Supplement of Geosci. Model Dev., 15, 6637–6657, 2022
https://doi.org/10.5194/gmd-15-6637-2022-supplement

Geoscientific Model Development
Open Access

Supplement of

Pixel-level parameter optimization of a terrestrial biosphere model for improving estimation of carbon fluxes with an efficient model–data fusion method and satellite-derived LAI and GPP data

Rui Ma et al.

Correspondence to: Rui Ma (marui6922@whu.edu.cn) and Shunlin Liang (shunlin@hku.hk)

Table S1 Site information

No.	Site	Latitude	Longitude	PFT	Year_Start	Year_End	Reference
1	US-Bar^{A}	44.06	-71.29	DBF	2004	2005	Ouimette et al. (2018)
2	US-Cwt^{A}	35.06	-83.43	DBF	2011	2015	Oishi et al. (2020)
3	US-Dk2^{A}	35.97	-79.10	DBF	2003	2005	Oishi et al. (2018)
4	US-MOz^{A}	38.74	-92.20	DBF	2004	2017	Gu et al. (2016)
5	US-LPH^{A}	42.54	-72.19	DBF	2002	2004	Hadley (2016)
6	US-Wi1^{A}	46.73	-91.23	DBF	2003		Chen et al. (2020a)
7	US-Wi8^{A}	46.72	-91.25	DBF	2002		Chen et al. (2020c)
8	US-Ha1	42.54	-72.17	DBF	2000	2016	Urbanski et al. (2007)
9	US-MMS	39.32	-86.41	DBF	2000	2014	Schmid et al. (2000)
10	US-Oho	41.55	-83.84	DBF	2004	2013	Noormets et al. (2008)
11	US-UMB	45.56	-84.71	DBF	2000	2014	Gough et al. (2013)
12	US-UMd	45.56	-84.70	DBF	2007	2014	Gough et al. (2013)
13	US-WCr	45.81	-90.08	DBF	2000	2014	Cook et al. (2004)
14	US-Wi3	46.63	-91.10	DBF	2002	2004	Chen et al. (2020b)

DBF, deciduous broadleaf forests; The superscript 'A' means sites from AmeriFlux network.

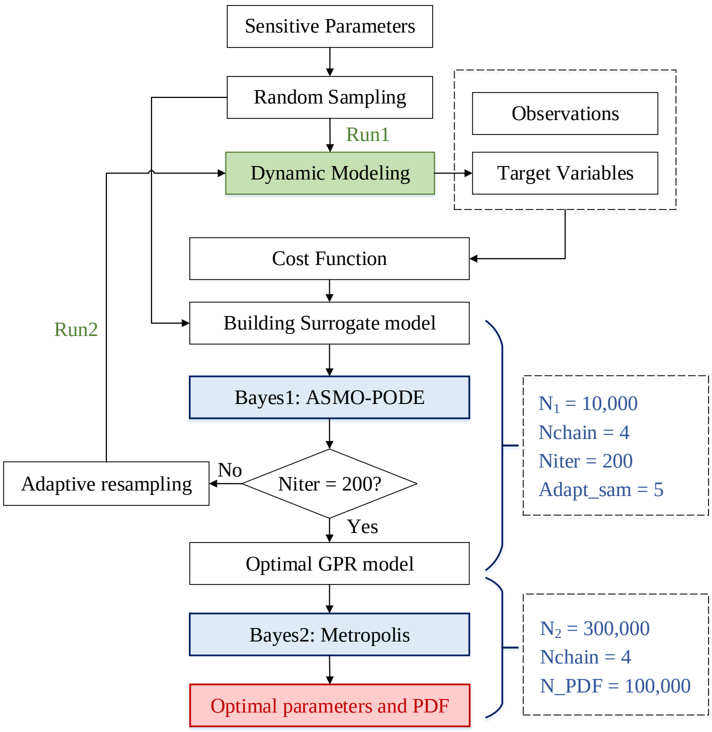

Figure S1. Framework of modified adaptive surrogate modeling (MASM). N_1 and N_2 mean length of a Markov chain in the two steps respectively. Nchain represents the number of Markov chains, Niter is the number of iterations, and Adapt_sam is the number of parameter samples in each adaptive resampling iterations. N_PDF means using the last 100,000 combinations to calculate the posterior parameter distributions of all sensitive parameters.

Table. S2 Environmental variables for building XGBoost model. Variables with 'C' means constant values; 'Y' represnets using annual mean values from all years; 'GS' represents using growing season mean values from all years. 'fraction' refers to the percentage of deciduous forests in each grid based on land cover data from NLCF (30m).

Types	Variable name	Source	C	Y	GS	Types	Variable name	Source	C	Y	GS
Climate	temperature (T)	Daymet		√	√	Soil	sand content	GSDE	√		
	precipitation (prec)			√			clay content		√		
	vapor pressure deficit (vpd)			√	√	Grid	latitude (lat)	-	√		
	daylength (dayl)			√			longitude (lon)	-	√		
	shortwave radiation (srad)			√	√		elevation (dem)		√		
	pressure (press)	NLDAS		√			fraction (frac)	NLCF	√		
	wind speed (vs)	GRIDMET		√		Vegetation variables	GPP	GLASS		√	
	specific humidity (sph)			√			LAI			√	
	cloud cover (cld)	NARR		√			photosynthetically active radiation (PAR)			√	√

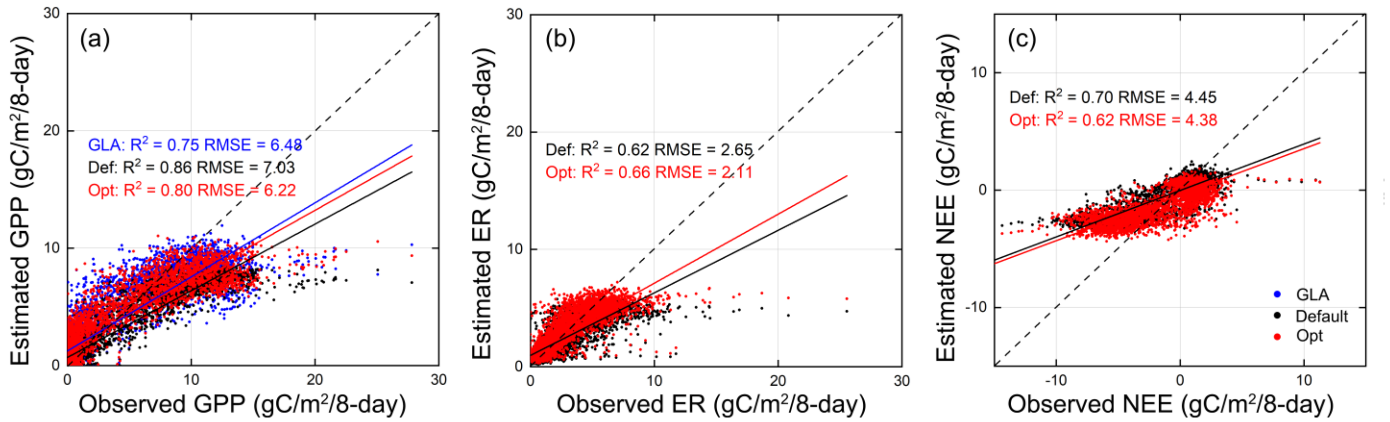

Figure S2. Scatter plots of observed 8-day carbon fluxes versus estimated 8-day carbon fluxes before and after optimization. 'GLA' represents GLASS GPP, 'Default' represents carbon fluxes estimated from previous model with default parameters, 'Opt' represents values from the optimized model with calibrated parameters.

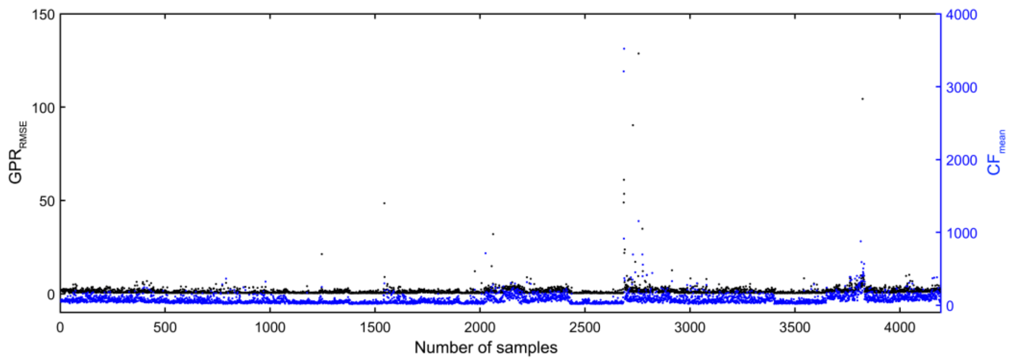

Figure S3. Errors of all the samples when running MASM method.

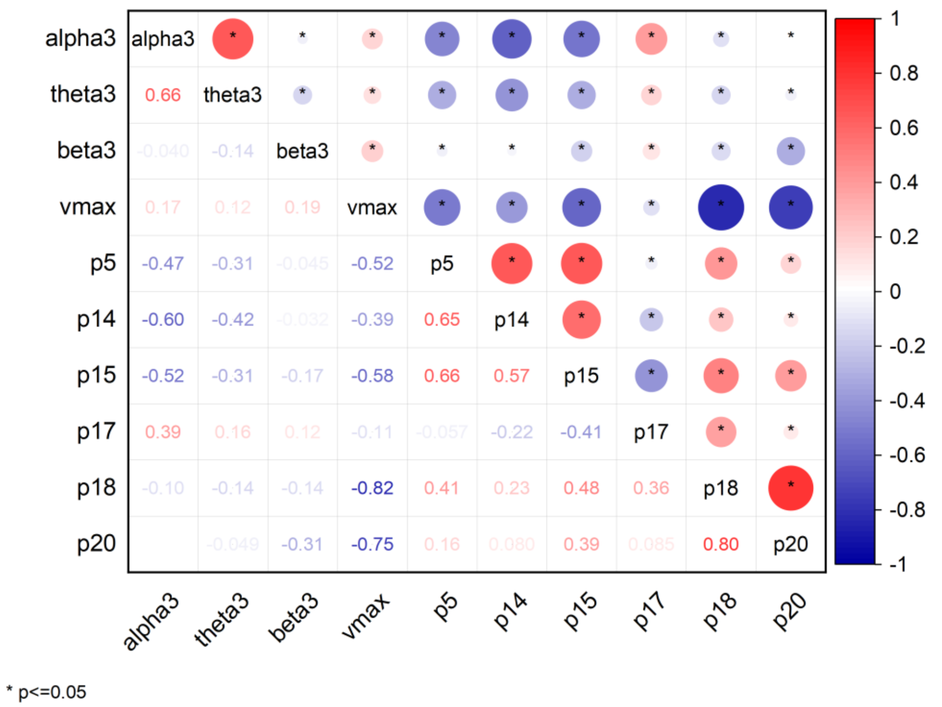

Figure S4. The 2D-correlation map between optimal parameters for the deciduous forests in the EUS. The upper triangle uses the circle size and color to indicate the degree of correlation between two parameters. Higher correlation is expressed as a darker and larger circle, with blue indicating negative correlation and red indicating positive correlation. The lower triangle shows the values of the correlation coefficient, and colors have the same meaning as the upper triangle. Asterisk indicates a significant correlation (p<=0.05).

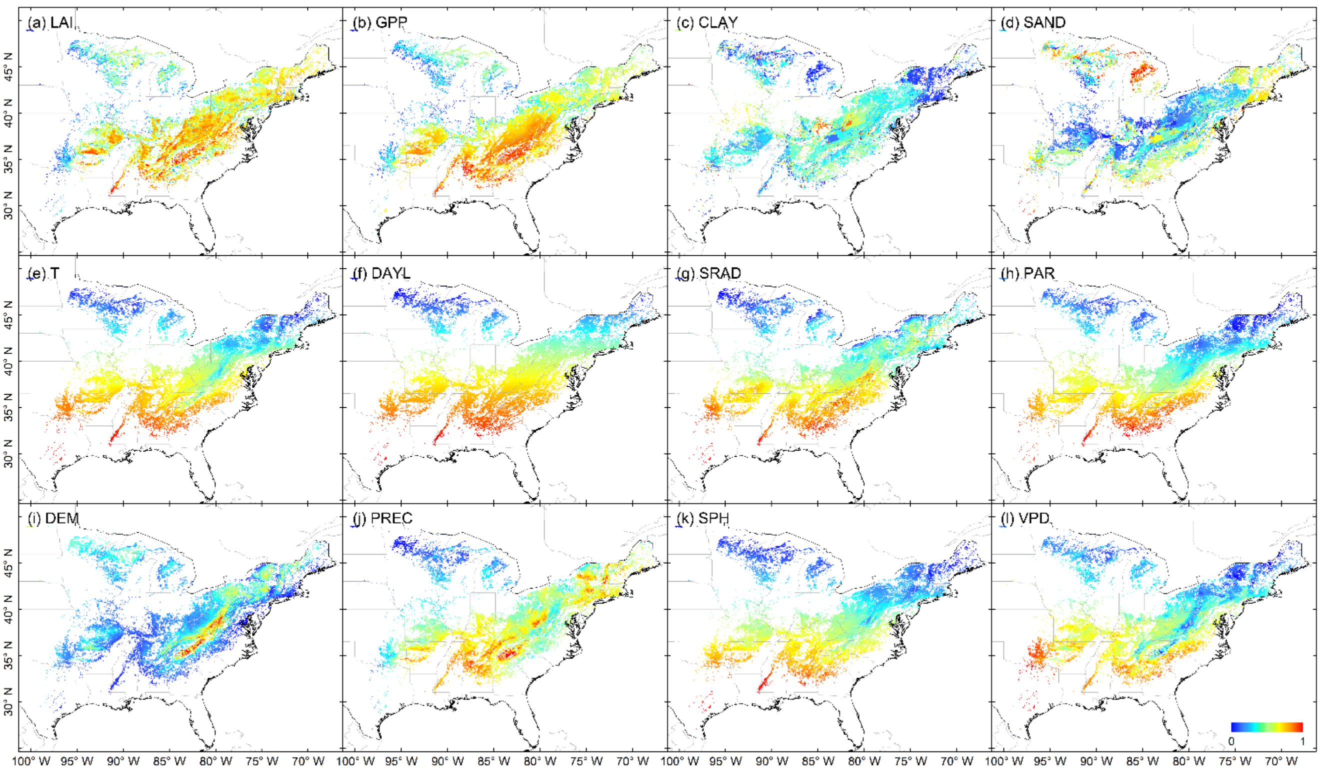

Figure S5. Spatial distribution of several key environmental variables used in the XGBoost approach. All variables have been normalized to 0-1. The explanation of variable abbreviations can be found in Table S2.

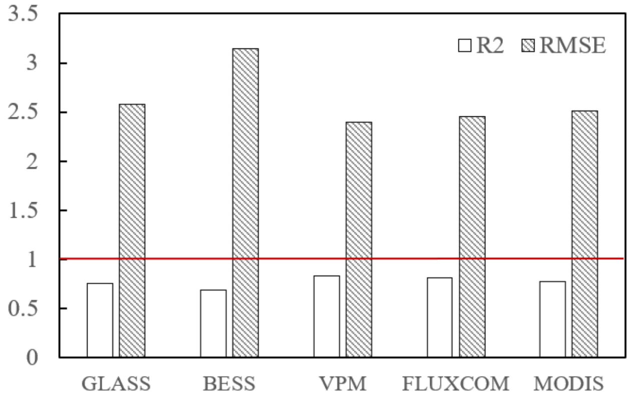

Figure S6. Accuracy comparison among different GPP products (8-day, 0.05°) for 14 flux sites from 2001 to 2015. BESS: GPP derived from Breathing Earth System Simulator (BESS) (Jiang and Ryu, et al); VPM: GPP derived from the Vegetation Photosynthesis Model (Zhang et al., 2017); FLUXCOM: GPP derived from the upscaling approaches based on machine learning methods (Jung et al., 2011); MODIS: GPP derived from the MOD17A2H V6 product (Running et al., 2004).

References:

Chen, J.: AmeriFlux BASE US-Wi1 Intermediate hardwood (IHW), Ver. 3-5, AmeriFlux AMP [dataset], https://doi.org/10.17190/AMF/1246015, 2020a.

Chen, J.: AmeriFlux BASE US-Wi3 Mature hardwood (MHW), Ver. 3-5, AmeriFlux AMP [dataset], https://doi.org/10.17190/AMF/1246018, 2020b.

Chen, J.: AmeriFlux BASE US-Wi8 Young hardwood clearcut (YHW), Ver. 3-5, AmeriFlux AMP [dataset], https://doi.org/10.17190/AMF/1246023, 2020c.

Cook, B. D., Davis, K. J., Wang, W., Desai, A., Berger, B. W., Teclaw, R. M., Martin, J. G., Bolstad, P. V., Bakwin, P. S., Yi, C., and Heilman, W.: Carbon exchange and venting anomalies in an upland deciduous forest in northern Wisconsin, USA, Agr. Forest Meteorol., 126, 271-295, 10.1016/j.agrformet. 2004.06.008, 2004.

Gough, C. M., Hardiman, B. S., Nave, L. E., Bohrer, G., Maurer, K. D., Vogel, C. S., Nadelhoffer, K. J., and Curtis, P. S.: Sustained carbon uptake and storage following moderate disturbance in a Great Lakes forest, Ecol. Appl., 2013.

Gu, L., Pallardy, S. G., Yang, B., Hosman, K. P., Mao, J., Ricciuto, D., Shi, X., and Sun, Y.: Testing a land model in ecosystem functional space via a comparison of observed and modeled ecosystem flux responses to precipitation regimes and associated stresses in a Central U.S. forest, J. Geophys. Res.: Biogeosciences, 121, 1884-1902, 10.1002/2015jg003302, 2016.

Hadley, J.: AmeriFlux BASE US-LPH Little Prospect Hill, Ver. 1-1, AmeriFlux AMP [dataset], https://doi.org/10.17190/AMF/1246072, 2016.

Jiang, C. and Ryu, Y.: Multi-scale evaluation of global gross primary productivity and evapotranspiration products derived from Breathing Earth System Simulator (BESS), Remote Sens. Environ., 186, 528-547, 10.1016/j.rse.2016.08.030, 2016.

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B. E., Lindroth, A., Merbold, L., Montagnani, L., Moors, E. J., Papale, D., Sottocornola, M., Vaccari, F., and Williams, C.: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, J. Geophys. Res., 116, 10.1029/2010jg001566, 2011.

Ma, H. and Liang, S.: Development of the GLASS 250-m leaf area index product (version 6) from MODIS data using the bidirectional LSTM deep learning model, Remote Sens. Environ., 273, 10.1016/j.rse.2022.112985, 2022.

Noormets, A., McNulty, S. G., DeForest, J. L., Sun, G., Li, Q. L., and Chen, J. Q.: Drought during canopy development has lasting effect on annual carbon balance in a deciduous temperate forest, New Phytol., 10.1111/j.1469-8137.2008.02501, 2008.

Oishi, A. C.: AmeriFlux BASE US-Cwt Coweeta, Ver. 1-5, AmeriFlux AMP [dataset], https://doi.org/10.17190/AMF/1671890, 2020.

Oishi, C., Kim, N., Paul, S.: meriFlux BASE US-Dk2 Duke Forest-hardwoods, Ver. 4-5, AmeriFlux AMP [dataset], https://doi.org/10.17190/AMF/1246047, 2018.

Ouimette, A. P., Ollinger, S. V., Richardson, A. D., Hollinger, D. Y., Keenan, T. F., Lepine, L. C., Vadeboncoeur, M. A.: Carbon Fluxes And Interannual Drivers In A Temperate Forest Ecosystem Assessed Through Comparison Of Top-Down And Bottom-Up Approaches, Agr. Forest Meteorol., 256-257, 420-430, 2018.

Running, S. W., Nemani, R. R., Heinsch, F. A., Zhao, M. S., Reeves, M., and Hashimoto, H.: A continuous satellite-derived measure of global terrestrial primary production, Bioscience, 54, 547–560, 2004.

Schmid, H. P., Grimmond, C. S. B., Cropley, F., Offerle, B., and Su, H. B.: Measurements of CO_2 and energy fluxes over a mixed hardwood forest in the mid-western United States, Agr. Forest Meteorol., 2000.

Urbanski, S., Barford, C., Wofsy, S., Kucharik, C., Pyle, E., Budney, J., McKain, K., Fitzjarrald, D., Czikowsky, M., and Munger, J. W.: Factors controlling CO_2 exchange on timescales from hourly to decadal at Harvard Forest, J. Geophys. Res., 112, 10.1029/2006jg000293, 2007.

Zhang, Y., Xiao, X., Wu, X., Zhou, S., Zhang, G., Qin, Y., and Dong, J.: A global moderate resolution dataset of gross primary production of vegetation for 2000-2016, Sci. data, 4, 170165, 10.1038/sdata.2017.165, 2017.

Zheng, Y., Shen, R., Wang, Y., Li, X., Liu, S., Liang, S., Chen, J. M., Ju, W., Zhang, L., and Yuan, W.: Improved estimate of global gross primary production for reproducing its long-term variation, 1982–2017, Earth Syst. Sci. Data, 12, 2725-2746, 10.5194/essd-12-2725-2020, 2020.