Global Change Biology

Global Change Biology (2016) **22**, 2198–2215, doi: 10.1111/gcb.13217

Constraining future terrestrial carbon cycle projections using observation-based water and carbon flux estimates

STEFANOS MYSTAKIDIS[1,2], EDOUARD L. DAVIN[1], NICOLAS GRUBER[2,3] and SONIA I. SENEVIRATNE[1,2]

[1]*Institute for Atmospheric and Climate Science, ETH Zurich, Zurich 8092, Switzerland,* [2]*Center for Climate Systems Modeling, ETH Zurich, Universitätstrasse 16, Zurich 8092, Switzerland,* [3]*Environmental Physics, Institute of Biogeochemistry and Pollutant Dynamics, ETH Zurich, Universitätstrasse 16, Zurich 8092, Switzerland*

Abstract

The terrestrial biosphere is currently acting as a sink for about a third of the total anthropogenic CO_2 emissions. However, the future fate of this sink in the coming decades is very uncertain, as current earth system models (ESMs) simulate diverging responses of the terrestrial carbon cycle to upcoming climate change. Here, we use observation-based constraints of water and carbon fluxes to reduce uncertainties in the projected terrestrial carbon cycle response derived from simulations of ESMs conducted as part of the 5th phase of the Coupled Model Intercomparison Project (CMIP5). We find in the ESMs a clear linear relationship between present-day evapotranspiration (ET) and gross primary productivity (GPP), as well as between these present-day fluxes and projected changes in GPP, thus providing an emergent constraint on projected GPP. Constraining the ESMs based on their ability to simulate present-day ET and GPP leads to a substantial decrease in the projected GPP and to a ca. 50% reduction in the associated model spread in GPP by the end of the century. Given the strong correlation between projected changes in GPP and in NBP in the ESMs, applying the constraints on net biome productivity (NBP) reduces the model spread in the projected land sink by more than 30% by 2100. Moreover, the projected decline in the land sink is at least doubled in the constrained ensembles and the probability that the terrestrial biosphere is turned into a net carbon source by the end of the century is strongly increased. This indicates that the decline in the future land carbon uptake might be stronger than previously thought, which would have important implications for the rate of increase in the atmospheric CO_2 concentration and for future climate change.

Keywords: CMIP5 projections, emergent constraints, evaluation, evapotranspiration, gross primary productivity, net biome productivity, terrestrial carbon cycle

Received 24 March 2015; revised version received 20 October 2015 and accepted 12 November 2015

Introduction

The carbon cycle response to climate change constitutes one of the largest sources of uncertainty in future climate projections (Bodman *et al.*, 2013; Collins *et al.*, 2013). A major part of this uncertainty is associated with terrestrial ecosystems and the so-called land carbon sink. About a third of the total annual anthropogenic emissions of CO_2 is currently removed from the atmosphere by this land sink (Sarmiento *et al.*, 2010; Pan *et al.*, 2011; Ballantyne *et al.*, 2012; Le Quéré *et al.*, 2015; Sitch *et al.*, 2015), but there is a very large spread in earth system model (ESM) results concerning its future evolution in a changing climate (Ciais *et al.*, 2013). The fate of the land sink depends on the ecosystems' response to climate change and a myriad of other factors including nutrient limitations (Ciais *et al.*, 2013) and changes in extreme events such as droughts and heat waves (Seneviratne *et al.*, 2012).

Reducing uncertainty in future projections of the land carbon cycle will require further efforts to improve ESMs and their land surface models (LSMs), in particular by addressing incomplete process representation and alleviating biases identified in these models (Anav *et al.*, 2013; Piao *et al.*, 2013; Hoffman *et al.*, 2014). This necessity is underscored by the fact that uncertainties in simulated features of the terrestrial biosphere might trigger uncertainties in other aspects of the earth system (e.g., atmospheric CO_2 concentration, air temperature) due to the feedbacks involving the carbon cycle and the climate system (Huntingford *et al.*, 2009; Booth *et al.*, 2012; Arora *et al.*, 2013; Bodman *et al.*, 2013; Friedlingstein *et al.*, 2014a).

A complementary approach to reduce uncertainties in future climate projections is to constrain the existing projections using observed features of the earth system,

Correspondence: Stefanos Mystakidis, tel. +41 44 632 82 85, fax +41 44 632 13 11, e-mail: stefanos.mystakidis@env.ethz.ch

for example, based on relationships between the short-term (interannual) and long-term (decadal) variations of some climate variables (Allen & Ingram, 2002; Hall & Qu, 2006; Knutti *et al*., 2006; Qu & Hall, 2014). These emergent constraints are based on the establishment of statistical relationships between modeled quantities across a range of models, with the condition that at least one of these quantities needs to have an observational constraint. This promising approach to reduce uncertainties in climate model projections is implicitly assuming that models showing good performance in present climate are expected to perform well in future climate, but this assumption might not hold in all cases (Knutti, 2008). In addition, one assumes that the models represent independent realizations of reality and that their initial spread represents a good measure of our current level of understanding with regard to the process under investigation. Thus, there are clear limitations to the emergent constraints approach, which need to be recognized when interpreting the results. But at the same time, it is a powerful approach that allows for the combination of observational constraints with future projections.

In the context of carbon cycle projections, Cox *et al*. (2013) have constrained the sensitivity of tropical carbon to climate change as simulated by different coupled climate–carbon cycle models used in the Coupled Climate Carbon Cycle Model Intercomparison Project (C^4MIP) project using the interannual variability in atmospheric CO_2 as the observational constraint. This model-derived relationship was based on the fact that short-term changes in the simulated tropical land carbon sink control to a large extent the simulated variations in atmospheric CO_2. The result suggested that the Amazon forest dieback might be less severe than previously thought. Nevertheless, changes in other climate drivers (e.g., precipitation, radiation) might potentially determine the fate of the tropical land sink in a future climate. Also, an important question about the future fate of the terrestrial carbon sink is whether this lower response to climate change is counteracted by lower response to CO_2 (i.e., CO_2 fertilization) in the models as the balance between the two feedbacks determines the change in the carbon storage on land (Friedlingstein *et al*., 2006). Following the same approach, Wenzel *et al*. (2014) found a similar effect of this observational constraint on the simulated sensitivity of tropical carbon to climate change in the CMIP5 models.

In this study, observation-based estimates of evapotranspiration (ET) and gross primary productivity (GPP) are used to constrain long-term terrestrial carbon cycle projections in CMIP5 models with a primary focus on GPP. This is motivated by the availability of new globally gridded observation-based products of these fluxes (Jung *et al*., 2011; Mueller *et al*., 2013), enabling us to constrain features of the terrestrial carbon cycle not merely globally, but also regionally. Furthermore, the net rate of photosynthesis (GPP) is the initial driver of the carbon uptake by terrestrial ecosystems, thus providing a biogeochemistry-based constraint, while ET provides a complementary hydrologically-based constraint on ecosystems' functioning. We first explore the statistical relationships between these fluxes and the future evolution of GPP and net biome productivity (NBP) in CMIP5 models. The latter is defined in the models as the difference between the carbon uptake by photosynthesis and the carbon release by respiration and disturbance processes (Prentice *et al*., 2001). Given the identified relationships, we then apply the observation-based products to constrain future changes in GPP and NBP regionally and globally and we test the robustness of our results by showing the sensitivity of the constrained fluxes to different constraint strategies (global or spatially explicit and GPP, ET or double constraint).

Materials and methods

CMIP5 models

We use gridded estimates of GPP, NBP and ET from an ensemble of 19 earth system models (Table S1) used in the framework of the CMIP5 project (Taylor *et al*., 2012). We analyze the concentration-driven experiments in which atmospheric CO_2 concentrations were prescribed as an input to the different ESMs. We analyze both the historical period (1989–2005) and the 21st century based on the RCP8.5 scenario (van Vuuren *et al*., 2011). For those models that provided results from several ensemble simulations, we used only the results from the first ensemble member.

All CMIP5 data were been interpolated bilinearly to a common grid with a resolution of $0.5 \times 0.5°$, corresponding to the resolution of the observation-based dataset for GPP and ET. When integrating GPP over different regions, we also accounted for the land fraction in each grid cell.

Observation-based products for GPP, ET, Temperature and Precipitation

We use various observation-based datasets for GPP, ET, temperature and precipitation. The products for GPP and ET are used for the application of the constraints investigated in this study.

Significant uncertainties remain in contemporaneous GPP estimates ranging from estimates based on diagnostic models (Beer *et al*., 2010; Jung *et al*., 2011) to oxygen isotopes (Ciais *et al*., 1997; Welp *et al*., 2011) and short-period satellite products (Zhao *et al*., 2005, 2006; Frankenberg *et al*., 2011). In this study, we use for GPP the product of Jung *et al*. (2011), which is based on the upscaling of local GPP estimates from a global

eddy covariance sites network (FLUXNET; Baldocchi *et al*., 2001). The dataset is available for the period 1982–2010 with a spatial resolution of 0.5°. The upscaling was performed using a machine learning approach based on an algorithm called model tree ensembles (MTE) described by Jung *et al.* (2009, 2011). The standard deviation of the 25 ensemble members generated by the model tree algorithm is provided as a measure of the spread in the MTE-GPP (Jung *et al*., 2011). In view of the large uncertainties in GPP estimates reported in the literature, we doubled this uncertainty range to define a more conservative estimate of the likely range for GPP.

In one of our constraint methodologies, we use the global GPP estimate reported in Beer *et al.* (2010). This study provides an observation-based estimate of GPP during the period 1998–2005 of 123 ± 8 PgC yr^{-1} combining different data-driven approaches to estimate GPP (e.g., model tree ensembles, artificial neural networks, Köppen–Geiger cross Biome approach). This global GPP estimate by Beer *et al.* (2010) was used as reference in the Intergovernmental Panel on Climate Change Fifth Assessment Report (Ciais *et al*., 2013) and has the advantage of considering several independent estimates and methods to quantify global GPP. As with the MTE-GPP product, we doubled the initial uncertainty range to define a more conservative estimate of the likely range.

As reference dataset for ET, we use the newly compiled LandFlux-Eval synthesis product which merges 14 different individual ET datasets (observation-based, satellite products and model estimates of ET) over the period 1989–2005 with a spatial resolution of 0.5° (Mueller *et al*., 2013). The standard deviation of the different ET datasets used in the LandFlux-Eval synthesis product is used as a measure of the spread in ET. As a sensitivity test, we also used a version of the LandFlux-Eval product that includes only diagnostic ET estimates, that is, excluded the estimates based on land surface models.

We further use as reference datasets for temperature and precipitation, the products from the University of Delaware (Willmott & Robeson, 1995) with a spatial resolution of 0.5° and from the Global Precipitation Climatology Project (GPCP) with a spatial resolution of 2.5° (Adler *et al*., 2003), respectively.

Observational constraint approach

If a relationship exists between an observable characteristic of the earth system and the future evolution of a given variable in this system, then observations can be used to constrain this future evolution. Here, we use GPP and ET observation-based datasets to constrain the future evolution of GPP and NBP in CMIP5 projections, based on the relationships we establish in the result section. The constraint approach consists of excluding models not lying within the observational likely range (Fig. S1). The full ensemble of CMIP5 models before applying the constraint is referred to as the 'prior ensemble', while the model ensemble resulting from the exclusion process is referred to as the 'constrained ensemble'. We apply the observational constraint based on the 1989–2005 period because of the availability of the LandFlux-Eval ET over this period only. Two different types of constraint approaches are used here. We present results from these two types of approaches (in total four individual constraint approaches; see Table 1 for a summary) to test the sensitivity of the constrained fluxes to the used constraint approach.

In the first approach, we constrain the models based on their ability to simulate global annual GPP ('*GPPglobal*'). In this approach, we consider only models with GPP lying within the uncertainty range of the global GPP estimate reported by Beer *et al*. (2010). This global estimate is of 123 ± 16 PgC yr^{-1} (±2SD), and as a result GFDL, IPSL and MPI are outside this range and are excluded from the 'GPPglobal' constrained ensemble. Anav *et al*. (2013) in their comparison of simulated CMIP5-GPP with the MTE-GPP during the period 1986–2005 also identified that these models are simulating higher GPP compared to the observation-based estimate by Jung *et al*. (2011).

In the second set of approaches, we apply a spatially explicit (grid point scale), monthly based constraint. The rationale for developing such a constraint is that models might simulate the correct present-day GPP and/or ET as a result of compensating regional (or seasonal) biases. Conversely, some models might perform relatively well in some specific regions/seasons, thus providing relevant regional/seasonal information while having unrealistic global values. We present three different constraints of this type. In the first one, we constrain the models based on their performance in simulating present-day GPP (constraint '*GPPcon*', models with present-day GPP not lying within the likely range of the MTE-GPP product are excluded; Fig. S1a). In the second one, models are constrained based on their performance in simulating ET (constraint '*ETcon*', models with present-day ET not lying within the likely range of the LandFlux-Eval ET product are excluded; Fig. S1b). The third one combines both the GPP and ET constraints (constraint '*GPP&ETcon*', models with present-day GPP and ET not lying within the likely range of the MTE-GPP and LandFlux-Eval ET products, respectively, are excluded; Fig. S1c). For all of these three approaches, the constraint is applied for each month and grid cell individually (using the

Table 1 Summary of the various constraint approaches used in the present study

Name	Reference dataset(s)	Spatial scale
ETcon	LandFlux-Eval synthesis dataset (Mueller *et al*., 2013)	Application on grid point, integration on global scale
GPPcon	MTE product (Jung *et al*., 2011)	Application on grid point, integration on global scale
GPP&ETcon	Combined ET and GPP constraints (see ETcon and GPPcon)	Application on grid point, integration on global scale
GPPglobal	Beer *et al*. (2010)	Application on global scale

likely range of GPP and/or ET for this particular month and grid cell), meaning that a different set of models can be selected for different months/grid cells. If at a given grid point the constraint results in the exclusion of all models, then this grid point is excluded from the analysis (even if it is the case only for 1 month) in order not to bias the annual values due to missing months. Moreover, if one grid cell is excluded based on one of the constraint approaches, we exclude it from all other constraint approaches to ensure a consistent comparison between different approaches. As a consequence, the global mean results presented here are representative for ca. 80% of the land area.

Moreover, to test the robustness of our results, we performed additional analyses and included the results in the Supporting Information. For instance, we used the multimodel mean to summarize the results of the prior and the constrained ensembles, but we also tested the use of the multimodel median instead of the multimodel mean to insure that our results are robust, independently of the chosen statistics (Table S2). We also constrained the models directly based on annual mean values instead of monthly values (Table S3). In the case of the annual-based constraints, we also constrained the models using only the diagnostic datasets of the LandFlux-Eval ET (Table S3, line 6).

Analysis for most productive month

Our various constrained approaches consider all months to derive annual results. However, when presenting maps, we chose to focus only on the most productive month, defined for each grid cell as the month having the maximum GPP, considering a climatology for the period 1989–2005 based on the MTE-GPP. This choice is motivated by the fact that the most productive month has a disproportionally large role in determining the annual mean GPP, even though the growing season length and other factors are clearly important as well (Piao *et al.*, 2007; Xia *et al.*, 2015). The spatial distribution of the most productive month is shown in Fig. 1.

Propagation of uncertainty

We define the model spread at a given grid point and for a given month as the standard deviation of the prior or constrained multimodel ensembles. When providing regional and global means, we aggregated the model spread in space using the following formula taking into account the spatial covariance.

Let a variable x that is a function of n variables,

$$x = f(x_1, x_2, x_3, x_4, x_5, \ldots x_n) \tag{1}$$

The covariance matrix ($n \times n$) can therefore be defined as:

$$\sum = \begin{bmatrix} \mathrm{cov}(x_1,x_1) & \mathrm{cov}(x_1,x_2) & \cdots & \mathrm{cov}(x_1,x_n) \\ \mathrm{cov}(x_2,x_1) & \mathrm{cov}(x_2,x_2) & \cdots & \mathrm{cov}(x_2,x_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{cov}(x_n,x_1) & \mathrm{cov}(x_n,x_2) & \cdots & \mathrm{cov}(x_n,x_n) \end{bmatrix} \tag{2}$$

When x equals to the sum of the n variables, then the total variance of x is given by:

$$\mathrm{Var}(x) = s^2_{x_1} + s^2_{x_2} + \ldots + s^2_{x_n} + \mathrm{cov}(x_1,x_2) + \ldots + \mathrm{cov}(x_1,x_n) + \mathrm{cov}(x_2,x_1) + \ldots + \mathrm{cov}(x_2,x_n) + \ldots \tag{3}$$

or

$$\mathrm{Var}(x) = \sum_{i}^{n} s^2_{x_n} + \sum_{j\neq k}^{n}\sum_{k\neq j}^{n} \mathrm{cov}(x_j,x_k) \tag{4}$$

where the first term of [Eqn (4)] is the sum of the variances of the n components of x [Eqn (1)] and the second term is the sum of the covariances between the n components of x.

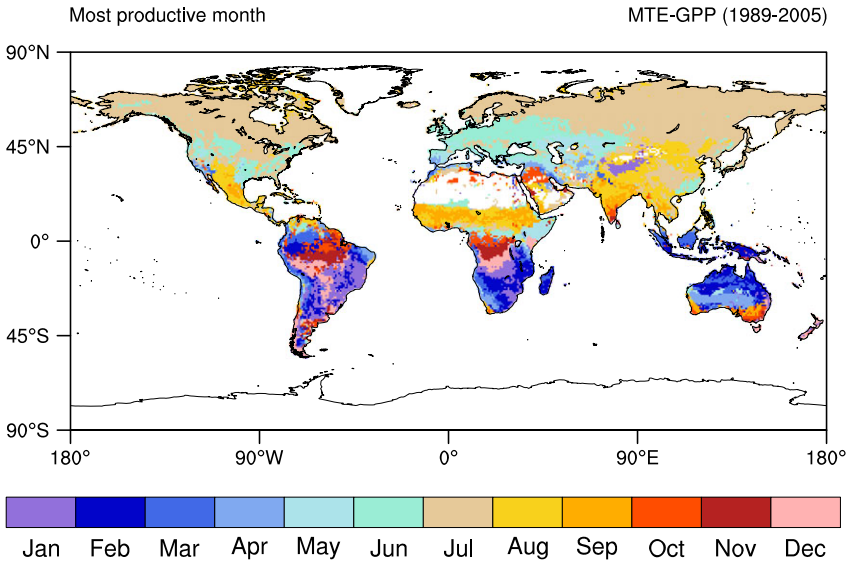

Fig. 1 Spatial distribution of the most productive month based on the MTE-GPP (1989–2005).

Results

Evaluation of CMIP5 models (GPP, ET, Precipitation and Temperature)

The spatial distribution of present-day GPP as well as the CMIP5 multimodel mean bias in GPP over the 1989–2005 period is shown in Fig. 2a, e for the most productive month. Over most of the world's areas, the CMIP5 models generally tend to overestimate GPP as compared to the MTE product, as already reported in previous studies for the global average (Anav *et al.*, 2013; Piao *et al.*, 2013). This overestimation of GPP is more pronounced over Western North America (WNA), Central and Southern Africa and Eastern Asia (EAS), where the models tend to overestimate GPP by

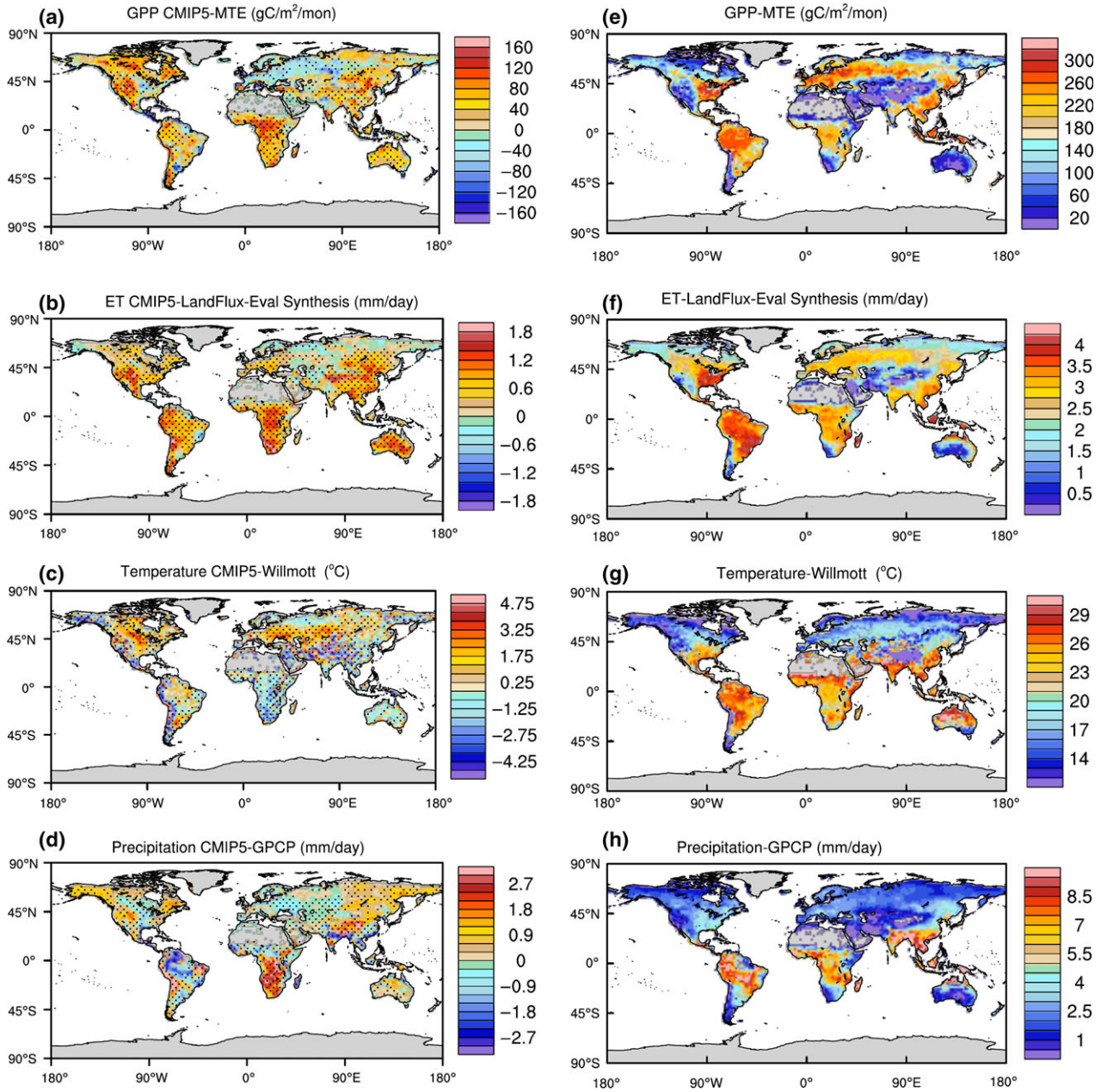

Fig. 2 Mean bias between (a) CMIP5 (multimodel mean) and MTE upscaled GPP (gC m^{-2} $month^{-1}$), (b) CMIP5 (multimodel mean) and LandFlux-Eval ET (mm day^{-1}), (c) CMIP5 (multimodel mean) and University of Delaware temperature (°C) and (d) CMIP5 (multimodel mean) and GPCP precipitation (mm day^{-1}). Stippling is applied over regions where at least 66% of the models (i.e., 13 of 19) agree on the sign of the difference (1989–2005). The spatial distribution of GPP, ET, temperature and precipitation based on reference products is also shown (e–h). Results are shown only for the most productive month.

more than 50%. But we find also regions with negative bias, that is, the models tend to underestimate GPP by more than 20% over Central North America (CNA) and Mexico, in the Southern Amazon and Eurasia. Interestingly, in most regions, the sign of the bias is consistent across models, indicating that these regional biases are systematic features of the ESMs. The bias in ET (Fig. 2b) follows a similar pattern as GPP, that is, ET is generally overestimated with the notable exception of Eurasia where the models tend to overestimate ET and underestimate GPP. The identified regional differences in the sign and the magnitude of the bias support our approach on constraining at grid point scale as some models might simulate the correct global GPP because of compensating regional biases. Overall, the systematic overestimation of GPP and ET on global scale in CMIP5 models possibly points to shortcomings in current climate models (Anav *et al*., 2013; Mueller & Seneviratne, 2014).

The biases in simulated GPP and ET are caused by any combination of wrong parameters, deficiencies in model structure or biases in the climate forcing. In particular, biases in climate (Fig. 2e–h) play an important role as indicated by the regional correlations between climate biases (temperature and precipitation) and biases in land fluxes (GPP and ET). For instance, the cold bias in temperature over Alaska and Northern Asia (Fig. 2c) might explain the underestimation of GPP. This interpretation is strengthened by the fact that temperature is the main limiting factor of terrestrial ecosystem productivity in northern latitudes (Myneni *et al*., 1997; Menzel & Fabian, 1999; Zhou *et al*., 2001; Lucht *et al*., 2002; Gong & Ho, 2003; Linderholm, 2006; Piao *et al*., 2007, 2008; Jeong *et al*., 2011). Also, Fig. 2d shows that CMIP5 models underestimate precipitation over large parts of Eurasia, Central North America, Central America and Mexico and Amazonia. It is therefore likely that the underestimation of GPP in these regions, where productivity is limited by water (Churkina & Running, 1998; Nemani *et al*., 2003), is related to underestimation of precipitation. Moreover, wet biases in precipitation over Africa might partially explain the overestimation of GPP, because in this region, the productivity of ecosystems is limited by water availability (Camberlin *et al*., 2007). Interestingly, the biases in GPP and precipitation over Amazon have the opposite sign despite the strong water dependence of productivity (Phillips *et al*., 2009; Gatti *et al*., 2014) (results for the annual mean show similar behavior, not shown). Enhanced light availability induced by decreased cloud cover might explain the overestimation of GPP over parts of Amazon (Graham *et al*., 2003). Also, other important limitations on productivity like nutrient availability could possibly explain the biases in GPP over Amazon, although this result should not be overinterpreted as the MTE product is only weakly constrained by observations in that region given the lack of reference FLUXNET sites (there and in the tropics in general). As for nutrient limitations, we note that models including nitrogen limitations on productivity (CESM, CCSM4 and NorESM) do not exhibit a significantly different behavior compared to other models not including this process (Fig. S2). Also, crop representation in LSMs (Bondeau *et al*., 2007) drastically affects the terrestrial carbon cycle and insufficient representation of agricultural processes might explain the underestimation of GPP over parts of Europe and Central North America, where croplands are a dominant part of the landscape (Leff *et al*., 2004).

It should be noted that despite the regional correlations between biases in climate and biases in GPP (Fig. 2 and also shown by Anav *et al*., 2015), there is no clear relationship across models between absolute temperature/precipitation and global GPP for present day ($r = 0.08$ and $r = -0.53$ for temperature and precipitation, respectively). Thus, the existing biases in the physical forcing do not appear to have a major imprint on the simulated intermodel spread in GPP. Moreover, offline LSM simulations driven with observation-based forcing still show a large intermodel spread and large biases (e.g., Piao *et al*., 2013; Anav *et al*., 2015; Sitch *et al*., 2015). Together, these two findings imply that the present-day intermodel spread in GPP in the ESMs is largely driven by differences in the land surface models and only to a lesser degree by climate (although the climate biases certainly amplify this spread). Also, it should be noted that in the majority (ca. 70%) of the land grid points, there is consistency in the CMIP5 models in terms of their present-day (1989–2005) performance in GPP and precipitation (see Appendix S1 and Fig. S9 in the Supporting Information). In other words, models that are doing well with regard to GPP should also do well with regard to the physical forcing, but this finding is certainly not enough to explain the large intermodel variations in GPP, given the issues discussed above.

Emergent Relationships

Most regions of the world exhibit positive intermodel correlations between present-day GPP and ET (Fig. 3a). In other words, models simulating higher present-day GPP also tend to simulate higher present-day ET. This positive correlation between GPP and ET is expected because of the tight coupling between transpiration and photosynthesis at the plant level (Cowan & Farquhar, 1977; Collatz *et al*., 1991; Sellers *et al*., 1997 Berry *et al*., 2010), despite the relatively simplified representation of

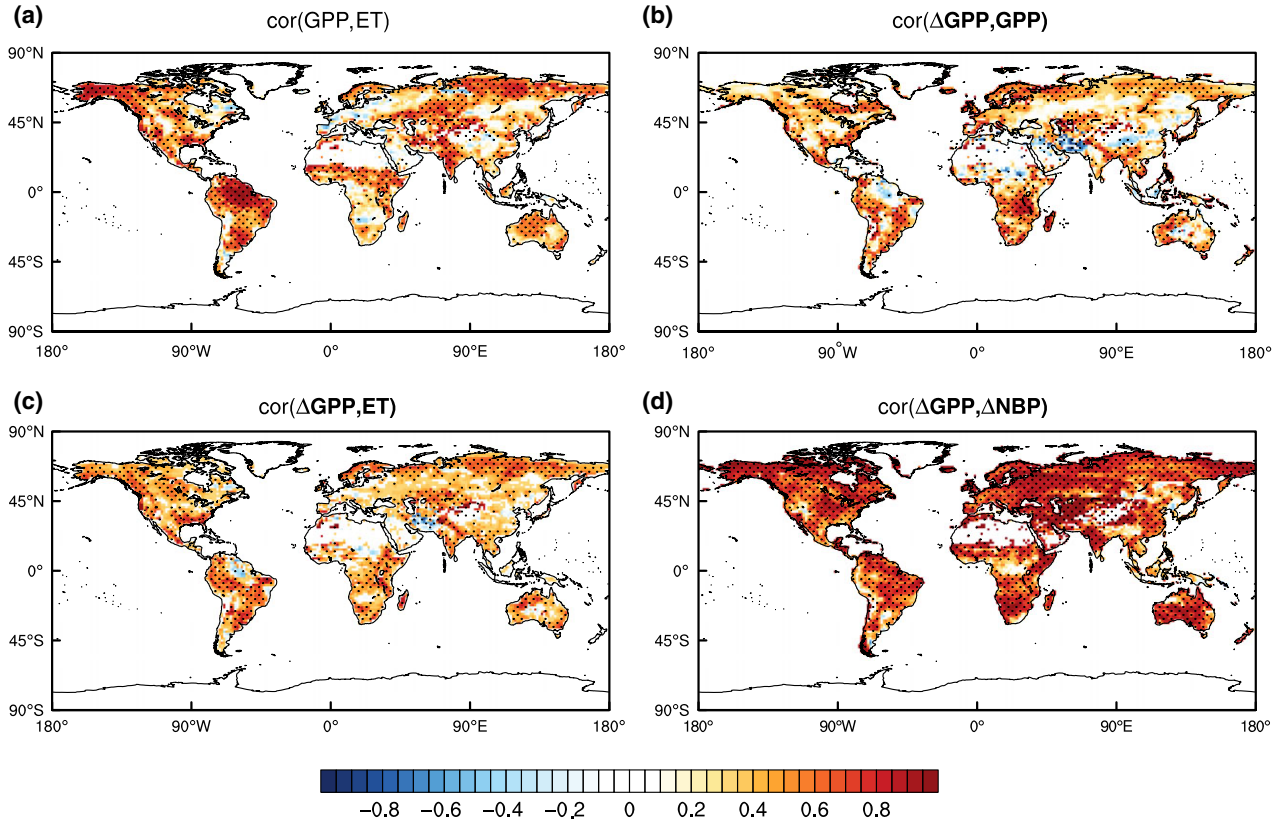

Fig. 3 Intermodel correlation between: (a) GPP and ET during the most productive month over the period 1989–2005, (b) ΔGPP (2084–2100 minus 1989–2005) and GPP (1989–2005), (c) ΔGPP and ET, and (d) ΔGPP and ΔNBP. The correlation is estimated based on a sample of 19 models. Correlation is shown only over regions where annual GPP is more than 10 gC m^{-2} yr^{-1}. Dots indicate regions where the correlation is statistically significant (5% significance level).

stomatal behavior used in current ESMs (Ball–Berry model). Over many regions, models simulating higher present-day GPP and ET tend to simulate also a higher increase in future GPP (Fig. 3b, c). Moreover, a near linear relationship exists in the models between ΔGPP and ΔNBP, that is, the change in GPP or NBP between 1989–2005 and 2084–2100 (Fig. 3d). This relationship means that models simulating higher increase in future GPP tend to simulate higher change in future NBP. This means that, at the regional scale, the future evolution of the land carbon sink in the CMIP5 models primarily depends on the evolution of photosynthetic assimilation implying that observation-based constraints on ΔGPP will also strongly constrain ΔNBP. The strong relationship between ΔGPP and ΔNBP also holds for other months (Fig. S3), although it tends to vanish in winter over a few northern regions. This is likely because of the high sensitivity of respiration to temperature change, which tends to decouple ΔNBP from ΔGPP. That said, the correlation between annual ΔGPP and ΔNBP at the global scale (Fig. S4) is still substantial and statistically significant ($r = 0.7$; $P < 0.0001$).

These emergent relationships between present-day GPP and ET and future changes in GPP (and NBP) combined with a definition of the uncertainty range in present-day GPP and ET can thus be used to constrain long-term projections of GPP (and NBP).

Global patterns of constrained GPP

We present here the results of the four constraint approaches (ETcon, GPP&ETcon, GPPcon and GPPglobal) we applied focusing on the most productive month. The characteristics of the constrained models (i.e., the percentage of land grid points where the simulated GPP and ET are not lying within the likely range of the MTE-GPP and the LandFlux-Eval ET) are shown in Fig. S5, while the number of retained models for the GPPcon and ETcon ensembles is shown in Fig. S6.

The spatial pattern of the difference between prior and constrained future change in GPP is displayed in Fig. 4 for the four different constrained ensembles. Over many regions, a remarkable although spatially quite heterogeneous impact of the constraint on the

future change in GPP is seen in the all spatially explicit constraint methods, that is, the ETcon, GPP&ETcon and GPPcon ensembles (Fig. 4a–c). The effect of the constraint is positive (i.e., higher increase in future GPP in the constrained ensemble compared to the prior) mainly over northern and mid-latitudes, while the effect in the tropics is generally negative (i.e., smaller change in future GPP compared to the prior ensemble). One feature worth mentioning is the strong increase signal over Central Europe and Eurasia. Despite regional differences, the different constraint approaches lead to qualitatively similar results. This finding reflects the strong correlation between GPP and ET at the regional scale discussed in the previous section.

Given the correlations highlighted in the previous section, the sign of the effect of the constraint on the future change in GPP directly reflects the sign of the bias in present-day GPP. In other words, the effect of the constraint tends to be positive (higher increase in future GPP) over regions where the models tend to underestimate the present-day GPP based on the MTE dataset.

Constraining the future change in GPP based on the models' performance in simulating global GPP (Fig. 4d) shows a spatially smoother picture. In this case, the effect of the constraint is negative in most of the word's regions and the models are simulating on average a decrease in the future change in GPP (i.e., the increase in future GPP is homogeneously lower in the constrained ensemble – also shown in Fig. S7).

Global upscaling-GPP

To constrain annual GPP on a global scale, we applied the various constraints on each month individually and then aggregated the results to the annual scale (results based on constraining directly the annual mean values instead of each month individually are also given in Table S3 as a sensitivity test and indicate similar results). The prior and the constrained global annual GPP based on the different constraint approaches (ETcon, GPP&ETcon, GPPcon and GPPglobal) are shown in Fig. 5a, b.

The constraints result in a relatively consistent and discernible reduction in both mean GPP and the

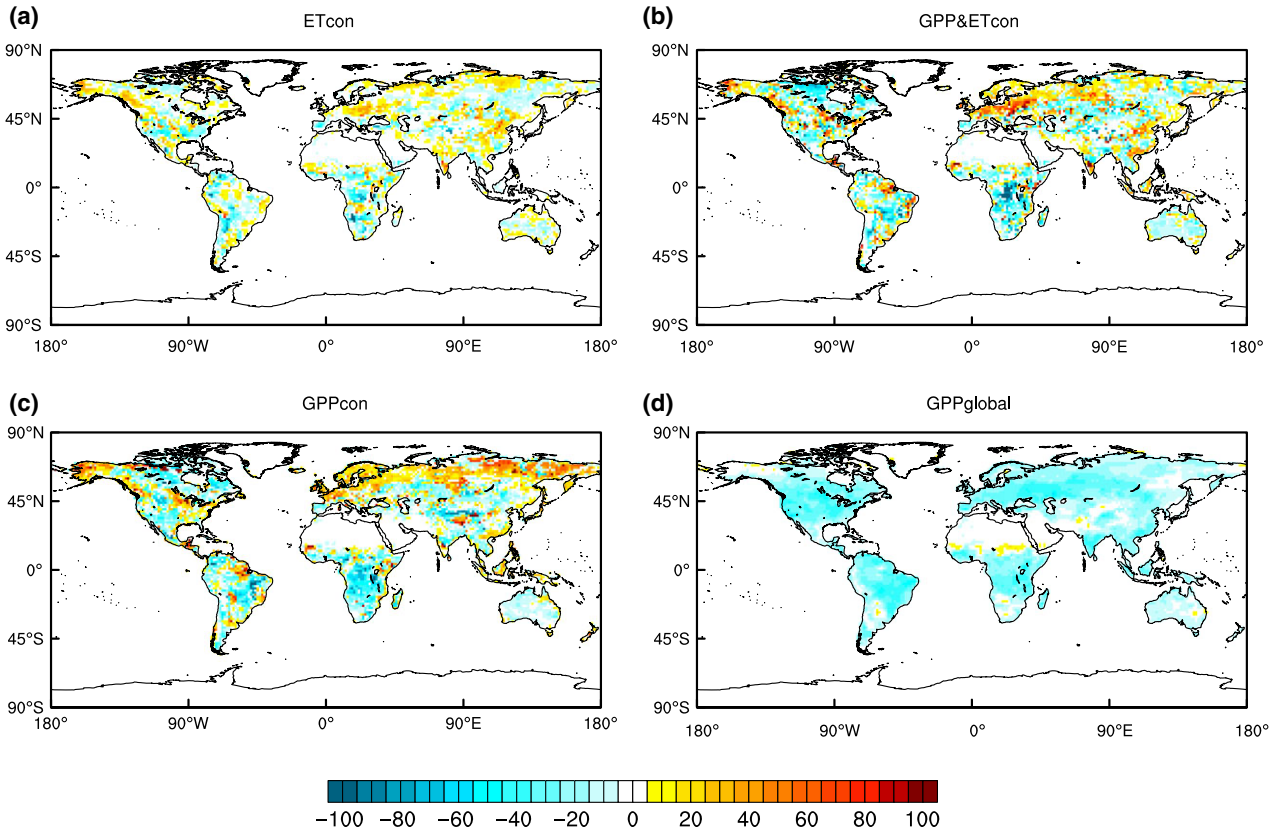

Fig. 4 Effect of the different constraints on the future change in GPP_{max} (gC m^{-2} $month^{-1}$). The effect is defined as the difference in the future change (2084–2100 minus 1989–2005) in GPP_{max} between the prior and the constrained ensembles. Positive values indicate higher increase in future GPP in the constrained ensemble compared to the prior ensemble.

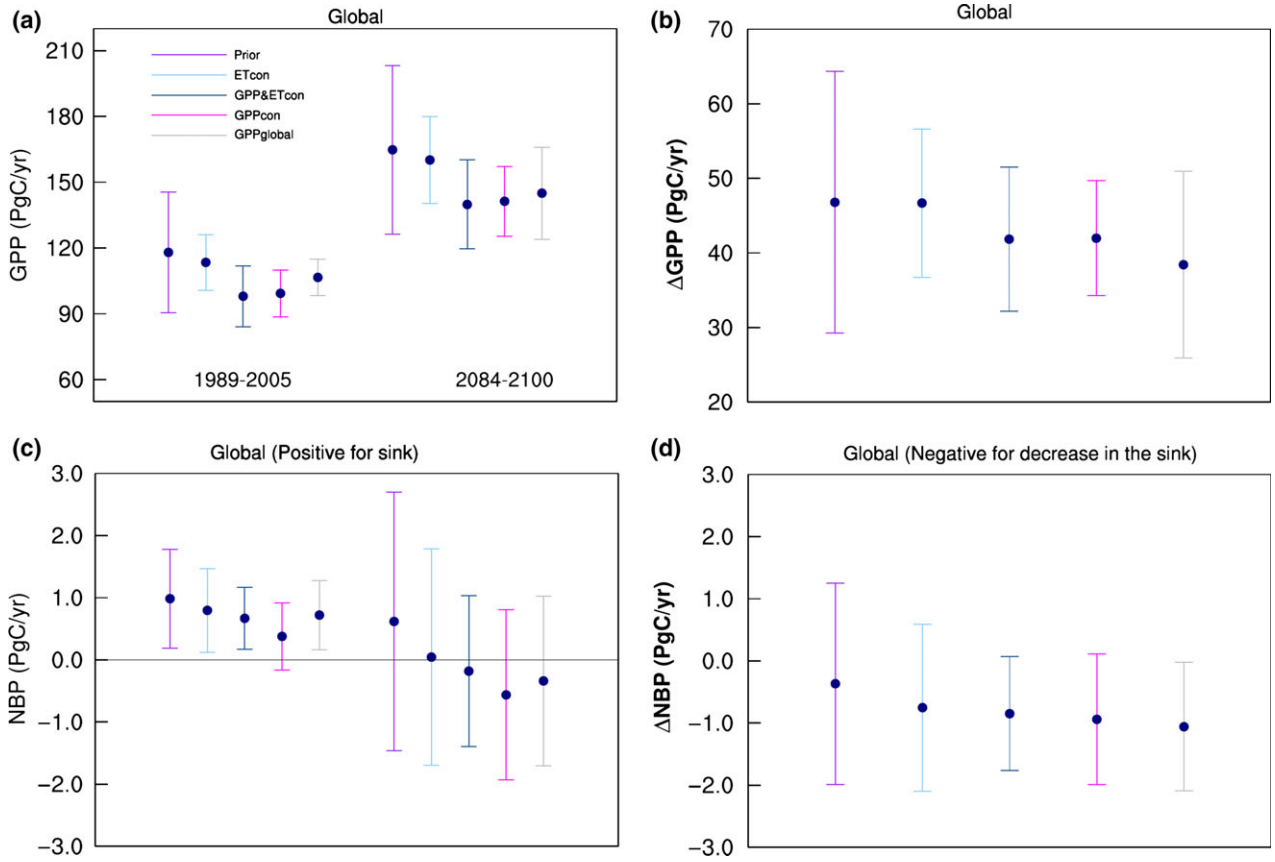

Fig. 5 (a) Prior and constrained mean annual global GPP (PgC yr^{-1}) for two time horizons: 1989 to 2005 and 2084 to 2100, (b) prior and constrained change in annual mean global GPP (ΔGPP, PgC yr^{-1}), (c) prior and constrained mean annual global NBP (PgC yr^{-1}) for two time horizons: 1989 to 2005 and 2084 to 2100 and (d) prior and constrained change in annual mean global NBP (ΔNBP, PgC yr^{-1}). The change is defined as the difference between (2084–2100) and (1989–2005). Error bars correspond to model uncertainty expressed as 1std of the multimodel mean. Negative numbers for NBP and ΔNBP correspond to carbon source and decrease in future land sink, respectively. (The estimates are representative only for about 80% of the total land area as the same number of grid points is compared for the different constrained ensembles).

Table 2 Prior and constrained estimates of global annual GPP, ΔGPP and trend in GPP

Name	GPP_{20th} (PgC yr^{-1})	GPP_{21st} (PgC yr^{-1})	ΔGPP (PgC yr^{-1})	$Trend_{2071–2100}$(PgC yr^{-2})
Prior	118 ± 27.5	164.8 ± 38.4	46.8 ± 17.5	0.49 ($P < 0.01$)
ETcon	113.4 ± 12.7	160 ± 19.8	46.6 ± 9.9	0.49 ($P < 0.01$)
GPP&ETcon	99.3 ± 10.6	139.8 ± 20.2	40.5 ± 9.6	0.45 ($P < 0.01$)
GPPcon	98 ± 13.8	141.3 ± 15.9	43.3 ± 7.7	0.45 ($P < 0.01$)
GPPglobal	106.6 ± 8.37	145 ± 21	38.4 ± 12.5	0.41 ($P < 0.01$)

intermodel spread for both periods, that is, 1989–2005 and 2084–2100. While the contemporary annual global GPP based on the CMIP5 (prior) ensemble is 118 ± 27.5 PgC yr^{-1} (Table 2), the constrained values over the same period are 113.4 ± 12.7 PgC yr^{-1}, 99.3 ± 10.6 PgC yr^{-1}, 98 ± 13.8 PgC yr^{-1} and 106.6 ± 8.37 PgC yr^{-1} for the ETcon, GPP&ETcon, GPPcon and GPPglobal constraint approaches, respectively, meaning that the prior multimodel mean GPP is reduced by 5–15% depending on the constraint approach. This is in line with the fact that the multimodel prior ensemble overestimates global GPP. Moreover, the intermodel spread is reduced by about 50% in all constrained ensembles.

Over the future period (2084–2100), the prior multimodel ensemble simulates a global mean annual GPP of

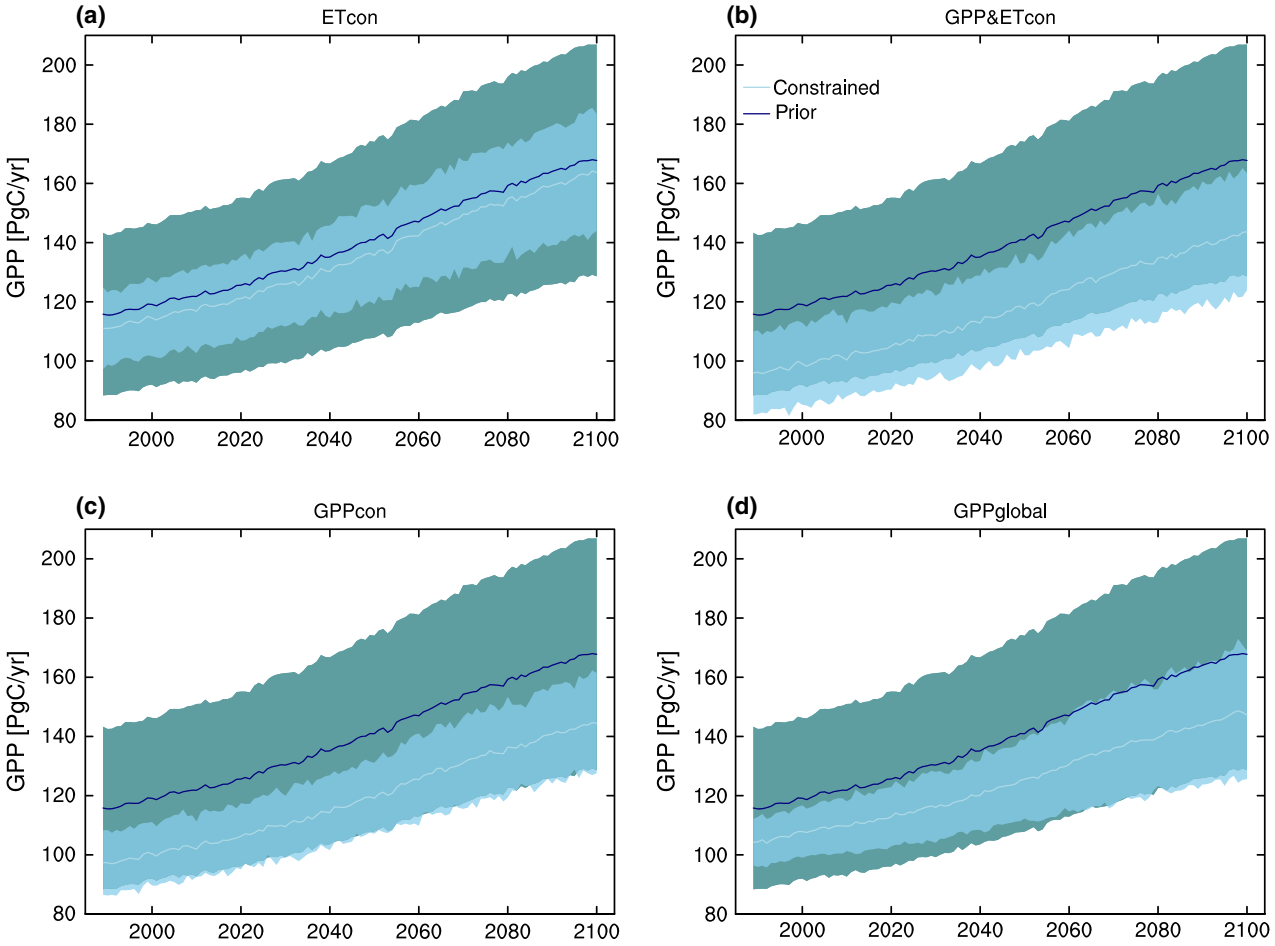

Fig. 6 Global annual GPP over the period 1989–2100 (RCP8.5) for the prior (navy blue) and the four different constrained (light blue) ensembles: (a) ETcon, (b) GPP&ETcon, (c) GPPcon and (d) GPPglobal. Shaded areas represent the model spread expressed as one standard deviation of the multimodel mean for the prior (cadet blue) and theconstrained (sky blue) ensemble.

Table 3 Prior and constrained estimates of global annual NBP (positive for sink), ΔNBP (negative for decrease in the sink) and trend in NBP (the trend is estimated over the period 2071–2100)

Name	NBP_{20th}(PgC yr^{-1})	NBP_{21st} (PgC yr^{-1})	ΔNBP (PgC yr^{-1})	$Trend_{2071–2100}$ (PgC yr^{-2})
Prior	0.98 ± 0.79	0.62 ± 2.1	−0.36 ± 1.62	−0.04 ($P < 0.01$)
ETcon	0.79 ± 0.49	0.04 ± 1.73	−0.75 ± 0.92	−0.04 ($P < 0.01$)
GPP&ETcon	0.66 ± 0.67	−0.18 ± 1.21	−0.84 ± 1.34	−0.037 ($P < 0.01$)
GPPcon	0.37 ± 0.54	−0.56 ± 1.36	−0.93 ± 1.05	−0.027 ($P < 0.01$)
GPPglobal	0.72 ± 0.56	−0.34 ± 1.37	−1.06 ± 1.03	−0.028 ($P < 0.01$)

164.8 ± 38.4 PgC yr^{-1}. Thus, all models agree that GPP would increase in a future climate with higher CO_2 concentration, but the model spread in GPP is considerably larger during the future period. The constrained GPP over the future period is 160 ± 19.8 PgC yr^{-1}, 139.8 ± 20.2 PgC yr^{-1}, 141.3 ± 15.9 PgC yr^{-1} and 145 ± 21 PgCyr for the ETcon, GPP&ETcon, GPPcon and GPPglobal constraint approaches, respectively. The effect of the constraints on mean GPP and model spread in future is therefore of the same magnitude as in the historical period. The increase in GPP (future minus past) is 46.8 ± 17.5 PgC yr^{-1} in the prior ensemble and slightly smaller in the constrained ensembles, that is, 46.6 ± 9.9 PgC yr^{-1}, 40.5 ± 9.6 PgC yr^{-1}, 43.3 ± 7.7 PgC yr^{-1} and 38.4 ± 12.5 PgC yr^{-1} in the ETcon, the GPP&ETcon, GPPcon and GPPglobal constrained

ensembles, respectively (Fig. 5b). The latter result shows that whereas there might be regional differences on the effect of the constraint on the future change in GPP between the GPP&ETcon and the GPPglobal constraint methodology (as seen on Fig. 4), the overall effect on the future change is comparable between the two different approaches. Moreover, the prior spread of 17.5 PgC yr^{-1} in the projected increase in GPP is reduced between about 30% and 56% depending on the constraint approach.

The temporal evolution of global annual GPP for the prior and the constrained ensembles over the period 1989–2100 is displayed in Fig. 6. Over the period 2071–2100, Fig. 6 shows a significant increasing trend (5% significance level) in GPP in both the prior (0.49 PgC yr^{-2}, $P < 0.01$) and the different constrained ensembles. The constrained ensembles indicate a decreased but still positive significant trend in GPP. Specifically, the trend is slightly decreased by 8% (0.45 PgC yr^{-2}, $P < 0.01$), 8% (0.45 PgC yr^{-2}, $P < 0.01$) and 16% (0.41 PgC yr^{-2}, $P < 0.01$) in the GPP&ETcon, GPPcon and GPPglobal constrained ensembles, respectively, while the ETcon ensemble shows no overall

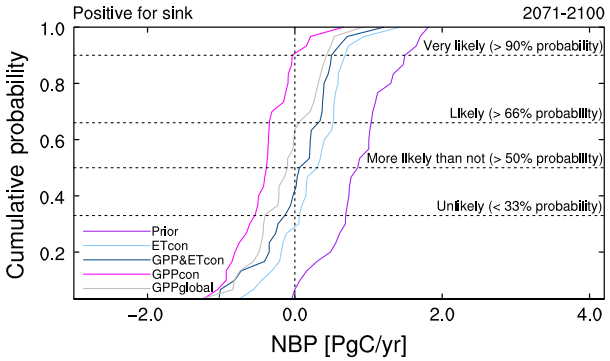

Fig. 8 Cumulative probability distribution for global annual NBP for the prior and the constrained ensembles for the period 2071–2100. Negative numbers correspond to carbon source. The likelihood ranges are defined according to IPCCs guidelines.

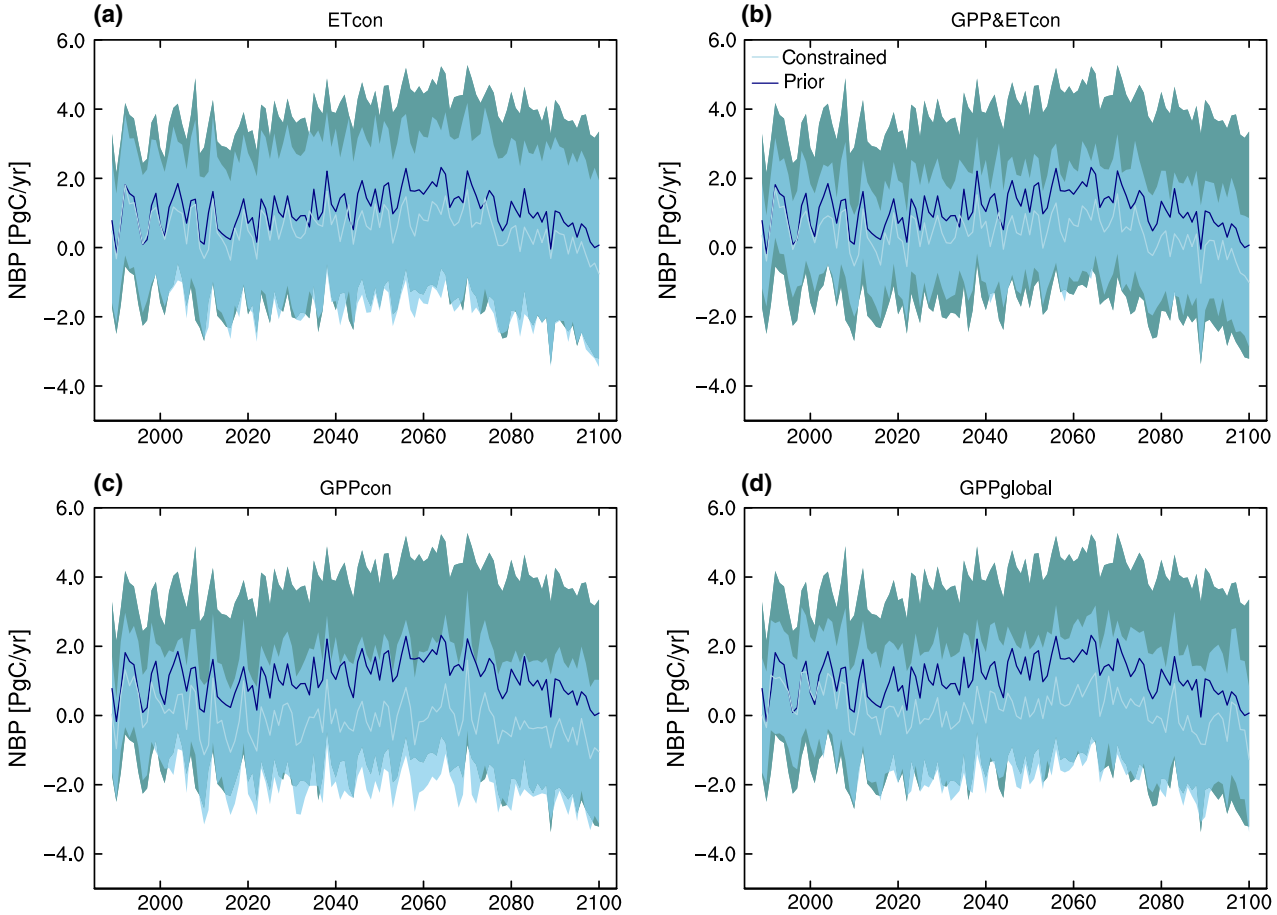

Fig. 7 Global NBP over the period 1989–2100 (RCP8.5) for the prior (navy blue) and the four different constrained (light blue) ensembles: (a) ETcon, (b) GPP&ETcon, (c) GPPcon and (d) GPPglobal. Shaded areas represent the model spread expressed as one standard deviation of the multimodel mean for the prior (cadet blue) and the constrained (sky blue) ensemble. Negative numbers correspond to carbon source.

change in the trend (0.49 PgC yr^{-2}, $P < 0.01$). Our results show remarkable agreement between the constrained ensembles on the reduction in the intermodel spread. This spread is at least 45–50% smaller in the constrained ensembles than the one in the prior, while the magnitude of this reduction is almost constant in time.

Results based on constraining the annual mean instead of each monthly individually are given in Table S3. These results are similar to the results based on the monthly based constraint approach.

Global upscaling-NBP

Given the strong positive correlation between future changes in GPP and in NBP inferred from the CMIP5 models, we applied the observational constraint approaches to NBP as well. The CMIP5 multimodel mean annual NBP over the period 1989–2005 is 0.98 ± 0.79 PgC yr^{-1} globally (carbon sink) (Table 3). Fig. 5c shows that the magnitude of this sink is reduced in all the constrained ensembles. Specifically, the constrained global NBP over the period 1986–2005 is 0.79 ± 0.49 PgC yr^{-1}, 0.66 ± 0.67 PgC yr^{-1}, 0.37 ± 0.54 PgC yr^{-1} and 0.72 ± 0.56 PgC yr^{-1} for the ETcon, GPP&ETcon, GPPcon and GPPglobal constraint approaches, respectively. Over the period 2084–2100, the prior multimodel ensemble indicates a net carbon sink of 0.62 ± 2.1 PgC yr^{-1}. Interestingly, the terrestrial biosphere seems to shift from a net carbon sink to a net carbon source in most of the constrained ensembles. Specifically, NBP based on GPP&ETcon is -0.18 ± 1.21 PgC yr^{-1}, -0.56 ± 1.36 PgC yr^{-1} for the GPPcon and -0.34 ± 1.37 PgC yr^{-1} for the GPPglobal constraints. In contrast, the terrestrial biosphere still acts as a small carbon sink in the ETcon (0.04 ± 1.73 PgC yr^{-1}) despite the substantial reduction in the magnitude of the net carbon uptake in this constrained ensemble. The model spread over the future period is reduced between 15 and 35% depending on the constraint approach, but it is larger compared to the historical period.

The change in NBP from the historical to the future period in the prior multimodel ensemble is -0.36 ± 1.62 PgC yr^{-1} (Fig. 5d) implying a decreasing land carbon sink. The constrained ensembles show a stronger decrease in the land carbon sink by the end of the 21st century. Specifically, the future change in NBP is -0.75 ± 0.92 PgC yr^{-1}, -0.84 ± 1.34 PgC yr^{-1}, -0.93 ± 1.05 PgC yr^{-1} and -1.06 ± 1.03 PgC yr^{-1} for the ETcon, GPP&ETcon, GPPcon and GPPglobal constrained ensembles. Furthermore, the prior spread of 1.62 PgC yr^{-1} in the future change in NBP is reduced between about 17% and 45% depending on the constraint approach. The magnitude of the land sink is substantially decreased in the 21st century in the different constrained ensembles and the terrestrial biosphere in most of the constrained ensembles is even turned into a net carbon source by the end of the century. The intermodel spread in ΔNBP, however, remains large in the GPP&ETcon constrained ensemble despite the substantial reduction in the magnitude of the net carbon sink.

Figure 7 shows time series of the net land to atmosphere carbon flux for the prior and the constrained ensembles over the period 1989–2100. Over the period 2071–2100, the prior model ensemble simulates a small significant decreasing trend (5% significance level) in NBP of -0.04 PgC yr^{-2} ($P < 0.01$), while over the same period, the constrained ensembles simulate a weaker decreasing trend of -0.037 PgC yr^{-2} ($P < 0.01$), -0.027 PgC yr^{-2} ($P < 0.01$) and -0.028 PgC yr^{-2} ($P < 0.01$) for the GPP&ETcon, GPPcon and GPPglobal constrained ensembles, while the ETcon ensemble shows (as in GPP) no change in the trend (-0.04 PgC yr^{-2}, $P < 0.01$). Therefore, the sign in the NBP trend over the future period in the constrained ensembles is consistent with the sign in the prior trend. Interestingly, the constrained ensembles show a less dramatic decreasing trend in future NBP despite the massive decrease in the magnitude of the land sink in these ensembles.

The cumulative probability distributions of the annual NBP over the period 2071–2100 in Fig. 8 reveal an increased probability of the land becoming a source of carbon to the atmosphere in the different constrained ensembles. While the likelihood of land source of carbon to the atmosphere is less than 10% in the prior ensemble, this likelihood increases to about 30% for the ETcon case and all the way up to 93% for the GPPcon case.

Discussion

In this study, we use observation-based fluxes of ET and GPP to constrain 21st century terrestrial carbon cycle projections from an ensemble of 19 earth system models used in the framework of the CMIP5 project. In contrast with previous studies (Cox *et al.*, 2013; Wenzel *et al.*, 2014), the use of global gridded products enables us to evaluate the models regionally and then constrain the land carbon projections both regionally and globally. The rationale underlying this strategy is that a given model might capture relatively well the globally averaged present-day GPP due to compensating regional biases, whereas another globally biased model might still perform well in some regions. On the basis of a globally averaged constraint, the first model would

be kept, despite having 'a right answer for a wrong reason' and the second model would be rejected, despite holding relevant information regionally. Applying the constraint in a spatially explicit way alleviates this issue and still enables us to aggregate results globally, although this approach also introduces additional challenges (e.g., characterization of the 'observational uncertainty' at the regional scale; possible mismatch in vegetation types between observation-based products and models). Moreover, it should be highlighted that this type of constraint makes use of 'more information' than the standard global approach where one model is either in or out of the ensemble, that is, none of the models is entirely discarded, each model has some skill in at least some regions, and this information contributes to the final results. Interestingly, the two independent spatially explicit constraints (GPPcon and ETcon) show in several ways qualitatively similar effects on the future change in terrestrial carbon cycle projections in terms of both regional changes and intermodel spread, although the ET constraint leads to a smaller reduction in the projected mean changes in GPP and NBP. We also show that constraining the models following commonly applied methodologies based on their performance in simulating quantities on global scale (e.g., global GPP) might lead to spurious spatially distributed constrained fluxes as the models might share regional compensating biases in simulating global GPP.

Uncertainties related to the observation-based products

One of the main challenges faced by such approaches relates to the reliability of the observation-based products. For instance, the MTE product is subject to uncertainties related to the representativeness of the FLUXNET network and the partitioning approach of net ecosystem exchange (NEE) to GPP and respiration (Reichstein *et al*., 2005; Lasslop *et al*., 2010). Its most robust features are believed to be the representation of spatial and seasonal variations, while there is less confidence in the interannual variability and trends of the product (Jung *et al*., 2011) as well as in the global integral. We doubled the uncertainty range in the reference products for GPP to define a more conservative estimate of GPP as it is anticipated that the published uncertainty range is likely underestimated. We decided against the use of alternative products for GPP for several reasons. For instance, while the satellite-based GPP from the Moderate Resolution Imaging Spectroradiometer (MODIS) (Zhao *et al*., 2005, 2006; Mao *et al*., 2012) is a potentially interesting alternative estimate for GPP, it covers an overly short period. For the same reason, we also did not use a recent spatially explicit estimate of GPP based on chlorophyll fluorescence (Frankenberg *et al*., 2011). Estimates of global GPP based on oxygen isotopes (Welp *et al*., 2011) are higher (150–175 PgC yr^{-1}) compared to the observation-based products from Jung *et al.* (2011) and Beer *et al.* (2010), but they do not provide spatiotemporal patterns of the flux which are necessary to our spatially explicit constraint approach. Moreover, the estimate from Welp *et al.* (2011) is subject to large uncertainties as they had only a limited number of observations available and also used a relative simple model (Anav *et al*., 2013). Bearing in mind the above uncertainties in the various estimates of GPP, we used the products from Jung *et al.* (2011) and Beer *et al.* (2010) as reference datasets for GPP. The LandFlux-Eval ET product (Mueller *et al*., 2013) has a large spread, as it considers different model products, diagnostic and reanalysis ET datasets. This large spread adds confidence to our methodology as the uncertainties associated with ET are rather over than underestimated. One of the major sources of uncertainties in this product might be related to the inclusion of model-based datasets. It should be noted that these model-based estimates are not found to be systematically less realistic than diagnostic ET datasets (Greve *et al*., 2014). Nevertheless, even when using the diagnostic datasets only as a constraint, the results go in the same direction as when using the full LandFlux-Eval ET dataset (Table S3).

Constrained GPP projections

The application of the constraints reduces not only the intermodel range in GPP by about 50%, but it reduces also the increase in the simulated future GPP. This elevated GPP under the future scenario (RCP8.5) compared to present-day GPP reflects a fertilization effect due to higher atmospheric CO_2 concentration depending on the nature of future CO_2 scenario applied (here RCP8.5) and possibly also the positive effect of climate change in some regions (e.g., warmer and longer growing season at high latitudes) (Cramer *et al*., 2001; Friedlingstein *et al*., 2006; Piao *et al*., 2007; Schimel *et al*., 2015; Xia *et al*., 2015). The finding of a reduction in the constrained global GPP is consistent with results from recent studies (Anav *et al*., 2013; Piao *et al*., 2013) confirming that the state-of-the-art LSMs tend to overestimate GPP – at least when compared to the MTE product. It is not possible to identify potential drivers of this overestimation with the prescribed CO_2 simulations used in this study as these simulations include both the effect of increased CO_2 and the effect of climate change on the terrestrial carbon cycle. Nevertheless, nutrient availability might be considered as a prime candidate for the high response to CO_2 and the

overestimation of GPP in CMIP5 models as most of the CMIP5 models (except CCSM4.0, CESM and NorESM) are ignoring nutrient limitations on plant productivity. The availability of these nutrients substantially affects the carbon uptake by plants (Norby *et al*., 2010; Vitousek *et al*., 2010; De Vries, 2014; Fernández-Martínez *et al*., 2014) and may even accelerate future climate change (Zaehle *et al*., 2010). Also, it is found that nutrient availability reduces the response to CO_2 in contemporary carbon cycle simulations (Thornton *et al*., 2007; Zhang *et al*., 2011, 2013) although these constrained simulated responses might still subject to large uncertainties (Thomas *et al*., 2013). It should be noted that Sun *et al.* (2014) argued recently that GPP might be in fact underestimated because of the inadequate representation of CO_2 diffusion inside leaves. But results from free-air CO_2 enrichment (FACE) experiments (Ainsworth & Long, 2005; Norby, 2011) in a limited number of temperate forest sites suggest the opposite (Piao *et al*., 2013). Thus, for the moment, we consider the MTE product as the best available estimate for constraining GPP, recognizing its uncertainties.

The effects on the projected mean change in GPP are less consistent between ET- vs. GPP-based constraints, with almost not change for ETcon and a ca. 20% reduction for the GPP-based constraints (GPPcon, GPP&ETcon and global). Using the median of the models instead of the mean shows similar results of the effect of the different constraints on the global annual GPP (Table S2). Consequently, our results are statistically robust independently of the chosen statistics.

Constrained NBP projections

We find a high correlation in CMIP5 models between future changes in GPP and in NBP implying that simulated future changes in NBP are primarily driven by changes in GPP. This model-derived correlation between ΔGPP and ΔNBP is also supported by experimental data suggesting that interannual variations in GPP drive anomalies in NBP (Luyssaert *et al*., 2007; Reichstein *et al*., 2007; Le Maire *et al*., 2010; Jung *et al*., 2011). However, this correlation is remarkable given the fact that NBP includes the effect of various disturbances (e.g., land use changes, and fires, grazing, harvesting), which are represented in very different ways across models and are known to play a critical role in determining the magnitude of the terrestrial carbon balance (Sitch *et al*., 2005; Bond-Lamberty *et al*., 2007; Chambers *et al*., 2007; Kurz *et al*., 2008; Frolking *et al*., 2009; Lindroth *et al*., 2009; Deb Richter & Houghton, 2011; Houghton *et al*., 2012; Brovkin *et al*., 2013). It should be noted here that land use trends are lower in all future scenarios compared to the historical period. In fact, scenarios have several land management changes affecting NBP and GPP (e.g., bioenergy forests in RCP8.5), but most ESMs did not distinguish between managed and primary forests and simply prescribed maps of forest area. This might also contribute to the high positive correlation between ΔGPP and ΔNBP. Moreover, future changes in NBP may be influenced by increased decomposition owing to warmer temperature (Kätterer *et al*., 1998; Cox *et al*., 2000; Rustad *et al*., 2001; Knorr *et al*., 2005; Davidson & Janssens, 2006; King *et al*., 2006) and by changes in the turnover time of carbon in ecosystems (Carvalhais *et al*., 2014; Friend *et al*., 2014).

We need to consider also the caveats. The identified relationship between ΔGPP and ΔNBP is purely based on models and cannot be evaluated by observations. Moreover, it is clear that the existence of a near linear relationship in the models does not guarantee that this relationship also holds in reality. For instance, Wang *et al.* (2014) found that current land carbon cycle models do not capture the observed increase in the sensitivity of tropical ecosystems to interannual variations in temperature over the last 50 years, possibly due to underestimation of soil moisture effects on the simulated plant productivity. However, this is a fundamental limitation associated with any proposed emergent constraint linking long-term changes in the earth system to short-term observations of different aspects of the earth system (Allen & Ingram, 2002; Hall & Qu, 2006; Cox *et al*., 2013; Wenzel *et al*., 2014). The confidence in these linkages emerges only through their robustness across models of very different nature and construction, employing very different assumptions, and through process-based arguments. In the case of the relationship between ΔGPP and ΔNBP, we note that the correlation is very strong across all models (Fig. 3d) and that the there is a direct mechanistic link between the two processes, with changes in GPP naturally leading to changes in NBP. Thus, we consider the relationship as robust and thus proceed to the discussion of the implications of the observation-based constraints on NBP, nevertheless keeping in mind the limitations.

During the historical period (1989–2005), the constrained ensembles suggest weaker carbon sink compared to the prior. The magnitude of this sink is in agreement with estimates of the contemporary net land carbon uptake and it is lying in the range of 1.1 ± 0.9 PgC yr^{-1} (period 1990–1999; Ciais *et al*., 2013). The ensemble means presented here (prior and constrained) are indeed contained within (the lower half of) this range. Note, however, the different periods considered and the different land area (our global estimates are representative for only 80% of the land area

due to methodological reasons explained in the method section), which can also explain some differences. Furthermore, it should also be noted that our constraint approach does not induce any imbalance in the land carbon budget as the same set of models is retained for both the GPP and NBP estimates presented here, thus preserving the internal land carbon balance. In this sense, the lower NBP in the constrained estimates is a direct consequence of the lower GPP.

Our analysis indicates that the expected decline in the future land carbon sink (Friedlingstein *et al*., 2006; Ciais *et al*., 2013) may in fact be higher than previously thought, associated with a larger probability for the terrestrial biosphere turning into a net carbon source by the end of the century. This decrease in the ability of terrestrial ecosystems to absorb carbon in the constrained ensembles can be tracked back to the lower increase in GPP simulated by the constrained models.

It is important to note that the estimates of the constrained fluxes might still be conservative, while the uncertainties are only conditional to the processes actually represented in the models as many processes are in fact missing in current ESMs. These may include: (i) uncertainties in future vegetation cover (Arneth, 2015; Campioli *et al*., 2015), (ii) cropland representation and terrestrial ecosystem management practices including soil fertility (Vicca *et al*., 2012), (iii) accelerated atmospheric nitrogen deposition (Galloway *et al*., 2004), (iv) competition processes (Arora & Boer, 2006), (v) heterogeneous shift in the geographical range of plants (Kelly & Goulden, 2008), (vi) changes in the carbon assimilation pathway (Gowik & Westhoff, 2011), (vii) changes in carbon allocation to the different vegetation and soil pools in response to climate change (Sevanto & Dickman, 2015), (viii) permafrost carbon feedbacks (Koven *et al*., 2011, 2015) and (ix) plant–insect interactions related to response of population size and phenology to climate change (Scriber, 2011; Peñuelas *et al*., 2013). Some of these processes (e.g., nitrogen deposition, land cover changes) have determined the current net land carbon fluxes (Ciais *et al*., 2013), and they also critically have the potential to determine the future fate of the land sink.

Outlook

Recent estimates suggest that we have already used about 2/3 of our carbon quota for the 2 °C warming target relative to preindustrial levels and at the moment the anthropogenic carbon emissions track one of the highest Representative Concentration Pathway (RCP8.5) (Peters *et al*., 2012; Friedlingstein *et al*., 2014b; Raupach *et al*., 2014; Le Quéré *et al*., 2015). The identified decreased magnitude of the future land carbon sink in the constrained ensembles implies that the carbon emissions might need to be reduced more than previously thought in order to keep global warming below the 2 °C targets (Meinshausen *et al*., 2009; Zickfeld *et al*., 2009).

We conclude that observation-based constraints are a promising avenue to narrow down the spread in future land carbon projections. This approach is complementary (and not a surrogate) to model evaluation and improvement and is encouraged by the apparent lack of reduction in climate projection uncertainty despite decades of efforts to improve climate models. The constrained terrestrial carbon fluxes might potentially help in constraining other features of the carbon cycle (e.g., ocean carbon sink, atmospheric CO_2) and might offer a window in reducing uncertainties in other features of future climate (e.g., air temperature). Therefore, our findings have several important implications on the terrestrial carbon balance, the fate of atmospheric CO_2 concentrations and future climate change.

Acknowledgments

We acknowledge funding from the Swiss National Science Foundation (SNSF) as part of the 'CarboCount CH' Sinergia Project (grant CRSII2 136273), from EU projects 283080 (GEOCARBON (N.G.)) and 282672 (EMBRACE (S.I.S., E.D.)), both funded by the European Commission's 7th Framework Programme and by the ERC Consolidator Grant DROUGHT-HEAT project (S.I.S.). We would like to thank Nuno Carvalhais, Pierre Friedlingstein and Philippe Ciais for providing constructive comments on the manuscript. We also thank Urs Beyerle (IAC) for support with downloading and storage of CMIP5 data. We also would like to thank the numerous data providers, especially Martin Jung for producing and providing the MTE dataset. The GPCP combined precipitation data were developed by the NASA/Goddard Space Flight Center's Laboratory for Atmospheres as a contribution to the GEWEX Global Precipitation Climatology Project. Willmott air temperature (University of Delaware) data were provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at http://www.esrl.noaa.gov/psd/.

References

Adler RF, Huffman GJ, Chang A *et al.* (2003) The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979-present). *Journal of Hydrometeorology*, **4**, 1147–1167.

Ainsworth EA, Long SP (2005) What have we learned from 15 years of free-air CO2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO_2. *New Phytologist*, **165**, 351–372.

Allen MR, Ingram WJ (2002) Constraints on future changes in climate and the hydrologic cycle. *Nature*, **419**, 224–232.

Anav A, Friedlingstein P, Kidston M *et al.* (2013) Evaluating the land and ocean components of the global carbon cycle in the CMIP5 earth system models. *Journal of Climate*, **26**, 6801–6843.

Anav A, Friedlingstein P, Beer C *et al.* (2015) Spatiotemporal patterns of terrestrial gross primary production: a review. *Reviews of Geophysics.*, **53**, 785–818.

Arneth A (2015) Climate science: uncertain future for vegetation cover. *Nature*, **524**, 44–45.

Arora VK, Boer GJ (2006) Simulating competition and coexistence between plant functional types in a dynamic vegetation model. *Earth Interactions*, **10**, 1.

Arora VK, Boer GJ, Friedlingstein P *et al.* (2013) Carbon-concentration and carbon-climate feedbacks in CMIP5 earth system models. *Journal of Climate*, **26**, 5289–5314.

Baldocchi D, Falge E, Gu L *et al.* (2001) FLUXNET: a new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bulletin of the American Meteorological Society*, **82**, 2415–2434.

Ballantyne AP, Alden CB, Miller JB, Tans PP, White JWC (2012) Increase in observed net carbon dioxide uptake by land and oceans during the past 50 years. *Nature*, **487**, 70–72.

Beer C, Reichstein M, Tomelleri E *et al.* (2010) Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. *Science*, **329**, 834–838.

Berry JA, Beerling DJ, Franks PJ (2010) Stomata: key players in the earth system, past and present. *Current Opinion in Plant Biology*, **13**, 233–240.

Bodman RW, Rayner PJ, Karoly DJ (2013) Uncertainty in temperature projections reduced using carbon cycle and climate observations. *Nature Climate Change*, **3**, 725–729.

Bondeau A, Smith PC, Zaehle S *et al.* (2007) Modelling the role of agriculture for the 20th century global terrestrial carbon balance. *Global Change Biology*, **13**, 679–706.

Bond-Lamberty B, Peckham SD, Ahl DE, Gower ST (2007) Fire as the dominant driver of central Canadian boreal forest carbon balance. *Nature*, **450**, 89–92.

Booth BBB, Jones CD, Collins M *et al.* (2012) High sensitivity of future global warming to land carbon cycle processes. *Environmental Research Letters*, **7**, 1–8.

Brovkin V, Boysen L, Arora VK *et al.* (2013) Effect of anthropogenic land-use and land-cover changes on climate and land carbon storage in CMIP5 projections for the twenty-first century. *Journal of Climate*, **26**, 6859–6881.

Camberlin P, Martiny N, Philippon N, Richard Y (2007) Determinants of the interannual relationships between remote sensed photosynthetic activity and rainfall in tropical Africa. *Remote Sensing of Environment*, **106**, 199–216.

Campioli M, Vicca S, Luyssaert S *et al.* (2015) Biomass production efficiency controlled by management in temperate and boreal ecosystems. *Nature Geoscience*, **8**, 843–846.

Carvalhais N, Forkel M, Khomik M *et al.* (2014) Global covariation of carbon turnover times with climate in terrestrial ecosystems. *Nature*, **514**, 213–217.

Chambers JQ, Fisher JI, Zeng H, Chapman EL, Baker DB, Hurtt GC (2007) Hurricane Katrina's carbon footprint on U.S. Gulf Coast forests. *Science*, **318**, 1107.

Churkina G, Running SW (1998) Contrasting climatic controls on the estimated productivity of global terrestrial biomes. *Ecosystems*, **1**, 206–215.

Ciais P, Denning AS, Tans PP *et al.* (1997) A three-dimensional synthesis study of $\delta^{18}O$ in atmospheric CO_2 1. Surface fluxes. *Journal of Geophysical Research: Atmospheres*, **102**, 5857–5872.

Ciais P, Sabine C, Bala G, Bopp L *et al.* (2013) Carbon and other biogeochemical cycles. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Stocker TF, Qin D, Plattner G-K *et al.*), pp. 465–570. Cambridge University Press, Cambridge, UK and New York, NY, USA.

Collatz GJ, Ball JT, Grivet C, Berry JA (1991) Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a laminar boundary layer. *Agricultural and Forest Meteorology*, **54**, 107–136.

Collins M, Knutti R, Arblaster J *et al.* (2013) Long-term climate change: projections, commitments and irreversibility. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Stocker TF, Qin D, Plattner G-K *et al.*), pp. 1029–1136. Cambridge University Press, Cambridge, UK and New York, NY, USA.

Cowan IR, Farquhar GD (1977) Stomatal function in relation to leaf metabolism and environment. *Symposia of the Society for Experimental Biology*, **31**, 471–505.

Cox PM, Betts RA, Jones CD, Spall SA, Totterdell IJ (2000) Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature*, **408**, 184–187.

Cox PM, Pearson D, Booth BB, Friedlingstein P, Huntingford C, Jones CD, Luke CM (2013) Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability. *Nature*, **494**, 341–344.

Cramer W, Bondeau A, Woodward FI *et al.* (2001) Global response of terrestrial ecosystem structure and function to CO_2 and climate change: results from six dynamic global vegetation models. *Global Change Biology*, **7**, 357–373.

Davidson EA, Janssens IA (2006) Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. *Nature*, **440**, 165–173.

De Vries W (2014) Forest ecology: nutrients trigger carbon storage. *Nature Climate Change*, **4**, 425–426.

Deb Richter D Jr, Houghton RA (2011) Gross CO_2 fluxes from land-use change: implications for reducing global emissions and increasing sinks. *Carbon Management*, **2**, 41–47.

Fernández-Martínez M, Vicca S, Janssens IA *et al.* (2014) Nutrient availability as the key regulator of global forest carbon balance. *Nature Climate Change*, **4**, 471–476.

Frankenberg C, Fisher JB, Worden J *et al.* (2011) New global observations of the terrestrial carbon cycle from GOSAT: patterns of plant fluorescence with gross primary productivity. *Geophysical Research Letters*, **38**, 1–6. doi:10.1029/2011GL048738.

Friedlingstein P, Cox P, Betts R *et al.* (2006) Climate-carbon cycle feedback analysis: results from the C4MIP model intercomparison. *Journal of Climate*, **19**, 3337–3353.

Friedlingstein P, Meinshausen M, Arora VK, Jones CD, Anav A, Liddicoat SK, Knutti R (2014a) Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *Journal of Climate*, **27**, 511–526.

Friedlingstein P, Andrew RM, Rogelj J *et al.* (2014b) Persistent growth of CO_2 emissions and implications for reaching climate targets. *Nature Geoscience*, **7**, 709–715.

Friend AD, Lucht W, Rademacher TT *et al.* (2014) Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO_2. *Proceedings of the National Academy of Sciences of the USA*, **111**, 3280–3285.

Frolking S, Palace MW, Clark DB, Chambers JQ, Shugart HH, Hurtt GC (2009) Forest disturbance and recovery: a general review in the context of spaceborne remote sensing of impacts on aboveground biomass and canopy structure. *Journal of Geophysical Research: Biogeosciences*, **114**, 1–27.

Galloway J, Dentener F, Capone D *et al.* (2004) Nitrogen cycles: past, present, and future. *Biogeochemistry*, **70**, 153–226.

Gatti LV, Gloor M, Miller JB *et al.* (2014) Drought sensitivity of Amazonian carbon balance revealed by atmospheric measurements. *Nature*, **506**, 76–80.

Gong DY, Ho CH (2003) Detection of large-scale climate signals in spring vegetation index (normalized difference vegetation index) over the Northern Hemisphere. *Journal of Geophysical Research D: Atmospheres*, **108**, ACL 4-1–ACL 4-12.

Gowik U, Westhoff P (2011) The path from C3 to C4 photosynthesis. *Plant Physiology*, **155**, 56–63.

Graham EA, Mulkey SS, Kitajima K, Phillips NG, Wright SJ (2003) Cloud cover limits net CO_2 uptake and growth of a rainforest tree during tropical rainy seasons. *Proceedings of the National Academy of Sciences of the USA*, **100**, 572–576.

Greve P, Orlowsky B, Mueller B, Sheffiled J, Reichstein M, Seneviratne SI (2014) Global assessment of trends in wetting and drying over land. *Nature Geoscience*, **7**, 716–721.

Hall A, Qu X (2006) Using the current seasonal cycle to constrain snow albedo feedback in future climate change. *Geophysical Research Letters*, **33**, 1–4. doi:10.1029/2005GL025127.

Hoffman FM, Randerson JT, Arora VK *et al.* (2014) Causes and implications of persistent atmospheric carbon dioxide biases in Earth System Models. *Journal of Geophysical Research: Biogeosciences*, **119**, 141–162.

Houghton RA, House JI, Pongratz J *et al.* (2012) Carbon emissions from land use and land-cover change. *Biogeosciences*, **9**, 5125–5142.

Huntingford C, Lowe JA, Booth BBB, Jones CD, Harris GR, Gohar LK, Meir P (2009) Contributions of carbon cycle uncertainty to future climate projection spread. *Tellus, Series B: Chemical and Physical Meteorology*, **61 B**, 355–360.

Jeong SJ, Ho CH, Gim HJ, Brown ME (2011) Phenology shifts at start vs. end of growing season in temperate vegetation over the Northern Hemisphere for the period 1982-2008. *Global Change Biology*, **17**, 2385–2399.

Jung M, Reichstein M, Bondeau A (2009) Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model. *Biogeosciences*, **6**, 2001–2013.

Jung M, Reichstein M, Margolis HA *et al.* (2011) Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *Journal of Geophysical Research: Biogeosciences*, **116**, 1–16. doi:10.1029/2010JG001566.

Kätterer T, Reichstein M, Andrén O, Lomander A (1998) Temperature dependence of organic matter decomposition: a critical review using literature data analyzed with different models. *Biology and Fertility of Soils*, **27**, 258–262.

Kelly AE, Goulden ML (2008) Rapid shifts in plant distribution with recent climate change. *Proceedings of the National Academy of Sciences of the USA*, **105**, 11823–11826.

King AW, Gunderson CA, Post WM, Weston DJ, Wullschleger SD (2006) Plant respiration in a warmer world. *Science*, **312**, 536–537.

Knorr W, Prentice IC, House JI, Holland EA (2005) Long-term sensitivity of soil carbon turnover to warming. *Nature*, **433**, 298–301.

Knutti R (2008) Should we believe model predictions of future climate change? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **366**, 4647–4664.

Knutti R, Meehl GA, Allen MR, Stainforth DA (2006) Constraining climate sensitivity from the seasonal cycle in surface temperature. *Journal of Climate*, **19**, 4224–4233.

Koven CD, Ringeval B, Friedlingstein P *et al.* (2011) Permafrost carbon-climate feedbacks accelerate global warming. *Proceedings of the National Academy of Sciences of the USA*, **108**, 14769–14774.

Koven CD, Lawrence DM, Riley WJ (2015) Permafrost carbon–climate feed-back is sensitive to deep soil carbon decomposability but not deep soil nitro-gen dynamics. *Proceedings of the National Academy of Sciences of the USA*, **112**, 3752–3757.

Kurz WA, Dymond CC, Stinson G *et al.* (2008) Mountain pine beetle and forest carbon feedback to climate change. *Nature*, **452**, 987–990.

Lasslop G, Reichstein M, Papale D *et al.* (2010) Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation. *Global Change Biology*, **16**, 187–208.

Le Maire G, Delpierre N, Jung M *et al.* (2010) Detecting the critical periods that underpin interannual fluctuations in the carbon balance of European forests. *Journal of Geophysical Research: Biogeosciences*, **115**, 1–16. doi:10.1029/2009JG001244.

Le Quéré C, Moriarty R, Andrew RM *et al.* (2015) Global carbon budget 2015. *Earth System Science Data*, **7**, 349–396.

Leff B, Ramankutty N, Foley JA (2004) Geographic distribution of major crops across the world. *Global Biogeochemical Cycles*, **18**, GB1009 1001–1027.

Linderholm HW (2006) Growing season changes in the last century. *Agricultural and Forest Meteorology*, **137**, 1–14.

Lindroth A, Lagergren F, Grelle A, Klemedtsson L, Langvall O, Weslien P, Tuulik J (2009) Storms can cause Europe-wide reduction in forest carbon sink. *Global Change Biology*, **15**, 346–355.

Lucht W, Prentice IC, Myneni RB *et al.* (2002) Climatic control of the high-latitude vegetation greening trend and Pinatubo effect. *Science*, **296**, 1687–1689.

Luyssaert S, Janssens IA, Sulkava M *et al.* (2007) Photosynthesis drives anomalies in net carbon-exchange of pine forests at different latitudes. *Global Change Biology*, **13**, 2110–2127.

Mao J, Thornton PE, Shi X, Zhao M, Post WM (2012) Remote sensing evaluation of CLM4 GPP for the period 2000-09. *Journal of Climate*, **25**, 5327–5342.

Meinshausen M, Meinshausen N, Hare W *et al.* (2009) Greenhouse-gas emission targets for limiting global warming to 2 C. *Nature*, **458**, 1158–1162.

Menzel A, Fabian P (1999) Growing season extended in Europe. *Nature*, **397**, 659.

Mueller B, Seneviratne SI (2014) Systematic land climate and evapotranspiration biases in CMIP5 simulations. *Geophysical Research Letters*, **41**, 128–134.

Mueller B, Hirschi M, Jimenez C *et al.* (2013) Benchmark products for land evapotranspiration: landFlux-EVAL multi-data set synthesis. *Hydrology and Earth System Sciences*, **17**, 3707–3720.

Myneni RB, Keeling CD, Tucker CJ, Asrar G, Nemani RR (1997) Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature*, **386**, 698–702.

Nemani RR, Keeling CD, Hashimoto H *et al.* (2003) Climate-driven increases in global terrestrial net primary production from 1982 to 1999. *Science*, **300**, 1560–1563.

Norby RJ (2011) Ecological and evolutionary lessons from free air carbon enhancement (FACE) experiments. *Annual Review of Ecology, Evolution, and Systematics.*, **42**, 181–203.

Norby RJ, Warren JM, Iversen CM, Medlyn BE, McMurtrie RE (2010) CO_2 enhancement of forest productivity constrained by limited nitrogen availability. *Proceedings of the National Academy of Sciences of the USA*, **107**, 19368–19373.

Pan Y, Birdsey RA, Fang J *et al.* (2011) A large and persistent carbon sink in the world's forests. *Science*, **333**, 988–993.

Peñuelas J, Sardans J, Estiarte M *et al.* (2013) Evidence of current impact of climate change on life: a walk from genes to the biosphere. *Global Change Biology*, **19**, 2303–2338.

Peters GP, Davis SJ, Andrew R (2012) A synthesis of carbon in international trade. *Biogeosciences*, **9**, 3247–3276.

Phillips OL, Aragão LEOC, Lewis SL *et al.* (2009) Drought sensitivity of the amazon rainforest. *Science*, **323**, 1344–1347.

Piao S, Friedlingstein P, Ciais P, Viovy N, Demarty J (2007) Growing season extension and its impact on terrestrial carbon cycle in the Northern Hemisphere over the past 2 decades. *Global Biogeochemical Cycles*, **21**, 1–11. doi:10.1029/2006GB002888.

Piao S, Ciais P, Friedlingstein P *et al.* (2008) Net carbon dioxide losses of northern ecosystems in response to autumn warming. *Nature*, **451**, 49–52.

Piao S, Sitch S, Ciais P *et al.* (2013) Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO_2 trends. *Global Change Biology*, **19**, 2117–2132.

Prentice IC, Farquhar GD, Fasham MJR *et al.* (2001) The carbon cycle and atmospheric carbon dioxide. In: *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC)*, (eds Houghton JT, Ding Y, Griggs DJ, Noguer M, der van Linden PJ, Dai X, Maskell K, Johnson CA), pp. 183–237. Cambridge University Press, Cambridge a.o, Cambridge.

Qu X, Hall A (2014) On the persistent spread in snow-albedo feedback. *Climate Dynamics*, **42**, 69–81.

Raupach MR, Davis SJ, Peters GP *et al.* (2014) Sharing a quota on cumulative carbon emissions. *Nature Climate Change*, **4**, 873–879.

Reichstein M, Falge E, Baldocchi D *et al.* (2005) On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. *Global Change Biology*, **11**, 1424–1439.

Reichstein M, Papale D, Valentini R *et al.* (2007) Determinants of terrestrial ecosystem carbon balance inferred from European eddy covariance flux sites. *Geophysical Research Letters*, **34**, 1–5. doi:10.1029/2006GL027880.

Rustad LE, Campbell JL, Marion GM *et al.* (2001) A meta-analysis of the response of soil respiration, net nitrogen mineralization, and aboveground plant growth to experimental ecosystem warming. *Oecologia*, **126**, 543–562.

Sarmiento JL, Gloor M, Gruber N *et al.* (2010) Trends and regional distributions of land and ocean carbon sinks. *Biogeosciences*, **7**, 2351–2367.

Schimel D, Stephens BB, Fisher JB (2015) Effect of increasing CO_2 on the terrestrial carbon cycle. *Proceedings of the National Academy of Sciences of the USA*, **112**, 436–441.

Scriber JM (2011) Impacts of climate warming on hybrid zone movement: geographically diffuse and biologically porous "species borders". *Insect Science*, **18**, 121–159.

Sellers PJ, Dickinson RE, Randall DA *et al.* (1997) Modeling the exchanges of energy, water, and carbon between continents and the atmosphere. *Science*, **275**, 502–509.

Seneviratne SI, Nicholls N, Easterling D *et al.* (2012) Changes in climate extremes and their impacts on the natural physical environment. In: *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation* (eds Field CB, Barros V, Stocker TF, Qin D, Dokken DJ, Ebi KL, Mastrandrea MD, Mach KJ, Plattner G-K, Allen SK, Tignor M, Midgley PM). *A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC)*, pp. 109–230. Cambridge University Press, Cambridge, UK.

Sevanto S, Dickman LT (2015) Where does the carbon go?—Plant carbon al-location under climate change. *Tree Physiology*, **35**, 581–584.

Sitch S, Brovkin V, von Bloh W, van Vuuren D, Eickhout B, Ganopolski A (2005) Impacts of future land cover changes on atmospheric CO_2 and climate. *Global Biogeochemical Cycles*, **19**, 1–15.

Sitch S, Friedlingstein P, Gruber N *et al.* (2015) Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences*, **12**, 653–679.

Sun Y, Gu L, Dickinson RE, Norby RJ, Pallardy SG, Hoffman FM (2014) Impact of mesophyll diffusion on estimated global land CO_2 fertilization. *Proceedings of the National Academy of Sciences of the USA*, **111**, 15774–15779.

Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, **93**, 485–498.

Thomas RQ, Zaehle S, Templer PH, Goodale CL (2013) Global patterns of nitrogen limitation: confronting two global biogeochemical models with observations. *Global Change Biology*, **19**, 2986–2998.

Thornton P, Lamarque J-F, Rosenbloom N, Mahowald N (2007) Influence of carbon-nitrogen cycle coupling on land model response to CO2 fertilization and climate variability. *Global Biogeochemical Cycles*, **21**, GB4018.

Vicca S, Luyssaert S, Peñuelas J *et al.* (2012) Fertile forests produce biomass more efficiently. *Ecology Letters*, **15**, 520–526.

Vitousek PM, Porder S, Houlton BZ, Chadwick OA (2010) Terrestrial phosphorus limitation: mechanisms, implications, and nitrogen-phosphorus interactions. *Ecological Applications*, **20**, 5–15.

van Vuuren DP, Edmonds J, Kainuma M *et al.* (2011) The representative concentration pathways: an overview. *Climatic Change*, **109**, 5–31.

Wang X, Piao S, Ciais P *et al.* (2014) A two-fold increase of carbon cycle sensitivity to tropical temperature variations. *Nature*, **506**, 212–215.

Welp LR, Keeling RF, Meijer HAJ *et al.* (2011) Interannual variability in the oxygen isotopes of atmospheric CO_2 driven by El Niño. *Nature*, **477**, 579–582.

Wenzel S, Cox PM, Eyring V, Friedlingstein P (2014) Emergent constraints on climate-carbon cycle feedbacks in the CMIP5 Earth system models. *Journal of Geophysical Research: Biogeosciences*, **119**, 794–807.

Willmott CJ, Robeson SM (1995) Climatologically aided interpolation (CAI) of terrestrial air temperature. *International Journal of Climatology*, **15**, 221–229.

Xia J, Niu S, Ciais P, Janssens IA *et al.* (2015) Joint control of terrestrial gross primary productivity by plant phenology and physiology. *Proceedings of the National Academy of Sciences of the USA*, **112**, 2788–2793.

Zaehle S, Friedlingstein P, Friend AD (2010) Terrestrial nitrogen feedbacks may accelerate future climate change. *Geophysical Research Letters*, **37**, 1–5.

Zhang Q, Wang YP, Pitman AJ, Dai YJ (2011) Limitations of nitrogen and phosphorous on the terrestrial carbon uptake in the 20th century. *Geophysical Research Letters*, **38**, 1–5. doi:10.1029/2011GL049244.

Zhang Q, Pitman AJ, Wang YP, Dai YJ, Lawrence PJ (2013) The impact of nitrogen and phosphorous limitation on the estimated terrestrial carbon balance and warming of land use change over the last 156 yr. *Earth System Dynamics*, **4**, 333–345.

Zhao M, Heinsch FA, Nemani RR, Running SW (2005) Improvements of the MODIS terrestrial gross and net primary production global data set. *Remote Sensing of Environment*, **95**, 164–176.

Zhao M, Running SW, Nemani RR (2006) Sensitivity of Moderate Resolution Imaging Spectroradiometer (MODIS) terrestrial primary production to the accuracy of meteorological reanalyses. *Journal of Geophysical Research: Biogeosciences*, **111**, 1–13. doi:10.1029/2004JG000004.

Zhou L, Tucker CJ, Kaufmann RK, Slayback D, Shabanov NV, Myneni RB (2001) Variations in northern vegetation activity inferred from satellite data of vegetation index during 1981 to 1999. *Journal of Geophysical Research: Atmospheres*, **106**, 20069–20083.

Zickfeld K, Eby M, Damon Matthews H, Weaver AJ (2009) Setting cumulative emissions targets to reduce the risk of dangerous climate change. *Proceedings of the National Academy of Sciences of the USA*, **106**, 16129–16134.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Conceptual diagrams of the different constraint approaches: (a) GPPcon, (b) ETcon and (c) GPP&ETcon. The likely range for GPP and ET is defined based on 2std and 1std windows for the mean respectively.
Figure S2. Grid points where models simulate GPP within the likely range (2std) of the MTE GPP for the most productive month over the period (1989–2005).
Figure S3. Emergent relationships between GPP and ET (1st column), ΔGPP and GPP (2nd column), ΔGPP and ET (3rd column) and ΔGPP and ΔNBP (4th column) for every month individually.
Figure S4. Intermodel correlation between future changes in global annual GPP and NBP. The change is defined as the difference between the period (2084–2100) and (1989–2005).
Figure S5. Percentage of land grid points where simulated GPP and ET are lying within the 2std range and the 1std range of the MTE and the LandFlux-Eval products respectively.
Figure S6. Number of retained models for the most productive month for the GPPcon (top) and ETcon (bottom) constrained ensembles
Figure S7. Prior and constrained change in future GPP_{max} (gC m^{-2} $month^{-1}$) during the most productive month for the prior and the constrained ensembles. The change is defined as the difference between 2084–2100 and 1989–2005.
Figure S8. Mean bias in precipiation between CMIP5 (Multi-model mean) and different reference precipitation datasets. The reuslts are shown for the most productive month (1989–2005).
Figure S9. Percentage of grid points with estimates of both GPP and precipitation based on the CMIP5 models not lying within the range of the reference products. Results are shown for the most productive month.
Table S1. List of the CMIP5 models used in this study.
Table S2. Prior and constrained estimates of global annual mean GPP and ΔGPP based on a monthly-based contraint. Numbers in parenthesis correspond to the median.
Table S3. Prior and constrained estimates of global annual mean GPP and ΔGPP based on an annual-mean based contraint. Numbers in parenthesis correspond to the median.
Table S4. Precipitation datasets. All datasets were interpolated to a common regular 0.5° grid.
Appendix S1. Constraints on precipitation.