

1 **Classification:**

2 **Biological Science – Environmental Sciences**

3

4 **Title:**

5 **Reducing uncertainties in decadal variability of the**  
6 **global carbon budget with multiple data sets**

7 **Wei Li<sup>a,1</sup>, Philippe Ciais<sup>a</sup>, Yilong Wang<sup>a</sup>, Shushi Peng<sup>a</sup>, Gregoire Broquet<sup>a</sup>, Ashley P.**  
8 **Ballantyne<sup>b</sup>, Josep G. Canadell<sup>c</sup>, Leila A. Cooper<sup>b</sup>, Pierre Friedlingstein<sup>d</sup>, Corinne Le**  
9 **Quéré<sup>e</sup>, Ranga B. Myneni<sup>f</sup>, Glen Peters<sup>g</sup>, Shilong Piao<sup>h</sup>, Julia Pongratz<sup>i</sup>**

10

11 *Authors from A.P.B. onward are in the alphabetical order of last names.*

12

13 <sup>a</sup>Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ,  
14 Université Paris-Saclay, F-91191 Gif-sur-Yvette, France

15 <sup>b</sup>Department of Ecosystem and Conservation Science, University of Montana, Missoula, MT,  
16 USA

17 <sup>c</sup>Global Carbon Project, CSIRO Oceans and Atmosphere, GPO Box 3023, Canberra, ACT 2601,  
18 Australia

19 <sup>d</sup>College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter EX4  
20 4QF, UK

21 <sup>e</sup>Tyndall Centre for Climate Change Research, University of East Anglia, Norwich Research Park,  
22 Norwich NR4 7TJ, UK

23 <sup>f</sup>Department of Earth and Environment, Boston University, Boston, MA 02215, USA

24 <sup>g</sup>Center for International Climate and Environmental Research – Oslo (CICERO), Oslo, Norway

25 <sup>h</sup>Department of Ecology, College of Urban and Environmental Sciences, Peking University,  
26 Beijing 100871, China

27 <sup>i</sup>Max Planck Institute for Meteorology, Hamburg, Germany

28

29 <sup>1</sup>Corresponding author phone: 0033-6-58276925; email: wei.li@lsce.ipsl.fr

30

31 **Keywords:**

32 *global carbon budget; decadal variations; Bayesian fusion*

33 **Abstract**

34 Conventional calculations of the global carbon budget infer the land sink as a residual between  
35 emissions, atmospheric accumulation and the ocean sink. Thus, the land sink accumulates the  
36 errors from the other flux terms and bears the largest uncertainty. Here, we present a Bayesian  
37 fusion approach that combines multiple observations in different carbon reservoirs to optimize the  
38 land (B) and ocean (O) carbon sinks, land-use change emissions (L), and indirectly fossil fuel  
39 emissions (F) from 1980 to 2014. Compared to the conventional approach, Bayesian optimization  
40 decreases the uncertainties in B by 41% and in O by 46%. The L uncertainty decreases by 47%  
41 whereas F uncertainty is marginally improved through the knowledge of natural fluxes. Both  
42 ocean and net land uptake (B+L) rates have positive trends of  $29 \pm 8$  and  $37 \pm 17$  Tg C yr<sup>-2</sup> since  
43 1980, respectively. Our Bayesian fusion of multiple observations reduces uncertainties thereby  
44 allowing us to isolate important variability in global carbon cycle processes.

45

46 **Significance Statement**

47 The conventional approach of calculating the global carbon budget makes the land sink the most  
48 uncertain of all budget terms. This is because rather than being constrained by observations it is  
49 inferred as a residual in the budget equation. Here, we overcome this limitation by performing a  
50 Bayesian fusion of different available observation-based estimates of decadal carbon fluxes. This  
51 approach reduces the uncertainty in the land sink by 41% and in the ocean sink by 46%. These  
52 results are significant because they give unprecedented confidence in the role of the increasing  
53 land sink in regulating atmospheric CO<sub>2</sub>, and shed light on the past decadal trend.

54

55

56 \body

57 The land and ocean carbon sinks provide a vital climate mitigation ‘service’ by absorbing on  
58 average about 55% of anthropogenic CO<sub>2</sub> emissions from fossil fuel combustion and land-use  
59 change. Research has focused on understanding the relationships between year-to-year variability  
60 in carbon sinks and climate (1, 2), as well as the long term trend over the full instrumental period  
61 of CO<sub>2</sub> monitoring at the Mauna Loa station (3). Quasi-decadal variations of emissions and sinks  
62 have received comparatively less attention. Yet, significant climate variation occurs at this  
63 specific time scale (4). Since 1980, the variable occurrence of different ENSO events, two large  
64 volcanic eruptions (El Chichón and Pinatubo) and the recent slow-down of land surface warming  
65 have modulated the strength of carbon sinks. There are also decadal-scale changes in the rate at  
66 which human activities perturb the natural carbon cycle, in particular the recent acceleration of  
67 fossil fuel and cement emissions in the 2000s (5) and the slow-down in global land-use change  
68 emissions (LUC) in the mid-2000s, which appears to be partly driven by reduced deforestation in  
69 Brazil (6).

70 Here, we provide a data-driven assessment of global CO<sub>2</sub> emissions and sinks at 5-year intervals  
71 for the period of 1980-2014. We use a new Bayesian fusion approach whereby different data-  
72 streams of ocean and land uptake, LUC emissions, are optimally combined, and their uncertainty  
73 reduced from prior knowledge. This approach estimates the land sink constrained by data, which  
74 is a major improvement over the “conventional” method for calculating the global carbon budget  
75 by Ciais et al. (7) and Le Quéré et al. (8), hereafter [LQ15](#), where the unknown land sink was  
76 determined as a residual from the other components (emissions, atmospheric increase, ocean  
77 uptake). Most of the data-streams used in this analysis start in 1980, and about half of them give  
78 decadal mean values of natural sinks and thus do not allow us to tackle the reconstruction of  
79 interannual variability. Our choice of applying a Bayesian fusion approach to optimize 5-year

80 average component fluxes of the global carbon budget is therefore a compromise that maximizes  
81 the use of available observations of decadal average fluxes.

82 The principle of the Bayesian fusion approach is to combine an *a priori* imperfect knowledge of  
83 fluxes with observations and their uncertainties to infer optimized estimates of fluxes. Here, we  
84 define *a priori* values of terms in the global carbon budget that are not from observations.  
85 Specifically, we set prior fossil fuel and cement emissions (F) from inventories and the simulated  
86 land, ocean and land-use change carbon fluxes from process-based models (Table S1). The fluxes  
87 in this study are only anthropogenic fluxes, assuming a mean CO<sub>2</sub> growth rate of zero during pre-  
88 industrial times. Observational data sets independent from those prior values are applied to  
89 constrain land-use change emissions (L), the ocean uptake of anthropogenic CO<sub>2</sub> (O), the land-  
90 biosphere sink (B) in ecosystems not affected by land-use change, and the net land flux (B+L)  
91 (Table S2). We select the constraining data from peer-reviewed publications and evaluate their  
92 reported uncertainties and possible error correlations with each other (see details in Table S2).

## 93 **Results**

94 In the optimization of the global anthropogenic carbon budget (Fig. 1), the prior value of F and its  
95 uncertainty (Table S1) were defined from the mean value and the range of different fossil fuel and  
96 cement emission inventories, namely from CDIAC (9), IEA (10), EDGAR (11) and BP (12).  
97 These emission inventories are not treated as direct observations of emissions, and there is  
98 currently no independent observation to verify F. The prior values of O are from seven ocean  
99 biogeochemistry models (8), and the prior values of B are from the nine TRENDY land carbon  
100 models (13). These prior values from state-of-the-art models are without direct observational  
101 constraints. Some of these models could possibly be tuned using similar observation-based data  
102 but which observational data were used for model tuning was not explicitly reported. The prior  
103 estimates of L are derived from the difference of simulated land carbon fluxes with and without

104 LUC in the TRENDY carbon models (13). All fluxes are defined as positive if CO<sub>2</sub> is emitted to  
105 the atmosphere by the land or the ocean reservoir. Uncertainties in the prior estimates of 5-yearly  
106 O, B and L, are set to the maximum between those reported by LQ15 and the standard deviations  
107 across models. All uncertainties here refer to 1- $\sigma$  Gaussian errors. In this context, the prior  
108 uncertainties are 0.5 for O, 0.9 for B, and 0.8 Pg C yr<sup>-1</sup> for L, thus not smaller than the values of  
109 0.5, 0.8 and 0.5 Pg C yr<sup>-1</sup> from LQ15. It is important that the prior uncertainties are not too small,  
110 so that adding observations can adjust and constrain the sought fluxes.

111 Several independent data-streams, each with their specific uncertainty and temporal averaging  
112 period (Table S2), are combined in the Bayesian optimization with the above prior knowledge.  
113 These data-streams are: 1) the atmospheric CO<sub>2</sub> growth rate (CGR) from the NOAA/ESRL  
114 atmospheric network (14) which constrains the sum of all fluxes and is determined very  
115 accurately from more than 60 monitoring stations; 2) the atmospheric 5-year mean (negative)  
116 growth rate of O<sub>2</sub>/N<sub>2</sub> in the atmosphere from the Scripps O<sub>2</sub> Program (15) which relates to the  
117 combined effect of B+L and F changes, while being insensitive to changes in O (note that O<sub>2</sub>/N<sub>2</sub>  
118 has a negative trend in the atmosphere); 3) a set of yearly mean estimates of O from observational  
119 products based on *in situ* partial pressure of CO<sub>2</sub> (pCO<sub>2</sub>) surveys corrected for natural ocean CO<sub>2</sub>  
120 outgassing from carbon delivered by rivers (16) and using a neural network approach (17) and a  
121 diagnostic mixed-layer approach (18), and a set of decadal-mean estimates of O from inventories  
122 of carbon change in the ocean deduced indirectly from chlorofluorocarbons (CFCs) (19)  
123 combined with <sup>14</sup>C (20), and observed atmospheric mean CO<sub>2</sub> level and oceanic CO<sub>2</sub> and  
124 dissolved inorganic carbon observations (21); 4) ten-year mean estimates of B from a global  
125 synthesis of changes in forest carbon stocks (22); 5) decadal mean B+L from inventory-based  
126 land carbon storage change from the RECCAP publications on regional budgets (Table S3); 6)  
127 five-year mean LUC emissions from two independent bookkeeping approaches constrained by  
128 observed carbon stocks (23, 24). The uncertainties in each data-stream are either derived directly

129 from the original publications (when reported) or estimated from expert judgments (see details in  
130 [Table S2](#)). The optimization is performed for seven consecutive 5-year windows between 1980  
131 and 2014.

132 In the Bayesian optimization, observations that describe mean fluxes during intervals longer than  
133 5 years are still useful to infer 5-yearly fluxes. For example, the mean ocean sink observation for  
134 the 1990s (19) constrains the mean 5-yearly O during 1990-1999, while other independent  
135 observations (O<sub>2</sub>/N<sub>2</sub> and CGR) help to further allocate O values between the periods 1990-1994  
136 and 1995-1999. Despite no direct observation of F, this flux is found to be slightly improved in  
137 the Bayesian fusion, through knowledge of the other terms, and because the sum of all fluxes is  
138 very well constrained from CGR observations. We are aware that some observation-based land  
139 sink estimates have systematic errors in the way they are included in the optimization. In  
140 particular, the estimate of B from [ref. \(22\)](#) is only for forests and ignores other biomes. However,  
141 the RECCAP studies (25–27) and other estimates (22, 28) of the carbon stock change in non-  
142 forest biomes suggest that the forest sink alone accounts for most of the global land sink B.

143 The improved global budget of anthropogenic CO<sub>2</sub> is shown in [Fig. 2](#), and all data are given in  
144 [Table S1](#). After optimization, the *a posteriori* uncertainty in each flux is reduced. Compared to  
145 the conventional method applied by [LQ15](#) and IPCC-AR5 (7), uncertainties in B and O are  
146 reduced by 41% and 46% in this study. In the Bayesian data fusion, the land sink is no longer  
147 solely inferred as a residual that accumulates uncertainties from all other terms, and it exhibits a  
148 large reduction in uncertainty. The uncertainty in L decreases by 47% but the uncertainty in F is  
149 marginally improved (by 1%) through the indirect constraints of other terms. In the absence of  
150 direct constraint on F, this small reduction in the F uncertainties compared to [LQ15](#) and IPCC-  
151 AR5 (7) is also because we use multiple emission inventories (while [LQ15](#) and IPCC-AR5 (7)  
152 only used CDIAC (9)) and start at relatively higher prior uncertainties in F ([Table S1](#)) than in  
153 [LQ15](#). Despite their improved (smaller) uncertainties, the 5-year mean fluxes shown in [Table S1](#)

154 do not differ statistically in their mean values from LQ15. This indicates that each flux of the  
155 Bayesian carbon budget is fully consistent with LQ15 even though we used an array of data with  
156 different measurement methods and with uncertainties estimated in different ways. Specifically,  
157 we obtain emissions from fossil fuel burning and cement production that are smaller than LQ15  
158 by  $0.18 \pm 0.19$  Pg C yr<sup>-1</sup> during 1980-2014 (Fig. 2). A downward revision of global F is consistent  
159 with the correction of the emissions for China based on evidence of the lower carbon content for  
160 coal burned in that country (29). Compared to LQ15, the optimized ocean sink during 2000-2004  
161 is larger by  $0.22$  Pg C yr<sup>-1</sup> but lower by  $0.18 \pm 0.10$  Pg C yr<sup>-1</sup> during all the other periods. In the  
162 past decade (2005-2014), both ocean sink and land sink from our optimization are smaller than  
163 LQ15. The optimized fluxes of L are similar to or lower than those from LQ15. The trend of F for  
164 the seven 5-year periods is positive ( $p = 0.003$ ), with a probability of a positive trend for ocean  
165 and net land (B+L) uptake rates of 93%; the trend of B or L individually is not statistically  
166 significant ( $p = 0.23$  for both). The increasing rate of ocean and net land uptake are  $29 \pm 8$  and  
167  $37 \pm 17$  Tg C yr<sup>-2</sup> since 1980, respectively. Similar statistically positive trends were also found in  
168 the 5-year mean ocean and net land uptake rates between 1980 and 2014 calculated from the  
169 yearly budget updated by LQ15 (Fig. S1A). Given the robustness of O inferred by our  
170 optimization (see also Fig. S2), and in view of the many observations constraining this flux, there  
171 is a high confidence that the ocean sink has been increasing over time since 1980. The optimized  
172 land sink is less variable between different 5-year periods than in LQ15 (Fig. S1A). But the ocean  
173 sink is more variable, with a standard deviation of  $0.34$  Pg C yr<sup>-1</sup> compared to  $0.27$  Pg C yr<sup>-1</sup> by  
174 LQ15 (standard deviations across the seven periods analyzed).

175 From 1980 to 2014, the average fractions of F+L emission re-absorbed by the land and ocean  
176 carbon reservoirs are  $29 \pm 5\%$  (mean  $\pm 1 \sigma$  of inter 5-year variability) and  $26 \pm 2\%$ , respectively. The  
177 ratios of both O and B to F+L emission do not exhibit any significant trends ( $p > 0.05$ , Fig. S1B).  
178 Even with their reduced uncertainty in this study compared to LQ15, the variability of O and B



179 between 5-year intervals prevents us from assessing the very small trends in their ratios to  
180 emissions. Similarly, we found no significant trend in the ratio of O or B to fossil fuel emission  
181 (F). The larger variability of the B-to-(F+L) ratios compared to the O-to-(F+L) ones (Fig. S1B)  
182 suggests that the efficiency of the land sink at absorbing emissions is more variable than that of  
183 the ocean sink. For instance, during the period that followed the cooling from the Pinatubo  
184 eruption in 1990-1994 (30–32), the B-to-(F+L) ratio increased by 46% above its long-term mean.  
185 This ratio was also higher than normal during 2005-2009, possibly due to the absence of El Niño  
186 and to the occurrence of a cooler and wetter La Niña event in 2008-2009 manifested by lower  
187 than normal CGR (3).

## 188 **Discussion**

189 In our Bayesian fusion, the CGR constrains the sum of the different fluxes, thus an  
190 overestimation of a component flux due to the use of a single published estimate of that flux  
191 would result in an underestimation of another flux, leading to negative correlations between  
192 uncertainties in the different components of the posterior fluxes. These negative correlations are  
193 clear between uncertainties in B and F (ranging from -0.80 to -0.35), in B and O (from -0.60 to -  
194 0.30), and in B and L (from -0.53 to -0.24) (Fig. S3). This indicates that although the  
195 uncertainties for each flux can be significantly decreased, the remaining (posterior) uncertainties  
196 are hard to be decoupled. In comparison, we also calculate the correlations between uncertainties  
197 in B and other fluxes by classical error propagation rule from mass balance equation  $B = CGR -$   
198  $F - O - L$ . The “conventional” error budget calculation gives a typical uncertainty of 0.76 Pg C  
199  $yr^{-1}$  for B. Negative error correlations also exist in that approach between B and F (-0.39),  
200 between B and O (-0.79), and between B and L (-0.39). Thus in our study, not only the posterior  
201 uncertainties of B (0.42-0.53 Pg C  $yr^{-1}$ ) are smaller than those calculated by the “conventional”  
202 approach, but also the negative error correlation between B and O is smaller, indicating a better

203 separation between the carbon sinks of B and O. In addition, although the negative correlations  
204 between B and F, and between B and L are slightly larger than those calculated by the  
205 “conventional” approach, it does not mean our approach has poorer ability to separate these  
206 components because we have succeed in significantly decreasing the absolute uncertainties for all  
207 the fluxes thus allowing us to draw more definitive conclusions.

208 The Bayesian fusion of different observational components of the anthropogenic CO<sub>2</sub> budget  
209 proposed in this study provides the most robust estimate to date of the strength and evolution of  
210 the land sink on 5-year intervals and brings a more robust picture of the current perturbation of  
211 the carbon cycle. Future work could apply the same fusion approach to regional carbon budget  
212 estimates and to gross CO<sub>2</sub> fluxes of photosynthesis, respiration and fire emissions.

213

## 214 **Materials and Methods**

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

$$224 \quad \mathbf{A} = (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \quad (1)$$

$$225 \quad \mathbf{x}^a = \mathbf{x}^b + \mathbf{A} \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{y}^o - \mathbf{H} \mathbf{x}^b) \quad (2)$$

226 where the superscripts T and “-1” denote the transpose and inverse of a matrix, respectively.

227 In the optimization, we update the estimate of the mean fluxes of fossil fuel and cement emissions  
228 (F), ocean sink (O), land sink (B) and land-use change emissions (L) for each 5-year interval from  
229 1980 to 2014 (Fig. 1 and Table S1). The observation vector contains estimates of 5-year mean  
230 global value for: the atmospheric growth rates of CO<sub>2</sub> (CGR, in Pg C yr<sup>-1</sup>), atmospheric growth  
231 rates of O<sub>2</sub>/N<sub>2</sub> (CGR-O<sub>2</sub>, per meg unit), observation-based estimates of ocean sinks, land sinks,  
232 land-use change emissions and net land sink (B+L) (the data sources for these components of  $\mathbf{y}^o$   
233 are summarized in Fig. 1 and provided in Table S2).  $\mathbf{H}$  is defined for each 5-year interval by:

$$\begin{array}{l}
234 \\
\mathbf{H}: \begin{bmatrix} \mathbf{F} \\ \mathbf{O} \\ \mathbf{B} \\ \mathbf{L} \end{bmatrix} \rightarrow \begin{bmatrix} \text{CGR} \\ \text{CGR} - \text{O}_2 \\ \mathbf{O} \\ \mathbf{B} \\ \mathbf{L} \\ \mathbf{B} + \mathbf{L} \end{bmatrix} = \begin{bmatrix} \mathbf{F} + \mathbf{O} + \mathbf{B} + \mathbf{L} \\ \alpha_{\mathbf{F}}\mathbf{F} + \alpha_{\mathbf{B}}(\mathbf{B} + \mathbf{L}) + \mathbf{Z}_{\text{O}_2} \\ \mathbf{O} \\ \mathbf{B} \\ \mathbf{L} \\ \mathbf{B} + \mathbf{L} \end{bmatrix}
\end{array} \quad (3)$$

235 where  $\alpha_{\mathbf{F}}$ ,  $\alpha_{\mathbf{B}}$  and  $\mathbf{Z}_{\text{O}_2}$  are constant coefficients from ref. (15).

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

243 **Data used to derive the Prior statistics of the control variables.** To define the prior estimate of  
244 F, individual country data from four emission inventories (CDIAC (9), IEA (10), EDGAR (11)  
245 and BP (12)) are grouped into geographic regions as specified by the United Nations Statistics  
246 Division (<http://unstats.un.org/unsd/methods/m49/m49regin.htm>). Cement emissions from  
247 EDGAR are added into the IEA and BP data sets that do not include cement emissions.  
248 Uncertainties for each country (35) are used to create regional uncertainty distributions for each  
249 emission index using a bootstrapping method, with the uncertainties of the highest emitters in  
250 each region contributing the most to the uncertainty distributions. This effect is achieved by  
251 weighting the sampling probability ( $P_s$ ) by the relative contribution of each country's emissions  
252 ( $E_C$ ) to the total emissions within the region ( $E_R$ ):

$$253 \quad P_s = \frac{E_C}{E_R} \quad (4)$$

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

$$259 \quad \varepsilon_{F(t)} = 0.95 \times \varepsilon_{F(t-1)} + \varepsilon_{(t)} \quad (5)$$

260 where emission error factors for any given year  $\varepsilon_{F(t)}$  are correlated with the emission errors from  
261 the previous year  $\varepsilon_{F(t-1)}$  by an autoregressive coefficient of 0.95 with  $\varepsilon_{(t)}$  as random error. The  
262 autocorrelated time series are then multiplied and added to the fossil fuel emissions for each  
263 country, and subsequently 500 samples of global fossil fuel emissions are taken for each 5-year  
264 bin. The means and standard deviations of each bin for each inventory are calculated from these  
265 500 samples. Additionally, the correlation in global uncertainty is calculated between 5-year bins  
266 and inventories to produce an error-covariance matrix. The maximum between the uncertainties  
267 calculated above and the standard deviations of the 5-year means across four emission inventories  
268 were adopted as the uncertainties in the prior estimate of F.

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

277 Prior values of L and B are set from simulations in the TRENDY (v2) model intercomparison  
 278 project (13). The simulations in TRENDY (v2) are up to 2012, and thus the priors for the period  
 279 of 2010-2012 were used for the period of 2010-2014. All the prior flux values are summarized in  
 280 [Table S1](#).

281 **Uncertainty Correlations between Optimized Variables.** The correlations of the uncertainties  
 282 between optimized variables are shown in [Fig. S3](#). In comparison, we also calculate the  
 283 correlations between uncertainties in the fluxes when deriving the estimate of B using the  
 284 classical mass balance equation  $B = CGR - F - O - L$ . First, the uncertainties in an estimate of the  
 285 flux  $i$  is calculated using inverse variance method, based on the variances associated with all the  
 286 estimates of this flux ( $j=1,2,\dots,m$ ):

$$287 \quad \sigma_i = \sqrt{\frac{1}{\sum_{j=1}^m \frac{1}{\sigma_{i,j}^2}}} \quad (6)$$

288 The resulting uncertainties for CGR, F, O and L are: 0.2, 0.3, 0.6 and 0.3 Pg C yr<sup>-1</sup>, respectively.  
 289 Based on a simple error propagation:  $\varepsilon_B = \varepsilon_{CGR} - \varepsilon_F - \varepsilon_O - \varepsilon_L$  and the independence of the  
 290 estimates of CGR, F, O and L, the variance of the uncertainty in B is given by:

$$291 \quad \sigma_B^2 = E(\varepsilon_B^T \varepsilon_B) = \sigma_{CGR}^2 + \sigma_F^2 + \sigma_O^2 + \sigma_L^2 \quad (7)$$

292 where  $E(\cdot)$  is the expectation of a variable. The correlation between the uncertainties in B and in  
 293 another variable, e.g. O in the equation below, is given by:

$$294 \quad r(\varepsilon_B, \varepsilon_O) = \frac{E(\varepsilon_B^T \varepsilon_O)}{\sigma_B \sigma_O} = -\frac{\sigma_O}{\sigma_B} \quad (8)$$

295 **Sensitivity Tests.** Four sets of sensitivity tests were conducted ([Fig. S2](#)): (a) using F from  
 296 different inventory data sets (IEA, EDGAR, CDIAC and BP, respectively) as priors; (b) using  
 297 different ocean constraining data sets (only data from ref. (17), only data from ref. (18), data from  
 298 ref. (17) and (18) but without other ocean data, and all ocean data except ref. (19), respectively);

299 (c) using different tiers of constraining data (Tier 1 only, Tier 1 + Tier 2, and Tier1 + Tier 2 +  
300 Tier 3, respectively); and (d) using enlarged prior uncertainties ( $10^6$  Pg C yr<sup>-1</sup> except F), prior B  
301 and O from CMIP5 models (36) instead of TRENDY (v2) models, and a subset of constraining  
302 data including CGR, O<sub>2</sub>/N<sub>2</sub>, L from ref. (24) and O from pCO<sub>2</sub> and inventories (17, 18, 20). The  
303 constraining data tiers (Table S2) were defined as: Tier 1, direct carbon observations (e.g. CGR);  
304 Tier 2, indirect carbon observations unambiguously related to carbon quantities (e.g. O<sub>2</sub>/N<sub>2</sub>); Tier  
305 3, direct carbon observations with an empirical (data-driven) model used to obtain global  
306 estimates, e.g. the use of geo-statistics to up-scale local data into global values; Tier 4, indirect  
307 carbon observations not simply related to global carbon flux quantities.

308 The ocean sink is rather consistent in all sensitivity tests (Fig. S2), implying it is very robustly  
309 constrained in the optimization because of the sufficient number and consistency of constraining  
310 observational data (see the O constraining data in Table S2). Small changes in optimized F were  
311 found (Fig. S2A) in the sensitivity tests using prior F from different data sets, and the land sink is  
312 more dependent on the prior F choice. In the sensitivity tests using different ocean constraining  
313 data (Fig. S2B), all fluxes are generally consistent, and very slight changes for O appear in the  
314 tests using ocean data from ref. (17) only and from ref. (18) only during 2000-2004 and 2005-  
315 2009. The decadal estimates (19–21) only have small impact on their corresponding periods. The  
316 important point for the sensitivity tests using different tiers of constraining data is that there is no  
317 inconsistency (i.e. shift in posterior estimate) between the assimilations of different “observation  
318 based” tiers although L may change due to lack of individual constrains in Tier 1 and Tier 1 +  
319 Tier 2 (Fig. S2C). With enlarged prior uncertainties ( $10^6$  Pg C yr<sup>-1</sup> except F), the mean flux values  
320 do not shift in general but the posterior uncertainties increase (Fig. S2D). Results are also very  
321 consistent by replacing prior B and O from TRENDY v2 models with from CMIP5 models (Fig.  
322 S2D). The sensitivity test using a subset of constraining data underestimate B and L compared to

323 the original optimized results (Fig. S2D) because the L estimates in ref. (23) (used in the original  
324 optimization but not in this sensitivity test) is higher than that in ref. (24) (Table S2).

325 **Trend Test.** A Mann-Kendall statistical test was applied as a trend test.

326

### 327 **Acknowledgements**

328 This paper builds on an analysis started by the late Michael R. Raupach and is a contribution to  
329 the work of the Global Carbon Project to understand and better constrain the human perturbation  
330 of the carbon budget. We thank Alessandro Anav for providing the data from CMIP5 and Peter  
331 Landschützer and Christian Rödenbeck for sharing their ocean flux data. We are grateful to the  
332 scientists from NOAA/ESRL and the Scripps O<sub>2</sub> Program for making the invaluable data from the  
333 long-term measurements of CGR and O<sub>2</sub>/N<sub>2</sub> available to the community.

334 W.L. is supported by the European Commission-funded project LUC4C (No. 603542). P.C. and  
335 S.Pe. acknowledge support from the European Research Council through Synergy grant ERC-  
336 2013-SyG-610028 “IMBALANCE-P”. J.G.C. is supported by the Australian Climate Change  
337 Science Program, G.P. is supported by the Norwegian Research Council (236296), and J.P. is  
338 supported by the German Research Foundation's Emmy Noether Program.

339



340 **References**

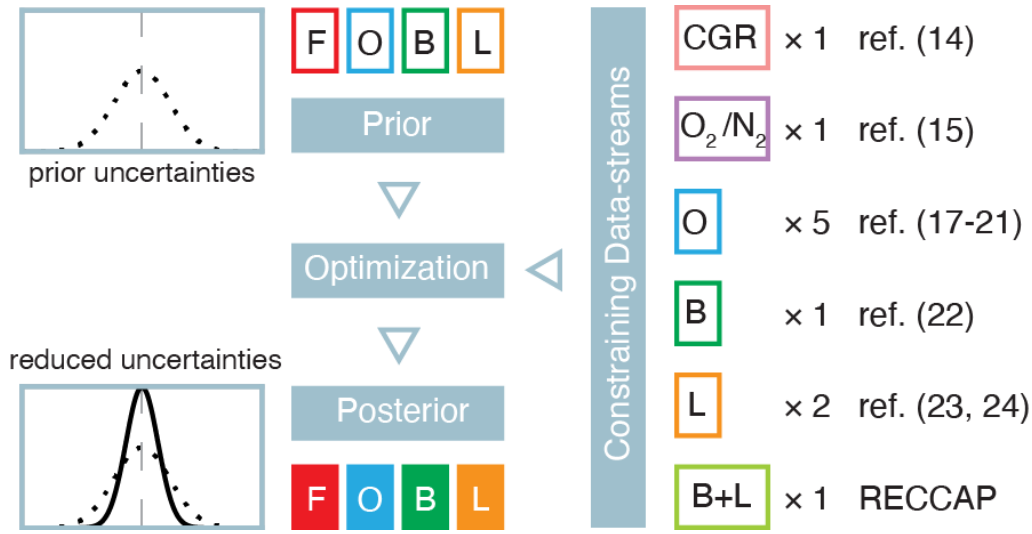
- 341 1. Wang X, et al. (2014) A two-fold increase of carbon cycle sensitivity to tropical temperature  
342 variations. *Nature* 506(7487):212–5.
- 343 2. Poulter B, et al. (2014) Contribution of semi-arid ecosystems to interannual variability of the global  
344 carbon cycle. *Nature* 509(7502):600–3.
- 345 3. Le Quéré C, et al. (2009) Trends in the sources and sinks of carbon dioxide. *Nat Geosci* 2(12):831–  
346 836.
- 347 4. Ghil M, Vautard R (1991) Interdecadal oscillations and the warming trend in global temperature  
348 time series. *Nature* 350(6316):324–327.
- 349 5. Raupach MR, et al. (2007) Global and regional drivers of accelerating CO<sub>2</sub> emissions. *Proc Natl*  
350 *Acad Sci U S A* 104(24):10288–93.
- 351 6. Song X-P, Huang C, Saatchi SS, Hansen MC, Townshend JR (2015) Annual Carbon Emissions  
352 from Deforestation in the Amazon Basin between 2000 and 2010. *PLoS One* 10:e0126754.
- 353 7. Ciais P, et al. (2013) 2013: Carbon and Other Biogeochemical Cycles. *Clim Chang 2013 Phys Sci*  
354 *Basis Contrib Work Gr I to Fifth Assess Rep Intergov Panel Clim Chang*:465–570.
- 355 8. Le Quéré C, et al. (2015) Global carbon budget 2014. *Earth Syst Sci Data*:47–85.
- 356 9. Boden TA, Marland G, Andres RJ (2013) Global, regional, and national fossil-fuel CO<sub>2</sub> emissions.  
357 *Carbon Dioxide Inf Anal Center, Oak Ridge Natl Lab USA Oak Ridge, TN Dep Energy*.  
358 doi:10.3334/CDIAC/00001.
- 359 10. International Energy Agency (2013) CO<sub>2</sub> Emissions From Fuel Combustion Highlights. *IEA*  
360 *Stat*:158.
- 361 11. Olivier JGJ, Janssens-Maenhout G, Peters JAHW (2014) *Trends in global CO<sub>2</sub> emissions: 2014*  
362 *Report* (PBL Netherlands Environmental Assessment Agency Hague).
- 363 12. BP (2015) BP Statistical Review of World Energy 2015.
- 364 13. Sitch S, et al. (2015) Recent trends and drivers of regional sources and sinks of carbon dioxide.  
365 *Biogeosciences* 12(3):653–679.
- 366 14. NOAA/ESRL NOAA/ESRL calculation of global means. Available at:  
367 [http://www.esrl.noaa.gov/gmd/ccgg/about/global\\_means.html](http://www.esrl.noaa.gov/gmd/ccgg/about/global_means.html) [Accessed August 9, 2015].
- 368 15. Keeling RF, Manning a. C (2014) *Studies of Recent Changes in Atmospheric O<sub>2</sub> Content* (Elsevier  
369 Ltd.). 2nd Ed. doi:10.1016/B978-0-08-095975-7.00420-4.
- 370 16. Jacobson AR, Fletcher SEM, Gruber N, Sarmiento JL, Gloor M (2007) A joint atmosphere-ocean  
371 inversion for surface fluxes of carbon dioxide: 1. Methods and global-scale fluxes. *Global*  
372 *Biogeochem Cycles* 21(1). doi:10.1029/2005GB002556.
- 373 17. Landschützer P, Gruber N, Bakker DCE, Schuster U (2014) Recent variability of the global ocean  
374 carbon sink. *Global Biogeochem Cycles* 28(9):927–949.
- 375 18. Rödenbeck C, et al. (2014) Interannual sea–air CO<sub>2</sub> flux variability from an observation-driven  
376 ocean mixed-layer scheme. *Biogeosciences* 11(17):4599–4613.
- 377 19. McNeil BI, Matear RJ, Key RM, Bullister JL, Sarmiento JL (2003) Anthropogenic CO<sub>2</sub> uptake by  
378 the ocean based on the global chlorofluorocarbon data set. *Science* 299(5604):235–9.
- 379 20. Khatiwala S, Primeau F, Hall T (2009) Reconstruction of the history of anthropogenic CO<sub>2</sub>(<sub>2</sub>)  
380 concentrations in the ocean. *Nature* 462(7271):346–9.
- 381 21. Steinkamp K, Gruber N (2013) A joint atmosphere-ocean inversion for the estimation of seasonal

- 382 carbon sources and sinks. *Global Biogeochem Cycles* 27(3):732–745.
- 383 22. Pan Y, et al. (2011) A large and persistent carbon sink in the world's forests. *Science*  
384 333(6045):988–993.
- 385 23. Hansis E, Davis SJ, Pongratz J (2015) Relevance of methodological choices for accounting of land  
386 use change carbon fluxes. *Global Biogeochem Cycles* 29(8):1230–1246.
- 387 24. Houghton R a., et al. (2012) Carbon emissions from land use and land-cover change.  
388 *Biogeosciences* 9(12):5125–5142.
- 389 25. Piao SL, et al. (2012) The carbon budget of terrestrial ecosystems in East Asia over the last two  
390 decades. *Biogeosciences* 9(9):3571–3586.
- 391 26. Luysaert S, et al. (2012) The European land and inland water CO<sub>2</sub>, CO, CH<sub>4</sub> and N<sub>2</sub>O balance  
392 between 2001 and 2005. *Biogeosciences* 9(8):3357–3380.
- 393 27. King AW, et al. (2015) North America's net terrestrial CO<sub>2</sub> exchange with the atmosphere 1990–  
394 2009. *Biogeosciences* 12(2):399–414.
- 395 28. Liu YY, et al. (2015) Recent reversal in loss of global terrestrial biomass. *Nat Clim Chang*  
396 5(5):470–474.
- 397 29. Liu Z, et al. (2015) Reduced carbon emission estimates from fossil fuel combustion and cement  
398 production in China. *Nature* 524(7565):335–338.
- 399 30. Lucht W, et al. (2002) Climatic control of the high-latitude vegetation greening trend and Pinatubo  
400 effect. *Science* 296(5573):1687–9.
- 401 31. Mercado LM, et al. (2009) Impact of changes in diffuse radiation on the global land carbon sink.  
402 *Nature* 458(7241):1014–7.
- 403 32. Frölicher TL, Joos F, Raible CC, Sarmiento JL (2013) Atmospheric CO<sub>2</sub> response to volcanic  
404 eruptions: The role of ENSO, season, and variability. *Global Biogeochem Cycles* 27(1):239–251.
- 405 33. Tarantola A (2005) *Inverse Problem Theory and Methods for Model Parameter Estimation*  
406 (SIAM).
- 407 34. Ballantyne AP, et al. (2015) Audit of the global carbon budget: estimate errors and their impact on  
408 uptake uncertainty. *Biogeosciences* 12(8):2565–2584.
- 409 35. Andres RJ, Boden TA, Higdon D (2014) A new evaluation of the uncertainty associated with  
410 CDIAC estimates of fossil fuel carbon dioxide emission. *Tellus B* 66.  
411 doi:10.3402/tellusb.v66.23616.
- 412 36. Anav A, et al. (2013) Evaluating the Land and Ocean Components of the Global Carbon Cycle in  
413 the CMIP5 Earth System Models. *J Clim* 26(18):6801–6843.
- 414 37. Rödenbeck C, et al. (2015) Data-based estimates of the ocean carbon sink variability - First results  
415 of the Surface Ocean pCO<sub>2</sub> Mapping intercomparison (SOCOM). *Biogeosciences* 12(23):7251–  
416 7278.
- 417 38. Rödenbeck C, et al. (2013) Global surface-ocean p<sup>CO<sub>2</sub></sup> and sea–air CO<sub>2</sub> flux variability from an  
418 observation-driven ocean mixed-layer scheme. *Ocean Sci* 9(2):193–216.
- 419 39. Hurtt GC, et al. (2011) Harmonization of land-use scenarios for the period 1500–2100: 600 years of  
420 global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Clim*  
421 *Change* 109(1–2):117–161.
- 422 40. King AW, et al. (2015) North America's net terrestrial CO<sub>2</sub> exchange with the atmosphere 1990–  
423 2009. *Biogeosciences* 12(2):399–414.
- 424 41. Peters GP, Davis SJ, Andrew R (2012) A synthesis of carbon in international trade. *Biogeosciences*

- 425 9(8):3247–3276.
- 426 42. Hartmann J, Jansen N, Dürre HH, Kempe S, Köhler P (2009) Global CO<sub>2</sub>-consumption by chemical  
427 weathering: What is the contribution of highly active weathering regions? *Glob Planet Change*  
428 69(4):185–194.
- 429 43. Dolman AJ, et al. (2012) An estimate of the terrestrial carbon budget of Russia using inventory-  
430 based, eddy covariance and inversion methods. *Biogeosciences* 9(12):5323–5340.
- 431 44. Lauerwald R, Laruelle GG, Hartmann J, Ciais P, Regnier PAG (2015) Spatial patterns in CO<sub>2</sub>  
432 evasion from the global river network. *Global Biogeochem Cycles* 29(5):2014GB004941.
- 433 45. Gloor M, et al. (2012) The carbon balance of South America: a review of the status, decadal trends  
434 and main determinants. *Biogeosciences* 9(12):5407–5430.
- 435 46. Valentini R, et al. (2014) A full greenhouse gases budget of Africa: synthesis, uncertainties, and  
436 vulnerabilities. *Biogeosciences* 11(2):381–407.
- 437 47. Haverd V, et al. (2013) The Australian terrestrial carbon budget. *Biogeosciences* 10(2):851–869.  
438

439 **Figure captions:**

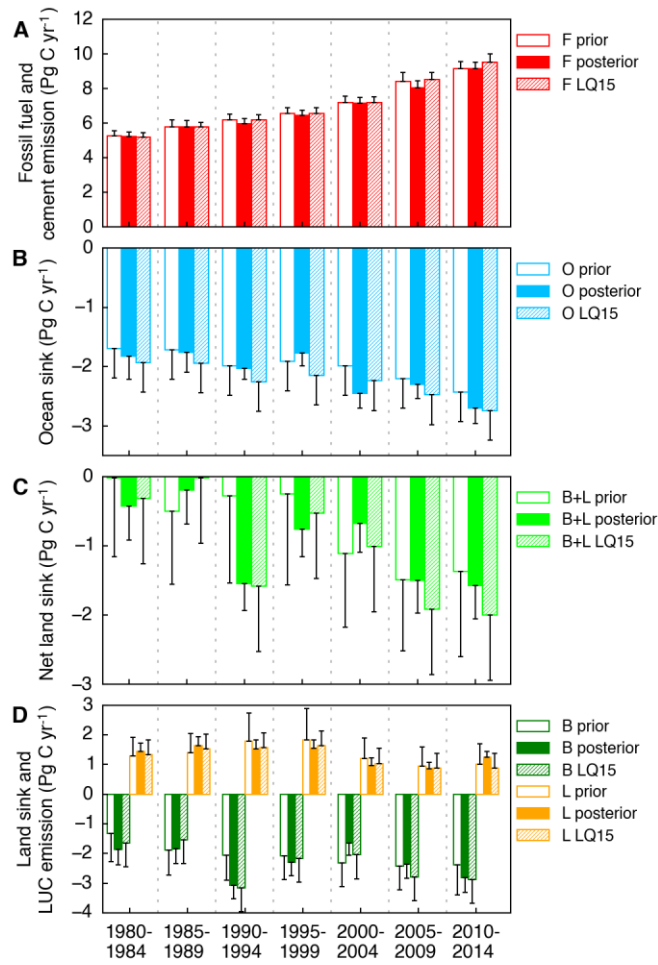
440 **Fig. 1.** The framework of our optimization. The number of constraining data-streams and the  
441 specific data sources are marked on the right. The fluxes that are optimized are 5-year averages of  
442 F, O, B and L representing fossil fuel and cement emissions, ocean sink, land sink, and land-use  
443 change emissions, respectively. The observations used to constrain these fluxes are the five-year  
444 averaged growth rates of CO<sub>2</sub> and O<sub>2</sub>/N<sub>2</sub> in the atmosphere, observations of O and B from carbon  
445 measurements made in these two reservoirs, and inventory-based estimates of L and the net land  
446 sink (B+L). In this framework, the CO<sub>2</sub> growth rate constrains the sum of all the fluxes. The  
447 O<sub>2</sub>/N<sub>2</sub> growth rate allows us to separate O and B+L and bring some constraint on F as well.



448

449

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



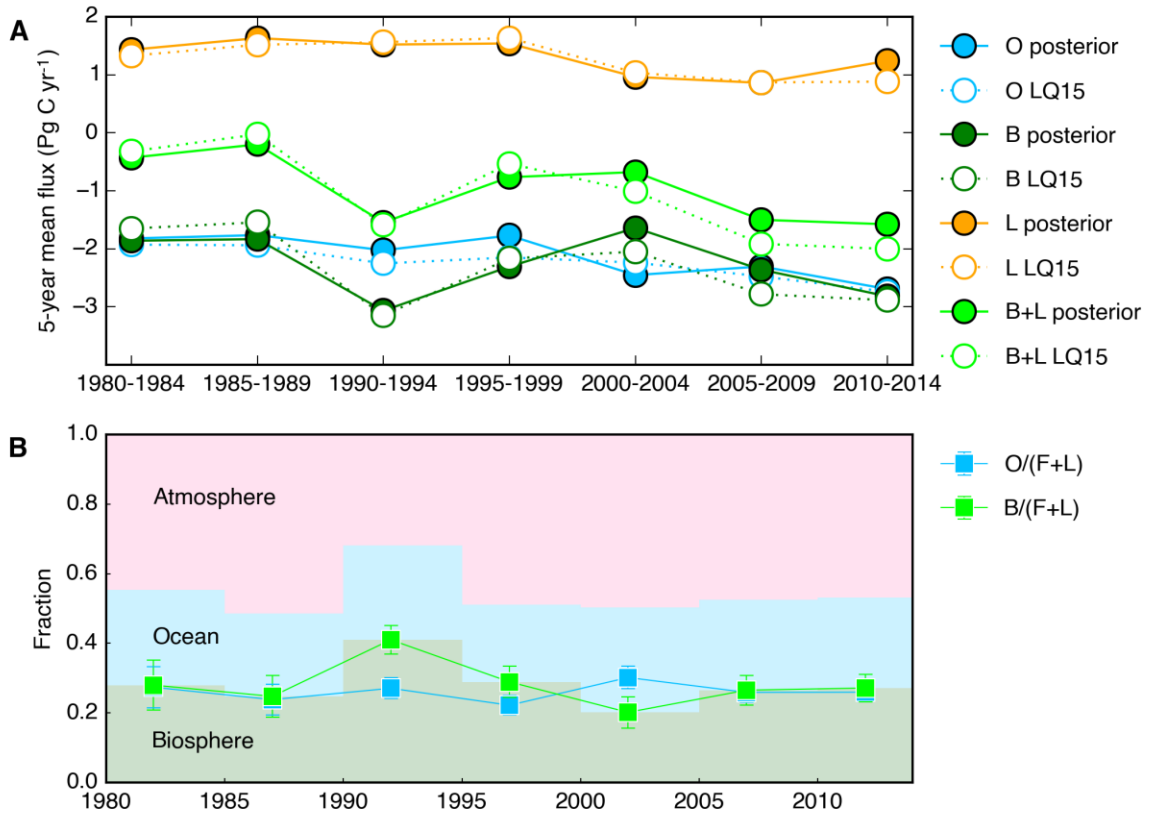
454

455

456

457 **SI Figure captions:**

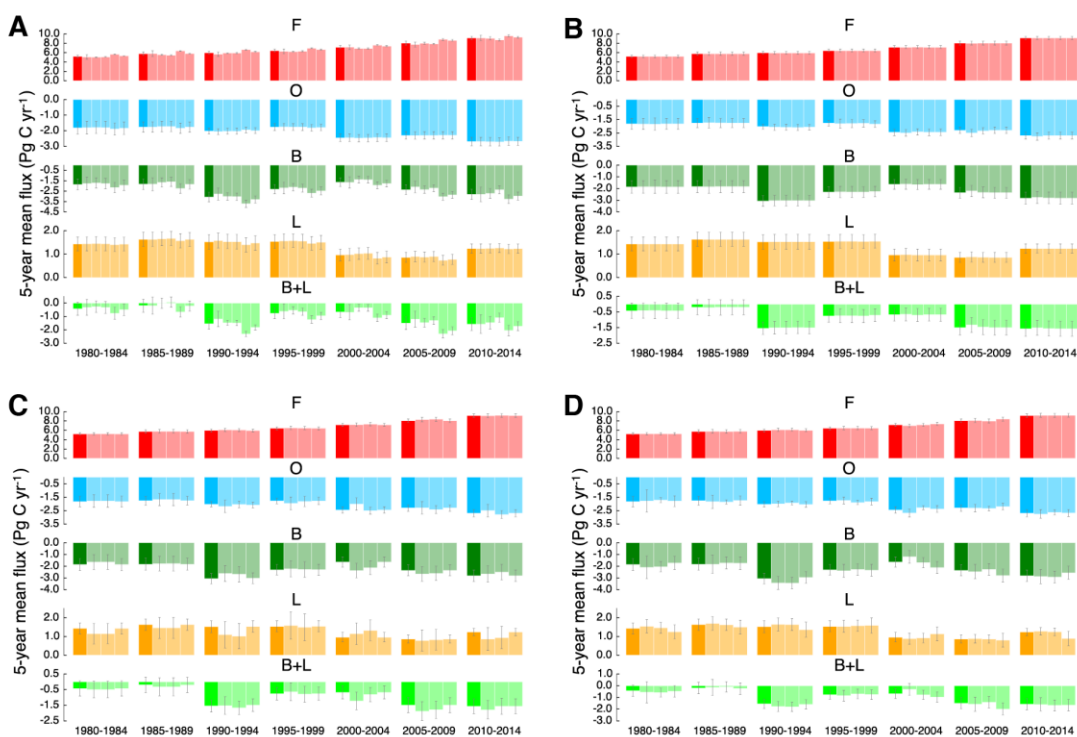
458 **Fig. S1.** (A) The temporal trend of ocean sink (O), land sink (B), land-use change emissions (L)  
 459 and net land flux (B+L) from our optimization and LQ15. (B) The partitioning of global total  
 460 emissions (fossil fuel and cement emissions and land-use change emissions, F+L) into  
 461 atmosphere, ocean (O) and biosphere (B, background shade). The markers are absolute absorbed  
 462 fractions and uncertainties in ocean (blue squares) and biosphere (green squares).



463

464

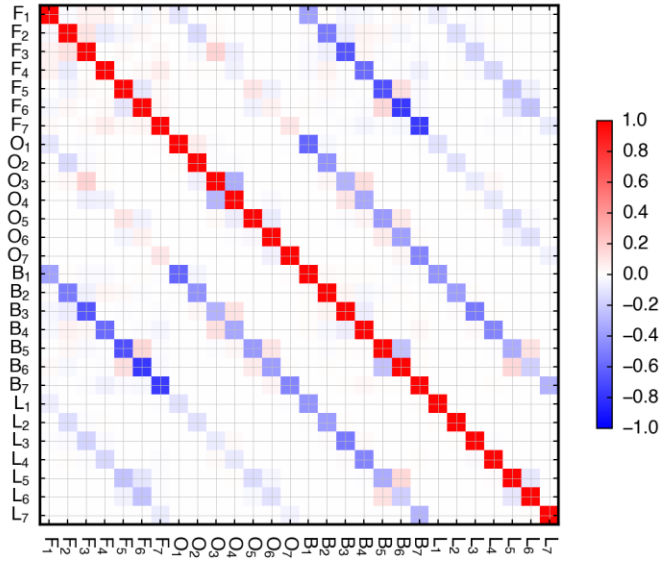
465 **Fig. S2.** Sensitivity tests using different data sets. F, O, B, L and B+L indicate fossil fuel and  
 466 cement emissions, ocean sink, land sink, land use change emissions and net land sink,  
 467 respectively. (A) Using F from different inventory data sets as priors. The bars around each flux  
 468 in each period from left to right represent the original optimization, double prior uncertainties,  
 469 prior F from IEA, prior F from EDGAR, prior F from CDIAC and prior F from BP, respectively.  
 470 (B) Using different ocean constraining data sets. The bars from left to right represent the original  
 471 optimization (all ocean data), only data from ref. (17), only data from ref. (18), data from ref. (17)  
 472 and (18) but without other ocean data, all ocean data except ref. (19), respectively. (C) Using  
 473 different tiers (Tier number shown in Table S2) of constraining data. The bars from left to right  
 474 represent the original optimization (all tiers), Tier 1 only, Tier 1 + Tier 2, Tier1 + Tier 2 + Tier 3,  
 475 respectively. (D) The bars from left to right represent the original optimization, enlarging prior  
 476 uncertainties to  $10^6$  Pg C yr<sup>-1</sup> except F, prior B and O from CMIP5, and using a subset of  
 477 constraining data including CGR, O<sub>2</sub>/N<sub>2</sub>, L from ref. (24) and O based on pCO<sub>2</sub> and inventories  
 478 (17, 18, 20).



479

480

481 **Fig. S3.** Correlations between the posterior uncertainties. The subscripts 1 to 7 represent the 5-  
482 year periods from 1980-1984 to 2010-2014 in sequence.



483

484



485 **Table S1** Prior and posterior values and uncertainties (Pg C yr<sup>-1</sup>). The adjusted O values (O<sub>adj</sub>)  
 486 from [LQ15](#) that are not used in the optimization are also shown for a comparison with the  
 487 unadjusted prior O values. The adjusted O values are on average 0.26±0.03 Pg C yr<sup>-1</sup> smaller than  
 488 the unadjusted ones.

	80-84	85-89	90-94	95-99	00-04	05-09	10-14	
prior	F	5.27±0.28	5.79±0.41	6.18±0.35	6.55±0.35	7.19±0.39	8.40±0.53	9.16±0.41
	O	-1.70±0.50	-1.71±0.50	-1.98±0.50	-1.91±0.50	-1.98±0.50	-2.20±0.50	-2.43±0.50
	O <sub>adj</sub>	-1.93±0.50	-1.94±0.50	-2.25±0.50	-2.15±0.50	-2.24±0.50	-2.48±0.50	-2.74±0.50
	B	-1.31±0.95	-1.89±0.83	-2.06±0.85	-2.08±0.80	-2.31±0.80	-2.43±0.80	-2.37±1.01
	L	1.30±0.62	1.39±0.65	1.78±0.94	1.83±1.04	1.20±0.70	0.94±0.65	1.00±0.70
posterior	F	5.22±0.27	5.77±0.38	5.96±0.31	6.43±0.32	7.17±0.33	8.03±0.40	9.14±0.37
	O	-1.82±0.40	-1.76±0.33	-2.02±0.19	-1.77±0.22	-2.45±0.24	-2.30±0.24	-2.69±0.26
	B	-1.86±0.53	-1.83±0.52	-3.07±0.45	-2.31±0.44	-1.64±0.42	-2.36±0.47	-2.82±0.50
	L	1.43±0.29	1.63±0.30	1.52±0.31	1.54±0.29	0.96±0.26	0.86±0.22	1.24±0.19
	B+L	-0.43±0.49	-0.20±0.49	-1.55±0.39	-0.77±0.40	-0.68±0.42	-1.50±0.48	-1.58±0.48

489

490

491 **Table S2** Sources and values of different constraining data-streams used in the optimization. The blue and orange shades indicate the length of time covered. The  
 492 constraining data tiers were defined as: Tier 1, direct carbon observations (e.g. CGR); Tier 2, indirect carbon observations unambiguously related to carbon  
 493 quantities (e.g. O<sub>2</sub>/N<sub>2</sub>); Tier 3, direct carbon observations with an empirical (data-driven) model used to obtain global estimates, e.g. the use of geo-statistics to up-  
 494 scale local data into global values; Tier 4, indirect carbon observations not simply related to global carbon flux quantities.

	80-84	85-89	90-94	95-99	00-04	05-09	10-14	Tier	Methods and References
CGR (Pg C yr <sup>-1</sup> )	2.94±0.22	3.81±0.22	2.35±0.18	3.89±0.18	3.93±0.16	4.12±0.16	4.80±0.16	1	The atmospheric CO <sub>2</sub> growth rate (CGR) is direct carbon observations from more than 60 monitoring stations from the NOAA/ESRL atmospheric network (14). Uncertainties are from “Table 1” in Ballantyne et al. (34), which are “the means of the decadal standard deviations”. These uncertainties represent our ability to detect a change.
O <sub>2</sub> /N <sub>2</sub> (per meg unit)			-6.10±0.22	-7.37±0.69	-8.89±0.30	-9.15±0.35	-10.64±0.37	2	Data are from Scripps O <sub>2</sub> program ( <a href="http://scrippso2.ucsd.edu/">http://scrippso2.ucsd.edu/</a> ), and the weighted average of data from stations of Alert, La Jolla and Cape Grim are taken as the global means as in Keeling et al. (15). The standard deviations across data from these stations are used as uncertainties. The Scripps O <sub>2</sub> program records changes in atmospheric oxygen levels by collecting air samples at monitoring stations over the world. The change of atmospheric O <sub>2</sub> is related to the change of net land flux (B+L) and fossil fuel emission (F) but it is insensitive the change of ocean flux (O). The relationship of $\Delta O_2 = -\alpha_F F + \alpha_B(B+L) + Z_{O_2}$ from Keeling et al. (15) is used in this study.  Here, only the growth rate of O <sub>2</sub> /N <sub>2</sub> was used to constrain the sum of F and B+L here, in order not to duplicate the use of CGR.
O (Pg C yr <sup>-1</sup> )	-1.79±1.07	-1.85±1.07	-1.88±1.07	-1.53±1.07	-1.59±1.07	-2.02±1.07	-2.28±1.07	3	Values from Landschützer et al. (17) are added to an outgassing of 0.45 Pg C yr <sup>-1</sup> and the corresponding uncertainties of outgassing (0.18 Pg C yr <sup>-1</sup> ) from Jacobson et al. (16). The estimates from Landschützer et al. (17) used a neural network approach primarily based on surface ocean pCO <sub>2</sub> . Note that Landschützer et al. (17) used SOCATv2 data that starts in 1968, and there is data in their analysis for the 1980-1984 period although not much.  Both Landschützer et al., (17) and Rödenbeck et al. (18) used pCO <sub>2</sub> data, but they conducted independent approaches to upscale flux and get very different regional results (see the Surface Ocean pCO <sub>2</sub> Mapping (SOCOM) intercomparison paper by Rödenbeck et al. (37)). In our inversion, we estimated the potential error correlations between these two O data sets from the series of annual fluxes of the two products by assuming that the correlation in annual fluxes within each 5-year period is an approximation of the 5-year mean flux error correlation. The temporal correlations between the uncertainties in different 5-year intervals for these two data sets were also estimated using the method from Ballantyne et al. (34). In addition, sensitivity tests using only one of these two O data sets were also performed.
O (Pg C yr <sup>-1</sup> )		-1.82±0.68	-2.03±0.68	-1.68±0.68	-1.76±0.68	-1.56±0.68	-2.39±0.68	3	Values from Rödenbeck et al. (18) are added to an outgassing of 0.45 Pg C yr <sup>-1</sup> and the corresponding uncertainties of outgassing (0.18 Pg C yr <sup>-1</sup> ) from Jacobson et al. (16). Rödenbeck et al. (18) estimated the ocean carbon fluxes using a diagnostic mixed-layer approach based on surface pCO <sub>2</sub> observation and mass conservation. The uncertainty of estimate of Rödenbeck et al. (18) is calculated from in Rödenbeck et al. (38) by assuming that range of the sensitivity tests represents its 1-σ range.
O (Pg C yr <sup>-1</sup> )			-2.00±0.40					4	See above for the possible relationship with O estimates from Landschützer et al., (17). Means and uncertainties are from McNeil et al. (19) to constrain O from 1990 to 1999. This decadal O estimate is based on CFCs data.  McNeil et al. (19) estimated O using CFC-12 data only. Khatiwala et al. (20) used CFC-12, CFC-11 and <sup>14</sup> C for surface-to-interior transit time, complemented by temperature, salinity, oxygen and phosphate to get information on end-member to determine the types of waters that are being mixed. The method of Khatiwala et al. (20) thus uses much more information although there is a little overlap. A sensitivity test without using data from McNeil et al. (19) was conducted, and the results are rather consistent.
O (Pg C yr <sup>-1</sup> )			-1.91±0.60	-2.05±0.60				4	Means and uncertainties are from Khatiwala et al. (20). These estimates were based on field measurements of oceanic tracers, such as chlorofluorocarbons (CFCs), natural <sup>14</sup> C, temperature and salinity from the GLODAP and World Ocean Atlas databases.
O (Pg C yr <sup>-1</sup> )			-1.80±0.20					4	See above for the possible relationship with O estimates from McNeil et al. (19). Means and uncertainties are from Steinkamp and Gruber (21) to constrain O from 1990 to 1999. This O estimate was

B (Pg C yr <sup>-1</sup> )																					
B+L (Pg C yr <sup>-1</sup> )																					
L (Pg C yr <sup>-1</sup> )	1.33±0.50	1.52±0.50	1.56±0.50	1.64±0.50	1.04±0.50	0.87±0.50	0.88±0.50														
L (Pg C yr <sup>-1</sup> )	1.76±0.49	1.89±0.52	1.82±0.50	1.54±0.43	1.27±0.40	1.13±0.28	1.36±0.22														

derived from an atmosphere-ocean inversion using atmospheric CO<sub>2</sub>, ocean interior dissolved inorganic carbon (DOC) and surface pCO<sub>2</sub> data, and inventory-based land fluxes. It should be noted they only used the mean CO<sub>2</sub> level in the atmosphere, not the CGR, and thus there is no double use of data.

Although this estimate of O included atmospheric CO<sub>2</sub>, but it only used the mean CO<sub>2</sub> level in the atmosphere rather than the variations of CGR used in this study. Thus, there is no correlation with CGR.

3 Means are from Pan et al. (22) and corrected for river outgassing fluxes (16). They estimated global forest carbon stock and fluxes using a bottom-up approach based on inventories and long-time field observations. Given that B here only accounts for the carbon sink in global established forests, the uncertainties of global residuals in “Table 3” from Pan et al. (22) are also included, which may represent the uncertainty of land carbon sinks from other ecosystems like grassland.

No correlation with other data except for Africa. RECCAP used B estimate in Africa from Pan et al. (22), and we used the sum of regional net ecosystem exchange (NEE) from RECCAP to constrain B+L. But the contribution of NEE in Africa to the global NEE is very small (see Table S3), and thus the correlation between B from Pan et al. (22) and B+L from RECCAP was not included here.

3 RECCAP (see Table S3).

Net ecosystem exchange (NEE) from RECCAP are data-driven products estimated based on regional carbon cycle using various data sources (see Table S3). The B+L used here refers to the global (sum of regional values) anthropogenic NEE (corrected for river outgassing fluxes of 0.45 ±0.18 Pg C yr<sup>-1</sup> from Jacobson et al.(16)) in Table S3.

Potentially small correlation with B from Pan et al. (22) was ignored (see above).

3 Means and uncertainties are from the bookkeeping model by Houghton et al. (24), which is also used in the global carbon budget from Le Quéré et al. (8). The bookkeeping model from Houghton et al. (24) used inventory data from FAO to track the area of land use change, and combined carbon density in different carbon pools (e.g. vegetation and soil) to estimate the land use change fluxes.

Although both Houghton et al. (24) and Hansis et al. (23) used bookkeeping methods, they used different land use change maps and included different processes. Houghton et al. (24) used FAO/FRA land use change data as input while Hansis et al. (23) used Land-Use Harmonization (LUH) data (39). Therefore, there is no error correlation between the two L estimates.

Means and uncertainties are from the BLUE model by Hansis et al. (23). BLUE model is a spatially explicit bookkeeping model and tracks histories of successive land use and land cover change and their interactions in each grid cell. It also accounts for the gross land use and land cover transitions rather than net ones. The uncertainties here refer to the difference with the alternative estimates using different carbon density.

See above for possible relationship with L from Houghton et al. (24)

495 **Table S3** The estimates of net ecosystem exchange (NEE) from Regional Carbon Cycle  
 496 Assessment and Processes (RECCAP) international research project. In brief, the observed  
 497 carbon fluxes in the nine regions were evaluated and summed to give the global fluxes. NEE is  
 498 calculated from the mass balance of the change in territorial carbon stocks and the lateral carbon  
 499 fluxes (soil carbon export to river, river outgassing to ocean, crop and wood export) in each  
 500 region. The global NEE corrected for river outgassing flux (16) is used to constrain B+L in this  
 501 study. Negative values indicate a flux from the atmosphere to the land.

region	flux	period	Mean (Tg C yr <sup>-1</sup> )	Uncertainty (Tg C yr <sup>-1</sup> )	References in the mass balance calculation
North America	NEE	2000-2009	-418	288	King et al. (40); Peters et al. (41)
Europe	NEE	2000-2007	-234	166	Luyssaert et al. (26); Peters et al. (41); Hartmann et al. (42)
Russia	NEE	2007-2009	-600	224	Dolman et al. (43); Hartmann et al. (42)
South Asia	NEE	2000-2009	-86	34	Liu et al. (28); Peters et al. (41); Hartmann et al. (42)
East Asia	NEE	2000-2009	-256	35	Piao et al. (25); Lauerwald et al. (44); Peters et al. (41); Hartmann et al. (42)
Southeast Asia	NEE	undefined	-97	186	Liu et al. (28); Peters et al. (41); Hartmann et al. (42)
South America	NEE	2000-2009	11	284	Gloor et al. (45); Peters et al. (41); Hartmann et al. (42)
Africa	NEE	2000-2007	12	292	Pan et al. (22); Valentini et al. (46); Peters et al. (41); Hartmann et al. (42)
Australia	NEE	1990-2011	-71	36	Haverd et al. (47); Hartmann et al. (42)
Globe	NEE	2000-2009	-1740	604	sum of regional NEE
	B+L	2000-2009	-1290	611	NEE corrected for river outgassing to ocean (16)

502

503

504