Title: A fractional land use change model for ecological applications

Authorship: Simon Kapitza[1*], Nick Golding[2,3], & Brendan A. Wintle[1]

[1] School of Ecosystem and Forest Sciences, The University of Melbourne, Parkville VIC 3010, Australia
[2] Telethon Kids Institute, Perth Children's Hospital, 15 Hospital Ave, Nedlands WA 6009, Australia
[3] Curtin University, Kent St, Bentley WA 6102, Australia

Corresponding author*: simon.kapitza.research@gmail.com; +49 171 328 748 27

Authorship contribution statement

SK, BW conceived the idea of providing a fractional land use model. NG, and SK and BW designed the model and validation steps. SK coded the model and analyzed the data. SK led the manuscript with edits from NG and BW.

Data accessibility

All input data required to repeat this work will be made available and receive a permanent DOI through FigShare upon publication. An R package containing the model source code is available on GitHub (kapitzas/flutes). All code for data preprocessing and analysis is also provided through a GitHub repository (kapitzas/frac_lumodel). Both repositories will be receive permanent DOI through Zenodo upon publication.

Highlights

- A model to predict fine-resolution fractional land use change, based on future land use demands.
- Development version implemented as R package, making the model accessible to R-trained users.
- Validation shows high prediction accuracy in Amazon basin, but the model can be fitted anywhere to predict future land use change.

Abstract

By mapping land use under projections of socio-economic change, ecological changes can be predicted to inform conservation decision-making. We present a land use model that enables the fine-scale mapping of land use change under future scenarios. Its predictions can be used as input to virtually all existing spatially-explicit ecological models. Our model maps the fractional cover of land use within each grid cell, providing higher information content than discrete classes at the same spatial resolution. The method accurately reproduced land use patterns observed in the Amazon, both in terms of the allocated fractional amounts and also the direction of predicted land use changes. A small case study showcases the application of our model to reproduce patterns of agricultural expansion and natural habitat declines. The model source code is provided as an open-source R package, making this new method available to bridge the gap

between socio-economic, land use and biodiversity modelling.

Keywords: Land use forecasting, fractional land cover, continuous fields, agricultural expansion, socio-economic change, biodiversity conservation.

1 Introduction

1.1 Accounting for land use change in biodiversity assessments

Land use change is a key driver of global environmental change, causing global declines in biodiversity, species extinctions and resulting in the deterioration of ecosystem services (Foley, 2005; IPBES, 2019). There is mounting evidence of adverse impacts of land use change on biodiversity. The need for global assessments of future biodiversity change in response to land use change has been increasingly acknowledged (Urban et al., 2016; Kim et al., 2018; Powers and Jetz, 2019). However, work concerned with understanding future biodiversity change tends to focus on climate change (Titeux et al., 2016; Struebig et al., 2015; Urban et al., 2016), or other aggregated effects of socio-economic change, such as forest loss (Pérez-Vega, 2012; Margono et al., 2014) and urban expansion (Seto et al., 2012). This is despite the fact that land use change is highly driven by dynamic bio-physical and socio-economic processes (Lambin et al., 2011). Climate change will likely result in global shifts and declines of land suitable for agricultural production, with projected depletion of land reserves in the first half of the 21st century (Lambin et al., 2011). Food production and international trade of goods will significantly increase (O'Neill et al., 2014, 2017) and even under lowest impact scenarios crop and livestock production are still likely to be higher and occupy a larger land area than they do today (O'Neill et al., 2014).

Consequently, future predictions of biodiversity change will benefit from explicit accounting of the drivers and effects of land use change at the level of individual types of use. Detailed, large-scale maps of future land use under competing future scenarios provide useful insights for researchers and policy makers, particularly in terms of informing conservation planning and preventing future biodiversity loss.

1.2 Overview of land use modelling approaches

Different modelling approaches have been developed to determine the overall amount of future land use and allocate changes across the landscape. Artificial neural networks and markov chain models learn and infer total amounts and spatial patterns of land use change from historic time series (Tayyebi and Pijanowski, 2014; Pijanowski et al., 2002). Markov chain models have been frequently combined with cellular automata (CA-Markov models, see Hyandye and Martz, 2017; Aburas et al., 2017; van Schrojenstein Lantman et al., 2011). In cellular automata the transition probability of a cell to another land use depends on its current state and the state of neighbouring cells, both of which are the result of historic changes (van Schrojenstein Lantman et al., 2011). Cellular automata have been used successfully to simulate strongly auto-correlated changes, such urban sprawl (Verburg et al., 2004b; Fang et al., 2005; Shafizadeh Moghadam and Helbich, 2013; Sun et al., 2007).

Many modelling approaches apply regression analysis and other techniques to identify associations between various environmental conditions and observed land use patterns (van Schrojenstein Lantman et al., 2011; Lambin et al., 2000; Verburg et al., 2004b). Available models applying this approach include SLEUTH (Slope, Land Use, Exclusion, Urban, Transportation, Hillshade, Dietzel and Clarke, 2007), Dinamica EGO (Environment for Geoprocessing Objects, Soares-Filho et al., 2009), LCM in TerrSet (Land Change Modeler, Eastman and Toledano, 2018) and, perhaps most prominently, the CLUE model series (Conversion of Land Use and its Effects, Verburg and Overmars, 2009) (Table 1). CLUE models have found application in the prediction of spatially-explicit patterns of land use at national and continental scales (Veldkamp and Fresco, 1996; Verburg and Overmars, 2009; Verburg et al., 1999, 2002; Kapitza et al., 2021). Exogenously determined future changes in area demands for different land uses, often predicted by an economic model (Aguiar et al., 2016), may be downscaled by establishing statistical relationships between observed land use and a set of socio-economic and bio-physical drivers of land use and land use change. Predicted land use suitability surfaces inform local competition for different land uses (Verburg et al., 2002; Meiyappan et al., 2014). Models can be further parametrized by including transition rules at local (cell) and landscape levels and constraints on overall turn-over through time. More simplistic models based on statistical analysis use an ordered allocation algorithm, in which competition between land uses is handled by ordering allocations in terms of perceived socio-economic value (Fuchs et al., 2013).

Land use change allocation algorithms are agnostic to the type of statistical analysis conducted to estimate land use suitability surfaces. Nevertheless, most models apply binary logistic regression to model the cell-wise probabilities of occurrence for each land use category, independent of the probabilities of other land uses. The resulting probability of land use occurrence at a site produced by separate models is an incomplete representation of the underlying structure of land use probability, because it omits that occurrence probabilities are dependent between land use types, and that the probabilities of all discrete classes must sum to one. For example, when a site has very high probability for urban land use, this implies relatively low probabilities for primary natural habitat, which separate, independent logistic regressions do not fully capture. One step toward explicitly modelling competition between land uses is to apply multinomial regression, thus allowing for the prediction of conditional binary probabilities of multiple classes (Noszczyk, 2019).

1.3 Continuous land use fractions

Categorical land use data sets are increasingly available at spatial resolutions of finer than 1km. Three prominent examples include the CORINE (Coordination of Information on the Environment) Land Cover inventory (Bossard et al., 2000), which contains several time steps between 1990 and 2018 at 100m resolution for the European continent, global land cover maps produced for the year 2010 through Copernicus Land Monitoring Service (European Union, 2019) at the same resolution, as well as global maps of land cover in annual time steps between 1992 and 2018, produced under the European Space Agency's (ESA) Climate Change Initiative Land Cover (CCI-LC) project (European Space Agency, 2019), available at 300m resolution. However, the spatial variables that represent drivers of land use and biodiversity change are often not available over large spatial extents at fine resolutions better than 1km (Dendoncker et al., 2006). Therefore, it is necessary to resample finely resolved land use data to match the coarser resolution of driver covariates. Lowering the resolution has the additional advantage of improving computational efficiency due to the smaller number of pixels in the coarser map. Resampling fine-resolution maps by assigning a single category of land use on each coarser pixel effectively eliminates sub-pixel information on land use (Seo et al., 2016), so this approach is not desirable (Fig. 1). In order to maximise the retained information contained in the coarser map, it is preferable to calculate the fractions of land use covering each new pixel, producing continuous fields of information and keeping information at sub-pixel level (Seo et al., 2016) (Fig. 1).

The higher information content retained in fractional land use representations has high utility in ecological modelling. For example, many species may be able to persist in an area if only a small proportion of the area is made up of a suitable land class, such as remnant vegetation (Wintle et al., 2019). Many wide-ranging species may persist in landscapes if a certain proportion of the landscape is comprised of old forest. It has been shown that continuous fields of land use allow better estimation of biomass and biomass change (Xian et al., 2015) and are better able to explain variation in home range sizes (Bevanda et al., 2014) than categorical land use data. Continental-scale biodiversity assessments have shown that patterns are associated with high spatial-resolution fractional land use measures such as the regional aggregation of land use types, land cover diversity and land use covariates including land use intensity (Mouchet et al., 2015) and actual evapo-transpiration (Mouchet et al., 2015; Whittaker et al., 2006). Creating maps of some of these covariates requires fine-scale maps of fractional land use as principal input (Plutzar et al., 2016). The intensification of agriculture and forest harvesting are crucial factors shaping biodiversity (Levers et al., 2014, 2016) that require inputs of crop type and vegetation composition within each spatial unit. These ecological considerations of the utility of fractional land cover and land use representations are underpinned by recent advancements in algorithms to produce high resolution maps of fractional land cover from satellite data (Allred et al., 2021; Hill and Guerschman, 2020).

However, only few land use modelling approaches are capable of predicting continuous fractions of land use directly (see Hasegawa et al., 2017; Meiyappan et al., 2014), by providing fine-resolution categorical representations that can be resampled to coarser-resolution continuous representations (see Future Land Use Simulation Model (FLUS), Liu et al., 2017) or as part of integrated assessment frameworks (see CLIMSAVE Integrated Assessment Platform (CLIMSAVE IA), Harrison et al., 2013). However, some of these documented approaches are not available in a usable package suited to regional-continental scale (Hasegawa et al., 2017; Meiyappan et al., 2014), or provide user interfaces that do not allow seamless, reproducible integration into programmatic workflows (Harrison et al., 2013; Liu et al., 2017) (Table 1).

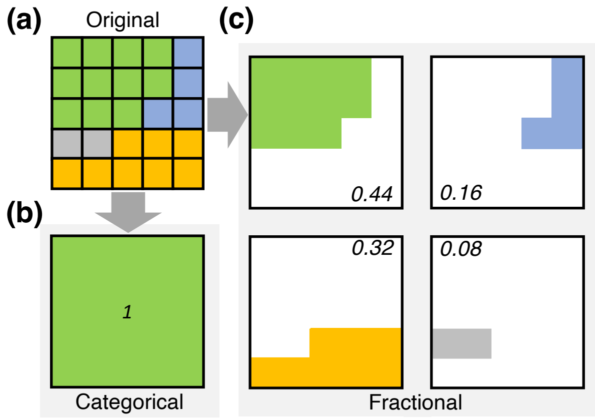

Figure 1: Illustration of methods to reduce the resolution of fine-resolution land use data. The original categorical land use grid (a) is resampled to a coarser resolution by assigning the class occupying the largest fraction in the new cell (b), effectively eliminating sub-pixel information. c) More information on a continuous scale is retained when resampling the data to a fractional representation at the coarseer resolution.

1.4 Objectives of this paper

Our new land use model FLUTES (Fractional Land Use Transitions in Ecological Systems) provides a readily available means to incorporate fractional land use change into ecological modelling. An advantage of FLUTES compared to existing fractional land use modelling approaches is its implementation in R (R Development Core Team, 2008), making use of a development environment for which high expertise already exists among ecological modellers. FLUTES can be fitted at different scales with minimal parametrization requirements and performs efficiently at various resolutions and extents. The source code for our method is freely available as a small open source R package hosted on GitHub (kapitzas/flutes). As such, our approach contributes a new open method toward bridging the gap between socio-economic, land use and biodiversity modelling.

We provide a mathematical description of the developed fractional land use model and evaluate FLUTES according to its ability to correctly estimate the direction and intensity of observed land use changes using a case study in the Brazilian Amazon.

Table 1: Comparison of land-use modelling approaches. Compared models were chosen based on their perceived feasibility in ecological research conducted by researchers with limited expertise in land-use modelling.

model	resolution	response	interface	source	license	reference
FLUS	**flexible**	categorical	GUI	C++	**freeware, open source**	Liu et al. (2017)
Dyna-CLUE	**flexible**	categorical	GUI, **R**	C++	**freeware, open source**	Verburg and Overmars (2009)
Dinamica EGO	**flexible**	categorical	GUI	C++, Java	freeware, closed source	Soares-Filho et al. (2009)
CLUMondo	flexible	categorical	GUI	C++	**freeware, open source**	Van Asselen and Verburg (2013)
SLEUTH	**flexible**	categorical	GUI	C	**freeware, open source**	Dietzel and Clarke (2007)
LCM in TerrSet	**flexible**	categorical	GUI	n/a	limited	Eastman and Toledano (2018)
CLIMSAVE IA	**flexible**	**continuous**	GUI	various (DLL)	freeware, closed source	Harrison et al. (2013)
FLUTES	**flexible**	**continuous**	**R**	**R**	**freeware, open source**	

2 Materials and methods

2.1 Model description

The model consists of two main components (Fig. 2). First, statistical analysis is used to determine how the suitability of the landscape for different land uses relates to a set of environmental drivers of land use change, producing a suitability surface for each land use class (Fig. 2a). Second, fractional changes in additional land use demands are allocated iteratively in the landscape, scaling with the land use suitability surfaces (Fig. 2b). We utilize a cellular automaton to introduce cell-level allocation decisions that constrain the location and direction of land use changes according to three rules.

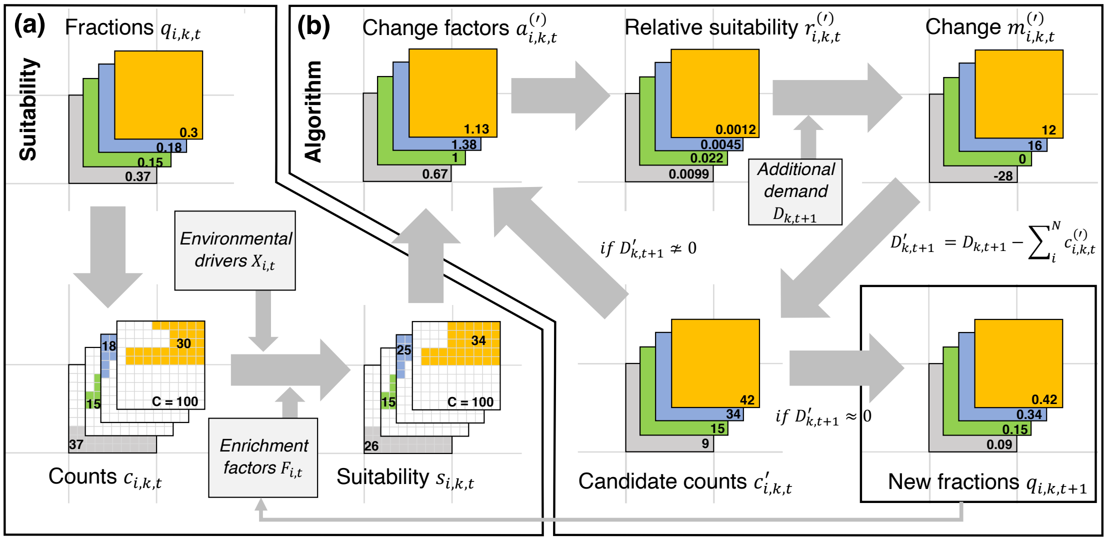

Figure 2: "Conceptual diagram of land use modelling approach. a) Land use suitability model. Observed fractions of land use are first converted to integer counts through multinomial draws and their relationship with environmental drivers and neighbourhood covariates (derived from previous time step's land-use distribution) is quantified. b) Allocation algorithm. First, it is estimated by how much each cell has to change to achieve the modelled ideal distribution of land uses. Change factors are then converted to relative suitabilities that serve to distribute land use supply required to satisfy the additional demand in the landscape. Multinomial draws ensure that each cell's land use class probabilities sum to 1. The resulting difference of the current supply and the total additional demand is recalculated to support allocation in the next iteration. The cycle repeats until the difference between the current supply and total additional demand is very close to zero, meaning that all additional demand has been allocated. At this point, the integer counts representing the land use fractions on each cell are converted back to fractional representation. The new fractions are used to calculate neighbourhood covariates in the next time step."

First, future land use supply must meet additional demand. Projections of land use demands may be provided through external models, such as Computational General Equilibrium (CGE) models (i.e. GTAP, Aguiar et al. (2016)), or through the analysis and extrapolation of historic patterns (Moulds et al., 2015). The model allocates additional demand by adding cell-level supply of that time step $d_{i,k,t+1}$ in cell i, land use k and time step $t+1$ to the fractions of the current time step $q_{i,k,t}$ (Fig. 2b). The first model objective can be formulated:

$$\sum_{i=1}^{N} q_{i,k,t+1} = \sum_{i=1}^{N}(q_{i,k,t} + d_{i,k,t+1})$$
$$\sum_{i=1}^{N} d_{i,k,t+1} = D_{k,t+1}$$

$D_{k,t+1}$ is the additional landscape-wide supply and is at equilibrium with additional demand after the algorithm converges.

Second, supply of all land-use types in a cell $d_{i,k,t+1}$ is allocated across cells so it adds up to one ($\sum_{k=1}^{K} q_{i,k,t+1} = 1$) (Fig. 2b).

Third, cell-level supply $d_{i,k,t+1}$ has to be distributed in such a way that the allocated amounts in each cell scale with a predicted probability surface s, by modelling $q_{i,k,t=0} \approx s_{i,k} = f^k(\mathbf{X_i})$, where X_i is a set of demographic and bio-physical drivers related to land use. f^k is a multinomial, multi-response model (Fig. 2a). The parameter estimation of this model is based on the first time step and predicted to the conditions of subsequent time steps. Accordingly, while the model assumes stationarity of the modelled statistical relationships, it implements temporal dynamics based on changing demand and changing environmental conditions. Changing environmental conditions are represented as changes to independent model variables.

The land use status in a cell's neighbourhood has been shown to play an important role in determining a cell's land use (Dendoncker et al., 2007; Mustafa et al., 2018; van Vliet et al., 2013; Verburg et al., 2004a). Our suitability model applies neighbourhood interactions by calculating autocovariates (Verburg et al., 2004a) and including these in the multinomial regression of the land use suitability model. Following Verburg et al. (2004a), our autocovariates measure the amount of clustering of land uses in a user-defined cell neighbourhood when compared to the entire landscape. We calculate autocovariates as enrichment factors $F_{d,i,k,t} = \frac{\sum_{i \in d}(q_{i,k,t})/N_d}{\sum_{i=1}^{N}(q_{i,k,t})/N}$. The numerator is the average fraction of land use k in the neighbourhood d of each central cell i and the denominator is the average fraction of land use k in the entire landscape N. Here, we only included neighbourhood characteristics in the $3x3$ neighbourhood around each central cell, but other neighbourhoods are possible (Verburg et al., 2004a). When predicting suitability at each time step, the autocovariates are recalculated based on the assigned fractions from the previous timestep.

Our response variable is a fractional land use value, not discrete classes normally required in multinomial regression. Therefore, we assume that underlying the land use fractions for each cell is a vector of counts $c_{i,k,t}$ that sums to a total number of counts C in each cell (e.g. $C = 1e6$). We derive these counts through $c_{i,k,t} \approx q_{i,k,t} * C$. In integer representation, the data are approximately proportional to the original fractions. When fitting the suitability model, parameter uncertainty depends on the assumption of C. C should be chosen to be small enough for fast model convergence and large enough to represent the degree of numerical precision in the observed fractions. For example, if there are only 2 decimal places, setting $C = 100$ results in counts that represent all of the information contained in the original fractions. Accordingly, the multinomial logit model takes the form

$$s_{i,k,t} = P(Y_i = k) = \frac{e^{\beta_k * X_{i,t} + \gamma_{d,k} * F_{d,i,k,t}}}{\sum_{k=1}^{k} e^{\beta_k * X_{i,t} + \gamma_{d,k} * F_{d,i,k,t}}}$$

where k is the reference land use class, β_k the estimated parameters in each class for covariates $X_{i,t}$ and $\gamma_{d,k}$ the estimated parameters for autocovariates $F_{d,i,k,t}$. We estimated parameters using R's 'nnet' package (Venables and Ripley, 2002). Predicted fractions satisfy $\sum_{k=1}^{K} s_{i,k,t} = 1$.

All software development and model validation was conducted in R (version 4.0.1) (R

Development Core Team, 2008).

2.2 Data

We developed and tested FLUTES using land use and environmental data from the Amazon basin. We downloaded 7 time steps (1992, 1997, 2003, 2008, 2013, 2015 and 2018) of the global land cover map provided through the European Space Agency's Climate Change Initiative Land Cover (CCI-LC) project (European Space Agency, 2019). These data are available at a grid resolution of 300m. We combined the recorded 31 land cover classes to 9 new classes of land use we deemed crucial to identify processes leading to agricultural expansion and declines in habitat (Table 2). We aggregated the resolution 10km^2 squares, calculating fractions of land use from the cell counts of each land use class on the original map present in each new cell. Fractional land use in K classes is mapped over N raster cells, with fractions $q_{i,k,t}$ in cell i in each land use class k always satisfying $0 \leqslant q_{i,k,t} \leqslant 1$ and $\sum_{k=1}^{K} q_{i,k,t} = 1$.

Table 2: Mapping of original land use classes to new classes applied in this study

	New class	Abbr.	CCI-LC class	Description
1	Cropland	Cro	10, 11, 12, 20, 30	Rainfed and irrigated cropland, mosaic cropland with >50% cropland and natural vegetation (tree, shrub, grass)
2	Cropland mosaic	CrM	40	Mosaic cropland with <50% cropland and natural vegetation (tree, shrub, grass)
3	Forest	For	50, 60-62, 70, 80, 90, 100, 160, 170	Forest, closed to open, with >15% canopy cover, Mosaic tree/shrub (>50%) / herbacious cover, Flooded tree cover
4	Grassland	Gra	110, 130	Grassland and mosaic herbacious cover (>50%) / tree/shrub
5	Shrubland	Shr	180	Closed to open and open shrubland
6	Wetland	Wet	190	Flooded shrub or herbacious cover
7	Urban	Urb	120	Settlement, Urban land uses
8	Other	Oth	140, 150, 151-153, 200-202, 220	Lichen/mosses, sparse trees/shrubs/herbaceous vegetation, bare areas, snow/ice
9	Inland water	Wat	210	Natural and artificial inland water bodies

We downloaded a set of spatially explicit climate, topographic soil and human covariates

(Table 3 for a full list of covariates), derived neighbourhood covariates from observed land use in the first time step and estimated observed demand change by calculating the landscape-wide mean fraction for each land use class in each observed time step. All explanatory covariates were standardized to have mean 0 and standard deviation 1. We removed covariates from correlated pairs (Spearman's rank correlation coefficient > 0.7), always retaining the covariate with the smaller average correlation with all other covariates in order to maximise the amount of independent information in the final data set used for fitting.

2.3 Model constraints

Analysing time series data, we determined that only very small percentages of cells change from being devoid of a particular land use to containing that land use within one time step (Table 4). To control unrealistic dispersal of land uses into areas where they have not previously existed, we added a user-defined constraint that land use increases are more likely to be applied to cells where the land use is already present. The constraint parameter was the percentage of cells in which a non-existent land use was newly established between time steps. For example, setting the constraint to 100% would allow increases of a land use in all cells that did not contain that land use in the previous time step.

We parametrized the constraint by determining on how many cells (expressed as a percentage) we could observe the new establishment of a land use from one time step to the next (Table 4). To account for annual variation, we calculated the mean of these percentages for each land use throughout the entire observed time series. For example, throughout the simulation, we allowed *Cro* increases in 1.35% of the cells in which *Cro* was not present in the preceding time step (Table 4). We selected those cells for new

Table 3: List of covariates that were included in land use suitability model

Type	Covariate name	Source
climate	Annual mean temperature	Fick and Hijmans (2017)
	Mean diurnal range	
	Isothermality	
	Temperature seasonality	
	Max. temperature of warmest month	
	Min. temperature of coldest month	
	Temperature annual range	
	Mean temperature of wettest quarter	
	Mean temperature of driest quarter	
	Mean temperature of warmest quarter	
	Mean temperature of coldest quarter	
	Annual precipitation	
	Precipitation of wettest week	
	Precipitation of driest week	
	Precipitation of driest month	
	Precipitation of wettest quarter	
	Precipitation of driest quarter	
	Precipitation of warmest quarter	
	Precipitation of coldest quarter	
topographic	Roughness	Hijmans et al. (2005)
	Slope	
	Elevation	
	Distance to coast	Wessel and Smith (1996)
	Distance to lake	
soil	Nitrogen Content	Global Soil Data Task Group (2000)
	Available Water Content	
	Carbon Density	
	Bulk Density	
human	Distance to built-up areas	FAO (1997)
	Distance to highways	CIESIN (2013)
	Distance to private roads	
	Distance to trails	
	Protected areas	IUCN and UNEP-WCMC (2014)

establishment of a land use that had the highest predicted suitability for that land use (see Appendix B for more information on this constraint).

We masked category I and II protected areas established up until 1992 from land use changes as has been shown previously (see Fig. 3 for a map of protected areas) (Verburg et al., 2002; IUCN and UNEP-WCMC, 2014; Kapitza et al., 2021). To reflect the high initial investment of urban infrastructure, we did not allow reductions in urban land (Verburg and Overmars, 2009).

Table 4: Share of cells (%) containing a land use that were completely devoid of that land use in the preceding time step. Values derived from observed time series.

Land use	1996	2001	2006	2011	2016	2018	mean
Cro	1.75	1.66	4.49	0.08	0.06	0.07	1.35
CrM	2.39	2.37	7.24	0.05	0.03	0.05	2.02
For	0	0	0	0	0	0	0
Gra	0.40	0.62	0.94	0.15	0.04	0.04	0.37
Shr	0.62	0.90	1.44	0.15	0.07	0.06	0.54
Wet	0.62	0.68	2.60	0.26	0.13	0.11	0.73
Urb	0.36	0.61	1.12	0.16	0.28	0.02	0.43
Oth	0.02	0.06	0.12	0.05	0.02	0.01	0.05
Wat	0.81	0.35	1.19	0.02	0.01	0.01	0.40

2.4 Validating the intensity and direction of predicted changes

First, we examined the accuracy of the multinomial suitability model and how it is affected by spatial resolution and the included covariates. To account for spatial autocorrelation in the environmental covariates and land use time series, we conducted spatial-blocks cross-validation (Valavi et al., 2019) by separating the landscape into 9 equal-sized spatial blocks. We fitted models using data from 8 of the 9 blocks and predicted the model to the withheld block, until predictions were made for the entire study area. We cross-validated suitability models at 1km and 10km, including 1) only environmental covariates, 2) only neighbourhood covariates and 3) both covariate types combined. For each of the three models we measured predictive performance by estimating cell-level suitability Root Mean Squared Error ($\text{RMSE}_{\text{suit}}$) between the predicted suitability surfaces $s_{m,i,k,t}$ and the observed fractions $o_{i,k,t}$, following $RMSE_{suit,m,i,t} = \sqrt{\frac{1}{K}\sum_{k=1}^{K}(o_{i,k,t} - s_{m,i,k,t})^2}$ for each suitability model m.

Second, to validate the intensity of changes predicted by the allocation algorithm, we assessed the accuracy of predictions of cell-level fractions under competing models predicted throughout the observed time series. 1) Under the **null model**, we assumed no change of land use through time. The null model served as reference to measure the improvements provided by each additional model component. 2) Under the **naive model** we only allocated additional demands, but scaled cell-level allocations with the average supply observed across the entire landscape. This model assumes that suitability is not informative about where a change will happen and that allocations are equally likely to be anywhere in the landscape. 3) Under the **semi-naive model**, cell-level allocations were additionally scaled with the predicted suitability surfaces $s_{i,k,t}$ (as illustrated in Fig. 2). 4) Under the **full model**, allocations were scaled with suitability surfaces $s_{i,k,t}$ and all constraints (constraining most increases to cells where land use type already exists and masking protected areas from changes) were applied.

We calculated $\text{RMSE}_{\text{alloc}}$ under each allocation model w to estimate how well the different model components simulated each cell-level vector of land use fractions $q_{m,i,k,t}$ compared to the respective observed vectors $o_{i,k,t}$, following $RMSE_{alloc,w,i,t} = \sqrt{\frac{1}{K}\sum_{k=1}^{K}(o_{i,k,t} - q_{w,i,k,t})^2}$.

Due to the squared term, RMSE cannot inform on whether the models correctly identified the direction of change. Therefore, we estimated and validated the direction of cell-level changes (decreases, no change, increases) separately. We mapped these transitions for each class between the time steps of the observed time series and the time steps of the time series simulated under each model. We calculated *overall difference* of each pair of corresponding maps to obtain an interpretable measure of similarity of predicted and observed direction of changes (Pontius and Millones, 2011; Pontius and Santacruz, 2014). Achieving high accuracy in these first two model goals would suggest that simulated patterns of land use change closely resemble observed patterns.

2.5 Case study: agricultural expansion in the Amazon Basin

The Amazon catchment is largest river basin in the world and occupies over one third of the South American land mass (Fig. 3a). As the world's most diverse tropical forest area, the basin hosts at least 10% of the world's known species (Da Silva et al., 2005).

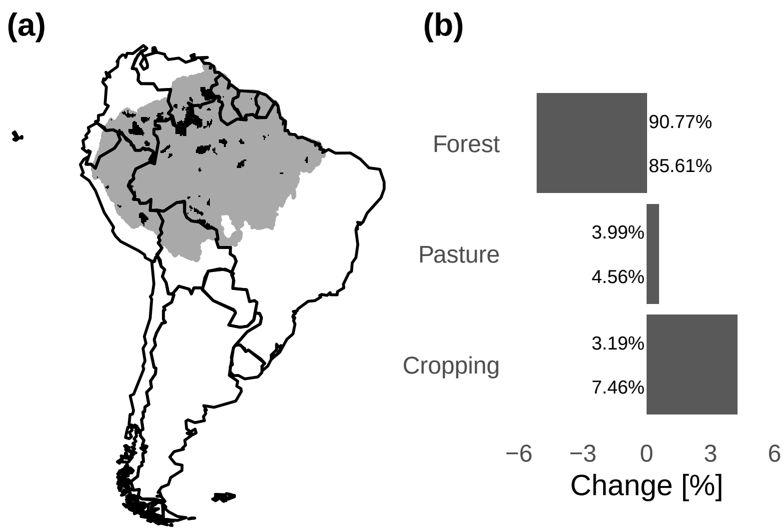

Figure 3: Overview of the study area. a) Location of the amazon catchment in South America (grey-shaded area), including IUCN protected areas (categories I and II) which were used to constrain land use changes (black shaded areas). b) Changes in selected land uses, derived from observed land use maps. Pasture includes Gra and Shr, Cropping includes Cro and CrM and forest includes For. Beside the bars are percentage cover in 1992 (top) and percentage cover in 2018 (bottom). Land use classes are specified in Table 2.

The Amazon biome is threatened by a multitude of interacting factors. Ecosystem services, such as water supply, carbon storage and provision of species habitat are directly threatened by the effects of climate change and the increasing pressure on land, with projected severe reductions in water yields, carbon content and species habitat, which is particularly affected by changes in natural vegetation cover (Prüssmann et al., 2016). The primary uses for cleared forest land are pasture for cattle farming and industrial soy cropping (Nepstad et al., 2014; FAO, 2015). Between 1992 and 2018, the biome has seen significant increases in land required for cropping and pasture, as well as significant decreases in forest cover (Fig. 3b).

Using a broad reclassification of the predicted and observed land use classes into cropland, pasture and habitat, we were able to specifically validate FLUTES's ability to predict agricultural expansion and habitat declines as aggregated threats to ecosystems and biodiversity. First, we determined areas of agricultural (pasture or cropland) expansion with simultaneous declines in classes containing natural habitats (*For*, *Wet* and *Oth*). We categorized the observed and predicted maps into 1) areas with no cropland increase, 2) areas where cropland increase led to mostly forest declines (net replacement of forest), and 3) areas where cropland increase led to mostly declines in other natural habitat classes (net replacement of other habitat). Similarly, we categorized the landscape into 1) areas with no pasture increase, 2) areas where pasture increase led to mostly forest declines, and 3) areas where pasture increase led to mostly declines in other natural habitat classes. From the resulting reclassified time series we assessed the difference between the respective observed and predicted maps by overlaying them and identifying where no agricultural increase was observed and predicted (persistence predicted as persistence), where agricultural increase was correctly predicted and led to decreases in the correct habitat class, where agricultural increase was correctly predicted but resulted in decreases in the incorrect habitat class, where no agricultural increase was observed, but agricultural increase was predicted, and where agricultural increase was observed, but not predicted (Pontius et al., 2011).

3 Results

3.1 Predicting land use change intensity

Results of the cross-validation of the suitability model component show that including neighbourhood covariates resulted in substantial predictive performance improvements across spatial blocks at both resolutions (Fig. 4c, Fig. S1 for predicted suitability maps of all 9 land use classes); models using neighbourhood covariates alone were approximately as good as the model using the full covariate set. Including only environmental variables resulted in less accurate predictions at both resolutions, with predictions under the fine resolution comparatively worse than under the coarse resolution.

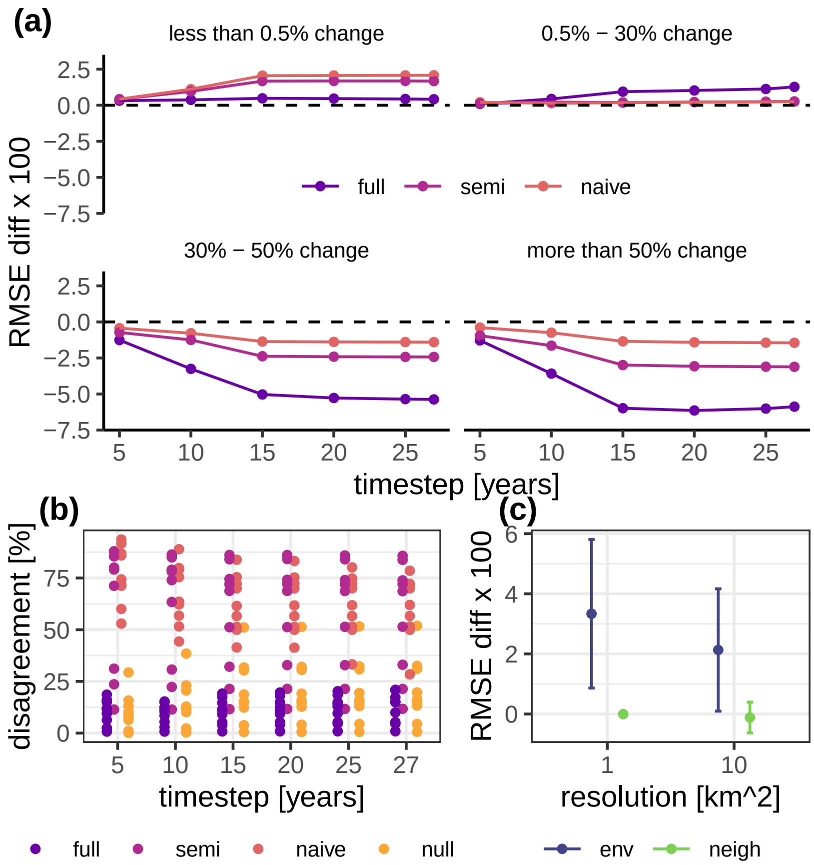

Figure 4: Validation of predicted land use change intensity and direction of change and cross-validation of suitability model. a) The difference between RMSE for each model (naive, semi-naive, full) and RMSE of the null model. The null model assumes that land use is static through time, the naive model assumes completely random allocations, the semi-naive model assumes that allocations are scaled with land use suitability and the full model assumes that allocations are both scaled with land use suitability and subject to model constraints (no changes in areas under high protection status and no land use increases in areas completely devoid of that land use). All RMSE were calculated at cell-level, using the predicted and observed vectors of land use fractions in each cell. Plotted are means across cells. Positive values indicate better fits under the null model, negative values indicate better fit under more highly parametrised models. Data on validation outcomes are grouped by the magnitude of the largest observed proportional change in any land use within a cell. In general, the larger the observed change in land use, the better the parameterized models did compared with the null model. b) The proportional disagreement between predictions of the cell-level direction of change (no change, decrease, increase) for each land use and the observed direction of change at each time step. Smaller values indicate lower overall difference and higher similarity between corresponding maps. c) Difference between cross-validated RMSE estimated for suitability models containing only environmental covariates and only neighbourhood covariates and models containing both covariate types combined. Positive values indicate a poorer fit than the model containing both covariate types.

Under all tested models (naive, semi-naive, full), the accuracy of cell-level allocations improved with the intensity of observed changes (Fig. 4a). This implies that FLUTES makes good predictions under scenarios with high expected overall changes.

Where observed changes were large (Fig. 4a, bottom two panels), including land use suitability and constraints (full model) resulted in substantial increases of predictive performance. In these areas, the null model's assumption of no spatial variation in reallocation of land use introduced very high bias, which our constraints were able to reduce.

When observed changes were small (Fig. 4a, top two panels), the null model made near perfect predictions. Given how close the null model already was to the truth, improvements by allocating demand (naive model) and accounting for land use suitability (semi-naive model) were difficult to achieve; in the smallest change category (Fig. 4a, top left panel), the naive and semi-naive predictions were in fact slightly worse than the null. In these areas the largest observed changes were below 0.5%, making the assumption of no change under the null model highly plausible. Under the full model, the applied constraint limited the areas that could be flagged for increases. Accordingly, where observed changes were small, this model made better predictions than the semi-naive and naive models, in which this constraint was not applied.

3.2 Predicting the direction of land use changes

The worst predictions of cell-level direction of change were made by the naive and semi-naive models and the best predictions under the full model (Fig. 4b), with overall difference consistently less than 25%. Predictions became more accurate the more model components were applied. Under the full model we achieved the highest prediction accuracy. Overall, the semi-naive model performed slightly better than the naive model, demonstrating the utility of scaling allocations with land use suitability surfaces.

3.3 Predicting agricultural expansion and habitat declines

FLUTES achieved high accuracy when predicting agricultural (cropland and pasture) expansion on forest and other land use types containing natural habitats (Fig. 5). In more than 80 % of cells FLUTES predicted correctly whether agricultural land (cropland or pasture) increased, or persisted at current levels or decreased, and which habitat type decreased due to increases in agricultural classes (Fig. 5b, d).

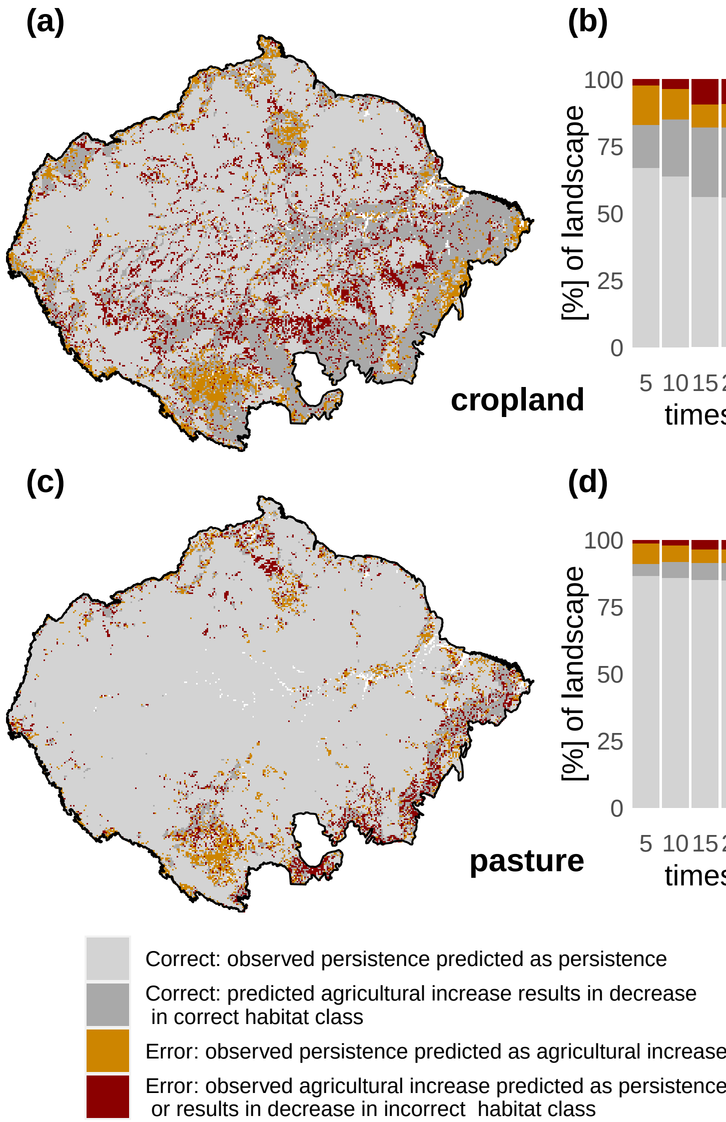

Figure 5: Validation of modelling agricultural expansion in cropland and pasture on forest and other natural habitat types in the Amazon basin. (a, b) Spatial configuration of correct and erroneous predictions of cropland in the last time step of the validation time series (2018) (a) and the relative size of the landscape where predictions matched observations (correct) and where predictions deviated from observations (error) in cropland (b). (c, d) Spatial configuration of correct and erroneous predictions of pasture in the last time step of the validation time series (2018) (c) and the relative size of the landscape where predictions matched observations (correct) and where predictions deviated from observations (error) in pasture (d).

The percentage of the landscape in which we correctly predicted cropland increase at the expense of the correct habitat class increased through the time series, suggesting that

FLUTES was good at identifying not only where cropland did not change or decreased (persistence), but also where it increased and on which habitat type that increase took place (Fig. 5b). FLUTES made some incorrect predictions of cropland increase in areas where no increase was observed in the southern tip and the central north of the study area, although these areas were very small compared to surrounding areas in which increases were correctly predicted and occurred on the appropriate habitat classes (Fig. 5a). Perhaps the most severe type of error in terms of ecological considerations was the prediction of no cropland increase (persistence) in areas where increase was observed. However, these areas were small (increasing from 2.3% of the landscape at the beginning to 9.6% at the end of the time series).

Pasture expansion (Fig. 5c, d) was much smaller than cropland expansion overall, with much larger areas of the basin correctly predicted as not increasing in pasture land (persistence) through time (Fig. 5d). Some small areas in the southern tip, the central north and along the western boundary of the basin were correctly predicted to increase in pasture land, with decreases in the appropriate natural habitat class. Pasture expansion was underpredicted in very small areas in the south, north and along the eastern boundary of the basin (Fig. 5c).

4 Discussion and conclusions

We have presented a new fractional land use change allocation model to predict land use fractions, thus retaining information at sub-pixel level. The model is able to accurately allocate fractions of land use through time, especially under scenarios of more extreme land use change. We explicitly accounted for competition between land use types and land use suitability in response to environmental drivers by means of a multinomial logistic model and could show that this aspect brings substantial improvements to predictions, when compared to the assumption that land use does not change at all (null model).

FLUTES made accurate predictions in areas in which only small land use changes were observable, but also in areas where land use changes were observed to be high. This suggests that FLUTES provides a suitable method to produce future land use maps under contrasting scenario settings. In scenarios where demand changes are expected to be high, FLUTES allocates supply to match aggregated demand, changing the total area allocated to different land uses and also allowing land uses to be established in new areas. In scenarios with small expected demand changes, land use changes, including the establishment of land uses in new areas, remain small.

We assumed that the initial land use distribution we used to calibrate FLUTES resulted from long time periods of optimizing behaviour and we have not yet implemented a parameter allowing to specify land use elasticity (the propensity of land uses to shift across the landscape without net changes to their total areas at the study area level), as is implemented, for example, in Dyna-CLUE (Verburg and Overmars, 2009). For this reason, in FLUTES land use cannot change to match predicted land use suitability alone. For example, if the modelled cropland suitability in an area is 0.8, but the observed cropland fraction is 0.2, FLUTES would only allow a local increase in cropland if the aggregated demand for cropland at the study area level increased. While the discrepancy between an area's potential for a certain land use measured by the predicted suitability and the realized fraction of that land use implicitly captures processes that cannot be captured by the suitability model, in this first version of FLUTES this only occurs when triggered by changes in external demand for that land use.

Similar to CLUE, our constraint on turn-over allowed us to account for conversion effort. Here, data from the observed validation time series allowed us to extract a raw estimate of the constraint parameter to tune FLUTES. We estimated the parameter using long-term observed means. We assume this to be similarly informative as informal expert knowledge, which has been suggested as a primary means to parametrize land use conversion effort in previous land use models (Van Asselen and Verburg, 2013; Overmars et al., 2007).

We could show that FLUTES is very easily adaptable to specific ecological study contexts. When validating our model's performance in the context of agricultural expansion on natural habitat, we mapped the model's ability to reproduce where agricultural expansion occurs in both pasture and cropland and which natural habitat classes decreased in their place. Consistently more than 80% of the landscape where correctly classified as no change or a decrease (persistence) or increase of agricultural land with decreases in the correct habitat types. Crucially, underprediction of agricultural expansion with possible negative implications for conservation management remained very small throughout. This demonstrates that FLUTES is a useful tool to predict the spatial configuration of land use change impacts that are driven by agricultural expansion into different habitat types.

Validating the suitability model component of our model approach, we found that neighbourhood covariates explained much of the suitability patterns across the landscape. This is a common effect of including flexible spatial correlation terms in models with other spatially-varying covariates (spatial confounding) (Hodges and Reich, 2010). The models describe the spatial pattern with the spatial correlation term, but this effect does not imply causation and other drivers included in the model may still drive changes in the response, particularly over long time periods. Here, similar to what was shown by Dendoncker et al. (2007), including neighbourhood covariates lead to the most highly fitted models. Allowing spatial autocorrelation to drive patterns seems a sensible choice for predictions in this case study because the model only predicts three decades. However, for longer time spans, spatial autocorrelation probably becomes less important and continental-scale environmental driving factors acting homogeneously across the whole landscape may dominate patterns in reality. When making such longer-term predictions, this could be captured by fitting the suitability model with several time steps of data, thus ensuring that land use suitability is less reliant on the present land use state, but more weight is given to long-term and large-scale environmental processes.

The results of our validation also strongly indicate that in the case of FLUTES, adding constraints (decision rules) in terms of where and how land use changes are allowed to occur, are responsible for the majority of increases in predictive performance. While we provide initial steps in parametrising these constraints, more specific knowledge of bottom-up processes that drive land use stasis and change across the landscape could further consolidate the accuracy of FLUTES. For example, this could be achieved by including data on the expected behaviour of economic agents who seek to maximise returns on their productive land. One example includes the Land Use Trade-offs (LUTO) model (Bryan et al., 2014; Connor et al., 2015), which includes pixel-wise optimisation of cost and return of alternative land uses. However, such models are difficult to parametrise in data-scarce regions and require significant computational power. Bottom-up processes, such as price feedbacks, also tend to act at very fine spatial resolutions, but have little effect when seen at a continental scale, where scenario uncertainty and global processes dominate predictions (Connor et al., 2015). Depending on scale, including very fine-scale dynamics of agent behaviour may simply not pay off, or it might be more appropriate to merely downscale them to the study area extent (Van Asselen and Verburg, 2013; Connor et al., 2015).

In order to allow scaling FLUTES to global applications, we only used drivers that were available at global scales. However, improvements to the land use suitability model can be achieved by including more proximate drivers of land use change, such as market accessibility (Meiyappan et al., 2014; Verburg et al., 2011), by fitting the land use suitability model for individual subsets of the study area to improve local fit, or by creating more land use classes for which particular biophysical constraints are known. Including location-dependent drivers and models and raising the resolution may substantially improve the accuracy of land use suitability maps, increasing the contribution of this model component to overall prediction accuracy.

Developments of FLUTES and expanding application could include the estimation of use intensity of different land use types, which has been shown to be an important driver of biodiversity change (Newbold et al., 2015, 2016). Such developments could enhance efforts to tailor macroeconomic and land use modelling to assess the fate of future biodiversity (Kapitza et al., 2021).

5 Acknowledgements

We are grateful for contributions made throughout the research phase by J. Elith and D. Zurell and helpful comments by Robert G. Pontius Jr on a preprint of this manuscript. Funding: this work received funding under the Australian Research Council Discovery grant DP170104795. NG was supported by an ARC DECRA fellowship (DE180100635). SK was supported by the Melbourne International Research Scholarship (MIRS).

References

Aburas, M. M., Ho, Y. M., Ramli, M. F., and Ash'aari, Z. H. (2017). Improving the capability of an integrated CA-Markov model to simulate spatio-temporal urban growth trends using an Analytical Hierarchy Process and Frequency Ratio. *International Journal of Applied Earth Observation and Geoinformation*, 59:65–78.

Aguiar, A., Narayanan, B., and McDougall, R. (2016). An Overview of the GTAP 9 Data Base. *Journal of Global Economic Analysis*, 1(1):181–208.

Allred, B. W., Bestelmeyer, B. T., Boyd, C. S., Brown, C., Davies, K. W., Duniway, M. C., Ellsworth, L. M., Erickson, T. A., Fuhlendorf, S. D., Griffiths, T. V., Jansen, V., Jones, M. O., Karl, J., Knight, A., Maestas, J. D., Maynard, J. J., McCord, S. E., Naugle, D. E., Starns, H. D., Twidwell, D., and Uden, D. R. (2021). Improving Landsat predictions of rangeland fractional cover with multitask learning and uncertainty. *Methods in Ecology and Evolution*, 12(5):841–849.

Bevanda, M., Horning, N., Reineking, B., Heurich, M., Wegmann, M., and Mueller, J. (2014). Adding structure to land cover – using fractional cover to study animal habitat use. *Movement Ecology*, 2(1):26.

Bossard, M., Feranec, J., and Otahel, J. (2000). CORINE land cover technical guide: Addendum 2000. Technical Report Technical report No 40, European Environment Agency.

Bryan, B. A., Nolan, M., Harwood, T. D., Connor, J. D., Navarro-Garcia, J., King, D., Summers, D. M., Newth, D., Cai, Y., Grigg, N., Harman, I., Crossman, N. D., Grundy, M. J., Finnigan, J. J., Ferrier, S., Williams, K. J., Wilson, K. A., Law, E. A., and Hatfield-Dodds, S. (2014). Supply of carbon sequestration and biodiversity services from Australia's agricultural land under global change. *Global Environmental Change*, 28(1):166–181.

CIESN (2013). Center for International Earth Science Information Network - CIESIN -

Columbia University, Information Technology Outreach Services - ITOS - University of Georgia, Global Roads Open Access Data Set, Version 1 (gROADSv1).

Connor, J. D., Bryan, B. A., Nolan, M., Stock, F., Gao, L., Dunstall, S., Graham, P., Ernst, A., Newth, D., Grundy, M., and Hatfield-Dodds, S. (2015). Modelling Australian land use competition and ecosystem services with food price feedbacks at high spatial resolution. *Environmental Modelling and Software*, 69:141–154.

Da Silva, J. M. C., Rylands, A. B., and da Fonseca, G. A. B. (2005). The Fate of the Amazonian Areas of Endemism. *Conservation Biology*, 19(3):689–694.

Dendoncker, N., Bogaert, P., and Rounsevell, M. (2006). A statistical method to downscale aggregated land use data and scenarios. *Journal of Land Use Science*, 1(2-4):63–82.

Dendoncker, N., Rounsevell, M., and Bogaert, P. (2007). Spatial analysis and modelling of land use distributions in Belgium. *Computers, Environment and Urban Systems*, 31(2):188–205.

Dietzel, C. and Clarke, K. C. (2007). Toward Optimal Calibration of the SLEUTH Land Use Change Model. *Transactions in GIS*, 11(1):29–45.

Eastman, J. R. and Toledano, J. (2018). A short presentation of the Land Change Modeler (LCM). In *Geomatic Approaches for Modeling Land Change Scenarios*, pages 499–505. Springer, Cham.

European Space Agency (2019). Land Cover CCI Version 2.0 and Version 2.1.1. https://www.esa-landcover-cci.org/?q=node/164.

European Union (2019). Copernicus Land Monitoring Service. European Enviornment Agency (EEA). https://land.copernicus.eu.

Fang, S., Gertner, G. Z., Sun, Z., and Anderson, A. A. (2005). The impact of interactions in spatial simulation of the dynamics of urban sprawl. *Landscape and Urban Planning*, 73(4):294–306.

FAO (1997). Built-up Areas of the World (Vmap0). Food and Agriculture Organization of the United Nations (FAO).

FAO (2015). AQUASTAT Transboundary River Basin Overview – Amazon. Food and Agriculture Organization of the United Nations (FAO).

Fick, S. and Hijmans, R. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12):4302–4315.

Foley, J. A. (2005). Global Consequences of Land Use. *Science*, 309(5734):570–574.

Fuchs, R., Herold, M., Verburg, P. H., and Clevers, J. G. (2013). A high-resolution and harmonized model approach for reconstructing and analysing historic land changes in Europe. *Biogeosciences*, 10(3):1543–1559.

Global Soil Data Task Group (2000). Global gridded surfaces of selected soil characteristics (IGBP-DIS). https://doi.org/10.3334/ornldaac/569.

Harrison, P. A., Holman, I. P., Cojocaru, G., Kok, K., Kontogianni, A., Metzger, M. J., and Gramberger, M. (2013). Combining qualitative and quantitative understanding for exploring cross-sectoral climate change impacts, adaptation and vulnerability in Europe. *Regional Environmental Change*, 13(4):761–780.

Hasegawa, T., Fujimori, S., Ito, A., Takahashi, K., and Masui, T. (2017). Global land-use allocation model linked to an integrated assessment model. *Science of The Total Environment*, 580:787–796.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15):1965–1978.

Hill, M. J. and Guerschman, J. P. (2020). The MODIS Global Vegetation Fractional Cover Product 2001–2018: Characteristics of Vegetation Fractional Cover in Grasslands and Savanna Woodlands. *Remote Sensing*, 12(3):406.

Hodges, J. S. and Reich, B. J. (2010). Adding Spatially-Correlated Errors Can Mess Up the Fixed Effect You Love. *The American Statistician*, 64(4):325–334.

Hyandye, C. and Martz, L. W. (2017). A Markovian and cellular automata land-use change predictive model of the Usangu Catchment. *International Journal of Remote Sensing*, 38(1):64–81.

IPBES (2019). Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. S. Díaz, J. Settele, E. S. Brondízio E.S., H. T. Ngo, M. Guèze, J. Agard, A. Arneth, P. Balvanera, K. A. Brauman, S. H. M. Butchart, K. M. A. Chan, L. A. Garibaldi, K. Ichii, J. Liu, S. M. Subramanian, G. F. Midgley, P. Miloslavich, Z. Molnár, D. Obura, A. Pfaff, S. Polasky, A. Purvis, J. Razzaque, B. Reyers, R. Roy Chowdhury, Y. J. Shin, I. J. Visseren-Hamakers, K. J. Willis, and C. N. Zayas (eds.). IPBES secretariat, Bonn. 56 pages.

IUCN and UNEP-WCMC (2014). The World Database on Protected Areas (WDPA). UNEP World Conservation Monitoring Centre. Cambridge, UK. https://www.protectedplanet.net.

Kapitza, S., Van Ha, P., Kompas, T., Golding, N., Cadenhead, N. C. R., Bal, P., and

Wintle, B. A. (2021). Assessing biophysical and socio-economic impacts of climate change on regional avian biodiversity. *Scientific Reports*, 11(3304).

Kim, H., Rosa, I. M. D., Alkemade, R., Leadley, P., Hurtt, G., Popp, A., van Vuuren, D. P., Anthoni, P., Arneth, A., Baisero, D., Caton, E., Chaplin-Kramer, R., Chini, L., De Palma, A., Di Fulvio, F., Di Marco, M., Espinoza, F., Ferrier, S., Fujimori, S., Gonzalez, R. E., Gueguen, M., Guerra, C., Harfoot, M., Harwood, T. D., Hasegawa, T., Haverd, V., Havlík, P., Hellweg, S., Hill, S. L. L., Hirata, A., Hoskins, A. J., Janse, J. H., Jetz, W., Johnson, J. A., Krause, A., Leclère, D., Martins, I. S., Matsui, T., Merow, C., Obersteiner, M., Ohashi, H., Poulter, B., Purvis, A., Quesada, B., Rondinini, C., Schipper, A. M., Sharp, R., Takahashi, K., Thuiller, W., Titeux, N., Visconti, P., Ware, C., Wolf, F., and Pereira, H. M. (2018). A protocol for an intercomparison of biodiversity and ecosystem services models using harmonized land-use and climate scenarios. *Geoscientific Model Development*, 11(11):4537–4562.

Lambin, E., Rounsevell, M., and Geist, H. (2000). Are agricultural land-use models able to predict changes in land-use intensity? *Agriculture, Ecosystems & Environment*, 82(1-3):321–331.

Lambin, E. F., Meyfroidt, P., E. F. Lambin, and P. Meyfroidt (2011). Global land use change, economic globalization, and the looming land scarcity. *Proceedings of the National Academy of Sciences of the United States of America*, 108(9):3465–72.

Levers, C., Butsic, V., Verburg, P. H., Müller, D., and Kuemmerle, T. (2016). Drivers of changes in agricultural intensity in Europe. *Land Use Policy*, 58:380–393.

Levers, C., Verkerk, P. J., Müller, D., Verburg, P. H., Butsic, V., Leitão, P. J., Lindner, M., and Kuemmerle, T. (2014). Drivers of forest harvesting intensity patterns in Europe. *Forest Ecology and Management*, 315:160–172.

Liu, X., Liang, X., Li, X., Xu, X., Ou, J., Chen, Y., Li, S., Wang, S., and Pei, F. (2017). A future land use simulation model (FLUS) for simulating multiple land use scenarios by coupling human and natural effects. *Landscape and Urban Planning*, 168:94–116.

Margono, B. A., Potapov, P. V., Turubanova, S., Stolle, F., and Hansen, M. C. (2014). Primary forest cover loss in Indonesia over 2000–2012. *Nature Climate Change*, 4(8):730–735.

Meiyappan, P., Dalton, M., O'Neill, B. C., and Jain, A. K. (2014). Spatial modeling of agricultural land use change at global scale. *Ecological Modelling*, 291:152–174.

Mouchet, M., Levers, C., Zupan, L., Kuemmerle, T., Plutzar, C., Erb, K., Lavorel, S., Thuiller, W., and Haberl, H. (2015). Testing the Effectiveness of Environmental Variables to Explain European Terrestrial Vertebrate Species Richness across Biogeographical Scales. *PLOS ONE*, 10(7):e0131924.

Moulds, S., Buytaert, W., and Mijic, A. (2015). An open and extensible framework for spatially explicit land use change modelling: The lulcc R package. *Geoscientific Model Development*, 8(10):3215–3229.

Mustafa, A., Heppenstall, A., Omrani, H., Saadi, I., Cools, M., and Teller, J. (2018). Modelling built-up expansion and densification with multinomial logistic regression, cellular automata and genetic algorithm. *Computers, Environment and Urban Systems*, 67:147–156.

Nepstad, D., McGrath, D., Stickler, C., Alencar, A., Azevedo, A., Swette, B., Bezerra, T., DiGiano, M., Shimada, J., Seroa da Motta, R., Armijo, E., Castello, L., Brando, P., Hansen, M. C., McGrath-Horn, M., Carvalho, O., and Hess, L. (2014). Slowing Amazon deforestation through public policy and interventions in beef and soy supply chains. *Science*, 344(6188):1118–1123.

Newbold, T., Hudson, L. N., Arnell, A. P., Contu, S., De Palma, A., Ferrier, S., Hill, S. L. L., Hoskins, A. J., Lysenko, I., Phillips, H. R. P., Burton, V. J., Chng, C. W. T., Emerson, S., Gao, D., Pask-Hale, G., Hutton, J., Jung, M., Sanchez-Ortiz, K., Simmons, B. I., Whitmee, S., Zhang, H., Scharlemann, J. P. W., and Purvis, A. (2016). Has land use pushed terrestrial biodiversity beyond the planetary boundary? A global assessment. *Science*, 353(6296):288–291.

Newbold, T., Hudson, L. N., Hill, S. L., Contu, S., Lysenko, I., a Senior, R., Börger, L., Bennett, D. J., Choimes, A., Collen, B., Day, J., De Palma, A., Dıáz, S., Echeverria-Londoño, S., Edgar, M. J., Feldman, A., Garon, M., Harrison, M. L. K., Alhusseini, T., Ingram, D. J., Itescu, Y., Kattge, J., Kemp, V., Kirkpatrick, L., Kleyer, M., Laginha Pinto Correia, D., Martin, C. D., Meiri, S., Novosolov, M., Pan, Y., Phillips, H. R. P., Purves, D. W., Robinson, A., Simpson, J., Tuck, S. L., Weiher, E., White, H. J., Ewers, R. M., Mace, G. M., Scharlemann, J. P., and Purvis, A. (2015). Global effects of land use on local terrestrial biodiversity. *Nature*, 520(7545):45–50.

Noszczyk, T. (2019). A review of approaches to land use changes modeling. *Human and Ecological Risk Assessment: An International Journal*, 25(6):1377–1405.

O'Neill, B. C., Kriegler, E., Ebi, K. L., Kemp-Benedict, E., Riahi, K., Rothman, D. S., van Ruijven, B. J., van Vuuren, D. P., Birkmann, J., Kok, K., Levy, M., and Solecki, W. (2017). The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century. *Global Environmental Change*, 42:169–180.

O'Neill, B. C., Kriegler, E., Riahi, K., Ebi, K. L., Hallegatte, S., Carter, T. R., Mathur, R., and van Vuuren, D. P. (2014). A new scenario framework for climate change research: The concept of shared socioeconomic pathways. *Climatic Change*, 122(3):387–400.

Overmars, K. P., Verburg, P. H., and Veldkamp, T. A. (2007). Comparison of a deductive and an inductive approach to specify land suitability in a spatially explicit land use model. *Land Use Policy*, 24(3):584–599.

Pérez-Vega, A. (2012). Comparing two approaches to land use/cover change modeling and their implications for the assessment of biodiversity loss in a deciduous tropical forest. *Environmental Modelling and Software*, 29(1):11–23.

Pijanowski, B. C., Brown, D. G., Shellito, B. A., and Manik, G. A. (2002). Using neural networks and GIS to forecast land use changes: A Land Transformation Model. *Computers, Environment and Urban Systems*, 26(6):553–575.

Plutzar, C., Kroisleitner, C., Haberl, H., Fetzel, T., Bulgheroni, C., Beringer, T., Hostert, P., Kastner, T., Kuemmerle, T., Lauk, C., Levers, C., Lindner, M., Moser, D., Müller, D., Niedertscheider, M., Paracchini, M. L., Schaphoff, S., Verburg, P. H., Verkerk, P. J., and Erb, K.-H. (2016). Changes in the spatial patterns of human appropriation of net primary production (HANPP) in Europe 1990–2006. *Regional Environmental Change*, 16(5):1225–1238.

Pontius, R. G. and Millones, M. (2011). Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing*, 32(15):4407–4429.

Pontius, R. G., Peethambaram, S., and Castella, J.-C. (2011). Comparison of Three Maps at Multiple Resolutions: A Case Study of Land Change Simulation in Cho Don District, Vietnam. *Annals of the Association of American Geographers*, 101(1):45–62.

Pontius, R. G. and Santacruz, A. (2014). Quantity, exchange, and shift components of difference in a square contingency table. *International Journal of Remote Sensing*, 35(21):7543–7554.

Powers, R. P. and Jetz, W. (2019). Global habitat loss and extinction risk of terrestrial vertebrates under future land-use-change scenarios. *Nature Climate Change*, 9:323–329.

Prüssmann, J., Suárez, C., Guevara, O., and Vergara, A. (2016). Vulnerability and climate risk analysis of the Amazon biome and its protected areas. Amazon Vision, REDPARQUES, WWF, UICN, FAO, PNUMA.

R Development Core Team (2008). R: A language and environment for statistical computing. Foundation for Statistical Computing.

Seo, B., Bogner, C., Koellner, T., and Reineking, B. (2016). Mapping Fractional Land Use and Land Cover in a Monsoon Region: The Effects of Data Processing Options. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(9):3941–3956.

Seto, K. C., Guneralp, B., and Hutyra, L. R. (2012). Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proceedings of the National Academy of Sciences*, 109(40):16083–16088.

Shafizadeh Moghadam, H. and Helbich, M. (2013). Spatiotemporal urbanization processes in the megacity of Mumbai, India: A Markov chains-cellular automata urban growth model. *Applied Geography*, 40:140–149.

Soares-Filho, B., Rodrigues, H., and Costa, W. (2009). Modeling environmental dynamics with Dinamica EGO. Instituto de Geociências - Centro de Sensoriamento Remoto. 114-114.

Struebig, M. J., Fischer, M., Gaveau, D. L. A., Meijaard, E., Wich, S. A., Gonner, C., Sykes, R., Wilting, A., and Kramer-Schadt, S. (2015). Anticipated climate and land-cover changes reveal refuge areas for Borneo's orang-utans. *Global Change Biology*, 21(8):2891–2904.

Sun, H., Forsythe, W., and Waters, N. (2007). Modeling Urban Land Use Change and Urban Sprawl: Calgary, Alberta, Canada. *Networks and Spatial Economics*, 7(4):353–376.

Tayyebi, A. and Pijanowski, B. C. (2014). Modeling multiple land use changes using ANN, CART and MARS: Comparing tradeoffs in goodness of fit and explanatory power of data mining tools. *International Journal of Applied Earth Observation and Geoinformation*, 28(1):102–116.

Titeux, N., Henle, K., Mihoub, J.-B., Regos, A., Geijzendorffer, I. R., Cramer, W., Verburg, P. H., and Brotons, L. (2016). Biodiversity scenarios neglect future land-use changes. *Global Change Biology*, 22(7):2505–2515.

Urban, M. C., Bocedi, G., Hendry, A. P., Mihoub, J.-B., Peer, G., Singer, A., Bridle, J. R., Crozier, L. G., De Meester, L., Godsoe, W., Gonzalez, A., Hellmann, J. J., Holt, R. D., Huth, A., Johst, K., Krug, C. B., Leadley, P. W., Palmer, S. C. F., Pantel, J. H., Schmitz, A., Zollner, P. A., and Travis, J. M. J. (2016). Improving the forecast for biodiversity under climate change. *Science*, 353(6304):1–9.

Valavi, R., Elith, J., Lahoz-Monfort, J. J., and Guillera-Arroita, G. (2019). blockCV: An r package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. *Methods in Ecology and Evolution*, 10(2):225–232.

Van Asselen, S. and Verburg, P. H. (2013). Land cover change or land-use intensification: Simulating land system change with a global-scale land change model. *Global Change Biology*, 19(12):3648–3667.

van Schrojenstein Lantman, J., Verburg, P. H., Bregt, A., and Geertman, S. (2011). Core Principles and Concepts in Land-Use Modelling: A Literature Review. In

Koomen, E. and Borsboom-van Beurden, J., editors, *Land-Use Modelling in Planning Practice*, volume 101, pages 35–57. Springer Netherlands, Dordrecht.

van Vliet, J., Naus, N., van Lammeren, R. J., Bregt, A. K., Hurkens, J., and van Delden, H. (2013). Measuring the neighbourhood effect to calibrate land use models. *Computers, Environment and Urban Systems*, 41:55–64.

Veldkamp, A. and Fresco, L. (1996). CLUE: A conceptual model to study the Conversion of Land Use and its Effects. *Ecological Modelling*, 85(2-3):253–270.

Venables, W. N. and Ripley, B. D. (2002). *Modern Applied Statistics with S.* Springer, New York, fourth edition.

Verburg, P. H., A Veldkamp, T., and Bouma, J. (1999). Land use change under conditions of high population pressure: The case of Java. *Global Environmental Change*, 9(4):303–312.

Verburg, P. H., de Nijs, T. C., Ritsema van Eck, J., Visser, H., and de Jong, K. (2004a). A method to analyse neighbourhood characteristics of land use patterns. *Computers, Environment and Urban Systems*, 28(6):667–690.

Verburg, P. H., Ellis, E. C., and Letourneau, A. (2011). A global assessment of market accessibility and market influence for global environmental change studies. *Environmental Research Letters*, 6(3):034019 (12pp).

Verburg, P. H. and 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.

Verburg, P. H., Schot, P. P., Dijst, M. J., and Veldkamp, A. (2004b). Land use change modelling: Current practice and research priorities. *GeoJournal*, 61(4):309–324.

Verburg, P. H., Soepboer, W., Veldkamp, A., Limpiada, R., Espaldon, V., and Mastura, S. S. A. (2002). Modeling the spatial dynamics of regional land use: The CLUE-S model. *Environmental Management*, 30(3):391–405.

Wessel, P. and Smith, W. H. F. (1996). A global, self-consistent, hierarchical, high-resolution shoreline database. *Journal of Geophysical Research: Solid Earth*, 101(B4):8741–8743.

Whittaker, R. J., Nogués-Bravo, D., and Araújo, M. B. (2006). Geographical gradients of species richness: A test of the water-energy conjecture of Hawkins et al. (2003) using European data for five taxa. *Global Ecology and Biogeography*, 16(0):76–89.

Wintle, B. A., Kujala, H., Whitehead, A., Cameron, A., Veloz, S., Kukkala, A., Moilanen, A., Gordon, A., Lentini, P. E., Cadenhead, N. C. R., and Bekessy, S. A. (2019). Global synthesis of conservation studies reveals the importance of small habitat patches for biodiversity. *Proceedings of the National Academy of Sciences*, 116(3):909–914.

Xian, G., Homer, C., Rigge, M., Shi, H., and Meyer, D. (2015). Characterization of shrubland ecosystem components as continuous fields in the northwest United States. *Remote Sensing of Environment*, 168:286–300.