A multi-data assessment of land use and land cover emissions from Brazil during 2000–2019

ENVIRONMENTAL RESEARCH
LETTERS

OPEN ACCESS

RECEIVED
11 December 2020

REVISED
2 June 2021

ACCEPTED FOR PUBLICATION
7 June 2021

PUBLISHED
22 June 2021

LETTER

A multi-data assessment of land use and land cover emissions from Brazil during 2000–2019

Thais M Rosan[1,*], Kees Klein Goldewijk[2,3], Raphael Ganzenmüller[4], Michael O'Sullivan[5], Julia Pongratz[4,6], Lina M Mercado[1,7], Luiz E O C Aragao[1,8], Viola Heinrich[9], Celso Von Randow[10], Andrew Wiltshire[1,11], Francesco N Tubiello[12], Ana Bastos[13], Pierre Friedlingstein[5] and Stephen Sitch[1]

1 Department of Geography, College of Life and Environmental Sciences, University of Exeter, Exeter, United Kingdom
2 Copernicus Institute of Sustainable Development, Utrecht University, Utrecht, The Netherlands
3 PBL Netherlands Environmental Assessment Agency, The Hague, The Netherlands
4 Department of Geography, Ludwig-Maximilians-Universität, Munich, Germany
5 College of Engineering, Mathematics, and Physical Sciences, University of Exeter, Exeter, United Kingdom
6 Max Planck Institute for Meteorology, Hamburg, Germany
7 UK Centre for Ecology and Hydrology, Wallingford, United Kingdom
8 Earth Observation and Geoinformatics Division, National Institute for Space Research (INPE), São José dos Campos, Brazil
9 School of Geographical Sciences, University of Bristol, Bristol, United Kingdom
10 General Coordination of Earth Sciences, National Institute of Space Research, Sao Jose dos Campos, Brazil
11 Met Office Hadley Centre, Exeter, United Kingdom
12 FAO, Statistics Division, Rome, Italy
13 Department Biogeochemical Integration, Max Planck Institute for Biogeochemistry, Jena, Germany
* Author to whom any correspondence should be addressed.

E-mail: T.Rosan@exeter.ac.uk

Keywords: land-use and land cover change, deforestation, land-use emissions, global carbon budget

Supplementary material for this article is available online

Abstract

Brazil is currently the largest contributor of land use and land cover change (LULCC) carbon dioxide net emissions worldwide, representing 17%–29% of the global total. There is, however, a lack of agreement among different methodologies on the magnitude and trends in LULCC emissions and their geographic distribution. Here we perform an evaluation of LULCC datasets for Brazil, including those used in the annual global carbon budget (GCB), and national Brazilian assessments over the period 2000–2018. Results show that the latest global HYDE 3.3 LULCC dataset, based on new FAO inventory estimates and multi-annual ESA CCI satellite-based land cover maps, can represent the observed spatial variation in LULCC over the last decades, representing an improvement on the HYDE 3.2 data previously used in GCB. However, the magnitude of LULCC assessed with HYDE 3.3 is lower than estimates based on MapBiomas. We use HYDE 3.3 and MapBiomas as input to a global bookkeeping model (bookkeeping of land use emission, BLUE) and a process-based Dynamic Global Vegetation Model (JULES-ES) to determine Brazil's LULCC emissions over the period 2000–2019. Results show mean annual LULCC emissions of 0.1–0.4 PgC yr^{-1}, compared with 0.1–0.24 PgC yr^{-1} reported by the Greenhouse Gas Emissions Estimation System of land use changes and forest sector (SEEG/LULUCF) and by FAO in its latest assessment of deforestation emissions in Brazil. Both JULES-ES and BLUE now simulate a slowdown in emissions after 2004 (-0.006 and -0.004 PgC yr^{-2} with HYDE 3.3, -0.014 and -0.016 PgC yr^{-2} with MapBiomas, respectively), in agreement with the Brazilian INPE-EM, global Houghton and Nassikas book-keeping models, FAO and as reported in the 4th national greenhouse gas inventories. The inclusion of Earth observation data has improved spatial representation of LULCC in HYDE and thus model capability to simulate Brazil's LULCC emissions. This will likely contribute to reduce uncertainty in global LULCC emissions, and thus better constrains GCB assessments.

1. Introduction

Brazilian ecosystems and especially forests play a fundamental role in regional and global carbon stocks and natural land C sinks. The Amazon forest is estimated to contain around 229–280 PgC in living biomass and soils (Malhi *et al* 2006, Gloor *et al* 2012), representing ∼10% of global land C stocks (Ciais *et al* 2013), and approximately 60% of its area is in Brazil. Carbon stocks in Brazilian ecosystems have been negatively impacted by significant land use and land cover change (LULCC) associated with demographic and agricultural expansion, resulting in large land-use emissions to the atmosphere (Houghton 2012, Aide *et al* 2013). Globally, in the last decade (2009–2018), a total of 1.5 ± 0.7 PgC yr^{-1} was released to the atmosphere due to LULCC (Friedlingstein *et al* 2019). Despite a significant slowdown in deforestation in Brazil after a peak in 2004, mainly due to policy introduced to curb deforestation (Arima *et al* 2014, Godar *et al* 2014, West and Fearnside 2021), Brazil is still contributing with between 17% and 29% of global LULCC emissions (ELUC) (Friedlingstein *et al* 2019). However, divergent ELUC estimates for Brazil in the global carbon cycle budget have contributed to a large fraction of the corresponding overall global uncertainty (Bastos *et al* 2020).

In the global carbon budget (GCB 2020), ELUC is defined as the net anthropogenic LULCC flux and includes removals (e.g. from forest regrowth after harvest and agricultural abandonment) and emissions e.g. from clearing natural vegetation and transitions (Friedlingstein *et al* 2020). These definitions are different from those used by countries and FAO to estimate and report emissions from LULCC within the IPCC LULUCF category of the national greenhouse gas inventory (NGHGI). Within the IPCC guidelines, LULCC is limited to emissions associated only to anthropogenic-related processes driven changes in land use and land cover (see SI table 3 (available online at stacks.iop.org/ERL/16/074004/mmedia)). In the GCB, differently, the ELUC is estimated by two different bottom-up approaches, namely process-based and bookkeeping models (Friedlingstein *et al* 2019). An ensemble of process-based, Dynamic Global Vegetation Models (DGVMs) from the Trends in Net Carbon Exchange Project (TRENDY) (Sitch *et al* 2015), are applied using observed historical CO_2 in atmosphere, climate and LULCC fields.

An additional method is the carbon stock change approach of the IPCC (2003), followed by FAO for its ELUC estimates (Tubiello *et al* 2020). The advantage of this method is the possibility to perform complex calculations using a very small set of input data, while on the other hand it cannot distinguish necessarily between natural and anthropogenic fluxes. Bookkeeping models include Houghton and Nassikas (H&N, 2017) and bookkeeping of land use emissions (BLUE) (Hansis *et al* 2015). Both TRENDY and BLUE use the same LULCC dataset as spatially explicit input based on the History Database of the Global Environment (HYDE 3.2) (Goldewijk *et al* 2017) for annual change in pasture, rangeland and cropland area. There are many uncertainties related to these global LULCC datasets, since they use FAO statistics as input, which are provided as national aggregates (FAO 2020a) and then rely on a suite of methods to disaggregate that information spatially. With the increasing availability of land cover datasets based on Earth observation (EO) covering the last 30 years, it has been possible to integrate time varying remote sensing data with the FAO national statistics to generate new and improved spatially explicit global LULCC datasets. Several EO-based products of LULCC and deforestation have been developed in Brazil, such as the Amazon Deforestation Monitoring Project (PRODES) from the National Institute for Space Research (INPE), and the MapBiomas dataset, which was developed specifically for Brazilian biomes and provide annual LULCC maps for the whole of Brazil from 1985 up to present. However, these datasets have hitherto not been used to assess the impact of LULCC in global C-cycle assessments. Therefore there is an urgent need to improve estimates of ELUC for Brazil to better represent the spatio-temporal trends in future GCB annual assessments. Furthermore, accurate estimates of ELUC contribute to the quantification of emissions and removals from LULUCF processes that are needed to guide global and national policies to achieve the overarching goal of the Paris Agreement (UNFCCC 2015).

Here we present a critical analysis and evaluation of LULCC in Brazil and associated C emissions over the 21st century (2000–2019). We use two versions of the HYDE land use dataset, HYDE 3.2 (Goldwijk *et al* 2017; used in GCB 2019) and the new HYDE 3.3 based on updated FAOSTAT statistics (FAO 2020a, 2020b) and time-varying land cover data from the European Space Agency Climate Change Initiative (ESACCI-LC 2017). First, we evaluate the two versions of HYDE using the national LULCC MapBiomas product. Then, we compare simulated ELUC for Brazil based on the two HYDE versions and MapBiomas using the process-based JULES-ES and BLUE bookkeeping models. Finally, we discuss simulated ELUC dynamics from BLUE and JULES-ES using HYDE 3.3 and MapBiomas with the other published estimates.

2. Methods

2.1. Land use datasets

2.1.1. HYDE

HYDE is a spatially explicit dataset of historical population estimates and time-dependent weighting maps of land use categories (Goldewijk *et al* 2017). The period covered is 10 000 Before the Common Era to 2019 Common Era. The land use maps produced by HYDE are based on an allocation algorithm that uses country totals from FAOSTAT statistical data of 'Cropland' and 'Permanent meadows and pastures' (henceforth indicated as grazing in HYDE dataset) available from 1961 up to present (FAO 2020b). In addition, HYDE includes the ESA CCI Land Cover maps to spatially allocate the FAO land use areas.

HYDE is available at 5 arc minutes (approximately 9 km at the equator) of spatial resolution. The grazing land use category has a distinction based on the intensity of use; grazing is divided in rangelands (extensive grazing on natural grasslands, shrublands, woodlands, wetlands and deserts), managed pastures (intensive grazing or mowing, on any natural vegetation type) and converted rangeland (located in forest biomes in areas with low human population density, and assumed to have undergone a conversion of natural vegetation, such as the Amazon biome).

There are two major updates from HYDE version 3.2 to HYDE version 3.3. First, the ESA CCI land cover data is now used for allocation of cropland and grazing land on a yearly basis for the period 1992–2018, instead of only the base year 2010 (HYDE 3.2). The method of reclassifying the ESA-CCI classes into cropland and grazing land remained the same and is described in Goldewijk *et al* (2017). Second, updated FAO statistics for cropland and grazing land are used and extended to the year 2018 (the last available FAOSTAT year).

Importantly, major FAOSTAT revisions made in 2019 for agricultural area in Brazil, reflecting new data from the national Census 2017 (e.g. see country notes in FAOSTAT 2020) impacted HYDE 3.3 total areas.

2.1.2. MapBiomas

The Brazilian Annual Land Use and Land Cover Mapping Project (MapBiomas) is an initiative to produce annual LULCC maps for Brazil for the Greenhouse Gas Emissions Estimation System (SEEG) (De Azevedo *et al* 2018) from the Brazil Climate Observatory's. This dataset is produced using the Google Earth Engine platform and the historical Landsat satellite images (Souza *et al* 2020). The classification process consists of using annual Landsat mosaics composed of pixels filtered by cloud coverage and ancillary information to classify each year. The LULCC maps produced by MapBiomas have a spatial resolution of 30 m and span over the 1985–2019 period for the collection (5) used in this research (http://mapbiomas.org/). The overall accuracy reported for entire Brazil classification was 89%, the method and explanation of the validation process is on Souza *et al* (2020). Given its higher spatial resolution of 30 m compared to the global products, such as ESA CCI LC at 300 m (with change detection at 1 km resolution), MapBiomas enables relatively small changes in LULCC to be detected across the whole country in both space and time.

2.2. Estimates of ELUC

2.2.1. Process-based approach

Process-based DGVMs simulate dynamics of carbon pools in vegetation, soil and wood products, and their response to changing environmental conditions. A consortium of international research groups (TRENDY) contributes annually to GCB with an ensemble of DGVMs, applying their models with common meteorological forcing and LULCC datasets to estimate the natural land sink and ELUC, and to attribute changes in the carbon cycle to individual environmental drivers at multiple temporal and spatial scales (Sitch *et al* 2015). In this study we use the JULES-ES (Joint UK Land Environment Simulator—Earth System configuration) (Sellar *et al* 2019) model, which also contributes to TRENDY and GCB (Friedlingstein *et al* 2019). JULES-ES has detailed representation of land surface processes (e.g. surface energy balance, coupled carbon and water cycle) and includes recent developments in surface physical processes (Wiltshire *et al* 2020b), the representation of plant physiology and plant functional types (Harper *et al* 2016, 2018), land use and nitrogen cycling (Wiltshire *et al* 2020a), dynamic vegetation (Cox 2001, Harper *et al* 2018), and wood products (Jones *et al* 2011). Additionally, it simulates natural vegetation cover, and human activities (e.g. land-use) can be prescribed with anciliary data representing annual cropland and pasture fractions, for example from the HYDE dataset.

2.2.2. Bookkeeping approach

Bookkeeping models track changes in the carbon stored in vegetation, soils and products before and after LULCC using prescribed rates of growth and decay through time. Unlike DGVMs they do not include the effect of changing environmental factors on vegetation growth rates (e.g. climate and CO_2 fertilization). Instead of simulating carbon stocks, bookkeeping models use directly observational data for carbon densities, such as literature-based biome-level values and from inventory data. Two main global bookkeeping models are used in GCB: the H&N model (Houghton and Nassikas 2017) and the BLUE model (Hansis *et al* 2015). There is also a regional bookkeeping model developed for Brazil by the National Institute for Space Research (INPE), INPE-EM bookkeeping model (Aguiar *et al* 2012, Assis *et al* 2020).

In this study we focus on BLUE to test the impact of the new HYDE 3.3 and MapBiomas, as it calculates ELUC on a spatially explicit basis for transitions from natural vegetation types to agricultural lands. Specifically, it considers transformations of natural vegetation to agriculture (cropland, pasture) and back, including gross transitions at the sub-grid scale ('shifting cultivation'), transitions between crop and pasture, and wood harvesting (Hansis *et al* 2015). Biome-level carbon densities are based on literature values and provided in Hansis *et al* (2015) (SI figure 2). Similarly, the temporary evolution of carbon gain or loss, i.e. how fast carbon pools decay or regrow following a land-use change, is based on response curves derived from literature (Hansis *et al* 2015). The response curves describe decay of vegetation and soil carbon, including transfer to product pools of different lifetimes, as well as carbon uptake due to regrowth of vegetation and subsequent refilling of soil carbon pools.

2.3. Analyses

2.3.1. Land use change analysis

To assess the differences between the LULCC products for Brazil and to understand LULCC dynamics, HYDE was evaluated against the national MapBiomas dataset. MapBiomas was processed as follows: first each category is reclassified with the proportion of crop and pasture at 30 m spatial resolution. The following class aggregation applies to cropland for MapBiomas: the categories, Annual and Perennial Crop and Semi-Perennial Crop were defined to contain 100% cropland in each pixel; the category Mosaic of Agriculture and Pasture as 40% cropland. For pastures we consider Pasture category as 100% and Mosaic of Agriculture and Pasture to contain 60% of pasture as most of these mosaic categories are used in extensive cattle ranching and small-scale agriculture. Then we re-grid to 0.5 degree to generate gridded maps of cropland and pasture cover fraction.

The resulting maps were compared against ESA CCI cropland categories 10, 11, 12, 20 which contain 100% cropland (table 1 SI description). Further categories were included, such as mosaic categories 30 (cropland >50%), 40 (cropland <50%), assuming a cropland area proportion of 60% and 40%, respectively, based on (Liu *et al* 2018). Note, ESA CCI does not have an explicit 'pasture' category, rather it is included in the 'cropland' category. For representation, we report the sum of cropland and pasture categories. These maps between land cover and land use categories introduce a level of uncertainty that is currently poorly quantified. Furthermore, the choice of the underlying land cover map introduces uncertainties, for instance results discussed herein are likely different from those that would be obtained using MODIS land cover maps rather the ESA CCI (e.g. see FAO 2020b).

To test the spatial similarity of the LULCC maps from HYDE against MapBiomas we apply the fuzzy numerical method implemented in the Map Comparison Kit 3 application (Visser and De Nijs 2006). It is a cell-by-cell comparison method for numerical maps that also considers the neighbourhood to show the similarity of each pair of grid-cells in a range between 0 (distinct) and 1 (identical). Here, we adopt the default settings provided by the algorithm, with an exponential decay function called Halving distance equal to 2 and neighbourhood distance equal to 4 grid-cells. This provides a spatial assessment showing the location and severity of the disagreement between two maps.

Since agricultural area change is currently used to infer tree cover loss in DGVMs, we calculated the pairwise Pearson correlation between increase of net land cover (cropland + pasture) with the vegetation cover loss (deforestation) from MapBiomas c5 including both primary and secondary vegetation and the following natural vegetation categories in MapBiomas: forest, savanna and natural grassland. We performed this correlation analysis for the changes between 2000 and 2019 comparing the grid-cells with increase in LULCC from HYDE 3.2, HYDE 3.3 and MapBiomas to grid-cells with vegetation cover loss for the same period.

2.3.2. Carbon emissions from LULCC

HYDE 3.3 and MapBiomas fraction maps were regridded for using in JULES-ES at N96 resolution (1.25° latitude × 1.875° longitude) with Climate Data Operators (CDOs) using first-order conservative remapping method. For this analysis, we also considered converted rangeland fraction from HYDE 3.3 in order to better represent the LULCC changes across the Amazon arc of deforestation (low human population density but transitions to extensive pastures), which was not considered in JULES-ES, TRENDY-v9. We run JULES-ES using the LULCC fields (cropland, pasture and converted rangeland) from HYDE 3.3 and MapBiomas (cropland and pasture) with the same configuration as in TRENDY-v9 (GCB2020), including time varying climate, CO_2, and nitrogen deposition following the TRENDY protocol. The simulations using HYDE 3.3 extend over the period 1700–2019 whereas the MapBiomas simulations are over the more recent period 1986–2019. JULES-ES is spun-up to steady state conditions and then we perform the following experiments: S2 (CO_2 and climate forcing varying, land-use constant at 1700 for HYDE 3.3 and 1986 for MapBiomas) and S3 (CO_2, climate and land-use time variant) over 1700–2019 for HYDE 3.3 and 1986–2019 for MapBiomas. ELUC is diagnosed as the difference in net biome productivity at a grid-cell level between the two runs, i.e. S2–S3:

$$(\text{S2})\ \text{NBP} = \text{NPP} - \text{RH} - \text{HARVEST}$$

$$(\text{S3})\ \text{NBP} = \text{NPP} - \text{RH} - \text{HARVEST} - \text{PRODUCT\_DECAY}$$

where NPP is net primary productivity, RH is the heterotrophic respiration, and the human disturbances in this study are represented by Harvest which is the crop harvest and Product decay which is the decay flux from wood product pools (Jones *et al* 2011). Further explanation on the MapBiomas legacy flux estimation can be found in the supplementary information and in SI table 5.

BLUE used the same configuration as in GCB2019, however land-use forcing (LUH2) was replaced by the two HYDE versions and wood harvest as well as sub-grid scale transitions (i.e. shifting cultivation) were consequently not considered because HYDE only provides net area changes per land cover category, and does not consider gross transitions within our large grid-cell areas (>27 km). Both BLUE simulations were initialized in 1961 for HYDE and 1986 for MapBiomas. In order to minimize differences to the GCB2019 setup to allow a cleaner comparison, HYDE 3.2 and HYDE 3.3 were re-gridded to a spatial resolution of 0.25° × 0.25° with CDO using first-order conservative remapping method. The re-gridded HYDE data was processed to match the pre-processing of the land-use forcing data used in other BLUE simulations. In particular rangeland areas were considered to imply clearing of natural vegetation only when the forest/non-forest map of LUH2-original land cover data set in BLUE-indicates forest. Potential vegetated grid cell fractions not in cropland or pasture were split into primary and secondary land according to the proportion in LUH2. The set-up of BLUE simulations can be found at SI table 5.

Finally, we discuss our new ELUC results using HYDE 3.3, MapBiomas and other published estimates. We compared ELUC estimates from JULES-ES and BLUE (2000–2019) for the whole of Brazil with the H&N global bookkeeping model (Houghton and Nassikas 2017), the FAOSTAT (Tubiello *et al* 2020), and national inventory datasets: the 4th National Communication of GHG to the United Nations Framework Convention on Climate Change (UNFCCC) (Brazil MCTI 2020) and provides estimates up to 2016, and the SEEG LULUCF estimates based on MapBiomas dataset (De Azevedo *et al* 2018). Comparison of ELUC for Amazonia and Cerrado is done using estimates from INPE-EM for the biome boundaries. The INPE-EM ELUC (Aguiar *et al* 2012) estimate is based on the official deforestation data from the National Institute for Space Research (INPE) for both biomes using the 2nd order estimates (i.e. instantaneous and legacy C emissions). A table with the model components considered to estimate ELUC is in SI table 2.

3. Results

3.1. Land use and land cover changes in Brazil

We find temporal agreement between both remote sensing-based products (ESA CCI and MapBiomas) and HYDE 3.3 at country level. The three datasets agreed on the peak LULCC in Brazil between 2003 and 2005 with a negative trend thereafter opposite to the positive trend shown by HYDE 3.2 (figure 1(a)). This period corresponds to the peak of deforestation in the Brazilian Amazon between 2003 and 2004 and is followed by a slowdown in deforestation rates due to the implementation of governmental regulations to reduce deforestation thereafter. Additional economic factors (West and Fearnside 2021), and improvement in the use of verification and detection tools based on remote sensing, e.g. PRODES and the DETER program for near-real time deforestation detection also contributed to the deforestation slowdown in Brazil. However, both ESA CCI and consequently HYDE 3.3 showed lower LULCC after 2005 compared to MapBiomas. This may partially be due to the pervasive increase in small-scale deforestation (<1 ha) (Kalamandeen *et al* 2018), which may remain undetected using the 1 km change detection implemented in ESA CCI LC methodology (ESACCI-LC 2017). In addition, the total LULCC in HYDE 3.3 reflected the updated cropland and land under permanent meadow and pasture area in FAOSTAT (FAO 2020b) based on the decadal Brazilian Agriculture census. To balance the total land area, FAO increased the residual area in another category called 'other land' which is not used by the HYDE dataset and may include a proportion of the LULCCs associated with deforestation. Further explanation of the other land category can be found in the supplementary information. All these changes contribute to lower net LULCC observed using HYDE 3.3 data when compared to MapBiomas (figure 1(a)).

Spatially, there were large differences between the global LULCC products and MapBiomas (figure 2). HYDE 3.2 had the largest LULCC located in SE Brazil associated with cropland expansion and lower changes in pasture area. These changes are not consistent with MapBiomas (figure 2), were large losses of natural vegetation are found in Amazonia and the Cerrado biomes, and an intensification of cropland areas mostly concentrated in repurposed pastures in the SE Brazil (Zalles *et al* 2019). HYDE 3.2 used ESA CCI LC baseline for 2010 but to estimate year-to-year changes it employs an algorithm to allocate transitions within a country giving preference for conversion of lands near existing agriculture and with high NPP. This may not be sufficient to capture deforestation in more remote regions in a large country like Brazil, hence the LULCC allocation in HYDE 3.2 was centred in SE Brazil in consolidated areas.

Although HYDE 3.3 showed lower magnitude change compared to MapBiomas, the main gain with the updated version was the spatial allocation of LULCC. The similarity analysis showed that 20.2% of the grid-cells from HYDE 3.2 are dissimilar (similarity = 0) when compared to MapBiomas in grid-cells mainly located in the SE and NW Brazil (SI figure 3) indicating a larger spatial inconsistency in the LULCC for this region. HYDE 3.3 on the other

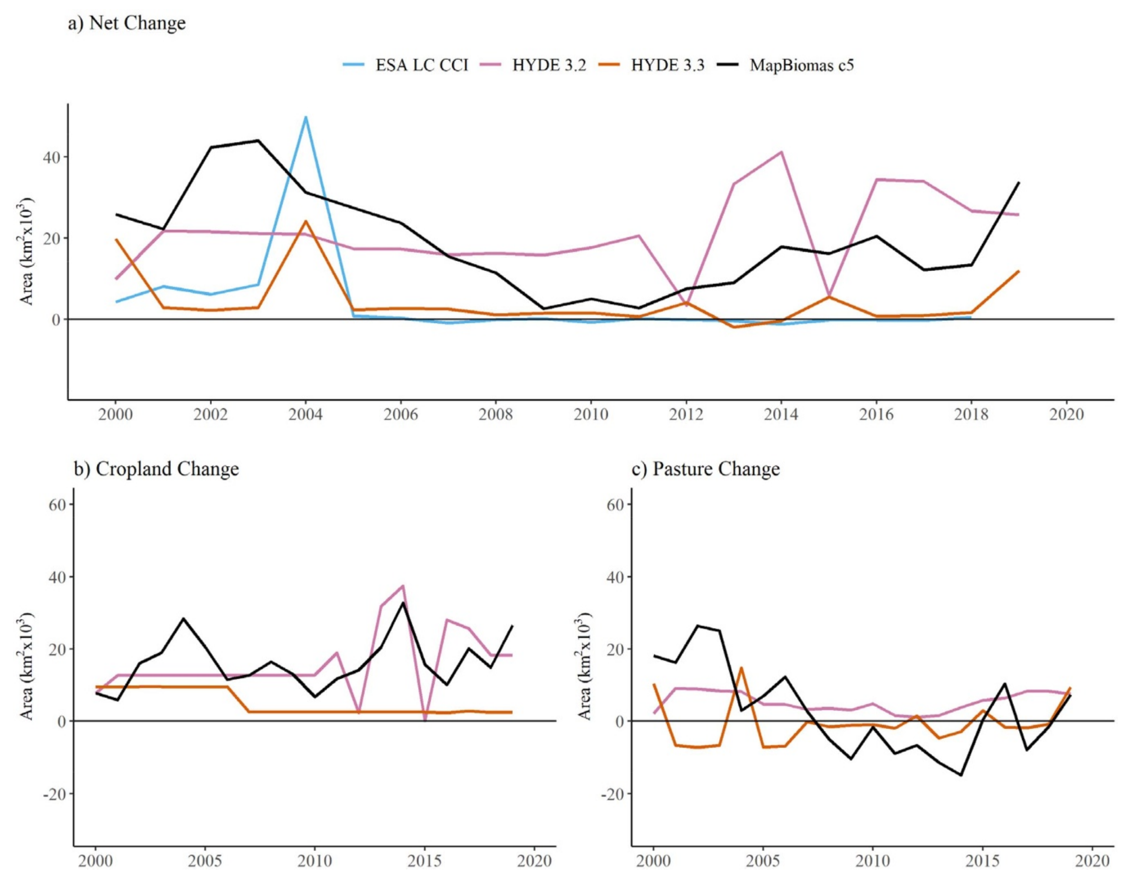

Figure 1. Annual land use change in Brazil (km^2/year) considering only the two main forcing: cropland and pasture* for HYDE 3.2, HYDE 3.3, ESA LC CCI and MapBiomas c5. (a) Net change (sum of cropland and pasture); (b) cropland category change and (c) pasture category change. *Pasture for both HYDE versions include pasture + converted rangeland categories.

hand, had only 2.2% of the grid-cells with similarity index equal to 0 when compared with MapBiomas and those grid-cells tend to be spatially spread, indicating an improvement on the spatial allocation. The spatial distribution of the similarity index and the frequency histogram is shown in SI figure 3. Additionally, the net changes from MapBiomas and natural vegetation cover loss showed a strong correlation ($R = 0.94$; SI figure 4), with net changes mainly due at expense of forest loss. This indicates that net observed LULCC change can be used as a proxy of the deforestation process. A pixelwise correlation comparison between both HYDE net LULCC versions and MapBiomas vegetation loss, indicated a superior performance in HYDE 3.3 ($R = 0.55$, SI figure 4) compared with HYDE 3.2 ($R = 0.094$; SI figure 4). This result demonstrates that although HYDE 3.3 still underestimates the changes, it is able to better reproduce and allocate spatially the deforestation pattern than HYDE 3.2.

3.2. Land use and land cover change emissions (ELUC) in Brazil

Simulated average ELUC for Brazil after the peak in 2004 was 0.34 PgC yr^{-1}, 0.18 PgC yr^{-1} and 0.32 PgC yr^{-1} in JULES-ES simulations with HYDE 3.2 (2005–2017), HYDE 3.3 and MapBiomas (2005–2019), respectively. ELUC with the BLUE model was 0.19 PgC yr^{-1}, 0.11 PgC yr^{-1} and 0.39 PgC yr^{-1} for HYDE 3.2. HYDE 3.3 and MapBiomas, respectively (SI table 4). Although simulations based on HYDE 3.2 showed higher emissions than HYDE. 3.3-based, both models when forced with HYDE 3.3 showed a negative trend after 2004. This is in agreement with the trend from our reference run based on MapBiomas LULCC forcing. Therefore, differences between MapBiomas and HYDE 3.3 simulations can reach up to 0.3 PgC yr^{-1} in some years due the higher LULCC in MapBiomas than HYDE datasets (figure 3(a)).

At the biome level, simulations with HYDE 3.3 and MapBiomas agreed on a downward trend after 2004 ($p < 0.05$, figure 3(b)) for the Amazon biome which is also consistent with INPE-EM bookkeeping model based on official deforestation data. However, for the same period simulations with HYDE 3.2 had an opposite direction and showed an increase of ELUC in both the Amazon and Cerrado biomes with the greatest increase in the latter (figure 3(c)). The update in HYDE also reflects an improvement in the contribution of each biome to the country total ELUC. Our results indicated that with the HYDE 3.2 version the Cerrado biome (Brazilian savannas)

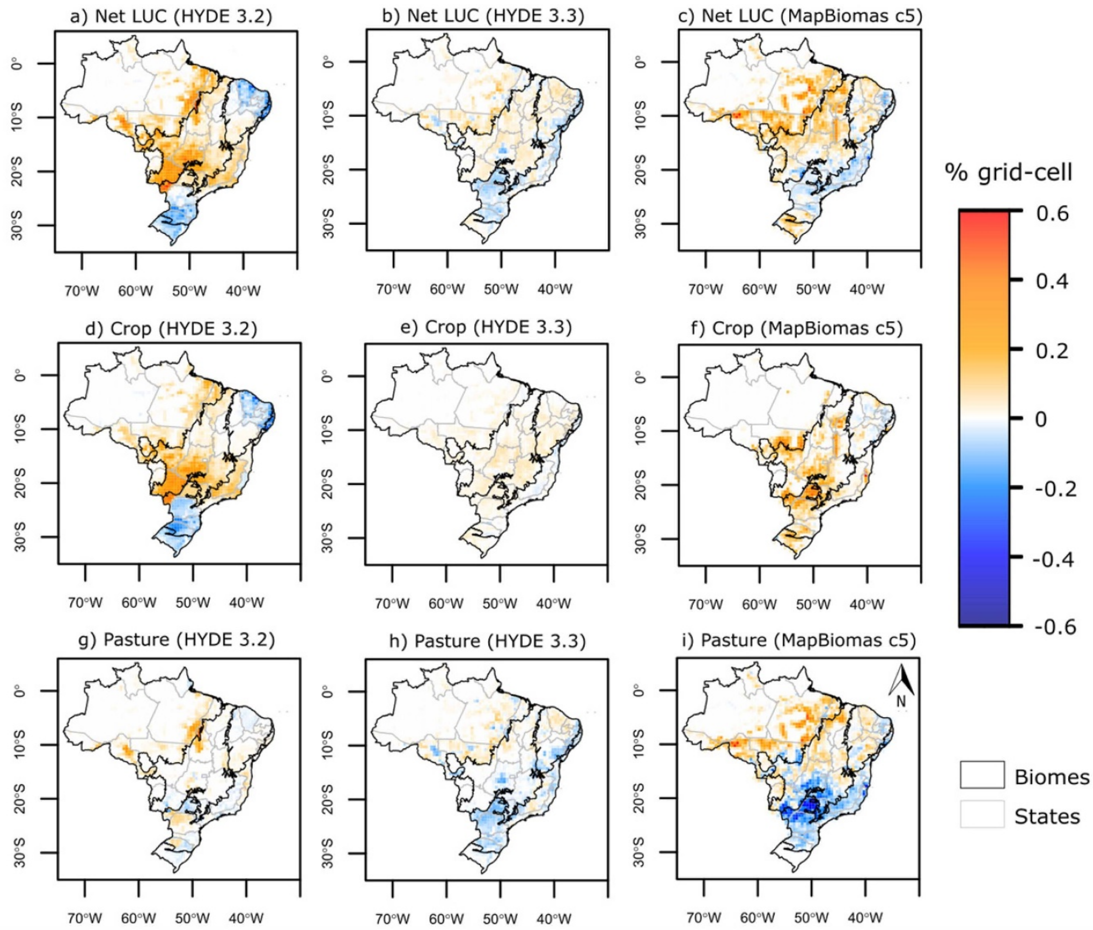

Figure 2. Spatial land use change fraction (0%–100% of grid-cell converted to 0–1) between 2000 and 2017 for Brazil. (a) net land use change in HYDE 3.2; (b) net land use change in HYDE 3.3; (c) net land use change MapBiomas c5; (d) cropland change HYDE 3.2; (e) cropland change HYDE 3.3; (f) cropland change MapBiomas c.5; (g) pasture change HYDE 3.3; (h) pasture change HYDE 3.3; (i) pasture change MapBiomas v.4. *Net land use change = cropland + pasture. **Pasture for both HYDE versions include pasture + converted rangeland categories.

was the highest contributor (40.7%–61%) in contrast to Amazon (14.6%–22.5%) (SI figure 5). This is opposite to data from INPE-EM, which confirms that Amazon emissions are on average 59% higher than Cerrado emissions. The spatial improvement in the allocation method of HYDE 3.3 increased the contribution of Amazon biome to the total ELUC, now responsible for about 40% and the Cerrado 24%–32% of the country's total (SI figure 5). These improvements are important to better represent spatially the human disturbances across the Brazilian biomes.

Spatially, the main differences between the HYDE-based simulations are in SE Brazil and the arc of deforestation in Amazonia (figures 3(d)–(i)). Both models using HYDE 3.2 simulate higher emissions concentrated in the SE Brazil (i.e. Sao Paulo, Mato Grosso do Sul and Parana states) (figures 3(b) and (d)) when compared to the Amazon region, which is unrealistic given the dynamic of recent trends in deforestation hotspots in Brazil (i.e. Amazon and north of Cerrado regions as shown in SI figure 4). Therefore, both HYDE-based simulations still spatially underestimate the emissions in the Amazon 'Arc of deforestation' as shown by the spatial simulated emissions based on MapBiomas (figures 3(f) and (i)).

4. Discussion

The goal of this study was to evaluate global land use products used to estimate Brazil ELUC emissions comparing with additional country-specific data (MapBiomas) and estimate the impact of the new HYDE 3.3 dataset on the simulated ELUC. One of the factors of uncertainty and disagreement in the mean and trends in ELUC for Brazil is the driving LULCC dataset (Bastos *et al* 2020, Gasser *et al* 2020). Our results show that the HYDE 3.3 including updated FAO statistics and an improvement in the allocation method using multi-annual remote sensing-based ESA CCI land cover) product better distributed the LULCC and consequently ELUC across Brazil compared to the previous version. Unlike ELUC based on HYDE 3.2, our new HYDE 3.3-based estimates now agree on an overall negative trend after

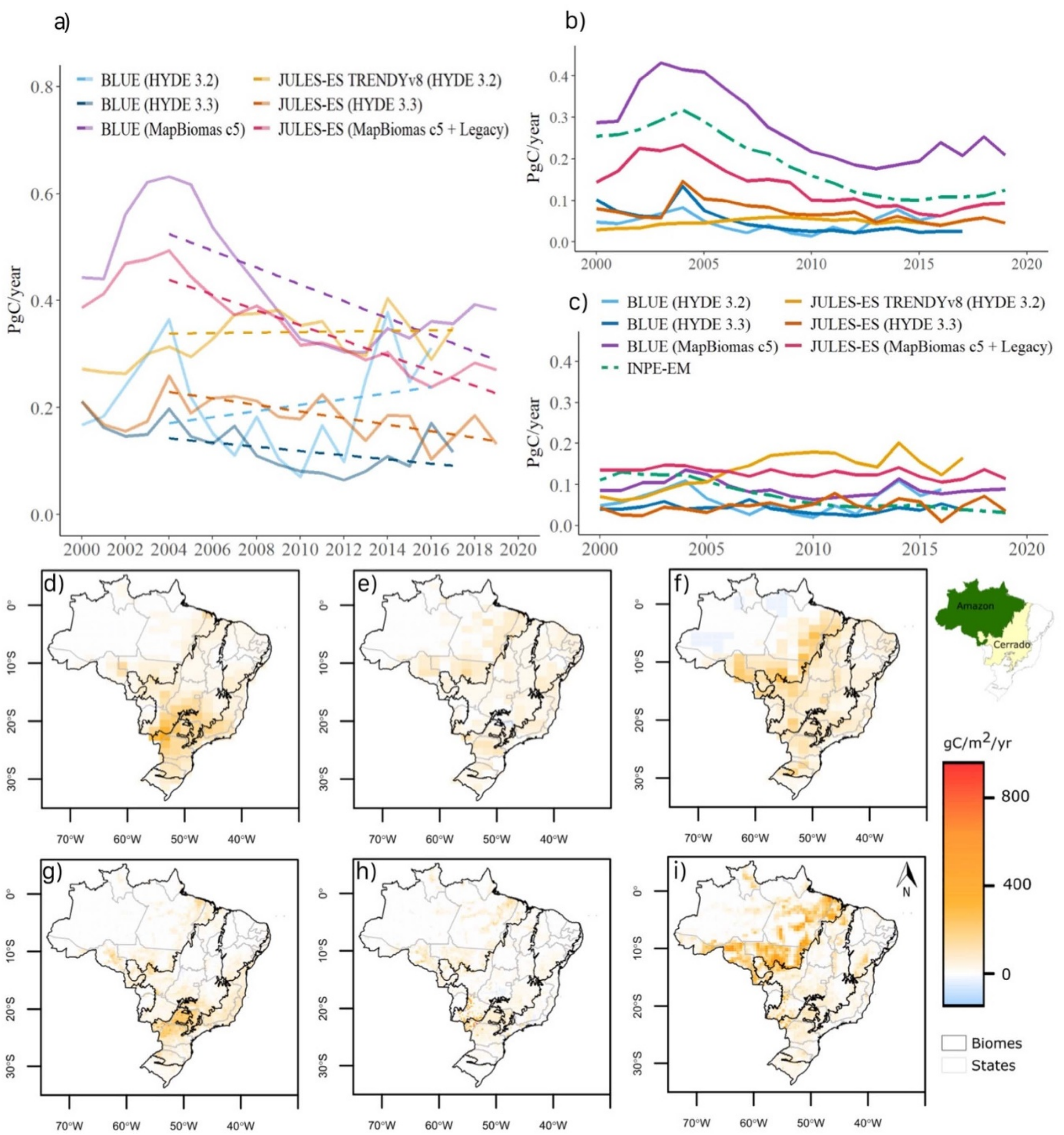

Figure 3. (a) Brazil total ELUC (PgC yr^{-1}) from 2000 and trendline from 2004 onwards; (b) Amazon biome ELUC (PgC yr^{-1}); (c) Cerrado biome ELUC (PgC yr^{-1}); (d) Simulated spatial explicit mean ELUC (gC m^{-2} yr^{-1}) for Brazil (2000–2016*) by JULES-ES TRENDYv8 with HYDE 3.2; (e) JULES-ES with HYDE 3.3; (f) JULES-ES with MapBiomas c5 + Legacy; (g) BLUE with HYDE 3.2; (h) BLUE with HYDE 3.3; (i) BLUE with MapBiomas c5. *Here we spatially show 2000–2016 mean and difference because BLUE (HYDE 3.2) was only available up to 2016.

2004 as shown in our reference MapBiomas simulation and other national datasets (NGHGI and SEEG LULUCF), FAOSTAT and H&N global bookkeeping models (figure 4 and SI table 4). Although there are still differences in the scale of emissions between the estimates in figure 4, these can be attributed to different methodological approaches, processes considered, and input data used to estimate ELUC (SI table 3).

Challenges remain in order to further improve representation of LULCC and ELUC in regional and global assessments. In general, DGVMs and some bookkeeping models use change in agricultural areas based on statistical data reported by countries to FAO as a forcing for tree-cover loss. By using only the agricultural areas based on FAOSTAT such as in HYDE dataset, our results demonstrated that we underestimate the ELUC compared to remote sensing-based data. This occurs because of the limitation of relying on the country reports and the assumption that all changes in natural vegetation loss can be inferred from changes in agricultural land without consideration of part of the 'other land' category which may be associated to the deforestation process. A recent study showed that about 13% of the tree cover loss in Brazil goes into a long transitional land category which is land not converted to be used by agricultural activities and may be associated to land grabbing (Zalles *et al* 2021). Thus, these areas are potentially included as a residual in the 'other land' category from FAO and not being considered in the ELUC estimates based on HYDE dataset. When using a remote sensing-based

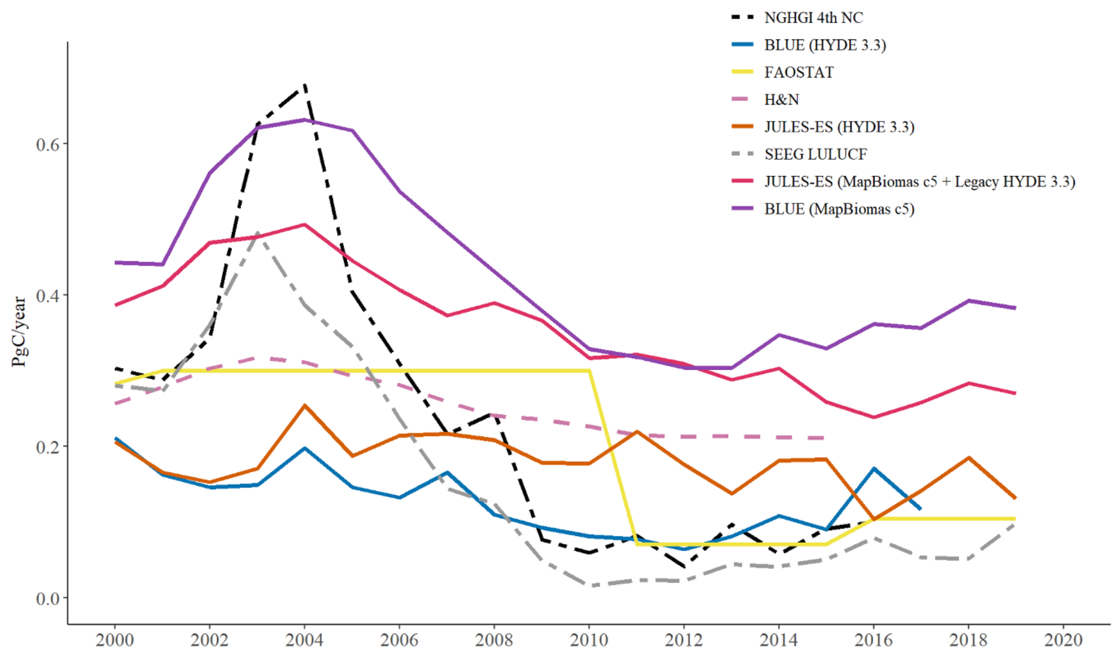

Figure 4. (a) Brazil ELUC from both models used in this study with HYDE 3.3 and MapBiomas and other estimates including the Houghton & Nassikas (H&N), the FAOSTAT, the 4th National Communication of the Brazilian estimates on GHG estimates (NGHGI 4th NC) and 8th edition of SEEG LULUCF. A table with the average estimates for Brazil and the trend are in SI table 4.

product to extract the observed changes in LULCC areas, such as the MapBiomas, the models were able to reproduce the overall pattern of vegetation loss and consequently ELUC trends over Brazil. Another caveat is that as we do not consider gross land cover changes and expect lower simulated ELUC in the early 2000s using the HYDE 3.3 dataset, also the regrowth sink in the late 2000s will likely be underestimated (Arneth *et al* 2017). A future challenge to the DGVMs and bookkeeping models such as BLUE will be to incorporate tree-loss directly from remote-sensing products. Nonetheless, this presents other challenges, such as the representation of legacy fluxes which occur on a timescale longer than the current availability of remote sensing datasets (Pongratz *et al* 2014). We believe this shortcoming will be overcome on the medium-term as longer time-series become available from satellite data. Moreover, global LULCC products with medium to coarse spatial resolution may not capture the increasing small-scale deforestation in Amazon (Kalamandeen *et al* 2018) which can contribute to an underestimate in the LULCC and its interannual variability. Variations in vegetation cover loss are associated with economic changes and environmental policy (Macedo *et al* 2012, West and Fearnside 2021), climatic events such as ENSO (El Niño–Southern Oscillation), which can contribute to spread of fires in intact forests (Alencar *et al* 2006, Aragão *et al* 2018), facilitating land cover conversion through fire, and also government decisions (Barlow *et al* 2020, Cardil *et al* 2020), which can lead to a high interannual variability. Using datasets based on semi- and decadal scale values such as in the updated version of FAO land-use statistical data for Brazil will result in lower year-to-year variation in vegetation loss and ELUC variability from DGVMs and Bookkeeping models. As shown in SI table 3 approaches vary in terms of LU processes included, e.g. shifting cultivation/sub-grid transitions, wood harvest, each of which could lead to an increase in ELUC estimates (Arneth *et al* 2017). In addition, forest degradation (selective logging, forest fire, edge-effects and fragmentation) is a growing threat, and may surpass deforestation in terms of both area and C emissions in several recent years (Aragão *et al* 2018, Assis *et al* 2020, Bullock *et al* 2020, Matricardi *et al* 2020, Silva Junior *et al* 2020), but is still not included in global DGVMs, bookkeeping models nor in national inventories. Hence improvements in DGVMs and global Bookkeeping models to explicitly use remote sensing derived tree-cover loss dataset and represent degradation processes are needed to further improve the representation of human disturbances on ecosystems and deliver better estimates of ELUC. Further efforts are still needed to better align (or map) the concepts between ELUC and LULUCF (Grassi *et al* 2018).

5. Conclusion

This study used a new global LULCC dataset based on the integration of time varying remote sensing data with updated national statistics to generate a new global land use dataset. HYDE 3.3 is shown to be superior in the spatial allocation to an earlier version based on only a single year satellite baseline. In particular, it reproduces the general spatial pattern of LULCC across Brazil when compared with national datasets, however it still underestimates the

LULCC changes due to limitations associated with the assumption of a one-to-one correspondence between natural vegetation loss and changes in agricultural land based on statistical reports. When HYDE 3.3 was applied as input to a processed-based DGVM and a global bookkeeping model, both simulated a negative trend in ELUC for Brazil, in agreement with national and other global estimates. The simulations with both HYDE 3.3 and MapBiomas also identify the Brazilian Amazon as the largest contributor to the total country ELUC. In summary, improvements in LULCC datasets have resulted in consistent estimates of ELUC trends across different methodologies for Brazil. These advances will likely improve GCB estimates, contributing to reduced uncertainty in the global estimates and improve our understanding of the global carbon cycle.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

This research has been supported by the Newton Fund through the Met Office Climate Science for Service Partnership Brazil (CSSP Brazil), RECCAP2 project which is part of the ESA Climate Change Initiative (contract no. 4000123002/18/I-NB), and the H2020 European Institute of Innovation and Technology (4C; Grant No. 821003). V H was supported by a NERC GW4+ Doctoral Training Partnership studentship from the Natural Environment Research Council (NE/L002434/1). R G acknowledges support from the European Commission through Horizon 2020 Framework Programme (VERIFY, Grant No. 776810). FAOSTAT is produced with the support of its member states and FAO regular programme funding. L M M acknowledges the UK Natural Environment Research Council funding (UK Earth System Modelling Project, UKESM, Grant NE/N017951/1). L E O C A was supported by CNPq (processes 305054/2016-3 and 442371/2019-5). C vR was supported by CNPq grant 314780/2020-3.

T M R and S S conceived the ideas and designed the methodology, T M R, R G, J P, K K G and M O S compiled the data. K K G developed new HYDE3.3 datasets iterating with T M R on Brazil. R G applied BLUE, and T M R and M O S applied JULES-ES using HYDE3.3 and MapBiomas. T M R analyzed the data, and results discussed with all authors. T M R and S S led the writing of the manuscript to which all authors contributed. F N T provided support in terms of FAO data on land use and GHG emissions as well as in providing mapping between land use classes. The views expressed in this publication are the author's and do not necessarily reflect the views or policies of FAO. We thank Renan Milagres Lage Novaes from EMBRAPA and Giorgia De Santis from FAO to help clarify the nature of FAOSTAT land use data used in the manuscript. We also thank Clement Albergel from ESA for his helpful comments on the manuscript.

MapBiomas data is freely available at http://mapbiomas.org/ and ESA CCI LC at www.esa-landcover-cci.org. HYDE 3.3 will be available upon request to (kees.kleingoldewijk@pbl.nl). The simulations will be available upon reasonable request to the correspondent author.

ORCID iDs

Thais M Rosan https://orcid.org/0000-0003-0155-1739
Raphael Ganzenmüller https://orcid.org/0000-0002-2337-0915
Michael O'Sullivan https://orcid.org/0000-0002-6278-3392
Lina M Mercado https://orcid.org/0000-0003-4069-0838
Viola Heinrich https://orcid.org/0000-0003-0501-0032
Francesco N Tubiello https://orcid.org/0000-0003-4617-4690
Ana Bastos https://orcid.org/0000-0002-7368-7806

References

Aguiar A P D *et al* 2012 Modeling the spatial and temporal heterogeneity of deforestation-driven carbon emissions: the INPE-EM framework applied to the Brazilian Amazon *Glob. Change Biol.* **18** 3346–66

Aide T M, Clark M L, Grau H R, López-Carr D, Levy M A, Redo D, Bonilla-Moheno M, Riner G, Andrade-Núñez M J and Muñiz M 2013 Deforestation and reforestation of Latin America and the Caribbean (2001–2010) *Biotropica* **45** 262–71

Alencar A, Nepstad D and Del Carmen Vera Diaz M 2006 Forest understory fire in the Brazilian Amazon in ENSO and non-ENSO years: area burned and committed carbon emissions *Earth Interact.* **10** 1–17

Aragão L E O C *et al* 2018 21st Century drought-related fires counteract the decline of Amazon deforestation carbon emissions *Nat. Commun.* **9** 1–12

Arima E Y, Barreto P, Araújo E and Soares-Filho B 2014 Public policies can reduce tropical deforestation: lessons and challenges from Brazil *Land Use Policy* **41** 465–73

Arneth A *et al* 2017 Historical carbon dioxide emissions caused by land-use changes are possibly larger than assumed *Nat. Geosci.* **10** 79–84

Assis T O, De Aguiar A P D, Von Randow C, Melo De Paula Gomes D, Kury J N, Ometto J P H B and Nobre C A 2020 CO2 emissions from forest degradation in Brazilian Amazon *Environ. Res. Lett.* **15** 104035

Barlow J, Berenguer E, Carmenta R and França F 2020 Clarifying Amazonia's burning crisis *Glob. Change Biol.* **26** 319–21

Bastos A *et al* 2020 Sources of uncertainty in regional and global terrestrial CO2 exchange estimates *Glob. Biogeochem. Cycles* **34** e2019GB006393

Brazil M C T I 2020 Fourth national communication of Brazil to the united nations framework convention on climate change

Bullock E L, Woodcock C E, Souza C and Olofsson P 2020 Satellite-based estimates reveal widespread forest degradation in the Amazon *Glob. Change Biol.* **26** 2956–69

Cardil A *et al* 2020 Recent deforestation drove the spike in Amazonian fires *Environ. Res. Lett.* **15** 121003

Ciais P *et al* 2013 The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change *Change IPCC Clim.* 465–570 (www.ipcc.ch/report/ar5/wg1/docs/review/WG1AR5_SOD_Ch06_All_Final.pdf%5Cn, http://ebooks.cambridge.org/ref/id/CBO9781107415324A023)

Cox P M 2001 Description of the 'T RIFFID' D ynamic G lobal V e g etation M odel (available at: www.metoffice.gov.uk/media/pdf/9/h/HCTN_24.pdf)

De Azevedo T R *et al* 2018 SEEG initiative estimates of Brazilian greenhouse gas emissions from 1970 to 2015 *Sci. Data* **5** 1–43

ESACCI-LC 2017 LAND COVER CCI product user guide 2 (available at: https://climate.esa.int/media/documents/CCI_Land_Cover_PUG_v2.0.pdf)

FAO 2020a FAOSTAT Forest Land emissions (available at: https://climate.esa.int/media/documents/CCI_Land_Cover_PUG_v2.0.pdf)

FAO 2020b Land statistics: global, regional and country trends, 1961–2018 (available at: www.fao.org/economic/ess/environment/data/land-use-and-land-cover/en/)

Friedlingstein P *et al* 2019 Global carbon budget 2019 *Earth Syst. Sci. Data* **11** 1783–838

Friedlingstein P *et al* 2020 Global carbon budget 2020 *Earth Syst. Sci. Data* **12** 3269–340

Gasser T, Crepin L, Quilcaille Y, Houghton R A, Ciais P and Obersteiner M 2020 Historical CO_2 emissions from land use and land cover change and their uncertainty *Biogeosciences* **17** 4075–101

Gloor M *et al* 2012 The carbon balance of South America: a review of the status, decadal trends and main determinants *Biogeosciences* **9** 5407–30

Godar J, Gardner T A, Tizado E J and Pacheco P 2014 Actor-specific contributions to the deforestation slowdown in the Brazilian Amazon *Proc. Natl Acad. Sci.* **111** 15591–6

Goldewijk K K, Beusen A, Doelman J and Stehfest E 2017 Anthropogenic land use estimates for the Holocene—HYDE 3.2 *Earth Syst. Sci. Data* **9** 927–53

Grassi G *et al* 2018 Reconciling global-model estimates and country reporting of anthropogenic forest CO2 sinks *Nat. Clim. Change* **8** 914–20

Hansis E, Davis S J and Pongratz J 2015 Relevance of methodological choices for accounting of land use change carbon fluxes *Glob. Biogeochem. Cycles* **29** 1230–46

Harper A B, Cox P M, Friedlingstein P, Duran-Rojas C, Wiltshire A J, Jones C D, Williams K, Mercado L M and Sitch S 2018 Vegetation distribution and terrestrial carbon cycle in a carbon cycle configuration of JULES4.6 with new plant functional types *Geosci. Model Dev.* **11** 2857–73

Harper A B *et al* 2016 Improved representation of plant functional types and physiology in the joint UK land environment simulator (JULES v4.2) using plant trait information *Geosci. Model Dev.* **9** 2415–40

Houghton R A 2012 Carbon emissions and the drivers of deforestation and forest degradation in the tropics *Curr. Opin. Environ. Sustain.* **4** 597–603

Houghton R A and Nassikas A A 2017 Global and regional fluxes of carbon from land use and land cover change 1850–2015 *Glob. Biogeochem. Cycles* **31** 456–72

Jones C D *et al* 2011 The HadGEM2-ES implementation of CMIP5 centennial simulations *Geosci. Model Dev.* **4** 543–70

Kalamandeen M, Gloor E, Mitchard E, Quincey D, Ziv G, Spracklen D, Spracklen B, Adami M, Aragão L E O C and Galbraith D 2018 Pervasive rise of small-scale deforestation in Amazonia *Sci. Rep.* **8** 1600

Liu X, Yu L, Li W, Peng D, Zhong L, Li L, Xin Q, Lu H, Yu C and Gong P 2018 Comparison of country-level cropland areas between ESA-CCI land cover maps and FAOSTAT data *Int. J. Remote Sens.* **39** 6631–45

Macedo M N, DeFries R S, Morton D C, Stickler C M, Galford G L and Shimabukuro Y E 2012 Decoupling of deforestation and soy production in the southern Amazon during the late 2000s *Proc. Natl Acad. Sci* **109** 1341–6

Malhi Y *et al* 2006 The regional variation of aboveground live biomass in old-growth Amazonian forests *Glob. Change Biol.* **12** 1107–38

Matricardi E A T, Skole D L, Costa O B, Pedlowski M A, Samek J H and Miguel E P 2020 Long-term forest degradation surpasses deforestation in the Brazilian Amazon *Science* **369** 1378–82

Pongratz J, Reick C H, Houghton R A and House J I 2014 Terminology as a key uncertainty in net land use and land cover change carbon flux estimates *Earth Syst. Dyn.* **5** 177–95

Sellar A A *et al* 2019 UKESM1 : description and evaluation of the U.K. Earth system model *J. Adv. Model. Earth Syst.* **11** 4513–58

Silva Junior C H L *et al* 2020 Persistent collapse of biomass in Amazonian forest edges following deforestation leads to unaccounted carbon losses *Sci. Adv.* **6** eaaz8360

Sitch S *et al* 2015 Recent trends and drivers of regional sources and sinks of carbon dioxide *Biogeosciences* **12** 653–79

Souza C M *et al* 2020 Reconstructing three decades of land use and land cover changes in brazilian biomes with landsat archive and earth engine *Remote Sens.* **12** 2735

Tubiello F N, Pekkarinen A, Marklund L, Wanner N, Conchedda G, Federici S, Rossi S and Grassi G 2020 Carbon emissions and removals by forests: new estimates 1990–2020 *Earth Syst. Sci. Data Discuss.* in review

UNFCCC 2015 *Adoption of the Paris Agreement—Conference of the Parties COP 21 vol Adopt. Par*

Visser H and De Nijs T 2006 The map comparison kit *Environ. Model. Softw.* **21** 346–58

West T A P and Fearnside P M 2021 Brazil's conservation reform and the reduction of deforestation in Amazonia *Land Use Policy* **100** 105072

Wiltshire A J *et al* 2020a JULES-CN : a coupled terrestrial carbon-nitrogen scheme (JULES vn5. 1) *Geosci. Model Dev.* **14** 2161–86

Wiltshire A J, Carolina Duran Rojas M, Edwards J M, Gedney N, Harper A B, Hartley A J, Hendry M A, Robertson E and Smout-Day K 2020b JULES-GL7: the global land configuration of the joint UK land environment simulator version 7.0 and 7.2 *Geosci. Model Dev.* **13** 483–505

Zalles V *et al* 2019 Near doubling of Brazil's intensive row crop area since 2000 *Proc. Natl Acad. Sci. USA* **116** 428–35

Zalles V *et al* 2021 Rapid expansion of human impact on natural land in South America since 1985 *Sci. Adv.* **7** eabg1620