Please note that line numbers refer to the 'track changes' version of the manuscript.

Reviewer 1 – Helen Briassoulis

GENERAL COMMENTS

The paper adopts the obvious and established thesis that differences in the underlying theory and assumptions of regional and larger scale LUC models (in fact, of any model) produce different projections of land use change (LUC) patterns, under alternative future scenarios, with different implications for pertinent land use decision making and planning. It presents a comparison of two large scale pan-European integrated land use models, a constrained optimising economic-equilibrium model and a stochastic agent-based model (ABM), which the authors consider as being representative of two different modeling paradigms, under the same set of alternative future scenarios.

This comparison is argued to be necessary to help model users learn about the possibilities and limitations of each modeling paradigm (and particular model) and use them intelligently. In its present form, the main issues the paper should address are: (a) the need for a structured and articulate conceptual/theoretical framework and an associated robust methodology for model comparison to warrant the validity of the results, strengthen their analysis and support their informed, comprehensive interpretation, (b) a clear and transparent presentation (definition, explanation) and use of certain terms and (c) issues concerning the use and users of the models. Selected important issues are detailed below that the authors might want to consider in revising the paper with the aim to improve and enhance its contribution to LUC modeling and discourse as well as to make it less esoteric and idiosyncratic and more accessible to a wider audience than it is presently the case.

We appreciate the detailed comments and suggestions, many thanks.

WRITING STYLE. Although the writing style of the paper is acceptable, it might be written more clearly, coherently and solidly. Several comments below indicate problematic sentences and expressions.

Thank you for identifying sections that need improvement; these suggestions have been followed in the revision and we also carefully revised the manuscript with this comment in mind.

TERMINOLOGY. The authors use, but do not explicitly define, the term 'model paradigm' to refer to the set of underlying model theory, assumptions and structure. Given its central role in the paper, 'model paradigm' should be defined and explained. It seems, though, that the literature uses the term 'modeling paradigm' more often. This term is encountered twice in this paper only (page 3 and page 8). In my opinion, the term 'paradigm' should be used with caution given its strong meaning (Kuhn 1962) and should be justified based on the literature. In the present case, both models considered (optimizing and ABM), in one sense, belong to the same (reductionist) paradigm.

We agree that 'modelling paradigm' is more widely used and appropriate, and have adopted this term throughout the revision. We have also noted the particular meaning of the term paradigm in the modelling literature (lines 75-77; 192-200).

Based on the definition of the modeling paradigm, the paper should justify and explain why the two models compared belong to different modeling paradigms. Part of the text in Discussion and Conclusions (that concerns the models as representative of different paradigms) should be placed in the 'Methods' section.

Good points, and have done as suggested in the revision, introducing the models and their paradigms earlier and in a more structured way (as also suggested in the other reviews) (in particular lines 185-200, also 75-77).

Explain the terms: Representative Concentration Pathways (RCPs) and Shared Socioeconomic Pathways, spatial and aggregate land use change, aggregate comparison, spatial comparison.

We have explained or altered each of these terms.

CONCEPTUAL/THEORETICAL FRAMEWORK FOR THE COMPARISON OF MODELS. It is absolutely necessary to describe the framework/schema used for the comparison of the models; i.e. define the main factors influencing the performance and output of a model (e.g. spatial and temporal frame of reference and resolution, model specification – definition and operationalization of variables, relationships among them, missing variables –, solution techniques, data), exogenous factors and conditions, contingencies (e.g. during data collection). This conceptual framework should serve as a systematic template for the comparison (what is being compared) and constitute the basis for (a) designing the methodology and (b) presenting and interpreting the results.

We have described and used this framework as suggested, adding a model comparison table (the new Table 3) to present these factors (many of which are the same or similar in both models), a model structure diagram (the new Figure 1), and additional text as necessary (in particular Tables 1 and 2 in addition to Table 3, and lines 372-379, 418-425 amongst further clarifications detailed below). We have also used the framework to structure the methods and results (lines 185-218, 239-244).

COMPARISON METHODOLOGY. This should draw on the conceptual/theoretical framework and used as a basis for the design of the comparison of the models so that the factors that are not taken into account in the present comparison are held constant (controlled).

Yes, we agree, and have presented the methodology in this way in the revision (key text in lines 186-191, 239-244).

DESCRIPTION OF MODELS AND SCENARIOS. In the present form of the paper, the models and scenarios are not completely and systematically described; e.g. aim, land classes, land users, ecosystem services, other variables, etc. This makes understanding of the results and the discussion difficult.

We are sympathetic to this point, and have included substantially more detail in the text, new tables (1-3) and diagram (Fig. 1) mentioned above. We have also added new scenario descriptions in the new Table 1. We must also note that it would be impossible to fully describe two such (previously published) models in each paper that uses them, not least because we would have to pay substantial page fees to reproduce text already available in open access publications (notably the linked article of Brown et al. 2019 in the same journal).

The similarities and differences between the two models must be clearly shown (in a Table?) – variables considered, data, etc.

We agree this would be helpful and have added a table as suggested (Table 3, with key differences in outcomes in Table 1).

The limitations of ABMs should be mentioned; the emphasis on behavioural (and indirectly cultural) issues is not enough. Institutional and political issues play an important role as they are the quintessence of LUC decisions.

We’re not entirely sure what limitations are referred to here, but the absence of institutional and political behaviours is neither a universal feature of ABMs nor unique to them, and indeed they are included in this ABM but not used here (clarified in lines 133-136). We have added text to highlight the importance of ‘missing’ factors (lines 428-430, 439-442).

It should be made clear that this modeling exercise concerns food security, rural land use and associated environmental concerns. Other model aims are possible that may not be well represented by the particular models compared and/or may not yield the same projections of LUC patterns.

A good point, and we have revised as suggested (lines 16-17 and added relevant details throughout as described above).

The instrumental assumptions (e.g. homogeneity, uniformity and similarity of land classes, land users, decision makers, etc. across Europe) of the comparison exercise should be made clear and taken into account in the discussion and interpretation of the results.

Yes, we agree this is important and have clarified and accounted for these assumptions (incidentally they do not include homogeneity and uniformity of land classes, as also clarified) (Table 3, lines 123-127, 223-225, 345-346, 428-430, 439-442).

The issue of aggregation (spatial, temporal definitional) of model inputs and outputs is not discussed.

We have added explanation and discussion of these issues in Table 3 and lines 223-225, in addition to Tables 4 and A1.

PRESENTATION OF RESULTS. At present, it is descriptive and technical and does not account for the instrumental assumptions of the comparison.

The interpretation and explanation of the results should be made within the limitations of the models (not simply noted at the end); in particular, their aggregate and reductionist nature, the very coarse spatial and temporal scale, the use of a grid that may not coincide with ecological and socio-political boundaries and conditions over the study area, the jurisdictions within decisions are made, MAUP issues, the effects of the unstructured clusters (IMPRESSIONS), instrumental assumptions, etc.

We have acknowledged and provided more detail about these issues (and other limitations) more prominently (Table 3, new text throughout the Methods and Discussion sections, and in particular in the new section 2.3), and present and interpret the results using the framework identified above (focusing on the differences between the models, both in terms of capabilities and limitations).

It should be underlined (and explained) that the study findings may not be applicable to lower scales.

Revised as suggested (lines 428-430, 439-442, as well as clarification of resolution issues in Table 3 and elsewhere in the text).

MODEL USE AND USERS. The paper should dedicate some space (Introduction, Conclusions) to the discussion of the users and uses of these models. Who are, or might be, the current and prospective users of these European scale aggregate models? Are they interested in these large scales and very long time horizons, especially under conditions of serious uncertainty? Do they have authority to make land use decisions and guide LUC at this scale and over long time horizons? Are they really interested in providing ecosystem services? Which services in particular (use the MEA classification). Have these models been used in real world decision making and by whom? At some point, the authors state: “Optimising models have the advantage of representing idealised conditions”. Who decides what these ideal conditions are and for whom?

We have summarised these issues in Table 3 and in lines 201-210, and clarified the 'idealised' results (line 432). The models are primarily intended and used for exploratory modelling in academic research, with the IAP in particular having also been used for capacity building (with stakeholders and students).

SPECIFIC COMMENTS TITLE Two terms in the current title "How model paradigms affect our representation of future land-use change" may have to be modified. Model paradigms or modeling paradigms? Representation? The paper discusses estimates of future LUC, not representation which is something different and, in any event, it is reflected in the pre-defined land use classes and patterns used in the models.

Yes these are good points. We now use 'modelling paradigms' instead and changed 'representation', with a new title of 'How modelling paradigms affect simulated future land-use change'

"optimisation may be appropriate in scenarios that allow for coherent political and economic control of land systems, but not in scenarios where economic and other scenario conditions prevent the normal functioning of price signals and responses." This is correct: model results are valid if model assumptions hold… but caution is needed here. Reality may change and render model assumptions invalid…

Yes we entirely agree and have reworded to clarify our meaning (lines 22-25, 444-445).

What is the 'normal' functioning of price signals?

This was poorly phrased. We now use 'equilibrium' instead.

"structured comparisons of parallel, transparent but paradigmatically distinct models are an important method for better understanding the potential scope and uncertainties of future land use change" It is not clear that the present comparison is structured, or, at least, its presentation is not adequate and clear (see comments on conceptual framework and methodology). Parallel? Explain. Caution: understanding model results is not tantamount to understanding reality… Who wants to understand the potential scope and uncertainties of future LUC? (the issue of users mentioned above).

The structure of the comparison is now clarified as described above, as are the respects in which the models are 'parallel'. We agree that model results are distinct from reality, and in addition to the new text/details on model uses and users we now emphasise this point in lines 201-210.

"The optimisation model, in contrast, maintains food supply through intensification of agricultural production in the most profitable areas, sometimes at the expense of active management in large, contiguous parts of Europe."

This is unclear… active management? large, contiguous parts of Europe?

Now rephrased (lines 21-22).

INTRODUCTION
"Computational models of the land system are essential in supporting efforts to limit climate change and reverse biodiversity loss (Harrison et al. 2018; Rogelj et al. 2018)"
These are not the only reasons for using these models. I suggest to broaden this sentence to encompass environmental and socio-economic change.

Revised as suggested.

"The need to radically alter human land use to avert social-ecological breakdowns".

This is unclear. Is it supposed to describe the aim of planned and/or unplanned land use change? If yes, the statement should be obviously modified.
Moreover, LUC modeling is used to analyse the impacts of past and current LUC, not only to project LUC under future scenarios.

This sentence refers simply to the unsustainable impacts of current land use. We have clarified this and now mention other model uses (lines 34-36).

“Because other methods are not available to generate alternative findings”
Of course, there are other methods, both quantitative and qualitative (e.g. Delphi dating back to the 1970s), as well as mixed methods.

Yes this was poorly phrased and has now been altered accordingly (line 44).

“This could be particularly misleading in social systems such as those underpinning human land use, where no universal laws or predictable patterns exist to guide model development, and modellers must instead choose between a range of contested theoretical foundations, practical designs and evaluation strategies (Brown et al. 2016; Meyfroidt et al. 45 2018; Verburg et al. 2019).”
This is a confusing sentence, especially the first part. It mixes up several notions and issues. It should be simplified and clarified.
Comment: predictable patterns are rare in nature also as complexity theory underlines and experience reveals.

We have split and revised this sentence (lines 46-49). We agree that predictability is rare in natural systems, but regard it as a particularly important issue in human/social systems where human behaviour introduces extra challenges beyond complexity.

“In this complex context, the proper analysis and interpretation of model outputs is just as important as proper model design”
Irrespective of context, the good modeling practice starts from the theory (however instrumental this may be) about the problem/situation modeled, that guides the development of the methodology, model design (model specification) and implementation (analysis), and, of course, the interpretation of model results.

We agree, and have emphasised that the current study aims to improve interpretation in particular.

“Steps such as standardised model descriptions, open access to model code, robust calibration and evaluation, benchmarking, uncertainty and sensitivity analyses are all necessary to ensure that model results are used appropriately (Baldos and Hertel 2013; Sohl and Claggett 2013).”
What is the meaning of the word ‘used’ here? Model users are not model builders and vice versa. Clarify and modify. Moreover, the real test of a model’s usefulness is model verification, not simply validation, however difficult to carry out (see, O’ Sullivan).

We have changed ‘used’ to ‘interpreted’. Model uses certainly extend beyond the uses model builders put them to, and accurate interpretation underpins them all. We also agree that model verification is crucial and involves more than the steps we mention here, and have amended the sentence to reflect this.

“However, while comparisons of model outputs have been made (Lawrence et al. 2016; Prestele et al. 2016; Alexander et al. 2017), their ability to link particular outputs to particular methodological choices has been limited.”
This is absolutely reasonable because the factors affecting model performance and results are interdependent, important factors may be missing and/or intangible, data are unavailable or inadequate, contingencies modify system characteristics and relationships, etc.

There are certainly limits to linking model design to model outputs, but we would suggest that these limits have not yet been reached, partly because model comparisons have been relatively few and relatively limited in their scope, for instance not including detailed comparisons of quite distinct but parallel models (in the sense of sharing application coverage, resolution, contextual data etc.) of the kind that we make here. We now more clearly present this as the motivation for the study (lines 64-74, Section 2.3), and carefully define the basis of the comparison in terms of model similarities and differences as described above.

"Conceptual research suggests that large areas of system behaviour remain under-explored as a result (Brown et al. 2016; Huber et al. 2018; Meyfroidt et al. 2018), with the likely consequence that established findings have implicit biases and blind spots."
Yes, this is very true and should problematize efforts at comparing different models.

We have extended this point to acknowledge that comparisons can illuminate some but not all such biases and blind spots (line 73).

METHODS
This section should start with a description of the conceptual framework of the comparison.

Now added in Section 2.3 (we do not place this at the start of the Methods section as we feel the need to establish some of the model details first, but also explain the conceptual basis of the comparison in the new text in the Introduction and earlier in the Methods).

IMPRESSIONS IAP. "Within this cross-sectoral modelling chain, rural land use is allocated within each 30-year timeslice according to a constrained optimisation algorithm that maintains equilibrium between the supply and demand for food and (as a secondary objective) timber, through iterating agricultural commodity prices (cereals, oilseeds, vegetable protein, milk, meat etc.) to promote agricultural expansion or contraction (Audsley et al. 2015)."
So, IMPRESSIONS' aim is to optimized food supply? This should be mentioned from the outset.

We now clarify from the outset that the model aims to satisfy food demand (taking account of net imports), and does so optimally subject to constraints imposed by biophysical and socio-economic conditions (lines 118-119, Table 3).

"similar production conditions (based on soil and agroclimate),"
Production conditions include many more factors, such as economic, fiscal, technological, institutional, etc.

Yes, we were referring more specifically to biophysical conditions here, and have amended as such ('similar biophysical conditions (based on soil and agroclimate)') (line 120, also 120-127).

"profitability thresholds used to determine which land use and management intensity is allocated to each cluster."
There are other factors (planning, policy, cultural…) that affect land use allocation and management intensity.
Moreover, caution is needed to interpret these findings given the spatio-temporal and conceptual (land use classes) aggregation of large-scale models. The 'managers' are not real people… so, what is the meaning of profitability? Profitable to whom?

We describe land allocation in both models more fully in the revision, and in particular the role of these and other additional factors (many of which affect allocation) – Tables 1-3 and lines 118-127, 157-158, 186-191, 201-218, 223-224, 345-346, 428-430, 439-442. Profitability is used here to mean the simulated profit available for a particular level of production in a particular cell, and indeed does not refer to profit to real land managers, which we now emphasise.

"Land use proportions within each 10' x 10' grid cell represent the aggregations of the solutions for each (up to 40) associated cluster."
What is the meaning of so aggregate results?

We now clarify the derivation and interpretation of the aggregated results in Table 3 and lines 119-127. The aggregation is the result of spatial weighting of the optimised land use solution for each cluster containing the grid cell in question. The clustering recognises that different biophysical conditions (soil and agroclimate) differentially influence the suitability, productivity and profitability of different crops and different agricultural systems (arable, dairy etc.), leading to heterogeneity in agricultural land use within a grid cell.

"Modelled land manager agents compete for land on the basis of their abilities to produce a range of ecosystem services that society is assumed to require…"
How do you know that this is the aim of the land decision makers, especially at such coarse level of aggregation? I.e. to produce ecosystem services?
Psychological (emotional, political and institutional factors) regulate their relationships.
Power relations are also important determinants of land managers behaviour (however coarse their representation is).

This text describes processes in the model, as distinct from reality – the competition for land on the basis of ecosystem service provision is a modelling assumption analogous to the allocation of land uses on the basis of profitability in the IAP. We describe the inclusion/exclusion of particular factors in more detail in the new model description table (Table 3) and also address missing factors and processes in the text as described above, as well as making it clear that we address differences between the models here and not differences between them and observed land use change.

"Land use productivities"
These should be defined taking into account the very high level of aggregation of the models.
One question is: given the high level of aggregation, how much sense do they make as goals of land managers and decision makers?
Another question is: which factors influence these very aggregate productivities?

We have rephrased for clarity (this refers to the yields / ecosystem service provision levels of the different simulated land use systems (crops, grassland, forestry etc.) under the agronomic scenario conditions) (lines 152-153).

"In CRAFTY-EU, these services are crops, meat, timber, carbon sequestration, recreation and landscape diversity. We therefore also compare ecosystem service production levels, which account for exact forms of management simulated in each cell"
The ecosystem services should be first defined in terms of the 4 main groups defined in MEA (2005) and then shown how they are operationalized.
Moreover, the sentence should be edited (crops, etc. are not services…).

We have revised as suggested (lines 138-142).

"In this case, these functions are linear and equivalent for all services, meaning that the benefit of production of each service increases equally per unit of unmet demand."
A very restrictive assumption indeed…

This is deliberately restrictive, yes. It allows us to compare an equal weighting of service provision (in CRAFTY) with a focus on food production (in the IAP), and avoids a more complex but equally arbitrary weighting that would make results harder to interpret. We now highlight this in line 145.

“Importantly for this study, CRAFTY-EU is parameterised on the basis of the IAP, taking IAP outputs as exogenous conditions and replacing only the land allocation component to provide alternative land use projections under identical driving conditions.”
This is unclear.
Comparison of two models when one takes input from then other?

We have now extensively clarified this point, presenting the relationship and inputs/outputs linking the two models in the new Fig. 1, and giving further description in Tables 2-3 and throughout the Methods section.

“For ecosystem services with economic values (meat, crops and timber), agents in CRAFTY therefore make production choices consistent with this basic level of economic rationality.”
First question: who are these agents at such a high level of aggregation and what is the meaning of economic rationality in this case?
Second question: what about non-economic benefits?

We have revised this sentence for clarity (lines 160-162). As described in the text, modelled agents do not correspond to real-world actors, but are used to capture elements of their behaviour within localised land systems. Non-economic benefits are also included in CRAFTY.

A note regarding NUTS2: They do not represent a uniform, EU-wide spatial division system and they differ significantly among countries.

We agree. There is no ideal resolution at which to make this comparison, but we chose NUTS2 as an established system to complement the results we present at cell and European scales.

Subsection: 3.1. Aggregate comparison
The presentation of the results is descriptive and technical.
The discussion is rather loose and tiresome with reference being made to the scenarios that have not been adequately described. The presentation of the results is rather boring and may not make a lot of sense to the reader.

We have added description of the scenarios and results in the revision in Table 1. We have also edited the text to ensure clarity and interest where we can.

One question here is: What is being interpreted? Model land use classes or real world land use?
Also, the processes that produce LUC differ between countries and for each land use class, among other factors.

Model land use classes are being interpreted. We now clarify this prominently in describing the conceptual framework of the exercise (section 2.3). We agree that processes differ and have added discussion of this important point as described above.

“because of the gradual decision-making of agents “
What does gradual decision making mean?

This refers to agents’ decisions having some probability of being delayed across multiple timesteps (representing years), rather than taking immediate effect. We have explained this in lines 282 and Table 3.

“Conversely, CRAFTY responds most strongly to scenarios in which agricultural productivity decreases because its design emphasises changes in capitals that support production (climatic or socio-economic), as is particularly clear in SSP3.”
The meaning is unclear.

Now rephrased (lines 287-294).

Subsection: 3.2. Spatial comparison
This is not spatial but 'geographical' comparison because it refers to geographic areas in Europe. The term 'spatial' is a general term and NOT identical to 'geographical'.

We now label this as a geographical comparison.

"In SSP4, the IAP projects substantially more very extensive and forest management than CRAFTY's more intensive results,"
extensive WHAT?
more intensive WHAT?

Now revised (line 300).

Subsection 3.3. Convergence experiment
The convergence/divergence of the results of different models owes to a host of factors, several of which were not examined here.
So, I wonder what is the meaning of carrying out this experiment?

We now clarify the purpose of this experiment in the text: it is indeed intended to identify these factors in this particular case (lines 178-180, 243-244, 320-324). The observed divergence in this scenario is partly due to conditions differing in the models (because food prices rise higher in the IAP and production levels fall lower in CRAFTY). By controlling these differences, we are able to identify additional factors that cause divergence – and in this case they reflect basic modelling assumptions, the effects of which would otherwise remain obscure.

4. Discussion and conclusions
I would have preferred a separate and proper Conclusions section.
The discussion of the results might be more meaningful to be combined and integrated with the presentation of results (previous section).

We have written a stand-alone Conclusions section (section 5) but prefer to keep results and discussion separate as we find it important to establish technical findings before interpreting them.

"Understanding the contributions of different modelling paradigms to land use projections is important for two main reasons. The first reason is that almost all large- to global-scale land system models share a single paradigm (economic optimisation of land uses). The second reason is that different paradigms are known to produce very different outcomes, but for reasons that remain unclear"
The question regarding those 'unclear reasons' is: what were the reasons hypothesized in this study? Otherwise, how was the comparison of model results carried out?

In this case, we hypothesise that the decision making / allocation paradigm is one dominant source of uncertainty in land use modelling, as opposed to uncertainty in crop yields, biophysical conditions etc. Hence we keep the latter factors common between the models and explore how different factors that influence decision making (profitability; demand; socio-economic conditions) affect the models. We set this hypothesis out in the revision as described above and use it to structure the methods and results section (in particular the new section 2.3).

"The focused comparison presented here is therefore intended to identify and explain key differences between models representing major, distinct paradigms to project land system dynamics on the basis of complex and integrated processes founded on a small number of key, transparent assumptions".
What are these key, transparent assumptions? Were these key differences explained? The issue of the conceptual/theoretical framework underpinning the comparison is critical here.

We elucidate the underlying framework in the revision with particular emphasis on these key assumptions that differ between the models (section 2.3 and Table 3).

"An overarching distinction is apparent between the basic assumptions underlying the models. The IAP is an example of a 'topdown' model that simulates change at the system-level – in this case through an assumption of constrained economic optimisation - while CRAFTY is an example of a 'bottom-up' model that simulates change at the level of individual decisionmakers – in this case through an assumption of behavioural choices made at the level of local land systems (Brown et al. 2016).
This basic difference affects the rate, extent and pattern of simulated land use change. The consequences of top-down and bottom-up perspectives is apparent in the main forms of land use change as the models respond to scenario conditions…. "
I am not sure if any adequate explanation of the implications of different model assumptions is offered in the above excerpt.

Explanation of implications is in the text that follows the quoted sentences. We believe the additional structured comparison of the models in the revision helps to explain the assumptions and their implications.

"This difference is also apparent in our convergence experiment, where increased imports in the IAP lead to reduced agricultural area,"
But, doesn't it happen the other way around? Land is abandoned, then production drops and necessitates increase in imports… Do I miss something?

We now clarify this in lines 320-324 and Fig. 5. The convergence experiment involved pre-emptively increasing imports in the IAP to mimic the lower European production levels generated by CRAFTY.

"One consequence of simulating demand and supply of a range of ecosystem services is that the relative economic support available for food production becomes a key determinant of the balance of different land uses"
What is the theoretical explanation offered?
The case may be that, because agriculture is the most extensive land use and occupies a larger number of cells, it leads to the results obtained. In other words, the results may owe to technicalities and not to real world market and social behaviour and responses.

An interesting point, and one we address in the revision in lines 391-393. The result is certainly due to technicalities, in the sense that the model is sensitive to the relative valuation because that is the basis for simulated land competition, but reflects the reality that land use, as primarily economically-driven, is subject to the relative economic support for food production and for other ecosystem service provision.

"In both models, the simulation of the European land system as distinct from the rest of the world requires implicit assumptions about conditions in other regions and their relationships to Europe. As conceptual alternatives, therefore, neither of these necessarily capture the true dynamics of food prices and production levels, which remains a major challenge for land system modelling (Pedde et al. 2019; Müller et al. 2020)."
This is a correct remark. The exogenous factors have been incompletely modeled. Their inclusion may have further differentiated the results of the two models.

Yes, we agree. We have added a sentence to emphasise that different approaches to modelling exogenous factors would likely introduce even greater differences (lines 403-404).

"Cell level decisions"
Do cells decide? ( ). Which theory concerns cells?

We have rephrased as ‘simulated decisions affecting individual cells’ (line 410).

“Indeed, their primary strength may be their ability to use theory as a guide to processes and conditions that empirical data and optimising models do not cover (Gostoli and Silverman 2020). “
This is partly true re ABMs. The question is: which theory do they use?

We have added a comment emphasising the choice of theory (line 438), although we see this as a distinct issue to the inclusion of a wider range of theories as an end in itself.

“The greatest value of these two approaches may therefore lie in their ability to provide alternatives; a value that is realised only in the (currently rare) cases when model assumptions are clearly communicated and when analogous models such as those used here are available for comparison (Polhill and Gotts 2009; Müller et al. 2014; Rosa et al. 2014).”
This sentence is unclear.

Now rephrased (lines 444-448).

THE LAST PARAGRAPH of the paper is a rather unstructured list of open issues and future research directions that does not flow directly from the preceding analysis and does not offer much direction centered around a concrete model aim…

This paragraph has now been edited and is followed by a new conclusions section.

The question is: why is it necessary to keep these modeling paradigms when alternatives are already tested and more meaningful? E.g. multi-paradigm modeling.

We are not entirely sure in what sense ‘multi-paradigm modeling’ is being used here, but the models and paradigms represented here have also been tested and found to be meaningful in a number of ways (and are of course widely used). In any case, to the extent that different paradigms are present within ‘multi-paradigm modeling’, we see our basic premise of understanding how underlying assumptions influence model outputs as still relevant.

It might be useful to discuss in the conclusions, issues of model users, use and usefulness that might further justify pertinent future research

Yes, we find this a good suggestion and added some discussion of models users/uses as described above.

TECHNICAL COMMENTS
Table headings should be placed at the top of the Table.
P. 23
The heading of the Table is long … it should be much shorter!
ecosystem service supply … SERVICES

We have placed Figure and Table headings according to the journal’s template, and retain the information in the heading of Table A1 as we find it important for interpretation. Ecosystem services has been changed.

<u>Reviewer 2:</u>

Overall this is a good paper with some really interesting results. With some additional improvements to the figures/analysis it could be excellent and make an excellent contribution.

Thanks for the positive comments and the suggestions.

I am mostly viewing this as a scholar who uses land use projections, and the discussion of the different approaches and how they differ is really important. I like the introduction in general but a bit more detail would be appreciated.

We have now added more detail, in common with the responses to R1.

I would also like the authors to discuss how observational data is incorperated into this models, or not. The usual standard for earth system science, is a lot of comparison to observations, so please explain how each of these paradigms try to make sure they actually compare well to observations, especially of historical land use change trajectories, or if they do not do such comparisons. If currently there is no comparison, this could be a way to differentiate these different approaches to see which is more accurate.

Thanks, an important point. We now detail specific and general use of observational data in the revision, particularly in the model description/comparison table (Table 3) and in lines 420-425, with the new Section 2.3 also being relevant in places. In general, optimising models have indeed been more often compared against data but we provide specific details for these models.

I also think a bit more analysis could be helpful in the figures to synthesize a bit more. Details below.
Figure 2a: the dark (IAP) vs. lighter (CRAFTY) symbol isn't clear enough here: I recommend you make more of a matrix with left being light colors and right being dark colors and showing then that the right ones are IAP and left ones are CRAFTY. I stared at the plot for awhile before I understood what the dark grey and light grey was in the figure.

Thanks for the suggestion; we have taken this approach.

Figure 2b: the colors aren't really different enough here, and the same issue with the dark vs. light colors.

The colours have been changed.

Figure 2 in general: Would a difference plot work better for this? Or a % difference? There are so many similar bar graphs?

We were unsure how well this worked, so have retained the bar charts. While there's a lot of information in them we think it's probably easier to extract all the information in this format, especially after the colour and legend changes.

Figure 3: white means two things here: not included, and not land? Please try use grey for one of those so this is clearer. Maybe you want difference plots here instead of these contrasting, but similar plots? Are there patterns of where in particular the differences are important that you can find and call out?

We have changed the white 'not included' (no difference) to grey. We also agree there was too much here, so have made a new replacement plot showing the total differences between the models across all scenarios (Fig. 4). This highlights the most persistent and informative differences, and we retain the full scenario plots in the new Appendix B for information.

Figure 4: this graph is not self standing enough: describe what is on the left versus the right, why the arrow goes in the opposite direction on the bottom, everything needs to be explained. Describe the alternative parameterizations briefly here in the figure

caption.

We have amended this figure as suggested, thanks.

Reviewer 3 (Robert Huber):

I think this is a valuable contribution to the discussion on how computational models can inform political and social efforts toward more sustainable land systems on a large spatial scale. I find it very important to explicitly discuss underlying paradigms in models of land-use change and I think this contribution is an important step to improve our understanding of how the paradigm affects the interpretability and validity of such models. In my view, however, the current version of the manuscript could be improved by describing the model paradigms more explicit and by a more careful presentation of the input-output relations in the result Section. In addition, I think the contribution would gain from discussing the implications of the different paradigms to inform what the authors call "efforts to limit climate change and reverse biodiversity loss". I would like to specify my general comment below.

Many thanks for the positive feedbacks and useful suggestions.

1) I think it would be important to introduce the two model paradigms earlier in the manuscript. In my understanding, the first part of the discussion (lines 239-260) is something that defines the research design and should not appear as something the authors would like to discuss. In addition, I think it would help the reader if the authors would also discuss and classify/categorize these two paradigms a bit broader e.g. in the context of their own work on ABMs and their theoretical and philosophical background (Arneth et al., 2014; Brown et al., 2016).

We followed this suggestion, moving the text referred to back in the manuscript, detailing the paradigms more carefully in the new model descriptions (including the new Table 3) and relating them to earlier publications as suggested (lines 75-77 and Section 2.3).

2) In the same vein, I think that the discussion of the paradigm should also include implications from the different mathematical model implementations. If I understood the models correctly, IAP maintains equilibrium between the supply and demand for food while agents in CRAFTY compete for land-uses having a satisficing behavior including non-economic benefits. The point I'd like to make is that models based on rational economic behavior usually are characterized by switching from one corner of the mathematical solution space to another. I do not know whether this describes IAP adequately. However, the outputs seem to suggest that CRAFTY results are always more balanced than the economically driven IAP results. Thus, I would suspect that IAP jumps to corner solutions. If this is true, then this would also be known before the comparison. There might be other direct implications of the mathematical implementation of the models for the interpretation of the output. This could be introduced and discussed in the context of model paradigms.

A good point, and we highlight issues of this sort that are clear prior to the comparison in lines 211-218, 372-379 and Table 3). This attribution is slightly complex – for instance the IAP simulates and aggregates up to 40 clusters in each grid cell that produce different solutions, with changes also possible within a land use class (e.g. different crop selections) (as now clarified in the text), but there are certainly elements of model design that mathematically pre-define model outcomes.

3) I also found it difficult to follow the input-output description in the text but also the figures. The authors use the pre-defined abbreviations for the different climate and socio-economic scenarios. I understand that there are reasons not to give explicit names to these scenarios. Nevertheless, it

makes the presentation of the comparison in this contribution very demanding. As a reader unfamiliar with the exact definition of each of the socio-economic scenarios, I always had to cross check what SSP3 or SSP1 now exactly implies with respect to the input assumptions. Since there is no description of the socio-economic scenarios in Section 2.2, I had to use O'Neill et al. to be able to follow the result Section. In addition, I did not really understand how the convergence scenario was developed. I think the manuscript would profit from a concise description of the socio-economic scenarios and how these scenarios affect the underlying assumption in the model exercise e.g. production functions, demand levels etc. This would help the reader to connect input- and outputs in the different models. Personally, I would find it also helpful if there would not be abbreviations for the socio-economic scenarios. This could make it easier and more accessible to the reader e.g. in Figure 1.

Thanks for highlighting this. We have added a scenario description and implementation table (Table 1), and also use the acronyms alongside more meaningful names on most occasions. We've also provided substantially more detail about how the models represent the scenarios (Tables 2 and 3, Fig. 1) and fuller description of the convergence experiment (320-325, Fig. 5).

4) In this context, I also had the impression that the authors did not adequately address and discuss the uncertainty with respect to model inputs. For example, the author writes that there is a "greatly increased productivity" in the scenario RCP8.5-SSP5 and consequently, the IAP model suggests that the supply of crops and meat can increase more than 30% with less than a third of the area of intensive agriculture (comparing Figure 1 and Figure 2a). The increase in productivity, in contrast, did obviously not affect land-use in CRAFTY. However, I would expect that a change in the productivity increase would considerably lower the extreme solution in the SSP5 scenario. Maybe that is something the authors wanted to address with the "convergence" comparison: Look at the sensitivity to input parameters of specific importance. I think this would deserve more attention. Maybe the authors can include more than just two input variations (increase in imports and food values) and discuss the results in the context of input uncertainty that seems to have very different impacts in the two model paradigms.

An interesting point, and it's quite correct that we didn't deal with uncertainty/sensitivity in any depth. To rectify this we give details of uncertainties and sensitivities in the model descriptions and in the results & discussion (Table 3, Section 2.3, lines 372-380, 474-480). While we're wary of adding more experiments here in addition to substantial extra explanation as suggested by the reviewers, both models have been quite extensively assessed in sensitivity and uncertainty analyses in the past, including with respect to scenario conditions, and we therefore hope that the summaries of these findings we now provide improve interpretation of the findings.

5) With respect to the methods, I acknowledge that these are well documented and state-of-art models that are suitable for comparing the effect of different model paradigms on future land-use change. However, one sentence in the manuscript confused me. The authors write (lines 113ff): "CRAFTY-EU is parameterised on the basis of the IAP, taking IAP outputs as exogenous conditions and replacing only the land allocation component to provide alternative land use projections under identical driving conditions." What is implied here by taking the output of IAP as exogenous conditions for CRAFTY. It would not make sense to use outputs as inputs in another model and then conclude that the models have different outputs. I'm sure this is a misunderstanding (culpa mea). However, the authors should be clearer in what they do here. What are these conditions (except land-use) and how do they affect the comparison? Maybe the solution here goes hand in hand with the reply to my comment 3. However, I would suggest that the authors explain the data exchange between the models in more detail.

Yes agreed, and we now explain this properly in the revision, including via a model diagram (Fig. 1) that shows the relationship of the two models and their input/output sharing. We also give extra details about model inputs and outputs in Tables 2 and 3, as well as at several points in the text.

6)The last comment I'd like to make is probably also the most difficult to address. When looking at the results, I had the impression that the two model paradigms lead to really large differences despite using the same scenarios (e.g. in Figure 3). The authors also state that the "greatest value of these two approaches may therefore lie in their ability to provide alternatives". But if these models should inform "efforts to limit climate change and reverse biodiversity loss" what do these alternatives imply? Obviously one would come to very different conclusion what to do concerning e.g. biodiversity loss depending on the model paradigm (irrespectively of the scenario). Given the potentially contradicting (policy) conclusions from these "alternatives", critics of mathematical modelling could argue that this "invalidates" such simulations. One can get any result by choosing the "right" paradigm. I'm aware that the contribution does not attempt to address all of the caveats in model design, analysis and interpretation mentioned in the Introduction. However, I had the impression that the authors take refuge in discussing "technical integration" of models. But how could such a hybrid modelling approach solve potential contradictions? In climate change modeling, model ensembles are a way of addressing different underlying functionalities of models. However, it seems to be impossible when looking at the results of this exercise. I think this point should at least be discussed: what if model paradigms prevent instead of foster discussions on how to use modelling of more sustainable land systems on a large spatial scale? I have the impression that the authors should also discuss the value of theoretical underpinnings and conceptual frameworks (which may be more important in this context) than just "more data from another discipline on another spatial level" (which is my simplified interpretation of the last paragraph).

We find these excellent suggestions and fair criticisms. It is probably true that we take refuge in technical issues to some extent! This is partly because we wish to establish basic differences here, but we should have better addressed this overarching issue. We have therefore added text to link our findings to the motivating question of model uses (including in Table 3, and lines 201-210), and return to this issue in particular in the new Conclusions section (section 5).

Minor comments

What is the unit of the Y-axes in Figure 1?

Now defined in the figure legend.

I would prefer if the difference between IAP/CRAFTY in the figures would not be represented by the level of shading only. Maybe the authors can use a different pattern or something that makes it easier to distinguish the models.

We agree this was confusing. We've settled on the suggestion of Reviewer 2 of changing some colours and providing a more informative legend in this revision.

I found the caption in the Figures not self-explaining (and I have to say a bit cryptic in the beginning). I do not really understand why some specific information is given in one Figure but not in the other. I think that the authors should try to make the caption self-explaining (in a way).

We've tried to standardise and complete all figure captions and hope they are now self-explanatory.

On line 302ff, the authors state that "Conversely, (constrained) optimising models like the IAP produce idealized results that (. . .) can use flexible spatial dependencies as proxies for processes such as imitation, diffusion of knowledge or the formation of social norms (). Are you sure that knowledge diffusion and social norms fit into the economic framework of IAP? Not sure I understood this sentence.

It's true that these processes do not really fit, but they can be accounted for to some extent through the use of proxies; especially undefined spatial correlations in land use changes. We have now clarified this point in lines 413-417.

How modelling paradigms affect ~~our representation~~simulated ~~of~~ future land-use change

Calum Brown[1], Ian Holman[2], Mark Rounsevell[1,3]

[1] Institute of Meteorology and Climate Research, Atmospheric Environmental Research (IMK-IFU), Karlsruhe Institute of Technology, Kreuzeckbahnstraße 19, 82467 Garmisch-Partenkirchen, Germany
[2]School of Water, Energy and Environment, Cranfield University, Vincent Building, Bedford MK43 0AL, UK
[3]School of Geosciences, University of Edinburgh, Edinburgh EH8 9XP, UK

Correspondence to: Calum Brown (calum.brown@kit.edu)

Abstract. Land use models operating at regional to global scales are almost exclusively based on the single paradigm of economic optimisation. Models based on different paradigms are known to produce very different results, but these are not always equivalent or attributable to particular assumptions. In this study, we compare two pan-European land use models that are based on the same integrated modelling framework and utilise the same climatic and socio-economic scenarios, but which adopt fundamentally different modelling paradigms. One of these is a constrained optimising economic-equilibrium model and the other is a stochastic agent-based model. We run both models for a range of scenario combinations and compare their projections of ~~spatial~~ geographical and aggregate land use changes and ecosystem services supply levels in food, forest and associated environmental systems. We find that the agent-based model projects more multifunctional and heterogeneous landscapes in most scenarios, providing a wider range of ecosystem services at landscape scales, as agents make individual, time-dependent decisions that reflect economic and non-economic motivations. This tendency also results in food shortages under certain scenario conditions. The optimisation model, in contrast, maintains food supply through intensification of agricultural production in the most profitable areas, sometimes at the expense of ~~active management~~land abandonment in large~~, contiguous~~ parts of Europe. We relate the principal differences observed to underlying model assumptions, and hypothesise that optimisation may be appropriate in scenarios that allow for coherent political and economic control of land systems, but not in scenarios where economic and other scenario conditions prevent the changes in ~~normal functioning of~~ prices ~~signals~~ and responses required to approach economic equilibrium. In these circumstances, agent-based modelling allows explicit consideration of behavioural processes, but in doing so provides a highly flexible account of land system development that is harder to link to underlying assumptions. We suggest that structured comparisons of parallel, transparent but paradigmatically distinct models are an important method for better understanding the potential scope and uncertainties of future land use change.

1 Introduction

Computational models of the land system ~~are~~ make essential ~~in~~ contributions to the exploration of environmental and socio-economic changes, supporting efforts to limit climate change and reverse biodiversity loss (Harrison et al. 2018; Rogelj et al. 2018). ~~The~~Such models are ~~need to radically alter human land use to avert social-ecological breakdowns makes modelling~~ particularly useful for exploring conditions that do not currently exist and cannot therefore be observed, as well as for understanding past and present land use impacts ~~or otherwise understood~~ (Filatova et al. 2016; IPBES 2018; Smith et al. 2019). ~~In order to make this contribution~~As a result, the scope and complexity of land system models have been steadily increasing, with many now representing multiple land sectors (e.g. agriculture, forestry and urbanisation) within an Earth System context (e.g. incorporating economic, climatic, hydrological and energy systems) (Harrison et al. 2016; Kling et al. 2017; Pongratz et al. 2018). ~~These models are used not only to explore ranges of scenarios of future change, but also to develop pathways towards sustainability objectives, such as land-based climate change mitigation (Rogelj et al. 2018; Roe et al. 2019; Papadimitriou et al. 2019).~~

Nevertheless, simulating expected or desired future changes under novel circumstances remains a substantial challenge. Because ~~other methods are not available to generate~~ comparable, alternative findings are rare, model results often go unchallenged, and may be misinterpreted as predictions of how the future will develop rather than projections dependent upon underlying assumptions (Low and Schäfer 2020). This could be particularly misleading in social systems such as those underpinning human land use, where no universal laws or predictable patterns exist to guide models' ~~development~~ representation of human behaviour.~~, and m~~ Modellers must ~~instead~~ therefore choose between a range of contested theoretical foundations, practical designs and evaluation strategies (Brown et al. 2016; Meyfroidt et al. 2018; Verburg et al. 2019).

In this complex context, the proper analysis and interpretation of model outputs is just as important as proper model design, but has received less attention. Steps such as standardised model descriptions, open access to model code, robust calibration, ~~and~~ evaluation and verification, benchmarking, uncertainty and sensitivity analyses are all necessary to ensure that model results are ~~used~~ interpreted appropriately (Baldos and Hertel 2013; Sohl and Claggett 2013). Currently, few if any of these steps are taken universally and rigorously in land use science (van Vliet et al. 2016; Brown et al. 2017; Saltelli et al. 2019). This study focuses on one in particular; the comparison or benchmarking of independent land use models against one another. Comparison is especially important for land use models because a range of very different conceptual and technical approaches could be valid for simulating social-ecological dynamics (Filatova et al. 2013; Brown et al. 2016; Elsawah et al. 2020). In the absence of fair comparisons, it is impossible to objectively choose between these approaches or to identify the assumptions on which their outputs are most conditional. However, while comparisons of model outputs have been made (Lawrence et al. 2016; Prestele et al. 2016; Alexander et al. 2017), their ability to link particular outputs to particular methodological choices has been limited by the sheer number of differences between individual models. Alexander et al (2017), for instance, found that model type explained more variance in model results than did the climatic and socio-economic scenarios, but they were not able to determine exactly why.

These previous comparisons reveal a major challenge: ~~Perhaps the greatest challenge to land use model comparisons is~~ the shortage of models that take distinct approaches at similar geographical and thematic scales, and which would therefore allow for more controlled and informative comparison exercises. Most established models, especially those operating over large geographical extents, share a basic approach that optimises land use against economic, climatic and/or environmental objectives. Technical and geophysical constraints are often treated in detail, while social, institutional and ecological factors are rarely included (Brown et al. 2017; de Coninck et al. 2018; Obermeister 2019). Conceptual research suggests that large areas of system behaviour remain under-explored as a result (Brown et al. 2016; Huber et al. 2018; Meyfroidt et al. 2018), with the likely consequence that established findings have implicit biases and blind spots. These can be especially problematic for the simulation of future scenarios in which neglected aspects of land system change become prominent (Estoque et al. 2020), and can be partially if not fully revealed by structured comparison exercises.

In this article, we take advantage of the development of two conceptually distinct,~~,~~ but practically equivalent models of the European land system to make a direct comparison between alternative modelling paradigms. We~~.~~ use the term ‘modelling paradigm’ here to refer to a coherent methodological and theoretical approach, and specifically the ‘top-down’ and ‘bottom-up’ approaches frequently identified as paradigms in the literature (Brown, Brown, & Rounsevell, 2016; Couclelis, 2002). These models, an Integrated Assessment Platform (IAP) and an agent-based model (ABM) share input data to run under the same internally consistent scenario combinations. The former is a constrained optimising economic-equilibrium model and the latter is a stochastic behavioural model. We run both models for combinations of the Representative Concentration Pathways (RCP~~s~~) climate scenarios and Shared Socioeconomic Pathways (SSP~~s~~) socio-economic scenarios (O’Neill et al. 2017), and compare their projections of ~~spatial~~ geographical and ~~aggregate~~ overall land use change and ecosystem service provision. We use this analysis to understand the effects and importance of the different assumptions contained in each model for simulated land use futures, and draw general conclusions about the contributions of both approaches to understanding land system change.

2. Methods

This paper uses two contrasting models of the European land system: CRAFTY-EU (Brown, Seo, & Rounsevell, 2019) and the IMPRESSIONS Integrated Assessment Platform (IAP) (P. A. Harrison, Holman, & Berry, 2015; Paula A. Harrison et al., 2019). Both models cover all European Union Member States except Croatia, as well as the UK, Norway and Switzerland. The IAP’s simulated baseline land use map, land use productivities, scenario conditions and ecosystem service provision levels were used in CRAFTY-EU, making them uniquely equivalent examples of different modelling paradigms (Fig. 1). Both models were run for a subset of socio-economic and climatic scenario combinations, and their outputs systematically compared, as described below.

~~This paper uses two contrasting models of the European land system: CRAFTY-EU (Brown, Seo, & Rounsevell, 2019) and the IMPRESSIONS Integrated Assessment Platform (IAP) (P. A. Harrison, Holman, & Berry, 2015; Paula A. Harrison et al.,~~

~~2019). Both models cover all European Union Member States except Croatia, as well as the UK, Norway and Switzerland. The IAP's simulated baseline land use map, land use productivities, scenario conditions and ecosystem service provision levels were used in CRAFTY-EU, making them uniquely equivalent examples of different modelling paradigms. Both models were run for a subset of socio-economic and climatic scenario combinations, and their outputs systematically compared, as described below.An overarching distinction is apparent between the basic assumptions underlying the models. The IAP is an example of a 'top-down' model that simulates change at the system-level – in this case through an assumption of constrained economic optimisation - while CRAFTY is an example of a 'bottom-up' model that simulates change at the level of individual decision-makers – in this case through an assumption of behavioural choices made at the level of local land systems (Brown et al., 2016). This basic difference affects the rate, extent and pattern of simulated land use change. These paradigms usually have different uses and justifications: the (dominant) top-down approach is computationally efficient, tractable and more in line with economic theory, although it is rarely justified as an accurate representation of how land use decisions are made in practice (in fact the evidence tends to contradict it; e.g. Chouinard et al. 2008; Schwarze et al. 2014; Appel and Balmann 2019). The bottom-up approach, in contrast, is more exploratory and often criticised for producing uncertain results, but explicitly attempts to achieve greater process accuracy (Brown et al., 2016).~~

2.1~~1~~ Model descriptions

The *IMPRESSIONS IAP* is an online model of European land system change that incorporates sub-models of urban development, water resources, flooding, coasts, agriculture, forests and biodiversity. Within this cross-sectoral modelling chain, rural land use is allocated within ~~each~~ 30-year timeslices according to a constrained optimisation algorithm that maintains equilibrium between the supply and demand for food and (as a secondary objective) timber, through iterating agricultural commodity prices (cereals, oilseeds, vegetable protein, milk, meat etc.) to promote agricultural expansion or contraction (Audsley et al., 2015). This model therefore aims to satisfy food demand (taking account of net imports), and does so optimally subject to constraints imposed by biophysical and socio-economic conditions. Calculations are carried out across overlapping geographically unstructured clusters of cells with similar ~~production~~ biophysical conditions (based on soil and agroclimate), with profitability thresholds used to determine which land use and management intensity offer the greatest returns across each cluster ~~is allocated to each cluster~~. Land use proportions within each 10' x 10' grid cell represent the aggregations of the optimal solutions for each (up to 40) associated cluster. At cell level, ~~-~~this aggregation therefore represents the (spatially weighted) optimised land use solution for each cluster containing the grid cell in question. The clustering recognises that different biophysical conditions (soil and agroclimate) differentially influence the suitability, productivity and profitability of different crops and different agricultural systems (arable, dairy etc.), leading to heterogeneity in agricultural land use within a grid cell. The IAP runs from a present-day simulated baseline land use configuration to the mid-2080s under combined climatic and socio-economic scenarios. The IAP has been applied and evaluated in a large number of studies including sensitivity and uncertainty analyses (e.g. Brown et al. 2014; Harrison et al. 2015, 2016, 2019; Kebede et al. 2015; Holman et al. 2017a, b;

Fronzek et al. 2019). A full model description and the online model itself are available at http://www.impressions-project.eu/show/IAP2_14855.

CRAFTY-EU is an application of the *CRAFTY* framework for agent-based modelling of land use change (Brown, Seo, et al., 2019; Murray-Rust et al., 2014) that covers the same extent as the IAP at the same (10 arcminute) resolution. CRAFTY uses the concept of Agent Functional Types (AFTs) (Arneth, Brown, & Rounsevell, 2014) to simulate land use change over large geographical extents while capturing key behaviours of decision-making entities (agents) that include individual land managers, groups of land managers and institutions or policy bodies (Holzhauer, Brown, & Rounsevell, 2019). Modelled land manager agents compete for land on the basis of their abilities to produce a range of ecosystem services that society is assumed to require. In CRAFTY-EU, these services ~~are~~ include provisioning (food crops and meat, timber), regulating (carbon sequestration), cultural (recreation) and supporting services (habitat provision through landscape diversity) ~~crops, meat, timber, carbon sequestration, recreation and landscape diversity~~. The abilities of agents to supply these services under given biophysical and socio-economic conditions are derived either from IAP model results (Fig. 1) or from basic assumptions linking land uses to service levels, as explained in Brown et al. (2019). Satisfying demands for services brings economic and non-economic benefits to individual agents, with benefits quantified as functions of unsatisfied demand. In this case, these functions are linear and equivalent for all services, meaning that the benefit of production of each service increases equally per unit of unmet demand, providing a clear basis for model comparison. Economic benefit represents income from marketable goods and services, and non-economic benefit represents a range of motivations, from subsistence production to the maintenance of societal, cultural or personal values associated with particular services or land uses. Ecosystem services production levels are determined by the natural productivity of the land and the form and intensity of agents' land management~~, as described in detail in Appendix A~~. The outcome of the competitive process at each annual timestep is determined by agent-level decision-making that is not constrained to generate the greatest benefit, and agents are parameterised here to continue with land uses that provide some return rather than abandon their land, but to gradually adopt significantly more beneficial alternatives if available.

Importantly for this study, CRAFTY-EU is parameterised on the basis of the IAP, taking IAP outputs as exogenous conditions and replacing only the land allocation component to provide alternative land use projections under identical driving conditions (Fig. 1). CRAFTY-EU is initialised on the IAP's baseline map, and ~~only~~ is known to only diverge~~s~~ from that stable baseline 'solution' as scenario conditions change (Brown, Seo, et al., 2019). Land use productivities, in terms of potential yields and ecosystem service provision levels of the simulated land use systems under the agronomic scenario conditions at cell scale, are also calculated from IAP outputs dependent on land use allocation, with the result that productivities are set to zero where the IAP determines production to be economically infeasible. For ecosystem services with economic values (meat, crops and timber), agents in CRAFTY therefore make production choices that conform ~~consistent with~~to this basic level of economic ~~rationality~~feasibility, while still being able to select a range of economically optimal or sub-optimal land uses. A full description of the model can be found in Brown et al. (2019) and an online version with access to full model code at https://landchange.earth/CRAFTY.

2.2~~2~~ Climate and socio-economic scenarios

Seven combinations of climatic and socio-economic scenarios were simulated, based on the Representative Concentration Pathways (RCPs) and Shared Socioeconomic Pathways (SSPs) (O'Neill et al., 2017). The RCPs and SSPs were combined taking account of internal consistency with their associated greenhouse gas emissions; RCP2.6 was combined with SSP1 and 4; RCP4.5 with SSP1, 3 and 4; and RCP8.5 with SSP3 and 5 (Table 1). The SSPs have been further developed for Europe through a stakeholder-engagement process that included interpretation and quantification of key drivers of change in land-based sectors (Table 2a; Kok, Pedde, Gramberger, Harrison, & Holman, 2019). For this study, RCPs were simulated in the IAP using outputs from two global-regional climate models (EC_Earth/RCA4 for RCP2.6, and HADGEM2-ES/RCA4 for RCPs 4.5 and 8.5 (Table 2b; Paula A. Harrison et al., 2019)). Scenario outcomes are described for CRAFTY-EU in Brown et al. (2019b) and for the IAP in Harrison et al. (2019) and Papadimitriou et al. (2019). In addition to these established scenarios, one scenario combination (RCP4.5 – SSP3) was simulated with additional variations in model parameterisations. This scenario was chosen as producing particularly divergent results between the two models, and parameter values were altered to constrain the differences in model responses to the scenario and so to reveal the roles of underlying assumptions in producing the observed divergence. ~~assess whether analogous driving factors led to convergence between the models.~~ Specifically, we increased imports in the IAP by 40% (to mimic an observed under-production of food in CRAFTY), and increased the value of food production in CRAFTY by ten times (to compensate for reductions in supporting capital levels responsible for the under-production of food).

2.3 Conceptual framework

The model comparison presented here is motivated by the hypothesis that the nature of simulated land use allocation is one dominant source of uncertainty in land use modelling, as opposed to uncertainty in crop yields, biophysical conditions or other land system characteristics. The selected models therefore allow us to keep the latter factors common and explore how different factors that influence land use allocation, such as profitability, non-economic motivations, demand levels and socio-economic conditions, affect model outcomes. This is possible because the models used share much of their information and design features, but adopt distinct paradigms for modelling the process of land allocation itself (Fig. 1, Table 3).

The IAP and CRAFTY-EU belong to distinct paradigms in the sense that the IAP is an example of a 'top-down' model that simulates change at the system-level – in this case through an assumption of constrained economic optimisation - while CRAFTY is an example of a 'bottom-up' model that simulates change at the level of individual decision-makers – in this case through an assumption of behavioural choices made at the level of local land systems (Brown et al., 2016; Couclelis, 2002).

These paradigms usually have different uses and justifications: the (dominant) top-down approach is computationally efficient, tractable and more in line with economic theory, although it is rarely justified as an accurate representation of how land use decisions are made in practice (in fact the evidence tends to contradict it; e.g. Chouinard et al. 2008; Schwarze et al. 2014; Appel and Balmann 2019). The bottom-up approach, in contrast, is more exploratory and often criticised for producing uncertain results, but explicitly attempts to achieve greater process accuracy (Brown et al., 2016).

Neither of these models is intended to accurately predict real-world land use change, but to project land system dynamics on the basis of complex and integrated processes founded on a small number of key, transparent assumptions. This comparison is therefore intended first and foremost to explore the reasons for *simulated* land use changes, and does not speak directly to observed land use changes. Nevertheless, both models have been extensively used and evaluated, and both respond stably and predictably to driving conditions (Brown et al., 2014; Brown, Holzhauer, Metzger, Paterson, & Rounsevell, 2018; Brown, Seo, et al., 2019; Paula A. Harrison et al., 2019, 2016; I. P. Holman et al., 2017). Both also have similar uses, being intended to support academic research and education and, to some extent, capacity building with stakeholders to increase understanding of the importance of socio-economic and climatic changes, systemic inter-relationships in the land system, and geographic regions that may be particularly vulnerable or resilient to change. As a result, the comparison does not consider model purpose or the suitability of either model for direct policy-support, prediction or other unintended uses.

Further, some of the effects of the different land allocation mechanisms contained in these models are apparent *a priori*. As a bottom-up, agent-based model, CRAFTY is less constrained than the IAP, with multiple outcomes being possible from a given set of input conditions. At the same time, land use decisions are subject to behavioural inertia in CRAFTY, with agents unwilling to change existing land uses and motivated by non-economic factors that can counteract price signals. The IAP will always identify the optimal result subject to economic drivers and modelled constraints, and does so without reference to the previously simulated timepoint (i.e., is not path-dependent). It is therefore expected that the IAP responds to smaller changes in conditions than does CRAFTY, and that the models are likely to diverge as time goes on and as the magnitude of changes increases.

2.4~~3~~ Comparison

In this study, both models are run until the mid-2080s (defined as a 30-year timeslice in the IAP, and the year 2086 in *CRAFTY-EU*). Both use a spatial grid of resolution 10 arcmin x 10 arcmin (approximately 16km x 16km in Europe), but simulated land classes differ between the two models (as described in Brown et al. 2019b) and are standardised here as described in Table 4~~1~~, to focus on major, comparable forms of agricultural and forestry management. These aggregate land use classes are not homogeneous or uniform across the simulations as they allow for a range of management forms within them. We therefore also compare ecosystem service production levels, which account for exact forms of management simulated in each cell. ~~-Other forms of land use and management (e.g. u~~Urban land use is ~~s) are~~ not compared as ~~the~~its locations~~y~~ are shared by both models. ~~The labels assigned to these land use classes reflect the dominant form, but not the remaining range, of management within~~

~~them. We therefore also compare ecosystem service production levels, which account for exact forms of management simulated in each cell.~~

The comparison of these land use classes was made at two spatial resolutions: across the whole of the modelled domain (without reference to spatial configurations) and across 323 Nomenclature of Territorial Units for Statistics (NUTS2) regions. NUTS2 resolution was chosen for the spatially explicit comparison instead of the original 10' model resolution to limit the impact of relatively uninformative differences in the allocation of individual cells, and to focus instead on systematic differences in model responses to the simulated scenarios. This choice also reflects the fact that neither model is intended to predict cell-level outcomes, but to provide illustrative realisations of scenario outcomes, with the cell-level results of CRAFTY-EU differing between individual runs because the model is stochastic and path dependent. At NUTS2 level, only differences between the models affecting at least 5% of the relevant cells were included in the analysis. In the following sections (Results and Discussion), CRAFTY-EU is referred to simply as CRAFTY, for brevity.

The presentation of the results below is structured to reveal the effects of the paradigmatic differences between the models (and not to assess the models' shared characteristics). First, we compare outputs from each scenario at EU scale to identify the principal differences that arise in the simulations. Because the scenarios relate to the modelling paradigms in different ways (e.g. allowing for stronger or weaker economic signals), this allows us to link the results to particular modelling choices. We then compare results at NUTS2 level to identify relatively minor or hidden differences, before experimenting with forced convergence to test the role of particular parameters and assumptions in each model.

3. Results

3.1 ~~Aggregate~~ Overall EU-level comparison

The responses of the two models to scenario conditions are notably different in most cases (Figures 2~~1~~ & 3~~2~~, Table 1), albeit within similar broad limits (Fig. 2~~1~~). The greatest similarities in terms of aggregate land use classes occur in the SSP1 simulations, where both models produce land systems that remain similar to the baseline, with large areas of intensive agriculture and small areas of land not managed for agriculture or forestry. The IAP results include more dedicated pastoral land and the CRAFTY results more forestry~~, with the differences being greatest in RCP2.6-SSP1~~. In ~~both~~all ~~RCP2.6~~ simulations with very low climate change (RCP2.6), CRAFTY produces an under-supply of food and both models produce an under-supply of timber, ~~though~~ and these shortfalls ~~supply-demand gaps are smaller~~ reduce under intermediate climate change ~~in~~ (RCP4.5), where productivity is slightly higher (Fig. 3~~2~~). CRAFTY ~~also has~~produces smaller ~~differences~~ imbalances between food and timber supplies due to its equivalent valuation of all modelled services.

In other scenarios, the IAP responds most strongly to SSPs 4 and 5, while CRAFTY responds most strongly to SSP3. At aggregate level, CRAFTY produces similar results in the SSP4 and 5 simulations as in SSP1 (Fig. 2~~1~~), though with generally less intensive agriculture and higher supply levels (even exceeding demand in the higher climatic productivities of RCP4.5

and 8.5) (Fig. 3~~2~~a). In contrast, the IAP projects a dramatic move away from intensive agriculture in SSPs 4 and 5 as a consequence of greatly increased productivity requiring a smaller agricultural area to meet demand. This loss of agricultural management in previously intensively-managed areas is far more pronounced in the IAP than in CRAFTY, where the wider range of valued ecosystem services supports more management and, in some cases, oversupply of services (Fig. 3~~2~~). ~~As in SSP1, t~~The extent of agricultural abandonment is greatest in the IAP ~~in~~ under intermediate climate change (RCP4.5), where increased yields in some areas reduce the relative competitiveness of agricultural land in less productive areas.

SSP3 produces considerably smaller responses in the IAP, with some areas of all land use types going out of management and with far larger areas of the intensive agriculture class remaining than in SSP4. CRAFTY outcomes for SSP3 are highly dependent on climate scenario, with RCP4.5 producing the strongest response, most notably in terms of a large shortfall in the supply of crops (Fig. 3~~2~~a). In this case, widespread extensification of land use occurs, with little intensive agriculture remaining by the end of the simulation, and a slight increase in land going out of agricultural or forestry management. In RCP8.5 these changes are less pronounced, with only small changes from intensive agriculture to extensive and forestry management. These changes occur because SSP3 includes deteriorating inherent agricultural productivity and also substantial declines in capital values that support land management (particularly financial, human and manufactured capitals). In CRAFTY, these simultaneous changes make it difficult for agents to maintain intensive management against competition from extensive and less capital-dependent forms of management. The increased yields in some parts of Europe produced by climate change in RCP8.5 make this scenario more conducive to the maintenance of intensive management.

The models also respond very differently to the SSP5 scenario (paired only with RCP8.5). In the IAP, large areas switch to extensive and other/no management classes while there is very little overall change in CRAFTY. The differences between the models' responses are mainly due to the higher yields and improved technological conditions in SSP5 making large areas of intensive agriculture surplus to requirements.~~;~~ ~~these~~ These surplus areas are no longer intensively managed for agriculture in the IAP by the 2080s, but are ~~retained~~ in CRAFTY (resulting in over-supply of food) because they provide other services and because of the gradual decision-making of agents that spreads abandonment decisions over multiple timesteps.

Together, these scenario results show that the IAP responds most strongly to scenarios with conditions in which agricultural productivity increases, and which therefore lead to reduced need for agricultural land and, in this model, extensification and agricultural abandonment (which occurs over larger extents in the IAP than in CRAFTY). CRAFTY responds less strongly to such conditions because agents have a (parameterizable) unwillingness to change or abandon their land use in the absence of a more viable alternative, and because a wider range of services produce returns for those agents. Conversely, CRAFTY responds most strongly to scenarios in which conditions affecting agricultural productivity ~~decreases~~ worsen because agents rely ~~its design emphasises changes in~~ more strongly on a range of climatic and socio-economic conditions. Many of these conditions deteriorate ~~capitals that support production (climatic or socio-economic), as is particularly clear~~ in SSP3, making~~.~~ ~~In these circumstances,~~ intensive agriculture ~~is~~ less competitive than extensive agriculture or other multifunctional land uses, and causing intensive agents ~~are~~ to be easily replaced (competition is a more rapid process than abandonment in the CRAFTY parameterisation used here).

3.2 ~~Spatial~~ Geographical comparison

Within the overall differences between model results exist some consistent spatial and geographical patterns (Fig. 4~~3~~). Across scenarios, the IAP often places more pastoral and very extensive land use classes in western Europe in particular, while CRAFTY often has more intensive agriculture in mid-latitudes, and forest in eastern and northern areas (Figs. 4 & B1~~3~~). These differences are very scenario-dependent, however, and as with the aggregate summaries above, the ~~spatial~~ geographical patterns produced by one model in SSP3 resemble those produced by the other model in SSP4. In SSP4, the IAP projects substantially more very extensive agricultural management and forest management than CRAFTY~~'s more intensive results~~, while the near-inverse is true for SSP3 (reflecting implicit assumptions that over-production is not penalised~~,~~ in CRAFTY, and that intensive agriculture retains an efficiency advantage over extensive~~,~~ in the IAP). CRAFTY also produces a great deal more forest management in RCP2.6-SSP1, with intensive arable agriculture dominating only in the most productive parts of France, Germany and the UK. SSP1 is also the scenario in which the IAP produces the most concentrated areas of intensive pastoral agriculture, particularly in Ireland, the UK and France.

Notwithstanding the smaller-scale fragmentation of land uses in CRAFTY (see below), these results show that at this aggregate level, CRAFTY has a tendency (except in SSP3) to concentrate intensive agriculture in mid-latitudes, extensive agriculture in the southern Baltic states, and very extensive land uses at the European latitudinal extremes. Forestry is distributed in the western UK and central-eastern states in particular. The IAP results are less consistent, but show a tendency to produce pastoral agriculture in the west and forestry more widely. Many of these differences may reflect the valuation of a wider range of services in CRAFTY, leading to a concentration of intensive management in the most productive areas where it can maintain relative competitiveness. As above, they also reflect the differences in the conditions that the models respond to, with the IAP particularly sensitive to changes in demand that do not have spatial manifestations, and CRAFTY more sensitive to capitals that are maximised in climatically suitable, but also politically stable and affluent countries.

3.3 Convergence experiment

The scenario combination RCP4.5-SSP3 was chosen as having particularly different results from the two models, and so used to examine the potential for convergence in model settings and results. In this scenario, CRAFTY produces a highly fragmented land system with areas of abandoned or extensively managed land scattered throughout Europe, and a substantial shortfall in food production. The IAP, in contrast, produces large contiguous agricultural areas with far more intensive management (albeit of greatly reduced productivity) and less forestry, satisfying food demands. To control for the main differences in scenario conditions in each model, we increased food imports in the IAP to produce lower production levels in the EU, as observed in the CRAFTY result, and we increased food prices in CRAFTY to produce greater support for intensive agriculture, as observed in the IAP result. In the absence of these major differences, any remaining divergence in model outputs could be attributed to other factors.

In terms of overall land system composition the changes in the IAP (an increase of 40% in food imports) did not lead to a result approaching the original CRAFTY results (Fig. 5~~4~~). While the extent of intensive agricultural management did decrease, this led to widespread agricultural abandonment rather than additional extensive or forestry management (demand for which was already satisfied), with remaining food production being even more concentrated in certain intensively-managed parts of Europe (particularly the East). Large parts of southern and northern Europe fell out of agricultural management, with other regions and countries being managed only for forestry. Other results (above) suggest that the IAP would have more closely resembled the CRAFTY result had there been an explicit driver for extensification, rather than simply an effective decrease in demand levels.

From the more extensively-managed and fragmented initial result produced by CRAFTY, a ten-fold increase in food prices did come closer to the initial IAP result, although with more intensive agriculture and less land under other or no management. The distribution of land uses was strikingly different, however. Unmanaged land mainly occurred in the same areas, and concentrations of forestry overlapped to some extent, but the agricultural land in the CRAFTY result remained highly fragmented across much of Europe. In this case, CRAFTY produced sufficient food to satisfy demand.

4. Discussion~~ & conclusions~~

Understanding the contributions of different modelling paradigms to land use projections is important for two main reasons. The first reason is that almost all large- to global-scale land system models share a single paradigm (economic optimisation of land uses), raising the risk of biases in model results and resultant, unrecognised knowledge gaps (e.g. Verburg et al. 2019; Elsawah et al. 2020; Müller et al. 2020). The second reason is that different paradigms are known to produce very different outcomes, but for reasons that remain unclear (Alexander et al., 2017; Prestele et al., 2016). The focused comparison presented here is therefore intended to identify and explain key differences between models representing major, distinct paradigms. While conclusions are inevitably limited by the breadth of the comparison, and in particular by the many characteristics that are shared between the selected models (Table 3), o~~ Neither model is intended to be predictively accurate, but to project land system dynamics on the basis of complex and integrated processes founded on a small number of key, transparent assumptions. Both models have also been extensively used and evaluated, and both respond stably and predictably to driving conditions (Brown et al., 2014; Brown, Holzhauer, Metzger, Paterson, & Rounsevell, 2018; Brown, Seo, et al., 2019; Paula A. Harrison et al., 2019, 2016; I. P. Holman et al., 2017). As expected, o~~ur results do reveal large and consistent differences ~~between the two selected models~~ that emerge from the different ways in which those models represent land system change.

~~An overarching distinction is apparent between the basic assumptions underlying the models. The IAP is an example of a 'top-down' model that simulates change at the system-level – in this case through an assumption of constrained economic optimisation - while CRAFTY is an example of a 'bottom-up' model that simulates change at the level of individual decision-makers – in this case through an assumption of behavioural choices made at the level of local land systems (Brown et al., 2016). This basic difference affects the rate, extent and pattern of simulated land use change. These paradigms usually have different uses and justifications: the (dominant) top-down approach is computationally efficient, tractable and more in line~~

~~with economic theory, although it is rarely justified as an accurate representation of how land use decisions are made in practice (in fact the evidence tends to contradict it; e.g. Chouinard et al. 2008; Schwarze et al. 2014; Appel and Balmann 2019). The bottom-up approach, in contrast, is more exploratory and often criticised for producing uncertain results, but explicitly attempts to achieve greater process accuracy (Brown et al., 2016).~~

The consequences of top-down and bottom-up perspectives ~~is~~ are apparent in the ~~main~~ forms, extents, rates and patterns of land use change as the models respond to scenario conditions. The IAP's consistent profitability thresholds within a deterministic optimising framework respond strongly to increasing yields or decreasing demands, when the model produces widespread agricultural abandonment outside the most productive land. Conversely, CRAFTY's heterogeneous competition process within a stochastic agent-based framework responds more strongly to decreases in productivity, when the model produces extensification and expansion of agriculture. This difference is also apparent in our convergence experiment, where increased imports in the IAP lead to reduced agricultural area, ensuring efficient production where competitiveness is highest, rather than the extensification that CRAFTY produces. Increasing food prices in CRAFTY did generate aggregate land use proportions similar to those of the IAP, albeit with largely distinct spatial distributions, suggesting that agents become more 'optimal' in behaviour when greater competitive advantages are available.

To some extent these differences are traceable to the underlying mathematical structures of the models, with the IAP identifying any change in optimal configurations and CRAFTY maintaining existing and multifunctional land uses where possible. But the results are also subject to model sensitivity and uncertainty. Previous analyses show that the IAP responds most strongly to changes in demand levels and climate-driven yields, and that their effects outweigh those of socio-economic scenarios (Brown et al., 2014; Kebede et al., 2015). CRAFTY has similar sensitivities complemented but not overwhelmed by simulated agent behaviour (Brown, Holzhauer, et al., 2018; Brown, Seo, et al., 2019). Together these suggest that the effects and differences we find are robust and traceable to model design interacting primarily with climatic scenarios (RCPs), and with socio-economic scenarios (SSPs) to a lesser extent.

~~This~~ Particularly influential is the ~~fundamental difference in dominant land use change trajectories is accentuated by the~~ representation in CRAFTY of individual and societal desires for a range of ecosystem services, which means that extensive management practices that provide recreation, carbon sequestration or landscape diversity, for example, are adopted instead of land abandonment. This is not necessarily tied to modelling paradigm; optimisation can in principle be performed across a range of criteria, potentially accounting for many more (economically-valued) ecosystem services, although this remains conceptually and computationally challenging (Newland, Maier, Zecchin, Newman, & van Delden, 2018; Seppelt, Lautenbach, & Volk, 2013; Strauch et al., 2019). The non-optimising representation used in models such as CRAFTY is closer to the reality of how land use actually changes (Appel & Balmann, 2019; Schwarze et al., 2014), but still requires additional parameterisation and rigorous uncertainty analysis (Verburg et al., 2019). In either case, there is strong justification for including a wide range of ecosystem services, particularly those such as carbon sequestration that may gain distinct values in different future scenarios (Estoque, Ooba, Togawa, & Hijioka, 2020; Kay et al., 2019).

One consequence of simulating demand and supply of a range of ecosystem services is that the relative economic support available for food production becomes a key determinant of the balance of different land uses as agriculture, while still dominant in area, must compete with alternative management options. Models such as the IAP seek to maintain food supplies, even at the expense of other services such as timber production, while models such as CRAFTY allow supply levels to emerge from simulated decisions and so are capable of producing shortfalls. All of the models' results are affected by this basic assumption about whether equilibrium does or will exist in the food system, and further by the extent of disequilibrium that is tolerated and the mechanism by which that extent is defined. For instance, food prices in CRAFTY can respond to shortfalls in production through a number of parametric functions, while the in the IAP prices are automatically adjusted within broad limits to ensure that demand and supply match. However, shortfalls in food production in CRAFTY ~~are not linked to~~do not lead to simulated hunger, societal unrest or migration, and food prices in the IAP may become unrealistically high in scenarios where economic and social conditions are very challenging (Hamilton et al., 2020; Pedde et al., 2019). In both models, the simulation of the European land system as distinct from the rest of the world requires implicit but shared assumptions about conditions in other regions and their relationships to Europe. Alternative assumptions would inevitably lead to different outcomes and, perhaps, greater differences between the two models' results. As conceptual alternatives, therefore, neither of these necessarily capture the true dynamics of food prices and production levels, which remains a major challenge for land system modelling (Müller et al., 2020; Pedde et al., 2019).

Beyond differences at aggregate level, another notable feature of results shown above are that CRAFTY produces far more small-scale heterogeneity in land use than does the IAP. This heterogeneity is particularly pronounced in CRAFTY's SSP3 simulations (Fig. 5~~4~~) and reflects a basic modelling approach: the simulation of ~~distinct cell-level and~~ time-dependent decisions affecting individual cells, with agents parameterised here to abandon land only if it provides no returns, and then only gradually. This effectively precludes the system-level optimisation practised by the IAP, which does not account for individual land use decisions. Individual-level heterogeneity is, inevitably, very difficult to parameterise precisely, although participatory techniques have some promise in this respect (Elsawah, Guillaume, Filatova, Rook, & Jakeman, 2015). Conversely, (constrained) optimising models like the IAP produce ~~idealised~~ results that may not replicate observed rates or spatial structures of land use change (Brown, Alexander, Arneth, Holman, & Rounsevell, 2019; Low & Schäfer, 2020; Turner, Field, Lobell, Sanchez, & Mach, 2018), but can introduce ~~use flexible~~ spatial dependencies as further constraints on optimisation in order to approximate spatially-mediated social ~~proxies for~~ processes such as imitation~~, diffusion of knowledge or the formation of social norms~~ (Brown, Alexander, & Rounsevell, 2018; Meiyappan, Dalton, O'Neill, & Jain, 2014). Bottom-up models in general tend to be less precisely specified and so produce more variable results (or are more "skittish" as (Couclelis, 2002) put it). They are also generally less often compared against observational (or other modelled) data, and while their flexibility makes fitting-to-data notably feasible in principle, their inherent tendency to produce variable results means that the production of any one particular outcome does not have the apparent significance that it does for a more constrained model. Both models used here have been compared against 'observed' land use data to some extent, with an example application of CRAFTY compared and calibrated to MODIS land cover data (Seo, Brown, & Rounsevell, 2018) and the IAP (and hence, indirectly, CRAFTY) calibrated to match CORINE land cover and NUTS2 yields (P. A. Harrison et al., 2015).

Notwithstanding the gains to be made by better understanding the relative performance of different modelling paradigms, it is essential to recognise some hard limits. No land use model is intended or able to provide calibrated representations of all the mechanisms responsible for land use change, especially under imagined future conditions. Models of this kind are inevitably reductionist in nature and omit a large number of important factors and processes that occur in reality – particularly, in this case, those occurring at smaller spatial scales than are simulated here. Both alternatives must therefore be seen as providing realisations of assumptions that are useful in some ways but incorrect in others. Optimising models have the advantage of representing idealised conditions in that they maximise achievement of modelled criteria such as production levels, but do not necessarily reveal ~~the~~ pathways by which those conditions can be reached in reality (Ligmann-Zielinska, Church, & Jankowski, 2008; Low & Schäfer, 2020). Process- or agent-based approaches, meanwhile, can allow exploration of the large behavioural uncertainties involved in the simulation of human systems, and can be powerful tools for stakeholder engagement and understanding (Low & Schäfer, 2020; Millington, Demeritt, & Romero-Calcerrada, 2011) – but are unlikely to perform any better at predicting system outcomes than simpler, more tightly constrained models (Salganik et al., 2020). Indeed, their primary strength may be their ability to use theory (and so to allow a choice among theories) as a guide to processes and conditions that empirical data and optimising models do not cover (Gostoli & Silverman, 2020). Both types of model represent abstracted units, managers and characteristics of land, which do not match exactly to real-world conditions as experienced and determined by actors in the system (e.g. productivities and profits used to drive the models are not the same as those available to real-world land managers).

Fundamentally, no single modelling paradigm is 'correct', and future developments are likely to invalidate even those assumption that appear safest at the present time. The greatest value of these two approaches may therefore lie in their ability to provide alternatives. This ~~; a~~ value ~~that~~ is realised only in the (currently rare) cases when ~~model assumptions are clearly communicated and when~~ analogous models with similar driving conditions but different underlying assumptions, such as those used here, are available for comparison (Müller et al., 2014; Polhill & Gotts, 2009; Rosa, Ahmed, & Ewers, 2014). Further benefits can be drawn from combinations of the two modelling approaches, although this usually involves an artificial choice of systems or scales at which top-down optimisation and bottom-up emergence are assumed to occur (e.g. Castella and Verburg 2007; Verburg and Overmars 2009; Houet et al. 2014). In addition, the benefits of using each type of model can be maximised (and the differences between them potentially minimised) by flexible multi-criteria optimisation on one hand and behavioural uncertainty analysis on the other (Brown, Holzhauer, et al., 2018; Fonoberova, Fonoberov, & Mezić, 2013; Ligmann-Zielinska, Kramer, Cheruvelil, & Soranno, 2014; Newland et al., 2018). Nevertheless, ~~Both can also be advanced by new interdisciplinary approaches to better represent qualitative knowledge about land system change (Elsawah et al., 2020). Such interdisciplinary approaches could, for instance, allow integration across the individual, societal and even political levels, using different or flexible modelling approaches at each level to improve their representation (e.g. Andersen et al. 2017). Technically, integration~~

~~of this kind can utilise powerful forms of 'hybrid' modelling that allows model design and complexity to be tailored to requirements (Lippe et al., 2019; Parrott, 2011). In allowing parallel or integrated usage of different paradigms, all of these methods can provide insights that suffer less from individual weaknesses, and benefit more from individual strengths, than each model in isolation. S~~substantial efforts to increase both the diversity and coherence of land system modelling are likely to be necessary if these important gains are to be made.

5. Conclusions

In taking two particular models as representative of major modelling paradigms we can only make tentative conclusions about the consequences of those paradigms for model outputs. Nevertheless, we find large, consistent differences between the models that are robust to known model sensitivities and directly traceable to basic assumptions. In particular, we find that the 'bottom-up' agent-based model produces more heterogenous, multifunctional land systems than the 'top-down' model, as expected. We also find that the models respond most strongly to different scenario conditions, despite both being most sensitive to climatic effects on yields and socio-economic effects on demand levels. In particular, the constrained optimisation of the top-down model is able to capitalise on increases in productivity by utilising the best land, while the agent-based model is limited by inertia and path-dependency in simulated conditions. Conversely, reductions in productivity, including through socio-economic disruption, prompt widespread extensification of land management in the bottom-up model that is not replicated in the top-down model, as simulated agents diversify and rely on more varied or even non-economic benefits. Currently these two modelling paradigms are far apart in their projections of future change, suggesting huge uncertainty about the role the land system can and will play in societal challenges such as climate change and biodiversity loss. However, this comparison suggests that such divergence, and hence uncertainty, rests largely on a few key features: in particular the assumed extent of non-economic decision-making, the relative importance that society places on cultural and regulating ecosystem services compared to provisioning, and the likely rate of land use change, including abandonment and intensification, as outcomes of human decisions. Our findings show the importance of communicating these assumptions to model users, but also of identifying better-supported and more generally-accepted positions that narrow the gap between the current extremes of dominant paradigms.

Code and data availability

The full model code and data for CRAFTY-EU are available for download and visualisation via https://landchange.earth/CRAFTY

The IAP is available for interactive online runs at http://www.impressions-project.eu/show/IAP2_14855 but model code is not available because the IAP utilises meta-models of several other stand-alone models under different ownership.

Appendices

Appendix A: ~~Ecosystem service production in CRAFTY~~Land use class composition

Author contribution

CB performed the analysis and drafted the manuscript; IH & MR assisted with planning, interpretation and writing.

Competing interests

The authors declare that they have no conflict of interest

Acknowledgements

This research was supported by the EU Seventh Framework Programme project IMPRESSIONS (grant no. 603416) and the Helmholtz Association. We acknowledge support by the KIT-Publication Fund of the Karlsruhe Institute of Technology

References

Alexander, P., Prestele, R., Verburg, P. H., Arneth, A., Baranzelli, C., Batista e Silva, F., … Rounsevell, M. D. A. (2017). Assessing uncertainties in land cover projections. *Global Change Biology*, *23*(2), 767–781. doi: 10.1111/gcb.13447

Andersen, L. E., Groom, B., Killick, E., Ledezma, J. C., Palmer, C., & Weinhold, D. (2017). Modelling Land Use, Deforestation, and Policy: A Hybrid Optimisation-Heterogeneous Agent Model with Application to the Bolivian Amazon. *Ecological Economics*, *135*, 76–90. doi: 10.1016/j.ecolecon.2016.12.033

Appel, F., & Balmann, A. (2019). Human behaviour versus optimising agents and the resilience of farms – Insights from agent-based participatory experiments with FarmAgriPoliS. *Ecological Complexity*, *40*, 100731. doi: 10.1016/j.ecocom.2018.08.005

Arneth, A., Brown, C., & Rounsevell, M. D. A. (2014). Global models of human decision-making for land-based mitigation and adaptation assessment. *Nature Climate Change*, *4*(7), 550–557. doi: 10.1038/nclimate2250

Audsley, E., Trnka, M., Sabaté, S., Maspons, J., Sanchez, A., Sandars, D., … Pearn, K. (2015). Interactively modelling land profitability to estimate European agricultural and forest land use under future scenarios of climate, socio-economics and adaptation. *Climatic Change*, *128*(3–4), 215–227. doi: 10.1007/s10584-014-1164-6

Brown, C., Alexander, P., Arneth, A., Holman, I., & Rounsevell, M. (2019). Achievement of Paris climate goals unlikely due to time lags in the land system. *Nature Climate Change*, *9*(3), 203–208. doi: 10.1038/s41558-019-0400-5

Brown, C., Alexander, P., & Rounsevell, M. (2018). Empirical evidence for the diffusion of knowledge in land use change. *Journal of Land Use Science*, *13*(3), 269–283. doi: 10.1080/1747423X.2018.1515995

Brown, C., Brown, E., Murray-Rust, D., Cojocaru, G., Savin, C., & Rounsevell, M. (2014). Analysing uncertainties in climate change impact assessment across sectors and scenarios. *Climatic Change*, *128*(3–4), 293–306. doi: 10.1007/s10584-014-1133-0

Brown, C., Brown, K., & Rounsevell, M. (2016). A philosophical case for process-based modelling of land use change. *Modeling Earth Systems and Environment*, *2*(2), 50. doi: 10.1007/s40808-016-0102-1

Brown, C., Holzhauer, S., Metzger, M. J., Paterson, J. S., & Rounsevell, M. (2018). Land managers’ behaviours modulate pathways to visions of future land systems. *Regional Environmental Change*, *18*(3), 831–845. doi: 10.1007/s10113-016-0999-y

Brown, C., Seo, B., & Rounsevell, M. (2019). Societal breakdown as an emergent property of large-scale behavioural models of land use change. *Earth System Dynamics*, *10*(4), 809–845. doi: 10.5194/esd-10-809-2019

Castella, J.-C., & Verburg, P. H. (2007). Combination of process-oriented and pattern-oriented models of land-use change in a mountain area of Vietnam. *Ecological Modelling*, *202*(3–4), 410–420. doi: 10.1016/j.ecolmodel.2006.11.011

Chouinard, H. H., Paterson, T., Wandschneider, P. R., & Ohler, A. M. (2008). Will farmers trade profits for stewardship? Heterogeneous motivations for farm practice selection. *Land Economics*, *84*(1), 66–82. doi: 10.3368/le.84.1.66

Couclelis, H. (2002). Modeling frameworks, paradigms, and approaches. *Geographic Information Systems and Environmental Modelling, Prentice Hall, London*.

Elsawah, S., Filatova, T., Jakeman, A. J., Kettner, A. J., Zellner, M. L., Athanasiadis, I. N., … Lade, S. J. (2020). Eight grand challenges in socio-environmental systems modeling. *Socio-Environmental Systems Modelling*, *2,* 16226. doi: 10.18174/sesmo.2020a16226

Elsawah, S., Guillaume, J. H. A., Filatova, T., Rook, J., & Jakeman, A. J. (2015). A methodology for eliciting, representing, and analysing stakeholder knowledge for decision making on complex socio-ecological systems: From cognitive maps to agent-based models. *Journal of Environmental Management*, *151*, 500–516. doi: 10.1016/j.jenvman.2014.11.028

Estoque, R. C., Ooba, M., Togawa, T., & Hijioka, Y. (2020). Projected land-use changes in the Shared Socioeconomic Pathways: Insights and implications. *Ambio*, 1–10. doi: 10.1007/s13280-020-01338-4

Fonoberova, M., Fonoberov, V. A., & Mezić, I. (2013). Global sensitivity/uncertainty analysis for agent-based models. *Reliability Engineering and System Safety*, *118*, 8–17. doi: 10.1016/j.ress.2013.04.004

Fronzek, S., Carter, T. R., Pirttioja, N., Alkemade, R., Audsley, E., Bugmann, H., … Yoshikawa, M. (2019). Determining sectoral and regional sensitivity to climate and socio-economic change in Europe using impact response surfaces. *Regional Environmental Change*, *19*(3), 679–693. doi: 10.1007/s10113-018-1421-8

Gostoli, U., & Silverman, E. (2020). Sound behavioural theories, not data, is what makes computational models useful. *Review of Artificial Societies and Social Simulation*. Retrieved from https://rofasss.org/2020/04/22/sound-behavioural-theories/

Hamilton, H., Henry, R., Rounsevell, M., Moran, D., Cossar, F., Allen, K., … Alexander, P. (2020). Exploring global food system shocks, scenarios and outcomes. *Futures*. doi: 10.1016/j.futures.2020.102601

Harrison, P. A., Holman, I. P., & Berry, P. M. (2015). Assessing cross-sectoral climate change impacts, vulnerability and adaptation: an introduction to the CLIMSAVE project. *Climatic Change*, *128*(3–4), 153–167. doi: 10.1007/s10584-015-1324-3

Harrison, Paula A., Dunford, R. W., Holman, I. P., Cojocaru, G., Madsen, M. S., Chen, P. Y., … Sandars, D. (2019). Differences between low-end and high-end climate change impacts in Europe across multiple sectors. *Regional Environmental Change*, *19*(3), 695–709. doi: 10.1007/s10113-018-1352-4

Harrison, Paula A., Dunford, R. W., Holman, I. P., & Rounsevell, M. D. A. (2016). Climate change impact modelling needs to include cross-sectoral interactions. *Nature Climate Change*, *6*(9), 885–890. doi: 10.1038/nclimate3039

Holman, I., Audsley, E., Berry, P., Brown, C., Bugmann, H., Clarke, L., … Wimmer, F. (2017). *Modelling Climate Change Impacts, Adaptation and Vulnerability in Europe: IMPRESSIONS project deliverable*. Retrieved from http://www.impressions-project.eu/getatt.php?filename=D3B.2_European_Scale_CCIAV_Applications_FINAL_14212.pdf

Holman, I. P., Brown, C., Janes, V., & Sandars, D. (2017). Can we be certain about future land use change in Europe? A multi-scenario, integrated-assessment analysis. *Agricultural Systems*, *151*, 126–135. doi: 10.1016/j.agsy.2016.12.001

Holzhauer, S., Brown, C., & Rounsevell, M. (2019). Modelling dynamic effects of multi-scale institutions on land use change. *Regional Environmental Change*, *19*(3), 733–746. doi: 10.1007/s10113-018-1424-5

Houet, T., Schaller, N., Castets, M., & Gaucherel, C. (2014). Improving the simulation of fine-resolution landscape changes by coupling top-down and bottom-up land use and cover changes rules. *International Journal of Geographical Information Science*, *28*(9), 1848–1876. doi: 10.1080/13658816.2014.900775

Kay, S., Graves, A., Palma, J. H. N., Moreno, G., Roces-Díaz, J. V., Aviron, S., … Herzog, F. (2019). Agroforestry is paying off – Economic evaluation of ecosystem services in European landscapes with and without agroforestry systems. *Ecosystem Services*, *36*, 100896. doi: 10.1016/J.ECOSER.2019.100896

Kebede, A. S., Dunford, R., Mokrech, M., Audsley, E., Harrison, P. A., Holman, I. P., … Wimmer, F. (2015). Direct and indirect impacts of climate and socio-economic change in Europe: a sensitivity analysis for key land- and water-based sectors. *Climatic Change*, *128*(3–4), 261–277. doi: 10.1007/s10584-014-1313-y

Kok, K., Pedde, S., Gramberger, M., Harrison, P. A., & Holman, I. P. (2019). New European socio-economic scenarios for climate change research: operationalising concepts to extend the shared socio-economic pathways. *Regional Environmental Change*, *19*(3), 643–654. doi: 10.1007/s10113-018-1400-0

Ligmann-Zielinska, A., Church, R., & Jankowski, P. (2008). Spatial optimization as a generative technique for sustainable multiobjective land-use allocation. *International Journal of Geographical Information Science*, *22*(6), 601–622. doi: 10.1080/13658810701587495

Ligmann-Zielinska, A., Kramer, D. B., Cheruvelil, K. S., & Soranno, P. A. (2014). Using uncertainty and sensitivity analyses in socioecological agent-based models to improve their analytical performance and policy relevance. *PLoS ONE*, *9*(10). doi: 10.1371/journal.pone.0109779

Lippe, M., Bithell, M., Gotts, N., Natalini, D., Barbrook-Johnson, P., Giupponi, C., … Thellmann, K. (2019). Using agent-based modelling to simulate social-ecological systems across scales. *GeoInformatica*, *23*(2), 269–298. doi: 10.1007/s10707-018-00337-8

Low, S., & Schäfer, S. (2020). Is bio-energy carbon capture and storage (BECCS) feasible? The contested authority of integrated assessment modeling. *Energy Research and Social Science*, *60*, 101326. doi: 10.1016/j.erss.2019.101326

Meiyappan, P., Dalton, M., O’Neill, B. C., & Jain, A. K. (2014). Spatial modeling of agricultural land use change at global scale. *Ecological Modelling*, *291*, 152–174. doi: 10.1016/j.ecolmodel.2014.07.027

Millington, J. D. A., Demeritt, D., & Romero-Calcerrada, R. (2011). Participatory evaluation of agent-based land-use models. *Journal of Land Use Science*, *6*(2–3), 195–210. doi: 10.1080/1747423X.2011.558595

Müller, B., Balbi, S., Buchmann, C. M., de Sousa, L., Dressler, G., Groeneveld, J., … Weise, H. (2014). Standardised and transparent model descriptions for agent-based models: Current status and prospects. *Environmental Modelling and Software*, *55*, 156–163. doi: 10.1016/j.envsoft.2014.01.029

Müller, B., Hoffmann, F., Heckelei, T., Müller, C., Hertel, T. W., Polhill, J. G., … Webber, H. (2020). Modelling food security: Bridging the gap between the micro and the macro scale. *Global Environmental Change*, *63*, 102085. doi: 10.1016/j.gloenvcha.2020.102085

Murray-Rust, D., Brown, C., van Vliet, J., Alam, S. J., Robinson, D. T., Verburg, P. H., & Rounsevell, M. (2014). Combining agent functional types, capitals and services to model land use dynamics. *Environmental Modelling & Software*, *59*, 187–201. doi: 10.1016/j.envsoft.2014.05.019

Newland, C. P., Maier, H. R., Zecchin, A. C., Newman, J. P., & van Delden, H. (2018). Multi-objective optimisation framework for calibration of Cellular Automata land-use models. *Environmental Modelling and Software*, *100*, 175–200. doi: 10.1016/j.envsoft.2017.11.012

O'Neill, B. C., Kriegler, E., Ebi, K. L., Kemp-Benedict, E., Riahi, K., Rothman, D. S., … Solecki, W. (2017). The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century. *Global Environmental Change*, *42*, 169–180. doi: 10.1016/j.gloenvcha.2015.01.004

Papadimitriou, L., Holman, I. P., Dunford, R., & Harrison, P. A. (2019). Trade-offs are unavoidable in multi-objective adaptation even in a post-Paris Agreement world. *Science of the Total Environment*, *696*, 134027. doi: 10.1016/j.scitotenv.2019.134027

Parrott, L. (2011). Hybrid modelling of complex ecological systems for decision support: Recent successes and future perspectives. *Ecological Informatics*, *6*(1), 44–49. doi: 10.1016/J.ECOINF.2010.07.001

Pedde, S., Kok, K., Hölscher, K., Frantzeskaki, N., Holman, I., Dunford, R., … Jäger, J. (2019). Advancing the use of scenarios to understand society's capacity to achieve the 1.5 degree target. *Global Environmental Change*, *56*, 75–85. doi: 10.1016/J.GLOENVCHA.2019.03.010

Polhill, J. G., & Gotts, N. M. (2009). Ontologies for transparent integrated human-natural system modelling. *Landscape Ecology*, *24*(9), 1255–1267. doi: 10.1007/s10980-009-9381-5

Prestele, R., Alexander, P., Rounsevell, M. D. A., Arneth, A., Calvin, K., Doelman, J., … Verburg, P. H. (2016). Hotspots of uncertainty in land-use and land-cover change projections: a global-scale model comparison. *Global Change Biology*, *22*(12), 3967–3983. doi: 10.1111/gcb.13337

Rosa, I. M. D., Ahmed, S. E., & Ewers, R. M. (2014). The transparency, reliability and utility of tropical rainforest land-use and land-cover change models. *Global Change Biology*, *20*(6), 1707–1722. doi: 10.1111/gcb.12523

Salganik, M. J., Lundberg, I., Kindel, A. T., Ahearn, C. E., Al-Ghoneim, K., Almaatouq, A., … McLanahan, S. (2020).

Measuring the predictability of life outcomes with a scientific mass collaboration. *Proceedings of the National Academy of Sciences*, *117*(15), 201915006. doi: 10.1073/pnas.1915006117

Schwarze, J., Sophie Holst, G., & Mußhoff, O. (2014). Do farmers act like perfectly rational profit maximisers? Results of an extra-laboratory experiment. *International Journal of Agricultural Management*, *4*(1), 11–20. doi: 10.5836/ijam/2014-01-03

Seo, B., Brown, C., & Rounsevell, M. (2018). Evaluation and calibration of an agent based land use model using remotely sensed land cover and primary productivity data. *International Geoscience and Remote Sensing Symposium (IGARSS)*, *2018-July*, 7472–7475. Institute of Electrical and Electronics Engineers Inc. doi: 10.1109/IGARSS.2018.8518023

Seppelt, R., Lautenbach, S., & Volk, M. (2013, October 1). Identifying trade-offs between ecosystem services, land use, and biodiversity: A plea for combining scenario analysis and optimization on different spatial scales. *Current Opinion in Environmental Sustainability*, Vol. 5, pp. 458–463. Elsevier. doi: 10.1016/j.cosust.2013.05.002

Strauch, M., Cord, A. F., Pätzold, C., Lautenbach, S., Kaim, A., Schweitzer, C., … Volk, M. (2019). Constraints in multi-objective optimization of land use allocation – Repair or penalize? *Environmental Modelling and Software*, *118*, 241–251. doi: 10.1016/j.envsoft.2019.05.003

Turner, P. A., Field, C. B., Lobell, D. B., Sanchez, D. L., & Mach, K. J. (2018). Unprecedented rates of land-use transformation in modelled climate change mitigation pathways. *Nature Sustainability*, *1*(5), 240–245. doi: 10.1038/s41893-018-0063-7

Verburg, P. H., Alexander, P., Evans, T., Magliocca, N. R., Malek, Z., Rounsevell, M. DA, & van Vliet, J. (2019, June 1). Beyond land cover change: towards a new generation of land use models. *Current Opinion in Environmental Sustainability*, Vol. 38, pp. 77–85. Elsevier B.V. doi: 10.1016/j.cosust.2019.05.002

Verburg, P. H., & Overmars, K. P. (2009). Combining top-down and bottom-up dynamics in land use modeling: exploring the future of abandoned farmlands in Europe with the Dyna-CLUE model. *Landscape Ecology*, *24*(9), 1167–1181. doi: 10.1007/s10980-009-9355-7

	SSP1 socio-economic conditions gradually improve through economic growth, stable government, high social cohesion and international cooperation	**SSP3** social and economic conditions worsen, with limited and ineffective political responses	**SSP4** large economic inequalities and fluctuations develop, low social cohesion, but high technological investment & environmental protection	**SSP5** emphasis on social and economic development, fossil fuel exploitation and technology
RCP2.6 Very low climate change	IAP simulates more intensive and pastoral agriculture and very little forest. CRAFTY increases forest at the expense of intensive agriculture. Under-supply of timber (especially in the IAP) and under-supply of food (only in CRAFTY).		Widespread agricultural extensification and abandonment in the IAP, and more forestry, but with under-supply of timber (agriculture shifts to optimal areas). More intensive agriculture in CRAFTY, but with under-supply of food (agriculture persists in less optimal areas).	
RCP4.5 Intermediate climate change	Small differences, with the IAP having a slight shift towards pastoral and very extensive agriculture, with less forest.	Limited change in the IAP but dramatic loss of intensive management in CRAFTY, along with fragmentation, temporal dynamism and supply shortfalls.	Widespread agricultural abandonment in the IAP. CRAFTY supply levels exceed demand	
RCP8.5 High climate change		Limited change in both models, with more extensification, forest and multifunctional production in CRAFTY.		Widespread agricultural extensification and abandonment in the IAP. Limited change in CRAFTY, with supply levels exceeding demand.

Table 1: Climatic and socio-economic scenario identities, summaries, and main findings.

Socio-economic scenario	**SSP1**	**SSP3**	**SSP4**	**SSP5**
Climate change (RCP pairing)	Very low / intermediate	Intermediate/ high	Very low / Intermediate	High
EU population change (% change from 2010)	0.4	-38	-22	47
Food imports (absolute % change)	-13	-5	4	18
Increase in arable land used for biofuel production (% change from 2010)	9	19	9	14
Land allocated to agri-environment schemes (e.g. set-aside, buffer strips, beetle banks) (%; baseline is approx. 3%)	6	2	5	0
Change in dietary preferences for beef and lamb (% change from 2010)	-82	0	0	53
Change in dietary preferences for chicken and pork (% change from 2010)	-34	35	35	74
Change in agricultural mechanisation (% change from 2010)	133	-35	133	133
Change in agricultural yields (% change from 2010)	-19	-35	89	89
Change in irrigation efficiency (% change in water efficiency relative to 2010); -50% = water halved per unit food	-57	53	-57	-57
Reducing diffuse source pollution from agriculture by reduced inputs of fertilisers and pesticides (higher value = less inputs) (absolute value relative to optimum nitrogen)	1.9	0.9	0.9	0.9
Water savings due to behavioural change (% change from 2010)	52	0	0	-30
Water savings due to technological change (% change from 2010)	45	0	29	29
GDP (% change from 2010)	259	48	200	724
Change in energy price (oil; % of 2010)	162	350	267	75
Household externalities (preferences for lived environment: 1 = Urban; 5 = Country). Baseline = 3	5	4	2	5
Compact vs sprawled development (Low = Sprawl; Medium or High = Compact); Baseline = Med	High	Low	Medium	Low

Preference to live by the coast (Low – High); Baseline = Med	Low	Low	Med	High

Table 2a: Details of the socio-economic scenarios (Shared Socioeconomic Pathways, SSPs) as simulated by the IAP. Values are shown for the 2080s timeslice. Table adapted from Harrison et al. 2019.

Emission scenario	**RCP2.6**	**RCP4.5**	**RCP8.5**
Climate change	Very low	Intermediate	High
GCM	EC_Earth	HadGEM2-ES	HadGEM2-ES
RCM	RCA4	RCA4	RCA4
GCM sensitivity	Intermediate	High	High
European ΔT / ΔPr	1.4°C / 4%	3.0°C / 3%	5.4°C / 5%

Table 2b: Details of the climate scenarios used in both models. RCP denotes Representative Concentration Pathway, GCM: General Circulation Model, RCM: Regional Climate Model. Change in temperature (ΔT) and change in precipitation (ΔPr) are relative to 1961-1990, and affect productivities as simulated by meta-models in the IAP, which are then fed into the alternative land use models (Fig. 1). Further details are available in Harrison et al. (2019).

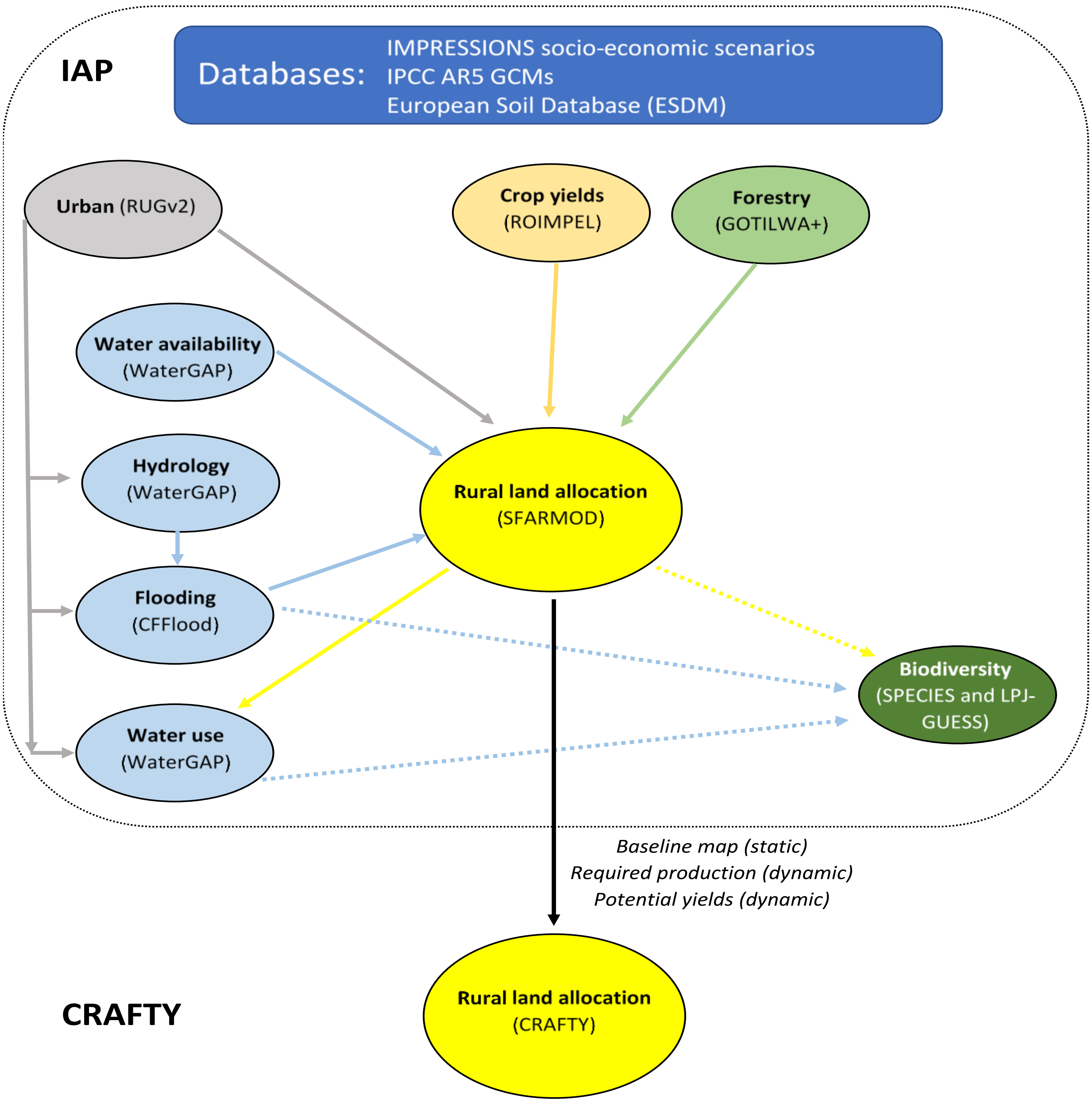
	IAP	**CRAFTY-EU**	**Key differences**
Modelling paradigm	‘Top-down’ model that represents land use change as single systemic response to drivers	‘Bottom-up’ model that represents land use change as emergent from responses of multiple entities within the system	Entirely distinct conceptualisation of land use change within shared reductionist (modelling) approach
Theoretical basis	Consistent with positivist and classical economic theories of system-level dynamic equilibrium under exogenous pressures (Brown et al. 2016)	Consistent with methodological individualism and subjective expected utility theory of decision-making given uncertainty and non-economic motivations (Murray-Rust et al. 2014; Brown et al. 2016)	Neither model explicitly theory-driven but are consistent with opposing theoretical movements.
Land allocation	Optimisation to satisfy food demand, subject to constraints imposed by biophysical and socio-economic conditions	Individual agent decisions based on competition to satisfy demands for ecosystem services	Land allocation is imposed in the IAP but emergent in CRAFTY, and therefore more variable
Variables considered (inputs)	Defined in Table 2	Potential and realised ES provision levels (derived from the IAP and dependent on the variables in Table 2) and agent abilities to produce ecosystem services, sensitivities to	Most inputs are shared directly or indirectly, although the IAP more explicitly includes biophysical conditions

		capital levels, willingness and time-dependent probability of abandoning their cells or relinquishing to other land uses when at a competitive disadvantage, and abilities to search for new cells to take over.	and CRAFTY human behaviour
Mathematical characteristics	Produces single, optimal results (subject to constraints) at each timeslice	Stochastic and path-dependent; produces sub-optimal and variable results	The IAP is more mathematically constrained, but complexity of 'option space' makes results of both models difficult to anticipate
Evaluation	Extensively evaluated, including uncertainty analyses and comparison to independent data and other models (e.g. Brown et al. 2014; Harrison et al. 2015, 2016, 2019; Kebede et al. 2015; Holman et al. 2017a, b; Fronzek et al. 2019)	Extensively evaluated, including uncertainty analyses and comparison to independent data and other models (Alexander et al., 2017; Brown, Holzhauer, Metzger, Paterson, & Rounsevell, 2018; Brown, Murray-Rust, et al., 2014; Holzhauer, Brown, & Rounsevell, 2019; Seo, Brown, & Rounsevell, 2018)	No significant difference, noting that neither models targets accurate reproduction of observed changes
Uncertainty & sensitivity	Well-understood, with land use outcomes most sensitive to temperature, precipitation, yields and import levels (Kebede et al. 2015)	Well-understood, with land use outcomes most sensitive to yields (including climate effects), import levels and (to lesser extent) agent behaviour (Brown et al. 2018)	CRAFTY has sensitivities to behavioural parameters not present in the IAP
Spatial resolution	10 arcminutes (approx. 16km in Europe), with up to 40 forms of land use and management proportionally distributed within each cell	10 arcminutes (approx. 16km in Europe), with continuous variation in characteristics within 17 forms of land use and management	Identical resolution for defined classes, but different forms and extents of variation within those classes
Temporal resolution	Timeslices: Baseline, 2020s, 2050s, 2080s	Annual 2016-2086	CRAFTY has higher temporal resolution
Principal uses	Research, education, capacity building (students and stakeholders)	Research, education	CRAFTY less used in stakeholder engagement

Table 3: Summary comparison of the two models used in this study across a range of characteristics, many of which stem from the distinct modelling paradigms used. Further details are provided in the text, and references cited there.

Land use classes for comparison	Explanation
Intensive agriculture	Intensive forms of agriculture primarily dedicated to crop production but including some grassland
Extensive agriculture	Extensive forms of arable and pastoral agriculture
Pastoral agriculture	Dedicated and primarily intensive pastoral agriculture
Very extensive management	Management for any service that is of the lowest intensity and leaves land in a near-natural state
Forestry	Active management for timber extraction and other forest services
Other/no management	Land that is not actively managed for agriculture or forestry, but which can have a range of natural or human-impacted land covers

Table 4~~1~~: Land use classes used in the comparison and their composition. Derivations from the full range of CRAFTY and IAP classes are given in Table A1.



695 **Figure 1: Simplified schematic showing the structure of the IAP in terms of its component metamodels, and its relationship to CRAFTY in this study. Results presented in this study are taken from the alternative land allocation models (yellow), and results from the biodiversity model are not used. The information transferred from the IAP to CRAFTY utilises all of the inputs to SFARMOD and describes initial and scenario-dependent conditions affecting agent decision-making in CRAFTY.**

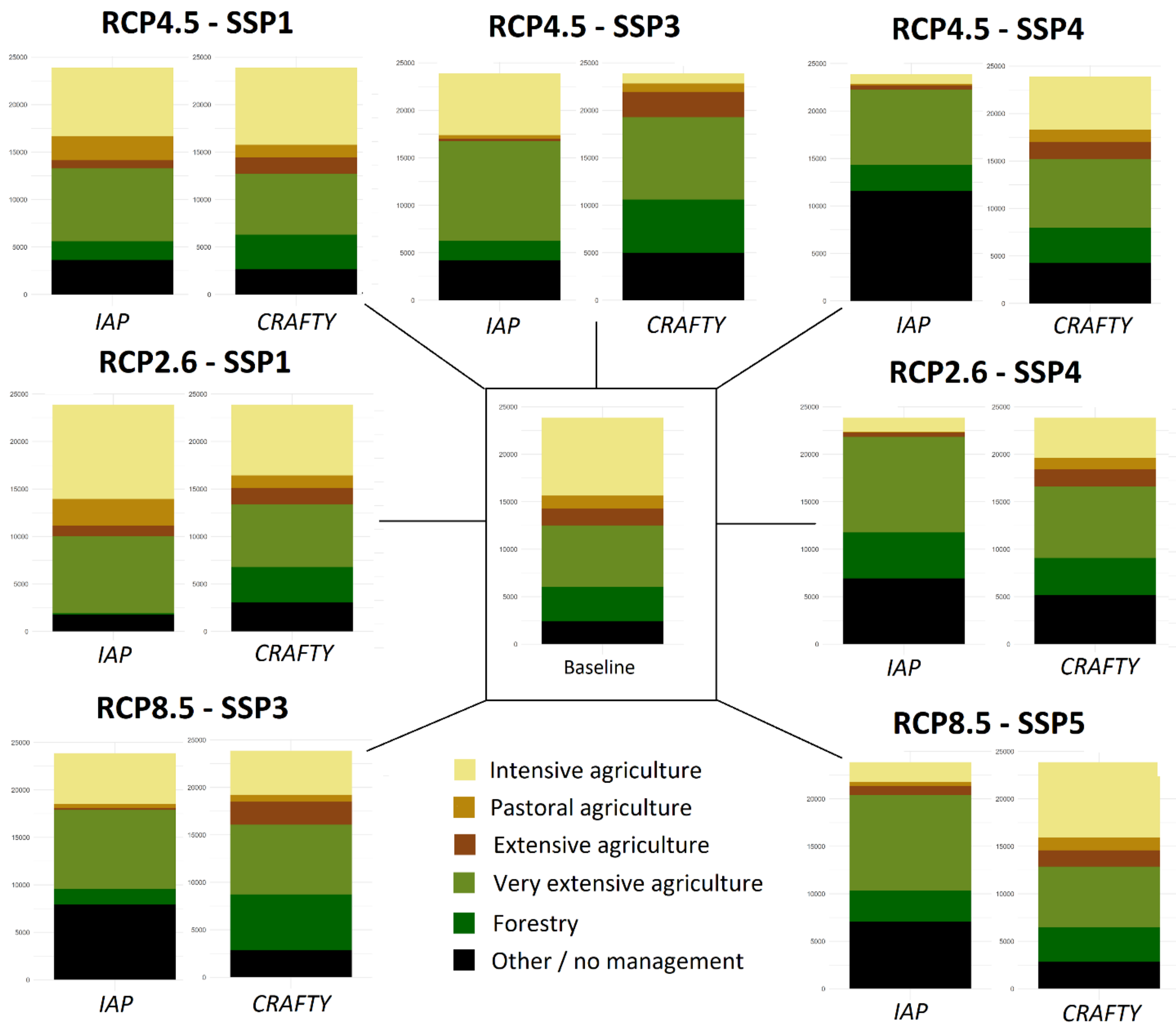
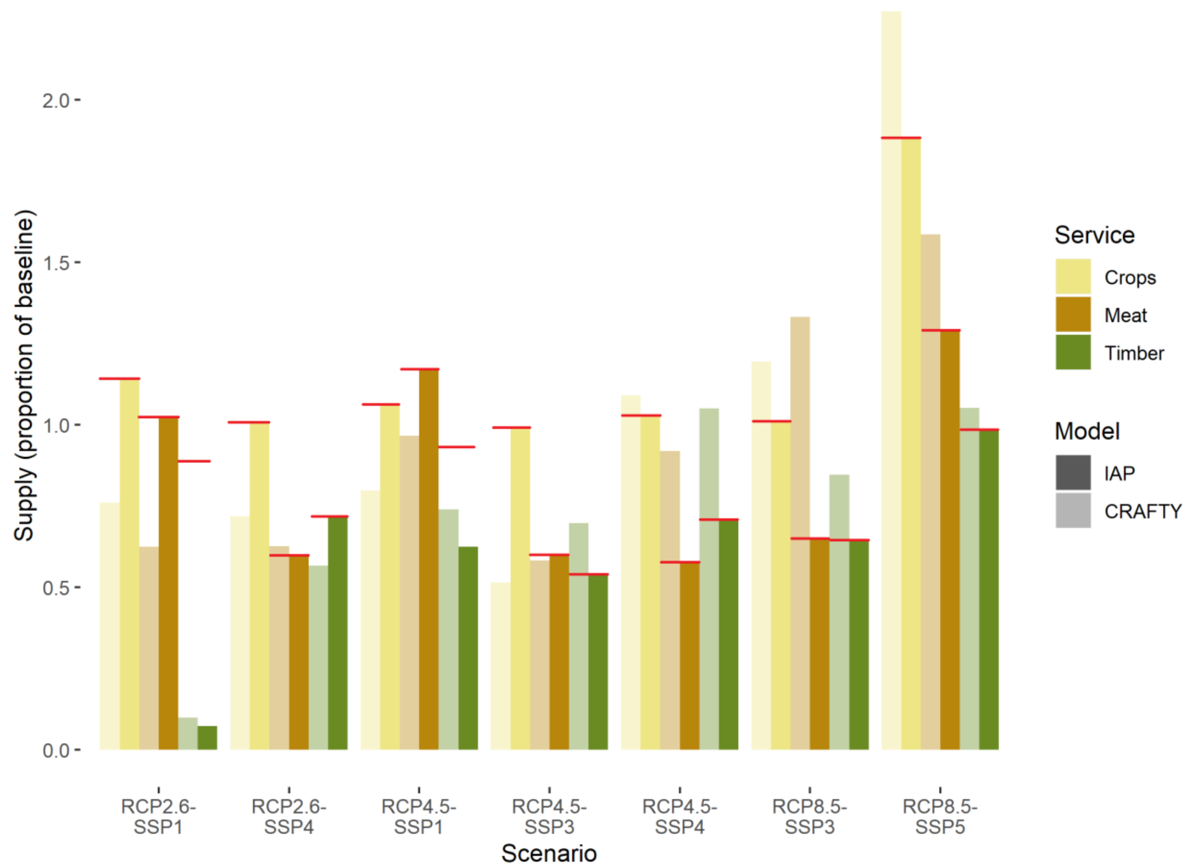
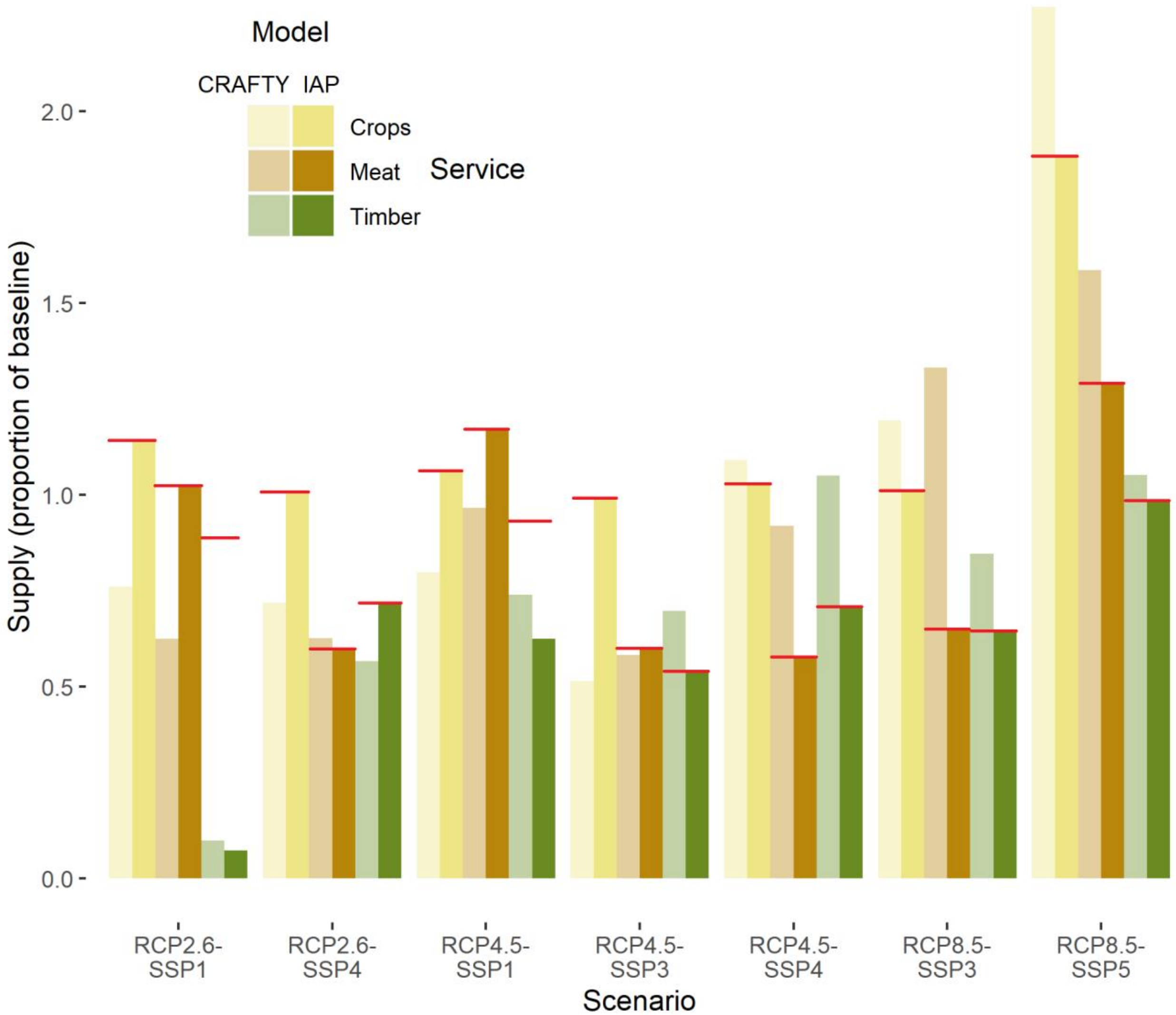
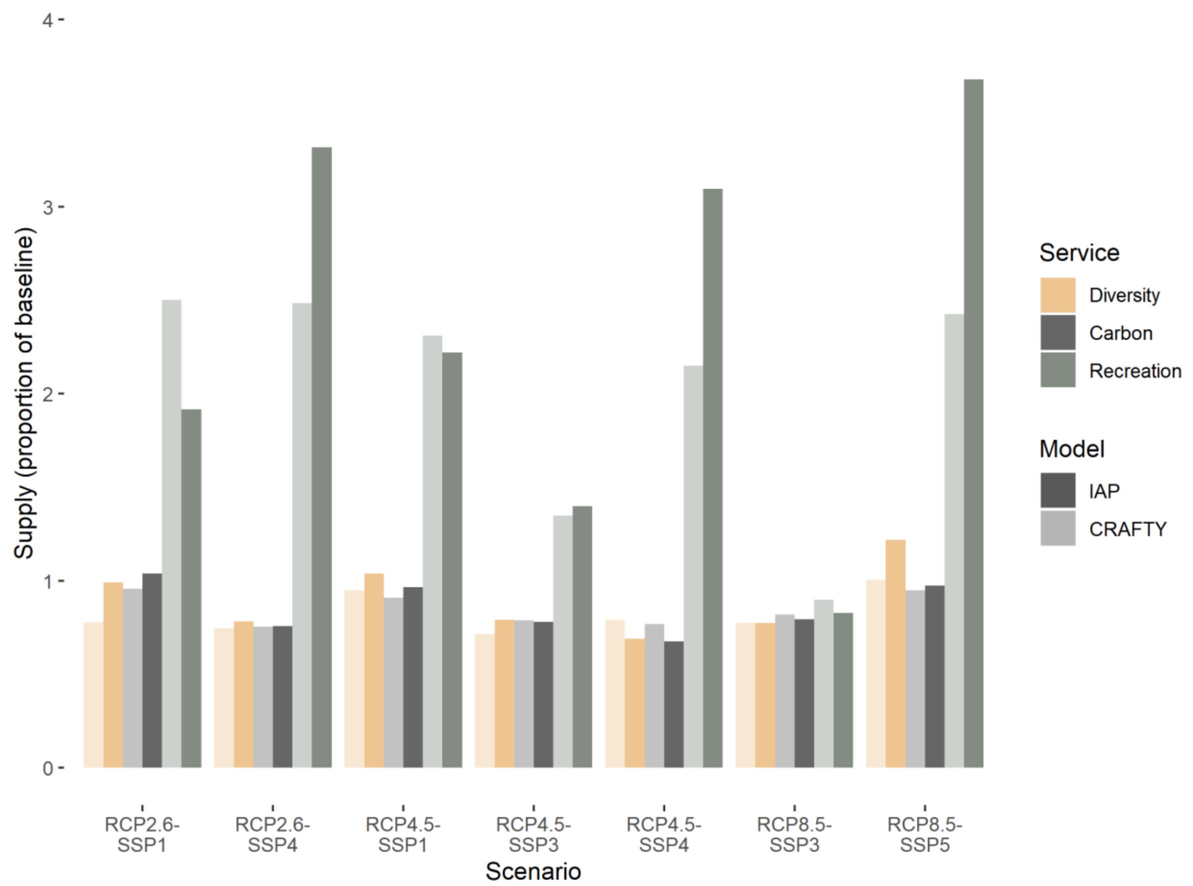


Figure 2~~1~~: Simulated land use classes for each scenario in each model in the mid-2080s. Bars show the number of cells occupied by each class, out of the total number of 23,871 cells (y-axes). Climate scenarios (RCPs) are arranged in rows. The baseline is identical in both models and so is only shown once.

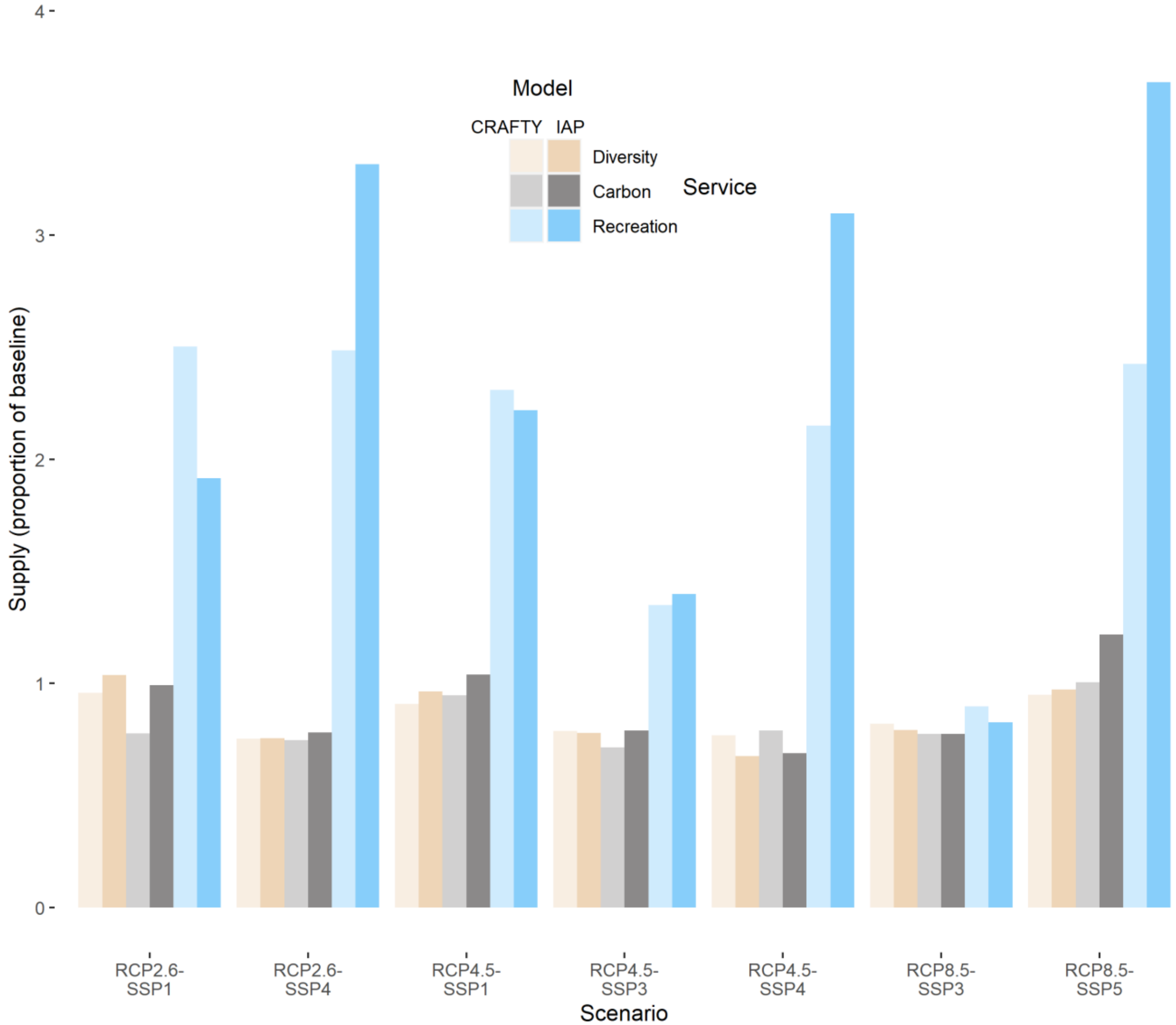

Supply (proportion of baseline)
2.0
1.5
1.0
0.5
0.0
RCP2.6-SSP1
RCP2.6-SSP4
RCP4.5-SSP1
RCP4.5-SSP3
RCP4.5-SSP4
RCP8.5-SSP3
RCP8.5-SSP5
Scenario
Service
Crops
Meat
Timber
Model
IAP
CRAFTY



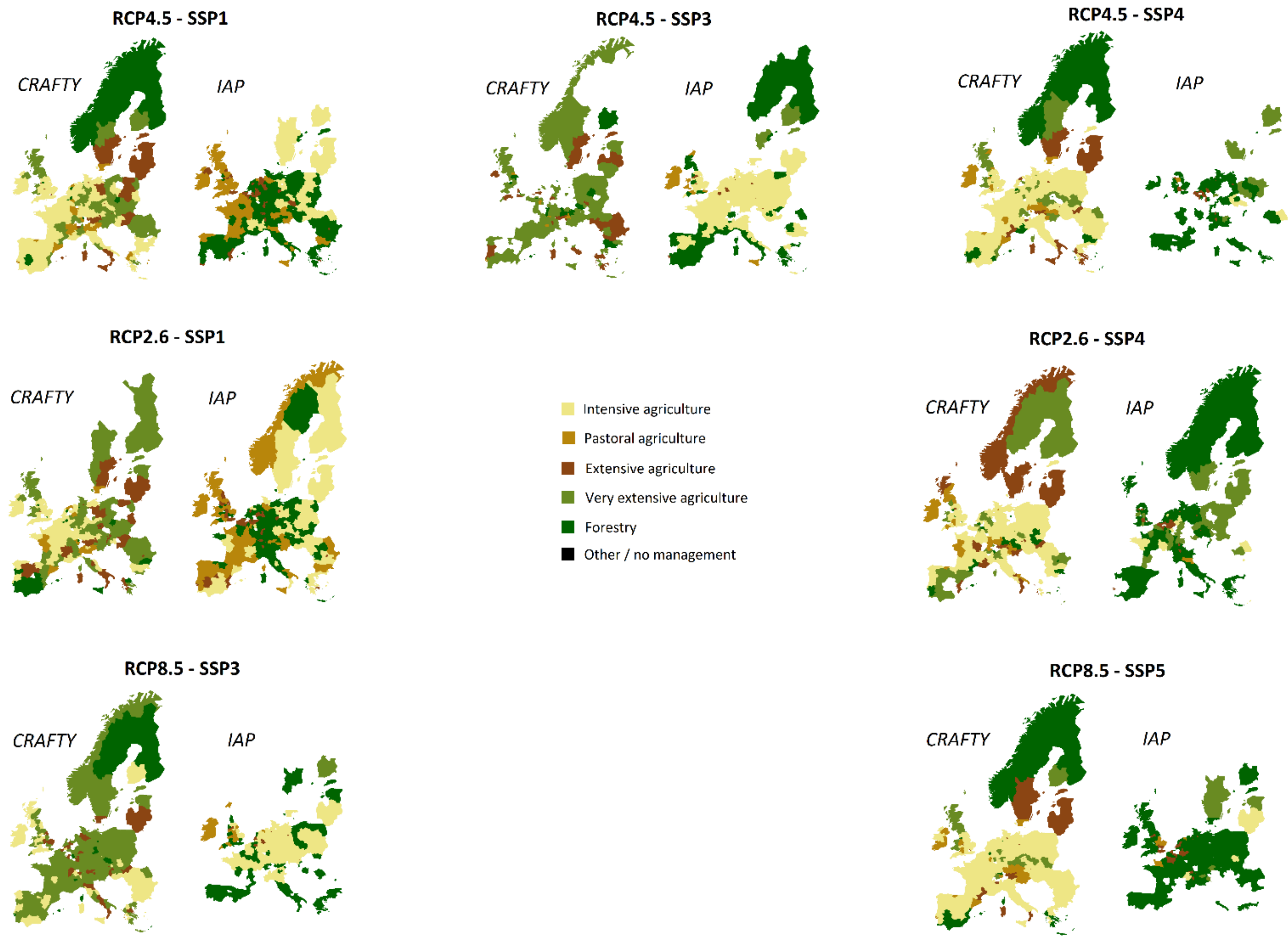
Figure~~.~~ 3~~2~~a: Supply levels of services ~~actively modelled in~~that both models attempt to satisfy demand for, in each scenario. Demand levels (derived from the IAP) are indicated by a red line for each service in each scenario. ~~supply levels are linked to scenario conditions and are set as demands for CRAFTY by default, after being calculated using CRAFTY production functions to ensure comparability.~~ IAP supplies are unequal to demand levels only where the IAP reports an underproduction of a particular service (in these results, timber in SSP1 simulations). A supply value of 1.0 (y-axis) is equal to baseline supply~~, and the scenario-specific demand level for each service is shown in red~~.



4
3
2
1
0
Supply (proportion of baseline)
RCP2.6-
SSP1
RCP2.6-
SSP4
RCP4.5-
SSP1
RCP4.5-
SSP3
RCP4.5-
SSP4
RCP8.5-
SSP3
RCP8.5-
SSP5
Scenario
Service
Diversity
Carbon
Recreation
Model
IAP
CRAFTY



Figure~~.~~ 3b~~2b~~: Supply levels of services that CRAFTY attempts to satisfy demands for, but the IAP does not. ~~with no defined demands in the IAP.~~ **IAP supply levels here are calculated using CRAFTY production functions and then set as demands for CRAFTY**~~, with production having equivalent value to the three primary services (Fig. 2a)~~**.** Demand levels are therefore equal to IAP supply by default and are not indicated by a line as in Fig. 3a. ~~The IAP therefore does not attempt to achieve particular supply levels for these services, while CRAFTY does.~~ A supply value of 1.0 (y-axis) is equal to baseline supply.


RCP4.5 - SSP1
CRAFTY
IAP
RCP4.5 - SSP3
CRAFTY
IAP
RCP4.5 - SSP4
CRAFTY
IAP
RCP2.6 - SSP1
CRAFTY
IAP
Intensive agriculture
Pastoral agriculture
Extensive agriculture
Very extensive agriculture
Forestry
Other / no management
RCP2.6 - SSP4
CRAFTY
IAP
RCP8.5 - SSP3
CRAFTY
IAP
RCP8.5 - SSP5
CRAFTY
IAP

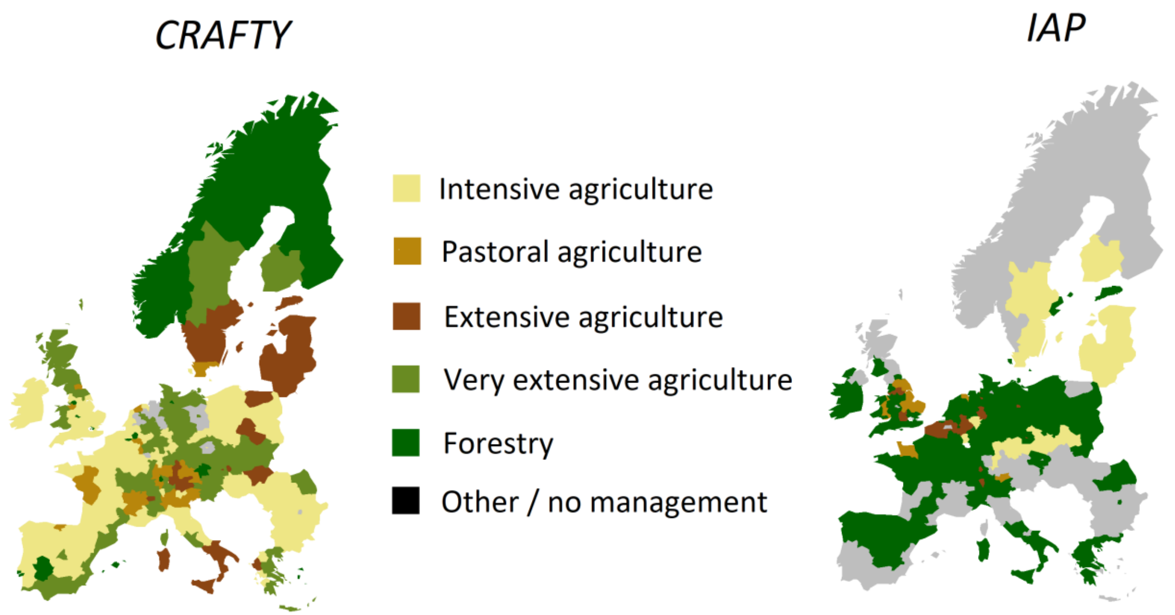


Fig 4~~3~~: ~~Spatialised~~ Geographical differences between the models' results ~~for each scenario~~across all scenarios, at NUTS2 level. Colours identify the most over-represented land use type in each region in the CRAFTY and IAP results, relative to the result of the other model (i.e. the land use with the biggest difference in occurrence in that region). ~~White~~ Grey is shown where no land use type has an over-representation of more than 5% of the region's cells. Scenario-specific results are shown in Fig. B1 (Appendix B).

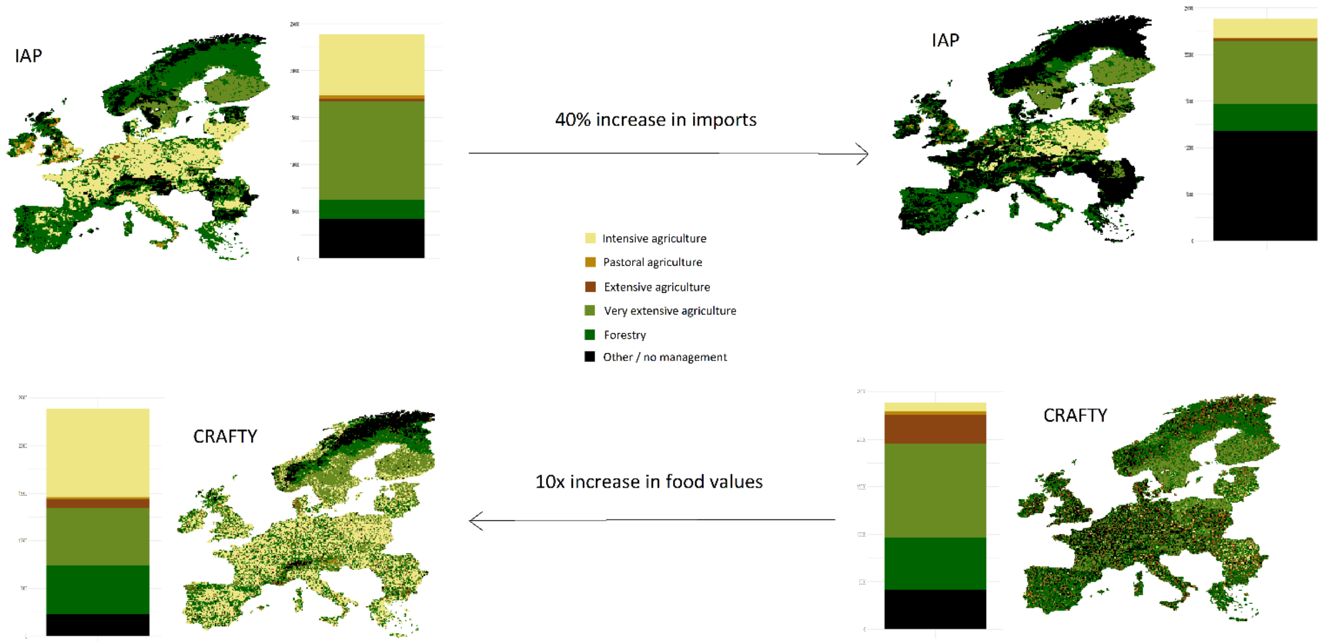


Fig. 54: Cell-level and ~~aggregate~~ EU-level results for the RCP4.5-SSP3 scenario with and without alternative parameterisations designed to introduce analogous driving conditions to each model in turn. The IAP experiment is shown on the top row, and the CRAFTY experiment on the bottom. The original IAP result (top-left) moves towards the original CRAFTY result (bottom-right) with a 40% increase in imports allowing less production of food within Europe, resulting in widespread land abandonment in the new IAP result (top-right). The original CRAFTY result (bottom-right) moves towards the original IAP result (top-left) with a 10-fold increase in food prices, used to stimulate production, resulting in far more intensive agriculture in the new CRAFTY result (bottom-left). In neither case do the new results reproduce the original extremes.

Appendix A: Land use class composition

Ecosystem services production in CRAFTY is derived from that of the IAP, which uses a suite of meta-models to simulate production levels as described in (Paula A. Harrison et al., 2019), and is presented in detail in Brown et al. (2019). CRAFTY-EU also shares a baseline map with the IAP, with the aggregated land use classes used here derived from CRAFTY's Agent Functional Types (AFTs) and the IAP's land use classes as described in Table A1.

Agent Functional Type	IAP Class	Aggregated class
Intensive arable farming	Intensively farmed	Intensive agriculture
Intensive agro-forestry mosaic	Combinations of: Intensively farmed, intensively grass, managed forest	
Intensive farming	Combinations of: Intensively farmed, intensively grass	
Mixed farming	Combinations of: Intensively farmed, intensively grass, extensively grass	
Managed forestry	Managed forest	Forestry
Mixed forest	Combinations of: Managed forest, unmanaged forest	
Mixed pastoral farming	Combinations of: intensively grass, extensively grass, very extensively grass	Extensive agriculture
Extensive agro-forestry mosaic	Combinations of: extensively grass, very extensively grass, managed forest	
Peri-urban	Any combination with > 40% urban area	
Intensive pastoral farming	Intensively grass	Pastoral agriculture
Extensive pastoral farming	Extensively grass	
Very extensive pastoral farming	Very extensively grass	Very extensive management
Multifunctional	4 or more land uses in uncommon combination	
Minimal management	Combinations of: very extensively grass, unmanaged forest, unmanaged land	
Unmanaged land	Unmanaged land	Other/no management
Unmanaged forest	Unmanaged forest	
Urban	Urban	

Table A1: The composition of the aggregated land use classes used here in terms of CRAFTY-EU's Agent Functional Types (AFTs) and the IAP's land use categories. In any case where the given IAP categories occupy more than 70% of a cell, that cell is allocated to the corresponding AFT in the baseline map of *CRAFTY-EU*, except in the case of the Peri-urban AFT, for which the threshold (of urban area) is 40%. The service production potentials of each AFT are calibrated to approximately match those within the IAP classes that constitute them, so that given the same productivities in a cell, the same levels of services will be produced. Names are therefore assigned in both cases on the basis of dominant land uses and do not account for minor variations in land use and production within them.

Appendix B: Full geographical scenario results

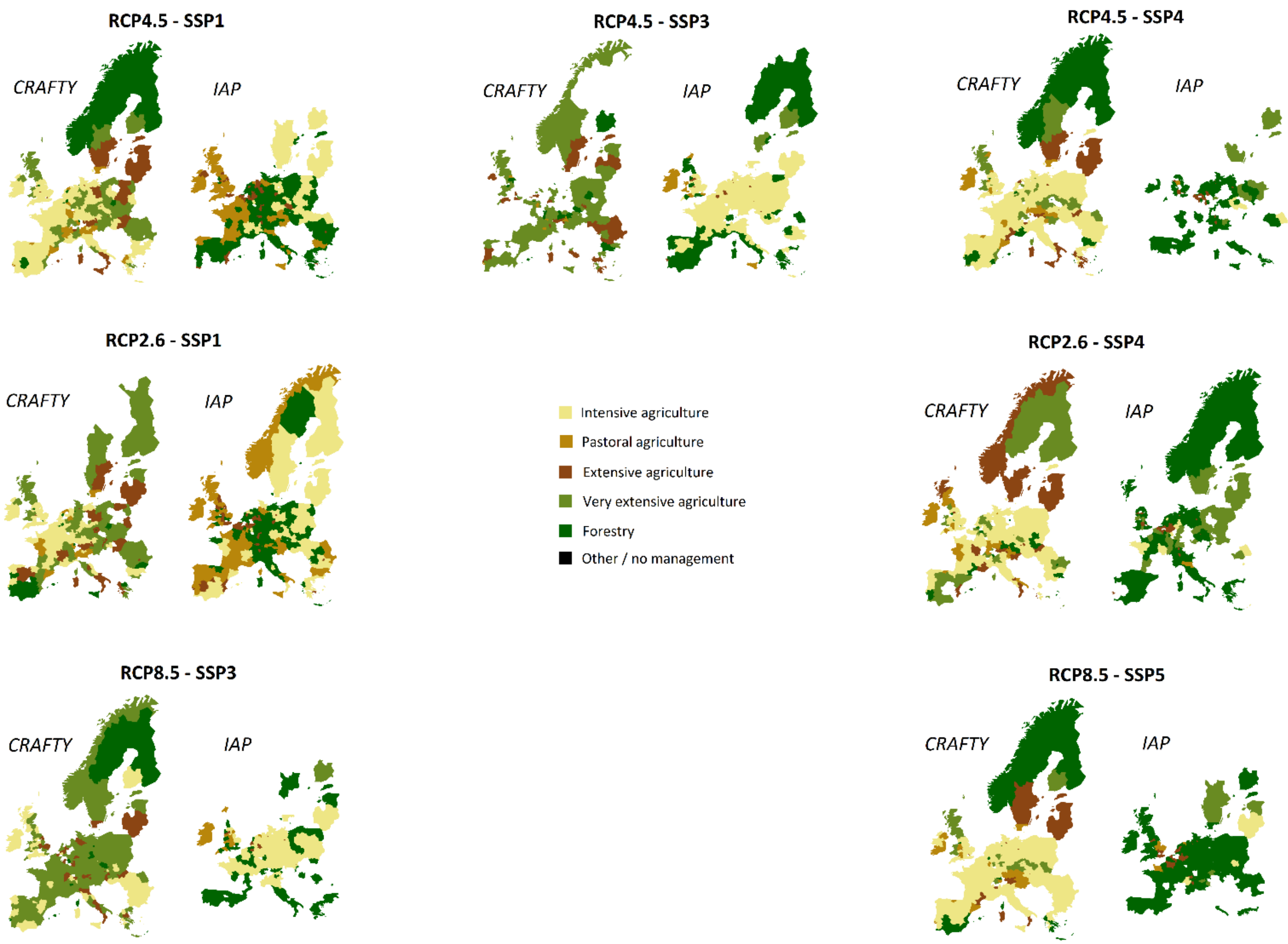


Fig B1: Spatialised differences between the models' results for each scenario, at NUTS2 level. Colours identify the most over-represented land use type in each region in the CRAFTY and IAP results, relative to the result of the other model. White is shown where no land use type has an over-representation of more than 5% of the region's cells.