Marenco R.A., & Antezana-Vera, S.A. (2021). Principal component regression analysis demonstrates the collinearity-free effect of sub estimated climatic variables on tree growth in the central Amazon. *Revista de Biología Tropical, 69*(2), 482-493. DOI 10.15517/rbt.v69i2.44489

DOI 10.15517/rbt.v69i2.44489

Principal component regression analysis demonstrates the collinearity-free effect of sub estimated climatic variables on tree growth in the central Amazon

Ricardo Antonio Marenco[1]*; Orcid: 0000-0002-9490-2624
Saul Alfredo Antezana-Vera[2]; Orcid: 0000-0001-7949-352X

1. Coordenação de Dinâmica Ambiental, Instituto Nacional de Pesquisas da Amazônia, Manaus, Amazonas, Brazil; rmarenco@inpa.gov.br (*Correspondence).
2. Programa de Pósgraduação em Botânica, Instituto Nacional de Pesquisas da Amazônia, Manaus, Amazonas, Brazil; saulantezana5@gmail.com

Received 07-XI-2020. Corrected 24-II-2021. Accepted 26-II-2021.

ABSTRACT

Introduction: Climatic variables show a seasonal pattern in the central Amazon, but the intra-annual variability effect on tree growth is still unclear. For variables such as relative humidity (RH) and air vapor pressure deficit (VPD), whose individual effects on tree growth can be underestimated, we hypothesize that such influences can be detected by removing the effect of collinearity between regressors. **Objective:** This study aimed to determine the collinearity-free effect of climatic variability on tree growth in the central Amazon. **Methods:** Monthly radial growth was measured in 325 trees from January 2013 to December 2017. Irradiance, air temperature, rainfall, RH, and VPD data were also recorded. Principal Component Regression was used to assess the effect of micrometeorological variability on tree growth over time. For comparison, standard Multiple Linear Regression (MLR) was also used for data analysis. **Results:** Tree growth increased with increasing rainfall and relative humidity, but it decreased with rising maximum VPD, irradiance, and maximum temperature. Therefore, trees grew more slowly during the dry season, when irradiance, temperature and VPD were higher. Micrometeorological variability did not affect tree growth when MLR was applied. These findings indicate that ignoring the correlation between climatic variables can lead to imprecise results. **Conclusions:** A novelty of this study is to demonstrate the orthogonal effect of maximum VPD and minimum relative humidity on tree growth.

Key words: Amazon rainforest; atmospheric dryness; dry season; relative humidity; wet season.

The Amazon rainforest has a significant impact on both water and carbon cycles, due to its enormous extension (~ 5.1×10^6 km^2) and the high amount of carbon stored in its vegetation, about 86 Pg (Saatchi, Houghton, Alvalá, Soares, & Yu, 2007). Tree growth can be defined as the increase of biomass through time and it is often estimated by measuring the increment of stem diameter over time – a proxy of biomass gains at the ecosystem level (Wagner, Rossi, Stahl, Bonal, & Herault, 2012; Wagner et al., 2014; Dias & Marenco, 2016; Antezana-Vera & Marenco, 2020). As tree growth is greatly affected by factors that affect photosynthesis, sun-induced fluorescence – a proxy of ecosystem photosynthesis has been used to estimate the effect of environmental factors on total carbon gain of the ecosystem (Lee et al., 2013; Yang et al., 2018; Green et al., 2020).

Tree growth can be affected by intrinsic factors (e.g., genetic make-up) and environmental factors, such as nutrient availability, irradiance, temperature, rainfall, and soil water content (SWC). The influence of environmental factors on tropical tree growth has been widely studied (Wagner et al., 2012; Mendes, Marenco, & Magalhães, 2013; Wagner et al., 2014; Dias & Marenco, 2016; Méndez, 2018). However, the drivers of tree growth are rather difficult to elucidate because they are often correlated (Bowman, Brienen, Gloor, Phillips, & Prior, 2013; Wagner et al., 2014). Therefore, the effects of climatic parameters, such as temperature, rainfall or SWC on tree growth, are still under investigation in the Amazon (Laurance et al., 2009; Wagner et al., 2014; Dias & Marenco, 2016). Rainfall and SWC seem to be the major factors that affect tree growth in the Amazon region, but there is still under debate whether Amazonian trees grow faster in the wet season than in the dry season. Although in most studies tree growth or ecosystem photosynthesis seems to decrease in the dry season (Méndez, 2018; Wagner et al., 2014; Yang et al., 2018; Antezana-Vera & Marenco, 2020), the opposite effect has also been reported (Laurance et al., 2009; Green, et al., 2020). Whereas Laurance et al. (2009) reported that tree growth was fastest during the dry period and positively correlated with maximum temperatures (T_{max}), Wagner et al. (2014) found no significant effect of T_{max} on tree growth.

Climatic parameters are often correlated, and hence collinearity can lead to imprecise results by increasing the Variance Inflation Factor (VIF), as an increase in VIF can lead to false non-significant effects. Moreover, because of collinearity, the sign of a regression coefficient may change (from positive to negative or *vice versa*, Montgomery, Peck, & Vining, 2012). This is important because the significance and sign of regression coefficients are crucial to understand the effect of climatic variables on tree growth. Principal Component Analysis (PCA) is commonly used to deal with the collinearity problem, whereby a new set of independent variables (orthogonal components) is computed from the original regressors. However, PCA's disadvantage is the lack of a direct association between a response variable and the extracted components. To overcome this difficulty, PCA's orthogonal components can be used to perform Principal Component Regression (PCR). An accurate estimate of the effect of climatic variability on tree growth is essential due to the influence of the Amazon forest on the global carbon balance and regional climate. Thus, this study aimed to determine the collinearity-free effect of climatic variability on tree growth in the central Amazon. We hypothesize that the influence of highly correlated climatic variables such as relative humidity and VPD on tree growth can only be detected after removing the effect of collinearity.

MATERIALS AND METHODS

Study site and plant material: The study was conducted from January 2013 to December 2017 (the experimental period) at the Tropical Forest Experiment Station (ZF2 Reserve), in central Amazonia, located 60 km North of Manaus (02°36'21" S & 60°08'11" W). The area is a *terra-firme* rainforest plateau at about 120 m above sea level. Annual rainfall is 2 420 mm, with a mild dry season, with the driest months from July through September (≤ 100 mm per month. The soil is an Oxisol with low fertility, clay texture and pH of 4.2 to 4.5 In this site, tree density is high, about 600 tree ha^{-1} (> 10 cm diameter at breast height-DBH), canopy height can reach 35-40 m, and most of trees have less than 30 cm in diameter, while leaf area index varies from 4.7 in dry season to 5.0 in the wet season. Mean wood density is about 0.75 g cm^{-3}, and species diversity is high (Dias & Marenco, 2016). At a site 30 km of Manaus, Prance, Rodrigues, and Silva (1976) recorded 179 species of trees in one hectare (≥ 15 cm DBH).

During the experimental period, air temperature (T), photosynthetically active radiation (PAR), RH, and rainfall data were daily recorded above the forest canopy, at the top

of a 40-m-tall observation tower (02°35'20" S & 06°06'55" W). Temperature, RH and PAR data were logged at 15 min (PAR) or 30 min intervals (T and RH) with specific sensors (Humitter 50y, Vaisala, Ov, Finland; LI-190SA, Li-Cor, Lincoln, NE, USA) connected to a data logger (Li-1400, Li-Cor), while daily rainfall data were collected with a tipping bucket gauge (ECR-100, Em5b, Decagon Devices, Pullman, WA, USA). The PAR data were integrated over a 24-h period to obtain a daily value (mol m^{-2} day^{-1}). We also computed VPD and potential evapotranspiration (EVT) and measured soil water content (SWC, %, v/v). VPD was obtained as $VP_{sat} - RH \times VP_{sat}$, where VP_{sat} is the saturation vapor pressure; VP_{sat} (kPa) = $0.61365\exp[17.502T/(240.97 + T)]$, being T (ºC) the air temperature (Buck, 1981). EVT was obtained as: $EVT = 0.0023R_a \times (T_{mean} + 17.8) (T_{max} - T_{min})^{0.5}$, where R_a denotes the extraterrestrial radiation (Hargreaves & Samani, 1985). Undisturbed soil samples were collected at a depth of 100 to 200 mm every two weeks to determine SWC after drying the samples at 105 °C, and then a mean monthly value was obtained.

In this study we collected data from 325 trees from more than 48 species (Digital Appendix), which had a mean diameter at breast height (DBH, diameter at 1.3 m from the ground) of 23.1 ± 11.8 cm. From tree diameter, tree height was estimated to be 22.5 ± 5.2 m (Nogueira, Nelson, Fearnside, França, & Oliveira, 2008). In these trees we measured radial growth at breast height at monthly intervals over 60 months (2013-2017) using stainless steel dendrometer bands, which had been installed three years before the beginning of the study.

Statistical analyses: To assess the effects of the monthly microclimatic variability on tree growth Principal Component Regression (PCR) was used. We used this approach to remove the effect of collinearity among climatic variables. In this analysis, we used detrended tree growth (T_{GC}, hereafter referred to as simply tree growth) instead of undetrended tree growth (T_{GR}, i.e., raw data), because a time-related trend in growth data can affect PCR results (Monserud & Marshall, 2001). This step was accomplished by using a first-order autoregression (Montgomery et al., 2012). Then the tree growth of the whole data set (N = 325) was randomly split into two subsets, one with 75 % of the trees (244 trees) was used to estimate the regression coefficients, and the remaining (25 %, i.e., 81 trees) was used for validation. Prior to PCR analysis the climatic data were standardized. In the PCR model, instead of including all the examined factors, we only used those that combined explained most of the variance (i.e., eigenvalues equal or greater than one). Also, for comparison the significance of the regression coefficients based on standard Multiple Linear Regression (MLR) were also computed. A MLR model can be represented by (Montgomery et al., 2012):

$$y_i = ß_o + ß_1x_1 + ß_2x_2 + \dots ß_kx_k + \epsilon_i$$

(Equation 1)

In Equation 1, y_i denotes the dependent variable, x_i the regressor, $ß_o$ the intercept, $ß_j$ the slope of the regression, and ϵ the error term, being $ß_o$ given by:

$$ß_o = \bar{y} - ß_1\bar{x}_1 - ß_2\bar{x}_2 - \dots ß_k\bar{x}_k.$$

(Equation 2)

For standardized regressors, with mean $\bar{x}_j$ and standard deviation s_j, y_i and $ß_o$ become:

$$y_i = ß_o + ß_1[(x_i - \bar{x}_1)/s_1] + ß_2[(x_i - \bar{x}_2)/s_2] + \dots ß_k[(x_i - \bar{x}_k)/s_k] + \epsilon_i$$

(Equation 3)

$$ß_o = ß^s_o - ß^s_1(\bar{x}_1/s_1) - ß^s_2(\bar{x}_2/s_2) - \dots ß^s_k(\bar{x}_k/s_k)$$

(Equation 4)

The coefficients obtained from standardized regressors (hereafter denoted by the superscript s) can be transformed back to the original regressor units, as follows:

$$b_o = b^s_o - \frac{\bar{x}_j(b^s_j)}{s_j}$$

(Equation 5)

$$b_j = \frac{b_j^s}{s_j}$$

(Equation 6)

Likewise, the variance of b^s_j and standard error (SE) of b^s_j can be transformed back to the original coefficients:

$$\mathrm{var}(b_j) = \mathrm{var}\left[\frac{b^s}{s_j}\right] = \frac{\mathrm{var}(b_j^s)}{s_j^2}$$

(Equation 7)

$$\mathrm{SE}(b_j) = \sqrt{\left[\frac{\mathrm{var}(b_j^s)}{s_j^2}\right]} = \frac{\mathrm{SE}(b_j^s)}{s_j}$$

(Equation 8)

In matrix notation Equation 1 can be represented by:

$$Y = X\beta + \epsilon$$

(Equation 9)

In Equation 9, Y represents the vector of observations (dependent variable); X, the matrix of the corresponding regressors, β the vector of coefficients, and ϵ the vector of random error terms. The normal equations of the linear regression are given in Equation 10, while the estimates of β (often termed $\hat{ß}$, and hereafter denoted by *b*) are given by Equation 11.

$$X'Xb = X'Y$$

(Equation 10)

$$b = (X'X)^{-1}\, X'Y$$

(Equation 11)

The sum of square (SS) of the model (Equation 12), SS of regression (Equation 13), SS of residual (Equation 14), and the covariance-matrix of *b* (Equation 15) are given by:

$$\text{The sum of square (SS) of the model} = b'X'Y$$

(Equation 12)

$$\text{The SS of regression} = b'X'Y - (\sum y^2)/n$$

(Equation 13)

$$\text{The SS of residual} = (Y'Y) - (b'X'Y) - (\sum y^2)/n$$

(Equation 14)

$$\text{The variance-covariance matrix of } b = \mathrm{var}(b) = (X'X)^{-1}\,\sigma^2$$

(Equation 15)

Being the mean square error (MSE = s^2) an estimator of σ^2, and $s = \sqrt{(\mathrm{MSE})}$. When the regressors are highly correlated principal components can be used for transforming those regressors into a new set of uncorrelated variables (orthogonal variables with each other). In terms of standardized regressors, the PCR can be computed, as follows (Montgomery et al., 2012):

$$Y = Z\alpha + \epsilon$$

(Equation 16)

$$Z = XT$$

(Equation 17)

$$\alpha = T'ß$$

(Equation 18)

$$T'X'XT = Z'Z = \Lambda$$

(Equation 19)

$$Z = [z_1, z_2 \ldots z_u]$$

(Equation 20)

In Equation 17, T is a matrix whose columns represent eigenvectors (derived from X data), while in Equation 20, the columns of Z represent a new set of orthogonal scores (i.e., the z-scores), which are termed principal components (Montgomery et al., 2012). The $\hat{\alpha}$ coefficients (Equation 21) and the covariance of $\hat{\alpha}$ (Equation22) are given by:

$$\hat{\alpha} = (Z'Z)^{-1}\, Z'Y = \Lambda^{-1}\, Z'Y$$

(Equation 21)

$$\mathrm{var}(\hat{\alpha}) = \sigma^2(Z'Z)^{-1} = \sigma^2\Lambda^{-1}$$

(Equation 22)

$$b = T\hat{\alpha}$$

(Equation 23)

$$\mathrm{var}(b) = \mathrm{var}(T\hat{\alpha}) = T\Lambda^{-1}T'\sigma^2$$

(Equation 24)

The standardized regressors, b_{pc} are given by Equation 25, while the variance and standard error (SE) of b_{pc} are given by Equation

26 and Equation 27, respectively. In Equation 25, the "pc" indicates that the principal components corresponding to near-zero eigenvalues have been removed from the analysis.

$$b_{\mathrm{pc}} = \mathrm{T}\hat{\alpha}_{\mathrm{pc}} = \sum_{j=1}^{u} \frac{t_{\mathrm{j}}' \mathrm{X}' y t_{\mathrm{j}}}{\lambda_{\mathrm{j}}}$$

(Equation 25)

$$\mathrm{var}(b_{\mathrm{j,pc}}) = \sigma^2 \times \sum_{j=1}^{u} \frac{t_{\mathrm{jm}}^2}{\lambda_{\mathrm{m}}}$$

(Equation 26)

$$\mathrm{SE}(b_{\mathrm{j,pc}}) = \sqrt{\mathrm{MSE}\left(\sum_{m=1}^{u} \frac{t_{\mathrm{jm}}^2}{\lambda_{\mathrm{m}}}\right)}$$

(Equation 27)

The step described in Equation 25 is important to obtain a new set of coefficients, after removing the smallest λ_{i}. This is a crucial step in principal component regression. The second part of Equation 25 shows the computation. Note that Z = XT, then Z´=T´X´, and Λ^{-1} Z´Y = $\hat{\alpha}$ (Equation 21).

In Equation 27, the MSE is obtained as the regression of Y on the *u*- principal components retained in the reduced model, while t_{jm} denotes the j$^{\mathrm{th}}$ element of the eigenvector t_{m} (m = 1 …u). The significance of the principal component estimator (b_{pc}) can be tested on individual coefficients by using the statistic $t_{\mathrm{n\text{-}k\text{-}1}}$, where k represents the number of principal components in the reduced model, as described in Equation 28.

$$t = \frac{b_{\mathrm{j,pc}}}{\mathrm{SE}(b_{\mathrm{j,pc}})}$$

(Equation 28)

Statistical analyses were carried out using Statistica 7.0 (Stat Soft Inc., 2004).

RESULTS

Monthly means of the climatic variables were 26.4 °C (temperature), 78.9 % (RH), and 28.9 mol m^{-2} day^{-1} (PAR). Mean rainfall was 213.7 mm month^{-1}, SWC 44.3 % (v,v), $\mathrm{VPD}_{\mathrm{mean}}$ 7.41 hPa, and EVT 120.7 mm month^{-1}. On the other hand, the mean values between seasons were (dry *vs* wet season): T_{mean} (27.0 *vs* 26.0 °C, P = 0.001), $\mathrm{RH}_{\mathrm{mean}}$ (74.5 *vs* 82.1 %, P < 0.001), $\mathrm{VPD}_{\mathrm{mean}}$ (9.2 *vs* 6.1 hPa, p < 0.001), and evapotranspiration (126.2 *vs* 117.7 mm month^{-1}, P = 0.034). Rainfall, PAR, and the other climatic variables varied within the wet and dry season as described in (Fig. 1).

The undetrended radial tree growth was 0.105 ± 0.11 mm per month (N = 325 trees), with a lower radial increment across the driest season (Fig. 1A). A preliminary analysis showed that by including all the 13 factors in the PCR model (full model), no effect on tree growth was observed, even when the regression explained 23.7 % of the total variance ($F_{(13,46)}$ = 1.10, P = 0.382, R^2= 0.237). In fact, the full PCR model corresponds to the MLR model of tree growth on all the climatic variables (full model MLR). The principal component analysis (PCA) showed that the first four factors out of the 13 factors extracted by PCA (in bold face in Fig. 2) together accounted for 92.9 % of the total variance and had eigenvalues higher than one (λ_{j} = 8.16, 1.86, 1.05, and 1.01). Whereas the values of λ_5 to λ_{13} were lower than 1.0 (Fig. 2). Therefore, we retained the first four factors (Kaiser criterion) and used their corresponding eigenvectors to obtain the z_{j} scores, hereafter referred to as principal components (z_{j}). In comparison with the full PCR model, the significance of the four-principal component model was improved ($F_{(4,55)}$ = 2.45, P = 0.056, R^2 = 0.151). By reducing the complexity of the model, the amount of variance on tree growth explained by the regression model was also reduced (23.7 against 15.1 %). Moreover, the four-factor model showed that only the principal components z_1 and z_3 had a significant effect on tree growth, whereas z_2 and z_4 did not (i.e., z_1: P = 0.03, z_2: P = 0.46, z_3: P = 0.04, and z_4: P = 0.86). Therefore, only z_1 and z_3 were retained for further analysis and hereafter referred to as the reduced model. As expected, significance of the reduced model was further improved by retaining only the two significant components, i.e., z_1 and z_3 ($F_{(2,57)}$ = 4.73, P = 0.012, R^2 = 0.142), with both regression

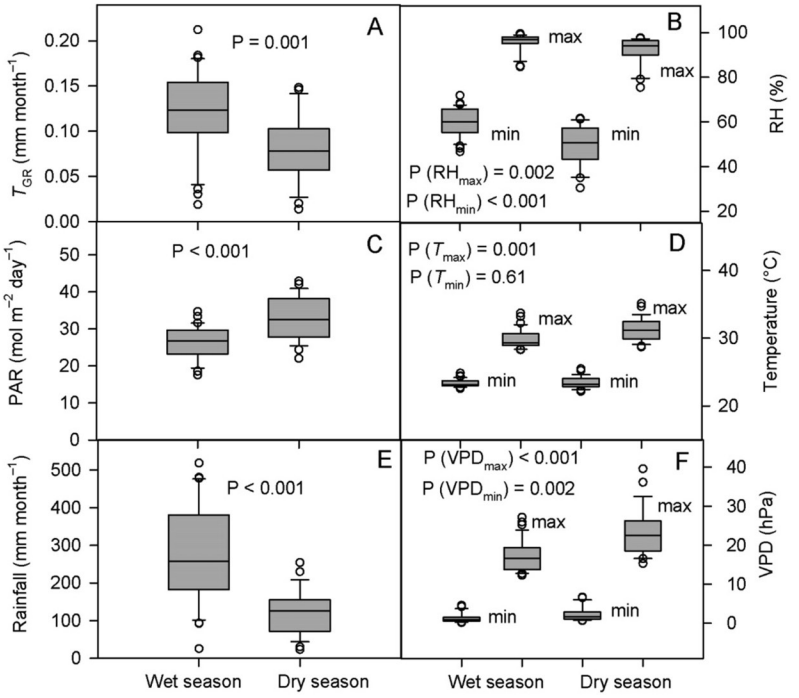

Fig. 1. Undetrended tree growth (T_{GR}) and climatic variables recorded in the wet season (November to May) and dry season (June to October) during the years of 2013 to 2017. **A.** Undetrended tree growth, **B.** RH_{min} and RH_{max}, **C.** PAR, **D.** T_{min} and T_{max}, **E.** Rainfall, and **F.** VPD_{min} and VPD_{max}. Abbreviations are shown in Table 1.

coefficients being significant (α_1 = -0.003927, P = 0.03) and α_3 = -0.010267, P = 0.04). As z_1 and z_3 were retained in the reduced PCR model, only Factor 1 and Factor 3 are shown (Fig. 2), and for further information tree growth (T_{GC}) is included in Fig. 2 as a supplementary variable. In Fig. 2, by taking T_{GC} as a reference point, climatic variables were separated into three groups. The first group comprised RH_{min}, RH_{mean}, rainfall, and SWC; which together with T_{GC} are in quadrant (III) on the factor plane. These results suggest a positive correlation between each of them and T_{GC}. The second group (PAR, EVT, T_{max} and VPD_{max}) is located in quadrant I (i.e., diagonally opposite to T_{GC}), and thereby indicating a negative correlation between the variables of this group and T_{GC}. The third group included variables located in adjacent quadrants, i.e., quadrant II (RH_{max}) and quadrant IV (T_{min}, T_{mean}, VPD_{min} and VPD_{mean}), and then indicating a low correlation between each of these variables and T_{GC}. We further investigated the relationship between the climatic variables and T_{GC} after removing the effect of collinearity (i.e., by PCR).

The PCR regression coefficients (α_1 and α_3) were used to compute the beta coefficients, $\boldsymbol{b}_{pc}$ (in $\boldsymbol{b}_{pc}$, the subscript *pc* stands for principal components) and their SE (Equation 25 and Equation 27). The coefficients $\boldsymbol{b}_{pc}$ based on standardized regressors are shown in Table 1, while those coefficients ($\boldsymbol{b}_j$) obtained by MLR are shown and Table 2.

After removing the effect of collinearity, we found that tree growth was significantly responsive to variation in PAR (x_1), rainfall (x_2), T_{max} (x_3), RH_{mean} (x_4), RH_{min} (x_5), VPD_{max} (x_6), SWC(x_7), and EVT (x_8). Whereas T_{min}, T_{mean}, RH_{max}, VPD_{min}, and VPD_{mean} had no significant effect on tree growth (Table 1). Tree growth increased with a rise in rainfall, SWC, RH_{mean}, and RH_{min}, whereas it decreased with

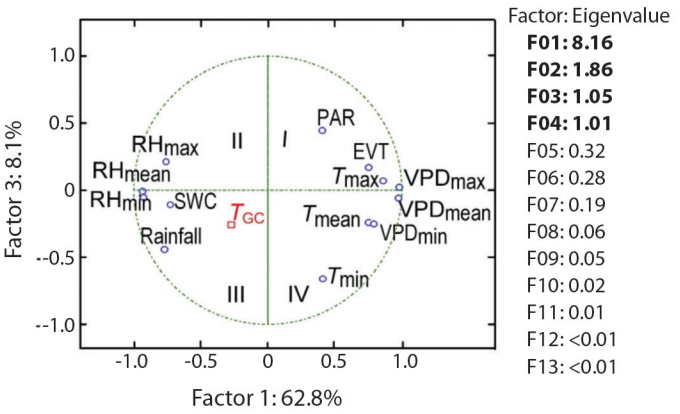

Fig. 2. Principal Component Analysis of climatic variables. The eigenvalues of the first four factors account for 92.9 % of total variance. In the factor plane, tree growth (T_{GC}) was included as a supplementary variable. Abbreviations are shown in Table 1.

TABLE 1
Regression coefficients of standardized climatic variables (b^S_{pc}), standard error (SE) of b^S_{pc}, variance inflation factor (VIF), and $t_{(57df)}$ and P values obtained by principal component regression (PCR) of tree growth (T_{GC}) on principal component z_1 and z_3 ($F_{(2.57)} = 4.73$, MSE = 0.001473, P = 0.0125, $R^2 = 0.142$)

Variable	Beta (b^s_{pc})	SE (b^s_{pc})	VIF	$t_{(57)}$	P value
PAR	-0.005019	0.002164	1.0	-2.348939	**0.022**
Rainfall	0.005498	0.002188	1.0	2.545052	**0.014**
T_{mean}	0.001380	0.001250	1.0	1.118301	0.268
T_{min}	0.006063	0.003198	1.0	1.919855	0.060
T_{max}	-0.001872	0.000628	1.0	-3.019680	**0.004**
RH_{mean}	0.001374	0.000583	1.0	2.387525	**0.020**
RH_{min}	0.001812	0.000632	1.0	2.905751	**0.005**
RH_{max}	-0.001084	0.001128	1.0	-0.972707	0.335
VPD_{mean}	-0.000745	0.000667	1.0	-1.131649	0.262
VPD_{min}	0.001446	0.001312	1.0	1.116269	0.269
VPD_{max}	-0.001571	0.000616	1.0	-2.582338	**0.012**
SWC	0.002088	0.000691	1.0	3.059548	**0.003**
EVT	-0.002726	0.000939	1.0	-2.938550	**0.005**

EVT: potential evapotranspiration, PAR: photosynthetically active radiation, T: temperature, T_{max}: mean maximum T, T_{mean}: mean T, T_{min}: mean minimum T, RH: relative humidity, RH_{max}: mean maximum RH, RH_{mean}: mean RH, RH_{min}: mean minimum RH, VPD: air vapor pressure deficit, VPD_{max}: mean maximum VPD, VPD_{min}: mean minimum VPD, VPD_{mean}: VPD mean, and MSE: mean square error. Climatic data were standardized prior to statistical analysis. For T_{GC}, N = 244. Significant P values are in bold face.

increasing PAR, T_{max}, EVT, and VPD_{max}, as shown in Equation 29 (based on standardized regressors, Table 1) and Equation 30, for regressors in the original scale. To validate the PCR model (Equation 30) we used growth data of 81 trees, which showed that the R^2 derived from the validation data set was even slightly higher than the R^2 of data used to build the model ($R^2 = 0.12$ vs. 0.17, Fig. 3).

$$T_{GC} = 0.0000 - 0.005019x_1 + 0.005498x_2 - 0.001872x_3 + 0.001374x_4 + 0.001812x_5 - 0.001571x_6 + 0.002088x_7 - 0.002726x_8$$

(Equation 29)

TABLE 2

Regression coefficients (b_j), standard error (SE), VIF, and $t_{(46\ df)}$ and P values obtained by standard multiple linear regression (MLR) of tree growth (T_{GC}) on climatic variables ($F_{(13,46)} = 1.10$, MSE = 0.001624, P = 0.382, $R^2 = 0.237$)

Variable	b_j	SE(b_j)	VIF	t value	P value
PAR	-0.000863	0.008981	2.9	-0.09609	0.923870
Rainfall	0.005264	0.011935	5.2	0.44107	0.661226
T_{mean}	-0.043624	0.038490	53.8	-1.13341	0.262915
T_{min}	0.020023	0.016432	9.8	1.21854	0.229230
T_{max}	0.060980	0.058528	124.5	1.04189	0.302907
RH_{mean}	0.229039	0.200442	1459.8	1.14267	0.259088
RH_{min}	-0.093012	0.117944	505.4	-0.78861	0.434383
RH_{max}	-0.295352	0.225920	1854.5	-1.30733	0.197599
VPD_{mean}	0.309690	0.237523	2049.8	1.30383	0.198778
VPD_{min}	-0.292876	0.228333	1894.3	-1.28267	0.206034
VPD_{max}	-0.201591	0.175032	1113.1	-1.15174	0.255380
SWC	0.007049	0.009000	2.9	0.78321	0.437519
EVT	0.001124	0.019793	14.2	0.05679	0.954959

Climatic data were standardized prior to statistical analysis. Note that in comparison with PCR coefficients (Table 1) T_{max}, RH_{min} and EVT had opposite sign. Abbreviations as described in Table 1.

$$T_{GC} = 0.027074 - 0.000819x_1 + 0.000043x_2 - 0.001135x_3 + 0.000178x_4 + 0.000202x_5 - 0.000275x_6 + 0.000474x_7 - 0.000158x_8$$

(Equation 30)

By applying the standard MLR approach, we found that none of the climatic parameters modified tree growth ($F_{(13,46)} = 1.10$, P = 0.382, R^2= 0.237, Table 2). Furthermore, in comparison with the results obtained with the PCR reduced model (only factor 1 and factor 3), the MLR coefficients (b_j) had much larger SE (Table 2). For instance, the SE of RH_{mean} and VPD_{mean} were more than two orders of magnitude higher than those obtained by PCR, due to the effect of collinearity. Besides having larger SE, some coefficients (T_{max}, RH_{min} and EVT) had opposite sign. In retrospect, using the result from PCR and hence discarding from the MLR model those climatic variables with no significant effect on tree growth (Table 1) did not yield any significant regression coefficient ($F_{(8,51)} = 1.55$, P = 0.16, $R^2 = 0.19$).

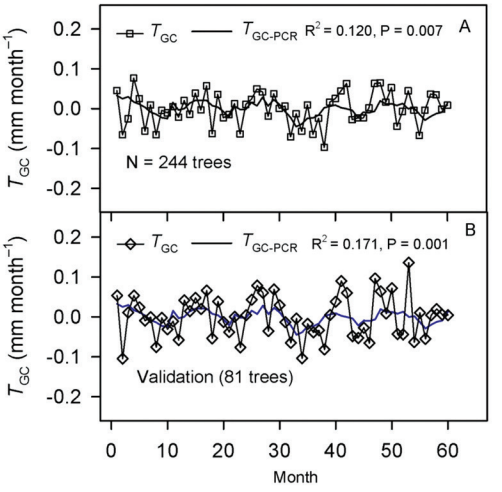

Fig. 3. A. Tree growth (T_{GC}, square, N = 244) and the regression line (solid line, T_{GC-PCR}) as a function time. **B.** The validation of the PCR model on growth data (N = 81 trees) is also shown. The solid blue line corresponds to the regression line, while the diamond represents the T_{GC} of validation trees. PCR: principal component regression, T_{GC-PCR}: PCR line fitted to data.

DISCUSSION

Most of the climatic variables assessed had a significant effect on tree growth. The exceptions were T_{min}, T_{mean}, RH_{max}, VPD_{min} and VPD_{mean} which did not modify tree growth. Thus, these results partially support our hypothesis, as highly correlated variables

such RH_{max}, VPD_{min} and VPD_{mean} influenced tree growth. We found that PCR explained 12 % of the total variance ($R^2 = 0.12$), which is not unexpected, as many factors can affect tree growth (Bowman et al., 2013). For instance, Wagner et al. (2012) found that only about 9 % of the variation in tree growth can be attributed to seasonal climate variability, which is slightly lower than the proportion of total variance explained by climatic variability in our study.

The standard MLR explained 23.7 % of total variance in tree growth. Nevertheless, due to the large standard error (SE) associated with each coefficient (Table 2), none of the regression coefficients significantly affected tree growth. On the other hand, even small-magnitude coefficients obtained by PCR, such as RH_{mean} and RH_{min}, showed a significant effect on tree growth. This finding supports our hypothesis, as the detrimental effect of collinearity becomes evident when the correlation between climatic variables was disregarded and the data subjected to MLR. The large SE of MLR coefficients undermined the predictive power of the MLR model, and therefore, the influence of the climatic variables on tree growth was underestimated. For instance, the VIF of RH_{mean}, RH_{max}, and VPD computed by MLR (Table 2) were over three orders of magnitude greater than that of genuinely independent orthogonal regressors (Table 1), which magnified the SE up to 200-300 times (e.g., RH_{max}, VPD_{mean} and VPD_{max}) as compared with the SE obtained by PCR. We found that some MLR coefficients (e.g., T_{max} and RH_{min}) had opposite sign.

The misleading effect of collinearity can occur because the variance of a regression coefficient (say b_1) is inversely proportional to the amplitude of the regressor [i.e., $\text{var}(b_1) = \sigma^2/\sum(x_i - \bar{x})^2$]. Hence, when the variance is so large, and the actual value of a coefficient is close to zero, a regression coefficient with opposite sign can result (Montgomery et al., 2012). This is remarkable because it can be concluded that a variable x_j has a positive (or negative) effect on Y, when in fact the opposite is true. Tree growth increased with rising mean and minimum RH, whereas it decreased with increasing VPD_{max} and EVT. Thus, by observing the VIF factor presented in Table 2, it is tempting to discard from the regression model not only RH but also VPD. Firstly, because a VIF value above 10 is an indicative of strong collinearity among regressors (Montgomery et al., 2012). Secondly, because it may be expected that the effects of these variables are already included within the effect of temperature. However, discarding these variables from the model may weaken its predictive strength as EVT, RH_{mean}, RH_{min} and VPD_{max} had a truly independent effect on tree growth. Collinearity dramatically increases the VIF, making it difficult to quantify the individual contribution of a regressor with little but real independent effect on a dependent variable, such as tree growth (Montgomery et al., 2012; Bowman et al., 2013).

Tree growth was positively responsive to an increase in rainfall intensity, whereas T_{max} and PAR had a negative effect on growth rates. The effect of T_{max} on tree growth found in this study agrees with the results of Way and Oren (2010), who reported that tree growth of tropical species can be negatively affected by warming. On the other hand, our results disagree with those of Wagner et al. (2014) and Laurance et al. (2009). Wagner et al. (2014) reported no effect of T_{max}, whereas Laurance et al. (2009) found a positive effect of maximum temperature on tree growth. This discrepancy can be ascribed to difference in environmental conditions during data collection. For instance, Green et al. (2020) reported that ecosystem photosynthesis increases in the central Amazon when VPD increased from 0.1 to 10 hPa. In tropical rainforests, the optimum temperature for photosynthesis is about 29 °C (Liu, 2020), with decreasing photosynthetic rates at higher temperatures. This can help explain the decline in tree growth with rising T_{max}. Beside the effect of temperature on photosynthesis, a raise in temperature has also an effect on transpiration via the effect of temperature on water viscosity (Darcy´s Law). In fact, in this experimental site, EVT can increase in the dry season

when temperatures are higher (Antezana-Vera & Marenco, 2020).

There are reports associating tree growth or ecosystem photosynthesis to variations in temporal rainfall variability in the Amazon region (Lee et al., 2013; Méndez, 2018; Yang et al., 2018) or VPD (Lee et al. 2013; Green et al., 2020). Some studies that aim to assess the effect of rainfall seasonality on tree growth in the Amazon have led to different results. Dias and Marenco (2016) and Silva et al., (2003) found no increase in tree growth during the wet season, whereas Wagner et al. (2014), Méndez (2018), and Antezana-Vera and Marenco (2020) reported that tree growth increased with an increase in rainfall intensity. Likewise, Lee et al. (2013) and Yang et al. (2018) reported a decline in photosynthesis-related activity during the dry season. Altogether these results indicate that the magnitude of the effect of drought on tree growth is related to the length of the dry season. In this study, we demonstrated that PCR could be a handy tool. We provide evidence that an increase in VPD_{max} (from 17 hPa – wet season to 23 hPa in the dry season) leads to a reduction in tree growth. Interestingly, such effect was only observed after removing the effect of collinearity. Marenco et al. (2014) showed that photosynthesis of canopy leaves (22-27 m tall trees) is closely related to stomatal conductance (g_s). They reported that g_s increased and reached its maximum value at a VPD of 16 hPa, and then it declined and became almost null at a VPD of 28 hPa. Likewise, Mendes and Marenco (2017) observed that g_s increased with increasing VPD in the range of 5 to 10 hPa. These results show that the effect of VPD on photosynthesis depended on the level of atmospheric moisture. Green et al. (2020) reported that ecosystem photosynthesis can increase at VPD values lower than 10 hPa, Lee et al. (2013), on the other hand, estimated that ecosystem photosynthesis declined as VPD progressively increased from 3.5 hPa (wet season) to 32 hPa in the dry season, which is in agreement with the results found in our study.

Solar radiation is intrinsically associated with tree growth via its effect on photosynthesis, and it has been reported that in tropical rainforests an increase in solar radiation can lead to an increase in tree growth (Wagner et al., 2014). On the contrary, we found that an increase in PAR leads to a decline in tree growth, which agrees with the results of Yang et al. (2018) and Méndez (2018). Yang et al. (2018) observed a decrease in solar-induced fluorescence during the drought of 2015-2016 in the Amazon region. Likewise, in a study carried out at the same experimental site, Antezana-Vera and Marenco (2020) found that transpiration significantly increased with an increase in PAR and VPD. An increase in transpiration rates does not mean an increase in g_s and photosynthesis. In fact, most of the time, g_s decreases as transpiration increases in response to an increase in VPD (Dai, Edwards, & Ku, 1992), which can explain the negative effect of PAR and VPD on tree growth reported in this study.

Our results are relevant because of the global importance of the Amazon forest and because of the effects of the ongoing climate changes, which have increased the temperature (about 0.16 °C per decade) and altered rainfall distribution, ranging from lower rainfall intensity (longer dry seasons) in Eastern and Southern Amazonia to higher rainfall intensity in the Northern Amazon (Marengo et al., 2018). The dry season is associated with a rise in solar radiation, temperature, and VPD (Lee et al., 2013; Green et al., 2020), which ultimately can lead to a decline in photosynthesis (Lee et al., 2013; Marenco et al., 2014; Yang et al., 2018). Because most of the climatic variables are correlated, assessing the collinearity-free effect is important to accurately quantify the climatic drivers of tree growth. Increased dry season length has been forecasted for some parts of the Amazon (Marengo et al., 2018), which ultimately may reduce tree growth, not only reducing soil water availability, but also by increasing VPD and reducing RH. Our results demonstrate that trees of the central Amazon grow more slowly during the dry season, not

only due the effect of a drop in rainfall intensity, but also in response to the effect of an increase in maximum temperature, evapotranspiration, and maximum vapor pressure deficit, and a decline in mean and minimum relative humidity. To our knowledge this is the first time the collinearity-free effect of RH_{min}, RH_{mean}, EVT and VPD_{max} on tree growth in the Amazon region has been evaluated. The novelty of this study is to demonstrate the orthogonal effect of VPD_{max} and RH_{min} on tree growth in the central Amazon, which contributes to enhance the current knowledge of the ecophysiology of Amazonian trees.

Ethical statement: authors declare that they all agree with this publication and made significant contributions; that there is no conflict of interest of any kind; and that we followed all pertinent ethical and legal procedures and requirements. All financial sources are fully and clearly stated in the acknowledgements section. A signed document has been filed in the journal archives.

ACKNOWLEDGMENTS

To Ministry of Science, Technology and Innovations (MCTI-INPA), the Foundation for Research Support of the State of Amazonas (FAPEAM), Coordination for the Improvement of Higher Education Personnel (CAPES code 0001) and the *National Council for Scientific and Technological Development* (CNPq, 303907/2018-5). We thank the Editors and anonymous reviewers for their valuable comments and suggestions.

RESUMEN

El análisis de regresión por componentes principales muestra el efecto libre de colinealidad de variables climáticas subestimadas sobre el crecimiento de los árboles en la Amazonía central

Introducción: Las variables climáticas muestran un patrón estacional en la Amazonía central, pero el efecto de la variabilidad intra-anual en el crecimiento de los árboles aún no está claro. Para variables como la humedad relativa (HR) y el déficit de presión de vapor (VPD), cuyo efecto individual en el crecimiento de los árboles puede ser subestimada, planteamos la hipótesis de que tales influencias pueden detectarse eliminando el efecto de colinealidad entre regresores. **Objetivo:** Este estudio tuvo como objetivo determinar el efecto libre de colinealidad de la variabilidad climática sobre el crecimiento de los árboles en la Amazonía central. **Métodos:** Se midió el crecimiento radial mensual en 325 árboles desde enero 2013 hasta diciembre 2017. También se registraron datos de irradiancia (PAR), temperatura del aire, lluvia, humedad relativa (RH) y déficit de presión de vapor de aire (VPD). Se utilizó la regresión de componentes principales para evaluar el efecto de la variabilidad micrometeorológica a lo largo del tiempo sobre el crecimiento de los árboles. Para comparación, también se utilizó la regresión lineal múltiple (MLR) estándar para el análisis de datos. **Resultados:** El crecimiento de los árboles incrementó con el aumento de las precipitaciones y la humedad relativa, y disminuyó con el aumento de la VPD máxima, la irradiancia y la temperatura máxima. Por lo tanto, los árboles crecieron más lentamente durante la estación seca, cuando la irradiancia, la temperatura y la VPD eran más altas. La variabilidad micrometeorológica no afectó el crecimiento de los árboles cuando se aplicó MLR. Estos hallazgos indican que ignorar la correlación entre las variables climáticas puede conducir a resultados imprecisos. **Conclusiones:** Una novedad de este estudio es demostrar el efecto ortogonal del VPD máximo y la humedad relativa mínima sobre el crecimiento de los árboles.

Palabras clave: floresta húmeda amazónica; sequedad atmosférica; estación seca; humedad relativa; temporada húmeda.

REFERENCES

Antezana-Vera, S.A., & Marenco, R.A. (2020). Sap flow rates of *Minquartia guianensis* in central Amazonia during the prolonged dry season of 2015-2016. *Journal Forestry Research*. DOI: 10.1007/s11676-020-01193-9

Bowman, D.M., Brienen, R.J., Gloor, E., Phillips, O.L., & Prior, L.D. (2013). Detecting trends in tree growth: not so simple. *Trends in Plant Science*, *18*, 11-17.

Buck, A.L. (1981). New equations for computing vapor-pressure and enhancement factor. *Journal of Applied Meteorology*, *20*, 1527-1532.

Dai, Z., Edwards, G.E., & Ku, M.S. (1992). Control of photosynthesis and stomatal conductance in *Ricinus communis* L. (castor bean) by leaf to air vapor pressure deficit. *Plant Physiology*, *99*, 1426-1434.

Dias, D.P., & Marenco, R.A. (2016). Tree growth, wood and bark water content of 28 Amazonian tree species in response to variations in rainfall and wood density. *iForest - Biogeosciences and Forestry*, *9*(3), 1-7.

Green, J.K., Berry, J., Ciais, P., Zhang, Y., & Gentine, P. (2020). Amazon rainforest photosynthesis increases in response to atmospheric dryness. *Science Advances*, *6*, eabb7232.

Hargreaves, G.H., & Samani, Z.A. (1985). Reference crop evapotranspiration from temperature. *Applied Engineering in Agriculture, 1*, 96-99.

Laurance, S.G., Laurance, W.F., Nascimento, H.E., Andrade, A., Fearnside, P.M., Rebello, E.R., & Condit, R. (2009). Long-term variation in Amazon forest dynamics. *Journal of Vegetation Science, 20*, 323-333.

Lee, J.E., Frankenberg, C., van der Tol, C., Berry, J.A., Guanter, L., Boyce, C.K., Fisher, J.B., Morrow, E., Worden, J.R., Asefi, S., Badgley, G., & Saatchi, S. (2013). Forest productivity and water stress in Amazonia: observations from GOSAT chlorophyll fluorescence. *Proceedings of the Royal Society Biological Sciences, 280*, 20130171.

Liu, Y. (2020). Optimum temperature for photosynthesis: from leaf-to ecosystem-scale. *Science Bulletin*, *65*, 601-604.

Marenco, R.A., Antezana-Vera, S.A., Santos, P.R.G., Camargo, M.A.B., Oliveira, M.F., & Santos, J.K.S. (2014). Fisiologia de espécies florestais da Amazônia: fotossíntese, respiração e relações hídricas. *Revista Ceres*, *61*, 786-799.

Marengo, J.A., Souza, C.M., Thonicke, K., Burton, C., Halladay, K., Betts, R.A., Alves, L.M., & Soares, W.R. (2018). Changes in climate and land use over the Amazon region: current and future variability and trends. *Frontiers in Earth Science, 6*, 228. DOI: 10.3389/feart.2018.00228

Mendes, K.R., & Marenco, R.A. (2017). Stomatal opening in response to the simultaneous increase in vapor pressure deficit and temperature over a 24-h period under constant light in a tropical rainforest of the central Amazon. *Theoretical and Experimental Plant Physiology*, *29*, 187-194.

Mendes, K.R., Marenco, R.A., & Magalhães, N.D.S. (2013). Growth and photosynthetic use efficiency of nitrogen and phosphorus in saplings of Amazonian tree species. *Revista Arvore*, *37*, 707-716.

Méndez, C.R. (2018). *Influência do El Niño 2015-2016 no incremento diamétrico das árvores da Amazônia Central* (Master's thesis). Instituto Nacional de Pesquisas da Amazônia, Manaus, Brazil.

Monserud, R.A., & Marshall, J.D. (2001). Time-series analysis of δ13C from tree rings. I. Time trends and autocorrelation. *Tree Physiology*, *21*, 1087-102.

Montgomery, D.C., Peck, E.A., & Vining, G.G. (2012). *Introduction to linear regression analysis*. Hoboken, USA: John Wiley & Sons.

Nogueira, E.M., Nelson, B.W., Fearnside, P.M., França, M.B., & Oliveira, A.C. (2008). Tree height in Brazil's 'arc of deforestation': shorter trees in south and southwest Amazonia imply lower biomass. *Forest Ecology and Management*, *255*, 2963-2972.

Prance, G.T., Rodrigues, W.A., & Silva, M.F. (1976). Inventario florestal de um hectare de mata de terra firme km 30 de Estrada Manaus-Itacoatiara. *Acta Amazonica*, *6*, 9-35.

Saatchi, S.S., Houghton, R.A., Alvalá, R.C.S., Soares, J.V., & Yu, Y. (2007). Distribution of aboveground live biomass in the Amazon basin. *Global Change Biology*, *13*, 816-837.

Silva, R.P., Nakamura, S., Azevedo, C.P., Chambers, J., Rocha, R.D., Pinto, A.C., Santos, J., & Higuchi, N. (2003). Use of metallic dendrometers for individual diameter growth patterns of trees at Cuieiras river basin. *Acta Amazonica, 33*, 67-84.

Stat Soft Inc. (2004). *STATISTICA* (Version 7, software). Tulsa, OK, USA.

Wagner, F., Rossi, V., Aubry-Kientz, M., Bonal, D., Dalitz, H., Gliniars, R., Stahl, C., Trabucco, A., & Herault, B. (2014). Pan-tropical analysis of climate effects on seasonal tree growth. *PLoS ONE*, *9*, e92337.

Wagner, F., Rossi, V., Stahl, C., Bonal, D., & Herault, B. (2012). Water availability is the main climate driver of neotropical tree growth. *PLoS ONE*, *7*, e34074.

Way, D.A., & Oren, R. (2010). Differential responses to changes in growth temperature between trees from different functional groups and biomes: a review and synthesis of data. *Tree Physiology*, *30*, 669-88.

Yang, J., Tian, H., Pan, S., Chen, G., Zhang, B., & Dangal, S. (2018). Amazon drought and forest response: Largely reduced forest photosynthesis but slightly increased canopy greenness during the extreme drought of 2015/2016. *Global Change Biology*, *24*, 1919-1934.