Past and Future Climate Variability Uncertainties in the Global Carbon Budget using the MPI Grand Ensemble

Tammas Francis Loughran^{1,1}, Lena R. Boysen^{2,2}, Ana Bastos^{3,3}, Kerstin Hartung^{4,4}, Felix Havermann^{5,5}, Hongmei Li^{2,2}, Julia Esther Marlene Sophia Nabel^{2,2}, Wolfgang A. Obermeier^{6,6}, and Julia Pongratz^{3,3}

¹Ludwig Maximilian University
²Max Planck Institute for Meteorology
³Ludwig-Maximilians Universität
⁴Deutsches Zentrum für Luft- und Raumfahrt
⁵Dept. of Geography, Ludwig Maximilian University
⁶Faculty of Geosciences, Ludwig Maximilian University

November 30, 2022

Abstract

Quantifying the anthropogenic fluxes of CO2 is important to understand the evolution of carbon sink capacities, on which the required strength of our mitigation efforts directly depends. For the historical period, the global carbon budget (GCB) can be compiled from observations and model simulations as is done annually in the Global Carbon Project's (GCP) carbon budgets. However, the historical budget only considers a single realization of the Earth system and cannot account for internal climate variability. Understanding the distribution of internal climate variability is critical for predicting the future carbon budget terms and uncertainties. We present here a decomposition of the GCB for the historical period and the RCP4.5 scenario using single model large ensemble simulations from the Max Planck Institute Grand Ensemble (MPI-GE) to capture internal variability, and by using this distribution, we investigate the likelihood of historical fluxes with respect to plausible climate states. Our results show these likelihoods have substantial fluctuations due to internal variability, which are partially related to ENSO. We find that the largest internal variability in the MPI-GE stems from the natural land sink and its increasing carbon stocks over time. The allowable fossil fuel emissions consistent with 3°C warming may be between 9–18 PgCyr-1. The MPI-GE is generally consistent with GCP's global budgets with the notable exception of land-use change emissions in recent decades, highlighting that human action is inconsistent with climate mitigation goals.

Past and Future Climate Variability Uncertainties in the Global Carbon Budget using the MPI Grand Ensemble

3 T. F. Loughran¹, L. Boysen³, A. Bastos^{1,2}, K. Hartung^{1,*}, F. Havermann¹, H. Li³, J. E. M. S.

4 Nabel³, W. A. Obermeier¹, and J. Pongratz^{1,3}

- ¹Dept. of Geography, Ludwig Maximilian University, Munich, Germany.
- 6 ²Max Planck Institute for Biogeochemistry, Department of Biogeochemical Integration, Jena,
- 7 Germany.
- **8** ³Max Planck Institute for Meteorology, Hamburg, Germany.
- 9 *Now at: Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre,
- **10** Oberpfaffenhofen, Germany.
- 11 Corresponding author: Tammas Loughran (t.loughran@lmu.de)

12 Key Points:

- We use a single-model large ensemble to estimate uncertainties from internal climate variability in the global carbon budget.
- The land sink accounts for most internal climate uncertainty which may permit 9–18
 PgCyr⁻¹ in allowable emissions by 2050 (for 3°C warming).

17 Abstract

- **18** Quantifying the anthropogenic fluxes of CO₂ is important to understand the evolution of carbon
- 19 sink capacities, on which the required strength of our mitigation efforts directly depends. For the
- 20 historical period, the global carbon budget (GCB) can be compiled from observations and model
- simulations as is done annually in the Global Carbon Project's (GCP) carbon budgets. However,
 the historical budget only considers a single realization of the Earth system and cannot account
- for internal climate variability. Understanding the distribution of internal climate variability is
- critical for predicting the future carbon budget terms and uncertainties. We present here a
- 25 decomposition of the GCB for the historical period and the RCP4.5 scenario using single model
- 26 large ensemble simulations from the Max Planck Institute Grand Ensemble (MPI-GE) to capture
- 27 internal variability. We calculate uncertainty ranges for the natural sinks and anthropogenic
- 28 emissions that arise from internal climate variability, and by using this distribution, we
- 29 investigate the likelihood of historical fluxes with respect to plausible climate states. Our results
- 30 show these likelihoods have substantial fluctuations due to internal variability, which are
- 31 partially related to ENSO. We find that the largest internal variability in the MPI-GE stems from
- 32 the natural land sink and its increasing carbon stocks over time. The allowable fossil fuel
- emissions consistent with 3°C warming may be between 9–18 PgCyr⁻¹. The MPI-GE is generally
- 34 consistent with GCP's global budgets with the notable exception of land-use change emissions in
- 35 recent decades, highlighting that human action is inconsistent with climate mitigation goals.

36 1 Introduction

- **37** The global carbon budget of CO₂ can be decomposed into anthropogenic emissions and natural
- 38 sinks. Anthropogenic emissions are mostly due to fossil fuel burning and fossil carbonates (E_{FF}),
- but also from land-use induced land cover change and land management ("land-use change"
- 40 emissions" in the following, E_{LUC}). The emitted CO₂ is then distributed into three natural sinks: it
- 41 is either assimilated by the land surface via ecosystem productivity (S_{LAND}), absorbed by the
- 42 ocean via diffusion and photosynthesis of marine organisms (S_{OCEAN}), or accumulated in the
- 43 atmosphere (atmospheric growth: G_{ATM}) leading to increased atmospheric CO₂ concentrations
- 44 (Le Quéré et al. 2013; Friedlingstein et al. 2020).

45 One of the key goals of the Global Carbon Project (GCP) is to evaluate anthropogenic 46 perturbations on the global carbon cycle and to understand the response of the natural carbon 47 sinks to increasing fossil emissions and land-use changes (e.g. Friedlingstein et al. 2020; Le 48 Quéré et al. 2018a,b). These global carbon budgets, conducted almost every year since 2007 49 (Canadell et al. 2008), provide an important understanding of the efficiency and potential 50 saturation of the natural sinks. This in turn is essential knowledge for predicting the future sink 51 capacities and, therefore, the required strength for future climate mitigation targets and of 52 "allowable" emissions under given climate targets. A comprehensive understanding of 53 uncertainties in these budgets is essential for guiding policy and decision-making.

The components of the GCP carbon budgets are associated with large uncertainties,
which are based on a combination of observation and model uncertainties. Fossil emissions are
based on energy and fuel consumption data whereby the uncertainties lie in the fuel
consumption, fuel carbon content, and combustion efficiency (Andres et al. 2012). The E_{LUC}
estimate is based on three bookkeeping models, in which estimates of land-use transitions are
combined with observation-based carbon densities to track terrestrial emissions and removals

- 60 according to empirical temporal response curves for each ecosystem (Hansis et al. 2015;
- 61 Houghton and Nassikas 2017). The corresponding estimates for E_{LUC} uncertainty have low
- 62 confidence and are based on expert knowledge, which considers the bookkeeping models and the
- **63** range of the 17 global dynamical vegetation models (DGVMs) (Friedlingstein et al. 2020). The
- 64 ocean sink estimate is based on the standard deviation of nine global ocean biogeochemical
- models and their consistency with observed CO₂ partial pressure-based flux estimates. The
 terrestrial sink in earlier budgets was estimated as a residual from all other terms or based on
- 67 DGVMs from the 2019 budget onwards. The estimates of both S_{LAND} and S_{OCEAN} are evaluated to
- 68 have medium confidence (Friedlingstein et al. 2020). When estimating the land sink with
- **69** DGVMs, the G_{ATM} cannot be matched, leading to a "budget imbalance" term of ~0.4 Pg C yr⁻¹.
- 70 While atmospheric measurements of CO_2 concentration are relatively more accurate, there are
- **71** substantial interannual variations (IAV) driven by natural climate variability (Dlugokencky and **72** Tang 2019; Converse et al. 1004)
- 72 Tans 2018; Conway et al. 1994).
- From such global carbon budgets, it is possible to quantify the future emissions to stay
 within a given trajectory of climate change (Rogelj et al. 2016, Millar et al. 2016). However,
 estimating these "allowable emissions" from historical budgets actually requires considering an
 additional source of uncertainty: the internal variability of the climate system. The uncertainties
 in the GCP budgets are related to observational and model uncertainties while uncertainties
 associated with internal climate variability are not directly addressed.
- 79 Much of the IAV in CO_2 concentration and its impacts on the regional (Zhu et al. 2018) 80 and global carbon sinks (Bastos et al. 2013, Ballantyne et al. 2012) is driven by internal 81 variability in the climate system. Internal variability arises from stochastic processes and 82 feedbacks in the coupled ocean-atmosphere system (e.g. El Niño–Southern Oscillation; ENSO) 83 and is difficult to predict due to high sensitivity to initial conditions and the chaotic evolution of 84 the Earth system (Deser et al. 2012). Traditionally, internal variability in weather and climate 85 forecasts is accounted for by performing ensemble forecasting, i.e. running multiple simulations 86 of the same (or several) models started from perturbed initial conditions, in order to estimate the 87 distribution of future climate states (Deser et al. 2012).
- 88 The importance of considering the full range of potential climate states due to internal 89 climate variability is particularly pertinent to future estimates of the carbon budget, where the 90 exact climate state (and consequently the strength of the natural sinks) in a given year is 91 unknown. Using only one realization may not robustly capture these future states. Furthermore, 92 we cannot assume that the variance of the natural CO₂ fluxes is stationary under increasing 93 atmospheric CO_2 . It is not possible to estimate the range of plausible carbon budget fluxes due to 94 internal climate variability using only one instance of historical observations or observationally 95 forced model simulations. Using ensemble simulations will allow for a more robust calculation 96 of future trends in the mean and variability of the carbon budget terms (e.g. Kay et al. 2015).
- 97 Since the historical observation-based carbon budget uncertainty only considers one
 98 realization of internal climate variability, the influence of internal climate variability on each
 99 budget term is unknown. Therefore, we ask the following research questions:
- How large is the uncertainty from internal climate variability in the global carbon budget terms and how does it compare to the variability of the latest global carbon budget (GCB2020) values?

- How likely were the historical carbon fluxes with respect to the distribution of possible fluxes from internal climate variability and what drove those anomalies?
- How will the carbon budget components and their internal variability change in the future (e.g. under RCP4.5)?

107 In this study, we estimate uncertainties associated with internal climate variability for 108 each component of the carbon budget using a large ensemble of single-model simulations from 109 the Max Planck Institute Grand Ensemble project (MPI-GE; Maher et al. 2019). We compare the 110 results of the estimates for internal climate variability uncertainties to the uncertainties of the 111 recent GCB2020 (Friedlingstein et al. 2020). Furthermore, we discuss the suitability and possible 112 limitations of using a large ensemble of simulations for better understanding variability and 113 uncertainties associated with E_{LUC} and S_{LAND} and how many ensemble members are required to

answer these questions.

115 2 Methods

116 2.1 Overview of models and simulations

117 The methods used to generate the ensemble of climate realizations as part of the MPI-GE project 118 are fully described in Maher et al. (2019). Therefore, we only give a summary here. The MPI-GE 119 is a single model large ensemble project that uses the Max Planck Institute Earth System Model 120 (MPI-ESM; for a full description see Giorgetta et al. 2013) version 1.1. The MPI-ESM is 121 composed of an atmospheric component provided by ECHAM 6.3.01p3 (Stevens et al. 2013) run 122 at T63L47 resolution (~1.8° and 47 vertical layers), an ocean component provided by MPIOM 123 1.6.1p1 (Marsland et al. 2003) run at GR15L40 resolution (~1.5°), the ocean biogeochemistry 124 model HAMOCC5.2 (Ilyina et al. 2013), and the land component JSBACH3 (Reick et al. 2013, 125 Goll et al. 2015). 100 ensemble members are generated by branched initialization (every ~6–24 126 years) from a sub-sample of years from a pre-industrial control (piControl) simulation. The 127 piControl as well as the subsequent historical and future simulations follow the protocol of 128 concentration-driven Earth system model runs of the Coupled Model Intercomparison projects 129 (CMIP), in this case specifically CMIP5 (Taylor et al. 2012).

130 The JSBACH3 component simulates transitions in land cover types with respect to both 131 natural vegetation dynamics and land-use changes. However, we utilize a smaller standalone 132 sub-component of JSBACH3 called Carbon Balance ALONE (CBALONE) to differentiate the 133 emissions due to land-use change from the remaining net land sink (as is done in e.g. Roeckner 134 et al. 2010). As in all Earth system model simulations that perform historical or scenario 135 simulations, anthropogenic and natural effects occur concurrently, i.e. the simulations only 136 provide the net land-atmosphere exchange (i.e. S_{LAND} + E_{LUC}). Only instantaneous emissions to 137 the atmosphere can be derived directly from the historical or scenario simulations (as, e.g., in 138 Lawrence et al. 2012). These, however, neglect legacy emissions that result in particular from the 139 slow decay of wood products, harvested material left on site, and the adjustment of soil carbon 140 stocks to the altered land-use over decades to centuries, but also comprise slow carbon uptake in 141 processes like forest regrowth. In order to capture all fluxes from land-use change (instantaneous 142 and legacy), additional simulations are essential that exclude the land-use change forcing, such 143 that by difference to the historical or scenario simulation E_{LUC} can be isolated (Pongratz et al., 144 2014). Note that effects of altered atmospheric CO_2 concentrations by E_{LUC} , with emissions 145 creating a compensating carbon sink in land and ocean (the "land-use feedback"), are excluded

146 in our concentration-driven feedback (Pongratz et al. 2014). Similarly, since CBALONE is

- driven with the climate from the coupled simulation, changes in surface climate due to land-use
- 148 change also act the same way in both simulations. Thus, the difference between the simulations
- 149 with (MPI-GE) and without land-use change (CBALONE) cancels these effects (apart from
- **150** secondary-order terms) and excludes resulting feedbacks. This is essential to make our estimates
- **151** consistent with the methodology used in the GCB2020 for the terrestrial budget terms.

152 CBALONE includes only the long-term dynamics associated with carbon turnover rates 153 and vegetation biogeography. We force CBALONE with daily data from 100 climate realizations 154 taken from the MPI-GE, both with and without anthropogenic land-use change (LUC and 155 noLUC simulations respectively) comparable to the approach taken by the GCP (Friedlingstein 156 et al. 2020). The land-use change transition data utilized by MPI-GE and CBALONE are taken 157 from the Land Use Harmonization 2 project (LUH2; Hurtt et al. 2011). While the carbon fluxes 158 from CBALONE did not exactly match JSBACH3 estimates, they consistently simulate 159 JSBACH3 fluxes to within 5% accuracy (Figure S6). Therefore, the CBALONE simulations with 160 land-use change are required so that E_{LUC} could be calculated independent of the small 161 CBALONE error (in absence of the error, the net land-atmosphere exchange could have been 162 directly provided by the MPI-GE simulations). 163 The climate realizations used to force CBALONE were taken from existing daily output

The climate realizations used to force CBALONE were taken from existing daily output from the MIP-GE historical and RCP4.5 scenarios (1850–2099; Table 1). We chose the RCP4.5 scenario as a scenario of medium climate change that estimates the CO₂ emissions under climate policies designed to limit global warming to no more than 3°C over present-day temperatures, allowing us to create uncertainty estimates of fossil emissions that are consistent with this goal. The daily model output variables that are used to force CBALONE include 2m air temperature, soil temperature, precipitation, net primary productivity (NPP) per plant functional type (PFT), leaf area index (also per PFT), and maximum wind. These variables are marked as

171 "environmental" in Figure 1.



172

173 Figure 1. Workflow schematic for simulations and carbon budget decomposition for each174 ensemble member. Variables from MPI-GE labeled *"environmental"* include leaf area index, net

primary productivity, topsoil temperature, maximum 10m wind speed, air temperature andprecipitation (see section 2.2)

- **176** precipitation (see section 2.2).
- 177

178

179 Table 1. Experiment simulations. Each experiment has 100 ensemble members. The MPI-GE
180 simulations have been labeled with the prefix "mpige", while the CBALONE simulations are
181 labeled as "cbal". The scenarios are labeled with the suffix "hist" for the historical scenario and
182 "rcp4.5" for the future scenario. Both scenarios for CBALONE are simulated with land-use
183 change (labeled with LUC) and without land-use change using 1850 land-use throughout the

simulation (labeled with noLUC). There are only 97 ensemble members for the CBALONERCP4.5 scenario because a few MPI-GE output files required by CBALONE contained

186 erroneous data.

	LUC	No LUC
Historical (1850–2005)	mpige-LUC-hist cbal-LUC-hist	cbal-noLUC-hist
RCP 4.5 (2006–2099)	mpige-LUC-rcp4.5 cbal-LUC-rcp4.5	cbal-noLUC-rcp4.5

187

188 2.2 Carbon budget decomposition

189 The carbon budget is decomposed here into various source and sink terms as in Friedlingstein et

190 al. (2019), utilizing output from the MPI-GE and the CBALONE simulations. The monthly

191 CBALONE output is aggregated to annual values for comparison to the GCB2020. The cbal-192 noLUC simulation provides land-atmosphere exchange that would occur without land-use

noLUC simulation provides land-atmosphere exchange that would occur without land-use
 changes, and thus S_{LAND} is calculated as the net biome productivity (NBP) from this simulation.

- 194 Equation 1 clarifies components of NBP taken from the model, where NPP is net primary
- 195 productivity, RH is heterotrophic respiration, fFire is carbon flux due to wildfires, fHarvest is
- 196 carbon flux due to crop and wood harvest, fGrazing is carbon flux due to herbivorous grazing,
- 197 and fLCC is the instantaneous emissions from land-use induced land cover changes. The fLCC
- **198** term is zero in the cbal-noLUC simulations.

$$NBP = S_{LAND} = NPP + RH + fFire + fHarvest + fGrazing + fLCC$$
(1)

199

E_{LUC} is calculated as the difference in NBP between the cbal-LUC and cbal-noLUC
 simulations (Equation 2; note that fluxes to the natural sinks are negative values and fluxes to the
 atmosphere are positive consistent with Friedlingstein et al. 2020). Correspondingly, the NBP
 from the cbal-LUC simulation is equivalent to the net land-atmosphere exchange (NET_{LAND}).

$$E_{LUC} = NBP|_{cbal-LUC} - NBP|_{cbal-noLUC} = NET_{LAND} - S_{LAND}$$
(2)

204

 G_{ATM} and S_{OCEAN} are taken directly from the MPI-GE output. The implied "compatible" emissions (also E_{FF}) are calculated as the residual of all other terms in the budget (Equation 3 & Figure 1), as described in Roeckner et al. (2010) and Jones et al. (2013). These are the emissions that would need to occur for CO_2 to be conserved given particular atmospheric concentration, land-use emissions, and natural sink fluxes. This is different from the GCB2020 approach, where all terms were determined independently based on model or observational estimates, which requires a budget imbalance term to be added.

$$E_{FF} + E_{LUC} = G_{ATM} + S_{OCEAN} + S_{LAND}$$
(3)

212

213 We calculated the full decomposition of the carbon budget for each ensemble member of 214 the historical and RCP4.5 scenarios and compare it to the GCB2020 (Friedlingstein et al. 2020) 215 as the best estimate of the real global carbon cycle. Decadal averages of the MPI-GE ensemble 216 mean and standard deviation are calculated for a direct comparison with the decadal mean and 217 uncertainties presented in the GCB2020. To assess the magnitude of the uncertainties due to 218 internal climate variability compared to the magnitude of the budget terms, we further calculate 219 the signal-to-noise ratio (SNR) of each term as the ensemble mean divided by the ensemble 220 standard deviation.

221 2.3 Interannual variability

222 While internal climate variability may contribute to interannual variations in carbon fluxes to the

223 natural sinks, there are also variations driven by non-internal climate related factors, for example

- 224 changes in anthropogenic activity ($E_{FF} + E_{LUC}$) and volcanism. An assessment of uncertainties
- 225 based on temporal standard deviations would be a mixture of internal and non-internal

- variability, while an ensemble standard deviation at a given time step would reflect variations
- only due to internal climate variability. In order to assess future uncertainties, it is important that
- the model can simulate historical IAV appropriately. Here we assess the ability of individual
- 229 MPI-GE and CBALONE ensemble members to adequately represent the temporal standard
- deviation of the historical year-to-year climate variations in each GCB2020 budget term.
- Therefore, we define a reference IAV as the temporal standard deviation of annual fluxes over
- the base period 1961–1990 (World Meteorological Organization standard reference period). All
 models have unique imperfections in their ability to simulate the statistical properties of the
- 234 carbon fluxes such as mean and standard deviation, which we refer to as model bias.
- Furthermore, each may have a different trend over the base period which would artificially alter
- the IAV. To remove the model biases in the ensemble mean of the MPI-GE, we detrend the
- budget terms of each ensemble member before calculating IAV using an ordinary least-squares
- regression (OLR) of the ensemble mean over the historical period 1959–2005. We also detrended
- each model used in the GCB2020 and calculate the IAV over the same period.
- 240 2.4 Probability of exceedance of past budget terms

To evaluate how likely past carbon fluxes were compared to the range of possible climate states
due to internal variability, we describe here a measure of the probability of exceedance.
Supposing a relatively small amount of CO₂ uptake by the land surface in a particular year, it is

- quite likely that under more favorable climate conditions for carbon storage this land CO₂ uptake
- would be exceeded. Therefore, we aim to calculate the probability that the MPI-GE members are
- **246** greater than the GCB2020 multi-model mean (which we assume to be the closest estimate to
- 247 historical CO₂ fluxes). Each budget term for the MPI-GE and GCB2020 is OLR detrended in the
- same way as described above (Section 2.3) except that we use the 1959–2018 period (i.e. the
- 249 longest available common period for GCB2020 and the MPI-GE simulations). For each year and
 250 budget term, we calculate the corresponding cumulative distribution functions ("exceedance") of
- the MPI-GE ensemble members using a kernel density estimator (Scott 2015). We then evaluate
 the GCB2020 terms on the complement of the cumulative distribution functions (1 Pr.) to find
- 253 their occurrence probability (e.g. see Figure S3). Since we use a cumulative distribution, the
- 254 complement probability is the "exceedance probability" of the ensemble spread being larger than
- the historical value. Unusually large historical fluxes will therefore have low probability ofexceedance. This is similar to the probability of exceedance calculations from studies on climate
- 257 extremes (e.g. Suarez-Gutierrez et al. 2020).

258 Finally, we assess the relationship of the GCB2020 exceedance probabilities for S_{LAND} 259 and S_{OCEAN} fluxes to ENSO, since this is the most prominent mode that drives internal climate 260 variability (Dannenberg et al. 2015; Zhang et al. 2019). We use the annual mean Niño 3.4 index 261 from the NOAA Climate Prediction Center (Climate Prediction Center 2017) which uses ERSST 262 V5 (Huang et al. 2017) sea surface temperatures averaged over the region 5°N–5°S, 170–120°W. 263 We then calculate the Pearson's correlation coefficient and the OLR between the exceedance 264 probabilities of the natural sinks and the Niño 3.4 index. We test the significance of this 265 correlation using a two-sided t-test under the null hypothesis that a relationship between the 266 exceedance probabilities of the GCB2020 fluxes and ENSO state can be rejected at the 95% confidence level. Since these methods assume normally distributed data, we beforehand tested 267 268 the normality of the budget terms and their probabilities using the Shapiro-Wilk test for 269 normality (Shapiro and Wilk, 1965). We found that all budget terms (except for G_{ATM}) are

270 normally distributed in the 1850–2018 period.

271 3 Results

3.1 Temporal evolution of budget components and internal climate variabilityuncertainties

274 The historical period and RCP4.5 scenario have globally increasing CO₂ fluxes from the

atmosphere to the land and ocean sinks until about 2040 before decreasing thereafter (see Figure

276 2) due to assumed RCP4.5 mitigation measures. The decrease in land and ocean sink is because

277 G_{ATM} in RCP4.5 decelerates after 2040 resulting in an atmospheric concentration of ~525 ppm

- **278** CO₂ by 2100 (Thomson et al. 2011). The compatible fossil emissions in the MPI-GE (E_{FF} in
- Figure 2) share similar temporal evolution of the natural sinks. On the other hand, E_{LUC} is driven
 by the LUH2 land-use data set and is independent of fossil emissions, which increases until
- about 1990 before becoming a weak net sink from around 2020 onward under the RCP4.5

scenario (Figure 2 and S1 b). Within the period 1970–2010, the ensemble means of the G_{ATM} and

- **283** E_{FF} terms show annual to decadal-scale variations, which are a known feature of the CO₂
- concentration forcing used in the historical period (caused by the introduction of additional CO₂
- observation stations in the 1960s, see Figure 10 of Meinshausen et al. 2017) and are not
- internally driven variations in the MPI-ESM. The S_{LAND} and S_{OCEAN} do not immediately respond
- to such rapid changes in G_{ATM} since they are dominated by the climate state and its variability. It
- 288 then follows that these variations are evident in the residual E_{FF} term.

289



- Figure 2. Stacked decomposition of the CO₂ budget terms from the MPI-GE for the historical
 (1850–2005) and RCP4.5 (2006–2099) scenarios (a) (unstacked plots of the individual terms can
 be found in Figure S1). Thick lines mark the ensemble mean and shading marks the range of the
 ensemble ±1 standard deviation. Overlaid are the GCB2020 budget terms for comparison.
 Vertical lines mark the end of the historical period (2006) and the end of the latest GCP budget
 (2019). An alternative budget using the CMIP5 E_{FF} taken from Andres et al. (2012) is also
 provided (b). The pink line shows the reflected net emissions, the difference with the net natural
- **298** sinks would give the simulated B_{IM} term in Figure S1 f.

299 The budget terms in Figure 2 are stacked for SLAND and GATM, and hence the shown 300 standard deviation of the ensemble members for these terms aggregates according to a normal 301 sum distribution (i.e., $\sigma(S_{OCEAN}+S_{LAND})=\sqrt{\sigma^2(S_{OCEAN}) + \sigma^2(S_{LAND})}$). The atmospheric 302 concentration is prescribed to be the same for all ensemble members, and so G_{ATM} has no 303 ensemble standard deviation. The standard deviation of residual E_{FF} is inherited directly from the 304 net natural sinks and E_{LUC} because it is calculated as a residual in the budget. S_{OCEAN} has a stable 305 standard deviation of ~0.15 Pg C yr⁻¹ (Figure 3 c), which does not have a trend. S_{LAND} has the 306 largest standard deviation throughout the historical period and the RCP4.5 scenario (see Figure 3 307 d), therefore the standard deviation of the net of natural sinks in Figure 2 (and consequently 308 residual E_{FF}) mostly originates from S_{LAND}. Standard deviation increases with time for residual 309 E_{FF} and S_{LAND} (Figure 3 a & d) from ~1 Pg C yr⁻¹ in 1850 to ~1.5 Pg C yr⁻¹ in 2100. E_{LUC} standard 310 deviation gradually increases from almost 0 to ~0.2 Pg C yr⁻¹ by 2010 and later (Figure 3 b).





Figure 3. Yearly ensemble standard deviation for each carbon budget term. The emissions are on the battern (a second se

313 the top (a residual E_{FF} & b E_{LUC}) and the natural sink terms are on the bottom (c S_{OCEAN} & d **314** S_{LAND}).

The importance of internal climate-driven variations (Figure 3) relative to the ensemble mean state can be better understood by analyzing the SNRs (Figure 4). Values greater than one indicate that the mean state dominates the signal, whereas values less than one indicate that the

- 318 internal climate variability uncertainty is the dominant factor in the carbon fluxes. For residual
- **319** E_{FF} and S_{LAND} (Figure 4 a & d), internal variations are more relevant up until 1970. After that, the
- mean carbon fluxes (i.e. the forced signal) are much larger than the variations due to internal
 climate variability, for example ~2.5–3 times greater for S_{LAND}. S_{OCEAN} generally follows the same
- 321 chinate variability, for example ~2.5–3 times greater for S_{LAND} . S_{OCEAN} generally follows the same same several times smaller than the mean
- 323 carbon flux to the ocean from about 1890 onward. On the other hand, the standard deviation in
- 324 E_{LUC} is as large as the mean from 2010 onward (Figure 4 b), however, this is likely a
- 325 consequence of the simulation setup: land-use changes begin in 1850 but the full range of
- 326 variation from the legacy emissions of land-use change does not manifest until several decades
- 327 later. This means the E_{LUC} SNR is effectively only valid under the future scenario when the mean
- **328** E_{LUC} is small.



329

Figure 4. Yearly signal-to-noise ratio for each budget term in the MPI-GE. Dashed lines

delineate ratio 1, where the standard deviation of the respective flux equals the mean flux. E_{LUC}has an inset plot with the post 2010 period zoomed in, when variations from legacy land-use

333 fluxes have fully established.

- 334 3.2 Comparison to GCB2020
- 335 3.2.1 Comparison of means

We compare here the GCB2020 mean of each budget term to the ensemble mean of the MPI-GE

337 for each decade, before comparing the variances in the following sections. Firstly, the residual

338 E_{FF} mean increases faster in the MPI-GE than observed in the GCB2020 (Figure 5 a). Initially,

339 MPI-GE residual E_{FF} in the 1960s is less than the GCB2020 estimate by 0.8 Pg C yr⁻¹ while it is

340 greater than it by 1.3 Pg C yr⁻¹ in the 2010–2018 decade. However, the range of GCB2020 means

is well within the range of values simulated by the MPI-GE. Secondly, there are large differences

in the mean E_{LUC} fluxes between MPI-GE and GCB2020 (Figure 5 b). MPI-GE E_{LUC} is larger
 compared to GCB2020 in decades prior to 2000, however, these values are also within the large

- 344 uncertainty ranges of the GCB2020. In recent decades, the MPI-GE estimates lower E_{LUC} than
- the GCB2020. Thirdly, S_{LAND} tends to be slightly higher in the MPI-GE for almost all decades
- **346** (Figure 5 e). Fourthly, S_{OCEAN} mean fluxes in MPI-GE and GCB2020 are very similar (Figure 5
- d). Lastly, G_{ATM} in MPI-GE has similar decadal variations as GCB2020, both displaying a dip in
- **348** the 1990s, and there is no consistent bias (Figure 5 f).



349

Figure 5. Decadal average of carbon flux budget terms (bars), and the uncertainty expressed as
±1 standard deviation from the mean (error whiskers). The MPI-GE uncertainties are ensemble
standard deviations and the GCB2020 uncertainties are multi model standard deviations. The
dark bars are the MPI-GE and the lighter bars are the GCB2020 values taken from Friedlingstein
et al. (2019). The top row (a and b) are the emissions and the simulated budget imbalance term

355 (c) as shown in Figure 2 b, and the bottom row (d, e and f) are the sink terms.

356 3.2.2 Un-bias-corrected comparison of uncertainties

The uncertainty ranges in Figure 5 are based on ensemble standard deviations for MPI-GE (and
therefore reflect internal climate variability uncertainties) and multi-model standard deviation for
GCB2020. These ranges can tell us two things: how realistic the MPI-GE range of fluxes is
compared to observations, and how large uncertainties associated with internal climate
variability are compared to other sources of uncertainty (e.g. from observational measurements
or the differing process representations in the diferent GCB2020 models). Therefore, we will
determine here whether the GCB2020 mean state lies outside the MPI-GE uncertainty ranges for

364 each budget term.

Residual E_{FF}, B_{IM} (based on the budget in Figure 2 b) and S_{LAND} (Figure 5 a, c & e) have larger standard deviations in the MPI-GE compared to GCB2020, i.e. internal variability is a larger source of error than observational and model uncertainty (more detail follows in 3.2.3). The GCB2020 mean for these budget terms falls within the uncertainty range due to internal climate variability, demonstrating the capability of MPI-GE to capture the observed carbon flux.

370 On the other hand, E_{LUC} and S_{OCEAN} have a narrower range of internal climate variability 371 uncertainty in the MPI-GE compared to the modeled uncertainty in the GCB2020 (Figure 5 b & 372 d). While the GCB2020 mean is within the MPI-GE uncertainty for S_{OCEAN} for most decades 373 (indicating consistency between the two), E_{LUC} GCB2020 means are outside the corresponding 374 MPI-GE ranges for nearly all decades. However, the uncertainty ranges of MPI-GE and 375 GCB2020 overlap for both S_{OCEAN} and E_{LUC}, i.e. certain ensemble members match certain 376 GCB2020 models. Only, the E_{LUC} 2009–2018 mean and standard deviation of the GCB2020 is 377 outside the standard deviation range of uncertainty due to internal climate variability, indicating

- 378 clear inconsistency (see discussion section 4.1).
- 379 There is no uncertainty range for G_{ATM} from MPI-GE (Figure 5 f) since all ensemble 380 members are prescribed with the same atmospheric CO_2 concentration. The error whiskers in the 381 G_{ATM} GCB2020 are derived from various observational uncertainties, which are very small 382 compared to the terms that are simulated by dynamical models (S_{LAND}, S_{OCEAN}, and E_{LUC}). 383 Because the MPI-GE CO₂ concentration starting 2006 is derived from the Global Change 384 Assessment Model (GCAM; Thomson et al. 2011), the difference in G_{ATM} between MPI-GE and 385 the GCB2020 for the last two decades may in part be due to the differences in carbon cycle 386 processes that are represented in MPI-ESM and GCAM.
- **387** 3.2.3 Bias-corrected comparison of uncertainties

388 To more directly evaluate the magnitude of the historical uncertainties associated with internal 389 climate variability compared to the GCB2020, Figure 6 shows the standard deviations where the 390 biases in the means have been removed (centered). The models used in the GCB2020 estimates 391 are forced by only one realization of the climate state—the actual historical climate evolution. 392 Therefore, the plausible carbon fluxes under different climate states cannot be inferred using 393 only the GCB2020, and while the models used in the GCB2020 do contain internal climate 394 variability, the multi-model standard deviations only account for model uncertainty, but not that 395 from natural variability. If we assume that there is no or negligible uncertainty due to internal 396 climate variability associated with the multi-model GCB2020 standard deviation and that the 397 standard deviation of the MPI-GE is entirely due to internal climate variability, then we can find 398 the proportion of the total uncertainty attributable to internal climate variability (i.e. the sum of 399 GCB2020 and MPI-GE uncertainties; red lines in Figure 6). The importance of internal climate 400 variability decreases with time for S_{LAND} and residual E_{FF} and the MPI-GE land sink uncertainty 401 increases faster than the multi-model uncertainty in the GCB2020. For the 2009–2018 decade the 402 contribution of internal climate variability to total uncertainty is 70% for the residual E_{FF} and 403 60% for SLAND. A constant multi-model uncertainty was assumed for ELUC in the GCB2020 and 404 therefore the MPI-GE E_{LUC} uncertainty increases gradually relative to it. By the 2009–2018 405 decade the uncertainty due to internal climate variability would account for 22% of the total E_{LUC} 406 uncertainty. Lastly, approximately 20% of total uncertainty is from internal climate variability

407 uncertainty for S_{OCEAN}.



- **409 Figure 6.** Centered standard deviation of carbon flux from the multi-model GCB2020 (solid
- 410 lines) and ensemble standard deviation from the MPI-GE (dashed lines). The relative
- 411 contribution of internal climate variability uncertainty is marked in red (dot-dashed lines
- 412 corresponding to the right-hand axis). The color coding is the same as that used in Figures 2–5.

413 3.2.4 Interannual variability

- 414 The ability of individual ensemble members to capture the IAV (in the base period 1961–1990)
- 415 for each term compared to the GCB2020 IAVs is shown in Figure 7. The ranges of the IAVs
- 416 generally have good overlap for the E_{LUC} and S_{OCEAN} budget terms. This means that individual
- 417 MPI-GE members can simulate a plausible range of IAV values that are not significantly
- 418 different from the published values from the GCB2020. S_{LAND}, however, shows some IAV bias in
- 419 the MPI-ESM compared to other models in the GCB2020. IAV in MPI-GE S_{LAND} tends to be on
- 420 average 0.4 Pg C yr^{-1} larger than other models. A higher IAV may contribute to the large
- 421 ensemble spread in the MPI-GE for S_{LAND} (compare to Figure 5). There are large differences
- **422** between MPI-GE and GCB2020 for E_{FF} , and G_{ATM} (Figure 7). Evaluation of G_{ATM} is difficult
- 423 because there is no associated uncertainty range; the GCB2020 only has one potential realization
- $424 \quad of \ past \ emissions \ and \ observed \ CO_2 \ concentration, \ and \ the \ MPI-GE \ atmospheric \ CO_2$
- 425 concentrations are prescribed. The observationally-based GCB2020 uncertainties are only 0.02

426 Pg C yr⁻¹ for G_{ATM} and at most 0.5 Pg C yr⁻¹ for residual E_{FF} and if we use these values as a range
427 on top of the GCB2020 IAV, MPI-GE is still outside these ranges.



428

429 Figure 7. Box and whisker plots of interannual variability (IAV) calculated as the standard
430 deviation over the base period 1961–1990 for the MPI-GE (blue) and the GCB2020 (red). The

431 ranges shown here are derived from the ensemble members for MPI-GE, and from multiple

432 model simulations for the GCB2020. The boxes mark the median and inter-quartile range, and

433 the whiskers mark the full range of values.

434 3.3 The relationship of historical probabilities to ENSO

435 To investigate a potential source of the IAV and uncertainty from internal climate variability, we 436 examine here the exceedance probabilities and the relationship to ENSO. Figure 8 shows the 437 probability of the magnitude of the past carbon fluxes in GCB2020 with respect to the 438 distribution of the MPI-GE. Higher values indicate years where the carbon flux for the respective 439 sink was unusually small compared to the MPI-GE distribution and thus were more likely to be exceeded under more favorable climate conditions. SLAND and SOCEAN have large annual variations 440 441 in exceedance probability. For example, since 1960 there were three years where the historical 442 S_{LAND} was so high, related to La Niña, that it had a chance of less than 20% to be exceeded and 443 five years with S_{LAND} so low that it had a chance of more than 80% to be exceeded (Figure 8 a). 444 This highlights the importance of using a large ensemble to capture the high variability in S_{LAND} 445 (see Section 4.5). The cause of these year-to-year variations may come from a variety of internal

446 climate variability modes. To investigate potential drivers, Figure 9 shows that there are

- $447 \quad \text{significant correlations between the Niño 3.4 index and S_{OCEAN} or S_{LAND} exceedance probability of S
- 448 -0.61 and 0.56 respectively (see also Supplementary Text and Figure S2).
- 449



- **450 Figure 8.** Probability of exceedance that the MPI-GE carbon fluxes are greater than the historical
- 451 GCB2020 mean. Lower values indicate years where the carbon flux to the respective sink was 452 *unusually* high compared to the MPI-GE *distribution (vice versa for large values)*. The vertical
- 453 lines mark El Niño (red) and La Niña (blue) years where Niño 3.4 index is greater than 1
- 454 standard deviation from the mean.

455





459 4 Discussion

460 In summary, S_{LAND} has the largest uncertainty, which emphasizes the dominant role of internal 461 climate variability on the land sink (Figure 3 d). This uncertainty gradually increases over time to 462 approximately ±1.5 Pg C yr⁻¹. While the global S_{LAND} flux and CO₂ concentration increases until 463 the middle of the 21st century (Figure 2), afterwards its signal-to-noise ratio of the mean flux 464 nevertheless decreases (Figure 4 b). The internal climate variability uncertainty in E_{LUC} is 465 relatively smaller at approximately ±0.2 Pg C yr⁻¹ (Figure 3 b). However, the trend in E_{LUC} 466 variability is likely due to a combination of sensitivity to initial conditions and the time delay 467 associated with legacy land-use change emissions. The S_{OCEAN} variations from internal climate 468 variability are similarly small as those in E_{LUC} but show almost no trend (Figure 3 c). The S_{LAND} 469 internal climate variability accounts for about 70% of the total uncertainty that results from both 470 internal variability and uncertainties from models and observations (Figure 6 d), much more than 471 for E_{LUC} (approximately 22%) and S_{OCEAN} (approximately 19%). The standard deviations of the 472 MPI-GE compare well with the uncertainty ranges of the GCB2020 for most budget terms: with 473 respect to the ensemble standard deviation against multi-model standard deviations (usually at 474 least an overlap, Figure 5), and with respect to individual ensemble IAV against individual 475 model IAV in the GCB2020 (Figure 7). Finally, we show that the effect of internal climate 476 variability on the historically observed exceedance probabilities of carbon fluxes to the land and 477 ocean have significant but moderate correlations to ENSO (Figure 9).

478 4.1 Differences between MPI-GE and GCB2020

479 One of the most striking differences between the MPI-GE and the GCB2020 estimates is in E_{LUC} ,

480 where the forced ensemble mean signal from land-use change in the RCP4.5 scenario differs

481 from the observed LUH2 data in the last historical decade. The MPI-GE E_{LUC} transitions to a net

- 482 sink at around 2020, while the forcing used in GCB2020 estimates sustained E_{LUC} until this
- 483 period (Friedlingstein et al. 2020, Bastos et al. 2020). Given that the variance of E_{LUC} ensemble

484 members is quite small compared to the forced mean response, the disparity between the RCP4.5 485 land-use change and the GCB2020 becomes evident. The RCP4.5 scenario is characterized by a 486 high CO₂ price that encourages investment into agricultural intensification rather than expansion. 487 Consequently, re-/afforestation would occur following widespread abandonment of agricultural 488 lands and substantial deforestation reduction since 2007 (Thomson et al. 2011). Despite the 489 process of forest regrowth (such as that in North America and Europe; Doelman et al. 2020) 490 being slow, the MPI-GE reduction in E_{LUC} associated with stopping deforestation globally (in 491 particular the Amazon and other tropical regions) is quick and modeling studies simulate 492 substantial carbon uptake by re-/afforestation and reduced deforestation. For example, Sonntag et 493 al. (2016) estimate an uptake of about 200 Pg C over the 21st century with RCP4.5 land-use 494 change in an RCP8.5 climate compared to unmitigated deforestation. However, the trajectory of 495 RCP4.5 land-use change has not been followed until now, and so the land-use-related mitigation 496 potential remains untapped. This explains the large divergence of our results from the GCB2020 497 estimates for the last 15 years.

498 There are also considerable differences in the "compatible" residual E_{FF} in the MPI-GE 499 compared to the GCB2020 values. If we assume the GCB2020 estimate to be the closest estimate 500 to the mean in reality, then the MPI-GE first underestimates the E_{FF} then overestimates it. The 501 discrepancy may arise due to the closure of the carbon balance and the consequent effect that 502 S_{LAND} has on the compatible emissions. On the other hand, the GCB2020 has an imbalance term 503 that includes carbon fluxes that remain unaccounted for. This term would include errors 504 introduced by the calculation of budget terms independently (e.g. model bias errors in E_{LUC} and 505 S_{LAND}, e.g. Dai and Fung, 1993), errors from incomplete coverage of observations, and minor 506 terms that are not included in the budget decomposition. For these reasons, we would not expect 507 the MPI-GE to accurately reproduce E_{FF} .

508 Lastly, another approach to evaluating the MPI-GE against the GCB2020 is to verify that 509 there are no trends in the budget imbalance relative to the GCB2020. If the compatible residual 510 E_{FF} in the MPI-GE budget is replaced with the CMIP5 E_{FF} values (Figure 2b), a budget imbalance term (B_{IM}) can be calculated that is the residual carbon flux that is not accounted for 511 512 under each ensemble member's climate state. This simulated B_{IM} term (Figure S1 f) derived from 513 the MPI-GE is largely consistent with the B_{IM} from the GCB2020 and shows no significant long-514 term trends over the analysis period. Both MPI-GE and GCB2020 show as a positive B_{IM} around 515 the 1950s and again more briefly in the 1990s (suggesting either an overestimate in the emissions 516 or underestimate in the sinks). While Friedlingstein et al. (2020) could not directly attribute a 517 cause to the B_{IM}, they suggest that its variations originate mostly from S_{LAND} and S_{OCEAN}. 518 Specifically, they suggest that it could originate from internal variability which models cannot 519 capture with a single realization. However, the multiple realizations in the MPI-GE B_{IM} range 520 also show positive values in the 1950's, which suggests that it is more likely from common 521 deficiencies in model physics, resolution, or forcing data. In particular, the land-use forcing 522 could explain the 1950s B_{IM}, as the LUH2 forcing creates large emissions in the 1950s (e.g., 523 Hansis et al. 2015) not captured by datasets based on other land-use forcing such as FAO 524 (Houghton and Nassikas 2017).

525 4.2 Allowable emissions under RCP4.5

526 The standard deviations in the MPI-GE (Figure 2) are derived either directly from the ensembles527 or are inferred from other budget terms, and therefore they should be interpreted with care. The

528 standard deviation of residual E_{FF} is mostly derived from S_{LAND} due to its calculation as a

- 529 residual. In this case, the ranges here are merely a range of emissions that are compatible with
- 530 the likely range of climate states, and the corresponding strengths of the ocean and land sinks.
- 531 Therefore, the residual E_{FF} uncertainty estimates from MPI-GE should not be interpreted as 532 variations in fossil fuel emissions due to internal climate variability-related global demand.
- 533 The net sinks and the corresponding compatible residual E_{FF} range are still useful when 534 deciding what the allowable future emissions may be. They indicate the allowable emissions 535 (accounting for internal climate variability) if appropriate policies are implemented to 536 successfully mitigate climate change in a manner that is consistent with the RCP4.5 scenario. 537 Therefore, the maximum and minimum ensemble ranges of 9–18 Pg C yr⁻¹ in residual E_{FF} at 2050 538 denote allowable emissions under this scenario (2019 was 9.95 Pg C yr⁻¹ as per the GCB2020). 539 In Fig 2 the ±1 standard deviation range of the ensemble is shown instead. In the comparison it is 540 clear that extreme outliers occur mainly at the maximum end. These maximum values may occur 541 before fossil emissions have to drop steeply in the MPI-GE and level off at around 5 Pg C yr⁻¹ if 542 the 3°C target is to be met by 2100. This evolution matches well the fossil emissions estimates 543 from GCAM (Thomson et al. 2011) but allows some higher peak emissions than the Integrated 544 Assessment Model assumed, suggesting smaller assumed sinks and slightly larger E_{LUC} in the 545 simplified carbon cycle of this assessment model (see Figure 2 to compare to E_{FF} and E_{LUC} from 546 GCAM).
- 547 As highlighted by Mankin et al. (2020), decision makers need to be provided the full 548 range of possible outcomes in order to make appropriate decisions. For example, policy 549 decisions based only on the most likely outcome may lead to a blowout of greenhouse gas 550 inventory targets, particularly if S_{LAND} performs poorly within a given 5-year accounting period 551 of the Paris Agreement's Global Stocktake (UNFCCC, 2015 and 2017). On the other hand, 552 caution should be taken when considering the efficacy of past decision making because internal 553 variability uncertainties can potentially obfuscate emission reduction efforts such as 554 re-/afforestation.
- 555 4.3 Trends in uncertainty
- 556 The increase in standard deviation in the ensemble members for S_{LAND} may be due to an increase 557 in the variability in the climate state as is expected under a warming climate. For example, 558 Maher et al. (2019) find an increase in the global mean precipitation variability in the MPI-GE 559 1% CO₂ scenario. The trend in S_{LAND} internal variability can also potentially arise from the 560 increase in the magnitude of fossil emissions, which is initially forced in the MPI-GE as the 561 prescribed atmospheric CO₂ concentration. Larger emissions would result in higher atmospheric 562 CO_2 concentrations and increased potential carbon uptake by vegetation via so-called CO_2 563 fertilization (Walker et al. 2021). This combined with the effect of unfavorable climatic 564 conditions (i.e. heat and drought stress) on the carbon uptake by plants acting on an increased 565 carbon stock, results in a larger variance depending on the climate conditions. The increasing 566 internal variability makes it more likely that SLAND becomes near-neutral by the end of the 567 century compared to the start of the historical period (Figure S1 d). This contrasts somewhat with 568 S_{OCEAN}, which has a relatively lower variance and does not have a trend in the historical or future 569 periods under the RCP4.5 scenario (a similar standard deviation is found by Li and Ilyina 2018). 570 However, under higher emissions scenarios S_{OCEAN} has been shown to also have increasing trends
- 571 in CO₂ flux standard deviation (see Figure 1 of Maher et al. 2019).

572 The trend in E_{LUC} may arise for several reasons. Firstly, the legacy effects of land-use 573 change (mostly from wood harvest) take time to manifest. The anthropogenic pools in which 574 CBALONE stores deforested biomass decay to the atmosphere at time scales of 1–100 years. 575 The variance of the ensemble members therefore not only depends on the climate variability of 576 the current year but also on that of preceding years. Consequently, it would take at least 100 577 years for the full variance due to land-use change to manifest. Similarly, the carbon pool of 578 woody, slowly-decomposing litter left on site after clearing or harvesting will build up over time 579 as land-use transitions occur. Thus, more litter is available to react to the climate-dependent 580 microbial decomposition. Note that while the study of Yue et al. (2020) included this effect in 581 their assessment of the contribution of land-use to the interannual variability of the land carbon 582 pools, their high IAV of E_{LUC} (30-45% of net land exchange IAV, compared to 15% in this 583 study) also originates from attributing part of S_{LAND} (the part on managed land) to E_{LUC}. Internal 584 variability alone, our study shows, is about 0.25 Pg C yr⁻¹ standard deviation for E_{LUC} in recent 585 decades (Figure 3) or 20% of the total uncertainty (model plus internal; Figure 6). IAV of E_{LUC} in 586 the MPI-GE is only slightly larger than in the GCB2020 (Figure 7), indicating that the main 587 driver is not internal climate variability, but land-use forcing.

588





591 While the data analyzed in this study is annual and much of the analysis concerns
592 interannual variations, we conducted simulations for several centuries, and therefore the longer
593 time scale variations must also be considered. There are centennial-scale internal variations in

the land carbon content in JSBACH3 and CBALONE (see Figure 2 in Schneck et al. 2013)

which could influence trends and variability of S_{LAND} and E_{LUC} for simulations that run for several

- **596** hundred years. These variations have a periodicity of ~250 years and consist of a change in the
- **597** total land carbon content of ~8 Pg C . This corresponds to an average land carbon flux of 0.03 Pg
- 598 C yr⁻¹ or roughly 2% of the MPI-GE S_{LAND} standard deviation. Schneck et al. (2013) suggest that 599 these long-duration variations in land carbon content are linked to variations in anthropogenic
- these long-duration variations in land carbon content are linked to variations in anthropogenicland cover changes, and the modulation of soil respiration by long-term changes in temperature
- 601 from volcanism and solar forcing. Since the duration of the MPI-GE and CBALONE simulation
- 602 in this study is 250 years, it is possible that these long-term variations may affect the estimates of
- 603 internal climate variability uncertainty in S_{LAND}.
- 604 4.4 ENSO as a potential source of variability
- 605 ENSO is positively correlated with S_{LAND} exceedance probabilities and negatively correlated with
- S_{OCEAN} exceedance probabilities, which is consistent with how ENSO affects CO₂ fluxes to the
 land surface and ocean. During La Niña, cool and moist mean global conditions tend to
- 608 encourage vegetative productivity on land and increase land carbon storage, while El Niño
- 609 drought conditions put widespread stress on ecosystems and suppress productivity (Gonsamo et
- 610 al. 2016; Jones et al. 2001). Meanwhile, over the ocean, stronger pacific equatorial up-welling
- 611 during La Niña brings dissolved inorganic carbon-rich subsurface water to the surface, thereby
- 612 favoring CO₂ out-gassing and reducing net CO₂ uptake (Jones et al. 2001; Feely et al. 1999). The
- 613 cooler sea surface temperatures during La Niña events can increase the dissolution of CO₂ and
- 614 can reduce CO₂ outgassing, but this is a smaller term relative to the up-welling-induced CO₂
- 615 outgassing. This could explain the diverging response of S_{OCEAN} to ENSO compared to that of
- 616 S_{LAND}. The moderate correlation suggests that while ENSO may have a considerable impact on
- 617 interannual variations in CO₂ fluxes, it is very likely that other climate modes and internal
 618 dynamics are also important. No significant correlations with other climate modes could be
- 618 dynamics are also important. No significant correlations with other climate modes could be found at the global scale, however the impacts of climate modes on regional budgets may be
- 619 found at the global scale, however the impacts of climate modes on regional budgets may be
- 620 considerable.

621 4.5 Importance of ensemble size

- 622 Lastly, it is important to discuss the effect of ensemble size on the results and whether or not
- 623 using 100 members is enough or more than necessary. A framework to assess this is
- 624 demonstrated in Milinski et al. (2020). In accordance with this framework, our goal is to quantify
- wariability using the metric of ensemble standard deviation, to within 5% accuracy of the full 100 entire to a standard deviation using 20 iterations of subsequences of the full 20
- 626 member variance. We estimate standard deviation using 30 iterations of subsample sizes from 3–
 627 100 members without replacement. Figure 10 suggests that at least 40 ensemble members are
- required to capture the standard deviation of S_{LAND} to within ±5% accuracy. Since S_{LAND} has the
- 629 largest standard deviation of all budget terms, the accuracy of a sub-sample of the carbon budget
- 630 decomposition would depend on this term. The other budget terms (Figure S5) do not display
- 631 variations as large as S_{LAND} and therefore 40 members are sufficient for those terms. Whether this
- 632 result is representative of other models that simulate internal variability through ensemble
- 633 simulations depends on the budget terms. In the absence of extensive multi-model large634 ensemble projects that provide the full suite of budget terms, including the split into SLAND and
- ensemble projects that provide the full suite of budget terms, including the split into S_{LAND} and
 E_{LUC}, we evaluated this based on the IAV in the models participating in the GCB2020
- 636 simulations that are forced with observed climate (Figure 7). A key assumption is that MPI-GE is
- 637 capable of accurately representing IAV, and the fact that MPI-GE slightly overestimates S_{LAND}

- **638** IAV by 0.4 Pg C yr⁻¹ compared to other models in the GCB2020 suggests that the minimum 40
- 639 ensemble members required here may be a conservative estimate.

640 5 Conclusion

641 In this study, we use a large ensemble of single-model simulations from the Max Planck Institute 642 Grand Ensemble and a sub-component of JSBACH3 (called CBALONE) to decompose the 643 global anthropogenic carbon budget into fossil and land-use change emissions, atmospheric 644 growth, and natural land and ocean sinks. Through its 100 ensemble members, the MPI-GE 645 captures the uncertainties associated with internal climate variability, which we compare to the 646 2020 global carbon budget's uncertainty and interannual variability, and calculate exceedance 647 probabilities of the past carbon fluxes with respect to a full range of climate variability states. 648 We estimate about 40 ensemble members are required to capture internal variability in S_{LAND} and 649 thus all budget components. Contrary to SLAND, to reduce uncertainty in SOCEAN and ELUC 650 estimates, we must prioritize reducing observational error and model spread rather than capturing 651 internal variability. Despite its high internal variability, S_{LAND} (or S_{OCEAN}) is likely not the reason 652 behind the high budget imbalance found in previous studies for the 1950s, which suggests

653 common model deficiencies or biases in the land-use forcing.

We also present a novel estimate of the uncertainty in land-use change emissions
associated with internal climate variability at approximately ±0.2 Pg C yr⁻¹, which we estimate
would account for about 20% of the total (internal and multi-model) land-use change emissions
uncertainties. Land-use change emissions thus contribute little to interannual variability of the
annual carbon budget and are driven rather by land-use forcing than by climate variability.

659 We investigate future changes in the global carbon budget under RCP4.5 and demonstrate 660 upper and lower bounds on the allowable future CO₂ emissions depending on climate variations. 661 The RCP4.5 scenario exemplifies a future where climate policies are implemented to limit 662 warming to less than 3°C over present-day conditions. Our study largely confirms that the 663 allowable emissions under the assumptions of the socioeconomic model GCAM are compatible 664 with RCP4.5, though slightly higher emissions of up to 13 Pg C yr⁻¹ on average would be 665 allowed in the MPI-ESM. The minimum of the full ensemble range is 9 Pg C yr⁻¹ and would be 666 the lower risk limit to ensure we stay below 3°C warming for all possible climate states, while 667 the maximum of 18 Pg C yr⁻¹ would be the higher risk limit for the climate states leading to 668 stronger land CO₂ uptake. Our results suggest that internal variability of the natural land sink 669 increases over the 21st century, which puts the steady persistence of carbon removal by land 670 ecosystems at risk. We also show that even when accounting for random variations in climate 671 and natural sinks, the emissions in recent decades for land-use change—characterized by 672 continuing global deforestation—are dangerously inconsistent with the RCP4.5 goals and further

673 erode our ability to successfully mitigate future warming.

674 Acknowledgments

675 The authors would like to acknowledge the generous help and comments from the Max Planck

676 Institute Grand Ensemble working group. We also thank the German Climate Computing Centre

677 (DKRZ), for providing computational resources. CMIP5 emissions data are available from RCP

678 Database <u>http://www.iiasa.ac.at/web-apps/tnt/RcpDb</u> and the GCP Global Carbon Budgets data

679 are available from <u>https://www.globalcarbonproject.org/carbonbudget/archive.htm</u>. H.L. was

supported by the European Union's Horizon 2020 research and innovation program under grantagreement no. 821003 (4C).

682

684 References

- Andres, R. J., et al. (2012), A synthesis of carbon dioxide emissions from fossil-fuel combustion. *Biogeosciences*, 9, 1845–1871, doi:<u>10.5194/bg-9-1845-2012</u>.
- Ballantyne, A. P., C. B. Alden, J. B. Miller, P. P. Tans, and W. C. White, (2012), Increase in
 observed net carbon dioxide uptake by land and oceans during the past 50 years. *Nature*,
 488, 70–72, doi:10.1038/nature11299.
- Bastos, A., S. W. Running, C. Gouveia, and R. M. Trigo (2013), The global NPP dependence on
 ENSO: La Niña and the extraordinary year of 2011: GLOBAL NPP AND ENSO: THE
 2011 LA NIÑA. *Journal of Geophysical Research: Biogeosciences*, **118**, 1247–1255,
 doi:10.1002/jgrg.20100.
- 694 Bastos, A., et al. (2018), Impact of the 2015/2016 El Niño on the terrestrial carbon cycle
 695 constrained by bottom-up and top-down approaches. *Philosophical Transactions of the*696 *Royal Society B Biological Sciences*, 373, 20170304, doi:10.1098/rstb.2017.0304.
- 697 Canadell, J. G. *et al.* (2007), Contributions to accelerating atmospheric CO2 growth from
 698 economic activity, carbon intensity, and efficiency of natural sinks. *Proceedings of the*699 *National Academy of Sciences* 104, 18866–18870, doi:10.1073/pnas.0702737104.
- Climate Prediction Center (2017), *Monthly Atmospheric SST Indices*,
 <u>https://www.cpc.ncep.noaa.gov/data/indices/</u>.
- Conway, T. J., P. P. Tans, L. S. Waterman, K. W. Thoning, D. R. Kitzis, K. A. Masarie, and N.
 Zhang, 1994: Evidence for interannual variability of the carbon cycle from the National
 Oceanic and Atmospheric Administration/Climate Monitoring and Diagnostics
 Laboratory Global Air Sampling Network. *Journal Geophysical Ressearch Atmospheres*,
 99, 22831–22855, doi:10.1029/94JD01951.
- 707 Dai, A. and Fung, I. Y. (1993), Can climate variability contribute to the 'Missing' CO2 sink?
 708 *Global Biogeochemical Cycles* 7, 599–609, doi:<u>10.1029/93GB01165</u>.
- Dannenberg, M. P., C. Song, T. Hwang, and E. K. Wise (2015), Empirical evidence of El Niño–
 Southern Oscillation influence on land surface phenology and productivity in the western
 United States. *Remote Sensing Environment*, **159**, 167–180,
 doi:10.1016/j.rse.2014.11.026.
- 713 Deser, C., Phillips, A., Bourdette, V. and Teng, H., (2012), Uncertainty in climate change
 714 projections: the role of internal variability. *Climate Dynamics* 38, 527–546,
 715 doi:10.1007/s00382-010-0977-x.
- 716 Dlugokencky, E., and P. Tans (2018), *Trends in atmospheric carbon dioxide, National Oceanic*717 & *Atmospheric Administration, Earth System Research Laboratory (NOAA/ESRL).*
- 718 Doelman, J. C., et al. (2020), Afforestation for climate change mitigation: Potentials, risks and
 719 trade-offs. *Global Change Biology*, 26, 1576–1591, doi:10.1111/gcb.14887.

- Feely, R. A., R. Wanninkhof, T. Takahashi, and P. Tans (1999), Influence of El Niño on the
 equatorial Pacific contribution to atmospheric CO2 accumulation. *Nature* 398, 597–601,
 doi:10.1038/19273.
- Friedlingstein, P., et al. (2019), Global Carbon Budget 2019. *Earth System Science Data*, 11, 1783–1838, doi:10.5194/essd-11-1783-2019.
- Giorgetta, M. A., et al. (2013), Climate and carbon cycle changes from 1850 to 2100 in MPIESM simulations for the Coupled Model Intercomparison Project phase 5. *Journal of*Advances in Modeling Earth Systems, 5, 572–597, doi:10.1002/jame.20038.
- Goll, D. S. *et al.* (2015), Strong dependence of CO2 emissions from anthropogenic land cover change on initial land cover and soil carbon parametrization. *Global Biogeochemical Cycles*, 29, 1511–1523, doi:10.1002/2014GB004988.
- 731 Gonsamo, A., J. M. Chen, and D. Lombardozzi (2016), Global vegetation productivity response
 732 to climatic oscillations during the satellite era. *Global Change Biology*, 22, 3414–3426,
 733 <u>https://doi.org/10.1111/gcb.13258</u>.
- Hansis, E., S. J. Davis, and J. Pongratz (2015), Relevance of methodological choices for
 accounting of land use change carbon fluxes. *Global Biogeochemical Cycles*, 29, 1230–
 1246, doi:10.1002/2014GB004997.
- 737 Houghton, R. A., and A. A. Nassikas (2017), Global and regional fluxes of carbon from land use
 738 and land cover change 1850-2015: Carbon Emissions From Land Use. *Global*739 *Biogeochemical Cycles*, **31**, 456–472, doi:10.1002/2016GB005546.
- Huang, B., et al. (2017), Extended Reconstructed Sea Surface Temperature, Version 5
 (ERSSTv5): Upgrades, Validations, and Intercomparisons. *Journal of Climate*, 30, 8179–8205, doi:10.1175/JCLI-D-16-0836.
- 743 Hurtt, G. C., et al. (2011), Harmonization of land-use scenarios for the period 1500-2100: 600
 744 years of global gridded annual land-use transitions, wood harvest, and resulting
 745 secondary lands. *Climate Change*, **109**, 117–161, doi:<u>10.1007/s10584-011-0153-2</u>.
- 746 Ilyina, T., K. D. Six, J. Segschneider, E. Maier-Reimer, H. Li, and I. Núñez-Riboni (2013),
 747 Global ocean biogeochemistry model HAMOCC: Model architecture and performance as
 748 component of the MPI-Earth system model in different CMIP5 experimental realizations:
 749 The Model Hamocc within Mpi-Esm in Cmip5. *Journal of Advances in Modeling Earth*750 *Systems*, 5, 287–315, doi:10.1029/2012MS000178.
- Jones, C., et al. (2013), Twenty-First-Century Compatible CO2 Emissions and Airborne Fraction
 Simulated by CMIP5 Earth System Models under Four Representative Concentration
 Pathways. *Journal of Climate*, 26, 4398–4413, doi:10.1175/JCLI-D-12-00554.
- Jones, C. D., M. Collins, P. M. Cox, and S. A. Spall (2001), The Carbon Cycle Response to
 ENSO: A Coupled Climate–Carbon Cycle Model Study. *Journal of Climate*, 14, 17,
 doi:10.1175/1520-0442(2001)014<4113:TCCRTE>2.0.CO;2.
- 757 Kay, J. E. et al. (2015), The Community Earth System Model (CESM) Large Ensemble Project:
 758 A Community Resource for Studying Climate Change in the Presence of Internal Climate
 759 Variability. Bulletin of the American Meteorological Society 96, 1333–1349,
 760 doi:10.1175/BAMS-D-13-00255.1.

- 761 Lawrence, P. J. et al, (2012), Simulating the Biogeochemical and Biogeophysical Impacts of
 762 Transient Land Cover Change and Wood Harvest in the Community Climate System
 763 Model (CCSM4) from 1850 to 2100. J. Climate, 25, 3071–3095, doi:10.1175/JCLI-D-11764 00256.1.
- 765 Le Quéré, C., et al. (2013), The global carbon budget 1959–2011. *Earth System Science Data*, 5, 165–185, doi:<u>10.5194/essd-5-165-2013</u>.
- 767 —, et al. (2018a), Global Carbon Budget 2017. *Earth System Science Data*, 10, 405–448, doi:10.5194/essd-10-405-2018.
- 769 —, et al. (2018b), Global Carbon Budget 2018. *Earth System Science Data*, 10, 2141–2194, doi:10.5194/essd-10-2141-2018.
- 1. Li, H., and T. Ilyina (2018), Current and Future Decadal Trends in the Oceanic Carbon Uptake
 Are Dominated by Internal Variability. *Geophysical Research Letters*, 45, 916–925,
 doi:10.1002/2017GL075370.
- Maher, N., et al. (2019), The Max Planck Institute Grand Ensemble: Enabling the Exploration of
 Climate System Variability. *Journal of Advances in Modeling Earth Systems*, **11**, 2050–
 2069, doi:10.1029/2019MS001639.
- 777 Mankin, J. S., F. Lehner, S. Coats, and K. A. McKinnon (2020), The Value of Initial Condition
 778 Large Ensembles to Robust Adaptation Decision-Making. Earth's Future 8,
 779 doi:10.1029/2020EF001610.
- 780 Marsland, S. J., H. Haak, J. H. Jungclaus, M. Latif, and F. Röske (2003), The Max-Planck781 Institute global ocean/sea ice model with orthogonal curvilinear coordinates. *Ocean*782 *Modelling*, 5, 91–127, doi:10.1016/S1463-5003(02)00015-X.
- 783 Meinshausen, M., et al. (2017), Historical greenhouse gas concentrations for climate modelling
 784 (CMIP6). *Geoscientific Model Development*, 10, 2057–2116, doi:10.5194/gmd-10-2057785 2017.
- 786 Millar, R. J. *et al.* (2017), Emission budgets and pathways consistent with limiting warming to
 787 1.5 °C. *Nature Geoscience* 10, 741–747, doi:10.1038/ngeo3031.
- 788 Milinski, S., N. Maher, and D. Olonscheck (2020), How large does a large ensemble need to be?
 789 *Earth System Dynamics*, **11**, 885–901, doi:<u>10.5194/esd-11-885-2020</u>.
- Pongratz, J., C. H. Reick, R. A. Houghton, and J. I. House (2014), Terminology as a key
 uncertainty in net land use and land cover change carbon flux estimates. Earth System
 Dynamics, 5, 177–195, doi:10.5194/esd-5-177-2014.
- 793
- Poulter, B., et al. (2014), Contribution of semi-arid ecosystems to interannual variability of the
 global carbon cycle. *Nature*, **509**, 600–603, doi:<u>10.1038/nature13376</u>.
- 796 Reick, C. H., T. Raddatz, V. Brovkin, and V. Gayler (2013), Representation of natural and
 797 anthropogenic land cover change in MPI-ESM. *Journal of Advances in Modeling Earth*798 *Systems*, 5, 459–482, doi:10.1002/jame.20022.
- Roeckner, E., Giorgetta, M. A., Crueger, T., Esch, M. and Pongratz, J. (2010), Historical and
 future anthropogenic emission pathways derived from coupled climate-carbon cycle
 simulations. *Climatic Change* 105, 91–108, doi:10.1007/s10584-010-9886-6.

- Rogelj, J. *et al.* (2016), Differences between carbon budget estimates unravelled. *Nature Clim Change* 6, 245–252, doi:10.1038/nclimate2868.
- Schneck, R., C. H. Reick, and T. Raddatz (2013) Land contribution to natural CO₂ variability on time scales of centuries: Land Contribution to CO₂ Variability. *Journal of Advances in Modeling Earth Systems*, 5, 354–365, doi:10.1002/jame.20029.
- Scott, D. W. (2015), *Multivariate density estimation: theory, practice, and visualization*. John
 Wiley & Sons.
- Shapiro, S. S. and Wilk (1965), M. B. An Analysis of Variance Test for Normality (Complete
 Samples). *Biometrika* 52, 591–611, doi:<u>10.2307/2333709</u>.
- Sonntag, S., J. Pongratz, C. H. Reick, and H. Schmidt (2016), Reforestation in a high-CO2world
 —Higher mitigation potential than expected, lower adaptation potential than hoped for.
 Geophysical Research Letters, 43, 6546–6553, doi:10.1002/2016GL068824.
- 814 Stevens, B., et al. (2013), Atmospheric component of the MPI-M Earth System Model:
 815 ECHAM6: ECHAM6. *Journal of Advances in Modeling Earth Systems*, 5, 146–172, doi:10.1002/jame.20015.
- Suarez-Gutierrez, L., Müller, W. A., Li, C. and Marotzke, J. (2020), Hotspots of extreme heat
 under global warming. *Climate Dynamics* 55, 429–447, doi:<u>10.1007/s00382-020-05263-</u>
 <u>w</u>.
- Taylor, K. E., Stouffer, R. J. and Meehl, G. A. (2012), An overview of CMIP5 and the
 experiment design. *Bulletin of the American Meteorological Society* 93, 485–498,
 doi:10.1175/BAMS-D-11-00094.1.
- 823 UNFCCC (2015), Paris Agreement, Article 14.
- 824 UNFCCC (2017), Report of the Subsidiary Body for Scientific and Technological Advice on its
 825 forty-fifth session, held in Marrakech from 7 to 15 November 2016.
- 826 Thomson, A. M., et al. (2011), RCP4.5: A pathway for stabilization of radiative forcing by 2100.
 827 *Climate Change*, 109, 77–94, doi:10.1007/s10584-011-0151-4.
- Walker, A. P. et al. (2021) Integrating the evidence for a terrestrial carbon sink caused by increasing atmospheric CO2. New Phytologist 229, 2413–2445 (2021).
- Yue, C., Ciais, P., Houghton, R. A. and Nassikas, A. A. (2020), Contribution of land use to the
 interannual variability of the land carbon cycle. *Nature Communications* 11, 3170,
 doi:10.1038/s41467-020-16953-8.
- Zhang, Y., M. P. Dannenberg, T. Hwang, and C. Song (2019), El Niño-Southern OscillationInduced Variability of Terrestrial Gross Primary Production During the Satellite Era. *Journal of Geophysical Research Biogeosciences*, **124**, 2419–2431,
 doi:<u>10.1029/2019JG005117</u>.
- Zhu, Z., S. Piao, T. Yan, P. Ciais, A. Bastos, X. Zhang, and Z. Wang (2018), The Accelerating
 Land Carbon Sink of the 2000s May Not Be Driven Predominantly by the Warming
 Hiatus. *Geophysical Research Letters*, 45, 1402–1409, doi:10.1002/2017GL075808.
- 840



Global Biogeochemical Cycles

Supporting Information for

Past and Future Climate Variability Uncertainties in the Global Carbon Budget using the MPI Grand Ensemble

T. F. Loughran¹, L. Boysen³, A. Bastos^{1,2}, K. Hartung^{1,*}, F. Havermann¹, H. Li³, J. E. M. S. Nabel³, W. A. Obermeier¹, and J. Pongratz¹

¹Dept. of Geography, Ludwig Maximilian University, Munich, Germany. ²Max Planck Institute for Biogeochemistry, Jena, Germany. ³Max Planck Institute for Meteorology, Hamburg, Germany. *Now at: Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany.

Contents of this file

Figures S1 Text S1 describing Figure S2 Figures S2 Figures S3 Figures S4 Figures S5 Figures S6 Figures S7

Introduction

The following document contains supporting text and figures for the analysis in the main article. The MPI-GE carbon budget is presented in various forms: there is an unstacked presentation of the budget terms (Figure S1), spatial maps of ensemble standard deviation (Figure S2), a budget composed using CMIP5 emissions (Figure S3), comparisons to the future emissions and the 2020 Global Carbon Budget (Figure S4), demonstration of the exceedance probability calculation (Figure S5 and S6) and the ensemble sub-sampling for terms not presented in the main text (Figure S7).



Figure S1. Unstacked MPI-GE carbon budget terms. The shaded region shows the $\pm 1\sigma$ uncertainty range around the MPI-GE ensemble mean. The fine gray lines mark the corresponding GCB2020 budget terms. Panel f) shows the simulated B_{IM} term using the budget in Figure S3.

Text S1.

The spatial distribution of the standard deviation may reveal the regions of SLAND and ELUC that are most sensitive to changes in the climate under the RCP4.5 scenario. Figure S2 shows the standard deviation of SLAND and ELUC averaged over the last decade of the RCP4.5 scenario. The regions of S_{LAND} with the largest variations are tropical regions that can store large masses of carbon in particular in plant biomass (such as the Amazon region in South America, the Congo region in equatorial Africa, and Southeast Asia), and are also strongly influenced by ENSO. Vegetation in all of these regions is known to be sensitive to variations in climate modes (Dannenberg et al. 2015: Poulter et al. 2014; Zhang et al. 2019; Bastos et al. 2018). There are also moderate variances found in extra-tropical regions that are affected by internal climate variability, such as North America, Europe and Australia. Regions that are not sensitive to climate variations are the highly arid regions of Saharan Africa, Central Asia, and the boreal tundra regions. The distribution patterns of sensitivity are similar for E_{LLC} (since the cleared biomass is affected by internal climate variations in the same way as the biomass contributed by S_{IAND} is, although the magnitude of the variations are much smaller, and the largest values are focused on regions with high land-use change (which are scenario dependent). The E_{FF} and E_{UC} are not directly affected by internal climate variations, but the historical exceedance probabilities are nonetheless presented in Figure S7.



Figure S2. Maps of MPI-GE S_{LAND} and E_{LUC} standard deviation averaged for the final decade of the RCP4.5 scenario.



Figure S3. Probability of occurrence calculation of S_{LAND} for a single year 1990. The probability distribution function of the MPI-GE is on the left and the cumulative distribution function is on the right. Dots mark the S_{LAND} values for individual ensemble members. The 1990 GCB2020 value is the vertical line and it's corresponding cumulative probability of occurrence is the horizontal line.



Figure S4. Probability of exceedance that the MPI-GE anthropogenic carbon fluxes are greater than the historical GCB2020 mean. The vertical lines mark El Niño (red) and La Niña (blue) years where Niño 3.4 index is greater than 1 standard deviation from the mean.



Figure S5. Range of standard deviation of the ensemble sub-samples for E_{FF} (top), E_{LUC} (middle) and S_{OCEAN} (bottom). Blue dashed lines mark the accuracy range of the subsample estimates for ±5% error and the black dotted lines mark the accuracy range of the subsamples estimates for ±10% error.



Figure S6. Net land-atmosphere exchange expressed as NBP for the historical period. MPI-GE and CBALONE with land use change are shown for comparison.