

¹ **Investigating Variability in the Intensity, Direction,**

² **and Spatial Distribution of Carbon Cycle Extremes**

³ **and Attribution to Climate Drivers Using**

⁴ **Observations and CMIP6 Earth System Models**

⁵ **Bharat Sharma^{1,2,*}, Jitendra Kumar³, Auroop R. Ganguly¹, and Forrest M. Hoffman^{2,4}**

⁶ ¹Sustainability and Data Sciences Laboratory, Department of Civil and Environmental Engineering, Northeastern
⁷ University, Boston, Massachusetts, USA

⁸ ²Computational Sciences & Engineering Division and the Climate Change Science Institute, Oak Ridge National
⁹ Laboratory, Oak Ridge, Tennessee, USA

¹⁰ ³Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

¹¹ ⁴Department of Civil and Environmental Engineering, University of Tennessee, Knoxville, Tennessee, USA

¹² ²Affiliation, department, city, postcode, country

¹³ *sharma.bha@northeastern.edu

¹⁴ **ABSTRACT**

Frequent climate extremes such as droughts, heatwaves, and fires driven by climate change have a potential to alter carbon cycle and gross primary productivity (GPP).

Understanding how extreme anomalies and interannual variability (IAV) in GPP are driven by climate change will help us forecast better model estimates of total carbon uptake.

Observations and models show that changes in the magnitude of GPP extremes are explained by IAV and detrended GPP anomalies, while total GPP and seasonality has negligible effect.

Both agree on water limitation being the dominant driver of negative GPP extremes.

Most climate models estimate that negative GPP extremes will increase at a rate larger than GPP, suggesting a relative weakening of total carbon uptake over time.

The disagreements in the characteristics of the GPP and extremes between observations and the models are significant and require an effort from the scientific community to address these uncertainties.

¹⁶ Please note: Abbreviations should be introduced at the first mention in the main text – no abbreviations lists. Suggested

¹⁷ structure of main text (not enforced) is provided below.

¹⁸ **Key Points**

- ¹⁹ • Agreement among models and observations on mean GPP is good and on extremes in GPP is poor.
- ²⁰ • The consistency among regions with high interannual variability in GPP and negative extremes in GPP is high.
- ²¹ • The biggest driver of negative extremes in GPP was dry climatic conditions.

²² **Introduction**

²³ Terrestrial ecosystems have sequestered a quarter of the anthropogenic carbon emissions over the last decade¹. However,
²⁴ the variability in the magnitude of carbon uptake is large^{2,3} and influenced by vegetation type, spatial location, and climate
²⁵ drivers^{4,5}. Climate extremes such as droughts, heatwaves, and fires have the potential to alter the carbon cycle by reducing
²⁶ the gross primary productivity (GPP), which is the total photosynthetic uptake of carbon⁶. Both the observations and models
²⁷ indicate that climate extremes are expected to increase with climate change. The increasing frequency and duration of climate
²⁸ extremes such as heatwaves and droughts and their compound occurrences could have a larger impact on carbon uptake than sum of
²⁹ individual impact of climate extremes⁷. Climate extremes increase the variability of the magnitude, duration, spatial distribution
³⁰ of carbon cycle anomalies, which is defined as the deviation from the expected carbon uptake. The key to understanding the
³¹ adverse shifts in carbon uptake, i.e. carbon cycle extremes defined as very large GPP anomalies, is essential to estimate the
³² variability and direction of climate-carbon feedback.

³³ With the availability of large amount ("big data") of observation, we are better equipped than before to test the agreement
³⁴ among observations flux products and with Earth System Models (ESMs). Long-term goals in global ecology include
³⁵ understanding the multifaceted functions of terrestrial ecosystems in the Earth system⁸. Ecological-level quantities such
³⁶ as GPP are derived from flux and biometric observations and often comparable to ecophysiological leaf-level quantities such as
³⁷ water-use efficiency, leaf conductance, light-use efficiency, or light-saturated photosynthetic CO₂ uptake⁸. Using statistical
³⁸ machine-learning methods using flux observations with remote sensing and meteorological drivers has allowed us to infer
³⁹ site-to-global fluxes of ecosystem functions such as GPP⁸. These fields from up-scaled observations are used as benchmark
⁴⁰ data for evaluating the capability of ESMs in representing same carbon-cycle metrics⁹. These benchmarking efforts are inclined
⁴¹ for evaluating the mean state of carbon fluxes and less importance is given to the tails of the probability distribution of the
⁴² carbon fluxes. However, some tail events have the potential to change the state of climate-carbon feedbacks for much longer
⁴³ time e.g., a large drought or heat wave or both can lead to drastic decline in GPP of a certain region and will further lead to
⁴⁴ physiological and biogeochemical changes. In this paper, we have focused on analysing the agreement of the characteristics of
⁴⁵ GPP, interannual annual variability (IAV) of GPP and GPP extremes among up-scaled gridded observation products and ESMs.

⁴⁶ The objectives of this study include, firstly, the quantification of the uncertainty within and among observations and models
⁴⁷ on the magnitude, frequency, and spatial distribution of GPP and IAV of GPP and negative extremes in GPP. Secondly, analysing

48 the expected changes in the above mentioned characteristics of GPP extremes in the future SSP585 scenario. Lastly, attribution
49 of negative carbon cycle extremes to individual and compound climate driver anomalies.

50 Most results are presented in the notation $\mu \pm \sigma$, where μ and σ denotes mean and standard deviation, respectively.

51 Results

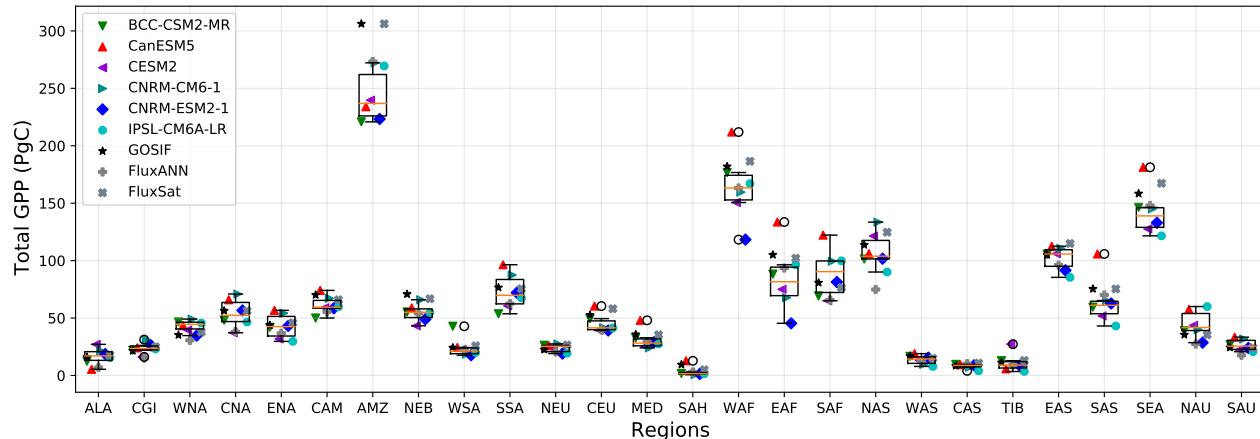
52 Comparison of Characteristics of GPP Extremes Among Observations and Models

53 During 2001–13, the distribution of total regional GPP among observations and models were consistent across most of the
54 SREX regions (Figure 1a) with exception in the tropical regions, especially Amazon (AMZ), where GPP from observations
55 was larger than models. The strong tropical carbon sink estimated by upscalled observation data is likely due to a combination
56 of systematic biases in underlying eddy covariance data in tall tropical forests¹⁰. Among CMIP6 models, CanESM5 estimated
57 the largest GPP across most regions. The global integrated timeseries of GPP was consistent among observations and models.
58 The mean and standard deviation of GPP estimates from observations, 127 ± 10.8 PgC/year, were within the range of GPP
59 from models, 128.5 ± 15.2 PgC/year (Figure [Extended Data1](#)).

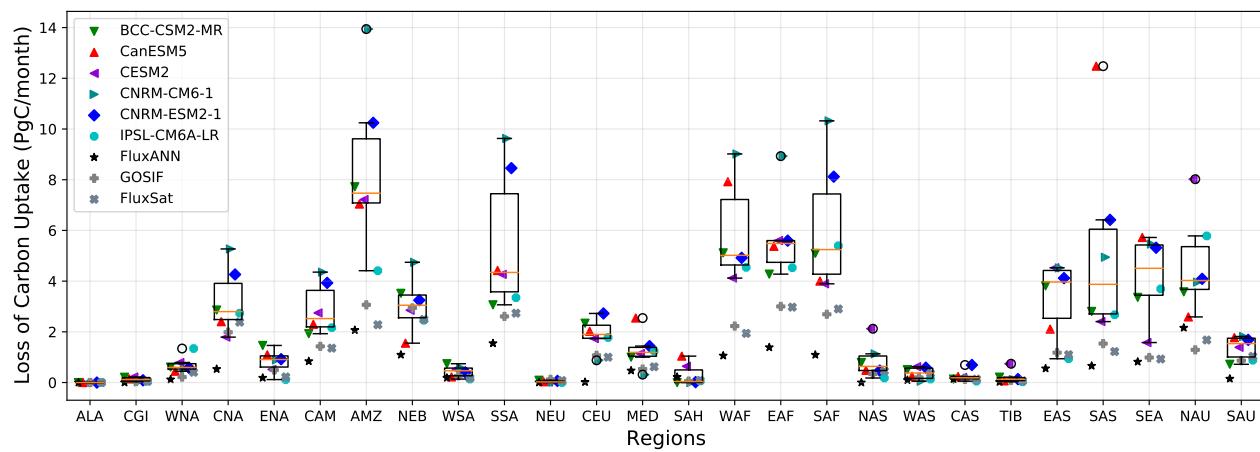
60 The regional magnitudes of negative carbon cycle extremes among models were about three times larger than observations
61 despite high consistency in regional and total GPP (Figure 1b). The difference in magnitude of losses in carbon uptakes in the
62 regions of AMZ, Southeastern South America (SSA), Southeast Asia (SEA), and East, West and South Africa (EAS, WAF,
63 SAF) between multi-model mean and multi-observation ranges from -5 to -8 PgC/month with large variability in the regions
64 of SEA, SSA, and Northeast Brazil (NEB). As a result, the disagreement in the magnitude of losses of carbon uptake (L_{GPP})
65 among models was much larger than observations (Figure 1b). The variation in the magnitude of L_{GPP} and total GPP (and
66 $L_{GPP} : GPP$) was large which suggests that L_{GPP} could increase or decrease disproportionately to the total regional GPP. For
67 example, the region of SSA and Northern Asia (NAS), with mean GPP of, show L_{GPP} of which iswas larger and
68 smaller relative to regional GPP, respectively.

69 The disagreement among models on the magnitude of negative carbon cycle extremes was much higher than observations
70 (Figure [Extended Data2](#)). The region of SAS showed highest standard deviation of total GPP of 3.5 PgC/month followed by
71 AMZ (2.3 PgC/month) and SSA (2.2 PgC/month). The difference between magnitude of L_{GPP} in models and observations
72 highlight, 24 out of 26 regions, estimate much larger losses in carbon uptake in models (Figure [Extended Data2e](#)). The biggest
73 disagreement was seen in tropical regions, especially in AMZ with -7 PgC/month.

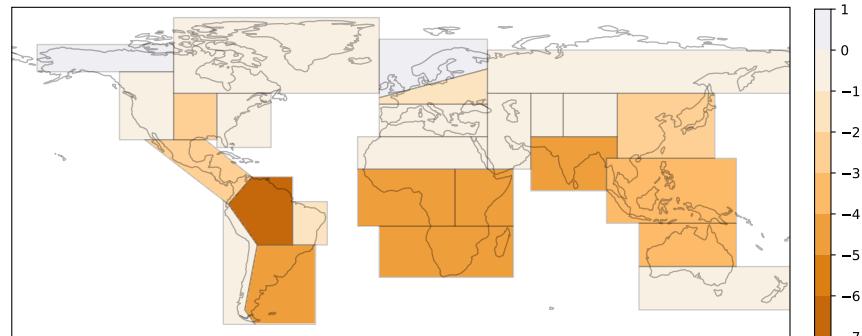
74 The agreement between GOSIF and FluxSat on the regional distribution of frequency of negative carbon cycle extremes
75 was higher than FluxANN (Figure [Extended Data3](#)). While the number of extremes estimated using FluxANN were larger in
76 tropics than mid-latitudes, GOSIF and FluxSat estimated the reverse. Among models, the agreement on frequency of negative
77 carbon cycle extremes across regions was high. **Add mean of std across regions** The negative extremes from the CESM2 and
78 BCC-CSM2-MR were higher in northern mid-latitudes, NAS, Canada and Greenland (CGI), and fewer in tropics compared to
79 the rest of the models (Figure [Extended Data3](#)). However, the agreement on the mean frequency of negative carbon extremes



(a) Obs Models GPP comparison SREX



(b) Loss of carbon uptake



(c) Regional differences in loss of carbon uptake, $\text{Models}_{\text{mean}} - \text{Observations}_{\text{mean}}$

Figure 1. Spatial distribution of total regional GPP and loss of carbon uptake from observation data and model simulations for 2001–13. (a) Total regional GPP are calculated by integrating the GPP for the period 2001–13. The boxplots are created using only the total GPP values of only models, shown in colored markers, for each SREX region (Table [Extended Data1](#)). Total GPP from the observation data are plotted later (i.e. not considered for creating boxplots) and shown in different shades of grey. (b) Total loss of carbon uptake (PgC/month) during top 50 percent of the negative extremes in GPP for the period 2001–13. The boxplots are created using the total carbon loss values of only models, observations are plotted later. (c) Regional difference, $\text{Models}_{\text{mean}} - \text{Observations}_{\text{mean}}$, in mean total loss of carbon uptake (PgC/month) among models and observations as shown in (b). The color range is from −9 to 9 PgC/month, the color bar ticker is shown for the range of regional data, for more details see Figure [Extended Data2](#).

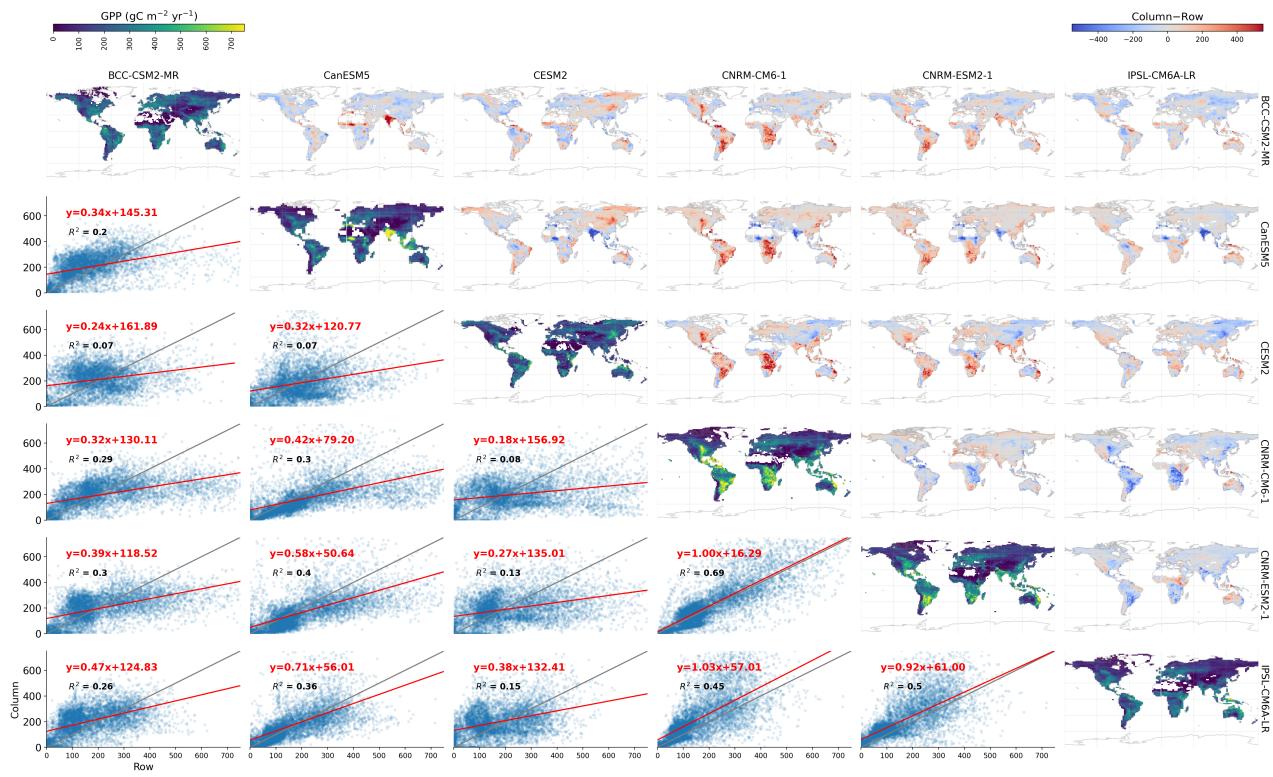


Figure 2. Comparison of interannual variability (IAV) of GPP, at 0.5° spatial resolution for the period 2001–2013, of 6 CMIP6 models. From left to right and top to bottom the order of models are BCC-CSM2-MR, CanESM5, CESM2, CNRM-CM6-1, CNRM-ESM2-1, and IPSL-CM6A-LR. The diagonal maps show the IAV GPP of every model. Maps above the diagonal show the difference of the IAV GPP of column dataset – row dataset. The plots below the diagonal show the point density in blue and 1:1 regression line in grey. Red line and equation represent the best fit line from total least-squares regression.

among observations and models is high except for NAS and CGI regions, where the standard deviation of frequency of extremes of observations is **two times** larger than models. The observations (SEA, SAF, WAF, EAS, and NAU) and models (SEA, SAF, WAF, NAS, and EAS) agree on four of top five regions with largest frequency of negative extremes. The regions of SEA, SAF, and WAF represented about 14%, 10%, and 6% of total negative carbon cycle extremes among observations and models, respectively.

To understand the characteristics and distribution of negative carbon cycle extremes among observations and models, we compared the interannual variability (IAV) in GPP for each data product. The possible reason of larger magnitude of negative carbon cycle extremes estimated from model outputs than observations are likely due to larger magnitude of IAV in GPP in models (Figure 2) than observations (Figure 3). The regions with large IAV in GPP are also associated with high L_{GPP} and vice versa. For example, CanESM5 estimated the largest L_{GPP} in the region of Southern Asia (SAS, Figure 1b) and largest IAV in GPP in SAS (Figure 2). Similarly, CNRM-CM6-1 has the largest L_{GPP} and largest IAV in GPP in the regions of Amazon (AMZ), West Africa (WAF), East Africa (EAF) and South Africa (SAF). CESM2 has the largest losses in carbon uptake and GPP IAV in Northern Australia (NAU) and lowest in Central North America (CNA), WAF and SAF. The differences in the magnitude of IAV in GPP across regions and models were large and were likely the reason for large uncertainty in the

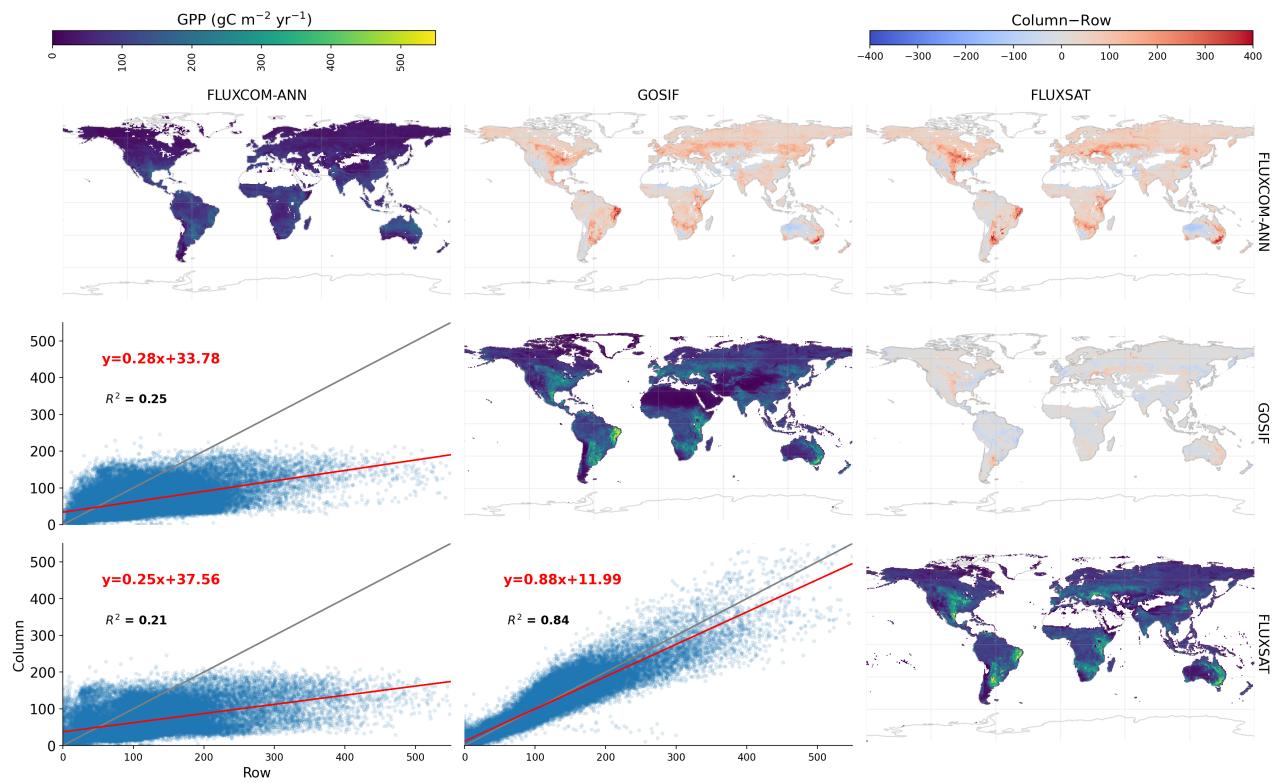


Figure 3. Comparison of interannual variability (IAV) of GPP, at 0.5° spatial resolution for the period 2001–2013, of 3 observation based upscaled GPP products. From left to right and top to bottom the order of models are FLUXCOM-ANN (FluxANN), GOSIF, and FLUXSAT (FluxSat). The diagonal maps show the IAV GPP of every dataset. Maps above the diagonal show the difference of the GPP of column dataset – row dataset. The plots below the diagonal show the point density in blue and 1:1 regression line in grey. Red line and equation represent the best fit line from total least-squares regression.

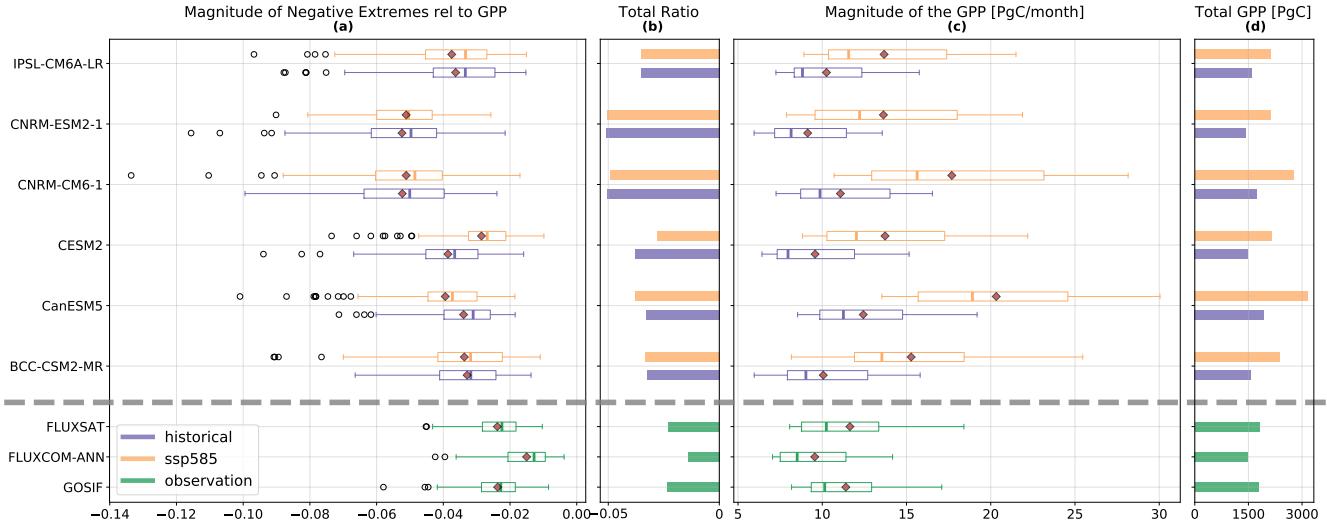


Figure 4. Change in the magnitude of negative carbon cycle extremes relative to GPP. The green, purple, and orange colors represent the statistics explained below for observation data (2001–13), historical model simulations (2001–13), and SSP585 model simulations (2081–93), respectively. Box plot of ratios of the monthly magnitude of negative carbon cycle extremes to total monthly GPP (a) and total monthly GPP (c) of observation data and model simulations. Ratio of total negative magnitude of negative extremes to to total GPP (b) and total GPP (d). The grey dashed line separates observations and models. Ratio of magnitude of negative extremes to total GPP is referred as $L : GPP$ (ratio) in the paper, which means loss in expected carbon uptake to total GPP.

characteristics of global and regional negative carbon cycle extremes. The disagreements among models on magnitude and spatial distribution are very high (Figure 2). Among the six models, BCC-CSM2-MR, CanESM5, and CESM2 showed the least consistent ($R^2 \leq 0.50$) distribution of IAV in GPP.

Similar patterns were seen in observation data products, where the losses in carbon uptake and GPP IAV are larger in FluxSat and GOSIF than FluxANN (Figure 3). The IAV in GPP in FluxANN was smaller than FluxSat and GOSIF, both of which possessed high consistency in comparing GPP IAV. The losses in carbon uptake among observations were also smaller than models likely due to low IAV in GPP among observations than models. Though the underestimation of IAV in GPP by observations is not fully clear¹⁰, the likely reason is the low IAV simulated by machine learning models. Latter is likely driven by low temporal resolution of observations (8day for MODIS), low sample size of eddy covariance (EC) sites in tropics and associated poor signal to noise, systematic errors of EC measurement, and errors of partitioning NEE to TER and GPP¹⁰. Improving the quality, standardization, and coverage of observation measurements will result in reduced uncertainties in the estimated carbon fluxes and better benchmark datasets for models' training and testing.

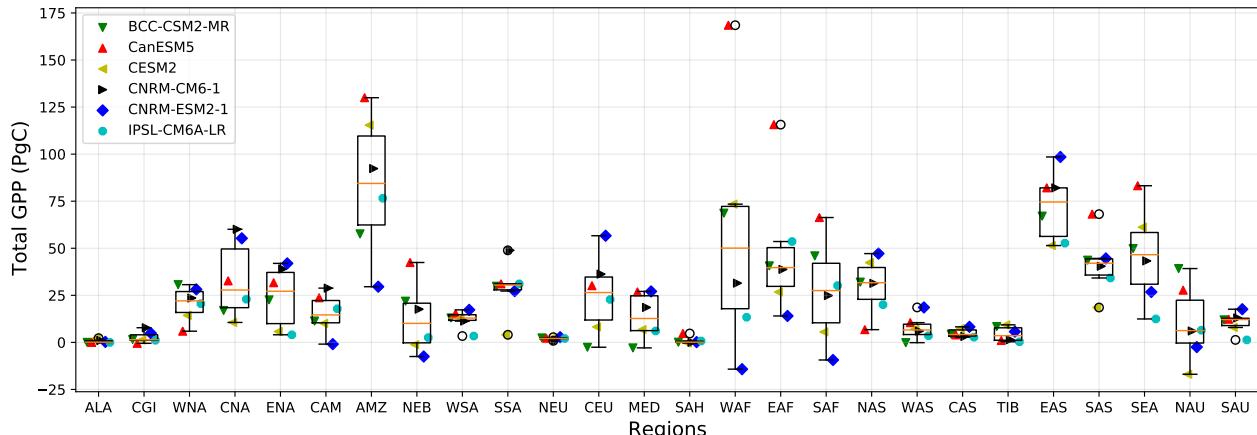
106 Changing Characteristics of GPP Extremes over time

107 All CMIP6 models showed an increase in total global GPP from the period 2001–13 (historical) to 2081–93 (ssp585), driven by
 108 increasing carbon fertilization and water use efficiency^{11–14}. The largest and smallest increase in GPP were shown by CanESM5
 109 (63% or 1231 PgC) and IPSL-CM6A-LR (33% or 533 PgC), respectively (Figure 4d). The overall increase in GPP is also

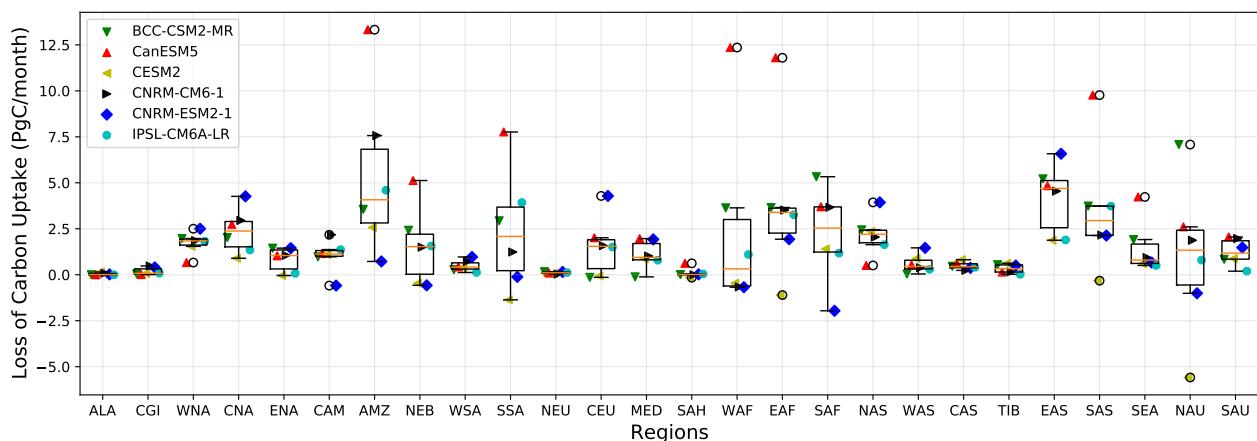
associated with large uncertainty (Figure 4c) and corresponding increase in IAV in GPP. Thus resulting in larger losses during negative carbon cycle extremes over time (Figure 5b and c) among all models and most regions. For the period of 2001–13, all observational datasets, except GOSIF and FluxSat, do not show any increase in total GPP (Figure Extended Data1). The cause of no-trend in GPP in FLUXCOM data products (and others) is due to lack of any treatment of CO₂ fertilization effect¹⁰. While the total GPP for the period 2001–13 for the models and observations are consistent (Figure 4d), the magnitude of negative carbon cycle extremes relative to total GPP is larger for models than observations (Figure 4b). This difference in the magnitude of negative carbon cycle extremes in models and observations is due to higher IAV in GPP in models than observations, which drives larger losses in carbon uptake during negative carbon cycle extremes in models than observations. The losses in carbon uptake (L_{GPP}) to total GPP, $L_{GPP}:GPP$, were associated with large variation for every dataset which highlight that changes in the monthly magnitude of negative extremes to total GPP are not proportional and may also be associated with larger carbon losses (shown in circle markers in Figure 4a).

The CNRM models have the largest $L_{GPP}:GPP$ for both the historical and ssp585 periods, indicating highest losses in carbon uptake relative to total GPP. CNRM models also have the largest IAV in GPP (Figure 2) which strongly correlates with large anomalies in GPP and losses in carbon uptake. Despite varied increases in GPP among models, the $L_{GPP}:GPP$ has been consistent for both periods which means that the negative carbon cycle extremes are expected to increase in proportion to the increase in GPP over time. The highest increase in $L_{GPP}:GPP$ was shown by CanESM5 followed by BCC-CSM2-MR which depicts that the negative carbon cycle extremes will rise faster than the total increase in GPP. The rate of $L_{GPP}:GPP$ of CanESM5 increased especially after 1950 and continued to further increase over time (see Figure Extended Data9 for details). The decreasing $L_{GPP}:GPP$ in CESM2 suggests the the relative magnitude of negative carbon extremes to total GPP will decline over time. However, the magnitude of total negative carbon cycle extremes in CESM2 are expected to increase over time but at a slower rate than increase in GPP (Figure Extended Data8). Five out of six models depict that the changes in $L_{GPP}:GPP$ from 2001–13 to 2081–93 are expected to remain consistent or increase with time.

The regional changes in GPP (Figure 5a) indicate that most regions are expected to witness an increase in total GPP from 2001–13 to 2081–93. The largest increase in GPP, based on multimodel mean GPP, is expected to occur in Amazon (AMZ) of about 85 ± 37 PgC, followed by East Asia (EAS) of about 75 ± 19 PgC. However, the uncertainty of GPP and IAV in GPP changes are large among the models. For the region of West Africa (WAF), CanESM5 estimates an increase of 170 PgC and CNRM-ESM2-1 a decrease of 20 PgC of total GPP (and standard deviation of 64 PgC). Similarly, the magnitude of negative carbon cycle extremes show large uncertainty and losses in carbon uptake show an increase overtime. Most regions and models that indicate an increase in losses in expected carbon uptake, predict GPP to increase over time which could help in compensating for the overall loss of vegetation during negative carbon cycle extreme events. A decrease in GPP has an opposite effect, such that the negative extremes weaken over time which is likely due to reduced IAV in GPP with reduction in GPP (e.g. NAU and CESM2 or SAF and CNRM-ESM2-1). The multimodel mean magnitude of negative carbon cycle extreme predicts an increase at most regions (Figure 5b,c). However, the spatial distribution of the change in the mean multimodel



(a) Change in Magnitude of GPP (2081–93 minus 2001–13)



(b) Change in Magnitude of Negative Carbon Cycle Extremes (2081–93 minus 2001–13)

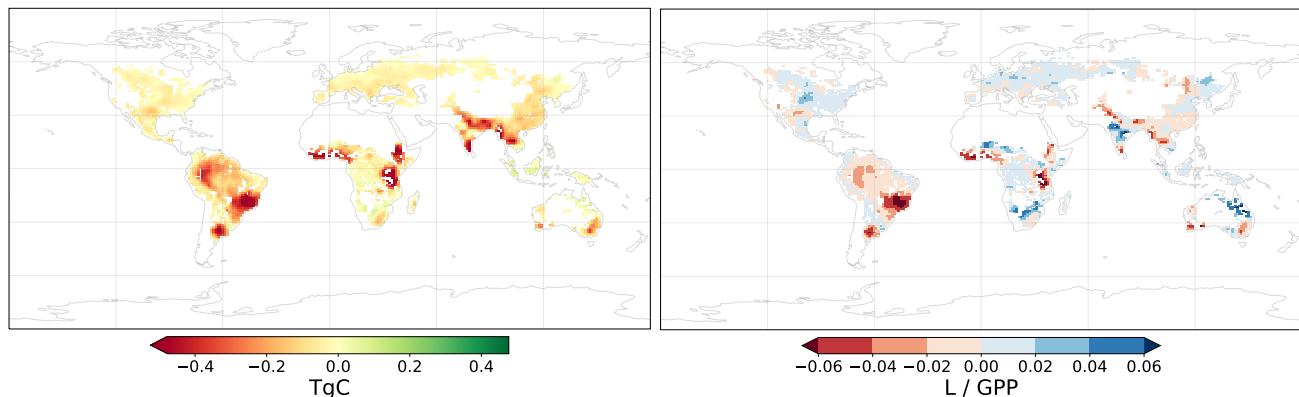


Figure 5. Changing magnitude of negative carbon cycle extremes and GPP from 2001–13 to 2081–93. (a) Total regional change in GPP is difference between total GPP of 2081–93 period and GPP of 2001–13 for each SREX region and model. (b) The total loss of carbon uptake (L_{GPP}) during top 50% of negative carbon cycle extremes is calculated for both periods. The regional change in L_{GPP} is the difference between L_{GPP} of 2081–93 and 2001–13, for each SREX region and models. (c) The spatial distribution of difference between the multimodel mean of L_{GPP} of 2081–93 and 2001–13 shows strengthening of negative carbon cycle extremes over time. (d) The spatial distribution of difference between the multimodel mean of $L_{GPP}:GPP$ of 2081–93 and 2001–13 shows a large variation of the strength of negative carbon cycle extremes relative to changing GPP.

143 $L_{GPP}:GPP$ varies, suggesting that relative to increase in GPP, the proportion of the magnitude of negative extreme may increase
144 or decrease over time (Figure 5d). Though the number of gridcells that show positive and negative $L_{GPP}:GPP$ are similar, but
145 total grid cells where $L_{GPP}:GPP < -0.04$ were more than twice the count of grid cells where $L_{GPP}:GPP > 0.04$. Thus, with
146 more regions that are expected to witness higher losses in carbon uptake indicate an overall strengthening of negative carbon
147 cycle extremes with respect to the change in GPP.

148 Figure 5: change in the magnitude of negative carbon extremes and total GPP

149 Attribution to Climate Drivers

150 Most negative carbon cycle extremes, 76% ($\pm 3\%$) and 56% ($\pm 5\%$), among GPP observational datasets were driven by *Dry*
151 events using GSWP3 and ERA5, respectively (Figure 6a, c). Compound *Hot & Dry* events drives 8% ($\pm 0.5\%$) and 19%
152 ($\pm 2.6\%$) of all negative carbon cycle extremes using GSWP3 and ERA5, respectively. Similarly, about 8.5% ($\pm 1.5\%$) and
153 12% ($\pm 1\%$) of all negative extremes were driven by *Hot* events using GSWP3 and ERA5, respectively. While the attributions
154 of negative carbon cycle extremes to GSWP3 and ERA5 were consistent, the models showed a large uncertainty among drivers.
155 The largest driver ($62\% \pm 23\%$) of negative carbon cycle extremes among models was *Dry* climatic condition, except for
156 CanESM5 where hot events caused most negative carbon cycle extremes (Figure 6e). *Hot* events and compound *Hot & Dry*
157 events caused 17% ($\pm 15\%$) and 13% ($\pm 13\%$) of all negative carbon cycle extremes.

158 The spatial distribution of climate drivers of negative carbon cycle extremes (Figure [Extended Data10](#)) depicts that *Dry*
159 and *Hot* events were dominant drivers across the globe. The regional distribution of gridcells from both observations and
160 models suggests that majority of negative carbon cycle extremes occurred in tropical regions of SEA and SAF and driven by
161 *Dry* events (Figure 6b,d, and f). Though, *Dry* events drive most negative carbon cycle extremes, the number of *Hot* events
162 were larger than *Dry* events in the latitudinal range of -23.27° to 23.27° for all except CNRM models. For example, the
163 spatial distribution of climate drivers for CESM2 (Figure [Extended Data10](#)) show a larger number of hot events in the region
164 of South Asia (SEA), Amazon (AMZ), North Eastern Brazil (NEB), and Central America (CAM) than dry events. The stark
165 differences in the regional distribution of magnitude (Figure 1b) and frequency (Figure 6b,d, and f) of negative carbon cycle
166 extremes indicate that either small number of extremes (e.g. AMZ, SSA) could cause large losses in most tropical regions or
167 many negative extremes in tropical regions or most negative extremes in tropics occur even during 2nd and 3rd quartile of
168 precipitation and temperature anomalies. The occurrence of large negative carbon cycle extremes during the non-extremes of
169 climate drivers is consistent with the earlier findings^{15,16} and such negative carbon extremes are likely driven by compound
170 effect of non-extreme climate drivers.

171 The terrestrial ecosystem has an ingrained capacity to buffer and push back the effects of climate change which results in
172 delayed response of climate extremes on vegetation productivity¹⁷. Using the CMIP6 models, we show the impact of no-lag
173 and 1 month lagged response of hot and dry events on negative carbon cycle extremes. First quartile of precipitation (*Dry-pr*)
174 and soil moisture (*Dry-sm*) anomalies were used to define dry events. The uncertainty of relative attribution to *Hot* and *Dry-sm*
175 events was less among the models than *Hot* and *Dry-pr* events (Figure [Extended Data11](#)). The percent of negative carbon

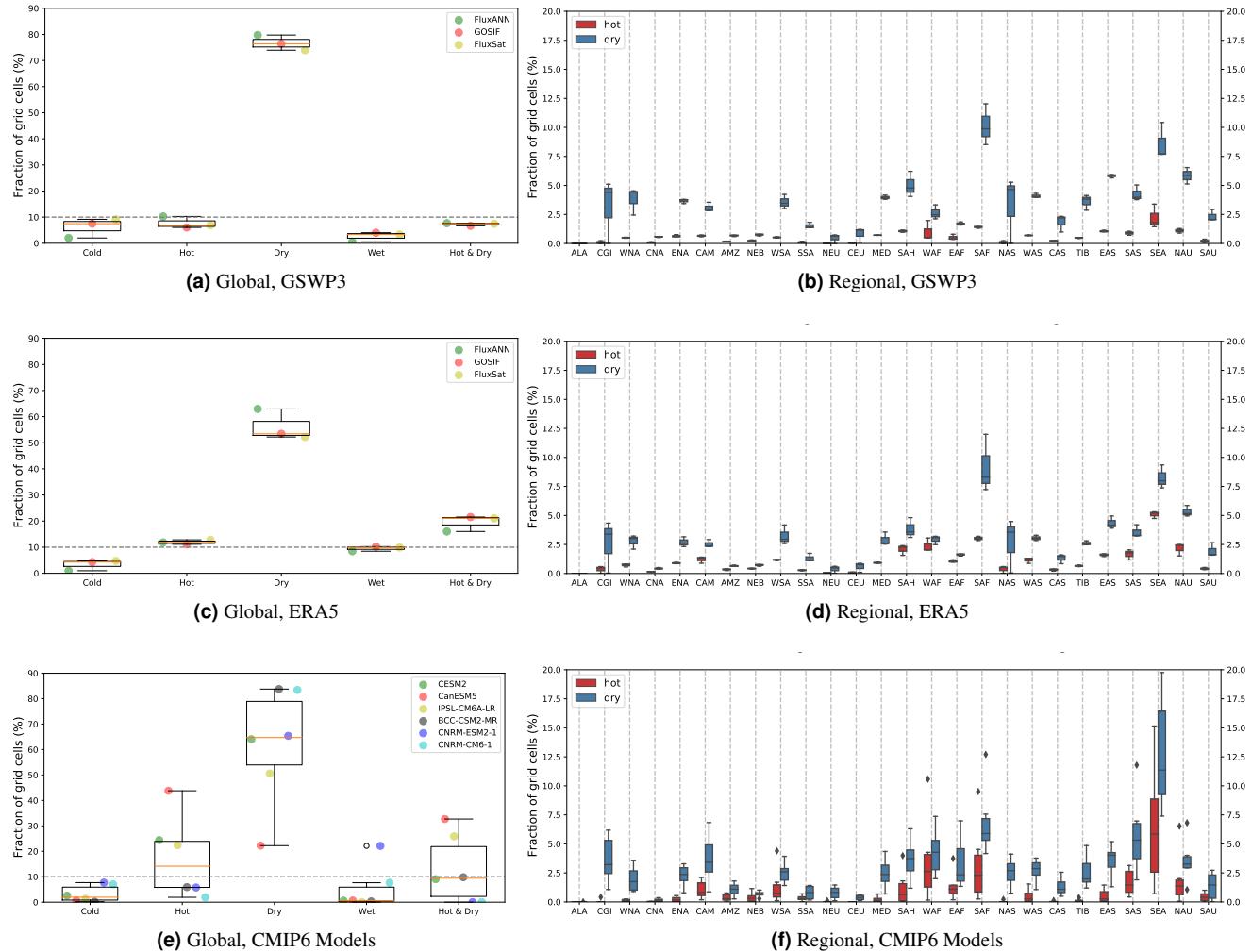


Figure 6. Global and regional distribution of climate drivers of negative carbon cycle extremes. The left panel (a, c, and e) shows fraction of grid cells where negative carbon cycle extremes are driven by exclusive individual and compound climate drivers aggregated globally. The right panel (b, d, and f) shows the fraction of grid cells driven either by hot (shown in red) or dry (shown in blue) events across 26 SREX regions. The attribution of negative carbon cycle extremes from observation flux data was performed with two observation based climate data, i.e. GSWP3 (a, b) and ERA5 (c, d). The attribution of negative carbon cycle extremes from CMIP6 models was performed with their own climate drivers (e, f). Cold and wet conditions rarely drive a negative carbon cycle extreme as they lie below the horizontal dashed line at 10%.

176 cycle extremes attributed to *Hot* and *Dry-pr* events at no-lag were 28.8% (\pm 17.4%) and 35.3% (\pm 21.7%), and at 1 month lag
177 were 17.5% (\pm 15.8%) and 61% (\pm 22.8%), which shows that lagged water scarcity caused higher losses in carbon uptake.
178 Total percent of negative carbon cycle extremes driven by *Hot*, *Dry-sm*, and compound *Hot & Dry-sm* were 8.8% (\pm 8.4%),
179 56% (\pm 12.6%), and 22.8% (\pm 16.6%) for no lag and 5.4% (\pm 4.5%), 63.8% (\pm 12.1%), and 18.2% (\pm 17.2%) for 1 month
180 lag, respectively. The percent of negative carbon extremes in CanESM5 attributed to *Dry-sm* were larger than both *Hot* and
181 *Dry-pr*. Consistent and increased attribution of negative extremes to decline in soil water content depicts that lack of soil
182 moisture impairs the vegetation productivity strongly than precipitation. Attribution of negative carbon cycle extremes to *Cold*
183 and *Wet* events were \leq 10% among both models and observations.

184 Discussion

185 The Discussion should be succinct and must not contain subheadings.

186 While global mean GPP values from observation estimates and model outputs are consistent, the disagreement on GPP at
187 spatial resolution is large, especially among models. The uncertainty in the GPP estimates among CMIP6 ESMs is about 1.5
188 times than of observations for the period 2001-13. The mean GPP of the FluxRSMetro (132.97 PgC/year) was 10% more than
189 FluxRS (121.69 PgC/year), which is similar to the finding of¹⁰, largely due to differences what differences in the tropics. The
190 difference in GPP in FluxRS and FluxRSMeteo is not likely because of FluxRS only dependence on remote sensing, rather
191 related to different predictor sets and machine learning methods, which requires more investigation¹⁰. Carbon flux trends of
192 FLUXCOM datasets are likely unrealistic since it does not account for CO₂ fertilization effects.

193 The order of dominance of regional magnitude of negative carbon cycle extremes and regional GPP were not consistent. The
194 patterns of losses in carbon uptake and IAV in GPP suggests that the largest negative extremes in the carbon cycle anomalies
195 are highly dependent on the magnitude of IAV in GPP. Since IAV is based on detrended GPP time series and anomalies
196 are calculated by removing seasonality from detrended anomalies, high correlation between the regional IAV and extremes,
197 suggests that seasonality of GPP has minimal effect on the magnitude of carbon cycle extremes. This suggests that spatial
198 patterns of negative carbon extremes are dominated by detrended anomalies and not seasonality and total GPP.

199 The AMZ showed the largest regional magnitude of negative carbon cycle extremes for models and observations. However,
200 SEA showed maximum frequency (14%) of negative carbon cycle extremes for both models and observations. Despite less
201 than 1% of all negative extremes occurred in the AMZ, the total losses in carbon uptake were highest in the AMZ. Large losses
202 represented in AMZ were due to high IAV in GPP in the AMZ which is highly correlated with total regional GPP, which
203 is highest for the region of AMZ. Thus, a few large negative carbon cycle extremes represented most of the IAV in carbon
204 cycle¹⁸. Moreover, the magnitude of negative carbon cycle extremes of observations were smaller than of models. One reason
205 of low variance of IAV in observed data products could be due to low resolution of time input in the machine learning model¹⁰.
206 Other machine learning methods can be applied which benefit from higher temporal variability by using high frequency remote
207 sensing measurements.

208 Most CMIP6 models show similar magnitude of $L_{GPP}:GPP$ for historical and ssp585 periods, indicating that increase in
209 magnitude of negative carbon cycle extremes will be proportional to increase in GPP. This also suggests that the increases
210 IAV in GPP in the future will most likely increase proportional to increase in GPP. While CESM2 suggests that $L_{GPP}:GPP$
211 will decrease, CanESM5 and BCC-CSM2-MR show an increase in $L_{GPP}:GPP$ over time. However, the uncertainties in the
212 magnitude of regional changes in GPP and IAV of GPP are very large and require us to improve climate-carbon modelling and
213 observations on how carbon uptake of terrestrial vegetation is affected by CO₂ enrichment of the atmosphere, and what are the
214 feedbacks from the vegetation to the atmosphere. The changes in $L_{GPP}:GPP$ from historical to ssp585 periods show a larger
215 number of gridcells with strengthening of negative carbon extremes, losses $\geq 4\%$ of GPP, than weakening of negative carbon
216 extremes. Stronger negative carbon cycle extremes could adversely impact terrestrial ecosystems, wildlife, and unirrigated
217 croplands.

218 Conclusions

219 a lot of repetition of results and Discussion, some of it should be moved to the introduction

220 With increasing human activities due to fossil fuel emissions and land use and land cover changes, the carbon emissions
221 have been increasing for decades. While the total carbon uptake by land and ocean was more than 50%, the rest of the carbon
222 emissions have caused an increase in the atmospheric carbon dioxide concentration. This increased atmospheric CO₂ has
223 resulted in increased vegetation productivity due to CO₂ fertilization and rising water use efficiency. The rising atmospheric CO₂
224 concentration is the largest driver of climate change and has led to a rise in the frequency and magnitude of climate extremes
225 such as heatwaves, droughts, and fires. As consequence of these rising extremes, the number of terrestrial carbon cycle extremes
226 have also increased over time. Higher rate of growth and magnitude of negative carbon extremes than positive extremes indicate
227 a net loss in carbon uptake. Since more than a quarter of carbon emissions are taken up by terrestrial ecosystems, further increases
228 in negative carbon cycle extremes could lead to reduction of expected carbon uptake which have a negative feedback on climate
229 change. Therefore, it is important to understand the variability of climate-carbon feedbacks using observation datasets and
230 state-of-the-art earth system models.

231 We found that the uncertainty in the estimation of GPP by multiple models (15%) was higher than multiple observations
232 (8.5%). Thought the mean GPP among multi-observations and multi-models was consistent, the uncertainty of the interannual
233 variability of GPP among multi-observations and multi-models was much larger. Even the agreement among multiple
234 observations on the interannual variability of GPP was very poor which poses a larger question on the reliability of the
235 benchmark datasets that are used to train the models. The magnitude of negative extremes in GPP was about 3 times higher
236 among models than observations. While the GPP is increasing over time, most models indicate that the rate of increase of the
237 magnitude of negative extremes (L) in GPP are expected to be proportional or larger than in rate of increase of GPP, globally.
238 The regions that show a weakening of negative extremes in GPP often witness a decline in total GPP (e.g. Sahel and Northern
239 Australia), thus lower vegetation productivity most likely causing reduced variance of GPP.

240 Both observations and models agree that the largest driver of negative carbon cycle extremes is water scarcity driven reduced
241 soil moisture and precipitation. The compound effect of dry and hot events drive the second largest proportion of negative
242 carbon cycle extremes. The hot temperature have the largest impact in the tropical regions of South America and South Asia,
243 while dry events drive most of the extremes in arid and semi-arid regions. Moreover, at the lag of 1 month the proportion of
244 negative carbon cycle extremes driven by dry events were higher than no-lag response, which implies that vegetation have a
245 buffer and capacity to push back the effects of climate drivers (e.g. water scarcity).

246 Methods

247 Data

248 We used the observations-based GPP up-scaled data products and CMIP6 Earth System Model simulation outputs. Table [Ex-](#)
249 [tended Data2](#) and Table [Extended Data3](#) list the product names, data source websites, and the abbreviations that are used in
250 this paper. The CMIP6 ESM simulation outputs were downloaded from the [ESGF website](#). For attribution analysis using
251 observational GPP, the meteorological data were used from reanalysis datasets of Global Soil Wetness Project Phase 3 ([GSWP3](#))
252 and [ERA5](#). For attribution analysis using ESMs, the meteorological data from the respective ESMs were used.

253 FLUXCOM provides global gridded carbon fluxes from two experimental setups, one with only remote sensing ("RS") input
254 drivers and other with RS and meteorological drivers ("RS+METEO"). The *FluxRS* data product represent the ensemble median
255 of 18 members of up-scaled GPP products which vary in machine learning methods (e.g. Random forest (RF), Artificial Neural
256 Network (ANN), Multivariate Adaptive Regression Splines (MARS), Model-Tree Ensemble (MTE), Kernel Ridge Regression
257 (KRR), Support Vector Regression(SVR))¹⁰. The fluxes in *FluxRS* are estimated only from Moderate Resolution Imaging
258 Spectroradiometer (MODIS) satellite data. The *FluxRSMeteo* data product represent the ensemble median of 30 members of
259 up-scaled GPP products which vary in machine learning methods (RF, ANN, and MARS) using mean seasonal cycle of satellite
260 data and daily meteorological data,¹⁰. The meteorological data sources of *FluxRSMeteo* were GSWP3, CERES_GPCP, WFDEI,
261 CRUJRA_v1, and ERA5. The *FluxANN* represents the up-scaled GPP based on "FLUXCOM RS+METEO" or *FluxRSMeteo*
262 with CRUNCEPv6 climate and using ANN¹⁹. The *GOSIF* dataset is a global Solar-induced chlorophyll fluorescence (SIF)
263 product based on discrete Orbiting Carbon Observatory-2 (OCO-2) SIF soundings, remote sensing data from the MODIS, and
264 meteorological reanalysis data²⁰. The resulting GPP product (are highly correlated with eddy covariance flux towers; $R^2 = 0.73$,
265 $p < 0.001$) has a spatial and temporal resolution 0.05° and 8-day over 2000–17. The *FluxSat* dataset consist of GPP estimates
266 based on models using Satellite-based geometry-adjusted reflectances from MODIS (without meteorological inputs) and Level
267 3 (gridded) SIF derived from the Global Ozone Monitoring Experiment 2 (GOME-2) onboard the MetOp-A (GOME-2A).
268 The GPP training model was calibrated with GPP estimates from eddy covariance flux towers and the resultant GPP was
269 evaluated using an independent subset of FLUXNET 2015 GPP data. The GPP estimates from *FluxSat* are highly correlated
270 with GPP estimates from GOSIF ($R^2 = 0.98$); similar comparison of IAV GPP yields $R^2 = 0.84$. The Water, Energy, and Carbon
271 with Artificial Neural Networks, *WECANN*, estimates latent heat, sensible heat and GPP using ANN models with 6 inputs

namely: GOME-2 SIF, net radiation, air temperature, soil moisture, precipitation, and snow water equivalent. The WECANN retrievals (at $1^\circ \times 1^\circ$ geographical grid and monthly resolution) were evaluated using eddy covariance tower estimates from the FLUXNET network across various climates and conditions²¹. For climate data we used precipitation and temperature from GSWP3 and ERA5. GSWP3 provides daily resolution observed climate data on a global $0.5^\circ \times 0.5^\circ$ from 1901 to 2010²². ERA5 provides climate data at hourly and 30 km spatial grid²³.

277 Data Preparations

The spatial resolution of GOSIF and FluxSat was 0.05° , ERA5 was 0.1° , and FluxANN and GSWP3 was 0.5° . To have a consistent grid level attribution, the GOSIF and FluxSat were conservatively regredded, using `tempestremap`, to 0.1° and 0.5° for attribution with climate drivers from ERA5 and GSWP3, respectively. Climate variables from ERA5 were conservatively regredded to 0.5° spatial resolution for attribution analysis with carbon cycle extremes from FluxANN.

282 Calculation of Extreme Events

Any variation from an expected value of a variable is called an anomaly. We computed anomalies in GPP by subtracting the non-linear trend and modulated annual cycle from the time series of GPP at every grid cell. Interannual variability of GPP is calculated by computing the standard deviation of detrended GPP time series²⁴. Trend and modulated annual cycle were extracted from the time series of GPP by using singular spectral analysis²⁵. Trend is computed by summing all signals that have a return period of 10 years or higher. Modulated annual cycle is composed of return periods of 12 months and harmonics. Therefore, the GPP anomalies comprise of intra-annual variability (periods < 12 months) and interannual variability (12 months $<$ periods < 10 years). To compute trend comprising of periods ≥ 10 years, the temporal duration of the data should be at least 20 years. Hence, only GOSIF, FluxSat, and FluxANN are selected for comparison of carbon cycle extremes with models. The common time period of GOSIF, FluxSat, and FluxANN is from the year 2001 to 2013 and the same period is selected for comparison of carbon cycle extremes with the CMIP6 models.

Increase in the magnitude, frequency and duration of negative extremes in GPP could have adverse impact on biodiversity, natural ecosystem, agriculture (especially rain-fed) and terrestrial carbon uptake. Thus, a negative extreme in GPP represents a large loss in carbon uptake than expected, which could be associated with loss of vegetation productivity driven by fire, droughts, heatwaves, and/or insect infestations. An extreme events occurs when the anomaly exceeds a certain threshold value. The negative and positive extremes in GPP, referred to as carbon negative and positive cycle extremes, are computed using percentile-based thresholds²⁶. A threshold (q) is computed on the probability density function (pdf) of absolute negative and positive GPP anomalies of all non-zero GPP land grid cells, such that extremes represent only 10% of all GPP anomalies (for schematic representation, see Figure [Extended Data4](#)). Therefore, the threshold value for calculation of GPP extremes is indifferent to the pdf of GPP anomalies. As a consequence, a threshold value, q can be calculate such that negative anomalies in GPP are always $< -q$ and positive anomalies are $> q$.

The thresholds were computed by selecting 10th percentile anomalies in GPP for period 2001–13 and 2081–93. Using the

304 singular spectral analysis²⁵, the trend (return period of 10 years and longer) and modulated annual cycle (return period of 12
305 months and harmonics) were calculated at every grid cell. Anomalies at every grid cell were calculated by subtracting trend and
306 annual cycle from the time series of GPP. Hence, anomalies represent the frequencies with time period less than 12 months
307 (intra-annual anomalies) and higher than 12 months but less than 10 years (inter-annual anomalies). The return period of ENSO
308 events varies in the range of 2–7 years, so anomalies comprise of ENSO events.

309 The magnitude of negative carbon extremes is highly depended on the interannual variability (IAV) of GPP at that grid cell,
310 which is defined as standard deviation of detrended GPP anomalies. The comparison of GPP IAV was done for both models
311 and observations to investigate the hot spots of high agreement and disagreement. The data preparation for comparison of GPP
312 among observation data products and models involves two basic steps. Firstly, all data products were regridded to a common
313 0.5° spatial resolution with monthly temporal resolution. Secondly, the time-period is chosen from 2007 to 2010 which is
314 constrained by the availability of WECANN starting in 2007 and the corresponding end year of the FluxRSMeteo data product,
315 i.e. 2010. The mean annual GPP values were calculated for all observation and model datasets. We report R^2 values are squared
316 Pearson's Correlation Coefficients as a measure of agreement between two datasets on GPP distribution.

317 Attribution to Climate Drivers

318 We attributed negative carbon cycle to climate drivers by comparing the distribution of climate drivers anomalies at a grid
319 cell during the negative carbon cycle extremes for the period 2001–13, such that if the median of the precipitation anomalies
320 during the negative carbon cycle extremes at a grid cell was less than the 1st quantile of precipitation anomalies it is assumed
321 that the negative carbon cycle extremes is driven by dry environmental conditions (see Figure [Extended Data4](#) for schematic
322 description). The main advantage of this method is that it allows for attribution of negative carbon cycle extremes to either
323 dry or wet and hot or cold conditions. Other most common method of attribution of two time series are linear or multi linear
324 regression. If the assumed independent variables are partially correlated with each other, the coefficients of multi-linear
325 regression equations could yield unreliable results. Using simple linear regression of dependent variable with each independent
326 variable and comparison of resultant coefficients could be used to quantify relative importance of individual drivers. However,
327 in order to get correlation coefficients with good significance values, one need to perform attribution on the combined time
328 series of negative and positive extreme anomalies. Thus, one needs to further study the correlation coefficients to determine
329 the direction of impact, e.g. if correlation coefficient of GPP and precipitation anomalies is greater than zero implies that
330 anomalous drop in GPP is attributed to drop in precipitation.

331 The common period chosen for attribution of compound climate drivers, precipitation and surface temperature, among
332 observation data and models is 2001–13. The spatio-temporal locations (mask) of negative carbon extremes were used to
333 compute the median value of climate driver anomalies, med_{dri_ext} , at every grid cell experiencing an extreme. The quartile
334 values at 25th and 75th percentiles of driver anomalies, dri_{q25} and dri_{q75} , were also calculated at every grid cell with extremes.
335 The attribution of negative carbon cycle extreme to climate driver is determined by comparing $med_{dri_du_ext}$ with dri_{q25} and
336 dri_{q75} . For example, if the 1st quartile of precipitation anomalies are indicative of dry climatic condition and 4th quartile of

337 surface temperature represents hot climate, then :

- 338 • $med_{pr_du_ext} < pr_{q25}$: represents extreme events driven by dry climate,
- 339 • $med_{tas_du_ext} > tas_{q75}$: represents extreme events driven by hot climate, and
- 340 • $med_{pr_du_ext} < pr_{q25} \& med_{tas_du_ext} > tas_{q75}$: represents extremes driven by compound effect of hot and dry climate.

341 The impact of lagged effect of climate drivers on carbon cycle extremes is analyzed by comparing $med_{dri_du_ext}^{t:n}$ with $dr{t}_{q25}^{0:n-t}$,
342 where t is the lag-months, n is total months in time series and $t : n$ represents the range of values from t^{th} to n^{th} index, akin to
343 python indexing.

344 The carbon cycle extremes could be driven by individual or compound climate drivers. When both hot and dry climatic
345 conditions are fulfilled for attribution, it is referred to as negative extremes driven by compound climate drivers. We selected
346 three observation GPP data products (GOSIF, FluxANN, and FluxSat) for attribution analysis with two observation climate
347 data products (GSWP3 and ERA5). Total precipitation and surface temperature were chosen as climate drivers causing negative
348 carbon cycle extreme events. Similarly, six CMIP6 models were analyzed for attribution of negative carbon cycle extremes to
349 precipitation and temperature. Additionally, attribution analysis was performed with soil moisture and surface temperature
350 for CMIP6 models and included in Additional Information section . The median quantile distribution of climate drivers for
351 the whole period and during extremes were compared to calculate the number of grid cells dominated by hot, cold, dry, wet,
352 and hot & dry climatic conditions. The attribution analysis is dependent on the number of extremes in the grid cell and fewer
353 number of extremes could give less meaningful results. Therefore, we report the attribution analysis using the grid cells that
354 have total number of negative carbon cycle extremes more than the median of global count of negative extremes. The global
355 aggregated results inform us of the most dominant climate driver(s) driving negative extremes but their spatial distribution
356 could vary. We used the SREX regions (see Table [Extended Data1](#)) for to compute the relative distribution of hot and dry
357 climate drivers across 26 regions.

358 References

- 359 1. Friedlingstein, P. *et al.* Global carbon budget 2021. *Earth Syst. Sci. Data* **14**, 1917–2005, DOI: [10.5194/essd-14-1917-2022](https://doi.org/10.5194/essd-14-1917-2022) (2022).
- 360 2. Ahlström, A., Xia, J., Arneth, A., Luo, Y. & Smith, B. Importance of vegetation dynamics for future terrestrial carbon
361 cycling. *Environ. Res. Lett.* **10**, 054019, DOI: [10.1088/1748-9326/10/5/054019](https://doi.org/10.1088/1748-9326/10/5/054019) (2015).
- 363 3. Chen, C., He, B., Yuan, W., Guo, L. & Zhang, Y. Increasing interannual variability of global vegetation greenness. *Environ.*
364 *Res. Lett.* **14**, 124005, DOI: [10.1088/1748-9326/ab4ffc](https://doi.org/10.1088/1748-9326/ab4ffc) (2019).
- 365 4. Chen, C., He, B., Yuan, W., Guo, L. & Zhang, Y. Increasing interannual variability of global vegetation greenness. *Environ.*
366 *Res. Lett.* **14**, 124005, DOI: [10.1088/1748-9326/ab4ffc](https://doi.org/10.1088/1748-9326/ab4ffc) (2019).

- 367 5. Bonan, G. *Ecological Climatology: Concepts and Applications* (Cambridge University Press, 2015), 3 edn.
- 368 6. Reichstein, M. *et al.* Climate extremes and the carbon cycle. *Nature* **500**, 287–295, DOI: [10.1038/nature12350](https://doi.org/10.1038/nature12350) (2013).
- 369 7. Zscheischler, J. *et al.* Future climate risk from compound events. *Nat. Clim. Chang.* **8**, 469–477, DOI: [10.1038/s41558-018-0156-3](https://doi.org/10.1038/s41558-018-0156-3) (2018).
- 370
- 371 8. Reichstein, M., Bahn, M., Mahecha, M. D., Kattge, J. & Baldocchi, D. D. Linking plant and ecosystem functional
372 biogeography. *Proc. Natl. Acad. Sci.* **111**, 13697–13702, DOI: [10.1073/pnas.1216065111](https://doi.org/10.1073/pnas.1216065111) (2014).
- 373 9. Hoffman, F. M. *et al.* 2016 international land model benchmarking (ilamb) workshop report. Tech. Rep., USDOE Office of
374 Science, Washington, DC (United States) (2017). DOI: [10.2172/1330803](https://doi.org/10.2172/1330803).
- 375 10. Jung, M. *et al.* Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the fluxcom approach.
376 *Biogeosciences* **17**, 1343–1365, DOI: [10.5194/bg-17-1343-2020](https://doi.org/10.5194/bg-17-1343-2020) (2020).
- 377 11. Kraemer, G., Camps-Valls, G., Reichstein, M. & Mahecha, M. D. Summarizing the state of the terrestrial biosphere in few
378 dimensions. *Biogeosciences* **17**, 2397–2424, DOI: [10.5194/bg-17-2397-2020](https://doi.org/10.5194/bg-17-2397-2020) (2020).
- 379 12. Mahowald, N. M. *et al.* Interactions between land use change and carbon cycle feedbacks. *Glob. Biogeochem. Cycles* **31**,
380 96–113, DOI: [10.1002/2016GB005374](https://doi.org/10.1002/2016GB005374) (2017).
- 381 13. Green, J. K. *et al.* Large influence of soil moisture on long-term terrestrial carbon uptake. *Nature* **565**, 476–479, DOI:
382 [10.1038/s41586-018-0848-x](https://doi.org/10.1038/s41586-018-0848-x) (2019).
- 383 14. Schimel, D., Stephens, B. B. & Fisher, J. B. Effect of increasing CO_2 on the terrestrial carbon cycle. *Proc.*
384 *Natl. Acad. Sci.* **112**, 436–441, DOI: [10.1073/pnas.1407302112](https://doi.org/10.1073/pnas.1407302112) (2015).
- 385 15. Frank, D. *et al.* Effects of climate extremes on the terrestrial carbon cycle: concepts, processes and potential future impacts.
386 *Glob. Chang. Biol.* **21**, 2861–2880, DOI: [10.1111/gcb.12916](https://doi.org/10.1111/gcb.12916) (2015).
- 387 16. Flach, M. *et al.* Vegetation modulates the impact of climate extremes on gross primary production. *Biogeosciences Discuss.*
388 **2020**, 1–20, DOI: [10.5194/bg-2020-80](https://doi.org/10.5194/bg-2020-80) (2020).
- 389 17. Zhang, T. *et al.* Lagged climatic effects on carbon fluxes over three grassland ecosystems in China. *J. Plant Ecol.* **8**,
390 291–302, DOI: [10.1093/jpe/rtu026](https://doi.org/10.1093/jpe/rtu026) (2014).
- 391 18. Zscheischler, J. *et al.* A few extreme events dominate global interannual variability in gross primary production. *Environ.*
392 *Res. Lett.* **9**, 035001, DOI: [10.1088/1748-9326/9/3/035001](https://doi.org/10.1088/1748-9326/9/3/035001) (2014).
- 393 19. Jung, M. & team FLUXCOM, F. global land carbon fluxes using cruncep climate data. *FLUXCOM Data Portal* (2016).
- 394 20. Li, X. & Xiao, J. A global, 0.05-degree product of solar-induced chlorophyll fluorescence derived from oco-2, modis, and
395 reanalysis data. *Remote. Sens.* **11**, DOI: [10.3390/rs11050517](https://doi.org/10.3390/rs11050517) (2019).

- 396 21. Alemohammad, S. H. *et al.* Water, energy, and carbon with artificial neural networks (wecann): a statistically based
397 estimate of global surface turbulent fluxes and gross primary productivity using solar induced fluorescence. *Biogeosciences*
398 **14**, 4101–4124, DOI: [10.5194/bg-14-4101-2017](https://doi.org/10.5194/bg-14-4101-2017) (2017).
- 399 22. Dirmeyer, P. A. *et al.* Gswp-2: Multimodel analysis and implications for our perception of the land surface. *Bull. Am.*
400 *Meteorol. Soc.* **87**, 1381–1398, DOI: [10.1175/BAMS-87-10-1381](https://doi.org/10.1175/BAMS-87-10-1381) (2006).
- 401 23. (C3S), C. C. C. S. Era5: Fifth generation of ecmwf atmospheric reanalyses of the global climate. *Copernicus Clim. Chang.*
402 *Serv. Clim. Data Store (CDS)* **15**, 2020 (2017).
- 403 24. Fatichi, S., Ivanov, V. Y. & Caporali, E. Investigating interannual variability of precipitation at the global scale: Is there a
404 connection with seasonality? *J. Clim.* **25**, 5512 – 5523, DOI: [10.1175/JCLI-D-11-00356.1](https://doi.org/10.1175/JCLI-D-11-00356.1) (2012).
- 405 25. Golyandina, N., Nekrutkin, V. & Zhigljavsky, A. A. *Analysis of Time Series Structure: SSA and Related Techniques*
406 (Chapman and Hall/CRC, 2001), 1 edn.
- 407 26. Seneviratne, S. I. *et al.* *Changes in Climate Extremes and their Impacts on the Natural Physical Environment*, 109–230
408 (Cambridge Univ. Press, Cambridge, U.K., and New York, 2012). [Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J.
409 Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley (eds.)]. A
410 Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC).

411 LaTeX formats citations and references automatically using the bibliography records in your .bib file, which you can edit via
412 the project menu. Use the cite command for an inline citation, e.g.[?].
413 For data citations of datasets uploaded to e.g. *figshare*, please use the howpublished option in the bib entry to specify
414 the platform and the link, as in the `Hao : gidmaps : 2014` example in the sample bibliography file.

415 Acknowledgements (not compulsory)

416 Acknowledgements should be brief, and should not include thanks to anonymous referees and editors, or effusive comments.
417 Grant or contribution numbers may be acknowledged.
418 *copy of ack of GCB* We acknowledge and thank the Community Earth System Model (CESM) Biogeochemistry Working
419 Group (BGCWG) for performing the CESM1(BGC) simulations analyzed in this study. The CESM1(BGC) model output used
420 here is available at <https://doi.org/10.5281/zenodo.5548153>. Data analysis was performed in Python, and the analysis codes will
421 be available on Github at https://github.com/sharma-bharat/Codes_Carbon_Extremes_2300 after the manuscript is accepted.
422 This research was supported by the Reducing Uncertainties in Biogeochemical Interactions through Synthesis and Compu-
423 tation (RUBISCO) Science Focus Area, which is sponsored by the Regional and Global Model Analysis (RGMA) activity of
424 the Earth & Environmental Systems Modeling (EESM) Program in the Earth and Environmental Systems Sciences Division
425 (EESSD) of the Office of Biological and Environmental Research (BER) in the US Department of Energy Office of Science.

⁴²⁶ This research used resources of the Compute and Data Environment for Science (CADES) at the Oak Ridge National Laboratory,
⁴²⁷ which is managed by UT-Battelle, LLC, for the US Department of Energy under Contract No. DE-AC05-00OR22725.

⁴²⁸ **Author contributions statement**

⁴²⁹ Must include all authors, identified by initials, for example: A.A. conceived the experiment(s), A.A. and B.A. conducted the
⁴³⁰ experiment(s), C.A. and D.A. analysed the results. All authors reviewed the manuscript.

431 **Additional information**

- 432 To include, in this order: **Accession codes** (where applicable); **Competing interests** (mandatory statement).
- 433 The corresponding author is responsible for submitting a [competing interests statement](#) on behalf of all authors of the paper.
- 434 This statement must be included in the submitted article file.

435 The mean GPP of observations for the period of 2001–13 are (Figure [Extended Data5](#)):

- 436 • GOSIF: 137.68 PgC/year
- 437 • FluxSat: 140.45 PgC/year
- 438 • FluxAnn: 114.87 PgC/year
- 439 • WECANN: 117.01 PgC/year
- 440 • FluxRS: 121.69 PgC/year
- 441 • FluxRSMeteo: 132.97 PgC/year

442 Figure [Extended Data7](#) represents the change in the $L : GPP$ for the period 2081–93 relative to the period 2001–13. It
443 shows that the increase in negative carbon cycle extremes grow at least in proportion to increase in GPP (e.g. CESM2 and
444 CNRM-ESM2-1 as $L : GPP$ remains similar) relative to period 2001–13 (Figure [Extended Data7](#)) or increase at a higher rate
445 (more than 20% for other models) with respect to total change in GPP.

446 While the change in mean and maximum $L : GPP$ ratio per decade is similar for all models, the patterns are very different
447 of models CESM2 and CNRM-ESM2-1 and other models. The decadal change in $L : GPP$ ratio has an increasing trend from

Table Extended Data1. SREX Reference Regions

Abreviation	Region's Full Name
ALA	Alaska/N.W. Canada
AMZ	Amazon
CAM	Central America/Mexico
CAS	Central Asia
CEU	Central Europe
CGI	Canada/Greenland/Iceland
CNA	Central North America
EAF	East Africa
EAS	East Asia
ENA	East North America
MED	South Europe/Mediterranean
NAS	North Asia
NAU	North Australia
NEB	North-East Brazil
NEU	North Europe
SAF	Southern Africa
SAH	Sahara
SAS	South Asia
SAU	South Australia/New Zealand
SEA	Southeast Asia
SSA	Southeastern South America
TIB	Tibetan Plateau
WAF	West Africa
WAS	West Asia
WNA	West North America
WSA	West Coast South America

Table Extended Data2. Observation GPP Datasets

Abbreviation	Product Name	#Lat × #Lon	Time
FluxRS ¹	FLUXCOM-RS	360 × 720	1-2001 to 12-2015
FluxRSMeteo ²	FLUXCOM-RS+METEO	360 × 720	1-2001 to 12-2010
FluxAnn ^{3,*}	FLUXCOM	360 × 720	1-1980 to 12-2013
GOSIF ^{4,*}	GOSIF	3600 × 7200	3-2000 to 12-2020
FluxSat ^{5,*}	FluxSat v2.0	3600 × 7200	3-2000 to 7-2020
WECANN ⁶	WECANN	180 × 360	1-2007 to 12-2015

¹GPP estimates from remote sensing (MODIS) only

²GPP estimates from remote sensing (MODIS) and meteorological forcing

³GPP based on FLUXCOM RS+METEO with CRUNCEPv6 climate using Artificial Neural Network.

⁴GPP derived from OCO-2 Solar-induced chlorophyll fluorescence (SIF)

⁵GPP derived from global MODIS and FLUXNET

⁶Water, Energy, and Carbon with Artificial Neural Networks (WECANN): Estimate of GPP using SIF.

*Used for comparison of carbon cycle extremes with models, for details refer to *Methods* section.

Table Extended Data3. CMIP6 Models' Data Details

Product Name	#Lat × #Lon
BCC-CSM2-MR	160 × 320
CanESM5	64 × 128
CESM2	192 × 288
CNRM-CM6-1	128 × 256
CNRM-ESM2-1	128 × 256
IPSL-CM6A-LR	143 × 144
MPI-ESM1-2-HR	96 × 192
MPI-ESM1-2-LR	192 × 384
NorESM2-LM	96 × 144

All the CMIP6 Earth System Models are fully coupled.

Downloaded from: <https://esgf-node.llnl.gov/search/cmip6/>

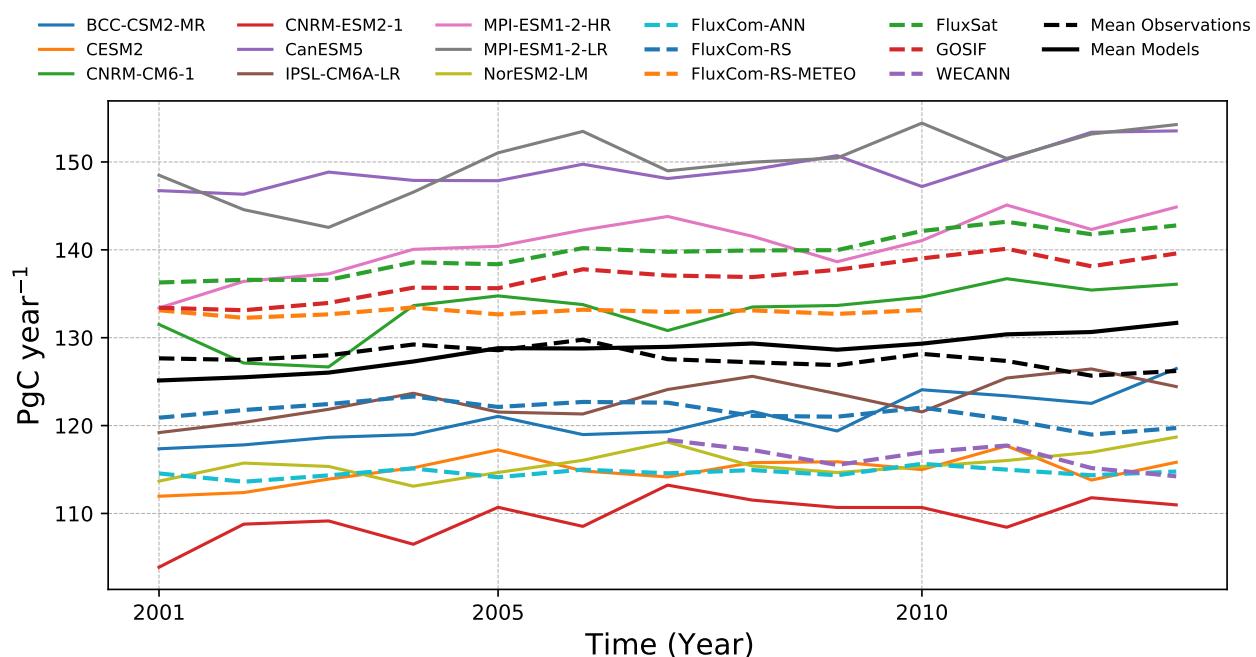


Figure Extended Data1. Timeseries of the integrated yearly GPP from the year 2001 to 2013 of observations and models. The observations and models are represented in dashed and solid lines respectively. For this period, observation yielded GPP estimates of 127 ± 10.8 PgC/year which are within the estimates of models, i.e. 128.5 ± 15.2 PgC/year.

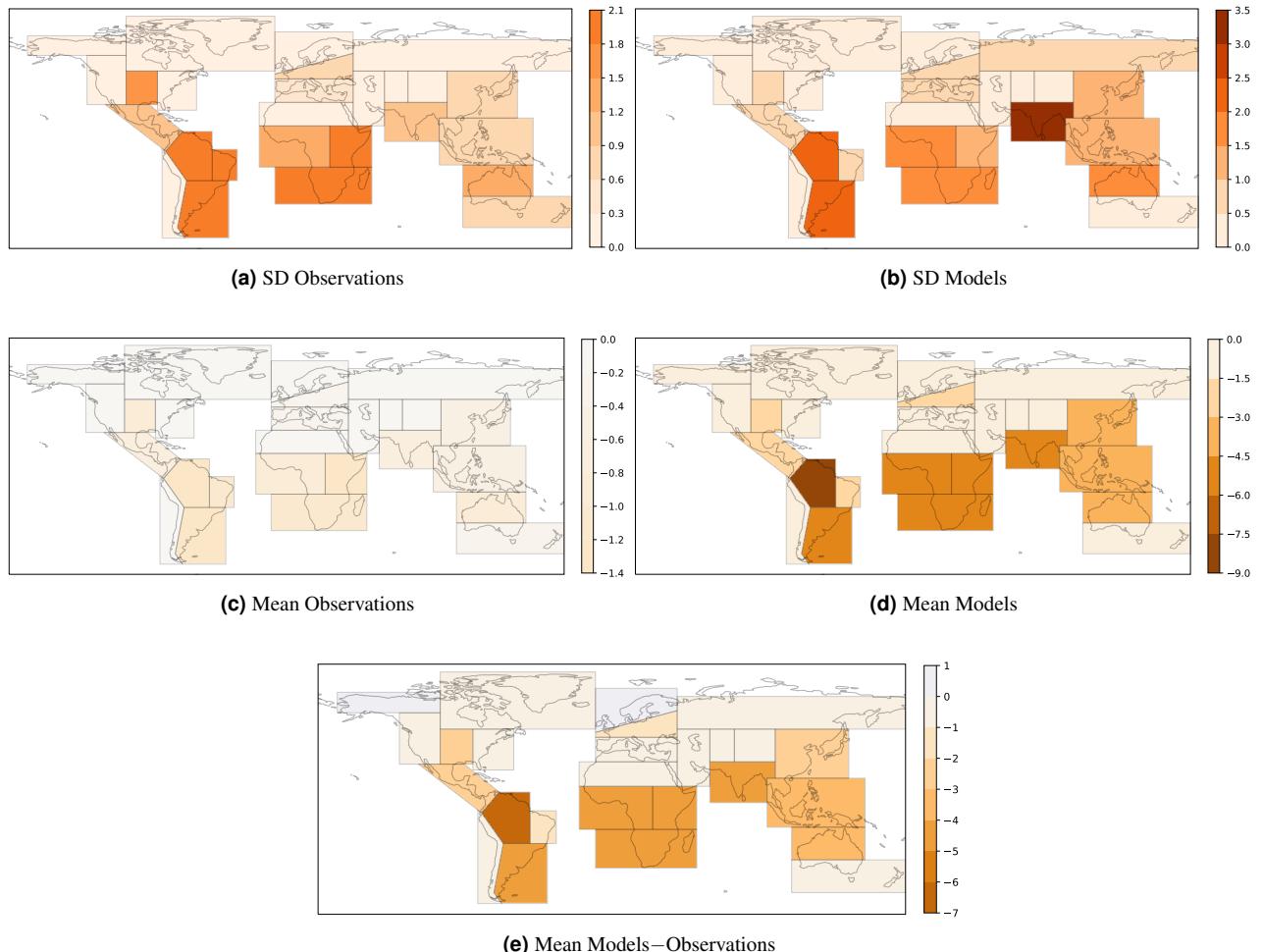


Figure Extended Data2. Spatial distribution of the magnitude of negative carbon cycle extremes from top 50 percent of grid cells experiencing most of the negative extremes. The colors of standard deviation of magnitude of negative GPP extremes of observations (a) and models (b) ranges from 0 to 3.5 PgC/month. The colors of mean magnitude of negative GPP extremes of observations (c), models (d), and models – observation ranges from -9 to 9 PgC/month.

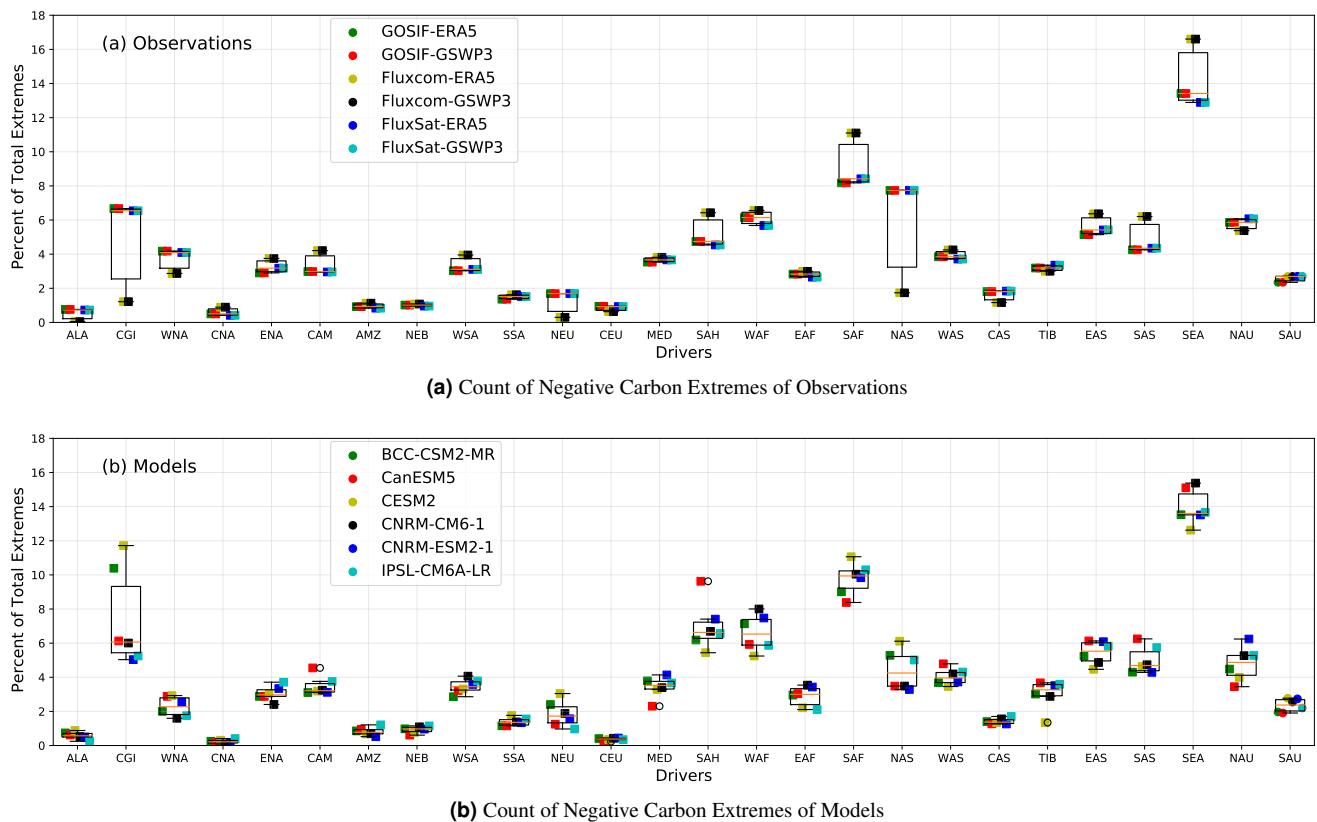
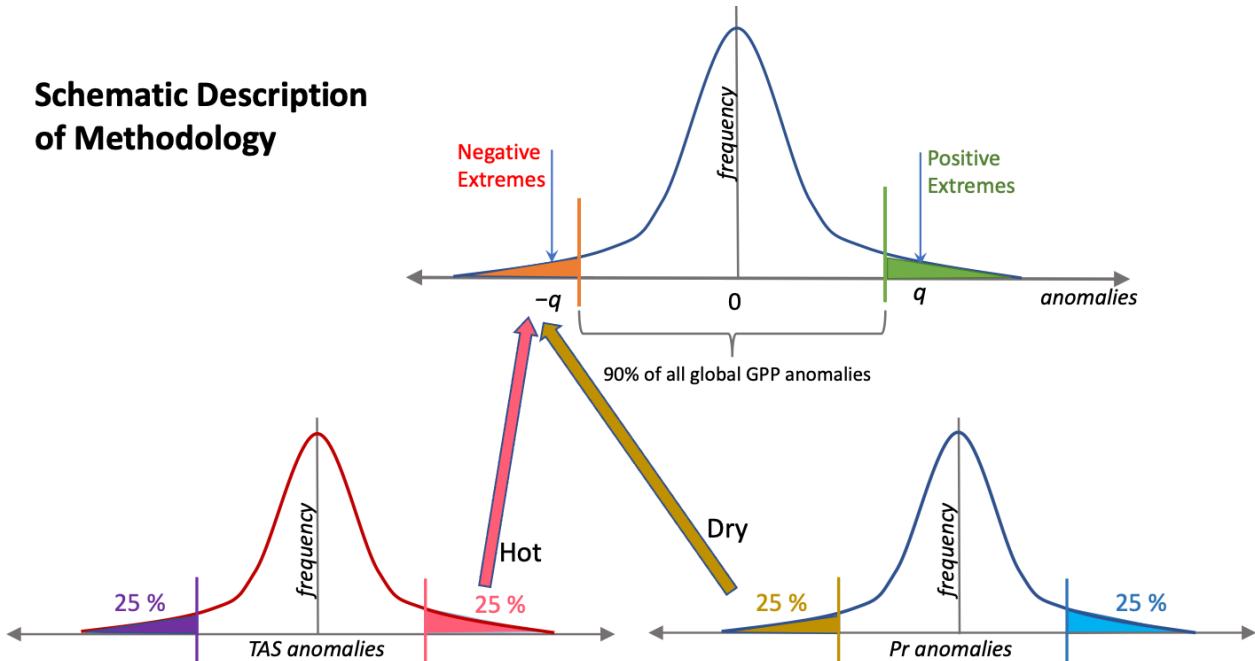


Figure Extended Data3. Spatial distribution of percent count of negative carbon cycle extremes. The box plots of percent count of negative carbon cycle extremes of observations (a) and models (b).

Schematic Description of Methodology



$med_{pr_du_ext} < pr_{q25}$

: carbon cycle extremes driven by dry conditions,

$med_{tas_du_ext} > tas_{q75}$

: carbon cycle extremes driven by hot condition,

$med_{pr_du_ext} < pr_{q25} \&$
 $med_{tas_du_ext} > tas_{q75}$

: carbon cycle extremes driven by compound effect of hot and dry conditions.

Figure Extended Data4. Schematic diagram of detection of GPP extremes and attribution to climate drivers. The shape of the plots is made Gaussian for illustration purposes only.

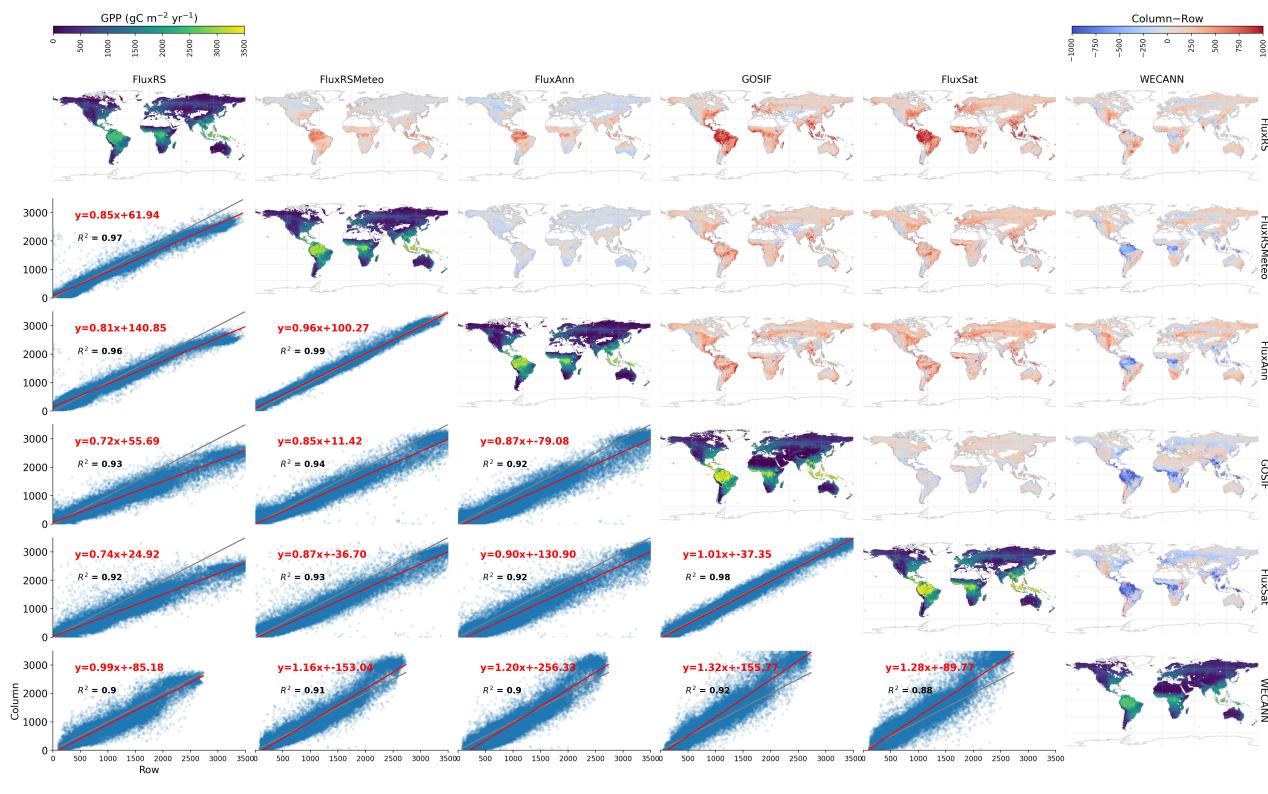


Figure Extended Data 5. Comparison of mean annual GPP at 0.5° spatial resolution for the period 2007–2010 among observations. From left to right and top to bottom the order of observation datasets are FluxRS, FluxRSMeteo, FluxANN, GOSIF, FluxSat, and WECANN. The diagonal maps show the mean annual GPP of every dataset. Maps above the diagonal show the difference of the GPP of column dataset – row dataset. The plots below the diagonal show the point density in blue and 1:1 regression line in grey. Red line and equation represent the best fit line from total least-squares regression.

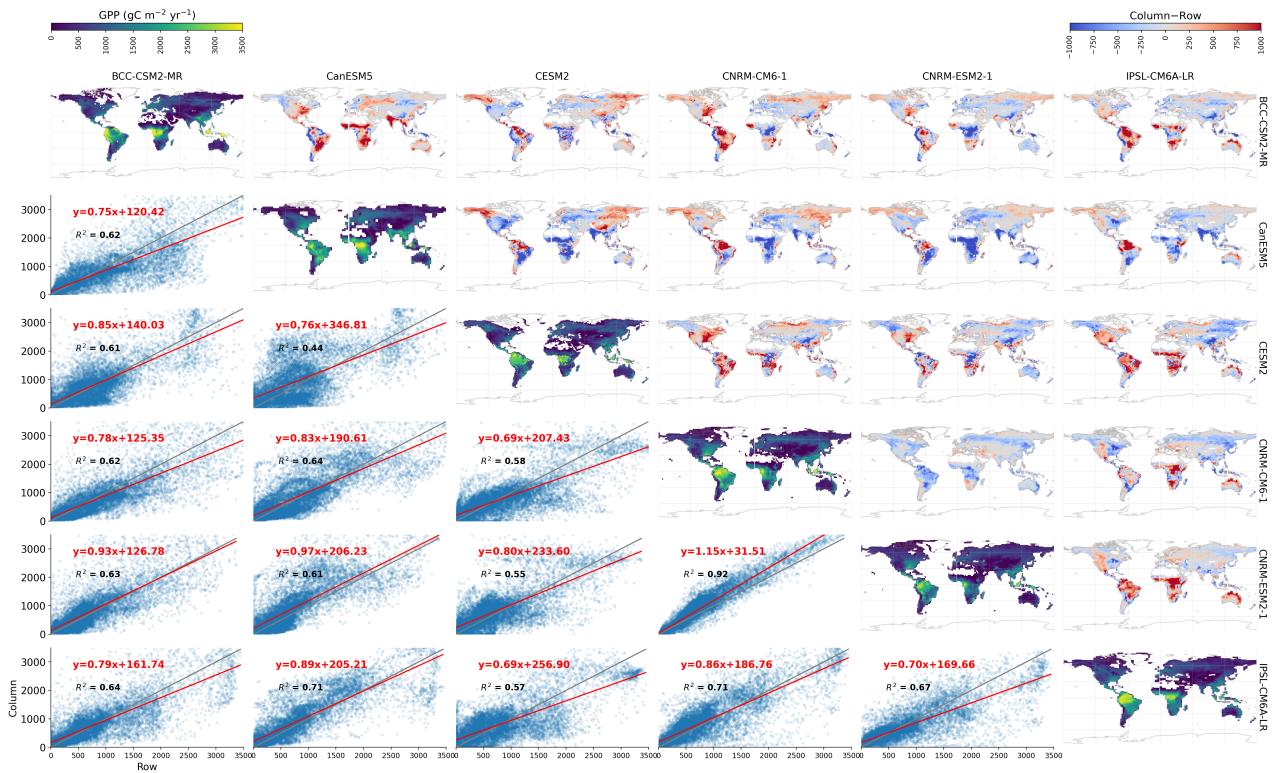


Figure Extended Data6. Comparison of mean annual GPP at 0.5° spatial resolution for the period 2007–2010 among CMIP6 models. From left to right and top to bottom the order of models are BCC-CSM2-MR, CanESM5, CESM2, CNRM-CM6-1, CNRM-ESM2-1, and IPSL-CM6A-LR. The diagonal maps show the mean annual GPP of every dataset. Maps above the diagonal show the difference of the GPP of column dataset – row dataset. The plots below the diagonal show the point density in blue and 1:1 regression line in grey. Red line and equation represent the best fit line from total least-squares regression.

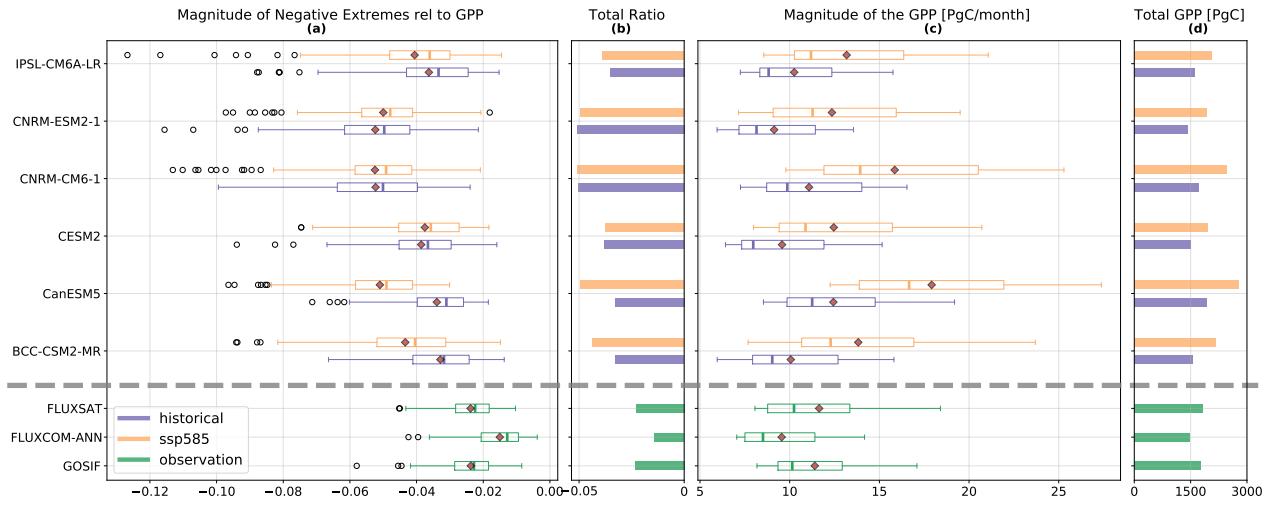


Figure Extended Data7. The green, purple, and orange colors represent the statistics explained below for observation data (2001–13), historical model simulations (2001–13), and SSP585 model simulations (2061–73), respectively. Box plot of ratios of the monthly magnitude of negative carbon cycle extremes to total monthly GPP (a) and total monthly GPP (c) of observation data and model simulations. Ratio of total negative magnitude of negative extremes to to total GPP (b) and total GPP (d). The grey dashed line separates observations and models. Ratio of magnitude of negative extremes to total GPP is referred as L/GPP (ratio) in the paper, which means loss in expected carbon uptake to total GPP. The negative extremes in GPP are calculated using same threshold as historical period.

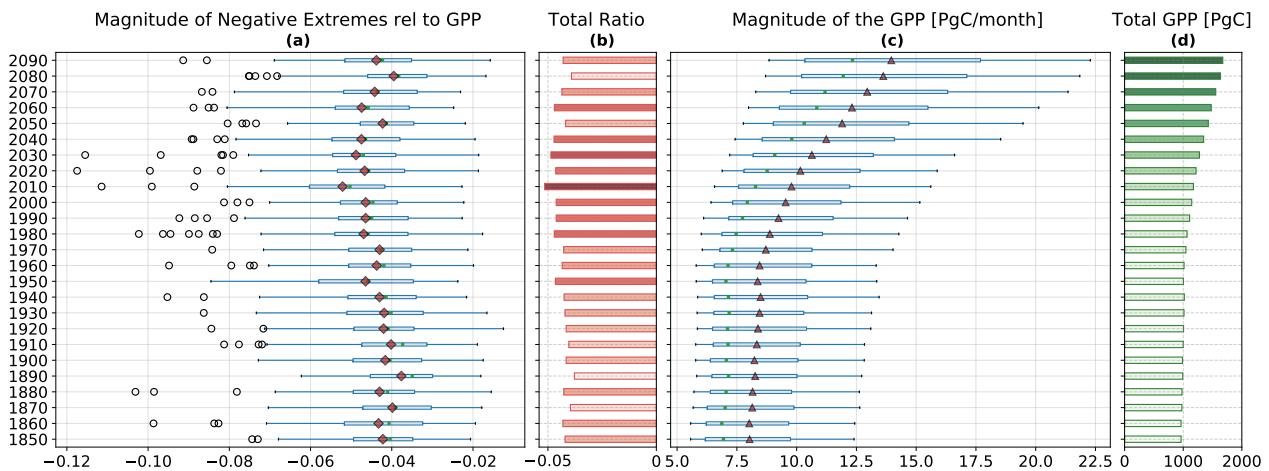


Figure Extended Data8. Decadal change in GPP and ratio of negative carbon cycle extremes to GPP ($L : GPP$) for the CESM2. The negative extremes are calculated with respect the extremes of the period 1850–60.

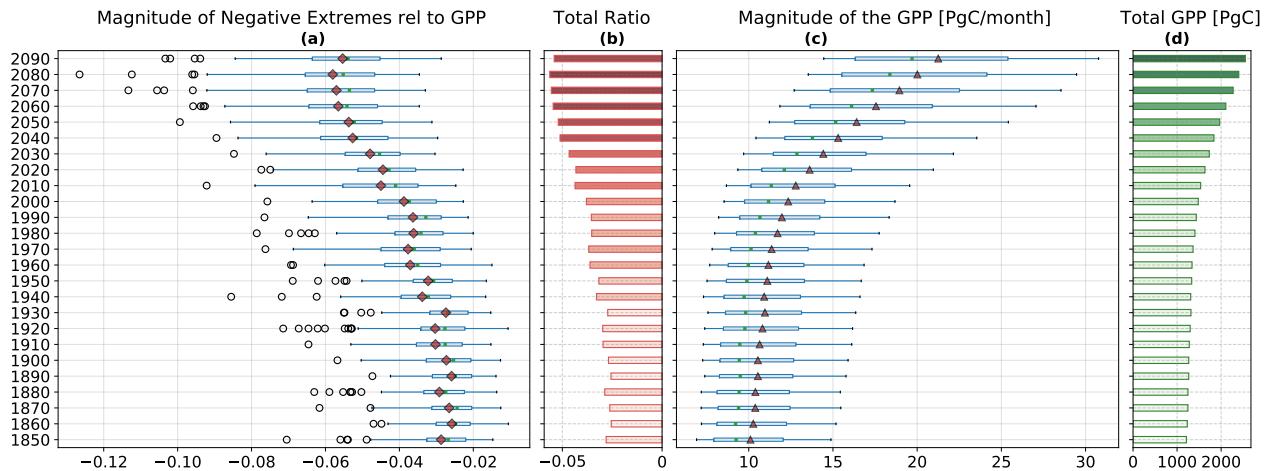


Figure Extended Data9. Decadal change in GPP and ratio of negative carbon cycle extremes to GPP ($L : GPP$) for the CanESM5. The negative extremes are calculated with respect the extremes of the period 1850–60.

448 1850 to 2020 followed by a slight reduction till 2100. Hence the overall increase in negative carbon cycle extremes in CESM2
 449 is proportional to increase in GPP (Figure Extended Data8). However, CanESM5 shows the increase of negative carbon cycle
 450 extremes at a rate higher than the rate of increase of GPP (Figure Extended Data9), which implied the expected losses in
 451 CanESM5 is accelerating over time.

452 Figure Extended Data10: Shows that the negative extremes are driven mostly by dry and hot events.

453 Figure Extended Data11 shows that as the dominance of dry climate conditions increases in driving negative carbon cycle
 454 extremes, especially when pr-tas attribution is analyzed. The attribution to climate drivers is exclusive.

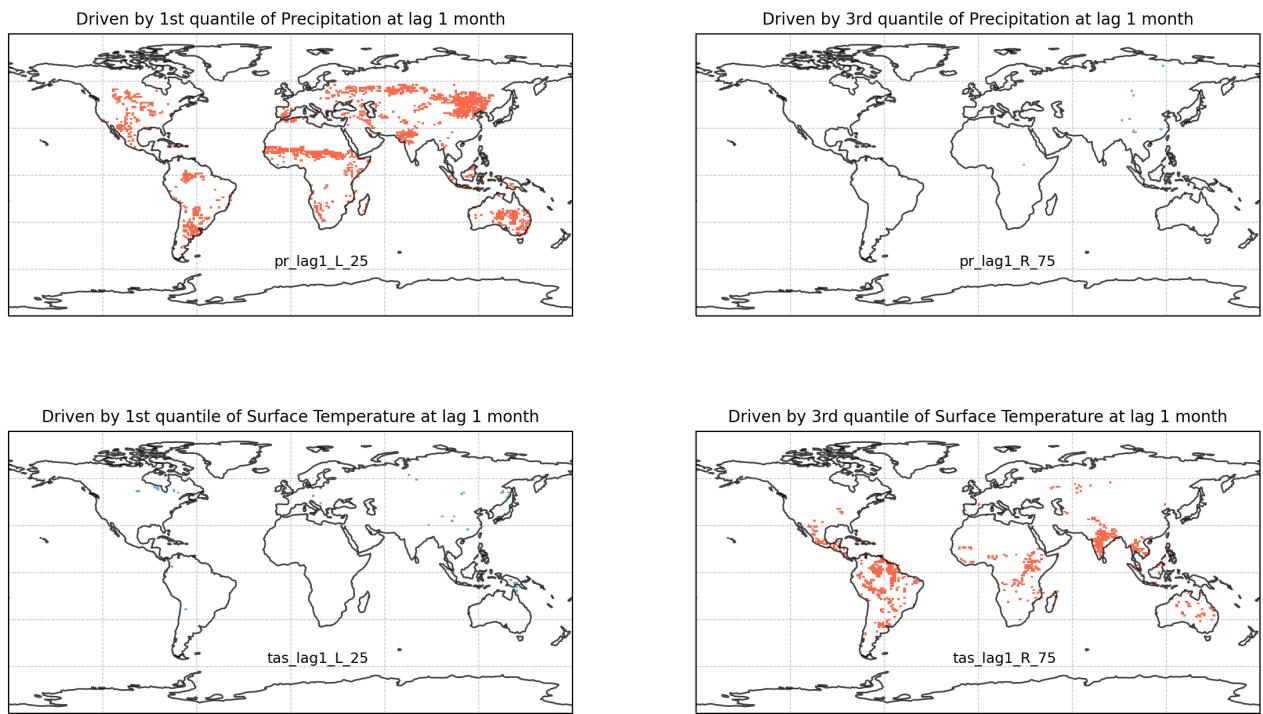


Figure Extended Data10. Spatial distribution of climate drivers, precipitation and surface temperature, causing negative carbon cycle extremes in CESM2 model at lag of 1 month.

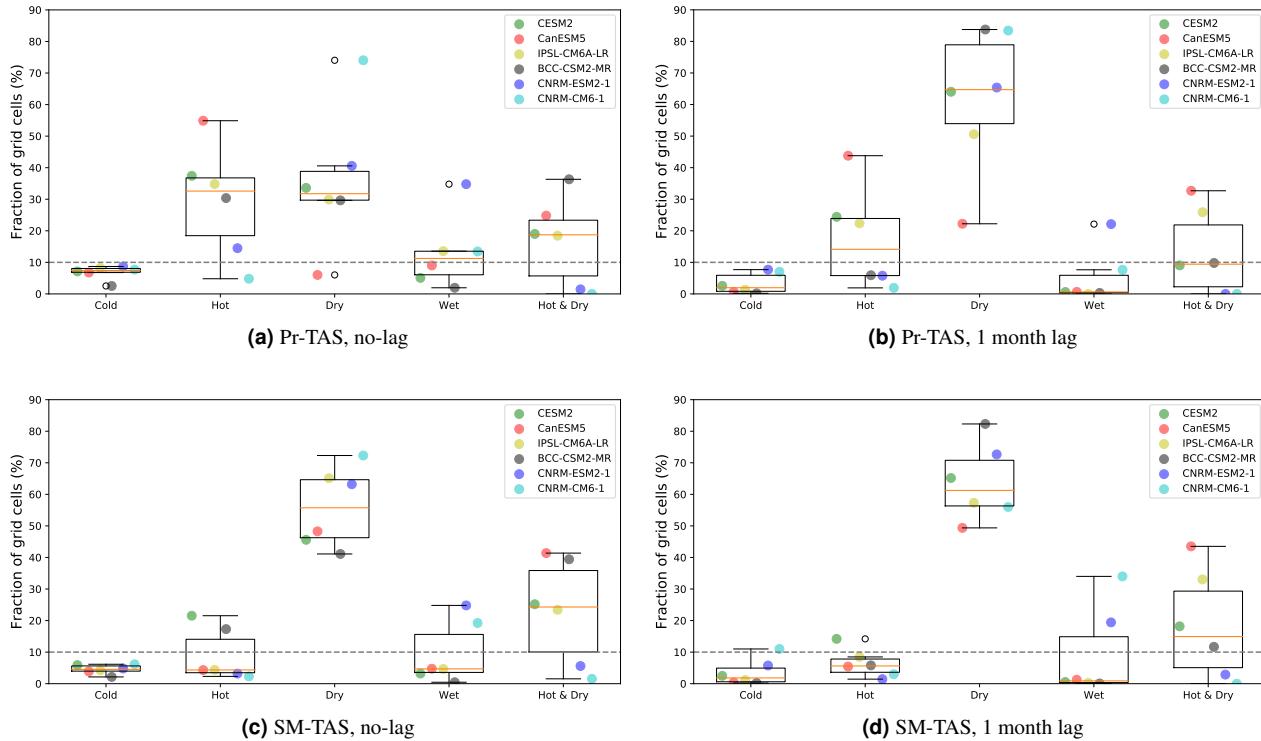


Figure Extended Data11. Attribution of negative carbon cycle extremes to precipitation and soil moisture at no-lag and 1-month lag, exclusively. Left panel shows no-lag and right panel shows lag 1. top shows pr-tas attribution bottom shows sm-tas attr. Cold and wet conditions rarely drive a negative carbon cycle extreme event as they lie below the horizontal dashed line at 10%.