CARI GPP model report

2023-01-30

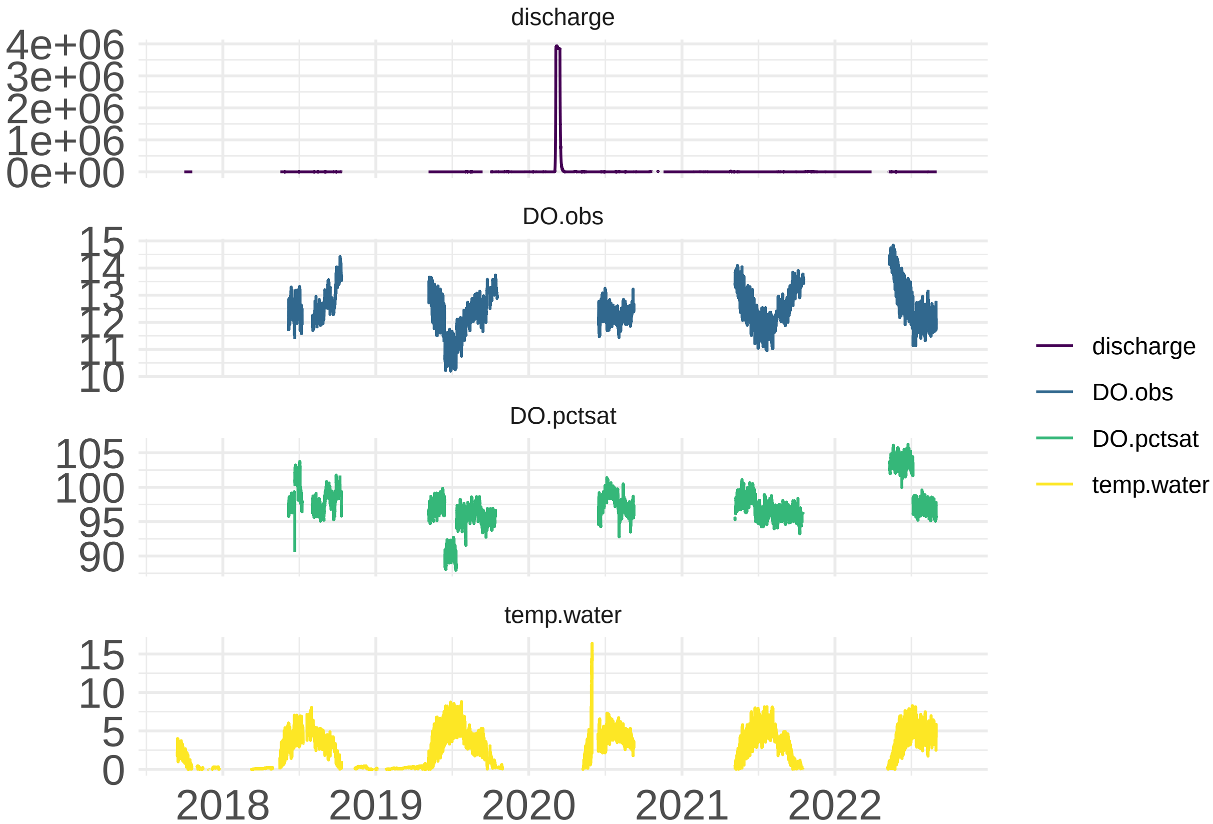


Figure 1: Full metabolism data streams.

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 (> 1.5x maximum annual values observed in Berhnardt et al. 2022 Figure 1, ~5000 g C m^{-2} 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: mm1_sat

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 GPP̃light relationship. Saturating models (e.g., 'mm1sat') assume a saturating GPP̃light relationship. Lastly, Q10 models (e.g., 'mm1satq10') assume a saturating function of GPP̃light and an exponential function of GPP̃Temp. '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	1.110000e+00	144	100	5354.857	0.309	442.051
mm1_sat	raw	9.600000e-01	69	54	2342.271	-0.797	783.429
mm1_satq10	raw	1.370000e+00	89	67	7841.527	-0.955	2116.862
mm2	loess	1.090000e+00	88	69	48.951	-0.169	169.330
mm2_sat	loess	1.030000e+00	67	49	48.951	-0.162	210.675
mm2_satq10	loess	1.360000e+00	77	52	7841.527	-0.961	1043.015
mm3	lm	1.120000e+00	94	68	121.330	-0.168	159.859
mm3_sat	lm	1.070000e+00	73	53	121.330	0.009	195.461
mm3_satq10	lm	1.330000e+00	73	55	7841.527	-0.961	1011.238
mm4	mean	1.130000e+00	91	68	40.149	NA	146.541
mm4_sat	mean	3.160000e+00	212	205	1.000	NA	-464.028
mm4_satq10	mean	1.330000e+00	74	57	7841.527	-0.961	1016.532
mm5	night	2.594357e+108	214	234	2316.524	NA	585.978
mm5_sat	night	4.879426e+106	114	111	2342.271	NA	118.690
mm5_satq10	night	5.502336e+85	111	102	2349.568	NA	220.068
mm6	empirical-gam	6.457000e+01	434	67	21549.759	-0.159	18821.102
mm6_sat	empirical-gam	6.368000e+01	4	68	21549.759	-0.097	2241.656
mm6_satq10	empirical-gam	2.390000e+01	35	64	21549.759	-0.020	26755.465

[a] GPPmean = g C m-2 y-1

The most distant from best is: mm1_satq10

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:

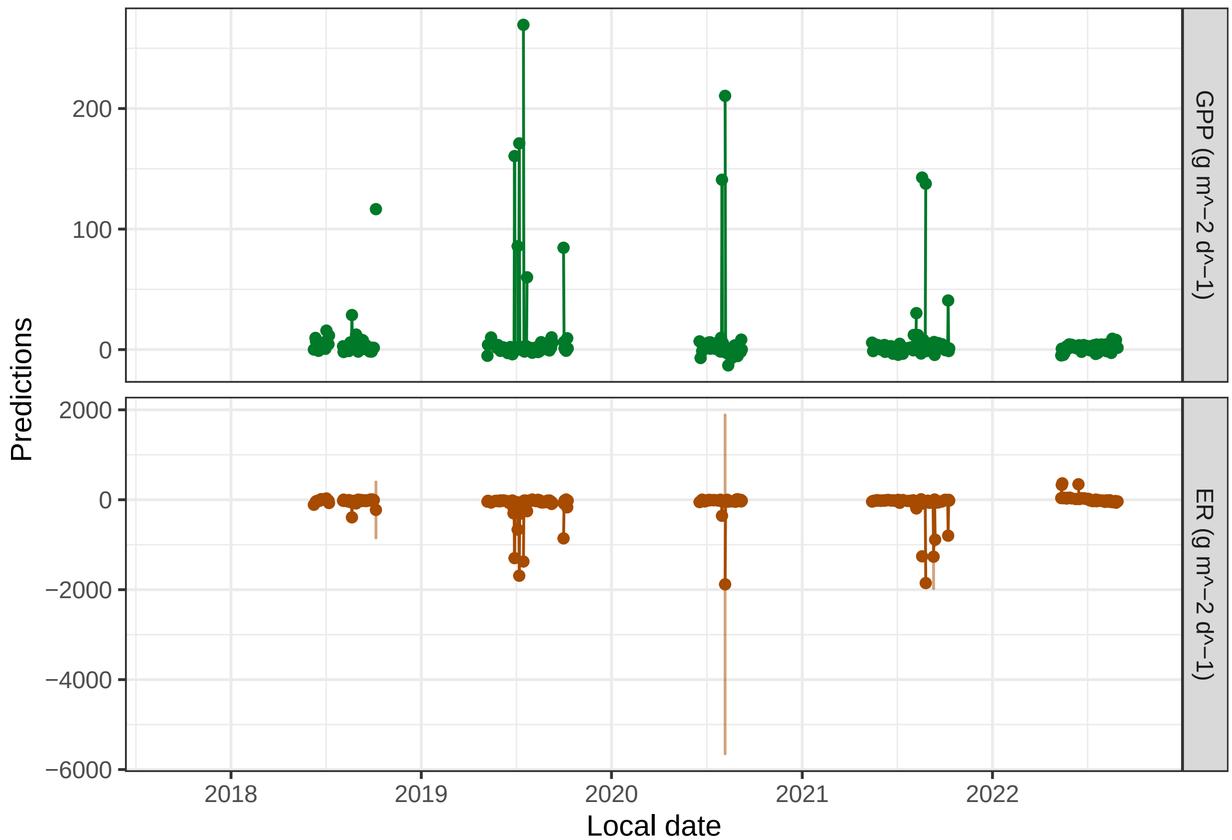


Figure 2: 'Best' model fit.

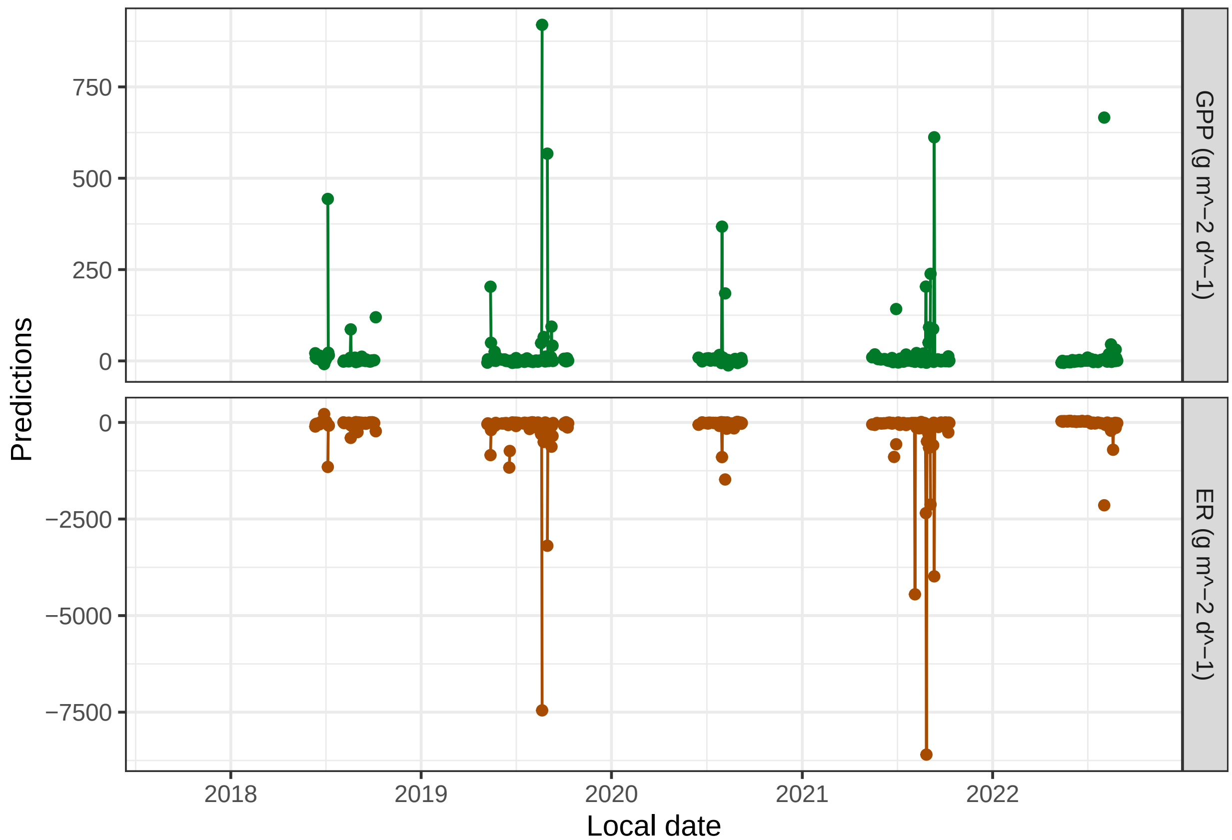


Figure 3: Most distant model.

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 GPP̃light relationship. Saturating models (e.g., 'mm1sat') assume a saturating GPPlight relationship. Lastly, Q10 models (e.g., 'mm1satq10') assume a saturating function of GPP̃light 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	1.11	144	100	5354.857	0.309	442.051
mm1_sat	raw	0.96	69	54	2342.271	-0.797	783.429
mm1_satq10	raw	1.37	89	67	7841.527	-0.955	2116.862
mm2	loess	1.09	88	69	48.951	-0.169	169.330
mm2_sat	loess	1.03	67	49	48.951	-0.162	210.675
mm2_satq10	loess	1.36	77	52	7841.527	-0.961	1043.015
mm3	lm	1.12	94	68	121.330	-0.168	159.859
mm3_sat	lm	1.07	73	53	121.330	0.009	195.461
mm3_satq10	lm	1.33	73	55	7841.527	-0.961	1011.238
mm4	mean	1.13	91	68	40.149	NA	146.541
mm4_satq10	mean	1.33	74	57	7841.527	-0.961	1016.532

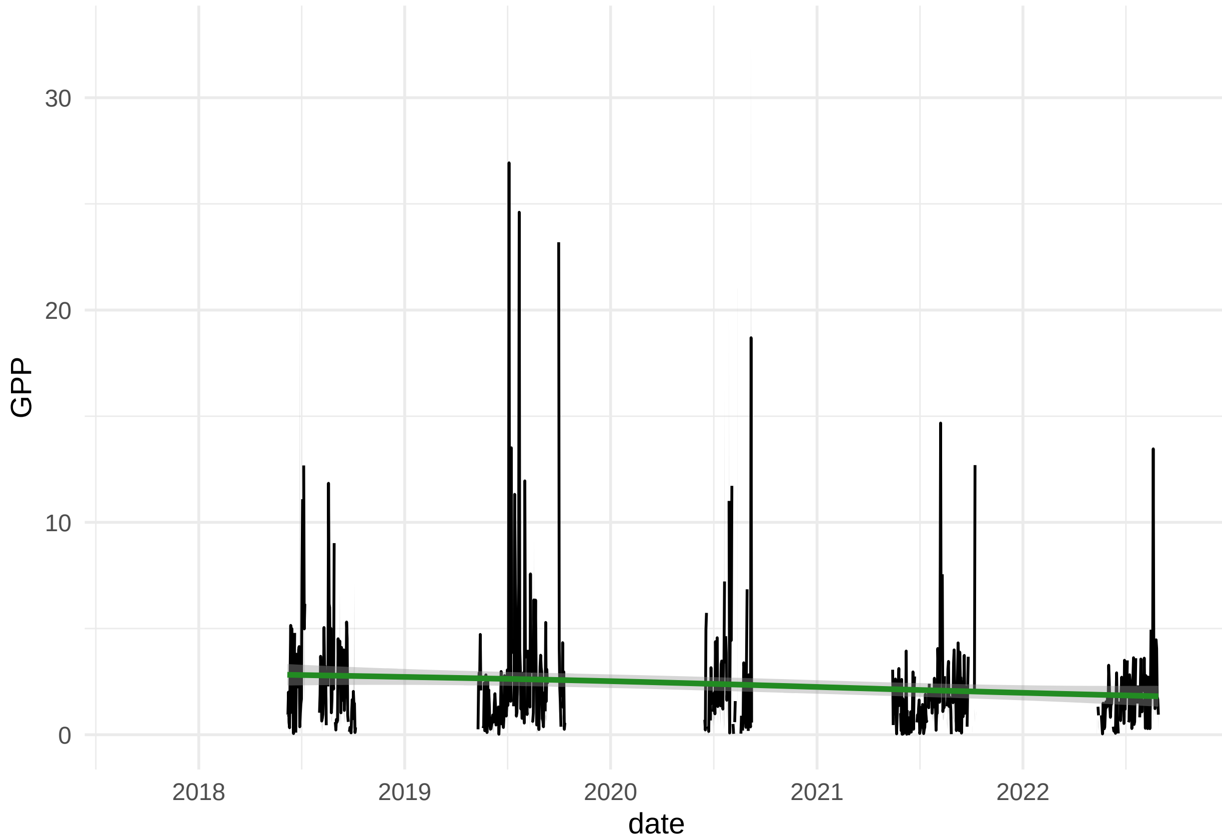
[a] GPPmean = g C m-2 y-1

lowGPP	lowGPPperc	highGPP	highGPPperc	NAs
132	21.639	24	3.934	3050

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

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	raw	1.11	144	100	5354.857	0.309	442.051	-0.15	0.140
mm1_sat	raw	0.96	69	54	2342.271	-0.797	783.429	0.00	0.151
mm2	loess	1.09	88	69	48.951	-0.169	169.330	-0.13	0.142
mm2_sat	loess	1.03	67	49	48.951	-0.162	210.675	-0.07	0.146
mm3	lm	1.12	94	68	121.330	-0.168	159.859	-0.16	0.139
mm3_sat	lm	1.07	73	53	121.330	0.009	195.461	-0.11	0.143
mm4	mean	1.13	91	68	40.149	NA	146.541	-0.17	0.139


30
20
10
0
GPP
2018
2019
2020
2021
2022
date