

1
2 **Supplementary Information for “Assimilation of multiple different datasets results in large**
3 **differences in regional to global-scale NEE and GPP budgets simulated by a terrestrial**
4 **biosphere model”**

5
6 **Cédric Bacour, Natasha MacBean, Frédéric Chevallier, Sébastien Léonard, Ernest N. Koffi, Philippe**
7 **Peylin**
8

9 **Supplementary Text S1: Data assimilation experiments: differences with the study of Peylin et al.**
10 **(2016)**

11 There are few differences in the DA experimental set-up compared to the present study: *i*) the study of
12 Peylin et al. (2016) considers three years of atmospheric CO₂ data (from 2002 to 2004), while this study
13 uses ten years of observations (2000-2009); *ii*) the set of optimized parameters is not strictly identical:
14 the step-wise study did not optimize the parameters controlling the maximum LAI value per PFT nor the
15 root profile, which are included in this study, but instead did include two additional parameters, one
16 controlling the albedo of vegetation and the other reducing the hydric limitation of photosynthesis,
17 which are not considered here; *iii*) for optimization of the phenology using satellite NDVI data, the C4
18 grass PFT was calibrated in Peylin et al. (2016), which is not the case in this study (MacBean et al., 2015)
19 found that phenology for semi-arid PFTs was not well captured by the model and further improvements
20 to the phenology schemes for these PFTs are needed); *iv*) the selection of the eddy-covariance sites is
21 more selective in the present study (a few sites for which the model-data inconsistency was too
22 important were discarded), which slightly reduces the number of site-years available for some PFTs; *v*)
23 finally, the *a priori* errors on model parameters (at the first and second steps) were set to 40% of the
24 parameter variation range in Peylin et al. (2016), and were hence larger than what is prescribed in this
25 study as a result of the consistency checks performed in Section 2.3.4.2.

26
27 **Supplementary Text S2: Processing of atmospheric CO₂ data**

28 In order to analyze the fit to the atmospheric CO₂ concentrations in terms of trend and seasonal cycle
29 (magnitude and phase), the measured and modeled monthly time series are fitted using the CCGCRV
30 package (<ftp://ftp.cmdl.noaa.gov/user/thoning/ccgcrv/>) following Thoning et al. (1989). It decomposes
31 the time series into a first-order polynomial term (that represents the trend) and four harmonics, and
32 then filters the residuals of that function in frequency space using a low-pass filter (cutoff frequency of
33 65 days). The seasonal cycle corresponds to the harmonics plus the filtered residuals. For a given time
34 series, we calculate the magnitude of the seasonal cycle for each year as the difference between the
35 maximum and minimum value, and the carbon uptake period (CUP) as the sum of the days when the

36 values of the seasonal cycle extracted from the CO₂ concentration time series are negative (plant
 37 removing CO₂ from the atmosphere by convention) (Peylin et al., 2016).

38

39 **Supplementary Text S3: Consistency diagnostics on the errors**

40 ***Desroziers et al. (2005) tests***

41 Several attempts were performed to specify the errors on model parameters in order to approach this
 42 goal considering each data-stream independently. With an initial definition of the parameter error
 43 corresponding to 40% of their variation range, the diagnostics on the **R** matrix, show a strong
 44 overestimation for all data streams (ratios about 3 for NEE and LE, 2 for NDVI and 12 for atmospheric
 45 CO₂), while the diagnostics on **B** were more consistent with ratios slightly higher than 1 but for NDVI
 46 (2.5). These results led us to revise the definition of **B** by decreasing the error for all parameters such
 47 that it corresponds to about 20% of the variation range for phenological parameters, and 12% for the
 48 other parameters (a value close to what was prescribed in Kuppel et al. (2013)).

49

50 ***Reduced chi-square***

51 For all experiments but those involving atmospheric CO₂ measurements, the values of the reduced chi-
 52 square (after optimization) over all data are below 1 (Table S1), which corroborates the overestimation
 53 of the model-data and parameter errors observed previously. For fluxes and satellite data, this
 54 overestimation of the model-data error was expected, and even desired, given that the covariances in **R**
 55 were neglected by construction (off-diagonal elements set to zero). For CO₂, the large value of χ^2
 56 expresses a strong underestimation of the observation error not highlighted by the consistency
 57 diagnostics. Indeed, when determining **R**_{CO₂}, we purposely did not account for the structural error in
 58 ORCHIDEE that largely explains the strong bias between observed and simulated CO₂ temporal profiles
 59 by about 1 ppm.yr⁻¹. This underestimation is even inflated in the joint assimilation experiments, even
 60 though the reduced chi-square over all data remains close to 1.

61

experiment	Data - stream			
	F	VI	CO2	all data
F	0.91			0.91
VI		0.78		0.78
CO2			8.57	8.57
F+VI	0.95	0.50		0.73
F+CO2	0.97		11.56	1.4
VI+CO2		0.74	11.72	1.18
F+VI+CO2	0.99	0.75	11.3	1.09
F+VI+CO2-2steps	0.96	0.67	7.88	0.97

62 **Table S1: Values of the reduced chi-square determined after model calibration for the various assimilation**

63 experiments, for each data-stream.

64

65 **Supplementary Text S4: Optimisation performances**

66 All optimizations but two (F and VI) reached a pre-defined maximum number of iterations (set to 35 for
67 L-BFGS-B), therefore causing a hard stopping of the optimization (cf, Table S2, which also provides the
68 values of the misfit functions for all assimilation experiments, relative to the background and to the
69 observations). For the last iterations however, the variations of the misfit functions were low in all these
70 cases, indicating that the final iterations were close to the minimum. The comparison between the
71 observation and parameter terms of the posterior cost function shows how the total cost function is
72 dominated by the weight of the model-data misfit.

73 The highest rate of change of the total cost function related to the observation term is obtained for the
74 CO₂ assimilation with a reduction of the misfit between model outputs and measurements by about 46.
75 This is directly related to the correction of the large bias in the prior model with carbon pools at
76 equilibrium relative to the prescribed prior error. Noticeably, the strong model improvement reached
77 for CO₂ comes with only a small variation in the model parameters as depicted by the posterior value of
78 J_b . For the assimilation of the fluxes and satellite data alone (F and VI respectively), the model
79 improvement is smaller, about 1.1, but shows a stronger departure of the parameters from their prior
80 values compared to CO₂ (Figure 3). The ratio of the norm of the gradient of the misfit function is also
81 the highest for the CO₂ experiment. On the opposite, it is slightly lower than one for VI which may
82 indicate a possible issue of convergence towards the solution.

83

84 The two-step approach for the assimilation involving the three data-streams results in an enhanced
85 agreement of the model with all data as compared to the one-step optimization. In parallel, the change
86 in parameter values (departure from the background) is also higher for the two-step approach (Figure 3).

87

experiment	Number of iterations	Jo prior	Jo post	Jo prior/ Jo post (obs part)	Jo(F) post	Jo(CO ₂) post	Jb post	Ratio norm grad J (prior/post)
F	34	75396	68305	1.10	68305		117.6	3.95
VI	29	65696	58517	1.12			37.9	0.94
CO ₂	35	1256783	27238	46.14		27238	7.8	759.5
F+VI	35	142118	108961	1.30	71353		79.3	0.97
F+CO ₂	35	1332190	109994	12.11	73232	36763	1.05	27.7
VI+CO ₂	35	1323494	92543	14.30		37257	1.3	132.3
F+VI+CO ₂	35	1398901	166797	8.39	74435	35918	1.6	168.7
F+VI+CO ₂ -2steps	35	1398901	148206	9.43	72654	25002	44.6	-

88 **Table S2: Characteristics of the various assimilation experiments: number of iterations, value of the cost**
89 **functions related to the observation (Jo) and parameter terms (Jb) prior and posterior to the assimilation (as**
90 **well as ratio of the posterior to prior values for Jo), ratio of norm of the gradient of the misfit functions (prior vs**

91 posterior).

92

93

94

95 **Supplementary Text S5: Analysis of the reduction of the model-data misfit**

96 ***Mono-data stream assimilations***

97 The increased consistency between model and flux data achieved after assimilation of F data is usually
98 higher for NEE (median RMSD reduction of 10.4%, ranging from -69% to 38%) than for LE (0.3%; -42% /
99 28% range). This is largely explained by the higher number of optimized parameters related to the
100 carbon cycle relative to the water cycle, and by the optimization of the multiplicative factor of the soil
101 carbon pools that corrects the bias in the ecosystem respiration inherent to the model spin-up
102 (Carvalhais et al., 2010; Kuppel et al. 2012). The strong model improvement for FAPAR in the VI
103 assimilation (22.2% median; -32% / 36% range) follows a strong decrease of the simulated growing
104 season length for deciduous PFTs in better accordance with the satellite observations, as discussed in
105 MacBean et al. (2015). It mainly results from an earlier senescence for the several PFTs while the change
106 of leaf onset depends on the type of vegetation. Both for the F and VI experiments, the reduction of the
107 model-data misfit can be negative for some sites/pixels. This reflects how the assimilation may degrade
108 the model performance at some sites/pixels by seeking for a common parameter set. This is not
109 observed for atmospheric CO₂ data for which the optimized model is always closer to the observations
110 than the prior model at all stations. Assimilating atmospheric CO₂ concentration measurements corrects
111 the strong overestimation of the prior model (as also described in Peylin et al. (2016)), with a median
112 RMSD reduction of 76% (ranging from 10% at HUN to 90% at SPO). This improvement corresponds to an
113 increase of the net land carbon sink at the global scale in order to correct the strong mismatch between
114 the observed trend and the *a priori* model. It is mainly realized by the optimization of the multiplicative
115 factor of the soil carbon pools. As seen in Figure 2 from the detrended seasonal cycles of atmospheric
116 CO₂ data (light red box), the changes in the modelled amplitude and phasing is smaller but still in better
117 agreement with the observed data (median value of RMSD reduction of 14.4%; -21% / 55% range).

118

119 ***Multiple-data stream assimilations***

120 The simultaneous assimilation of flux measurements and satellite NDVI data leads to enhanced model
121 improvement as compared to when these data are assimilated alone: the median RMSD reductions are
122 10.8% for NEE (10.4% in the F case) and 36.7% for FAPAR/NDVI (22.2% in the VI case). In the
123 simultaneous assimilations involving atmospheric CO₂ data, the most of the model improvement is
124 attributed to CO₂ while the benefit relative to fluxes and FAPAR/NDVI is weak: for NEE, the median
125 RMSD reductions are only of 2.5% and 2.6% in the F+CO₂ and F+VI+CO₂ cases (as compared to 10% in

126 the F case); for FAPAR, the median values are 1.2% and 1.4% for the VI+CO₂ and F+VI+CO₂ experiments
127 (22% in the VI case).

128 The 2-steps assimilation F+VI+CO₂ results in a higher model improvement regarding both NEE and
129 FAPAR (respectively 5.5% and 11.2%) than the one-step approach.

130 Regarding the raw atmospheric CO₂ data, the median improvements are 76.1% for CO₂, 76.3% for
131 F+CO₂, 73.6% for VI+CO₂, 72.9% for F+VI+CO₂ and only 25.6% for F+VI+CO₂-2steps.

132 More pronounced differences between experiments are obtained for the de-trended CO₂ time series:
133 while the median RMSD reduction is of 14% in the CO₂ experiment, it is decreased to 7.8% in F+CO₂,
134 8.4% in VI+CO₂, and 10.6% in F+VI+CO₂; at the opposite the RMSD reduction is increased to 15.4% in
135 F+VI+CO₂-2steps.

136

137 **Supplementary Text S6: Global budget and uncertainty reduction**

138 For NEE, the global scale budget is about -2.4 GtC.yr⁻¹ for all experiments using atmospheric CO₂ as a
139 constraint: the lower value of -2.28 GtC.yr⁻¹ is found for F+CO₂; the higher values of -2.49GtC.yr⁻¹ and -
140 2.48 GtC.yr⁻¹ are obtained for CO₂ / F+VI+CO₂-2steps.

141 In the northern and southern hemispheres, the CO₂ assimilation results in the largest C sinks (-1.65 / -
142 0.04 GtC.yr⁻¹ for NH/SH) while the 2step assimilation induces the lowest one (-0.41 / 0.003 GtC.yr⁻¹); the
143 opposite result is obtained in the southern hemisphere with lower (-0.79 GtC.yr⁻¹) / higher (-2.06 GtC.yr⁻¹)
144 budgets found for CO₂ / F+VI+CO₂-2steps.

145 The reduction of the global scale GPP budget is respectively of -19.61 GtC.yr⁻¹ and -17.91 GtC.yr⁻¹ for the
146 F and VI experiments, which correspond to the largest corrections obtained among the various
147 assimilations considered.

148 The averaged change in GPP is about -7.33 GtC.yr⁻¹ globally for the CO₂ assimilation experiment. The
149 corrections for the joint assimilations involving CO₂ data is even lower: the mean global change are -
150 1.07 GtC.yr⁻¹ for VI+CO₂, -1.35 GtC.yr⁻¹ for F+CO₂ and -1.98 GtC.yr⁻¹ for F+VI+CO₂. For the F+VI+CO₂ 2-
151 step experiment, the constraint on GPP is close to that obtained when CO₂ data are assimilated alone (-
152 7.70 GtC.yr⁻¹).

153

154 For the joint assimilations, the posterior errors on NEE is about 0.9 GtC.yr⁻¹ globally and about 0.3
155 GtC.yr⁻¹ for the three regions considered. The lowest posterior errors on GPP are obtained for the two
156 experiments that combine the three data streams (about 0.09 GtC.yr⁻¹ at the global scale, and about
157 0.04 GtC.yr⁻¹ depending on the region). The values are close to the ones obtained with F+CO₂.

158

Name	TropEBF ^F	TropBRF ^{VI}	TempENF ^F	TempEBF ^F	TempDBF ^{F,VI}	BorENF ^F	BorDBF ^{F,VI}	BorDNF ^{VI}	C3Grass ^{F,VI}
<i>Photosynthesis</i>									
V_{cmax}	65 [35;95] <i>10</i>	65 [35;95] <i>10</i>	35 [19;51] <i>5.3</i>	45 [25;65] <i>6.7</i>	55 [30;80] <i>8.3</i>	35 [19;51] <i>5.3</i>	45 [25;65] <i>6.7</i>	35 [19;51] <i>5.3</i>	70 [38;102] <i>10.7</i>
G_{s,slope}	9 [6;12] <i>1</i>	9 [6;12] <i>1</i>	9 [6;12] <i>1</i>	9 [6;12] <i>1</i>	9 [6;12] <i>1</i>	9 [6;12] <i>1</i>	9 [6;12] <i>1</i>	9 [6;12] <i>1</i>	9 [6;12] <i>1</i>
T_{opt}	37 [29;45] <i>2.7</i>	37 [29;45] <i>2.7</i>	25 [17;33] <i>2.7</i>	32 [24;40] <i>2.7</i>	26 [18;34] <i>2.7</i>	25 [17;33] <i>2.7</i>	25 [17;33] <i>2.7</i>	25 [17;33] <i>2.7</i>	27.25 [19.2;35.2] <i>2.7</i>
SLA	0.0154 [0.007;0.03] <i>0.0038</i>	0.0260 [0.013;0.05] <i>0.0062</i>	0.0093 [0.004;0.02] <i>0.0027</i>	0.02 [0.01;0.04] <i>0.005</i>	0.026 [0.013;0.05] <i>0.0062</i>	0.0093 [0.004;0.02] <i>0.0027</i>	0.026 [0.013;0.05] <i>0.0062</i>	0.019 [0.009;0.04] <i>0.0052</i>	0.026 [0.013;0.05] <i>0.0062</i>
<i>Soil water availability</i>									
H_{um,cste}	0.8 [0.2;3] <i>0.47</i>	0.8 [0.2;3] <i>0.47</i>	1 [0.25;4] <i>0.62</i>	0.8 [0.2;3] <i>0.47</i>	0.8 [0.2;3] <i>0.47</i>	1 [0.25;4] <i>0.62</i>	1 [0.25;4] <i>0.62</i>	0.8 [0.2;3] <i>0.47</i>	4 [1;10] <i>1.5</i>
<i>Phenology</i>									
LAI_{MAX}	7 [4;10] <i>1</i>	7 [4;10] <i>1</i>	5 [3;8] <i>0.8</i>	5 [3;8] <i>0.8</i>	5 [3;8] <i>0.8</i>	4.5 [2.5;6.5] <i>0.7</i>	4.5 [2.5;6.5] <i>0.7</i>	3 [1.5;4.5] <i>0.5</i>	2.5 [1.5;3.5] <i>0.3</i>
K_{pheno,crit}		1 [0.7; 1.8] <i>0.18</i>			1 [0.7; 1.8] <i>0.18</i>		1 [0.7; 1.8] <i>0.18</i>	1 [0.7; 1.8] <i>0.18</i>	1 [0.7; 1.8] <i>0.18</i>
T_{senes}					12 [2;22] <i>3.3</i>		7 [-3;17] <i>3.3</i>	2 [-8;12] <i>3.3</i>	-1.375 [-11.4;9.4] <i>3.5</i>
L_{age,crit}	730 [490;970] <i>80</i>	180 [120;240] <i>20</i>	910 [610;1210] <i>100</i>	730 [490;970] <i>80</i>	180 [90;240] <i>25</i>	910 [610;1210] <i>100</i>	180 [90;240] <i>27.5</i>	180 [90;240] <i>27.5</i>	120 [30;180] <i>25</i>
K_{LAI,happy}	0.5 [0.35;0.7] <i>0.06</i>	0.5 [0.35;0.7] <i>0.06</i>	0.5 [0.35;0.7] <i>0.06</i>	0.5 [0.35;0.7] <i>0.06</i>	0.5 [0.35;0.7] <i>0.06</i>	0.5 [0.35;0.7] <i>0.06</i>	0.5 [0.35;0.7] <i>0.06</i>	0.5 [0.35;0.7] <i>0.06</i>	0.5 [0.35;0.7] <i>0.06</i>
<i>Respiration</i>									
Q10					1.9937 [1;3] <i>0.33</i>				
HR_{H,c}					-0.29 [-0.59;0.01] <i>0.1</i>				
MR_c					1 [0.5;2] <i>0.25</i>				
K_{soilC,site}					1 [0.5;2] <i>0.1</i>				
K_{soilC,reg}					1 [0.7;1.3] <i>0.1</i>				

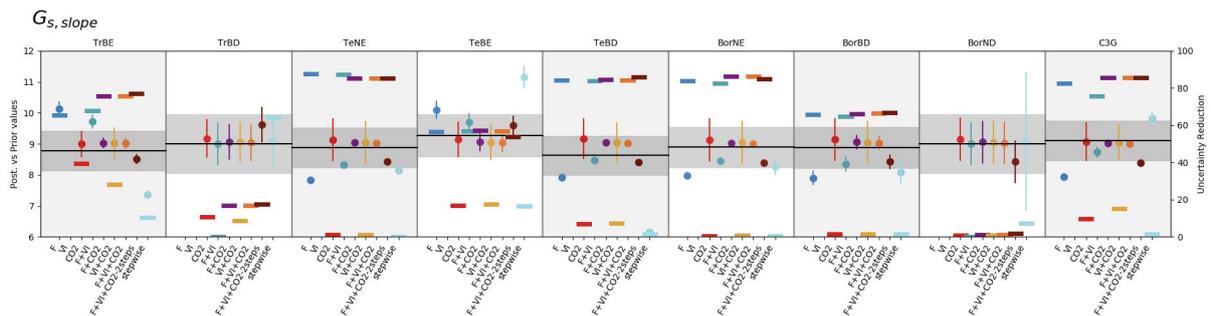
159 **Table S3: Prior value, interval of variation (in square brackets) and 1-sigma prior error (italic), of the optimized**
160 **parameter. Except for those related to respiration, all parameters are PFT-dependent. The exponents F and VI**
161 **associated to each PFT name indicate the availability of flux (F) and satellite (VI) data.**

162

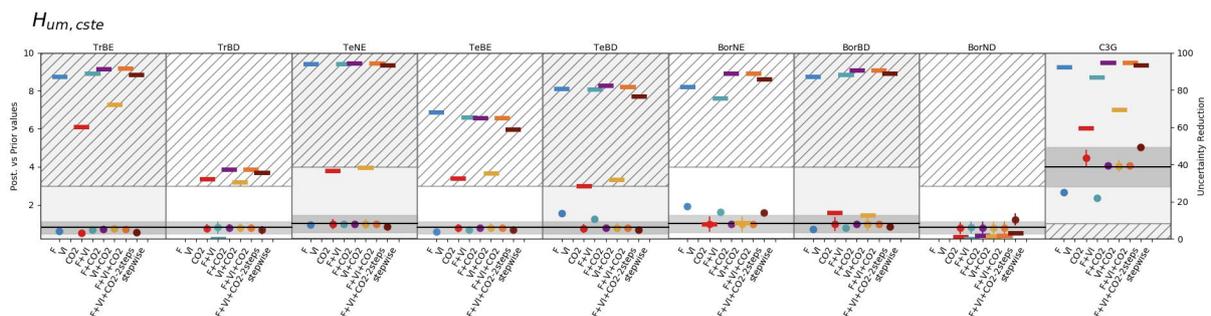
163

164
 165 **Figure S1: Prior and posterior parameter values and uncertainties for a set of optimized parameters (eight PFT-**
 166 **dependent parameters and four non-PFT dependent). The prior value is shown as the horizontal black line and**
 167 **the prior uncertainty (standard deviation) as the gray area encompassing it along the x-axis. For the PFT-**
 168 **dependent parameters, each box corresponds to a given PFT; empty boxes indicate that this parameter was not**
 169 **constrained for the corresponding PFTs. The white zone (non-dashed area) corresponds to the allowed range of**
 170 **variation. The optimized values are provided for each assimilation experiment (the eight ones considered in this**
 171 **study and the one from Peylin et al. (2016) – "stepwise"); the corresponding posterior errors are displayed as**
 172 **the vertical bars. Note that the prior values presented here are those used in this study, and not those of the**
 173 **stepwise (which are higher/lower for the photosynthesis and respiration / phenological parameters). For each**
 174 **assimilation experiment is also provided the uncertainty reduction (right y-axis) as the thick opaque horizontal**
 175 **bars. For KsoilC_reg, the posterior values displayed here correspond to the mean over the eco-regions (without**
 176 **Antarctica) considered; the semi-transparent horizontal bars on either side of the posterior values correspond to**
 177 **the standard deviation of the estimates**

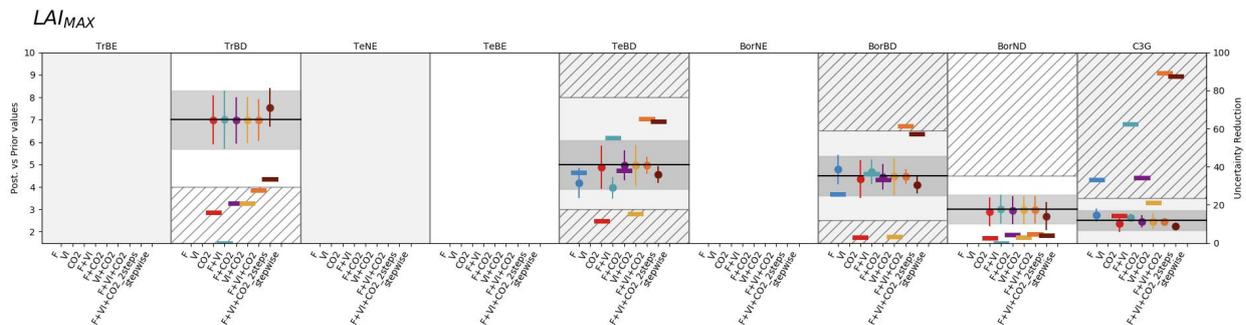
178



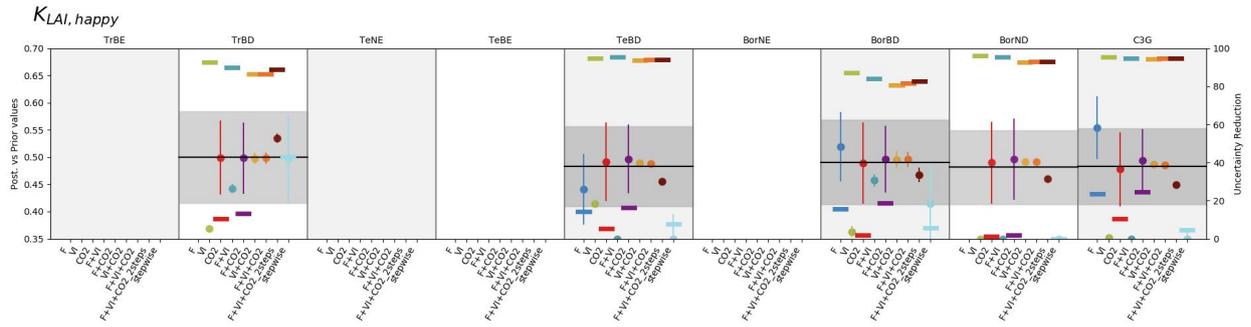
180



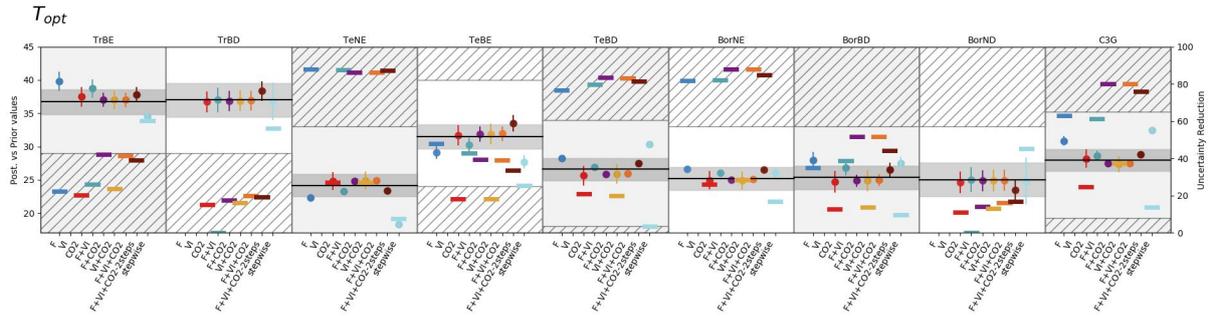
181



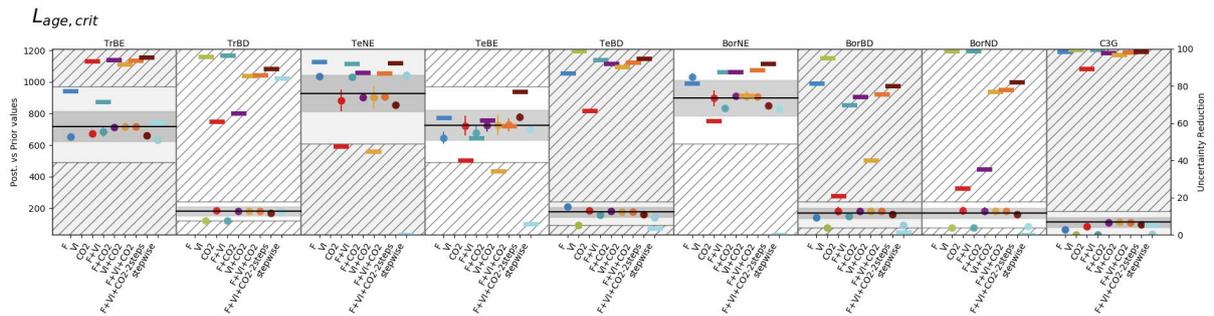
182



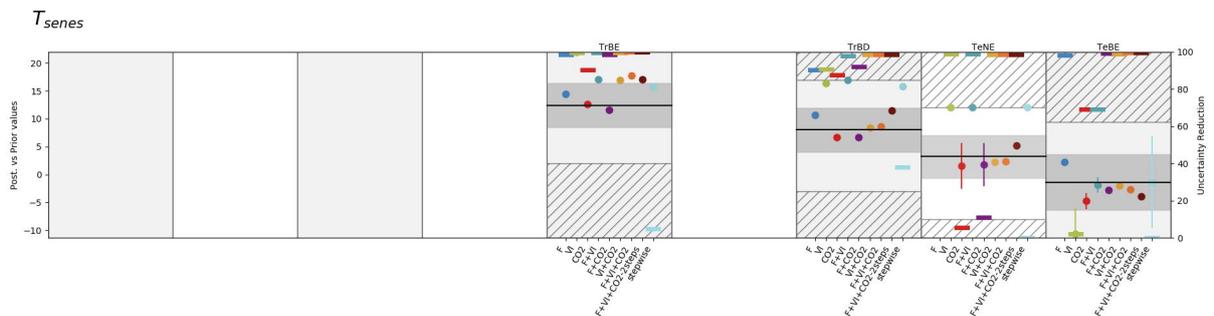
182



183



184



185

186

187 **References**

188

189 Carvalho, N., Reichstein, M., Ciais, P., Collatz, G. J., Mahecha, M. D., Montagnani, L., et al. (2010).

190 Identification of vegetation and soil carbon pools out of equilibrium in a process model via eddy
 191 covariance and biometric constraints. *Global Change Biology*, 16(10), 2813–2829.

192 Kuppel, S., Chevallier, F., & Peylin, P. (2013). *Quantifying the model structural error in carbon cycle data*
 193 *assimilation systems*, *Geosci. Model Dev.*, 6, 45–55, doi: 10.5194. gmd-6-45-2013.

- 194 Kuppel, Sylvain, Peylin, P., Chevallier, F., Bacour, C., Maignan, F., & Richardson, A. D. (2012).
195 Constraining a global ecosystem model with multi-site eddy-covariance data. *Biogeosciences*, 9(10),
196 3757–3776.
- 197 MacBean, N., Maignan, F., Peylin, P., Bacour, C., Bréon, F.-M., & Ciais, P. (2015). Using satellite data to
198 improve the leaf phenology of a global terrestrial biosphere model. *Biogeosciences*, 12(23), 7185–
199 7208.
- 200 Peylin, P., Bacour, C., MacBean, N., Leonard, S., Rayner, P., Kuppel, S., et al. (2016). A new stepwise
201 carbon cycle data assimilation system using multiple data streams to constrain the simulated land
202 surface carbon cycle. *Geoscientific Model Development (Online)*, 9(9).
- 203 Thoning, K. W., Tans, P. P., & Komhyr, W. D. (1989). Atmospheric carbon dioxide at Mauna Loa
204 Observatory: 2. Analysis of the NOAA GMCC data, 1974–1985. *Journal of Geophysical Research:*
205 *Atmospheres*, 94(D6), 8549–8565.
- 206