al.[27] and HYDE[34]. (8)
CLIMSAVE-IAP [9]	10 x 10 arc-minute grid	Europe (EU27+2)	2010 and 2050	Rule-based	SRES A1, A2, B1 and B2, each under current baseline and the socio-economic factors for the SRES scenario*. (8)
CLUMondo [35]	9,25 x 9,25 km grid	Global	2000 - 2040; decadal (yearly)	Allocation model using demand from CGE or PE model (Hybrid)	OECD scenario. (1)
CRAFTY [36]	1 x 1 km grid	Europe (EU27)	2010 - 2040; decadal	Agent-based model (Rule-based)	SRES A1 and B1. (2)
DynaCLUE [10]	1 x 1 km grid	Europe (EU27)	2000-2040; decadal	Allocation model using demand from CGE or PE model (Hybrid)	SRES A1, A2, B1 and B2. (4)
EcoChange [37]	250 x 250m grid	Europe (EU25+2)	2010, 2020, 2050, 2080	Rule-based	Three core socio-economic scenarios, growth and globalisation, BAU, and sustainable development, and three shock scenarios, climate, energy price and pandemic shocks. (6)
FABLE [38]	Global	Global	2005-2105; annual	PE	Baseline consistent with SRES A1B and RCP 2.6, with other scenarios adjusting population, climate to RCP 8.5, oil prices, economic growth, and more stringent GHG emission regulations (6)
FALAFEL [39]	Global	Global	2000 - 2050; decadal	Rule-based	SSP1, SSP2, SSP3, SSP4 and SSP5. (5)
FARM [40]	13 regions	Global	2005 - 2050; five year steps	CGE	SSP1, SSP2 and SSP3, each under the current climate and climate scenario RCP 4.5, RCP 6.0 and RCP 8.5, respectively*. (6)
GCAM [11]	32 regions	Global	2010 - 2100; decadal	PE	SSP1, SSP2, SSP3, SSP4 and SSP5. (5)
IMAGE [13]	0.5 x 0.5 degree grid (5 x 5 arc-minute)	Global	2010, 2030, 2050 and 2100 (annual)	Allocation model using demand from CGE model (Hybrid)	SSP2 reference and high bio-energy demand scenario under RCP 2.6. (2)
LandSHIFT [41]	5 x 5 arc-minute grid	Global	2005-2050; five year steps	Rule-based	Fuel and heat scenarios, with both BAU and regulation assumptions for each. (4)
LUISA [42]	100 x 100m grid	Europe (EU28)	2010 - 2050; decadal (annual)	Cellular-automata and statistical model (Rule-based)	Reference scenario. (1)
MAGNET [43]	26 regions	Global	2007, 2010, 2020, 2030, 2050 and 2100	CGE	SSP1, SSP2 and SSP3. (3)
MAgPIE [44]	0.5 x 0.5 degree grid	Global	1995-2100, five year steps	PE	Scenarios based on SSP2, with and without bioenergy CCS. (2)
PLUM [45]	157 countries	Global	1990-2100; annual	Rule-based	SRES A1, A2, B1 and B2 (4)

Notes:
+ EU27 is current 28 European Union members (EU28) less Croatia. EU25+2 additionally excludes member states of Romania and Bulgaria, i.e. EU25, but includes Norway and Switzerland.
* CLIMSAVE-IAP and FARM provided results for multiple climate models under otherwise the same scenario, the mean figure for each scenario/model combination was used.

5.2 Processing of model results

To provide a consistent dataset: points for missing decadal end years were linearly interpolated between the result-years provided to decadal end years from 2010 to 2100; the aggregate land cover areas for global and European (taken as EU27) extents were extracted; and the areas scaled to match historic data from FAOSTAT[26] at 2010. Further details are available in the SI.

5.3 Scenarios

Research groups submitted results for multiple scenarios, to explore the possible space of potential land cover results. A total of 70 scenarios were used (Table 1), including business-as-usual and those with mitigation measures. No attempt was made to align the inputs between models, and consequentially the results are not based on the same set of scenarios or parameterisation data. The majority of the scenarios were either SSP or SRES based, however in some cases parameters were adjusted away from the scenario baseline values, e.g. FABLE. Alternatively, some models have conducted experiments where either the socio-economic or climate scenario was held at current baseline levels, within an otherwise SSP or SRES scenario, e.g. FARM and CLIMSAVE-IAP. It is therefore not possible to fully describe the scenarios by mapping them onto a small number of similar categories (as done by Busch, 2006[46]). Additionally, there are difficulties in mapping between SRES and SSP/RCP[47]. Consequently, scenarios were described by a series of values, with default values obtainable for both SRES and SSP[15,48] (Table S1). The aim was to characterise the scenarios in a way that is consistent with the scenario, rather than specify the exact inputs used. Where a parameter differs from the default, the adjusted figure was used for that scenario. Table S2 gives the resultant characterisation.

5.4 Statistical analysis of model results

The aim of the statistical analysis of the model results was to identify the sources of variance associated with aspects of the models or scenarios. The analysis identified the significant variables with a multiple linear regression of the areas for each land cover type, year and spatial extent. The observed variance was then partitioned into components attributed to the selected variables in an analysis of variance approach (ANOVA), to quantify the sources of variability in the results.

The modelled area for each land cover type and year was assumed to be a multiple linear function of 10 variables (Table S3). The factors used can be classified into three groups, those associated with the model, the scenario or the initial conditions. The models were described by 3 variables; model type, number of cells, and the model extent. The scenarios were described by 5 socio-economic variables and the CO_2 concentration, as a proxy to the climate scenario. The initial condition delta represents the difference between the model result and historic baseline in 2010[26]. The regression fitting process was conducted for the three land cover types considered at the decadal end years 2010-2100. To avoid over-fitting, and to identify the predictive variables of the modelled areas, an Akaike information criterion (AIC) approach was used[49]. The least significant variables in the candidate regression model were iteratively reduced, to minimize the AIC score. This accounts for the trade-off between goodness of fit and the model complexity, and selects variables for the regression that are of higher significance. The regression results for global cropland at 2050 and 2100 are given as an example (Tables S4-5) and discussed in the SI.

ANOVA was used on the regression model to decompose the variability of the model[23,24]. The type II sum of squares values were calculated for each variable in the fitted regression model. This approach has the important advantage that, unlike Type I sums of squares, they do not depend on the order in which variables are considered, and has been suggested to be suitable for use with unbalanced data[50], although they are not constrained to sum to the total variance in the raw data. The interaction terms were not determined[23], and the variance associated with such interactions would be incorporated within the residual.

6 References

1. Houghton, R. a. *et al.* Carbon emissions from land use and land-cover change. *Biogeosciences* **9,** 5125–5142 (2012).
2. Quéré, C. Le *et al.* Global carbon budget 2014. 47–85 (2015). doi:10.5194/essd-7-47-2015
3. Rose, S. K. *et al.* Land-based mitigation in climate stabilization. *Energy Econ.* **34,** 365–380 (2012).
4. Parry, M. L., Rosenzweig, C., Iglesias, a., Livermore, M. & Fischer, G. Effects of climate change on global food production under SRES emissions and socio-economic scenarios. *Glob. Environ. Chang.* **14,** 53–67 (2004).
5. Howden, S. M. *et al.* Adapting agriculture to climate change. *Proc. Natl. Acad. Sci.* **104,** 19691–19696 (2007).
6. Foley, J. a *et al.* Solutions for a cultivated planet. *Nature* **478,** 337–42 (2011).
7. Tilman, D., Balzer, C., Hill, J. & Befort, B. L. Global food demand and the sustainable intensification of agriculture. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 20260–4 (2011).
8. Weinzettel, J., Hertwich, E. G., Peters, G. P., Steen-Olsen, K. & Galli, A. Affluence drives the global displacement of land use. *Glob. Environ. Chang.* **23,** 433–438 (2013).
9. Harrison, P. A., Holman, I. P. & Berry, P. M. Assessing cross-sectoral climate change impacts, vulnerability and adaptation: an introduction to the CLIMSAVE project. *Clim. Change* **128,** 153–167 (2015).
10. Verburg, P. H. & Overmars, K. P. Combining top-down and bottom-up dynamics in land use modeling: Exploring the future of abandoned farmlands in Europe with the Dyna-CLUE model. *Landsc. Ecol.* **24,** 1167–1181 (2009).
11. Calvin, K. *et al.* Implications of simultaneously mitigating and adapting to climate change: Initial experiments using GCAM. *Clim. Change* **117,** 545–560 (2013).
12. Meiyappan, P., Dalton, M., O'Neill, B. C. & Jain, A. K. Spatial modeling of agricultural land use change at global scale. *Ecol. Modell.* **291,** 152–174 (2014).
13. Stehfest, E. *et al.* *Integrated Assessment of Global Environmental Change with IMAGE 3.0. Model description and policy applications*. (PBL Netherlands Environmental Assessment Agency. The Hague, Netherlands, 2014). at

<http://www.pbl.nl/en/publications/integrated-assessment-of-global-environmental-change-with-IMAGE-3.0>

14. Fujimori, S., Masui, T. & Matsuoka, Y. AIM/CGE [basic] manual Discussion paper series (Vol. 2012-01). *Cent. Soc. Environ. Syst.* (2012).

15. IPCC. *Special report on emissions scenarios: a special report of working group III of the intergovernmental panel on climate change*. (Cambridge University Press, Cambridge, UK, 2000). at <http://www.ipcc.ch/ipccreports/sres/emission>

16. O'Neill, B. C. *et al.* The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century. *Glob. Environ. Chang.* (2015). doi:10.1016/j.gloenvcha.2015.01.004

17. Van Vuuren, D. P. *et al.* The representative concentration pathways: An overview. *Clim. Change* **109,** 5–31 (2011).

18. Von Lampe, M. *et al.* Why do global long-term scenarios for agriculture differ? An overview of the AgMIP Global Economic Model Intercomparison. *Agric. Econ.* **45,** 3–20 (2014).

19. Schmitz, C. *et al.* Land-use change trajectories up to 2050: insights from a global agro-economic model comparison. *Agric. Econ.* **45,** 69–84 (2014).

20. Nelson, G. C. *et al.* Climate change effects on agriculture: economic responses to biophysical shocks. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 3274–9 (2014).

21. Brovkin, V. *et al.* Effect of Anthropogenic Land-Use and Land-Cover Changes on Climate and Land Carbon Storage in CMIP5 Projections for the Twenty-First Century. *J. Clim.* **26,** 6859–6881 (2013).

22. Robinson, S. *et al.* Comparing supply-side specifications in models of global agriculture and the food system. *Agric. Econ. (United Kingdom)* **45,** 21–35 (2014).

23. Nishina, K. *et al.* Decomposing uncertainties in the future terrestrial carbon budget associated with emission scenario, climate projection, and ecosystem simulation using the ISI-MIP result. *Earth Syst. Dyn. Discuss.* **5,** 1197–1219 (2014).

24. Yip, S., Ferro, C. A. T. & Stephenson, D. B. A simple, coherent framework for partitioning uncertainty in climate predictions. *J. Clim.* **24,** (2013).

25. Bates, D., Mächler, M., Bolker, B. M. & Walker, S. C. Fitting linear mixed-effects models using lme4. *ArXIV e-print; Press. J. Stat. Softw.* (2014). at <http://arxiv.org/abs/1406.5823>

26. FAOSTAT. *Resources/Land (2014-04-15)*. (Food and Agriculture Organization of the United Nations, Rome, Italy, 2014).

27. Ramankutty, N., Evan, A. T., Monfreda, C. & Foley, J. a. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochem. Cycles* **22,** 1–19 (2008).

28. Verburg, P. H., Neumann, K. & Nol, L. Challenges in using land use and land cover data for global change studies. *Glob. Chang. Biol.* **17,** 974–989 (2011).

29. Walker, W. E. *et al.* Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. *Integr. Assess.* **4,** 5–17 (2003).

30. Kastner, T., Rivas, M. J. I., Koch, W. & Nonhebel, S. Global changes in diets and the consequences for land requirements for food. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 6868–6872 (2012).

31. Pontius, R. G. *et al.* Comparing the input, output, and validation maps for several models of land change. *Ann. Reg. Sci.* **42,** 11–37 (2008).

32. Solomon, S. *et al.* Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. *Cambridge Univ. Press* (2007).

33. Rounsevell, M. D. A. *et al.* Towards decision-based global land use models for improved understanding of the Earth system. *Earth Syst. Dyn.* **5,** 117–137 (2014).

34. Goldewijk, K. K. Estimating global land use change over the past 300 years: The HYDE Database. *Global Biogeochem. Cycles* **15,** 417 (2001).

35. Van Asselen, S. & Verburg, P. H. Land cover change or land-use intensification: simulating land system change with a global-scale land change model. *Glob. Chang. Biol.* **19,** 3648–67 (2013).

36. Murray-Rust, D. *et al.* Combining agent functional types, capitals and services to model land use dynamics. *Environ. Model. Softw.* **59,** 187–201 (2014).

37. Dendoncker, N., Bogaert, P. & Rounsevell, M. A statistical method to downscale aggregated land use data and scenarios. *J. Land Use Sci.* **1,** 63–82 (2006).

38. Steinbuks, J. & Hertel, T. W. Confronting the Food-Energy-Environment Trilemma : Global Land Use in the Long Run. *Environ. Resour. Econ.* 1–26 (2014). doi:10.1007/s10640-014-9848-y

39. Powell, T. Closing loops to rebalance the global carbon cycle: Biomass flows modelling of global agricultural carbon fluxes. *PhD Thesis Univ. Exet.* (2015).

40. Sands, R. D., Jones, C. A. & Marshall, E. *Global Drivers of Agricultural Demand and Supply. ERR-174*. (Economic Research Service, U.S. Department of Agriculture, 2014).

41. Schaldach, R. *et al.* An integrated approach to modelling land-use change on continental and global scales. *Environ. Model. Softw.* **26,** 1041–1051 (2011).

42. Baranzelli, C. *et al. The Reference scenario in the LUISA platform – Updated configuration 2014 Towards a Common Baseline Scenario for EC Impact Assessment procedures*. (Luxembourg Publications Office of the European Union, 2014). doi:10.2788/85104

43. Van Meijl, H., van Rheenen, T., Tabeau, a. & Eickhout, B. The impact of different policy environments on agricultural land use in Europe. *Agric. Ecosyst. Environ.* **114,** 21–38 (2006).

44. Popp, A. *et al.* Land-use protection for climate change mitigation. *Nat. Clim. Chang.* **4,** 2–5 (2014).

45. Baumanns, K. *et al.* Applying Occam's Razor to global agricultural land use change (Submitted). *Environ. Model. Softw.* (2015).

46. Busch, G. Future European agricultural landscapes—What can we learn from existing quantitative land use scenario studies? *Agric. Ecosyst. Environ.* **114,** 121–140 (2006).

47. Van Vuuren, D. P. & Carter, T. R. Climate and socio-economic scenarios for climate change research and assessment: Reconciling the new with the old. *Clim. Change* **122,** 415–429 (2014).

48. IIASA. *SSP Database (version 0.93)*. (International Institute for Applied Systems Analysis, Laxenburg, Austria, 2015). at <https://secure.iiasa.ac.at/web-apps/ene/SspDb>

49. Burnham, K. P. & Anderson, D. R. Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociol. Methods Res.* **33,** 261–304 (2004).

50. Langsrud, Ø. ANOVA for unbalanced data: Use Type II instead of Type III sums of squares. *Stat. Comput.* **13,** 163–167 (2003).