

Uncertainty in land use obscures global soil organic carbon stock estimates

Chengcheng Gang^{a,b,d,1}, Hao Shi^{c,d,1}, Hanqin Tian^{d,e,*}, Shufen Pan^d, Naiqing Pan^d,
Rongting Xu^d, Zhuonan Wang^d, Zihao Bian^d, Yongfa You^d, Yuanzhi Yao^d

^a Institute of Soil and Water Conservation, Northwest A&F University, Yangling 712100, China

^b Institute of Soil and Water Conservation, Chinese Academy of Sciences and Ministry of Water Resources, Yangling 712100, China

^c Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China

^d International Center for Climate and Global Change Research, College of Forestry & Wildlife Sciences, Auburn University, Auburn 36845, USA

^e Schiller Institute for Integrated Science and Society, Department of Earth and Environmental Sciences, Boston College, Chestnut Hill 02467, USA

ARTICLE INFO

Keywords:

Land use and land cover change
Soil organic carbon
Terrestrial ecosystem modeling
Remote sensing
Uncertainty analysis

ABSTRACT

The impact of land use and land cover change (LULCC) on soil organic carbon (SOC) stock is one of the most uncertain items in estimating the global C budget. Despite the improvements in satellite monitoring techniques and inventory data in recent decades, the uncertainty in modeled LULCC-induced SOC changes stemming from the choice of land use datasets remains largely unknown. Using a process-based model, the Dynamic Land Ecosystem Model (DLEM), we investigated global SOC changes during 1900–2018 driven by six LULCC datasets (i.e., LUH2-GCB2019, ESA CCI-LC, MODIS, GLASS-GLC, HH, and RF), which were generated by varied data sources and methodologies. The simulated global SOC stock was negatively affected by land conversions in all the LULCC datasets; however, the corresponding SOC loss was highly different ranging from 33.11 Pg C to 106.20 Pg C. Such significant differences in the global SOC stock estimates due to LULCC uncertainty were mainly located in boreal and temperate forests of the northern high latitudes and were most likely attributed to the LULCC-induced changes in vegetation net primary production. Meanwhile, regions exhibiting large divergence in relative changes of LULCC-induced SOC loss were mainly located in the low latitudes. When considering the interactive effects of LULCC with other environmental factors, the simulated SOC showed divergent trends, increasing in MODIS-, ESA CCI-LC-, and GLASS-GLC-based estimations, but decreasing in LUH2-GCB2019-, HH-, and RF-based estimations. These results highlight the importance of the accuracy of LULCC data in determining the global carbon budget. Future efforts are required for harmonizing satellite observations and inventory data both spatially and temporally to better represent the land conversion processes in terrestrial ecosystem modeling.

1. Introduction

Soil is the largest terrestrial carbon pool, of which the carbon stock is nearly four times greater than that of the vegetation biomass and three times greater than the atmosphere (Jobbágy and Jackson, 2000; Lal, 2004). Roughly two-thirds of soil carbon is held as soil organic carbon (SOC), which is the net balance between C input from decaying vegetation, fungal and bacterial necromass, and C output through erosion, leaching, and decomposition (Davidson and Janssens, 2006; Regnier et al., 2013; Smith, 2008; Tian et al., 2015). Due to its vast quantity, a slight change in SOC stock could lead to a drastic impact on the atmospheric CO₂ and CH₄ concentrations, and thereby, the climate system (Zhou et al., 2019). Meanwhile, SOC is a major determinant of soil

quality, providing substrates for microbial processes like respiration and nutrient cycling and thus further regulating plant growth and ecosystem sustainability (Lal, 2004; Ramesh et al., 2019; Reeves, 1997; Tian et al., 2015). Higher levels of SOC can improve soil structure, increase carbon sequestration capacity, and reduce the risk of erosion and nutrient leaching. Therefore, understanding the spatial and temporal patterns of SOC is crucial for global climate mitigation and sustainable soil management (Ramesh et al., 2019; Wang et al., 2019).

The magnitude of SOC is impacted by both abiotic and biotic factors, such as precipitation (Doetterl et al., 2015), temperature (Conant et al., 2011), topography (Seibert et al., 2007), soil properties (Tian et al., 2010), vegetation types (Jobbágy and Jackson, 2000), soil biota (Liang and Balser, 2011), and various human disturbances to the ecosystems,

* Corresponding author.

E-mail address: hanqin.tian@bc.edu (H. Tian).

¹ Chengcheng Gang and Hao Shi contributed equally to the article

particularly the land use and land cover change (LULCC). LULCC affects SOC stock in various pathways, including deforestation (Bautista-Cruz et al., 2012; Petrenko and Friedland, 2015), cultivation (Huo et al., 2013), afforestation (Poeplau and Don, 2013), cropland management (de Oliveira et al., 2015; Yu et al., 2018), and wetland/peatlands drainage (Gao et al., 2014), etc.

Quantifying the impacts of LULCC on the magnitude and variations of SOC has garnered increasing interest across the local (Tan et al., 2004), regional (Batjes, 2002; Guevara et al., 2020; Xie et al., 2007), and global scales (Hengl et al., 2017; Tian et al., 2015). Yue et al. (2020) estimated that LULCC accounted for 30–45% of the interannual variability of land carbon balance over 1959–2015. Machmuller et al. (2014) found that emerging land use practices had rapidly increased soil organic matter. Most research has attributed uncertainties in LULCC emissions to model structures and parameterization schemes (Luo et al., 2016; Scharlemann et al., 2014; Xu et al., 2020), whereas the role of LULCC data itself has been usually overlooked, especially in ensemble modeling initiatives (Tian et al., 2015; Todd-Brown et al., 2013).

Previous studies have demonstrated that the uncertainty in LULCC emissions is, in large part, related to the changing rate in land use, which underpins the establishment of present and historical land use maps (Hartley et al., 2017; Houghton et al., 2012; Hurtt et al., 2020; Jain et al., 2013; Peng et al., 2017; Shevliakova et al., 2009). Jain and Yang (2005) found that the difference in the changing rates of cropland area of two widely used land use datasets contributed significantly to the land use emission uncertainty. It was also reported that during 1750–2010, discrepancies between reconstructed and dynamically computed land covers led to a difference of 84–114 Pg C in historical LULCC emissions (Goll et al., 2015). Additionally, the diversity of LULCC definitions is also an uncertainty source (Goll et al., 2015; Houghton et al., 2012; Obermeier et al., 2021).

Therefore, understanding how LULCC uncertainties will influence the global soil carbon stock estimates is essential and will contribute to the management of existing carbon pools (Obermeier et al., 2021; Poeplau et al., 2011). This is particularly important as LULCC is estimated to have affected 32% of the global land area (Winkler et al., 2021). Here we use six different LULCC datasets to drive the process-based Dynamic Land Ecosystem Model (DLEM) to investigate how global SOC estimates would be influenced by the choice of LULCC dataset both spatially and temporally during 1900–2018. The outcomes of this research would contribute to improving LULCC cartography and quantifying the global carbon budget.

2. Materials and methods

2.1. The dynamic land ecosystem model (DLEM)

DLEM is a highly-integrated terrestrial biosphere model (TBM) that can quantify spatially explicit carbon, water, and nutrient fluxes and stocks at a daily time-step (Tian et al., 2011a, 2015; Wang et al., 2020; Yao et al., 2020). It contains five key components, including biophysics, plant physiology, soil biogeochemistry, dynamic vegetation, and land use and management (Ren et al., 2020; Tian et al., 2011a, 2015). DLEM has been extensively used to investigate the ecosystem responses to multiple natural and anthropogenic influences, such as climate change, atmospheric CO₂ concentration, tropospheric ozone, land-use change, and disturbances (e.g., fire, hurricane, and harvest), across a wide variety of biomes and spatiotemporal scales (Friedlingstein et al., 2020; Pan et al., 2014, 2020; Ren et al., 2020; Tian et al., 2011a, 2015; Wang et al., 2020; Yang et al., 2015; Yao et al., 2020; Zhang et al., 2016). The agricultural module in DLEM is developed by incorporating explicit and mechanistic representations of dynamic crop growth processes, including crop-specific phenological development, biomass accumulation and allocation, yield formation, and biogeochemical and hydrological processes, along with agricultural management practices such as nitrogen fertilization, irrigation, rotation, manure application, and

tillage (Zhang et al., 2018). By coupling detailed biogeochemical, biophysical, and hydrological processes, the agricultural module is capable of simulating and predicting the exchange of carbon, water, nutrient, and energy fluxes within the agriculture-climate-environment system. Additionally, we considered the cropland expansion, abandonment, and rotation when assessing the impacts of LULCC on SOC and other fluxes and stocks (Tian et al., 2011, 2015; Ren et al., 2020).

2.2. LULCC datasets

Six LULCC datasets were used, namely Land-Use Harmonization 2 Update for the Global Carbon Budget (LUH2-GCB2019; Chini et al., 2021a; Hurtt et al., 2020), Moderate-Resolution Imaging Spectroradiometer (MODIS) MCD12Q1 version 6 (Friedl and Sulla-Menashe, 2019), European Space Agency Climate Change Initiative-Land Cover (ESA CCI-LC; Bontemps et al., 2013, 2015), Global Land Surface Satellite-Global Land Cover (GLASS-GLC; Liu et al., 2020), data by Ramankutty and Foley (1999; hereafter RF), and data by Houghton and Hackler (2001; hereafter HH).

LUH2-GCB2019 (0.25° × 0.25°, 1700–2018) was built upon the updated data from the History Database of the Global Environment (HYDE 3.2) and the most recent Food and Agriculture Organization (FAO) statistics. It adopts corrected cropland and grazing areas, along with a significant sub-national reorganization of agricultural land-use patterns within Brazil as far back as 1950 (Chini et al., 2021b; Hurtt et al., 2020). This dataset has been used as a standard input for the multiple terrestrial biosphere models to assess the annual global carbon budget (Friedlingstein et al., 2020). The MCD12Q1 version 6 product (500m × 500m, 2001–2018) was generated by applying the supervised classification method to MODIS reflectance data (Friedl et al., 2010), with the classifier instability reduced dramatically through using Hidden Markov Models (Abercrombie and Friedl, 2015). The improved ESA CCI-LC maps (300m × 300m, 1992–2018) were produced based on the Advanced Very High Resolution Radiometer (AVHRR), the Système Pour l'Observation de la Terre-VEGETATION (SPOT-VEG), the Medium Resolution Imaging Spectrometer (MERIS), and the Project for On-Board Autonomy-Vegetation (PROBA-V) using the GlobCover unsupervised classification chain (ESA, 2017). The GLASS-GLC product (5 km × 5 km, 1982–2015) is built upon the latest version of AVHRR Global Land Surface Satellite climate data records (Liu et al., 2020). The RF dataset is mainly based on the FAOSTAT agricultural statistics and considers changes in other land use types to compensate for the agricultural area changes (Ramankutty et al., 2008; Ramankutty and Foley, 1999). Oppositely, the HH dataset relies on forest area and biomass from FAO Global Forest Resources Assessment (FRA) and assumes that the area changes of other land use (e.g. croplands) account for the forest area change (Houghton, 2008; Houghton and Hackler, 2001). Both the RF and HH datasets provide a fractional area for croplands and natural vegetation at a 0.5° × 0.5° spatial resolution. The RF dataset is available from 1700 to 2005, and HH is from 1700 to 2007 (Meiyappan and Jain, 2012). All these datasets, except RF and HH, were resampled to a 0.5° × 0.5° resolution.

The historical cropland distributions for MODIS (1700–2000), ESA CCI-LC (1700–1991), and GLASS-GLC (1700–1981) were reconstructed according to the annual changing fraction of cropland in each grid cell in LUH2-GCB2019 and the spatial patterns in their respective first available years. Natural vegetation types from MODIS, ESA CCI-LC, GLASS-GLC, HH, and RF were further re-categorized to be consistent with DLEM biomes classifications. In LUH2-GCB2019, the non-agricultural land was partitioned into natural biomes according to the spatial patterns of natural plant functional types (PFTs) in the Synergetic Land Cover Product (SYNMAP; Jung et al., 2006), which were then scaled proportionally with the annual cropland fraction within each grid cell.

2.3. Other model forcings

Daily climate data, including maximum, minimum, mean temperature, precipitation, and downward shortwave radiation, were obtained from the Climate Research Unit-Japanese reanalysis (CRU-JRA; [Harris, 2019](#)). Annual atmospheric CO₂ concentration was obtained from the NOAA GLOBALVIEW-CO₂ dataset from atmospheric and ice core measurements (<https://www.esrl.noaa.gov>). Annual atmospheric nitrogen deposition data was obtained from the International Global Atmospheric Chemistry (IGAC)/Stratospheric Processes and Their Role in Climate (SPARC) Chemistry-Climate Model Initiative (CCMI) ([Eyring et al., 2013](#)). Soil physical and chemical properties were acquired from the International Soil Reference and Information Centre-World Inventory of Soil Emission Potentials (ISRIC-WISE) Harmonized Global Soil Profile dataset ([Batjes, 2008](#)). Other auxiliary inputs, such as topography and river network data, were acquired from the previous studies ([Pan et al., 2014](#); [Tian et al., 2011b](#)).

2.4. Model calibration and simulation designing

DLEM was calibrated against GPP measurements from FLUXNET and the Long-Term Ecological Research (LTER) Network for each PFT ([Fig. 1](#)). Specifically, we first ran the model with the default parameters, of which most GPP-related parameters were tuned (within a $\pm 20\%$ range of default values) to obtain a close match between the observed and modeled GPP values. The parameter set with a minimal bias between the simulated and observed values was adopted. Then, the DLEM model was further tuned and validated using SOC measurements across various vegetation types, soil, and climate conditions, which were from the World Soil Information Service (WSIS; [Batjes et al., 2017](#)), the Harmonized World Soil Database (HWSD v1.2; [Wieder et al., 2014](#)), the Global Soil Organic Carbon Map (GSOC v1.2.0), and the International Geosphere-Biosphere Program Data and Information System (IGBP-DIS; [Global Soil Data Task, 2014](#)). In this step, considering the number of SOC measurements is sufficient to cover a wide variety of land conversion trajectories and it is difficult to judge which LULCC dataset is mostly near to the truth, it only requires ensuring the DLEM model using any one of the LULCC datasets can achieve an overall acceptable accuracy compared with those observation-based SOC products. Here, the LUH-GCB2019 dataset was used to drive the DLEM model, as it has the longest record of land conversion and is widely used in inter-model comparison projects (e.g., TRENDY) and the estimation of the global

carbon budget ([Chini et al., 2021a](#)).

“All-combined” simulations with historical time-varying climate, LULCC, N deposition, and atmospheric CO₂, were run to estimate SOC changes. DLEM was first run to reach equilibrium using the daily climatology during 1901–1930. When differences in carbon, nitrogen, and water pools between their current and previous averages in two consecutive 20-year moving windows were less than $0.1 \text{ g C m}^{-2} \text{ yr}^{-1}$, $0.1 \text{ g N m}^{-2} \text{ yr}^{-1}$, and $0.1 \text{ mm H}_2\text{O m}^{-2} \text{ yr}^{-1}$, respectively, the equilibrium state in a grid cell was considered to be reached ([Ren et al., 2020](#); [Zhang et al., 2018](#)). Before the transient simulations, the model was spun up for another 1,000 years to avoid potential bias resulting from the shift between the equilibrium state to the transient state. In this step, the climate data between 1901 and 1930 was randomly selected for each year. To evaluate the impacts of LULCC on SOC dynamics, the paired “without-LULCC” simulations, i.e., fixing LULCC in 1700 for each of the six LULCC datasets, were conducted. The LULCC-induced SOC change for each LULCC dataset was calculated as the difference between the “all-combine” and “without-LULCC” simulations.

3. Results

3.1. Historical changes in cropland area

The temporal dynamics of the global cropland area showed large differences across the six LULCC datasets. The largest cropland area was recorded in ESA CCI-LC, which had a cropland area of 19.47 million km² in the 2010s ([Fig. 2](#)), followed by LUH2-GCB2019 with a cropland area of 15.56 million km² in the 2010s. MODIS, GLASS-GLC, and RF had similar cropland areas of approximately 12.00 million km² in recent decades, while HH got the smallest cropland area of approximately 7.00 million km² in the 2000s.

The three long-record datasets agreed that cropland had expanded substantially since 1900 though they differed in the changing rate ([Fig. 2](#)). In recent decades, LUH2-GCB2019 showed the most pronounced increase in cropland area, whereas MODIS, GLASS-GLC, and RF showed barely cropland expansion or even recession ([Fig. 2](#)). Cropland expansion also occurred in ESA CCI-LC, which mainly took place in the 1990s, and has remained stable since 2000. This trend contrasts with that of LUH2-GCB2019, in which the cropland area sharply increased in the 2010s. It is worth noting that the cropland area in GLASS-GLC exhibited the greatest fluctuations among all the datasets.

The spatial distribution of cropland also exhibited great divergences

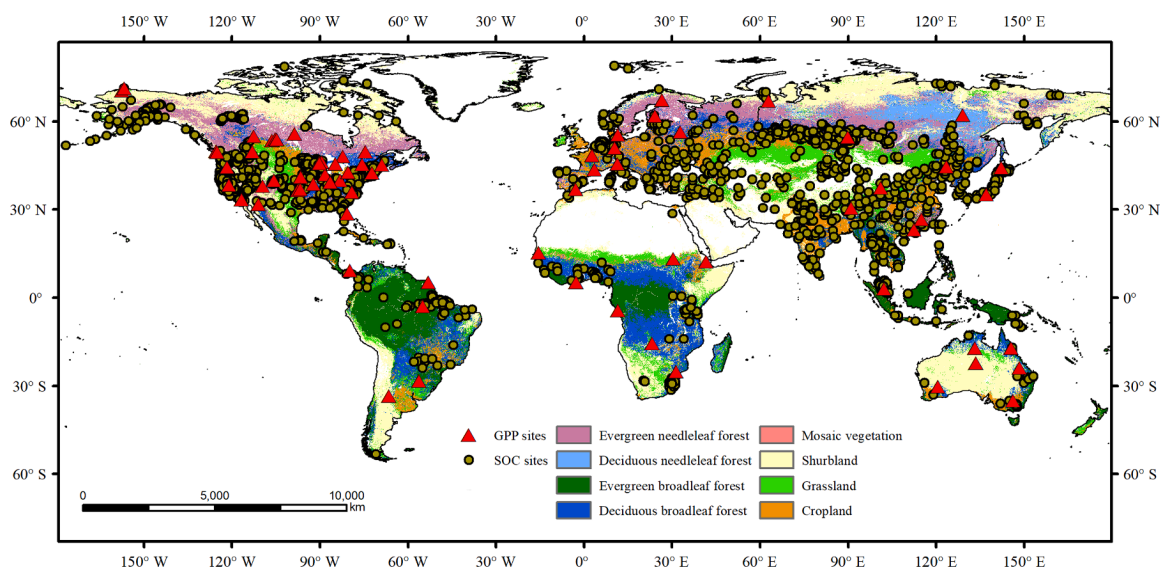


Fig. 1. Spatial distribution of site-level GPP and SOC measurements.

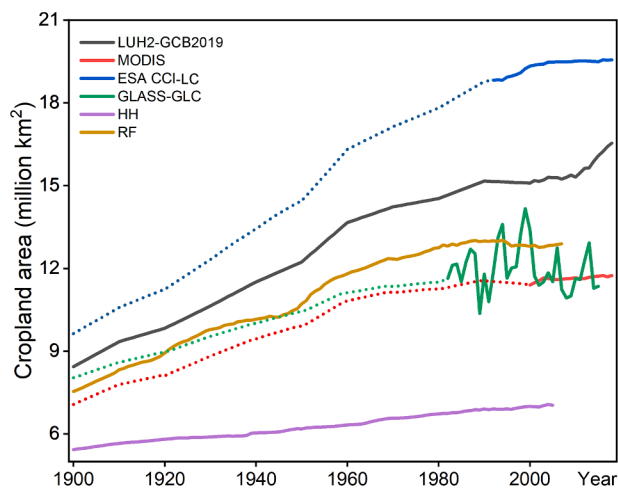


Fig. 2. The temporal trends of cropland area since 1900 for the six LULCC datasets. The dotted lines indicate the reconstructed periods.

among the six LULCC datasets (SI Fig. S2). LUH2-GCB2019 and RF had a similar spatial pattern, with RF having a smaller area in Eurasia. RF also had a larger cropland area than MODIS, particularly in South America and Africa. In GLASS-GLC, cropland was mainly concentrated in North America, southeastern South America, Southern Europe, India, and eastern Asia. HH had a larger cropland area in Africa than MODIS and GLASS-GLC. Across all datasets, cropland expansion was particularly notable in regions where it had already been intensively cultivated, such as North America, southeastern South America, and India (Fig. 3). Cropland abandonment was mainly located in eastern North America

and western Europe. Notably, a significant decrease in cropland area was observed in eastern China in HH and RF, at a much higher rate than in the other LULCC datasets (Fig. 3).

3.2. Validation of model results

The validation results revealed that modeled GPP agreed well with the observations ($R^2 = 0.72$; Fig. 4a). In terms of the modeled SOC, a relatively lower deviation from the observations was observed ($r = 0.40$), and it performed comparably or slightly better than HWSD, GSOC, an IGBP-DIS (Fig. 4b). The spatial patterns of the simulated SOC by DLEM were generally consistent with HWSD, GSOC, or IGBP-DIS (SI Fig. S1).

3.3. Impacts of LULCC on the global SOC changes

As shown in Fig. 5, the largest LULCC-induced SOC reduction from 1900 to 2018 was found in the LUH2-GCB2019-based estimation, which amounted to 106.20 Pg C. The LULCC-induced SOC losses estimated by MODIS, ESA CCI-LC, and GLASS-GLC were close, being 36.02, 37.80, and 33.48 Pg C, respectively. Meanwhile, HH and RF exhibited higher SOC losses, with values of 57.28 and 66.40 Pg C, respectively.

Spatially, the LULCC-induced SOC loss widely occurred across the globe, especially in the northern high latitudes (Fig. 6). In LUH2-GCB2019-based estimation, LULCC-induced SOC loss at a rate of $>5 \text{ kg C m}^{-2} \text{ yr}^{-1}$ was mainly distributed in the central and mid-east regions of North America. Such a loss rate was higher than those estimated by MODIS, ESA CCI-LC, and GLASS-GLC. For the HH- and RF-based estimations, regions exhibiting the LULCC-induced SOC loss were mainly concentrated in northern North America. In Europe, both LUH2-GCB2019- and RF-based estimations exhibited larger areas with LULCC-induced SOC loss compared to the other LULCC datasets. Regions with a

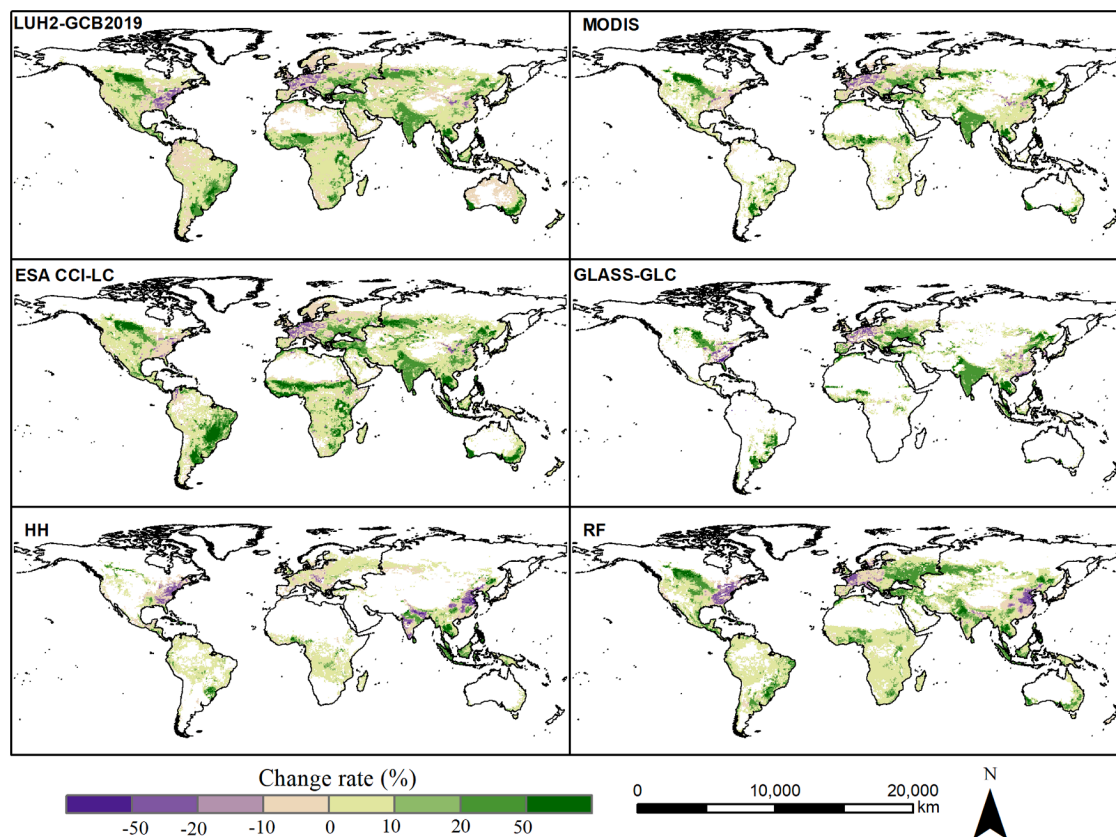


Fig. 3. Spatial patterns of the historical changes in cropland area since 1900 in the six LULCC datasets. The area of regions showing a significant trend in cropland coverage accounts for 95.31%, 93.94%, 80.94%, 78.86%, 93.89%, and 95.82% for LUH2-GCB2019, MODIS, ESA CCI-LC, GLASS-GLC, HH, and RF, respectively.

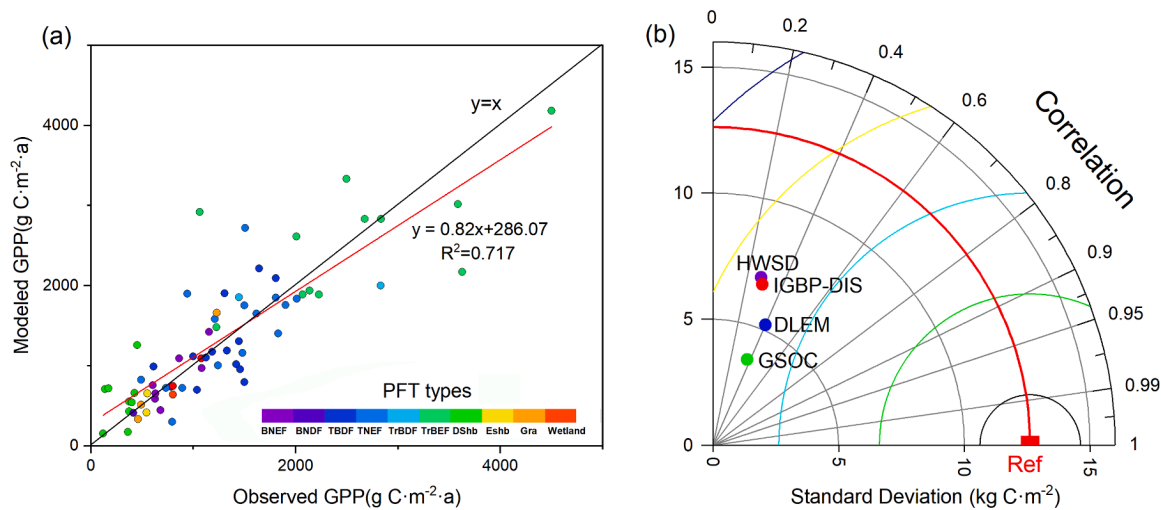


Fig. 4. Comparisons of modeled GPP (a) and SOC (b) with observations. The LUH2-GCB2019 is used in driving DLEM. The Ref data in Fig. 4b is the site-level data from WSIS. BNEF, Boreal needleleaf evergreen forest; BNDF, Boreal needleleaf deciduous forest; TBDF, Temperate broadleaf deciduous forest; TNEF, Temperate needleleaf evergreen forest; TrBDF, Tropical broadleaf deciduous forest; TrBEF, Tropical broadleaf evergreen forest; DShb, Deciduous shrubland; EShb, Evergreen shrubland; Gra: Grassland.

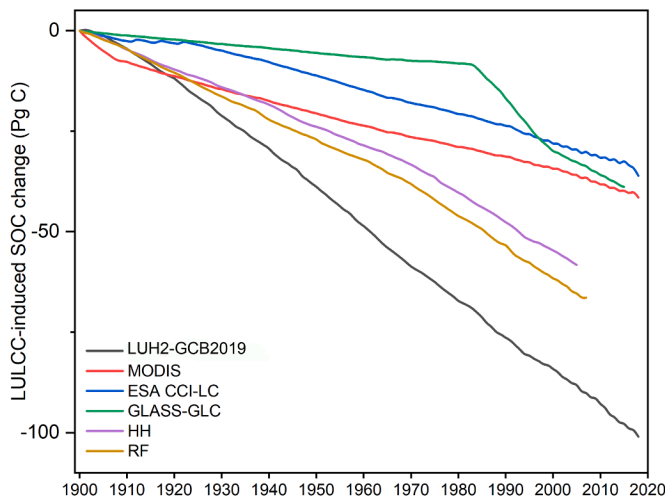


Fig. 5. The modeled accumulative LULCC-induced SOC loss using different LULCC datasets.

high LULCC-induced SOC loss rate were primarily located in northern Europe in the HH-based estimation but in southern Europe for estimations by MODIS, ESA-CCI, and GLASS-GLC. In Asia, LUH2-GCB2019-based estimation showed the largest area with high LULCC-induced SOC decreasing rates, followed by RF-, HH-, ESA CCI-LC-, and MODIS-based estimations. Most of these regions were distributed in boreal areas. In contrast, the smallest area experiencing LULCC-induced SOC loss was observed in GLASS-GLC estimation, mainly concentrated in northeastern and southern China. In South America, the largest area with a high SOC loss rate was found in the LUH2-GCB2019-based estimation (Fig. 6).

To further identify the hotspots with the large discrepancies in SOC resulting from LULCC differences, the LULCC-induced SOC dynamics in ten regions, including North America (NA), South America (SA), Europe (EU), Africa (AF), Russia (RU), Middle East (MIDE), East Asia (EAS), South Asia (SA), Southeast Asia (SEA), and Oceania (OC), were compared. As shown in Fig. 7, Russia experienced the largest LULCC-induced SOC loss (14.53 ± 11.48 Pg C), with the LUH2-GCB2019-based estimation yielding 34.39 Pg C, more than twice the amount simulated by other LULCC datasets. In NA, the second-largest decrease

in SOC caused by LULCC was observed, with a decrease of 14.16 ± 7.08 Pg C. The RF-, LUH2-GCB2019-, and HH-based simulations were twice as high as the estimates based on MODIS, ESA CCI-LC, and GLASS-GLC. In other regions, the reduction of SOC resulting from LULCC was less than 10.00 Pg C. The smallest decrease was detected in MIDE with 0.68 ± 0.69 Pg C. In OC, the largest discrepancy among the simulations was observed, with the LUH2-GCB2019-based estimation indicating a SOC loss of 6.80 Pg C, while the GLASS-GLC-based estimation indicated a SOC increase of 0.77 Pg C. The LULCC-induced SOC loss was 2.31 ± 1.67 Pg C and 1.60 ± 0.64 Pg C in SES and SAS, respectively. The regions with large divergence in relative changes of LULCC-induced SOC loss were primarily located in the low latitudes. The largest difference was observed in SAS, where the relative change of LUH2-GCB2019 was ten times higher than that estimated by other LULCC datasets. Smaller disparities in relative change were observed in RU, EU, and NA. Overall, LUH2-GCB2019-based estimation has the largest SOC loss resulting from LULCC in all regions except in NA, EU, and SAS, and the relative change of this loss deriving from different LULCC datasets varied across the ten regions.

3.4. Co-effects of LULCC and other factors on the global SOC

When considering both LULCC and other factors, the simulated dynamics of SOC driven by the six LULCC datasets showed greater differences than those by LULCC alone, both spatially and temporally. All the modeled SOC decreased from 1900 to the 1960s (Fig. 8). The highest and lowest decreasing rates were found in LUH2-GCB2019- and GLASS-GLC-based estimations, respectively. After the 1960s, ameliorations in SOC loss or even SOC increase were observed to different extents for all the LULCC datasets except RF. It is noted that global SOC showed a decreasing trend around the 1990s in estimates driven by GLASS-GLC. Over the entire study period, the largest SOC decrease of 39.11 Pg C was found in the LUH2-GCB2019-based estimation, followed by 12.36 Pg C in the HH-based estimation and 22.43 Pg C in the RF-based estimation. In contrast, MODIS-, ESA CCI-LC-, and GLASS-GLC-based estimations indicated that SOC increased by 7.80, 14.73, and 8.13 Pg C, respectively. This suggests that the contribution of other factors may have surpassed the impacts of LULCC during certain periods.

All the LULCC datasets revealed two hotspots of SOC stocks, located around the 0° and 60°N (Fig. 9). LUH2-GCB2019- and GLASS-GLC-based SOC showed similar spatial patterns, with higher SOC density in eastern and northern Russia compared to other LULCC datasets. The HH- and

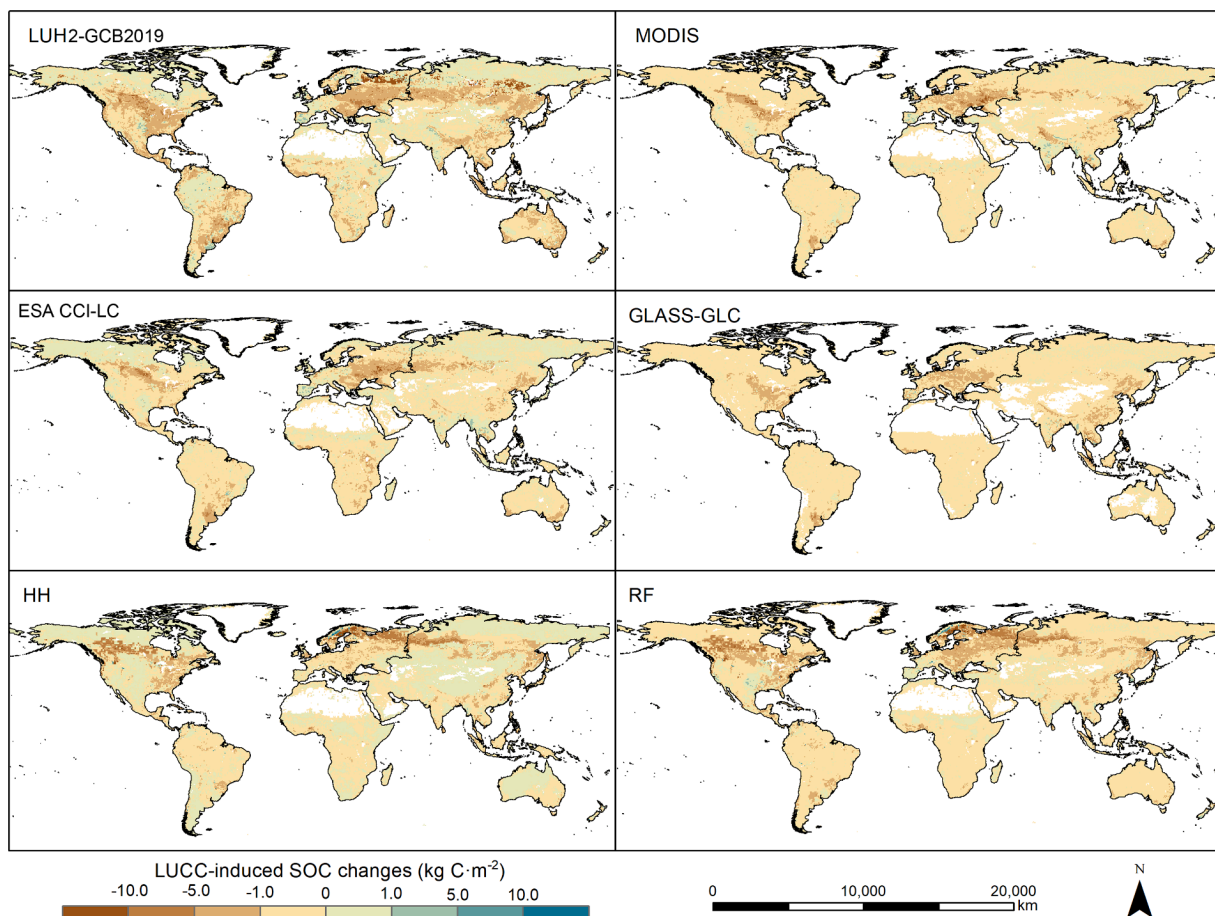


Fig. 6. Spatial dynamics of LULCC-induced SOC driven by different LULCC datasets. The SOC changes are derived as the differences between “All-combine” and “without-LULCC” simulations. The area of regions showing a significant trend in LULCC-induced SOC changes accounts for 96.14%, 93.13%, 94.02%, 94.17%, 95.42%, and 87.87% for LUH2-GCB2019, MODIS, ESA CCI-LC, GLASS-GLC, HH, and RF, respectively.

RF-based estimations, however, indicated that high SOC density in Russia was concentrated in the western regions, while it was widely distributed across central and eastern Russia in other LULCC dataset estimations. In central Africa, the MODIS-based SOC density was the highest among all six LULCC datasets (Fig. 9).

The spatial dynamics of SOC estimated by different LULCC datasets also exhibited large divergence (Fig. 10). Regions with SOC reduction were primarily concentrated in the northern high latitudes. Notably, in North America, the estimated SOC loss estimated by ESA CCI-LC, GLASS-GLC, and HH was markedly lower than those estimated by the other three LULCC datasets. Furthermore, the SOC loss estimated by LUH2-GCB2019 was primarily located in the central and mid-eastern regions, whereas HH and RF estimated SOC loss in the northern regions. In Europe, the high SOC loss rates were found in the southern region based on LUH2-GCB2019, MODIS, and ESA CCI-LC, while HH and RF showed similar loss rates in the northern regions. GLASS-GLC-based estimation indicated SOC gain in most regions of Europe. In Russia, LUH2-GCB2019, HH, and RF estimated widespread SOC loss with high rates, while MODIS, ESA CCI-LC, and GLASS-GLC showed lower rates of SOC loss in this area.

4. Discussion

4.1. Effects of LULCC on SOC dynamics

The LULCC effects on SOC dynamics are diverse and complex. When one ecosystem changes into another one, the previous state of soil would be disturbed, leading to either release or uptake of soil carbon until a

new equilibrium is eventually reached (Bolin, 2001; Guo and Gifford, 2002; Smith, 2008). Notably, the conversion from native vegetation to cropland would reduce the SOC pool due to the decreasing input of biomass and enhanced decomposition after disturbances (Poeplau et al., 2011). Oppositely, afforestation on cropland is supposed to increase the soil carbon pool (Deng et al., 2014; Korkanç, 2014; Lal, 2004). The development of secondary forests contributes to the recovery of SOC but may take a century-long time (Powers et al., 2011; Wright, 2005). However, there are also studies arguing that the re-vegetation of cropland did not necessarily lead to an increase or decrease in SOC (Farley et al., 2004; Smal and Olszewska, 2008; Vesterdal et al., 2002).

We found that the northern high latitudes experienced the largest SOC loss resulting from LULCC. In North America and Russia, the cropland area expanded the most in LUH2-GCB2019 among all the LULCC datasets, leading to a large LULCC-induced SOC loss. This finding is consistent with previous studies (Eglin et al., 2010; Tian et al., 2015). Eglin et al. (2010) demonstrated that LULCC-induced SOC loss outweighed the carbon accumulation caused by climate and CO₂ in North America, the Former Soviet, and Europe, due to forest clearing during the 20th century. During 1850–2015, cropland expansion dominated the net carbon balance in the Midwestern US (Yu et al., 2018). The cropland area decreased in North America in HH but a similar magnitude of SOC loss was observed, supporting the idea that the conversion of cropland to natural vegetation may not lead to a gain of SOC (Farley et al., 2004; Li et al., 2012; Smal and Olszewska, 2008). The cropland area in LUH2-GCB2019 experienced an increasing trend before the 1960s but a decreasing trend thereafter in Europe. However, the LULCC-induced SOC loss was persistent, implying a legacy effect of LULCC on SOC

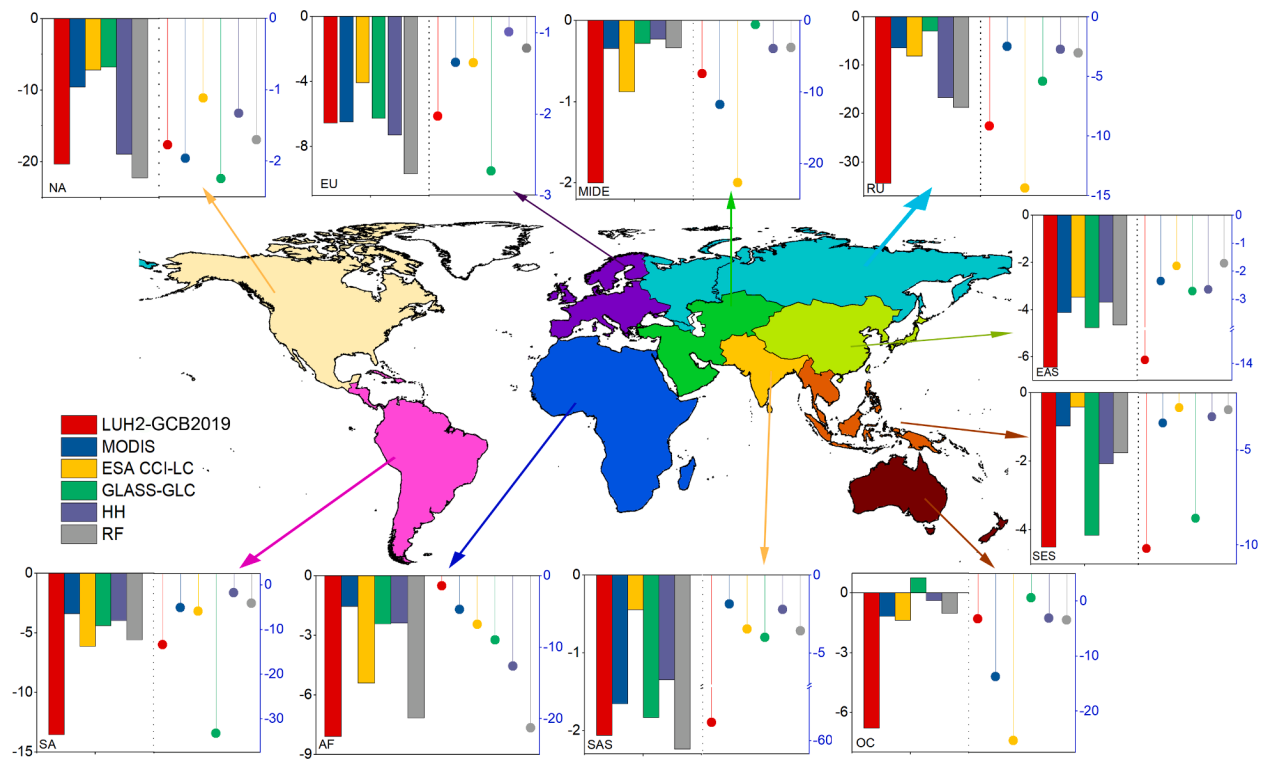


Fig. 7. The LULCC-induced SOC dynamics (The bar chart in each panel, Unit: Pg C) and the relative changes (The scatter chart in each panel) in the ten regions based on the six LULCC datasets.

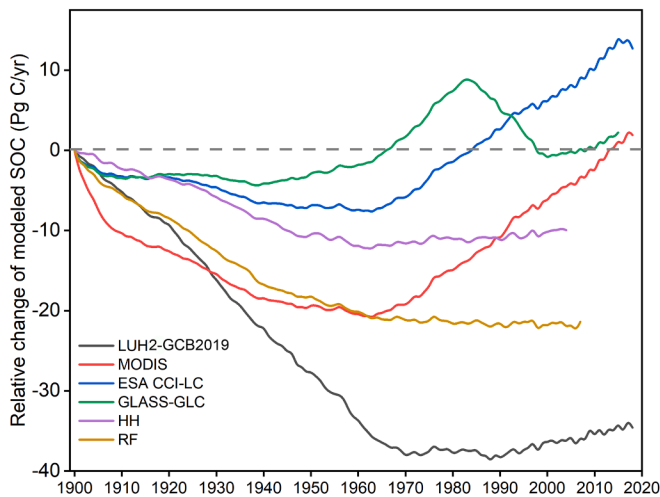


Fig. 8. The relative change of modeled SOC over the study period based on the six LULCC datasets. The annual relative SOC was calculated as the difference between SOC in each year and in 1900.

loss. A previous study showed that up to 60% of current LULCC emissions are attributed to the large legacy influence of LULCC in the past (Goll et al., 2015). Despite the cropland area in MODIS, ESA CCI-LC, and GLASS-GLC complies the same changing rate as in LUH2-GCB2019, they differed in their initial cropland map (SI Fig. S2), resulting in different carbon re-allocation following land conversion and diverse responses of SOC dynamics to LULCC (Bárcena et al., 2014; Houghton et al., 2012; Poeplau et al., 2011).

4.2. Mechanism for the contribution of LULCC to SOC changes

A change in the fraction of PFT resulting from LULCC would affect the amount of carbon that flows to vegetation, litter, and soil pools, leading to changes in NPP and the input of organic material, and thereby affecting the soil decomposition by controlling the C substrate availability (Hartley et al., 2017; Jain and Yang, 2005; Tian et al., 2015). Previous studies have demonstrated NPP is the primary source that determines the total amount and variation of SOC, implying that differences in NPP could account for the divergent modeled soil carbon stock estimates (Goll et al., 2015; Hartley et al., 2017; Tian et al., 2015; Xu et al., 2020). Therefore, we investigated the LULCC-induced NPP changes based on all the LULCC datasets. The correlation analysis revealed that the LULCC-induced NPP changing rates were closely related to the LULCC-induced SOC changing rates ($R^2=0.88$), supporting the hypothesis that the different SOC changes can be attributed to differences in NPP (Fig. 11). Generally, changes in NPP can be explained by the two main counterbalancing effects, including the decrease in NPP due to deforestation and the increase in NPP due to cropland abandonment and the following growth of secondary forests (Jain and Yang, 2005). In our study, the differences in the LULCC-induced NPP changing rates can be attributed to the varying initial land use maps and land conversion histories for different LULCC datasets. Changes in the natural vegetation maps resulting from cultivation or re-vegetation would affect terrestrial NPP and subsequently impact SOC dynamics (Guo and Gifford, 2002; Vesterdal et al., 2002; Brovkin et al., 2013).

4.3. Uncertainties

4.3.1. The current land use maps

The fractional forest area was found to have a relatively large uncertainty in Africa, South America, and continental Southeast Asia. Higher forest coverage was observed in Europe and southeastern China in MODIS and GLASS-GLC than in SYNMAP (SI Fig. S2). Moreover, large

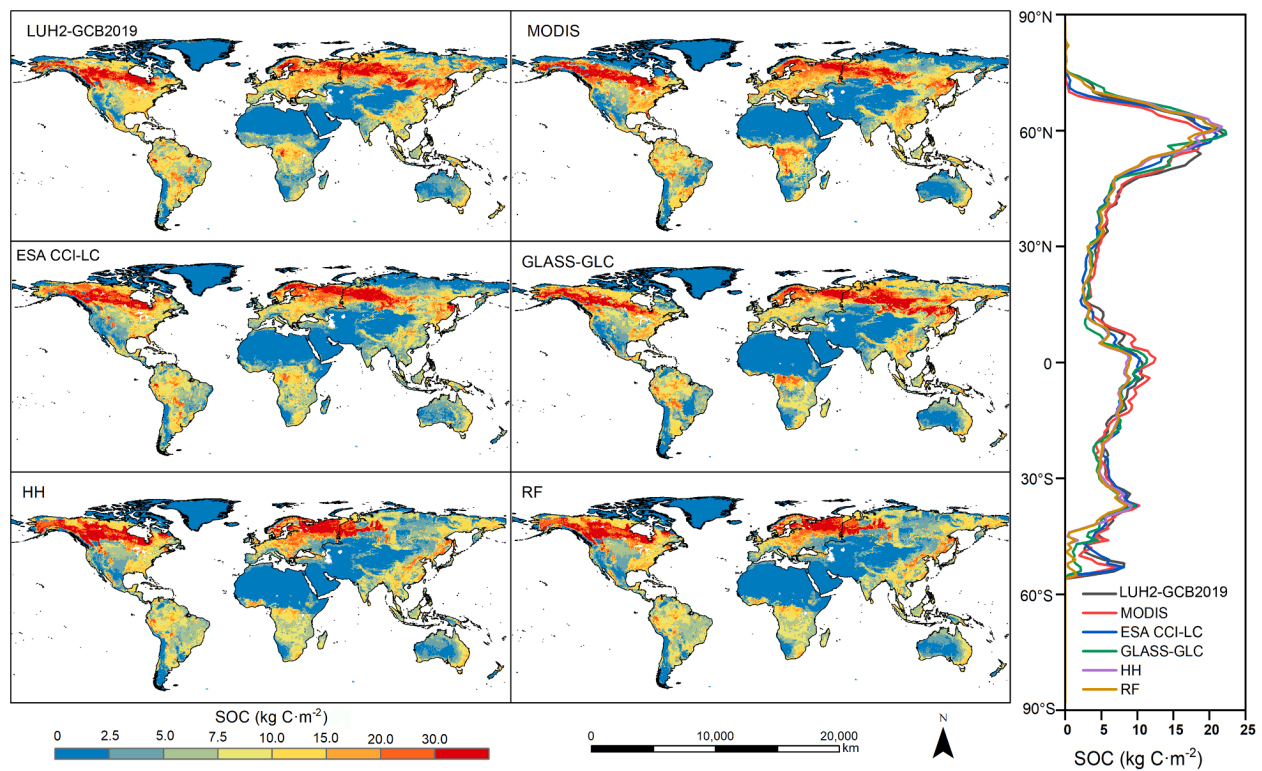


Fig. 9. The spatial patterns of modeled SOC stocks driven by different LULCC datasets (left) and their latitudinal distribution (right).

discrepancies in cropland fraction were identified in Europe, India, and North America (Li et al., 2018). Several reasons may help to explain the large inconsistency among these LULCC datasets. As the forest area in FAO is reported by the member countries, the clearance of forest may still be reported as a forest class in inventory data, whereas it may be detected as non-vegetated lands from satellite observations (Keenan et al., 2015; Li et al., 2018). In dryland Africa, it is challenging to distinguish forests from shrublands or savannas using medium- or coarse-resolution satellite remote sensing. The global dryland forest area was reported to be underestimated by 40% (Achard et al., 2014; Bastin et al., 2017). Additionally, different definitions for a particular land use type can lead to uncertainties. For instance, oil palm plantations and rubbers in tropical countries are considered as cropland in the FAO definition, but classified as tree cover in remote sensing datasets (Carlson et al., 2013; Liu et al., 2020). Our findings demonstrated that ESA CCI-LC has the largest cropland area among all the datasets, indicating that greater discrepancies may exist in other land use types. Another factor contributing to the uncertainty is the time lag between the real-time changes captured by remote sensing and the reporting of inventory data by individual countries (Hartley et al., 2017).

The advancement of remote sensing technology over the past few decades has provided a valuable opportunity to characterize the spatiotemporal pattern of vegetation distribution under natural and anthropogenic activities. However, satellite products obtained from different platforms typically differ in their land cover scheme, spatial resolution, time coverage, and accuracy, mainly due to the differences in sensors or validation methods (Bontemps et al., 2012). This uncertainty in the original satellite images would propagate in producing the satellite-derived land use products (Bayer et al., 2017; Li et al., 2018). The performance of the classification methods also contributes to the overall uncertainties (Hartley et al., 2017). For example, the canopy cover of forests is defined as >15% in ESA CCI-LC maps according to the Land Cover Classification System (LCCS) developed by the United Nations (UN) FAO, whereas it is 10% in MODIS land cover products with PFT classification. This difference may partly explain the higher forest

coverage observed in MODIS than in ESA CCI-LC.

Previous studies demonstrated that the land use maps face difficulty in separating the spectral signature of subsistence agriculture from natural vegetation between 15°S–5°N, which leads to a higher heterogeneity of mixed vegetation in these regions (Hartley et al., 2017). Generally, satellite-based land use products have the lowest accuracy in identifying regions with sub-pixel scales mixtures of plants, such as cropland/natural vegetation mosaics (Bontemps et al., 2015; Liu et al., 2020; Sulla-Menashe et al., 2019). Furthermore, each remotely sensed product has its own limitations. For example, wetlands are under-represented in MCD12Q1, and there are misclassifications among different forest types in Japan, the Pacific Northwest of North America, Chile, Australia, and parts of South America (Sulla-Menashe and Friedl, 2018). Similarly, land conversions between different forest types are not sufficiently captured in ESA CCI-LC because forest growth is a gradual process (Goll et al., 2015; Li et al., 2018). In addition, low-intensity of land use changes, such as shifting cultivation, may not be detectable using medium- or coarse-resolution satellite imagery, which can lead to errors in classification (ESA, 2017). The accuracy of these satellite products is also constrained by the contamination of persistent cloud cover in some regions, such as the Congo Basin and Insular Southeast Asia (Houghton et al., 2012). For such areas, the use of high-resolution satellite images, such as Landsat and Sentinel, can generate more precise maps of forest coverage. GLASS-GLC is reported to have lower classification accuracy in regions such as Africa, eastern and southern South America, southern Alaska, northern and eastern Australia, and southwestern Indonesia (Liu et al., 2020).

4.3.2. The historical land use maps

Historical land use maps provide a crucial baseline for projecting future vegetation maps and the atmosphere-land interactions (Brovkin et al., 2013; Hurtt et al., 2011). The uncertainty of historical LULCC is also one of the key uncertainties in estimating the terrestrial carbon flux since carbon flux resulting from LULCC depends on both land use types and carbon stocks prior- and post-change (Armeth et al., 2017; Bayer

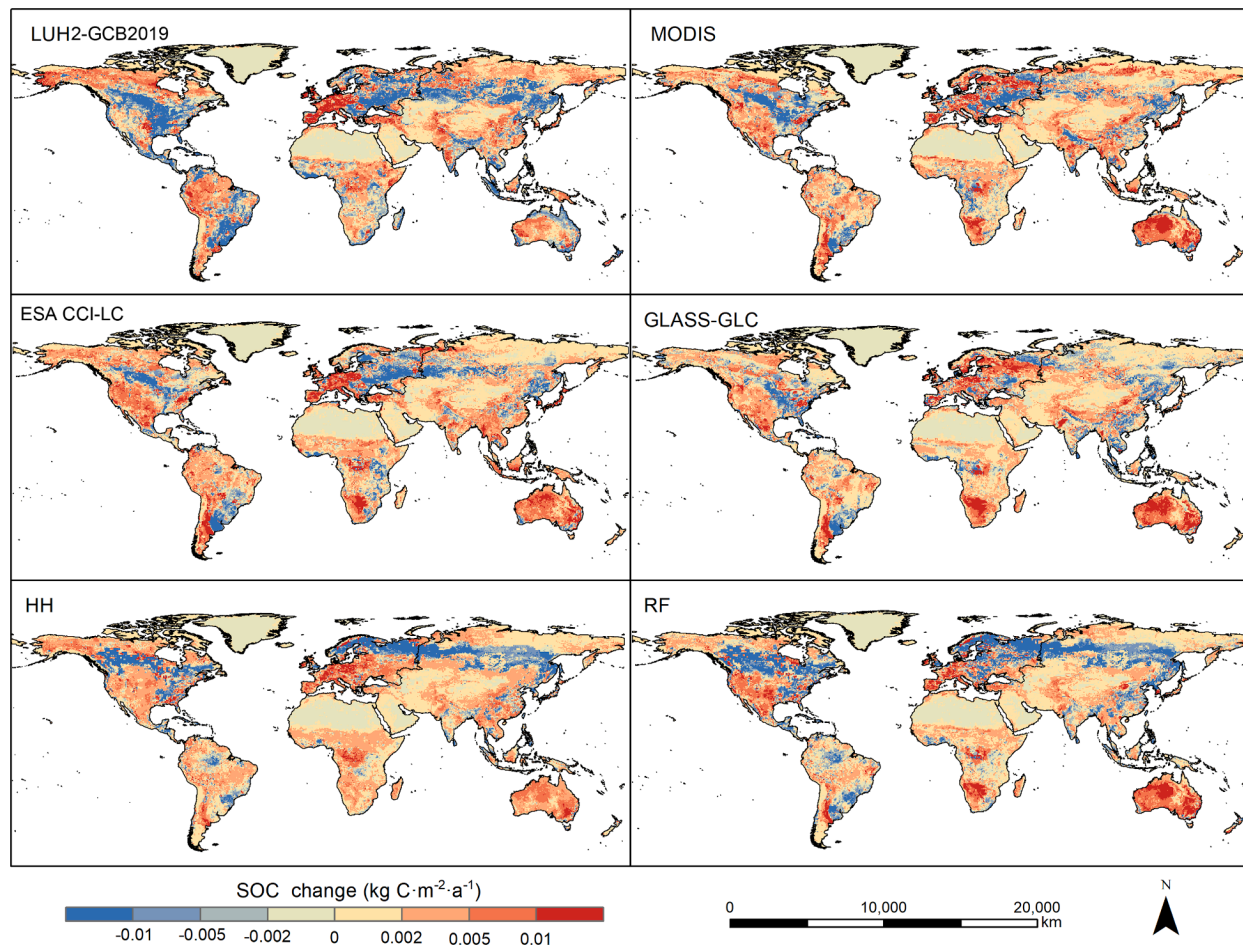


Fig. 10. The spatial dynamics of modeled SOC driven by different LULCC datasets. The area of regions showing a significant trend in SOC accounts for 95.91%, 95.12%, 95.37%, 93.52%, 96.32%, and 95.04% for LUH2-GCB2019, MODIS, ESA CCI-LC, GLASS-GLC, HH, and RF, respectively.

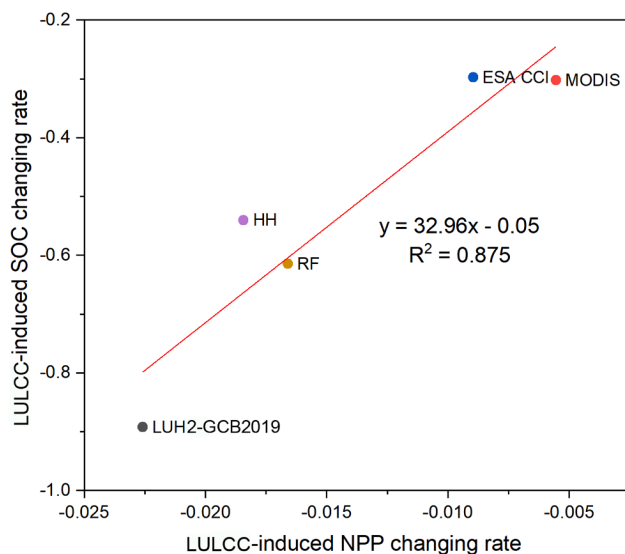


Fig. 11. Correlations between changing rates of LULCC-induced NPP and LULCC-induced SOC.

et al., 2017; Fuchs et al., 2015; Houghton and Nassikas, 2017; Di Vittorio et al., 2018). As remote sensing products are only available for recent decades and prone to large uncertainties from various sources, making it difficult to reconstruct consistent historical land-use maps (Houghton

et al., 2012; Peng et al., 2017). In this study, we did not rely on country-level statistics, population statistics, or model assumptions due to the limited available dataset. Instead, we used the change in cropland area of LUH2-GCB2019 as a reference, all the cropland area in satellite-based land use maps (i.e., MODIS, ESA CCI-LC, GLASS-GLC) shared the same changing rate as that in LUH2-GCB2019 between two-time steps (inter-annual) at a given grid. Consequently, the fraction area changes of PFTs are all the “net change” rather than “gross change”, which may underestimate the land use change, and thus, introduce the uncertainty in the estimation of SOC dynamic during the historical period (Lawrence et al., 2016; Fuchs et al., 2015).

Also, we did not account for the possible impact of the cross-walking table (CW) when converting land use types to PFTs. The CW table prescribes the fraction of each PFT that occurs within each land use class and is subject to the expert interpretation of the description of a given land use type, making it a potential source of uncertainty in model simulations (Hartley et al., 2017; Li et al., 2018). Hartley et al (2017) found that, based on the ESA CCI dataset, the uncertainty of tree fraction caused by the cross-walking table was particularly significant for simulating the albedo and ET in northern boreal forests. To minimize the uncertainty resulting from using different CWs, we assumed that each grid cell was covered by one “pure” vegetation class across all satellite-based land use products. Besides, the cohort approach used in DLEM allows the coexistence of four nature PFTs in a grid cell, ranking by the relative percentage of PFTs that occupy the total area of natural vegetation. This feature may help to reduce the uncertainty caused by CW to some extent. Higher uncertainties are supposed to occur in regions covered by mosaic/mixed vegetation types, which could lead to an

underestimation of the area of a certain PFT and an overestimation of another.

The transition rule used in reconstructing the historical land use maps is another source of uncertainty in estimating the impact of LULCC on the terrestrial carbon cycle (Di Vittorio et al., 2018). For instance, the RF dataset used in our study reported a forest decrease of approximately 8.5 million km² from 1901 to 2005, which is slightly lower than the 9.0 million km² reported by HH (Houghton, 2008). However, the global forest decreased by 5.3 million km² in LUH2-GCB2019 over the same period. In Dynamic Global Vegetation Models (DGVMs), the estimated global forest area decreases between 1900 and 2005 ranged from 2.2 million km² to 16.9 million km² in different models, even though they used the same HYDE cropland and pasture dataset (Peng et al., 2017). The conversion of forests with high carbon density into agricultural land by deforestation leads to much higher emissions compared to conversion to natural grassland. Due to the differences in the initial land use maps, the reconstructed MODIS, ESA CCI-LC, and GLASS-GLC show different magnitudes of forest area decrease over the same period, with reductions of 2.57 million, 3.82 million, and 0.91 million km², respectively.

Generally, modelers tend to use the same time-varying maps of cropland and pasture, however; however, the land use histories often vary due to the different rules to account for the expansion or abandonment of cropland in TBMs (Peng et al., 2017). Some models, including DLEM, adopt the rule that the fractional change of cropland area results in a proportional reduction or expansion of existing natural PFTs within the same grid (Jain et al., 2013; Meiyappan and Jain, 2012; Ren et al., 2020). Some other models use a preferential reduction or expansion of grassland for pasture and a proportional reduction or expansion of natural PFT for cropland, or the reduction of grassland for both cropland and pasture (Brovkin et al., 2013; Reick et al., 2013). As a result, the use of different allocation rules for the compensation of cropland inevitably leads to diverse LULCC histories, and consequently, to different historical LULCC-induced carbon dynamics.

In addition, we use the individual DLEM to estimate the response of the SOC dynamic to the LULCC uncertainty in the present study. The number and types of PFTs included in process-based models often differ, and some modeling groups even generate their own PFT maps required as input based on different land use products (Hartley et al., 2017; Tian et al., 2015; Pongratz et al., 2014; Di Vittorio et al., 2020). Therefore, the outputs of other models are expected to differ from ours in terms of both quantity and spatial distribution. Tian et al. (2015) used ten different TBMs and found that the effects of LULCC contributed -20 – 400% to the net SOC change during the 1901–2010 period, despite using the same land use maps as input. Our study demonstrated that the variability in LULCC-induced SOC dynamics resulting from different LULCC datasets was similar in magnitude to the uncertainty caused by different TBMs. This large divergence in the contribution of LULCC to SOC stocks calls for a more accurate representation of SOC dynamics associated with vegetation changes (e.g., cropland abandonment and afforestation) in TBMs.

5. Conclusions

The estimation uncertainty of SOC stock and changes stemming from the multiple LULCC datasets is an imperative uncertainty source when estimating the global carbon dynamics, which is even comparable to the uncertainty of environmental forcing to SOC changes by multi-models using the same land use map as input. This study not only provides insights into the possible range of SOC changes but also highlights the locations most sensitive to the uncertainty induced by LULCC. Moreover, the uncertainty of current land use maps is highly possible to propagate when reconstructing the historical land use maps. Therefore, accurate estimation of the LULCC-induced land carbon budget requires the precise representation of both present and historical PFTs derived from land use maps. Future efforts should focus on improving the

mapping of vegetation distribution in key regions, such as the boreal and arid regions. Fully reconciling this issue requires a range of approaches, including incorporating external datasets for specific land use types and using higher spatial resolution maps to reduce classification uncertainty. Integration of remotely sensed products is a plausible way for the modeling communities to better represent the PFTs, and consequently, the LULCC process.

Declaration of Competing Interest

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

Data availability

The data that support the findings of this study were derived from the following resources available in the public domain. MODIS land cover products (MCD12Q1) are available in LP DAAC at <https://lpdaac.usgs.gov/products/mcd12q1v006/>. ESA CCI-LC dataset is available in Climate Research Data Package (CRDP) at <http://maps.elie.ucl.ac.be/CCI>. GLASS-GLC land cover dataset is available at <https://doi.org/10.1594/PANGAEA.913496>. RF and HH are available in Dr. Atul Jain's Research Group at <https://www.atmos.illinois.edu/~meiyapp2/datasets.htm>. LUH2-GCB2019 is available in Oak Ridge National Laboratory at <https://doi.org/10.3334/ORNLDAAAC/1851>. The model outputs corresponding with the figures are publicly available. They are provided at <http://doi.org/10.5281/zenodo.7854916>. The model code and the code license can be accessed as per request to the corresponding author.

Acknowledgments

This work was supported partially by the National Natural Science Foundation of China (Grant No. 31602004), U.S. National Science Foundation (Grant numbers: 1903722, 1922687), and Andrew Carnegie Fellowship Program (Grant No. G-F-19-56910). We also appreciate the China Scholarship Council for the financial support, and MODIS LP DAAC, ESA CCI Land Cover project, and Global Soil Information System for sharing datasets.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.agrformet.2023.109585](https://doi.org/10.1016/j.agrformet.2023.109585).

References

- Abercrombie, S.P., Friedl, M.A., 2015. Improving the consistency of multitemporal land cover maps using a hidden Markov model. *IEEE Trans. Geosci. Electron.* 54 (2), 703–713.
- Achard, F., Beuchle, R., Mayaux, P., Stibig, H.J., Bodart, C., Brink, A., et al., 2014. Determination of tropical deforestation rates and related carbon losses from 1990 to 2010. *Global Change Biol.* 20 (8), 2540–2554.
- Arnell, A., Sitch, S., Pongratz, J., Stocker, B.D., Ciais, P., Poulter, B., et al., 2017. Historical carbon dioxide emissions caused by land-use changes are possibly larger than assumed. *Nat. Geosci.* 10 (2), 79–84.
- Bárceña, T.G., Kiær, L.P., Vesterdal, L., Stefánsdóttir, H.M., Gundersen, P., Sigurdsson, B. D., 2014. Soil carbon stock change following afforestation in Northern Europe: a meta-analysis. *Global Change Biol.* 20 (8), 2393–2405.
- Bastin, J.F., Berrahmouni, N., Grainger, A., Maniatis, D., Mollicone, D., Moore, R., et al., 2017. The extent of forest in dryland biomes. *Science* 356 (6338), 635–638.
- Batjes, N.H., 2002. Carbon and nitrogen stocks in the soils of Central and Eastern Europe. *Soil Use and Manage.* 18 (4), 324–329.
- Batjes, N. H. (2008). ISRIC-WISE Harmonized Global Soil Profile Dataset (Ver. 3.1): Wageningen: ISRIC - World Soil Information.
- Batjes, N.H., Ribeiro, E., van Oostrum, A., Leenaars, J., Hengl, T., Mendes De Jesus, J., 2017. WoSIS: providing standardised soil profile data for the world. *Earth Syst. Sci. Data* 9 (1), 1–14.
- Bautista-Cruz, A., Del Castillo, R.F., Etchevers-Barra, J.D., Del Carmen Gutiérrez-Castorena, M., Baez, A., 2012. Selection and interpretation of soil quality indicators

- for forest recovery after clearing of a tropical montane cloud forest in Mexico. *Forest Ecol. Manag.* 277, 74–80.
- Bayer, A.D., Lindeskog, M., Pugh, T.A.M., Anthoni, P.M., Fuchs, R., Arneth, A., 2017. Uncertainties in the land-use flux resulting from land-use change reconstructions and gross land transitions. *Earth Syst. Dynam.* 8 (1), 91–111.
- Bolin, B., 2001. Global perspective. IPCC Special Report on Land Use, Land-Use Change and Forestry. Cambridge University Press, Cambridge, UK, pp. 23–51.
- Bontemps, S., Boettcher, M., Brockmann, C., Kirches, G., Lamarche, C., Radoux, J., et al., 2015. Multi-year global land cover mapping at 300 m and characterization for climate modelling: achievements of the land cover component of the ESA climate change initiative. *International Archives of the Photogrammetry. Remote Sens. Spatial Inf. Sci.* 323–328.
- Bontemps, S., Defourny, P., Radoux, J., Van Bogaert, E., Lamarche, C., Achard, F., Kirches, G., 2013. Consistent global land cover maps for climate modelling communities: current achievements of the ESA's land cover CCI. In: Paper presented at the Proceedings of the ESA Living Planet Symposium, Edinburgh, pp. 9–13.
- Bontemps, S., Herold, M., Kooistra, L., Van, G.A., Hartley, A., Arino, O., et al., 2012. Revisiting land cover observation to address the needs of the climate modeling community. *Biogeosciences* 9 (6), 2145–2157.
- Brovkin, V., Boysen, L., Arora, V.K., Boieser, J.P., Weiss, M., 2013. Effect of anthropogenic land-use and land-cover changes on climate and land carbon storage in CMIP5 Projections for the twenty-first century. *J. Climate* 26 (18), 6859–6881.
- Carlson, K.M., Curran, L.M., Asner, G.P., Pittman, A.M., Trigg, S.N., 2013. Carbon emissions from forest conversion by Kalimantan oil palm plantations. *Nat. Clim. Change* 3 (3), 283–287.
- Chini, L., Hurtt, G., Sahajpal, R., Frolking, S., Klein Goldewijk, K., Sitch, S., et al., 2021a. Land-use harmonization datasets for annual global carbon budgets. *Earth Syst. Sci. Data* 13 (8), 4175–4189.
- Chini, L., Hurtt, G., Sahajpal, R., Frolking, S., Goldewijk, K., Sitch, S., et al., 2021b. LUH2-GCB2019: Land-Use Harmonization 2 Update for The Global Carbon Budget. ORNL DAAC, Oak Ridge, Tennessee, USA, pp. 850–2019. Retrieved from: https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1851.
- Conant, R.T., Ryan, M.G., Ågren, G.I., Birge, H.E., Davidson, E.A., Eliasson, P.E., et al., 2011. Temperature and soil organic matter decomposition rates—synthesis of current knowledge and a way forward. *Global Change Biol.* 17 (11), 3392–3404.
- Davidson, E.A., Janssens, I.A., 2006. Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. *Nature* 440 (7081), 165–173.
- de Oliveira, S.P., de Lacerda, N.B., Blum, S.C., Escobar, M.E.O., de Oliveira, T.S., 2015. Organic carbon and nitrogen stocks in soils of northeastern Brazil converted to irrigated agriculture. *Land Degrad. Dev.* 26 (1), 9–21.
- Deng, L., Liu, G., Shangguan, Z., 2014. Land-use conversion and changing soil carbon stocks in China's 'Grain-for-Green' Program: a synthesis. *Global Change Biol.* 20 (11), 3544–3556.
- Di Vittorio, A.V., Shi, X., Bond-Lamberty, B., Calvin, K., Jones, A., 2020. Initial land use/cover distribution substantially affects global carbon and local temperature projections in the integrated earth system model. *Global Biogeochem. Cy.* 34 (5), e2019G-e6383G.
- Di Vittorio, A.V., Mao, J., Shi, X., Chini, L., Hurtt, G., Collins, W.D., 2018. Quantifying the effects of historical land cover conversion uncertainty on global carbon and climate estimates. *Geophys. Res. Lett.* 45 (2), 974–982.
- Doetterl, S., Stevens, A., Six, J., Merckx, R., Van Oost, K., Pinto, M.C., et al., 2015. Soil carbon storage controlled by interactions between geochemistry and climate. *Nature Geosci* 8 (10), 780–783.
- Eglin, T., Ciais, P., Piao, S., Barre, P., Bellassen, V., Cadule, P., et al., 2010. Historical and future perspectives of global soil carbon response to climate and land-use changes. *Tellus B* 62 (5), 700–718.
- ESA (2017). *ESA: land cover CCI product user guide version 2.0*. [online] Available from: <http://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2.0.pdf>, 2017.
- Eyring, V., Lamarque, J., Hess, P., Arfeuille, F., Bowman, K., Chipperfield, M., et al., 2013. Overview of IGAC/SPARC Chemistry-Climate Model Initiative (CCMI) community simulations in support of upcoming ozone and climate assessments. *SPARC Newsletter* 40, 48–66.
- Farley, K.A., Kelly, E.F., Hofstede, R.G., 2004. Soil organic carbon and water retention after conversion of grasslands to pine plantations in the Ecuadorian Andes. *Ecosystems* 7 (7), 729–739.
- Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X., 2010. MODIS Collection 5 global land cover: algorithm refinements and characterization of new datasets. *Remote Sens. Environ.* 114 (1), 168–182.
- Friedl, M., Sulla-Menashe, D., 2019. MCD12Q1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC. <https://doi.org/10.5067/MODIS/MCD12Q1.006>.
- Friedlingstein, P., O'Sullivan, M., Jones, M.W., Andrew, R.M., Hauck, J., Olsen, A., et al., 2020. Global carbon budget 2020. *Earth Syst. Sci. Data* 12 (4), 3269–3340.
- Fuchs, R., Herold, M., Verburg, P.H., Clevers, J.G.P.W., Eberle, J., 2015. Gross changes in reconstructions of historic land cover/use for Europe between 1900 and 2010. *Global Change Biol.* 21 (1), 299–313.
- Gao, J., Zhang, X., Lei, G., Wang, G., 2014. Soil organic carbon and its fractions in relation to degradation and restoration of wetlands on the Zoige Plateau, China. *Wetlands* 34 (2), 235–241.
- Global Soil Data Task, 2014. Global soil data products CD-ROM contents (IGBP-DIS). from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A. <https://doi.org/10.3334/ORNLDAAC/565>. Retrieved from Data Set. Available online. <http://daac.ornl.gov>.
- Goll, D.S., Brovkin, V., Liski, J., Raddatz, T., Thum, T., Todd Brown, K.E.O., 2015. Strong dependence of CO₂ emissions from anthropogenic land cover change on initial land cover and soil carbon parametrization. *Global Biogeochem. Cy.* 29 (9), 1511–1523.
- Guevara, M., Arroyo, C., Brunsell, N., Cruz, C.O., Domke, G., Equihua, J., et al., 2020. Soil organic carbon across Mexico and the conterminous United States (1991–2010). *Global Biogeochem. Cy.* 34, e2019GB006219.
- Guo, L.B., Gifford, R.M., 2002. Soil carbon stocks and land use change: a meta analysis. *Global Change Biol.* 8 (4), 345–360.
- Harris, I. C. (2019). University of East Anglia Climatic Research Unit: CRU JRA: Collection of CRU JRA forcing datasets of gridded land surface blend of Climatic Research Unit (CRU) and Japanese reanalysis (JRA) data. Centre Environ. Data Anal. <http://catalogue.ceda.ac.uk/>.
- Hartley, A.J., MacBean, N., Georgievski, G., Bontemps, S., 2017. Uncertainty in plant functional type distributions and its impact on land surface models. *Remote Sens. Environ.* 203, 71–89.
- Hengl, T., de Jesus, J.M., Heuvelink, G.B., Gonzalez, M.R., Kilibarda, M., Blagotić, A., et al., 2017. SoilGrids250m: global gridded soil information based on machine learning. *PLoS One* 12 (2), e0169748.
- Houghton, R.A., 2008. Carbon Flux to the Atmosphere From Land-Use Changes: 1850–2005. TRENDS: A Compendium of Data on Global Change. Carbon Dioxide Information Analysis Center. Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A, pp. 1850–2005.
- Houghton, R.A., House, J.I., Pongratz, J., van der Werf, G.R., DeFries, R.S., Hansen, M.C., et al., 2012. Carbon emissions from land use and land-cover change. *Biogeosciences* 9 (12), 5125–5142.
- Houghton, R.A., Hackler, J.L., 2001. Carbon flux to the atmosphere from land-use changes: 1850 to 1990. ORNL/CDIAC-131. NDP-050/R1. <https://doi.org/10.3334/CDIAC/LUE.002>.
- Houghton, R.A., Nassikas, A.A., 2017. Global and regional fluxes of carbon from land use and land cover change 1850–2015. *Global Biogeochem. Cy.* 31 (3), 456–472.
- Huon, S., De Rouw, A., Bonté, P., Robain, H., Valentin, C., Lefèvre, I., et al., 2013. Long-term soil carbon loss and accumulation in a catchment following the conversion of forest to arable land in northern Laos. *Agr. Ecosyst. Environ.* 169, 43–57.
- Hurt, G.C., Chini, L.P., Frolking, S., Betts, R.A., Feddema, J., Fischer, G., et al., 2011. Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Clim. Change* 109 (1–2), 117.
- Hurt, G.C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B.L., Calvin, K., et al., 2020. Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6. *Geosci. Model Dev.* 13 (11), 5425–5464.
- Jain, A.K., Meiyappan, P., Song, Y., House, J.I., 2013. CO₂ emissions from land-use change affected more by nitrogen cycle, than by the choice of land-cover data. *Global Change Biol.* 19 (9), 2893–2906.
- Jain, A.K., Yang, X., 2005. Modeling the effects of two different land cover change data sets on the carbon stocks of plants and soils in concert with CO₂ and climate change. *Global Biogeochem. Cy.* 19 (2), B2015.
- Jobbágy, E.G., Jackson, R.B., 2000. The vertical distribution of soil organic carbon and its relation to climate and vegetation. *Ecol. Appl.* 10 (2), 423–436.
- Jung, M., Henkel, K., Herold, M., Churkina, G., 2006. Exploiting synergies of global land cover products for carbon cycle modeling. *Remote Sens. Environ.* 101 (4), 534–553.
- Keenan, R.J., Reams, G.A., Frédéric, A.C., De Freitas, J.V., Grainger, A., Lindquist, E., 2015. Dynamics of global forest area: results from the FAO Global Forest Resources Assessment 2015. *Forest Ecol. Manag.* 352, 9–20.
- Korkaç, S.Y., 2014. Effects of afforestation on soil organic carbon and other soil properties. *Catena* 123, 62–69.
- Lal, R., 2004. Soil carbon sequestration impacts on global climate change and food security. *Science* 304 (5677), 1623–1627.
- Lawrence, D.M., Hurtt, G.C., Arneth, A.K., Brovkin, V., Calvin, K.V., Jones, A.D., et al., 2016. The Land Use Model Intercomparison Project (LUMIP) contribution to CMIP6: rationale and experimental design. *Geosci. Model Dev.* 9, 2973–2998.
- Li, D., Niu, S., Luo, Y., 2012. Global patterns of the dynamics of soil carbon and nitrogen stocks following afforestation: a meta-analysis. *New Phytol.* 195 (1), 172–181.
- Li, W., MacBean, N., Ciais, P., Defourny, P., Lamarche, C., Bontemps, S., et al., 2018. Gross and net land cover changes in the main plant functional types derived from the annual ESA CCI land cover maps (1992–2015). *Earth Syst. Sci. Data* 10 (1), 219–234.
- Liang, C., Balser, T.C., 2011. Microbial production of recalcitrant organic matter in global soils: implications for productivity and climate policy. *Nat. Rev. Microbiol.* 9 (1), 75.
- Liu, H., Gong, P., Wang, J., Clinton, N., Bai, Y., Liang, S., 2020. Annual dynamics of global land cover and its long-term changes from 1982 to 2015. *Earth Syst. Sci. Data* 12 (2), 1217–1243.
- Luo, Y., Ahlström, A., Allison, S.D., Batjes, N.H., Brovkin, V., Carvalhais, N., et al., 2016. Toward more realistic projections of soil carbon dynamics by Earth system models. *Global Biogeochem. Cy.* 30 (1), 40–56.
- Meiyappan, P., Jain, A.K., 2012. Three distinct global estimates of historical land-cover change and land-use conversions for over 200 years. *Front. Earth Sci.* 6 (2), 122–139.
- Obermeier, W.A., Nabel, J., Loughran, T., Hartung, K., Pongratz, J., 2021. Modelled land use and land cover change emissions – A spatio-temporal comparison of different approaches. *Earth Syst. Dynam.* 12 (2), 635–670.
- Pan, S., Tian, H., Dangal, S., Zhang, C., Yang, J., Tao, B., et al., 2014. Complex spatiotemporal responses of global terrestrial primary production to climate change and increasing atmospheric CO₂ in the 21st century. *PLoS One* 9 (11), e112810.
- Pan, S., Yang, J., Tian, H., Shi, H., Chang, J., Ciais, P., et al., 2020. Climate extreme versus carbon extreme: responses of terrestrial carbon fluxes to temperature and precipitation. *J. Geophys. Res. Biogeod.* 125 (4), e2019J-e5252J.

- Peng, S., Ciais, P., Maignan, F., Li, W., Chang, J., Wang, T., Yue, C., 2017. Sensitivity of land-use change emission estimates to historical land-use and land-cover mapping. *Global Biogeochem. Cy.* 31 (4), 626–643.
- Petrenko, C.L., Friedland, A.J., 2015. Mineral soil carbon pool responses to forest clearing in Northeastern hardwood forests. *GCB-Bioenergy* 7 (6), 1283–1293.
- Poeplau, C., Don, A., Vesterdal, L., Leifeld, J., Van Wesemael, B., Schumacher, J., Gensior, A., 2011. Temporal dynamics of soil organic carbon after land-use change in the temperate zone—carbon response functions as a model approach. *Global Change Biol.* 17 (7), 2415–2427.
- Poeplau, C., Don, A., 2013. Sensitivity of soil organic carbon stocks and fractions to different land-use changes across Europe. *Geoderma* 192, 189–201.
- Pongratz, J., Reick, C.H., Houghton, R.A., House, J.I., 2014. Terminology as a key uncertainty in net land use and land cover change carbon flux estimates. *Earth Syst. Dynam.* 5 (1), 177–195.
- Powers, J.S., Corre, M.D., Twine, T.E., Veldkamp, E., 2011. Geographic bias of field observations of soil carbon stocks with tropical land-use changes precludes spatial extrapolation. *P. Natl. Acad. Sci.* 108 (15), 6318–6322.
- Ramankutty, N., Evan, A.T., Monfreda, C., Foley, J.A., 2008. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochem. Cy.* 22 (1), GB1003.
- Ramankutty, N., Foley, J.A., 1999. Estimating historical changes in global land cover: croplands from 1700 to 1992. *Global Biogeochem. Cy.* 13 (4), 997–1027.
- Ramesh, T., Bolan, N.S., Kirkham, M.B., Wijesekara, H., Kanchikerimath, M., Rao, C.S., et al., 2019. Soil organic carbon dynamics: impact of land use changes and management practices: a review. *Adv. Agron.* 1–170.
- Reeves, D.W., 1997. The role of soil organic matter in maintaining soil quality in continuous cropping systems. *Soil Till. Res.* 43 (1), 131–167.
- Regnier, P., Friedlingstein, P., Ciais, P., Mackenzie, F.T., Gruber, N., Janssens, I.A., et al., 2013. Anthropogenic perturbation of the carbon fluxes from land to ocean. *Nat. Geosci.* 6 (8), 597–607.
- Reick, C.H., Raddatz, T., Brovkin, V., Gayler, V., 2013. Representation of natural and anthropogenic land cover change in MPI-ESM. *J. Adv. Model Earth Sy.* 5 (3), 459–482.
- Ren, W., Banger, K., Tao, B., Yang, J., Huang, Y., Tian, H., 2020. Global pattern and change of cropland soil organic carbon during 1901–2010: roles of climate, atmospheric chemistry, land use and management. *Geography Sustain.* 1 (1), 59–69.
- Scharlemann, J.P., Tanner, E.V., Hiederer, R., Kapos, V., 2014. Global soil carbon: understanding and managing the largest terrestrial carbon pool. *Carbon Manag* 5 (1), 81–91.
- Seibert, J., Stendahl, J., Sørensen, R., 2007. Topographical influences on soil properties in boreal forests. *Geoderma* 141 (1–2), 139–148.
- Shevliakova, Pacala, S.W., Malyshev, Hurtt, G.C., Milly, P.C.D., Caspersen, J.P., et al., 2009. Carbon cycling under 300 years of land use change: importance of the secondary vegetation sink. *Global Biogeochem. Cy.* 23, 1–16.
- Smal, H., Olszewska, M., 2008. The effect of afforestation with Scots pine (*Pinus silvestris* L.) of sandy post-arable soils on their selected properties. II. Reaction, carbon, nitrogen and phosphorus. *Plant Soil* 305 (1–2), 171–187.
- Smith, P., 2008. Land use change and soil organic carbon dynamics. *Nutri. Cycl. Agroecos.* 81 (2), 169–178.
- Sulla-Menashe, D., Gray, J.M., Abercrombie, S.P., Friedl, M.A., 2019. Hierarchical mapping of annual global land cover 2001 to present: the MODIS Collection 6 Land Cover product. *Remote Sens. Environ.* 222, 183–194.
- Sulla-Menashe, D., Friedl, M. A. (2018). User Guide to Collection 6 MODIS land Cover (MCD12Q1 and MCD12C1) Product. USGS: Reston, VA, USA, 1–18. https://lpdaac.usgs.gov/documents/101/MCD12_User_Guide_V6.pdf.
- Tan, Z.X., Lal, R., Smeck, N.E., Calhoun, F.G., 2004. Relationships between surface soil organic carbon pool and site variables. *Geoderma* 121 (3–4), 187–195.
- Tian, H., Chen, G., Lu, C., Xu, X., Hayes, D.J., Ren, W., et al., 2015. North American terrestrial CO₂ uptake largely offset by CH₄ and N₂O emissions: toward a full accounting of the greenhouse gas budget. *Clim. Change* 129 (3–4), 413–426.
- Tian, H., Chen, G., Zhang, C., Melillo, J.M., Hall, C.A., 2010. Pattern and variation of C:N:P ratios in China's soils: a synthesis of observational data. *Biogeochemistry* 98 (1–3), 139–151.
- Tian, H., Lu, C., Yang, J., Banger, K., Huntzinger, D.N., Schwalm, C.R., et al., 2015. Global patterns and controls of soil organic carbon dynamics as simulated by multiple terrestrial biosphere models: current status and future directions. *Global Biogeochem. Cy.* 29 (6), 775–792.
- Tian, H., Melillo, J., Lu, C., Kicklighter, D., Liu, M., Ren, W., et al., 2011a. China's terrestrial carbon balance: contributions from multiple global change factors. *Global Biogeochem. Cy.* 25 (1), 775–792.
- Tian, H., Xu, X., Lu, C., Liu, M., Ren, W., Chen, G., et al., 2011b. Net exchanges of CO₂, CH₄, and N₂O between China's terrestrial ecosystems and the atmosphere and their contributions to global climate warming. *J. Geophys. Res.* 116, G2011.
- Todd-Brown, K.E.O., Randerson, J.T., Post, W.M., Hoffman, F.M., Tarnocai, C., Schuur, E. A.G., Allison, S.D., 2013. Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations. *Biogeosciences* 10 (3), 1717–1736.
- Vesterdal, L., Ritter, E., Gundersen, P., 2002. Change in soil organic carbon following afforestation of former arable land. *Forest Ecol. Manag.* 169 (1), 137–147.
- Wang, Z., et al., 2020. Coupling of phosphorus processes with carbon and nitrogen cycles in the dynamic land ecosystem model: model structure, parameterization, and evaluation in tropical forests. *J. Adv. Model Earth Sy.* 12 (10), e2020M-e2123M.
- Wang, M., et al., 2019. Patterns and drivers of soil carbon stock in southern China's grasslands. *Agr. Forest Meteorol.* 276, 107634.
- Wieder, W. R., Boehner, J., Bonan, G. B., Langseth, M. (2014). RegridDED harmonized world soil database v1. 2. Retrieved from Data set. Available on-line [<http://daac.ornl.gov>] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA. 10.3334/ORNLDAAAC/1247.
- Winkler, K., Fuchs, R., Rounsevell, M., Herold, M., 2021. Global land use changes are four times greater than previously estimated. *Nat. Commun.* 12 (1), 2501.
- Wright, S.J., 2005. Tropical forests in a changing environment. *Trends Ecol. Evol.* 20 (10), 553–560.
- Xie, Z., Zhu, J., Liu, G., Cadisch, G., Hasegawa, T., Chen, C., et al., 2007. Soil organic carbon stocks in China and changes from 1980s to 2000s. *Global Change Biol.* 13 (9), 1989–2007.
- Xu, W., Chang, J., Ciais, P., Guenet, B., Viovy, N., Ito, A., et al., 2020. Reducing uncertainties of future global soil carbon responses to climate and land use change with emergent constraints. *Global Biogeochem. Cy.* 34 (10).
- Yang, J., Tian, H., Tao, B., Ren, W., Lu, C., Pan, S., et al., 2015. Century-scale patterns and trends of global pyrogenic carbon emissions and fire influences on terrestrial carbon balance. *Global Biogeochem. Cy.* 29 (9), 1549–1566.
- Yao, Y., Tian, H., Shi, H., Pan, S., Xu, R., Pan, N., Canadell, J.G., 2020. Increased global nitrous oxide emissions from streams and rivers in the Anthropocene. *Nat. Clim. Chang.* 10 (2), 138–142.
- Yu, Z., Lu, C., Cao, P., Tian, H., 2018. Long-term terrestrial carbon dynamics in the Midwestern United States during 1850–2015: roles of land use and cover change and agricultural management. *Global Change Biol.* 24 (6), 2673–2690.
- Zhang, B., Tian, H., Ren, W., Tao, B., Lu, C., Yang, J., et al., 2016. Methane emissions from global rice fields: Magnitude, spatiotemporal patterns, and environmental controls. *Global Biogeochem. Cy.* 30 (9), 1246–1263.
- Zhang, J., Tian, H., Yang, J., Pan, S., 2018. Improving representation of crop growth and yield in the Dynamic Land Ecosystem Model and its application to China. *J. Adv. Model. Earth Syst.* 10 (7), 1680–1707.
- Zhou, G., Xu, S., Ciais, P., Manzoni, S., Fang, J., Yu, G., et al., 2019. Climate and litter C/N ratio constrain soil organic carbon accumulation. *Natl. Sci. Rev.* 6 (4), 746–757.