

Introducing LandScaler: A novel method for spatial downscaling of land use projections

Tamsin L. Woodman^{a,*}, Cristina Rueda-Uribe^a, Roslyn C. Henry^a, David F.R.P. Burslem^a, Justin M.J. Travis^a, Peter Alexander^b

^a University of Aberdeen, UK

^b University of Edinburgh, UK

ARTICLE INFO

Handling Editor: Daniel P Ames

Keywords:

Colombia
Downscaling
Land use and land cover
Land use models

ABSTRACT

Land use and land cover (LULC) projections do not always have sufficient spatial resolution to allow them to be used by environmental models that project how LULC impacts a range of variables, including ecosystem services, biodiversity, and hydrology. We present a downscaling method designed to generate the high resolution LULC projections often required for environmental modelling. LULC change is allocated to a high-resolution reference map based on the density of LULC classes in neighbouring grid cells. Increasing a parameter that controls the likelihood of cells adjacent to existing LULC classes being converted to the same class generated less spatially aggregated landscapes that better represented historic LULC patterns in Colombia between 1960 and 2019. This new downscaling method is available as an R package and will enable the reconciliation of the spatial resolution of LULC projections and key processes that are embedded in a range of environmental models.

1. Introduction

Humans have been altering the land surface for thousands of years through activities such as clearing natural vegetation, growing crops, and extensive grazing of animals (ArchaeoGLOBE Project, 2019). Around 75% of the Earth's terrestrial surface has been impacted in some way by anthropogenic activities (Arnell et al., 2019; Luyssaert et al., 2014) and approximately 10% is intensively managed (Erb et al., 2017). The rate of land use and land cover (LULC) change has increased over much of the last century (Winkler et al., 2021) and LULC change is expected to continue in the future due to population growth, shifts in diet, and the need for energy production (Foley et al., 2011; Tilman et al., 2011). Land use has wide-ranging impacts on the climate system, biodiversity, hydrology, natural hazards such as landslides, and other components of the Earth System (Forzieri et al., 2017; Guo et al., 2022, 2023a, 2023b; Jaureguiberry et al., 2022; Meier et al., 2021; Newbold et al., 2015; Rigby et al., 2022; Salehpour Jam et al., 2023), for instance through greenhouse gas emissions and conversion of species' habitats (Jia et al., 2019; Pereira et al., 2012). Understanding the impacts of LULC change and modifying land use will be key to solving many of the challenges currently facing society, including climate change, food security, and the biodiversity crisis (Foley et al., 2011).

Projections of how LULC will change in the future are used to investigate the effects of LULC on the environment. Land use models make spatially explicit predictions of how the area of LULC classes will change over time (Heistermann et al., 2006), often based on a particular storyline of socio-economic development such as a Shared Socio-economic Pathway (O'Neill et al., 2014; Popp et al., 2017). Land use projections from models have been used to investigate the impacts of LULC on the environment; for example, global-scale LULC projections have been employed to investigate the effects of future LULC change on health outcomes, carbon storage, habitat loss, and soil erosion (Baisero et al., 2020; Borrelli et al., 2020; Henry et al., 2022; Molotoks et al., 2020). Similarly, Dullinger et al. (2020) used LULC projections from a local-scale agent-based model to simulate the distributions of plant species in the Austrian Alps, finding that LULC change was on average the most important predictor of species' current distributions.

Although there is a wide range of land use models available across local to regional scales (for example, Dullinger et al., 2020; Millington et al., 2021), global-scale land use models are important because some drivers of LULC change occur at a global scale, and geographically distant regions may be linked via mechanisms such as international trade (Heistermann et al., 2006), sometimes termed telecoupling (Hull and Liu, 2018). Although there have been a few global-scale land use

* Corresponding author.

E-mail address: tlw36@cantab.ac.uk (T.L. Woodman).

projections generated at fine scales (Chen et al., 2020; Li et al., 2017), there are many global-scale land use models which produce LULC projections that do not match the scale of environmental processes. For example, the outputs from many global land use models are at coarse resolutions (typically 0.25 or 0.5°), which are not equivalent to the scale on which organisms are affected by land use (de Chazal and Rounsevell, 2009; Titeux et al., 2016) or at which management decisions are made. Moreover, modelling a process at too coarse a resolution can lead to over- or underestimation of key model outputs. Suh et al. (2020) found that using land use projections at 5 arc minute resolution led to an over-estimation of carbon dioxide emitted from agricultural expansion in most regions, compared to using land use projections at 10 arc second resolution. Additionally, using coarser resolution land cover data as input to a mechanistic biodiversity model resulted in the model predicting substantially larger population sizes and rates of range expansion (Bocedi et al., 2012) than when the greater detail of finer resolution land cover was incorporated. There is strong evidence that the resolution of input LULC data is important for effectively modelling a range of key environmental processes.

Spatial downscaling, where variables at a coarse spatial resolution are converted to a higher resolution (van Vuuren et al., 2010), offers a method of incorporating LULC projections from global-scale models into studies that require fine resolution LULC data. Downscaling has been applied to global-scale LULC projections to investigate the effects of future LULC change on ecosystem services, extinction risk of plant communities, and soil erosion (Borrelli et al., 2020; Di Marco et al., 2019; Johnson et al., 2021; von Jeetze et al., 2023). A variety of methods are available for downscaling that range in complexity from simple algorithms, such as proportional downscaling, to fully coupled multi-scale models (de Chazal and Rounsevell, 2009; van Vuuren et al., 2010). Hasegawa et al. (2017), for example, converted regional land use outputs from the AIM/CGE integrated assessment model to gridded maps. While simple downscaling algorithms may not capture the local-scale processes which influence LULC patterns, more complex approaches tend to require more input data and are less transparent (van Vuuren et al., 2010).

Downscaling methods commonly use statistical relationships between LULC classes and covariates, such as human population density and soil nutrient availability, to determine the likelihood of fine resolution grid cells being covered by different LULC classes (for example, Cao et al., 2019; Le Page et al., 2016; van Asselen and Verburg, 2013; von Jeetze et al., 2023). One disadvantage of using statistical models for downscaling is that they assume the underlying processes of LULC change remain the same through time (Veldkamp and Lambin, 2001), whereas it has been shown that LULC change drivers vary temporally and spatially (Alexander et al., 2015). Statistical relationships between covariates and LULC can be combined with additional constraints that influence the probability of a grid cell containing a particular LULC class, such as neighbourhood rules, which govern how the LULC in surrounding cells affects LULC in the cell of interest (van Vliet et al., 2013). Regression models containing neighbourhood-based predictors only have been found to perform better at predicting LULC distributions in Belgium than those containing non-neighbourhood or mixed predictors (Dendoncker et al., 2007). However, a downscaling method based on neighbourhood-only regression models caused greater clustering of LULC classes compared to the observed LULC patterns (Dendoncker et al., 2006), suggesting that neighbourhood-based variables alone might not be sufficient for generating accurate landscape patterns through downscaling.

We introduce LandScaleR, a downscaling method that allocates LULC change to a fine resolution reference map based on neighbourhood rules and stochasticity in the placement of new LULC areas. LandScaleR is more likely to allocate LULC close to existing patches of the same LULC class, as has previously been observed (Dendoncker et al., 2007). We mitigate the clustering of LULC classes that can occur during downscaling by introducing a parameter that controls the degree of

stochasticity in the placement of new LULC areas, and which also impacts the likelihood that new LULC patches appear in the landscape. The method is generic and requires minimal input data aside from the LULC projections to be downscaled and a fine resolution reference map. We have implemented the method as an R package (R Core Team, 2022) due to the wide range of models and methods for handling LULC and environmental data that are already available in R (for example, García Molinos et al., 2019; Kapitzka et al., 2022; Malchow et al., 2021). The spatial arrangement of LULC is vitally important for environmental processes such as the persistence of a species in a landscape (Fischer and Lindenmayer, 2007), so we validated the landscape patterns generated by the downscaling method against historic LULC data for Colombia between 1960 and 2019 (Winkler et al., 2021). Our method will facilitate the integration of LULC with environmental models, such as biodiversity and geohazard models, by generating fine resolution projections that match the scale on which key environmental processes occur.

2. Methods

2.1. Overview of the approach

LandScaleR takes LULC projections from a land use model and allocates the change in LULC projected from one time point to the next to grid cells on a finer resolution reference map. The algorithm does not require input covariates to determine where LULC change is allocated on the reference map; instead, grid cells that are adjacent to an existing LULC patch are more likely to be converted to the same class. The resulting downscaled map becomes the reference for downscaling the subsequent timestep of LULC change, so multiple timesteps of LULC output from a land use model can be downscaled in a single run (Fig. 1). LandScaleR is designed to use gridded LULC change projections and a discrete reference map where each grid cell contains a single LULC class. The LandScaleR R package and source code are available via GitHub (TamsinWoodman/LandScaleR).

The first step in LandScaleR is to assign all fine resolution grid cells to their nearest coarse resolution grid cell, which determines where LULC change from the coarse cells will be allocated to on the reference map. Next, the areas of LULC change from the LULC projections are adjusted to match the total area of reference map grid cells per coarse resolution cell (Le Page et al., 2016). The LULC classes in the LULC projections are also matched to the reference map LULC classes. Given that the distribution of LULC in neighbouring cells influences LULC in a cell of interest (van Vliet et al., 2013), a kernel density is calculated for each fine resolution grid cell and LULC class (Le Page et al., 2016; West et al., 2014). The kernel density values give a measure of the density of each LULC class surrounding a focal cell and are used to determine the order in which fine resolution cells are converted to a new LULC class.

After calculating kernel densities for each fine resolution cell, LandScaleR begins the LULC change allocation process. Three methods are provided for allocating LULC change to the reference map: the quasi-deterministic method where LULC change is allocated to grid cells from highest to lowest kernel density, except where cells have a kernel density of zero in which case new LULC is allocated randomly; a “fuzzy” stochastic method where each kernel density is adjusted by adding a deviation drawn from a Normal distribution, and a random method where cells are chosen at random to receive new LULC. Following LULC allocation the new downscaled map is harmonised, meaning that the algorithm takes any LULC change that was not placed on the reference map during allocation and attempts to allocate it within neighbouring coarse grid cells instead. Harmonisation gives the final downscaled map, which is then used as the reference map to downscale the next timestep of LULC change.

2.1.1. Input maps

LandScaleR requires two sets of input maps: gridded, coarse

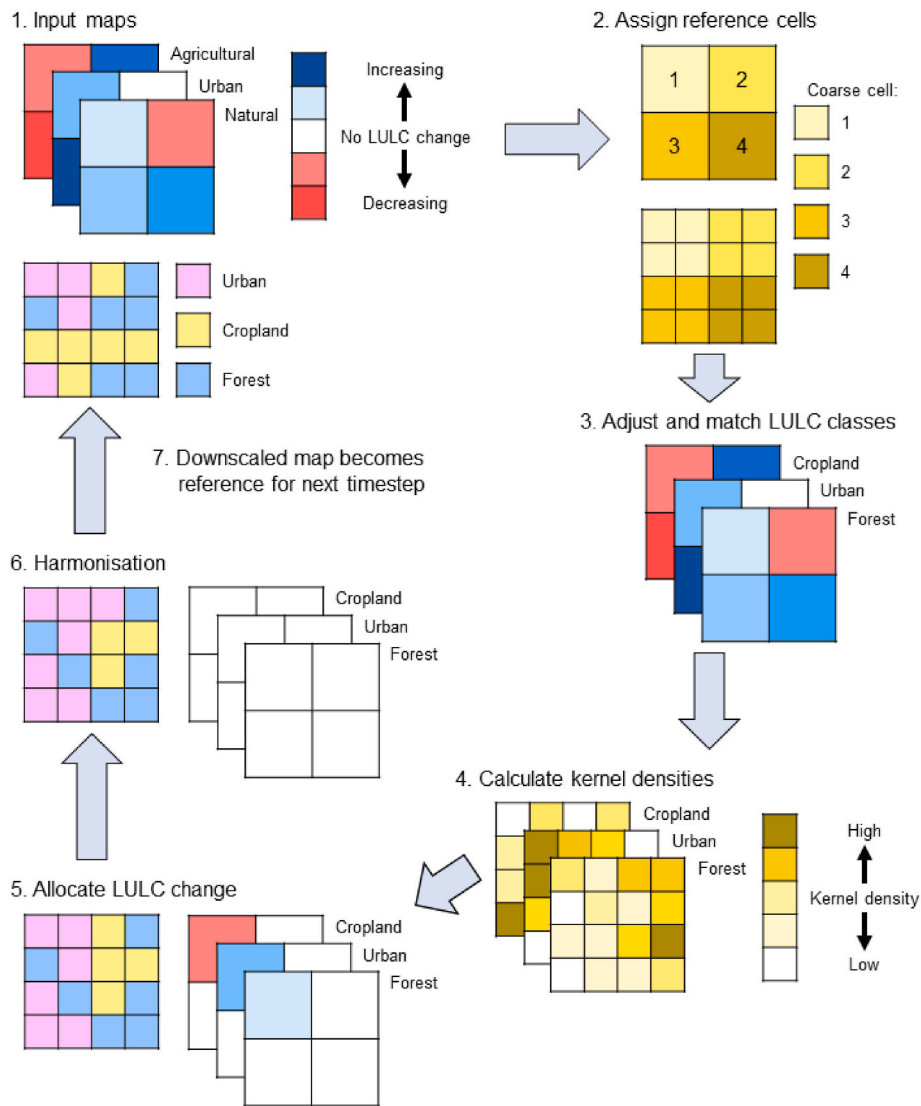


Fig. 1. Flow diagram showing the stages in the LandScaleR downscaling algorithm. 1. Input maps of LULC change and a high resolution reference map. 2. Cells from the reference map are each assigned to the nearest coarse resolution cell from the LULC change map. 3. LULC change areas in the LULC change map are adjusted and matched to those in the reference. 4. A kernel density is calculated for each LULC class and reference map grid cell based on the area of that LULC class in neighbouring cells. 5. LULC change is allocated to grid cells in the reference map according to the kernel density maps. 6. Unallocated LULC change is harmonised with the reference map until all change has been allocated. 7. The downscaled map becomes the input reference map for the next time step of LULC change.

resolution maps of LULC change, and a reference map at the resolution to which the LULC change maps will be downscaled. The input maps must have the same coordinate reference system and be on a regular grid, although the fine resolution cells do not need to nest exactly within the coarse resolution cells (Fig. 2). The coarse resolution LULC change maps provide the total area of change for each LULC class in each grid cell to LandScaleR, and the values of change can be positive or negative. There must be one map of LULC change per timestep. The fine resolution reference map should contain observed LULC in the timestep before the first timestep of the sequence of LULC change maps; for example, if the first timestep of LULC change provided is the change from 2010 to 2011 the reference map should give observed LULC in 2010. The LULC change between 2010 and 2011 will then be applied to the reference map to give downscaled LULC in 2011. The reference map can either contain a single LULC class per cell or the fraction of each LULC class per cell, although here we only validate LandScaleR with reference maps that contain one LULC class per cell. The LULC change and reference maps can have different LULC classes as these will be matched later in the downscaling process.

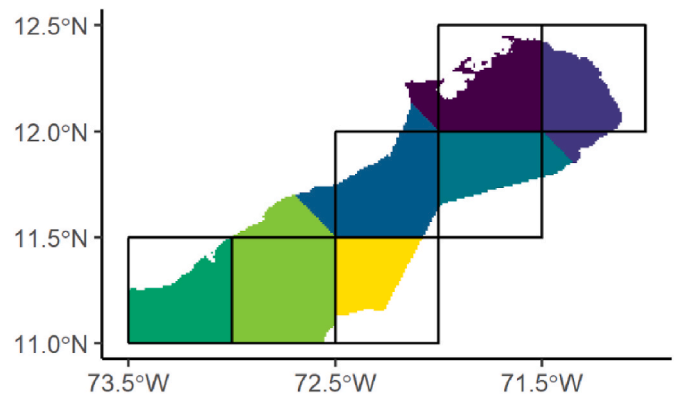


Fig. 2. Overlap of fine-scale grid cells at approximately 0.01° resolution (1 km at the equator) from the HILDA+ dataset and a set of hypothetical 0.5° resolution coarse cells in northern Colombia. Fine resolution cells are shaded according to their nearest coarse resolution cell, which they would be assigned to during downscaling. Black lines indicate 0.5° resolution grid cells.

2.1.2. Assign reference map grid cells to coarse resolution cells

The first stage in LandScaleR is to define where the LULC change from each coarse resolution grid cell will be allocated on the reference map, by assigning each fine resolution reference map cell to the nearest coarse resolution cell. Assigning grid cells from the reference map to the nearest coarse cell ensures that any cells not contained within a coarse cell can still be allocated new LULC, which accounts for cases where grid cells in the LULC change map do not fully cover the fine resolution cells (Fig. 2). Each coarse resolution cell can be assigned a different number of reference map cells, which is accounted for later in the method by adjusting the areas of LULC change by the total area of reference map cells assigned to each coarse cell (section 2.1.3). We use the ‘get.knnx’ function from the ‘FNN’ R package with the default kd tree algorithm to find the nearest coarse resolution grid cell to each reference map grid cell (Beygelzimer et al., 2022).

2.1.3. Adjust land use and land cover change areas

The areas of LULC change within each coarse resolution grid cell are adjusted by the total area of reference map cells assigned to each coarse cell, to prevent more LULC being allocated than the total area of assigned reference map cells (Le Page et al., 2016). The equation used by LandScaleR to adjust the LULC change areas by the corresponding area of assigned reference map grid cells is the same as equation (1) in Le Page et al. (2016). However, Le Page et al. (2016) downscaled LULC projections from the Global Change Assessment Model (GCAM), which produces regional rather than gridded LULC cover maps. Therefore, regions from GCAM output are replaced with coarse resolution grid cells in the adjustment equation in LandScaleR. The LULC area for a given LULC class and coarse resolution cell is adjusted by the ratio between the total area of assigned reference map cells and the area of the coarse cell:

$$\Delta A_{k,LC,t} = \delta A_{k,LC,t} \cdot \frac{\sum_{m=1}^J A_m}{A_k} \tag{1}$$

Where A is area, ΔA is adjusted LULC change area and δA is the original LULC change area, k indicates a coarse resolution grid cell, LC is one LULC class, t is timestep, m is a reference map grid cell, and J is the number of m assigned to one k . The area of each coarse resolution grid cell can either be provided by the user as an extra layer in the input map or calculated within LandScaleR using the ‘cellSize’ function in the ‘terra’ R package (Hijmans, 2022). The user can specify the units for calculating cell areas, which can be ‘m’ for metres, ‘km’ for kilometres, or ‘ha’ for hectares (Hijmans, 2022).

2.1.4. Match land use and land cover classes

The downscaling procedure in LandScaleR next matches the LULC classes from the map of LULC change to those in the reference map using a similar procedure to that in Le Page et al. (2016) and Vernon et al. (2018). LandScaleR currently converts the LULC classes from the LULC change map into the same classes as the reference map according to user-defined fractions. The user inputs a table specifying the proportion of each LULC class from the LULC change map to be converted into each of the reference map LULC classes (for example, Table 1). The

Table 1

Example showing how LULC classes from a map of LULC change could be matched to classes in a reference map by LandScaleR. Row names contain LULC classes from the LULC change map, while column names give classes from the reference map. Values in each row are specified by the user and give the proportions by which each LULC class from the LULC change map are divided between the reference map classes.

		LULC classes from the reference map					
		Primary vegetation	Secondary vegetation	Forest plantation	Cropland	Pasture	Urban
LULC classes from the LULC change map	Managed forest	0	0	1	0	0	0
	Unmanaged forest	1	0	0	0	0	0
	Other natural	0.5	0.5	0	0	0	0
	Cropland	0	0	0	1	0	0
	Pasture	0	0	0	0	1	0
	Urban	0	0	0	0	0	1

proportions of LULC classes in Table 1 are usually straightforward to determine; for example, the cropland class appears in both LULC maps. However, there may be definitional differences in LULC classes between the reference and LULC maps, in which case expert judgement can be used to determine the values in Table 1. The column names must contain the reference map LULC classes, and row names must be the LULC classes in the LULC change map. The values in the input table must be between zero and one and sum to one per each row. For managed forest in the example in Table 1, there is a value of one in the forest plantation column, meaning that if managed forest increased by 100 km² all the increase would be allocated to the reference map as forest plantation. Similarly, other natural land cover has a value of 0.5 in the primary and secondary vegetation columns. This indicates that a decrease of 50 km² in other natural land cover would be treated by LandScaleR as a 25 km² decrease in both primary and secondary vegetation.

2.1.5. Calculate kernel density values for each land use and land cover class

New areas of LULC have been previously found to occur near existing areas of the same class (Dendoncker et al., 2007). Based on this assumption, the LandScaleR downscaling algorithm calculates a kernel density for each reference map grid cell and LULC class, which is the distance-weighted sum of the area of a LULC class in neighbouring cells, divided by the number of neighbouring cells (Le Page et al., 2016; West et al., 2014). The kernel densities are used later in the downscaling process to select cells for LULC change allocation. We implemented equation (2) from Le Page et al. (2016) using the ‘focal’ function from the ‘terra’ R package (Hijmans, 2022):

$$kd_{LC,m} = \frac{1}{n} \cdot \sum_{i=1}^n \frac{A_{LC,i}}{(D_i)^2} \tag{2}$$

Where $kd_{LC,m}$ is the kernel density for LULC class LC in the focal reference map cell m , n is the number of neighbouring cells used in the calculation, $A_{LC,i}$ is the area of LULC class LC in neighbour cell i , and D_i is the distance between the neighbour cell i and the focal cell m . The user can specify the radius of cells that are used to calculate kernel density, which controls the number of neighbour cells n used in the kernel density equation; for example, a radius of one would mean eight neighbour cells are used. The distance D between a neighbour cell and a focal cell is calculated by setting the shortest possible distance between centroids in the reference map to one. Thus, the distance between a focal cell and its four connected neighbour cells would be one, and the distance between the focal cell and the four diagonally connected cells would be 1.41.

2.1.6. Allocation of land use and land cover change to the baseline map

The next step in LandScaleR is to allocate LULC change areas from the coarse resolution LULC change map to the reference map. Allocation of LULC change areas occurs independently for each coarse resolution cell because LULC change from one coarse cell can only be allocated to its assigned reference map cells. For each coarse cell, the first step in allocating LULC change is to define the magnitude of transitions

between each LULC class. Next, LULC change is allocated in order from the largest to the smallest LULC transitions in a cell. Three methods are described below for selecting where to place new LULC on a reference map: quasi-deterministic, the ‘fuzzy’ stochastic method, and random. The three LULC allocation methods lie on a continuum of stochasticity, with the quasi-deterministic method being the most deterministic and random the least.

2.1.6.1. Defining land use and land cover transitions. Land use models may not output the exact area of LULC that was converted between each class. Instead, a land use model might generate the total area or the change in each LULC class per timestep. Therefore, before allocating LULC change to the reference map LandScaleR calculates the area of each LULC class to be converted to each other class within a coarse resolution grid cell. For each LULC class that increases in representation within a coarse cell in the current timestep, the increase is divided proportionally between the LULC classes that decrease within the same grid cell. For example, assume that one LULC change map grid cell has an 80 km² increase in cropland, a 20 km² increase in pasture, a 60 km² decrease in forest and a further 40 km² decrease in grassland. LandScaleR would convert 48 km² of forest and 32 km² of grassland into cropland, while 12 km² of forest and 8 km² of grassland would become pasture.

LULC change in each coarse resolution cell is allocated sequentially to the reference map, starting with the LULC class with the largest increase in area in a cell. The LULC class with the highest increase in area is allocated to the reference map in order from the LULC class that decreases the most to the one that decreases the least. The algorithm then moves to the LULC class with the second largest increase in area, and so on until all LULC change has been allocated. In the example above the largest increase in LULC is 80 km² of cropland, so the algorithm begins by converting 48 km² of forest into cropland before replacing 32 km² of grassland with cropland. The algorithm then moves to the next largest increasing LULC class, which in our example is pasture. A total of 12 km² of forest is converted into pasture, and lastly the remaining 8 km² of grassland is changed to pasture.

2.1.6.2. Selection of cells to receive new land use and land cover. LandScaleR begins the process of converting LULC in the reference map once the magnitude and order of LULC conversions within one coarse resolution grid cell have been calculated. To convert one LULC class to another, the algorithm first selects all reference map grid cells that are eligible to receive the increasing LULC class. Grid cells are eligible if they contain the LULC class to be converted; for instance, in our example the first LULC transition is a conversion of 48 km² of forest into cropland, so all cells containing forest cover are suitable for conversion to cropland.

2.1.6.3. Quasi-deterministic method of land use and land cover allocation. The quasi-deterministic method of allocating LULC change places new LULC areas in reference map grid cells according to the kernel density of each grid cell for the new LULC class. Once all of the decreasing LULC class in cells with a kernel density value greater than zero has been converted, new LULC is randomly allocated to grid cells with a kernel density of zero. This provides a mechanism for new patches of a LULC class to occur in the landscape. The first step in the quasi-deterministic method is to obtain kernel densities for all cells that are eligible for conversion to the new LULC class and sort them from highest to lowest value. Any cells with a kernel density of zero are randomly sorted at the end of the list of grid cells. The algorithm works through the sorted grid cells and in each cell replaces all of the decreasing LULC class with the new LULC class, until the total area for that LULC transition has been converted.

For example, for a LULC transition of 48 km² from forest to cropland the algorithm would find and sort the cropland kernel densities for all reference map grid cells that contain forest cover. In order from highest

to lowest kernel density, the algorithm converts all forest area in each grid cell to cropland until 48 km² of forest has become cropland. If all forest area in grid cells with a cropland kernel density of greater than zero has been converted to cropland, the algorithm begins changing forest to cropland in the randomly sorted cells with cropland kernel density equal to zero, until 48 km² of forest has been replaced by cropland. Therefore, for each grid cell the area of forest converted to cropland is the minimum of the area of forest within that cell and the remaining forest to be converted.

2.1.6.4. Fuzzy method of land use and land cover allocation. The fuzzy method of allocating LULC change to a reference map introduces stochasticity into the placement of new areas of LULC to account for factors that affect LULC change other than distance from existing LULC patches, such as soil properties and topography. As in the quasi-deterministic method, the first step in the fuzzy method of LULC change allocation is to obtain kernel densities for all cells that are eligible for conversion to the new LULC class. The standard deviation σ of these kernel densities is then calculated as below, where $kd_{LC,c}$ is the kernel density for a single eligible baseline map cell c and LULC class LC , and G is the number of cells eligible to receive the new LULC class.

$$\sigma = \sqrt{\frac{\sum_{c=1}^G kd_{LC,c} - \overline{kd_{LC,G}}}{G - 1}} \quad (3)$$

For each cell c that is eligible for conversion, a deviation E is drawn from a Normal distribution with a mean of zero and standard deviation equal to the standard deviation of kernel densities (σ), multiplied by the user-specified parameter f :

$$E_{LC,c} \sim N(0, f \bullet \sigma) \quad (4)$$

The higher the value of f specified by the user, the more variation there will be in the values of E . The kernel density $kd_{LC,c}$ and random deviation $E_{LC,c}$ are summed for each cell to get the adjusted kernel density $KD_{LC,c}$.

$$KD_{LC,c} = kd_{LC,c} + E_{LC,c} \quad (5)$$

Using an f value of zero with the fuzzy method should generate the same outcomes as the quasi-deterministic method, while very high values of f will tend towards random LULC allocation. After adjusted kernel densities have been generated for all cells that are eligible to receive the new LULC class, they are sorted from highest to lowest. Starting with the cell with the highest adjusted kernel density, LandScaleR goes through the sorted list of cells and replaces all of the decreasing LULC class with the new LULC class, until the total LULC transition between the two classes has been met.

2.1.6.5. Random method of land use and land cover allocation. The random method of LULC allocation uses no information to select where to place new areas of LULC change on a reference map. LandScaleR randomly selects a reference map cell that is eligible to receive the new LULC class and replaces all of the decreasing LULC class in that cell with the new LULC class. The algorithm then randomly selects another eligible cell and continues the process of LULC conversion until the total area of the LULC transition has been converted.

2.1.7. Harmonisation of land use and land cover

There may be occasions when the downscaling algorithm cannot convert enough of one LULC class into another in a coarse resolution grid cell to meet the calculated LULC transition. For example, the algorithm might calculate that 48 km² of forest should be converted into cropland in a coarse grid cell. However, if the reference map grid cells assigned to that coarse cell contain only 10 km² of forest cover it will only be possible to convert 10 km² of forest to cropland, leaving 38 km² of cropland unallocated. Similarly, if the reference map grid cells are

completely covered by cropland it will not be possible to increase the area of cropland within them further; this would result in 48 km² of unallocated cropland. The downscaling algorithm therefore attempts to place any unallocated LULC change in nearby grid cells in a process known as harmonisation, to prevent large areas of unallocated LULC change and ensure consistency between the reference and LULC change maps (Rabin et al., 2020).

Our approach to harmonisation follows that of Rabin et al. (2020) in that LandScaleR searches neighbouring grid cells for a place where it can allocate the remaining LULC change. For a coarse resolution grid cell with unallocated LULC, the algorithm first redefines the LULC transitions within that cell as described in section 2.1.6.1. LandScaleR then selects the diagonally connected neighbour cell to the northwest of the focal cell and attempts to meet the LULC transitions by converting LULC within that cell. If there is still unallocated LULC change remaining after this, the algorithm moves to the first-neighbour cell to the north of the focal cell, recalculates the LULC transitions, and tries again to meet these transitions by converting LULC within the cell. The algorithm continues attempting to place unallocated LULC change in neighbouring coarse resolution cells until all LULC has been allocated, working through first-neighbour cells, followed by second-neighbour cells, in order from north-west to south-east and moving across each row from east to west. Harmonisation of a single coarse resolution grid cell stops when all unallocated LULC change has been placed on the reference map or once the algorithm has attempted to allocate LULC in all first- and second-neighbour cells, so up to a maximum of 24 alternative cells. If there is any unallocated LULC change after harmonisation, LandScaleR will generate a warning and print out the remaining unallocated LULC change within each grid cell.

2.1.8. Output maps

LandScaleR generates one output file per timestep containing a downscaled map with the area of each LULC class per grid cell, at the same resolution as the input reference map. There is also an option to output a second map file per timestep that contains the LULC class with largest area in each grid cell of the downscaled map. All output maps are in the GeoTIFF file format.

2.2. Method validation: downscaling land use and land cover change in Colombia

Colombia is a highly biodiverse country in northern South America, encompassing two of the world's biodiversity hotspots (Myers et al., 2000) and containing over 75,000 species of animals and plants (SiB Colombia, 2022). Large changes in LULC have occurred in Colombia historically (Etter et al., 2008), particularly in the Andean and Caribbean regions and in the last two decades also in the regions of the Amazon and Orinoco (Correa Ayram et al., 2020). Although the main drivers of changes in LULC vary by region, they include clearing of land for cattle grazing, crops, mining, and urbanization (Armenteras et al., 2011; Etter et al., 2008; González-González et al., 2021). Moreover, a peace agreement reached in 2016 between the government and the FARC (Fuerzas Armadas Revolucionarias de Colombia) has led to increased deforestation in areas that were previously inaccessible due to the conflict but now lack state control and have become more open to land grabbing and colonization (Armenteras et al., 2019). Forests are being converted into pastures for cattle ranching, licit and illicit crops, and human settlements (Ganzenmüller et al., 2022; Murillo-Sandoval et al., 2021; Van Dexter and Visseren-Hamakers, 2020).

Downscaling models have previously been applied to historic LULC data to assess their accuracy and validity (Cao et al., 2019; Le Page et al., 2016). Here, we applied LandScaleR to maps of LULC change in Colombia from 1960 to 2019 to test how well the algorithm can recreate historic land use patterns. Maps of LULC change in Colombia were generated from the HILDA+ dataset (Winkler et al., 2020, 2021), which provides global historic LULC data between 1899 and 2019 at

approximately 0.01° spatial resolution (1 × 1 km at the equator).

2.2.1. Input maps for downscaling land use and land cover change in Colombia

First, the global HILDA+ maps from 1960 to 2019 were cropped using an outline of Colombia from the 'rnatuarearth' package (South, 2017) to generate maps with approximately 0.01° spatial and yearly temporal resolution. The map of LULC in Colombia in 1960 at 0.01° resolution (Fig. 3a) was used as the baseline map for downscaling LULC change to 2019.

Maps of LULC change in Colombia from 1960 to 2019 were generated from the cropped HILDA+ maps of LULC in Colombia. A 0.5° resolution grid was created from 82.5°W to 66.5°W longitude and 4.5°S to 14°N latitude, which covered the entire terrestrial surface of Colombia. For each year between 1960 and 2019, the grid was overlaid with the 0.01° resolution map of LULC in Colombia and the area of each LULC class in each 0.5° grid cell was calculated to generate a map of LULC in Colombia at 0.5° spatial resolution. LULC change areas for each year were calculated by subtracting the area of each LULC class in each 0.5° grid cell from the area in the previous year. This process generated 59 maps of LULC change in Colombia at 0.5° spatial and yearly temporal resolution from 1961 to 2019, which were saved as GeoTIFF files for use with LandScaleR. All maps were handled in R (R Core Team, 2022) using the 'terra' R package version 1.6–17 (Hijmans, 2022). An example of LULC change in Colombia from 1960 to 1961 for four LULC classes is shown in Fig. 3b.

2.2.2. Parameters for downscaling land use and land cover change in Colombia

The 0.5° resolution maps of LULC change in Colombia produced from the HILDA+ dataset were downscaled using seven different methods, to test which was the most effective in recreating historic LULC patterns in Colombia. The five methods were: quasi-deterministic; five variations of the fuzzy method with the *f* parameter ranging from 1.0 to 2.0 at intervals of 0.25, and random. Ten simulations were run for each method to assess the variability generated by the downscaling algorithm. Aside from the method used to allocate LULC change, all other parameters were kept the same. The radius used to calculate kernel density was set to 1 as initial tests suggested that values greater than this generated very highly spatially aggregated landscapes. The unit for cell areas was kilometres. The LULC classes were the same in the LULC maps and the reference map of Colombia, so each LULC class from the LULC change maps was matched entirely to the same LULC class in the reference map.

2.2.3. Analysis

We tested how well LandScaleR was able to recreate historic LULC patterns in Colombia from the HILDA+ dataset. LandScaleR was intended to generate realistic landscape patterns whilst using minimal covariates to predict the location of LULC change, so it was not expected to generate maps that exactly matched LULC in Colombia over time. Moreover, the HILDA+ dataset describes gross LULC change whereas the downscaling algorithm works with net LULC change, so there should be more LULC change occurring in the HILDA+ dataset than in the downscaled maps. Based on these factors and the importance of landscape patterns for a range of environmental processes such as the movements of individual organisms (Fischer and Lindenmayer, 2007), fire ignition and spread (Pais et al., 2021; Ryu et al., 2007), hydrological dynamics (Amiri et al., 2018), and accumulation of soil organic carbon (Liu et al., 2022), we calculated landscape pattern metrics at both the landscape- and class-levels to test whether LandScaleR could accurately reproduce historic landscape configurations. We also calculated the proportion of grid cells in downscaled maps that contained the same LULC class as the HILDA+ map, known as overall accuracy (Mas et al., 2022), to give a measure of the similarity between the observed and predicted maps and how this similarity decayed over time. Additionally, Figure of Merit (FM; Paegelow et al., 2022; Pontius et al., 2008) was

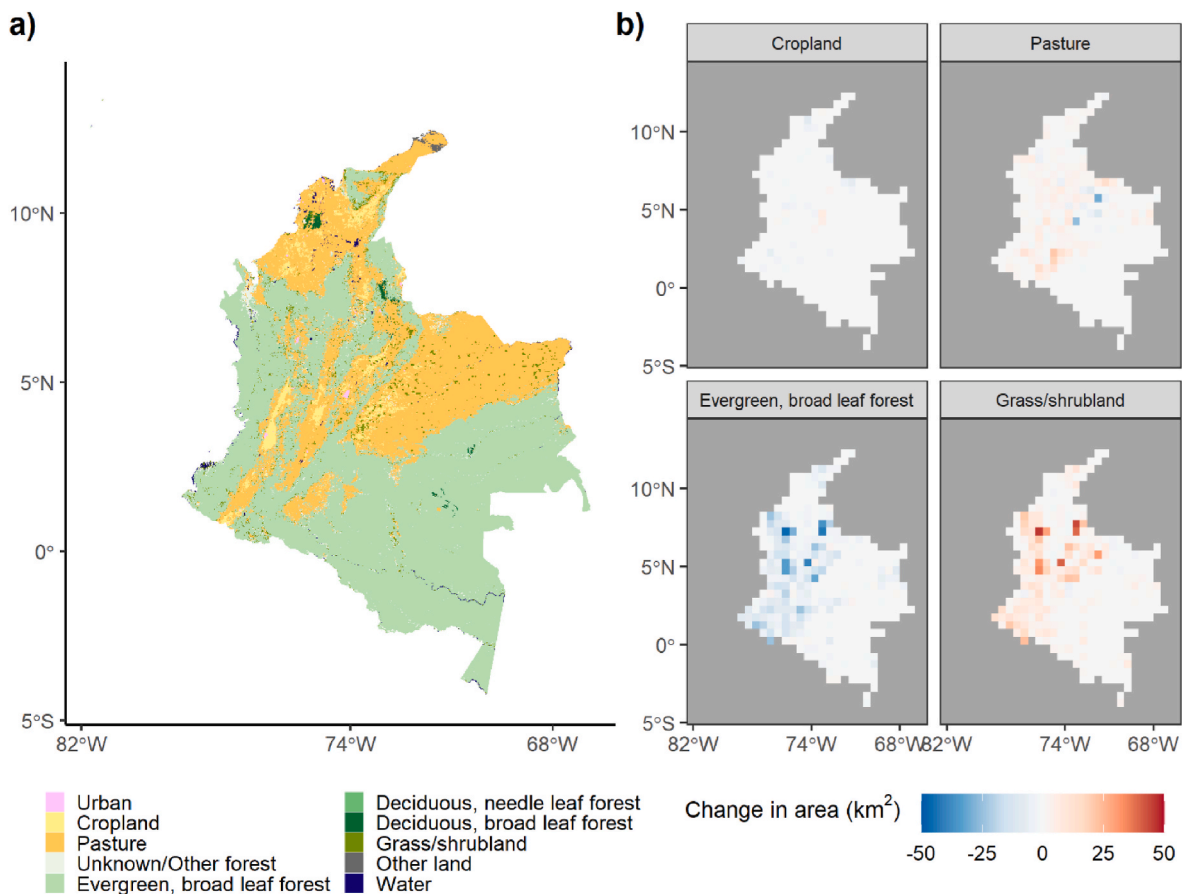


Fig. 3. LULC in Colombia in 1960 (a) and LULC change for four LULC classes between 1960 and 1961 (b), derived from the HILDA+ historic LULC dataset.

computed at yearly intervals and across the entire study period to measure the overlap of change between downscaled and HILDA+ maps. All analyses were performed in R version 4.1.3.

2.2.3.1. Overall accuracy. The proportion of grid cells in each downscaled map that contained the same LULC class as the corresponding grid cell from the observed HILDA+ dataset was calculated for each year of the study period to test the accuracy of LandScaleR at grid cell level. This statistic is known as overall accuracy (Mas et al., 2022) and was calculated for each simulation and time step by summing the number of cells that had the same LULC class in both the downscaled and HILDA+ maps, then dividing by the total number of terrestrial grid cells. Version 1.5–21 of the ‘terra’ R package (Hijmans, 2022) was used to calculate the proportion of cells with the expected LULC class.

2.2.3.2. Figure of Merit. FM is a measurement of the overlap between change that was observed in a reference map and change predicted by a land use model (Paegelow et al., 2022; Pontius et al., 2008). FM was calculated as:

$$FM = \frac{B}{(A + B + C + D)} \times 100\% \quad (6)$$

Where A is the number of grid cells which changed LULC class in the HILDA+ map but not in the downscaled map; B is the number of grid cells which changed class in both maps and were correctly predicted by LandScaleR; C is the number of grid cells which changed in both maps but the LULC class was incorrectly predicted by LandScaleR, and D is the number of grid cells which did not change class in HILDA+ but were predicted to change by LandScaleR. A FM metric of 0% would indicate that the location and class of LULC change was entirely incorrectly

predicted by the downscaling algorithm, while a value of 100% would show that the location and class of all LULC change was correctly predicted (Paegelow et al., 2022; Pontius et al., 2008). We calculated FM for each downscaling method in two ways. First, FM was calculated at yearly time steps across the study period by comparing the transitions that occurred in each year during downscaling to the transitions observed in the same year in HILDA+, to test the accuracy of LandScaleR in predicting yearly change. Second, because incorrect LULC transitions in one year might be compensated for by the same change in HILDA+ in another year, we calculated FM across the entire study period using the difference in LULC between 1960 and 2019.

2.2.3.3. Landscape pattern analysis. Five landscape pattern metrics were chosen to compare the downscaled maps with the HILDA+ maps: percentage of landscape covered by each LULC class, mean patch area, number of patches, Aggregation Index, and edge density. We calculated the percentage of Colombia covered by each LULC class to ensure that LandScaleR generated the correct area of each class, as this was not prescribed by the model. A previous study found that downscaling based on neighbourhood-based variables alone generated larger, more aggregated patches with shorter edge length than observed (Dendoncker et al., 2006), hence we selected mean patch area, Aggregation Index, and edge density to test whether LandScaleR generated patches of the expected size and aggregation. Moreover, mean patch area is important in many environmental processes; for example, the species-area relationship states that larger habitat patches will hold more species (Fischer and Lindenmayer, 2007). The Aggregation Index measures habitat aggregation and distinguishes the effects of habitat fragmentation from habitat loss in a landscape (Wang et al., 2014). Aggregation Index scales from zero to one hundred and is calculated as the number of edges shared by grid cells containing the same LULC class divided by the

hypothetical maximum possible number of edges that could be shared between those cells (Hesselbarth et al., 2019). Edge density is determined as the length of edges in the landscape divided by the total area of the landscape (Hesselbarth et al., 2019) and was recently found to be a suitable metric for calibrating a cellular automata model of LULC change (Lin et al., 2020). Finally, the number of patches provides a measure of the fragmentation of a landscape (Wang et al., 2014) and would be expected to increase as aggregation decreases. Number of patches has also been employed in previous studies to assess the accuracy of LULC models (Dezhkam et al., 2016).

The five landscape pattern metrics were calculated for the entire landscape and for each LULC class using the 'landscapemetrics' R package, version 1.5.5 (Hesselbarth et al., 2019), because class-level and landscape-level metrics can show different results in terms of the accuracy of a land use model (Dezhkam et al., 2016). Prior to calculating landscape statistics, the HILDA+ maps of Colombia at 0.01° resolution from 1960 to 2019 and the downscaled maps across the same period were converted from the EPSG:4326 coordinate reference system to the equal area World Eckert IV system, as the 'landscapemetrics' package recommends that maps are provided in units of metres. Landscape pattern metrics were calculated for the downscaled maps generated from every method of allocating LULC change and every simulation, so a total of 59 maps for 70 individual downscaling simulations. The resulting landscape pattern metrics were then summarised by calculating the mean and standard deviation across the ten simulations for each LULC change allocation method.

3. Results

3.1. Percentage of landscape in each land use and land cover class

LandScaleR accurately recreated the percentage of Colombia covered by each LULC class for each year of the simulations and for all LULC change allocation methods (Fig. S 1). The largest LULC class in Colombia in every year from 1960 to 2019 was evergreen broad leaf forest, which decreased consistently across the study period from 60.1% of Colombia in 1960 to 52.4% in 2019. Pasture was the second major LULC class in Colombia and it increased between 1960 and 2019, likely at the expense of evergreen broad leaf forest as no other LULC classes decreased substantially. The percentage of Colombia covered by grass/shrubland was very low (2.3%) at the beginning of the study period but increased from around 1990 onwards to reach 5.3% in 2019. The other LULC classes covered a small percentage of Colombia in the HILDA+ dataset and deciduous needle leaf forest disappeared from HILDA+ and all downscaling simulations just after the year 2000. Therefore, dividing increases in one LULC class proportionally between any decreasing LULC classes in one year appeared to be an adequate strategy for defining LULC transitions during downscaling.

3.2. Overall accuracy

The overall accuracy of LandScaleR, calculated as the proportion of grid cells in a downscaled map that contained the same LULC class as the corresponding cell in the HILDA+ map from the same time step, varied little between the quasi-deterministic and fuzzy methods (Fig. 4a). The quasi-deterministic and fuzzy with f of 1.0 LULC allocation methods generated downscaled maps with the highest overall accuracy (means and standard deviations of $0.88 \pm 1.4 \times 10^{-4}$ and $0.88 \pm 2.8 \times 10^{-3}$ in 2019 for quasi-deterministic and fuzzy with f of 1.0, respectively). Overall accuracy decreased very slightly as the f parameter for the fuzzy method increased, while the random LULC allocation method consistently produced downscaled maps with the lowest accuracy ($0.85 \pm 1.5 \times 10^{-4}$ in 2019), suggesting that both the quasi-deterministic and fuzzy methods of LULC allocation were more accurate than allocating LULC at random.

Overall accuracy generally decreased over time as errors

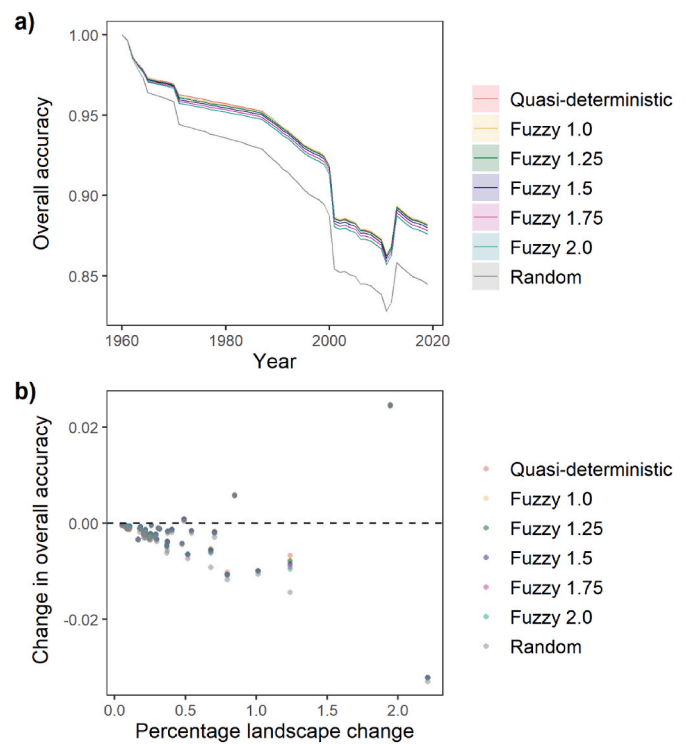


Fig. 4. Overall accuracy decreases as the LULC allocation method becomes more stochastic. a) Mean overall accuracy decays over time. Lines give the mean and shaded areas the standard deviation of the proportion of cells containing the expected LULC class across ten simulations per LULC allocation method. Standard deviations were small (range of 1.9×10^{-6} to 3.9×10^{-4}) and were not sufficiently large to appear on the plot. b) Difference in the mean overall accuracy between years according to the percentage of land area in Colombia that changed LULC class in the same year.

compounded across each simulation (Fig. 4a). Larger decreases in overall accuracy were observed in years when more LULC change occurred (Fig. 4b), which was expected as there will be a higher likelihood of the downscaling algorithm making errors when more LULC is added to a reference map. The overall accuracy was very similar for all methods, except the random method, in most years of the study period. A particularly large decrease in overall accuracy of the downscaling method occurred in 2001 when 23,757 km² of grass/shrubland was added to the reference map in regions that previously had very little grass/shrubland cover (Fig. S 2), meaning that there was little information available in the downscaling algorithm to guide where the new grass/shrubland was allocated on the reference map. Unexpectedly, there were three years (2003, 2012 and 2013) when overall accuracy increased for all LULC allocation methods. These increases were associated with a loss of grass/shrubland cover, generating larger, more uniform patches of LULC that were more similar across the downscaled and HILDA+ maps than the small, scattered patches of grass/shrubland that they replaced (Fig. S 2).

3.3. Figure of Merit

FM also showed little variation across the seven LULC allocation methods when calculated at yearly time steps over the study period (Fig. 5a). However, the random method of LULC allocation had a consistently lower FM index compared to the quasi-deterministic and fuzzy methods, suggesting that these methods perform better than allocating LULC at random. FM was often relatively low at yearly intervals, with mean values ranging from 0.3% to 21.1%. Comparatively, FM was generally higher when calculated across the entire study period from 1960 to 2019 (Fig. 5b), so incorrect LULC transitions during

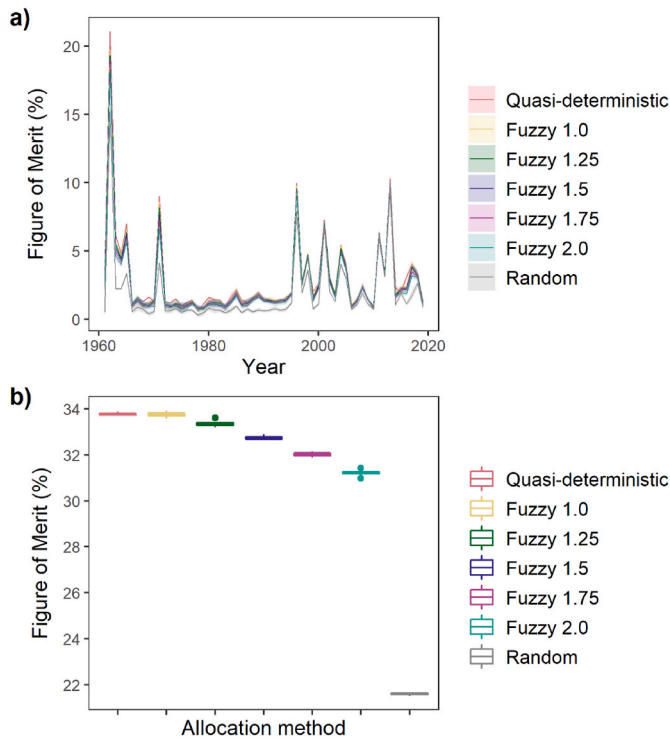


Fig. 5. Figure of Merit decreases slightly as the LULC allocation method becomes more stochastic. a) Mean Figure of Merit varied across years in the study period. Lines give the mean and shaded areas the standard deviation of Figure of Merit across ten simulations per LULC allocation method. b) Figure of Merit index as calculated for the entire study period (between 1960 and 2019).

downscaling in one year might be correct in HILDA+ in another year. FM was found to decrease slightly with increasing stochasticity in LULC allocation when calculated across the whole study period. FM was highest for the quasi-deterministic and fuzzy with f of 1.0 methods ($33.8\% \pm 0.056$ and $33.8\% \pm 0.10$ for the quasi-deterministic and fuzzy f of 1.0 methods, respectively) and lowest for the random method ($21.6\% \pm 0.048$). The fuzzy method with f of 2.0 had the second lowest FM value of $31.2\% \pm 0.11$, which was only slightly less than the highest FM ratios. Therefore, there was little difference in the accuracy of LandScaleR between the quasi-deterministic and fuzzy methods in terms of both overall accuracy and FM, although there was a tendency for accuracy to decrease with increasing stochasticity in LULC allocation. The random method of LULC allocation consistently performed worse than the quasi-deterministic and fuzzy methods.

3.4. Landscape pattern analysis

The seven LULC change allocation methods used to downscale LULC change in Colombia produced maps with differing landscape patterns. For all four landscape pattern metrics in Fig. 6, the fuzzy LULC change allocation method with an f -value of 2.0 was closest to the observed patterns in HILDA+ in 2019. Downscaling LULC change with the quasi-deterministic method generated landscapes with much larger mean patch area, higher Aggregation Index, and lower number of patches and edge density, than the actual HILDA+ landscapes. Increasing stochasticity in the placement of LULC change on reference maps by increasing the f parameter in the fuzzy method reduced the tendency of LandScaleR to produce very aggregated landscapes and generated maps with more similar patterns to HILDA+. The random method of LULC allocation led to landscapes with a much smaller mean patch area and Aggregation Index than observed in Colombia from the HILDA+ data, suggesting that increasing stochasticity in LULC change allocation will generate landscapes more similar to observed ones up to an optimal value. The

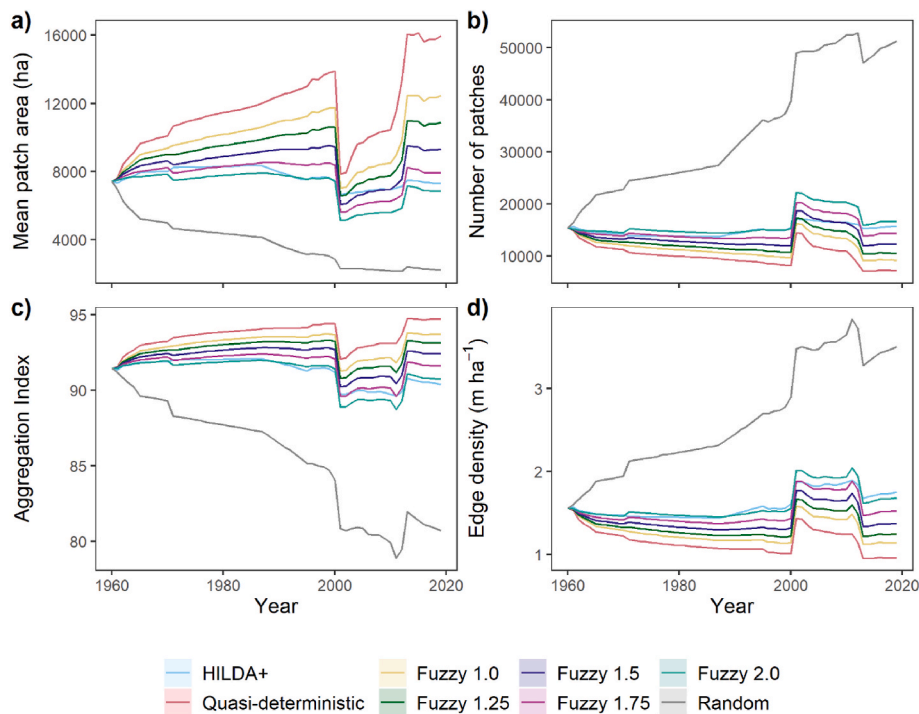


Fig. 6. Using a more stochastic LULC allocation method can better recreate landscape patterns observed in HILDA+. Mean patch area (ha) (a), number of patches (b), Aggregation Index (c), and edge density ($m\ ha^{-1}$) (d) by year. Lines give the mean and shaded areas the standard deviation of each landscape pattern metric across ten simulations where LULC change in Colombia was downscaled using one method of LULC change allocation. Standard deviations were small (range of 0 to 73 for mean patch area; 0 to 182 for number of patches; 0 to 0.039 for Aggregation Index, and 0 to 6.9×10^{-3} for edge density) so may not be sufficiently large to appear on the plot.

optimal value of f for downscaling LULC change in Colombia varied over time as the landscape properties in Colombia were not constant over the study period (Fig. 6).

The optimal method for downscaling LULC change in Colombia varied by LULC class as well as through time (Fig. 7). For example, mean patch area of pasture in 2019 in Colombia from the HILDA+ dataset was 24,332 ha and the most similar downscaled mean patch area was 23,954 ha from the fuzzy method with an f -value of 1.5. Comparatively, mean patch area of evergreen broad leaf forest in Colombia from HILDA+ was 26,277 ha and the fuzzy method with f -value of 2.0 generated the most similar mean patch area (33,981 ha). Moreover, while all methods of LULC change allocation apart from random produced downscaled landscapes with a comparable Aggregation Index to HILDA+ for the pasture and evergreen broad leaf forest classes, the quasi-deterministic and fuzzy methods all overestimated the Aggregation Index for grass/shrubland. Therefore, in the case of grass/shrubland a higher f -value than tested here might be optimal for reproducing the Aggregation Index. The quasi-deterministic and fuzzy methods all tended to produce more similar patterns of pasture, evergreen broad leaf forest and grass/shrubland to HILDA+ than the random method of LULC change allocation.

A region of Colombia is shown in Fig. 8 to demonstrate how the pattern of the downscaled landscapes compared to HILDA+ for three LULC change allocation methods in 2019: quasi-deterministic, fuzzy with f of 2.0, and random. The quasi-deterministic method generated a more aggregated landscape with fewer patches compared to the HILDA+ landscape. Comparatively, the random method led to a disaggregated landscape with many small patches of grass/shrubland and

visible outlines of at least two of the individual coarse-scale 0.5° grid cells. The downscaled landscape produced by the fuzzy method of allocation with f of 2.0 was most similar to the HILDA+ landscape in terms of the number and size of patches, but sometimes the patches were more aggregated than observed in HILDA+ and occurred in different locations. Landscapes from one simulation per LULC allocation method only are shown in Fig. 8, and although landscapes from simulations using the same allocation method have very similar average landscape patterns (Fig. 6) they may vary in the placement of LULC patches (Fig. S 3).

4. Discussion

The LandScaleR downscaling algorithm was able to reproduce LULC patterns as observed in Colombia between 1960 and 2019. The quasi-deterministic method of LULC allocation generated landscapes that were more spatially aggregated than HILDA+. Previous work demonstrated that modelling LULC change in Belgium using only neighbourhood variables led to much higher mean patch areas and lower total edge length than observed LULC (Dendoncker et al., 2006). Introducing stochasticity into LandScaleR, which is based on neighbourhood rules alone, generated more similar landscape patterns to observed ones by allowing for new patches of LULC to appear, and thereby decreasing mean patch area and increasing edge length. Increasing the LandScaleR f -parameter in the fuzzy method produces increasingly less aggregated LULC patterns, with an f -parameter value of 2.0 producing patterns similar to those observed in Colombia. Downscaling methods which incorporate both neighbourhood and non-neighbourhood variables (Cao

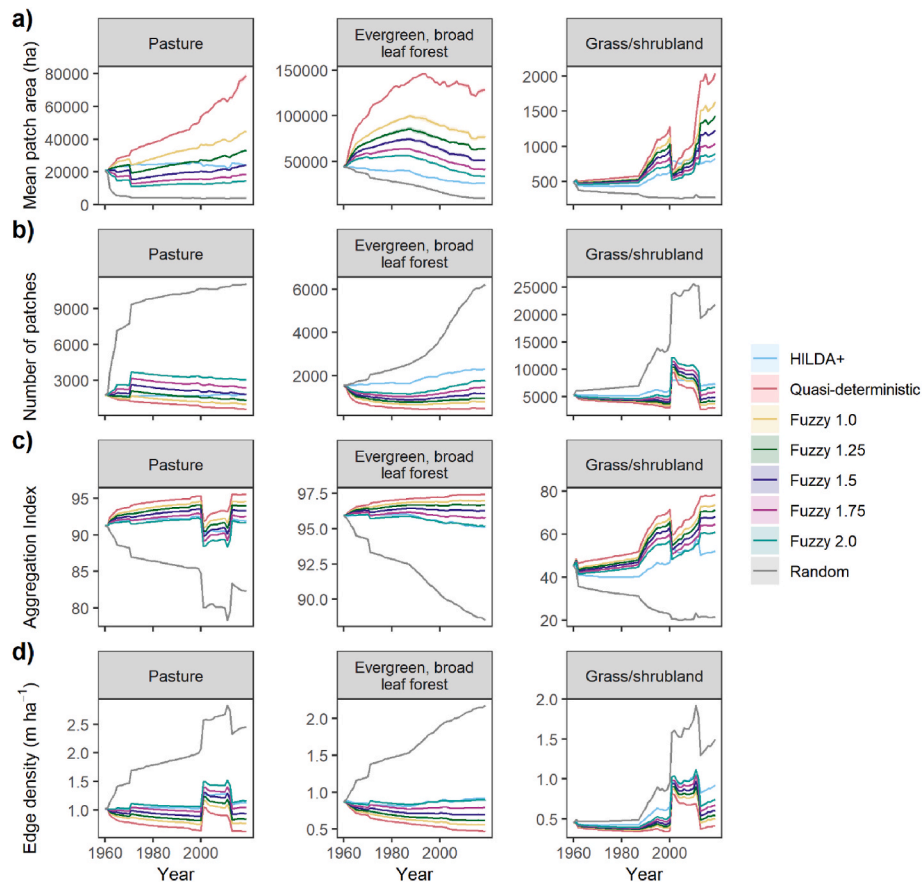


Fig. 7. Landscape pattern metrics for HILDA+ and downscaled maps of Colombia for the three largest LULC classes in Colombia in 2019. Mean patch area (ha) (a), number of patches (b), Aggregation Index (c) and edge density (m ha^{-1}) (d) by year and LULC class. Lines give the mean and shaded areas the standard deviation for a landscape pattern metric for one method of LULC change allocation across ten simulations. Standard deviations were small (range of 0 to 3378 for mean patch area; 0 to 122 for number of patches; 0 to 1.4 for Aggregation Index, and 0 to 7.4×10^{-3} for edge density) so may not be sufficiently large to appear on the plot.

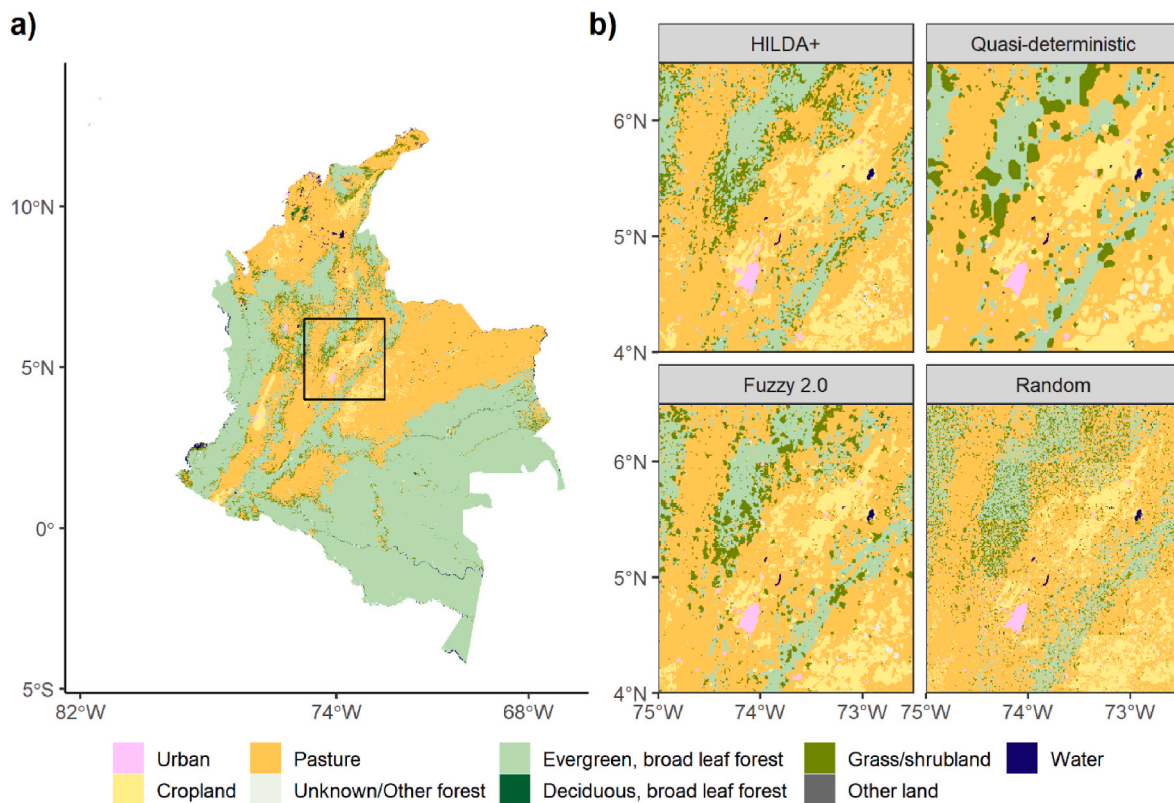


Fig. 8. The LULC change allocation method used in downscaling can be adjusted to better recreate historic LULC patterns. LULC in Colombia in 2019 from the HILDA+ dataset (a) with a black square indicating the inset in b). LULC in 2019 is shown in the inset (b) for the HILDA+ dataset and three methods of allocating LULC change during downscaling: quasi-deterministic, fuzzy with f -parameter of 2.0, and random.

et al., 2019; Le Page et al., 2016; von Jeetze et al., 2023) enable the appearance of new LULC patches in a landscape through correlations between LULC classes and environmental variables, so the f -parameter in LandScaleR is likely emulating some of the processes that drive LULC patterns, such as soil properties, topography, and land management.

In addition to landscape pattern metrics we also calculated the overall accuracy and FM of the downscaling method for recreating historic LULC in Colombia. LandScaleR was much more effective at correctly predicting LULC transitions in some years than others, as yearly FM varied from 0.3% to 21.1%. Pontius et al. (2008) found that FM of a range of LULC models increased with increasing net LULC change, so it may be that the low yearly FM ratios we observed were due to the low net LULC change observed in Colombia in HILDA+ in several years of the study period. The results for FM calculated over the entire period (31.2–33.8% for the quasi-deterministic and fuzzy LULC allocation methods) were comparable to other LULC models, which often have FM indices of less than 50% (Cao et al., 2019; Hu et al., 2022; Pontius et al., 2008). All quasi-deterministic and fuzzy methods had a higher FM than allocating LULC at random and there was little difference in FM between these methods, suggesting that any of the quasi-deterministic and fuzzy methods would be sufficient in terms of grid cell-based accuracy for downscaling LULC in Colombia from HILDA+.

Overall accuracy was approximately 0.88 in 2019 when downscaling was run with the quasi-deterministic and fuzzy with f of 1.0 LULC allocation methods, which was relatively high and comparable to the accuracy of a multinomial logistic regression model used to predict LULC in Belgium with neighbourhood variables only (Dendoncker et al., 2006). The errors made by the downscaling algorithm accumulated over time, which is a feature of other downscaling models (Chen et al., 2019), and higher decreases in accuracy were observed in years when more LULC change occurred.

The largest decrease in the overall accuracy of LandScaleR occurred

in 2001 when many small patches of grass/shrubland appeared in HILDA+ in the Llanos region, which is a globally important area of tropical savannas that covers 17 million hectares of land in Colombia (Romero-Ruiz et al., 2010). Conversely, twelve years later there was an increase in overall accuracy as the small patches of grass/shrubland in the Llanos region disappeared. HILDA+ is a time series of global LULC maps that integrates data from a number of other LULC products. From 2001 to 2013 HILDA+ incorporated yearly MODIS data at a global scale (Winkler et al., 2020, 2021), which corresponds to the timings of the substantial fluctuations in grass/shrubland cover in the Llanos region and suggests that differences in LULC classifications between MODIS and the other LULC products used by HILDA+ likely caused these apparent changes in grass/shrubland cover. Differences in LULC classifications might be created by difficulties in distinguishing between natural grass/shrubland and pasture areas, especially as extensive cattle grazing is one of the dominant land uses in the Llanos region (Romero-Ruiz et al., 2012). Therefore, the particularly large decrease and subsequent increase in the overall accuracy of LandScaleR in 2001 and 2013, respectively, were caused by inconsistencies in the input data. Downscaling the inconsistencies in grass/shrubland cover showed that LandScaleR may be less accurate at the grid cell level when there are few or no existing patches of a LULC class in a region, because any new areas of that LULC class will be allocated in a highly stochastic manner.

Validation of LandScaleR with historic LULC data from Colombia demonstrated that there is a trade-off between overall accuracy and landscape patterns in the method, as has been found previously when assessing the accuracy of a set of cellular automata models with cell-based and pattern-based metrics (Lin et al., 2020). Overall accuracy of the downscaling algorithm decreased marginally with higher stochasticity in LULC allocation, whereas landscape patterns became increasingly similar to those observed in Colombia up to an optimal f value. LandScaleR includes minimal covariate data for LULC allocation and is

intended to generate landscapes that have realistic LULC patterns rather than high predictive accuracy at the grid cell level. LandScaleR can be calibrated against historic LULC datasets to find the optimal values of the f -parameter and kernel radius for a region and LULC class to be downscaled; for example, based on our results it would be appropriate to select the fuzzy method with an f -value of 2.0 to downscale future LULC change in Colombia. This LULC allocation method best recreated landscape-level LULC patterns in Colombia but was not optimal for all years or LULC classes. For example, using an f -value of 2.0 tended to produce more aggregated grass/shrubland patches than observed in HILDA+, so higher values of f could be tested in future to assess whether they better reproduce the characteristics of grass/shrubland patches in Colombia. Future work could also investigate the effects of scale on calibration of the f -parameter and kernel radius, and the accuracy of downscaled maps, as scale has been shown to be important in land use modelling (Blanchard et al., 2015). The f -parameter and kernel radius could also be chosen to depict specific socio-economic scenarios. For instance, a future scenario where land systems become dominated by a few land managers might have more spatially aggregated patterns of LULC than a scenario where there are many land managers who control smaller areas.

While some LULC downscaling algorithms have been developed for downsampling LULC outputs from a specific land use model (for example, Hasegawa et al., 2017; Le Page et al., 2016; Vernon et al., 2018), LandScaleR is generic and can be applied to any gridded LULC projections. The algorithm is provided in the 'LandScaleR' R package and requires minimal input data because it allocates LULC based on the density of LULC classes in surrounding grid cells and the user-defined f -parameter, and does not rely on statistical relationships between LULC and environmental covariates that might change over time. LandScaleR only has two input parameters (f and kernel radius) compared to some downscaling models which might have ten or more parameters (von Jeetze et al., 2023). Both the f -parameter and kernel radius parameter can be calibrated using historic LULC data or adjusted for specific socio-economic scenarios. Further development of the LandScaleR algorithm could allow the f -parameter to vary by LULC class, through space, and over time, which would enable the algorithm to better recreate LULC patterns of individual classes, although at the cost of greater effort required for calibration. Developing approaches to formally fit the f -parameter to spatio-temporal patterns of LULC change, for example using Bayesian approaches, would be a natural direction to progress this work in and would open up opportunities for improving understanding of land-use change dynamics as well as improving forecasting capability.

5. Conclusions

We present LandScaleR, a spatial downscaling algorithm that is able to generate realistic spatial landscape configurations with minimal input data. LandScaleR controls the placement of new areas of LULC on a reference map via the density of LULC classes in neighbouring grid cells and a stochasticity parameter that alters the likelihood of grid cells receiving new LULC. The introduction of stochasticity into LandScaleR allows for the appearance of new LULC patches in the landscape and the stochasticity parameter can be calibrated by the user to best recreate historic LULC patterns in a particular region, or to represent a socio-economic scenario. LandScaleR was able to recreate historic LULC patterns in Colombia, a country with high biodiversity and ongoing loss of natural land covers. Increasing stochasticity of LULC allocation generated landscape patterns which better matched historic patterns in Colombia, up to an optimum amount of stochasticity. The LandScaleR algorithm is provided as an R package with the aim of furthering the integration of LULC with models that investigate the impact of LULC on key environmental processes such as hydrology, biodiversity, carbon emissions, and landslide risk. Downscaled projections from LandScaleR could also be applied in natural hazard models that require LULC

projections (Guo et al., 2022, 2023a, 2023b). Furthermore, LandScaleR could aid in the design of land management strategies with high spatial resolution that minimise the impacts of LULC on environmental processes.

Software and data availability

Software name: LandScaleR.

Developer: Tamsin L. Woodman.

Developer contact information: tlw36@cantab.ac.uk.

First year available: 2023.

Hardware requirements: PC/Mac.

Software requirements: R statistical environment and language;

Software availability: <https://github.com/TamsinWoodman/LandScaleR>;

Cost: free.

Program language: R.

Program size: 3.56 MB.

The code and HILDA+ dataset that were used to validate the LandScaleR method are available at <https://github.com/TamsinWoodman/LandScaleR-validation> and <https://doi.pangaea.de/10.1594/PAN-GAEA.921846> (Winkler et al., 2020), respectively.

Declaration of competing interest

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

Acknowledgements

TW was funded by an EASTBIO UKRI BBSRC grant number BB/T00875X/1. Downscaling simulations and calculation of landscape pattern metrics were performed on the University of Aberdeen HPC, Maxwell.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.envsoft.2023.105826>.

References

- Alexander, P., Rounsevell, M.D.A., Dislich, C., Dodson, J.R., Engström, K., Moran, D., 2015. Drivers for global agricultural land use change: the nexus of diet, population, yield and bioenergy. *Global Environ. Change* 35, 138–147. <https://doi.org/10.1016/j.gloenvcha.2015.08.011>.
- Amiri, B.J., Junfeng, G., Fohrer, N., Mueller, F., Adamowski, J., 2018. Regionalizing flood magnitudes using landscape structural patterns of catchments. *Water. Res. Manag.* 32, 2385–2403. <https://doi.org/10.1007/s11269-018-1935-3>.
- ArchaeoGLOBE Project, 2019. Archaeological assessment reveals Earth's early transformation through land use. *Science* 365, 897–902. <https://doi.org/10.1126/science.aax1192>.
- Armenteras, D., Rodríguez, N., Retana, J., Morales, M., 2011. Understanding deforestation in montane and lowland forests of the Colombian Andes. *Reg. Environ. Change* 11, 693–705. <https://doi.org/10.1007/s10113-010-0200-y>.
- Armenteras, D., Schneider, L., Dávalos, L.M., 2019. Fires in protected areas reveal unforeseen costs of Colombian peace. *Nat. Ecol. Evol.* 3, 20–23. <https://doi.org/10.1038/s41559-018-0727-8>.
- Arneth, A., Denton, F., Agus, F., Elbehri, A., Erb, K., Osman Elasha, B., Rahimi, M., Rounsevell, M., Spence, A., Valentini, R., 2019. Framing and Context, Climate Change and Land: an IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems. <https://doi.org/10.1017/9781009157988.003>.
- Baisero, D., Visconti, P., Pacifici, M., Cimatti, M., Rondinini, C., 2020. Projected global loss of mammal habitat due to land-use and climate change. *One Earth* 2, 578–585. <https://doi.org/10.1016/j.oneear.2020.05.015>.
- Beygelzimer, A., Kakadet, S., Langford, J., Arya, S., Mount, D., Li, S., 2022. FNN: Fast Nearest Neighbor Search Algorithms and Applications. R package version 1.1.3.1. <https://cran.r-project.org/package=FNN>.
- Blanchard, S.D., Pontius Jr., R.G., Urban, K.M., 2015. Implications of using 2 m versus 30 m spatial resolution data for suburban residential land change modeling. *J. Environ. Inform.* 25 <https://doi.org/10.3808/jei.201400284>.

- Bocedi, G., Pe'er, G., Heikkinen, R.K., Matsinos, Y., Travis, J.M.J., 2012. Projecting species' range expansion dynamics: sources of systematic biases when scaling up patterns and processes. *Methods Ecol. Evol.* 3, 1008–1018. <https://doi.org/10.1111/j.2041-210X.2012.00235.x>.
- Borrelli, P., Robinson, D.A., Panagos, P., Lugato, E., Yang, J.E., Alewell, C., Wuepper, D., Montanarella, L., Ballabio, C., 2020. Land use and climate change impacts on global soil erosion by water (2015–2070). *Proc. Natl. Acad. Sci. U.S.A.* 117, 21994–22001. <https://doi.org/10.1073/pnas.2001403117>.
- Cao, M., Zhu, Y., Quan, J., Zhou, S., Lü, G., Chen, M., Huang, M., 2019. Spatial sequential modeling and predication of global land use and land cover changes by integrating a global change assessment model and cellular automata. *Earth's Future* 7, 1102–1116. <https://doi.org/10.1029/2019EF001228>.
- Chen, M., Vernon, C.R., Huang, M., Calvin, K.V., Kraucunas, I.P., 2019. Calibration and analysis of the uncertainty in downscaling global land use and land cover projections from GCAM using Demeter (v1.0.0). *Geosci. Model Dev. (GMD)* 12, 1753–1764. <https://doi.org/10.5194/gmd-12-1753-2019>.
- Chen, M., Vernon, C.R., Graham, N.T., Hejazi, M., Huang, M., Cheng, Y., Calvin, K., 2020. Global land use for 2015–2100 at 0.05° resolution under diverse socioeconomic and climate scenarios. *Sci. Data* 7, 320. <https://doi.org/10.1038/s41597-020-00669-x>.
- Colombia, SiB., 2022. Biodiversidad en Cifras, Sistema de Información sobre Biodiversidad de Colombia [WWW Document]. URL <https://biodiversidad.co/cifras>, 12.22.22.
- Correa Ayram, C.A., Etter, A., Díaz-Timoté, J., Rodríguez Buritica, S., Ramírez, W., Corzo, G., 2020. Spatiotemporal evaluation of the human footprint in Colombia: four decades of anthropic impact in highly biodiverse ecosystems. *Ecol. Indic.* 117, 106630. <https://doi.org/10.1016/j.ecolind.2020.106630>.
- de Chazal, J., Rounsevell, M.D.A., 2009. Land-use and climate change within assessments of biodiversity change: a review. *Global Environ. Change* 19, 306–315. <https://doi.org/10.1016/j.gloenvcha.2008.09.007>.
- Dendoncker, N., Bogaert, P., Rounsevell, M., 2006. A statistical method to downscale aggregated land use data and scenarios. *J. Land Use Sci.* 1, 63–82. <https://doi.org/10.1080/17474230601058302>.
- Dendoncker, N., Rounsevell, M., Bogaert, P., 2007. Spatial analysis and modelling of land use distributions in Belgium. *Comput. Environ. Urban Syst.* 31, 188–205. <https://doi.org/10.1016/j.compenvurbsys.2006.06.004>.
- Dezhkam, S., Jabbarian Amiri, B., Darvishsefat, A.A., Sakieh, Y., 2016. Performance evaluation of land change simulation models using landscape metrics. *Geocarto Int.* 1–23. <https://doi.org/10.1080/10106049.2016.1167967>.
- Di Marco, M., Harwood, T.D., Hoskins, A.J., Ware, C., Hill, S.L.L., Ferrier, S., 2019. Projecting impacts of global climate and land-use scenarios on plant biodiversity using compositional-turnover modelling. *Global Change Biol.* 25, 2763–2778. <https://doi.org/10.1111/gcb.14663>.
- Dullinger, I., Gattlinger, A., Wessely, J., Moser, D., Plutzer, C., Willner, W., Egger, C., Gaube, V., Haberl, H., Mayer, A., Bohner, A., Gilli, C., Pascher, K., Essl, F., Dullinger, S., 2020. A socio-ecological model for predicting impacts of land-use and climate change on regional plant diversity in the Austrian Alps. *Global Change Biol.* 26, 2336–2352. <https://doi.org/10.1111/gcb.14977>.
- Erb, K.H., Luysaert, S., Meyfroidt, P., Pongratz, J., Don, A., Kloster, S., Kuemmerle, T., Fetzel, T., Fuchs, R., Herold, M., Haberl, H., Jones, C.D., Marin-Spiotta, E., McCallum, I., Robertson, E., Seufert, V., Fritz, S., Valade, A., Wiltshire, A., Dolman, A.J., 2017. Land management: data availability and process understanding for global change studies. *Global Change Biol.* 23, 512–533. <https://doi.org/10.1111/gcb.13443>.
- Etter, A., McAlpine, C., Possingham, H., 2008. Historical patterns and drivers of landscape change in Colombia since 1500: a regionalized spatial approach. *Ann. Assoc. Am. Geogr.* 98, 2–23. <https://doi.org/10.1080/00045600701733911>.
- Fischer, J., Lindenmayer, D.B., 2007. Landscape modification and habitat fragmentation: a synthesis. *Global Ecol. Biogeogr.* 16, 265–280. <https://doi.org/10.1111/j.1466-8238.2007.00287.x>.
- Foley, J.A., Ramankutty, N., Brauman, K.A., Cassidy, E.S., Gerber, J.S., Johnston, M., Mueller, N.D., O'Connell, C., Ray, D.K., West, P.C., Balzer, C., Bennett, E.M., Carpenter, S.R., Hill, J., Monfreda, C., Polasky, S., Rockström, J., Sheehan, J., Siebert, S., Tilman, D., Zaks, D.P.M., 2011. Solutions for a cultivated planet. *Nature* 478, 337–342. <https://doi.org/10.1038/nature10452>.
- Forzieri, G., Alkama, R., Miralles, D.G., Cescatti, A., 2017. Satellites reveal contrasting responses of regional climate to the widespread greening of Earth. *Science* 356, 1180–1184. <https://doi.org/10.1126/science.aal1727>.
- Genzenmüller, R., Sylvester, J.M., Castro-Nunez, A., 2022. What peace means for deforestation: an analysis of local deforestation dynamics in times of conflict and peace in Colombia. *Front. Environ. Sci.* 10, 803368. <https://doi.org/10.3389/fevs.2022.803368>.
- García Molinos, J., Schoeman, D.S., Brown, C.J., Burrows, M.T., 2019. VoCC: an R package for calculating the velocity of climate change and related climatic metrics. *Methods Ecol. Evol.* 10, 2195–2202. <https://doi.org/10.1111/2041-210X.13295>.
- González-González, A., Clerici, N., Quesada, B., 2021. Growing mining contribution to Colombian deforestation. *Environ. Res. Lett.* 16, 064046. <https://doi.org/10.1088/1748-9326/abcf8>.
- Guo, Z., Torra, O., Hürlimann, M., Abancó, C., Medina, V., 2022. FSLAM: a QGIS plugin for fast regional susceptibility assessment of rainfall-induced landslides. *Environ. Model. Software* 150, 105354. <https://doi.org/10.1016/j.envsoft.2022.105354>.
- Guo, Z., Ferrer, J.V., Hürlimann, M., Medina, V., Puig-Polo, C., Yin, K., Huang, D., 2023a. Shallow landslide susceptibility assessment under future climate and land cover changes: a case study from southwest China. *Geosci. Front.* 14, 101542. <https://doi.org/10.1016/j.gsf.2023.101542>.
- Guo, Z., Tian, B., Li, G., Huang, D., Zeng, T., He, J., Song, D., 2023b. Landslide susceptibility mapping in the Loess Plateau of northwest China using three data-driven techniques—a case study from middle Yellow River catchment. *Front. Earth Sci.* 10, 1033085. <https://doi.org/10.3389/feart.2022.1033085>.
- Hasegawa, T., Fujimori, S., Ito, A., Takahashi, K., Masui, T., 2017. Global land-use allocation model linked to an integrated assessment model. *Sci. Total Environ.* 580, 787–796. <https://doi.org/10.1016/j.scitotenv.2016.12.025>.
- Heistermann, M., Müller, C., Ronneberger, K., 2006. Land in sight? Achievements, deficits and potentials of continental to global scale land-use modeling. *Agric. Ecosyst. Environ.* 114, 141–158. <https://doi.org/10.1016/j.agee.2005.11.015>.
- Henry, R.C., Arneeth, A., Jung, M., Rabin, S.S., Rounsevell, M.D., Warren, F., Alexander, P., 2022. Global and regional health and food security under strict conservation scenarios. *Nat. Sustain.* 5, 303–310. <https://doi.org/10.1038/s41893-021-00844-x>.
- Hesselbarth, M.H.K., Sciaini, M., With, K.A., Wiegand, K., Nowosad, J., 2019. landscapemetrics: an open-source R tool to calculate landscape metrics. *Ecography* 42, 1648–1657. <https://doi.org/10.1111/ecog.04617>.
- Hijmans, R., 2022. Terra: Spatial Data Analysis. R Package Version 1.5-21 and Version 1.6-17. <https://rspace.org/terra/>.
- Hu, X., Liu, F., Qi, Y., Zhang, J., Li, X., 2022. A novel strategy to assimilate category variables in land-use models based on Dirichlet distribution. *Environ. Model. Software* 149, 105324. <https://doi.org/10.1016/j.envsoft.2022.105324>.
- Hull, V., Liu, J., 2018. Telecoupling: a new frontier for global sustainability. *Ecol. Soc.* 23, 41. <https://doi.org/10.5751/ES-10494-230441>.
- Jaureguiberry, P., Titeux, N., Wiemers, M., Bowler, D.E., Coscieme, L., Golden, A.S., Guerra, C.A., Jacob, U., Takahashi, Y., Settle, J., Díaz, S., Molnár, Z., Purvis, A., 2022. The direct drivers of recent global anthropogenic biodiversity loss. *Sci. Adv.* 2022, eabm9982. <https://doi.org/10.1126/sciadv.abm9982>.
- Jia, G., Shevliakova, E., Artaxo, P., De Noblet-Ducoudré, N., Houghton, R., House, J., Kitajima, K., Lennard, C., Popp, A., Sirin, A., Sukumar, R., Verchot, L., 2019. Land-climate interactions. In: Shukla, P.R., Skea, J., Calvo Buendia, E., Masson-Delmotte, V., Pörtner, H.-O., Roberts, D.C., Zhai, P., Slade, R., Connors, S., van Diemen, R., Ferrat, M., Haughey, E., Luz, S., Neogi, S., Pathak, M., Petzold, J., Portugal Pereira, J., Vyas, P., Huntley, E., Kissick, K., Belkacemi, M., Malley, J. (Eds.), *Climate Change and Land: an IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems*, pp. 131–248. <https://doi.org/10.1017/9781009157988.004>.
- Johnson, J.A., Ruta, G., Baldos, U., Cervigni, R., Chonabayashi, S., Corong, E., Gavrilyuk, O., Gerber, J., Hertel, T., Nootenboom, C., Polasky, S., 2021. *The Economic Case for Nature. The World Bank*.
- Kapitza, S., Golding, N., Wintle, B.A., 2022. A fractional land use change model for ecological applications. *Environ. Model. Software* 147, 105258. <https://doi.org/10.1016/j.envsoft.2021.105258>.
- Le Page, Y., West, T.O., Link, R., Patel, P., 2016. Downscaling land use and land cover from the Global Change Assessment Model for coupling with Earth system models. *Geosci. Model Dev. (GMD)* 9, 3055–3069. <https://doi.org/10.5194/gmd-9-3055-2016>.
- Li, X., Chen, G., Liu, X., Liang, X., Wang, S., Chen, Y., Pei, F., Xu, X., 2017. A new global land-use and land-cover change product at a 1-km resolution for 2010 to 2100 based on human–environment interactions. *Ann. Assoc. Am. Geogr.* 107, 1040–1059. <https://doi.org/10.1080/24694452.2017.1303357>.
- Lin, J., Li, X., Li, S., Wen, Y., 2020. What is the influence of landscape metric selection on the calibration of land-use/cover simulation models? *Environ. Model. Software* 129, 104719. <https://doi.org/10.1016/j.envsoft.2020.104719>.
- Liu, X., Li, S., Wang, S., Bian, Z., Zhou, W., Wang, C., 2022. Effects of farmland landscape pattern on spatial distribution of soil organic carbon in Lower Liaohe Plain of northeastern China. *Ecol. Indic.* 145, 109652. <https://doi.org/10.1016/j.ecolind.2022.109652>.
- Luysaert, S., Jammot, M., Stoy, P.C., Estel, S., Pongratz, J., Ceschia, E., Churkina, G., Don, A., Erb, K., Ferlicq, M., Gielen, B., Grünwald, T., Houghton, R.A., Klumpp, K., Knohl, A., Kolb, T., Kuemmerle, T., Laurila, T., Lohila, A., Loustau, D., McGrath, M. J., Meyfroidt, P., Moors, E.J., Naudts, K., Novick, K., Otto, J., Pilegaard, K., Pio, C.A., Rambal, S., Reibmann, C., Ryder, J., Suyker, A.E., Varlagin, A., Wattenbach, M., Dolman, A.J., 2014. Land management and land-cover change have impacts of similar magnitude on surface temperature. *Nat. Clim. Change* 4, 389–393. <https://doi.org/10.1038/nclimate2196>.
- Malchow, A.K., Bocedi, G., Palmer, S.C.F., Travis, J.M.J., Zurell, D., 2021. RangeShiftR: an R package for individual-based simulation of spatial eco-evolutionary dynamics and species' responses to environmental changes. *Ecography* 44, 1443–1452. <https://doi.org/10.1111/ecog.05689>.
- Mas, J.F., García-Álvarez, D., Paegelow, M., Domínguez-Vera, R., Castillo-Santiago, M.Á., 2022. Metrics based on a cross-tabulation matrix to validate land use cover maps. In: García-Álvarez, D., Camacho Olmedo, M.T., Paegelow, M., Mas, J.F. (Eds.), *Land Use Cover Datasets and Validation Tools: Validation Practices with QGIS*. Springer International Publishing, Cham, pp. 127–151. <https://doi.org/10.1007/978-3-030-90998-7>.
- Meier, R., Schwaab, J., Seneviratne, S.I., Sprenger, M., Lewis, E., Davin, E.L., 2021. Empirical estimate of forestation-induced precipitation changes in Europe. *Nat. Geosci.* 14, 473–478. <https://doi.org/10.1038/s41561-021-00773-6>.
- Millington, J.D.A., Katerinchuk, V., da Silva, R.F.B., de Castro Victoria, D., Batistella, M., 2021. Modelling drivers of Brazilian agricultural change in a telecoupled world. *Environ. Model. Software* 139, 105024. <https://doi.org/10.1016/j.envsoft.2021.105024>.
- Molotoks, A., Henry, R., Stehfest, E., Doelman, J., Havlik, P., Kriztint, T., Alexander, P., Dawson, T.P., Smith, P., 2020. Comparing the impact of future cropland expansion on global biodiversity and carbon storage across models and scenarios. *Philos. Trans. R. Soc. B: Biol. Sci.* 375, 20190189. <https://doi.org/10.1098/rstb.2019.0189>.

- Murillo-Sandoval, P.J., Gjerdseth, E., Correa-Ayram, C., Wrathall, D., Van Den Hoek, J., Dávalos, L.M., Kennedy, R., 2021. No peace for the forest: rapid, widespread land changes in the Andes-Amazon region following the Colombian civil war. *Global Environ. Change* 69, 102283. <https://doi.org/10.1016/j.gloenvcha.2021.102283>.
- Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots for conservation priorities. *Nature* 403, 853–858. <https://doi.org/10.1038/35002501>.
- Newbold, T., Hudson, L.N., Hill, S.L.L., Contu, S., Lysenko, I., Senior, R.A., Börger, L., Bennett, D.J., Choimes, A., Collen, B., Day, J., De Palma, A., Díaz, S., Echeverría-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., Correia, D.L.P., 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.W., Purvis, A., 2015. Global effects of land use on local terrestrial biodiversity. *Nature* 520, 45–50. <https://doi.org/10.1038/nature14324>.
- O'Neill, B.C., Krieger, E., Riahi, K., Ebi, K.L., Hallegatte, S., Carter, T.R., Mathur, R., van Vuuren, D.P., 2014. A new scenario framework for climate change research: the concept of shared socioeconomic pathways. *Clim. Change* 122, 387–400. <https://doi.org/10.1007/s10584-013-0905-2>.
- Paegelow, M., Mas, J.-F., Gallardo, M., Camacho Olmedo, M.T., García-Álvarez, D., 2022. Pontius Jr. Methods based on a cross-tabulation matrix to validate land use cover maps. In: *Land Use Cover Datasets and Validation Tools: Validation Practices with QGIS*. Springer International Publishing, Cham, pp. 153–187. <https://doi.org/10.1007/978-3-030-90998-7>.
- Pais, C., Miranda, A., Carrasco, J., Shen, Z.-J.M., 2021. Deep fire topology: understanding the role of landscape spatial patterns in wildfire occurrence using artificial intelligence. *Environ. Model. Software* 143, 105122. <https://doi.org/10.1016/j.envsoft.2021.105122>.
- Pereira, H.M., Navarro, L.M., Martins, I.S., 2012. Global biodiversity change: the bad, the good, and the unknown. *Annu. Rev. Environ. Resour.* 37, 25–50. <https://doi.org/10.1146/annurev-environ-042911-093511>.
- Pontius, R.G., Boersma, W., Castella, J.-C., Clarke, K., de Nijs, T., Dietzel, C., Duan, Z., Fotsing, E., Goldstein, N., Kok, K., Koomen, E., Lippitt, C.D., McConnell, W., Mohd Sood, A., Pijanowski, B., Pithadia, S., Sweeney, S., Trung, T.N., Veldkamp, A.T., Verburg, P.H., 2008. Comparing the input, output, and validation maps for several models of land change. *Ann. Reg. Sci.* 42, 11–37. <https://doi.org/10.1007/s00168-007-0138-2>.
- Popp, A., Calvin, K., Fujimori, S., Havlik, P., Humpenöder, F., Stehfest, E., Bodirsky, B.L., Dietrich, J.P., Doelmann, J.C., Gusti, M., Hasegawa, T., Kyle, P., Obersteiner, M., Tabeau, A., Takahashi, K., Valin, H., Waldhoff, S., Weindl, I., Wise, M., Krieger, E., Lotze-Campen, H., Fricko, O., Riahi, K., van Vuuren, D.P., 2017. Land-use futures in the shared socio-economic pathways. *Global Environ. Change* 42, 331–345. <https://doi.org/10.1016/j.gloenvcha.2016.10.002>.
- R Core Team, 2022. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.
- Rabin, S.S., Alexander, P., Henry, R., Anthoni, P., Pugh, T.A.M., Rounsevell, M., Arneth, A., 2020. Impacts of future agricultural change on ecosystem service indicators. *Earth Syst. Dyn.* 11, 357–376. <https://doi.org/10.5194/esd-11-357-2020>.
- Rigby, A.M.F., Butcher, P.W.S., Ritsos, P.D., Patil, S.D., 2022. LUCST: a novel toolkit for Land Use Land Cover change assessment in SWAT+ to support flood management decisions. *Environ. Model. Software* 156, 105469. <https://doi.org/10.1016/j.envsoft.2022.105469>.
- Romero-Ruiz, M., Etter, A., Sarmiento, A., Tansey, K., 2010. Spatial and temporal variability of fires in relation to ecosystems, land tenure and rainfall in savannas of northern South America. *Global Change Biol.* 16, 2013–2023. <https://doi.org/10.1111/j.1365-2486.2009.02081.x>.
- Romero-Ruiz, M.H., Flantua, S.G.A., Tansey, K., Berrio, J.C., 2012. Landscape transformations in savannas of northern South America: land use/cover changes since 1987 in the Llanos Orientales of Colombia. *Appl. Geogr.* 32, 766–776. <https://doi.org/10.1016/j.apgeog.2011.08.010>.
- Ryu, S.-R., Chen, J., Zheng, D., Lacroix, J.J., 2007. Relating surface fire spread to landscape structure: an application of FARSITE in a managed forest landscape. *Landsc. Urban Plann.* 83, 275–283. <https://doi.org/10.1016/j.landurbplan.2007.05.002>.
- Salehpour Jam, A., Mosaffaie, J., Tabatabaei, M.R., 2023. Raster-based landslide susceptibility mapping using compensatory MADM methods. *Environ. Model. Software* 159, 105567. <https://doi.org/10.1016/j.envsoft.2022.105567>.
- South, A., 2017. Rnaturalearth: world map data from natural Earth. R package version 0.1.0. <https://cran.r-project.org/package=rnaturalearth>.
- Suh, S., Johnson, J.A., Tambjerg, L., Sim, S., Broeckx-Smith, S., Reyes, W., Chaplin-Kramer, R., 2020. Closing yield gap is crucial to avoid potential surge in global carbon emissions. *Global Environ. Change* 63, 102100. <https://doi.org/10.1016/j.gloenvcha.2020.102100>.
- Tilman, D., Balzer, C., Hill, J., Befort, B.L., 2011. Global food demand and the sustainable intensification of agriculture. *Proc. Natl. Acad. Sci. U.S.A.* 108, 20260–20264. <https://doi.org/10.1073/pnas.1116437108>.
- Titeux, N., Henle, K., Mihoub, J.B., Regos, A., Geijzendorffer, I.R., Cramer, W., Verburg, P.H., Brotons, L., 2016. Biodiversity scenarios neglect future land-use changes. *Global Change Biol.* 22, 2505–2515. <https://doi.org/10.1111/gcb.13272>.
- van Asselen, S., Verburg, P.H., 2013. Land cover change or land-use intensification: simulating land system change with a global-scale land change model. *Global Change Biol.* 19, 3648–3667. <https://doi.org/10.1111/gcb.12331>.
- Van Dexter, K., Visseren-Hamakers, I., 2020. Forests in the time of peace. *J. Land Use Sci.* 15, 327–342. <https://doi.org/10.1080/1747423X.2019.1699614>.
- van Vliet, J., Naus, N., van Lammeren, R.J.A., Bregt, A.K., Hurkens, J., van Delden, H., 2013. Measuring the neighbourhood effect to calibrate land use models. *Comput. Environ. Urban Syst.* 41, 55–64. <https://doi.org/10.1016/j.compenvurbysys.2013.03.006>.
- van Vuuren, D.P., Smith, S.J., Riahi, K., 2010. Downscaling socioeconomic and emissions scenarios for global environmental change research: a review. *Wiley Interdiscip. Rev. Clim. Change* 1, 393–404. <https://doi.org/10.1002/wcc.50>.
- Veldkamp, A., Lambin, E.F., 2001. Predicting land-use change. *Agric. Ecosyst. Environ.* 85, 1–6. [https://doi.org/10.1016/S0167-8809\(01\)00199-2](https://doi.org/10.1016/S0167-8809(01)00199-2).
- Vernon, C.R., Le Page, Y., Chen, M., Huang, M., Calvin, K.V., Kraucunas, I.P., Braun, C.J., 2018. Demeter - a land use and land cover change disaggregation model. *J. Open Res. Software* 6, 15. <https://doi.org/10.5334/jors.208>.
- von Jeetze, P.J., Weindl, I., Johnson, J.A., Borrelli, P., Panagos, P., Molina Bacca, E.J., Karstens, K., Humpenöder, F., Dietrich, J.P., Minoli, S., Müller, C., Lotze-Campen, H., Popp, A., 2023. Projected landscape-scale repercussions of global action for climate and biodiversity protection. *Nat. Commun.* 14, 2515. <https://doi.org/10.1038/s41467-023-38043-1>.
- Wang, X., Blanchet, F.G., Koper, N., 2014. Measuring habitat fragmentation: an evaluation of landscape pattern metrics. *Methods Ecol. Evol.* 5, 634–646. <https://doi.org/10.1111/2041-210X.12198>.
- West, T.O., Le Page, Y., Huang, M., Wolf, J., Thomson, A.M., 2014. Downscaling global land cover projections from an integrated assessment model for use in regional analyses: results and evaluation for the US from 2005 to 2095. *Environ. Res. Lett.* 9, 064004. <https://doi.org/10.1088/1748-9326/9/6/064004>.
- Winkler, K., Fuchs, R., Rounsevell, M.D.A., Herold, M., 2020. HILDA+ Global Land Use Change between 1960 and 2019. <https://doi.org/10.1594/PANGAEA.921846>.
- Winkler, K., Fuchs, R., Rounsevell, M., Herold, M., 2021. Global land use changes are four times greater than previously estimated. *Nat. Commun.* 12, 1–10. <https://doi.org/10.1038/s41467-021-22702-2>.