



*Journal of Geophysical Research Biogeosciences*

Supporting Information for

**Optimizing Phenology Parameters Drastically Improves Terrestrial Biosphere Model  
Underestimates of Dryland Net CO<sub>2</sub> Flux Inter-Annual Variability**

**K. Mahmud<sup>1</sup>, R. L. Scott<sup>2</sup>, J. A. Biederman<sup>2</sup>, M. E. Litvak<sup>3</sup>, T. Kolb<sup>4</sup>, T. P. Meyers<sup>5</sup>, P.  
Krishnan<sup>5,6</sup>, V. Bastrikov<sup>7,8</sup>, and N. MacBean<sup>1</sup>**

<sup>1</sup>Department of Geography, Indiana University, Bloomington, IN 47405, USA

<sup>2</sup>Southwest Watershed Research Center, United States Department of Agriculture, Agricultural Research Service,  
Tucson, AZ 85719, USA

<sup>3</sup>Department of Biology, University of New Mexico, Albuquerque, NM, 87131, USA

<sup>4</sup>School of Forestry, Northern Arizona University, Flagstaff, AZ, 86011, USA

<sup>5</sup>NOAA/ARL Atmospheric Turbulence and Diffusion Division, Oak Ridge, TN, 37830, USA

<sup>6</sup>Oak Ridge Associated Universities, Oak Ridge, TN, 37830, USA

<sup>7</sup>Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, F-91191, France

<sup>8</sup>Now at: Science Partners, Paris, 75010, France

**Contents of this file**

Tables S1 to S2

Figures S1 to S11

Deleted: 0

**Table S1.** Prior information for all ORCHIDEE parameters optimized in this study: prior value, uncertainty and maximum and minimum bounds for the different plant functional types (temperate needleleaf/broadleaf evergreen (TeNE, TeBE) forests, temperate broadleaf deciduous (TeBD) forest, C4 grassland (GC4)).

Parameter Name	Description (Unit)	Plant functional type			
		TeNE	TeBE	TeBD	GC4
<b>Photosynthesis parameters</b>					
ARJV	a coefficient of the linear regression (a+bT) defining the Jmax25/Vcmax25 ratio (mu mol e- (mu mol CO <sub>2</sub> )-1)	2.59 ± 0.4 2, 3	2.59 ± 0.4 2, 3	2.59 ± 0.4 2, 3	1.715 ± 0.48 1, 2.2
aSJ	a coefficient of the linear regression (a+bT) defining the Entropy term for Jmax (J K-1 mol-1)	659.7 ± 264 330, 990	659.7 ± 264 330, 990	659.7 ± 264 330, 990	630 ± 252 315, 945
aSV	a coefficient of the linear regression (a+bT) defining the Entropy term for Vcmax (J K-1 mol-1)	668.39 ± 267.6 334, 1003	668.39 ± 267.6 334, 1003	668.39 ± 267.6 334, 1003	641.64 ± 256.4 321, 962
BRJV	b coefficient of the linear regression (a+bT) defining the Jmax25/Vcmax25 ratio (mu mol e- (mu mol CO <sub>2</sub> )-1)	-0.035 ± 0.028 -0.07, 0	-0.035 ± 0.028 -0.07, 0	-0.035 ± 0.028 -0.07, 0	-0.01 ± 0.028 -0.035, 0.035
bSJ	b coefficient of the linear regression (a+bT) defining the Entropy term for Jmax (J K-1 mol-1 C-1)	-0.75 ± 0.6 -1.5, 0	-0.75 ± 0.6 -1.5, 0	-0.75 ± 0.6 -1.5, 0	0.01 ± 0.6 -0.75, 0.75
bSV	b coefficient of the linear regression (a+bT) defining the Entropy term for Vcmax (J K-1 mol-1 C-1)	-1.07 ± 0.8 -2, 0	-1.07 ± 0.8 -2, 0	-1.07 ± 0.8 -2, 0	0.1 ± 0.856 -1.07, 1.07
CN	C/N ratio	40 ± 32 20, 100	40 ± 32 20, 100	40 ± 32 20, 100	-
D_Jmax	Energy of deactivation for Jmax (J/mol)	200000 ± 16000 180000, 220000	200000 ± 16000 180000, 220000	200000 ± 16000 180000, 220000	192000 ± 15200 173000, 211000
D_Vcmax	Energy of deactivation for Vcmax (J/mol)	200000 ± 16000 180000, 220000	200000 ± 16000 180000, 220000	200000 ± 16000 180000, 220000	192000 ± 15200 173000, 211000
E_gamma_star	Energy of activation for gamma_star (J mol-1)	37830 ± 8000 27830, 47830			
E_Jmax	Energy of activation for Jmax (J mol-1)	49884 ± 8000 39884, 59884	49884 ± 8000 39884, 59884	49884 ± 8000 39884, 59884	77900 ± 8000 67900, 87900
E_KmC	Energy of activation for KmC (J mol-1)	79430 ± 8000 69430, 89430			
E_KmO	Energy of activation for KmO (J mol-1)	36380 ± 8000 26380, 46380			
fpseudo	Fraction of electrons at PSI that follow pseudocyclic transport	-	-	-	0.1 ± 0.032 0.06, 0.14
fpsir	Fraction of PSII e-transport rate partitioned to the C4 cycle	-	-	-	0.4 ± 0.16 0.4, 0.6
FRAC_GRO_WTHRESP	Fraction of GPP which is lost as growth respiration	0.28 ± 0.064 0.2, 0.36			

fQ	Fraction of electrons at reduced plastoquinone that follow the Q-cycle	-	-	-	$1 \pm 0.24$ 0.7, 1.3
gamma_star2_5	Ci-based CO <sub>2</sub> compensation point in the absence of Rd at 25C (ubar)	$42.75 \pm 8$ 22.75, 62.75	$42.75 \pm 8$ 22.75, 62.75	$42.75 \pm 8$ 22.75, 62.75	$42.75 \pm 8$ 22.75, 62.75
gbs	Bundle-sheath conductance (mol m <sup>-2</sup> s <sup>-1</sup> bar <sup>-1</sup> )	-	-	-	$0.003 \pm 0.0008$ 0.001, 0.005
HYDROL_H_UMLCSTE	Root profile (m) in empirical plant water stress function calculation	$1 \pm 1.5$ 0.25, 4	$0.8 \pm 1.12$ 0.2, 3	$0.8 \pm 1.12$ 0.2, 3	$1 \pm 1.5$ 0.25, 4
KmC25	Michaelis-Menten constant of Rubisco for CO <sub>2</sub> at 25C (ubar)	$404.9 \pm 160$ 204.9, 604.9	$404.9 \pm 160$ 204.9, 604.9	$404.9 \pm 160$ 204.9, 604.9	$650 \pm 160$ 450, 850
KmO25	Michaelis-Menten constant of Rubisco for O <sub>2</sub> at 25C (ubar)	$278400 \pm 80000$ 178400, 378400	$278400 \pm 80000$ 178400, 378400	$278400 \pm 80000$ 178400, 378400	$450000 \pm 80000$ 350000, 550000
kp	Initial carboxylation efficiency of the PEP carboxylase (mol m <sup>-2</sup> s <sup>-1</sup> bar <sup>-1</sup> )	-	-	-	$0.7 \pm 0.24$ 0.4, 1
LAI_MAX	Maximum LAI (m <sup>2</sup> /m <sup>2</sup> )	$5 \pm 2$ 3, 8	$5 \pm 2$ 3, 8	$5 \pm 2$ 3, 8	$2.5 \pm 0.8$ 4, 10
Sco25	Relative CO <sub>2</sub> /O <sub>2</sub> specificity factor for Rubisco at 25C (bar bar <sup>-1</sup> )	$2800 \pm 800$ 1800, 3800	$2800 \pm 800$ 1800, 3800	$2800 \pm 800$ 1800, 3800	$2590 \pm 800$ 1590, 3590
SLA	Specific leaf area (m <sup>2</sup> /gC)	$0.00926 \pm 0.0005$ 0.004, 0.02	$0.02 \pm 0.012$ 0.01, 0.04	$0.026 \pm 0.0148$ 0.013, 0.05	$0.026 \pm 0.0148$ 0.013, 0.05
theta	Convexity factor for response of J to irradiance	$0.7 \pm 0.18$ 0.5, 0.95			
TPHOTO_M_AX	Maximum photosynthesis temperature (deg C)	$55 \pm 4$ 50, 60			
TPHOTO_MI_N	Minimum photosynthesis temperature (deg C)	$-4 \pm 4$ -9, 1			
VCMAX25	Maximum rate of Rubisco activity-limited carboxylation at 25C (micromol/m <sup>2</sup> /s)	$35 \pm 10$ 19, 51	$45 \pm 16$ 25, 65	$55 \pm 20$ 30, 80	$70 \pm 25.6$ 38, 102
VMAX_OFF_SET	Offset (minimum relative vcmax)	$0.3 \pm 0.048$ 0.24, 36			
<b>Post C uptake parameters - autotrophic and heterotrophic respiration, C allocation, biomass and soil C turnover</b>					
HCRIT_LITTER	Scaling depth for litter humidity (m)	$0.08 \pm 0.192$ 0.02, 0.5			
KSOILC	Scalar on the active soil C pool content (to account for uncertainty in spin-up)	$1 \pm 0.6$ 0.5, 2			
MAINT_RES_P_COEFF	Coefficient to calculate maintenance respiration as a fraction of biomass	$1.4 \pm 0.84$ 0.7, 2.4			
MAINT_RES_P_SLOPE_C	Slope of maintenance respiration coefficient (1/K), constant c of aT <sup>2</sup> +bT+c, tabulated	$0.16 \pm 0.064$ 0.08, 0.24	$0.16 \pm 0.064$ 0.08, 0.24	$0.16 \pm 0.064$ 0.08, 0.24	$0.12 \pm 0.06$ 0.04, 0.2
MAX_LTOL_SR	Extrema of leaf allocation fraction	$0.5 \pm 0.08$ 0.4, 0.6			
MIN_LTOLS_R	Extrema of leaf allocation fraction	$0.2 \pm 0.08$ 0.1, 0.3			
MOIST_COE_FF_1	Coefficient to calculate moisture control for litter and soil C decomposition	$1.1 \pm 0.24$ 0.8, 1.4			
MOIST_COE_FF_2	Coefficient to calculate moisture control for litter and soil C decomposition	$2.4 \pm 0.24$ 2.1, 2.7			
MOIST_COE_FF_3	Coefficient to calculate moisture control for litter and soil C decomposition	$0.29 \pm 0.232$ 0.01, 0.59			
MOISTCON_T_MIN	Minimum soil wetness to limit the heterotrophic respiration	$0.25 \pm 0.2$ 0.1, 0.6			

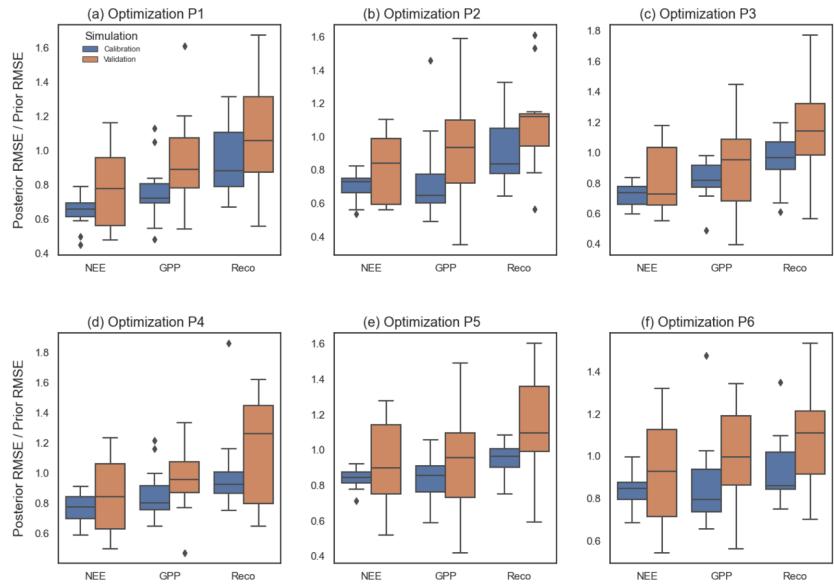
RESIDENCE TIME	Residence time of trees (years)	40 ± 24 30, 90	40 ± 24 30, 90	40 ± 24 30, 90	0 ± 0 0, 0
SOIL_Q10	Temperature dependency factor for heterotrophic respiration (Note: actual Q10 = $\exp^{SOIL\_Q10}$ )	0.69 ± 0.44 0, 1.1			
TAU_FRUIT	Fruit lifetime (days)	90 ± 24 60, 120	90 ± 24 60, 120	90 ± 24 60, 120	-
TAU_META_BOLIC	A coefficient to calculate residence times in metabolic litter pools (days)	0.066 ± 0.0112 0.052, 0.08			
TAU_SAP	Sapwood heartwood conversion time (days)	730 ± 144 550, 910	730 ± 144 550, 910	730 ± 144 550, 910	-
TAU_STRUC_T	A coefficient to calculate residence times in structural litter pools (days)	0.245 ± 0.04 0.2, 0.3			
<b>Phenology parameters</b>					
GDD_THRESHOLD	Temperature threshold used in the calculation of number of growing degree day, GDD (days)	5 ± 0.8 4, 6			
GDDNCD_CURVE	Constant in the computation of critical GDD	0.0091 ± 0.00112 0.0072, 0.01			
GDDNCD_OFFSET	Constant in the computation of critical GDD (days)	64 ± 11.2 50, 78			
GDDNCD_REFERENCE	Reference value used in the computation of critical GDD (days)	603 ± 96.8 482, 724			
HUM_FRAC	Critical humidity (relative to min/max) for phenology (%)	-	-	-	0.5 ± 0.2 0.25, 0.75
HUM_MIN_TIME	Minimum time elapsed since moisture minimum (days)	-	-	-	35 ± 12 20, 50
LAI_MAX_TO_HAPPY	Threshold of LAI below which plant uses carbohydrate reserves	0.5 ± 0.14 0.35, 0.7	0.5 ± 0.14 0.35, 0.7	0.5 ± 0.14 0.35, 0.7	0.5 ± 0.14 0.35, 0.7
LEAF_AGE_CRIT_COEF_1	A coefficient to calculate critical leaf age (days)	1.5 ± 0.24 1.2, 1.8			
LEAF_AGE_CRIT_COEF_2	A coefficient to calculate critical leaf age (days)	0.75 ± 0.12 0.6, 0.9			
LEAF_AGE_CRIT_COEF_3	A coefficient to calculate critical leaf age (days)	10 ± 1.6 12, 8			
LEAF_AGE_CRIT_TREF	Reference temperature used to calculate of critical leaf age (days)	20 ± 4 15, 25			
LEAFAGE_FIRSTMAX	Leaf age at which vmax attains vcmax_opt (in fraction of critical leaf age)	0.03 ± 0.0048 0.024, 0.036			
LEAFAGE_LASTMAX	Leaf age at which vmax falls below vcmax_opt (in fraction of critical leaf age)	0.5 ± 0.08 0.4, 0.6			
LEAFAGE_CRITICAL	Critical leaf age, tabulated (days)	910 ± 200 610, 1210	730 ± 192 490, 970	180 ± 60 120, 240	120 ± 60 30, 180
LEAFFALL	Length of death of leaves, tabulated (days)	-	-	10 ± 4 5, 15	10 ± 4 5, 15
LEAFLIFE_TAB	Leaf longevity (years)	0.33 ± 0.1 0.2, 0.75	1 ± 0.668 0.33, 2	2 ± 0.9 0.75, 3	2 ± 0.9 0.75, 3
MAX_TURN_OVER_TIME	Maximum turnover time for grass (days)	-	-	-	80 ± 4 75, 85
MIN_GROWTHINIT_TIME	Minimum time since last beginning of a growing season (days)	300 ± 24 270, 330			
MIN_LEAF_AGE_FOR_SENESCENCE	minimum leaf age to allow senescence (days)	-	-	90 ± 8 80, 100	30 ± 4 25, 35
MIN_TURN_OVER_TIME	Minimum turnover time for grass (days)	-	-	-	10 ± 4 5, 15

NCD_MAX_YEAR	A coefficient to calculate maximum possible number of chilling days (NCD)	3 ± 0.8 2, 4				
NCDGDD_T_EMP	Critical temperature for the ncd vs. gdd function in phenology (C)	-	-	5 ± 4 0, 10	-	
NOSENESCE_NCE_HUM	Relative moisture availability above which there is no humidity-related senescence	-	-	-	0.3 ± 0.12 0.15, 0.45	
PHENO_GD_D_CRIT_A	Critical gdd tabulated constant a	-	-	-	0 ± 0 0, 0	
PHENO_GD_D_CRIT_B	Critical gdd constant b	-	-	-	0 ± 0 0, 0	
PHENO_GD_D_CRIT_C	Critical gdd constant c	-	-	-	400 ± 64 320, 480	
PHENO_MOI_GDD_T_CRI_T	Average temperature threshold for C4 grass used (C)	-	-	-	22 ± 8 12, 32	
SENECENC_E_HUM	Critical relative moisture availability for senescence	-	-	-	0.2 ± 0.08 0.1, 0.3	
SENECENC_E_TEMP_A	Critical temperature for senescence (C), constant a of aT^2+bT+c, tabulated	-	-	0 ± 0 0, 0	0 ± 0 0, 0	
SENECENC_E_TEMP_B	Critical temperature for senescence (C), constant b of aT^2+bT+c, tabulated	-	-	0 ± 0 0, 0	0 ± 0 0, 0	
SENECENC_E_TEMP_C	Critical temperature for senescence (C), constant c of aT^2+bT+c, tabulated	-	-	12 ± 8 2, 22	5 ± 4.8 -1, 11	
TAU_CLIMA_TOLOGY	tau for climatologic variables (days)	20 ± 8 10, 30				
TAU_GDD	Time scales for phenology and other processes (days)	40 ± 16 20, 60				
TAU_GPP_WEEK	Time scales for phenology and other processes (days)	6 ± 1 5, 7				
TAU_HUM_MONTH	Time scales for phenology and other processes (days)	20 ± 8 10, 30				
TAU_HUM_WEEK	Time scales for phenology and other processes (days)	6 ± 1 5, 7				
TAU_LEAFI_NIT	Time to attain the initial foliage using the carbohydrate reserve (days)	10 ± 10 5, 30	10 ± 10 5, 30	10 ± 10 5, 30	10 ± 10 5, 30	
TAU_NGD	Time scales for phenology and other processes (days)	50 ± 20 25, 75				
TAU_SOILHUM_MONTH	Time scales for phenology and other processes (days)	20 ± 8 10, 30				
TAU_T2M_MONTH	Time constant for the “monthly” 2-meter temperature (days)	20 ± 8 10, 30				
TAU_T2M_WEEK	Time constant for the “weekly” 2-meter temperature (days)	6 ± 1 5, 7				
TAU_TSOL_MONTH	Time constant for the “monthly” soil temperature (days)	20 ± 8 10, 30				
<b>Conductance parameters - included in initial optimization sensitivity test but not in final optimizations</b>						
A1	empirical factor involved in the calculation of fvpd	0.85 ± 0.04 0.8, 0.9	0.85 ± 0.04 0.8, 0.9	0.85 ± 0.04 0.8, 0.9	0.85 ± 0.04 0.8, 0.9	
B1	empirical factor involved in the calculation of fvpd	0.14 ± 0.032 0.1, 0.18	0.14 ± 0.032 0.1, 0.18	0.14 ± 0.032 0.1, 0.18	0.2 ± 0.032 0.15, 0.25	
CHOISNEL_RSOL_CSTE	Constant in the computation of resistance for bare soil evaporation (s/m <sup>2</sup> )	3.3E4 ± 19400 1.75E4, 6.6E4				
CONDVEG_Z0	Surface roughness (m)	0.15 ± 0.12 0, 0.3				
DEFC_MUL_T	Constant in the computation of surface resistance (KW <sup>-1</sup> )	1.5 ± 0.9 0.75, 3				
DEFC_PLUS	Constant in the computation of surface resistance (KW <sup>-1</sup> )	0.023 ± 0.016 0.003, 0.043				

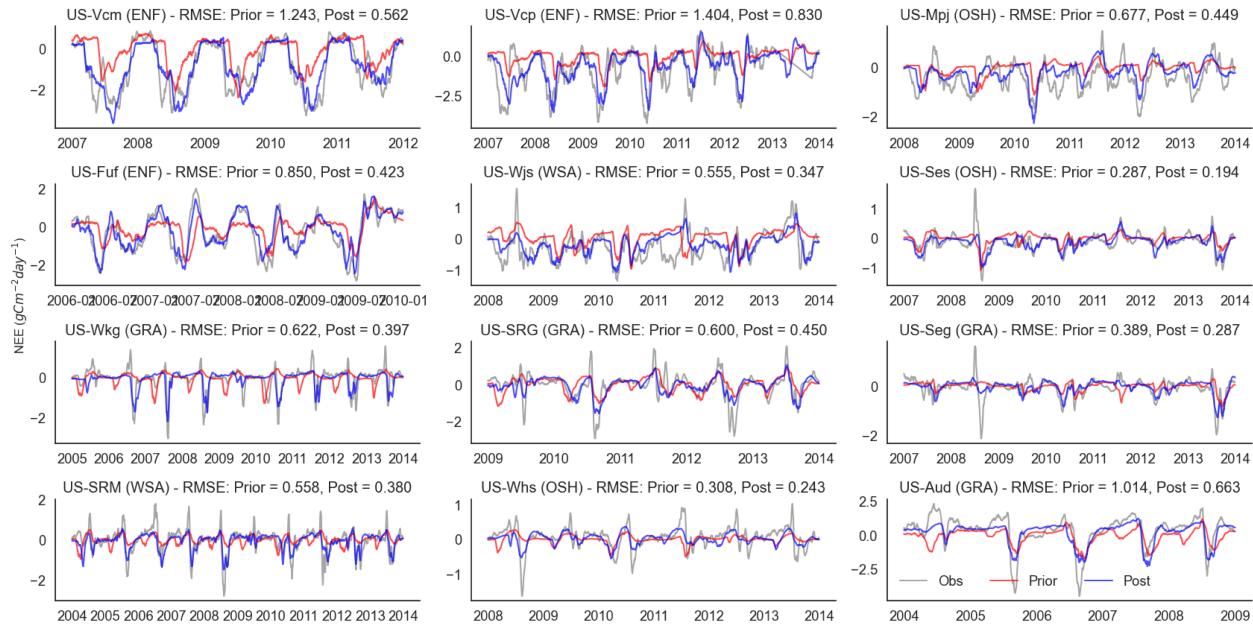
g0	Residual stomatal conductance when irradiance approaches zero (mol m <sup>-2</sup> s <sup>-1</sup> bar <sup>-1</sup> )	0.00625 ± 0.00048 0.00565, 0.00685	0.00625 ± 0.00048 0.00565, 0.00685	0.00625 ± 0.00048 0.00565, 0.00685	0.01875 ± 0.0016 0.01675, 0.02075
GB_REF	Leaf bulk boundary layer resistance (s m <sup>-1</sup> )	0.04 ± 0.032 0, 0.08			
KZERO	A vegetation dependent constant used in the calculation of the surface resistance (kg/m <sup>2</sup> /s)	0.00012 ± 0.000016 0.0001, 0.00014	0.00012 ± 0.000016 0.0001, 0.00014	0.00025 ± 0.00004 0.0002, 0.0003	0.0003 ± 0.00004 0.00025, 0.00035
RATIO_Z0M_Z0H	Ratio between z0m and z0h	1 ± 0.4 0.5, 1.5	1 ± 0.4 0.5, 1.5	1 ± 0.4 0.5, 1.5	1 ± 0.4 0.5, 1.5
Z_DECOMP	Scaling depth for soil activity	0.2 ± 0.6 0, 1.5			
Z0_BARE	Bare soil roughness length (m)	0.01 ± 0.0016 0.008, 0.012			
Z0_OVER_H_EIGHT	To get z0 from height	0.0625 ± 0.032 0.02, 0.1			

**Table S2.** Daily GPP and ecosystem respiration ( $R_{eco}$ ) model-data fit when assimilating NEE observations with all parameters (P1) in terms of prior and posterior root mean square error (RMSE) for all twelve sites. The reduction in daily GPP RMSE varies between 0.03 to 0.7 gCm<sup>-2</sup>d<sup>-1</sup>, and the reduction in daily  $R_{eco}$  RMSE varies between 0 to 0.7 gCm<sup>-2</sup>d<sup>-1</sup>. The sites are listed in order from largest mean annual C sink (US-Vcm) to mean annual C source (US-Aud).

Site	Daily root mean square error (RMSE)			
	GPP		$R_{eco}$	
	Prior	Posterior	Prior	Posterior
US-Vcm	1.392	1.366	2.009	1.315
US-Vcp	1.822	1.116	0.979	0.927
US-Mpj	1.312	0.903	1.051	0.867
US-Fuf	1.074	0.828	0.534	0.629
US-Wjs	0.995	0.73	0.778	0.672
US-Ses	0.259	0.229	0.233	0.255
US-Wkg	1.177	0.809	0.617	0.49
US-SRG	1.345	0.982	0.95	0.825
US-Seg	0.699	0.449	0.441	0.324
US-SRM	1.158	0.617	0.8	0.572
US-Whs	0.698	0.531	0.507	0.437
US-Aud	1.028	0.707	0.561	0.602

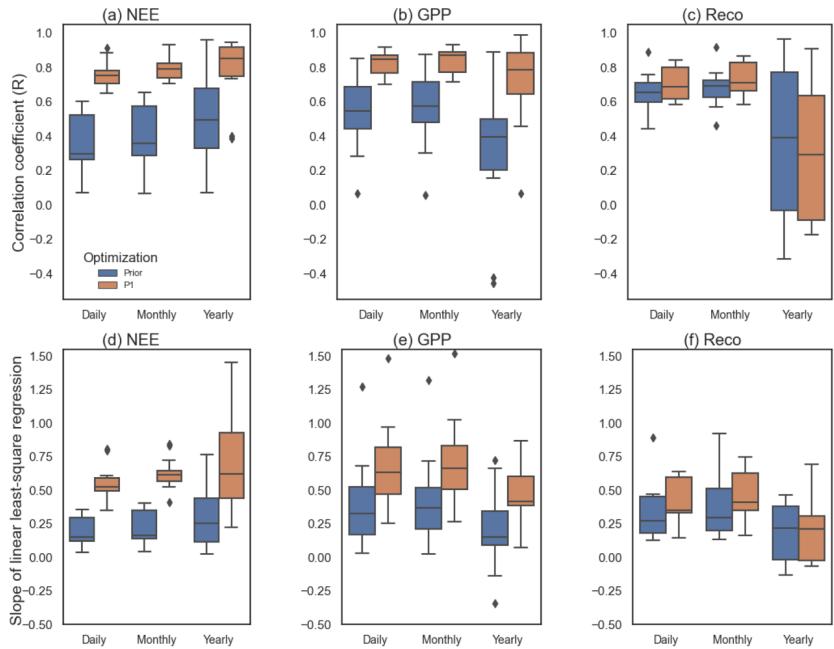


**Figure S1.** Comparison between the ratios of posterior RMSE to prior RMSE of daily flux data for both calibration and validation exercise. The last year of flux time series for each site is used for the validation purpose. The box-whiskers indicate the spread of RMSE ratios due to all 12 sites.



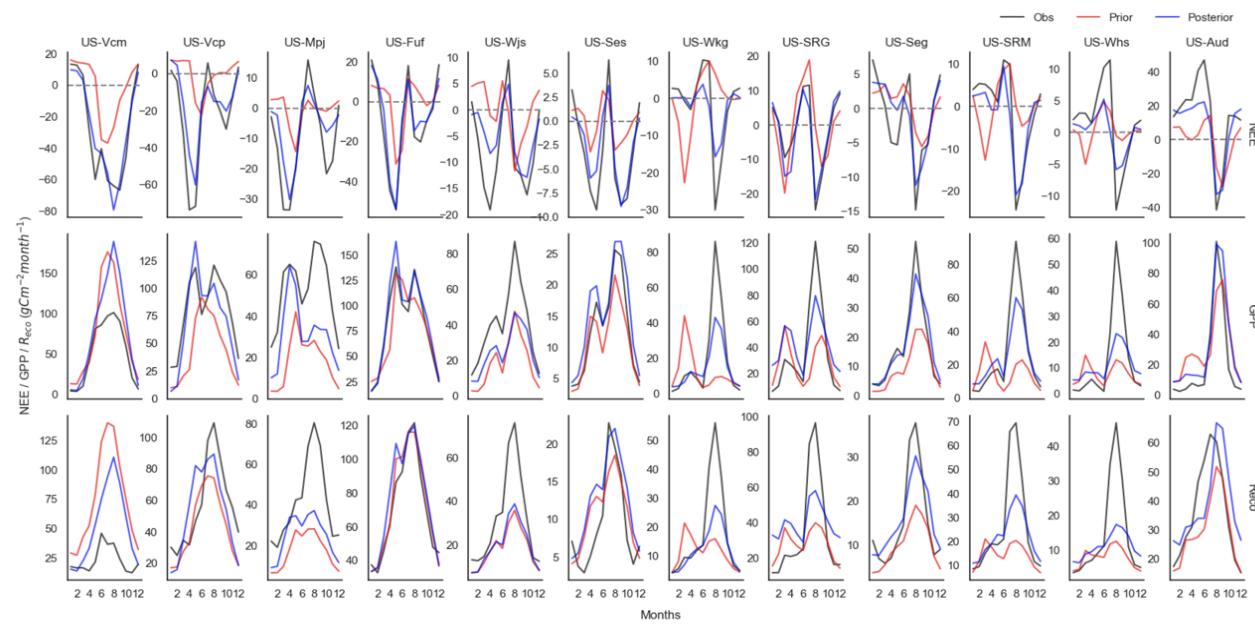
**Figure S2.** Comparison of NEE observations (grey) with corresponding ORCHIDEE model simulations before (red line) and after assimilation (blue line) for assimilating NEE observations with all parameters (P1). The vegetation types are listed within brackets for each site. The RMSE measures the fit of the model prior and posterior simulations with the corresponding observations. Across all sites, the prior and posterior NEE RMSEs vary between 0.291-1.377 and 0.196-0.788, respectively. The sites are listed in order from largest mean annual C sink (US-Vcm) to mean annual C source (US-Aud).

Deleted: 1  
Deleted: g  
Deleted: re



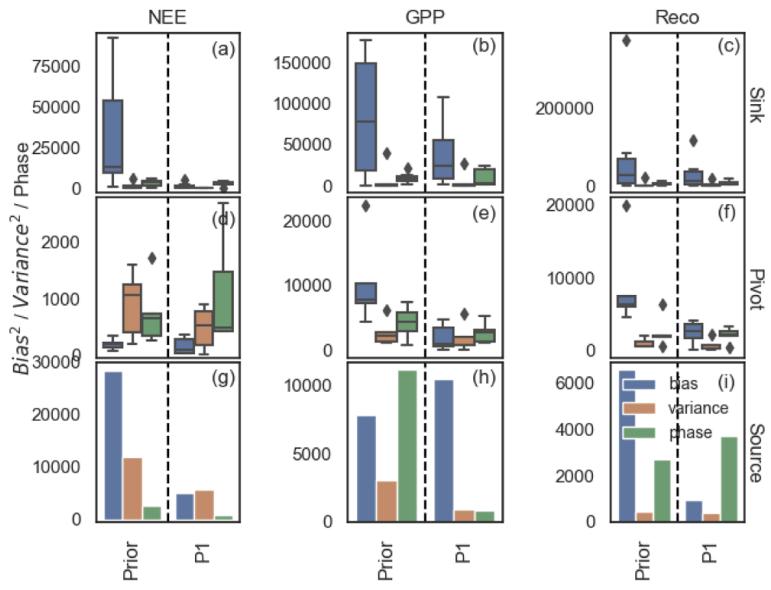
**Figure S3.** Daily, monthly and annual NEE (a, d), GPP (b, e) and  $R_{\text{eco}}$  (c, f) prior and posterior Pearson correlation coefficients ( $R$ ) and slope values for the linear regression between model and observed fluxes for assimilating NEE observations and optimizing all phenology, photosynthesis and post C uptake parameters (P1). The  $R$  between observed and modeled NEE at daily, monthly and annual timescales for optimizing all parameters (P1) increase by up to 0.50, 0.55, 0.65 respectively. Note that the y axis limits for both  $R$  and slope are the same and therefore 3 sites fall outside the y-axis upper limit for the  $R_{\text{eco}}$  slope.

Deleted: 2



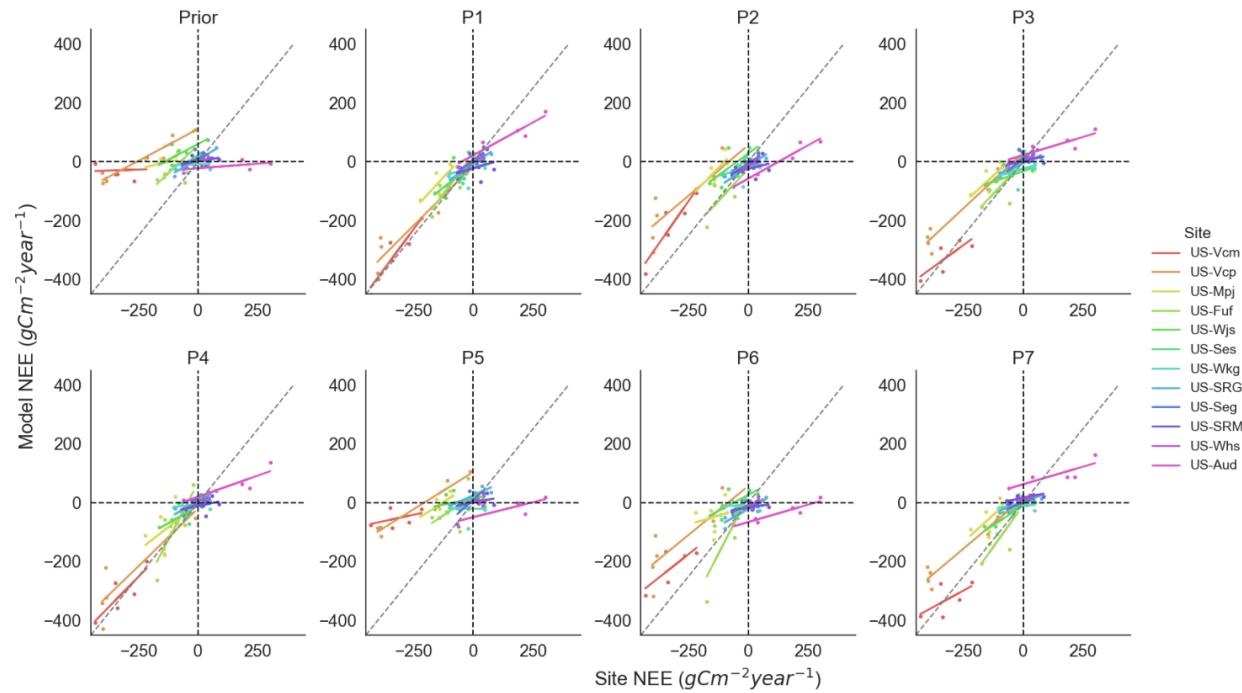
**Figure S4.** Seasonal cycle with mean monthly total fluxes. Comparison of flux observations with corresponding ORCHIDEE model simulations (prior and posterior) for assimilating NEE observations and optimizing all phenology, photosynthesis and post C uptake parameters (P1). The sites are listed from left to right according to C sink to source.

Deleted: 3



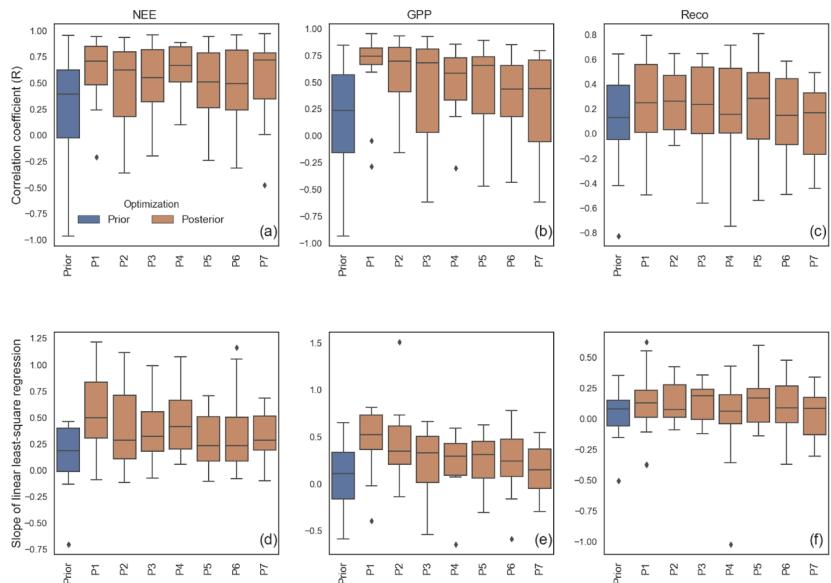
**Figure S5.** Annual NEE, GPP and  $R_{\text{eco}}$  mean square deviation (MSD) decomposition into bias, variance, and phase between simulations and observations for assimilating NEE observations and optimizing all phenology, photosynthesis and post C uptake parameters (P1). Different rows separate the sites as sink (a-c), pivot (d-f) and source (g-i) based on total annual C flux. The sink sites are: US-Vcm, US-Vcp, US-Mpj, US-Fuf, US-Wjs and US-Ses; the pivot sites are: US-Wkg, US-SRG, US-Seg, US-SRM and US-Whs; and the source site ia: US-Aud. The x axes display the optimization scenarios (Prior and P1). The box whiskers show the spread of bias, variance and phase for all 12 sites considered in this study. The bias, variance and phase indicate the mean difference in flux magnitude, the mismatch in terms of flux fluctuation magnitude scales with the mean seasonal amplitude, and the seasonality in flux time series, respectively. Note that the y axis limits are different for all fluxes and site types.

Deleted: 4



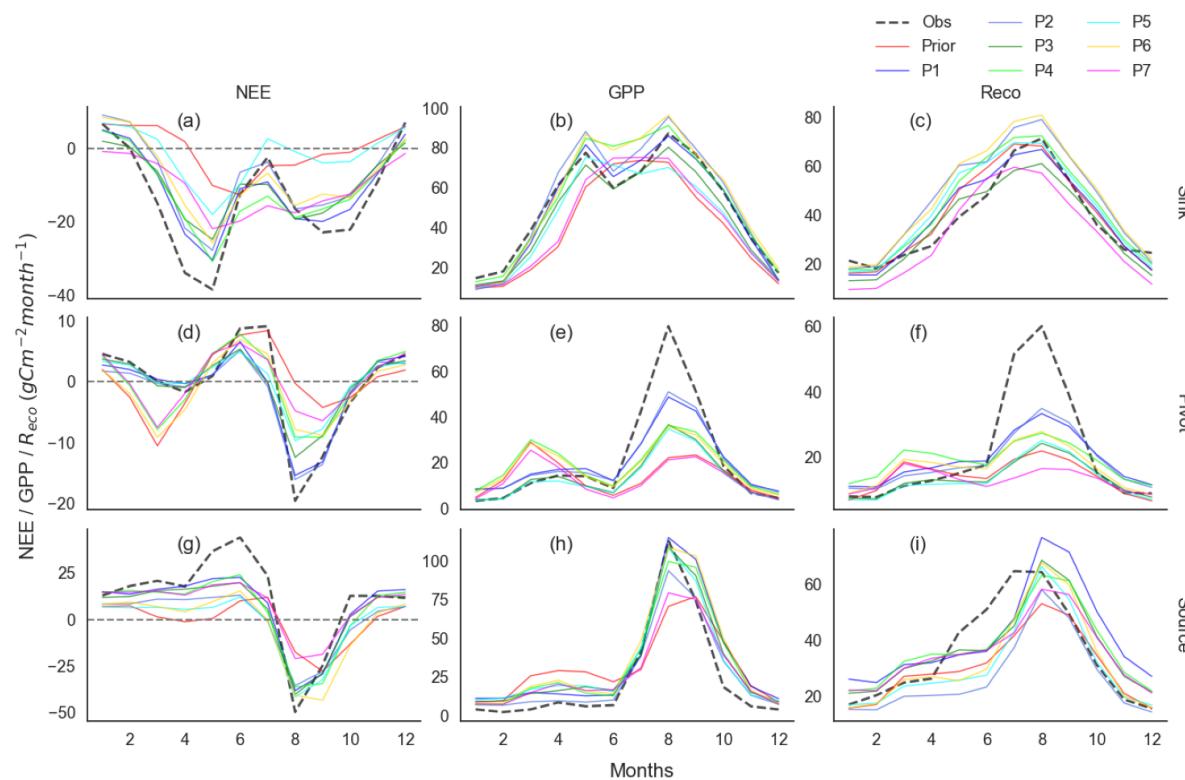
**Figure S6.** Annual NEE scatter plots for prior and all posterior simulations for assimilating NEE observations with various parameter sets (P1–P7). Different colour legends represent various sites, ordered from the largest mean sink (US-Vcm) to the largest mean source (US-Aud). The middle of the trend line should sit on the 1:1 line if the accurate mean annual source/sink behavior for a site is well captured by the model. A slope value close to or equal to 1 demonstrates the model is better at capturing the IAV. The sink sites are: US-Vcm, US-Vcp, US-Mpj, US-Fuf, US-Wjs and US-Ses; the pivot sites are: US-Wkg, US-SRG, US-Seg, US-SRM and US-Whs; and the source site is: US-Aud.

Deleted: 5



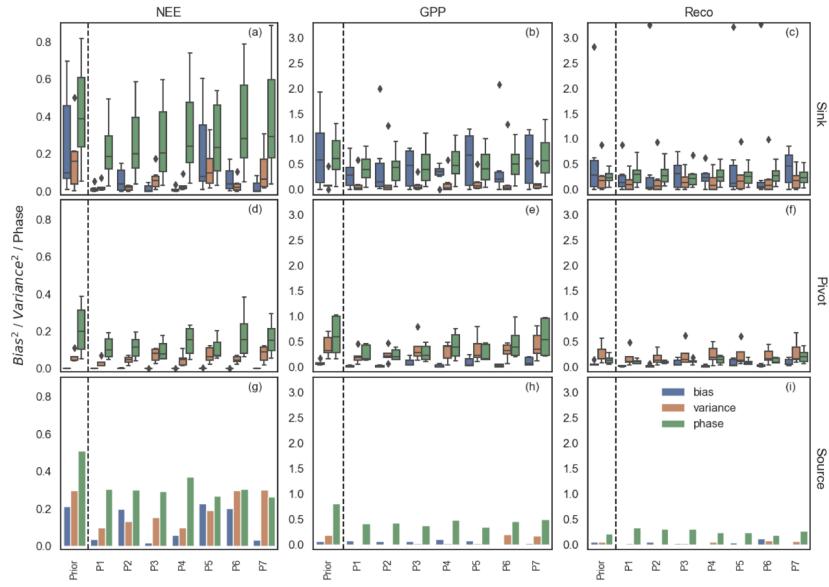
**Figure S7.** NEE (a, d), GPP (b, e) and  $R_{\text{eco}}$  (c, f) annual anomaly prior and posterior Pearson correlation coefficients (R) and slope values for the linear regression between model and observed fluxes across all assimilation scenarios with different parameter combinations (P1-P7). The legend represents various assimilation scenarios (Prior - blue bars, and posterior P1-P7 - orange bars).

Deleted: 6



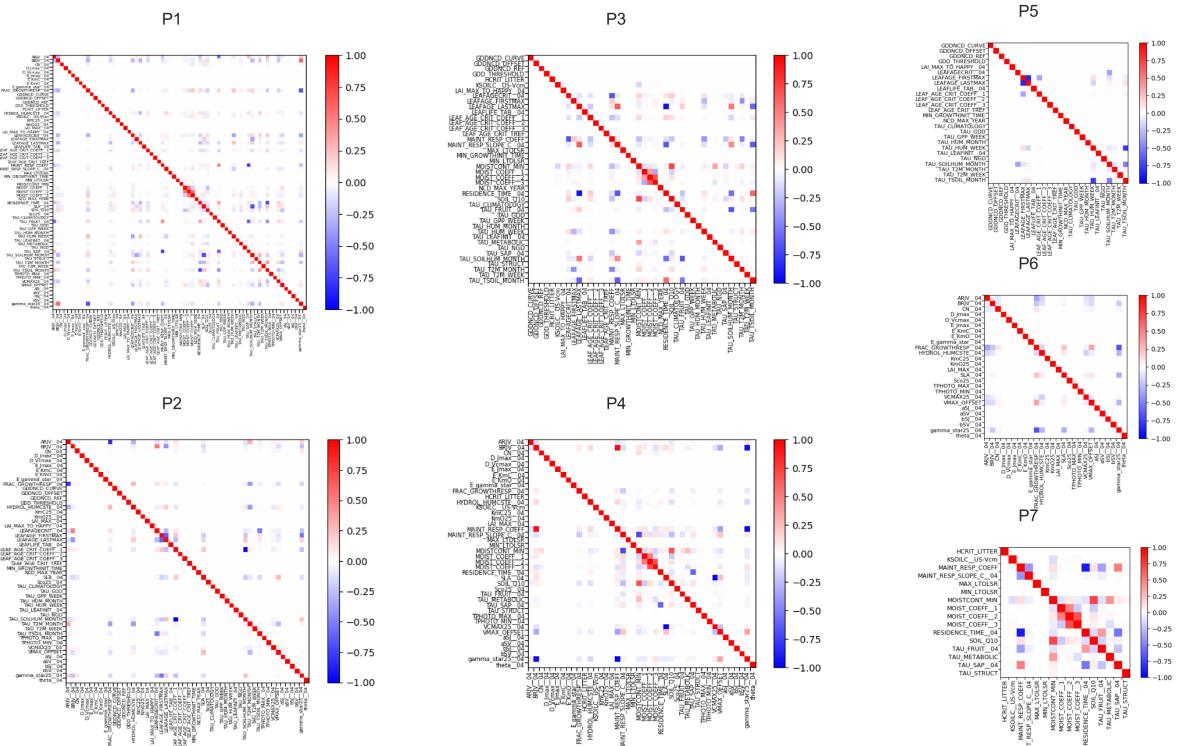
**Figure S8.** Mean monthly seasonal cycles comparing observations (thick black dashed curve), prior (red curve), and posterior simulations for assimilation scenarios (P1 to P7 - blue to magenta curves) for NEE (left column), GPP (middle column), and  $R_{\text{eco}}$  (right column) averaged across site C balance types (sink - top row; pivot - middle row; and the source site, US-Aud, on the bottom row).

Deleted: 7



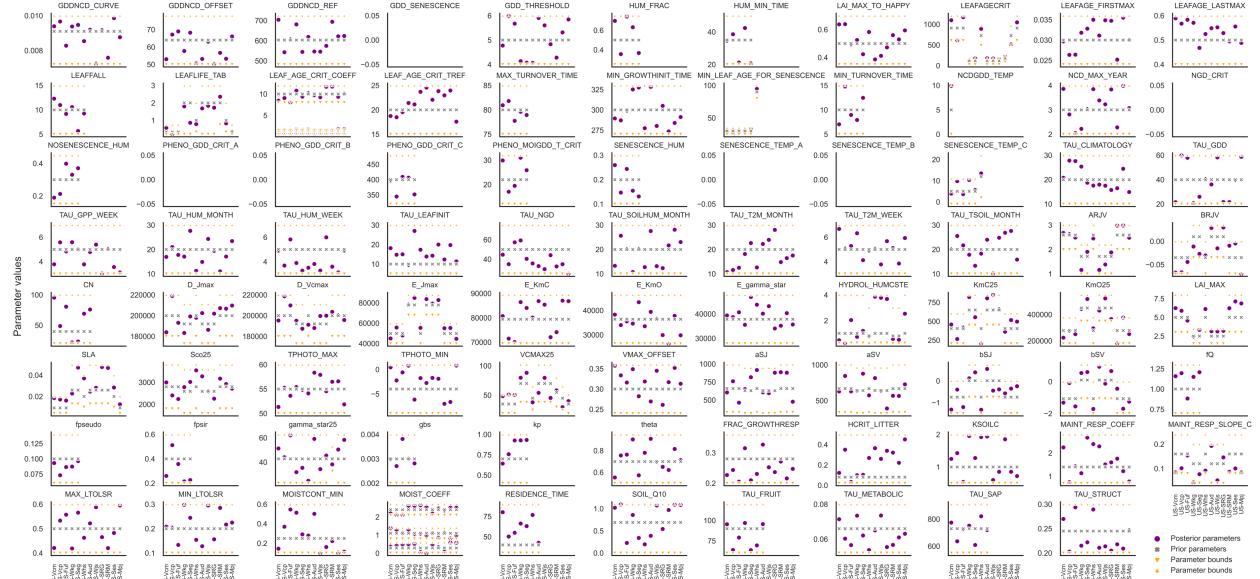
**Figure S9.** Daily NEE, GPP and  $R_{\text{eco}}$  mean square deviation (MSD) decomposition into bias, variance, and phase between simulations and observations for assimilating NEE observations with various parameter sets (P1-P7). Different rows separate the sites as sink (a-c), pivot (d-f) and source (g-i) based on total annual C flux. The sink sites are: US-Vcm, US-Vcp, US-Mpj, US-Fuf, US-Wjs and US-Ses; the pivot sites are: US-Wkg, US-SRG, US-Seg, US-SRM and US-Whs; and the only source site is: US-Aud. The x axes display various optimization scenarios (Prior, P1-P7). The parameters included in each optimization are: P1: all parameters; P2: phenology and photosynthesis; P3: phenology and post C uptake; P4: photosynthesis and post C uptake; P5: phenology; P6: photosynthesis and P7: post C uptake. The box whiskers show the spread of bias, variance and phase for all 12 sites considered in this study. The bias, variance and phase indicate the mean difference in flux magnitude, the mismatch in terms of flux fluctuation magnitude scales with the mean seasonal amplitude, and the seasonality in flux time series, respectively. Note that the y axis limits for both gross fluxes are the same.

Deleted: 8



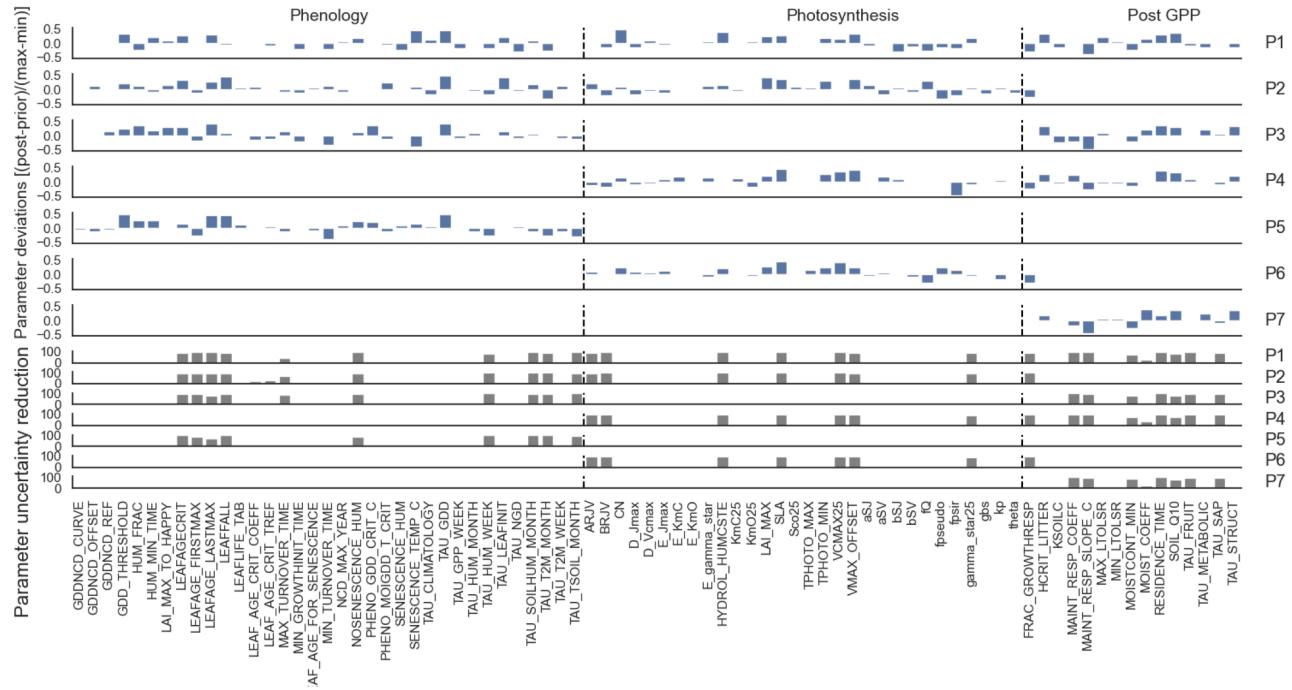
**Figure S10.** Parameter posterior error covariance matrix for US-Vcm for various assimilation scenarios (P1-P7).

**Deleted: 9**



**Figure S11.** Values of all optimized parameters related to phenology, photosynthesis and post C uptake when assimilating NEE and optimizing all parameters (P1) for all 12 sites. For each parameter, the range of variation (corresponding to yellow arrows), the prior and the posterior values are provided for all sites. For the mixed-PFT sites, only the parameters for the majority PFT fraction are presented, although the other PFT parameters are also optimized. For example, this figure shows the parameters associated with PFT=4 (TeNE) for site US-Mpj, however the optimization is performed with all the parameters of both PFT=4 (TeNE) and 11 (C4G). Note that the soil Q10 parameter is the exponent of the actual Q10 value used to calculate heterotrophic respiration temperature sensitivity (see **Table S1** for more information).

Deleted: 0



**Figure S12.** Optimized median parameter deviations  $[(\text{posterior} - \text{prior}) / (\text{max} - \text{min})]$  (blue bars) and associated median parameter uncertainty reductions (grey bars) for all parameters controlling phenology, photosynthesis and post C uptake assimilating NEE data (P1-P7). Bars represent the median across all *sink* sites.

Deleted: 1