Modelling drivers of Brazilian agricultural change in a telecoupled world

James D.A. Millington[1], Valeri Katerinchuk[1], Ramon Felipe Bicudo da Silva[2,3], Daniel de Castro Victoria[4], Mateus Batistella[3,4]

1. Department of Geography, King's College London, London, UK
2. Center for Systems Integration and Sustainability, Department of Fisheries and Wildlife, Michigan State University, United States of America
3. Center for Environmental Studies and Research, State University of Campinas (NEPAM/UNICAMP), Campinas, Brazil
4. Embrapa Agricultural Informatics. Brazilian Agricultural Research Corporation (EMBRAPA), Campinas, Brazil

Corresponding Author: James Millington, james.millington@kcl.ac.uk

Abstract

Increasing global demand for agricultural commodities has driven local land use/cover change (LUCC) and agricultural production across Brazil during the 21st Century. Tools to understand the range of possible outcomes due to these 'telecoupled' global-to-local relationships are needed, given future political, economic and environmental uncertainties. We present progress on the development of one such tool; a spatially-explicit LUCC model representing production of multiple agricultural commodities, structured to be linked to a System Dynamics model of global trade. The model accounts for spatially explicit (e.g., land access) and temporally contingent (e.g., agricultural debt) processes of importance across our nearly four million km^2 Brazilian study area. We present model calibration and evaluation for 2001-2018, before examining scenarios of commodity demand, agricultural yields, climate change, and policy decisions for 2018-2035. Results indicate greater confidence in modelled time-series than spatial allocation and provide insights into appropriate approaches to modelling telecoupled systems.

Keywords

land use/cover; telecoupling; Brazil; CRAFTY; simulation

Software and Data Availability

Code for both the simulation model and our data analysis is freely available online; we refer to the relevant GitHub repositories in the text at the appropriate points. The model can be deployed via Docker using Lane and Millington (2021). Also see Victoria *et al*. (2021).

1. Introduction

It is now well understood that local land use/cover changes in many regions of the world are influenced by international demand for agricultural commodities and that socio-ecological systems are now often 'telecoupled' over great distances (Liu *et al.* 2013, 2018). For example, increased Chinese demand for soybean over the last several decades has contributed to increased soybean production in Brazil, making the country a key production region in the global food system and driving change in local land use (Silva *et al.* 2017, Sun *et al.* 2017). Increased production has been achieved through a combination of expansion of agricultural land, increases in yields and changes in farming practices, including the development of a double-crop system with maize (predominantly) as a second crop. Improvements in yields have come through improved seed varieties (including genetic modification), increases in agricultural inputs (fertilisers, pesticides, machinery) and economies of scale (Wesz 2016). These changes have come in tandem with significant economic changes in the Brazilian farming system, meaning that many farmers are faced by tough economic decisions to ensure the future viability of their businesses (Silva *et al.* 2020). Future uncertainty is further exacerbated by the spectre of climate change which may bring increased frequency of drought during the second crop and other conditions unfavourable to consistent production from year-to-year (Heinemann *et al.* 2017, Hampf *et al.* 2020).

To create a tool for examining the range of possible outcomes given such a range of drivers and uncertainties, we set out to develop a spatially-explicit land use/cover change model capable of representing both production and associated land use of multiple agricultural commodities that could be subsequently linked to a system dynamics model of global trade (see Millington *et al.* 2017). The Telecoupling framework within which we developed our model emphasises the importance of agents and flows as drivers of change in coupled human-natural systems that are linked across long distances. To represent agency in land use and agricultural production decision-making we use the previously developed Competition for Resources between Agent Functional Types (CRAFTY) modelling framework (Murray-Rust *et al.* 2014), adapting it to improve representation of spatially explicit (e.g., land access) and temporally contingent (e.g., agricultural debt) processes of importance in our Brazilian study area. The CRAFTY framework has been designed specifically with the intention of simulating broad-scale land use change over large spatial extents (national to continental). For example, Blanco *et al.* (2017) parameterised CRAFTY to examine ecosystem services and decision-making under scenarios of climate change for the entire land area of Sweden over many decades. Brown *et al.* (2019) used CRAFTY to simulate land use change across the entire European Union to investigate land manager behaviour at the continental scale. In work similar to the one we report on here, investigating the telecoupled effects of global food commodity trade between China and Brazil on land use, Dou *et al.* (2019, 2020) developed a bespoke agent-based model. Whereas the aims of that approach were to understand land use impacts at a fine scale for a single municipality, our work aims to understand land use across much broader extents and hence the use of CRAFTY is more appropriate (Millington *et al.* 2017).

In this paper we present the description and first results from the application of the CRAFTY framework to simulate land use/cover change over several decades for ten Brazilian states, an implementation we call CRAFTY-Brazil. We provide an overview of the endogenous and exogenous processes represented, the data used to parameterize and calibrate the model, and results from using CRAFTY-Brazil to simulate scenarios of future change. Importantly, we developed and tested CRAFTY-Brazil using empirically-grounded data in a fashion that has not previously been achieved for other applications of the CRAFTY framework. This empirically-grounded approach aims to ensure internal model consistency, but also identify key areas of uncertainty. The model structure and results presented here are independent of the System Dynamics model (based on Warner *et al.* 2013) we ultimately intend to couple CRAFTY-Brazil with. However, by examining counterfactual scenarios of commodity demand, agricultural yields, climate change, and policy decisions over land use rights, we are able to both identify important uncertainties in the model but also shed light on important processes influencing land use change in Brazil. Subsequently, we reflect on what we have learned from this initial use of the model and discuss future directions in which this work should proceed.

2. Methods

2.1 Study Area and Data

CRAFTY-Brazil was designed with the intention of subsequently connecting this spatial explicit model of land use/cover change with a System Dynamics model representing the international trade of three primary agricultural commodities: soybean, maize and beef. With this focus, our study area was designated as the 10 states in Brazil that have been the dominant producers of soybean and maize in recent years and for which pasture area is widespread (Figure 1). Our aim to simulate this region of 3,850,000 km^2 for several decades required compromises on model spatial and temporal resolution to ensure feasible execution times while representing sufficient variation to explore system dynamics. Thus, the model operates on a raster grid at a 5 km spatial resolution (for a total of 162,026 simulated cells) and we aggregate results to municipality level (for our study area, mean and median municipality areas are 1,040 and 398 km^2 respectively). These resolutions are appropriate given the available data needed to calibrate the model, which comes from a variety of sources and with a range of original resolutions including many aggregated at municipality level (Table 1). Furthermore, the 5km resolution is comparable to other applications of CRAFTY across large spatial extents (e.g., 1 km for Blanco *et al*. 2017 and 10 arcminute - approx. 18 km at the equator - for Brown *et al*. 2019). The model runs with an annual timestep and we calibrate using data for 2001-2018.

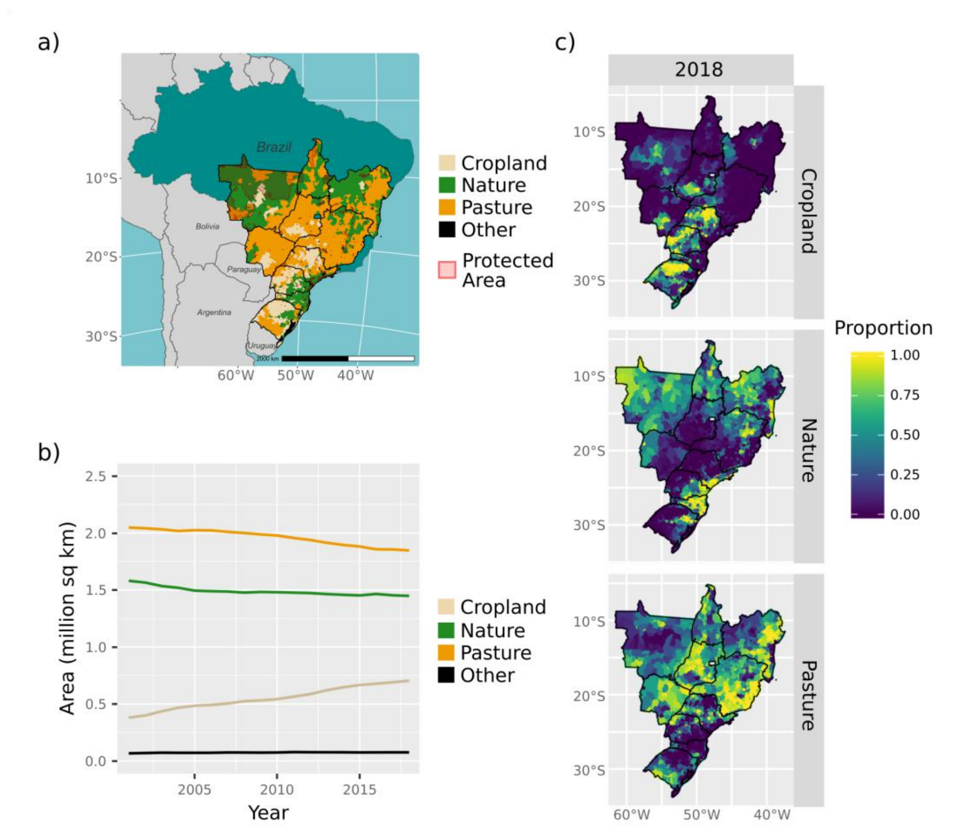

Figure 1. Study area. a) Location of the ten Brazilian states composing the study area and their modal municipality land cover for 2018, b) Study area LUCC through time for the 2001-2018 calibration period, c) Municipality proportion of Cropland, Nature and Pasture for 2018 (scenario results are compared against these proportions).

Table 1. Data used for model calibration.

Variable	*Type*	*Spatial Res.*	*Temporal Res.*	*Source*	*Use*
Land Cover/Use	Raster	30 m	Annual	MapBiomas (2019)	Model initialization and calibration
Climate	Raster	0.5°	Annual	Harris (2020)	Moisture Capitals
Transport Network	Vector	NA	Quinquennial	DNIT (2019), ANTAQ (2019)	Transport Capital
Protected Areas	Vector	NA	Annual	MMA (2019)	Protection Capitals
Land Price	Vector	Municipality	Annual	IEG/FNP (2017)	Land Value Capital
Commodity Production	Statistical	Municipality	Annual	IBGE (2019)	Study area selection, Tech-Soy-Maize and Tech-Pasture Capitals
Commodity Planted Area	Statistical	Municipality	Annual	IBGE (2019)	Land cover/use map disaggregation
Commodity Exports	Statistical	Municipality	Annual	IBGE (2019)	Commodity demand estimation

For model initialisation and calibration, we use land use/cover (LUC) data from the MapBiomas project (version 4; MapBiomas 2018). Original 30m data were resampled to 5km (modal pixel class; see Millington 2019) and LUC classes were reclassified from 27 classes to four: ‘Cropland’, ‘Pasture’, ‘Nature’ and ‘Other’ (Appendix A). These LUC classes are associated with the production of ‘services’ by different Agent Functional Types (as described below, Section 2.2.1). Cropland represents land used to produce soybean, maize and other crops (e.g., rice, sugarcane), while Pasture is assumed to represent all land area used for beef production. Nature represents vegetation cover not used for agriculture (i.e., crops or pasture); processes of natural vegetation dynamics are not explicitly represented (although see use of Nature capital in section 2.2.1 below) but simulated land that is no longer needed for human use reverts to the Nature class (and Nature cells can be converted to all other classes). Finally, Other represents land covers such as urban and inland water bodies and is assumed to not change in simulations that project into the future.

Given uncertainty in the MapBiomas data and the multiple alternative ways those data might be re-classified, we examined several re-classifications of the MapBiomas v4.0 data and used commodity planted area data (from IBGE) to disaggregate ambiguous classes. The overall accuracy for the MapBiomas v4.0 data for the period 2001-2018 (at the Level 2 classification) is estimated to be 88%, with the Grassland class consistently the most poorly classified of all classes (high rates of Natural Forest and Pasture commission error; see MapBiomas 2019b). On comparing the implied beef pasture yields for LUC classifications including/excluding Grassland in our Pasture class, we find a closer match with observed yields when Grassland is included (see Millington 2019). The MapBiomas class 'Mosaic of Agriculture and Pasture' is also problematic given that our focus is on distinguishing between pasture and cropland. To address this, we used planted-area data (IBGE 2019) to allocate pixels in the MapBiomas 'Mosaic of Agriculture and Pasture' class into either Cropland or Pasture classes (see Millington 2019).

2.2 Model Description

2.2.1 Services, Agents and Capitals

The CRAFTY framework represents the production of land 'services' by agent functional types (AFTs) through production functions for multiple 'capitals' available in each cell representing a land unit (Murray-Rust *et al.* 2014). Cobb Douglas production functions are used:

$$p_s = \prod_c c_i^{\lambda_{c,a}} \tag{1}$$

where $\lambda_{c,a}$ is a weighting factor specific to capital *c* and agent-type *a*, and p_s is the productivity for service *s* (in abstract 'production units', later converted to kg for agricultural commodity services). Competition between agents is represented by calculating 'competitiveness' using a utility function that accounts for production and marginal demand for each service. In each timestep, the competitiveness of different AFTs is calculated for cells based on current capital values (which CRAFTY assumes are scaled 0 to 1) and service demands (specified in abstract 'production units', which in our case represent demand for agricultural commodities in domestic and international markets); the AFT with the greatest competitiveness is allocated to that cell (subject to the 'giving-in' threshold of the currently occupying agent). Land may also be abandoned from agent use if competitiveness falls below an AFT's 'giving-up' threshold. For full details on the CRAFTY framework, see Murray-Rust *et al.* (2014).

Given our focus on soybean, maize and beef, CRAFTY-Brazil represents provision of the following 'services': Soybean, Maize, Beef, Other Crops (OCrops), Nature and Other. The Nature service represents the value of land not under human influence, while Other represents all other land not in other classes (e.g., urban, water). The AFTs represented are soybean producers, maize producers, soybean-maize double-crop producers (producing both services), beef producers, other crop producers, and other land managers (e.g., to represent urban land cover). A Nature AFT is also required for cells not under control of any other AFT and representing land not under direct human land use (regardless of whether that is primary or secondary vegetation). Representing double-crop farmers as an individual AFT is important as in recent years land managers have begun implementing the practice of sowing and planting a soybean crop early in the season followed by a late season crop (usually maize). The double-cropping system of soybean and maize was initiated in Brazil in the late 1990s, as raising maize following soybean harvest provides protective cover for soils and the system eventually improves soil quality (Silva *et al.*, 2017). Initial yields using this method were low, but since the early 2000s many improvements have been made in management practices such as no-tillage agriculture, Nitrogen biological fixation serviced by soybeans, hybrid and GMO maize varieties, which boosted yields for the second crop. Consequently, this double-cropping practice has grown through the 21st century, mainly pushed by economic demand than by a desire for improved soil management (Silva *et al.*, 2017). Although this system provides potential for both agronomic and economic gains over single-crop systems, the second crop growth is exposed to the risk of late-season drought and consequent production losses (Brunini *et al.*, 2001; Gonçalves *et al.*, 2002). For the soybean, the double-cropping system has pushed its production into a very short growing season forcing producers to adopt short-cycle GMO varieties (about 90 days from planting to harvest) that are impacting soybean yields (Oliveira Neto, 2017). We represent 'Other Crops' and 'Other' land managers as grouped AFTs although there may be multiple management strategies within these types, our focus here is on the production of soybean, maize, and beef.

2.2.2 Exogenous Processes

We use multiple CRAFTY 'capitals' to represent the numerous influences on the production of services (Table 2). For example, important determinants of agricultural production in Brazil have been found to include human capital, technology generation and dissemination, climate conditions, and transport networks and land access (Pereira 2012; Rada 2013). All capitals vary spatially across the study area except for Tech capitals, which are spatially uniform as these represent improvements in technology that lead to broad-scale yield improvements through time (including due to improved seed varieties, machinery and fertilisers that are available widely). Values for capitals are provided exogenously (i.e., from ancillary data sources), except for the three 'Access' (Nature, Soybean-Maize, Other Crops) capitals and the Conservation capital, which are calculated endogenously (i.e., during a simulation run) from the dynamic spatial configuration of AFTs. All endogenous and many exogenous capitals are updated in each timestep (i.e., annually), although some are updated less frequently (see Table 2) either because of data availability (e.g., transport network) or because the process they represent does not occur on an annual basis (e.g., the soybean moratorium occurred in a given year, see below). All scripts to create files for initializing and updating capital values are available online (Millington 2020a). We discuss exogenous capitals in the remainder of this section, and endogenous capitals and representation of other processes in the following section (section 2.2.3).

Table 2. 'Capitals' influencing modelled services.

Capital	*Description*	*Services Influenced*	*Update Years*	*Data Source**
Moisture-Main	Main growing season (Oct-Mar climate)	Soybean, Maize, Beef, Other Crops	All	Climate
Moisture-Second	Maize second-crop growing season (Jan – Jun climate)	Soybean, Maize	All	Climate
Transport	Import and export costs due to transportation	Soybean, Maize, Beef, Other Crops	2005,2010, 2017	Transport Network
Land Value	Land attractiveness for establishing new agriculture	Soybean, Maize, Beef, Other Crops	All	Land Price
Conservation	Conservation value of natural land	Nature	All	Endogenous
Tech-Soy-Maize	Technology and resources influencing yield	Soybean, Maize	All	Commodity Production
Tech-Pasture	Technology and resources influencing yield	Beef	All	Commodity Production
Other	Incentive for 'Other' uses	Other	All	Land Cover/Use
Protection-Soybean	Prevents Soybean production in given cell	Soybean, Other	2006	Protected Areas
Protection-Maize	Prevents Maize	Maize, Other	2009	Protected Areas

	production in given cell			
Protection-Beef	Prevents Beef production in given cell	Beef, Other	None	Protected Areas
Protection-OCrops	Prevents OCrops production in given cell	Other Crops, Other	None	Protected Areas
Access-Nature	Spatial: proximity to natural land	Soybean, Maize	All	Endogenous
Access-Soy-Maize	Spatial: proximity to soybean or maize cells	Soybean, Maize	All	Endogenous
Access-OCrops	Spatial: proximity to other crop cells	Other Crops	All	Endogenous

*Data Sources correspond to Variables in Table 1

The Moisture capitals are derived from the monthly mean temperature and precipitation variables from the CRU TS v. 4.03 high-resolution gridded datasets (see Harris *et al*. 20202) and represent the role of climate on agricultural production. To understand climatic limitations associated with plant growth and agricultural production, we used the dryness index to describe the relation between water deficit and potential evapotranspiration (Pereira and Pruitt, 2004), both obtained from the Thornthwaite and Matter (1955) climatic water balance, as implemented by Victoria *et al.* (2007). This index represents the water deficit in percentage of potential evapotranspiration and is calculated by the equation:

$$DI = 100 * DEF / PET \tag{2}$$

where DI (%) is the dryness index; DEF is the water deficit; and PET is potential evapotranspiration (see Table 2, Victoria *et al*. 2007 and Millington 2020a for full definition). We calculate mean monthly DI for two different growing seasons (Oct–Mar and Jan–Jun) to calculate two sets of moisture capital values to represent climate influence on single- vs double-crop production. The use of these Moisture capitals also allows us to investigate the possible influence of climate change on agricultural productivity in simulations of alternative future scenarios (see section 2.3 below).

Transport infrastructure is a key variable influencing the spatial distribution and volume of agricultural production, related to land conversion (e.g., Soares-Filho *et al*. 2006; Weinhold and Reis 2008) and both imports of agricultural inputs and exports of commodities to markets (Rada 2013). The Transport capital uses data on the national road network (DNIT 2019) with locations (and operating years) of ports (ANTAQ 2019) to derive a spatial cost surface at a broad scale. This cost surface weights the quality of transport infrastructure such that paved roads present lower cost than unpaved roads for access to land (see Victoria *et al.* Submitted Data in Brief for full derivation). The use of the Transport capital allows us to investigate the possible influence of future alternative infrastructure development on land use change and agricultural productivity at a broad scale.

The Land Value capital is used to represent the incentive for agents to convert Nature land to Agriculture in places agricultural potential and infrastructure (as represented by the previous capitals) are poor, but which we know historically have seen conversion (e.g., based on Mapbiomas LULCC between 1985 and 2018). In these 'frontier regions', land prices are lower compared to other developed regions of Southern Brazil, reflecting the challenges of making a profit from agricultural production in frontier lands. These lower land prices provide an incentive to those willing to take a risk on developing land for grain production, assuming that future improvements (e.g., logistics, infrastructure) in the region will improve yields to pay-off the risk. Previous studies have shown that as the agricultural sector takes over in a given region, land prices tend to increase alongside infrastructure and social standards (Rezende, 2002; Ferro and Castro, 2013; Martinelli *et al.*, 2017). To develop this capital we used data from (IEG/FNP, 2017) to represent the relative cost of land for new development. As noted above, improvement in agricultural productivity (yields) through time due to advances in technological resources – such as improved machinery and seed varieties (e.g., Pereira 2012) – are represented using the 'Tech' Capitals. Because these capitals represent the aggregation of multiple sources of improvements in yield, we calibrate their values by combining our observed land use/cover data with commodity production data to calculate yields that are internally consistent within the model (see Millington 2020b).

The four Protection capitals represent areas of land that cannot be used for soybean, maize, beef or other crops. This exclusion may be because an area is designated as National or State Parks, indigenous lands or because of a policy that excludes production of a given commodity (e.g., Soybean Moratorium). For example, the Protection-Soybean capital is used in calibration runs (section 2.3) to represent the Soybean Moratorium policy implemented in 2006 to discourage deforestation and limiting the market for soybean grown on deforested lands (Gibbs *et al.* 2015; Dou *et al.*, 2018). These capitals therefore also allow examination of simulation runs that implement similar policies. Finally, the 'Other' capital is used to drive a high probability of cells being in the Other land cover category (e.g., urban and water), based on observed land cover change for calibration and potentially for representation of future expansion of this land type.

In the CRAFTY modelling framework, demand for services is provided exogenously, rather than incorporated as an endogenous process. Furthermore, this demand is specified in the same abstract units used to represent services production (section 2.2.1 above). As our focus here is on soybean, maize and beef we derive the abstract demand for these services from real production units (kg), while for Other Crops, Other and Nature we derive demand based on land area (ha; see Millington 2020b).

2.2.3 Endogenous Processes

Processes driven by the spatial configuration or historical contingency of land resources and agent actions are represented in the model endogenously by updating cell- or agent-states dynamically during a simulation, dependent on their circumstances in each timestep. Specifically, we represent the spatial agglomeration effect of agricultural economies, the tendency of land conversion to be spatially contagious, vegetation regeneration processes, and producer debt. Representing these processes required additions to the CRAFTY source code (see Millington 2020c).

The importance of spatial proximity for driving land use change due to efficiencies afforded by agglomeration economies is well known (e.g., Fujita and Krugman 1995; Porter 2000) and has been shown to be important in Brazil (Vera-Diaz *et al.* 2008; Garrett *et al.* 2013; Picoli *et al.* 2020). To represent this in the model at a local level, the Access-Soy-Maize and Access-OCrops capitals are updated in each cell in each timestep during simulation runs based on whether one of these agent-types is present within the eight neighbouring cells (Moore neighbourhood; with value 0.05 if the target agent-type is not present, 0.95 if the agent-type is present, and 1.0 if the cell is occupied by the specified AFT). Similarly, conversion of natural land for agriculture is known to be well-modelled as a contagious process of spread from existing cultivated areas at the edge of natural lands (e.g., Rosa *et al.* 2013). To represent spatial access to natural land at a local level the Access-Nature capital is updated in each cell in each timestep during simulation runs based on the adjacency of nature and non-nature land covers. For any given cell, if the Moore neighbourhood is composed entirely of nature cells the capital takes a value of 1.0, if between 1 and 7 cells in the Moore neighbourhood are nature cells the cell takes a value of 0.75, and finally a Nature Access capital value of 0.0 is taken if all neighbouring cells are non-nature.

Through time, we represent regeneration of natural vegetation following land abandonment by modifying Conservation capital cell values during a simulation run based both on the time since last human disturbance but also the type of disturbance. Several recent studies support the hypothesis that forest regeneration rate is related to 'intensity' of previous land use (Mesquita *et al.* 2015; Jakovac *et al.* 2015; Martines-Ramos *et al.* 2016), and here we assume that the rate of regeneration is faster following extensive pasture land-use than intensive crop (soybean, maize) land uses. Hence, for cells in a simulation run that have never been disturbed (i.e., have always had a Nature land cover) the Conservation capital will have value 1.0. On conversion, the Conservation capital value is reduced, to 0.4 if to pasture and to 0.0 for the other non-nature land covers. Following abandonment of a non-nature use, Conservation capital is increased by a value by 0.01 each timestep. The final endogenous process we represent is the accrual and repayment of debt by producers changing land uses. New producers often need to take out loans to pay for land, new machinery, seed, and other start-up costs. Producers can be 'trapped' into activities needed to earn profits to make repayments (e.g., Silva *et al.* 2020) and changes in land use are unlikely during the repayment period. To represent this inertia following conversion, we prevent new agricultural agents from changing land use until the debt is repaid. Debt is measured in years (the number of years to pay off the debt) and is specified for transitions as shown in Appendix B.

2.3 Calibration and Scenarios

Previous implementations of the CRAFTY framework for modelling real-world regions have calibrated model parameters using methods that ensure internal consistency and produce expected system trajectories but without comparing model outputs to empirical observations (e.g., Blanco *et al.* 2017; Brown *et al.* 2019). In contrast, here we use empirical data for land cover/use and agricultural commodity production (Table 1) to parameterise AFT production functions and capital conversions (e.g., from climate dryness index to Moisture capitals), identifying values that reproduce trends and patterns observed over the period 2001-2018. This approach aims to both ensure internal consistency, but also identify key areas of uncertainty in the model and is one that has not been employed in previous CRAFTY modelling applications. Understanding this uncertainty is important and useful for assessing model outputs for scenarios that project future land cover/use and agricultural production. Here, we compare simulated land cover/use and agricultural production to observed data for the same variables, aggregated for the entire study area. We also compare observed and simulated municipality-level proportions of land cover/use for snap-shots in time (5-year intervals). Final calibrated production function values are shown in Appendix C and capital conversions are shown in Millington (2020a).

Once calibrated, we use the model to examine how alternative scenarios of exogenous processes influence simulated land use/cover and agricultural production. Specifically, we examine scenarios of commodity demand, agricultural yields, climate change and land protection designations (Table zz) for the period 2019-2035. Examining outputs from these scenarios also helps us to better understand the drivers and dynamics in the model and identify most expected important drivers of future change. For scenarios of commodity demand we use trends of change over the period 2001-2018 to specify three alternative scenarios: i) business as usual that simply continues the observed mean rate of change in demand for different commodities linearly into the future, ii) increase demand over observed rates of change by 10% and iii) decreased demand below observed rates of change by 10%. These scenarios allow us to understand the sensitivity of the model to varying demand, whether due to changes in food consumption patterns, population growth or both. For agricultural yield scenarios we use projections of future yield change derived from two different observation periods as yield improvements have changed over time. In *YieldA* scenario we use annual growth rate of yields for 2001-2018 (i.e., 1.75% for Crops and 3.84% for Pasture), while in *YieldB* scenario we use annual growth rate of yields for 2013-2018 (i.e., 3.47% for Crops and -0.2% for Pasture). For climate change scenarios we use regionally-downscaled projections of temperature and precipitation from the WCRP Coordinated Regional Downscaling Experiment (CORDEX). Our chosen driving model was HAD-GEM2 as this model has been shown to have better agreement with observed rainfall in Atlantic Forest, Caatinga and Cerrado biomes than other models, although biases do remain (Rosolem *et al.* 2018). We use monthly data for Daily Minimum Near-Surface Air Temperature (*tamin*), Daily Maximum Near-Surface Air Temperature (*tasmax*) and Precipitation, (*pr*) from ensemble r1i1p1 and CORDEX region SAM44 (downscaling realisation v3). Projections were for representative concentration pathways RCP4.5 and RCP 8.5 and were accessed via the ESGF-CEDA project (CEDA 2019). Scripts that convert these data to Moisture Capitals values are available in Millington (2020a). Finally, we also examine a scenario including possible change in protected areas to examine how important policies that restrict agricultural land-use and production are for model dynamics, a likely scenario given the environmental policy direction the current Brazilian Federal government is taking (e.g., Abessa *et al.* 2019).

Table 3. Specification of counter-factual scenarios (2019-2035).

Scenario [label]	*Demand Conditions*	*Climate Conditions*	*Yield Conditions*	*Variations*
Constant [*Const*]	Invariant from 2018	Invariant from 2018	Invariant from 2018	None
Observed Demand [*ObsD*]	Observed rate of change over 2001-2018	Invariant from 2018	Invariant from 2018	None
High Soybean-Maize Demand [*HighD*]	Soybean & Maize 10% *greater* than annual Observed Demand	Invariant from 2018	Invariant from 2018	None
Low Soybean-Maize Demand [*LowD*]	Soybean & Maize 10% *lower* than annual Observed Demand	Invariant from 2018	Invariant from 2018	None
Yield 2001-18 [*YieldA*]	Invariant from 2018	Invariant from 2018	Observed 2001-2018 yield rate increase	None
Yield 2013-18 [*YieldB*]	Invariant from 2018	Invariant from 2018	Observed 2013-2018 yield rate increase	None
RCP 4.5 Climate [*RCP45*]	Invariant from 2018	Moisture capitals from GCM for *RCP45*	Invariant from 2018	None
RCP 8.5 Climate [*RCP85*]	Invariant from 2018	Moisture capitals from GCM for *RCP85*	Invariant from 2018	None
Combined 1 [*BAU*]	As for *ObsD*	Moisture capitals from GCM for *RCP45*	As for *YieldB*	None
Combined 2 [*EXT*]	As for *HighD*	Moisture capitals from GCM for *RCP85*	As for *YieldB*	None
Combined 3 [*NoPro*]	As for *HighD* plus 10% *lower* Nature demand	Moisture capitals from GCM for *RCP85*	As for *YieldB*	Protected Areas removed

3. Results

3.1 Calibration

Results from model calibration indicate that CRAFTY-Brazil is able to reproduce observed time-series of total area and production for the entire study area, but performs less well in terms of spatial allocation across the study area. Observed trends of decreasing Pasture and Nature area combined with increases in cropland area are reproduced well, with small year-to-year variation (Figure 2a). The general trends of increases in production of Soybean and Maize are reproduced, although much of the large inter-annual variability in production is not captured (Figure 2b). Interestingly, there is also a slight lag in the rate of increase in Maize production 2011-2014, and dramatic decreases in recent years are not captured. Correspondence between the two sets of time series can also be noted, for example with the under-estimation of cropland area 2003-2005 linked to under-estimation of Soybean production in these years.

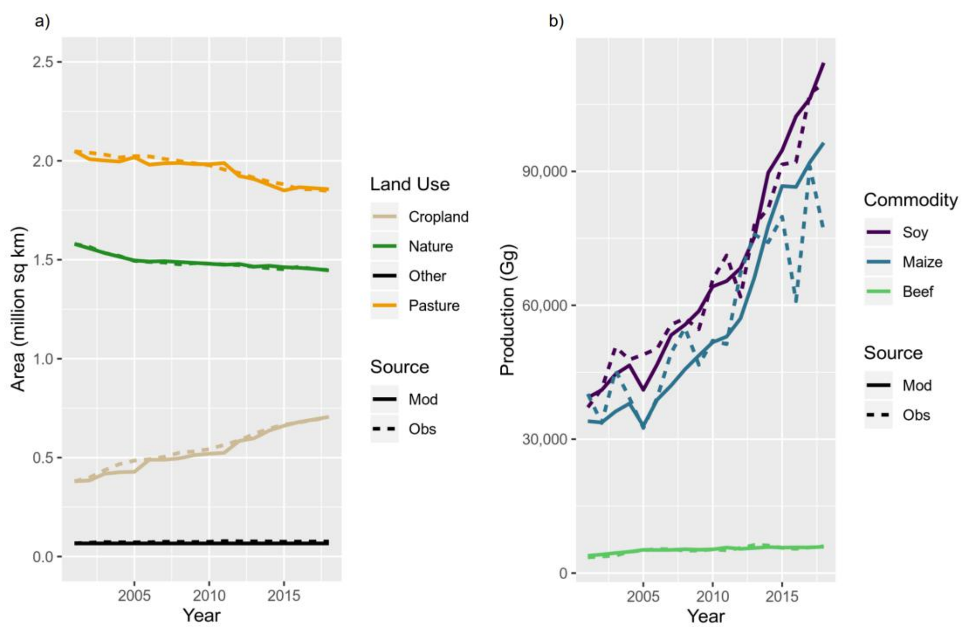

Figure 2. Calibration time-series. a) land use/cover and b) commodity production

Although time-series of observed trends are reasonably well reproduced, there are some disparities between observed and modelled locations of land use/cover (Figure 3). For example, the model tends to locate more Cropland in the north east of the study area (Bahia state) than has been observed, with correspondingly less Pasture than observed in this area. Conversely, in the central part of the study area (São Paulo state), the model produces more Pasture than observed, at the expense of cropland. Nature is reasonably well modelled across the study area, although with some over estimation (at the expense of cropland) in the north west of the study area (Mato Grosso state). While we see generally consistent variation from observations in the simulated time-series, accuracy in spatial allocation of land use seems to deteriorate through time. For example, while the modal land use/cover was incorrectly modelled for 9.8% of municipalities in 2009, this had risen to 16.4% by 2018.

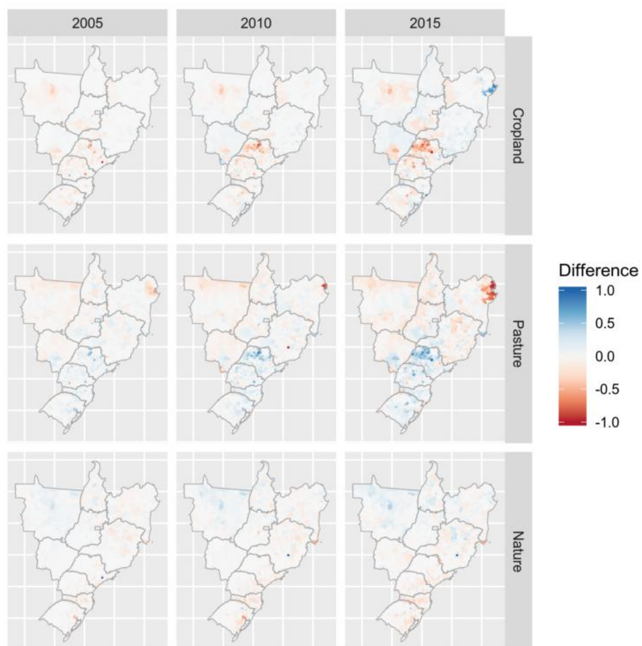

Figure 3. Municipality proportional difference from observed by land use/cover. Red shades indicate underprediction and blue shades overprediction relative to observed data (i.e., Figure 1c).

3.2 Scenarios

Assuming constant 2018 conditions into the future (*Const* scenario, Figure 4a), all land covers level-off to a steady state by 2022, with the exception of Maize and Double-cropping; DC replacing Maize as it is more competitive. Hence, Soybean production continues to rise while Maize production declines slightly.

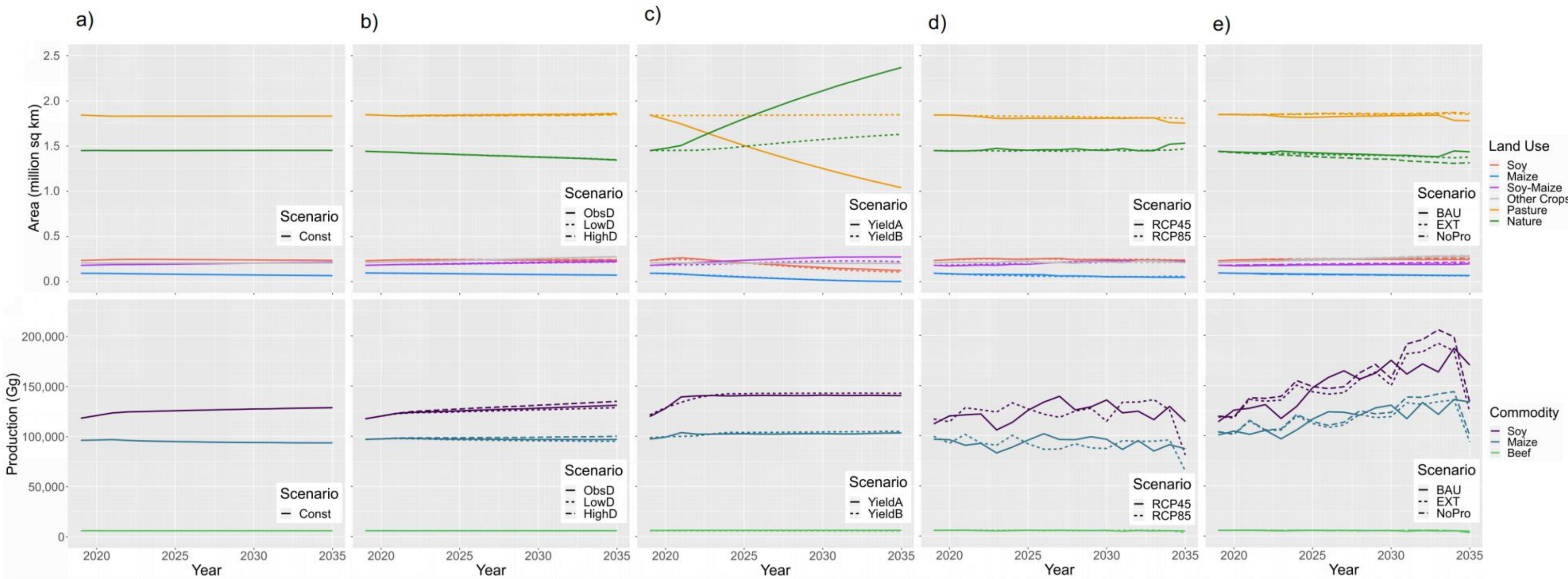

Figure 4. Time series of land use and production for all scenarios.

Results for scenarios of future commodity demand based on observed trends (*ObsD*, *LowD* and *HighD*, Figure 4b), show that land cover and production are relatively insensitive with all other conditions constant. For example, projected demand for Maize and Soybean in the *LowD* and *HighD* scenarios (10% lower and 10% higher than the annually observed 2001-2018 rate of increase respectively) would expect ~4% lower/higher production by the end of the simulation runs, yet simulated production outputs are only ~2% and ~3% lower/higher. As would be expected given this insensitivity and the corresponding limited deviation from constant conditions through time, spatially there is limited variation with differences close to zero across the study area (Figure 5).

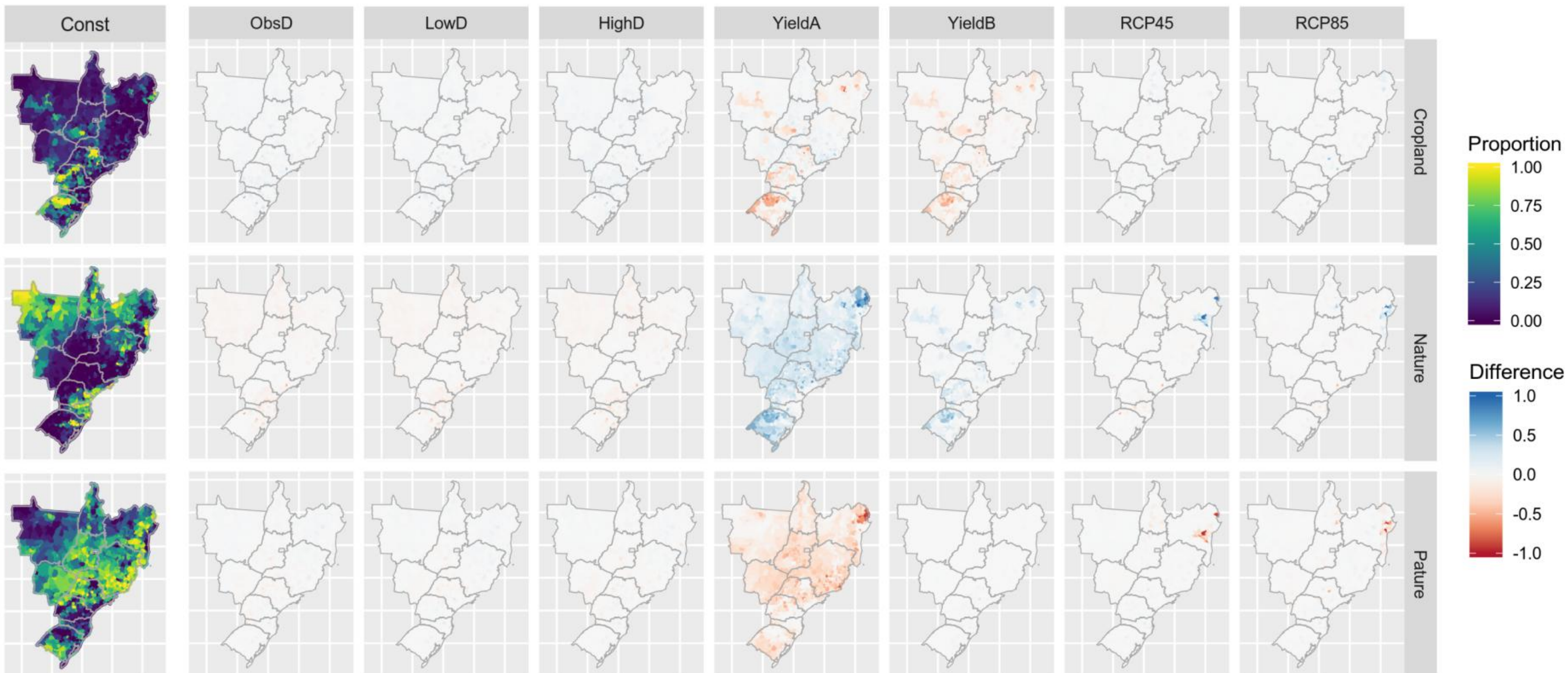

Figure 5. Spatial variation in outputs for individual scenarios. Municipality Differences are calculated from Proportions for the *Const* scenario (shown at left) for simulated year 2033.

Yield scenarios (*YieldA* and *YieldB*, Figure 4c) have the largest impact on land cover, with some lesser impact on production. The *YieldA* scenario (which projects future yield-increase rates based on observed 2001-2018 rates; 1.75% for Cropland, 3.84% for Pasture) results in a significant decrease in land devoted to Pasture and larger increases in Nature, plus a smaller (but still large) increase in double-cropped Soybean-Maize land. Spatially, this can be observed as large areas of Cropland and Pasture with negative Cropland and Pasture differences (compared to the Constant scenario), and positive Nature differences (red and blue hues respectively in Figure 5). The *YieldB* scenario (using 2013-2018 rates of yield change; 3.47% for Cropland and -0.2% for Pasture) results in very similar decreases for Cropland uses, but Pasture land remains constant as production struggles (and fails) to meet demand. Consequently, only Cropland land is abandoned, reverting to Nature (as shown by negative Cropland, positive Nature and neutral Pasture differences in Figure 5).

Results from climate scenarios based on outputs from GCM simulations (*RCP45* and *RCP85*, Figure 4d) produce greatest inter-annual variation in production, although with much lower inter-annual variation for land use (which is similar to the other, non-climate, scenarios). Variation in production is a result of inter-annual variation in precipitation and temperature which directly influence Moisture Capitals, and therefore agricultural service provision. Of particular note is the large drop in production for scenario *RCP85* in the final year of the simulation (2035, with a similar but lesser drop for scenario *RCP45*), the result of a deep and widespread decrease in annual precipitation in GCM outputs. This drop in production does not produce a commensurate change in land cover (in the same year) as there is a lag in agent decision-making (e.g., to abandon land). This lag in decision-making can be seen in the noticeable decrease in Pasture and increase in Nature land covers (indicating Pasture abandonment) simulated in 2034 for scenario *RCP45*. This abandonment of Pasture is the result of consecutive years of low precipitation which also caused a period of relatively low Soybean and Maize production for 2030-2033 (compared to the overall increasing trend) although with no effect on Cropland cover. The 2034 Pasture abandonment is focused in the north east of the study area (Figure 5, *RCP45*).

The results for combined scenarios (Figure 4e and Figure 6) exhibit combinations of the trends and patterns seen in the other scenarios that vary individual driving factors. All three simulations result in increased production through time, in response to improved yields and increased demand (for *EXT* and *NoPro*). All three scenarios exhibit the same inter-annual variation due to climate change. Greatest increases in production are found in the *NoPro* scenario, although corresponding increases in Cropland and Pasture land are not spatially confined to formerly protected areas (e.g., blue/red shades for Cropland/Nature respectively in Figure 6 are found across the entire study area). As found in the climate change scenarios, switches from Pasture to Nature land in the north east of the study area are also observed (especially for the *BAU* scenario which uses the RCP 4.5 climate projection).

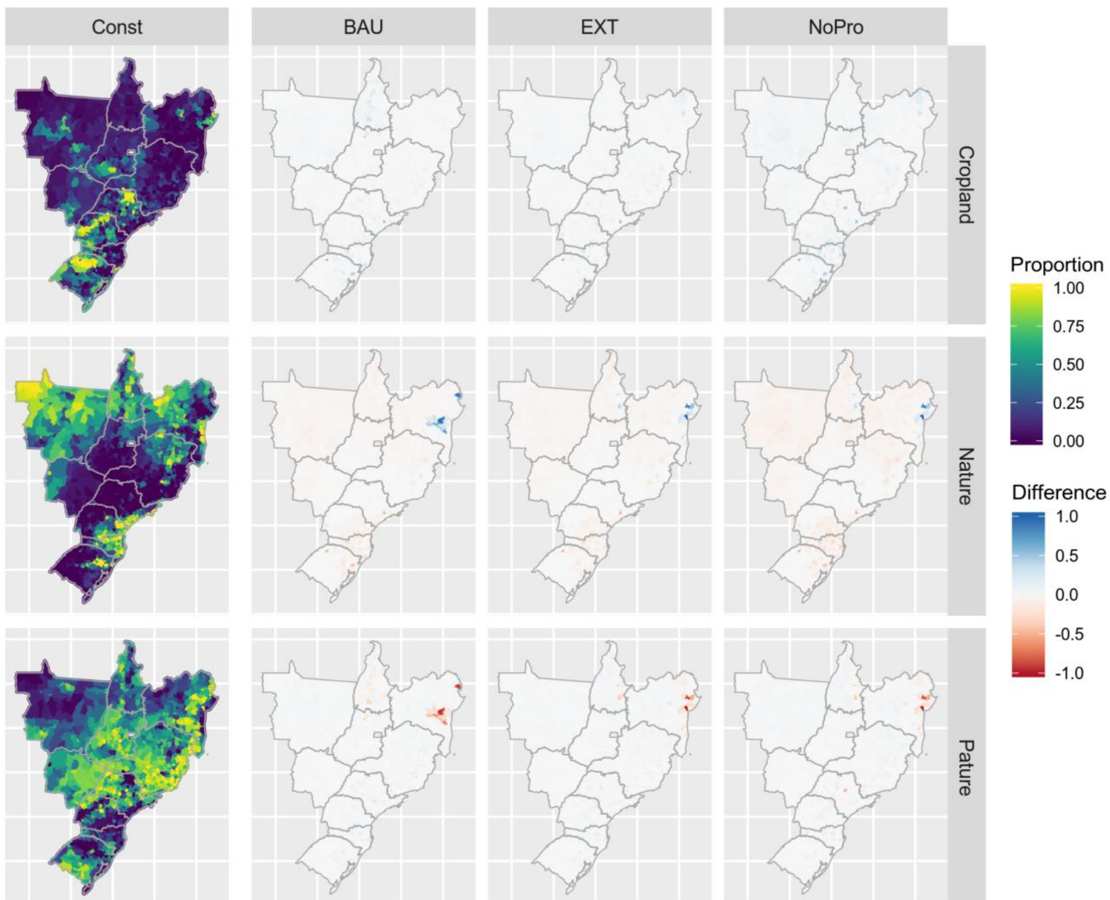

Figure 6. Spatial variation in outputs for combined scenarios. Municipality Differences are calculated from Proportions for the *Const* scenario (shown at left) for simulated year 2033.

4. Discussion and Conclusions

4.1 Calibration

This paper represents one of the first attempts to calibrate the CRAFTY land use/cover modelling framework against observed data. Such an approach has not been used in the past, often because comprehensive data describing Capitals, Demand and spatial distribution of land use are unavailable at the broad (national to continental) scales CRAFTY is designed for (Brown *et al*. 2019). Other approaches to calibrate CRAFTY have used 'stability checks' with baseline inputs to ensure outputs do not deviate or oscillate wildly from expected behaviour (e.g., Blanco *et al*. 2017; similar to our Constant conditions scenario) or to ensure 'sensible' outputs are produced when starting from a blank map of null land use (e.g., Brown *et al*. 2019). As described in Section 2.1, we have utilised multiple observed data sets to enable our approach but have also needed to make assumptions about how data represent different processes. For example, commodity demands in any given year are difficult to assess and for our calibration here we have assumed continuous market clearing such that observed (2001-2018) production perfectly met demand in each year. The value of such a strong but simple assumption during calibration is that it enables clear understanding of the meaning of demand in future scenarios, and matches data that can be readily produced by the System Dynamics model of global trade (based on Warner *et al*. 2013) that we plan to link with CRAFTY-Brazil. In the hybrid model produced, demand will be modelled endogenously by the System Dynamics model, adding further variation to CRAFTY-Brazil inputs that will need to be appropriately assessed (e.g. via sensitivity analyses).

Our calibration of CRAFTY-Brazil was more successful in reproducing observed aggregate land cover and production values than for the spatial distribution of variables (Figures 2 and 3), a situation similar to previous applications of CRAFTY (Blanco *et al*. 2017; Brown *et al*. 2019). In particular, results from our calibration show Pasture outcompeting Cropland in the centre of the study area and vice versa in the north east, neither of which were observed historically. This is surprising as the north east of the study area is relatively marginal for Cropland uses, while the conditions further south and centrally are better. It seems that our calibration allows the marginal utility of the Beef service to increase at a rate faster than Cropland services, pushing the latter to less productive land. Such issues are likely further exacerbated by uncertainty in the land use/cover maps against which we calibrate our model (see section 2.1). Although the MapBiomas data are the best available, the uncertainty in classifying Grassland and 'Mosaics of Agriculture and Pasture' into the classes required for CRAFTY-Brazil may also contribute some level of 'error' in our calibration (although not at the magnitude of the spatial differences we observed; Figure 3).

4.2 Scenarios

Results of the scenarios that project change for 2019-2035 demonstrate that qualitative differences in spatio-temporal variation of future drivers generally have commensurate impacts on land cover and production (Figures 4, 5 and 6). For example, inputs for demand and yield scenarios are specified as a simple, temporally-uniform and spatially-invariant rates of change based on observed rates of mean annual change (e.g., annual 1.75% increase in crop yields in scenario *YieldA*). Output production time series for these scenarios generally change smoothly over time in a similar manner to the smoothly changing inputs (Figures 4a, 4b and 4c). In contrast, the precipitation and temperature outputs from GCMs used as inputs for the climate scenarios have much greater inter-annual and spatial variation. Consequently, output production time-series for scenarios using GCM data as inputs share this much greater inter-annual variability (Figures 4d and 4e) compared to scenarios using simple rates of change. These patterns highlight a strong influence of climate inputs on production outputs, but not on land cover change (yield is also important for production, but in a different way as we discuss below). The inter-annual variability in production we see here is not sufficient for land cover change to occur through abandonment (as discussed by Silva *et al.* 2020), but many fine details of farm-level financing that may be vital for individual farm viability are not represented in this model and so we cannot conclusively argue that land cover change would not occur under the climate projects we have examined.

Spatially, the distinction between scenarios is less clear. Land cover outputs for demand scenarios (that use spatially-uniform inputs) exhibit broad and diffuse change, while land cover change for scenarios using spatial GCM data (that use spatial inputs) show intense and spatially-focused change (Figures 5 and 6). However, yield scenarios (that use spatially-uniform inputs) exhibit a range of spatial patterns (Figure 5), with some areas of spatially uniform change (e.g., central areas of the study area for Nature and Pasture under scenario *YieldA*), some areas of intense and spatially-focused change (e.g., in the north-east of the study area for Nature and Pasture under scenario *YieldA*), and other regions of less intense by still localised areas of change (e.g., the south-western region for Cropland and Nature in both yield scenarios). These spatial variations reflect different aspects of land cover that are influenced by high yields: in the centre of the study area Pasture dominates at initialisation of simulations but is abandoned (i.e., switches to Natural land) gradually when yield-increase rates are high (i.e., *YieldA*); in the north east areas of Pasture land initially are the most marginal and so entirely abandoned when yields increase rates are high (*YieldA*); and the south-west is initially dominated by dense Cropland which is abandoned to varying degrees for different yield increase rates (both scenarios).

Time series of land cover change indicate differences in the sensitivity of land cover between inputs. In particular, land cover seems to be insensitive to changes in demand (i.e., there is a smaller % change in land cover outputs than the % change in demand specified; Figure 4b), but much more sensitive to changes in yield (with some very large shifts in land cover observed for yield scenarios; Figure 4c). The insensitivity of land cover in the demand scenarios presented here is because competition for land with other services (such as Nature) and constant 2018 yields means that commodity production never reaches the required demand in these scenarios. Hence, production and land cover time series differ little from those seen in the Constant scenario as production is already at its limit at the start of simulations given the calibrated yield values. In contrast, when yields increase through time (as observed over 2001-2018; scenario *YieldA*), less land is needed for Pasture land to meet Beef demand and much is abandoned (reverting to Nature land cover). Similarly, increases in yield for crops (although not as fast) also allow decreased Cropland area and a switch to more intensive double-crop Soybean-Maize land use (resulting in Maize-only land decreasing to zero) In both situations, led by yield increases in pasture and crops, we have what previous studies defined as 'land sparing' where the increased volume of production per land unit (i.e., agricultural intensification) leads to a decrease in cropland, or at least alleviating the pressure for cropland expansion (Angelsen and Kaimowitz 2001; Hertel *et al.* 2014). In this scenario, production of all products is able to meet (the constant 2018) demand because of the high yields, resulting in production time series that flatline. However, when Pasture yields change very little, but crop yields change more quickly (as was the case 2013-2018; scenario *YieldB*), the area of Pasture land required remains relatively constant and increases in Nature land are due only to decreased area of Cropland. Hence, spatially we see decreases in Cropland in areas that have historically been intensively farmed, with corresponding increases in Nature land (see red and blue hues, respectively, in Figure 5).

Combined scenarios were designed to enable examination both of the effects of variations in multiple input factors but also potential counterfactual futures. For example, for all three combined scenarios we see both inter-annual variability in Soybean and Maize production (due to climate) but with increasing trend (due to increases in yield and demand), a combined pattern we do not see in the other scenarios. For land cover time series, although all three scenarios use the *YieldB* input values for yield, we do not see the same increases and decreases in Nature and Cropland (respectively) that we see for that individual scenario. This is because Cropland and Pasture demands increase through time (and in the *NoPro* scenario, Nature demand also decreases) ensuring pressure for agricultural land remains (whereas for increasing yield with constant demand the pressure is reduced). The *NoPro* scenario results in greatest production (and lowest Nature land area) as the ability to farm (formerly) protected areas plus decreased demand for Nature land allow greatest shifts in land from Nature to Cropland and Pasture. However, increases in Cropland and Pasture land in the *NoPro* scenario are not spatially confined to formerly protected areas as might have been expected. This is likely because these protected areas have relatively limited infrastructure, which is considered invariant through time in the *NoPro* scenario. With many of the indigenous and park lands are some distance from the ‘core’ agricultural production areas and much pastures and other non-protected land available for conversion (processes also represented in the model), even under the scenarios we examine there is little pressure on the current protected areas. However, if restrictions on land use in protected areas really are relaxed (e.g. Abessa *et al*. 2019) we might expect improvements in infrastructure (e.g., road building), which may in-turn lead to a positive feedback and greater exploitation of these areas over the longer term (e.g., Weinhold and Reis 2008).

4.3 Future Development

The tradeoffs necessary in spatial agent-based modelling of land-use systems have been well identified in the literature, in particular with regard a perceived spectrum from empirically-grounded and complicated models to theoretically-focused and simple models (e.g., O’Sullivan *et al*. 2015, Sun *et al*. 2016). Here, our approach has been to build on the theoretical structure provided by the CRAFTY framework and remain relatively conceptually simple, but also incorporating empirical data where possible to ground our application for Brazil. In doing so, this version of CRAFTY-Brazil limited the number of agent-functional types and services represented (eight services aggregated to four land use/cover types) and yet this still required the representation of a greater number of capitals (15) and processes (multiple, including the accrual of debt) than we had initially expected. While model calibration and performance are currently as good as we might hope, our limited number of agent-function types may actually be limiting model performance by creating what Sayer (1992, p.138) termed a ‘chaotic conception’, a group of agent-types that artificially “lump together the unrelated and the inessential” and inadequately represent differences between real world actors that are needed to reproduce empirical events. Given the scales at which we are modelling and the broader aims to link CRAFTY-Brazil with representation of global trade we argue that the conceptualisation presented is the most appropriate for current purposes (and see Millington *et al.* 2017), but using more refined agent-types may help improve the spatial distribution of simulated outcomes.

The results of the scenarios we have simulated have helped us to understand where the greatest uncertainties are in the model and to identify processes that could be better represented. For example, the design of the scenarios we examined has been useful to show that the specification of yield is particularly important and limits the influence of demand. Agricultural yields are known to be a key but uncertain parameter in land use and global trade modelling (Plevin *et al.* 2015), and future use of CRAFTY-Brazil should account for this by considering a range of possible yield values in scenarios to understand behaviour more robustly. Furthermore, results of the *NoPro* scenario (limited loss of Nature in former protected areas, likely due to absence of infrastructure) highlight the importance for future uses of the model to consider change in multiple (all) capitals simultaneously. Each of these issues will be needed to be considered as we dynamically couple CRAFTY-Brazil to a system dynamics model to create a hybrid simulation model for examining the land use impacts of telecoupled global trade (Millington *et al.* 2017, Liu et al 2018).

Appendix A: MapBiomas v4.0 Reclassification

Code and Description are from MapBiomas, Reclassification is used here. Note that Cropland are further disaggregated into Soy, Maize and Other Crops using planted area data. Scripts used to resample, reclassify and disaggregated the MapBiomas data are available in Millington (2019).

Code	*Description*	*Reclassification*
1	Forest Formations	Nature
1.1	Natural Forest Formations	Nature
1.1.1	Dense Forest	Nature
1.1.2	Open Forest	Nature
1.1.3	Mangrove	Nature
1.2	Forest Plantations	Nature
2	Non-Forest Natural Formations	Nature
2.1	Non-Forest Formations in Wetlands	Nature
2.2	Grassland	Pasture (Nature in protected areas)
2.3	Salt Flat	Nature
2.4	Rocky Outcrop	Other
2.5	Other non-forest natural formations	Other
3	Farming	Cropland
3.1	Pasture	Pasture
3.2	Agriculture	Cropland
3.2.1	Annual and Perennial Crop	Cropland
3.2.2	Semi-perennial Crop	Cropland
3.3	Mosaic of Agriculture and Pasture	Cropland
4	Non-Vegetated Areas	Other
4.1	Beaches and Dunes	Other
4.2	Urban Infrastructure	Other
4.3	Mining	Other
4.4	Other Non-Vegetated Area	Other
5	Water Bodies	Other
5.1	River, Lake and Ocean	Other
5.2	Aquaculture	Other
6	Not Observed	Other

Appendix B: Debt

Debt incurred by agents following land use change. Units are years.

Previous Land Use	*Cropland Agent*	*Pasture Agent*
Nature or Other	5	3
Soybean, Maize or Other Crops	3	3
Double-Cropping	0	3
Pasture	4	NA

Appendix C: Production Functions

Capitals, Agent-Functional Types and their production weighting factors for each Service

a) Soybean AFT

	Soybean	*Maize*	*Nature*	*OCrops*	*Other*	*Beef*
Moisture-Main	0.8	0	0	0	0	0
Moisture-Second	0	0	0	0	0	0
Transport	0.5	0	0	0	0	0
Land Value	0	0	0	0	0	0
Conservation	0	0	0	0	0	0
Tech-Soy-Maize	0.8	0	0	0	0	0
Tech-Pasture	0	0	0	0	0	0
Other	0	0	0	0	0	0
Protection-Soy	1	0	0	0	0	0
Protection-Maize	0	0	0	0	0	0
Protection-Beef	0	0	0	0	0	0
Protection-OCrop	0	0	0	0	0	0
Access-Nature	1	0	0	0	0	0
Access-Soy-Maize	0.4	0	0	0	0	0
Access-Orop	0	0	0	0	0	0
Production	***1***	***0***	***0***	***0***	***0***	***0***

b) Maize AFT

	Soybean	*Maize*	*Nature*	*OCrops*	*Other*	*Beef*
Moisture-Main	0	0.8	0	0	0	0
Moisture-Second	0	0	0	0	0	0
Transport	0	0.5	0	0	0	0
Land Value	0	0	0	0	0	0
Conservation	0	0	0	0	0	0
Tech-Soy-Maize	0	0.8	0	0	0	0
Tech-Pasture	0	0	0	0	0	0
Other	0	0	0	0	0	0
Protection-Soy	0	0	0	0	0	0
Protection-Maize	0	1	0	0	0	0
Protection-Beef	0	0	0	0	0	0
Protection-OCrop	0	0	0	0	0	0
Access-Nature	0	1	0	0	0	0
Access-Soy-Maize	0	0.4	0	0	0	0
Access-Orop	0	0	0	0	0	0
Production	***0***	***1***	***0***	***0***	***0***	***0***

c) Double-Crop AFT

	Soybean	*Maize*	*Nature*	*OCrops*	*Other*	*Beef*
Moisture-Main	0.8	0	0	0	0	0
Moisture-Second	0	0.5	0	0	0	0
Transport	0.5	0.5	0	0	0	0
Land Value	0	0	0	0	0	0
Conservation	0	0	0	0	0	0
Tech-Soy-Maize	0.8	0.8	0	0	0	0
Tech-Pasture	0	0	0	0	0	0
Other	0	0	0	0	0	0
Protection-Soy	1	0	0	0	0	0
Protection-Maize	0	1	0	0	0	0
Protection-Beef	0	0	0	0	0	0
Protection-OCrop	0	0	0	0	0	0
Access-Nature	1	1	0	0	0	0
Access-Soy-Maize	0.4	0.4	0	0	0	0
Access-Orop	0	0	0	0	0	0
Production	***0.8***	***0.75***	***0***	***0***	***0***	***0***

d) Nature AFT

	Soybean	*Maize*	*Nature*	*OCrops*	*Other*	*Beef*
Moisture-Main	0	0	0	0	0	0
Moisture-Second	0	0	0	0	0	0
Transport	0	0	0	0	0	0
Land Value	0	0	1	0	0	0
Conservation	0	0	1	0	0	0
Tech-Soy-Maize	0	0	0	0	0	0
Tech-Pasture	0	0	0	0	0	0
Other	0	0	0	0	0	0
Protection-Soy	0	0	0	0	0	0
Protection-Maize	0	0	0	0	0	0
Protection-Beef	0	0	0	0	0	0
Protection-OCrop	0	0	0	0	0	0
Access-Nature	0	0	0	0	0	0
Access-Soy-Maize	0	0	0	0	0	0
Access-Orop	0	0	0	0	0	0
Production	***0***	***0***	***1***	***0***	***0***	***0***

e) Other Crops AFT

	Soybean	*Maize*	*Nature*	*OCrops*	*Other*	*Beef*
Moisture-Main	0	0	0	0	0	0
Moisture-Second	0	0	0	0	0	0
Transport	0	0	0	0	0	0
Land Value	0	0	1	0	0	0
Conservation	0	0	1	0	0	0
Tech-Soy-Maize	0	0	0	0	0	0
Tech-Pasture	0	0	0	0	0	0
Other	0	0	0	0	0	0
Protection-Soy	0	0	0	0	0	0
Protection-Maize	0	0	0	0	0	0
Protection-Beef	0	0	0	0	0	0
Protection-OCrop	0	0	0	0	0	0
Access-Nature	0	0	0	0	0	0
Access-Soy-Maize	0	0	0	0	0	0
Access-Orop	0	0	0	0	0	0
Production	***0***	***0***	***1***	***0***	***0***	***0***

f) Other AFT

	Soybean	*Maize*	*Nature*	*OCrops*	*Other*	*Beef*
Moisture-Main	0	0	0	0	0	0
Moisture-Second	0	0	0	0	0	0
Transport	0	0	0	0	0	0
Land Value	0	0	0	0	0	0
Conservation	0	0	0	0	0	0
Tech-Soy-Maize	0	0	0	0	0	0
Tech-Pasture	0	0	0	0	0	0
Other	0	0	0	0	1	0
Protection-Soy	0	0	0	0	1	0
Protection-Maize	0	0	0	0	1	0
Protection-Beef	0	0	0	0	1	0
Protection-OCrop	0	0	0	0	1	0
Access-Nature	0	0	0	0	0	0
Access-Soy-Maize	0	0	0	0	0	0
Access-Orop	0	0	0	0	0	0
Production	***0***	***0***	***0***	***0***	***1***	***0***

g) Pasture AFT

	Soybean	*Maize*	*Nature*	*OCrops*	*Other*	*Beef*
Moisture-Main	0	0	0	0	0	0.2
Moisture-Second	0	0	0	0	0	0
Transport	0	0	0	0	0	0.5
Land Value	0	0	0	0	0	0
Conservation	0	0	0	0	0	0
Tech-Soy-Maize	0	0	0	0	0	0
Tech-Pasture	0	0	0	0	0	1
Other	0	0	0	0	0	0
Protection-Soy	0	0	0	0	0	0
Protection-Maize	0	0	0	0	0	0
Protection-Beef	0	0	0	0	0	1
Protection-OCrop	0	0	0	0	0	0
Access-Nature	0	0	0	0	0	0.2
Access-Soy-Maize	0	0	0	0	0	0
Access-Orop	0	0	0	0	0	0
Production	***0***	***0***	***0***	***0***	***0***	***0.85***

Author Contributions

Conceptualization, All authors; Software, JM and VK; Formal Analysis, JM, RBS and DCV; Data Curation, JM, RBS and DCV; Writing – Original Draft Preparation, JM; Writing – Review & Editing, JM, DCV, RBS and MB.

Funding

This work was supported by the UK Natural Environment Research Council, under grant reference NE/M021548/1. We gratefully acknowledge the funding support by the São Paulo Research Foundation (FAPESP), processes 14/50628-9, 15/25892-7, and 18/08200-2.

Acknowledgements

We are grateful to numerous colleagues for ideas and understanding gained from discussion on the subjects presented here. These include, but are not limited to, Steve Peterson, Jem Woods, Hang Xiong, Yue Dou and Jack Liu plus several local stakeholders in Brazil. We are grateful to Paul McCord for his early work with MapBiomas data.

Conflict of Interest

The authors declare no conflict of interest.

References

Abessa, D., Famá, A. and Buruaem, L. (2019) The systematic dismantling of Brazilian environmental laws risks losses on all fronts. *Nature Ecology & Evolution*, 3(4), 510-511.

Angelsen, A. and Kaimowitz, D. (2001) *Agricultural Technologies and Tropical Deforestation*. New York, NY: CABi.

ANTAQ (2019) Portos Brasileiros - Agência Nacional de Transportes Aquaviários. [Online] Available at: http://portal.antaq.gov.br/index.php/portos/portos-brasileiros/

Blanco, V., Holzhauer, S., Brown, C., Lagergren, F., Vulturius, G., Lindeskog, M. and Rounsevell, M.D. (2017) The effect of forest owner decision-making, climatic change and societal demands on land-use change and ecosystem service provision in Sweden. *Ecosystem Services*, 23, 174-208.

Brown, C., Seo, B. and Rounsevell, M.D. (2019) Societal breakdown as an emergent property of large-scale behavioural models of land use change. *Earth System Dynamics*, 10, 809–845

Brunini, O., Zullo, J., Pinto, H.S., Assad, E., Sawazaki, E., Duarte, A.P. and Patterniani, M.E.Z. (2001) Riscos climáticos ara a cultura de milho no estado de São Paulo, Brasil. *Revista Brasileira de Agrometeorologia, Passo Fundo*, 9, 519–526.

CEDA (2019) *CEDA ESGF Search Portal* [Online] Available at: https://esgf-index1.ceda.ac.uk/projects/esgf-ceda/

DNIT (2019) *PNV e SNV - Departamento Nacional de Infraestrutura de Transportes*. [Online] Available at: http://www.dnit.gov.br/sistema-nacional-de-viacao/sistema-nacional-de-viacao

Dou, Y., da Silva, R.F.B., Yang, H. and Liu, J. (2018) Spillover effect offsets the conservation effort in the Amazon. *Journal of Geographical Sciences*, 28(11), 1715-1732.

Dou, Y., Millington, J.D., Bicudo Da Silva, R.F., McCord, P., Viña, A., Song, Q., Yu, Q., Wu, W., Batistella, M., Moran, E. and Liu, J. (2019). Land-use changes across distant places: design of a telecoupled agent-based model. *Journal of Land Use Science*, 14(3), 191-209.

Dou, Y., Yao, G., Herzberger, A., Bicudo Da Silva, R., Song, Q., Hovis, C., Batistella, M., Moran, E., Wu, W. and Liu, J. (2020) Land-Use Changes in Distant Places: Implementation of a Telecoupled Agent-Based Model. *Journal of Artificial Societies and Social Simulation*, 23(1), p.11.

Ferro, A.B. and Castro, E.R.D. (2013) Determinantes dos preços de terras no Brasil: uma análise de região de fronteira agrícola e áreas tradicionais. *Revista de Economia e Sociologia Rural*, 51(3), pp.591-609.

Fujita, M. and Krugman, P. (1995) When is the economy monocentric?: von Thünen and Chamberlin unified. *Regional Science and Urban Economics*, 25(4), pp.505-528.

Garrett, R.D., Lambin, E.F., Naylor, R.L. (2013) The new economic geography of land use change: Supply chain configurations and land use in the Brazilian Amazon. *Land Use Policy*, 34, 265-275.

Gibbs, H.K., Rausch, L., Munger, J., Schelly, I., Morton, D.C., Noojipady, P., Soares-Filho, B., Barreto, P., Micol, L., Walker, N.F. (2015) Brazil's soybean moratorium. *Science*, 347, 377-378.

Gonçalves, S.L., Caramori, P.H., Wrege, M.S., Shioga, P. and Gerage, A.C. (2002) Épocas de semeadura do milho "safrinha", no Estado do Paraná, com menores riscos climáticos. Acta Scientiarum. Agronomy, 24, 1287-1290.

Hampf, A.C., Stella, T., Berg-Mohnicke, M., Kawohl, T., Kilian, M. and Nendel, C. (2020) Future yields of double-cropping systems in the Southern Amazon, Brazil, under climate change and technological development. *Agricultural Systems*, 177, 102707.

Harris, I., Osborn, T.J., Jones, P. and Lister, D. (2020) Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data*, 7(1), 1-18.

Heinemann, A.B., Ramirez-Villegas, J., Stone, L.F. and Didonet, A.D. (2017) Climate change determined drought stress profiles in rainfed common bean production systems in Brazil. *Agricultural and Forest Meteorology*, 246, 64-77.

Hertel, T.W., Ramankutty, N. and Baldos, U.L.C. (2014) Global market integration increases likelihood that a future African Green Revolution could increase crop land use and CO2 emissions. *Proceedings of the National Academy of Sciences*, 111(38), 13799-13804.

IBGE (2019) *Instituto Brasileiro de Geografia e Estatística*. [Online] Available at: https://www.ibge.gov.br/

IEG/FNP (2017) *AGRIANUAL 2017: Da visão do produtor ao empreendedor agrícola: Por que o produtor rural deve enxergar sua propriedade como uma organização lucrativa?* São Paulo: IEG/FBP

Jakovac, C.C., Peña-Claros, M., Kuyper, T.W., Bongers, F. (2015) Loss of secondary-forest resilience by land-use intensification in the Amazon. *Journal of Ecology*, 103(1), 67-77.

Lane, A. and Millington, J.D.A. (2021) *Maestro Solo (Version v1.0.1)* [Online] Available at: http://doi.org/10.5281/zenodo.4570115

Liu, J., Dou, Y., Batistella, M., Challies, E., Connor, T., Friis, C., Millington, J.D., Parish, E., Romulo, C.L., Silva, R.F.B. and Triezenberg, H. (2018) Spillover systems in a telecoupled Anthropocene: Typology, methods, and governance for global sustainability. *Current Opinion in Environmental Sustainability*, 33, 58-69.

MapBiomas (2019a) *Collection 4.0 of Brazilian Land Cover & Use Map Series*. [Online] Available at: http://mapbiomas.org/en

MapBiomas (2019b) *Collection 4.0 of Brazilian Land Cover & Use Map Series: Accuracy Analysis*. [Online] Available at: https://mapbiomas.org/en/accuracy-analysis

Martinelli, L.A., Batistella, M., Silva, R.F.B.D. and Moran, E. (2017) Soy expansion and socioeconomic development in municipalities of Brazil. *Land*, 6(3), 62.

Martínez-Ramos, M., Pingarroni, A., Rodríguez-Velázquez, J., Toledo-Chelala, L., Zermeño-Hernández, I., Bongers, F. (2016) Natural forest regeneration and ecological restoration in human-modified tropical landscapes. *Biotropica*, 48(6), 745-757.

Mesquita, R.D.C.G., Massoca, P.E.D.S., Jakovac, C.C., Bentos, T.V., Williamson, G.B. (2015) Amazon rain forest succession: Stochasticity or land-use legacy? *BioScience*, 65(9), 849-861.

Millington, J.D.A. (2019) *CRAFTY-Brazil Input Maps (Version v.1.0.0)*. [Online] Available at: http://doi.org/10.5281/zenodo.3549788

Millington, J.D.A. (2020a) *CRAFTY-Brazil Inputs (Version v1.0.0)*. [Online] Available at: http://doi.org/10.5281/zenodo.3746050

Millington, J.D.A. (2020b) *Brazil Agri Analysis (Version v1.0.0)*. [Online] Available at: http://doi.org/10.5281/zenodo.3746125

Millington, J.D.A. (2020c) *CRAFTY-Brazil (Version v1.0.1)*. [Online] Available at: http://doi.org/10.5281/zenodo.3746072

Millington, J.D.A., Xiong, H., Peterson, S. and Woods, J. (2017) Integrating modelling approaches for understanding telecoupling: global food trade and local land use, *Land*, 6(3) 56.

MMA (2019) *i3Geo - Ministério do Meio Ambiente* [Online] Available at: https://www.mma.gov.br/governanca-ambiental/geoprocessamento

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

Neto, A.D.O. (2017) *A produtividade da soja: analise e perspectivas. techreport Compendio de Estudos Conab 10*. Brasilia: Diretoria de Politica Agricola e Informacoes, Superintendencia de Informacoes do Agronegocio, Companhia Nacional de Abastecimento.

O'Sullivan, D., Evans, T., Manson, S., Metcalf, S., Ligmann-Zielinska, A. and Bone, C., (2016) Strategic directions for agent-based modeling: avoiding the YAAWN syndrome. *Journal of Land Use Science*, 11(2), 177-187.

Pereira, A.R. and Pruitt, W.O., (2004) Adaptation of the Thornthwaite scheme for estimating daily reference evapotranspiration. *Agricultural Water Management*, 66(3), 251-257.

Pereira, P.A.A., Martha, G.B., Santana, C.A. and Alves, E. (2012) The development of Brazilian agriculture: future technological challenges and opportunities. *Agriculture & Food Security*, 1(1), 4.

Plevin, R, Beckman, J, Golub, A, Witcover, J, & O'Hare, M. (2015) Carbon accounting and economic model uncertainty of emissions from biofuels-induced land use change. *Environmental Science and Technology*, 49, 2656-64.

Porter, M.E. (2000) Location, competition, and economic development: local clusters in a global economy. *Economic Development Quarterly*, 14(1), 15–34.

Picoli, M.C., Rorato, A., Leitão, P., Camara, G., Maciel, A., Hostert, P. and Sanches, I.D.A. (2020) Impacts of Public and Private Sector Policies on Soybean and Pasture Expansion in Mato Grosso—Brazil from 2001 to 2017. *Land*, 9(1), 20.

Rada, N. (2013) Assessing Brazil's Cerrado agricultural miracle. *Food Policy*, 38, 146-155.

Rezende, G.C.D. (2002) *Ocupação agrícola e estrutura agrária no cerrado: o papel do preço da terra, dos recursos naturais e da tecnologia*. IPEA: Rio de Janeiro.

Rosa, I.M., Purves, D., Souza Jr, C., Ewers, R.M. (2013) Predictive modelling of contagious deforestation in the Brazilian Amazon. *PloS one*, 8(10), 77231.

Rosolem, R., Almagro, A., Oliveira, P.T.S. and Hagemann, S. (2018) Performance Evaluation of HadGEM2-ES and MIROC5 Downscaled Rainfall Simulations over Brazil. *AGU Fall Meeting Abstracts*, A21L-2890.

Sayer, A. (1992) *Method in Social Science: A Realist Approach*. Abingdon: Routledge.

Silva, R.F.B., Batistella, M., Moran, E., Celidonio, O.L.D.M. and Millington, J.D.A. (2020) The Soybean Trap: Challenges and Risks for Brazilian Producers. *Frontiers in Sustainable Food Systems*, 4, 12.

Silva, R.F.B., Batistella, M., Dou, Y., Moran, E., Torres, S.M., Liu, J. (2017) The Sino-Brazilian Telecoupled Soybean System and Cascading Effects for the Exporting Country. *Land*, 6(3), 53.

Soares-Filho, B.S., Nepstad, D.C., Curran, L.M., Cerqueira, G.C., Garcia, R.A., Ramos, C.A., Voll, E., McDonald, A., Lefebvre, P., Schlesinger, P. (2006) Modelling conservation in the Amazon basin. *Nature*, 440, 520-523.

Sun, J., TONG, Y.X. and Liu, J. (2017) Telecoupled land-use changes in distant countries. *Journal of Integrative Agriculture*, 16(2), 368-376.

Sun, Z., Lorscheid, I., Millington, J.D.A., Lauf, S., Magliocca, N.R., Groeneveld, J., Balbi, S., Nolzen, H., Müller, B., Schulze, J. and Buchmann, C.M. (2016) Simple or complicated agent-based models? A complicated issue. *Environmental Modelling & Software*, 86, 56-67.

Thornthwaite, C.W., Mather, J. R. (1955) *The Water Balance*. Drexel Institute of Technology: Philadelphia, PA, USA.

Vera-Diaz, M.C., Kaufmann, R.K.R.K., Nepstad, D.C.D.C., Schlesinger, P (2008) An interdisciplinary model of soybean yield in the Amazon Basin: the climatic, edaphic, and economic determinants. *Ecological Economics*, 65 (2), 420–431.

Victoria, D.C., Santiago, A.V., Ballester, M.V.R., Pereira, A.R., Victoria, R.L., Richey, J.E. (2007) Water balance for the Ji-Paraná river basin, western Amazon, using a simple method through Geographical Information Systems and Remote Sensing. *Earth Interactions*, 11(5), 1-22.

Victoria, D.C., Silva, R.F.B., Millington, J.D.A., Katerinchuk, V. and Batistella, M. (2021) Transport Cost to Port though the Brazilian Federal Roads Network: Dataset for Years 2000, 2005, 2010 and 2017 *Data In Brief*

Warner, E., Inman, D., Kunstman, B., Bush, B., Vimmerstedt, L., Peterson, S., Macknick, J., Zhang, Y. (2013) Modeling biofuel expansion effects on land use change dynamics. *Environmental Research Letters*, 8(1), 015003

Wesz, V. J. (2016) Strategies and hybrid dynamics of soy transnational companies in the Southern Cone. *Journal of Peasant Studies*, 43, 286–312.

Weinhold, D. and Reis, E. (2008) Transportation costs and the spatial distribution of land use in the Brazilian Amazon. *Global Environmental Change*, 18(1), 54-68.