Supplement of Biogeosciences, 22, 1597–1614, 2025
https://doi.org/10.5194/bg-22-1597-2025-supplement

Biogeosciences Open Access EGU

Supplement of

Precipitation–fire functional interactions control biomass stocks and carbon exchanges across the world's largest savanna

Mathew Williams et al.

Correspondence to: Mathew Williams (mat.williams@ed.ac.uk)

Supplement:

Causal effect estimation using Wright's path method:

To estimate causal effects (Runge et al., 2015) we first need to define a causal graph describing all direct causal links between the variables, based on expert domain knowledge. While most works conduct causal analysis on time series, in this work we apply a similar framework to a spatial analysis on ensemble-median, 12-year annual means, over the SAW region. For this the spatial 2d-fields are flattened (1d array) and then treated as a "time series". Thus, our proposed causal graph cannot have any lagged links, which would represent spatial correlations. This also requires our graph to be stationary in space, i.e., the graphs include the impact of climate and fire on variations in vegetation type and function across the SAW region. This means that the causal links represented in the networks reflect both mechanistic process sensitivities between the drivers and specific C fluxes that are coded in the model (e.g. GPP is dependent on soil moisture, which is directly tied to the history of precipitation and evaporation), and emergent spatial effects that reflect functional variations driven by the SAW climate gradient, as inferred from our spatially explicit parameter retrieval. Finally, our graph needs to be resolved to not contain any hidden confounding, the direction of the links must be known, and the graph cannot present any feedback loops.
To describe the causal relations between the n variables in the graph, a structural causal model (SCM) is defined:

$$X_j(t) \coloneqq f_j(pa\left(X_j(t)\right), U_j(t))$$

For the linear case, which is assumed in this work, this is given through:

$$X_j(t) = \sum_{X_i(t-\tau) \in pa\left(X_j(t)\right)} \alpha_{X_i,X_j,\tau} X_i(t-\tau) + U_j(t)$$

for $j \in [1: n]$. f_j is a linear function, that estimates the value of $X_j(t)$ based on its direct causal parents $pa(X_j(t)) = (X_i(t - \tau))_i$. $\alpha_{X_i,X_j,\tau}$ are referred to as the path coefficients or linear direct causal effects (in the linear case) and describe the direct causal effect of the parent variable $X_i(t)$ on the variable $X_j(t)$ at lag τ. U_j is a jointly independently distributed noise variable.

To quantify the linear direct causal effect Wright's method (Wright (1921), Wright (1934), Runge et al (2015)) is used, which only applies in the linear case. Direct causal effects $\alpha_{X_i,X_j,\tau}$ are estimated as the partial regression slopes of the multiple linear regression of $X_j(t)$ on its parents $pa(X_j(t))$. This allows to control for any source of confounding. The total causal effect X on Y can then be estimated through:

$$\beta_{X,Y,\tau} = \sum_{\substack{causal\ paths \\ from\ X(t-\tau) \\ to\ Y(t)}} \left(\prod_{\substack{X_i \to X_j \\ arrow\ in \\ path\ for\ \tau}} \alpha_{X_i,X_j,\tau} \right)$$

The freely available python package tigramite (https://github.com/jakobrunge/tigramite) was used for the causal effect quantification. To compare causal effects along different mediating paths, the effect variables were standardised before the causal analysis, by subtracting the mean and dividing the anomalies by the standard deviation. $\tau = 0$ is assumed for the full analysis, to prevent the influence of long distance spatial effects. Uncertainties of $\alpha_{X_i,X_j,0}$ were estimated using a bootstrapping method with 500 members.

SI figures:

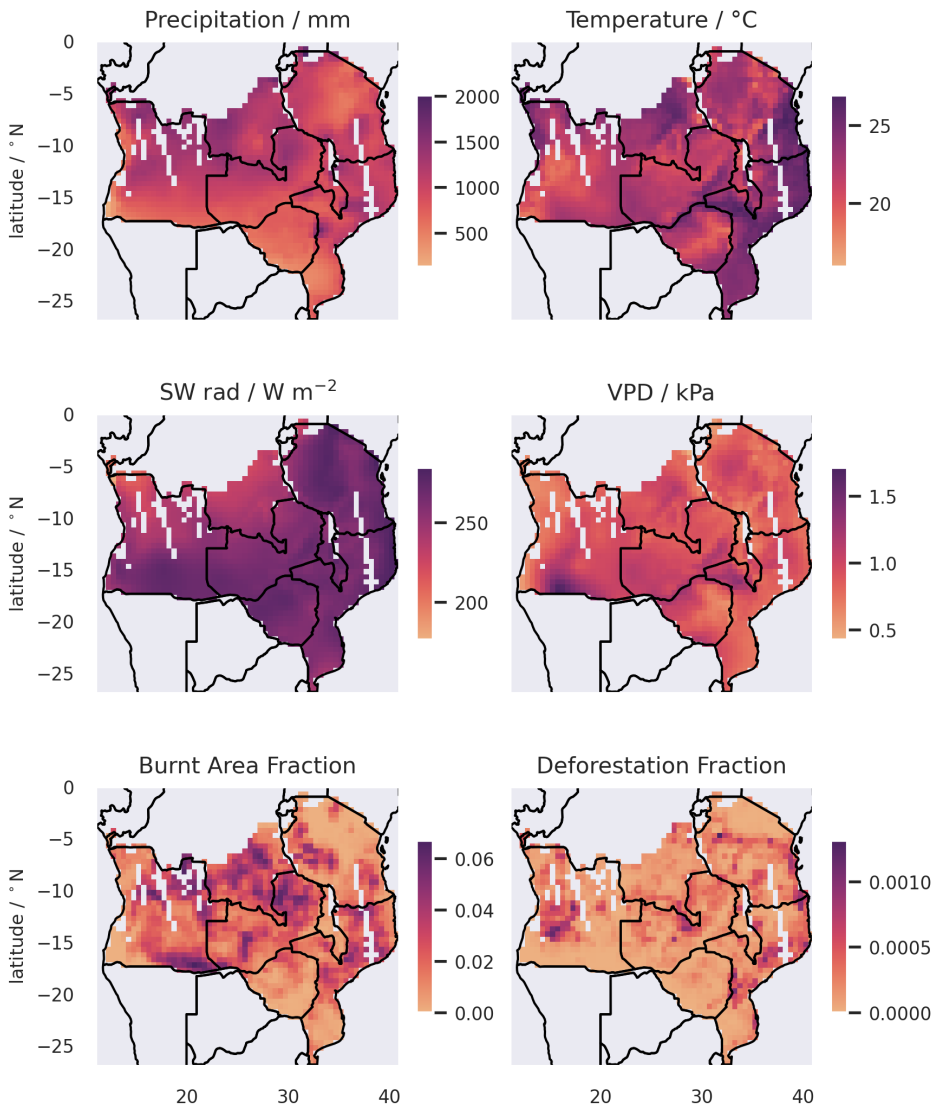


Figure S 1. Mapping of annual mean climate and disturbance drivers over the study period (2006-2017) across the SAW region. See text for data sources. Blanks relate to areas not analysed due to earth observation gaps in the biomass data.

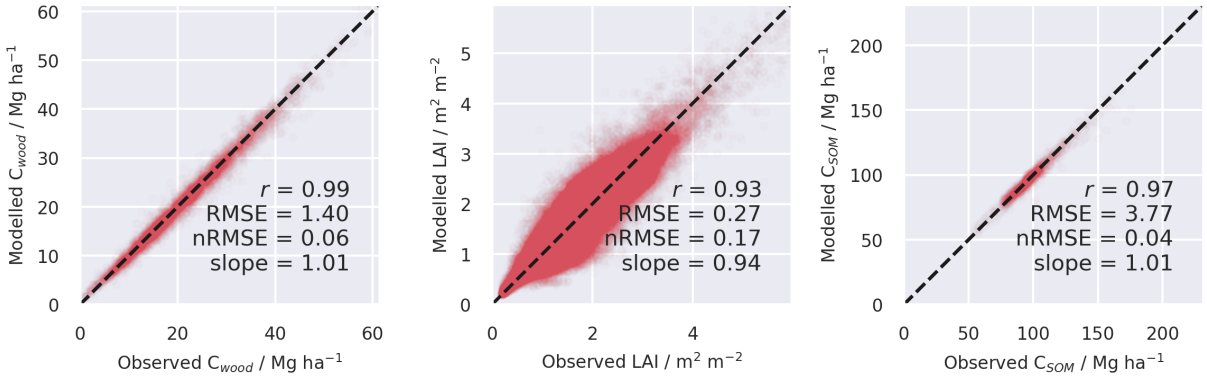


Figure S 2. Evaluation of the DALEC calibration for C_{wood} (total coarse wood carbon stock), LAI and C_{SOM} (total soil organic carbon). Scatterplots show the correspondence between estimates from the calibrated model against observations for all time points and pixel locations. R = correlation coefficient; RMSE = root mean square error; nRMSE = RMSE normalised by the mean; slope is linear regression slope. Dashed lines = 1:1 lines.

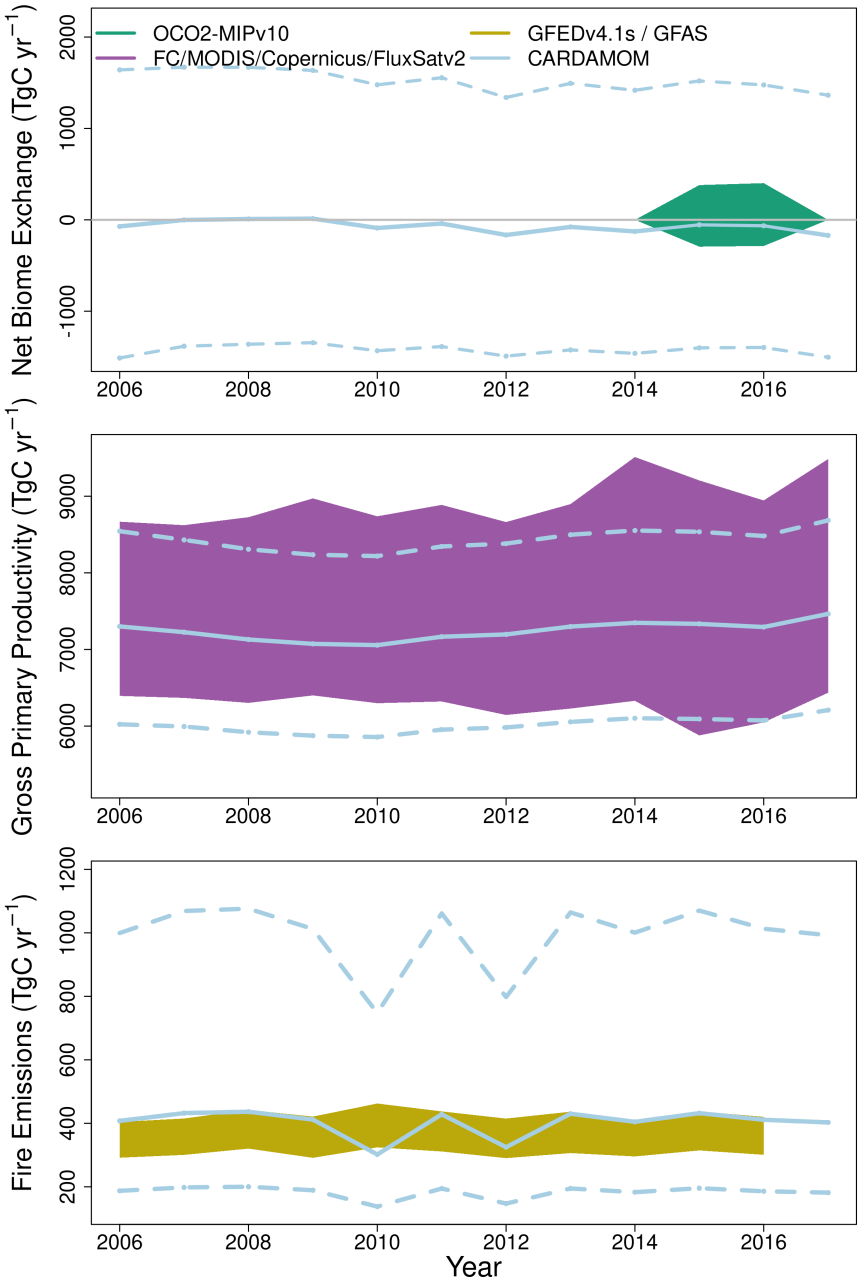


Figure S 3. A comparison of CARDAMOM analyses of major C fluxes (top NBE; middle GPP; bottom fire emissions) against alternative data sources (see legend in top panel) for the SAW region. CARDAMOM estimates are shown in blue; solid line for the median estimate and dashed lines for the 95% confidence intervals.

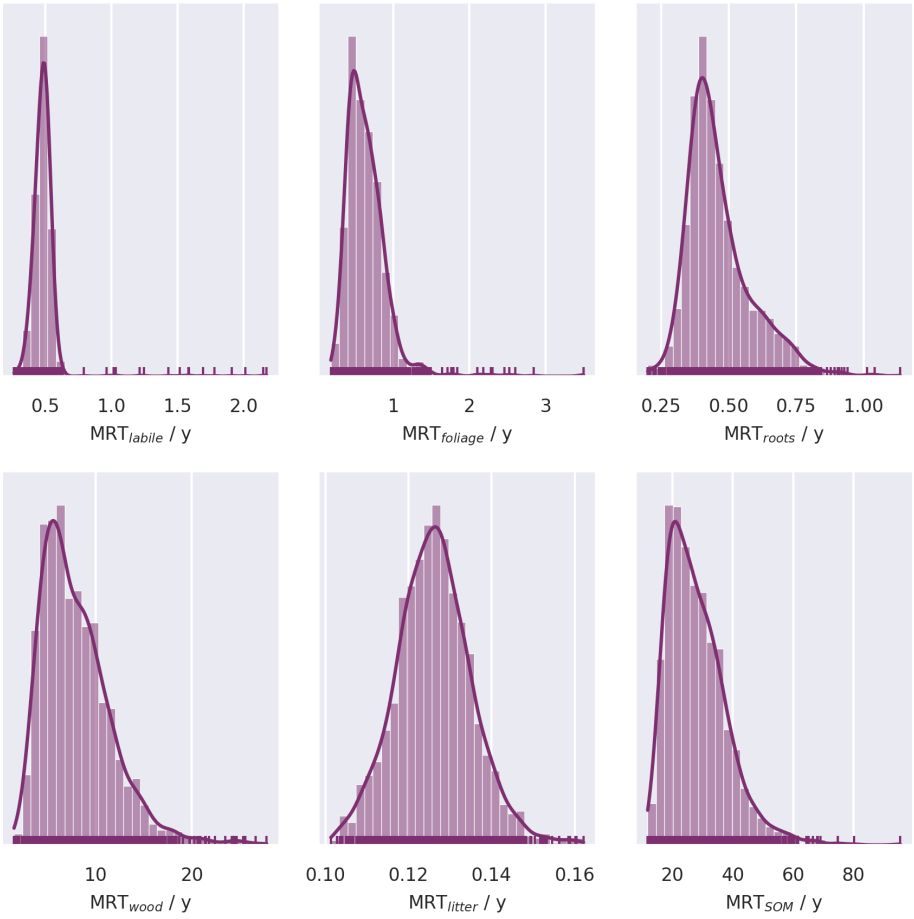


Figure S 4. Probability distributions of mean residence times (MRT) for vegetation and dead organic C pools across the SAW region at 0.5° resolution, across 2006-2017, as determined by diagnostic analysis.

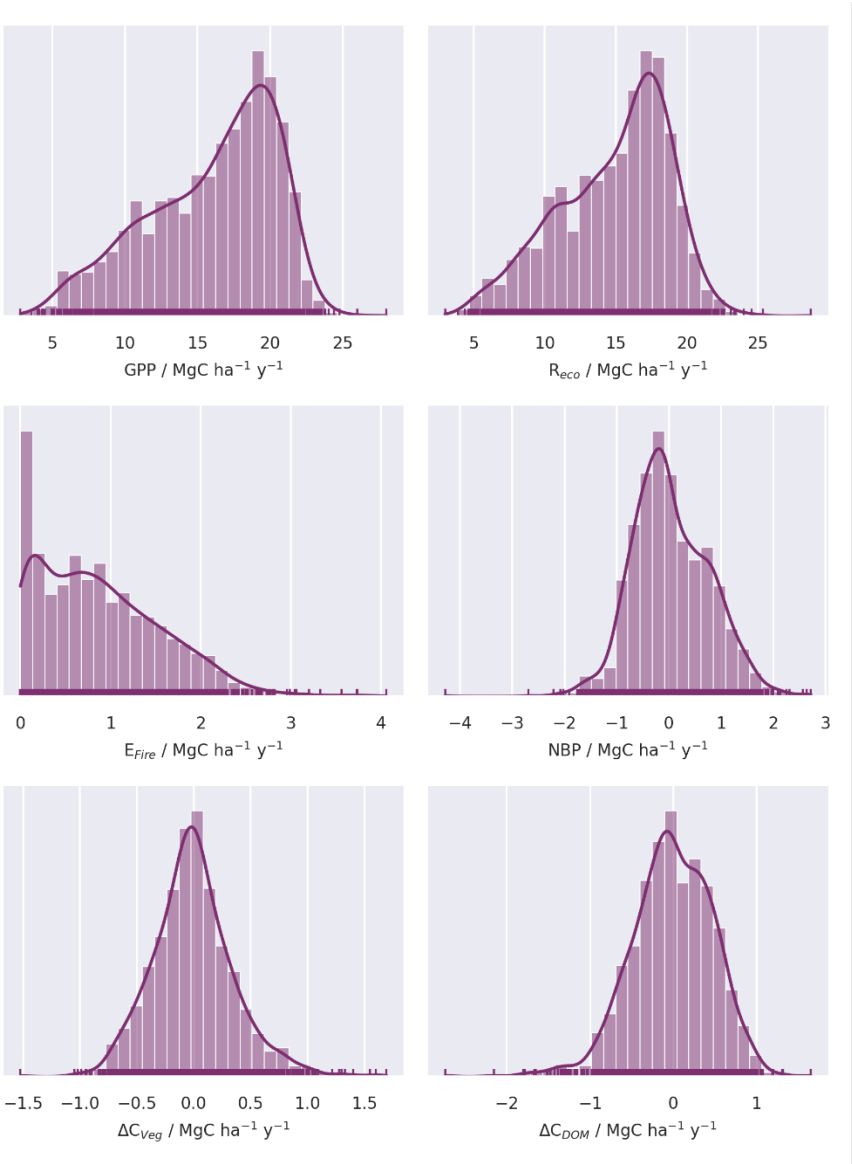


Figure S 5. Probability distributions of median gross fluxes, NBP and C stock changes across the SAW region at 0.5° resolution, 2006-2017, as determined by diagnostic analysis

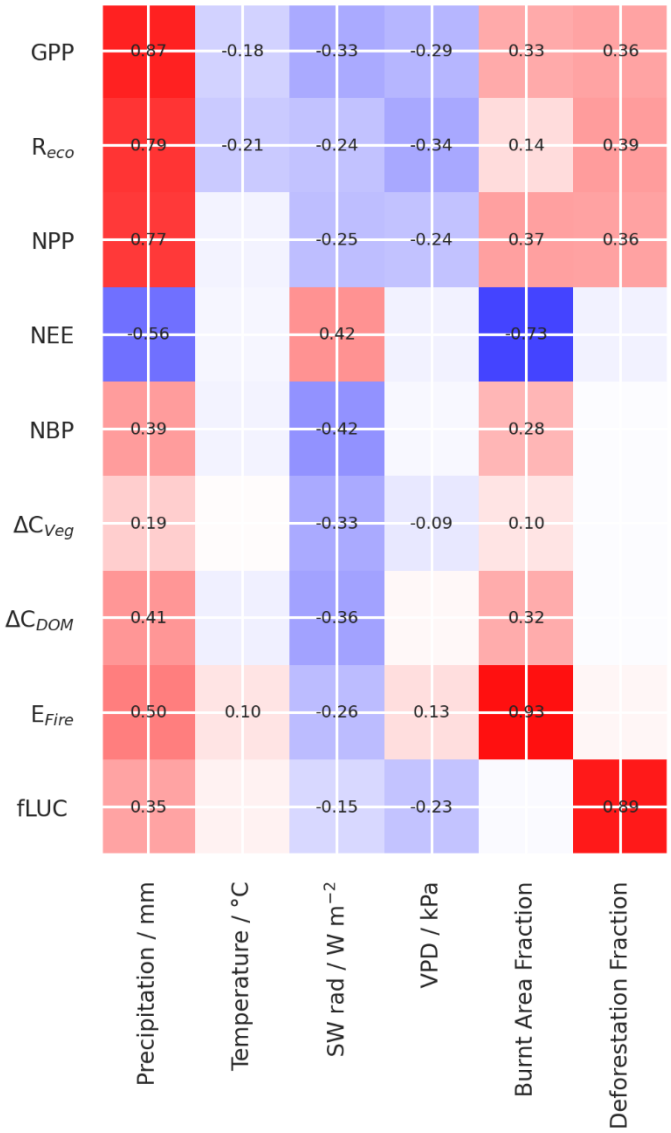


Figure S 6. Correlations between mean biosphere-atmosphere fluxes (annual medians) across SAW and external forcing factors. Colours indicate positive (red) and negative (blue) correlations, with intensity linked to magnitude of the correlation coefficient. SW rad is short wave radiation. Correlation coefficients are only reported when the correlation is statistically significant ($p<0.001$).

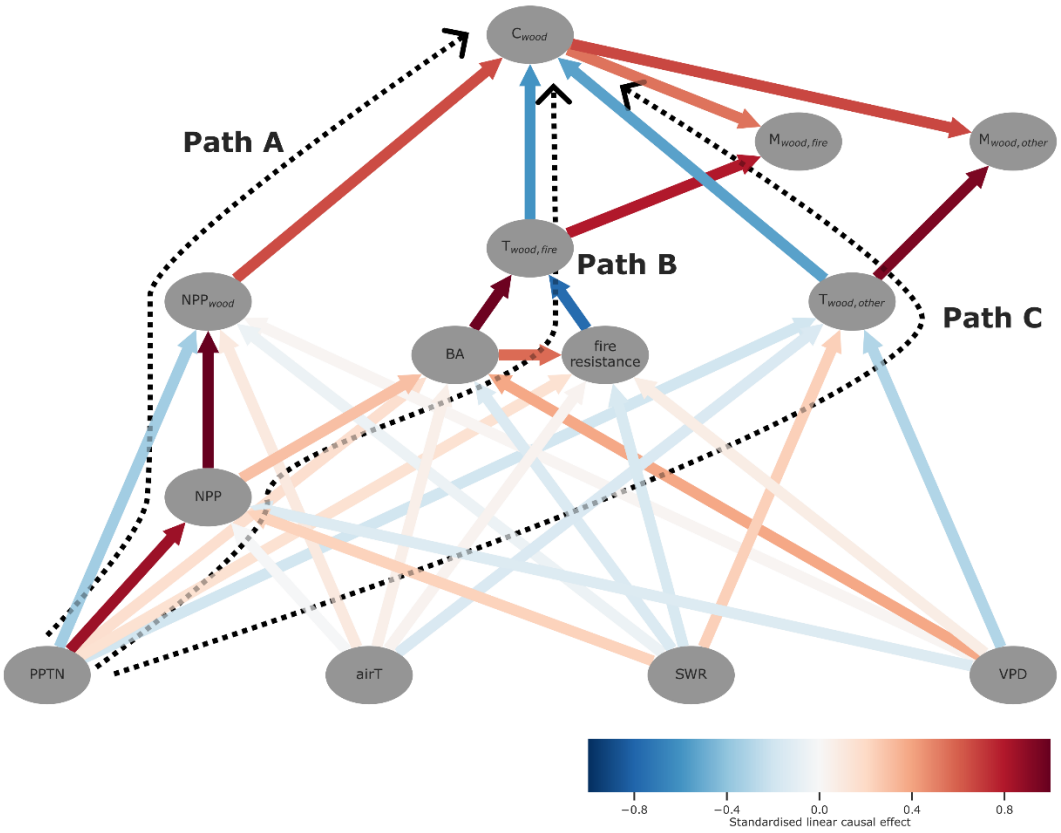


Figure S 7. Causal graph of direct linear causal effect resolving the influence of environmental drivers on the spatial distribution of relative mortality rates and C wood stocks (RQ2). Causal effects were estimated using Wright's path approach. Nodes depict causal variables, arrows show the direction of the direct causal effect, and the colour of the arrows indicates the sign of the direct causal effect (red: positive, blue: negative). To aid interpretation of the causal graph, we break down the overall causal effect of a given driver on the spatial distribution of C_{wood}, we break the overall causal effect into three mediating pathways, representing the total standardised linear effect mediated of a given driver on C_{wood} via: (A) woody productivity, NPP_{wood}, (B) turnover driven by fire, $T_{wood,fire}$, and (C) via non-fire turnover, $T_{wood,other}$. The dotted arrows illustrate these causal pathways for the precipitation driver (PPTN). The total standardised linear effects mediated along these paths are used to summarise the causal effects in Figure 7.

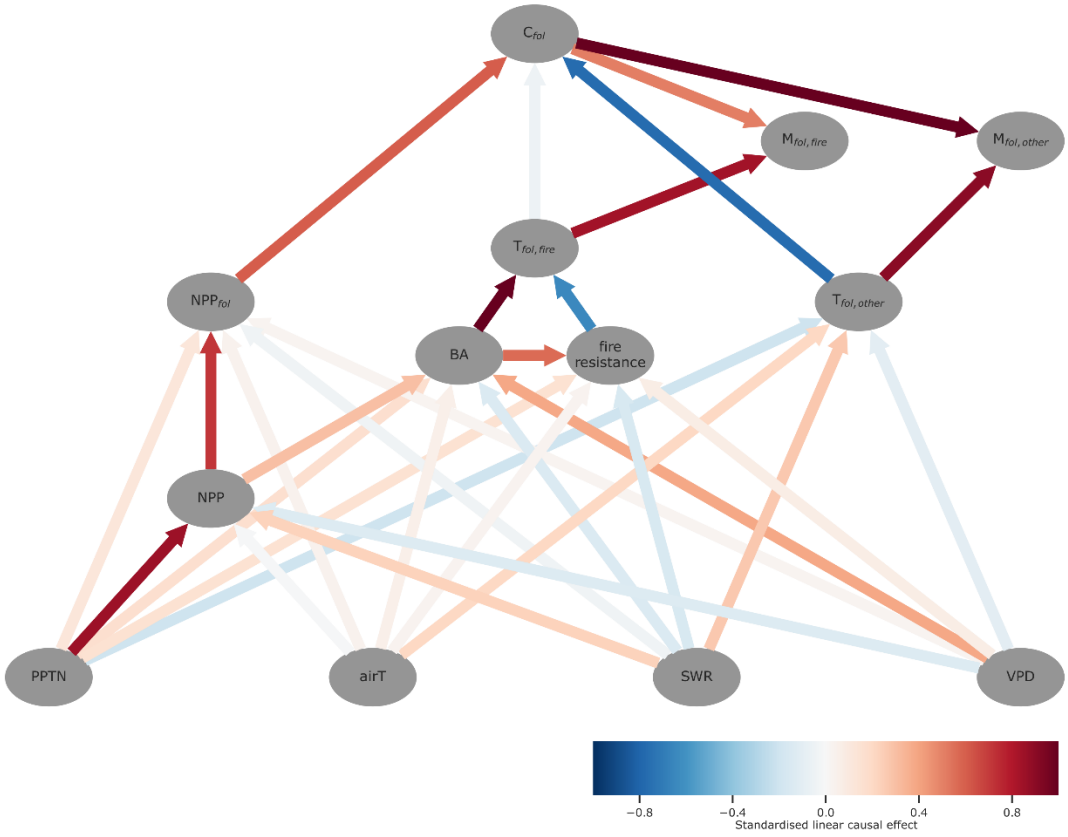


Figure S 8. Causal graph of direct linear causal effect resolving the influence of environmental drivers on the spatial distribution of relative mortality rates and foliage C stocks ($C_{foliage}$) (RQ2). Causal effects were estimated using Wright's path approach. Nodes depict causal variables, arrows show the direction of the direct causal effect, and the colour of the arrows indicates the sign of the direct causal effect (red: positive, blue: negative). Non-fire mortality (i.e. phenology) has a stronger impact on turnover of the foliage pool.

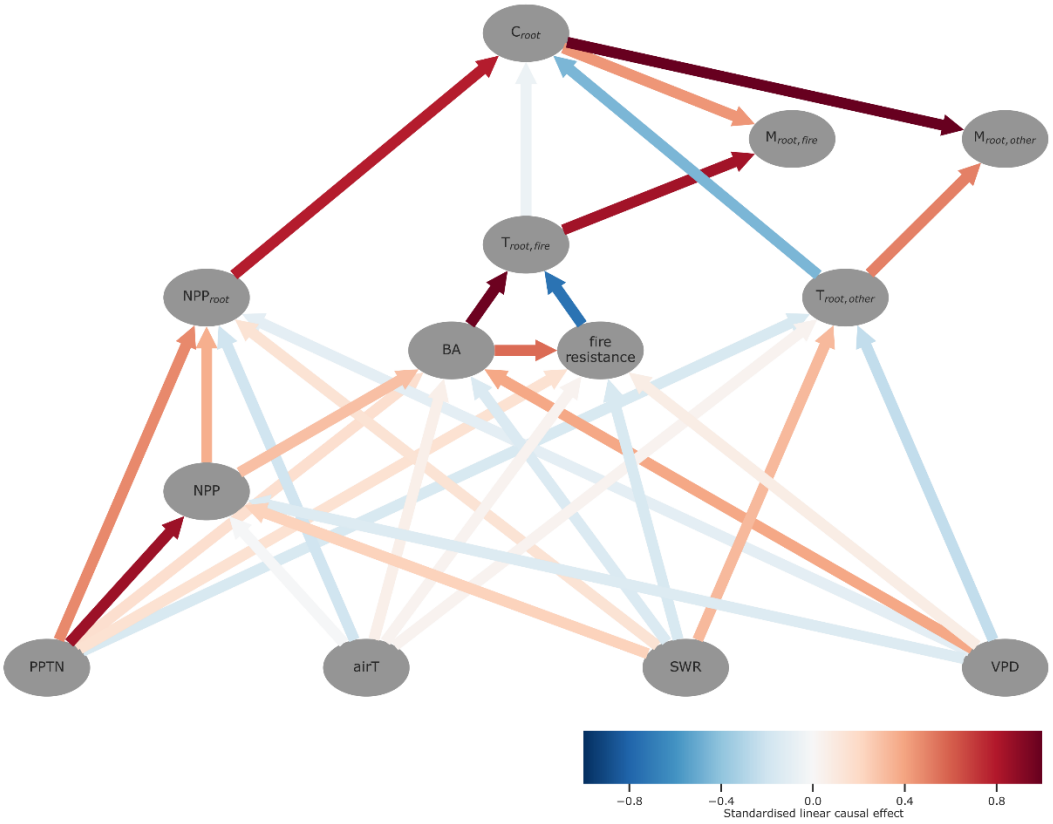


Figure S 9. Causal graph of direct linear causal effect resolving the influence of environmental drivers on the spatial distribution of relative mortality rates and fine root C stocks (C_{roots}) (RQ2). Causal effects were estimated using Wright's path approach. Nodes depict causal variables, arrows show the direction of the direct causal effect, and the colour of the arrows indicates the sign of the direct causal effect (red: positive, blue: negative). Non-fire mortality (i.e. phenology) has a stronger impact on turnover of the fine roots pool.

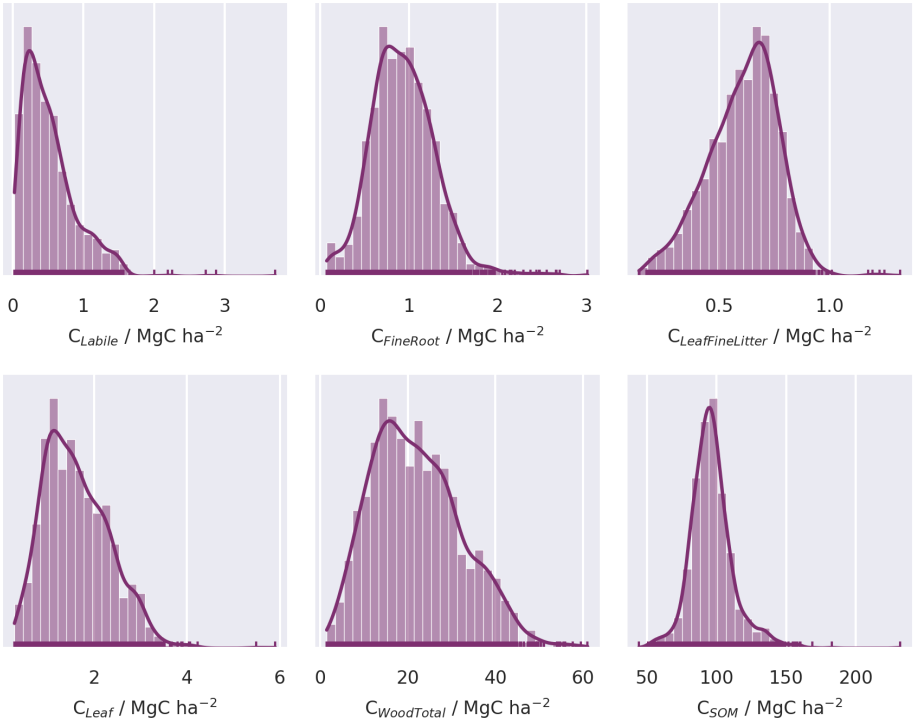


Figure S 10. Probability distributions of live and dead C pools across the SAW region at 0.5º resolution, 2006-2017, as determined by diagnostic analysis.

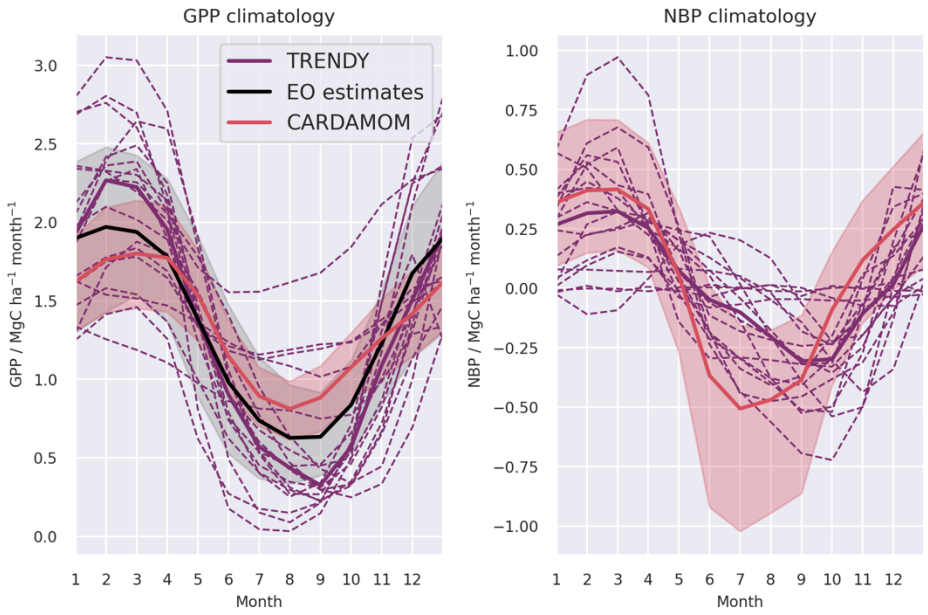


Figure S 11 Comparison of estimated climatologies (mean seasonal cycles) for gross primary production (left) and net biome production (right) over the SAW region 2006-2017. The mean of the Trendy ensemble (solid line) and individual Trendy models (dashed lines) are compared to CARDAMOM (with a 95% confidence interval, red shading). The GPP panel also includes the mean (solid line) and range (shaded area) for the GPP climatology derived from four earth observation orientated GPP products.

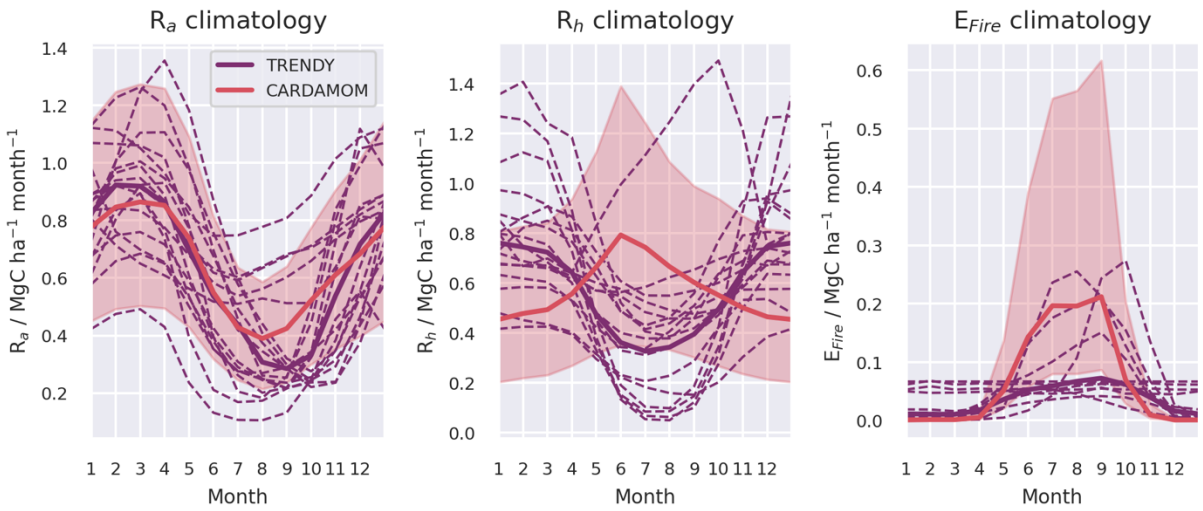


Figure S 12. Mean annual climatologies for main emissions sources from SAW as estimated by CARDAMOM analysis (orange line = median, shading for 95% confidence intervals) and Trendy (purple line =ensemble mean, dashed lines are individual models). R_a is autotrophic respiration, R_h is heterotrophic respiration, E_{Fire} is fire-related C emissions.

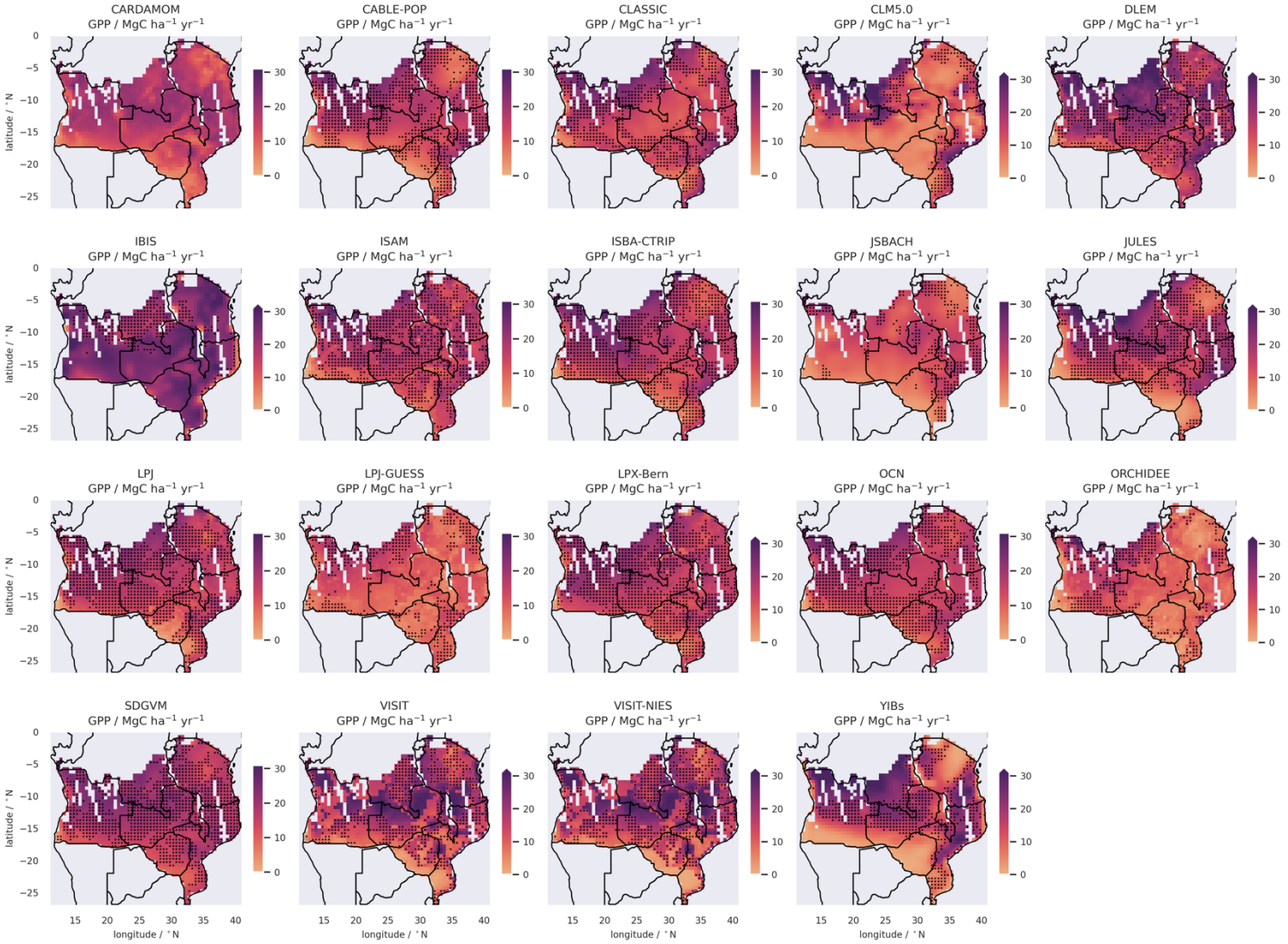
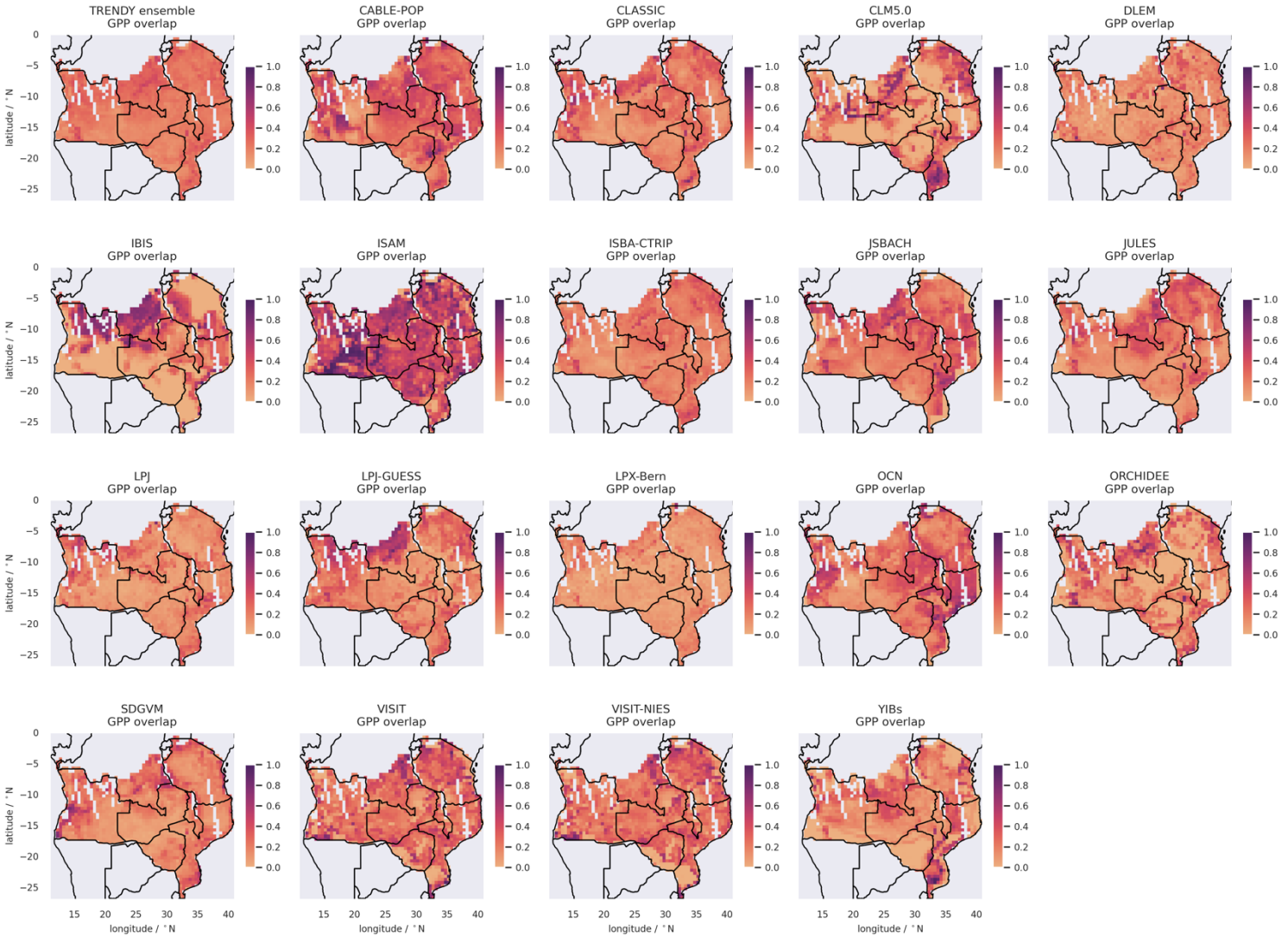
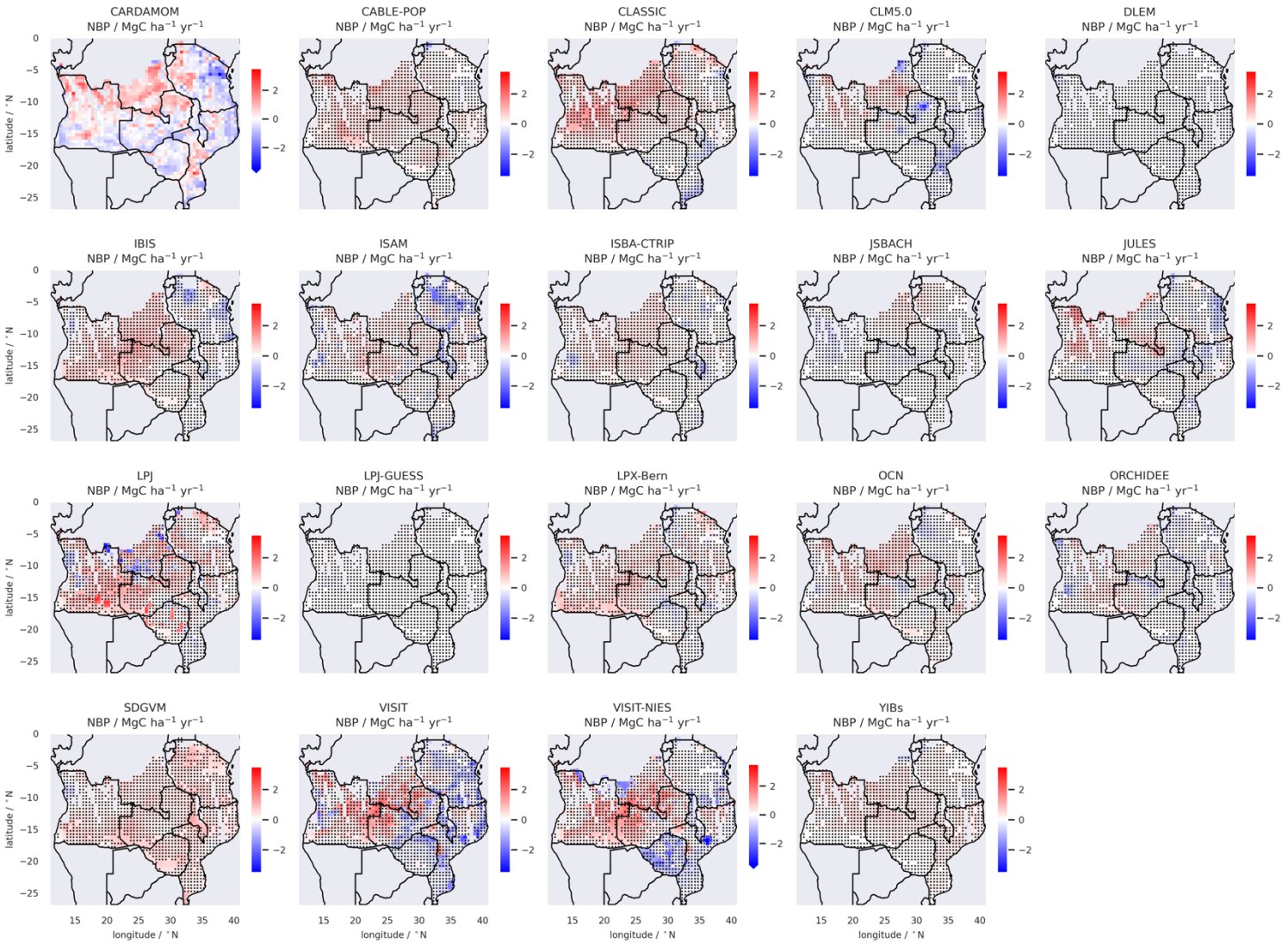

CARDAMOM
CABLE-POP
CLASSIC
CLM5.0
DLEM
IBIS
ISAM
ISBA-CTRIP
JSBACH
JULES
LPJ
LPJ-GUESS
LPX-Bern
OCN
ORCHIDEE
SDGVM
VISIT
VISIT-NIES
YIBs
GPP / MgC ha^{-1} yr^{-1}
latitude / °N
longitude / °N

Figure S 13. Time-averaged GPP (2006-2017) simulated by CARDAMOM and the TRENDY LSMs. Stippling indicates that the time-averaged LSM estimates fall within the 95% CI of the data-constrained CARDAMOM estimate.


TRENDY ensemble
GPP overlap
CABLE-POP
GPP overlap
CLASSIC
GPP overlap
CLM5.0
GPP overlap
DLEM
GPP overlap
IBIS
GPP overlap
ISAM
GPP overlap
ISBA-CTRIP
GPP overlap
JSBACH
GPP overlap
JULES
GPP overlap
LPJ
GPP overlap
LPJ-GUESS
GPP overlap
LPX-Bern
GPP overlap
OCN
GPP overlap
ORCHIDEE
GPP overlap
SDGVM
GPP overlap
VISIT
GPP overlap
VISIT-NIES
GPP overlap
YIBs
GPP overlap
latitude / ˚N
longitude / ˚N
0
−5
−10
−15
−20
−25
15
20
25
30
35
40
1.0
0.8
0.6
0.4
0.2
0.0

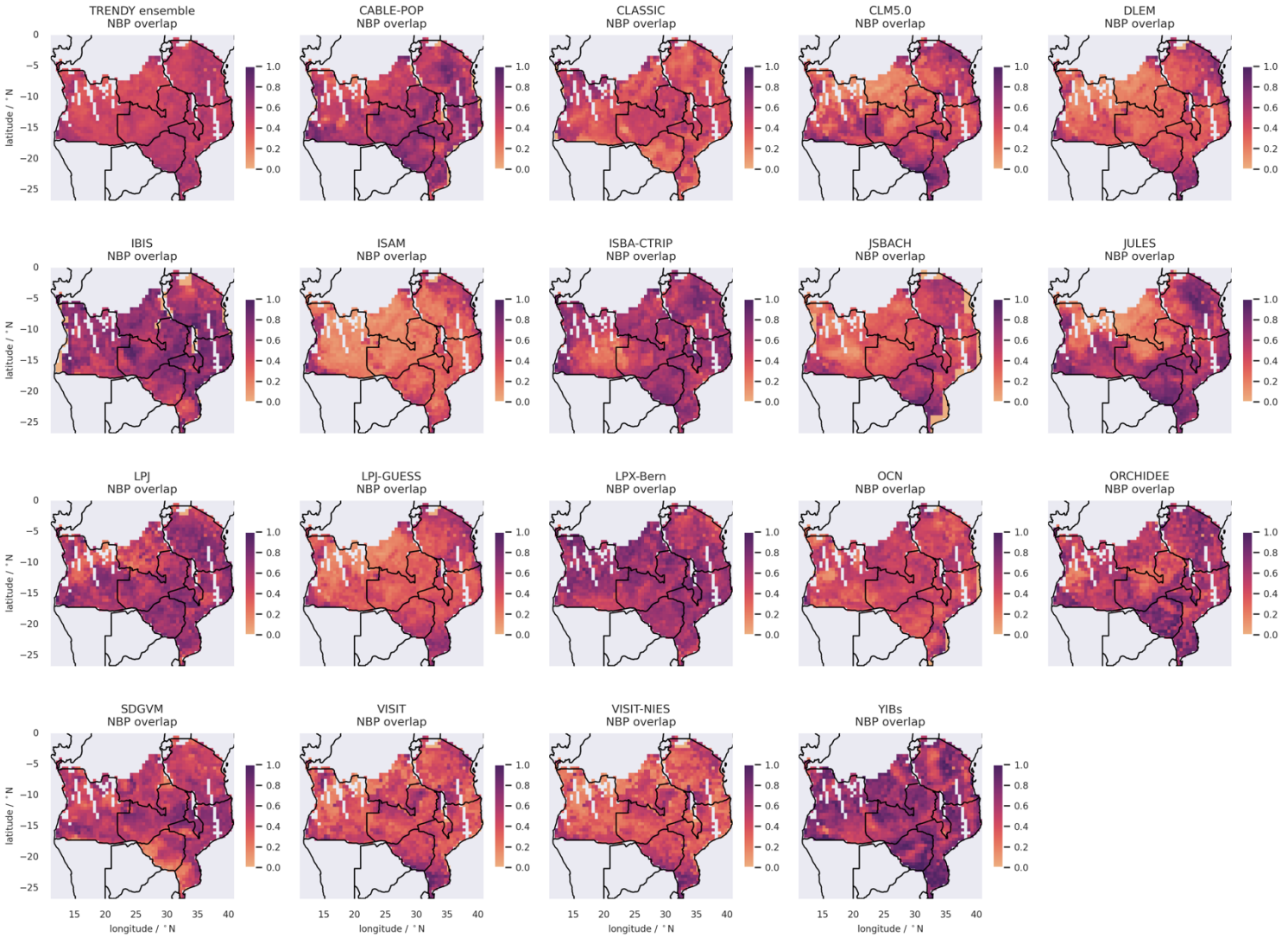
1

Figure S 14. Average overlap fraction (fraction of time per pixel for which flux estimate falls within 95% CI from CARDAMOM) for GPP, averaged across all TRENDY LSMs, and for each of the individual LSMs.


CARDAMOM
CABLE-POP
CLASSIC
CLM5.0
DLEM
IBIS
ISAM
ISBA-CTRIP
JSBACH
JULES
LPJ
LPJ-GUESS
LPX-Bern
OCN
ORCHIDEE
SDGVM
VISIT
VISIT-NIES
YIBs
NBP / MgC ha^{-1} yr^{-1}
latitude / ˚N
longitude / ˚N

1

Figure S 15. Time-averaged NBP (2006-2017) simulated by CARDAMOM and the TRENDY LSMs. Stippling indicates that the time-averaged LSM estimates fall within the 95% CI of the data-constrained CARDAMOM estimate.


TRENDY ensemble
NBP overlap
CABLE-POP
NBP overlap
CLASSIC
NBP overlap
CLM5.0
NBP overlap
DLEM
NBP overlap
IBIS
NBP overlap
ISAM
NBP overlap
ISBA-CTRIP
NBP overlap
JSBACH
NBP overlap
JULES
NBP overlap
LPJ
NBP overlap
LPJ-GUESS
NBP overlap
LPX-Bern
NBP overlap
OCN
NBP overlap
ORCHIDEE
NBP overlap
SDGVM
NBP overlap
VISIT
NBP overlap
VISIT-NIES
NBP overlap
YIBs
NBP overlap
latitude / °N
longitude / °N

1

Figure S 16. Average overlap fraction (fraction of time per pixel for which flux estimate falls within 95% CI from CARDAMOM) for NBP, averaged across all TRENDY LSMs, and for each of the individual LSMs.

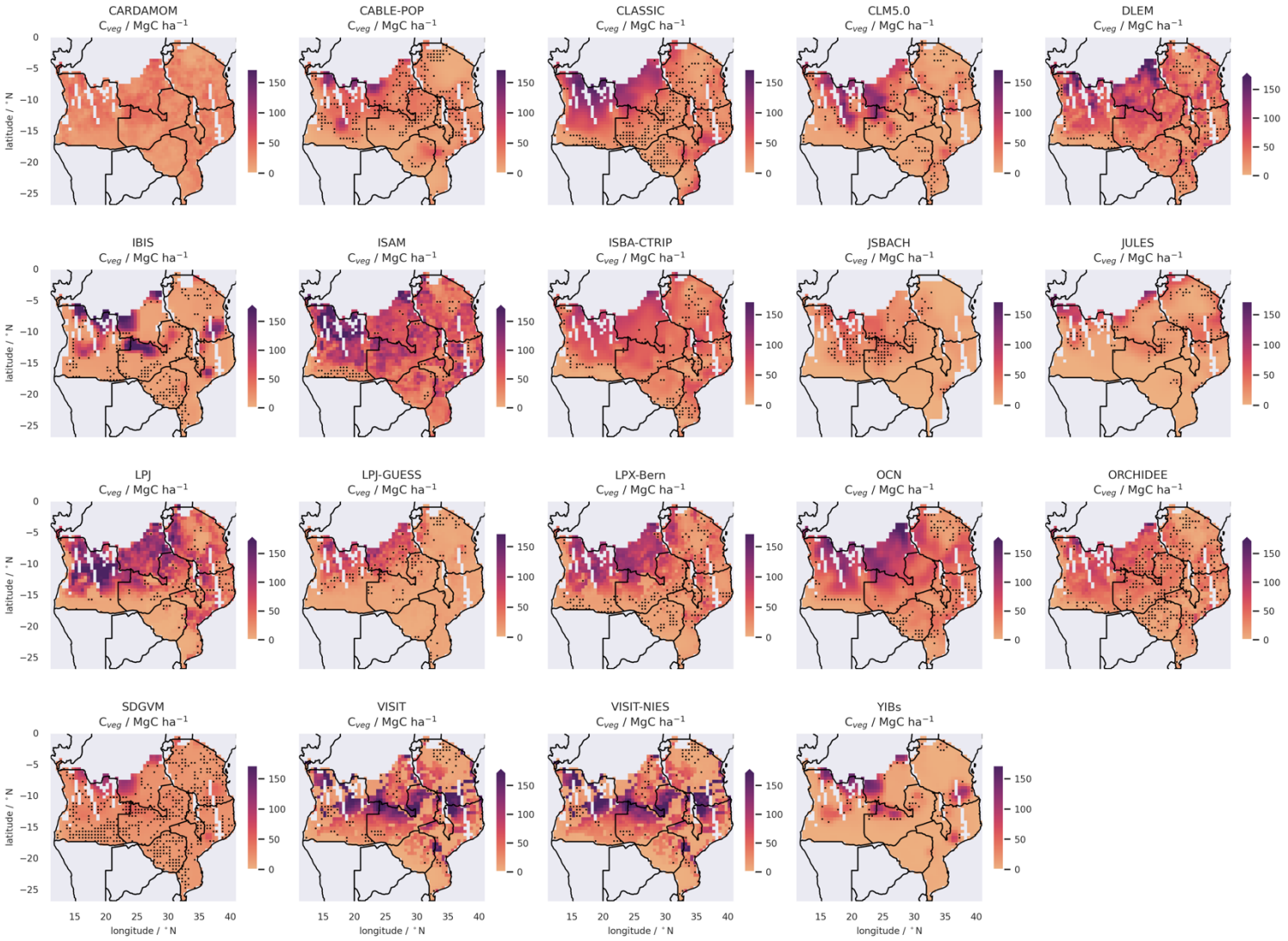

CARDAMOM
CABLE-POP
CLASSIC
CLM5.0
DLEM
IBIS
ISAM
ISBA-CTRIP
JSBACH
JULES
LPJ
LPJ-GUESS
LPX-Bern
OCN
ORCHIDEE
SDGVM
VISIT
VISIT-NIES
YIBs
C_{veg} / MgC ha^{-1}
latitude / °N
longitude / °N
0
−5
−10
−15
−20
−25
15
20
25
30
35
40
150
100
50
0

Figure S 17. Time-averaged live vegetation C stock (2006-2017) across the SAW simulated by CARDAMOM and the TRENDY LSMs. Stippling indicates that the time-averaged C stocks for a given LSM fall within the 95% CI of the data-constrained CARDAMOM estimate.

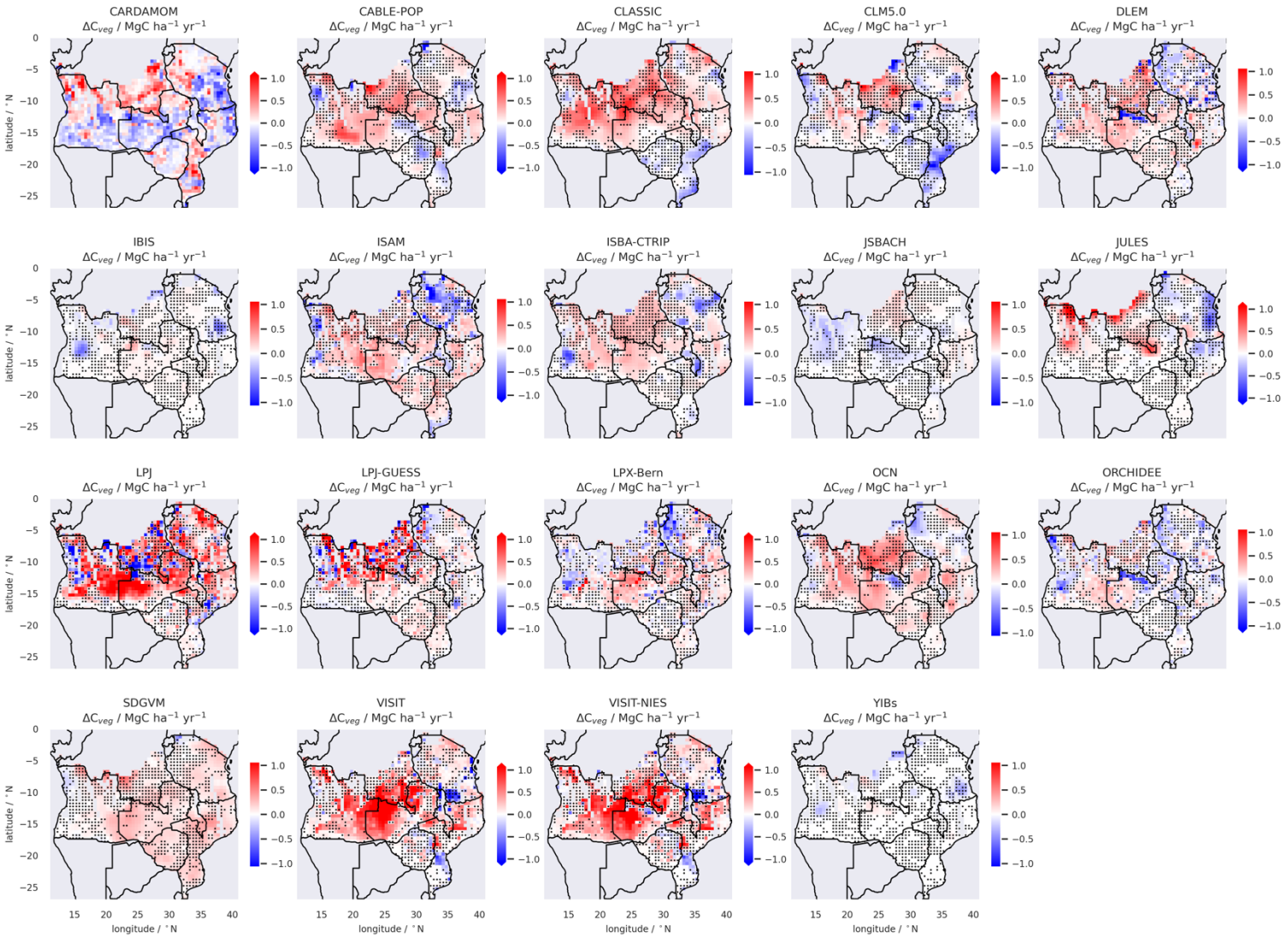


Figure S 18. Time-averaged rate of change in the live vegetation C stock (2006-2017) simulated by CARDAMOM and the TRENDY LSMs. Stippling indicates that the time-averaged LSM estimates fall within the 95% CI of the data-constrained CARDAMOM estimate.

Model	CARDAMOM median (95% CI)	Trendy median (range)	CABLE-POP*	CLASS IC	CLM5.0	DLEM*	IBIS	ISAM*	ISBA-CTRIP	JSBACH	JULES	LPJ	LPJ-GUESS	LPX-Bern	OCN*	ORCHIDEE*	SDGVM	VISIT	VISIT-NIES	YIBs*	EO estimates mean (range)
C_{veg} (MgC ha^{-1})	26 (22-31)	37 (15-66)	**26**	43	**27**	48	32	66	43	15	17	55	21	42	51	32	**26**	51	51	17	-
C_{dom} (MgC ha^{-1})	98 (58-142)	93 (31-194)	**75**	**102**	**59**	**77**	**84**	**93**	**98**	31	**137**	**81**	**101**	194	**111**	**72**	**92**	154	**93**	**90**	-
GPP (MgC ha^{-1} y^{-1})	16.00 (13.06-18.73)	15.78 (10.52-23.22)	**14.59**	**15.17**	12.13	19.33	23.22	**16.00**	**15.99**	11.56	**15.76**	**15.80**	10.52	**16.15**	**15.49**	10.62	**17.94**	**17.11**	**16.98**	**14.88**	**15.69** (10.63-20.76)
GPP seasonality (MgC ha^{-1} y^{-1})	10.38	17.96 (5.66-31.85)	18.22	23.39	5.66	31.25	8.02	5.77	23.59	14.25	20.78	26.14	14.33	31.85	17.70	14.64	28.28	7.14	6.29	20.08	15.23
NBP (MgC ha^{-1} y^{-1})	0.05 (-1.69-1.64)	0.14 (-0.04-0.39)	**0.26**	**0.35**	**-0.00**	**-0.04**	**0.25**	**0.04**	**0.12**	**0.02**	**0.07**	**0.32**	**-0.00**	**0.20**	**0.18**	**0.02**	**0.39**	**0.04**	**0.16**	**0.19**	-
NBP seasonality (MgC ha^{-1} y^{-1})	10.20	5.96 (-1.97-16.56)	9.77	16.56	1.51	0.04	10.33	-1.97	11.98	3.07	4.58	5.45	0.01	8.58	7.02	6.47	10.08	1.97	0.98	10.19	-
E_{Fire} (MgC ha^{-1} y^{-1})	0.89 (0.36-2.52)	0.72 (0.31-1.04)	-	1.04	0.87	-	NA	-	1.04	0.32	0.72	0.75	0.58	0.65	-	-	0.80	0.46	0.31	-	-
ΔC_{veg} (MgC ha^{-1} y^{-1})	-0.00 (-0.38-0.36)	0.06 (-0.02-0.24)	**0.12**	**0.19**	**-0.01**	**0.05**	**0.00**	**0.06**	**0.07**	**-0.02**	**0.05**	**0.21**	**0.10**	**0.03**	**0.13**	**-0.02**	**0.14**	**0.23**	**0.24**	**-0.01**	-
ΔC_{dom} (MgC ha^{-1} y^{-1})	-0.01 (-1.51-1.44)	0.06 (-0.03-0.30)	**0.09**	**0.06**	**-0.02**	**0.02**	**0.10**	**-0.03**	**-0.01**	**0.02**	**-0.02**	**0.09**	**0.00**	**0.11**	**0.05**	**0.01**	**0.09**	**0.30**	**0.21**	**0.16**	-

Table S1: Summary comparison of key components of the terrestrial C cycle for the SAW region estimated by CARDAMOM and the Trendy v11 models regarding. Values indicate spatially and temporally averaged estimates. Models marked with a * are members of the Trendy v11 ensemble that do not include process representation of fire. Seasonality is the absolute difference in fluxes between the wet (JFM) and dry (JAS) season months.

Model	Trendy	CABLE-POP*	CLASSIC	CLM5.0	DLEM*	IBIS	ISAM*	ISBA-CTRIP	JSBACH	JULES	LPJ	LPJ-GUESS	LPX-Bern	OCN*	ORCHIDEE*	SDGVM	VISIT	VISIT-NIES	YIBs*	Overlap category
C_{veg} overlap	0.14	0.21	0.22	0.13	0.12	0.17	0.07	0.12	0.12	0.10	0.07	0.10	0.15	0.13	0.25	0.35	0.08	0.08	0.05	A
C_{dom} overlap*	0.62	0.72	0.80	0.25	0.74	0.92	0.83	0.93	0.00	0.47	0.85	0.78	0.14	0.86	0.64	0.88	0.37	0.69	0.34	A
GPP Overlap	0.23	0.31	0.22	0.19	0.15	0.22	0.48	0.20	0.26	0.22	0.16	0.22	0.10	0.27	0.17	0.20	0.27	0.26	0.17	B
NBP Overlap	0.49	0.57	0.42	0.44	0.38	0.64	0.29	0.59	0.41	0.51	0.60	0.39	0.64	0.45	0.57	0.53	0.42	0.39	0.68	B
E_{Fire} overlap	0.11	0.00	0.20	0.34	-	-	-	0.22	0.16	0.21	0.17	0.16	0.15	-	-	0.15	0.14	0.13	-	B
ΔCveg Overlap	0.53	0.53	0.48	0.57	0.56	0.60	0.51	0.57	0.59	0.59	0.39	0.50	0.52	0.52	0.54	0.56	0.45	0.44	0.63	C
ΔC_{dom} Overlap	0.96	0.96	0.97	0.98	0.98	0.93	0.98	0.98	0.93	0.98	0.94	0.98	0.94	0.95	0.98	0.96	0.91	0.92	0.97	C

Table S2: Summary comparison of the consistency between key components of the terrestrial C cycle for the SAW region estimated by CARDAMOM and the Trendy v11 models regarding. Consistency is calculated based on the fraction of pixels (in space and time) for which the trendy estimate falls within the CARDAMOM 95% confidence interval. The column: Overlap category distinguishes components for which the overlap was estimated based on: (A) the annual estimates, (B) monthly estimates, and (C) time-integrated estimates. Models marked with a * are members of the Trendy v11 ensemble that do not include process representation of fire.