Reducing uncertainties in decadal variability of the global carbon budget with multiple datasets

Wei Li[a,1], Philippe Ciais[a], Yilong Wang[a], Shushi Peng[a], Grégoire Broquet[a], Ashley P. Ballantyne[b], Josep G. Canadell[c], Leila Cooper[b], Pierre Friedlingstein[d], Corinne Le Quéré[e], Ranga B. Myneni[f], Glen P. Peters[g], Shilong Piao[h], and Julia Pongratz[i]

[a]Laboratoire des Sciences du Climat et de l'Environnement/Institut Pierre Simon Laplace, Commissariat à l'Énergie Atomique et aux Énergies Alternatives–CNRS–Université de Versailles Saint-Quentin, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France; [b]Department of Ecosystem and Conservation Science, University of Montana, Missoula, MT 59812; [c]Global Carbon Project, Commonwealth Scientific and Industrial Research Organization Oceans and Atmosphere, Canberra, ACT 2601, Australia; [d]College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter EX4 4QF, United Kingdom; [e]Tyndall Centre for Climate Change Research, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, United Kingdom; [f]Department of Earth and Environment, Boston University, Boston, MA 02215; [g]Center for International Climate and Environmental Research–Oslo, 0349 Oslo, Norway; [h]Department of Ecology, College of Urban and Environmental Sciences, Peking University, Beijing 100871, China; and [i]Max Planck Institute for Meteorology, 20146 Hamburg, Germany

Edited by Douglas C. Morton, NASA Goddard Space Flight Center, Greenbelt, MD, and accepted by Editorial Board Member Ruth S. DeFries September 15, 2016 (received for review March 9, 2016)

Conventional calculations of the global carbon budget infer the land sink as a residual between emissions, atmospheric accumulation, and the ocean sink. Thus, the land sink accumulates the errors from the other flux terms and bears the largest uncertainty. Here, we present a Bayesian fusion approach that combines multiple observations in different carbon reservoirs to optimize the land (B) and ocean (O) carbon sinks, land use change emissions (L), and indirectly fossil fuel emissions (F) from 1980 to 2014. Compared with the conventional approach, Bayesian optimization decreases the uncertainties in B by 41% and in O by 46%. The L uncertainty decreases by 47%, whereas F uncertainty is marginally improved through the knowledge of natural fluxes. Both ocean and net land uptake (B + L) rates have positive trends of 29 ± 8 and 37 ± 17 Tg C·y^{-2} since 1980, respectively. Our Bayesian fusion of multiple observations reduces uncertainties, thereby allowing us to isolate important variability in global carbon cycle processes.

global carbon budget | carbon cycle | decadal variations | Bayesian fusion

The land and ocean carbon sinks provide a vital climate mitigation "service" by absorbing on average about 55% of anthropogenic CO_2 emissions from fossil fuel combustion and land use change. Research has focused on understanding the relationships between year-to-year variability in carbon sinks and climate (1, 2), as well as the long-term trend over the full instrumental period of CO_2 monitoring at the Mauna Loa station (3). Quasidecadal variations of emissions and sinks have received comparatively less attention. However, significant climate variation occurs at this specific timescale (4). Since 1980, the variable occurrence of different El Niño–Southern Oscillation events, two large volcanic eruptions (El Chichón and Pinatubo) and the recent slowdown of land surface warming have modulated the strength of carbon sinks. There are also decadal-scale changes in the rate at which human activities perturb the natural carbon cycle, in particular the recent acceleration of fossil fuel and cement emissions in the 2000s (5) and the slowdown in global land use change emissions (LUC) in the mid-2000s, which appears to be partly driven by reduced deforestation in Brazil (6).

Here, we provide a data-driven assessment of global CO_2 sources and sinks at 5-y intervals for the period of 1980–2014. We use a Bayesian fusion approach whereby different data streams of ocean and land uptake, LUC emissions, are optimally combined, and their uncertainty reduced from prior knowledge. This approach estimates the land sink constrained by data, which is a major improvement over the conventional method for calculating the global carbon budget by Ciais et al. (7) and Le Quéré et al. (8), hereafter LQ15, where the unknown land sink was determined as a residual from the other components (emissions, atmospheric increase, ocean uptake). Most of the data streams used in this analysis start in 1980, and about one-half of them give decadal mean values of natural sinks and thus do not allow us to tackle the reconstruction of interannual variability. Our choice of applying a Bayesian fusion approach to optimize 5-y average component fluxes of the global carbon budget is therefore a compromise that maximizes the use of available observations of decadal average fluxes.

The principle of the Bayesian fusion approach is to combine an a priori imperfect knowledge of fluxes with observations and their uncertainties to infer optimized estimates of fluxes. Here, we define a priori values of terms in the global carbon budget that are not from observations. Specifically, we set prior fossil fuel and cement emissions (F) from inventories and the simulated land, ocean, and land use change carbon fluxes from process-based models (Table S1). The fluxes in this study are only anthropogenic fluxes, assuming a mean CO_2 growth rate of zero during preindustrial times. Observational datasets independent from those prior values are applied to constrain land use change emissions (L), the ocean uptake of anthropogenic CO_2 (O), the land-biosphere sink (B) in ecosystems not affected by land use change, and the net land flux (B + L) (Table S2). We select the constraining data from peer-reviewed publications and evaluate

Significance

The conventional approach of calculating the global carbon budget makes the land sink the most uncertain of all budget terms. This is because, rather than being constrained by observations, it is inferred as a residual in the budget equation. Here, we overcome this limitation by performing a Bayesian fusion of different available observation-based estimates of decadal carbon fluxes. This approach reduces the uncertainty in the land sink by 41% and in the ocean sink by 46%. These results are significant because they give unprecedented confidence in the role of the increasing land sink in regulating atmospheric CO_2, and shed light on the past decadal trend.

Author contributions: P.C. designed research; W.L. performed research; Y.W., G.B., A.P.B., L.C., C.L.Q., R.B.M., and J.P. contributed to design the method and choose the datasets; W.L., Y.W., S. Peng, and C.L.Q. analyzed data; and W.L., P.C., G.B., J.G.C., L.C., P.F., C.L.Q., R.B.M., G.P.P., S. Piao, and J.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. D.C.M. is a Guest Editor invited by the Editorial Board.

[1]To whom correspondence should be addressed. Email: wei.li@lsce.ipsl.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1603956113/-/DCSupplemental.

their reported uncertainties and possible error correlations with each other (see details in Table S2).

Results

In the optimization of the global anthropogenic carbon budget (Fig. 1), the prior value of F and its uncertainty (Table S1) were defined from the mean value and the range of different fossil fuel and cement emission inventories, namely from the Carbon Dioxide Information Analysis Center (CDIAC) (9), International Energy Agency (IEA) (10), Emissions Database for Global Atmospheric Research (EDGAR) (11), and BP Statistical Review of World Energy (12). These emission inventories are not treated as direct observations of emissions, and there is currently no independent observation to verify F. The prior values of O are from seven ocean biogeochemistry models (8), and the prior values of B are from the nine TRENDY land carbon models (13). These prior values from state-of-the-art models are without direct observational constraints. Some of these models could possibly be tuned using similar observation-based data, but which observational data were used for model tuning was not explicitly reported. The prior estimates of L are derived from the difference of simulated land carbon fluxes with and without LUC in the TRENDY carbon models (13). All fluxes are defined as positive if CO_2 is emitted to the atmosphere by the land or the ocean reservoir. Uncertainties in the prior estimates of 5-yearly O, B, and L, are set to the maximum between those reported by LQ15 and the SDs across models. All uncertainties here refer to 1-σ Gaussian errors. In this context, the prior uncertainties are 0.5 for O, 0.9 for B, and 0.8 Pg C·y^{-1} for L, thus not smaller than the values of 0.5, 0.8, and 0.5 Pg C·y^{-1} from LQ15. It is important that the prior uncertainties are not too small, so that adding observations can adjust and constrain the sought fluxes.

Several independent data streams, each with their specific uncertainty and temporal averaging period (Table S2), are combined in the Bayesian optimization with the above prior knowledge. These data streams are as follows: (*i*) the atmospheric CO_2 growth rate (CGR) from the National Oceanic and Atmospheric Administration (NOAA)/Earth System Research Laboratory (ESRL) atmospheric network (14), which constrains the sum of all fluxes and is determined very accurately from more than 60 monitoring stations; (*ii*) the atmospheric 5-y mean (negative) growth rate of O_2/N_2 in the atmosphere from the Scripps O_2 Program (15), which relates to the combined effect of B + L and F changes, while being insensitive to changes in O (note that O_2/N_2 has a negative trend in the atmosphere); (*iii*) a set of yearly mean estimates of O from observational products based on in situ partial pressure of CO_2 (pCO_2) surveys corrected for natural ocean CO_2 outgassing from carbon delivered by rivers (16) and using a neural network approach (17) and a diagnostic mixed-layer approach (18), and a set of decadal-mean estimates of O from inventories of carbon change in the ocean deduced indirectly from chlorofluorocarbons (CFCs) (19) combined with ^{14}C (20), and observed atmospheric mean CO_2 level and oceanic CO_2 and dissolved inorganic carbon observations (21); (*iv*) 10-y mean estimates of B from a global synthesis of changes in forest carbon stocks (22); (*v*) decadal mean B + L from inventory-based land carbon storage change from the RECCAP publications on regional budgets (Table S3); (*vi*) 5-y mean LUC emissions from two independent bookkeeping approaches constrained by observed carbon stocks (23, 24). The uncertainties in each data stream are either derived directly from the original publications (when reported) or estimated from expert judgments (see details in Table S2). The optimization is performed for seven consecutive 5-y windows between 1980 and 2014.

In the Bayesian optimization, observations that describe mean fluxes during intervals longer than 5 y are still useful to infer 5-yearly fluxes. For example, the mean ocean sink observation for the 1990s (19) constrains the two mean 5-yearly O during 1990–1999, whereas other independent observations (O_2/N_2 and CGR) help to further allocate O values between the periods 1990–1994 and 1995–1999. Despite no direct observation of F, this flux is found to be slightly improved in the Bayesian fusion, through knowledge of the other terms, and because the sum of all fluxes is very well constrained from CGR observations. We are aware that some observation-based land sink estimates have systematic errors in the way they are included in the optimization. In particular, the estimate of B from ref. 22 is only for forests and ignores other biomes. However, the Regional Carbon Cycle Assessment and Processes (RECCAP) studies (25–27) and other estimates (22, 28) of the carbon stock change in nonforest biomes suggest that the forest sink alone accounts for most of the global land sink B.

The improved global budget of anthropogenic CO_2 is shown in Fig. 2, and all data are given in Table S1. After optimization, the a posteriori uncertainty in each flux is reduced. Compared with the conventional method applied by LQ15 and IPCC-AR5 (7), uncertainties in B and O are reduced by 41% and 46% in this study. In the Bayesian data fusion, the land sink is no longer solely inferred as a residual that accumulates uncertainties from all other terms, and it exhibits a large reduction in uncertainty. The uncertainty in L decreases by 47%, but the uncertainty in F is marginally improved (by 1%) through the indirect constraints of other terms. In the absence of direct constraint on F, this small reduction in the F uncertainties compared with LQ15 and IPCC-AR5 (7) is also because we use multiple emission inventories [whereas LQ15 and IPCC-AR5 (7) only used CDIAC (9)] and start at relatively higher prior uncertainties in F (Table S1) than in LQ15. Despite their improved (smaller) uncertainties, the 5-y mean fluxes shown in Table S1 do not differ statistically in their mean values from LQ15. This indicates that each flux of the Bayesian carbon budget is fully consistent with LQ15 even though we used an array of data with different measurement methods and with uncertainties estimated in different ways. Specifically, we obtain emissions from fossil fuel burning and cement production that are smaller than LQ15 by 0.18 ± 0.19 Pg C·y^{-1} during 1980–2014 (Fig. 2). A downward revision of global F is consistent with the correction of the emissions for China based on evidence of the lower carbon content for coal burned in that country (29). Compared with LQ15, the optimized ocean sink during 2000–2004 is larger by 0.22 Pg C·y^{-1} but lower by 0.18 ± 0.10 Pg C·y^{-1} during

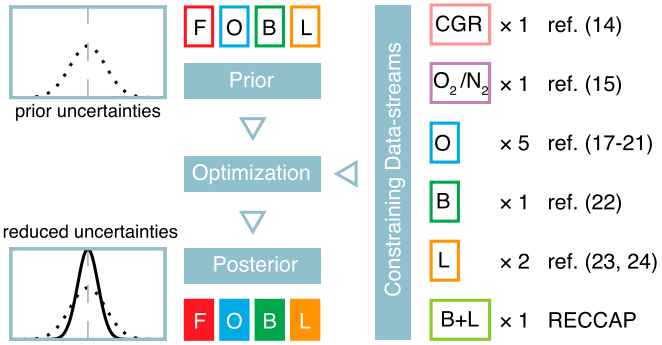

Fig. 1. The framework of our optimization. The number of constraining data streams and the specific data sources are marked on the *Right*. The fluxes that are optimized are 5-y averages of F, O, B, and L, representing fossil fuel and cement emissions, ocean sink, land sink, and land use change emissions, respectively. The observations used to constrain these fluxes are the 5-y averaged growth rates of CO_2 and O_2/N_2 in the atmosphere, observations of O and B from carbon measurements made in these two reservoirs, and inventory-based estimates of L and the net land sink (B + L). In this framework, the CO_2 growth rate constrains the sum of all of the fluxes. The O_2/N_2 growth rate allows us to separate O and B + L and bring some constraint on F as well.

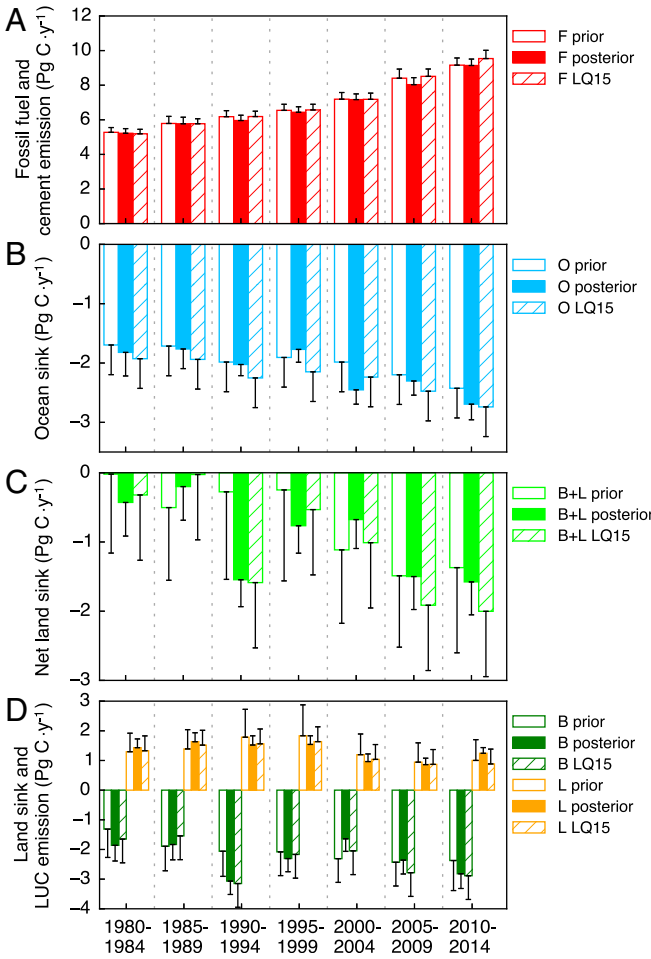

Fig. 2. (*A*) The fossil fuel and cement emissions (F), (*B*) ocean sink (O) and (*C*) net land flux (B + L), and (*D*) land sink (B) and land use change emissions (L) from prior knowledge, posterior results, and LQ15. All of the fluxes are 5-y means in each period. The error bars represent the 1-σ uncertainties.

all of the other periods. In the past decade (2005–2014), both ocean sink and land sink from our optimization are smaller than LQ15. The optimized fluxes of L are similar to or lower than those from LQ15. The trend of F for the seven 5-y periods is positive ($P = 0.003$), with a probability of a positive trend for ocean and net land (B + L) uptake rates of 93%; the trend of B or L individually is not statistically significant ($P = 0.23$ for both). The increasing rate of ocean and net land uptake are 29 ± 8 and 37 ± 17 Tg C·y^{-2} since 1980, respectively. Similar statistically positive trends were also found in the 5-y mean ocean and net land uptake rates between 1980 and 2014 calculated from the yearly budget updated by LQ15 (Fig. S1*A*). Given the robustness of O inferred by our optimization (Fig. S2), and in view of the many observations constraining this flux, there is a high confidence that the ocean sink has been increasing over time since 1980. The optimized land sink is less variable between different 5-y periods than in LQ15 (Fig. S1*A*). However, the ocean sink is more variable, with a SD of 0.34 Pg C·y^{-1} compared with 0.27 Pg C·y^{-1} by LQ15 (SDs across the seven periods analyzed).

From 1980 to 2014, the average fractions of F + L emission reabsorbed by the land and ocean carbon reservoirs are 29 ± 5% (mean ± 1 σ of inter–5-y variability) and 26 ± 2%, respectively. The ratios of both O and B to F + L emission do not exhibit any significant trends ($P > 0.05$; Fig. S1*B*). Even with their reduced uncertainty in this study compared with LQ15, the variability of O and B between 5-y intervals prevents us from assessing the very small trends in their ratios to emissions. Similarly, we found no significant trend in the ratio of O or B to fossil fuel emission (F). The larger variability of the B-to-(F + L) ratios compared with the O-to-(F + L) ones (Fig. S1*B*) suggests that the efficiency of the land sink at absorbing emissions is more variable than that of the ocean sink. For instance, during the period that followed the cooling from the Pinatubo eruption in 1990–1994 (30–32), the B-to-(F + L) ratio increased by 46% above its long-term mean. This ratio was also higher than normal during 2005–2009, possibly due to the absence of El Niño and to the occurrence of a cooler and wetter La Niña event in 2008–2009 manifested by lower than normal CGR (3).

Discussion

In our Bayesian fusion, the CGR constrains the sum of the different fluxes; thus, an overestimation of a component flux due to the use of a single published estimate of that flux would result in an underestimation of another flux, leading to negative correlations between uncertainties in the different components of the posterior fluxes. These negative correlations are clear between uncertainties in B and F (ranging from −0.80 to −0.35), in B and O (from −0.60 to −0.30), and in B and L (from −0.53 to −0.24) (Fig. S3). This indicates that, although the uncertainties for each flux can be significantly decreased, the remaining (posterior) uncertainties are hard to be decoupled. In comparison, we also calculate the correlations between uncertainties in B and other fluxes by classical error propagation rule from mass balance equation B = CGR − F − O − L. The conventional error budget calculation gives a typical uncertainty of 0.76 Pg C·y^{-1} for B. Negative error correlations also exist in that approach between B and F (−0.39), between B and O (−0.79), and between B and L (−0.39). Thus, in our study, not only the posterior uncertainties of B (0.42–0.53 Pg C·y^{-1}) are smaller than those calculated by the conventional approach, but also the negative error correlation between B and O is smaller, indicating a better separation between the carbon sinks of B and O. In addition, although the negative correlations between B and F, and between B and L are slightly larger than those calculated by the conventional approach, it does not mean our approach has poorer ability to separate these components because we have succeeded in significantly decreasing the absolute uncertainties for all of the fluxes, thus allowing us to draw more definitive conclusions.

The Bayesian fusion of different observational components of the anthropogenic CO_2 budget proposed in this study provides the most robust estimate to date of the strength and evolution of the land sink on 5-y intervals and brings a more robust picture of the current perturbation of the carbon cycle. Future work could apply the same fusion approach to regional carbon budget estimates and to gross CO_2 fluxes of photosynthesis, respiration, and fire emissions.

Materials and Methods

Bayesian Estimation System. Each estimate of the 5-y mean carbon fluxes, called hereafter the "control variables" $\mathbf{x}$, is based on the update from a prior estimate of these variables $\mathbf{x}^b$, using some observation-based estimates $\mathbf{y}^o$ of the fluxes that are connected to the control variables through the relationships H: $\mathbf{x}\rightarrow\mathbf{y} = H[\mathbf{x}]$. We follow a Bayesian statistical approach for this estimation. Assuming that the distributions of uncertainties in $\mathbf{x}^b$ and $\mathbf{y}^o$ are unbiased and Gaussian, being characterized by the prior and observation uncertainty covariance matrices $\mathbf{B}$ and $\mathbf{R}$, respectively, and that H is linear (denoted as a matrix $\mathbf{H}$), the statistical estimate of $\mathbf{x}$, given $\mathbf{x}^b$ and $\mathbf{y}^o$, is unbiased and Gaussian, and the corresponding optimal estimate $\mathbf{x}^a$ and uncertainty covariance matrix $\mathbf{A}$ are given (33) as the following:

$$\mathbf{A} = \left(\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\right)^{-1}, \quad [1]$$

$$\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{b}} + \mathbf{A}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\left(\mathbf{y}^{\mathrm{o}} - \mathbf{H}\mathbf{x}^{\mathrm{b}}\right), \quad [2]$$

where the superscripts T and "−1" denote the transpose and inverse of a matrix, respectively.

In the optimization, we update the estimate of the mean fluxes of fossil fuel and cement emissions (F), ocean sink (O), land sink (B), and land use change emissions (L) for each 5-y interval from 1980 to 2014 (Fig. 1 and Table S1). The observation vector contains estimates of 5-y mean global value for the following: the atmospheric growth rates of CO_2 (CGR, in petagrams of carbon per year), atmospheric growth rates of O_2/N_2 (CGR-O_2, per meg unit), observation-based estimates of ocean sinks, land sinks, land use change emissions, and net land sink (B + L) (the data sources for these components of $\mathbf{y^o}$ are summarized in Fig. 1 and provided in Table S2). **H** is defined for each 5-y interval by the following:

$$\begin{array}{ccccc} \mathbf{x} & \rightarrow & \mathbf{y^o} & = & \mathbf{Hx} \\ & & \mathrm{CGR} & & \mathrm{F+O+B+L} \\ \mathbf{H}: \begin{bmatrix} \mathrm{F} \\ \mathrm{O} \\ \mathrm{B} \\ \mathrm{L} \end{bmatrix} & \rightarrow & \begin{bmatrix} \mathrm{CGR-O_2} \\ \mathrm{O} \\ \mathrm{B} \\ \mathrm{L} \\ \mathrm{B+L} \end{bmatrix} & = & \begin{bmatrix} \alpha_F \mathrm{F}+\alpha_B(\mathrm{B+L})+\mathrm{Z_{O2}} \\ \mathrm{O} \\ \mathrm{B} \\ \mathrm{L} \\ \mathrm{B+L} \end{bmatrix}, \end{array} \quad [3]$$

where α_F, α_B, and Z_{O2} are constant coefficients from ref. 15.

The prior estimates for the different control variables are built with independent datasets so that there are no correlations between the prior uncertainties in the different control variables. When setting up the observation covariance matrix, the temporal correlations between the uncertainties in different 5-y intervals for CGR and O are estimated using the method in ref. 34. The correlations between the uncertainties in the two data-driven estimates of O (17, 18) are estimated from the series of annual fluxes of the two products by assuming that the correlation in annual fluxes within each 5-y period is an approximation of the 5-y mean flux error correlation.

Data Used to Derive the Prior Statistics of the Control Variables. To define the prior estimate of F, individual country data from four emission inventories [CDIAC (9), IEA (10), EDGAR (11), and BP (12)] are grouped into geographic regions as specified by the United Nations Statistics Division (unstats.un.org/unsd/methods/m49/m49regin.htm). Cement emissions from EDGAR are added into the IEA and BP datasets that do not include cement emissions. Uncertainties for each country (35) are used to create regional uncertainty distributions using a bootstrapping method, with the uncertainties of the highest emitters in each region contributing the most to the uncertainty distributions. This effect is achieved by weighting the sampling probability (P_s) by the relative contribution of each country's emissions (E_C) to the total emissions within the region (E_R) as follows:

$$P_s = \frac{E_C}{E_R}. \quad [4]$$

To constrain the temporal component of the emission errors, 10 random samples are drawn from the corresponding regional uncertainty distribution for each country, producing 10 random uncertainties for each country. These country-level uncertainties are used to constrain a random error time series covering 1980–2014, which is then run through an algorithm incorporating autocorrelated random noise, such that

$$\varepsilon_{F(t)} = 0.95 \times \varepsilon_{F(t-1)} + \varepsilon_{(t)}, \quad [5]$$

where emission error factors for any given year $\varepsilon_{F(t)}$ are correlated with the emission errors from the previous year $\varepsilon_{F(t-1)}$ by an autoregressive coefficient of 0.95 with $\varepsilon_{(t)}$ as random error. The autocorrelated time series are then multiplied and added to the fossil fuel emissions for each country, and subsequently 500 samples of global fossil fuel emissions are taken for each 5-y bin. The means and SDs of each bin for each inventory are calculated from these 500 samples. Additionally, the correlation in global uncertainty is calculated between 5-y bins and inventories to produce an error-covariance matrix. The maximum between the uncertainties calculated above and the SDs of the 5-y means across four emission inventories were adopted as the uncertainties in the prior estimate of F.

Prior O values are set from the ocean biogeochemistry model values used in LQ15, which represent state-of-the-art ocean models and are generally consistent with estimates from data-based products (8). Note that LQ15 adjusted their simulated O so as to match ocean observations during the decade of the 1990s and then used these bias-corrected ocean models outside this period (see adjusted values in Table S1). Here, for setting the prior O estimate and uncertainty, we consider simply the spread and the mean of ocean models without any adjustment, because the adjustment performed by LQ15 is already a kind of model data-fusion approach based on the ocean observations used in our study.

Prior values of L and B are set from simulations in the TRENDY (version 2) model intercomparison project (13). The simulations in TRENDY (version 2) are up to 2012, and thus the priors for the period of 2010–2012 were used for the period of 2010–2014. All of the prior flux values are summarized in Table S1.

Uncertainty Correlations Between Optimized Variables. The correlations of the uncertainties between optimized variables are shown in Fig. S3. In comparison, we also calculate the correlations between uncertainties in the fluxes when deriving the estimate of B using the classical mass balance equation B = CGR − F − O − L. First, the uncertainties in an estimate of the flux i is calculated using inverse variance method, based on the variances associated with all of the estimates of this flux (j = 1, 2, …, m):

$$\sigma_i = \sqrt{\frac{1}{\sum_{j=1}^{m} \frac{1}{\sigma_{i,j}^2}}}. \quad [6]$$

The resulting uncertainties for CGR, F, O, and L are as follows: 0.2, 0.3, 0.6, and 0.3 Pg C·y^{-1}, respectively. Based on a simple error propagation: $\varepsilon_B = \varepsilon_{CGR} - \varepsilon_F - \varepsilon_O - \varepsilon_L$ and the independence of the estimates of CGR, F, O, and L, the variance of the uncertainty in B is given by the following:

$$\sigma_B^2 = E\left(\varepsilon_B^T \varepsilon_B\right) = \sigma_{CGR}^2 + \sigma_F^2 + \sigma_O^2 + \sigma_L^2, \quad [7]$$

where E(.) is the expectation of a variable. The correlation between the uncertainties in B and in another variable, namely, O in the equation below, is given by the following:

$$r(\varepsilon_B, \varepsilon_O) = \frac{E\left(\varepsilon_B^T \varepsilon_O\right)}{\sigma_B \sigma_O} = \frac{\sigma_O}{\sigma_B}. \quad [8]$$

Sensitivity Tests. Four sets of sensitivity tests were conducted (Fig. S2): (*i*) using F from different inventory datasets (IEA, EDGAR, CDIAC, and BP, respectively) as priors; (*ii*) using different ocean constraining datasets (only data from ref. 17, only data from ref. 18, data from refs. 17 and 18 but without other ocean data, and all ocean data except ref. 19, respectively); (*iii*) using different tiers of constraining data (tier 1 only, tier 1 plus tier 2, and tier1 plus tier 2 plus tier 3, respectively); and (*iv*) using enlarged prior uncertainties (10^6 Pg C·y^{-1} except F), prior B and O from CMIP5 models (36) instead of TRENDY (version 2) models, and a subset of constraining data including CGR, O_2/N_2, L from ref. 24 and O from pCO_2 and inventories (17, 18, 20). The constraining data tiers (Table S2) were defined as follows: tier 1, direct carbon observations (e.g., CGR); tier 2, indirect carbon observations unambiguously related to carbon quantities (e.g., O_2/N_2); tier 3, direct carbon observations with an empirical (data-driven) model used to obtain global estimates, for example, the use of geostatistics to up-scale local data into global values; tier 4, indirect carbon observations not simply related to global carbon flux quantities.

The ocean sink is rather consistent in all sensitivity tests (Fig. S2), implying it is very robustly constrained in the optimization because of the sufficient number and consistency of constraining observational data (see the O constraining data in Table S2). Small changes in optimized F were found (Fig. S2*A*) in the sensitivity tests using prior F from different datasets, and the land sink is more dependent on the prior F choice. In the sensitivity tests using different ocean constraining data (Fig. S2*B*), all fluxes are generally consistent, and very slight changes for O appear in the tests using ocean data from ref. 17 only and from ref. 18 only during 2000–2004 and 2005–2009. The decadal estimates (19–21) only have small impact on their corresponding periods. The important point for the sensitivity tests using different tiers of constraining data are that there is no inconsistency (i.e., shift in posterior estimate) between the assimilations of different "observation-based" tiers, although L may change due to the lack of individual constraints in tier 1 and tier 1 plus tier 2 (Fig. S2*C*). With enlarged prior uncertainties (10^6 Pg C·y^{-1} except F), the mean flux values do not shift in general but the posterior uncertainties increase (Fig. S2*D*). Results are also very consistent by replacing prior B and O from TRENDY, version 2, models with from CMIP5 models (Fig. S2*D*). The sensitivity test using a subset of constraining data underestimate B and L compared with the original optimized results (Fig. S2*D*) because the L estimates in ref. 23 (used in the original optimization but not in this sensitivity test) is higher than that in ref. 24 (Table S2).

Trend Test. A Mann–Kendall statistical test was applied as a trend test.

ACKNOWLEDGMENTS. This paper builds on an analysis started by the late Michael R. Raupach and is a contribution to the work of the Global Carbon Project to understand and better constrain the human perturbation of the carbon budget. We thank Alessandro Anav for providing the data from CMIP5 and Peter Landschützer and Christian Rödenbeck for sharing their ocean flux data. We are grateful to the scientists from NOAA/ESRL and the Scripps O_2 Program for making the invaluable data from the long-term measurements of CGR and O_2/N_2 available to the community. W.L. is supported by the European Commission-funded project LUC4C (603542). P.C. and S. Peng acknowledge support from the European Research Council through Synergy Grant ERC-2013-SyG-610028 "IMBALANCE-P." J.G.C. is supported by the Australian Climate Change Science Program, G.P.P. is supported by the Norwegian Research Council (236296), and J.P. is supported by the German Research Foundation's Emmy Noether Program.

1. Wang X, et al. (2014) A two-fold increase of carbon cycle sensitivity to tropical temperature variations. *Nature* 506(7487):212–215.
2. Poulter B, et al. (2014) Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle. *Nature* 509(7502):600–603.
3. Le Quéré C, et al. (2009) Trends in the sources and sinks of carbon dioxide. *Nat Geosci* 2(12):831–836.
4. Ghil M, Vautard R (1991) Interdecadal oscillations and the warming trend in global temperature time series. *Nature* 350(6316):324–327.
5. Raupach MR, et al. (2007) Global and regional drivers of accelerating CO_2 emissions. *Proc Natl Acad Sci USA* 104(24):10288–10293.
6. Song X-P, Huang C, Saatchi SS, Hansen MC, Townshend JR (2015) Annual carbon emissions from deforestation in the Amazon Basin between 2000 and 2010. *PLoS One* 10(5):e0126754.
7. Ciais P, et al. (2013) Carbon and other biogeochemical cycles. *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, et al. (Cambridge Univ Press, Cambridge, UK), pp 465–570.
8. Le Quéré C, et al. (2015) Global carbon budget 2014. *Earth Syst Sci Data* 7:47–85.
9. Boden TA, Marland G, Andres RJ (2013) *Global, Regional, and National Fossil-Fuel CO_2 Emissions* (Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory Oak Ridge, TN; US Department of Energy, Oak Ridge, TN).
10. International Energy Agency (2013) *CO_2 Emissions from Fuel Combustion: Highlights.* IEA Statistics (International Energy Agency, Paris).
11. Olivier JGJ, Janssens-Maenhout G, Peters JAHW (2014) *Trends in Global CO_2 Emissions: 2014 Report* (PBL Netherlands Environmental Assessment Agency, The Hague).
12. BP (2015) *BP Statistical Review of World Energy 2015* (BP, London).
13. Sitch S, et al. (2015) Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences* 12(3):653–679.
14. NOAA/ESRL, NOAA/ESRL calculation of global means. Available at www.esrl.noaa.gov/gmd/ccgg/about/global_means.html. Accessed August 9, 2015.
15. Keeling RF, Manning C (2014) *Studies of Recent Changes in Atmospheric O_2 Content* (Elsevier, Amsterdam), 2nd Ed.
16. Jacobson AR, Fletcher SEM, Gruber N, Sarmiento JL, Gloor M (2007) A joint atmosphere-ocean inversion for surface fluxes of carbon dioxide: 1. Methods and global-scale fluxes. *Global Biogeochem Cycles* 21(1):GB1019.
17. Landschützer P, Gruber N, Bakker DCE, Schuster U (2014) Recent variability of the global ocean carbon sink. *Global Biogeochem Cycles* 28(9):927–949.
18. Rödenbeck C, et al. (2014) Interannual sea–air CO_2 flux variability from an observation-driven ocean mixed-layer scheme. *Biogeosciences* 11(17):4599–4613.
19. McNeil BI, Matear RJ, Key RM, Bullister JL, Sarmiento JL (2003) Anthropogenic CO_2 uptake by the ocean based on the global chlorofluorocarbon data set. *Science* 299(5604):235–239.
20. Khatiwala S, Primeau F, Hall T (2009) Reconstruction of the history of anthropogenic CO_2 concentrations in the ocean. *Nature* 462(7271):346–349.
21. Steinkamp K, Gruber N (2013) A joint atmosphere-ocean inversion for the estimation of seasonal carbon sources and sinks. *Global Biogeochem Cycles* 27(3):732–745.
22. Pan Y, et al. (2011) A large and persistent carbon sink in the world's forests. *Science* 333(6045):988–993.
23. Hansis E, Davis SJ, Pongratz J (2015) Relevance of methodological choices for accounting of land use change carbon fluxes. *Global Biogeochem Cycles* 29(8):1230–1246.
24. Houghton RA, et al. (2012) Carbon emissions from land use and land-cover change. *Biogeosciences* 9(12):5125–5142.
25. Piao SL, et al. (2012) The carbon budget of terrestrial ecosystems in East Asia over the last two decades. *Biogeosciences* 9(9):3571–3586.
26. Luyssaert S, et al. (2012) The European land and inland water CO_2, CO, CH_4 and N_2O balance between 2001 and 2005. *Biogeosciences* 9(8):3357–3380.
27. King AW, et al. (2015) North America's net terrestrial CO_2 exchange with the atmosphere 1990–2009. *Biogeosciences* 12(2):399–414.
28. Liu YY, et al. (2015) Recent reversal in loss of global terrestrial biomass. *Nat Clim Chang* 5(5):470–474.
29. Liu Z, et al. (2015) Reduced carbon emission estimates from fossil fuel combustion and cement production in China. *Nature* 524(7565):335–338.
30. Lucht W, et al. (2002) Climatic control of the high-latitude vegetation greening trend and Pinatubo effect. *Science* 296(5573):1687–1689.
31. Mercado LM, et al. (2009) Impact of changes in diffuse radiation on the global land carbon sink. *Nature* 458(7241):1014–1017.
32. Frölicher TL, Joos F, Raible CC, Sarmiento JL (2013) Atmospheric CO_2 response to volcanic eruptions: The role of ENSO, season, and variability. *Global Biogeochem Cycles* 27(1):239–251.
33. Tarantola A (2005) *Inverse Problem Theory and Methods for Model Parameter Estimation* (SIAM, Philadelphia).
34. Ballantyne AP, et al. (2015) Audit of the global carbon budget: Estimate errors and their impact on uptake uncertainty. *Biogeosciences* 12(8):2565–2584.
35. Andres RJ, Boden TA, Higdon D (2014) A new evaluation of the uncertainty associated with CDIAC estimates of fossil fuel carbon dioxide emission. *Tellus B Chem Phys Meterol* 66:23616.
36. Anav A, et al. (2013) Evaluating the land and ocean components of the global carbon cycle in the CMIP5 earth system models. *J Clim* 26(18):6801–6843.
37. Rödenbeck C, et al. (2015) Data-based estimates of the ocean carbon sink variability—first results of the surface ocean pCO_2 mapping intercomparison (SOCOM). *Biogeosciences* 12(23):7251–7278.
38. Rödenbeck C, et al. (2013) Global surface-ocean $p^{CO}{}_2$ and sea–air CO_2 flux variability from an observation-driven ocean mixed-layer scheme. *Ocean Sci* 9(2):193–216.
39. Hurtt GC, et al. (2011) Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Clim Change* 109(1-2):117–161.
40. Peters GP, Davis SJ, Andrew R (2012) A synthesis of carbon in international trade. *Biogeosciences* 9(8):3247–3276.
41. Hartmann J, Jansen N, Dürr HH, Kempe S, Köhler P (2009) Global CO_2-consumption by chemical weathering: What is the contribution of highly active weathering regions? *Global Planet Change* 69(4):185–194.
42. Dolman AJ, et al. (2012) An estimate of the terrestrial carbon budget of Russia using inventory-based, eddy covariance and inversion methods. *Biogeosciences* 9(12):5323–5340.
43. Lauerwald R, Laruelle GG, Hartmann J, Ciais P, Regnier PAG (2015) Spatial patterns in CO_2 evasion from the global river network. *Global Biogeochem Cycles* 29(5):2014GB004941.
44. Gloor M, et al. (2012) The carbon balance of South America: A review of the status, decadal trends and main determinants. *Biogeosciences* 9(12):5407–5430.
45. Valentini R, et al. (2014) A full greenhouse gases budget of Africa: Synthesis, uncertainties, and vulnerabilities. *Biogeosciences* 11(2):381–407.
46. Haverd V, et al. (2013) The Australian terrestrial carbon budget. *Biogeosciences* 10(2):851–869.