

# BIGC GPP model report

2023-01-30

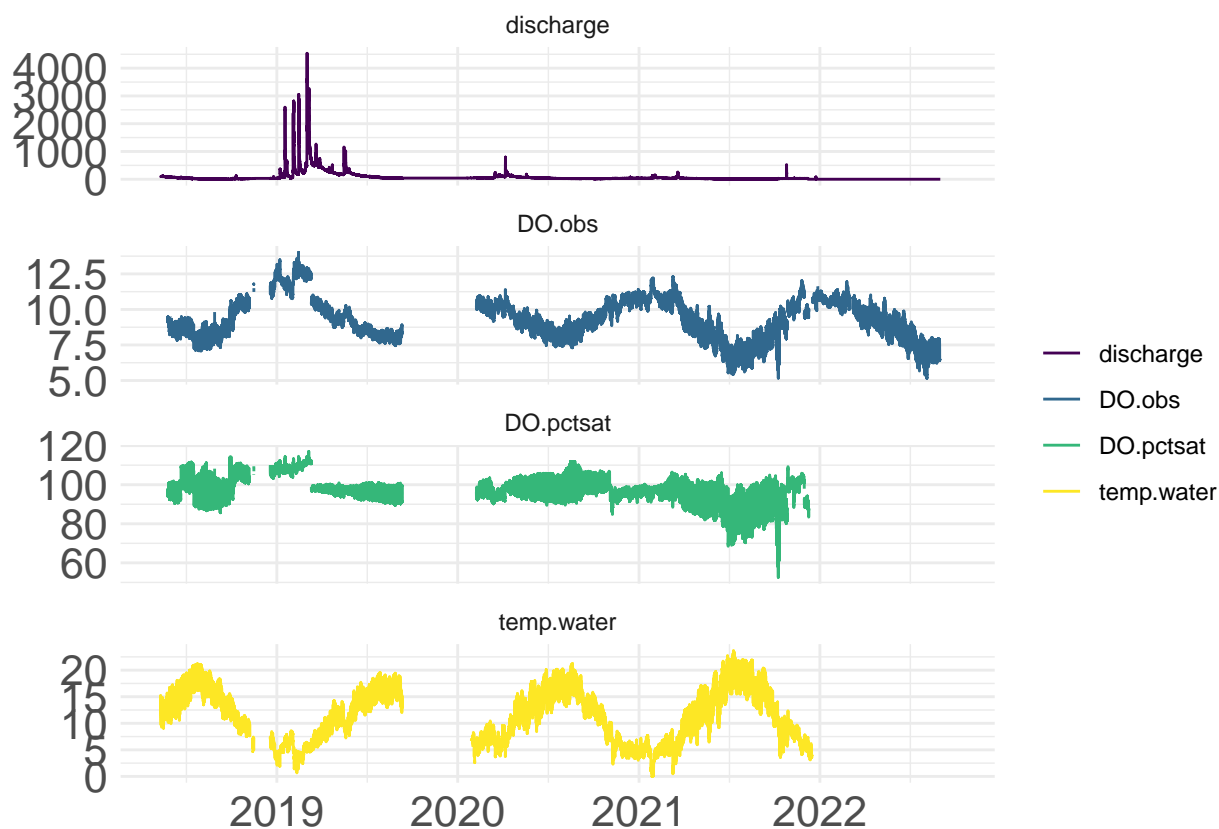


Figure 1: Full metabolism data streams.

The process for building a model suite and selecting a model is iterative. More details for data cleaning will be added here later.

We undergo a process to slim these models down based on a number of criteria. First, we throw out values that have very high estimates of GPP ( $> 1.5\times$  maximum annual values observed in Bernhardt et al. 2022 Figure 1,  $\sim 5000 \text{ g C m}^{-2} \text{ y}^{-1}$ ). Next, any models that have especially poor fits are excluded. In this case I automatically exclude any models that are in the upper 30% percentile of RMSE based on the model set, unless this is fewer than 5 models. This slimmed model set is the group I choose from after one final visual assessment.

I do a visual assessment of the slimmed model set to see how they compare and identify any days where the fit is especially poor across dates. To do this, I first chose the ‘best’ model of the set:

The ‘best’ model is: mm3\_satq10

I then compute the model that is most different than the ‘best’ model and plot the metabolism estimates for each to visually assess.

Table 1: Full model output for all MLE models. 'modelID' represents the model identifier and also information regarding the assumed equations for GPPER. Unmodified models (e.g., 'mm1') represent the default 'mm\_specs' of linear GPPlight relationship. Saturating models (e.g., 'mm1sat') assume a saturating GPPlight relationship. Lastly, Q10 models (e.g., 'mm1satq10') assume a saturating function of GPPlight and an exponential function of GPPTemp. 'modelType' represents how daily K600 values are calculated, whether they are modeled simultaneously with GPP and ER (e.g., 'raw'), modeled as a function of daily discharge with varying relationships (e.g., 'loess', 'lm', 'mean'), modeled from observed night observation (e.g., 'night'), or from empirical gas release estimates (e.g., 'empiricalgam'). The empirical models are further identified by the model used to relate discharge to k, whether GAM or LM. 'RSME' represents the estimated relative root mean square error between observed and modeled dissolved oxygen concentrations standardized to the mean. 'negativeGPP' and 'positiveER' represent the number of days when daily estimates of GPP or ER are negative and positive, respectively. 'meanGPP' is the estimated mean GPP. This value can be scaled to different timeframes or units (i.e., carbon or oxygen) and must be check based on attributes of 'meanGPP'. 'maxK' is the maximum estimated daily k600 value.

modelID	modelType	RMSE	negativeGPP	positiveER	maxK	ER.Kcorr	meanGPP
mm1	raw	5.11	40	117	105902.939	-0.951	2908.132
mm1_sat	raw	5.99	15	65	15456.802	0.538	3297.993
mm1_satq10	raw	5.29	24	59	26832.061	0.974	1735.546
mm2	loess	4.96	14	122	64.501	0.087	505.628
mm2_sat	loess	6.35	14	72	51.312	-0.349	289.672
mm2_satq10	loess	5.06	24	55	26832.061	1.000	362.407
mm3	lm	4.95	23	122	37.300	0.471	450.112
mm3_sat	lm	6.06	14	72	25.237	-0.543	260.504
mm3_satq10	lm	4.67	24	54	425.913	0.688	216.193
mm4	mean	4.94	37	122	23.039	NA	464.475
mm4_sat	mean	6.02	15	74	25.029	NA	263.207
mm4_satq10	mean	5.10	24	52	425.913	0.688	222.186
mm5	night	3950146.97	258	253	56.604	-0.536	197.770
mm5_sat	night	6.08	79	66	56.604	-0.319	126.396
mm5_satq10	night	6251359.48	69	70	56.604	-0.387	112.263
mm6	empirical-lm	167.19	286	131	19394336.727	-0.414	915623.768
mm6_sat	empirical-lm	164.03	69	134	19394336.727	-0.484	29265.001
mm6_satq10	empirical-lm	179.04	103	128	19394336.727	-0.379	32149.449

<sup>a</sup> GPPmean = g C m<sup>-2</sup> y<sup>-1</sup>

The most distant from best is: mm1

These plots are used to assess any dates that are especially bad (e.g., negative GPP, positive ER, very high GPP, etc.) across all models. This dates identified below for the 'top' model and are excluded from the analysis.

The most distant of the best models. Days between this model and above that are both bad are likely very difficult days to fit. Depending on how many of them there are we need to remove them or take a look at the QC flags or the timeseries data to decide if they should be thrown out.

The number of dates and proportion of bad dates within the best model are:

From here we assess any data removal and rerun until reasonably clean model outputs are achieved and a final model is chosen based on minimizing 'negativeGPP' and 'postiveER' and assessing the reasonableness of 'maxK'.

Lastly, we then take the top model suite and build an average model based on model fit, RMSE. To do so, we first remove models that have very high correlations between ER and K (i.e.,  $r \geq 0.80$ ) as these models are largely driven by variability in the estimates in K. Then we create a model weight based on relative model fit,  $dRMSE$ , the difference in relative root mean squared error from the best fitting model based on  $RMSE$  for a set of models,  $m$ , is:

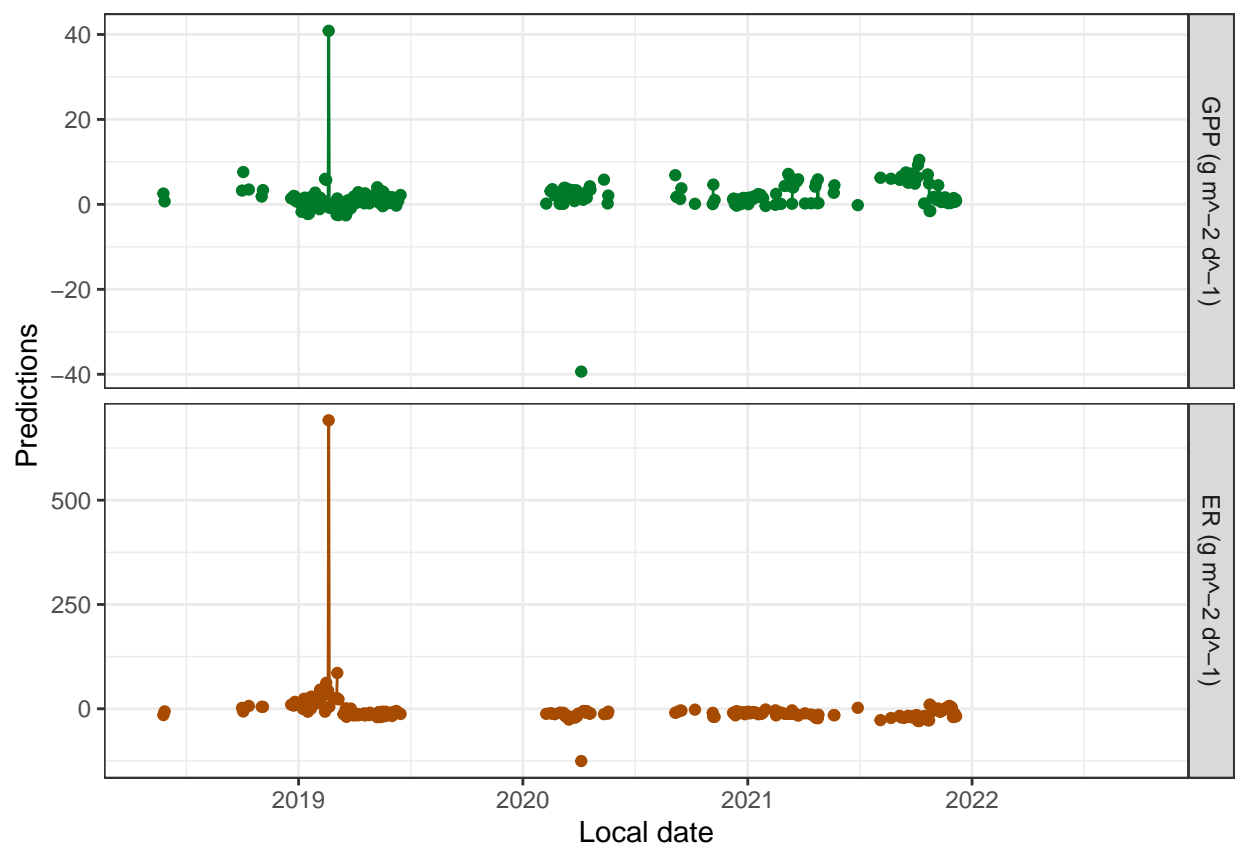


Figure 2: 'Best' model fit.

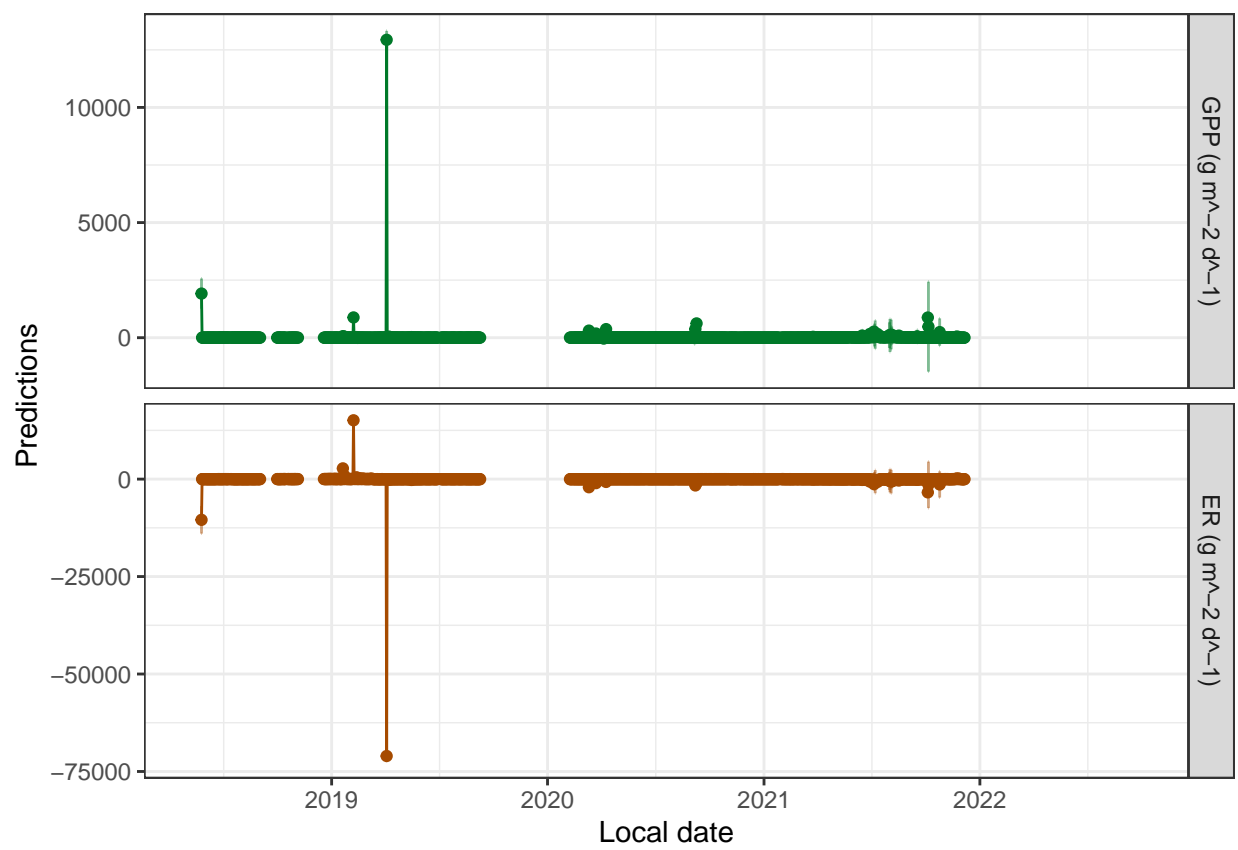


Figure 3: Most distant model.

Table 2: Slimmed model output for all MLE models. 'modelID' represents the model identifier and also information regarding the assumed equations for GPPER. Unmodified models (e.g., 'mm1') represent the default 'mm\_specs' of linear GPPlight relationship. Saturating models (e.g., 'mm1sat') assume a saturating GPPlight relationship. Lastly, Q10 models (e.g., 'mm1satq10') assume a saturating function of GPPlight and an exponential function of GPPTemp. 'modelType' represents how daily K600 values are calculated, whether they are modeled simultaneously with GPP and ER (e.g., 'raw'), modeled as a function of daily discharge with varying relationships (e.g., 'loess', 'lm', 'mean'), modeled from observed night observation (e.g., 'night'), or from empirical gas release estimates (e.g., 'empirical-gam'). The empirical models are further identified by the model used to relate discharge to k, whether GAM or LM. 'RSME' represents the estimated relative root mean square error between observed and modeled dissolved oxygen concentrations standardized to the mean. 'negativeGPP' and 'positiveER' represent the number of days when daily estimates of GPP or ER are negative and positive, respectively. 'meanGPP' is the estimated mean GPP. This value can be scaled to different timeframes or units (i.e., carbon or oxygen) and must be check based on attributes of 'meanGPP'. 'maxK' is the maximum estimated daily k600 value.

modelID	modelType	RMSE	negativeGPP	positiveER	maxK	ER.Kcorr	meanGPP
mm1	raw	5.11	40	117	105902.939	-0.951	2908.132
mm1_sat	raw	5.99	15	65	15456.802	0.538	3297.993
mm1_satq10	raw	5.29	24	59	26832.061	0.974	1735.546
mm2	loess	4.96	14	122	64.501	0.087	505.628
mm2_satq10	loess	5.06	24	55	26832.061	1.000	362.407
mm3	lm	4.95	23	122	37.300	0.471	450.112
mm3_sat	lm	6.06	14	72	25.237	-0.543	260.504
mm3_satq10	lm	4.67	24	54	425.913	0.688	216.193
mm4	mean	4.94	37	122	23.039	NA	464.475
mm4_sat	mean	6.02	15	74	25.029	NA	263.207
mm4_satq10	mean	5.10	24	52	425.913	0.688	222.186
mm5_sat	night	6.08	79	66	56.604	-0.319	126.396

<sup>a</sup> GPPmean = g C m<sup>-2</sup> y<sup>-1</sup>

lowGPP	lowGPPperc	highGPP	highGPPperc	NAs
46	9.957	2	0.433	2394

$$w_i = \exp(\frac{1}{2}dRMSE_i) / \sum_{j=1}^m \exp(\frac{1}{2}dRMSE_j)$$

From this, we create a weighted average model for GPP estimates, plotted below.

modelID	modelType	RMSE	negativeGPP	positiveER	maxK	ER.Kcorr	meanGPP	dRMSE	RMSEwt
mm1_sat	raw	5.99	15	65	15456.802	0.538	3297.993	-1.32	0.080
mm2	loess	4.96	14	122	64.501	0.087	505.628	-0.29	0.134
mm3	lm	4.95	23	122	37.300	0.471	450.112	-0.28	0.135
mm3_sat	lm	6.06	14	72	25.237	-0.543	260.504	-1.39	0.078
mm3_satq10	lm	4.67	24	54	425.913	0.688	216.193	0.00	0.155
mm4	mean	4.94	37	122	23.039	NA	464.475	-0.27	0.136
mm4_sat	mean	6.02	15	74	25.029	NA	263.207	-1.35	0.079
mm4_satq10	mean	5.10	24	52	425.913	0.688	222.186	-0.43	0.125
mm5_sat	night	6.08	79	66	56.604	-0.319	126.396	-1.41	0.077

