



Supplement of

Resolving temperature limitation on spring productivity in an evergreen conifer forest using a model–data fusion framework

Stephanie G. Stettz et al.

Correspondence to: Stephanie G. Stettz (sstettz@uci.edu)

The copyright of individual parts of the supplement might differ from the article licence.

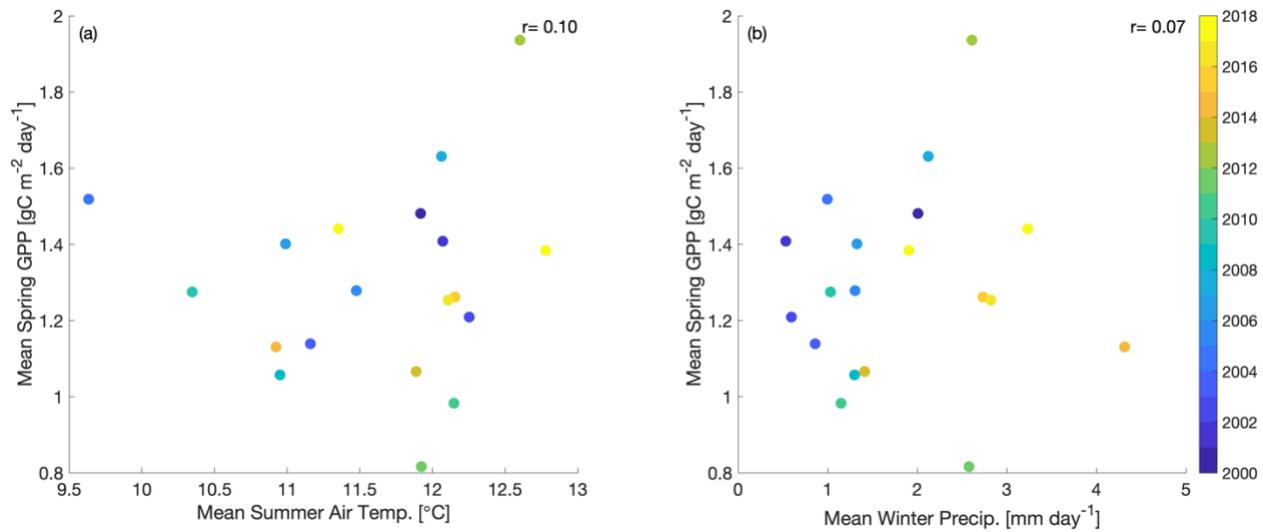


Figure S1. Scatter plots showing the relationship between observed a.) mean summer (June-September) air temperature and b.) mean winter (December-February) precipitation and mean spring (March-May) GPP at Niwot Ridge (US-NR1). Dot color varies by year (2000-2018). 'r' is the Pearson's linear correlation coefficient.

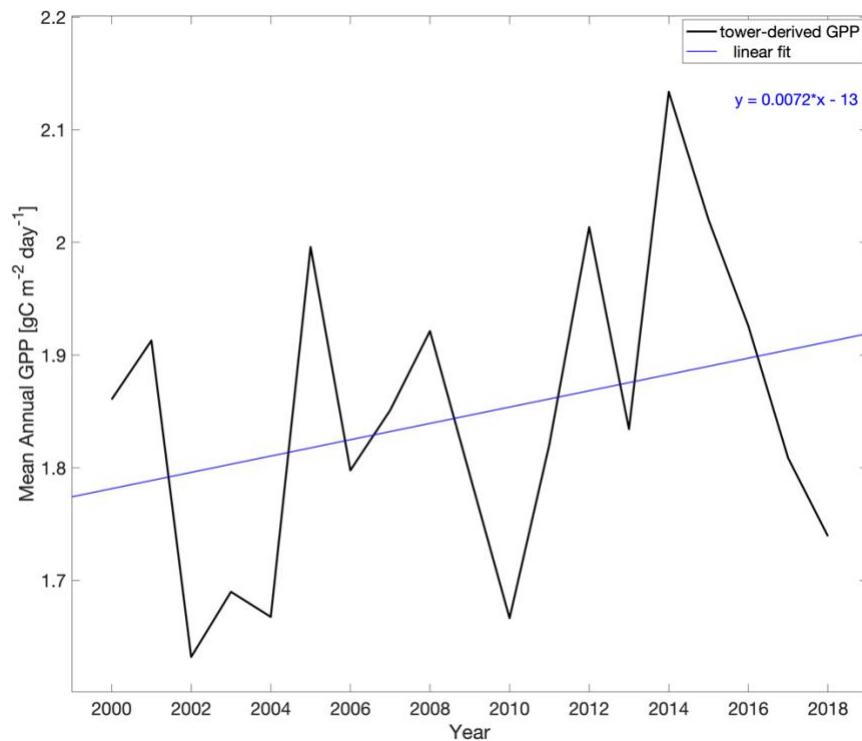
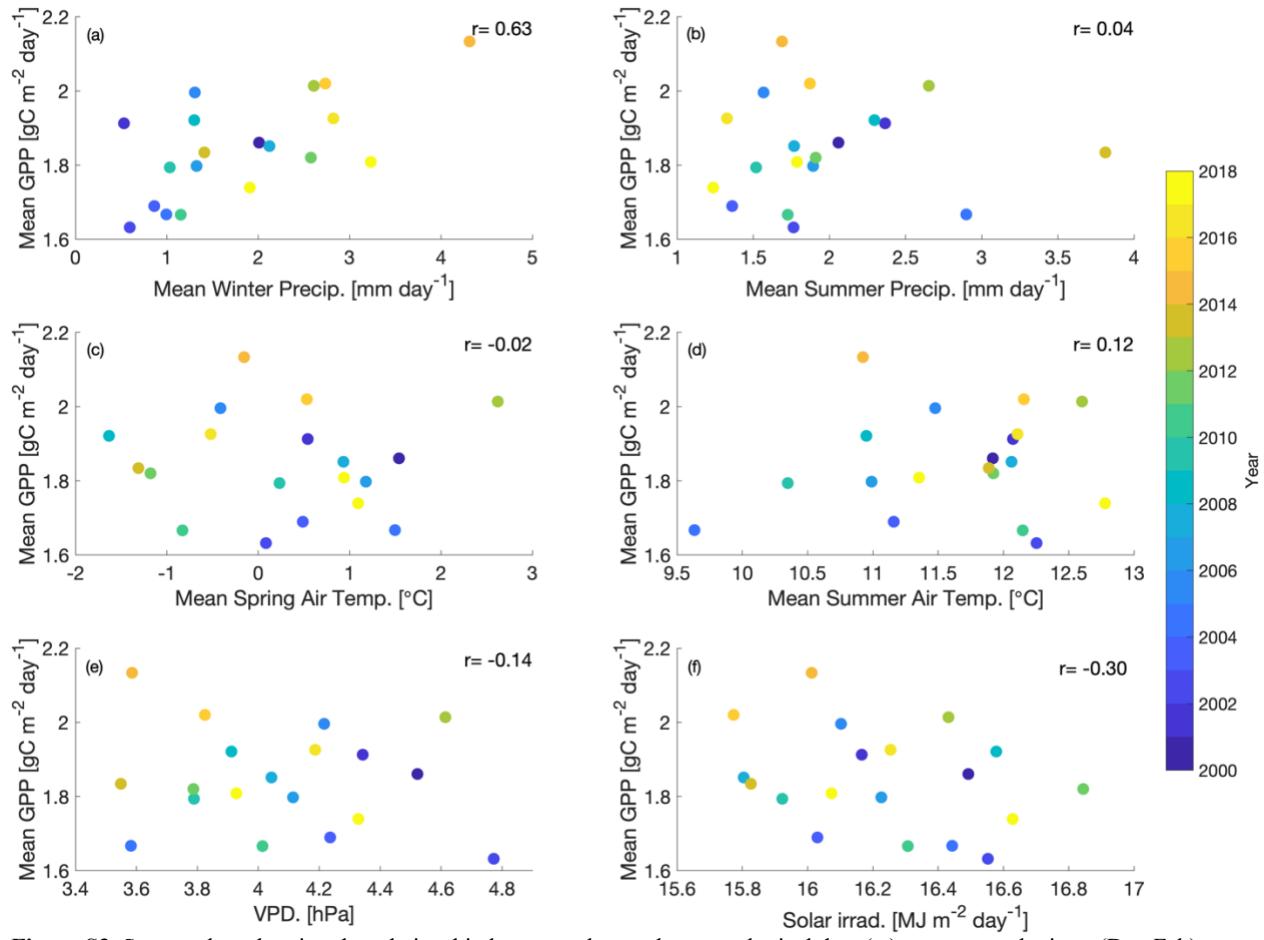


Figure S2. Timeseries of mean annual GPP derived from the AmeriFlux Niwot Ridge eddy covariance tower (US-NR1) from 2000-2018.

53



54

55

56

57

58

59

60

61

62

63

64

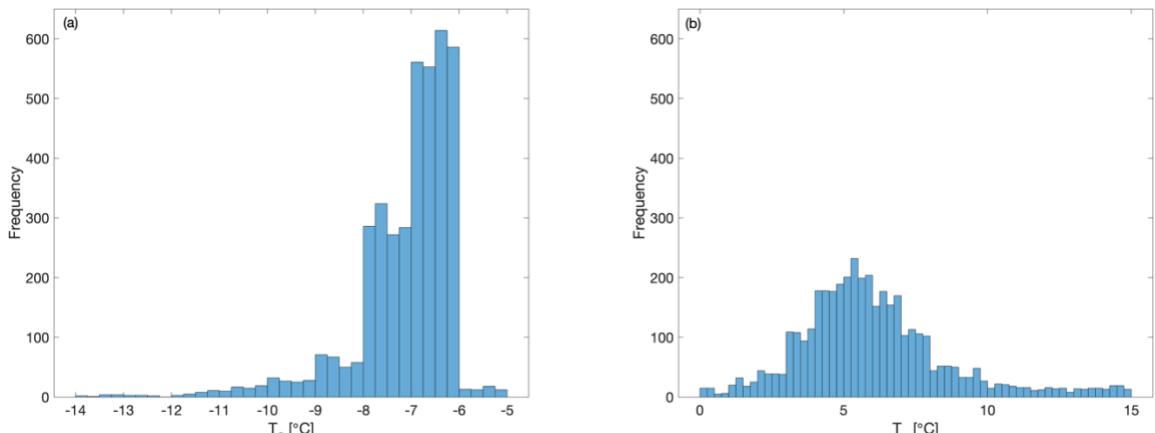
65

66

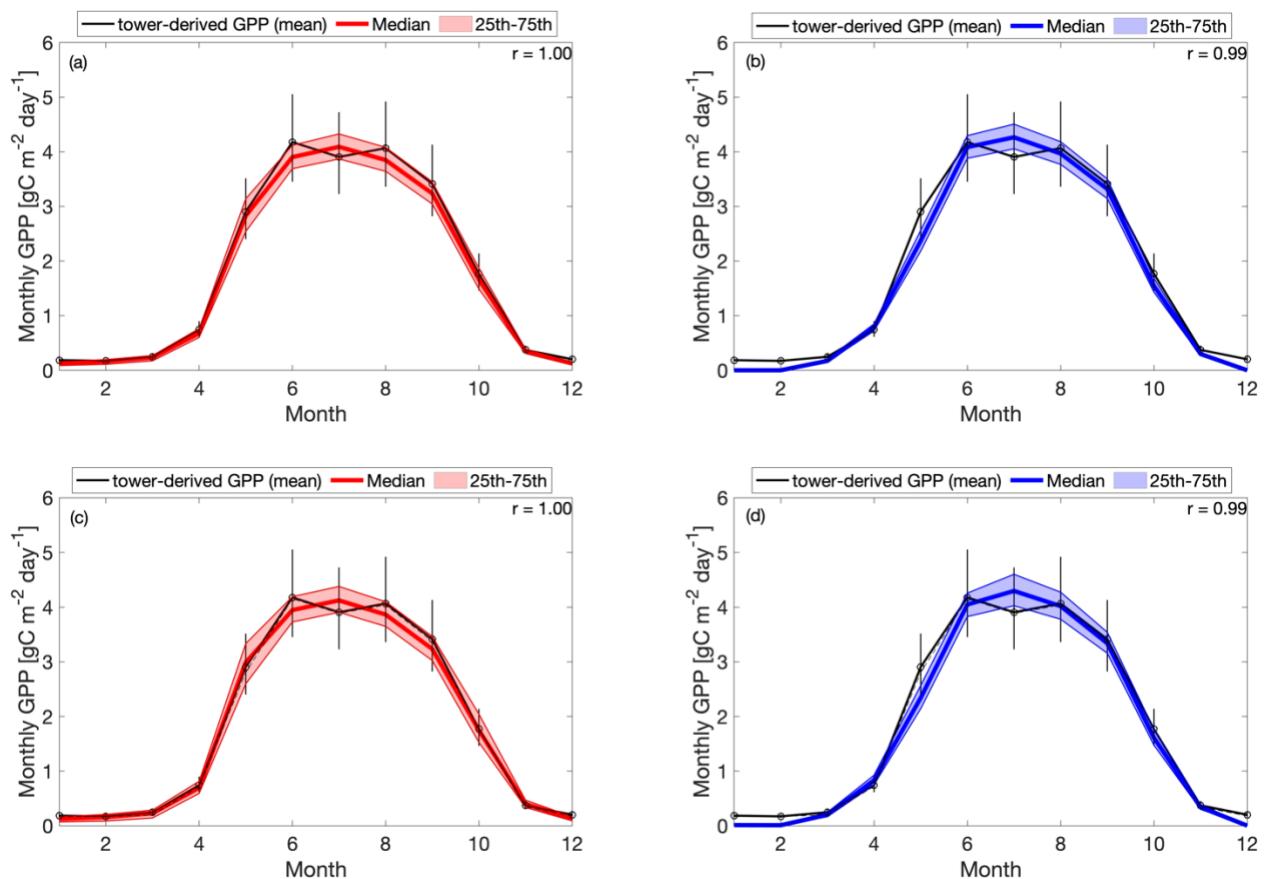
67

68

Figure S3. Scatter plots showing the relationship between observed meteorological data (a.) mean annual winter (Dec-Feb) precipitation, b.) mean summer (Jun-Sep) precipitation, c.) mean spring (Mar-May) air temperature, d.)mean summer (Jun-Sep) air temperature, e.) mean annual vapor pressure deficit (VPD) and f.) mean annual shortwave irradiance, and mean annual GPP at Niwot Ridge (US-NR1). Dot color varies by year (2000-2018). 'r' is the Pearson's linear correlation coefficient.



69
70
71
72
73
Figure S4. Histograms of a.) T_0 (photosynthesis shutdown air temperature) and b.) T_g (photosynthesis initiation air temperature) parameters included in CARDcold experiments ($N = 4000$ ensembles).



74
75
76
77
78
79
80
81
82
Figure S5. Tower-derived (black line) mean monthly GPP with simulated seasonal cycles at US-NR1 for 2000-2018, for a.) CARD, b.) CARDcold, c.) CARD-Half and d.) CARDcold-Half experiments. Model outputs include the median value of each experiment (bold color line) with the 25th-75th percentiles of the ensembles (shaded area). The median is plotted instead of the mean to avoid impact of outlier ensemble members ($N = 4000$ members). Error bars = tower-derived GPP multiplied/divided by $\exp(\sqrt{\log(2)^2 * n}) / n$, $n = \#$ of years in average ($n = 19$). ‘r’ is Pearson’s r coefficient. Experiments are described in more detail in Table 1.

83
84
85
86
87
88

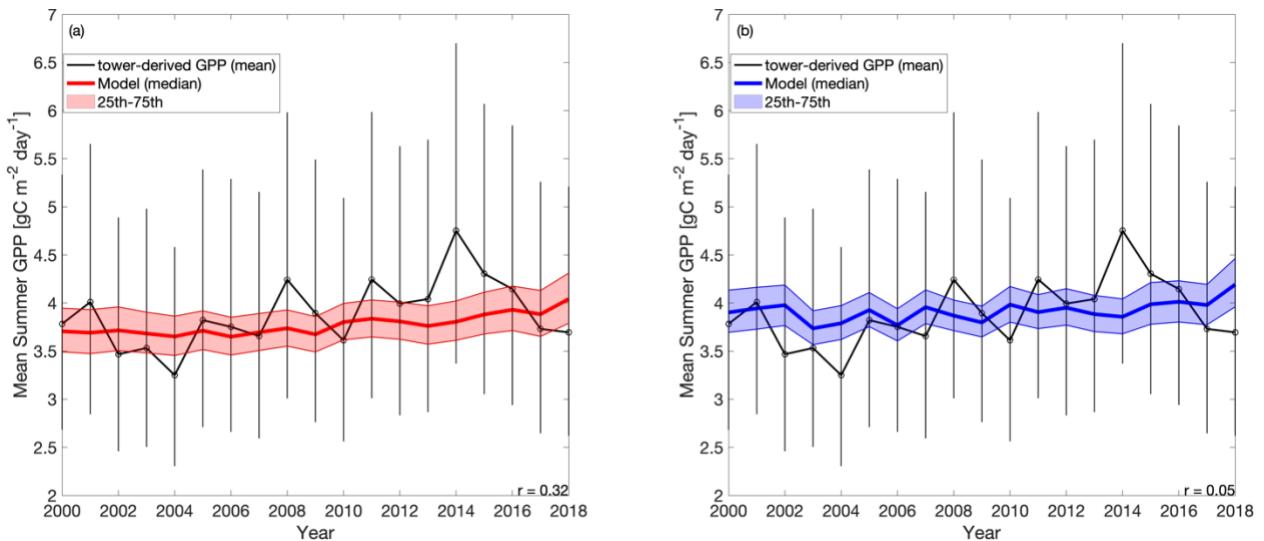
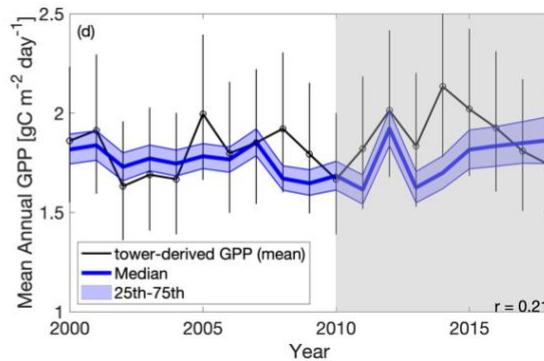
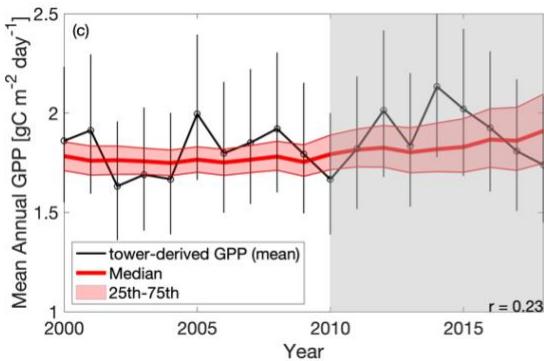
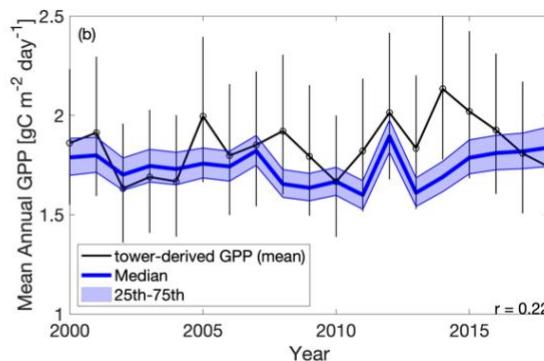
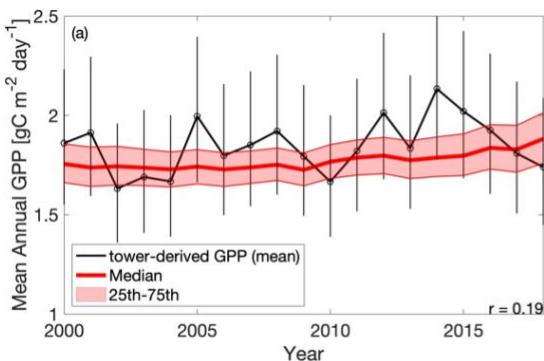


Figure S6. Tower-derived (black line) mean summer (June–September) GPP with model interquartile range (shaded area) and median (bold color line) of summer GPP outputs for a.) CARD and b.) CARDcold experiments. Error bars = tower-derived GPP multiplied/divided by $\exp(\sqrt{\log(2)^2 n}/n)$, $n = \#$ of months in average ($n = 4$).



89
90
91
92
93
94

Figure S7. Tower-derived (black line) mean annual GPP with model interquartile range (color shaded area) and median (bold color line) of annual GPP outputs for a.) CARD, b.) CARDcold, c.) CARD-Half, and d.) CARDcold-Half experiments. The grey regions indicate no data assimilation (i.e. testing window). Error bars = tower-derived GPP multiplied/divided by $\exp(\sqrt{\log(2)^2 n}/n)$, $n = \#$ of years in average ($n = 19$).

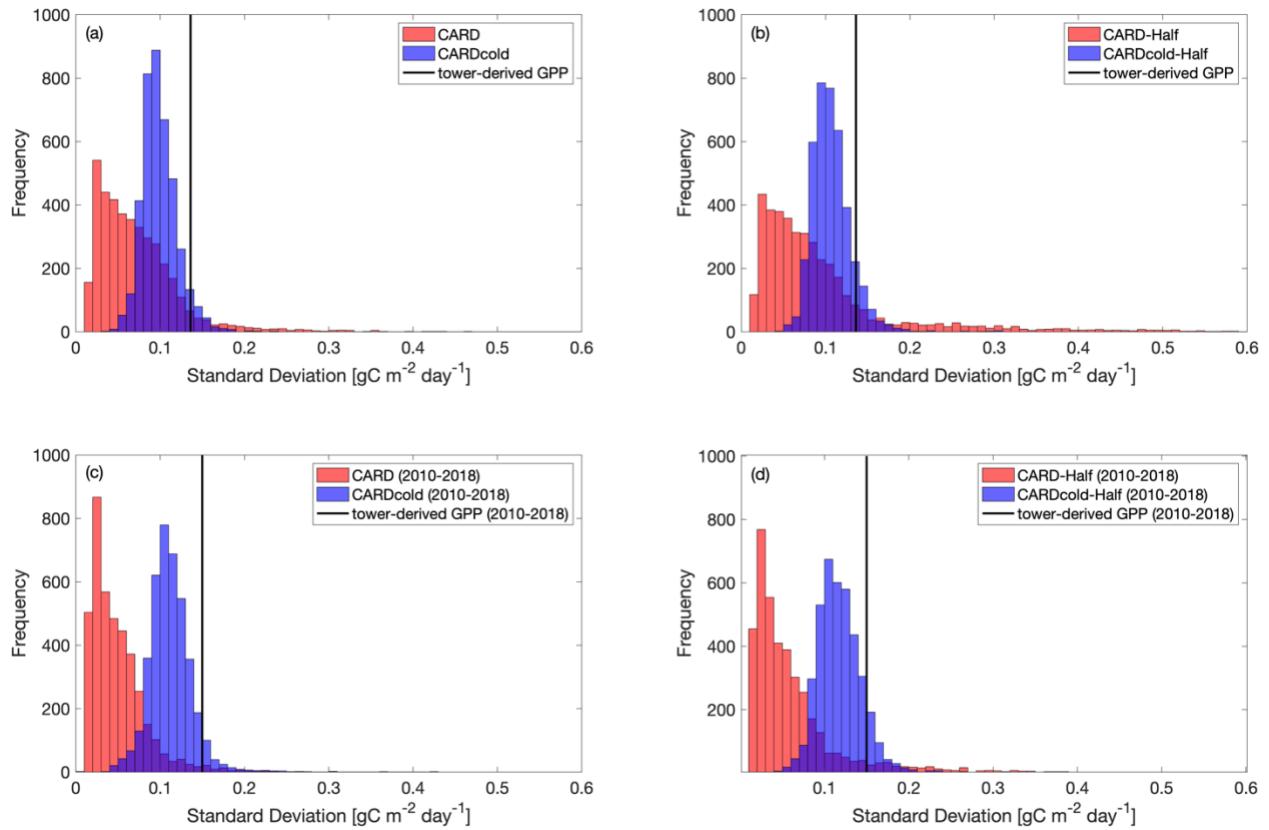
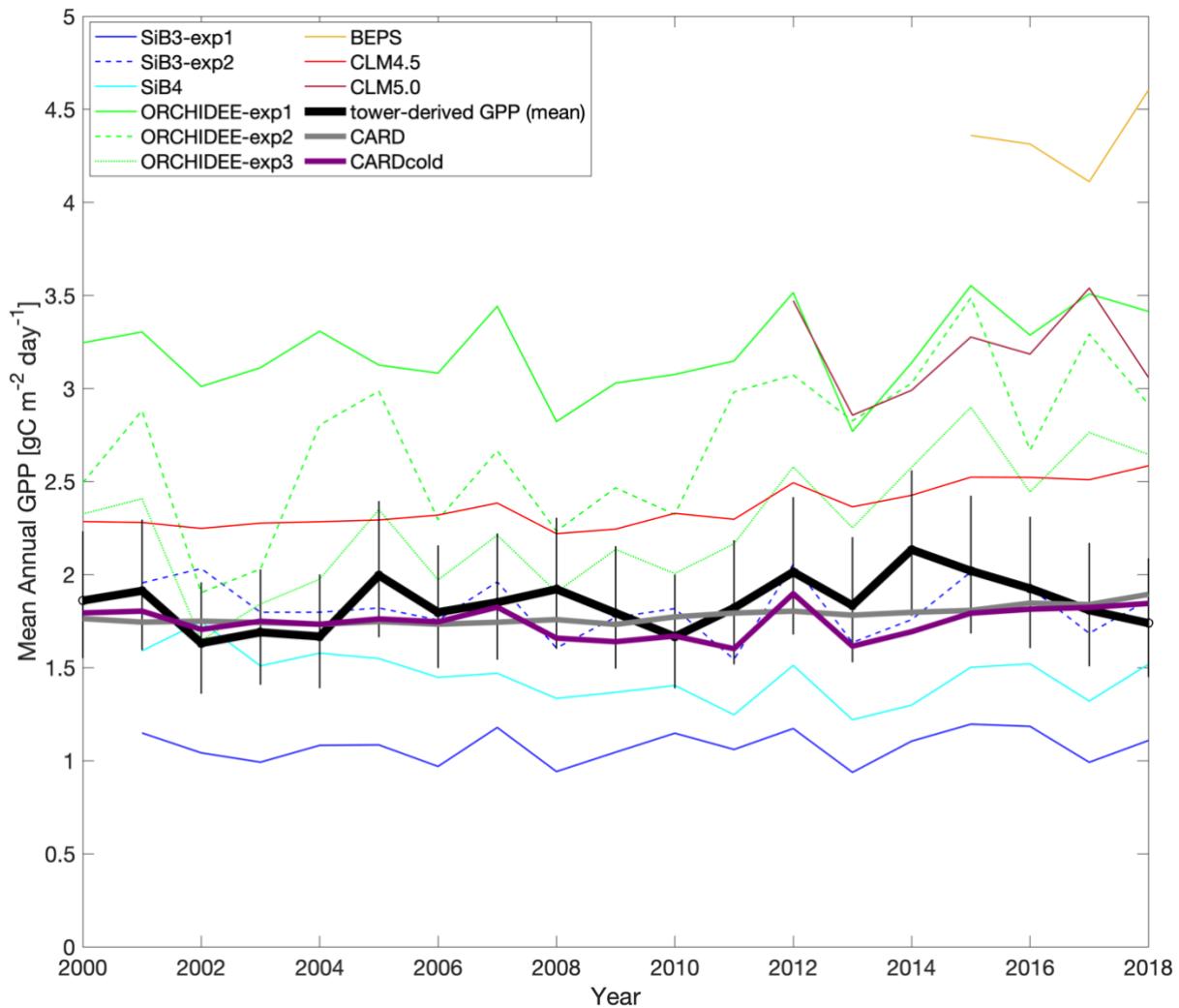


Figure S8. Histograms comparing standard deviation in mean annual GPP across all ensembles ($N=4000$) for CARD (red bars) and CARDcold (blue bars) experiments with a.) full assimilation, b.) half assimilation, c.) full assimilation for the second decade (2010-2018), and d.) half assimilation for the second decade (2010-2018). Black line indicates standard deviation in tower-derived mean annual GPP (std = $0.14 \text{ gC m}^{-2} \text{ day}^{-1}$ for full period (a-b), std = $0.15 \text{ gC m}^{-2} \text{ day}^{-1}$ for 2010-2018 (c-d)).

95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114



115
116 **Figure S9.** Comparison of TBM-MIP models to CARD and CARDcold experiments for mean annual GPP for 2000-2018.
117 Uncertainty = $\exp(\sqrt{\log(2)^2 * n}/n)$, where $n = \#$ years in average ($n = 19$).
118

119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136

137
138
139

Table S1. Pearson's linear r, R-squared, p-value, standard deviation, root mean square error (RMSE), and mean bias error (MBE) for TBM-MIP and all CARDAMOM experiments to Niwot Ridge tower-derived mean spring GPP for 2015-2018. All relevant statistics calculated at 5% significance level.

model	r-value	R-squared	p-value ($\alpha = 0.05$)	RMSE (gC m ⁻² d ⁻¹)	MBE (gC m ⁻² d ⁻¹)	standard deviation (gC m ⁻² d ⁻¹)
CARD-Half	0.82	0.67	0.18	0.08	0.06	0.05
CARD	0.82	0.67	0.18	0.13	0.12	0.04
CARDcold-Half	0.78	0.61	0.22	0.22	0.20	0.18
CARDcold	0.76	0.58	0.24	0.25	0.22	0.19
SiB3-exp1	0.44	0.20	0.56	1.24	1.24	0.08
SiB3-exp2	0.42	0.17	0.58	1.16	1.16	0.09
SiB4	0.57	0.33	0.43	1.01	1.01	0.06
ORCHIDEE-exp1	0.67	0.45	0.33	1.43	-1.42	0.19
ORCHIDEE-exp2	-0.09	0.01	0.91	1.38	-1.37	0.20
ORCHIDEE-exp3	-0.26	0.07	0.74	0.93	-0.92	0.08
BEPS	0.84	0.70	0.16	2.46	-2.46	0.16
CLM4.5	0.74	0.55	0.26	0.36	-0.35	0.15
CLM5.0	0.91	0.82	0.09	1.10	-1.09	0.24

140
141
142
143
144

Table S2. Pearson's linear r, R-squared, p-value, standard deviation, root mean square error (RMSE), and mean bias error (MBE) for TBM-MIP and all CARDAMOM experiments to Niwot Ridge tower-derived mean annual GPP for 2001-2018. All relevant statistics calculated at 5% significance level.

model	r-value	R-squared	p-value ($\alpha = 0.05$)	RMSE (gC m ⁻² d ⁻¹)	MBE (gC m ⁻² d ⁻¹)	standard deviation (gC m ⁻² d ⁻¹)
CARD-Half	0.23	0.05	0.35	0.14	0.05	0.05
CARD	0.19	0.04	0.44	0.15	0.07	0.04
CARDcold-Half	0.21	0.04	0.41	0.17	0.08	0.09
CARDcold	0.22	0.05	0.37	0.18	0.11	0.08
SiB3-exp1	0.32	0.10	0.19	0.78	0.77	0.09
SiB3-exp2	0.10	0.01	0.69	0.19	0.02	0.15
SiB4	-0.25	0.06	0.32	0.45	0.39	0.14
ORCHIDEE-exp1	0.21	0.04	0.40	1.38	-1.36	0.23
ORCHIDEE-exp2	0.60	0.36	0.01	0.94	-0.87	0.43
ORCHIDEE-exp3	0.64	0.41	0.00	0.50	-0.42	0.34
BEPS	-0.25	0.06	0.75	2.48	-2.47	0.20
CLM4.5	0.34	0.12	0.17	0.54	-0.52	0.11
CLM5.0	0.03	0.00	0.96	1.30	-1.27	0.25

145
146
147
148
149
150
151
152
153
154
155

156
157
158

Table S3. Pearson's linear r, R-squared, p-value, standard deviation, root mean square error (RMSE), and mean bias error (MBE) for TBM-MIP and all CARDAMOM experiments to Niwot Ridge tower-derived mean monthly GPP for 2015-2018. All relevant statistics calculated at 5% significance level.

model	r-value	R-squared	p-value ($\alpha = 0.05$)	RMSE (gC m ⁻² d ⁻¹)	MBE (gC m ⁻² d ⁻¹)	standard deviation (gC m ⁻² d ⁻¹)
CARD-Half	1.00	1.00	0.00	0.09	0.007	1.76
CARD	1.00	1.00	0.00	0.09	0.04	1.75
CARDcold-Half	0.99	0.98	0.00	0.23	0.04	1.79
CARDcold	0.99	0.98	0.00	0.22	0.07	1.80
SiB3-exp1	0.91	0.84	0.00	1.02	0.76	1.50
SiB3-exp2	0.92	0.84	0.00	1.06	0.01	2.47
SiB4	0.94	0.88	0.00	0.73	0.42	1.77
ORCHIDEE-exp1	0.98	0.95	0.00	1.80	-1.55	2.58
ORCHIDEE-exp2	0.95	0.91	0.00	1.48	-1.20	2.40
ORCHIDEE-exp3	0.96	0.93	0.00	0.95	-0.81	1.90
BEPS	0.98	0.95	0.00	2.56	-2.46	2.32
CLM4.5	0.99	0.98	0.00	1.20	-0.65	2.76
CLM5.0	0.97	0.95	0.00	1.60	-1.38	2.45

159