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.

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

Underrepresented controls of aridity in climate sensitivity of carbon cycle models

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

[1]Max Planck Institute for Biogeochemistry, Jena, Germany.

[2]Met Office Hadley Centre, Exeter, EX1 3PB, UK

[3]Technische Universität Dresden, Institute of Photogrammetry and Remote Sensing, Helmholtzstr. 10, 01069, Dresden, Germany

[4]Max Planck Institute for Meteorology, Hamburg, Germany.

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

[6]NORCE Norwegian Research Centre, Bjerknes Centre for Climate Research, Bergen, Norway

[7]Center for the Environment, Faculty of Arts & Sciences, Harvard University, Cambridge, MA, USA

[8]Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC), Bologna, Italy

[9]Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ CE Orme des Merisiers, Gif sur Yvette, France

[10]Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden

[11]Departamento de Ciências e Engenharia do Ambiente, DCEA, Faculdade de Ciências e Tecnologia, FCT, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

Corresponding Author: Sujan Koirala (skoirala@bgc-jena.mpg.de)

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

Introduction

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

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

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 to climate change have been extensively studied [15], but mostly focused on temperature, as the first principles dictate the key role of temperature in determining the current and future carbon cycle dynamics [16]. The studies focusing only on energy-limited humid and cold regions, understandably, have shown strong associations of spatial variations of carbon stocks and temperature [17]. When other climatic regions globally are considered, though, temperature alone only explains ~10% of the total spatial variation of the observation-based soil organic carbon [10] suggesting a larger influence of additional environmental factors. In particular, precipitation, which is the primary source of moisture for all land processes, is potentially a strong driver of spatial variation of turnover times, the τ [18], as demonstrated by the significance of moisture effects on decomposition [19] and the observation-based linkages between water and carbon cycles [20]. The moisture availability is especially critical in arid to semi-arid regions, which cover ~39% of the global vegetated land, and substantially influence the interannual variabilities of global biosphere-atmosphere carbon exchange [21–23], but the effects on modelled τ has been largely ignored [24].

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 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 τ. The climate, though, has differential influences on and association with *GPP* and τ, which result in a unique spatial heterogeneity in C_{total} compared to those of either *GPP* or τ. It is, therefore, critical that the sensitivities of *GPP*, τ, and C_{total} to climatic variations are simultaneously evaluated, especially in ESMs where the uncertainties in C_{total} may be related to not only *GPP* [2] but to both τ and *GPP* [9].

Here, based on observation-based estimates of the carbon cycle, we first evaluate the apparent spatial sensitivities of climatological mean *GPP* and τ to two primary climatic drivers: mean annual air temperature (*MAT*) and mean annual precipitation (*MAP*). To infer effective or apparent climatological τ of an ecosystem, we assume that the ecosystem carbon

reaches a steady-state over multiple decades, and calculate τ as the ratio of long-term averages of total ecosystem carbon stock and *GPP* [17,18,25,26]. The apparent spatial sensitivities of the climatological carbon cycle are assessed in a climate phase-space of temperature and precipitation for different climatic regions (arid to humid) that are delineated using an aridity index, defined as the ratio of mean annual precipitation and potential evapotranspiration (see Methods and **Figure A1**). While aridity itself is not directly mechanistically connected to physiological and decomposition processes within an ecosystem, it is a rather clear indicator of relative availability of limiting moisture and energy resources that define the soil moisture regime (**Figure A2**), and consequently, the vegetation responses and patterns [27–29]. The climate sensitivities of τ and *GPP* are represented using non-linear relationships with 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 CRESCENDO project (see Methods), which participated in the Sixth Phase of Coupled Model Intercomparison Project (CMIP6).

Observation-based climate sensitivities

Both τ and *GPP* exhibit unique sensitivities to the spatial variabilities of precipitation and temperature in all climatic regions despite large heterogeneity within a climatic region (**Figure 1**, **Table A4**). First, τ varies strongly across both temperature and precipitation gradients in all climatic regions. The sensitivity of τ to temperature is well-documented for humid regions [17]. Interestingly, there is a significant offset in τ as the climate gets drier with the longest τ in arid regions (τ=178 years at 0°C) compared to humid regions (76 years). In fact, for any given temperature, τ is longer in the arid regions compared to humid regions, 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 1**a, are all within ~17% across different climatic regions (**Table A3**). It should be noted that a significant number of grid cells, with a mean temperature less than 0°C, have longer τ than under warmer temperatures within the same climatic region. Yet, the non-linear increase in temperature sensitivity with decreasing temperature is not as strong as reported previously for τ of soil organic carbon [17] and the linear coefficients dominate the non-linear coefficients in different climatic regions (inset of **Figure 1**, and **Table A3**). While several factors may lead to such differences (e.g., differences in observed data, methodology, etc.), one key reason may also be the selection of climatic regions. For instance, in regions with temperature >10°C,

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

The apparent sensitivity of τ to precipitation is large when the precipitation is lower than ~1000 mm/year across all climatic regions (**Figure 1**b). Only within 0-1000 mm/year range of precipitation, τ varies by more than an order of magnitude. Contrasting to temperature, the τ-precipitation relationship saturates at higher precipitation (>1000 mm/year) in all regions except arid regions where precipitation never crosses the threshold. The overall apparent sensitivity of τ to precipitation is, therefore, hyper-logarithmically nonlinear, as indicated by a clear positive curvature. While the functional relationship of precipitation is almost the same for all but energy-limited humid regions, the sensitivity of τ to precipitation is smaller for humid regions compared to arid regions, indicating a saturation of the precipitation effect on τ. The difference of τ across different climatic regions is the smallest in the lowest precipitation ranges. Similar to the τ-temperature relationship, though, there is still a large variability of τ for given precipitation within all climatic regions (as indicated by low r^2 and 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 climatic region (**Figure 1**c). The temperature sensitivity is the largest in humid and the smallest in arid climatic region. In the humid region, *GPP* sensitivity to temperature increases with temperature. The *GPP* increases steadily for increasing precipitation with very small differences across different climatic regions (**Figure 1**d). The increase in *GPP* with precipitation is quasi-linear in all climates, as *GPP* saturates at the highest precipitation values, especially in the humid regions. This, again, suggests that precipitation plays an important role in determining *GPP* when the vegetation activities are likely to be limited by moisture. While a strong relationship of *GPP* with precipitation has been shown previously [30–32], the minimal difference of *GPP* across different climatic regions for given precipitation shows that the spatial variation of *GPP* is mainly related to precipitation alone.

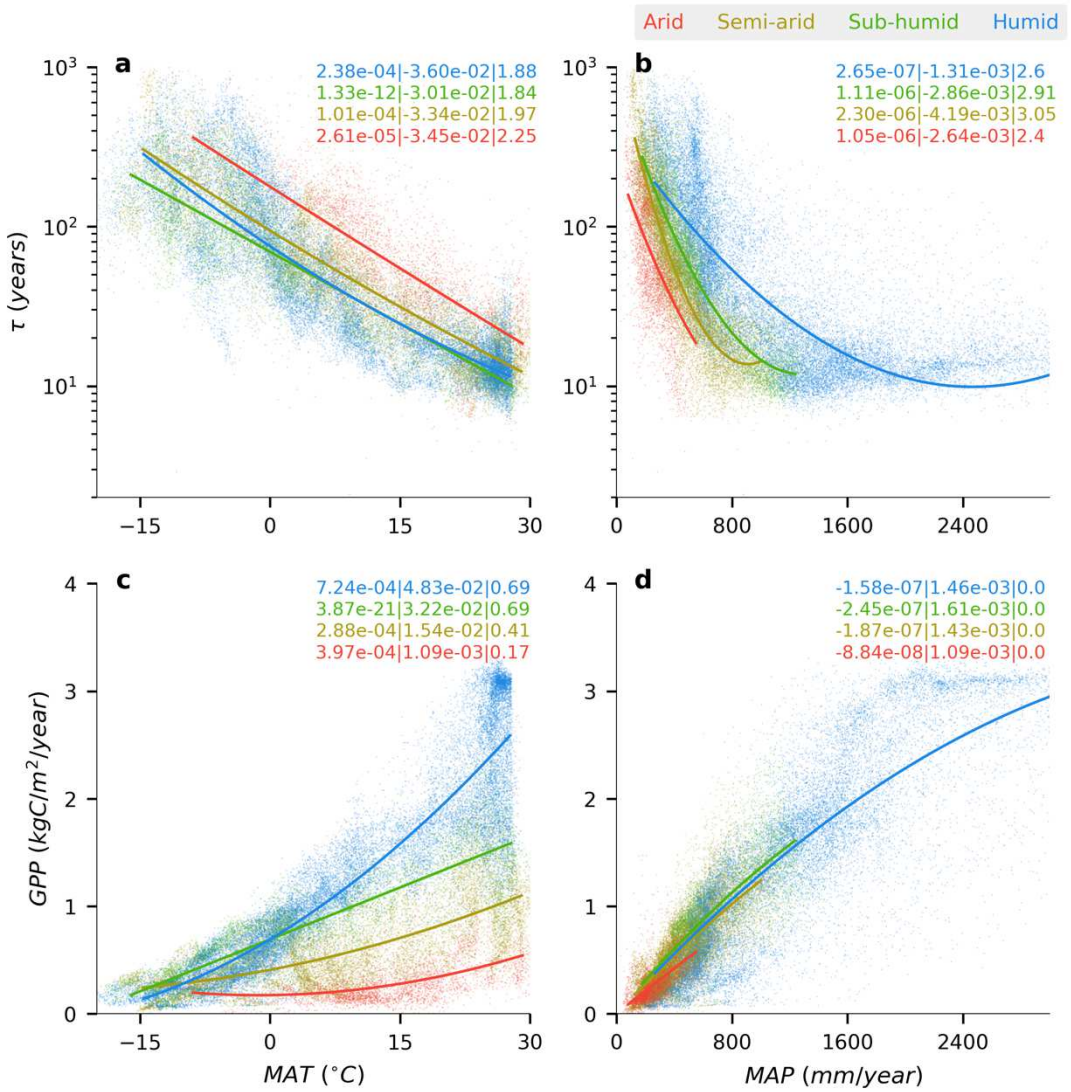

Figure 1. Relationship between observation-based ecosystem turnover time of carbon (τ, years) and gross primary productivity (*GPP*, kgC/m^2/year) with climate. In the top row, the relationship of τ with a) mean annual temperature (MAT, °C) and b) precipitation (MAP, mm/year) are plotted. The bottom row shows the same for *GPP* (c and d). The different colours indicate different climatic regions defined by the aridity index, and individual grid-cells are plotted as dots. The lines show a non-linear least square fit (see Methods) for the variation of τ and *GPP* with MAT and MAP within each climatic region. The inset text shows the parameters of the fit (non-linear coefficient | linear coefficient | constant) while the fitting performances are provided in **Table A4**. Note the logarithmic vertical axis for τ in **a** and **b**.

Climate sensitivities in Carbon Cycle models

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

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

176 Similar to observation, the models also exhibit a tendency of non-linear increase in τ for
177 decreasing precipitation, with the largest sensitivity in arid regions and the smallest
178 sensitivity in humid regions (second column of **Figure 2**). Most models, though, show a
179 smaller sensitivity to precipitation than the observation with a consistent underestimation bias
180 across all climate, with the largest difference in the arid regions. In fact, only CLM4.5 and
181 SURFEX show a significant increase of τ 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
184 regions (third column of **Figure 2**) except for CLM4.5 (**Figure 2**c), which shows a smaller
185 increase in *GPP* at a higher temperature in all but the humid region, and LPJ-GUESS (**Figure**
186 **2**s) with much larger *GPP* in arid regions. In general, all models represent the sensitivity of
187 *GPP* to temperature fairly well with a relatively smaller biases than the same of τ. The
188 sensitivity of *GPP* to precipitation also shows a small spread among the models (fourth
189 column of **Figure 2**). Only in the humid regions, a large number of models, e.g., CLM4.5 and
190 JSBACH, over-predict *GPP* in the highest precipitation range, where the observation-based
191 *GPP* shows a saturating tendency.

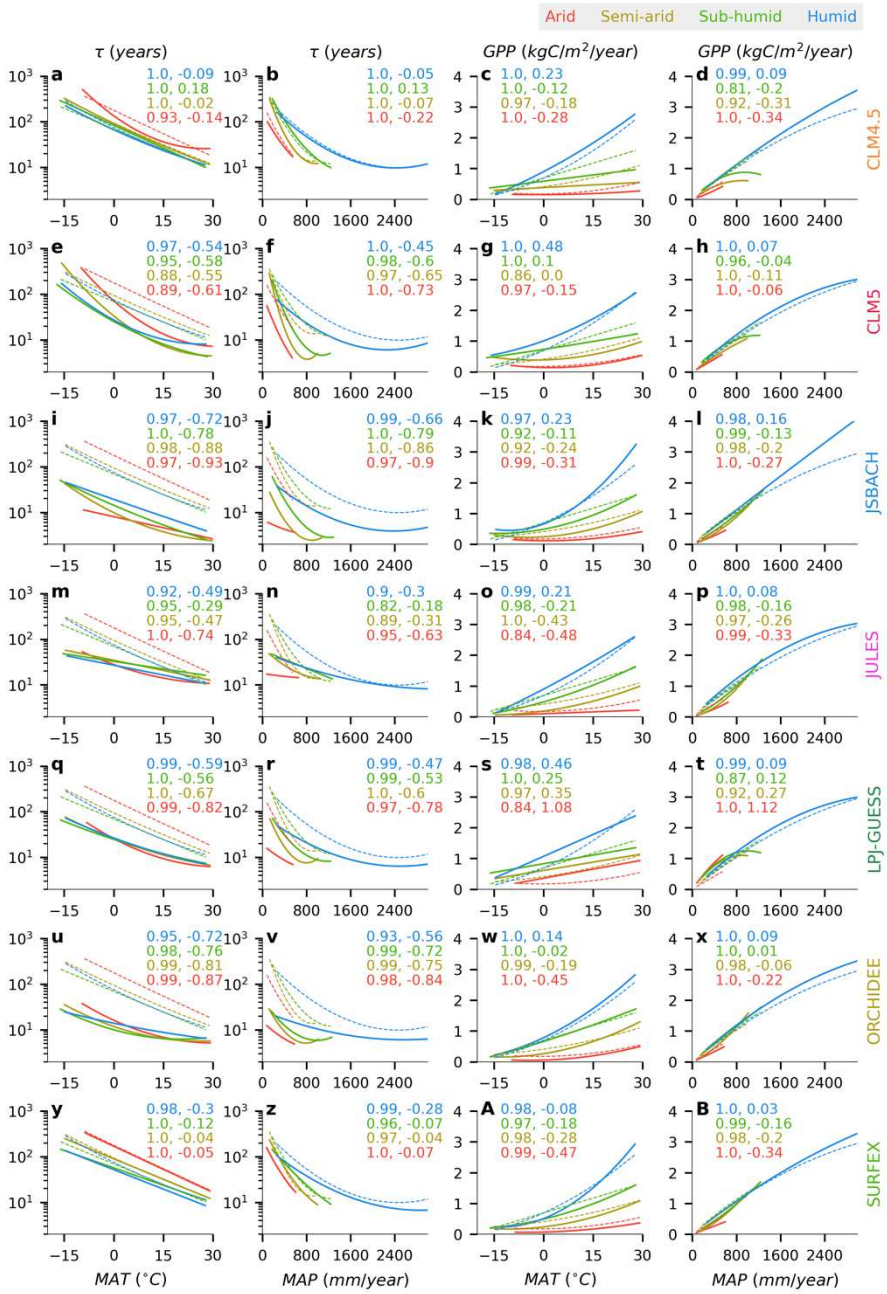

Figure 2. Evaluation of sensitivities of climatological carbon cycle dynamics to climate in CRESCENDO models. In the first and second columns, the relationships between ecosystem turnover time of carbon (τ, years) versus mean annual temperature (*MAT*) and precipitation (*MAP*) are presented, followed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

Partial correlations with climate

To elucidate the regional variations of τ, *GPP* and the resulting C_{total}, and eliminate collinearity between precipitation and climate, here, we evaluate the zonal variations and their partial associations with precipitation and temperature.

In general, the large-scale zonal distribution of τ follows temperature, and that of *GPP* follows precipitation (first column of **Figure 3**, and **Figure A4**). The τ is the shortest in tropical regions with the highest temperature and the longest in high latitudes with the lowest temperature (**Figure 3**a). Similarly, *GPP* is the largest in the wettest tropical regions with the highest precipitation (**Figure 3**d). The zonal distribution of C_{total} reveals a smaller variation across the latitude compared to τ and *GPP* (**Figure 3**g). Nonetheless, the largest C_{total} occur in either the humid tropics with the largest *GPP* or in the high latitudes with the longest τ. In general, the CRESCENDO models produce zonal variations similar to observation-based estimates, with the smallest bias for *GPP*, and a consistent underestimation bias of τ and C_{total} across the latitude.

The observation-based τ has a stronger negative correlation with temperature in the energy-limited regions such as the tropics and high latitudes of both hemispheres compared to the moisture-limited regions (**Figure 3**b). In general, the latitudinal variation of τ-temperature correlation is much stronger in the models than in the observation. This is especially true for sub-tropical mid-latitudes (40ºS to 40ºN) except the humid tropical regions. In the same regions, the local variation of τ has a stronger correlation with precipitation than with temperature (**Figure 3**c). The τ-precipitation relationships in the models have a larger spread, and most models have a weaker τ-precipitation correlation than in the observation. This is especially clear in the subtropical southern hemisphere and temperate northern hemisphere. This weaker τ-precipitation correlation is concurrent with a stronger τ-temperature correlation in the models suggesting a dominant temperature control on carbon cycle sensitivity to climate. In these regions, only SURFEX shows a strong local-scale τ-precipitation correlation that is consistently stronger than the observation, but it also exhibits a stronger bias in the high latitudes.

Further, both the observations and models show a much stronger correlation of *GPP* with precipitation (**Figure 3**f) than with temperature (**Figure 3**e) across most of the latitudes. Only in the northern high latitudes over ~50ºN (with MAT < 0ºC), *GPP* has a stronger correlation with temperature than with precipitation. In fact, in these regions, the low winter temperature

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 correlation of *GPP* with climate than of τ with climate. This shows that the models, in general, reproduce both the global and local responses of *GPP* to temperature and precipitation much better than those of τ.

Further, as expected, the observation-based relationship of C_{total} and climate is a complex amalgam of those of *GPP* and τ. For example, C_{total}-temperature correlation is relatively small in magnitude (**Figure 3**h) due to the opposing τ-temperature (negative) and *GPP*-temperature (positive) correlations. The C_{total}-precipitation correlation follows the same zonal variation as that of *GPP* but with a lower magnitude (**Figure 3**i) due to predominantly negative τ-precipitation correlation. The CRESCENDO models generally exhibit a stronger positive correlation of C_{total} with climate than in the observation. This strong C_{total}-precipitation correlation in the models is due to a strong correlation between C_{total} and *GPP* (see **Figure A6**).

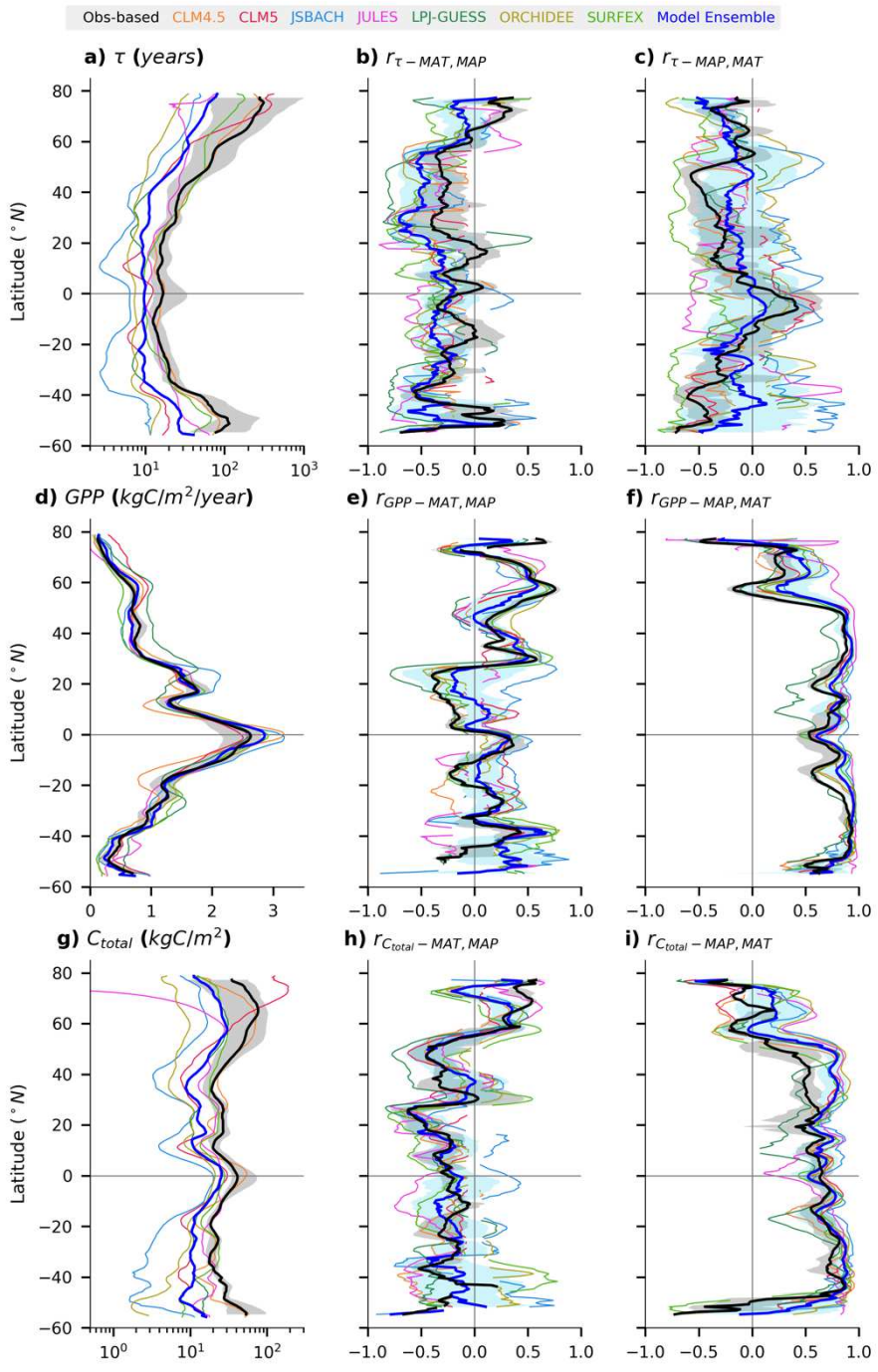

Figure 3. Zonal distribution of ecosystem turnover time of carbon (τ, years), gross primary productivity (*GPP*, kgC/m^2/year), and total ecosystem carbon stock (C_{total}, kgC/m^2) and their relationships with climate. The zonal means are plotted in the left column, and the correlations with temperature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

Model biases and agreement

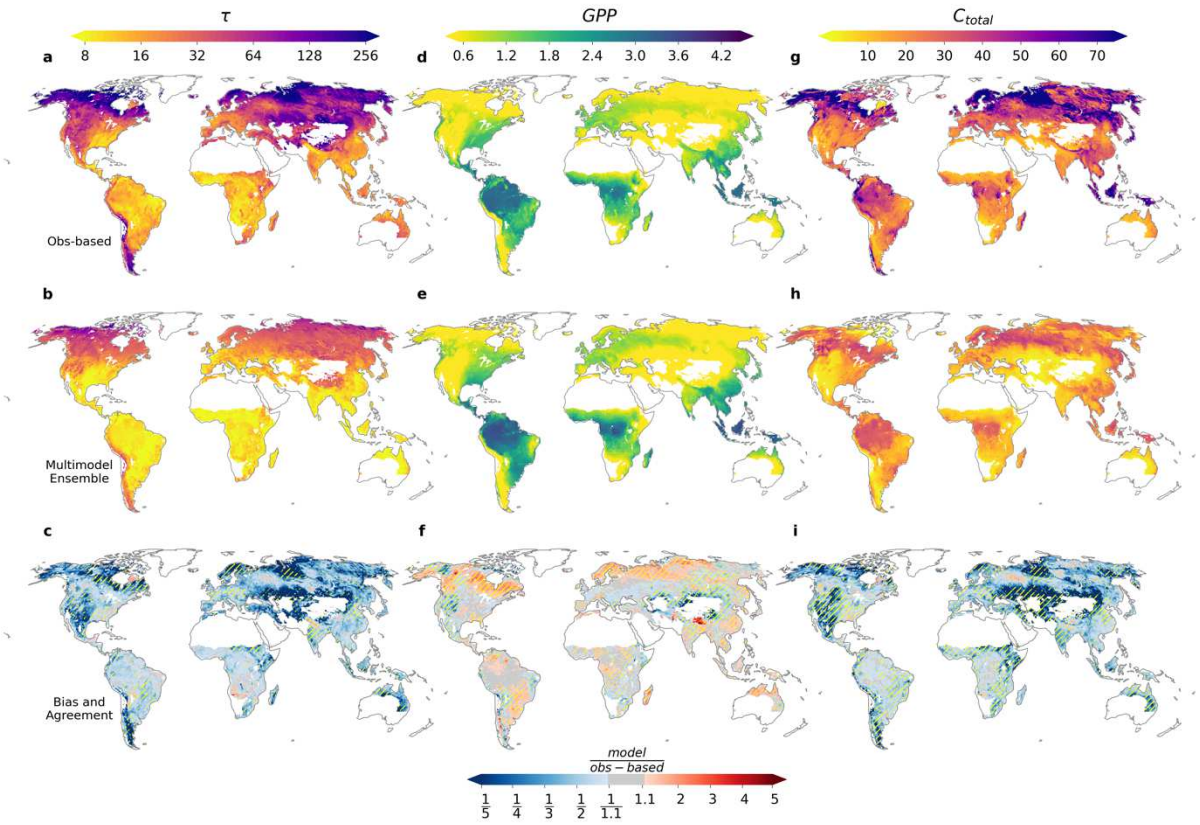

Figure 4. Global distributions of ecosystem turnover time of carbon (τ, years), gross primary productivity (*GPP*, kgC/m^2/year), and total ecosystem carbon stock (C_{total}, kgC/m^2) and their biases. The multimodel ensemble, observation-based estimate, and the bias and agreement are presented in the first, second, and third row, respectively. Multimodel ensemble is calculated as the median of the seven CRESCENDO models. The bias is 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 models fall within the range of observational uncertainties (5th and 95th percentiles).

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

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

Both the spatial variability and magnitude of *GPP* in CRESCENDO models compare better with the corresponding observation than in the case of τ (see **Figure A9** as well). In the tropical humid and semi-arid regions, the *GPP* bias is small in relative terms (**Figure 4**f). In semi-arid regions of south America and Australia, the relative bias in *GPP* is high due to relatively lower *GPP* values in the observation. But the largest overestimation bias can be seen in the northern high latitudes which have low GPP. The models also agree less in the regions with the largest bias with almost ~35% of global grid cells having a low model agreement. Note that such low agreement is also due to small uncertainty range in the ensemble of GPP observations (bar charts in **Figure A7**b). When the global and regional *GPP* are compared, almost all models are within the observational uncertainty globally due to better performance in the humid and sub-humid climate compared to other regions (**Figure A7**b).

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

316 cells), and biases, mostly similar with those of τ, are widespread across different climatic
317 regions (**Figure A7**c).

318 To attribute the bias in C_{total}, we compare the biases in *GPP*, C_{total}, and τ in each
319 CRESCENDO model (**Figure 5**). Note that an ideal model would have low biases in all *GPP*,
320 τ and C_{total}. The contrast between the biases in *GPP* and C_{total} reveals a dominant control of τ
321 on the C_{total} 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
323 of τ, as the biases in *GPP* are relatively lower than those in C_{total} (larger density of points
324 along the axis of no bias in GPP). Other models perform better for both *GPP* and τ in all but
325 arid climatic region. Only CLM4.5 and SURFEX show a good agreement of τ in arid region
326 (red lines close to 1:1 line). CLM4.5, though, also has a slight overestimation bias in C_{total} in
327 all other climatic regions arising from overestimation of τ. In fact, in CLM5, the biases in τ
328 are reduced across all climatic regions. Only JULES and SURFEX produce a larger density
329 of grid cells where biases are low for *GPP*, τ, and C_{total}. In JULES, τ and C_{total} are
330 underestimated mainly in the arid and humid region. Globally, only CLM4.5 has larger range
331 of spatial variability than in the observation (the slope of the fitted global line greater than 1).

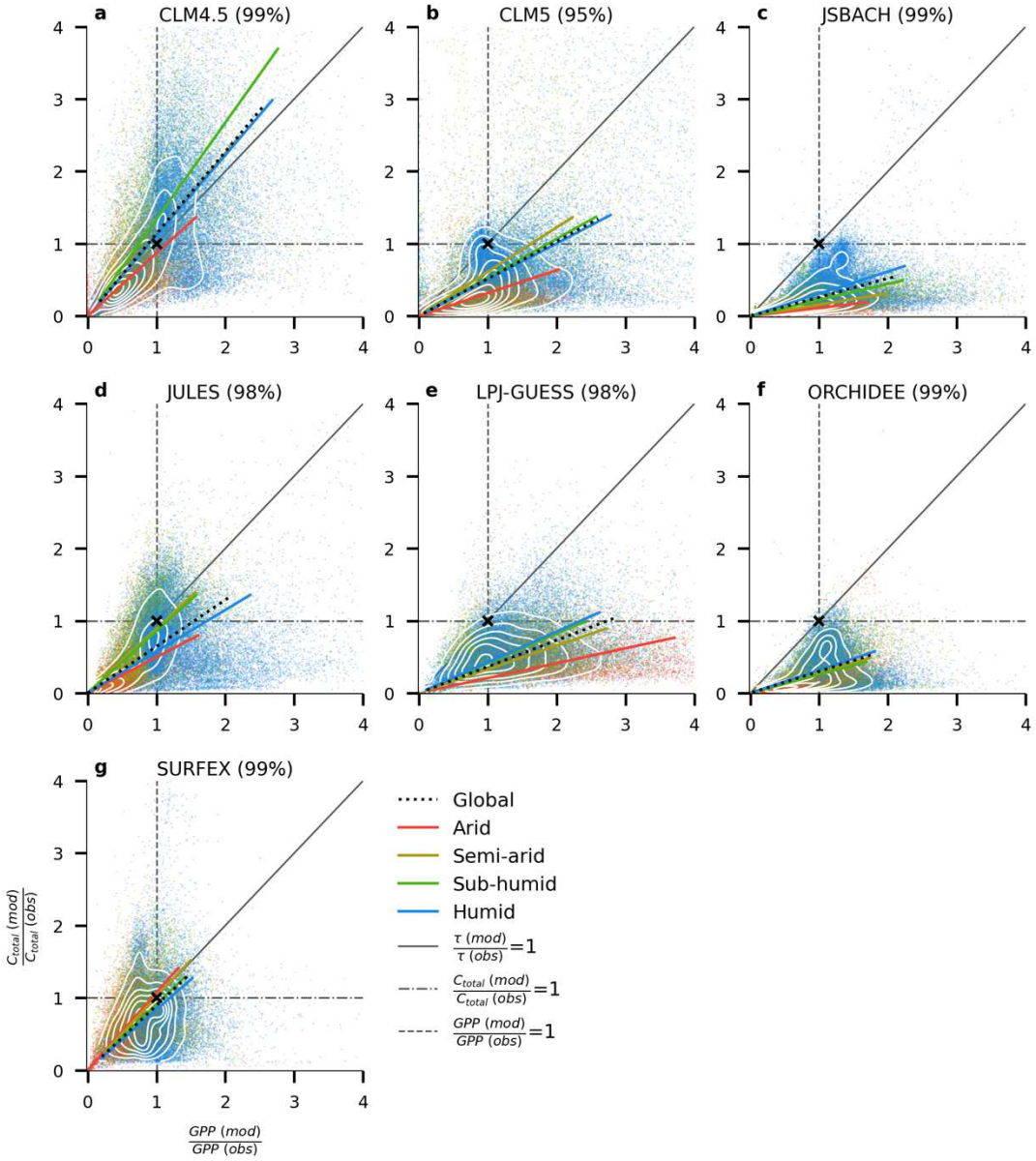

Figure 5. Relationship of biases in gross primary productivity (*GPP*) and total ecosystem carbon stock (C_{total}) in CRESCENDO model simulations. The different colours indicate different climate defined by the aridity classes, and individual grid-cells are plotted as dots. The percentage values in the parenthesis of the title shows the fraction of the global grid cells that are within the range of bias between 0.1 and 4. The dashed vertical line shows the axis of no bias in *GPP* (ratio of *GPP* from model to that from observation = 1), the same for C_{total} is shown by a horizontal dashed line. The grey solid line indicates the axis along which there is no bias in τ. An ideal point of no biases in *GPP*, C_{total}, and τ is indicated by a cross 'x'. The coloured lines indicate the main axis of variations of biases in *GPP* and C_{total} (linearly fitted lines that passes through origin). The white contour lines indicate the density of grid cells, with higher density of lines showing larger occurrence of grid cells.

Conclusions and discussion

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

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

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

377 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,
381 and to bridge the differences between, the apparent moisture controls presented in this study.

382 Lastly, the weaker performance of precipitation-turnover relationships leads to large
383 uncertainties of carbon stocks especially in arid and semi-arid climatic regions, as the biases
384 in C_{total} are largely associated with biases in turnover time rather than those in *GPP*.
385 Additionally, we found that biases of *GPP* and turnover time manifest to a larger bias in
386 C_{total}. This has serious implications in the predictions of changes in carbon stock under
387 climate change, as the changes in precipitation, temperature, and thus aridity [36] will have a
388 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
390 a broader scope of improving the hydrological influences on the terrestrial carbon cycle.

391

Methods

Ecosystem carbon turnover time:

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

$$\tau = \frac{cTotal}{GPP} \quad (1),$$

where, τ is the ecosystem turnover time (years), C_{total} is the long-term average total ecosystem carbon storage per unit area (kgC/m^2) and *GPP* is the annual mean gross primary productivity (kgC/m^2/year). As stated and used in previous studies [17,18,25,26], the τ 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}.

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 [38], solar-induced fluorescence (SIF) based GOSIF *GPP* [39], light use efficiency based *GPP* from vegetation photosynthesis model (VPM) [40], and lastly, an independent machine learning prediction from FluxSat [41] that uses corrected satellite-based reflectance. All *GPP* products are either upscaled from or extensively evaluated against eddy covariance

observations of carbon fluxes from FLUXNET (www.fluxdata.org) sites (Baldocchi et al., 2001). For example, the MTE, FLUXCOM, and FluxSat products use different combinations and corrections of training data and drivers from satellite remote sensing and meteorological dataset to upscale site-level observation to the global scale using different machine learning algorithms. VPM, a light use efficiency model, is similar to MODIS *GPP* but with an additional validation for FLUXNET sites, and GOSIF *GPP* uses a set of *GPP*-SIF relationships to produce an ensemble of high resolution global *GPP* fields that are independent of reflectance-based remote sensing products. While the spatial variability of the mean *GPP* is one of the most robust features when comparing different observation-based *GPP* products, it is desirable to use an ensemble as no single data is superior to all others universally [43].

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

The C_{veg} ensemble includes four different RADAR and LIDAR satellite-based estimates of aboveground biomass with a corresponding estimate for belowground biomass [48]. As different data are based on different satellites and algorithms, and have been validated with observations whenever and wherever available, they are indicative of global vegetation biomass and associated uncertainties in observation-based estimates.

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

In this study, for *GPP*, C_{total} and τ, the ensemble median across data products was used as the representative observation-based estimate, and the 5th and 95th percentiles were used as uncertainty range.

Lastly, for an additional evaluation of CRESCENDO models, the ensemble estimates of evapotranspiration from FLUXCOM remote sensing products [50] was used.

Climatic regions based on aridity index:

The aridity index, calculated as the ratio of mean precipitation and potential evapotranspiration (PET), is used to delineate different climatic regions based on relative availability of moisture and energy. For consistency, the CRU-NCEP precipitation, used to force the CRESCENDO model simulations, is also used for calculating the aridity index. The PET data is based on an estimate from the latest climate datasets at a high resolution [51]. Different climatic regions were then delineated based on the aridity index ranges suggested by UNEP [52] 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 mm/year (0.9°C) in humid regions. These regions constitute 8.2%, 24.3%, 15.1%, and 52.4% of the total area considered in this study. In relative terms, the arid and semiarid regions contribute ~18% of the global *GPP* and hold ~24% of the global carbon stock (**Table A1**). Combined, these regions have a τ of ~43 years, which is ~34% longer than the global τ of ~32 years.

Climate sensitivities:

The overarching aim of this study is to investigate the climate sensitivities of climatological mean carbon cycle dynamics across all major climatic regions globally and assess the apparent sensitivities of spatial covariations of carbon cycle with climate. Under humid climate, spatial sensitivity of τ is itself dependent on temperature [17]. But, to what extent the same applies to other climatic regions and carbon cycle variables is unclear. This is especially 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 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 τ/GPP and climate are non-linear and assume it to be of a second-degree polynomial form as,

$$y = ax^2 + bx + c \quad (2),$$

where, y is the response/dependent variable, x is the independent variable, and a, b and c are the parameters of the quadratic equation. The terms a and b are the coefficients of the non-linear and linear terms, respectively, and c is the constant term that defines the offset/intercept when the independent variable is zero. As τ is expected to have a non-linear exponential relationship with temperature, we log-normalize it and use $\log(\tau)$ as the dependent variable.

The parameters of Eqn (2) for each climatic region are estimated using a robust non-linear least square regression with Huber loss function [53] that is less sensitive to outliers in the data.

CRESCENDO model simulations:

We use the observation-based climate sensitivities to evaluate the process representations in land surface schemes of (used in) the ESMs from seven different European research institutes. The stand-alone offline land simulations were carried out as a part of the CRESCENDO project (https://ukesm.ac.uk/crescendo/) 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 intercomparison project (CMIP6). Following the TRENDY protocol [54], the models were forced by CRUNCEP v7 [55] forcing dataset (https://vesg.ipsl.upmc.fr/thredds/catalog/work/p529viov/cruncep/V8_1901_2016/catalog.html), observed atmospheric CO_2 concentration, nitrogen deposition, and annual land-use changes. The offline simulations forced with an observation-based dataset are suitable for evaluating the response of the land carbon cycle processes to the climate forcing, as they are less prone to biases and uncertainties arising from differences in carbon-climate feedbacks in the coupled simulations of ESM.

The data of CRESCENDO model simulations were last accessed in January, 2019. While the simulations are available for longer time period, the simulation results from 2001 to 2010 inclusive, consistent with the time period across different observation-based estimates of 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, τ was calculated as the ratio of C_{total} and GPP.

All models were forced in an offline mode with the same meteorological drivers, but at a different resolution with a different experimental and process setup (see **Table A2** for an 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 total of 500-10000 years. It should be noted that, in essence, irrespective of the initial condition, the JULES model produces a consistent model state at the end of spinup period [56], and, in general, by definition, the spin-up and initial condition are not systematically related to the biases in the model states. There are further differences in soil physics and model structure therein. For example, CLM5, JULES, and SURFEX have a finer discretization of soil layers and/or represent soil physics with freeze-thaw dynamics.

CLM4.5, the terrestrial component of the CMCC coupled model version 2 [57] within CMIP6, was used in its Biogeochemical (BGC) configuration [58,59] for the CRESCENDO simulations. 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, oxygen and depth.

CLM5, the terrestrial component of NorESM, builds on CLM4.5, with major updates on soil hydrology and carbon coupling [60]. Soil hydrology is based on variable soil thickness with high resolution layers and spatially varying root profile. Photosynthesis is based on the Medlyn-model and includes nitrogen limitation using optimality principles. The representation of the decomposition process is updated with a new metric for the apparent soil carbon turnover [17] in which the temperature affects the moisture limitation of decomposition through freezing of liquid water.

JSBACH version 3.2 [61] is the terrestrial component of MPI-ESM1.2. As compared to its predecessor, it includes novel components of soil carbon, nitrogen limitation, a five-layer hydrology scheme, the wildfire model SPITFIRE, as well as improved land use representations [62]. The soil carbon decomposition processes are based on the YASSO model [63] and include five different soil carbon pools according to the chemical quality of the organic 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 decomposition rates.

JULES [56], the land component of UKESM1, includes the improvements in parameterization of the vegetation dynamics, canopy structural properties, and the parameters affecting the photosynthesis and respiration fluxes [64]. The nutrient availability limits the photosynthesis indirectly by controlling the biomass and leaf area index. The litter and exudate fluxes convert to the soil organic carbon through controls of nitrogen availability, and the soil decomposition process are controlled further by temperature and moisture.

LPJ-GUESS [65] is the dynamic vegetation model of EC-Earth used in CMIP6. LPJ-GUESS 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 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.

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 improved thermodynamics and conduction process [66]. 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 soil organic matter.

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 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 specific leaf area.

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 τ was estimated as

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

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

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 5th and 95th 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.

Acknowledgments

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

Author Contributions

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 the discussion of the results and completion of the manuscript.

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 to analyse the data and plot the results are available publicly at https://github.com/koir-su/c_cycle_eval.

Extended Tables

Table A1. Summary of the median and the range of observation-based gross primary productivity (*GPP* in pgC/year), total carbon storage (C_{total} in pgC) and turnover time (τ 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 5th and 95th percentile of the ensemble members.

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

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

SN	Model Name (LSM/ESM)	Institution	Spinup (years)	Initial carbon stock	Forcing for spinup (repeat years)	Native spatial resolution (° latitude x ° longitude)	Dynamic vegetation process	Human activities	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°x0.9375°	no	LUC, W, D	15 layers, 35.1776 m	RH
2	CLM5/NorESM	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°x1.875° (T63)	not activated	LUC, NDep, W, H	Soil physics: 5 layers, 9.834 m; soil carbon: 1 layer	RH
4	JULES/UKESM	Met Office Hadley Centre, UK NERC	10000	C stock from independent simulation	20 years	~1.25°x1.875° (N96)	yes	LUC, NDep, H	4 layers, 3 m	RH
5	LPJ-GUESS/EC-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	ORCHIDEE/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

634

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

		Arid			Semi-arid			Sub-humid			Humid		
		a	**b**	**c**	**a**	**b**	**c**	**a**	**b**	**c**	**a**	**b**	**c**
τ-MAT	**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
	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
	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
	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
	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
τ-MAP	**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
	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
	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
***GPP*-MAT**	**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
	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
	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
	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
	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
***GPP*-MAP**	**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
	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
	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
	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
	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

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

639

		MAT								MAP							
		Arid		Semi-arid		Sub-humid		Humid		Arid		Semi-arid		Sub-humid		Humid	
		r^2	r_{mad}	r^2	r_{mad}	r^2	r_{mad}	r^2	r_{mad}	r^2	r_{mad}	r^2	r_{mad}	r^2	r_{mad}	r^2	r_{mad}
τ	**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
	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
	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
GPP	**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
	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

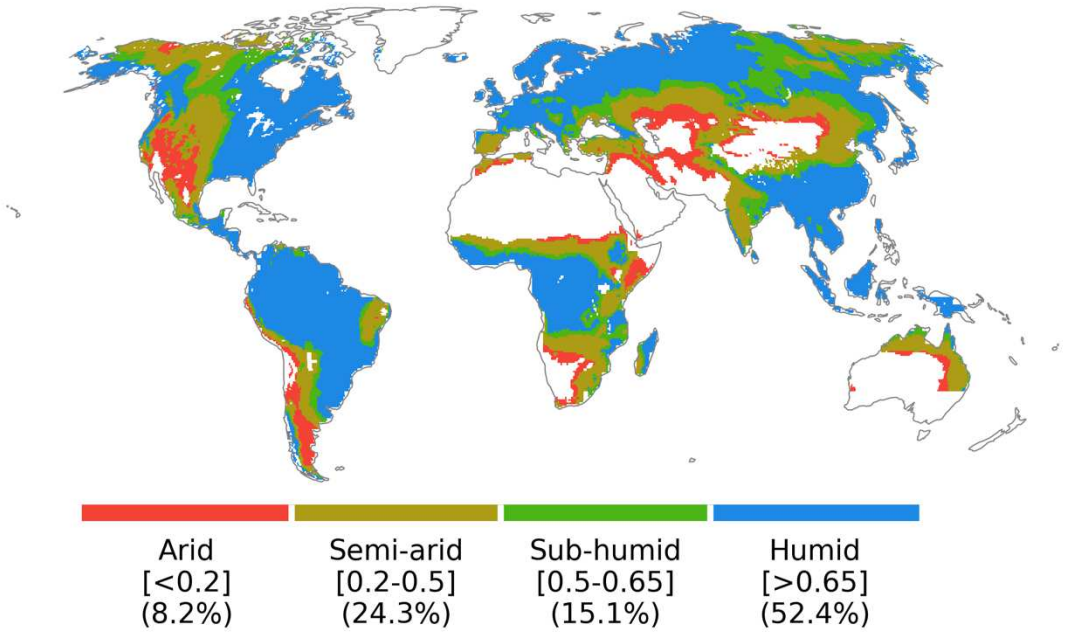

Figure A1. Global distribution of climatic regions. The climatic regions are defined using the ranges of aridity index, which is calculated as the ratio of mean annual precipitation and potential evapotranspiration. The ranges of aridity (in square brackets) that define different climate regions are 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.

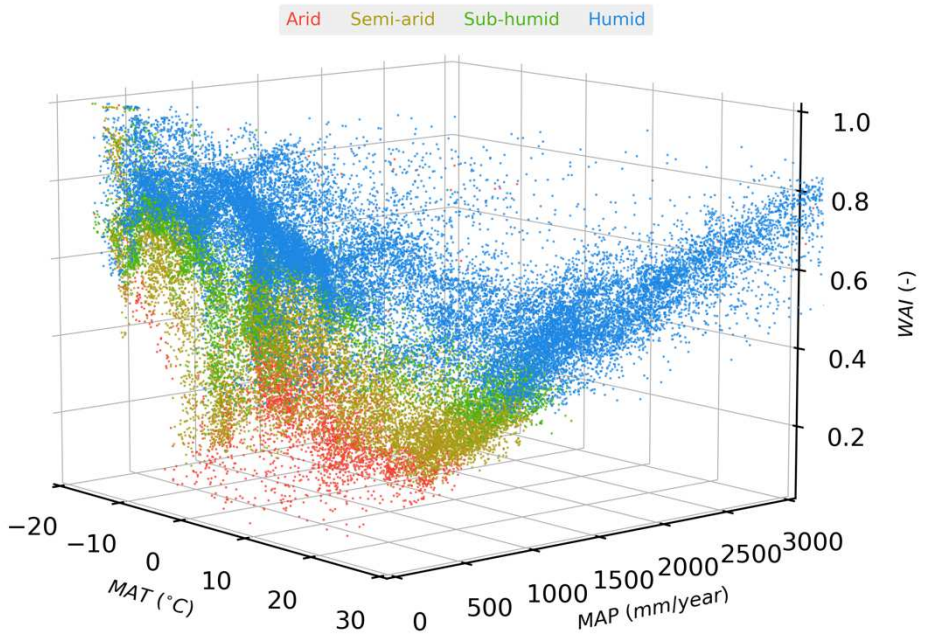

Figure A2. Covariation of climate and moisture. Mean annual precipitation (MAP, mm/year), mean annual temperature (MAT, °C) and water availability index (WAI, -) [50] are plotted in X, Y and Z axes respectively. The different colours indicate different climatic regions defined by aridity index, and individual 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).

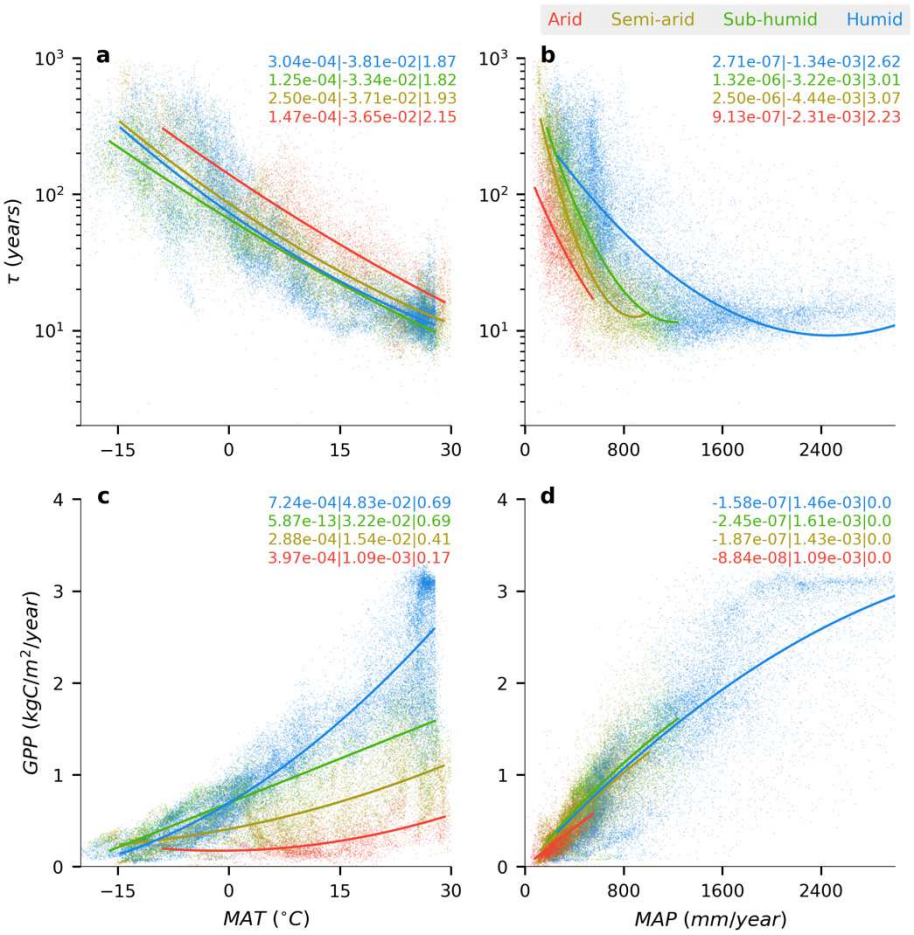

Figure A3. Same and **Figure 1**, but using the LANDGIS [69] instead of SoilGrids [47] dataset in the observation-based ensemble of soil organic carbon estimates.

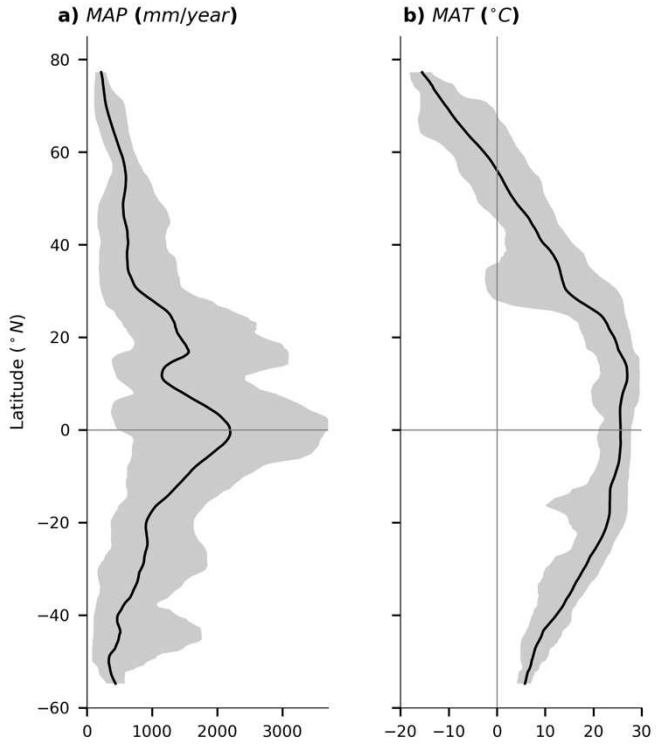

Figure A4. Zonal distributions of mean annual precipitation (mm/year) and mean air temperature (℃**).** The values for each latitude are calculated as mean within a moving window of 10 grid cells along the latitude (5°). The shaded region indicates the 5th and 95th percentiles within the latitudinal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

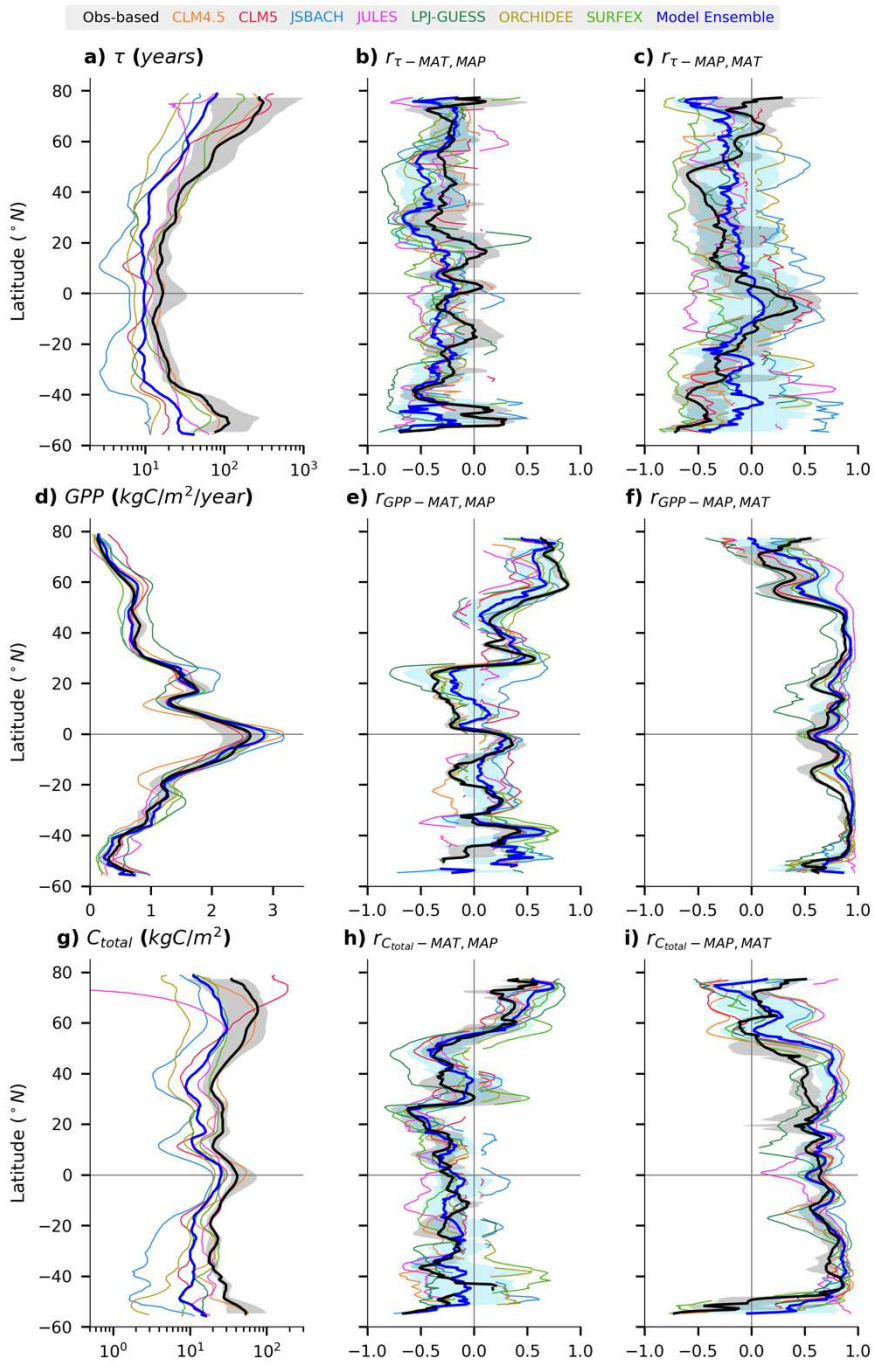

Figure A5. Same as **Figure 3** but using the average temperature that is corrected for low winter temperature below freezing point. For the correction, the temperature of months below 0°C are set at 0°C before calculating the mean annual temperature, which results in an estimate that only includes physiologically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

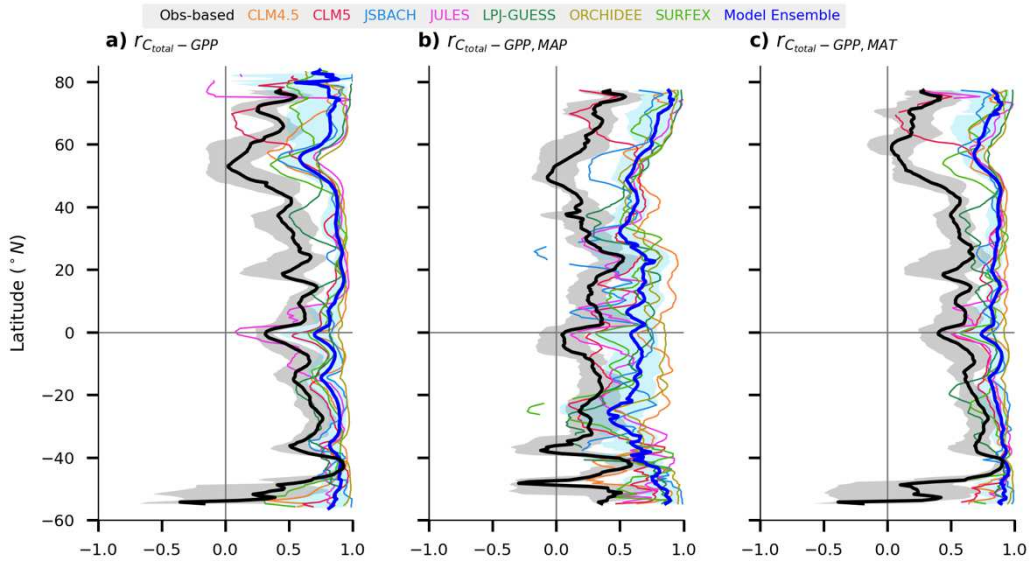

Figure A6. Zonal correlation between total ecosystem carbon stock (C_{total}, kgC/m^2) and gross primary productivity (*GPP*, kgC/m^2/year). In a) correlation between C_{total} and *GPP*, b) controlled for precipitation, and c) controlled for temperature are presented. The Pearson's correlation coefficient 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 5th and 95th percentiles.

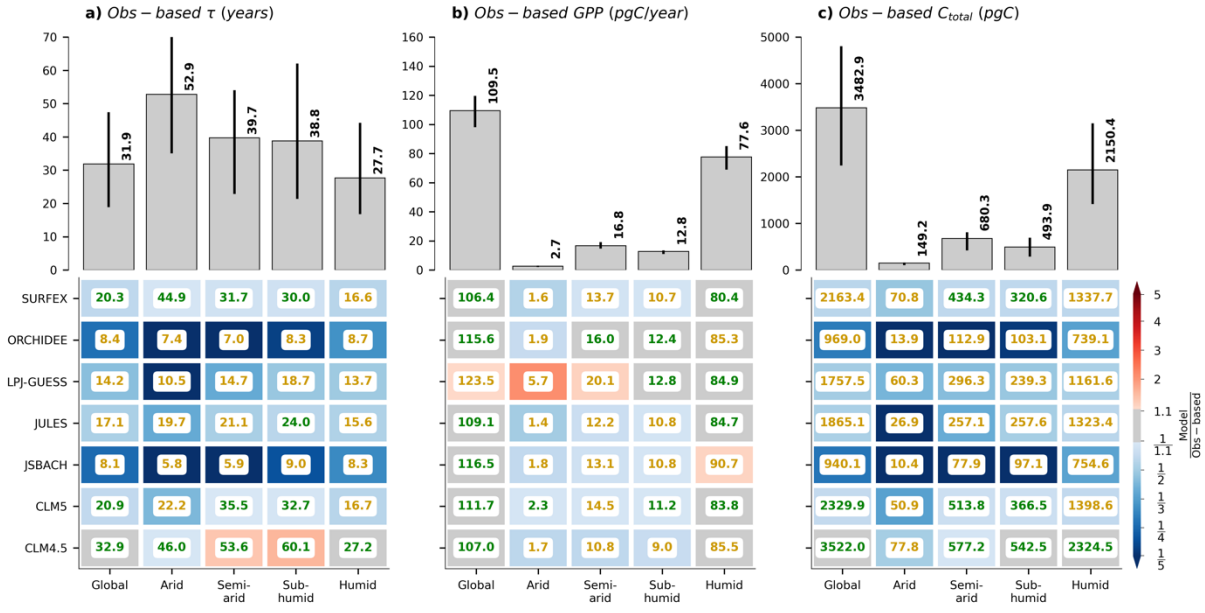

Figure A7. Evaluation of a) ecosystem turnover time (τ, years), b) gross primary productivity (*GPP*, pgC/year), and c) total carbon storage (C_{total}, pgC) over different climatic regions. In the bar charts, the observation and their uncertainties (5th and 95th percentiles) are plotted. In the colourmap 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 τ 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.

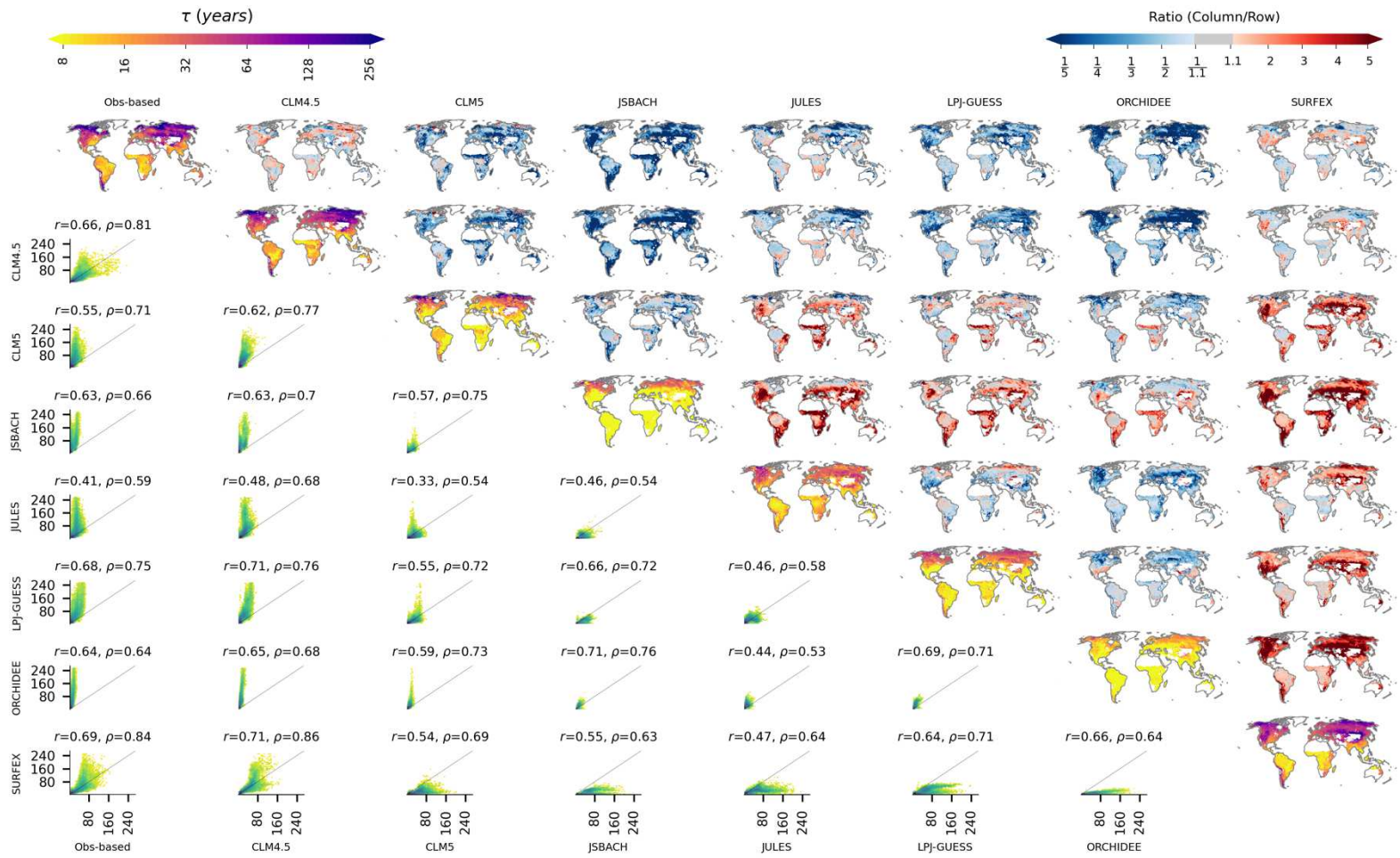

Figure A8. Comparison of the global distribution of ecosystem turnover times of carbon (τ, years) from observation-based estimate and CRESCENDO model simulations. Along the diagonal, the maps of τ from observation and model simulations are plotted. Above the diagonal, the biases (ratio of observation/model in 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 ρ indicate the spatial Pearson's and Spearman's correlation coefficient, respectively.

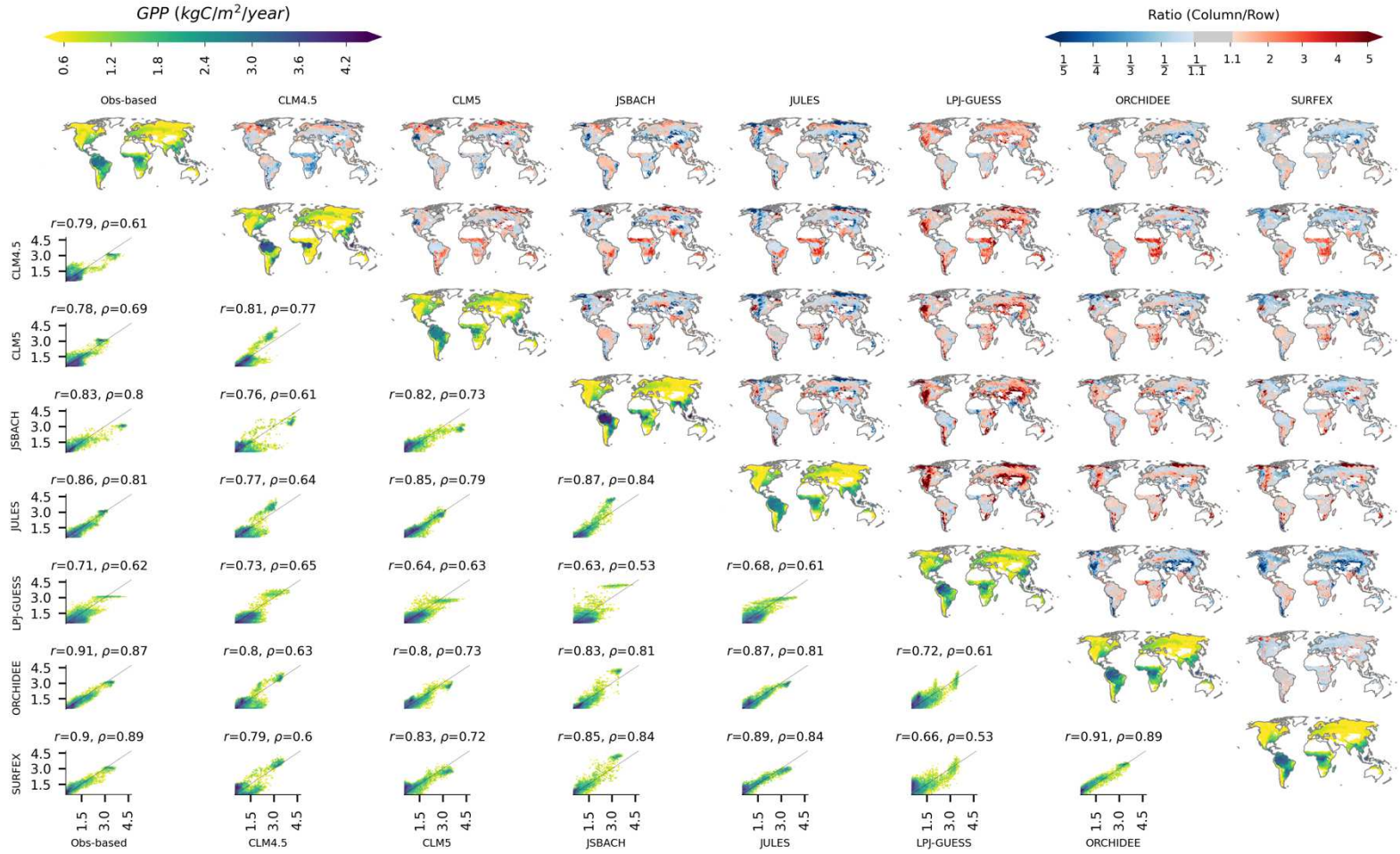

Figure A9. Comparison of the global distribution of gross primary productivity (*GPP*, kgC/m^2/year) from observation-based estimate and CRESCENDO model simulations. Along the diagonal, the maps of *GPP* from observation and model simulations are plotted. Above the diagonal, the biases (ratio of observation/model 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 ρ indicate the spatial Pearson's and Spearman's correlation coefficient, respectively.

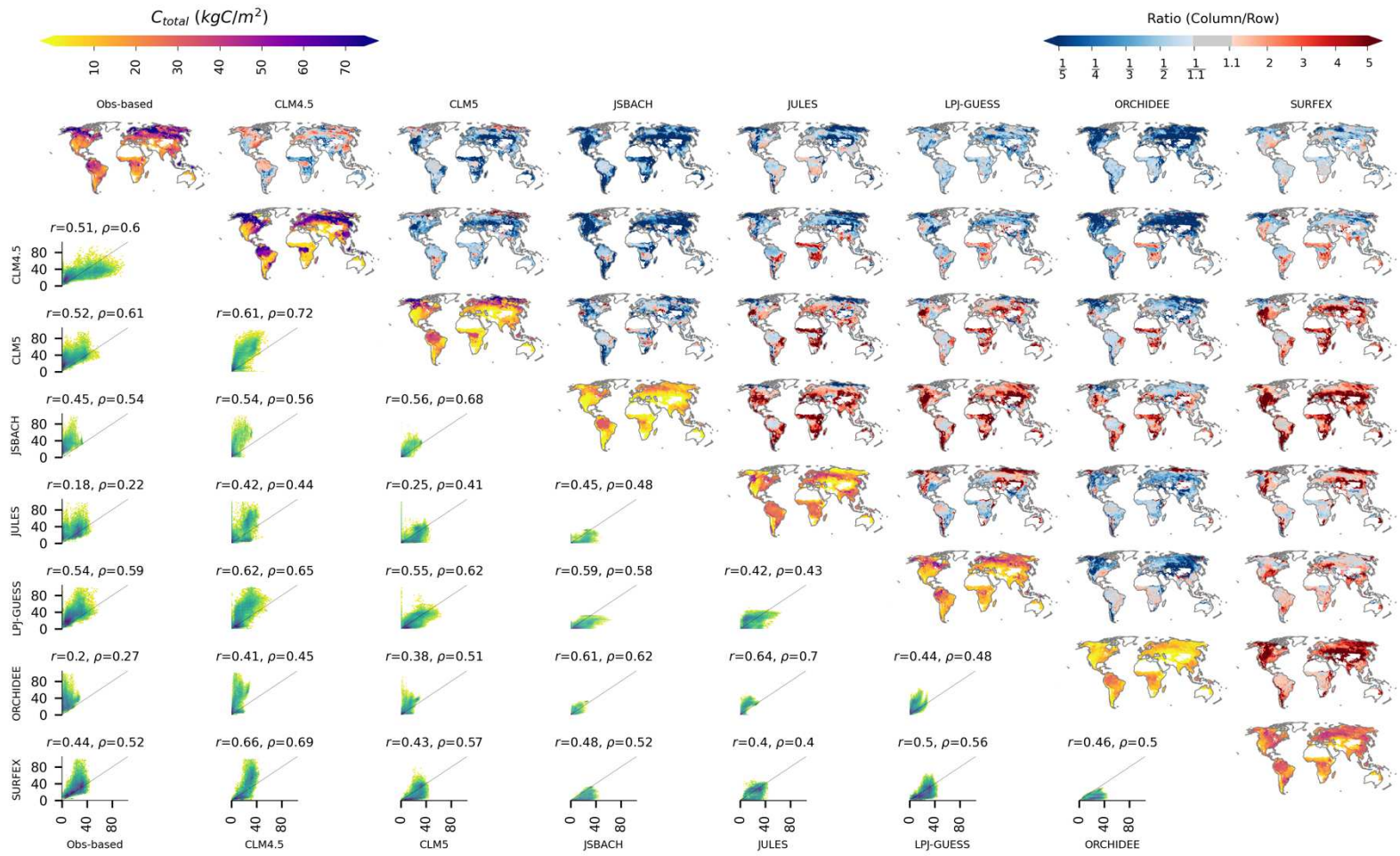

Figure A10. Comparison of the global distribution of total ecosystem carbon storage (C_{total}, kgC/m^2) from observation-based estimate and CRESCENDO model simulations. Along the diagonal, the maps of C_{total} from observation and model simulations are plotted. Above the diagonal, the biases (ratio 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 ρ indicate the spatial Pearson's and Spearman's correlation coefficient, respectively.

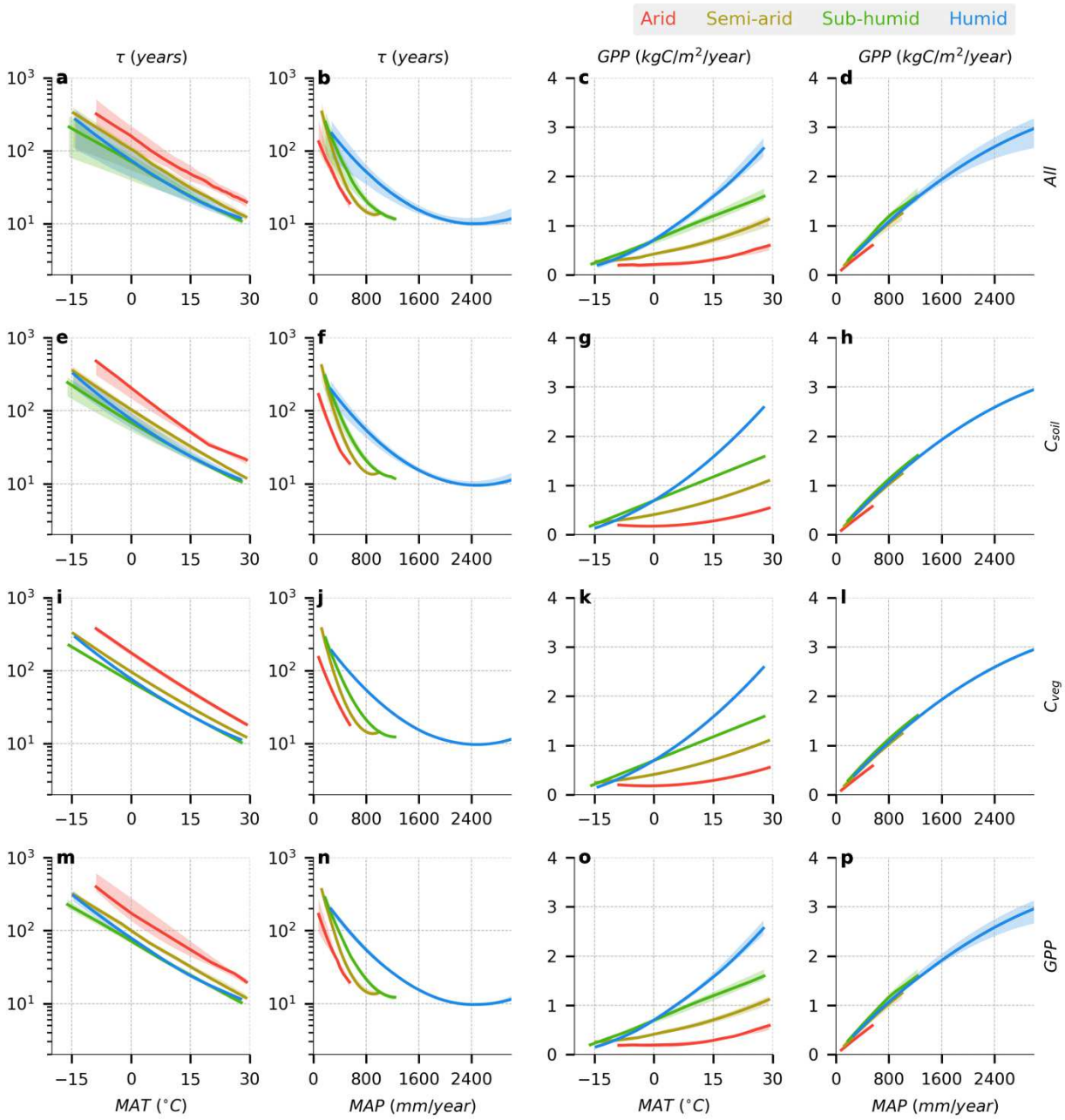

Figure A11. Same as **Figure 1** of the main text but using the state-of-the-art observation-based datasets of soil carbon (cSoil) and vegetation carbon (cVeg) stocks, and gross primary productivity (*GPP*). In total, 3 different soil carbon datasets, 4 different vegetation carbon datasets, and 6 different *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 τ in the first and second columns.

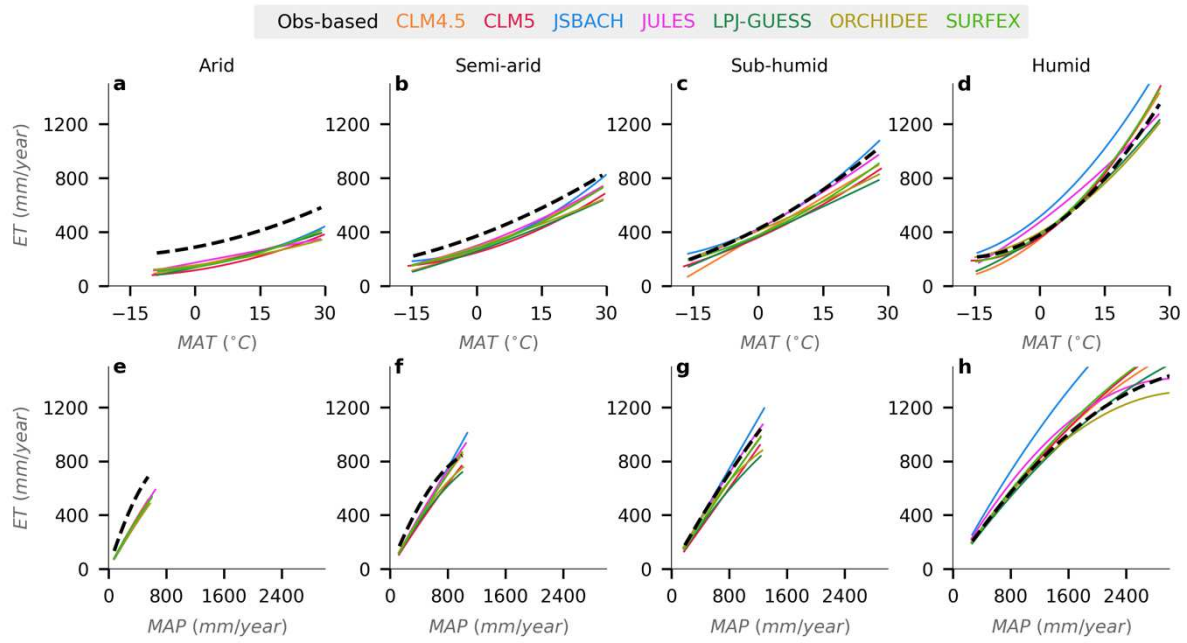

Figure A12. Evaluation of sensitivities of evapotranspiration to climate in CRESCENDO models. The relationships between total evapotranspiration (ET, in mm/day) versus mean annual temperature (MAT) and precipitation (MAP) are presented in the first and second column, respectively. Along the row, the comparisons for different models against the observation [50] 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.

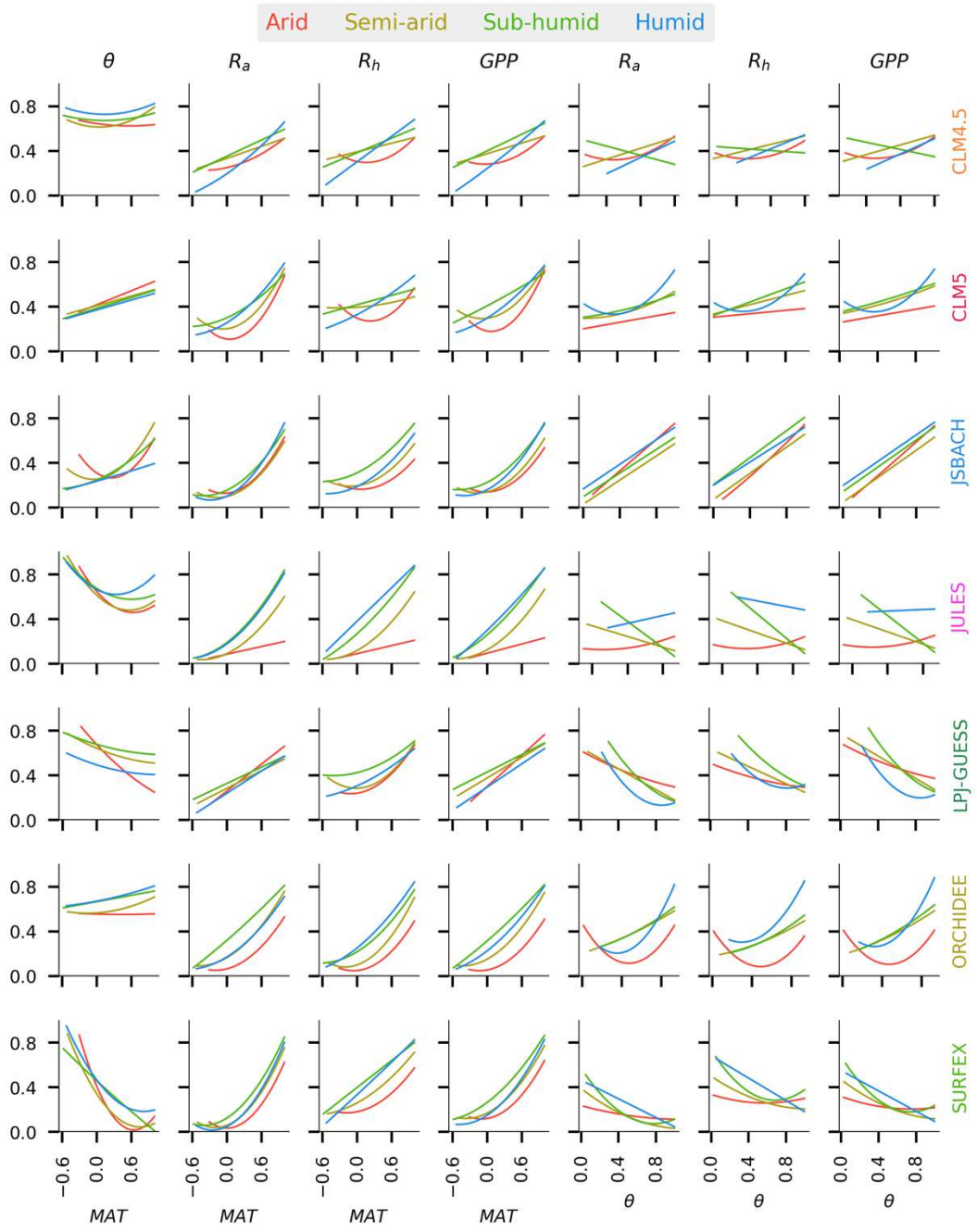

Figure A13. Spatial sensitivities of climatological total soil moisture state and carbon fluxes in CRESCENDO model simulations. In the first column, the relationship between the spatial variability of mean annual temperature (*MAT*) and total soil moisture (θ) is presented. In the second, third and fourth 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 θ. 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 models against the observation are presented. For a consistent comparison across different variables and differences in sizes of soil column in different models, all variables are scaled by the 98th percentile within each climate region so that the spatial variability is roughly normalized between 0 and 1.

References:

1. Taylor, P. G. *et al.* Temperature and rainfall interact to control carbon cycling in tropical forests. *Ecol. Lett.* **20**, 779–788 (2017).
2. 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).
3. Friedlingstein, P. *et al.* Uncertainties in CMIP5 Climate Projections due to Carbon Cycle Feedbacks. *J. Clim.* **27**, 511–526 (2014).
4. 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).
5. 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).
6. Heimann, M. & Reichstein, M. Terrestrial ecosystem carbon dynamics and climate feedbacks. *Nature* **451**, 289–92 (2008).
7. 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).
8. 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 U A* **113**, 1285–90 (2016).
9. 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).
10. 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).
12. 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).
13. Booth, B. B. B. *et al.* High sensitivity of future global warming to land carbon cycle processes. *Environ. Res. Lett.* **7**, 024002 (2012).

14. 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).
16. 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).
18. Carvalhais, N. *et al.* Global covariation of carbon turnover times with climate in terrestrial ecosystems. *Nature* **514**, 213–7 (2014).
19. 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).
20. 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.
21. 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).
22. Jung, M. *et al.* Compensatory water effects link yearly global land CO2 sink changes to temperature. *Nature* **541**, 516–520 (2017).
23. 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).
25. Jiang, L. *et al.* Scale-Dependent Performance of CMIP5 Earth System Models in Simulating Terrestrial Vegetation Carbon*. *J. Clim.* **28**, 5217–5232 (2015).
26. 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).
28. 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).
30. 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).
32. 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).
33. 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).
34. 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).
35. 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).
36. 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).
37. 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).
38. 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).
39. 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).
40. 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).
42. 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).
45. 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).
46. Hengl, T. *et al.* SoilGrids250m: Global gridded soil information based on machine learning. *PLOS ONE* **12**, e0169748 (2017).
47. 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).
48. 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).
50. Jung, M. *et al.* The FLUXCOM ensemble of global land-atmosphere energy fluxes. *Sci. Data* **6**, 74 (2019).
51. 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.
52. 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).

54. 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).
56. 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).
57. 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).
58. 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).
59. Oleson, K. *et al.* Technical description of version 4.5 of the Community Land Model (CLM). (2013) doi:10.5065/D6RR1W7M.
60. 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).
61. 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.
62. 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).
63. Tuomi, M. *et al.* Leaf litter decomposition—Estimates of global variability based on Yasso07 model. *Ecol. Model.* **220**, 3362–3371 (2009).
64. 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).
66. 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.