New Phytologist **Supporting Information**

Article title: **Hydraulic traits explain differential responses of Amazonian forests to the 2015 El Nino -induced drought**

Authors: Fernanda V. Barros, Paulo R. Bittencourt, Mauro Brum, Natalia Restrepo-Coupe, Luciano Pereira, Grazielle S. Teodoro, Scott Saleska, Laura S. Borma, Bradley O. Christoffersen, Deliane Penha, Luciana F. Alves, Adriano J.N. Lima, Vilany M.C. Carneiro, Pierre Gentine, Jung-Eun Lee, Leila S. M. Leal, Alessandro C. Araujo, Luiz E. O. C. Aragão, Valeriy Ivanov, Rafael S. Oliveira.

Article acceptance date: Click here to enter a date.

The following Supporting Information is available for this article

Table S1. List of species name, family and mean hydraulic traits value for all studied species at low seasonal forest (LSF) and high seasonal forest (HSF).

Table S2. Summary of hydraulic traits and statistical results of hypothesis 1, that HSF has more drought resistant hydraulic traits than LSF.

Table S3. General mixed model result from the test of hypothesis 2, that species from HSF are less sensitive to ENSO than species from LSF.

Table S4. General mixed model result from the test of our hypothesis 3, that HSF forest is less sensitive to atmospheric drought and soil drought than the LSF forest.

Table S5. General mixed site-specific model results for evapotranspiration (ET) varying in function of atmospheric drought, soil drought.

Methods S1 Species dominance and trait distribution in the communities.

Methods S2. The biogeographic dry affiliation index as a trait to differ LSF and HSF community composition.

Methods S3 Eddy covariance flux measurements.

Methods S4 Canopy conductance calculation.

Methods S5 Statistic functions and packages

Figure S1. Times series of (a) cumulative water deficit (CMWD; mm); (b) rainfall (mm); (c) air temperature (^{o}C); (d) vapour pressure deficit (VPD; kPa); (e) canopy conductance (mm s^{-1}); and (f) evapotranspiration (mm); for the low seasonality forest (LSF, Manaus, blue) and the high seasonality forest (HSF, Tapajos, red). Boxes are the data distribution of the time series for LSF and HSF. Whiskers in a) are either maximum/minimum value or, when outliers are present, 1.5 interquartile range above/bellow the quartiles 2 and 3.

Figure S2. Monthly mean soil water content (%) as a function of monthly cumulative water deficit (CMWD; mm) for (a) low seasonality forest (LSF, Manaus) and, (b) high seasonality forest (HSF, Tapajos). Soil water content time series, available for (c) LSF in 2015 and, (d) HSF. Red lines are the soil water content at a depth of 0.8m for LSF and 0.5m for HSF, green lines are 1.6m in LSF and 1m in HSF and blue line is 2.4m in LSF and 2m in HSF. Black lines in (a) and (b) are the best-fit linear model for LSF ($F(1,9) = 25.6$; $r^2 = 0.74$; $p < 0.001$) and HSF ($F(1,52) = 154.5$; $r^2 = 0.75$; $p < 0.001$). Soil water content data was not calibrated, so while patterns are reliable, absolute values should be evaluated with care.

Figure S3. The relationship between monthly mean canopy conductance (G_S) and $CMWD_r$ in the (a) low seasonality forest (LSF, Manaus), and (b) high seasonality forest (HSF, Tapajos); and canopy conductance (Gs) and VPD_r for (c) LSF and (d) HSF. The $CMWD_r$, and VPD_r correspond to the residuals of monthly mean VPD (vapour pressure deficit) and monthly CMWD (cumulative water deficit) after removing the correlation between VPD and CMWD. The data correspond to the period from July to December. The colour of the data points is proportional to the VPD_r value in plots (a) and (b) and to the $CMWD_r$ values in plots (c) and (d), according to the colour charts bellow the panels. Note both sites were modelled together and the p-value in panel (a) also applies to panel (b) and p and R^2m values in panel (c) also apply to panel (d). Triangles points are data from 2015 ENSO. Circles points are monthly data from 1999 to 2016, excluding the 2015 ENSO period.

Figure S4. The relationship between monthly mean evapotranspiration (ET) and CMWD in the

(a) low seasonality forest (LSF, Manaus), and (b) high seasonality forest (HSF, Tapajos); and evapotranspiration (ET) and VPD for (c) LSF and (d) HSF. The data correspond to the period from July to December. The colour of the data points is proportional to the VPD value in plots (a) and (b) and to the CMWD values in plots (c) and (d), according to the colour charts bellow the panels. Note both sites were modelled together and the p-value in panel (a) also applies to panel (b) and p and R^2m values in panel (c) also apply to panel (d). Triangles are data from 2015 ENSO. Circles are monthly data from 1999 to 2016, excluding the 2015 ENSO period.

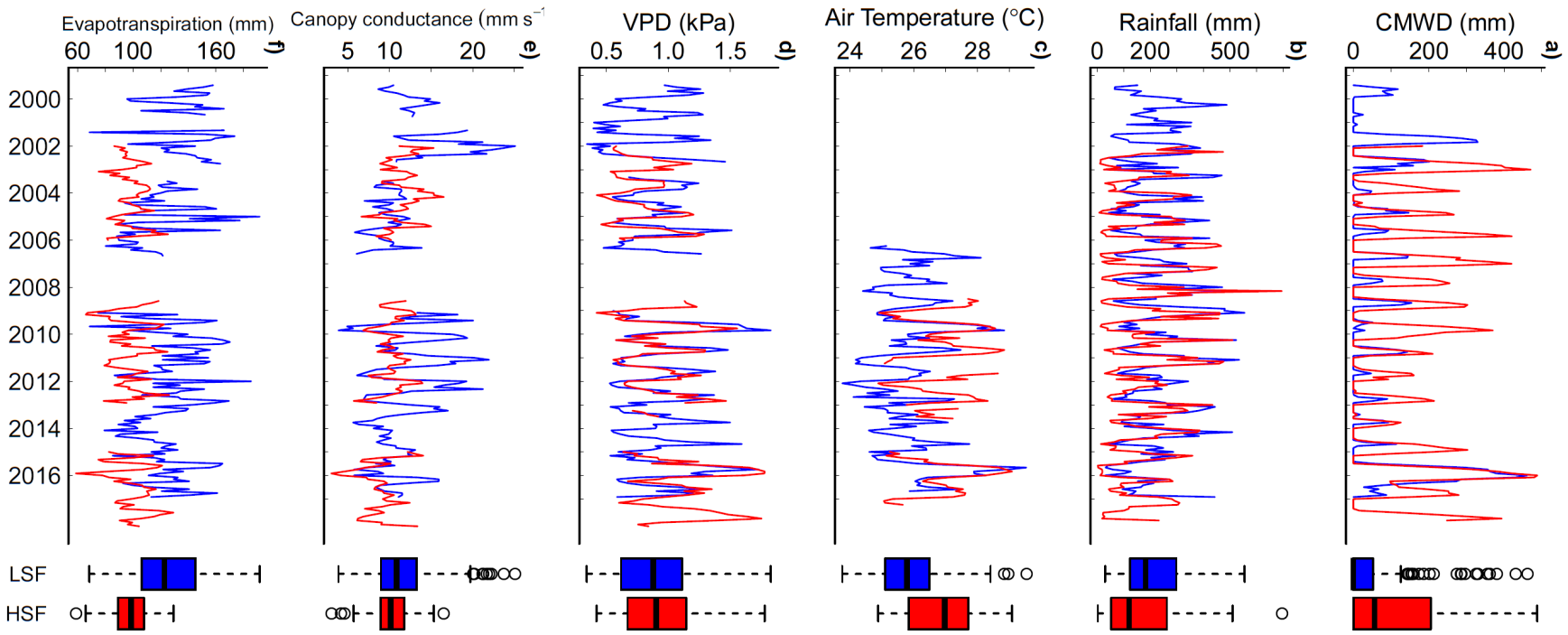
Evapotranspiration (mm)
60 100 160
f)
Canopy conductance (mm s⁻¹)
5 10 20
e)
VPD (kPa)
0.5 1.0 1.5
d)
Air Temperature (°C)
24 26 28
c)
Rainfall (mm)
0 200 500
b)
CMWD (mm)
0 200 400
a)
2000
2002
2004
2006
2008
2010
2012
2014
2016
LSF
HSF

Figure S1. Times series of (a) cumulative water deficit (CMWD; mm); (b) rainfall (mm); (c) air temperature (^{o}C); (d) vapour pressure deficit (VPD; kPa); (e) canopy conductance (mm s^{-1}); and (f) evapotranspiration (mm); for the low seasonality forest (LSF, Manaus, blue) and the high seasonality forest (HSF, Tapajos, red). Boxes are the data distribution of the time series for LSF and HSF. Whiskers in a) are either maximum/minimum value or, when outliers are present, 1.5 interquartile range above/bellow the quartiles 2 and 3.

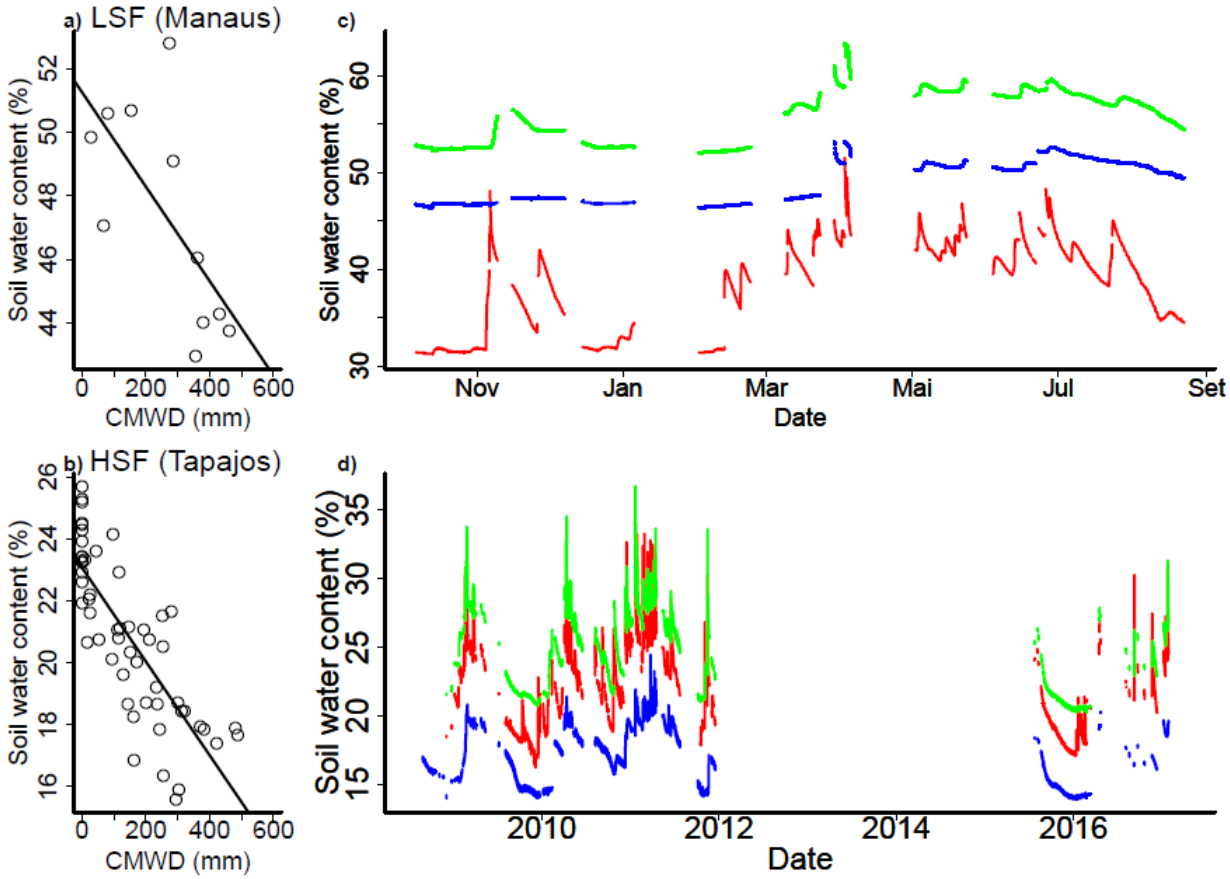

Figure S2. Monthly mean soil water content (%) as a function of monthly cumulative water deficit (CMWD; mm) for (a) low seasonality forest (LSF, Manaus) and, (b) high seasonality forest (HSF, Tapajos). Soil water content time series, available for (c) LSF in 2015 and, (d) HSF. Red lines are the soil water content at a depth of 0.8m for LSF and 0.5m for HSF, green lines are 1.6m in LSF and 1m in HSF and blue line is 2.4m in LSF and 2m in HSF. Black lines in (a) and (b) are the best-fit linear model for LSF ($F_{(1,9)} = 25.6$; $r^2 = 0.74$; $p < 0.001$) and HSF ($F_{(1,52)} = 154.5$; $r^2 = 0.75$; $p < 0.001$). Soil water content data was not calibrated, so while patterns are reliable, absolute values should be evaluated with care.

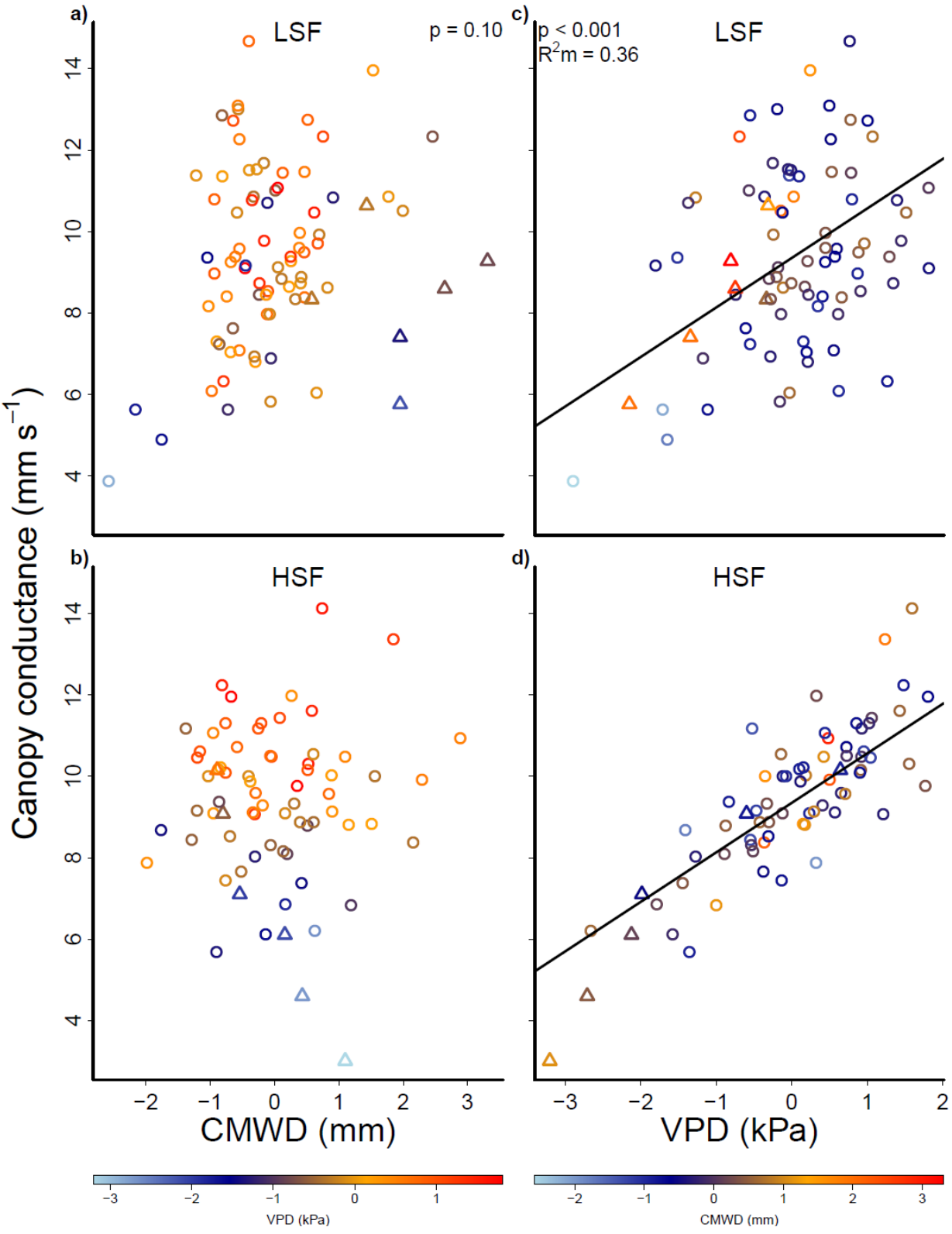

Figure S3. The relationship between monthly mean canopy conductance (G_S) and $CMWD_r$ in the (a) low seasonality forest (LSF, Manaus), and (b) high seasonality forest (HSF, Tapajos); and

canopy conductance (Gs) and VPD_r for (c) LSF and (d) HSF. The $CMWD_r$, and VPD_r correspond to the residuals of monthly mean VPD (vapour pressure deficit) and monthly CMWD (cumulative water deficit) after removing the correlation between VPD and CMWD. The data correspond to the period from July to December. The colour of the data points is proportional to the VPD_r value in plots (a) and (b) and to the $CMWD_r$ values in plots (c) and (d), according to the colour charts bellow the panels. Note both sites were modelled together and the p-value in panel (a) also applies to panel (b) and p and R^2m values in panel (c) also apply to panel (d). Triangles points are data from 2015 ENSO. Circles points are monthly data from 1999 to 2016, excluding the 2015 ENSO period.

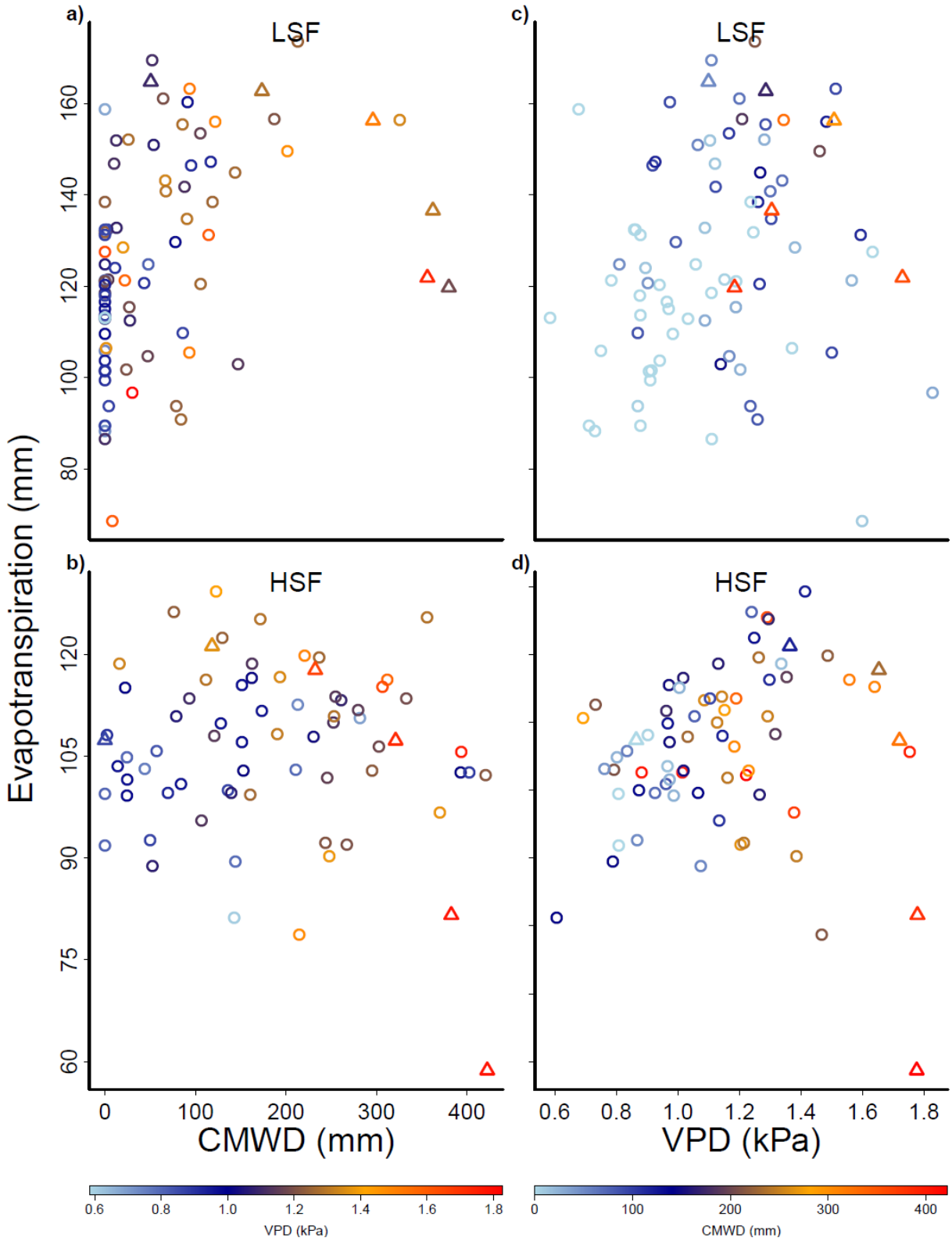

Figure S4. The relationship between monthly mean evapotranspiration (ET) and CMWD in the (a) low seasonality forest (LSF, Manaus), and (b) high seasonality forest (HSF, Tapajos); and

evapotranspiration (ET) and VPD for (c) LSF and (d) HSF. The data correspond to the period from July to December. The colour of the data points is proportional to the VPD value in plots (a) and (b) and to the CMWD values in plots (c) and (d), according to the colour charts bellow the panels. Note both sites were modelled together and the p-value in panel (a) also applies to panel (b) and p and R^2m values in panel (c) also apply to panel (d). Triangles are data from 2015 ENSO. Circles are monthly data from 1999 to 2016, excluding the 2015 ENSO period.

Table S1. List of species name, family and mean hydraulic traits value for all studied species at low seasonal forest (LSF) and high seasonal forest (HSF). The traits abbreviations are described below (*). The sample size (n) for the hydraulic traits evaluated in this study is represented for each species. NA indicates not available data.

Site	Species	Family	*n*	RD	Ψ50	Ψ88	$\Psi_{nonENSO}$	Ψ_{ENSO}	VD	VA	Kh	Dh
	Caryocar glabrum	Caryocaraceae	2	0.01	-1.78	-3.10	-0.71	-1.1	578.2	5.9	22.6	34.6
	Dypterix odorata	Fabaceae	2	0.14	-4.47	-6.22	-0.95	-2.61	285.6	13.2	6.5	30.1
	Eschweilera coriaceae	Lecitidaceae	4	1.77	-1.57	-1.59	-1.41	-1.54	465.8	8.2	10.6	30.1
	Eschweilera cyathiformis	Lecitidaceae	1	0.26	-3.05	-4.74	-0.8	-2.895	NA	NA	NA	NA
	Eschweilera sp.	Lecitidaceae	1	0.54	-2.47	-6.20	-1.71	-2.545	303.0	19.9	3.1	19.4
	Eschweilera wachenheimii	Lecitidaceae	2	3.02	-2.19	-2.86	-1.51	-1.59	337.4	12.1	7.7	29.3
	Goupia glabra	Celastraceae	3	1.2	-2.2	-3.80	-0.57	-1.14	495.2	7.6	16.7	31.0
	Gustavia elliptica	Lecitidaceae	3	0.09	-2.75	-6.64	-1.43	-1.63	890.6	6.2	5.1	20.8
LSF	*Lecyths prancei*	Lecitidaceae	3	0.76	-1.8	-2.13	-2.09	-1.93	384.2	10.0	9.5	30.2
	Maquira sclerophylla	Moraceae	3	0.24	-2.21	-3.82	-1.07	-1.72	773.5	5.9	15.8	26.5
	Minquartia guianensis	Olacaceae	2	1.21	-2.16	-4.56	-1.28	-1.8	611.2	11.1	2.0	16.3
	Ocotea sp.	Lauraceae	2	0.16	-1.84	-3.60	-0.91	-2.16	206.8	17.8	12.8	33.5
	Pouteria anomala	Sapotaceae	3	1.05	-1.01	-1.47	-0.9	-1.35	770.4	5.9	6.9	23.9
	Pouteria erythrochrysa	Sapotaceae	1	0.38	-3.92	-6.96	-0.88	-1.43	NA	NA	NA	NA
	Protium hebetatum	Burseraceae	3	0.64	-1.49	-3.52	-0.71	-1.16	697.9	8.5	5.8	19.3
	Scleronema micranthum	Bombacaceae	3	1.37	-1.77	-1.97	-1.16	-1.13	602.3	5.8	23.9	34.5
	Zygia racemosa	Fabaceae	3	0.85	-3.02	-6.19	-0.49	-1.16	188.1	27.6	3.5	23.0

	Amphyrrhox longifolia	Violaceae	5	1.08	-2.28	-5.77	-2.12	-1.93	1005.1	7.1	1.5	15.1
	Chamaecrista xinguensis	Fabaceae	3	6.15	-3.14	-6.05	-2.68	-2.58	300.5	11.4	9.7	32.9
	Coussarea albescens	Rubiaceae	3	4.61	-4.86	-6.64	-2.25	-3.21	1314.0	5.7	1.6	14.5
	Endopleura uchi	Humiriaceae	2	1.01	-1.52	-4.83	-1.1	-1.62	594.3	7.2	7.9	26.4
	Erisma unsinatum	Vochysiaceae	3	11.07	-2.13	-3.24	-1.22	-1.06	391.7	11.2	5.6	26.5
	Manilkara huberi	Sapotaceae	3	26.96	-1.75	-4.51	NA	NA	NA	NA	NA	NA
HSF												
	Mezilaurus itauba	Lauraceae	2	1.2	-2.98	-5.54	-1.64	-1.98	992.1	4.3	12.3	25.8
	Miconia lepidota	Melastomataceae	3	0.09	-5.02	-6.76	-1.96	-3.01	1337.1	4.9	4.0	18.1
	Minquartia guianensis	Olacaceae		0.06	-2.37	-6.03	NA	NA	NA	NA	NA	NA
	Protium apiculatum	Burseraceae	3	2.15	-1.94	-2.12	-1.38	-1.29	956.1	5.0	7.5	23.3
	Rinourea passourea	Violaceae	5	7.59	-2.99	-7.30	-2.58	-4.43	909.4	9.2	0.6	12.5
	Tachigali chrysophylla	Fabaceae	2	15.77	-3.79	-5.13	NA	NA	NA	NA	NA	NA

* RD: relative dominance (percentage stem basal area of the species in relation to forest tree stem basal area), Ψ_{50} (MPa), Ψ_{88} (MPa), $\Psi_{nonENSO}$: minimum water potential for the non-ENSO year (MPa), Ψ_{ENSO}: minimum water potential for the ENSO 2015 (MPa), VD: Vessel density (number of vessels per mm^2 of xylem area), VA: Vessel area (percentage vessel area xylem area), K_h: Potential specific conductance (kg MPa^{-1} s^{-1} m^{-1}), D_h: Hydraulic diameter (μ)

Table S2. Summary of hydraulic traits and statistical results of hypothesis 1, that HSF has more drought resistant hydraulic traits than LSF: mean (μ), standard deviation (sd), Dominance Weighted mean (DWM) for each forest (LSF - low seasonal forest, HSF - high seasonal forest); and the statistical results of one tailed Welch`s t-test to assess if mean traits from LSF species are less hydraulically resistant than HSF species (p value, df- degrees of freedom, t value). $p_{dwm}$ is the p-value for the same hypothesis as tested by Welch`s t-test, but testing whether the community level trait, estimated from DWM, differs between the two forests. The p_{dwm} is the result of a Monte Carlo method probability distribution with the test statistic being difference in the DWM trait between the two communities (see Analysis section for details). Values in bold represent significant differences at the 95% confidence level ($p<0.05$). The traits abbreviation is described on Table S1.

Traits	**LSF (Manaus; K34)**				**HSF (Tapajos; K67)**				**Statistics**			
	mean	*sd*	*DWM*	*95%*	*mean*	*sd*	*DWM*	*95%*	*p*	*df*	*t*	p_{dwm}
Ψ_{50}	-2.34	0.89	-2.07	0.34	-2.90	1.15	-2.78	0.65	0.09	19.72	1.42	0.058
Ψ_{88}	-4.08	1.83	-3.30	0.87	-5.33	1.49	-5.10	1.08	**0.027**	26.33	2.02	**0.026**
Ψ_{50} - Ψ_{88}	1.74	1.2	1.22	0.61	2.43	1.2	2.32	0.84	0.07	23.7	-1.50	0.055
$\Psi_{nonENSO}$	-1.09	0.43	-1.20	0.23	-1.88	0.58	-1.96	0.55	**0.002**	12.80	3.59	**0.030**
Ψ_{ENSO}	-1.70	0.56	-1.55	0.22	-2.35	1.07	-2.44	1.11	0.06	10.39	1.69	0.122
HSM_{P50}	1.24	1.05	0.87	0.44	1.10	1.05	0.91	0.62	0.38	16.41	0.32	0.20
HSM_{P88}	2.99	1.95	2.10	0.97	3.48	1.32	3.25	1.07	0.77	22.24	-0.76	0.20
VD	506	220	476	100	867	369	692	309	**0.011**	11.48	-2.67	0.053
VA	11.04	6.35	10.87	3.25	7.35	2.69	9.22	2.09	**0.031**	20.44	1.98	0.18

k_h	10.16	6.87	9.76	3.69	5.64	4.05	4.99	3.06	**0.027**	22.00	2.03	0.17
D_h	26.8	6	27.1	3.2	21.7	7	22.4	6.5	**0.04**	15.09	1.86	0.18

Table S3. General mixed model result from the test of hypothesis 2 that species from HSF are less sensitive to ENSO than species from LSF. Model testing p values are the log likelihood significance test of the effect of removing the variable from model. Final significant model with its parameters are presented below model testing. Ψ_{min} is dry season minimum leaf water potential. (1|species) indicates species is a random fixed effect on intercept. R^2m and R^2c are, respectively, marginal and conditional pseudo-R^2. The sample number is 58 for all models.

Model testing	**Response**	**Predictor 1**	**Predictor 2**	**Interaction**
	Ψ_{min}	*ENSO*	*Site*	*ENSO: Site*
p-value		< 0.001	0.003	0.57

Final model: Ψ_{min} ~ ENSO + Site + (1|species)

Parameter	*Value*	*Standard error*
Intercept	-1.12	0.15
ENSO	-0.56	0.12
Site (HSF)	-0.72	0.23
Species	0.49	
R^2m	0.34	
R^2c	0.68	

Model testing	**Response**	**Predictor 1**	**Predictor 2**	**Interaction**
	Ψ_{min}	*VPD*	*Site*	*VPD: site*
p-value		< 0.001	0.001	0.008

Final model: Ψ_{min} ~ VPD + Site + (1|species)

Parameter	*Value*	*Standard error*
Intercept	0.74	0.5
VPD	-1.41	0.32
Site (HSF)	-2.32	0.62
Site (HSF): VPD	1.1	0.41
Species	0.4	

	R^2m	0.32
	R^2c	0.66

Model testing	**Response**	**Predictor 1**	**Predictor 2**	**Interaction**
	Ψ_{min}	*CWD*	*Site*	*CWD: Site*
p-value		< 0.001	0.11	0.31

Final model: Ψ_{min} ~ ENSO + Site + (1|species)

Parameter	*Value*	*Standard error*
Intercept	-0.99	0.17
CWD	-0.0025	0.000489
Species	0.48	
R^2m	0.23	
R^2c	0.64	

Table S4. General mixed model result from the test of our hypothesis 3, that HSF forest is less sensitive to atmospheric drought and soil drought than the LSF forest. Model testing p-values are the log likelihood significance test of the effect of removing the variable from model. Final significant model with its parameters are presented below model testing. CMWD is cumulative water deficit (mm) and VPD is vapour pressure deficit (kPa). $CMWD_r$ are the residual of CMWD after removing the effect of VPD on it and VPD_r are the residual of VPD after removing the CMWD effect on it. (1|Month) indicates month of the year is a random fixed effect on intercept. R^2m and R^2c are, respectively, marginal and conditional pseudo-R^2. Sample number is 158 for all models.

Model testing	**Response**	**Predictor 1**	**Predictor 2**	**Interaction**
	VPD	*CWD*	*Site*	*CWD: Site*
p-value		< 0.001	0.002	0.50

Final model: VPD ~ CWD + Site + (1|Month)

Parameter	*Value*	*Standard error*
Intercept	1.05	0.05
ENSO	-0.0012	0.0001
Site (HSF)	-0.12	0.04
Month	0.11	
R^2m	0.22	
R^2c	0.38	

Model testing	**Response**	**Predictor 1**	**Predictor 2**	**Interaction**
	CWD	*VPD*	*Site*	*VPD: site*
p-value		< 0.001	< 0.001	0.57

Final model: CWD ~ VPD + Site + (1|Month)

Parameter	*Value*	*Standard error*
Intercept	-144.9	37.9
VPD	186.4	28.6
Site (HSF)	113.0	13.4
Month	43.6	
R^2m	0.40	
R^2c	0.53	

Model testing	**Response**	**Predictor 1**	**Predictor 2**	**Interaction**
	Gs	*VPD_r*	*Site*	*VPD_r: Site*
p-value		< 0.001	0.95	0.34

Final model: Gs ~ VPD_r + Site + (1|Month)

Parameter	*Value*	*Standard error*
Intercept	9.35	0.19
VPD_r	-1.28	0.13
Month	0.34	
R^2m	0.36	
R^2c	0.39	

Model testing	**Response**	**Predictor 1**	**Predictor 2**	**Interaction**
	G_s	*CWD_r*	*Site*	*CWD_r: Site*
p-value		0.10	0.95	0.16

Final model: Gs ~ (1|Month)

Parameter	*Value*	*Standard error*
Intercept	9.35	2.04
Month	0.32	
R^2m	0.00	
R^2c	0.02	

Model testing	**Response**	**Predictor 1**	**Predictor 2**	**Interaction**
	G_s	*VPD*	*Site*	*VPD: site*
p-value		< 0.001	0.91	0.08

Final model: CWD ~ VPD + (1|Month)

Parameter	*Value*	*Standard error*
Intercept	15.11	0.61
VPD	-5.08	0.52
Month	0.22	
R^2m	0.40	
R^2c	0.42	

Table S5. General mixed site-specific model results for evapotranspiration (ET) varying in function of atmospheric drought, soil drought. Model testing p-values are the log likelihood significance test of the effect of removing the variable from model. Final significant model with its parameters are presented below model testing. CMWD is cumulative water deficit (mm) and VPD is vapour pressure deficit (kPa). $CMWD_r$ are the residual of CMWD after removing the effect of VPD on it and VPD_r are the residual of VPD after removing the CMWD effect on it. (1|Month) indicates month of the year is a random fixed effect on intercept. R^2m and R^2c are, respectively, marginal and conditional pseudo-R^2. For all models sample number is 154.

Model testing	**Response**	**Predictor 1**	**Predictor 2**	**Interaction**
	ET	VPDr	Site	VPDr: Site
p-value		0.32	<0.001	0.38
Final model: ET ~ Site + (1|Month)				
	Parameter	*Value*	*Standard error*	
	Intercept	127.0	3.14	
	Site	-21.0	2.95	
	Month			
	R^2m	0.23		
	R^2c	0.31		

Model testing	**Response**	**Predictor 1**	**Predictor 2**	**Interaction**
	ET	*$CMWD_r$*	*Site*	*$CMWD_r$: Site*
p-value		0.29	<0.001	0.39
Final model: ET ~Site + (1|Month)				
	Parameter	*Value*	*Standard error*	
	Intercept	127.0	3.14	
	Site	-21.0	2.95	
	Month			
	R^2m	0.23		
	R^2c	0.31		

Methods S1. Species dominance and trait distribution in the communities.

The low seasonal forest (LSF) and the high seasonal forest (HSF) differ in species dominance homogeneity, as described in the main text. The consequence of such difference in our sampling was an unbalance species sampling number and the total basal area correspondent: while in LSF we sampled 17 species, which correspond to 13.7% of this forest basal area, in HSF, only 9 species had a greater representation in the community, 35% of total basal area. The low representability in basal area of the LSF is here discussed in terms of whether the community-weighted mean could be used to represent the LSF community (**the properties of community weighted mean**), and whether an increase in our sampling size (i.e., including mores species) would change the results here presented (if there are no changes in our estimates when **a larger data sample is considered**). In this case we used the data set published by Oliveira et al. 2018, to show that traits are randomly distributed across species at this community.

Community-weighted mean (CWM) is a way to scale species trait for a community by weighting the trait for its representativeness, in terms of biomass, basal stem area, or another index. Estimation of CWM usually requires high coverage of the biomass or basal area of species in a community. However, community-weighted means only differ significantly from regular, non-weighted, community means when: (1) few species dominate the stand and their trait value differs from the community (non-weighted) mean. In this situation, CWM is biased towards dominant species and non-weighted means are more likely to differ from CWM; (2) or there is no dominant species (i.e. when situation (1) does not occur) but there is a relationship between species dominance and the trait evaluated. For example, suppose wood density increase with dominance (i.e. rarer species have denser wood). In this case, the traits of the more dominant species are different from the rare species. In a non-weighted community mean, random sampling of species would equally consider the traits of rare and dominant species, however dominant species have a distinct different trait value from those of rare species, making CWM and non-weighted mean differ.

To exemplify case 1), consider a community with 100 species, each with trait $X \sim N(10,10)$. This community has a dominance $D \sim \text{Beta}(\alpha, \beta)$, where $000.1 <= \alpha, \beta <= 10$. The Beta distribution can have both positive and negative skewness depending on the values of its parameters and changing its α and β parameters generate a range of dominance distributions with

different skewness. For this simulated community, we generated its dominance and trait X value according to the above definitions 100000 times and calculate trait X is CWM and non-weighted mean, as well as the cumulative dominance of the five most dominant species (CD5) (Figure a).

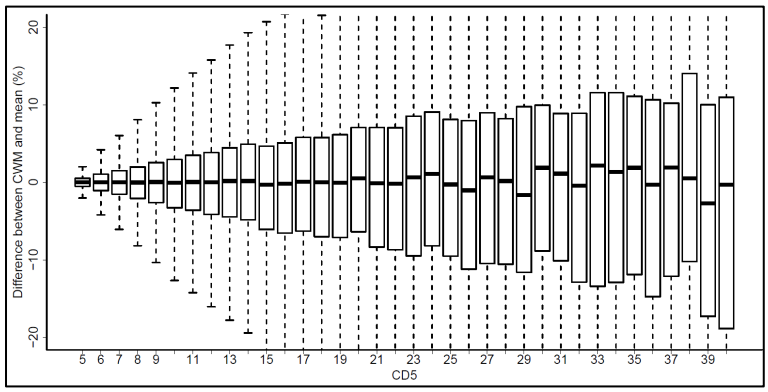

Figure a. Difference between community-weighted mean (CWM) and the non-weighted mean (expressed as a percentage of true CWM), as a function of five most dominant species (CD5) for all simulations.

This figure illustrates how the CWM is most likely to differ from the mean the more the weighting is dominated by a few species. The more the CWM is independent from a few species, the more likely it will be equal to the non-weighted community mean, which can be reasonably estimated from the mean of a subsample. As for the low seasonality forest (LSF), the cumulative dominance of the five dominant species is 10.7%, which suggests CWM is not much different from the regular community mean, as no single species, even if it is an outlier in the analysed trait, would heavily bias the CWM towards its value.

The above example presents a situation where $X \sim N(10,10)$, that is, the standard deviation equals the mean. In the LSF we studied, P50 equals -2.34 ± 0.89 MPa and the standard deviation is 0.38 the mean. Using those values for trait X, and repeating the calculations above, the lower the variability of a trait, the less the CWM differs from the community regular mean, as there is less variability in total in the trait (i.e. it is less likely that a combination of a trait outlier species also is a dominant species) (Figure b).

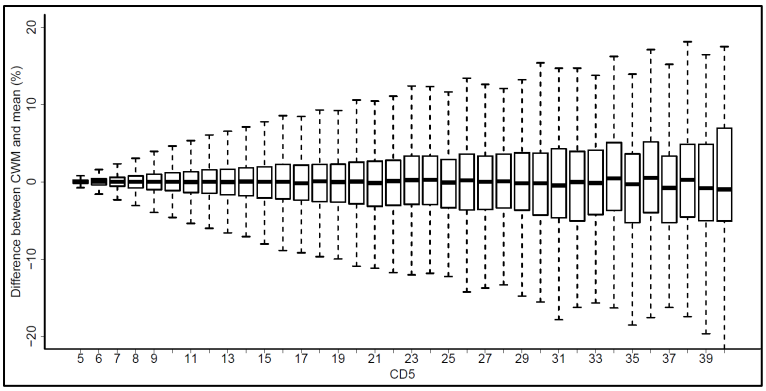

Figure b. Difference between community weighted mean (CWM) and the non-weighted mean (expressed as a percentage of true CWM) as a function of five most dominant species (CD5) for all simulations, for a trait with lower variability.

Finally, in Garnier et al. (2004), one of the key papers to first use CWM to infer ecosystem processes, the community with more species had a value of 12 species. Repeating the first procedure of Figure a, with trait X ~ N (10,10) and a total of 12 species is represented at Figure c. In this case, the CD5 cannot be lower than 42, the situation when the dominances are the most similar possible and when the CWM is most likely to equal the non-weighted mean (i.e. weighted mean equals non-weighted mean when all weights are equal). In the above situation, slightly more dominant species makes the CWM strongly differ from the community regular mean.

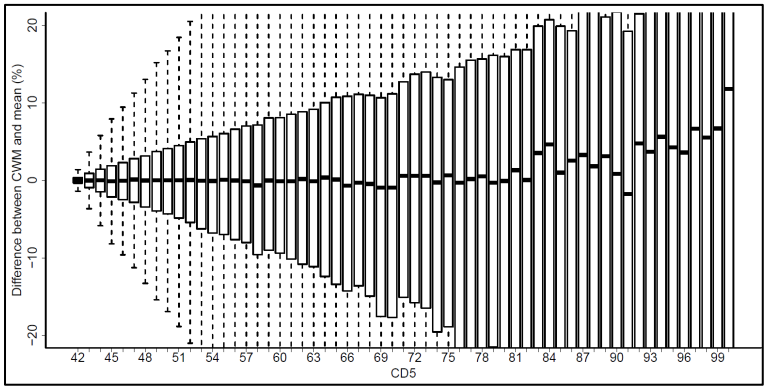

Figure c. Difference between community weighted mean (CWM) and the non-weighted mean (expressed as a percentage of true CWD) as a function of five most dominant species (CD5) for all simulations. Repeating the first procedure, with trait X ~ N (10,10) and a total of 12 species.

In summary, for case (1), we want to highlight that in rich communities where no few species dominate, the CWM and the community mean are unlikely to differ, particularly if the standard deviation of the trait being analysed is low compared to its mean. This is the case for the LSF, which is why we believe our coverage of 13% of the dominance, but 17 species, is a good indicator of the ecosystem function.

Regarding case (2), there is another situation when CWM and the non-weighted mean of a community can differ even if the conditions highlighted in (1) are fulfilled (i.e. species rich community with low trait variability and no few dominant species). This condition is when there is a relationship between the trait analysed and the dominance of the species. Applying the same analysis done before, lets consider the trait value of each species X = Da + N(10,5), where D is the dominance (D ~ Beta (α, β), again) and a represents the intensity of the relationship between dominance and trait X. For a community with 100 species and no very dominant species (CD5 <= 10), we have:

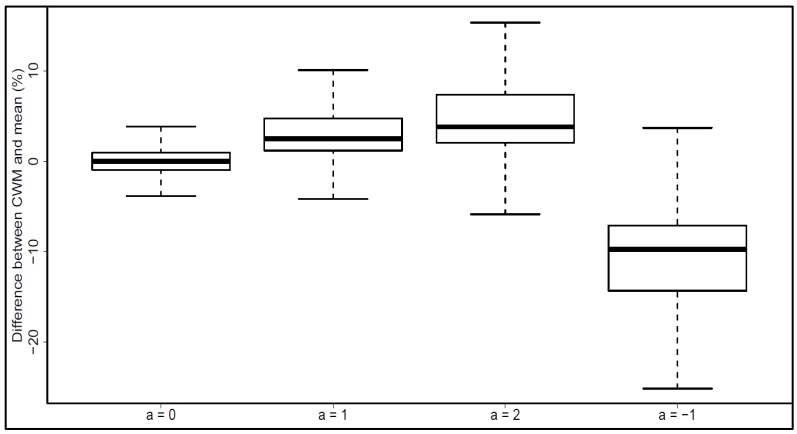

Figure d. Difference between community-weighted mean (CWM) and mean (%) for different situations of dependence between the trait and species dominance.

As can be seen, even for a situation with high trait variability (X ~ N (10,5)), if conditions pointed in (1) hold and dominance and the analysed trait is independent of dominance (a = 0), the difference between CWM and community mean is small. However, even if conditions highlighted in (1) hold, if the analysed trait is related to the dominance of the species, CWM and non-weighted community mean will systematically differ positively or negatively, depending on the whether the relationship between trait and dominance is positive or negative (Figure d). In our analysis, dominance is not related to any of the analysed traits for both, LSF or High

Seasonality Forest (HSF). Thus we conclude that, even if the community mean value we used for up scaling differs from the true CWM, it should not differ much.

Adding dataset of Ducke Reserve (Oliveira et al. 2018) to LSF measurements. The other topic of this method section includes the additional analysis showing that there is no change in our estimates when a larger data sample is considered. Here, we used the embolism data for Ducke Reserve at Manaus, Brazil (Oliveira et al. 2018), an area close to the studied LSF forest (~100 km distance), with the same climate and species composition. We paired this dataset with the species dominance in the LSF site, assuming the species trait is similar in both areas, to test whether increasing the basal area coverage would affect our traits estimative. The current dataset (n = 17, total dominance = 13.7%) has a mean P50 of -2.33 MPa (P50 CWM = -2.07 MPa). The larger database (n = 41, total dominance = 23.23%) has a mean P50 of -2.50 MPa (P50 CWM = -2.38 MPa) and it is not significantly different from the smaller dataset (T-test p = 0.47). According to this analysis, including more information is unlikely to change our results and conclusions.

Finally, we can say the LSF has very few dominant species, with the 5 most dominant species summing 10.67% of total basal area. To have a total basal area of 50% would require sampling at least the 53 most dominant species. We found no evidence that there is a relationship between relative dominance and the studied traits for LSF (linear model p = 0.57) (Figure e). Our data shows P_{50} in LSF is randomly distributed along species with different dominances. In this case, we can say the average is a good estimator of the community-weighted mean, which gives us confidence in our analysis.

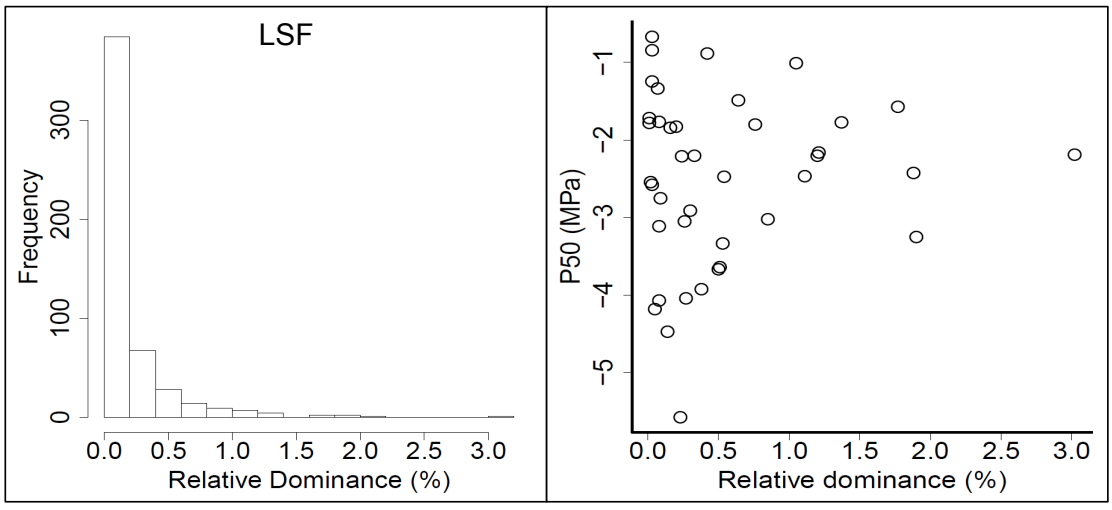

Figure e. Distribution of LSF dominance (left panel) and relationship between relative dominance and the studied traits (here showing P50; MPa) (right panel).

Methods S2. The biogeographic dry affiliation index as a trait to differ LSF and HSF community composition.

In addition to the evaluation of hydraulic trait to differ LSF and HSF communities, we used a new approach with the data published by Esquivel-Muelbert *et al.* (2017). In their study, the authors obtained an index, which considered the genus distribution across other Amazon sites, and represents the probability (p-value) of recording a higher dry-affiliated precipitation centre of gravity (PCG) value than the observed by chance for each different taxa (Esquivel-Muelbert *et al.*, 2017). Thus, if a certain genus has a p-value closer to zero, it meant it is nearly improbable to find another genus that has a higher dry-affiliation, indicating that the genus under consideration was found in drier environments. On the other hand, if a genus had a high p-value (closer to 1), it meant any other genus could have a higher dry-affiliation than it, indicating the genus under consideration is now found is wetter environments.

We used this dry-affiliation index (PCG 2−tail p−value) as a biogeographic trait of the genera. Figure 4a shows how this trait value was related with the gradient of CMWD across Amazonian sites (data from Esquivel-Muelbert *et al.* 2017). As lower is the genus PCG 2−tail p−value, higher is it's dry-affiliation, as it seems to occur in drier environments (higher water deficit - CWD). For each genus inventoried in our HSF and LSF communities, we obtained the respective dry affiliation index from Esquivel-Muelbert *et al.* (2017), and we extended this index as biogeographic trait for the entire community as suggested by Esquivel-Muelbert *et al.* (2018). Most of the studied species for LSF showed higher values of PCG 2−tail p−value, which is associated with low dry-affiliation and wet environments. The figure 4b shows this difference for the whole HSF and LSF community, we can see the lower dry-affiliation of LSF genera (higher PCG 2−tail p−value, p-value < 0.001), when compared with the HSF genera.

Methods S3. Eddy covariance flux measurements.

The ET_m was measured by the eddy-covariance method (Araújo *et al*., 2002; Restrepo-Coupe *et*

al., 2013). Tower observations, during the El Niño dry period, were of good quality with high continuity and daytime ET data completeness from September 2015 through March 2016 was 71.7 % at the K67 tower (HSF) and 63.4% at the K34 (LSF), from a total of 2544 daytime values. The data gaps were filled using the linear regression between incoming shortwave radiation (SW_{in}; W m-2) and ET_m observations by the eddy covariance tower (R^2~0.6, $p<0.01$) (Restrepo-Coupe *et al*., 2016). The ET-SWin relation is stably linear and consistent between dry and wet seasons at these sites (Hasler and Avissar, 2007). Although the relation between ET and SWin changes during extreme droughts, this did not significantly affect our analysis of ENSO because few points were filled by this method during the ENSO. Alternatively, we used satellite derived SW_{in} (Shortwave Flux – All-Sky) from the Clouds and the Earth's Radiant Energy System (CERES) at 1-degree resolution for the 2003-2016 period (Kato et al., 2012; NASA, 2017b) and its relation to ET_m monthly calculations to fill values when neither in situ fluxes or SWin were available. Any remaining missing ET_m values (prior to 2003) were calculated as the mean monthly value from the available ET_m measurements. Mean monthly precipitation was obtained from the Tropical Rainfall Measuring Mission (TRMM) 1999-2016 product (Huffman et al., 2007; NASA, 2017a). A single 0.25 x 0.25 degree cell was considered as representative of the study site.

Methods S4. Canopy conductance calculation.

The G_S was obtained by the inversion of the Penman–Monteith (PM) equation for daytime hours only:

$$ET = \frac{\varepsilon . A + \left(\frac{c_p . \rho_a}{\gamma}\right)(e_s - e_a).G_a}{\varepsilon + 1 + \left(\frac{G_a}{G_s}\right)}$$

where A is the available energy absorbed by the surface (W m^{-2}), the net absorbed radiation minus the soil heat flux, here assumed to be equivalent to the sum of sensible and latent heat flux (H+LE); e_a is the actual vapour pressure (kPa), e_s is the saturation vapour pressure (kPa), ($VPD=e_s - e_a$), γ is the psychrometric coefficient (kPa $ºC^{-1}$), ρ_a is the mean air density (kg m^{-3}), C_p is the specific heat of air at constant pressure (J kg^{-1} $ºC^{-1}$), G_a and G_s, are the aerodynamic and surface conductance (m s^{-1}), respectively; and ε is the unitless ratio between the slope of the saturation vapour pressure versus temperature curve (s; kPa $ºC^{-1}$) and γ ($\varepsilon = s/\gamma$).

The canopy aerodynamic conductance (G_a; m s^{-1}) was calculated as the inverse of aerodynamic resistance (r_a), which was calculated using the resistance to momentum transfer analogy and hence calculated with the expression proposed by Allen et al. (1998) and Verma (1989):

$$r_a = \frac{ln\left(\frac{z-d}{z_{oH}}\right)}{ku^*} = \frac{1}{k^2 u} ln\left(\frac{z-d}{z_o}\right) ln\left(\frac{z-d}{z_{oH}}\right)$$

$$G_a = \frac{1}{r_a}$$

where r_a is the aerodynamic resistance (s m^{-1}), u^* is the friction velocity (m s^{-1}), u the wind velocity (m s^{-1}) measured at the EC height (z = 64 m at HSF and 52 m at LSF), d is the zero-plane displacement at z_o and z_{oH} are the roughness lengths for momentum and heat respectively (m). The quantities d, z_o and z_{oH} were estimated as d=2/3 h, 1/8 h and 1/80 h respectively, where h is canopy height (40 m).

The inversion of the Penman-Monteith equation implies that the available energy equals the sum of latent and sensible heat exchange (energy balance closure, EBC). As at most eddy covariance sites around the world, closure here is incomplete, typically around 80%, due in part to a mismatch between the footprints of the radiation and EC sensors, which is enhanced when low turbulence or advection is present (Leuning et al., 2012; Wilson et al., 2002) or by instrument malfunction (e.g. dirty net radiometer). We addressed this issue by removing monthly flux measurements for those periods when the total turbulent energy (LE and H) deviated from the overall linear regression estimate of LE + H *versus* R_n by 3 standard deviations or more (Barraza et al., 2015). We calculated monthly values using the mean daily cycle of daytime hours for the period of aggregation to reduce the over/under sampling of certain times of day. On the monthly series we expect the energy storage terms (soil and air space between the EC and the surface) to approach to zero, thereby increasing the EBC.

Methods S5 Statistic functions and packages

For all statistical analyses, data processing, and curve fitting, we used R (R Core Team 2018, version 3.5). We used the "t.test" function for the Welch`s t-test (i.e. default of the function), which is more reliable than Student T-Test for samples of different sizes while not requiring variances of the two populations to be equal. We evaluated normality of the data using quantile-

quantile plots and the Shapiro-Wilk test ("shapiro.test" function). We used the "lm" function (base package) for general fixed linear models and the "lme" function ("nlme" package; Pinheiro *et al*. 2014) for general mixed effect models. We followed Zuur *et al*. (2009) and Thomas *et al*. (2017) guidelines for evaluating significance of model terms and validating models assumptions: i) we started with the more complex model and tested the importance of the terms evaluating whether dropping a term significantly affected the model using log-likelihood test with a threshold p value of 0.05; and ii) we evaluated the model assumptions (normality and homogeneity of residuals, collinearity of predictors and bias of influential measured) using diagnostic plots and sample cooks distance and dfbeta. We assessed mode performance using marginal and conditional pseudo-R^2 (R^2m and R^2c, calculate using "r.squaredGLMM" function from the "MuMIn" package; Barton 2016).

References

Allen RG, Pereira LS, Raes D, Smith M, *et al.* 1998. Crop evapotranspiration- Guidelines for computing crop water requirements-FAO. *Irrigation and drainage* **300**: 6541.

Araújo AC, Nobre AD, Kruijt B, Elbers JA, Dallarosa R, Stefani P, Randow C von, Manzi AO, Culf AD, Gash JHC, *et al.* 2002. Comparative measurements of carbon dioxide fluxes from two nearby towers in a central Amazonian rainforest: The Manaus LBA site. *Journal of Geophysical Research* **107**: LBA 58-1–LBA 58-20.

Barraza V, Restrepo-Coupe N, Huete A, Grings F, Van Gorsel E. 2015. Passive microwave and optical index approaches for estimating surface conductance and evapotranspiration in forest ecosystems. *Agric. For. Meteorol.* **213**: 126–137.

Barton K. 2016. MuMIn: multi-model inference. https://cran.r-project.org/web/packages/MuMIn/index.html

Esquivel-Muelbert A, Baker TR, Dexter KG, Lewis SL, Brienen RJW, Feldpausch TR , *et al.* 2018. Compositional response of Amazon forests to climate change. *Global Change Biology* **25:** 39-56.

Esquivel-Muelbert A, Baker TR, Dexter KG, Lewis SL, ter Steege H, Lopez-Gonzalez G, *et al.* 2017. Seasonal drought limits tree species across the Neotropics. *Ecography* **40**: 618–629.

Garnier E, Cortez J, Billès G, Navas ML, Roumet C, Debussche M. *et al.* 2004. Plant Functional Markers Capture ecosystem properties during secondary succession. *Ecology* **85**: 2630–2637.

Hasler N, Avissar R. **2007**. What Controls Evapotranspiration in the Amazon Basin? *Journal of Hydrometeorology* **8**: 380–395.

Huffman GJ, Bolvin DT, Nelkin EJ, Wolff DB, Adler RF, Gu G *et al.* 2007. The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. J. *Hydrometeorol.* **8**: 38–55.

Kato S, Loeb NG, Rose FG, Doelling DR, Rutan DA, Caldwell TE *et al.* 2012. Surface Irradiances Consistent with CERES-Derived Top-of-Atmosphere Shortwave and Longwave Irradiances. *J. Clim.* **26**: 2719–2740.

Leuning R, van Gorsel E, Massman WJ, Isaac PR. 2012. Reflections on the surface energy imbalance problem. *Agric. For. Meteorol.* **156**: 65–74.

NASA 2017a. Tropical Rainfall Measuring Mission Project (TRMM), 3B43: Monthly 0.25x0.25 degree merged TRMM and other estimates v7 [WWW Document]. NASA Distrib Act. Arch Cent Goddard Space Flight Cent Earth Sci Greenbelt Md. URL http://mirador.gsfc.nasa.gov/cgi-bin/mirador/

NASA, 2017b. Clouds and the Earth's Radiant Energy System Information and Data (CERES) [WWW Document]. URL http://ceres.larc.nasa.gov/

Oliveira RS, Costa FRC, Baalen E, Jonge A, Bittencourt PR, Almanza Y *et al.* 2019. Embolism resistance drives the distribution of Amazonian rainforest tree species along hydro-topographic gradients. *New Phytologist* **221**: 1457-1465.

Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team. 2014. Nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-118. http://CRAN.R-project.org/package=nlme.

Restrepo-Coupe N, da Rocha HR, Hutyra LR, da Araujo AC, Borma LS, Christoffersen B *et al.* 2013. What drives the seasonality of photosynthesis across the Amazon basin? A cross-site analysis of eddy flux tower measurements from the Brasil flux network. *Agricultural and Forest Meteorology* **182**:128–144.

Restrepo-Coupe N, Levine NM, Christoffersen BO, Albert LP, Wu J, Costa MH, *et al.*

2016. Do dynamic global vegetation models capture the seasonality of carbon fluxes in the Amazon basin? A data-model intercomparison. *Glob. Change Biology* **23**:191-208.

Thomas R, Lello J, Medeiros R, Pollard A, Robinson P, Seward A *et al*. 2017. Data analysis with R Statistical Software: a guidebook for scientists. Eco-Explore, United Kindgom.

Verma SB. 1989. Aerodynamic resistances to transfers of heat, mass and momentum. *Estim. Areal Evapotranspiration* **177**: 13–20.

Wilson K, Goldstein A, Falge E, Aubinet M, Baldocchi D, Berbigier P *et al*. 2002. Energy balance closure at FLUXNET sites. *Agric. For. Meteorol., FLUXNET 2000 Synthesis* **113**: 223–243.

Zuur A, Ieno E, Walker N, Saveliev A, Smith G. 2009. *Mixed Effects Models and Extensions in Ecology with R*. New York, US. Springer Verlar.