Uncertainty in Land Carbon Fluxes Simulated by CMIP6 Models from Treatments of Crop Distributions and Photosynthetic Pathways.

Joseph Ovwemuvwose[1], Ian C. Prentice[2, 3], and Heather. D. Graven[1]

[1]Department of Physics, Imperial College London, London, UK.
[2]Georgina Mace Centre for the Living Planet, Department of Life Sciences, Imperial College London, London, UK
[3]Department of Earth System Science, Tsinghua University, Beijing, China

Correspondence to: Joseph Ovwemuvwose (j.ovwemuvwose22@imperial.ac.uk & josovw@gmail.com)

Supplementary materials

Contents of this File:

Material and Method S1 - Description of implementation of vegetation cover in CMIP6 models

S1.1 ACCESS-ESM1.5

In the Australian Community Climate and Earth System Simulator Earth System Model version 1.5 (ACCESS-ESM1.5) land surface and biogeochemical processes are simulated using the Community Atmosphere Biosphere Land Exchange (CABLE) (Kowalczyk et al., 2013) model and a modified Carnegie–Ames–Stanford Approach (CASA) carbon (C) cycle model which incorporates nitrogen (N) and phosphorus (P) cycles (CASA-CNP module) (Ziehn et al., 2020). The area abundance of crops for the historical period is based on the LUH2 dataset (Hurtt et al., 2020). The 12 land-use types in the LUH2 dataset are mapped into 9 plant functional types (PFTs) in CABLE including C_3 grass, C_3 crops and C_4 crops. The rest of the PFTs (evergreen needleleaf forest, evergreen broadleaf forest, deciduous needleleaf forest, deciduous broadleaf forest, shrub and tundra) are not specified as either C_3 or C_4 but most of them would be C_3 vegetation since it is the most prevalent of these two photosynthetic pathways in the given PFTs. However, even though a C_4 crop PFT was specified in the model description (Ziehn et al., 2020), the variable cropFracC4 was not provided in the CMIP6 archive. The values for cropFracC3 and for the total crop fraction (cropFrac) were identical, so we specified the fraction of C_4 crops to be zero everywhere.

S1.2 CanESM5

The land component of the Canadian Earth System Model version 5 (CanESM5) is comprised of the Canadian Land Surface Scheme (CLASS) and the Canadian Terrestrial Ecosystem Model (CTEM) making up the CLASS-CTEM modelling framework. CLASS simulates the physical and CTEM simulates the biogeochemical land surface processes. Energy and water balances and land surface physical processes are calculated in CLASS using four vegetation categories: needleleaf trees, broadleaf trees, grasses and crops. Each of these is divided into sub-categories: needleleaf trees into evergreen and deciduous, broadleaf trees into three categories; cold deciduous, drought deciduous, and evergreen, crops into C_3 and C_4 types, and grasses into C_3 and C_4 types to form 9 PFTs (Swart et al., 2019). The increase in the area abundance of crops for the historical period is based on the LUH2 dataset (Hurtt et al., 2020).

S1.3 CESM2, CESM2-WACCM, CMCC-CM2-SR5 and CMCC-ESM2-HAM

The Community Earth System Model Version 2 (CESM2) and Whole Atmosphere Community Climate Model (WACCM) are developed by teams of individuals from the National Centre for Atmospheric Research (NCAR), and other research institutions and universities. The major difference between CESM2 and CESM2-WACCM is the vertical extent of their atmospheric configuration. In CESM2-WACCM the simulated properties extend up to 130-140 km in the upper atmosphere with 70 vertical levels (Gettelman et al., 2019) whereas CESM2 has only 32 vertical levels and a model top of 40 km, without a prognostic scheme for ozone and other stratospheric entities (Danabasoglu et al., 2020).

CESM2 and CESM2-WACCM use the Community Land Model Version 5 (CLM5) for the simulation of terrestrial biogeochemical cycles including interactions with the soil-plant-atmospheric continuum, anthropogenic land use change and vegetation coverage (Danabasoglu et al., 2020; Lawrence et al., 2019). CLM5 computes land unit weights through two methods: dataset input or prognostic representation. In the prognostic representation, the land surface heterogeneity is represented in several land units which are divided into columns and the columns are divided into patches. The land units are vegetated land, lake, urban, glacier, and cropland. The vegetated land unit is divided into patches of 14 plant or crop functional type (P/CFTs) which are (N = needleleaf; B = broadleaf; E = evergreen; D = deciduous; T = tree; S = shrub): NET temperate, NET boreal, NDT boreal, BET tropical, BET temperate, BDT tropical, BDT temperate, BDT boreal, BES temperate, BDS temperate, BDS boreal, C_3 Arctic grass, C_3 grass and C_4 grass. Crops are categorised into C_3 unmanaged rainfed crop, C_3

unmanaged irrigated crop, managed rainfed unirrigated crops and managed irrigated crops. Apart from the last three PFTs, it is not explicitly stated which of the other PFTs uses the C_3 or C_4 photosynthetic pathways. However, their output variables provided include the percentage area coverage of all C_3 and C_4 vegetation. The CLM uses its Land Use Data Tool to translate the LUH2 dataset transitions and information into its P/CFTs distribution and management (Hurtt et al., 2020; Lawrence et al., 2019). The CESM2 models did not provide cropFracC3 and cropFracC4 variables, but they did provide c3PftFrac or c4PftFrac (the fractions for total vegetation), so cropFracC3 and cropFracC4 were calculated by multiplying cropFrac by c3PftFrac or c4PftFrac.

The Euro-Mediterranean Centre on Climate Change (CMCC) Foundation model simulates biogeochemical cycles and land use change with the Community Land Model version 4.5 (CLM4.5) (Lovato et al., 2022), which is an earlier version of the model used in CESM2 and CESM2-WACCM. The land use implementation is similar to CESM2, which implies they used the LUH2 dataset as well (Lawrence et al., 2019). However, unlike CESM2, the CMCC models do provide cropFracC3, although not cropFracC4. Their cropFracC3 was equal to the total crop fraction (cropFrac), so we specified the fraction of C_4 crops to be zero everywhere.

S1.4 CNRM-CM6.1 and CNRM-ESM2.1

CNRM-CM6.1 is the sixth generation of the fully coupled atmosphere-ocean general circulation model jointly developed by Centre National de Recherches Météorologiques (CNRM) and Cerfacs. Its land surface is simulated with the Interaction Soil-Biosphere-Atmosphere-CNRM TRIP (ISBA-CTRIP) coupled system (Decharme et al., 2019). Its vegetation cover is based on the ECOCLIMAP-II database in which each land cover type is portioned into fractions of 4 main surface types or tiles (nature, water bodies, sea, urban areas). The nature tile is further divided into bare soil, bare rock or one of ten PFTs. CNRM-CM6-1 adopts a fixed vegetation cover map from ECOCLIMAP-II (Faroux et al., 2013; Voldoire et al., 2019). It does not incorporate LUH2 datasets.

CNRM-ESM2.1 and CNRM-CM6.1 share the same code, physical parameterization and grid resolution but vegetation coverage in CNRM-ESM2.1 is dynamic, incorporating the LUH2 crop coverage dataset. In CNRM-CM6-1 and CNRM-ESM2.1, the cropFracC3 and cropFracC4 variables were identical to the c3PftFrac and c4PftFrac variables, which appears to be an error in the output. So, in CNRM-CM6-1 and CNRM-ESM2.1 cropFracC3 and cropFracC4 were calculated by multiplying cropFrac by c3PftFrac or c4PftFrac.

S1.5 MPI-ESM1.2

The Max Planck Institute Earth System Model version 1.2 (MPI-ESM1.2) uses the land surface scheme JSBACH which focuses on the competition between two vegetation types: graminoids (grass-like) and woody species (trees and shrubs). These two broad categories are further divided into different PFTs. To integrate the LUH2 dataset into their JSBACH simulation, they aggregated the four LUH2 natural vegetation and grasslands states into the natural vegetation type, the five LUH2 crop variables into crops, the LUH2 pasture and urban into pasture. These three broad types: natural vegetation, crops and pastures are then proportionally adjusted to sum up to one (Mauritsen et al., 2019). JSBACH dynamic vegetation component DYNVEG has 8 PFTs, 6 of which are woody (tropical evergreen trees, tropical deciduous trees, extratropical evergreen trees, extratropical deciduous trees, raingreen shrubs, deciduous shrubs) and 2 of which are grass (C_3 grasses and C_4 grasses). In the model, prevalent climate conditions determine the extent of the presence or absence of vegetation in hot and cold regions (Reick et al., 2013).

S1.5 UKESM1-0-LL

The UK Earth System Model version 1 (UKESM1) (Mulcahy et al., 2023; Sellar et al., 2019) uses a physical core built on a prior model, HadGEM3-GC3.1 (Kuhlbrodt et al., 2018; Williams et al., 2018). It simulates the terrestrial ecosystem biogeochemistry and land surface dynamics using the Joint UK Land Environment Simulator (JULES) (Clark et al., 2011; Littleton et al., 2020) which is built on the Top-down Representation of Interactive Foliage and Flora Including Dynamics (TRIFFID) which uses the LUH2 crop cover data with some modifications including simulation of competition between C_3 and C_4 (Cox et al., 2001). It uses Rothamsted carbon model soil carbon scheme (RothC) (Falloon et al., 1998). TRIFFID is a dynamic vegetation model that divides the earth surface into five components: vegetated land, urban and lake surface types, bare soil, and land covered permanently by ice. The vegetated component is divided into five PFTs : broadleaf trees, needleleaf trees, C_3 (temperate) grasses, C_4 (tropical) grasses and shrubs. These are assumed to sufficiently represent the phenotypic and genotypic variations needed to capture biophysical and biogeochemical roles of vegetation in earth system models (Sellar et al., 2019, Clark et al., 2011).

Table S1. Global land area in each category in LUH2 and in CMIP6 models for 1970 and 2014.

	Area Abundance (Million km^2)			
	C3 crop	**C4 Crop**	**C3 Natural**	**C4 Natural**
LUH2 1970 (2014)	**11.59 (12.75)**	**2.68 (3.34)**		
Luo (2014)		**(3.32)**		**(17.02)**
Ensemble mean 1970 (2014)	10.71 (11.55)	1.81 (2.31)	66.76 (66.41)	15.33 (15.29)
ACCESS-ESM1-5 1970 (2014)	12.41 (13.76)	0.00 (0.00)	67.66 (67.35)	9.97 (8.95)
CanESM5 1970 (2014)	11.43 (12.41)	2.68 (3.30)	64.96 (63.92)	6.88 (6.33)
CESM2 1970 (2014)	9.42 (9.94)	3.04 (3.92)	71.34 (70.48)	19.44 (18.91)
CESM2-WACCM 1970 (2014)	9.42 (9.94)	3.04 (3.92)	71.34 (70.48)	19.44 (18.91)
CMCC-CM2-SR5 1970 (2014)	13.21 (14.68)	0.00 (0.00)	68.99 (67.79)	19.63 (19.35)
CMCC-ESM2 1970 (2014)	13.21 (14.68)	0.00 (0.00)	68.99 (67.79)	19.63 (19.35)
CNRM-CM6.1 Fixed	9.71 (9.71)	3.00 (3.00)	59.82 (59.82)	19.22 (19.22)
CNRM-ESM2.1 1970 (2014)	9.48 (10.03)	2.26 (3.10)	65.01 (65.37)	20.09 (19.56)
MPI-ESM1.2-HAM 1970 (2014)	11.89 (13.19)	1.35 (1.96)	70.72 (71.18)	8.75 (10.18)
MPI-ESM1-2-LR 1970 (2014)	11.96 (13.03)	1.14 (1.60)	68.33 (68.59)	8.79 (9.91)
UKESM1-LL 1970 (2014)	5.68 (5.65)	3.36 (4.62)	57.12 (57.76)	16.84 (17.47)

Table S2. Ranges of simulated terrestrial biosphere carbon flux variables for 2014 simulated by CMIP6 models: GPP is Gross Primary Production; C_{veg} is vegetation carbon content; dC_{veg}/dt is vegetation carbon accumulation from 1850 to 2014; C_{land} is terrestrial biosphere carbon content; dC_{land}/dt is terrestrial biosphere carbon accumulation from 1850 to 2014.

Mean and ranges across CMIP6 models' variables in 2014 (Mean [Minimum, Maximum])								
Variable	**Area (Million km^2)**	**GPP (PgCyr^{-1})**	**Cveg (PgC)**	**dCveg/dt (PgC)**	**Cland (PgC)**	**dCland/dt (PgC)**	**Discrimination (‰)**	**Discrimination change (‰)**
Total	95 [85, 105]	116 [93, 150]	490 [375, 672]	-10 [-70, 40]	2304 [1382, 3445]	20 [-47, 170]	18.12 [16.98, 19.15]	-0.23 [-0.68, 0.20]
Total Natural	82 [71, 90]	100 [80, 126)	416 [328, 573]	-27 [-78, 25]	1826 [1111, 2903]	-108 [-181, 3]	18.15 [17, 19.42]	-0.17 [-0.80, 0.22]
Total Crop	14 [10, 15.4]	17 [13, 26]	40 [25, 66]	18 [-18, to 36]	221 [130, 342]	135 [82, 191]	-0.06 [-0.7, 0.33]	-0.07 [-0.50, 0.20]
C3 total	78 [63, 84]	96 [73, 126]	407 [316, 593]	-9 [-68, 38]	1845 [1131, 2996]	2 [-57, 124]		
C3 Natural	66 [57, 71]	82 [62, 105]	375 [288, 539]	-24 [-84, 26]	1656 [993, 2654]	-111 [-171, -3]		
C3 crop	12 [6, 15]	14 [6, 21]	32 [13, 56]	14 [-9, 28]	189 [96, 342]	113 [58, 191]		
C4 total	18 [9, 13]	21 [12, 30]	47 [29, 77]	1 [-13, 9]	198 [86, 336]	23 [-8, 43]		
C4 Natural	15 [6, 19]	18 [15, 26]	41 [25, 58]	-3 [-9, 6]	170 [86, 249]	3 [-27, 33]		
C4 crop	2 [0, 4.6]	3 [0, 8]	6 [0, 24]	3 [-6, 18]	28 [0, 100]	20 [0, 67]		

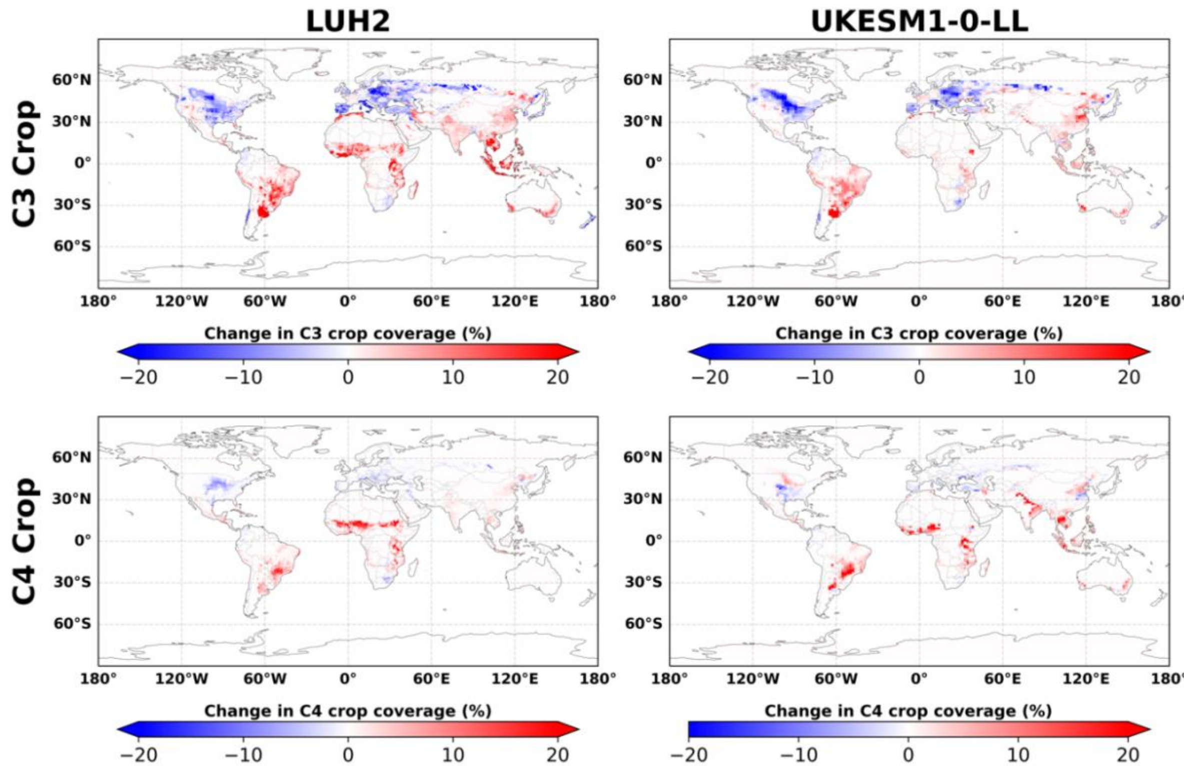

Figure S1: Percentage change in crop abundance between 1970 and 2014 for C_3 and C_4 in LUH2 and UKESM1-0-LL. All the other models not shown use the LUH2 crop variables for their crop coverage and change. CNRM-CM6-1 is not shown here because it has fixed crop coverage.

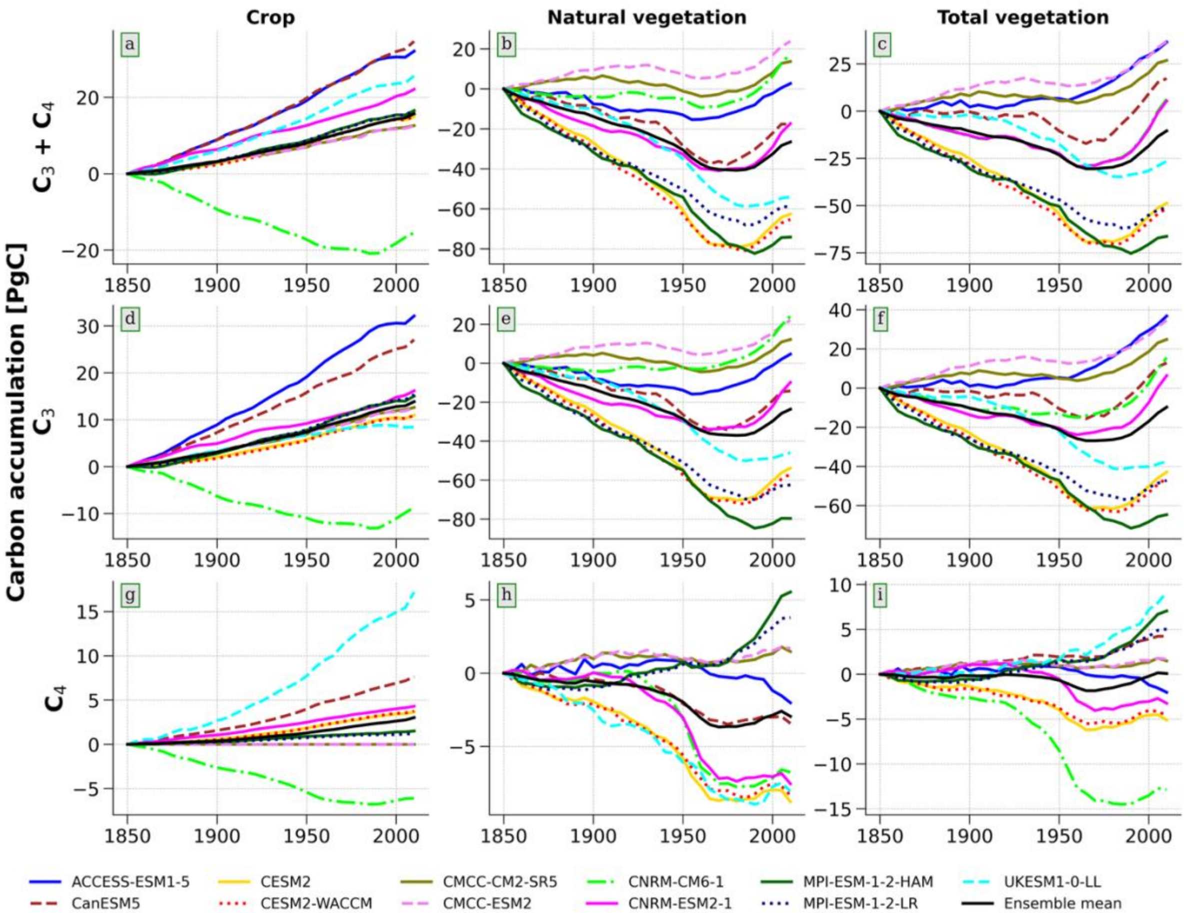

Figure S2: Global vegetation carbon accumulation between 1850 and 2014. Values of carbon accumulation in (a) croplands, (b) natural vegetation and (c) total vegetation. (d) C_3 crop, (e) natural C_3 vegetation, (f) total C_3 vegetation, (g) C_4 crops, (h) natural C_4 vegetation and (i) total C_4 vegetation

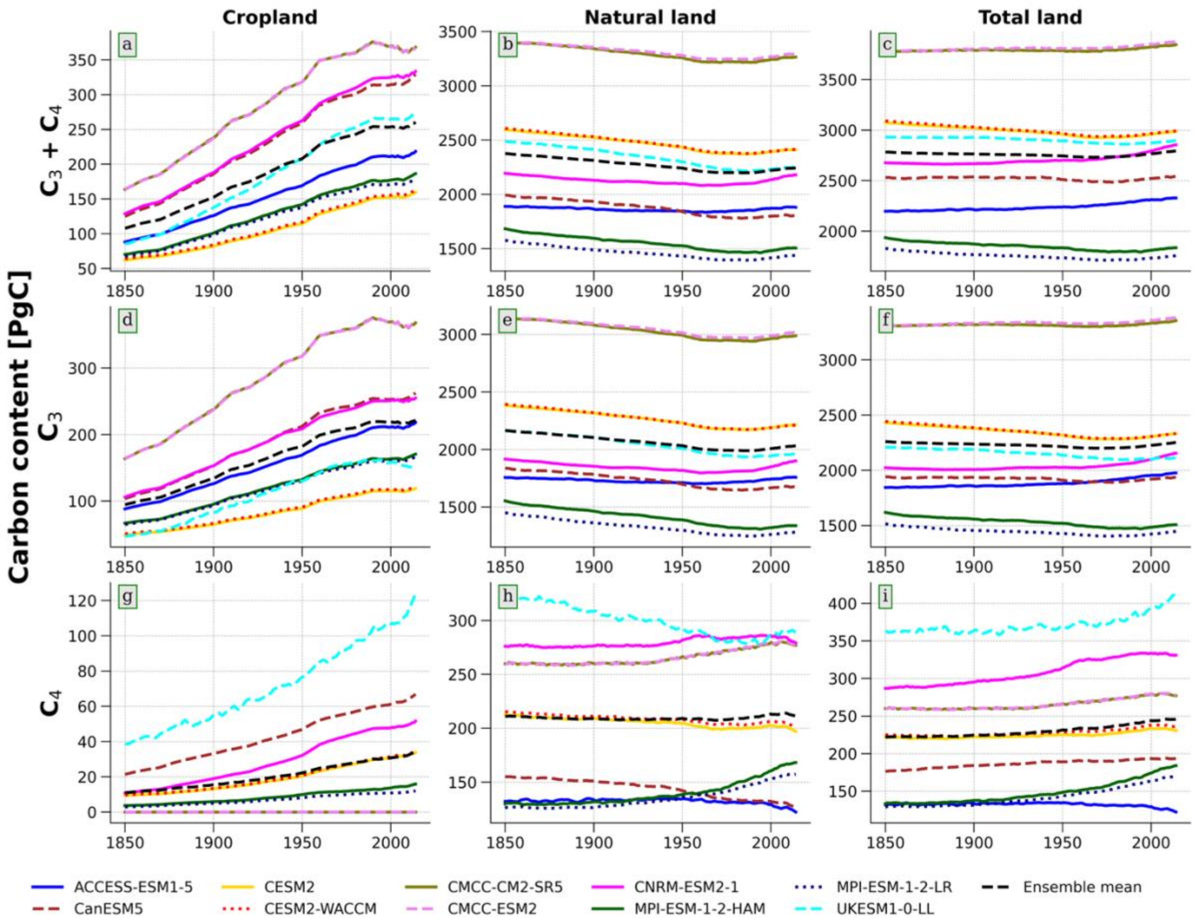

Figure S3: Global land carbon content between 1850 and 2014. Absolute values of carbon content in (a) croplands, (b) natural vegetated land (c) total vegetated land, (d) C_3 croplands, (e) natural C_3 vegetated land, (f) total C_3 vegetated land, (g) C_4 croplands, (h) natural C_4 vegetated land and (i) total C_4 vegetated lands.

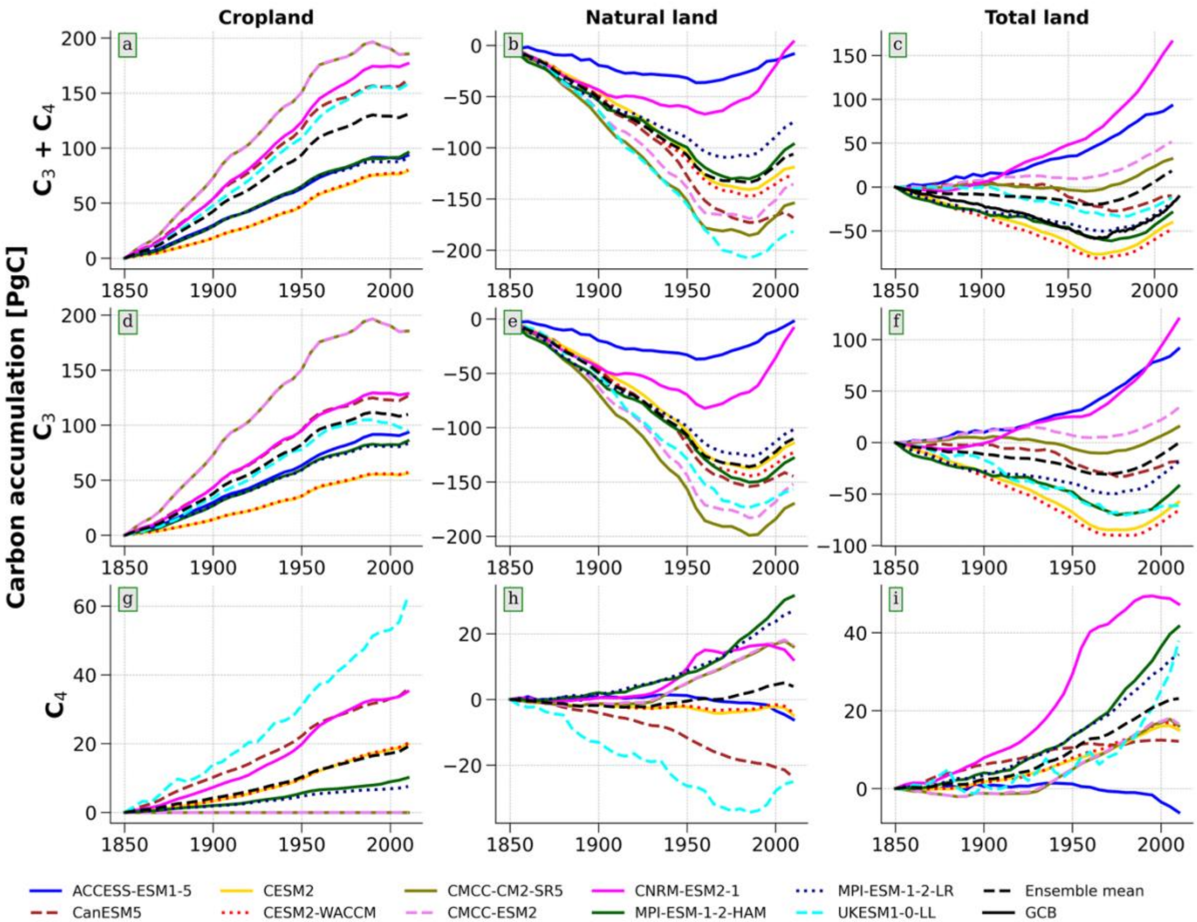

Figure S4: Global land carbon accumulation between 1850 and 2014. Absolute values of carbon content in (a) croplands, (b) natural vegetated land (c) total vegetated land, (d) C_3 croplands, (e) natural C_3 vegetated land, (f) total C_3 vegetated land, (g) C_4 croplands, (h) natural C_4 vegetated land and (i) total C_4 vegetated lands. Estimated global land carbon accumulation from the Global Carbon Budget by quantifying five components of the global carbon cycle. Global land carbon accummulaion provided here are based on 'land-use and land-use change data and bookkeeping models.' (Friedlingstein et al 2025).

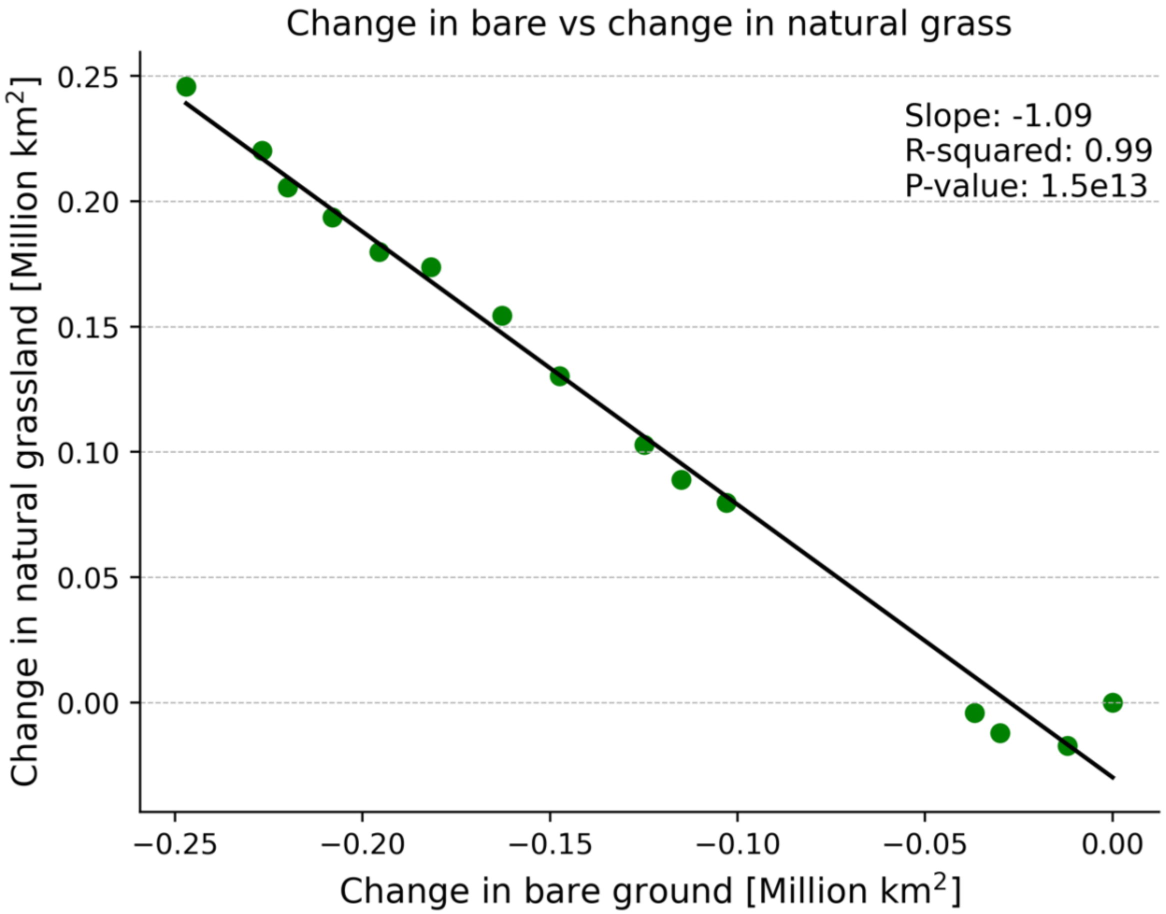

Figure S5: Relationship between the change in bare ground and natural grass in the ESA CCI land cover dataset between 2000 and 2014. The change is calculated as the difference between the values of the variables in 2000 and 2014.

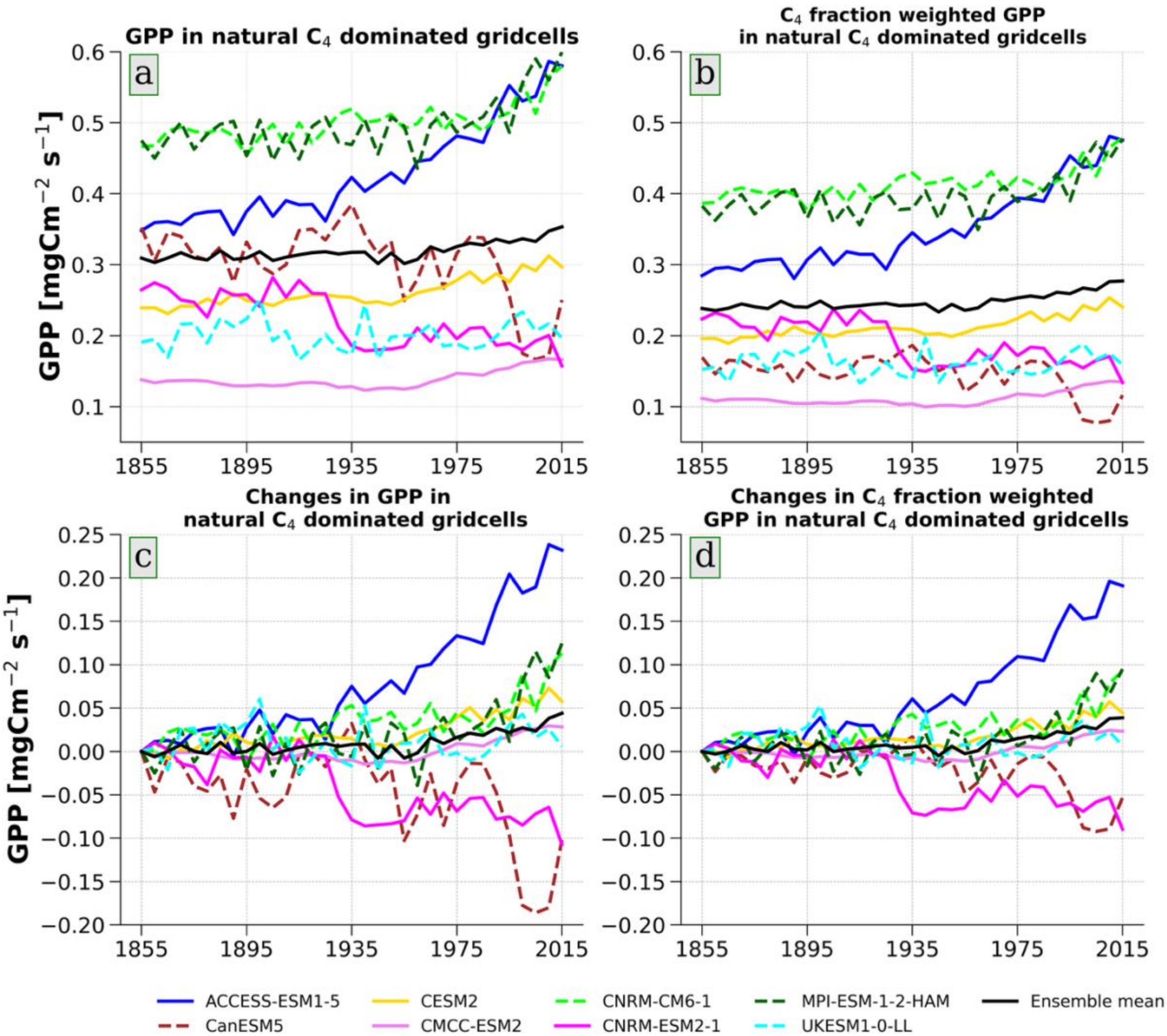

Figure S6: Temporal trend of GPP in natural C_4 dominated grid cells between 1850 and 2014. Here natural C_4 => 75% except in CanESM5 where there are no grid cells with natural C_4 => 75% and as such a threshold of natural C_4 => 45.5% is used instead. (a) GPP in natural C_4 dominated grid cells, (b) GPP in natural C_4 dominated grid cells weighted by the natural C_4 vegetation fraction, (c) Changes in the GPP in natural C_4 dominated grid cells, (d) changes in the GPP in natural C_4 dominated grid cells weighted by the natural C_4 vegetation fraction.

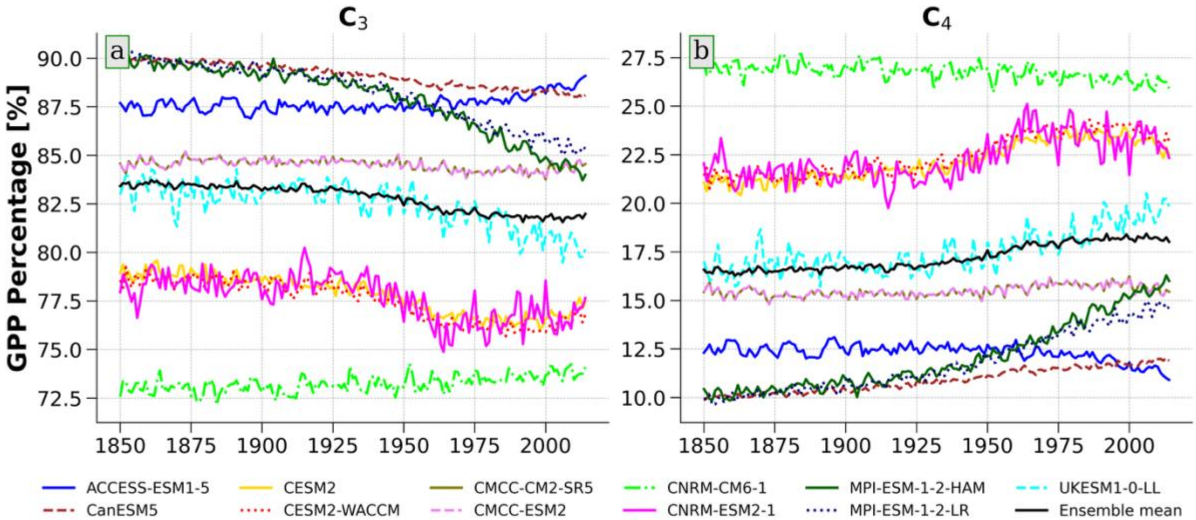

Figure S7: Temporal trend of C_3 and C_4 GPP percentage in CMIP6 models between 1850 and 2014. The model means show a decreasing C_3 and an increasing C_4 contribution to the total GPP.

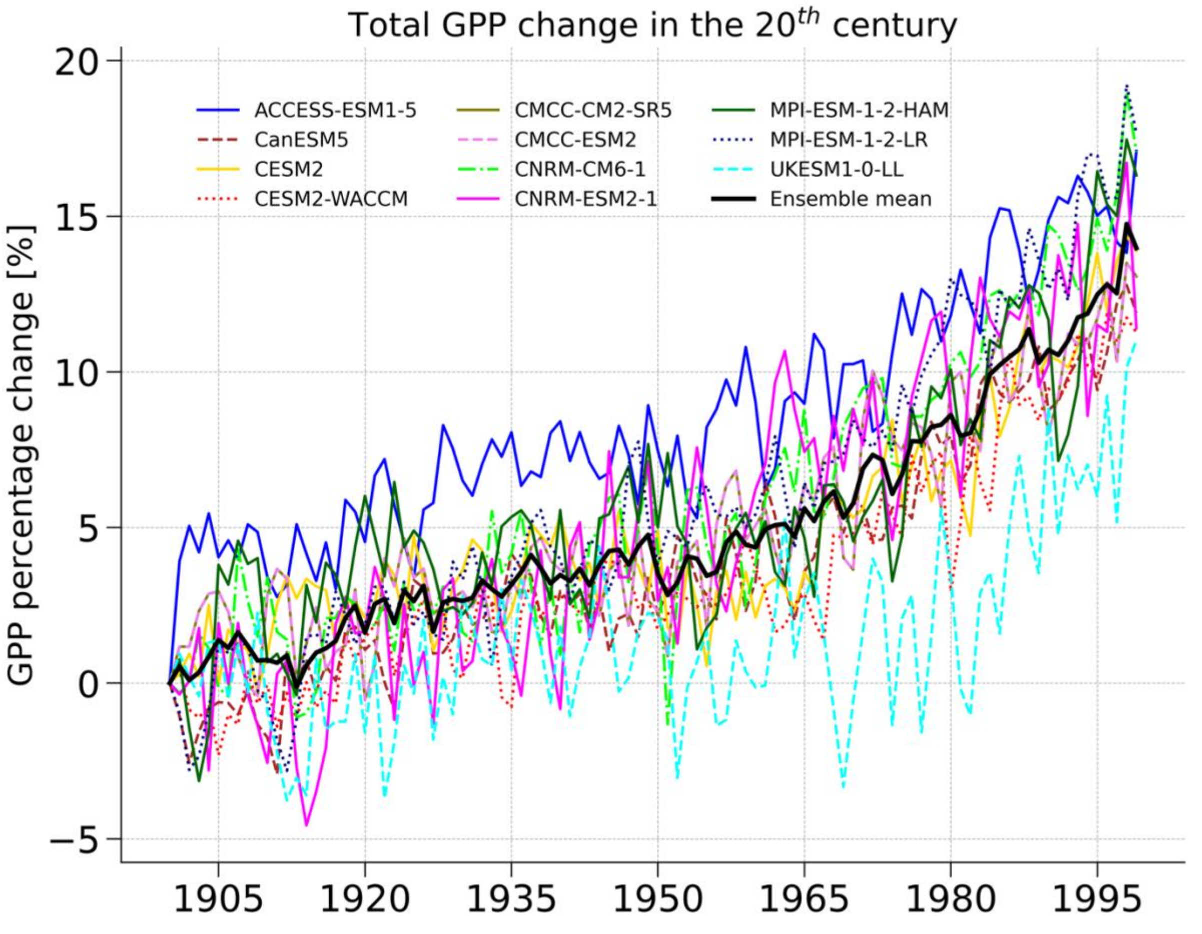

Figure S8: CMIP6 models percentage change in total GPP between 1900 and 2000. The value of the percentage change in the GPP is weak (11 to 17%) compared to 30% observed in other studies in the 20th century.

References

190 Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Landschützer, P., Le Quéré, C., Li, H., Luijkx, I. T., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., … Zeng.: Global Carbon Budget 2024, Earth System Science Data, 17(3), 965–1039. https://doi.org/10.5194/ESSD-17-965-2025, 2025.