1 *Classification:*

2 **Biological Science – Environmental Sciences**

3

4 *Title:*

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

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

10

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

12

13 [a]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 [b]Department of Ecosystem and Conservation Science, University of Montana, Missoula, MT,
16 USA

17 [c]Global Carbon Project, CSIRO Oceans and Atmosphere, GPO Box 3023, Canberra, ACT 2601,
18 Australia

19 [d]College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter EX4
20 4QF, UK

21 [e]Tyndall Centre for Climate Change Research, University of East Anglia, Norwich Research Park,
22 Norwich NR4 7TJ, UK

23 [f]Department of Earth and Environment, Boston University, Boston, MA 02215, USA

24 [g]Center for International Climate and Environmental Research – Oslo (CICERO), Oslo, Norway

25 [h]Department of Ecology, College of Urban and Environmental Sciences, Peking University,
26 Beijing 100871, China

27 [i]Max Planck Institute for Meteorology, Hamburg, Germany

28

29 [1]Corresponding author phone: 0033-6-58276925; email: wei.li@lsce.ipsl.fr

30

31 *Keywords:*

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

Abstract

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 to the conventional approach, Bayesian optimization decreases the uncertainties in B by 41% and in O by 47%. The L uncertainty decreases by 46% whereas F uncertainty marginally improves through the knowledge of all the other terms. Both ocean and net land uptake (B+L) rates have positive trends of 28±8 and 43±17 Tg C yr^{-2} since 1980, respectively. We explore the possibility of separating the net land flux into gross primary production (GPP), terrestrial ecosystem respiration (TER) and fire emissions by using proxies for GPP and TER trends, namely satellite vegetation greenness and the time-series from a global database of field-scale soil respiration. Between 1980 and 1994, GPP grew faster than TER, causing an increasing net land sink. After 2000, both GPP and TER growth stalled, leading to the maintenance of the land sink. This suggests that a different regime of GPP and TER trends explains the land sink after 2000. Our Bayesian fusion of multiple observations reduces uncertainties thereby allowing us to isolate important variability in global carbon cycle processes.

52   **Significance Statement**

53   The conventional approach of calculating the global carbon budget makes the land sink the most

54   uncertain of all budget terms. This is because rather than being constrained by observations it is

55   inferred as a residual in the budget equation. Here, we overcome this limitation by performing a

56   Bayesian fusion of different available observation-based estimates of decadal carbon fluxes. This

57   approach reduces the uncertainty in the land sink by 41% and in the ocean sink by 47%. These

58   results are significant because they give unprecedented confidence in the role of the increasing

59   land sink in regulating atmospheric CO_2, and shed light on the past decadal trend where the land

60   sink continued to grow despite plant carbon uptake having stalled.

61

62

63   The land and ocean carbon sinks provide a vital ecosystem service by absorbing on average about

64   55% of anthropogenic CO_2 emissions from fossil fuel combustion and land-use change. Research

65   has focused on understanding the relationships between year-to-year variability in carbon sinks

66   and climate (1, 2), as well as the long term trend over the full instrumental period of CO_2

67   monitoring at the Mauna Loa station (3). Quasi-decadal variations of emissions and sinks have

68   received comparatively less attention. Yet, significant climate variation occurs at this specific

69   time scale (4). Since 1980, the variable occurrence of different ENSO events, two large volcanic

70   eruptions (El Chichón and Pinatubo) and the recent slow-down of land surface warming (hiatus)

71   have modulated the strength of natural carbon sinks. There are also decadal-scale changes in the

72   rate at which human activities perturb the natural carbon cycle, in particular the recent

73   acceleration of fossil fuel and cement emissions in the 2000s (5) and the slow-down in global

74   land-use change emissions (LUC) in the mid-2000s, which appears to be partly driven by reduced

75   deforestation in Brazil (6).

76   Here, we provide a data-driven assessment of global CO_2 emissions and sinks at 5-year intervals

77   for the period of 1980-2014. We use a new Bayesian fusion approach whereby different data-

78   streams of ocean and land uptake, LUC emissions, are optimally combined, and their uncertainty

79   reduced from prior knowledge. This approach estimates the land sink constrained by data, which

80   is a major improvement over the "conventional" method for calculating the global carbon budget

81   by Ciais et al. (7) and Le Quéré et al. (8), hereafter LQ15, where the unknown land sink was

82   determined as a residual from the other components (emissions, atmospheric increase, ocean

83   uptake). Most of the data-streams used in this analysis start in 1980, and about half of them give

84   decadal mean values of natural sinks and thus do not allow us to tackle the reconstruction of

85   interannual variability. Our choice of applying a Bayesian fusion approach to optimize 5-year

86   average component fluxes of the global carbon budget is therefore a compromise that maximizes

87   the use of available observations of decadal average fluxes.

88   The principle of the Bayesian fusion approach is to combine an *a priori* imperfect knowledge of

89   fluxes with observations and their uncertainties to infer optimized estimates of fluxes. Here, we

90   define *a priori* values of terms in the global carbon budget that are not from observations.

91   Specifically, we set prior fossil fuel and cement emissions (F) from inventories and the simulated

92   land, ocean and land-use change carbon fluxes from process-based models (Table S1).

93   Observational data sets independent from those prior values are applied to constrain land-use

94   change emissions (L), the ocean uptake of anthropogenic CO_2 (O), the land-biosphere sink (B) in

95   ecosystems not affected by land-use change, and the net land flux (B+L) (Table S2). A more

96   uncertain attempt to extend the Bayesian optimization approach is proposed, where the net land

97   flux is decomposed into three gross carbon fluxes: Gross Primary Production (GPP), Terrestrial

98   Ecosystem Respiration (TER) and fire emissions (FR). Here the trends of GPP and TER are only

99   indirectly constrained by the proxies of satellite-based vegetation greenness trends for GPP and of

100 in-situ soil respiration data for TER.

Results

102 In the optimization of the global carbon budget (Opt-A, Fig. 1), the prior value of F and its

103 uncertainty (Table S1) were defined from the mean value and the range of different emission

104 inventories, namely from CDIAC (9), IEA (10), EDGAR (11) and BP (12) (Methods). These

105 inventories are not treated as direct observations of emissions, and there is currently no

106 independent observation to verify F. The prior values of O are from seven ocean biogeochemistry

107 models (8), and the prior values of B are from the nine TRENDY land carbon models (13)

108 (Methods). These prior values from state-of-the-art models are without direct observational

109 constraints. The prior estimates of L are derived from the difference of simulated land carbon

110 fluxes with and without LUC in the TRENDY carbon models (13). All fluxes are defined as

111 positive if CO_2 is lost to the atmosphere by the land or the ocean reservoir. Uncertainties in the

112 prior estimates of 5-yearly O, B and L, are set to the maximum between those reported by LQ15

113 and the standard deviations across models. All uncertainties here refer to 1-σ Gaussian errors. In

114 this context, the prior uncertainties are 0.5 for O, 0.9 for B and 0.8 Pg C yr^{-1} for L, thus not

115 smaller than the values of 0.5, 0.8 and 0.5 Pg C yr^{-1} from LQ15. It is important that the prior

116 uncertainties are not too small, so that adding observations can adjust and constrain the sought

117 fluxes.

118 Several independent data-streams, each with their specific uncertainty and temporal resolution

119 (Fig. S1, Table S2), are combined in the Bayesian optimization with the above prior knowledge.

120 These data-streams are: 1) the atmospheric CO_2 growth rate (CGR) from the NOAA/ESRL

121 atmospheric network (14) which constrains the sum of all fluxes and is determined very

122 accurately from more than 60 monitoring stations; 2) the atmospheric 5-year mean (negative)

123 growth rate of O_2/N_2 in the atmosphere from the Scripps O_2 Program (15) which relates to the

124 combined effect of B+L and F changes, while being insensitive to changes in O (note that O_2/N_2

125 has a negative trend in the atmosphere); 3) a set of yearly mean observation-based estimates of O

126 from shipboard partial pressure of CO_2 surveys corrected for natural outgassing (16, 17) and

127 decadal-mean observations of O from inventories of carbon change in the ocean (18–20); 4) ten-

128 year mean estimates of B from a global synthesis of changes in forest carbon stocks (21) and

129 inventory-based land carbon storage change from the RECCAP publications (Table S3); 5)

130 decadal mean B+L based on microwave remote sensing of Vegetation Optical Depth (22); 6)

131 five-year mean LUC emissions from two independent bookkeeping approaches constrained by

132 observed carbon stocks (23, 24) and from the RECCAP publications (Table S3). The

133 uncertainties in each data-stream are either derived directly from the original publications (when

134 reported) or estimated from expert judgments (details in Table S2). The optimization is performed

135 for seven consecutive 5-year windows between 1980 and 2014.

136 In the Bayesian optimization, observations that describe mean fluxes during intervals longer than

137 5-years are still useful to infer 5-yearly fluxes. For example, the mean ocean sink observation for

138   the 1990s (19) constrains the mean 5-yearly O during 1990-1999, while other independent

139   observations (O_2/N_2 and CGR) help to further separate O values between 1990-1994 and 1995-

140   1999. Despite no direct observation of F, this flux is found to be slightly improved in the

141   Bayesian fusion, through knowledge of the other terms, and because the sum of all fluxes is very

142   well constrained from CGR observations (Methods). We are aware that some observation-based

143   land sink estimates have systematic errors in the way they are included in the optimization. In

144   particular, the estimate of B from ref. (21) is only for forests and ignores other biomes. However,

145   the RECCAP studies (25–27) and other estimates (21, 22) of the carbon flux in non-forest biomes

146   suggest that the forest sink alone accounts for most of the global land sink B.

147   The improved global budget of anthropogenic CO_2 in Opt-A is shown in Fig. 2, and all data are

148   given in Table S1. After optimization, the *a posteriori* uncertainty in each flux is reduced.

149   Compared to the conventional method applied by LQ15 and IPCC-AR5 (7), uncertainties in B

150   and O are reduced by 41% and 47% in this study. In the Bayesian data fusion, the land sink is no

151   longer solely inferred as a residual that accumulates uncertainties from all other terms, and it

152   exhibits a large reduction in uncertainty. The uncertainty in L decreases by 46% but the

153   uncertainty in F is marginally improved (by 2%) through the indirect constraints of other terms.

154   In the absence of direct constraint on F, this small reduction in the F uncertainties compared to

155   LQ15 and IPCC-AR5 (7) is also because we use multiple emission inventories (while LQ15 and

156   IPCC-AR5 (7) only used CDIAC (9)) and start at relatively higher prior uncertainties in F (Table

157   S1) than in LQ15. Despite their improved (smaller) uncertainties, the 5-year mean fluxes shown

158   in Table S1 (Opt-A) do not differ statistically in their mean values from LQ15. This indicates that

159   each flux of the Bayesian carbon budget is fully consistent with LQ15 even though we used an

160   array of data with different measurement methods and with uncertainties estimated in different

161   ways. Specifically, we obtain emissions from fossil fuel burning and cement production that are

162   smaller than LQ15 by 0.24±0.16 Pg C yr^{-1} during 1980-1999 and 2005-2014 (Fig. 2), and higher

163   by 0.14 Pg C yr^{-1} during 2000-2004. A downward revision of global F during 1980-1999 is

164   consistent with the correction of the emissions for China based on evidence of the lower carbon

165   content for coal burned in that country (28). Compared to LQ15, the optimized ocean sink during

166   2000-2004 is larger by 0.16 Pg C yr^{-1} but lower by 0.20±0.12 Pg C yr^{-1} during all the other

167   periods. In the past decade (2005-2014), both ocean sink and land sink from our optimization are

168   smaller than LQ15. The optimized fluxes of L are similar to or lower than those from LQ15. The

169   trend of F for the seven 5-year periods is positive ($p = 0.007$), with a probability of a positive

170   trend for O and B+L of 93%; the trend of B or L individually is not significant for ($p = 0.23$ and

171   0.13). The increasing rate of O and B+L are 28±8 and 43±17 Tg C yr^{-2} since 1980, respectively.

172   Similar significantly positive trends were also found in the 5-year mean O and B+L between 1980

173   and 2014 calculated from the yearly budget updated by LQ15 (Fig. S2). Given the robustness of

174   O inferred by Opt-A (see also Fig. S3), and in view of the many observations constraining this

175   flux, there is a high confidence that the ocean sink has been increasing over time since 1980. The

176   land sink from Opt-A is less variable between different 5-year periods than in LQ15 (Fig. S2).

177   But the ocean sink is more variable, with a standard deviation of 0.36 Pg C yr^{-1} compared to 0.29

178   Pg C yr^{-1} by LQ15 (standard deviations across the seven periods analyzed).

179   From 1980 to 2014, the average fractions of F+L emission re-absorbed by the land and ocean

180   carbon reservoirs are 29 ±6% and 26 ±2%, respectively. The ratios of both O and B to F+L

181   emission do not exhibit any significant trends ($p > 0.05$, Fig. S4). Even with their reduced

182   uncertainty in this study compared to LQ15, the variability of O and B between 5-year intervals

183   prevents us from assessing the very small trends in their ratios to emissions. Similarly, we found

184   no significant trend in the ratio of O or B to fossil fuel emission (F). The larger variability of the

185   B-to-(F+L) ratios compared to the O-to-(F+L) ones (Fig. S4) suggests that the efficiency of the

186   land sink at absorbing emissions is more variable than that of the ocean sink. For instance, during

187   the period that followed the cooling from the Pinatubo eruption in 1990-1994 (29–31), the B-to-

188   (F+L) ratio increased by 41% above its long-term mean. This ratio was also higher than normal

189   during 2005-2009, possibly due to the absence of El Niño and to the occurrence of a cooler and

190   wetter La Niña event in 2008-2009 manifested by lower than normal CGR (3).

191   A second more exploratory optimization called Opt-B was attempted to further decompose

192   changes in the net land carbon flux (B+L) into changes of gross fluxes of GPP, TER and FR. This

193   second optimization (Fig. S5) uses observations to constrain the equation B + L = GPP + TER +

194   FR, and fire emissions, and proxies to constrain changes of GPP and TER on 5-year successive

195   intervals. The prior values of GPP, TER and FR are from the TRENDY land carbon models as in

196   Opt-A (Methods, Table S1) (13). The data constraints on GPP and TER included in Opt-B (Table

197   S2) are less stringent than those used for net fluxes in Opt-A because there are only indirect

198   proxies of these fluxes. Namely, we used: 1) the decadal mean value of GPP during 2000-2009

199   derived from MODIS satellite observations (32) and from a data-driven product using both

200   satellite and flux tower measurements (33); 2) a proxy of the GPP trend, given by the trend of

201   satellite global leaf area index (LAI) since 1982 (yearly average of three LAI data sets (34–36));

202   3) an indirect proxy of the global TER trend from the global trend observed in long-term field-

203   scale soil respiration measurements (37). For fire emissions, we used 5-year mean values from

204   three inventory- or satellite-based products (38–40). Only the decadal mean GPP from ref. (33)

205   during 2000-2009 is used and not the trend of this GPP data-product, because this trend has been

206   reported to be possibly underestimated (41).

207   The results of Opt-B are consistent with Opt-A for the net fluxes, which is not surprising because

208   both optimizations share the same constraints for the net fluxes (Fig. S6). The optimized 5-year

209   mean GPP and TER estimates are lower by about 10 Pg C yr^{-1} than their prior values (Fig. 3a),

210   giving a best value of GPP = -127 ±8 Pg C yr^{-1} and TER = 124 ±8 Pg C yr^{-1}. The optimized fire

211   emission (FR) is 2.3 Pg C yr^{-1} during 1980-2014, with a variability of 0.29 Pg yr^{-1} across the

212   different 5-year periods (Fig. 3b). Compared with the prior value, FR decreases after optimization

213   in all the seven 5-year periods, and its uncertainties are reduced by 54% from the prior (Fig. 3b).

214   The reductions in both mean values and uncertainties of FR are mainly a result of the constraints

215   of three inventory- and satellite-based fire emission data sets (38–40) (Tables S1 and S2).

216   In Opt-B, 5-year GPP and TER are found to increase from 1980 to 2000 (Fig. 3a). After 2000, a

217   different regime is found: the growth in GPP stalled, but TER did not increase either, so that a net

218   land sink was maintained (Fig. 3a, Table S1). The proxy data for GPP trends and TER trends used

219   as observational constraints are sufficient in Opt-B to change the sign of the prior trends after

220   2000. The break point change in the trend of GPP and TER after 2000 is found to be more robust

221   than the mean values of these fluxes, across a series of sensitivity tests (Methods, Figure S7). The

222   GPP trend changed from 1.4 ±0.56 before 2000 to -0.026 ±0.10 Pg C yr^{-2} after 2000, and

223   similarly, the TER trend decreased from 1.3 ±0.54 to -0.31 ±0.17 Pg C yr^{-2}. This reduction in the

224   positive trend of GPP and TER after 2000 was not present in the prior from TRENDY models,

225   where GPP and TER persistently increase during the last 35 years (13) (Fig. 3a). In Opt-B, the

226   inferred stalling of growth in GPP is constrained by the stalling of global satellite LAI trend

227   (consistent across three satellite LAI products (34–36)). Yet, it should not be over-interpreted

228   since a stalled LAI is no proof of a stalled GPP because the light-use efficiency of plant canopies

229   might have continued to increase while their fraction of intercepted solar radiation (directly

230   related to LAI) may have stalled.

Discussion

232   The Bayesian approach used in this study provides the most robust estimate to date of the strength

233   and evolution of the land sink and provides hints about how gross fluxes may have changed to

234   cause this sink. The trend of 5-year mean GPP and TER reflect both decadal climate variability

235   and long-term drivers such as the increasing concentration of CO_2, nitrogen deposition and land

236   management. During 1990-1994, the five-year period within which the Pinatubo eruption

237 occurred, we infer from Opt-B that GPP was abnormally high relative to TER (Fig. 3a, Table S1).

238 This result supports the conclusions of previous studies that a positive effect on GPP occurred

239 due to the increased fraction of diffuse light produced by volcanic aerosols (42) and TER was

240 suppressed due to the cooling that followed the eruption (29). The stalled TER since 2000-2004

241 may reflect the effect of a slow-down soil carbon decomposition in response to the winter cooling

242 of northern lands (43). It is not possible with our global approach to infer which region was

243 responsible for the stalling of GPP and TER after 2000. Possible explanations for the recent

244 stalling of global GPP could involve either emerging biogeochemical limitations (e.g., reduced

245 nitrogen deposition in some regions, or emerging nutrient limitations elsewhere (44, 45)) or the

246 slower warming at high latitudes resulting in a slower photosynthetic enhancement and

247 lengthening of the growing season. Combining spatially explicit observations to constrain

248 process-based carbon cycle models, as done e.g., in carbon cycle data assimilation systems (46,

249 47) should provide more insights about the regional processes and explain why GPP and TER

250 have stalled in the 2000s. If the driver of the GPP stalling proves not to be the slower warming of

251 the northern hemisphere after 2000, and should high warming rates resume in the coming decade,

252 then TER would increase while GPP remains stalled. This combination is expected to strongly

253 reduce the sink strength of land ecosystems, and would require stronger mitigation of

254 anthropogenic emissions to keep to the climate stabilization pathway of 2°C recently agreed by

255 187 nations (48).

256

269   **Additional information**

270   Supplementary information is available online.

271

272   **Methods**

273   **Bayesian Estimation System.** Each estimate of the 5-year mean carbon fluxes (and trends for

274   Opt-B), called hereafter the "control variables" **x**, is based on the update from a prior estimate of

275   these variables $\mathbf{x^b}$, using some observation-based estimates $\mathbf{y^o}$ of the fluxes (and trends for Opt-B)

276   that are connected to the control variables through the relationships H: $\mathbf{x} \rightarrow \mathbf{y} = H[\mathbf{x}]$. We follow a

277   Bayesian statistical approach for this estimation. Assuming that the uncertainties in $\mathbf{x^b}$ and $\mathbf{y^o}$ are

278   unbiased and Gaussian, characterized by the prior and observation uncertainty covariance

279   matrices **B** and **R** respectively, and that H is linear (denoted as a matrix **H**), the statistical estimate

280   of **x**, given $\mathbf{x^b}$ and $\mathbf{y^o}$, is unbiased and Gaussian, and the corresponding optimal estimate $\mathbf{x^a}$ and

281   uncertainty covariance matrix **A** are given (49) as:

282    $\mathbf{A} = \left(\mathbf{B^{-1}} + \mathbf{H^T R^{-1} H}\right)^{-1}$ (1)

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

284   where the superscripts T and "-1" are the transpose of a matrix and the inverse of a matrix,

285   respectively.

286   In the optimization, Opt-A, we update the estimate of the mean fluxes of fossil fuel and cement

287   emissions (F), ocean sink (O), land sink (B) and land-use change emissions (L) for each 5-year

288   interval from 1980 to 2014 (Fig. 1 and Table S1). The observation vector contains estimates of 5-

289   year mean global atmospheric growth rates of CO_2 (CGR, in Pg C yr^{-1}), atmospheric growth rates

290   of O_2/N_2 (CGR-O_2, per meg unit), observation-based estimates of ocean sinks, land sinks, land-

291   use change emissions and net land sink (B+L) (the data sources for these components of $\mathbf{y^o}$ are

292   summarized in Fig. 1 and provided in Table S2). The prior estimates for the different control

293   variables are built with independent data sets so that there are no correlations between the prior

294   uncertainties in the different control variables. The correlations between different 5-year intervals

295   for CGR and O, and potential correlations between two data-driven estimates of O are estimated

296   using the method in ref. (50). **H** in Opt-A is defined for each 5-year interval by:

297

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

298   where α_F, α_B and Z_{O2} are constant coefficients from ref. (15). The optimal estimate $\mathbf{x^a}$ discussed in

299   the main text is computed by Eqs. (1) and (2).

300   In the optimization, Opt-B, we solve for the mean fluxes of F, O, fire emissions (FR), and for the

301   linear relative trends of GPP (α) and TER (β) for each 5-year interval, as well as for the mean

302   value of GPP (G_0) and TER (T_0) in the year 2005 (Fig. S5 and Table S1). The observation vector

303   contains CGR, CGR-O_2, observation-based estimates of ocean sinks, net land sinks, fire

304   emissions and linear relative trends of GPP (α) and TER (β). The linear relative trend of GPP (α)

305   is not directly observable and we made the assumption that it is equal to the observable relative

306   LAI trends from satellites (34–36). The relative trends of LAI (34–36), soil respiration (37) and

307   G_0 (32, 33) are rather uncertain and sparse (Table S2). As for Opt-A, the prior covariance error

308   matrix in Opt-B is also diagonal, while the observation covariance error matrix accounts for the

309   potential correlations in the data-driven products (Table S1 and Table S2). The observation

310   operator H is defined by the following relationship:

311

$$H : \begin{bmatrix} F \\ O \\ G_0 \\ T_0 \\ \alpha \\ \beta \\ FR \end{bmatrix} \rightarrow \begin{bmatrix} CGR \\ CGR - O_2 \\ O \\ B+L \\ \alpha \\ GPP \\ \beta \\ FR \end{bmatrix} = \begin{bmatrix} F + O + GPP(G_0,\alpha) + TER\,(T_0,\beta) + FR \\ \alpha_F F + \alpha_B [GPP(G_0,\alpha) + TER(T_0,\beta) + FR] + Z_{O2} \\ O \\ GPP(G_0,\alpha) + TER(T_0,\beta) + FR \\ \alpha \\ GPP(G_0,\alpha) \\ \beta \\ FR \end{bmatrix}$$

(4)

312 where GPP(G_0, α) and TER (T_0, β) for a given 5-year interval are functions of G_0 and T_0 with

313 corresponding trends:

314
$$GPP_i(G_0,\alpha) = G_0\left(1 + 3\alpha_i + \sum_{j=0}^{i-1} 5 \times \alpha_j\right)$$
(5)

315
$$TER_i(T_0,\beta) = T_0\left(1 + 3\beta_i + \sum_{j=0}^{i-1} 5 \times \beta_j\right)$$
(6)

316 where the subscripts i and j are the number of 5-year intervals from the given 5-year to 2005. Eqs.

317 (5) and (6) and, thus, the observation operator are not linear. However, assuming unbiased and

318 Gaussian uncertainties in $\mathbf{x^b}$ and $\mathbf{y^o}$, the Bayesian inference indicates that the optimal estimate $\mathbf{x^a}$

319 of the control variables minimizes the following cost function (49):

320
$$J(\mathbf{x}) = \frac{1}{2}\left[\left(\mathbf{x}-\mathbf{x^b}\right)^T \mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x^b}\right) + \left(\mathbf{y^o} - H[\mathbf{x}]\right)^T \mathbf{R}^{-1}\left(\mathbf{y^o} - H[\mathbf{x}]\right)\right]$$
(7)

321 We minimize this cost function using a quasi-Newton iterative approach (49). At each iteration

322 $n+1$, the operator H is linearized against the best estimate of the control variables $\mathbf{x_n}$ given by the

323 previous iteration (or against $\mathbf{x^b}$ for the first iteration):

324
$$\mathbf{H_{x_n}} = \left(\frac{\partial H}{\partial \mathbf{x_n}}\right)_{\mathbf{x_n}}$$
(8)

325 $$H[\mathbf{x}] \approx H[\mathbf{x}_n] + \mathbf{H}_{\mathbf{x}_n}(\mathbf{x} - \mathbf{x}_n)$$ (9)

326 J is thus approximated by a quadratic function, whose minimum $\mathbf{x}_{n+1}$ is given by a revised version

327 of Eqs. (1) and (2):

328 $$\mathbf{A}_{n+1} = \left(\mathbf{B}^{-1} + \mathbf{H}_{\mathbf{x}_n}^{T}\mathbf{R}^{-1}\mathbf{H}_{\mathbf{x}_n}\right)^{-1}$$ (10)

329 $$\mathbf{x}_{n+1} = \mathbf{x}^{b} + \mathbf{A}_{n+1}\mathbf{H}_{\mathbf{x}_n}^{T}\mathbf{R}^{-1}\left(\mathbf{y}^{0} - H[\mathbf{x}_n] - \mathbf{H}_{\mathbf{x}_n}\left(\mathbf{x}^{b} - \mathbf{x}_n\right)\right)$$ (11)

330 We continue to derive our approximation of $\mathbf{x}^{a}$ until the series of $\mathbf{x}_n$ converges (relative difference

331 between $\mathbf{x}_{n+1}$ and $\mathbf{x}_n$ within 1×10^{-5}). We approximate the uncertainty in this final estimate of the

332 control variable using $\mathbf{A}_{n+1}$ (49).

333 **Prior Data.** To define the prior F, individual country data from four emission inventories

334 (CDIAC (9), IEA (10), EDGAR (11) and BP (12)) are grouped into geographic regions as

335 specified by the United Nations Statistics Division

336 (http://unstats.un.org/unsd/methods/m49/m49regin.htm). Cement emissions from EDGAR are

337 added into the IEA and BP data sets that do not include cement emissions. Uncertainties for each

338 country (51) are used to create regional uncertainty distributions for each emission index using a

339 bootstrapping method, with the uncertainties of the highest emitters in each region contributing

340 the most to the uncertainty distributions. This effect is achieved by weighting the sampling

341 probability (P_s) by the relative contribution of each country's emissions (E_C) to the total

342 emissions within the region (E_R):

343 $P_s = E_C / E_R$ (12)

344 To constrain the temporal component of the emission errors, ten random samples are drawn from

345 the corresponding regional uncertainty distribution for each country, producing ten random

346 uncertainties for each country. These country-level uncertainties are used to constrain a random

347   error time series covering 1980-2014, which is then run through an algorithm incorporating

348   autocorrelated random noise, such that:

349    $\varepsilon_{F(t)} = 0.95 \times \varepsilon_{F(t-1)} + \varepsilon_{(t)}$ (13)

350   where emission error factors for any given year $\varepsilon_{F(t)}$ are correlated with the emission errors from

351   the previous year $\varepsilon_{F(t-1)}$ by an autoregressive coefficient of 0.95 with $\varepsilon_{(t)}$ as random error.   The

352   autocorrelated time series are then multiplied and added to the fossil fuel emissions for each

353   country, and subsequently 500 samples of global fossil fuel emissions are taken for each 5-year

354   bin. The means and standard deviations of each bin for each inventory are calculated from these

355   500 samples. Additionally, the correlation in global uncertainty is calculated between 5-year bins

356   and inventories to produce an error-covariance matrix. The maximum between the uncertainties

357   calculated above and the standard deviations of the 5-year means across four emission inventories

358   were adopted as the uncertainties of prior F.

359   Prior O values are set from the ocean biogeochemistry model values used in LQ15. Note that

360   LQ15 adjusted their simulated O so as to match ocean observations during the decade of the

361   1990s and then used these bias-corrected ocean models outside this period. Here, for setting the

362   prior O values, we consider simply the spread and the mean of ocean models without any

363   adjustment.

364   Prior values of L and B are set from simulations in the TRENDY (v2) model intercomparison

365   project (13). The simulations in TRENDY (v2) are up to 2012, and thus the priors for the period

366   of 2010-2012 were used for 2013 to 2014. All the prior flux values are summarized in Table S1.

367   **Correlations between Optimized Fluxes.** Due to the very small uncertainties in CGR, there are

368   relatively strong correlations between the optimized flux components related to CGR in Eq. (3) in

369   the same period (Fig. S8a). To be specific, the fluxes B and O are negatively correlated, as are the

370   fluxes F and L. Positive correlations can be seen between F and B, and between L and B. The

371 correlations between fluxes in different periods can be attributed to the autocorrelations in

372 observed CGR and O and the decadal mean observations used to constrain two consecutive 5-

373 year periods. Similarly, the components related to CGR in Eq. (4) for Opt-B could also be

374 correlated, because they are all constrained by CGR (Fig. S8b).

375 **Sensitivity Tests.** Sensitivity tests were conducted by doubling the uncertainty in prior F, testing

376 prior F and its uncertainties from individual inventory data sets (IEA, EDGAR, CDIAC and BP),

377 and using prior B and O from CMIP5 models (52) instead of TRENDY (v2) models and ocean

378 carbon models from LQ15 in Opt-A (Fig. S3). The ocean sink is very robustly constrained in the

379 optimization because of the sufficiency of constraining observational data, but the land sink is

380 dependent on the prior F choice. By doubling the prior F uncertainties or replacing prior B and O

381 with data from CMIP5, the means of all fluxes remain stable, although posterior uncertainties

382 change slightly, emphasizing the robustness of this optimizing system. However, caution should

383 be applied when deciding the means of prior F due to their relatively greater impacts on B and

384 B+L.

385 Sensitivity tests for Opt-B were also performed, including: (a) using yearly GPP estimates during

386 1982-2011 from ref. (33) and removing LAI constraints as a proxy of the GPP relative trend (34–

387 36); (b) only using the average GPP from MODIS (32) to constrain G_0; (c) only using the average

388 GPP from ref. (33) to constrain G_0; (d) additionally using the 5-year mean GPP during 2000-

389 2004, 2005-2009 and 2010-2014 from MODIS (32) to constrain the corresponding 5-year mean

390 GPP; (e) setting prior relative trend of GPP (α) to 0; (f) setting prior trend of GPP (α) to 0 and

391 removing LAI constraints to the relative trend of GPP (34–36); (g) removing the constraint from

392 the trend of soil respiration (37); (h) setting prior trend of TER (β) to zero and removing the

393 constraint from the trend of soil respiration (37); and (i) setting both prior trends of GPP (α) and

394 TER (β) to 0 and removing LAI constraints (34–36). The optimized fluxes of F, O, B+L and FR

395 are very stable in these sensitivity tests. The optimized mean values of GPP and TER vary (Fig.

396   S7) due to the large discrepancy between the GPP from MODIS and from ref. (33). However, the

397   trends of GPP and TER are generally similar, but with slightly different scales (Fig. S7). This

398   indicates that the patterns of GPP and TER are very robust and mutually controlled by the prior

399   values, the LAI trend (34–36) and the soil respiration trend (37).

400   **Trend Test.** A Mann-Kendall statistical test was applied as a trend test. To remove

401   autocorrelations, pre-whitening (53) was performed using the correlations between the posterior

402   uncertainties (Fig. S8).

References

1. Wang X, et al. (2014) A two-fold increase of carbon cycle sensitivity to tropical temperature variations. *Nature* 506(7487):212–5.

2. Poulter B, et al. (2014) Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle. *Nature* 509(7502):600–3.

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 CO2 emissions. *Proc Natl Acad Sci U S A* 104(24):10288–93.

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:e0126754.

7. Ciais P, et al. (2013) 2013: Carbon and Other Biogeochemical Cycles. *Clim Chang 2013 Phys Sci Basis Contrib Work Gr I to Fifth Assess Rep Intergov Panel Clim Chang*:465–570.

8. Le Quéré C, et al. (2015) Global carbon budget 2014. *Earth Syst Sci Data*:47–85.

9. Boden TA, Marland G, Andres RJ (2013) Global, regional, and national fossil-fuel CO□ emissions. *Carbon Dioxide Inf Anal Center, Oak Ridge Natl Lab USA Oak Ridge, TN Dep Energy*. doi:10.3334/CDIAC/00001.

10. International Energy Agency (2013) CO2 Emissions From Fuel Combustion Highlights. *IEA Stat*:158.

11. Olivier JGJ, Janssens-Maenhout G, Peters JAHW (2014) *Trends in global CO2 emissions: 2014 Report* (PBL Netherlands Environmental Assessment Agency Hague).

12. BP (2015) BP Statistical Review of World Energy 2015.

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: http://www.esrl.noaa.gov/gmd/ccgg/about/global_means.html [Accessed August 9, 2015].

15. Keeling RF, Manning a. C (2014) *Studies of Recent Changes in Atmospheric O2 Content* (Elsevier Ltd.). 2nd Ed. doi:10.1016/B978-0-08-095975-7.00420-4.

16. Landschützer P, Gruber N, Bakker DCE, Schuster U (2014) Recent variability of the global ocean carbon sink. *Global Biogeochem Cycles* 28(9):927–949.

17. 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.

18. Khatiwala S, Primeau F, Hall T (2009) Reconstruction of the history of anthropogenic CO(2) concentrations in the ocean. *Nature* 462(7271):346–9.

19. McNeil BI, Matear RJ, Key RM, Bullister JL, Sarmiento JL (2003) Anthropogenic CO2 uptake by the ocean based on the global chlorofluorocarbon data set. *Science* 299(5604):235–9.

20. 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.

21. Pan Y, et al. (2011) A large and persistent carbon sink in the world's forests. *Science* 333(6045):988–993.

445   22.   Liu YY, et al. (2015) Recent reversal in loss of global terrestrial biomass. *Nat Clim Chang*
446   5(5):470–474.

447   23.   Hansis E, Davis SJ, Pongratz J (2015) Relevance of methodological choices for accounting of land
448   use change carbon fluxes. *Global Biogeochem Cycles* 29(8):1230–1246.

449   24.   Houghton R a., et al. (2012) Carbon emissions from land use and land-cover change.
450   *Biogeosciences* 9(12):5125–5142.

451   25.   Piao SL, et al. (2012) The carbon budget of terrestrial ecosystems in East Asia over the last two
452   decades. *Biogeosciences* 9(9):3571–3586.

453   26.   Luyssaert S, et al. (2012) The European land and inland water CO_2, CO, CH_4 and N_2O balance
454   between 2001 and 2005. *Biogeosciences* 9(8):3357–3380.

455   27.   King AW, et al. (2015) North America's net terrestrial CO_2 exchange with the atmosphere 1990–
456   2009. *Biogeosciences* 12(2):399–414.

457   28.   Liu Z, et al. (2015) Reduced carbon emission estimates from fossil fuel combustion and cement
458   production in China. *Nature* 524(7565):335–338.

459   29.   Lucht W, et al. (2002) Climatic control of the high-latitude vegetation greening trend and Pinatubo
460   effect. *Science* 296(5573):1687–9.

461   30.   Mercado LM, et al. (2009) Impact of changes in diffuse radiation on the global land carbon sink.
462   *Nature* 458(7241):1014–7.

463   31.   Frölicher TL, Joos F, Raible CC, Sarmiento JL (2013) Atmospheric CO2 response to volcanic
464   eruptions: The role of ENSO, season, and variability. *Global Biogeochem Cycles* 27(1):239–251.

465   32.   Zhao M, Running SW, Nemani RR (2006) Sensitivity of Moderate Resolution Imaging
466   Spectroradiometer (MODIS) terrestrial primary production to the accuracy of meteorological
467   reanalyses. *J Geophys Res* 111(G1):G01002.

468   33.   Jung M, et al. (2011) Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and
469   sensible heat derived from eddy covariance, satellite, and meteorological observations. *J Geophys*
470   *Res* 116:G00J07.

471   34.   Zhu Z, et al. (2013) Global Data Sets of Vegetation Leaf Area Index (LAI)3g and Fraction of
472   Photosynthetically Active Radiation (FPAR)3g Derived from Global Inventory Modeling and
473   Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI3g) for the Period
474   1981 to 2. *Remote Sens* 5(2):927–948.

475   35.   Xiao Z, et al. (2014) Use of General Regression Neural Networks for Generating the GLASS Leaf
476   Area Index Product From Time-Series MODIS Surface Reflectance. *IEEE Trans Geosci Remote*
477   *Sens* 52(1):209–223.

478   36.   Liu Y, Liu R, Chen JM (2012) Retrospective retrieval of long-term consistent global leaf area index
479   (1981–2011) from combined AVHRR and MODIS data. *J Geophys Res* 117(G4):G04003.

480   37.   Bond-Lamberty B, Thomson A (2010) Temperature-associated increases in the global soil
481   respiration record. *Nature* 464(7288):579–82.

482   38.   Giglio L, Randerson JT, van der Werf GR (2013) Analysis of daily, monthly, and annual burned
483   area using the fourth-generation global fire emissions database (GFED4). *J Geophys Res*
484   *Biogeosciences* 118(1):317–328.

485   39.   Mieville A, et al. (2010) Emissions of gases and particles from biomass burning during the 20th
486   century using satellite data and an historical reconstruction. *Atmos Environ* 44(11):1469–1477.

487   40.   Schultz MG, et al. (2008) Global wildland fire emissions from 1960 to 2000. *Global Biogeochem*
488   *Cycles* 22(2). doi:10.1029/2007GB003031.

489   41.   Huang M, et al. (2015) Change in terrestrial ecosystem water-use efficiency over the last three
490   decades. *Glob Chang Biol* 21(6):2366–78.

491   42.   Gu L, et al. (2003) Response of a deciduous forest to the Mount Pinatubo eruption: enhanced
492   photosynthesis. *Science* 299(5615):2035–8.

493   43.   Cohen JL, Furtado JC, Barlow M, Alexeev VA, Cherry JE (2012) Asymmetric seasonal
494   temperature trends. *Geophys Res Lett* 39(4). doi:10.1029/2011GL050582.

495   44.   Tørseth K, et al. (2012) Introduction to the European Monitoring and Evaluation Programme
496   (EMEP) and observed atmospheric composition change during 1972–2009. *Atmos Chem Phys*
497   12(12):5447–5481.

498   45.   Davidson EA, et al. (2011) Excess nitrogen in the U.S. environment: Trends, risks, and solutions.
499   *Issues Ecol* (15). Available at: https://pubs.er.usgs.gov/publication/70032270 [Accessed September
500   18, 2015].

501   46.   Rayner PJ, et al. (2005) Two decades of terrestrial carbon fluxes from a carbon cycle data
502   assimilation system (CCDAS). *Global Biogeochem Cycles* 19(2):n/a–n/a.

503   47.   Peylin P, et al. (2016) A new step-wise Carbon Cycle Data Assimilation System using multiple
504   data streams to constrain the simulated land surface carbon cycle. *Geosci Model Dev Discuss*:1–52.

505   48.   United Nations (2015) Adoption of the Paris Agreement. *Conf Parties its twenty-first Sess*
506   21932(December):32.

507   49.   Tarantola A (2005) *Inverse Problem Theory and Methods for Model Parameter Estimation*
508   (SIAM).

509   50.   Ballantyne AP, et al. (2015) Audit of the global carbon budget: estimate errors and their impact on
510   uptake uncertainty. *Biogeosciences* 12(8):2565–2584.

511   51.   Andres RJ, Boden TA, Higdon D (2014) A new evaluation of the uncertainty associated with
512   CDIAC estimates of fossil fuel carbon dioxide emission. *Tellus B* 66.
513   doi:10.3402/tellusb.v66.23616.

514   52.   Anav A, et al. (2013) Evaluating the Land and Ocean Components of the Global Carbon Cycle in
515   the CMIP5 Earth System Models. *J Clim* 26(18):6801–6843.

516   53.   Yue S, Pilon P, Phinney B, Cavadias G (2002) The influence of autocorrelation on the ability to
517   detect trend in hydrological series. *Hydrol Process* 16(9):1807–1829.

518   54.   Jacobson AR, Fletcher SEM, Gruber N, Sarmiento JL, Gloor M (2007) A joint atmosphere-ocean
519   inversion for surface fluxes of carbon dioxide: 1. Methods and global-scale fluxes. *Global
520   Biogeochem Cycles* 21(1). doi:10.1029/2005GB002556.

521   55.   Van Der Werf GR, et al. (2010) Global fire emissions and the contribution of deforestation,
522   savanna, forest, agricultural, and peat fires (1997-2009). *Atmos Chem Phys* 10(23):11707–11735.

523   56.   King AW, et al. (2007) *The North American Carbon Budget and Implications for the Global
524   Carbon Cycle*.

525   57.   SCHULZE ED, et al. (2010) The European carbon balance. Part 4: integration of carbon and other
526   trace-gas fluxes. *Glob Chang Biol* 16(5):1451–1469.

527   58.   Dolman AJ, et al. (2012) An estimate of the terrestrial carbon budget of Russia using inventory-
528   based, eddy covariance and inversion methods. *Biogeosciences* 9(12):5323–5340.

529   59.   Patra PK, et al. (2013) The carbon budget of South Asia. *Biogeosciences* 10:513–527.

530   60.   Piao SL, et al. (2012) The carbon budget of terrestrial ecosystems in East Asia over the last two
531   decades. *Biogeosciences* 9(9):3571–3586.

532   61.   Gloor M, et al. (2012) The carbon balance of South America: a review of the status, decadal trends

533   and main determinants. *Biogeosciences* 9(12):5407–5430.

534   62.   Valentini R, et al. (2014) A full greenhouse gases budget of Africa: synthesis, uncertainties, and
535   vulnerabilities. *Biogeosciences* 11(2):381–407.

536   63.   Haverd V, et al. (2013) The Australian terrestrial carbon budget. *Biogeosciences* 10(2):851–869.

537

538 **Figure 1** The framework of Opt-A. The number of constraining data-streams and the specific

539 data sources are marked on the right. The fluxes that are optimized are 5-year averages of F, O, B

540 and L representing fossil fuel and cement emissions, ocean sink, land sink, and land-use change

541 emissions, respectively. The observations used to constrain these fluxes are the five-year

542 averaged growth rates of CO_2 and O_2/N_2 in the atmosphere, observations of O, B and the net land

543 sink (B+L) from carbon measurements made in these two reservoirs, and inventory-based

544 estimates of L. In this framework, the CO_2 growth rate constrains the sum of all the fluxes. The

545 O_2/N_2 growth rate allows us to separate O and B+L and bring some constraint on F as well.
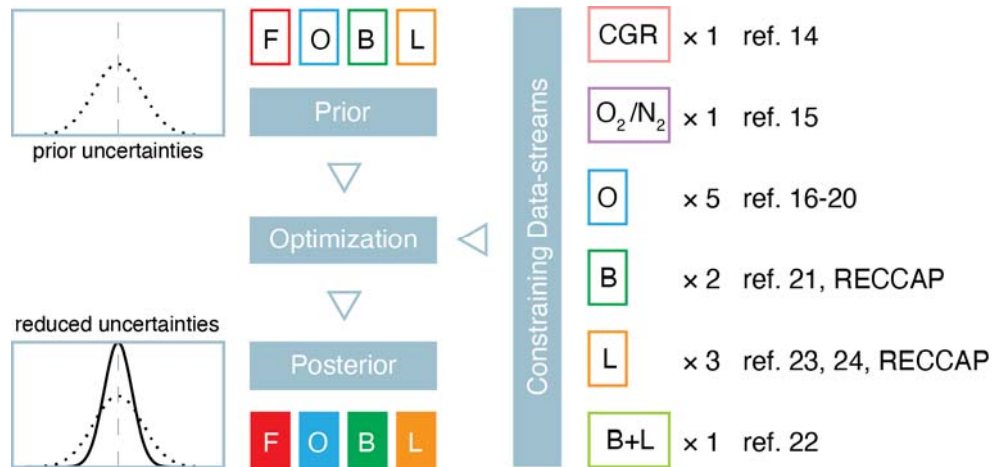

546
547
548
549
550
551
552
553

554 **Figure 2** (**a**) The fossil fuel and cement emissions (F), (**b**) ocean sink (O) and (**c**) net land flux

555 (B+L), and (**d**) land sink (B) and land-use change emissions (L) from prior knowledge, Opt-A

556 and LQ15. All the fluxes are 5-year means in each period. The error bars represent the 1-σ

557 uncertainties.
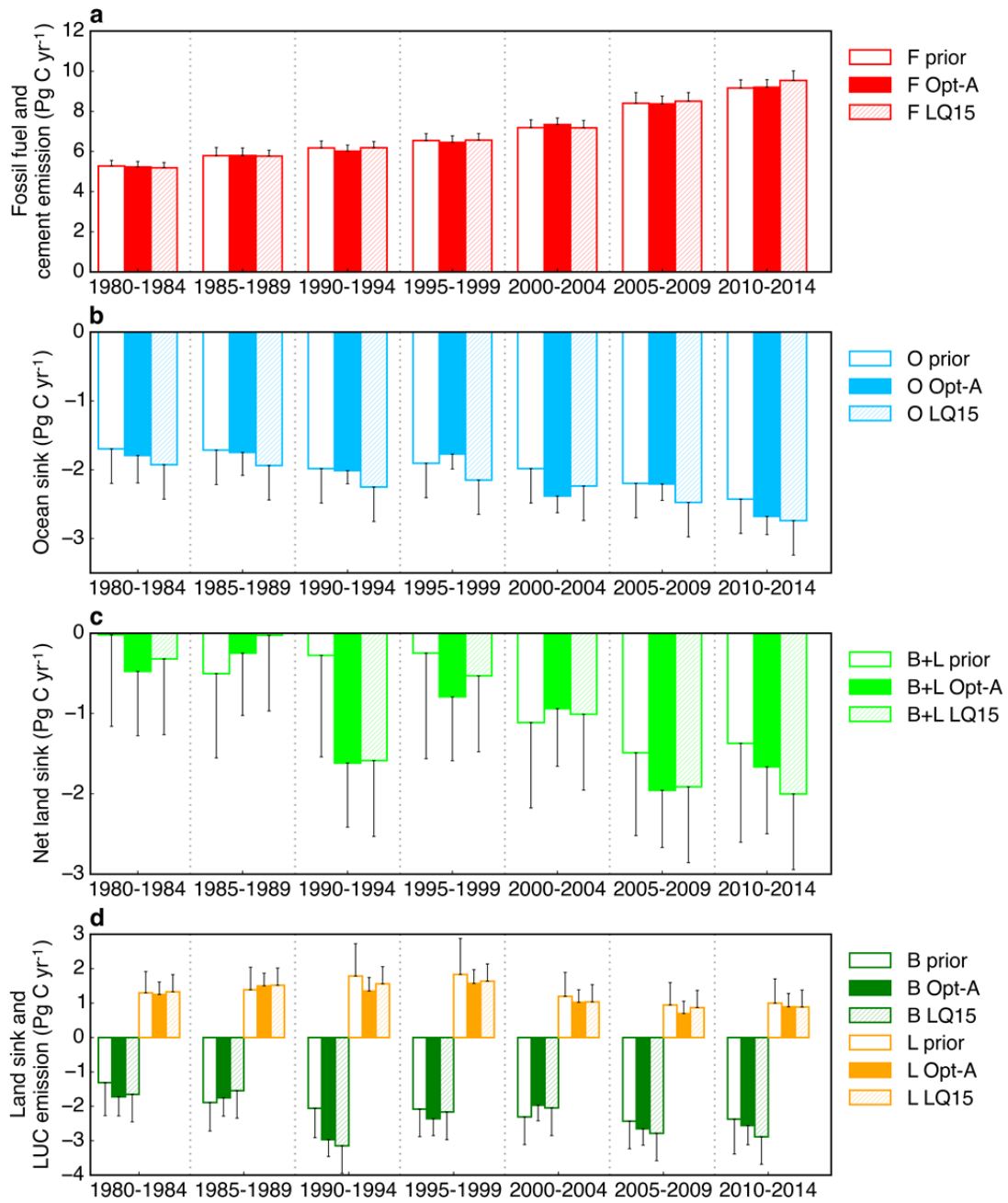
558
559

560 **Figure 3** (**a**) The prior and the optimized gross primary production (GPP, shown as reversed sign)

561 and terrestrial ecosystem respiration (TER) and (**b**) the prior and the optimized fire emissions

562 (FR) from Opt-B. The error bars represent the 1-σ uncertainties. Uncertainties in the prior GPP

563 and TER are not shown because of their large size (about 35 Pg C yr^{-1}).

564
565
566

Supporting information
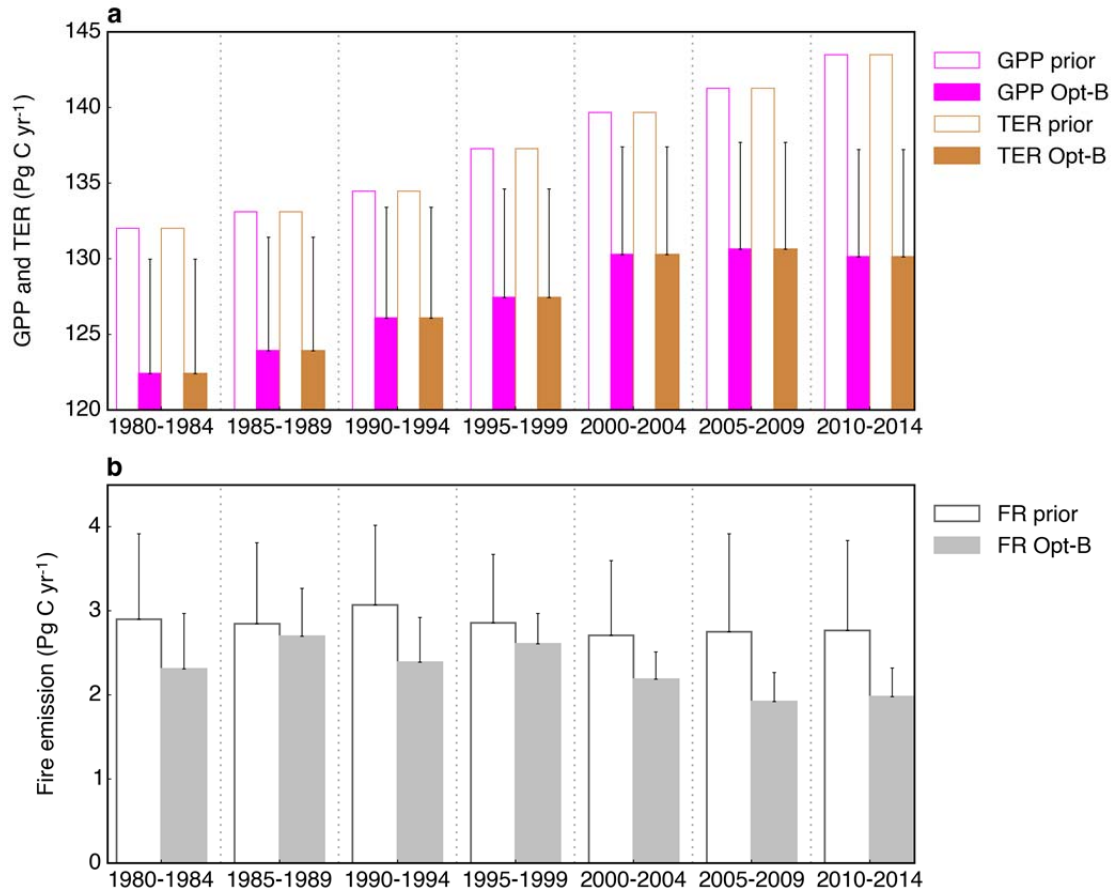
Table S1 Prior and optimized values and uncertainties.

		80-84	85-89	90-94	95-99	00-04	05-09	10-14
prior	F (Pg C yr^{-1})	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 (Pg C yr^{-1})	-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
	B (Pg C yr^{-1})	-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 (Pg C yr^{-1})	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
	α (Pg C yr^{-2})	0.08±0.56	0.61±0.52	-0.19±0.50	1.07±0.13	-0.11±0.55	0.37±0.12	-0.27±0.78
	β (Pg C yr^{-2})	-0.06±0.28	0.55±0.14	-0.17±0.56	0.89±0.42	-0.23±0.24	0.13±0.11	-0.36±0.01
	G_0 (Pg C yr^{-1})						-140.27±36.58	
	T_0 (Pg C yr^{-1})						129.54±44.21	
	FR (Pg C yr^{-1})	2.90±1.02	2.85±0.96	3.07±0.95	2.86±0.81	2.71±0.89	2.75±1.17	2.77±1.07
Opt-A	F (Pg C yr^{-1})	5.22±0.27	5.77±0.38	5.98±0.31	6.45±0.32	7.32±0.32	8.28±0.38	9.15±0.37
	O (Pg C yr^{-1})	-1.82±0.40	-1.76±0.33	-2.02±0.19	-1.77±0.22	-2.39±0.24	-2.23±0.24	-2.70±0.26
	B (Pg C yr^{-1})	-1.86±0.53	-1.83±0.52	-3.10±0.45	-2.34±0.44	-2.03±0.43	-2.80±0.45	-2.83±0.50
	L ((Pg C yr^{-1})	1.43±0.29	1.63±0.30	1.53±0.31	1.55±0.29	1.13±0.27	0.96±0.22	1.24±0.19
	B+L (Pg C yr^{-1})	-0.43±0.70	-0.20±0.69	-1.57±0.67	-0.79±0.64	-0.90±0.59	-1.83±0.56	-1.59±0.59
Opt-B	F (Pg C yr^{-1})	5.24±0.27	5.92±0.40	6.08±0.32	6.54±0.33	7.44±0.32	8.48±0.42	9.33±0.39
	O (Pg C yr^{-1})	-1.77±0.42	-1.66±0.34	-2.03±0.19	-1.74±0.22	-2.32±0.24	-2.15±0.24	-2.60±0.27
	α (Pg C yr^{-2})	-0.08±0.41	0.88±0.33	-0.25±0.39	1.04±0.12	-0.13±0.30	0.36±0.11	-0.79±0.43
	β (Pg C yr^{-2})	0.03±0.27	0.53±0.13	-0.16±0.40	1.18±0.28	-0.23±0.23	0.14±0.11	-0.36±0.01
	G_0 (Pg C yr^{-1})						-129.91±7.05	
	T_0 (Pg C yr^{-1})						126.29±7.08	
	FR (Pg C yr^{-1})	2.31±0.66	2.70±0.57	2.39±0.53	2.61±0.36	2.19±0.32	1.92±0.35	1.98±0.34
	GPP (Pg C yr^{-1})	-122.42±7.56	-123.93±7.49	-126.09±7.32	-127.42±7.21	-130.29±7.11	-130.63±7.05	-130.13±7.10
	TER (Pg C yr^{-1}))	119.59±7.53	120.75±7.47	122.02±7.33	123.89±7.21	126.98±7.10	126.56±7.08	126.26±7.08
	B+L (Pg C yr^{-1})	-0.52±0.54	-0.49±0.54	-1.69±0.41	-0.92±0.42	-1.11±0.41	-2.14±0.50	-1.89±0.51

569 **Table S2** Sources and values of different constraining data-streams used in Opt-A and Opt-B. The blue and orange shades indicate the length of
570 time covered.

	80-84	85-89	90-94	95-99	00-04	05-09	10-14	Opt-A	Opt-B	Reference
CGR (Pg C yr^{-1})	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	x	x	Means are from NOAA (14), and uncertainties are from "Table 1" in Ballantyne et al. (50).
O (Pg C yr^{-1})	-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	x	x	Values from Landschützer et al. (16) are added to an outgassing of 0.45 Pg C yr^{-1} and the corresponding uncertainties from Jacobson et al. (54).
O (Pg C yr^{-1})		-1.82±0.68	-2.03±0.68	-1.68±0.68	-1.76±0.68	-1.56±0.68	-2.39±0.68	x	x	Values from Rödenbeck et al. (17) are added to an outgassing of 0.45 Pg C yr^{-1} and the corresponding uncertainties from Jacobson et al. (54).
O (Pg C yr^{-1})			-2.00±0.40					x	x	Means and uncertainties are from McNeil et al. (19) to constrain O from 1990 to 1999.
O(Pg C yr^{-1})			-1.80±0.20					x	x	Means and uncertainties are from Steinkamp and Gruber (20) to constrain O from 1990 to 1999.
O (Pg C yr^{-1})			-1.91±0.60	-2.05±0.60				x	x	Means and uncertainties are from Khatiwala et al. (18).
B (Pg C yr^{-1})			-2.50±1.08		-2.30±1.12			x		Means are from Pan et al. (21). 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. (21) are also included, which may represent the uncertainty of land carbon sinks from other ecosystems like grassland.
B (Pg C yr^{-1})					-1.70±0.57			x		RECCAP (see **Table S3**)
L (Pg C yr^{-1})					1.00±0.19			x		RECCAP (see **Table S3**)
L (Pg C yr^{-1})	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	x		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).
L (Pg C yr^{-1})	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	x		Means and uncertainties are from the BLUE model by Hansis et al. (23). The uncertainties refer to the difference to the alternative estimate.
O$_2$/N$_2$[a] (per meg unit)			-6.10±0.22	-7.37±0.69	-8.89±0.30	-9.15±0.35	-10.64±0.37	x	x	Data are from Scripps O$_2$ program (http://scrippso2.ucsd.edu/), 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.
B+L[b] (Pg C yr^{-1})					-1.82±0.55			x	x	Values are from Liu et al. (22). They estimated the aboveground biomass carbon from Vegetation Optical Depth, extrapolated to the total biomass by the ratios between aboveground and belowground biomass in

									different regions, and then calculated the total carbon (including litter and soil carbon) trend based on the ratios between total carbon and total biomass.
LAI trend[c]	0.04±0.02	0.08±0.10	-1.66±0.48	-0.06±0.24	0.26±0.18	-0.04±0.05		x	The trend of LAI is calculated from the average of global mean LAI from 3 datasets: GIMMS LAI3G (Zhu et al. (34)), GLASS LAI (Xiao et al. (35)) and GLOBMAP LAI (Liu et al. (36)). The errors in the linear trend regression are taken as uncertainties.
β (Pg C yr^{-2})			0.10±0.49					x	An assumed linear trend of soil respiration from 1989 to 2008 is calculated from Bond-Lamberty and Thomson. (37). Note that the trend of soil respiration here is used as a proxy of the global TER trend, and the autotrophic respiration is omitted. The value of soil respiration is not used to constrain TER, because of the omission of autotrophic respiration.
GPP[d] (Pg C yr^{-1})					-110.79±21.17			x	GPP values are from MODIS by Zhao et al. (32). The average of global yearly GPP between 2001 and 2009 is used here. The uncertainties are estimated from the variations between different meteorological data required in the MODIS GPP calculation by Zhao et al. (32).
GPP[e] (Pg C yr^{-1})					-131.88±7.76			x	GPP values are from the upscaling of FLUXNET by Jung et al. (33). The average of global yearly mean GPP between 2001 and 2009 is used here. The uncertainties refer to the standard deviation across different model trees in Jung et al. (33).
FR (Pg C yr^{-1})				2.50±0.50	1.97±0.39	1.92±0.38	1.83±0.37	x	Means are from GFED4 by Giglio et al. (38), and uncertainties of 20% are assumed, as by van der Werf et al. (55).
FR (Pg C yr^{-1})	2.40±0.96		2.58±1.03	2.78±1.11	2.53±1.01			x	Means are from GICC data set, and the uncertainties are set as 40% based on the uncertainties on burned areas in Mieville et al. (39).
FR (Pg C yr^{-1})	2.14±1.07	1.84±0.92	2.42±1.21	2.35±1.17				x	Means are from RETRO data set, and the estimated uncertainties of 50% by Shultz et al. (40) is adopted here.

571   [a] The relationship of $\Delta O_2 = -\alpha_F F + \alpha_B(B+L) + Z_{O_2}$ from Keeling et al. (15) is used.

572   [b] B+L is corresponding to B+L in Opt-A and is corresponding to $(G_0+\alpha t) + (T_0+\beta t) + FR$ in Opt-B.

573   [c] The relative trend of LAI is used to constrain the relative trend of GPP.

574   [d] The average of yearly GPP values between 2001 and 2011 (the overlapped period of GPP from FLUXNET) is used to constrain G_0.

575   [e] The average of yearly GPP values from 2001 to 2011 (the overlapped period of GPP from MODIS) is used to constrain G_0.

576

577 **Table S3** The estimates of B and L from Regional Carbon Cycle Assessment and Processes
578 (RECCAP) international research project. In brief, the observed carbon fluxes in the nine regions
579 were evaluated and summed to give the global fluxes. The net ecosystem exchange (NEE)
580 calculated from the mass balance of the emissions and outgassing of carbon to the atmosphere,
581 net primary production (NPP), and soil heterotrophic respiration (SHR) is taken as B in this
582 study. Negative values indicate a flux from the atmosphere to the land.

| region	flux	period	Mean ($Tg\ C\ yr^{-1}$)	Uncertainty ($Tg\ C\ yr^{-1}$)	Reference
North America	L	2000s	-130	60	King et al. (56)
	B	2000-2009	-418	288	Mass balance
Europe	L	2003-2007	-2	1	Schulze et al. (57)
	B	2000-2007	-234	166	Mass balance
Russia	L	1990-2006	-34	2	Dolman et al. (58)
	B	2007-2009	-600	224	Mass balance
South Asia	L	2000-2009	-14	50	Patra et al. (59)
	B	2000-2009	-86	34	Mass balance
East Asia	L	2000-2009	-13	29	Piao et al. (60)
	B	2000-2009	-256	35	Mass balance
Southeast Asia	L	undefined	59	12	Liu et al. (22) only over LUC affected grid cells
	B	undefined	-97	186	Mass balance
South America	L	1990-2012	525	130	Gloor et al. (61) "Table 12"
	B	2000-2009	11	284	Mass balance
Africa	L	1990-2009	510	100	Valentini et al. (62) "Table 11"
	B	2000-2007	12	292	Mass balance
Australia	L	1990-2011	18	7	Haverd et al. (63)
	B	1990-2011	-71	36	Mass balance
Globe	L	2000-2009	919	184	sum
	B	2000-2009	-1740	604	sum

583

584   **Figure S1** The values and uncertainties of ocean sinks (O), land sinks (B) and land-use change
585   (LUC) emissions (L) in different data-streams. The dashed lines indicate the period covered.
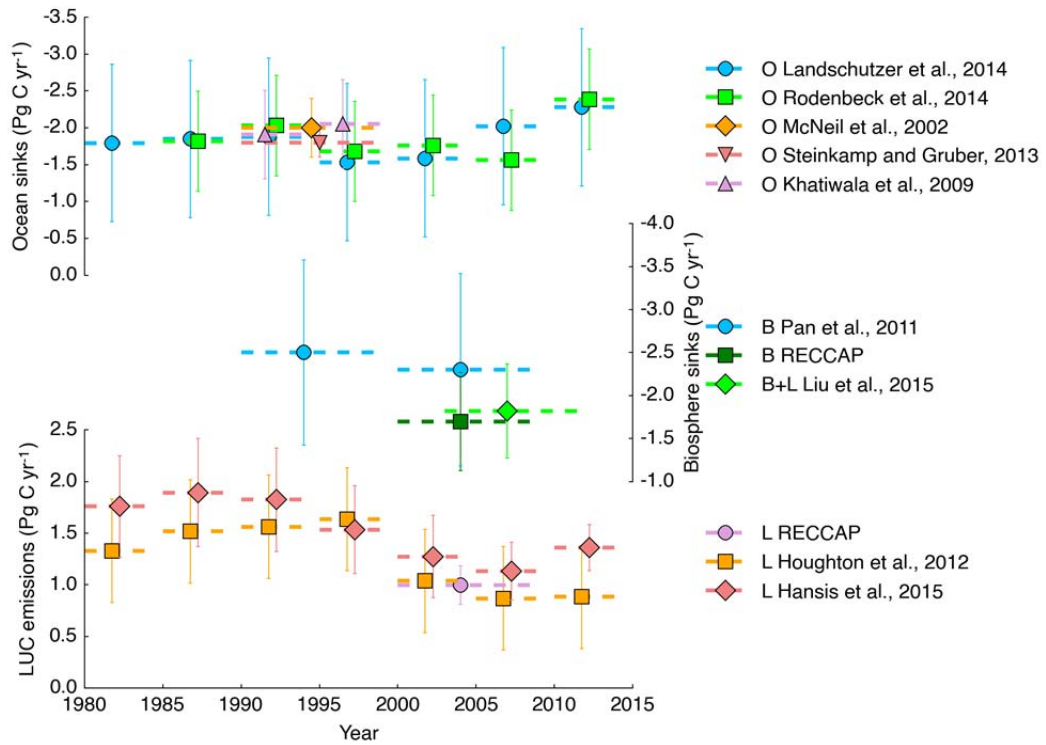

586
587

588 **Figure S2** The temporal trend of ocean sink (O), land sink (B), land-use change emissions (L)
589 and net land flux (B+L) from Opt-A and LQ15.
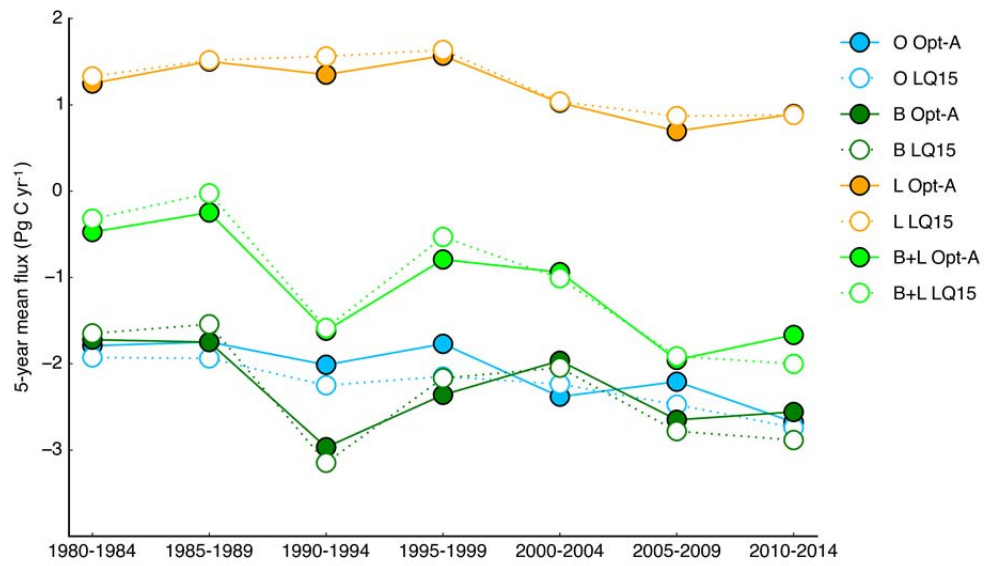
590

591 **Figure S3** The fluxes of fossil fuel and cement (F), and land-use change (L) emissions, ocean (O)
592 and land (B) sink, and net land sink (B+L) using different prior data. The bars around each flux in
593 each period from left to right represent Opt-A, double prior uncertainties, prior F from IEA, prior
594 F from EDGAR, prior F from CDIAC, prior F from BP and prior B and O from CMIP5,
595 respectively.
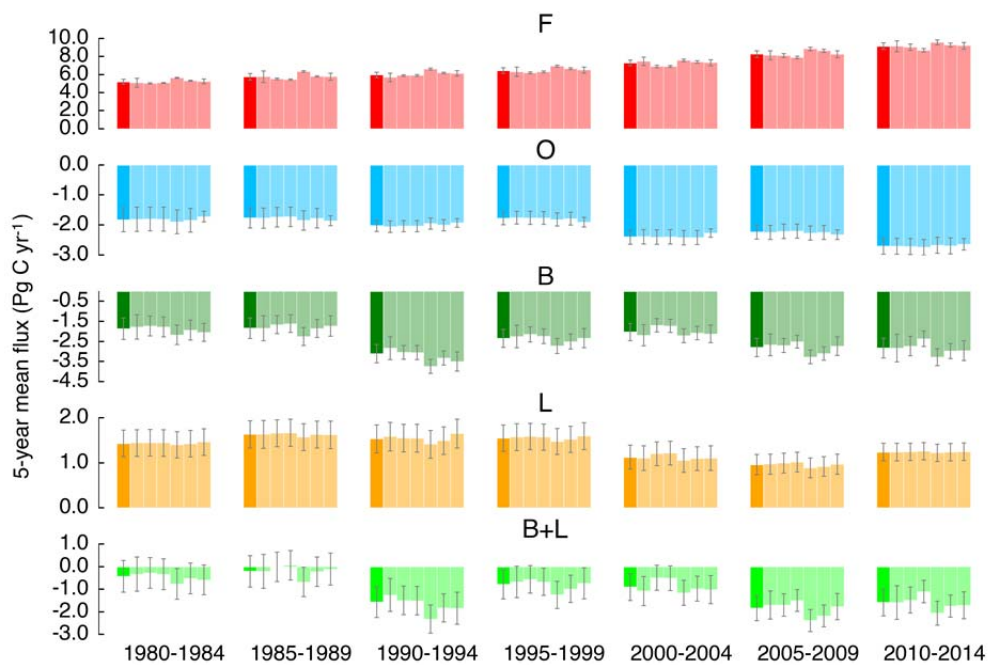
596
597

598   **Figure S4** The partitioning of global total emissions (fossil fuel and cement emissions and land-
599   use change emissions, F+L) into atmosphere, ocean (O) and biosphere (B, background shade).
600   The markers are absolute absorbed fractions and uncertainties in ocean (blue circles) and
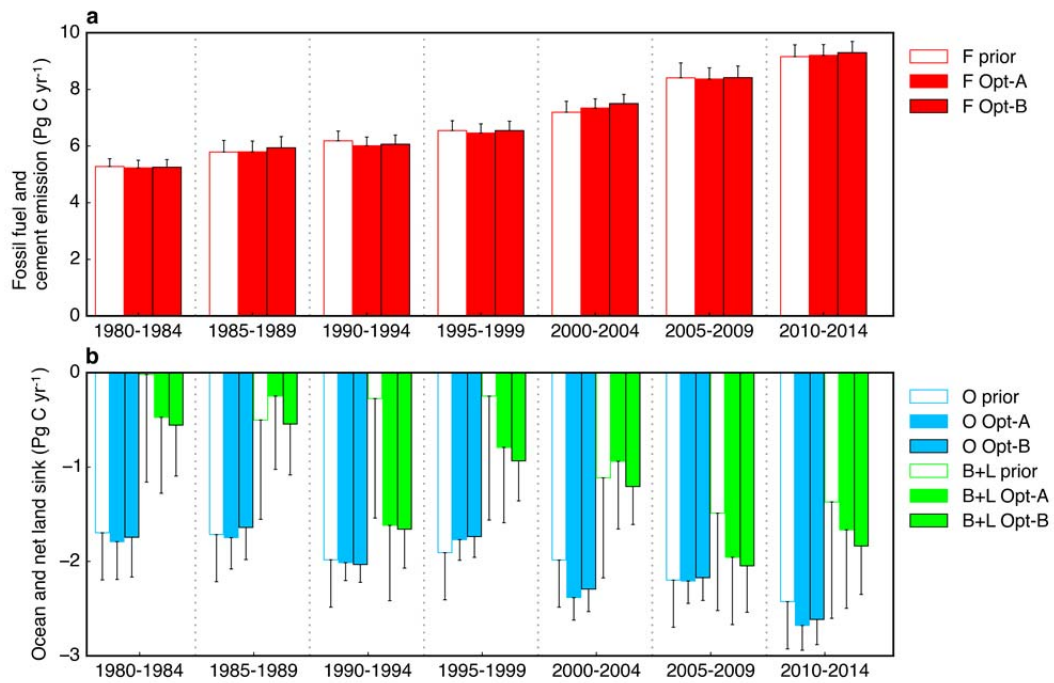601   biosphere (green squares).
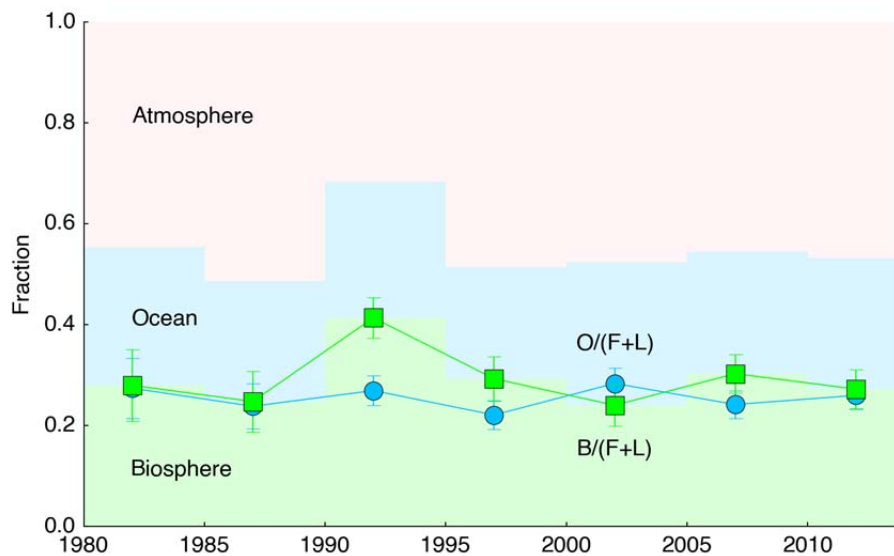
602

603

604   **Figure S5** The framework of Opt-B. The number of data-streams is marked on the right.


605
606
607
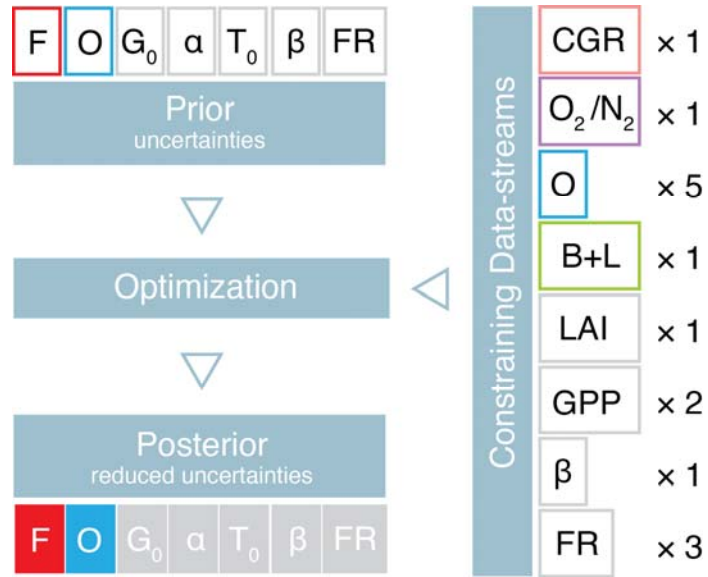
Figure S6 (**a**) The fossil fuel and cement emissions (F) and (**b**) the ocean sink (O) and net land
flux (B+L) from prior knowledge, Opt-A and Opt-B. All the fluxes are 5-year means over each
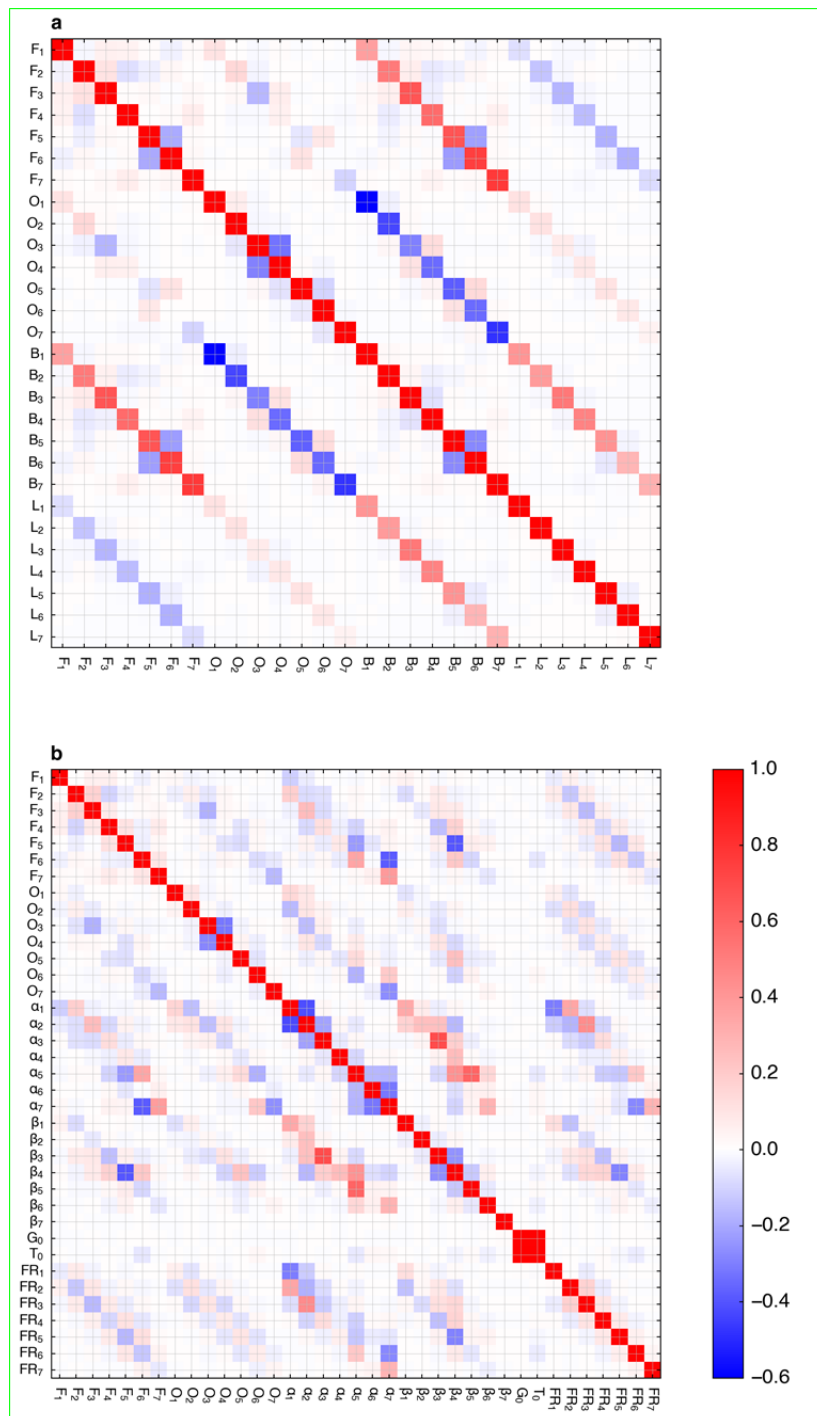period.

611

612

613

614 **Figure S7** The optimized gross primary production (GPP, shown as reversed sign) and terrestrial
615 ecosystem respiration (TER) in the sensitivity tests for Opt-B. The shaded areas represent their
616 uncertainties. The differences between sensitivity tests and Opt-B are: (a) using yearly GPP
617 estimates during 1982-2011 from ref. (33) and removing LAI constraints as a proxy of the GPP
618 relative trend(34–36); (b) only using the average GPP from MODIS(32) to constrain G_0; (c) only
619 using the average GPP from ref. (33) to constrain G_0; (d) additionally using the 5-year mean GPP
620 during 2000-2004, 2005-2009 and 2010-2014 from MODIS(32) to constrain the corresponding 5-
621 year mean GPP; (e) setting prior relative trend of GPP (α) to zero; (f) setting prior trend of GPP
622 (α) to 0 and removing LAI constraints to the relative trend of GPP(34–36); (g) removing the
623 constraint from the trend of soil respiration(37); (h) setting prior trend of TER (β) to zero and
624 removing the constraint from the trend of soil respiration(37); and (i) setting both prior trends of
625 GPP (α) and TER (β) to 0 and removing LAI constraints(34–36).

626
627

37
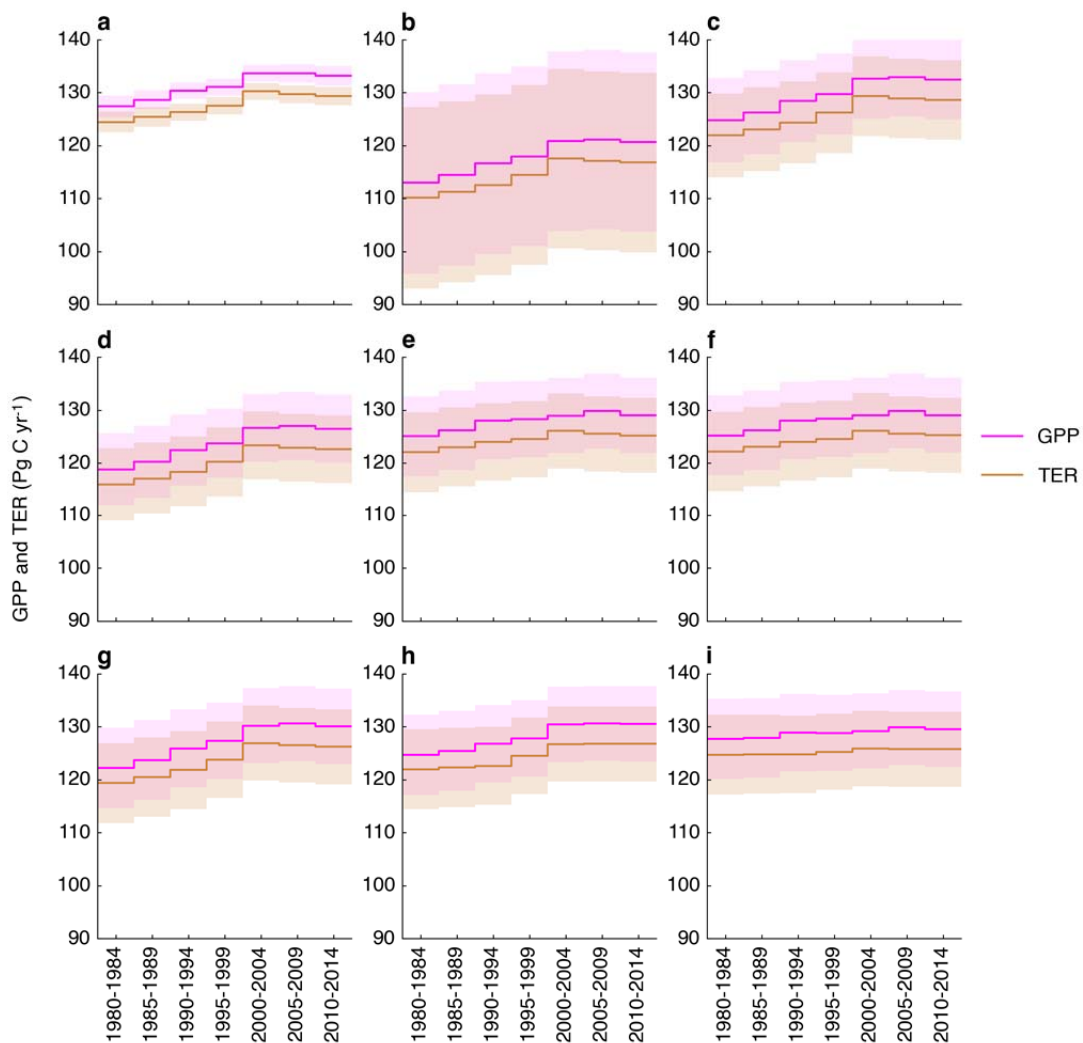
Figure S8 Correlations between the posterior uncertainties in (**a**) Opt-A and (**b**) Opt-B. The
subscripts 1 to 7 represent the 5-year periods from 1980-1984 to 2010-2014 in sequence.

630

631

38