

# Underrepresented controls of aridity in climate sensitivity of carbon cycle models

Sujan Koirala, Chris Jones, Bernhard Ahrens, Naixin Fan, Victor Brovkin, Christine Delire, Yuanchao Fan, Veronika Gayler, Emilie Joetzjer, Hanna Lee,

et al.

# ▶ To cite this version:

Sujan Koirala, Chris Jones, Bernhard Ahrens, Naixin Fan, Victor Brovkin, et al.. Underrepresented controls of aridity in climate sensitivity of carbon cycle models. 2022, 10.21203/rs.3.rs-2013805/v1. hal-04302319

# HAL Id: hal-04302319 https://hal.science/hal-04302319

Submitted on 23 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Preprints are preliminary reports that have not undergone peer review. They should not be considered conclusive, used to inform clinical practice, or referenced by the media as validated information.

# Underrepresented controls of aridity in climate sensitivity of carbon cycle models

# Sujan Koirala (Sujan.koirala@bgc-jena.mpg.de)

Max Planck Institute for Biogeochemistry

# **Chris Jones**

Met Office Hadley Centre https://orcid.org/0000-0002-7141-9285

## **Bernhard Ahrens**

Max Planck Institute for Biogeochemistry https://orcid.org/0000-0001-7226-6682

# Naixin Fan

Max Planck Institute for Biogeochemistry and Technische Universität Dresden, Institute of Photogrammetry and Remote Sensing

# Victor Brovkin

Max Plank Institute for Meteorology https://orcid.org/0000-0001-6420-3198

# **Christine Delire**

CNRM, Université de Toulouse, Météo-France, CNRS https://orcid.org/0000-0002-6114-3211

# Yuanchao Fan

NORCE Norwegian Research Centre, Bjerknes Centre for Climate Research and Center for the Environment, Faculty of Arts & Sciences, Harvard University

# Veronika Gayler

Max Plank Institute for Meteorology

## **Emilie Joetzjer**

Centre National de Recherches Météorologiques, Université de Toulouse/Météo-France/CNRS

## Hanna Lee

NORCE Norwegian Research Centre, Bjerknes Centre for Climate Research

## Stefano Materia

Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC)

## Julia Nabel

Max Planck Institute for Meteorology https://orcid.org/0000-0002-8122-5206

## **Daniele Peano**

Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC),

# **Pilippe Peylin**

CNRS

# David Wårlind

Lund University

## Andrew Wiltshire

Met Office Hadley Centre

# Sönke Zaehle

Max Planck Institute for Biogeochemistry https://orcid.org/0000-0001-5602-7956

# Markus Reichstein

Max Planck Institute for Biogeochemistry, Jena https://orcid.org/0000-0001-5736-1112

# Nuno Carvalhais

Max Planck Institute for Biogeochemistry

Article

Keywords:

Posted Date: September 1st, 2022

# DOI: https://doi.org/10.21203/rs.3.rs-2013805/v1

License: © ① This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

# Underrepresented controls of aridity in climate sensitivity of carbon cycle models

- <sup>4</sup> Sujan Koirala<sup>1</sup>, Chris Jones<sup>2</sup>, Bernhard Ahrens<sup>1</sup>, Naixin Fan<sup>1,3</sup>, Victor Brovkin<sup>4</sup>, Christine
- <sup>5</sup> Delire<sup>5</sup>, Yuanchao Fan<sup>6, 7</sup>, Veronika Gayler<sup>4</sup>, Emilie Joetzjer<sup>5</sup>, Hanna Lee<sup>6</sup>, Stefano Materia<sup>8</sup>,
- Julia E.M.S. Nabel<sup>1, 4</sup>, Daniele Peano<sup>8</sup>, Philippe Peylin<sup>9</sup>, David Wårlind<sup>10</sup>, Andy Wiltshire<sup>2</sup>,
- 7 Sönke Zaehle<sup>1</sup>, Markus Reichstein<sup>1</sup>, Nuno Carvalhais<sup>1,11</sup>
- 8

1

- <sup>9</sup> <sup>1</sup>Max Planck Institute for Biogeochemistry, Jena, Germany.
- <sup>10</sup> <sup>2</sup>Met Office Hadley Centre, Exeter, EX1 3PB, UK
- <sup>11</sup> <sup>3</sup>Technische Universität Dresden, Institute of Photogrammetry and Remote Sensing,
- 12 Helmholtzstr. 10, 01069, Dresden, Germany
- <sup>13</sup> <sup>4</sup>Max Planck Institute for Meteorology, Hamburg, Germany.
- <sup>14</sup> <sup>5</sup>Centre National de Recherches Météorologiques, Université de Toulouse/Météo-
- 15 France/CNRS, Toulouse, France
- <sup>16</sup> <sup>6</sup>NORCE Norwegian Research Centre, Bjerknes Centre for Climate Research, Bergen,
- 17 Norway
- <sup>18</sup> <sup>7</sup>Center for the Environment, Faculty of Arts & Sciences, Harvard University, Cambridge,
- 19 MA, USA
- <sup>20</sup> <sup>8</sup>Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC), Bologna, Italy
- <sup>21</sup> <sup>9</sup>Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ CE Orme
- 22 des Merisiers, Gif sur Yvette, France
- <sup>23</sup> <sup>10</sup>Department of Physical Geography and Ecosystem Science, Lund University, Lund,
- 24 Sweden
- <sup>25</sup> <sup>11</sup>Departamento de Ciências e Engenharia do Ambiente, DCEA, Faculdade de Ciências e
- <sup>26</sup> Tecnologia, FCT, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal
- 27 Corresponding Author: Sujan Koirala (<u>skoirala@bgc-jena.mpg.de</u>)

Terrestrial ecosystems respond to changes in environmental conditions, mainly via key 28 climatic controls of precipitation and temperature on vegetation activities and 29 decomposition processes <sup>1</sup>. Yet, the relationship between climate and the overall 30 spatiotemporal dynamics and uncertainties of the global carbon cycle, i.e., gross 31 primary productivity (GPP), effective ecosystem carbon turnover times  $(\tau)$ , and 32 consequently the total ecosystem carbon stock ( $C_{total}$ ), are unclear <sup>2–5</sup>. Using a global 33 observation-based synthesis, we first show that the apparent partial spatial climate 34 sensitivities of GPP and  $\tau$  are associated with relative availability of precipitation and 35 temperature, and are therefore modulated by aridity. The apparent sensitivity of GPP 36 to temperature increases from arid to humid climatic regions. In contrast, its sensitivity 37 to precipitation is invariant throughout different climatic regions. Simultaneously, the 38  $\tau$ -precipitation response is strongly non-linear resulting in ~2 times longer  $\tau$  in arid 39 regions compared to humid regions for a given temperature. Compared with these 40 observed patterns, the offline carbon cycle simulations of seven European Earth System 41 Models (ESMs), that participated in CMIP6, perform relatively better for climate 42 sensitivities of GPP than those of  $\tau$ . This leads to a large spread and bias in  $C_{total}$  in both 43 warm and cold semi-arid and arid regions where only a few models capture the 44 observed  $\tau$ -precipitation relationship. The emergence of the hydrological controls, 45 modulated by aridity, on global carbon cycle implies that the changes in precipitation 46 may moderate the temperature-driven climate feedback of the global carbon cycle 47 under climate change. 48

# 49 Introduction

50 The exchange of carbon between the biosphere and atmosphere represents a key mechanism controlling the effect of global changes on the carbon cycle <sup>6</sup>, as well as on the water and 51 energy cycles <sup>7</sup>. The land ecosystems, at the centre of biosphere-atmosphere interaction, store 52 carbon by integrating the differences in carbon captured through gross primary productivity 53 (GPP), and carbon released through decomposition processes. Due to the complex 54 dependencies of the photosynthesis and decomposition on climate, biomes, and carbon use 55 efficiency of vegetations<sup>8</sup>, the resulting carbon stocks vary significantly across space and 56 time. The differences in modelling these complex climate-biosphere dependencies, therefore, 57 lead to large differences in Earth system model (ESM) predictions of carbon stocks <sup>9,10</sup> that 58 are prevalent under the current climate and exacerbate under future climate change scenarios. 59

The ecosystem turnover and associated carbon losses and stocks, thus, represent major
sources of uncertainties regarding carbon cycle – climate feedbacks <sup>11,12</sup> and the projections
of future carbon budgets <sup>5,13,14</sup>, and modelling and constraining them with observations still
pose a significant challenge to ESMs <sup>2-4</sup>.

64 Various studies have, therefore, investigated the relationship between land carbon stocks and their climate drivers. For example, the uncertainties in soil organic carbon and their responses 65 to climate change have been extensively studied <sup>15</sup>, but mostly focused on temperature, as the 66 first principles dictate the key role of temperature in determining the current and future 67 carbon cycle dynamics <sup>16</sup>. The studies focusing only on energy-limited humid and cold 68 regions, understandably, have shown strong associations of spatial variations of carbon stocks 69 and temperature <sup>17</sup>. When other climatic regions globally are considered, though, temperature 70 alone only explains ~10% of the total spatial variation of the observation-based soil organic 71 carbon <sup>10</sup> suggesting a larger influence of additional environmental factors. In particular, 72 precipitation, which is the primary source of moisture for all land processes, is potentially a 73 strong driver of spatial variation of turnover times, the  $\tau^{18}$ , as demonstrated by the 74 significance of moisture effects on decomposition <sup>19</sup> and the observation-based linkages 75 between water and carbon cycles <sup>20</sup>. The moisture availability is especially critical in arid to 76 semi-arid regions, which cover ~39% of the global vegetated land, and substantially influence 77 the interannual variabilities of global biosphere-atmosphere carbon exchange <sup>21–23</sup>, but the 78 effects on modelled  $\tau$  has been largely ignored <sup>24</sup>. 79

- 80 In an ecosystem under a steady state, the temporal variations in total ecosystem carbon
- storage (hereafter  $C_{total}$ ) become negligible as the net input (*GPP*) and output (total ecosystem
- 82 respiration) carbon fluxes balance each other. The spatial variations of  $C_{total}$  across
- ecosystems are, then, essentially defined by the differences in *GPP* and the carbon loss
- determined by decomposition rates and  $\tau$ . The climate, though, has differential influences on
- and association with *GPP* and  $\tau$ , which result in a unique spatial heterogeneity in *C*<sub>total</sub>
- so compared to those of either GPP or  $\tau$ . It is, therefore, critical that the sensitivities of GPP,  $\tau$ ,
- and  $C_{total}$  to climatic variations are simultaneously evaluated, especially in ESMs where the
- <sup>88</sup> uncertainties in  $C_{total}$  may be related to not only  $GPP^2$  but to both  $\tau$  and  $GPP^9$ .
- <sup>89</sup> Here, based on observation-based estimates of the carbon cycle, we first evaluate the
- apparent spatial sensitivities of climatological mean *GPP* and  $\tau$  to two primary climatic
- 91 drivers: mean annual air temperature (*MAT*) and mean annual precipitation (*MAP*). To infer
- 92 effective or apparent climatological  $\tau$  of an ecosystem, we assume that the ecosystem carbon

reaches a steady-state over multiple decades, and calculate  $\tau$  as the ratio of long-term 93 averages of total ecosystem carbon stock and GPP <sup>17,18,25,26</sup>. The apparent spatial sensitivities 94 of the climatological carbon cycle are assessed in a climate phase-space of temperature and 95 precipitation for different climatic regions (arid to humid) that are delineated using an aridity 96 index, defined as the ratio of mean annual precipitation and potential evapotranspiration (see 97 Methods and Figure A1). While aridity itself is not directly mechanistically connected to 98 physiological and decomposition processes within an ecosystem, it is a rather clear indicator 99 of relative availability of limiting moisture and energy resources that define the soil moisture 100 regime (Figure A2), and consequently, the vegetation responses and patterns 27-29. The 101 climate sensitivities of  $\tau$  and *GPP* are represented using non-linear relationships with 102 temperature and precipitation (see Methods). The relationships are then used as the basis for evaluation of the offline carbon cycle model simulations of seven European ESMs from 104 CRESCENDO project (see Methods), which participated in the Sixth Phase of Coupled 105 Model Intercomparison Project (CMIP6). 106

# 107 Observation-based climate sensitivities

Both  $\tau$  and GPP exhibit unique sensitivities to the spatial variabilities of precipitation and 108 temperature in all climatic regions despite large heterogeneity within a climatic region 109 (Figure 1, Table A4). First,  $\tau$  varies strongly across both temperature and precipitation gradients in all climatic regions. The sensitivity of  $\tau$  to temperature is well-documented for 111 humid regions <sup>17</sup>. Interestingly, there is a significant offset in  $\tau$  as the climate gets drier with 112 the longest  $\tau$  in arid regions ( $\tau$ =178 years at 0°C) compared to humid regions (76 years). In 113 fact, for any given temperature,  $\tau$  is longer in the arid regions compared to humid regions, 114 115 suggesting an influence of, or covariation with, precipitation across climatic regions. The sensitivities to temperature, i.e., the slope of the fitted lines in Figure 1a, are all within ~17% 116 across different climatic regions (Table A3). It should be noted that a significant number of 117 grid cells, with a mean temperature less than 0°C, have longer  $\tau$  than under warmer 118 temperatures within the same climatic region. Yet, the non-linear increase in temperature 119 sensitivity with decreasing temperature is not as strong as reported previously for  $\tau$  of soil 120 organic carbon <sup>17</sup> and the linear coefficients dominate the non-linear coefficients in different 121 climatic regions (inset of Figure 1, and Table A3). While several factors may lead to such 122 differences (e.g., differences in observed data, methodology, etc.), one key reason may also 123 be the selection of climatic regions. For instance, in regions with temperature >10°C, 124

ignoring arid to semi-arid regions with a systematically longer  $\tau$  than humid region would skew the overall distribution of  $\tau$ , which would then lead to a hyper-logarithmically nonlinear  $\tau$ - temperature relationship.

The apparent sensitivity of  $\tau$  to precipitation is large when the precipitation is lower than 128 129 ~1000 mm/year across all climatic regions (Figure 1b). Only within 0-1000 mm/year range of precipitation,  $\tau$  varies by more than an order of magnitude. Contrasting to temperature, the 130  $\tau$ -precipitation relationship saturates at higher precipitation (>1000 mm/year) in all regions 131 except arid regions where precipitation never crosses the threshold. The overall apparent 132 sensitivity of  $\tau$  to precipitation is, therefore, hyper-logarithmically nonlinear, as indicated by a 133 clear positive curvature. While the functional relationship of precipitation is almost the same 134 135 for all but energy-limited humid regions, the sensitivity of  $\tau$  to precipitation is smaller for humid regions compared to arid regions, indicating a saturation of the precipitation effect on 136  $\tau$ . The difference of  $\tau$  across different climatic regions is the smallest in the lowest 137 precipitation ranges. Similar to the  $\tau$ -temperature relationship, though, there is still a large 138 variability of  $\tau$  for given precipitation within all climatic regions (as indicated by low  $r^2$  and 139

140 higher  $r_{mad}$  of fitted relationships in Figure 1 and Table A4).

GPP also has a clear relationship with temperature, the strength of which is dependent on the 141 climatic region (Figure 1c). The temperature sensitivity is the largest in humid and the 142 smallest in arid climatic region. In the humid region, GPP sensitivity to temperature increases 143 with temperature. The GPP increases steadily for increasing precipitation with very small 144 differences across different climatic regions (Figure 1d). The increase in GPP with 145 precipitation is quasi-linear in all climates, as GPP saturates at the highest precipitation 146 values, especially in the humid regions. This, again, suggests that precipitation plays an 147 important role in determining GPP when the vegetation activities are likely to be limited by 148 moisture. While a strong relationship of *GPP* with precipitation has been shown previously 149  $^{30-32}$ , the minimal difference of *GPP* across different climatic regions for given precipitation 150 shows that the spatial variation of GPP is mainly related to precipitation alone. 151



152



# <sup>162</sup> Climate sensitivities in Carbon Cycle models

- In the CRESCENDDO model simulations, the apparent sensitivities of  $\tau$  to climate have
- 164 larger biases than those of GPP to climate, when compared against respective observation-
- based sensitivities (Figure 2). First, the comparison of the  $\tau$ -temperature relationship reveals
- that all models, except CLM4.5 and SURFEX, have a shorter  $\tau$  than observation across all
- 167 climatic regions while the linear correlation between the model and observation-based

relationships are generally high (first column, **Figure 2**). This is accompanied by a large bias in  $\tau$  with most models, except CLM4.5, showing a clear underestimation. The differences between models and observation and among models also remain fairly similar across all climates. For example, the model with the longest  $\tau$  in the arid region also has the longest  $\tau$  in other climatic regions. This suggests that the temperature dependence of  $\tau$  in a given model remains fairly consistent across different climatic regions. Of the models, only CLM5 shows a stronger sensitivity of  $\tau$  to temperature with a hyper-logarithmically nonlinear increase in  $\tau$ when the temperature decreases below 0°C (Figure 2e).

176 Similar to observation, the models also exhibit a tendency of non-linear increase in  $\tau$  for

decreasing precipitation, with the largest sensitivity in arid regions and the smallest

sensitivity in humid regions (second column of Figure 2). Most models, though, show a

smaller sensitivity to precipitation than the observation with a consistent underestimation bias

across all climate, with the largest difference in the arid regions. In fact, only CLM4.5 and

181 SURFEX show a significant increase of  $\tau$  in the lower precipitation range that is akin to the

*182* observation.

183 For *GPP*, the models show increased sensitivity to temperature from arid to humid climatic

regions (third column of Figure 2) except for CLM4.5 (Figure 2c), which shows a smaller

increase in GPP at a higher temperature in all but the humid region, and LPJ-GUESS (Figure

2s) with much larger *GPP* in arid regions. In general, all models represent the sensitivity of

*GPP* to temperature fairly well with a relatively smaller biases than the same of  $\tau$ . The

sensitivity of *GPP* to precipitation also shows a small spread among the models (fourth

column of Figure 2). Only in the humid regions, a large number of models, e.g., CLM4.5 and

JSBACH, over-predict *GPP* in the highest precipitation range, where the observation-based

191 GPP shows a saturating tendency.



Figure 2. Evaluation of sensitivities of climatological carbon cycle dynamics to climate 193 in CRESCENDO models. In the first and second columns, the relationships between 194 ecosystem turnover time of carbon ( $\tau$ , vears) versus mean annual temperature (MAT) and 195 precipitation (MAP) are presented, followed by those for gross primary productivity (GPP, 196 kgC/m<sup>2</sup>/year) in the third and fourth columns. Along the row, the comparisons for different 197 models against the observation are presented. The different colours indicate different climatic 198 regions defined by the aridity index. The dotted lines indicate the observation-based 199 relationship, and the solid lines indicate relationship from the model simulations. The inset 200 text shows coefficient of determination  $(r^2)$  and mean relative bias between the fitted 201 202 relationships from the models and observation. Note the logarithmic vertical axis for  $\tau$  in the first and second columns. 203

#### Partial correlations with climate 204

To elucidate the regional variations of  $\tau$ , GPP and the resulting  $C_{total}$ , and eliminate 205 collinearity between precipitation and climate, here, we evaluate the zonal variations and 206 their partial associations with precipitation and temperature. 207

In general, the large-scale zonal distribution of  $\tau$  follows temperature, and that of *GPP* 208

follows precipitation (first column of Figure 3, and Figure A4). The  $\tau$  is the shortest in 209

tropical regions with the highest temperature and the longest in high latitudes with the lowest 210

temperature (Figure 3a). Similarly, GPP is the largest in the wettest tropical regions with the 211

highest precipitation (Figure 3d). The zonal distribution of  $C_{total}$  reveals a smaller variation 212

across the latitude compared to  $\tau$  and GPP (Figure 3g). Nonetheless, the largest  $C_{total}$  occur in 213

either the humid tropics with the largest *GPP* or in the high latitudes with the longest  $\tau$ . In 214

general, the CRESCENDO models produce zonal variations similar to observation-based 215

estimates, with the smallest bias for GPP, and a consistent underestimation bias of  $\tau$  and  $C_{total}$ 216

across the latitude. 217

219

222

The observation-based  $\tau$  has a stronger negative correlation with temperature in the energy-218

moisture-limited regions (Figure 3b). In general, the latitudinal variation of  $\tau$ -temperature 220

limited regions such as the tropics and high latitudes of both hemispheres compared to the

221 correlation is much stronger in the models than in the observation. This is especially true for sub-tropical mid-latitudes (40°S to 40°N) except the humid tropical regions. In the same

regions, the local variation of  $\tau$  has a stronger correlation with precipitation than with 223

temperature (Figure 3c). The  $\tau$ -precipitation relationships in the models have a larger spread, 224

and most models have a weaker  $\tau$ -precipitation correlation than in the observation. This is 225

especially clear in the subtropical southern hemisphere and temperate northern hemisphere. 226

This weaker  $\tau$ -precipitation correlation is concurrent with a stronger  $\tau$ -temperature correlation 227

in the models suggesting a dominant temperature control on carbon cycle sensitivity to 228

climate. In these regions, only SURFEX shows a strong local-scale  $\tau$ -precipitation correlation 229

that is consistently stronger than the observation, but it also exhibits a stronger bias in the 230

high latitudes. 231

Further, both the observations and models show a much stronger correlation of GPP with 232

233 precipitation (Figure 3f) than with temperature (Figure 3e) across most of the latitudes. Only

in the northern high latitudes over ~50°N (with MAT  $\leq$  0°C), GPP has a stronger correlation 234

with temperature than with precipitation. In fact, in these regions, the low winter temperature 235

- affects the MAT significantly, and correcting for winter reveals a clearer role of temperature
- in the high latitudes (see Figure A6). The model spread is also relatively smaller for the
- 238 correlation of *GPP* with climate than of  $\tau$  with climate. This shows that the models, in
- 239 general, reproduce both the global and local responses of *GPP* to temperature and
- 240 precipitation much better than those of  $\tau$ .
- Further, as expected, the observation-based relationship of  $C_{total}$  and climate is a complex
- amalgam of those of *GPP* and  $\tau$ . For example, *C*<sub>total</sub>-temperature correlation is relatively
- small in magnitude (Figure 3h) due to the opposing  $\tau$ -temperature (negative) and GPP-
- temperature (positive) correlations. The C<sub>total</sub>-precipitation correlation follows the same zonal
- variation as that of *GPP* but with a lower magnitude (**Figure 3**i) due to predominantly
- $^{246}$  negative  $\tau$ -precipitation correlation. The CRESCENDO models generally exhibit a stronger
- 247 positive correlation of  $C_{total}$  with climate than in the observation. This strong  $C_{total}$ -
- 248 precipitation correlation in the models is due to a strong correlation between  $C_{total}$  and GPP
- 249 (see Figure A6).



Figure 3. Zonal distribution of ecosystem turnover time of carbon ( $\tau$ , years), gross 251 primary productivity (*GPP*, kgC/m<sup>2</sup>/year), and total ecosystem carbon stock ( $C_{total}$ , 252  $kgC/m^2$ ) and their relationships with climate. The zonal means are plotted in the left 253 254 column, and the correlations with temperature and precipitation are plotted in the centre and right column, respectively. The correlation with precipitation is controlled for temperature 255 and vice-versa. The Pearson's correlation coefficient for each latitude is calculated for a 256 moving window of 10 grid cells along the latitude (5°). Only the coefficients that are 257 statistically significant at 5% significance level are shown. The individual models are plotted 258 in coloured thin lines, and the multimodel ensemble in thick blue lines. In the correlation 259 plots, the thick dashed blue lines show the normalized mean correlation of all models with 260 shades indicating variation within one standard deviation. The observation is plotted as a 261 thick black line with shade indicating the range within the 5<sup>th</sup> and 95<sup>th</sup> percentiles. 262

# <sup>264</sup> Model biases and agreement



Figure 4. Global distributions of ecosystem turnover time of carbon ( $\tau$ , years), gross 266 primary productivity (GPP, kgC/m<sup>2</sup>/year), and total ecosystem carbon stock (C<sub>total</sub>, 267  $kgC/m^2$ ) and their biases. The multimodel ensemble, observation-based estimate, and the 268 bias and agreement are presented in the first, second, and third row, respectively. Multimodel 269 ensemble is calculated as the median of the seven CRESCENDO models. The bias is 270 271 calculated as the ratio between multimodel ensemble and the corresponding observation. In the global maps of bias (third row), stippling indicates the regions where only two or fewer 272 models fall within the range of observational uncertainties (5<sup>th</sup> and 95<sup>th</sup> percentiles). 273

- 274 Spatially, the CRESCENDO multi-model ensembles exhibit similar global distributions
- 275 compared to the corresponding observation-based estimates of  $\tau$  (Figure 4a, b), *GPP* (Figure
- 4d, e), and  $C_{total}$  (Figure 4g, h). Globally, the observation-based  $\tau$  is the shortest in tropical
- regions with high GPP as well as a relatively high  $C_{total}$ , and the longest in high latitude
- regions with a relatively lower *GPP* but high  $C_{total}$ . The longer  $\tau$  is also prevalent in cold and
- dry arid regions, such as the surroundings of the Gobi Deserts, where both *GPP* and *C*<sub>total</sub> are
- 280 lower than other regions. The observation-based Ctotal shows a pattern of large carbon
- storages in either a region with high *GPP* or a longer  $\tau$ . Additionally, larger *C*<sub>total</sub> are also
- evident in the Arctic North America and Boreal Eurasia characterized by occurrences of peat.

On a broad scale, the CRESCENDO multimodel ensemble  $\tau$  also exhibits a similar spatial 283 gradient with shorter  $\tau$  in the tropics and longer  $\tau$  in the northern high latitudes. Such spatial 284 gradients are reproduced by individual models as well (Figure A8). Despite the similarity in 285 spatial variability, the multimodel ensemble  $\tau$  still has a substantial underestimation bias in 286 the semiarid regions (sub-Saharan Sahel, central Asia, northern Australia, and western United 287 States), where the multimodel ensemble is  $\sim 5$  times shorter than the observation (Figure 4c). 288 The models tend to agree with each other more in the northern high latitudes and the humid 289 tropics than they do in the arid and semi-arid regions with large biases, as also shown 290 previously for CMIP5 models (Carvalhais et al., 2014). Moreover, fewer than two models are 291 within the observational uncertainty (hereafter referred to as low agreement) in ~22% of the 292 total grid cells. In the semi-arid region, the underestimation of  $\tau$  (the overestimation of 293 turnover rate) is prevalent in all but three models (CLM4.5, JULES and SURFEX in Figure 294 A7). 295

Both the spatial variability and magnitude of GPP in CRESCENDO models compare better 296 with the corresponding observation than in the case of  $\tau$  (see Figure A9 as well). In the 297 tropical humid and semi-arid regions, the GPP bias is small in relative terms (Figure 4f). In 298 semi-arid regions of south America and Australia, the relative bias in GPP is high due to 299 relatively lower GPP values in the observation. But the largest overestimation bias can be 300 seen in the northern high latitudes which have low GPP. The models also agree less in the 301 regions with the largest bias with almost ~35% of global grid cells having a low model 302 agreement. Note that such low agreement is also due to small uncertainty range in the 303 ensemble of GPP observations (bar charts in Figure A7b). When the global and regional 304 GPP are compared, almost all models are within the observational uncertainty globally due to 305 306 better performance in the humid and sub-humid climate compared to other regions (Figure A7b). 307

Lastly, the spatial variability of Ctotal in most CRESCENDO models is similar to that from 308 observation-based estimate (Figure A10), even though the distinctly large storage in 309 peatlands is not reproduced by most models. The bias in Ctotal bears a striking similarity to the 310 311 spatial pattern of bias of  $\tau$ . There are, however, also regions where the opposing biases in  $\tau$ and GPP lead to unbiased Ctotal. For example, in the Iberian Peninsula and Hudson Bay, the 312 overestimation of GPP is compensated by an underestimation of  $\tau$  (overestimation of 313 turnover rate) resulting in a relatively unbiased  $C_{total}$ . Due to such compounding effects, there 314 is even more widespread occurrence of low model agreement for Ctotal (~40% of global grid 315

- $_{316}$  cells), and biases, mostly similar with those of  $\tau$ , are widespread across different climatic
- 317 regions (Figure A7c).
- To attribute the bias in  $C_{total}$ , we compare the biases in *GPP*,  $C_{total}$ , and  $\tau$  in each
- 319 CRESCENDO model (Figure 5). Note that an ideal model would have low biases in all GPP,
- 320  $\tau$  and  $C_{total}$ . The contrast between the biases in GPP and  $C_{total}$  reveals a dominant control of  $\tau$
- on the *C<sub>total</sub>* bias. The JSBACH, LPJ-GUESS, and ORCHIDEE models show consistently
- $_{322}$  lower  $C_{total}$  than observation-based estimate, that is likely associated with an underestimation
- of  $\tau$ , as the biases in *GPP* are relatively lower than those in  $C_{total}$  (larger density of points
- along the axis of no bias in GPP). Other models perform better for both *GPP* and  $\tau$  in all but
- arid climatic region. Only CLM4.5 and SURFEX show a good agreement of  $\tau$  in arid region
- (red lines close to 1:1 line). CLM4.5, though, also has a slight overestimation bias in  $C_{total}$  in
- all other climatic regions arising from overestimation of  $\tau$ . In fact, in CLM5, the biases in  $\tau$
- are reduced across all climatic regions. Only JULES and SURFEX produce a larger density
- of grid cells where biases are low for GPP,  $\tau$ , and  $C_{total}$ . In JULES,  $\tau$  and  $C_{total}$  are
- underestimated mainly in the arid and humid region. Globally, only CLM4.5 has larger range
- of spatial variability than in the observation (the slope of the fitted global line greater than 1).



332

# **Figure 5. Relationship of biases in gross primary productivity (***GPP***) and total**

ecosystem carbon stock (Ctotal) in CRESCENDO model simulations. The different colours 334 indicate different climate defined by the aridity classes, and individual grid-cells are plotted 335 as dots. The percentage values in the parenthesis of the title shows the fraction of the global 336 grid cells that are within the range of bias between 0.1 and 4. The dashed vertical line shows 337 the axis of no bias in GPP (ratio of GPP from model to that from observation = 1), the same 338 for  $C_{total}$  is shown by a horizontal dashed line. The grey solid line indicates the axis along 339 which there is no bias in  $\tau$ . An ideal point of no biases in GPP,  $C_{total}$ , and  $\tau$  is indicated by a 340 cross 'x'. The coloured lines indicate the main axis of variations of biases in GPP and Ctotal 341 (linearly fitted lines that passes through origin). The white contour lines indicate the density 342

of grid cells, with higher density of lines showing larger occurrence of grid cells.

# 344 Conclusions and discussion

Adding to a well-established association of carbon cycle with temperature, we show that 345 precipitation also plays a key role in defining primary climate controls on spatial variabilities 346 of GPP and carbon turnover times. The moisture control results in an emergent association of 347 turnover time with aridity (Figure 1) that encompasses a wide range of precipitation and 348 temperature regimes. Aridity regulates the moisture availability and, thus, modulates the 349 apparent climate sensitivities of climatological carbon cycle dynamics. Despite uncertainties 350 in the observation-based data, the presented observation-based climate sensitivities of carbon 351 cycle are robust across full factorial of the latest estimates of GPP, soil and vegetation carbon 352 stocks (Figure A11). We note that the apparent sensitivities presented here only reveals the 353 primary controls, which is also indicated by a large spread within a climatic region (Figure 354 1). The additional heterogeneities may be associated with differences in soil and land 355 properties, along with differences in secondary moisture effects such as photodegradation and 356 pulsing of microbial activity from rainfall, drought-related mortality <sup>33</sup>, disturbances and fire 357 regimes <sup>34,35</sup>, etc. 358

We found that the offline simulations of the latest generation of Earth System Models from 359 CRESCENDO project reproduce well the association of GPP with climate, but fall short for 360 turnover times, especially in semiarid and arid regions leading to large uncertainties in 361 simulated carbon stocks. The models with relatively lower biases in turnover time and carbon 362 stock in arid and semiarid regions produce a better zonal variation of  $\tau$ -precipitation 363 relationship, while the poorer model performance is characterized by biases in  $\tau$ -temperature 364 relationship. This implies an insufficient model representation of moisture control on 365 climatological carbon cycle dynamics even though all models include the moisture effects on 366 both GPP and decomposition rates (Table A2). 367

The identification of exact mechanisms of moisture control on carbon cycle, especially 368 without an observation-based soil moisture that extends beyond top few centimetres of soil, is 369 still challenging. For example, we cannot clarify if the modelled soil moisture is itself biased 370 or if the response of carbon cycle to soil moisture is misrepresented. Nevertheless, an 371 evaluation of evapotranspiration shows a very consistent performance across models in all 372 climatic regions (Figure A12), which should lead to a consistent soil moisture across models. 373 Yet, the spatial covariation of carbon fluxes-moisture-temperature reveals diverse moisture-374 temperature relationships across different models (Figure A13). The temperature responses 375

of *GPP* and respiration are highly consistent, but the moisture responses differ significantly

- across different models. The model performing better in arid and semi-arid regions (e.g.,
- 378 SURFEX) has contrasting responses than the models with poorer performance providing
- 379 secondary evidence of the potential weakness in the model representation of moisture
- 380 controls. Factorial experiments of each model are necessary to identify the reasons behind,
- and to bridge the differences between, the apparent moisture controls presented in this study.
- Lastly, the weaker performance of precipitation-turnover relationships leads to large
- <sup>383</sup> uncertainties of carbon stocks especially in arid and semi-arid climatic regions, as the biases
- in  $C_{total}$  are largely associated with biases in turnover time rather than those in *GPP*.
- Additionally, we found that biases of *GPP* and turnover time manifest to a larger bias in
- $C_{total}$ . This has serious implications in the predictions of changes in carbon stock under
- <sup>387</sup> climate change, as the changes in precipitation, temperature, and thus aridity <sup>36</sup> will have a
- <sup>388</sup> non-negligible influence. The reduction of uncertainties in modelled carbon stock under
- 389 global changes is, therefore, not a carbon-cycle only challenge and it must be addressed with
- a broader scope of improving the hydrological influences on the terrestrial carbon cycle.

#### 392 Methods

#### 393 Ecosystem carbon turnover time:

A turnover rate of a storage is broadly defined as the rate of fractional loss of storage per unit 394 time. For an ecosystem under natural condition, the turnover time, an inverse of turnover rate, 395 is, therefore, the time duration carbon stays in the ecosystem between carbon assimilation 396 through photosynthesis and loss to atmosphere through respiration and decomposition 397 processes. Under steady state assumption, the output carbon fluxes equilibrate with the input 398 primary productivity, and temporal changes in carbon storage become small. Thus, the 399 steady-state turnover time of an ecosystem can be effectively calculated as the ratio of carbon 400 storage and the input flux to the ecosystem as, 401

402

$$\tau = \frac{cTotal}{GPP} \tag{1},$$

where,  $\tau$  is the ecosystem turnover time (years),  $C_{total}$  is the long-term average total ecosystem carbon storage per unit area (kgC/m<sup>2</sup>) and *GPP* is the annual mean gross primary productivity (kgC/m<sup>2</sup>/year). As stated and used in previous studies <sup>17,18,25,26</sup>, the  $\tau$  in Eqn (1) is under the assumption of a steady state ecosystem and represents the apparent or effective turnover time of the whole ecosystem carbon storage. It emerges as the diagnostic property of an ecosystem, rather than the intrinsic property of decomposition processes that explicitly controls  $C_{total}$ .

#### 410 **Observation-based datasets:**

- Due to potential uncertainties in the global observation-based estimates, we used an ensemble of six GPP datasets, four soil carbon ( $C_{soil}$ ) and four vegetation carbon ( $C_{veg}$ ) stock datasets. The data ensembles, spanning a wide range of sources and methods, represent the current state-of-the-art of global observation-based estimates of the carbon cycle components and, likely, cover the full range of observational uncertainties. As we use only a single ensemble estimate from each product, we assume that the differences and uncertainties are larger across the datasets than within a single dataset.
- The *GPP* ensemble includes Model Tree Ensembles (MTE) (Jung et al., 2011) and its recent
  successors from FLUXCOM based on remote sensing and that with additional meteorological
  forcing <sup>38</sup>, solar-induced fluorescence (SIF) based GOSIF *GPP* <sup>39</sup>, light use efficiency based *GPP* from vegetation photosynthesis model (VPM) <sup>40</sup>, and lastly, an independent machine
  learning prediction from FluxSat <sup>41</sup> that uses corrected satellite-based reflectance. All *GPP*
- 423 products are either upscaled from or extensively evaluated against eddy covariance

observations of carbon fluxes from FLUXNET (www.fluxdata.org) sites (Baldocchi et al., 424 2001). For example, the MTE, FLUXCOM, and FluxSat products use different combinations 425 and corrections of training data and drivers from satellite remote sensing and meteorological 426 dataset to upscale site-level observation to the global scale using different machine learning 427 algorithms. VPM, a light use efficiency model, is similar to MODIS GPP but with an 428 additional validation for FLUXNET sites, and GOSIF GPP uses a set of GPP-SIF 429 relationships to produce an ensemble of high resolution global GPP fields that are 430 independent of reflectance-based remote sensing products. While the spatial variability of the 431 mean GPP is one of the most robust features when comparing different observation-based 432 GPP products, it is desirable to use an ensemble as no single data is superior to all others 433

434 universally <sup>43</sup>.

For C<sub>soil</sub>, we use the full depth soil organic carbon estimates including a data that combines 435 inventories from Harmonized World Soil Database <sup>44</sup> with the Northern Circumpolar Soil 436 Carbon Database <sup>45</sup> extrapolated to full depth <sup>18</sup>; from SoilGrids that uses machine learning 437 method to upscale soil profile measurements with land and climate characteristics as 438 predictors<sup>46</sup>; and from Sanderman et al. that corrects for land use and forest cover <sup>47</sup>. Due to 439 differences in number and locations of observed soil profiles used, predictors, and prediction 440 methods, the soil carbon estimates are uncertain, especially in high latitude regions, with a 441 large C<sub>soil</sub> but a relatively fewer measurement profile. Nonetheless, the global distributions of 442  $C_{soil}$  in these products are robust <sup>48</sup>, and their differences reflect potential uncertainties across 443 all observation-based global estimates. We note that the SoilGrids data has been recently 444 updated, but the new version of the data, the LANDGIS, has not been used in this analysis 445 owing to lack of extensive validation and potential over-estimation issue in the northern high 446 latitudes <sup>48,49</sup>. It should here be noted that using the LANDGIS instead of SoilGrids does not 447 significantly affect the main findings of this study (see Figure 1 and Figure A3). 448

449 The Cveg ensemble includes four different RADAR and LIDAR satellite-based estimates of

aboveground biomass with a corresponding estimate for belowground biomass <sup>48</sup>. As

451 different data are based on different satellites and algorithms, and have been validated with

452 observations whenever and wherever available, they are indicative of global vegetation

453 biomass and associated uncertainties in observation-based estimates.

- 454 An ensemble of observation-based  $\tau$  was then obtained using Eqn (1) for a full factorial
- 455 combination of all  $C_{soil}$ ,  $C_{veg}$ , and *GPP* products resulting in 96 different  $\tau$  maps.

- In this study, for *GPP*,  $C_{total}$  and  $\tau$ , the ensemble median across data products was used as the representative observation-based estimate, and the 5<sup>th</sup> and 95<sup>th</sup> percentiles were used as uncertainty range.
- 459 Lastly, for an additional evaluation of CRESCENDO models, the ensemble estimates of
- 460 evapotranspiration from FLUXCOM remote sensing products <sup>50</sup> was used.

#### 461 Climatic regions based on aridity index:

- 462 The aridity index, calculated as the ratio of mean precipitation and potential
- 463 evapotranspiration (PET), is used to delineate different climatic regions based on relative
- <sup>464</sup> availability of moisture and energy. For consistency, the CRU-NCEP precipitation, used to
- <sup>465</sup> force the CRESCENDO model simulations, is also used for calculating the aridity index. The
- $_{466}$  PET data is based on an estimate from the latest climate datasets at a high resolution  $^{51}$ .
- <sup>467</sup> Different climatic regions were then delineated based on the aridity index ranges suggested
- by UNEP <sup>52</sup> as: arid (<0.2), semi-arid, sub-humid, and humid (see **Figure A1**). The average
- annual precipitation (temperature) varies from 245 mm/year (7.6°C) in arid regions to 880.3
- 470 mm/year (0.9°C) in humid regions. These regions constitute 8.2%, 24.3%, 15.1%, and 52.4%
- 471 of the total area considered in this study. In relative terms, the arid and semiarid regions
- 472 contribute ~18% of the global *GPP* and hold ~24% of the global carbon stock (**Table A1**).
- 473 Combined, these regions have a  $\tau$  of ~43 years, which is ~34% longer than the global  $\tau$  of ~32
- 474 years.

#### 475 Climate sensitivities:

The overarching aim of this study is to investigate the climate sensitivities of climatological

- 477 mean carbon cycle dynamics across all major climatic regions globally and assess the
- 478 apparent sensitivities of spatial covariations of carbon cycle with climate. Under humid
- 479 climate, spatial sensitivity of  $\tau$  is itself dependent on temperature <sup>17</sup>. But, to what extent the
- same applies to other climatic regions and carbon cycle variables is unclear. This is especially
- 481 critical with regard to precipitation that represents the potential moisture supply for an
- ecosystem that determines both the carbon assimilation and decomposition processes.
- Further, we hypothesize that the relative availability of moisture (supplied by precipitation)
- and energy (determined by temperature) plays a central role in defining the carbon cycle
- 485 dynamics across spatiotemporal scales. We, therefore, investigate the relationships between
- the carbon turnover time and *GPP* with precipitation and temperature across different

climatic regions characterized by an aridity index. We assume that the relationship between  $\tau/GPP$  and climate are non-linear and assume it to be of a second-degree polynomial form as,

$$y = ax^2 + bx + c \tag{2},$$

490 where, y is the response/dependent variable, x is the independent variable, and a, b and c are

491 the parameters of the quadratic equation. The terms a and b are the coefficients of the non-

492 linear and linear terms, respectively, and c is the constant term that defines the

- 493 offset/intercept when the independent variable is zero. As  $\tau$  is expected to have a non-linear
- exponential relationship with temperature, we log-normalize it and use  $log(\tau)$  as the
- 495 dependent variable.
- <sup>496</sup> The parameters of Eqn (2) for each climatic region are estimated using a robust non-linear
- <sup>497</sup> least square regression with Huber loss function <sup>53</sup> that is less sensitive to outliers in the data.

#### 498 **CRESCENDO model simulations:**

- 499 We use the observation-based climate sensitivities to evaluate the process representations in
- <sup>500</sup> land surface schemes of (used in) the ESMs from seven different European research
- 501 institutes. The stand-alone offline land simulations were carried out as a part of the
- 502 CRESCENDO project (<u>https://ukesm.ac.uk/crescendo/</u>) to understand and evaluate the carbon
- and nitrogen cycle processes in the current state-of-the-art ESMs. The same land surface
- schemes were used in the ESM simulations for the sixth phase of coupled model
- <sup>505</sup> intercomparison project (CMIP6). Following the TRENDY protocol <sup>54</sup>, the models were
- 506 forced by CRUNCEP v7 <sup>55</sup> forcing dataset

- <sup>508</sup> <u>ml</u>), observed atmospheric CO<sub>2</sub> concentration, nitrogen deposition, and annual land-use
- 509 changes. The offline simulations forced with an observation-based dataset are suitable for
- <sup>510</sup> evaluating the response of the land carbon cycle processes to the climate forcing, as they are
- 511 less prone to biases and uncertainties arising from differences in carbon-climate feedbacks in
- 512 the coupled simulations of ESM.
- <sup>513</sup> The data of CRESCENDO model simulations were last accessed in January, 2019. While the
- 514 simulations are available for longer time period, the simulation results from 2001 to 2010
- 515 inclusive, consistent with the time period across different observation-based estimates of
- $_{516}$  carbon cycle variables, were used for evaluations in this study. The  $C_{total}$  was calculated as a
- sum of all the carbon pools respiring to the atmosphere. Similar to the observation,  $\tau$  was
- 518 calculated as the ratio of  $C_{total}$  and GPP.

<sup>507 (</sup>https://vesg.ipsl.upmc.fr/thredds/catalog/work/p529viov/cruncep/V8 1901 2016/catalog.ht

- 519 All models were forced in an offline mode with the same meteorological drivers, but at a
- 520 different resolution with a different experimental and process setup (see **Table A2** for an
- 521 overview). Regarding carbon cycle spinup, all models except JULES were initialized with
- *zero* carbo stock, and forced by the repeated meteorological data of the first few years for a
- 523 total of 500-10000 years. It should be noted that, in essence, irrespective of the initial
- <sup>524</sup> condition, the JULES model produces a consistent model state at the end of spinup period <sup>56</sup>,
- <sup>525</sup> and, in general, by definition, the spin-up and initial condition are not systematically related
- <sup>526</sup> to the biases in the model states. There are further differences in soil physics and model
- 527 structure therein. For example, CLM5, JULES, and SURFEX have a finer discretization of
- *soil layers and/or represent soil physics with freeze-thaw dynamics.*
- 529 CLM4.5, the terrestrial component of the CMCC coupled model version 2<sup>57</sup> within CMIP6,
- <sup>530</sup> was used in its Biogeochemical (BGC) configuration <sup>58,59</sup> for the CRESCENDO simulations.
- 531 CLM4.5 describes photosynthetic, hydrologic, and decomposition processes. Photosynthesis
- is based on Ball-Berry model, and hydrology includes soil moisture and groundwater
- processes. The decomposition rates are dependent on soil temperature, soil moisture, oxygenand depth.
- 535 CLM5, the terrestrial component of NorESM, builds on CLM4.5, with major updates on soil
- <sup>536</sup> hydrology and carbon coupling <sup>60</sup>. Soil hydrology is based on variable soil thickness with
- 537 high resolution layers and spatially varying root profile. Photosynthesis is based on the
- 538 Medlyn-model and includes nitrogen limitation using optimality principles. The
- representation of the decomposition process is updated with a new metric for the apparent
- <sup>540</sup> soil carbon turnover <sup>17</sup> in which the temperature affects the moisture limitation of
- 541 decomposition through freezing of liquid water.
- JSBACH version 3.2<sup>61</sup> is the terrestrial component of MPI-ESM1.2. As compared to its
- 543 predecessor, it includes novel components of soil carbon, nitrogen limitation, a five-layer
- 544 hydrology scheme, the wildfire model SPITFIRE, as well as improved land use
- <sup>545</sup> representations <sup>62</sup>. The soil carbon decomposition processes are based on the YASSO model
- <sup>63</sup> and include five different soil carbon pools according to the chemical quality of the organic
- 547 matter, each for woody and non-woody litter. The vertical distribution of soil carbon is not
- resolved and permafrost is not considered in the applied version. Decomposition rates depend
- on air temperature and precipitation, due to the foundation of YASSO on observed litter
- 550 decomposition rates.

JULES <sup>56</sup>, the land component of UKESM1, includes the improvements in parameterization

- of the vegetation dynamics, canopy structural properties, and the parameters affecting the
- <sup>553</sup> photosynthesis and respiration fluxes <sup>64</sup>. The nutrient availability limits the photosynthesis
- <sup>554</sup> indirectly by controlling the biomass and leaf area index. The litter and exudate fluxes
- <sup>555</sup> convert to the soil organic carbon through controls of nitrogen availability, and the soil
- <sup>556</sup> decomposition process are controlled further by temperature and moisture.
- 557 LPJ-GUESS <sup>65</sup> is the dynamic vegetation model of EC-Earth used in CMIP6. LPJ-GUESS
- <sup>558</sup> employs a two-layer leaky bucket soil hydrology scheme with percolation between layers and
- deep drainage. Photosynthesis is controlled by nitrogen limitation following whole plant
- optimal nitrogen content. The soil organic matter dynamics are based on the CENTURY soil
- <sup>561</sup> model in which the decay rates depend on the moisture content of the top soil layer, soil
- temperature, texture, lignin fractions, tillage and nitrogen limitation.
- 563 ORCHIDEE, the land surface scheme of the IPSL-CM5 ESM, consists of a multi-layer soil
- that accounts for transport of water by diffusion and deep drainage, and of heat with
- <sup>565</sup> improved thermodynamics and conduction process <sup>66</sup>. The resulting soil moisture and
- temperature are used for analytically solving soil organic matter dynamic that guarantees
- steady-state conditions. Photosynthesis is based on the Farquhar-model and is limited by the
- leaf nitrogen. The soil heterotrophic respiration is independent of the nitrogen content of soilorganic matter.
- 570 SURFEX/ISBA-CTRIP, the land surface scheme of the CNRM-ESM2-1 in CMIP6 67,68
- solves the energy and water balance of one vegetation canopy, 12 snow layers, and up to 14
   soil layers in deep soils and permafrost. Soil texture, albedo and carbon content determine the
- thermal and hydraulic properties of the soil. The litter and soil organic matter processes are
- <sup>574</sup> based on the soil carbon part of the CENTURY model, and the soil heterotrophic respiration
- is limited under high soil moisture. The nitrogen limitation leaf growth is empirically based
  on a meta-analysis of CO2 enrichment experiments that limit leaf nitrogen content and
- 577 specific leaf area.
- 578

# 579 Zonal variations and correlation:

The zonal variations and correlations are computed using a moving window of 10 grid cells along the latitude (5° coverage). The zonal means of the *GPP* and  $C_{total}$  are weighted by the

area of the grid cell which vary along the latitude, and the zonal mean of  $\tau$  was estimated as

the ratio of zonal  $C_{total}$  and GPP, and not as a zonal mean of  $\tau$  in different grid cells. The observational uncertainty band for the zonal means are calculated as 5<sup>th</sup> and 95<sup>th</sup> percentiles of the zonal value using different data products.

The correlations between  $\tau$  or *GPP* or *C*<sub>total</sub> and temperature or precipitation were calculated 586 as the partial Pearson's correlation coefficient controlling for the other. To accommodate for 587 the influence of the extreme cold winter temperature on the mean annual temperature in high 588 latitude regions, the monthly temperature below -5°C were considered physiologically (for 589 photosynthesis) and metabolically (for respiration) irrelevant, and these values were set to -590 5°C while calculating the mean annual temperature. Note that the processing is only relevant 591 for high latitude regions where the duration of and temperature in winter vary significantly 592 within a moving window (see Figure 3 and Figure A1. 593

Within a moving window, one percent grid cells were discarded to minimize the effects of outliers on the correlation coefficient. Unless otherwise mentioned in the captions, the uncertainty band around the observation are the ranges within 5<sup>th</sup> and 95<sup>th</sup> percentile of the values calculated from different data products. The partial correlation coefficients were tested for statistical significance at 5% level of significance, and all insignificant correlations are masked out when plotting.

# 601 Acknowledgments

- 602 The CRESCENDO simulations and analysis for this study were carried out under the support
- of "Coordinated Research in Earth Systems and Climate: Experiments, kNowledge,
- 604 Dissemination and Outreach (CRESCENDO)" project of the European Union's Horizon 2020
- 605 Framework Programme (grant agreement no. 641816). SK acknowledges the additional
- <sup>606</sup> support of the "Erdsystemforschung: Afrikanische Grundwasserressourcen im Zuge des
- 607 globalen Wandels" (Earth System Research: Groundwater Resources in Africa under Global
- 608 Change) project of the Max Planck Society.

# 609 Author Contributions

- 610 SK and NC conceptually designed the study. SK carried out the analysis and prepared the
- manuscript. CJ, VB, CD, YF, VG, EJ, HL, SM, JN, DP, PP, DW, AW contributed to and or
   provided CRESCENDO model simulations. All authors had intellectual and/or direct input to
- 613 the discussion of the results and completion of the manuscript.

# 614 Data and Code Availability

- All the datasets used in this study are available in the public domain, and can be obtained by
- following the original publications that are cited in respective sections. The python code used
- 617 to analyse the data and plot the results are available publicly at <u>https://github.com/koir-</u>
- 618 <u>su/c\_cycle\_eval</u>.

619

# 621 Extended Tables

622

# Table A1. Summary of the median and the range of observation-based gross primary

624 productivity (*GPP* in pgC/year), total carbon storage ( $C_{total}$  in pgC) and turnover time ( $\tau$  in

*years*) globally, and in different climatic regions. The median is calculated from the ensemble

members of the observation-based data. The range indicates the values within 5<sup>th</sup> and 95<sup>th</sup>

627 percentile of the ensemble members.

	Global	Arid	Semi-arid	Sub-humid	Humid
GPP	109.5	2.7	16.8	12.8	77.6
	(98.3-119.6)	(2.3-3.1)	(14.8-19.3)	(11.1-13.6)	(69.0-85.3)
Ctotal	3482.9	149.2	680.3	493.9	2150.4
	(2245.0-4804.8)	(105.3-162.4)	(424.7-811.5)	(289.3-698.0)	(1418.0-3151.5)
τ	31.9	52.9	39.7	38.8	27.7
	(18.9-47.4)	(35.1-73.2)	(22.9-54.1)	(21.4-62.1)	(16.8-44.3)

629

630 Table A2. An overview of the CRESCENDO models, the major processes therein, and the simulation setup. Human activities include human 631 need-driven processes that affects the vegetation carbon stock such as crop harvest (H), pasture (P), wood harvest (W), deforestation (D), land 632 use changes (LUC), and nitrogen deposition (NDep). All models include carbon losses from heterotrophic respiration (RH), and some models 633 include the carbon losses from soil due to leaching of dissolved organic carbon (DOC).

SN	Model Name (LSM/ES M)	Institution	Spinup (years)	Initial carbon stock	Forcing for spinup (repeat years)	Native spatial resolution (° latitude x ° longitude)	Dynamic vegetatio n process	Human activitie s	Soil physics (number of layers, total depth)	Soil carbon sink terms
1	CLM4.5/ CMCC- ESM	Euro Mediterranean Centre on Climate Change (CMCC) Foundation	1020 accelerated + 300 non- accelerated	0	30 years	1.25°x0.937 5°	no	LUC, W, D	15 layers, 35.1776 m	RH
2	CLM5/No rESM	NORCE Norwegian Research Centre, Bergen, Norway	1526	0	30 years	0.5°x0.5°	no	LUC, W, D	20 layers, 8.5 m	RH; carbon cost for nutrient processes
3	JSBACH/ MPI-ESM	Max Planck Institute for Meteorology	> 10000	C stock from TRENDY v6 simulation	20 years	~1.875°x1.8 75° (T63)	not activated	LUC, NDep, W, H	Soil physics: 5 layers, 9.834 m; soil carbon: 1 layer	RH
4	JULES/U KESM	Met Office Hadley Centre, UK NERC	10000	C stock from independent simulation	20 years	~1.25°x1.87 5° (N96)	yes	LUC, NDep, H	4 layers, 3 m	RH
5	LPJ- GUESS/E C-Earth	Department of Physical Geography and Ecosystem Science, Lund University	Soil: 40500; Vegetation: 500	0	30 years, no IAV of temperature	0.5°x0.5°	yes	LUC, NDep, H, P	Soil moisture: 2 layers, 1.5 m; soil carbon: 1 layer	RH; DOC
6	ORCHID EE/IPSL- CM5	Institut Pierre Simon Laplace (IPSL), France	340 for litter input; cSoil: until steady state	0	30 years	0.5°x0.5°	no	LUC, W, H, D	Soil moisture: 11 layers, 2 m; soil carbon: 1 layer	RH
7	SURFEX/ CNRM- CM5	CNRM, Meteo- France/CNRS/Université Fédérale de Toulouse	Soil: 8875; Vegetation: 470	0	20 years	1°x1°	no	LUC	Soil temperature: 14 layers, 12 m; soil moisture: 14 layers, till rooting depth (1m for grasses, 8m for tropical forests)	RH; DOC

Table A3. Summary of the parameters of the non-linear least square fitting of the relationships between  $\tau$  and *GPP* and climate: a: coefficient for the quadratic term, b: coefficient for the linear term, and c: constant.

			Arid	S	emi-arid		Su	ıb-humid	-	Humid			
		a b c		c	a b c		c	a b		c	a	b	c
	Obs-based	2.61E-05	-3.45E-02	2.25	1.01E-04	-3.34E-02	1.97	1.33E-12	-3.01E-02	1.84	2.38E-04	-3.60E-02	1.88
	CLM4.5	9.51E-04	-5.23E-02	2.13	1.36E-04	-3.48E-02	1.96	7.23E-16	-3.33E-02	1.92	1.92E-04	-3.32E-02	1.82
<u> </u>	CLM5	1.01E-03	-6.36E-02	1.86	9.70E-04	-5.82E-02	1.53	4.41E-04	-3.96E-02	1.40	7.31E-04	-3.93E-02	1.44
[Y]	JSBACH	1.04E-20	-1.62E-02	0.91	5.00E-04	-3.55E-02	1.00	1.56E-04	-3.11E-02	1.16	2.06E-05	-2.53E-02	1.29
ž	JULES	3.94E-04	-2.56E-02	1.45	9.43E-20	-1.50E-02	1.54	1.52E-18	-1.11E-02	1.51	7.76E-16	-1.44E-02	1.43
ı	LPJ-GUESS	5.86E-04	-3.81E-02	1.41	2.84E-04	-2.81E-02	1.41	1.70E-04	-2.36E-02	1.39	2.16E-04	-2.67E-02	1.42
	ORCHIDEE	5.55E-04	-3.32E-02	1.21	4.82E-04	-2.49E-02	1.07	4.93E-04	-2.01E-02	1.00	1.05E-04	-1.50E-02	1.15
	SURFEX	3.31E-14	-3.37E-02	2.23	1.48E-17	-2.98E-02	1.95	5.78E-20	-2.54E-02	1.75	2.85E-17	-2.84E-02	1.72
	Obs-based	1.05E-06	-2.64E-03	2.40	2.30E-06	-4.19E-03	3.05	1.11E-06	-2.86E-03	2.91	2.65E-07	-1.31E-03	2.60
	CLM4.5	9.25E-07	-2.26E-03	2.19	2.03E-06	-3.90E-03	2.96	8.09E-07	-2.57E-03	2.93	2.63E-07	-1.27E-03	2.52
•	CLM5	1.54E-06	-3.44E-03	2.01	3.97E-06	-6.45E-03	3.20	2.00E-06	-4.32E-03	3.00	2.74E-07	-1.24E-03	2.18
P	JSBACH	2.86E-07	-6.45E-04	0.84	1.79E-06	-3.21E-03	1.83	1.22E-06	-2.99E-03	2.28	2.20E-07	-1.05E-03	1.84
π-π	JULES	7.30E-08	-1.87E-04	1.25	4.86E-07	-1.16E-03	1.81	1.80E-31	-4.95E-04	1.76	7.74E-08	-5.02E-04	1.72
	LPJ-GUESS	4.23E-07	-1.02E-03	1.28	2.19E-06	-3.50E-03	2.28	9.63E-07	-2.25E-03	2.23	1.79E-07	-8.92E-04	1.92
	ORCHIDEE	5.30E-07	-1.15E-03	1.18	1.65E-06	-2.61E-03	1.74	9.30E-07	-1.84E-03	1.68	9.53E-08	-4.91E-04	1.41
	SURFEX	9.71E-07	-2.51E-03	2.38	9.44E-07	-2.69E-03	2.69	2.63E-07	-1.50E-03	2.41	1.76E-07	-1.01E-03	2.27
	Obs-based	3.97E-04	1.09E-03	0.17	2.88E-04	1.54E-02	0.41	3.87E-21	3.22E-02	0.69	7.24E-04	4.83E-02	0.69
	CLM4.5	1.35E-04	2.77E-04	0.15	3.68E-15	5.76E-03	0.38	3.85E-34	1.35E-02	0.59	4.17E-04	5.60E-02	0.90
<b>T</b> ▲	CLM5	5.20E-04	-2.19E-03	0.14	5.74E-04	3.08E-03	0.38	4.42E-05	1.66E-02	0.73	6.07E-04	3.89E-02	1.00
M	JSBACH	3.57E-04	-6.90E-04	0.11	7.05E-04	6.56E-03	0.24	7.18E-04	1.97E-02	0.49	1.89E-03	3.95E-02	0.66
-da	JULES	4.85E-15	4.02E-03	0.10	4.40E-04	1.50E-02	0.18	4.03E-04	3.02E-02	0.48	2.07E-04	5.62E-02	0.90
$\overline{\mathbf{G}}$	LPJ-GUESS	2.12E-18	1.97E-02	0.36	1.06E-15	1.80E-02	0.61	6.14E-22	1.87E-02	0.83	4.63E-15	4.74E-02	1.07
	ORCHIDEE	4.04E-04	3.40E-03	0.05	6.45E-04	1.67E-02	0.26	1.54E-04	3.36E-02	0.66	8.37E-04	5.06E-02	0.78
	SURFEX	2.61E-04	2.62E-03	0.07	5.66E-04	1.27E-02	0.24	5.08E-04	2.57E-02	0.49	1.64E-03	4.20E-02	0.51
	Obs-based	-8.84E-08	1.09E-03	0.00	-1.87E-07	1.43E-03	0.00	-2.45E-07	1.61E-03	0.00	-1.58E-07	1.46E-03	0.00
	CLM4.5	3.30E-07	5.79E-04	0.00	-7.28E-07	1.33E-03	0.00	-9.64E-07	1.84E-03	0.00	-1.01E-07	1.48E-03	0.00
₹L	CLM5	8.81E-08	9.87E-04	0.00	-2.28E-07	1.31E-03	0.00	-8.16E-07	1.96E-03	0.00	-2.27E-07	1.68E-03	0.00
M	JSBACH	2.44E-07	6.05E-04	0.00	6.06E-07	6.19E-04	0.00	3.00E-07	9.70E-04	0.00	-1.31E-08	1.40E-03	0.00
-de	JULES	3.77E-07	4.79E-04	0.00	9.44E-07	3.75E-04	0.00	5.15E-07	8.24E-04	0.00	-2.40E-07	1.73E-03	0.00
Б	LPJ-GUESS	-9.69E-07	2.55E-03	0.00	-1.33E-06	2.42E-03	0.00	-1.23E-06	2.47E-03	0.00	-2.34E-07	1.70E-03	0.00
	ORCHIDEE	1.74E-07	7.45E-04	0.00	7.27E-07	8.20E-04	0.00	-4.85E-08	1.48E-03	0.00	-1.77E-07	1.62E-03	0.00
	SURFEX	3.38E-08	6.52E-04	0.00	7.23E-07	6.34E-04	0.00	3.42E-07	9.49E-04	0.00	-1.09E-07	1.42E-03	0.00

<sup>637</sup> **Table A4**. Performance metrics for least square regression of the relationship between  $\tau$ , *GPP*, and climate:  $r^2$ : coefficient of determination.  $r_{mad}$ : <sup>638</sup> relative median absolute deviation calculated as the median of absolute deviation normalized by the interquartile range of the observation.

		МАТ								MAP								
		Arid		Semi-arid		Sub-humid		Humid		Arid		Semi-arid		Sub-humid		Humid		
		$r^2$	r <sub>mad</sub>	$r^2$	r <sub>mad</sub>	$r^2$	<b>r</b> <sub>mad</sub>	$r^2$	r <sub>mad</sub>	$r^2$	rmad	$r^2$	<b>r</b> <sub>mad</sub>	$r^2$	<b>r</b> <sub>mad</sub>	<b>r</b> <sup>2</sup>	rmad	
	Obs-based	0.06	0.25	0.45	0.14	0.44	0.14	0.37	0.13	0.00	0.27	0.58	0.14	0.57	0.14	0.00	0.18	
	CLM4.5	0.00	0.32	0.00	0.11	0.31	0.10	0.00	0.11	0.00	0.48	0.00	0.16	0.32	0.10	0.00	0.18	
	CLM5	0.00	0.36	0.00	0.18	0.00	0.25	0.00	0.24	0.00	0.37	0.00	0.30	0.01	0.26	0.00	0.35	
ь	JSBACH	0.23	0.42	0.40	0.22	0.34	0.14	0.65	0.16	0.01	0.44	0.40	0.29	0.48	0.15	0.62	0.20	
ι	JULES	0.00	0.39	0.00	0.31	0.01	0.30	0.20	0.17	0.00	0.49	0.00	0.38	0.01	0.32	0.22	0.17	
	LPJ-GUESS	0.17	0.31	0.10	0.19	0.23	0.14	0.11	0.14	0.00	0.47	0.11	0.22	0.25	0.14	0.00	0.19	
	ORCHIDEE	0.51	0.40	0.66	0.22	0.67	0.22	0.66	0.19	0.01	0.45	0.00	0.25	0.04	0.26	0.47	0.24	
	SURFEX	0.14	0.22	0.49	0.14	0.68	0.16	0.86	0.07	0.00	0.29	0.60	0.16	0.74	0.14	0.00	0.12	
	Obs-based	0.22	0.41	0.51	0.32	0.77	0.20	0.84	0.13	0.52	0.32	0.60	0.24	0.73	0.20	0.80	0.18	
	CLM4.5	0.03	0.47	0.11	0.54	0.23	0.51	0.54	0.30	0.69	0.32	0.23	0.47	0.29	0.48	0.69	0.31	
	CLM5	0.19	0.37	0.36	0.37	0.33	0.42	0.63	0.26	0.74	0.31	0.44	0.31	0.33	0.43	0.70	0.25	
GPP	JSBACH	0.14	0.49	0.43	0.45	0.63	0.27	0.74	0.16	0.70	0.39	0.59	0.36	0.68	0.28	0.83	0.15	
	JULES	0.02	0.71	0.49	0.26	0.77	0.23	0.84	0.18	0.69	0.52	0.82	0.16	0.89	0.18	0.90	0.14	
	LPJ-GUESS	0.39	0.37	0.38	0.40	0.29	0.40	0.46	0.36	0.33	0.35	0.43	0.36	0.36	0.37	0.49	0.36	
	ORCHIDEE	0.23	0.37	0.59	0.25	0.73	0.21	0.82	0.12	0.60	0.43	0.71	0.20	0.71	0.22	0.79	0.17	
	SURFEX	0.19	0.34	0.61	0.24	0.79	0.19	0.85	0.10	0.57	0.33	0.78	0.17	0.80	0.13	0.78	0.17	

640

# **Extended** Figures



**Figure A1. Global distribution of climatic regions**. The climatic regions are defined using the ranges of aridity index, which is calculated as the ratio of mean annual precipitation and potential evapotranspiration. The ranges of aridity (in square brackets) that define different climate regions are taken from UNEP (United Nations Environment, 1992). From the original UNEP classification, the hyper-arid climate was merged into arid climate because it only comprised 0.1% of grid cells after excluding the desert regions. The percentage values (in parenthesis) indicate the fraction of the valid grid cells with the climate.



**Figure A2. Covariation of climate and moisture**. Mean annual precipitation (MAP, mm/year), mean annual temperature (MAT, °C) and water availability index (WAI, -) <sup>50</sup> are plotted in X, Y and Z axes respectively. The different colours indicate different climatic regions defined by aridity index, and individual grid-cells are plotted as dots. The WAI is a proxy for soil moisture content based on the water balance concept, and it varies between 0 (dry) and 1 (wet).



**Figure A3**. Same and **Figure 1**, but using the LANDGIS <sup>69</sup> instead of SoilGrids <sup>47</sup> dataset in the observation-based ensemble of soil organic carbon estimates.



**Figure A4. Zonal distributions of mean annual precipitation (mm/year) and mean air temperature (°C).** The values for each latitude are calculated as mean within a moving window of 10 grid cells along the latitude (5°). The shaded region indicates the 5<sup>th</sup> and 95<sup>th</sup> percentiles within the latitudinal moving window.



Figure A5. Same as Figure 3 but using the average temperature that is corrected for low winter temperature below freezing point. For the correction, the temperature of months below  $0^{\circ}$ C are set at  $0^{\circ}$ C before calculating the mean annual temperature, which results in an estimate that only includes physiologically relevant months while considering the duration of the period with below freezing temperature.



Figure A6. Zonal correlation between total ecosystem carbon stock ( $C_{total}$ , kgC/m<sup>2</sup>) and gross primary productivity (*GPP*, kgC/m<sup>2</sup>/year). In a) correlation between  $C_{total}$  and *GPP*, b) controlled for precipitation, and c) controlled for temperature are presented. The Pearson's correlation coefficient for each latitude is calculated for a moving window of 10 grid cells along the latitude (5°). The thick dashed blue lines show the normalized mean correlation of all models with shades indicating variation within 1 standard deviation. The observation is plotted as a thick black line with shade indicating the range within the 5<sup>th</sup> and 95<sup>th</sup> percentiles.



Figure A7. Evaluation of a) ecosystem turnover time ( $\tau$ , years), b) gross primary productivity (*GPP*, pgC/year), and c) total carbon storage ( $C_{total}$ , pgC) over different climatic regions. In the bar charts, the observation and their uncertainties (5<sup>th</sup> and 95<sup>th</sup> percentiles) are plotted. In the colourmap matrix, horizontal axis shows different climatic regions and the global values, and the vertical axis shows different models or model ensembles. The colour indicates the bias, calculated as the ratio of modelled and observed values. The inset text shows  $\tau$  or *GPP or*  $C_{total}$  for a given model and climate. The colour of

the text indicates whether the modelled values are within the range of observational uncertainty with green for those within and orange for those outside.







Figure A9. Comparison of the global distribution of gross primary productivity (*GPP*, kgC/m<sup>2</sup>/year) from observation-based estimate and CRESCENDO model simulations. Along the diagonal, the maps of *GPP* from observation and model simulations are plotted. Above the diagonal, the biases (ratio of observation/model along column to observation/model along row) are plotted. Below the diagonal, density scatter plots are plotted with darker colour indicating larger density of points. The thin black line shows the 1:1 line, r and  $\rho$  indicate the spatial Pearson's and Spearman's correlation coefficient, respectively.



Figure A10. Comparison of the global distribution of total ecosystem carbon storage ( $C_{total}$ , kgC/m<sup>2</sup>) from observation-based estimate and CRESCENDO model simulations. Along the diagonal, the maps of  $C_{total}$  from observation and model simulations are plotted. Above the diagonal, the biases (ratio of observation/model along column to observation/model along row) are plotted. Below the diagonal, density scatter plots are plotted with darker colour indicating larger density of points. The thin black line shows the 1:1 line, r and  $\rho$  indicate the spatial Pearson's and Spearman's correlation coefficient, respectively.



**Figure A11.** Same as **Figure 1** of the main text but using the state-of-the-art observationbased datasets of soil carbon (cSoil) and vegetation carbon (cVeg) stocks, and gross primary productivity (*GPP*). In total, 3 different soil carbon datasets, 4 different vegetation carbon datasets, and 6 different *GPP* estimations were used (see Methods). In the first row, a full factorial of all datasets generating 72 members were used. The second, third and fourth rows consider the influence of using different cSoils (3), cVeg (4), and *GPP* (6), respectively, while the other two variables are set at the ensemble medians. The thick lines and shaded regions around them indicate the median and interquartile range of all fitted relationships, respectively. Different colours indicate different climatic regions. Note the logarithmic axis for  $\tau$  in the first and second columns.



Figure A12. Evaluation of sensitivities of evapotranspiration to climate in

**CRESCENDO models.** The relationships between total evapotranspiration (ET, in mm/day) versus mean annual temperature (MAT) and precipitation (MAP) are presented in the first and second column, respectively. Along the row, the comparisons for different models against the observation <sup>50</sup> are presented. The different colours indicate different climatic regions defined by the aridity index. The solid lines indicate the observation-based relationship, and the broken lines indicate relationship from the model simulations. The inset text shows the ratio between the linear coefficients and constants from the least square fits for the model and observation for each climatic region. Note that for ET-MAP relationship, the constant/intercept term is always 0.



Figure A13. Spatial sensitivities of climatological total soil moisture state and carbon fluxes in CRESCENDO model simulations. In the first column, the relationship between the spatial variability of mean annual temperature (*MAT*) and total soil moisture ( $\theta$ ) is presented. In the second, third and fourth columns, the relationships of spatial variabilities of heterotrophic respiration ( $R_h$ ), autotrophic respiration ( $R_a$ ), and gross primary productivity (*GPP*) with MAT are, respectively, presented. The fifth, sixth and seventh columns repeat the same but with  $\theta$ . The different colours indicate different climatic regions defined by the aridity index. The lines are the least square fits for each climatic region. Along the row, the comparisons for different variables and differences in sizes of soil column in different models, all variables are scaled by the 98<sup>th</sup> percentile within each climate region so that the spatial variability is roughly normalized between 0 and 1.

#### **References:**

- Taylor, P. G. *et al.* Temperature and rainfall interact to control carbon cycling in tropical forests. *Ecol. Lett.* 20, 779–788 (2017).
- Anav, A. *et al.* Evaluating the Land and Ocean Components of the Global Carbon Cycle in the CMIP5 Earth System Models. *J. Clim.* 26, 6801–6843 (2013).
- Friedlingstein, P. *et al.* Uncertainties in CMIP5 Climate Projections due to Carbon Cycle Feedbacks. J. Clim. 27, 511–526 (2014).
- Friend, A. D. *et al.* Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO2. *Proc Natl Acad Sci U A* 111, 3280–5 (2014).
- Jones, C. *et al.* Twenty-First-Century Compatible CO2 Emissions and Airborne Fraction Simulated by CMIP5 Earth System Models under Four Representative Concentration Pathways. *J. Clim.* 26, 4398– 4413 (2013).
- Heimann, M. & Reichstein, M. Terrestrial ecosystem carbon dynamics and climate feedbacks. *Nature* 451, 289–92 (2008).
- Ferguson, P. R. & Veizer, J. Coupling of water and carbon fluxes via the terrestrial biosphere and its significance to the Earth's climate system. *J. Geophys. Res.* 112, (2007).
- Bloom, A. A., Exbrayat, J. F., van der Velde, I. R., Feng, L. & Williams, M. The decadal state of the terrestrial carbon cycle: Global retrievals of terrestrial carbon allocation, pools, and residence times. *Proc Natl Acad Sci UA* 113, 1285–90 (2016).
- Koven, C. D. *et al.* Controls on terrestrial carbon feedbacks by productivity versus turnover in the CMIP5 Earth System Models. *Biogeosciences* 12, 5211–5228 (2015).
- Todd-Brown, K. E. O. *et al.* Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations. *Biogeosciences* 10, 1717–1736 (2013).
- 11. Luo, Y. Q. et al. A framework for benchmarking land models. Biogeosciences 9, 3857-3874 (2012).
- Todd-Brown, K. E. O. *et al.* Changes in soil organic carbon storage predicted by Earth system models during the 21st century. *Biogeosciences* 11, 2341–2356 (2014).
- Booth, B. B. B. *et al.* High sensitivity of future global warming to land carbon cycle processes. *Environ. Res. Lett.* 7, 024002 (2012).

- Huntingford, C. *et al.* Contributions of carbon cycle uncertainty to future climate projection spread. *Tellus B* 61, 355–360 (2009).
- 15. Varney, R. M. *et al.* A spatial emergent constraint on the sensitivity of soil carbon turnover to global warming. *Nat. Commun.* **11**, 5544 (2020).
- Braswell, B. H., Schimel, D. S., Linder, E. & Moore, B. The Response of Global Terrestrial Ecosystems to Interannual Temperature Variability. *Science* 278, 870 (1997).
- 17. Koven, C. D., Hugelius, G., Lawrence, D. M. & Wieder, W. R. Higher climatological temperature sensitivity of soil carbon in cold than warm climates. *Nat. Clim. Change* **7**, 817–822 (2017).
- Carvalhais, N. *et al.* Global covariation of carbon turnover times with climate in terrestrial ecosystems. *Nature* 514, 213–7 (2014).
- Falloon, P., Jones, C. D., Ades, M. & Paul, K. Direct soil moisture controls of future global soil carbon changes: An important source of uncertainty. *Glob. Biogeochem. Cycles* 25, (2011).
- Worden, J. *et al.* Satellite Observations of the Tropical Terrestrial Carbon Balance and Interactions with the Water Cycle During the 21st Century. *Rev. Geophys.* e2020RG000711 (2021) doi:https://doi.org/10.1029/2020RG000711.
- Ahlström, A. *et al.* The dominant role of semi-arid ecosystems in the trend and variability of the land CO2 sink. *Science* 348, 895–899 (2015).
- Jung, M. *et al.* Compensatory water effects link yearly global land CO2 sink changes to temperature. *Nature* 541, 516–520 (2017).
- Poulter, B. *et al.* Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle. *Nature* 509, 600–3 (2014).
- 24. Ahrens, B. & Reichstein, M. Depth of understanding. Nat. Clim. Change 7, 762-763 (2017).
- Jiang, L. *et al.* Scale-Dependent Performance of CMIP5 Earth System Models in Simulating Terrestrial Vegetation Carbon\*. *J. Clim.* 28, 5217–5232 (2015).
- Yan, Y., Zhou, X., Jiang, L. & Luo, Y. Effects of carbon turnover time on terrestrial ecosystem carbon storage. *Biogeosciences* 14, 5441–5454 (2017).
- 27. Berdugo, M. et al. Global ecosystem thresholds driven by aridity. Science 367, 787-790 (2020).
- Schulze, E.-D. *et al.* Rooting depth, water availability, and vegetation cover along an aridity gradient in Patagonia. *Oecologia* 108, 503–511 (1996).

- 29. Sohoulande Djebou, D. C., Singh, V. P. & Frauenfeld, O. W. Vegetation response to precipitation across the aridity gradient of the southwestern United states. *J. Arid Environ.* **115**, 35–43 (2015).
- Anav, A. *et al.* Spatiotemporal patterns of terrestrial gross primary production: A review. *Rev. Geophys.* 53, 785–818 (2015).
- 31. Beer, C. *et al.* Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. *Science* **329**, 834–8 (2010).
- Garbulsky, M. F. *et al.* Patterns and controls of the variability of radiation use efficiency and primary productivity across terrestrial ecosystems. *Glob. Ecol. Biogeogr.* 19, 253–267 (2010).
- Thurner, M. *et al.* Large-scale variation in boreal and temperate forest carbon turnover rate related to climate. *Geophys. Res. Lett.* 43, 4576–4585 (2016).
- Nyawira, S., Nabel, J. E. M. S., Brovkin, V. & Pongratz, J. Input-driven versus turnover-driven controls of simulated changes in soil carbon due to land-use change. *Environ. Res. Lett.* 12, 084015 (2017).
- Nyawira, S., Nabel, J. E. M. S., Don, A., Brovkin, V. & Pongratz, J. Soil carbon response to land-use change: evaluation of a global vegetation model using observational meta-analyses. *Biogeosciences* 13, 5661–5675 (2016).
- Roderick, M. L., Greve, P. & Farquhar, G. D. On the assessment of aridity with changes in atmospheric CO2. *Water Resour. Res.* 51, 5450–5463 (2015).
- Jung, M. *et al.* Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *J. Geophys. Res.* 116, (2011).
- Jung, M. *et al.* Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the FLUXCOM approach. *Biogeosciences* 17, 1343–1365 (2020).
- Li, X. & Xiao, J. Mapping Photosynthesis Solely from Solar-Induced Chlorophyll Fluorescence: A Global, Fine-Resolution Dataset of Gross Primary Production Derived from OCO-2. *Remote Sens.* 11, 2563 (2019).
- Zhang, Y. *et al.* A global moderate resolution dataset of gross primary production of vegetation for 2000–2016. *Sci. Data* 4, 170165 (2017).

- 41. Joiner, J. & Yoshida, Y. Satellite-based reflectances capture large fraction of variability in global gross primary production (GPP) at weekly time scales. *Agric. For. Meteorol.* **291**, 108092 (2020).
- Baldocchi, D. *et al.* FLUXNET: A New Tool to Study the Temporal and Spatial Variability of Ecosystem-Scale Carbon Dioxide, Water Vapor, and Energy Flux Densities. *Bull. Am. Meteorol. Soc.* 82, 2415–2434 (2001).
- 43. Sun, Z. *et al.* Evaluating and comparing remote sensing terrestrial GPP models for their response to climate variability and CO2 trends. *Sci. Total Environ.* **668**, 696–713 (2019).
- 44. *Harmonized World Soil Database HWSD (version 1.2)*. http:// webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/ (2012).
- Hugelius, G. *et al.* The Northern Circumpolar Soil Carbon Database: spatially distributed datasets of soil coverage and soil carbon storage in the northern permafrost regions. *Earth Syst. Sci. Data* 5, 3–13 (2013).
- Hengl, T. *et al.* SoilGrids250m: Global gridded soil information based on machine learning. *PLOS ONE* 12, e0169748 (2017).
- Sanderman, J., Hengl, T. & Fiske, G. J. Soil carbon debt of 12,000 years of human land use. *Proc. Natl. Acad. Sci.* 114, 9575–9580 (2017).
- Fan, N. *et al.* Apparent ecosystem carbon turnover time: uncertainties and robust features. *Earth Syst. Sci. Data* 12, 2517–2536 (2020).
- 49. Tifafi, M., Guenet, B. & Hatté, C. Large Differences in Global and Regional Total Soil Carbon Stock Estimates Based on SoilGrids, HWSD, and NCSCD: Intercomparison and Evaluation Based on Field Data From USA, England, Wales, and France. *Glob. Biogeochem. Cycles* 32, 42–56 (2018).
- Jung, M. *et al.* The FLUXCOM ensemble of global land-atmosphere energy fluxes. *Sci. Data* 6, 74 (2019).
- Trabucco, A. & Zomer, R. J. Global Aridity Index and Potential Evapotranspiration (ET0) Climate Database v2. (2019) doi:https://doi.org/10.6084/m9.figshare.7504448.v3.
- Barrow, C. J. World atlas of desertification (United nations environment programme), edited by N. Middleton and D. S. G. Thomas. Edward Arnold, London, 1992. isbn 0 340 55512 2, £89.50 (hardback), ix + 69 pp. *Land Degrad. Dev.* 3, 249–249 (1992).
- 53. Huber, P. J. Robust Estimation of a Location Parameter. Ann. Math. Stat. 35, 73–101 (1964).

- Sitch, S. *et al.* Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences* 12, 653–679 (2015).
- 55. Wei, Y. *et al.* The North American Carbon Program Multi-scale Synthesis and Terrestrial Model Intercomparison Project – Part 2: Environmental driver data. *Geosci. Model Dev.* 7, 2875–2893 (2014).
- Clark, D. B. *et al.* The Joint UK Land Environment Simulator (JULES), model description Part 2: Carbon fluxes and vegetation dynamics. *Geosci. Model Dev.* 4, 701–722 (2011).
- Cherchi, A. *et al.* Global Mean Climate and Main Patterns of Variability in the CMCC-CM2 Coupled Model. *J. Adv. Model. Earth Syst.* 11, 185–209 (2019).
- Koven, C. D. *et al.* The effect of vertically resolved soil biogeochemistry and alternate soil C and N models on C dynamics of CLM4. *Biogeosciences* 10, 7109–7131 (2013).
- Oleson, K. *et al.* Technical description of version 4.5 of the Community Land Model (CLM). (2013) doi:10.5065/D6RR1W7M.
- Lawrence, D. M. *et al.* The Community Land Model Version 5: Description of New Features,
   Benchmarking, and Impact of Forcing Uncertainty. *J. Adv. Model. Earth Syst.* 11, 4245–4287 (2019).
- Reick, C. H. *et al.* JSBACH 3 The land component of the MPI Earth System Model: documentation of version 3.2. (2021) doi:10.17617/2.3279802.
- Mauritsen, T. *et al.* Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO2. *J. Adv. Model. Earth Syst.* 11, 998–1038 (2019).
- Tuomi, M. *et al.* Leaf litter decomposition—Estimates of global variability based on Yasso07 model. *Ecol. Model.* 220, 3362–3371 (2009).
- Sellar, A. A. *et al.* UKESM1: Description and Evaluation of the U.K. Earth System Model. *J. Adv. Model. Earth Syst.* 11, 4513–4558 (2019).
- 65. Smith, B. *et al.* Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model. *Biogeosciences* **11**, 2027–2054 (2014).
- Vuichard, N. *et al.* Accounting for carbon and nitrogen interactions in the global terrestrial ecosystem model ORCHIDEE (trunk version, rev 4999): multi-scale evaluation of gross primary production. *Geosci. Model Dev.* 12, 4751–4779 (2019).

- 67. Decharme, B. *et al.* Recent Changes in the ISBA-CTRIP Land Surface System for Use in the CNRM-CM6 Climate Model and in Global Off-Line Hydrological Applications. *J. Adv. Model. Earth Syst.*11, 1207–1252 (2019).
- 68. Delire, C. *et al.* The Global Land Carbon Cycle Simulated With ISBA-CTRIP: Improvements Over the Last Decade. *J. Adv. Model. Earth Syst.* **12**, e2019MS001886 (2020).
- 69. Wheeler, I. & Hengl, T. Soil organic carbon stock (0–30 cm) in kg/m2 time-series 2001–2015 based on the land cover changes. (2018) doi:10.5281/zenodo.2529721.