Fire frequency alone does not explain forest - savanna transition: the role of dry season precipitation variability in northern South America

Santiago Valencia Cárdenas

Tesis de maestría presentada para optar al título de Magíster en Ingeniería Ambiental

Director
Juan Fernando Salazar, Doctor (PhD) Aprovechamiento de Recursos hidráulicos
Codirector
Juan Camilo Villegas, PostDoctor (PostDoc) en Natural resources emphasis Watershed management and ecohydrology

Universidad de Antioquia
Facultad de Ingeniería
Maestría en Ingeniería Ambiental
Medellín, Antioquia, Colombia
2022

Cita	(Valencia et al., 2022)
Referencia **Estilo APA 7 (2020)**	Valencia, S., Salazar, JF. & Villegas, JC. (2022). *Fire frequency alone does not explain forest - savanna transition: the role of dry season precipitation variability in northern South America* [Tesis de maestría]. Universidad de Antioquia, Medellín, Colombia.

Maestría en Ingeniería Ambiental, Cohorte Seleccione cohorte posgrado.

Grupo de Investigación Ingeniería y Gestión Ambiental (GIGA).

Seleccione centro de investigación UdeA (A-Z).

Agradecimientos

En primer lugar, agradezco a mi familia, mis abuelos, mi mamá, mi papá y mi hermana, quienes han sido mi apoyo y motivación permanente durante este proceso.

A Juliana Mejía por su infinito cariño y apoyo incondicional.

Agradezco a mis asesores, los profesores Juan Fernando Salazar y Juan Camilo Villegas por su permanente acompañamiento y motivación para el desarrollo de este trabajo. Siempre han sido un gran ejemplo a nivel personal y profesional para mí. De igual manera, extiendo mis agradecimientos a la profesora Natalia Hoyos por darme la oportunidad de participar en el proyecto que dio origen a la pregunta de investigación de este trabajo. Además, quiero agradecerles a los profesores Esnedy Hernández, Néstor Aguirre, Alejandro Martínez, Juan Pablo Serna y Lina Berrouet, quienes han sido muy importantes en mi formación profesional.

A mis amigos Derly Gómez, Maikol Córdoba, José Morales, Sergio Herazo, Mateo Parra, Karen Palacio y Benjamín Atehortúa, por su apoyo. Así como a todas las personas que de alguna manera estuvieron conmigo durante mi maestría. En especial, quiero darle las gracias a mi amigo Diver Marín por su permanente apoyo y compartir conmigo el mismo entusiasmo y curiosidad por la investigación.

A la Universidad de Antioquia por la financiación de mí maestría y formación a nivel profesional y personal que me ha brindado en los últimos 7 años.

Termino dedicando este trabajo a Álvaro Antonio Chica (q.e.p.d), quién fue mi primer profesor. Ojalá pudiera ver hasta donde me han permitido llegar sus enseñanzas.

ABSTRACT

The forest - savanna transition is the most widespread ecotone in the tropical regions and with important ecological, climatic, and biogeochemical implications at local to global scales. However, the processes and mechanisms that control this transition vary among regions and remain not fully understood in all of them. In general, this transition is influenced by interactions between vegetation and environmental factors such as climate, soil properties, fire, and herbivory. However, the importance of these effects can vary substantially across continents, which can result in different responses to environmental change. For this reason, more regional studies are needed to describe and understand the factors and interactions that control forest - savanna transition in different regions. Using remotely-sensed data, we examined the relationship between the tropical forest-savanna transition and several environmental factors in northern South America, in the Llanos ecoregion. We used several vegetation structure metrics, as well as multiple precipitation statistics, soil properties, and a fire regime descriptor. In addition, we developed a statistical analysis on the interactive effects of soil silt content, fire frequency as well as three dry season precipitation variability components (season length, wet day frequency, and precipitation intensity) on the forest - savanna transition, using tree cover, canopy cover and $PAVD_{max}$ as indicator variables that differentiate forest from savanna. Our results show that savannas in the Llanos region occur in mean annual precipitation (MAP) levels in which forest would be predicted based on previously proposed thresholds to other savanna regions. Our results also highlight that the MAP range in which both forest and savanna can occur in our study area correspond, almost exclusively, to forest in other South American regions and globally. Although both forest and savanna can also occur in a large interval of intermediate values of dry season precipitation variability (PV) components, forest dominates in areas with higher precipitation frequency and intensity than savanna. Savanna tends to occur in pixels where fires are present, while fires are absent in forest. However, a large proportion of pixels classified as savanna pixels have no fires in the analysis period, even those that occur in the same climatic or edaphic space of forest. Finally, our analysis shows that fire frequency and dry season precipitation are the most important variables to predict the forest-savanna transition. This highlights the role of fire regime and water availability in determining the limits between forest and the second largest area of savanna in South America. Further, our results support the importance of refining our understanding of the factors, relationships, and mechanisms that control forest-savanna transition at regional scales, as a requirement to assess the effects of environmental change on tropical forest and savanna distribution.

keywords: forest-savanna transition, precipitation variability, fire frequency, northern South America

1. INTRODUCTION

Tropical forests and savannas account for more than 60% of the terrestrial productivity (Oliveras & Malhi, 2016; Erb et al., 2018; Lipsett-Moore et al., 2018). Both kinds of tropical ecosystems are globally strategic, and their presence and dynamics have important ecological, climatic and biogeochemical implications, even at the global scale (Oliveras & Malhi, 2016; Xu et al., 2018; Armenteras et al., 2021a). Forest-savanna transition is the most widespread and, perhaps, the most dynamic ecotone in the tropics. However, the processes and mechanisms that control this transition vary among regions and remain not fully understood (Murphy et al., 2012; Lehmann et al., 2014; Archibald et al., 2019; Bernardino et al., 2021). More specifically, this transition is influenced by multiple interactions between vegetation and environmental factors such as climate, soil properties, fire, and herbivory, which operate at different spatio-temporal scales (Lehmann et al., 2011; Veenendaal et al., 2015; Staver et al., 2017; Newberry et al., 2020).

Tropical forests are characterized by a dense and structurally-complex canopy, typically with a high diversity of trees, lianas, and epiphytes, and competition for light primarily drives their dynamics and structural complexity (Murphy et al., 2012). Tropical savannas are characterized by a dominant C_4 grass layer, with discontinuous patches of woody vegetation, with lower structural complexity than forest canopies (Lehmann et al., 2014; Archibald et al., 2019; Stark et al., 2020). However, the transition between savannas and forests exhibits various degrees of woody-grass combinations worldwide, potentially similar to a grassland-forest continuum (Breshears, 2006). The extent of this transition varies regionally, depending on geographic location and regional environmental factors (Ratnam et al., 2011, Oliveras & Malhi, 2016).

In tropical regions, water availability --which results from vegetation, climate, and soil properties-- has been identified as one of the major determinants of ecosystem structure and distribution (Oliveras & Malhi, 2016; Ehbrecht et al., 2021). For instance, both in-situ and remotely-sensed observations show that tropical forests occur more often in regions with high mean annual precipitation (MAP) and short dry seasons (Lehmann et al., 2011, 2014; Staver et al., 2011; Viglizzo et al., 2015). However, there is not a simple precipitation threshold that defines the distribution or transition of forests and savannas (Murphy et al., 2012; Archibald et al., 2019; Ciemer et al., 2019; Staal et al., 2021). For example, both savanna and forests occur in regions with intermediate MAP values (between 1000 and 2500 mm globally). In those regions, other factors such as the interactions between climate, vegetation, soil, and disturbance regimes such as fire and herbivory appear to better explain the transition between ecosystems (Hirota et al., 2011; Staver et al., 2011). Several studies highlight how, in addition to MAP, precipitation variability (PV - indicated by intra-seasonal precipitation metrics at multiple temporal scales) can improve the prediction of tropical forest and savanna distribution (Good & Caylor, 2011; Zeng et al., 2014; Case & Staver, 2018; Xu et al., 2018; Hoyos et al., 2021) via its effects on water availability. This highlights the importance of understanding the effect of PV on ecosystem distribution, particularly when climate models predict changes in precipitation properties (such as precipitation frequency and intensity) in many areas of the world, even where MAP will not change (IPCC, 2013; 2021).

The effect of water availability on vegetation depends not only on the amount and seasonality of precipitation but also on soil properties (Rodriguez-Iturbe & Porporato, 2007). Sandier soils allow deeper infiltration, promoting deeper root distributions, which may be associated with higher tree cover (Case & Staver, 2018). Soil fertility has also been recognized as an important determinant of forest and savanna distribution at different spatial scales, especially in regions with similar precipitation regimes (Lloyd et al., 2015; Pellegrini, 2016). Savannas are often associated with lower soil fertility than forests (e.g., low cation exchange capacity, organic matter as well as macro- and micro-nutrients; Lloyd et al., 2013, 2015; Veenendaal et al., 2015; February et al., 2019). However, it is still unclear whether differences in soil fertility are a cause or a consequence of forest-savanna distribution (Pellegrini, 2016; Archibald et al., 2019) and their effects vary depending on climate (Lehmann et al., 2011).

Forest-savanna transition is not only determined by vegetation-climate-soil interactions but also by feedbacks with other factors such as fire, mainly in more mesic climates (Hoffman et al., 2012; Bernardino et al., 2021). Specifically, fire-vegetation feedbacks in savannas allow frequent burning that maintains savanna (open-canopy) where both climate and soil properties would otherwise lead to predicting the occurrence of a forest (closed-canopy; Bond, 2008; Newberry et al., 2020). Fire decreases tree cover, which subsequently favors the light-dependent grassy layer, promoting fuel for fire spread and maintaining an open-canopy state (Bernardino et al., 2021). However, the role of fire (either natural or anthropogenic) as a determinant of forest-savanna transition is not fully understood (Staal & Flores, 2015; Veenendaal et al., 2018), particularly in mesic climates (Archibald et al., 2019).

In South America, water availability proxies such as MAP, precipitation seasonality, mean precipitation in the dry season or dry season soil saturation index, have a lower explanation potential of tropical forest and savanna distribution than in Africa or Australia (Hirota et al., 2011; Murphy et al., 2012; Lehmann et al., 2011; 2014; Staver et al., 2011, Zheng et al., 2014; Xu et al., 2018). Indeed, there are extensive areas in South America with MAP levels that would predict a forest if they were located in Africa or Australia (Lehman et al., 2011, 2014). However, these regions in South America are savannas, which suggests a weaker climatic control (based on commonly used climate descriptors) on forest and savanna distribution in this continent (Murphy et al., 2012; Lehmann et al., 2014). In addition, South American savannas differ from African or Australian savannas in other fundamental factors such as soil properties, fire regime (frequency, season, extension, and intensity), and their feedbacks with vegetation (Staver et al., 2011; Zheng et al., 2014; Moncrieff et al. 2016; February et al., 2019), highlighting biogeographic differences among continents (Murphy et al., 2012; Lehmann et al., 2014). Notably, these biogeographic differences can also occur within continents. Such is the case of the two main savanna-dominated regions in South America: the Cerrado (South of the Amazon region in Brazil) and the Llanos (North of the Amazon region between Colombia and Venezuela) (Borguetti et al., 2019). Although both regions have a forest-savanna transition, they exhibit noticeable ecological differences, including vegetation structure and composition, climatic, and edaphic properties (Borguetti et al., 2019), leading to different relationships

between vegetation and environmental factors that describe the transition between forest and savanna. For example, when applying the climatic space thresholds proposed for the forest-savanna transition in the Cerrado region by Malhi et al. (2009; MAP < 1500 mm and the maximum climatological water deficit, MCWD < -300 mm) to the Llanos, most savannas in the region would be predicted to be forests (Section 3.1). This indicates the importance of refining our understanding of the factors, relationships, and mechanisms that control forest-savanna transition at the regional scale, to improve projections of global change effects on ecosystem distributions (Oliveras & Malhi, 2016; Archibald et al., 2019).

In this study, we examine the relationship between forest-savanna transition and several environmental factors in northern South America using remotely-sensed data. We use several vegetation structure metrics as indicators of forest-savanna occurrence and transition. Based on global and regional studies, we use multiple PV components at different time scales, soil characteristics (including texture and fertility), and fire regime as explanatory variables for the transition. We further discuss our results indicating that fire frequency and PV components -- particularly frequency of wet days during the dry season-- are more significantly associated with the forest-savanna transition than the other explanatory variables in this region.

2. MATERIAL AND METHODS

2.1. Study area

The study area is located in northern South America (62-72°W, 1°S-10°N), which corresponds to the Llanos ecoregion between Colombia (Llanos Orientales) and Venezuela (Orinoco Llanos), and the northwestern portion of the Amazon basin shared by Colombia, Venezuela, and Brazil (Figure 1). MAP ranges from 1000-1500 mm in the northern region near the Venezuelan border to 2500-3500 mm in the forest-savanna transition and forest areas in the south and southwest (Behling & Hooghiemstra, 2000; Borguetti et al., 2019). The climate in the region is seasonal, with a dry season that extends from 4 to 7 months (between November and April-May; Huber et al., 2006). This precipitation regime suggests that savannas in this region occur in MAP levels where forests would be expected to occur based on biogeographical models for southern South America (i.e., the Cerrado; Malhi et al., 2009; Ciemer et al., 2010), Africa, or Australia (Lehman et al., 2011). Further, even when tropical savannas are typically water-limited, large areas of savanna in the Llanos ecoregion do not show annual moisture deficit under average climatic conditions (AI=MAP/PET > 1; Figure 1).

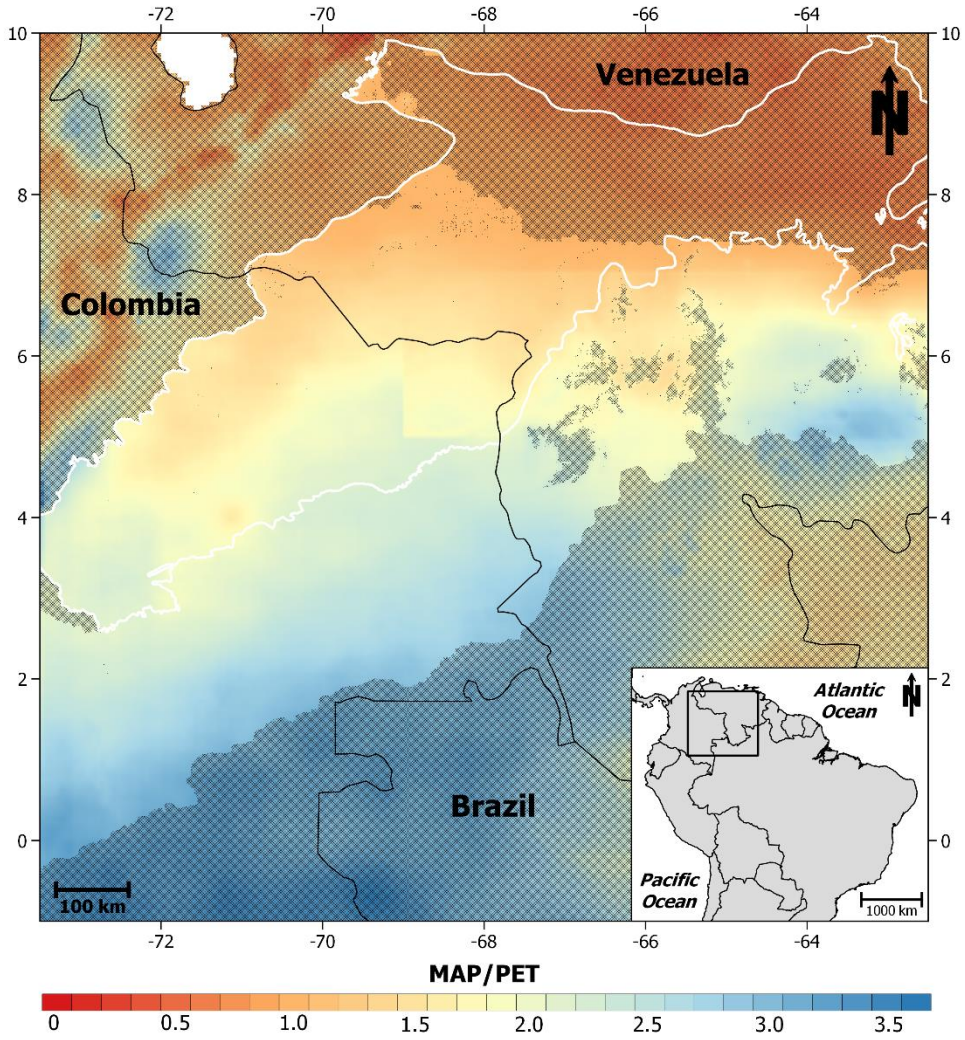

Figure 1. Study area in northern South America, including Aridity index (AI) values. The AI is defined as the ratio of MAP to mean annual potential evapotranspiration (PET; Zomer et al., 2008). The white line indicates the limits of the Llanos ecoregion. The hatching represents pixels excluded from the analysis, following the criteria defined in Section 2.3. MAP and PET data were obtained from CHIRPS and TerraClimate datasets for 1981-2010, respectively.

Major soil groups are Oxisols in the forest, and Oxisols, Ultisols, and Inceptisols in the savanna (Romero-Ruiz et al., 2010). Oxisols are typically highly weathered and have a high content of Fe and Al, low cation exchange capacity (CEC), and small amounts of exchangeable bases (Buol & Eswaran, 1999; Fageria & Nascente, 2014). Ultisols are less weathered but more acidic than oxisols, with moderate to high levels of clay but low organic matter and low base saturation (West et al., 1997; Huber et al., 2006). Inceptisols are poorly drained soils, which can be moisture saturated in the wet season. The occurrence of fires is high and relatively frequent (0.5 to 2.0 years; Borguetti et al., 2019) in the savannas, mainly during the dry season (Armenteras et al., 2005). However, although vegetation and climate largely explain fire occurrence (Barreto & Armenteras, 2020), a large portion of fires in the savanna is related to traditional management practices and cattle grazing (Armenteras et al., 2005; Romero-Ruiz et al., 2010). The landscape of the Llanos includes several types of savanna formations (e.g.,

permanently and seasonally flooded savannas and high plain savannas), riparian or gallery forests, palm-dominated forests, and wetland vegetation (Armenteras et al., 2021b). In these areas, the predominant vegetation is dominated by C_4 grasses, especially from the Poaceae and Cyperaceae families, in association with dispersed woody plants (Huber et al., 2006, February et al., 2019). Forests in the region correspond mainly to the Negro-Branco, Caquetá, Japorá-Solimões-Negro, and a part of the Guiana piedmont moisture forest ecoregions (Dinerstein et al., 2017).

2.2. Data sources

We collected information on vegetation, climate, soil properties, fire, land cover, and topography from multiple remotely-sensed data sources (Table 1). We obtained vegetation structure data from NASA's Global Ecosystem Dynamics Investigation (GEDI; Dubayah et al., 2020) and the Moderate Resolution Imaging Spectroradiometer (MODIS; DiMiceli et al., 2015). Unlike MODIS, GEDI provides not only vegetation cover data but also data on the vertical structure of the vegetation, which is key to characterize vegetation structural complexity and related ecosystem processes (e.g., energy balance; Stark et al., 2020). From the GEDI dataset (collected between April 19, 2019, and September 02, 2020, for the study area), we extracted canopy cover (for comparison with MODIS) and maximum Plant Area Volume Density ($PAVD_{max}$, a metric of vertical vegetation complexity; Decuyper et al., 2018; Meeussen et al., 2020). Canopy cover and $PAVD_{max}$ were gridded to 0.01° (~1.1 km) pixels, as described in the Supporting Information S1 and Figure S1a. We obtained percent tree cover for 2019 from the MODIS Collection 6 Vegetation Continuous Fields product at a 250 m resolution.

For soil texture and fertility, we extracted sand, clay, and silt content as well as cation exchange capacity (CEC) and soil organic carbon (SOC) from SoilGrid, available at a 250 m resolution (Hengl et al., 2017). Following Case & Staver (2018), we used a weighted average spanning three depth horizons (0-5 cm, 5-15 cm, and 15-30 cm) to obtain values representing soil properties in the top 30 cm, where most ecological processes occur. For fire information, we extracted burned areas from FIRECC51, a global monthly burned area product with 250 m resolution spanning 2001-2019 (Chuvieco et al., 2018). We calculated the number of times each pixel burned across the available period as an estimate of typical fire frequency (following Staver et al., 2011; Lehmann et al., 2014; Case & Staver, 2018), excluding pixels with more than 50% invalid observations (i.e., with a 70% or lower confidence level). Although herbivory has been identified as an important driver of savanna dynamics and transition to forest (Hempson et al., 2015), herbivore population data is not available for the scale and spatial extent of our analysis and, therefore, was not included in the analysis.

Table 1. Description of data and sources used in this study.

Category	Variables	Dataset	Spatial resolution	Temporal resolution	Temporal coverage	Reference
Vegetation	Tree cover (%)[1]	MODIS[3]	0.25 km	Annual	2000 - present	DiMiceli et al., 2015
	Canopy Cover (%)[2]	GEDI[4]	0.025 km	Depends on the ISS* trajectories	2019 - present	Dubayah et al., 2020
	Maximum Plant Area Vegetation Density ($PAVD_{max}$, m^2/m^3)					
Soil properties	Sand, clay, and silt content (%)	SoilGrids v2.0[5]	0.25 km	-	-	Hengl et al., 2017; de Sousa et al., 2020
	Cation exchange capacity (CEC, mmol(c)/kg)					
	Soil organic carbon (SOC, dg/kg)					
Fire	Burned area	FIRECCI51[6]	0.25 km	Monthly	2001 - 2019	Chuvieco et al., 2018
Climate	Mean annual precipitation (MAP, mm)	CHIRPSv2.0[7]	0.05° (~ 5.5 km)	Daily	1981 - present	Funk et al., 2015
	Mean total precipitation in dry (d) and wet (w) season ($MAP_{d/w}$, mm)					
	Mean length of dry (d) and wet (w) season ($T_{d/w}$, days)					
	Mean daily precipitation intensity in dry (d) and wet (w) season ($\alpha_{d/w}$, mm/day)					
	Mean daily precipitation frequency in dry (d) and wet (w) season ($\lambda_{d/w}$)					
	Mean length of dry spells during the dry season (days)					
	Mean frequency of wet days with daily precipitation < 10 and ≥10 mm/day during the dry season					
	Monthly minimum temperature (°C)	WorldClim v2.1	~ 1.0 km	Monthly	1970 - 2000	Fick & Hijmans, 2017
	Potential evapotranspiration (PET, mm)	TerraClimate	~ 4.0 km	Monthly	1981 - 2010	Abatzoglou et al., 2018
Land cover	Global land cover	ESA[8]	0.30 km	Annual	1992 - 2019	ESA, 2017
Topography	Elevation	SRTM[9]	0.25 km	-	-	Jarvis et al., 2008

[1]The percent tree cover refers to the amount of skylight obstructed by tree canopies equal to or greater than 5 m in height (DiMiceli et al., 2015); [2]The GEDI-derived canopy cover is the percent of the ground covered by the vertical projection of canopy material (i.e., leaves, branches and stems only; Tang et al., 2019); [3]MODIS: Moderate Resolution Imaging Spectroradiometer; [4]GEDI: Global Ecosystem Dynamics Investigation; [5]SoilGrids: Global Gridded Soil Information; [6]FIRECCI51: the ESA FireCCI project; [7]CHIRPS: Climate Hazards Group InfraRed Precipitation with Station data; [8]ESA: European Space Agency; [9]SRTM: Shuttle Radar Topographic Mission; *International Space Station.

We used daily precipitation data from the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS; Funk et al., 2015) available at 0.05° (~ 5.5 km) to calculate multiple precipitation statistics using the available period between 1981 and 2019. For each pixel in the CHIRPS raster layer, we calculated MAP (in mm), mean total dry (d) and wet (w) season precipitation ($MAP_{d/w}$, in mm), mean daily dry and wet season precipitation intensity ($\alpha_{d/w}$, in mm/day), mean frequency of wet days (daily precipitation > 0) in the dry and wet season ($\lambda_{d/w}$), and mean length of the dry and wet season ($T_{d/w}$, in days). To define the dry season, instead of setting a precipitation threshold, e.g., months with monthly precipitation less than 100 mm (e.g., Marengo et al., 2011; Anderson et al., 2021) or months with mean monthly potential evapotranspiration greater than precipitation (e.g., Yang et al., 2016), we used the global gridded dataset (Rainy and Dry Seasons (RADS) dataset) developed by Bombardi et al., (2019), previously used by Uribe & Dukes, (2021) and Rodrigues et al., (2021). We also calculated the mean length of dry spells in the dry season (following Hoyos et al., 2021). Further, we estimated the mean frequency of wet days classified into two intensity categories (< 10 and ≥10 mm/day) to characterize the intensity and frequency of precipitation in the dry season (i.e., precipitation pulses). Finally, we calculated the maximum climatological water deficit (MCWD; Aragão et al., 2007). Following Malhi et al. (2009) and for comparison with this study, we obtained MCWD from the mean annual cycle of precipitation (1981-2019) with fixed monthly evapotranspiration of 100 mm.

Although regional climatic patterns are reasonably well represented in a resolution of 0.05°, vegetation structure in the forest-savanna transition may vary considerably over a distance of 5.5 km (0.05°). Since aggregation or resampling techniques could cause loss of vegetation structure variability (Hirota et al., 2011; Supporting Information Figure S2), all data layers, except CHIRPS, were resampled to match the resolution of the GEDI data (0.01° x 0.01°) using bilinear interpolation in the projectRaster function of the raster R package (Hijmans, 2020). We rescaled CHIRPS data to 0.01° by dividing each pixel at 0.05° into 25 pixels without varying the original values. Before further statistical analysis, all the datasets were converted into mean temporal values because our focus is on the average transition state. Therefore, the differences in temporal durations among datasets should have little impact on our analysis, as Xu et al. (2018) suggested. Maps of all variables are shown in the Supporting Information Figure S3.

2.3. Study area delimitation and forest-savanna discrimination

To avoid the potential inclusion of Andean ecosystems, we limited our study area to pixels with an elevation lower than 800 m (to avoid high-altitude habitats) and a minimum monthly temperature higher than 15°C using the Shuttle Radar Topography Mission (SRTM; Jarvis et al., 2008) and WorldClim (Fick & Hijmans, 2017) datasets, respectively. We also used the 300-m ESA global land cover map (ESA CCI-LC, v1.6.1) for 2019 to mask out grid cells with more than 30% of the area covered by croplands, urban areas, water bodies (codes ≤40, 190, 210), or permanent snow or ice (code 220). Although riparian forests have contrasting soil properties

and water table depth compared to the grass layer in Llanos, we did not exclude these areas because riparian forests represent less than 5% (N= 4740) of the total 0.01° pixels into savanna ecoregion. Savannas can be seen as having a mesic boundary where they transition into the forest and an arid boundary transitioning into arid vegetation (Archibald et al., 2019). Hence, the datasets were split into arid and mesic transitions using the AI, which captures the interactive effects that climate has on water availability for plants (Pellegrini et al., 2016). Given our interest in the mesic transition, pixels with AI < 1 were masked to exclude savanna regions with an annual moisture deficit in average climatic conditions (Figure 1). Then, we selected an equivalent area within the forest ecoregions to obtain a similar number of pixels in both ecosystems. In addition, to focus the analysis on forest and savanna vegetation, we defined pixels as forest or savanna based on canopy cover (savanna < 40% and forest ≥ 40%), tree cover (savanna < 60% and forest ≥ 60%), and $PAVD_{max}$ (savanna < 0.10 m^2/m^3 and forest ≥ 0.10 m^2/m^3; details in Supplementary Information S1). Finally, all datasets were extracted at 0.01° resolution for GEDI available pixels (N=201474; Supporting Information Figure S3).

2.4. Spatial analysis of the forest-savanna transition

We estimated the importance of each predictor variable described in the previous sections to explain response variables based on the Pearson correlation coefficient (r) and the Spearman rank correlation analysis (r_s), indicating whether relations are linear or nonlinear, respectively. Only those variables with correlations higher than 0.2 or lower than -0.2 and p-value < 0.01 were considered for further analysis. Potential collinearity between predictor variables was also assessed with those metrics (r or $r_s \geq |0.70|$; Dormann et al., 2013). Then, we looked for spatial patterns in response and predictor variables across forest-savanna transects. We set up 1835 transects across forest-savanna transitions in the study area. Transects started in a defined forest region, crossed the transition, and ended in the savanna region. We sampled transects in adjacent 0.01° available pixels (details in Supplementary Information S2 and Supporting Information Figure S4a). Finally, to identify whether vegetation structure varied significantly over the forest-savanna transition, we conducted a changepoint analysis for median values of canopy cover, tree cover, and $PAVD_{max}$ transects. Changepoint identifies the point (i.e., the distance along the transect) at which statistical properties (e.g., mean, variance, or both) of a sequence of observations change (Killick & Eckley, 2014).

2.5. Statistical analyses

We analyzed the relationship between vegetation descriptors (i.e., canopy cover, tree cover, and $PAVD_{max}$) and the selected predictor variables using generalized linear models (GLMs; McCullagh & Nelder, 1989) with the stats package in R (R Development Core Team, 2021). We standardized the predictor variables by subtracting the mean of each variable and then dividing by their standard deviation, such that their coefficient magnitude was a measure of their importance in the model. The goodness-of-fit was evaluated as the fraction of deviance explained (pseudo-R^2, R^2 henceforth for brevity), equivalent to the explained variance in a linear least-squares regression model. It was computed as $R^2 = 1 - D_m/D_0$, where D_m is the

residual deviance, i.e., the deviance that remains unexplained by the fit, and D_0 is the deviance of the intercept-only model (D'onofrio et al., 2019). To take into account the potential effect of differences in spatial variability between PV components (0.05°) and other predictor and response variables (0.01°, Section 2.2), we fitted 1000 GLMs using a random sample of 5% (N=10074) of the available 0.01° pixels to determine the direction, strength, and significance of predictor variables on tree cover, canopy cover, and $PAVD_{max}$. See Supporting Information S4 for further details.

3. RESULTS

3.1. Forest-savanna transition climatic space in the Llanos versus the Cerrado regions

The climatic space --a set of climate-related variables ranges-- in the Llanos region does not correspond with the predicted climate space for savannas in the Cerrado region (yellow portion of Figure 2). A comparison between forest-savanna transition in Llanos and Cerrado regions show that the savanna (here defined as pixels with canopy cover < 40%) in Llanos occurs in regions with MAP and MCWD regimes that would be associated with forest, according to thresholds suggested by Malhi et al. (2009; MAP > 1500 mm and MCWD > -300 mm) for the Cerrado region (Figure 2b). In the Llanos region, forest (canopy cover ≥ 40%) dominates in pixels with MAP > 2360 mm and MCWD > -230 mm (ranges delineated by solid black lines). When MAP varies between 1665 mm and 3070 mm, savannas dominate if MCWD < -33 mm (ranges delineated by solid red lines), evidencing the high MAP and low dry season severity in the savannas of the Llanos. Interestingly, there is a broad climatic space in 2360 mm < MAP < 3070 mm and -230 mm > MCWD > -33 mm, where both forest and savanna can occur, highlighting the importance of including other variables in assessing savanna/forest occurrence in the region, as discussed below. It is important to note that these ranges could change with the canopy cover threshold for forest and savanna. However, these ranges do not greatly vary when using $PAVD_{max}$ or tree cover instead of canopy cover to define them (dashed and dotted lines in Figure 2a, respectively).

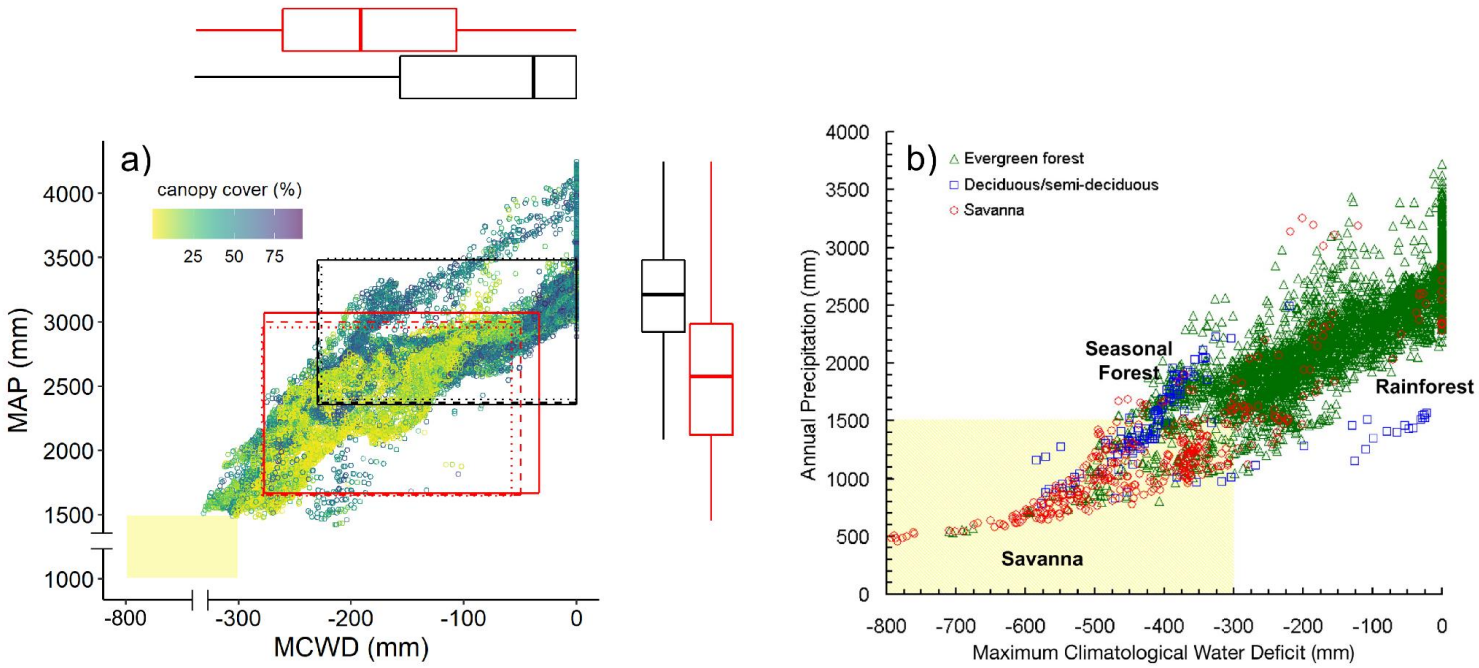

Figure 2. Relationship between a) canopy cover and climatic space (MAP, MCWD) for forest-savanna transition in the Llanos region. b) Relationship between vegetation type and precipitation regime for the Cerrado region (0°-20°S and 45°W-70°W) modified from Malhi et al. (2009). Ranges delineated by different lines indicate each variable's 10th and 90th percentiles for savanna (red) and forest (black) defined by canopy cover (savanna < 40% and forest ≥ 40%; solid line), $PAVD_{max}$ (savanna < 0.10 m^2/m^3 and forest ≥ 0.10 m^2/m^3; dashed line), and tree cover (savanna < 60% and forest ≥ 60%; dotted line). In a), the boxplots show the MAP and MCWD values for all savanna (red) and forest (black) pixels defined by canopy cover threshold (savanna < 40% and forest ≥ 40%).

3.2. Explanatory variables selection

As expected, canopy cover, tree cover, and $PAVD_{max}$ are highly correlated with each other (r_s and $r \geq 0.82$), particularly canopy cover and $PAVD_{max}$ that come from the same dataset. However, we observe differences in the representation of vegetation structure between these variables (Supporting Information Figure S5). For example, although canopy cover and $PAVD_{max}$ show almost a perfect fit for the linear regression (the regression $R^2 = 0.97$), the dispersion of points was lower for low canopy cover (< 25%) and $PAVD_{max}$ (< 0.1 m^2/m^3) than for high values of this vegetation descriptors (canopy cover > 50%; $PAVD_{max}$ > 0.2 m^2/m^3). This contrast evidenced the differences between horizontal (i.e., canopy cover) and vertical (i.e., $PAVD_{max}$) vegetation structure as shown in Supporting Information Figure S6a. Additionally, as reported by DiMiceli et al. (2021), GEDI's values are, on average, higher and lower than tree cover from MODIS at the higher and lower end of the range, respectively (Supporting Information Figure S5b), which is related to differences in these products' spatial resolution and methods. Therefore, our further analysis combines canopy cover, tree cover, and $PAVD_{max}$ and considers that the most robust patterns are those consistent across the three vegetation descriptors.

Our results suggest that precipitation during the dry season is a better predictor of forest-savanna transition than precipitation during the wet season. MAP_d (r_s and $r \geq 0.64$) and λ_d (r_s

and r ≥ 0.54), as well as other PV components during the dry season (r_s and r ≥ 0.26; Tables S1 and S2), exhibit a high positive correlation with vegetation descriptors with both the Spearman (r_s) and Pearson (r) correlation metrics. Additionally, length of dry spells (Ds_d) and frequency of wet days with precipitation < 10 mm/day within this season (λ_{10d}) show a high negative correlation with vegetation descriptors (r_s and r ≤ -0.58). Although $\alpha_w$is also highly correlated with vegetation descriptors (r_s and r ≥ 0.51), the other PV components for the wet season (MAP_w, λ_w, and T_w) show lower correlations (r_s and r ≤ -0.26) than those for the dry season.

Fire frequency has a high negative correlation with vegetation descriptors, with the non-linear correlation (r_s ≤ -0.64) being higher than the linear correlation (r ≤-0.51), which suggests that this relationship is not constant across different ranges of fire frequency. Overall, soil texture and fertility both have low correlations with vegetation descriptors (r_s and r < -0.22), except for silt soil content (r_s and r < -0.28). As a result, PV components during the dry season (MAP_d, λ_d, α_d, T_d, Ds_d, λ_{10d}), fire frequency, and silt soil content were selected for further analysis.

3.3. Forest - savanna transition

In our spatial analysis that considers the variation of several descriptors along 1853 parallel transects along the forest-savanna transition (analysis described in section 2.4 and Supporting information S2), we used changepoint analysis with canopy cover to define schematic limits or boundaries between forest (F) and savanna (S) with a transition (T) zone in between (dashed lines in Figure 3). In this analysis, we found that, as expected, vegetation descriptors exhibit a noticeable transition from the forest to the savanna regions (Figures 3a-c). However, the differentiation between regions is not consistent among vegetation descriptors. For instance, transition and savanna regions differed more clearly from each other in canopy cover and $PAVD_{max}$ than in tree cover (see inset boxplots in Figures 3a-c), which may be associated with the relatively large variation in the 10^{th}-90^{th} and 25^{th}-75^{th} percentile ranges, particularly in the forest and transition regions. In this case, the transition zone (almost 100 km wide) represents the distance to the forest edge where, in general, the mean canopy cover of all transects shows a significant change in both mean and variance canopy cover. For this reason, this transition zone shows a less sharp change in canopy cover values than individual transects (Supporting information Figure S2). For this reason, although this schematic transition zone allows a comparison of spatial patterns across the transition between vegetation descriptors (Figure 3a-c) and explicative variables (Figure 3d-i), it does not refer to the shape of the forest-savanna transition, which is quite sharp.

Our results show that although MAP_d and λ_d are highly correlated with vegetation descriptors (r_s and r ≥ 0.54; Tables S1 and S2), both variables exhibit a gradual (Figures 3d-e), less sharp transition than canopy cover, tree cover, or $PAVD_{max}$ transects (Figures 3a-c). Overall, MAP_d and λ_d are higher in the forest than in the transition and savanna regions, with median values of MAP_d=953 mm, λ_d=0.38, MAP_d=514 mm, λ_d=0.27, and MAP_d=360 mm, and λ_d=0.19, respectively. However, the forest region also exhibits larger spatial variability (i.e., 10^{th}-90^{th}

and 25th-75th percentile ranges) than the other regions, particularly when compared to the savanna region (Figure 3d and 3e; Supporting Information Figure S3d and S3e). Although α_d and T_d also decrease from the forest region to the savanna region (Figure 3f and 3g), differences in median or percentile ranges between forest (α_d=13.05 mm/day; T_d=199 days), transition (α_d=11.38 mm/day; T_d=172 days), and savanna (α_d=10.35 mm/day; T_d=178 days) regions were not evident. This indicates that precipitation frequency (instead of dry season length or precipitation intensity) explains the difference in MAP_d among regions (Figure 3d).

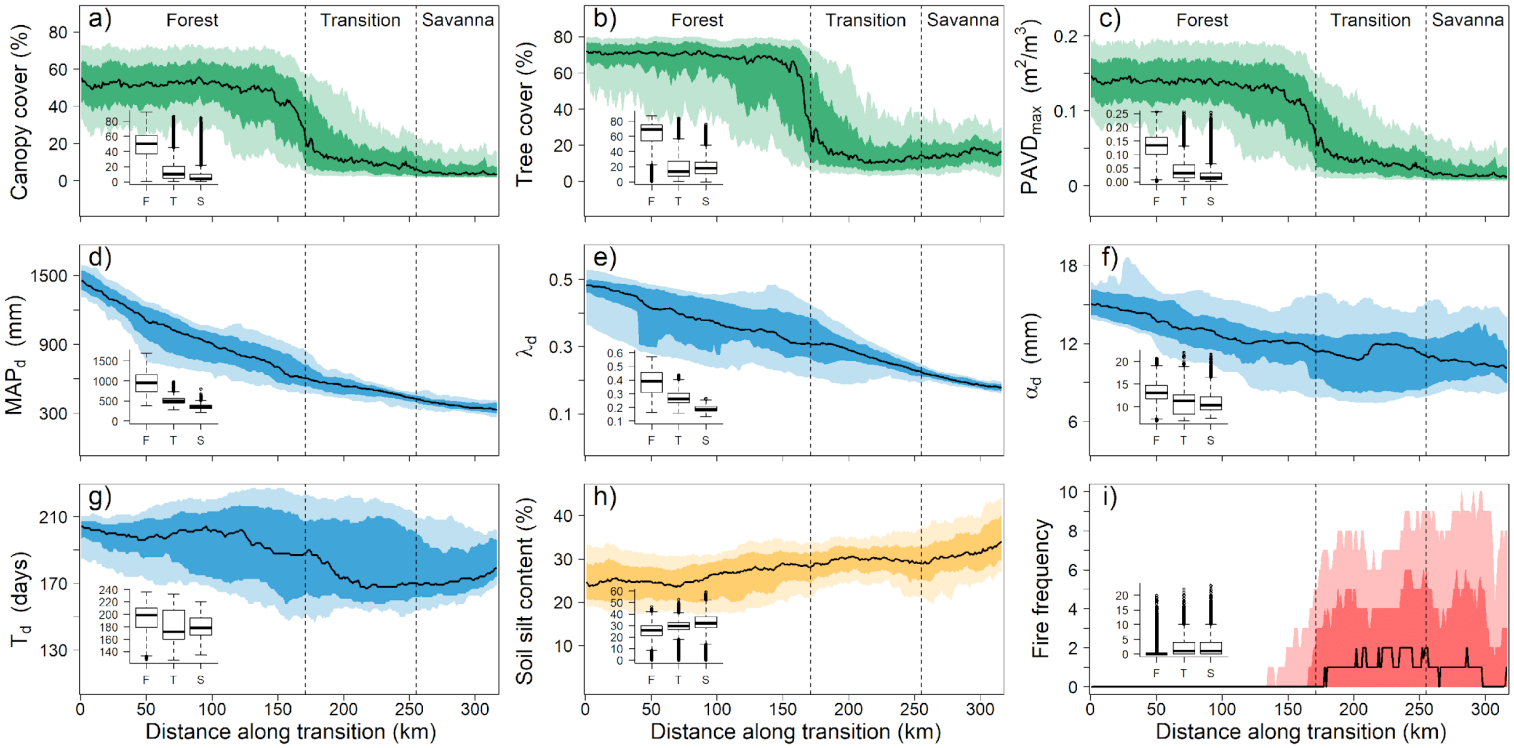

Figure 3. Transects across forest-savanna transition. a) canopy cover, b) tree cover, c) maximum plant area volume density ($PAVD_{max}$), d) mean total dry-season precipitation (MAP_d), e) frequency of wet days (precipitation > 0) within the dry season ($\boldsymbol{\lambda_d}$), f) intensity of wet days within the dry season ($\boldsymbol{\alpha_d}$), g) length of the dry season (T_d), h) soil silt content, and i) fire frequency. In all panels, the lightest shade is 10th - 90th percentile, the darker shape is 25th - 75th percentile, and the black line represents the median (50th percentile) for all transects analyzed (n=1835). Vertical dashed lines show the schematic limits of forest (F), transition (T), and savanna (S) regions quantified using a changepoint analysis for canopy cover as a function of distance along the transition. The inset boxplots show the values of each variable in the F, T, and S regions. The Supporting Information Figure S7 shows results for predictor variables that were not considered for further analysis.

Similar to α_d and T_d, soil silt content (0-30 cm) also exhibited low changes in median or percentile range along forest-savanna transition (Figure 3h), consistent with low correlation values (Supporting Information Tables S1 and S2). Fire frequency is the only predictor variable that shows a sharp transition between regions, similar to vegetation descriptor variables (Figure 3i). The transition and savanna regions have the highest fire frequency, which is close to zero in the forest region. However, there is a large spatial variability (i.e., 10th-90th and 25th-75th percentile ranges) in fire frequency in the transition and savanna regions (Supporting Information Figure S3i). These results coincide with several studies that reported that savanna

occurs more commonly where fires are present (e.g., Lehmann et al., 2011; Staver et al., 2011; Bernardino et al., 2021). However, the high spatial variability in fire frequency highlights the complex processes and feedbacks that determine fire regime, including human activities in the study region (Armenteras et al., 2005; Romero-Ruiz et al., 2010; Barreto & Armenteras, 2020).

3.4. Forest-Savanna transition climatic space

Following the approach of Figure 2, we plotted MAP vs. the selected predictor variables to analyze the ranges in which forest and savanna can occur in the Llanos region (Figure 4). We used tree cover (dotted line), $PAVD_{max}$ (dashed line), and canopy cover (solid line) to define forest and savanna climatic domains. We selected MAP as the independent variable because it has been widely used to analyze the forest and savanna distribution (e.g., Hirota et al., 2011; Staver et al., 2011; Aleman et al., 2017; Xu et al., 2018). The results show that rather than a single threshold in the dry season PV components (MAP_d, λ_d, α_d, and T_d; Figure 4a-d) and the soil silt content (Figure 4e), there is a broad range in each variable where both savanna (canopy cover < 40%) and forest (canopy cover ≥ 40%; ranges delineated by solid red and black lines, respectively) can occur. Notably, fire occurrence (Figure 4f) is almost exclusively linked to the presence of savanna, with scattered fire pixels in the forest outside of the distribution (10^{th}-90^{th} percentiles). Additionally, it is important to note that ranges do not differ greatly between vegetation descriptors, except MAP_d (90^{th} percentile between 822 - 980 mm) in savanna ranges, when savanna and forest are defined using tree cover or $PAVD_{max}$ instead of the canopy cover.

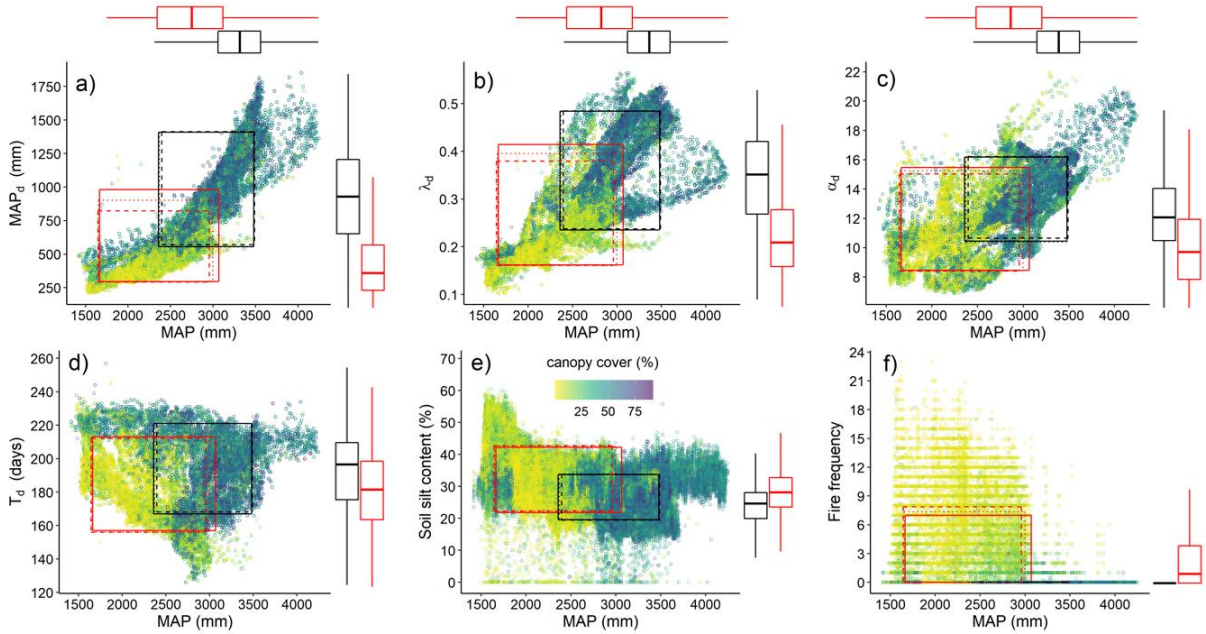

Figure 4. Relationships between mean annual precipitation (MAP) and a) mean total dry-season precipitation (MAP_d), b) frequency of wet days (precipitation > 0) within the dry season (λ_d), c) intensity of wet days within the dry season (α_d), d) length of the dry season (T_d), e) soil silt content (%), and f) fire frequency. Ranges delineated by different lines indicate the 10th and 90th percentiles for each variable for savanna (red) and forest (black) defined by canopy cover (savanna < 40% and forest ≥ 40%; solid line), $PAVD_{max}$ (savanna < 0.10 and forest ≥ 0.10; dashed line), and tree cover (savanna < 60% and forest ≥ 60%; dotted line). The boxplots show the values of each variable for all savanna (red) and forest (black) pixels defined by canopy cover threshold (savanna < 40% and forest ≥ 40%).

Forest dominates in pixels with MAP > 2360 mm, while savanna occurs in MAP levels between 1665 mm and 3070 mm, indicating that between 2360 mm and 3070 mm both savanna and forest can occur (Figure 4a). However, pixels within these MAP ranges that have different PV values (i.e., MAP_d, λ_d, α_d, or T_d) exhibit different canopy cover values. For example, in the MAP range where both forest and savanna can occur (i.e., 2360 to 3070 mm), savanna dominates if MAP_d and λ_d are lower than 556 mm and 0.23, respectively (Figure 4a and 4b). However, MAP_d and λ_d also exhibit a broad range where both forest and savanna can occur (556 mm ≤ MAP_d ≤ 981 mm and 0.23 < λ_d <0.41). α_d and T_d show lower variability (i.e., lower coefficient of variance) as well as lower difference between forest (10.4 mm/day < α_d <16.2 mm/day; 167 days < T_d < 221 days) and savanna (8.5 mm/day < α_d <15.5 mm/day; 157 days < T_d < 213 days) 10th-90th percentile ranges than MAP_d and λ_d (Figure 4c and 4d), consistent with results shown in Figure 2f and 2g. Indeed, both forest and savanna can occur when α_d and T_d vary between 10.4 mm/day < α_d <15.5 mm/day and 167 days < T_d < 213 days, respectively.

Soil silt content also shows little differences between savanna and forest ranges (Figure 4e). Overall, savanna dominates in pixels with higher silt content (21.8% to 42.2%) than the forest (19.6% to 33.7%), with a relatively narrow range where both occur. Finally, while savanna tends to occur in pixels where fires are present, fires are absent in the forest (Figure 2i and 4f).

However, our results also show that 46% of pixels (N=53271) classified as savanna (canopy cover < 40%) do not have fires in the period 2001-2019.

3.5. Dry season length of dry spells and frequency of wet days

To further explore the relationship between forest-savanna transition and the dry season PV components, we also plotted MAP versus the length of dry spells (i.e., the consecutive number of days without precipitation; Ds_d) and the frequency of wet days (with intensity < 10 mm; λ_{10d}). We selected this threshold because the events with 10 mm or less represent approximately 70% of the precipitation events during the dry season in the study area. Like other PV components, Ds_d and λ_{10d} exhibit a large range in which forest and savanna can occur. Savanna dominates in pixels with Ds_d> 3.4 days, while forest occurs in Ds_d periods between 2.6 and 6.2 days (Figure 5a). In addition to the lower precipitation frequency and intensity as well as long dry spells in savanna, most precipitation events are of lower intensity ($0.80 < \lambda_{10d} < 0.95$) than in the forest ($0.71 < \lambda_{10d} < 0.90$; Figure 5b). This evidences that savanna occurs in pixels with longer Ds_d and lower λ_{10d}, corresponding to pixels with lower MAP_d, λ_d, α_d, and more frequent fires than forest.

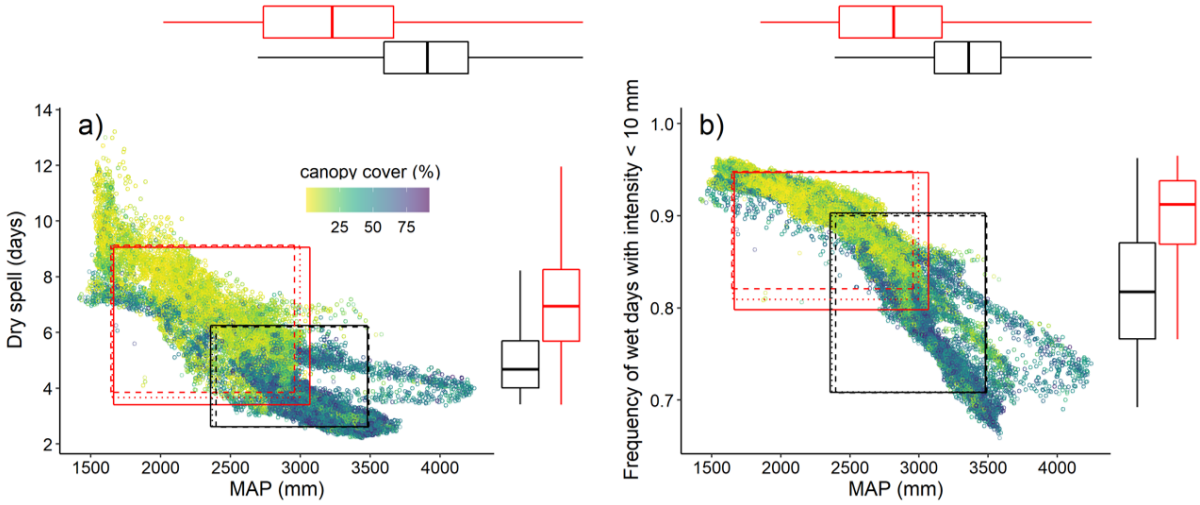

Figure 5. a) dry spells (Ds_d) and b) frequency of wet days with intensity < 10 mm/day during the dry season (λ_{10d}). Ranges by different lines indicate the 10th and 90th percentiles for each variable for savanna (red) and forest (black) defined by canopy cover (savanna < 40% and forest ≥ 40%; solid line), $PAVD_{max}$ (savanna < 0.10 and forest ≥ 0.10; dashed line), and tree cover (savanna < 60% and forest ≥ 60%; dotted line). The boxplots show the values of each variable for all savanna (red) and forest (black) pixels defined by canopy cover threshold (savanna < 40% and forest ≥ 40%).

3.6. Determinants of the forest-savanna transition

All selected predictor variables, except soil silt content, have a statistically significant effect on determining canopy cover in forest-savanna transition (results from GLMs; Figure 6). In particular, models can explain between 52.2% and 55.4% of the deviance of the data (Supporting Information Table S3). Our results indicate that fire frequency followed by λ_d are the most important predictor variables for canopy cover as indicated by the magnitude of

standardized estimates for both variables (-0.79 and 0.40 for fire frequency and λ_d, respectively). However, as expected, fire frequency has a negative effect on canopy cover while λ_d has a positive effect. Given the direct association between the PV components ($MAP_d = \lambda_d \alpha_d T_d$), when α_d and T_d are held constant, the increase of λ_d corresponds with an increase in MAP_d that results in increased canopy cover. Consequently, canopy cover and λ_d are higher in areas with higher MAP_d (Figure 3 and 4; Supporting Information Figure S3). Although also α_d and T_d are also statistically significant, these variables have a lower explanatory power for canopy cover (Figure 6 and Supporting Information Table S3). Finally, soil silt content has the lowest and non-statistically significant effect on canopy cover. Importantly, models are also informative when using both tree cover and $PAVD_{max}$ (Supporting Information Figure S8; Supporting Information Table S3). Overall, the magnitude and significance of standardized estimates are similar between vegetation descriptors. However, although fire frequency is the variable with the highest explanatory power with both canopy cover and $PAVD_{max}$, the magnitude of λ_d (0.48) is slightly higher than fire frequency (-0.42) with tree cover. These results highlight the importance of using multiple descriptors when assessing structurally-diverse ecotones such as the forest-savanna transition.

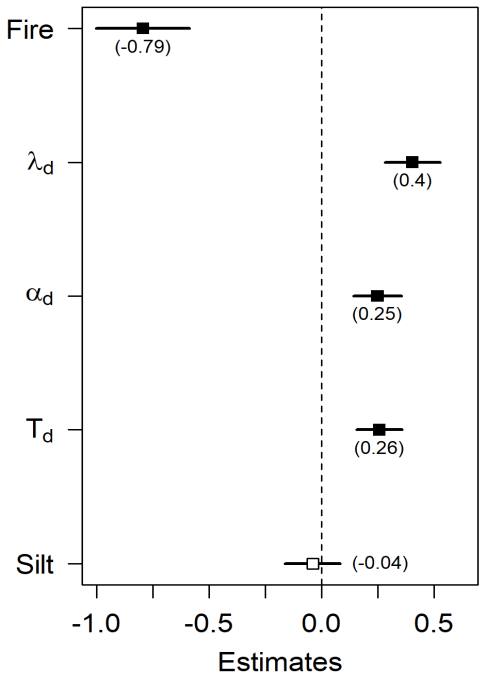

Figure 6. Effect of each predictor variable on canopy cover. Predictor variables were standardized such that in the GLM their coefficient magnitude is a measure of their importance in the model. The median estimate (points, values between parentheses) and 95% confidence interval (error bars) are based on 1000 GLMs (see Section 2.5 and Supporting information S4). Terms are not significant (open symbol) when the confidence interval includes zero (dashed vertical line). Predictors are: fire frequency (fire), mean daily precipitation frequency in dry season (λ_d), mean daily precipitation intensity in dry season (α_d), mean length of dry season (T_d), soil silt content (silt). See Supporting Information Figure S8 for equivalent analyses with tree cover and $PAVD_{max}$, and Supporting Information Table S3 for details of GLMs results.

4. DISCUSSION

4.1. Beyond a MAP threshold in forest-savanna transition

Multiple thresholds, mainly based on MAP, have been suggested to define tropical forest and savanna distribution at regional (Malhi et al., 2009; Ciemer et al., 2019), continental (Sankaran et al., 2005, Bucini & Hanan, 2007; Good & Caylor, 2011; Lehmann et al., 2011; Staal et al., 2020), and global scales (Staver et al., 2011; Archibald et al., 2019). Our results show a broad MAP range where both forest and savanna can occur, but this range (2360 < MAP < 3070 mm, Figure 4a) is larger than previously proposed for South America (1200 < MAP < 2100 mm, Ciemer et al., 2019; Staal et al., 2020) and globally (1000 < MAP < 2500 mm; Staver et al., 2011). This suggests that MAP ranges in which both forests and savannas occur in the Llanos region correspond, almost exclusively, to forest in other South American regions (e.g., the Cerrado, Figure 2) and elsewhere.

Extending this analysis to considering intra-seasonal precipitation characteristics, our results also indicate that both forest and savanna can occur in a large interval of intermediate values of dry season PV components (Figures 4a-d and Figure 5). Our results suggest that there are MAP levels in which forest (> 3070 mm) and savanna (< 2360 mm) dominate independent of the e.g., λ_d, α_d, T_d, or MAP_d. However, regions with similar MAP but different, e.g., λ_d, α_d or Ds_d, exhibit different canopy cover values (Figures 4 and 5; as discussed below). Indeed, various combinations of mean dry season lengths, as well as precipitation frequency and intensity, can result in a similar MAP (Good & Caylor, 2011) or MAP_d. For example, at intermediate MAP (2360 < MAP < 3070 mm), our results show that savanna (forest) dominates if $\lambda_d$andα_d are lower (higher) than 0.23 and 8.5 mm/day (0.41 and 15.5 mm/day), respectively. These results highlight that the present-day and future definition of forest and savanna distribution requires additional consideration of a climatic space with multiple precipitation characteristics, as suggested by Schwartz et al. (2020) for tropical ecosystems.

The transition extent highlights the complex interactions between vegetation and environmental factors that determine forest and savanna distribution at regional scales. Additionally, although previous studies (e.g., Malhi et al., 2009) have highlighted a potential effect of errors in vegetation descriptors or phenological classification, as well as in precipitation data on the absence of sharp climatic thresholds (Figures 3ad-g and Figures 4a-d), our results are robust across the different vegetation descriptors from GEDI and MODIS. Further, precipitation data from CHIRPS exhibits a high performance for our study area (Paredes-Trejo et al., 2016; Cavalcante et al., 2020; Valencia et al., in preparation; and Supporting Information Figure S9).

4.2. The role of fire frequency on forest-savanna transition

Our results coincide with several studies reporting that fire is one of the most important factors to explain the present-day (Bond, 2008; Hirota et al., 2011; Staver et al., 2011; Staver et al., 2017; Xu et al., 2018; Newberry et al., 2020) and the past (e.g., during the Last Glacial Maximum, Sato et al., 2021) forest and savanna distribution. Fire-vegetation feedbacks in

savannas allow frequent burning that maintains an open-canopy where both climate and soil could otherwise support forest, consistent with the idea of a fire-suppression threshold (Hoffmann et al., 2012; Bernardino et al., 2021). This can help to explain the observed savanna pixels (i.e., canopy cover < 40) in the same climatic space or soil silt content of forest pixels (i.e., canopy cover > 40; Figure 4a-e). However, our results also indicate that 46% of the savanna pixels (N=53271) do not present fires over 2001-2019 (Figure 4f and Supporting Information Figure S3i). More interestingly, between 30% and 42% of savanna pixels in the same climatic space or silt soil content ranges of forest pixels (i.e., intersections between the red and black rectangles in Figures 2 and Figures 4a-e) also do not present fires for this period (Supporting Information Figure S10). This suggests that fire frequency alone does not explain the occurrence of savanna pixels in the same climatic or edaphic space of forest in the Llanos region.

Our results also show that fires in the Llanos savannas occur in pixels with MAP levels (> 2000 mm, Figure 4f) where fires are unlikely in other savanna regions (Lehmann et al., 2011; Dantas et al., 2016, Staal et al., 2018), evidencing the complex interactions between fire, vegetation, climate, topography, and human activities in this region (Romero-Ruiz et al., 2010; Oliveras & Malhi, 2019). Notably, most fires in the Llanos region occur during the dry season (November to April-May), mostly associated with human activities such as traditional agricultural practices and cattle grazing (Armenteras et al., 2005; Romero-Ruiz et al., 2010; Armenteras et al., 2020). More specifically, Barreto & Armenteras (2020) show that vegetation (Normalized Difference Water Index) followed by mean monthly temperature and human alteration are the most important variables predicting the probability of fire occurrence in the Llanos ecoregion. This key role of human activities in present-day fire regimes has also been documented in the Cerrado (e.g., Conciani et al., 2021) and Africa savannas (e.g., Archibald, 2016) via fire ignition or suppression. Finally, despite burned area products being one of the best sources of data to estimate fire frequency at regional and global scales (Lizundia-Loiola et al., 2020), they have limitations associated with relatively narrow time coverage, precluding the identification of fires with long return intervals (> 20 years), potentially explaining the observed savanna pixels without fire present in our study area.

4.3. Precipitation variability as a determinant of forest-savanna transition

Our statistical analysis reveals that PV components, particularly λ_d, contribute significantly to explaining the vegetation variations in forest-savanna transition (Figure 6 and Supporting Information Table S3). More specifically, although we do not assess the PV effect on canopy cover at different MAP windows as Xu et al. (2018), our results suggest that PV can be more important at intermediate MAP levels (i.e., 2360 < MAP < 3070 mm) in which both forest and savanna occur (Figures 4a-d and Figure 5). Our results are consistent with previous studies showing that PV is a key determinant of forest and savanna dynamics and distribution at local and global scales (e.g., Good & Caylor, 2011; Kulmatiski & Beard, 2013; Guan et al., 2014; Case & Staver, 2018; Xu et al., 2018; D'onofrio et al., 2019). However, in contrast with those studies, our analysis indicates that dry season PV components are more related to forest-

savanna transition than PV components during the wet season, consistent with Zeng et al. (2014) and Hoyos et al. (2021).

Overall, our results show that areas have higher canopy cover if precipitation consists of more frequent and intense events (Figures 4b-c and Figure 6). More frequent and intense precipitation events (i.e., regular precipitation) can lead to low fluctuations in soil moisture during the dry season, which results in higher water availability and decreasing plant water stress (Knapp et al., 2008, 2015; Liu et al., 2020). More specifically, Ds_d and λ_{10d} highlight how shorter dry periods -- as suggested by Hoyos et al. (2021) -- and larger precipitation amounts in each event increase canopy cover (Figure 5). However, Good & Caylor (2011) and Xu et al. (2018) show that tree cover is also higher in areas where precipitation is more frequent, but less intense, evidencing that the response of tropical vegetation to precipitation frequency and intensity is heterogeneous, varying with the regional environment context, seasonality (wet versus dry season), and with differences in water-use strategies between grasses (i.e., savanna) and trees (i.e., forest) (Case & Staver, 2018). Although sandier soils can help to explain increases of canopy cover with precipitation intensity, at least in African savannas (e.g., Staver et al., 1017 and Case & Staver, 2018), we did not observe differences in soil sand, clay, or silt content between forest and savanna (Figure 4e and Supporting Information Figure S3). Additionally, soil silt content does not significantly affect canopy cover, tree cover, or $PAVD_{max}$ (Figure 6 and Supporting Information Figure S8). Our analysis suggests that soil properties (texture and fertility) do not provide an alternative mechanism to explain the forest-savanna transition or the effect of PV on vegetation in the Llanos. This is consistent with the results of Hoyos et al. (2021), who highlight the low explanatory power of soil units on the probability of forest occurrence in the Llanos, which may be related to similar long-term climate, parent material, and relief across the transition. However, it is possible that global soil databases are insufficiently accurate or fine-scaled to represent differences in soil properties between tropical forest and savanna as well as among savanna types in the Llanos (i.e., permanently and seasonally flooded savannas and high plain savannas) reported in previous studies (e.g., Armenteras et al., 2005; Romero-Ruiz et al., 2010; Sánchez & Armenteras, 2017).

Savannas occur more commonly in regions with longer T_d than forests (Archibald et al., 2019). However, our results show that despite the low explanatory power of T_d (Supporting Information Table S3), it has a positive effect on canopy cover (Figure 6 and Supporting Information Figure S8). This suggests that a long dry season is not necessary for the occurrence of savanna, consistent with Staver et al. (2011) for South America, highlighting how the precipitation distribution within the dry season may be more important than its duration. Finally, from a methodological perspective, a preliminary analysis suggests how the wet season length definition approach may lead to differences in its duration (Supporting Information Figure S11).

4.4. Alternative determinants of forest-savanna transition

In regions with high precipitation, long wet seasons, and high nutrients availability can result in open-canopy conditions (e.g., savanna) independent of the fire regime (Archibald et al.,

2019). This may also contribute to explaining savanna pixels without fire present in the same climatic or edaphic space of forest in our study area (Figure 4 and Supporting Information Figure S10). For example, seasonal flooding is common in some savanna regions of the Llanos in the wet season (Romero-Ruiz et al., 2012; Borguetti et al., 2019), which may maintain a lower canopy cover via waterlogged soils (Oliveras & Malhi, 2016; Daskin et al., 2019). However, although a preliminary analysis confirms that some savanna pixels show high water occurrence (Supporting Information Figure S12a), only ~2% (N=560) of savanna pixels in the same climatic or edaphic space of forest also exhibit water occurrence in the period 1984 - 2020 (Supporting Information Figure S12b). Therefore, alternative factors that may contribute to explaining forest-savanna transition in Llanos and require further exploration include: i) root depth (Langan et al., 2017; Sakschewski et al., 2021; Singh et al., 2022); ii) water table depth (Ferreira-Júnior et al., 2016; Ribeiro et al., 2021); iii) tree-grass competition (Xu et al., 2018).

The present-day tropical forest and savanna distribution is the result of both current (i.e., year and decades) and long-term (i.e., centuries and millennia) relationships and feedbacks between vegetation and environmental factors (Oliveras & Malhi, 2016; Jaramillo, 2019; Azevedo et al., 2020). For example, human activities (e.g., via agriculture and hunting) have led to fire regime changes and megafauna extinction, resulting in forest (savanna) expansion into savanna (forest) around the tropics (Berrio et al., 2012; Oliveras & Malhi, 2016 and references therein; Doughty et al., 2016; Dantas & Pausas, 2022) over the past centuries and millennia. For this reason, although we did not consider changes in the tropical forest and savanna limits in our study – due to limited historical land cover maps and low land cover changes across the transition for the study period (Romero-Ruiz et al., 2012) –, dynamical change in tropical forest and savanna limits is an important element of this transition. Finally, the atmospheric CO_2 concentration is another environmental factor that can help to explain the tropical forest - savanna transition (Oliveras & Malhi, 2016; Sato et al., 2021) and which have been largely changed over the past millennia, particularly, during the last 200 years (280 ppm in the pre-industrial era to around 400 ppm by 2015). For example, Higgins & Scheiter, (2012) shows that the probability of C4 (grassland or savanna) dominance increases at intermediate MAP levels (1000 mm - 1500 mm) from < 0.2 to > 0.7 when atmospheric CO_2 concentration decreases from 400 ppm to 170 ppm in Africa. Additional works need to be done to dress the effect of CO_2, including its feedbacks with precipitation variability and fire regime, on the past, present, and future tropical forest and savanna distribution, mainly due to atmospheric CO_2 concentration will continue changing at least for the next 50-100 years (IPCC, 2021).

4.5. Implications in the context of environmental change

Our results show that canopy cover, tree cover, and $PAVD_{max}$ are associated primarily with fire frequency and λ_d, which support that future projections of forest and savanna dynamics and distribution should not only consider MAP changes (e.g., Zeng et al., 2013; Anadón et al., 2014; Aleman et al., 2017; Anjos & Toledo, 2018; Ciemer et al., 2019; Staal et al., 2020; Anjos et al., 2021), but also changes in PV components (Guan et al., 2014; Xu et al., 2018) and fire-vegetation feedbacks (De Faria et al., 2021). In addition, there is high confidence in a dominant increase in the number of dry days (i.e., lower precipitation frequency) and drought frequency

in regions like northern South America (IPCC, 2021). The impacts of climate change on vegetation are both direct through changes in soil water availability (Knapp et al., 2008, 2015; Liu et al., 2020) and indirect via alterations in fire regimes, which are amplified by land-use dynamics (e.g., fire ignition or suppression) (Andela et al., 2017; Zheng et al., 2021). These alterations modify not only forest and savanna distribution (e.g., via savannization or woody/forest encroachment) but also their structure, composition, dynamics, and associated ecosystem services (Oliveras & Malhi, 2016, and references therein, including Lipsett-Moore et al., 2018, Brando et al., 2019; Rosan et al., 2019; Durigan, 2020; Stark et al., 2020; Armenteras et al., 2021a; De Faria et al., 2021). However, biogeographic differences at regional scale can lead to contrasting responses of tropical forest and savannas to environmental change as evidenced by several studies (e.g., Lehmann et al., 2014; Oliveras & Malhi, 2016; Stevens et al., 2017; Xu et al., 2018; Alvarado et al., 2019; Esquivel-Muelbert et al., 2020). Our results highlight how the response of forest-savanna transition to environmental change would also be different between the North (i.e., Llanos) and South (i.e., the Cerrado) Amazon region.

5. CONCLUSIONS

We found that savannas in the Llanos ecoregion occur in a MAP range that would be associated with tropical forest, according to MAP ranges proposed to other savanna regions (e.g., the Cerrado). In addition, our analysis shows that MAP ranges in which both forests and savannas occur in the Llanos region correspond, almost exclusively, to forest in other South American regions and globally. Although both forest and savanna can also occur in a large interval of intermediate values of dry season PV components, forest dominates in areas with higher precipitation frequency and intensity than savanna. Although fire frequency is higher in savanna than forest, a large proportion of pixels classified as savanna pixels had no fires in the analysis period, even those that occur in the same climatic or edaphic space of the forest. In summary, our analysis shows that fire frequency and dry season precipitation are the most important variables to predict canopy cover, as well as tree cover and $PAVD_{max}$, in forest-savanna transition. This highlights the role of fire regime and water availability in determining the limits between forest and the second largest area of savanna in South America. Further, our results highlight the importance of refining our understanding of the factors, relationships, and mechanisms that control forest-savanna transition at regional scales, as a requirement to assess the effects of environmental change on this ecologically, biogeochemically, and climatically important ecotone.

ACRONYM DICTIONARY

$\lambda_{d/w}$	Mean frequency of wet days (daily precipitation > 0) in the dry (d) and wet (w) season	MCWD	Maximum Climatological Water Deficit
$\alpha_{d/w}$	Mean daily dry (d) and wet (w) season precipitation intensity	MODIS	Moderate Resolution Imaging Spectroradiometer
λ_{10d}	Frequency of wet days with intensity < 10 mm/day during the dry season	$PAVD_{max}$	Maximum Plant Area Volume Density

AI	Aridity Index	PET	Potential Evapotranspiration
CEC	Cation Exchange Capacity	PV	Precipitation Variability
CHIRPS	Climate Hazards Group InfraRed Precipitation with Station data	r	Pearson correlation coefficient
Ds_d	Length of dry spells	R^2	Deviance explained
ESA	European Space Agency	RADS	Rainy and Dry seasons dataset
F	Forest	r_s	Spearman rank correlation analysis
FIRECC51	The ESA FireCCI project	S	Savanna
GEDI	Global Ecosystem Dynamics Investigation	SOC	Soil Organic Carbon
GLM	Generalized Linear Model	SRTM	Shuttle Radar Topographic Mission
ISS	International Space Station	T	Transition
MAP	Mean Annual Precipitation	$T_{d/w}$	Mean length of the dry (d) and wet (w) season
$MAP_{d/w}$	Mean total dry (d) and wet (w) season precipitation		

DATA ACCESSIBILITY

The original data used in this study are all publicly available from their sources: GEDI: https://gedi.umd.edu/, Terraclimate: http://www.climatologylab.org/terraclimate.html, ESA: http://maps.elie.ucl.ac.be/CCI/viewer/download.php, SoilGrids: https://soilgrids.org/, SRTM: https://srtm.csi.cgiar.org/srtmdata/, MODIS: https://ladsweb.modaps.eosdis.nasa.gov/, RADS: https://climatology.tamu.edu/research/. The MCWD data can be produced using data from CHIRPS combined with code available from Campanharo & Silva-Junior, (2019) at: https://doi.org/10.5281/zenodo.2652629. The dataset that supports the findings of this study will be available through a data-sharing repository.

ACKNOWLEDGMENTS

This work was funded by the Universidad de Antioquia through the Estudiante Instructor Program for graduate studies and the Colombian Ministry of Science, Technology and Innovation (MINCIENCIAS) through program “Sostenibilidad de sistemas ecológicos y sociales en la cuenca Magdalena-Cauca bajo escenarios de cambio climático y pérdida de bosques” (code 1115-852-70719) with funds from “Patrimonio Autónomo Fondo Nacional de Financiamiento para la Ciencia, la Tecnología y la Innovación, Fondo Francisco José de Caldas”. The authors are grateful to Diana María Agudelo (Universidad de Antioquia) and Alex Correa-Metrio (Universidad Nacional Autónoma de México) for their support with the statistical analysis. We also appreciate the comments and suggestions of Catalina González (Universidad de los Andes) and Carlos Jaramillo (Smithsonian Tropical Research Institute) as jury members of this thesis.

SUPPORTING INFORMATION

The Supplementary Material for this article can be found at the end of this document.

REFERENCES

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., & Hegewisch, K. C. (2018). TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Scientific data*, 5, 170191.

Aleman, J. C., Blarquez, O., Gourlet-Fleury, S., Bremond, L., & Favier, C. (2017). Tree cover in Central Africa: determinants and sensitivity under contrasted scenarios of global change. *Scientific reports*, *7*, 41393.

Alvarado, S. T., Andela, N., Silva, T. S., & Archibald, S. (2020). Thresholds of fire response to moisture and fuel load differ between tropical savannas and grasslands across continents. *Global Ecology and Biogeography*, *29*(2), 331-344.

Andela, N., Morton, D. C., Giglio, L., Chen, Y., van der Werf, G. R., Kasibhatla, P. S., ... & Randerson, J. T. (2017). A human-driven decline in global burned area. *Science*, *356*(6345), 1356-1362.

Anderson, L.O., Burton, C., dos Reis, J.B.C., Pessôa, A.C.M., Bett, P., Carvalho, N.S., et al. (2021) An alert system for Seasonal Fire probability forecast for South American Protected Areas. *Climate Resilience and Sustainability*, 1–19.

Aragão, L. E. O., Malhi, Y., Roman-Cuesta, R. M., Saatchi, S., Anderson, L. O., & Shimabukuro, Y. E. (2007). Spatial patterns and fire response of recent Amazonian droughts. *Geophysical Research Letters*, 34(7).

Archibald, S., Bond, W. J., Hoffmann, W., Lehmann, C., Staver, C., & Stevens, N. (2019). Distribution and determinants of savannas. *Savanna woody plants and large herbivores*, 1-24.

Archibald, S. (2016). Managing the human component of fire regimes: lessons from Africa. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1696), 20150346.

Armenteras, D., Romero, M., & Galindo, G. (2005). Vegetation fire in the savannas of the Llanos Orientales of Colombia. *World Resource Review*, 17(4), 531-543.

Armenteras, D., González, T. M., Vargas Ríos, O., Meza Elizalde, M. C., & Oliveras, I. (2020). Fire in the ecosystems of northern South America: advances in the ecology of tropical fires in Colombia, Ecuador and Peru. *Caldasia*, *42*(1), 1-16.

Armenteras, D., Dávalos, L. M., Barreto, J. S., Miranda, A., Hernández-Moreno, A., Zamorano-Elgueta, C., ... & Retana, J. (2021a). Fire-induced loss of the world's most biodiverse forests in Latin America. *Science Advances*, *7*(33), eabd3357.

Armenteras, D., Meza, M. C., González, T. M., Oliveras, I., Balch, J. K., & Retana, J. (2021b). Fire threatens the diversity and structure of tropical gallery forests. *Ecosphere*, *12*(1), e03347.

Azevedo, J. A., Collevatti, R. G., Jaramillo, C. A., Strömberg, C. A., Guedes, T. B., Matos-Maraví, P., ... & Antonelli, A. (2020). On the young savannas in the land of ancient forests. In *Neotropical diversification: Patterns and processes* (pp. 271-298). Springer, Cham.

Barreto, J. S., & Armenteras, D. (2020). Open Data and Machine Learning to Model the Occurrence of Fire in the Ecoregion of "Llanos Colombo–Venezolanos". *Remote Sensing*, 12(23), 3921.

Berrio, J. C., Wouters, H., Hooghiemstra, H., Carr, A. S., & Boom, A. (2012). Using paleoecological data to define main vegetation dynamics along the savanna–forest ecotone in Colombia: implications for accurate assessment of human impacts. In *Ecotones between forest and grassland* (pp. 209-225). Springer, New York, NY.

Behling, H., & Hooghiemstra, H. (2000). Holocene Amazon rainforest–savanna dynamics and climatic implications: high-resolution pollen record from Laguna Loma Linda in eastern Colombia. *Journal of Quaternary Science: Published for the Quaternary Research Association*, 15(7), 687-695.

Bernardino, P. N., Dantas, V. L., Hirota, M., Pausas, J. G., & Oliveira, R. S. (2021). Savanna–Forest Coexistence Across a Fire Gradient. *Ecosystems*, 1-12.

Bond, W. J. (2008). What limits trees in C4 grasslands and savannas?. *Annual review of ecology, evolution, and systematics*, *39*, 641-659.

Bombardi, R. J., Kinter III, J. L., & Frauenfeld, O. W. (2019). A global gridded dataset of the characteristics of the rainy and dry seasons. *Bulletin of the American Meteorological Society*, *100*(7), 1315-1328.

Borghetti, F., Barbosa, E., Ribeiro, L., Ribeiro, J. F., & Walter, B. M. T. (2019). South American Savannas. *Savanna Woody Plants and Large Herbivores*, 77-122.

Brando, P. M., Paolucci, L., Ummenhofer, C. C., Ordway, E. M., Hartmann, H., Cattau, M. E., ... & Balch, J. (2019). Droughts, wildfires, and forest carbon cycling: A pantropical synthesis. *Annual Review of Earth and Planetary Sciences*, *47*, 555-581.

Breshears, D. D. (2006). The grassland–forest continuum: trends in ecosystem properties for woody plant mosaics?. *Frontiers in Ecology and the Environment*, *4*(2), 96-104.

Bucini, G., & Hanan, N. P. (2007). A continental-scale analysis of tree cover in African savannas. *Global Ecology and Biogeography*, *16*(5), 593-605.

Buol, S. W., & Eswaran, H. (1999). Oxisols. *Advances in agronomy*, *68*, 151-195.

Campanharo, W. A., and Silva Junior, C. H. L. (2019). Maximum Cumulative Water Deficit - MCWD: a R language script. doi:10.5281/zenodo.2652629

Case, M. F., & Staver, A. C. (2018). Soil texture mediates tree responses to precipitation intensity in African savannas. *New Phytologist*, *219*(4), 1363-1372.

Cavalcante, R. B. L., da Silva Ferreira, D. B., Pontes, P. R. M., Tedeschi, R. G., da Costa, C. P. W., & de Souza, E. B. (2020). Evaluation of extreme rainfall indices from CHIRPS precipitation estimates over the Brazilian Amazonia. Atmospheric Research, 238, 104879.

Chuvieco, E., Lizundia-Loiola, J., Pettinari, M. L., Ramo, R., Padilla, M., Tansey, K., ... & Plummer, S. (2018). Generation and analysis of a new global burned area product based on MODIS 250 m reflectance bands and thermal anomalies. *Earth System Science Data*, *10*(4), 2015-2031.

Conciani, D. E., dos Santos, L. P., Silva, T. S. F., Durigan, G., & Alvarado, S. T. (2021). Human-climate interactions shape fire regimes in the Cerrado of São Paulo state, Brazil. *Journal for Nature Conservation*, *61*, 126006.

de Sousa, L. M., Poggio, L., Batjes, N. H., Heuvelink, G., Kempen, B., Riberio, E., & Rossiter, D. (2020). SoilGrids 2.0: producing quality-assessed soil information for the globe. *Soil Discussions*, 1-37.

Dantas, V. D. L., Hirota, M., Oliveira, R. S., & Pausas, J. G. (2016). Disturbance maintains alternative biome states. *Ecology letters*, *19*(1), 12-19.

Daskin, J. H., Aires, F., & Staver, A. C. (2019). Determinants of tree cover in tropical floodplains. *Proceedings of the Royal Society B*, *286*(1914), 20191755.

Decuyper, M., Mulatu, K. A., Brede, B., Calders, K., Armston, J., Rozendaal, D. M., ... & Bongers, F. (2018). Assessing the structural differences between tropical forest types using terrestrial laser scanning. *Forest Ecology and Management*, 429, 327-335.

De Faria, B. L., Staal, A., Silva, C. A., Martin, P. A., Panday, P. K., & Dantas, V. L. Climate change and deforestation increase the vulnerability of Amazonian forests to post-fire grass invasion. *Global Ecology and Biogeography*.

DiMiceli, C., Carroll, M., Sohlberg, R., Kim, D., Kelly, M., Townshend, J. (2015). MOD44B MODIS/Terra Vegetation Continuous Fields Yearly L3 Global 250m SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC. Accessed 2020-12-22 from https://doi.org/10.5067/MODIS/MOD44B.006

DiMiceli, C., Townshend, J., Carroll, M., & Sohlberg, R. (2021). Evolution of the representation of global vegetation by vegetation continuous fields. *Remote Sensing of Environment*, *254*, 112271.

Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N. D., Wikramanayake, E., ... & Hansen, M. (2017). An ecoregion-based approach to protecting half the terrestrial realm. *BioScience*, *67*(6), 534-545.

D'Onofrio, D., Sweeney, L., von Hardenberg, J., & Baudena, M. (2019). Grass and tree cover responses to intra-seasonal rainfall variability vary along a rainfall gradient in African tropical grassy biomes. *Scientific reports*, *9*(1), 1-10.

D'Onofrio, D., von Hardenberg, J., & Baudena, M. (2018). Not only trees: Grasses determine African tropical biome distributions via water limitation and fire. *Global ecology and biogeography*, *27*(6), 714-725.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., ... & Lautenbach, S. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36(1), 27-46.

Doughty, C. E., Faurby, S., & Svenning, J. C. (2016). The impact of the megafauna extinctions on savanna woody cover in South America. *Ecography*, *39*(2), 213-222.

Dubayah, R., Blair, J. B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., ... & Armston, J. (2020). The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of Remote Sensing*, 1, 100002.

Durigan, G. (2020). Zero-fire: Not possible nor desirable in the Cerrado of Brazil. *Flora*, *268*, 151612.

Ehbrecht, M., Seidel, D., Annighöfer, P., Kreft, H., Köhler, M., Zemp, D. C., ... & Ammer, C. (2021). Global patterns and climatic controls of forest structural complexity. *Nature communications*, *12*(1), 1-12.

Erb, K. H., Kastner, T., Plutzar, C., Bais, A. L. S., Carvalhais, N., Fetzel, T., ... & Luyssaert, S. (2018). Unexpectedly large impact of forest management and grazing on global vegetation biomass. *Nature*, *553*(7686), 73-76.

ESA, (2017). Land Cover CCI Product User Guide Version 2.0, available at: http://maps.elie.ucl.ac.be/CCI/viewer/download/ ESACCI-LC-Ph2-PUGv2_2.0.pdf, last access: 10 November 2017

Esquivel-Muelbert, A., Phillips, O. L., Brienen, R. J., Fauset, S., Sullivan, M. J., Baker, T. R., ... & Galbraith, D. (2020). Tree mode of death and mortality risk factors across Amazon forests. *Nature communications*, *11*(1), 1-11.

Fageria, N. K., & Nascente, A. S. (2014). Management of soil acidity of South American soils for sustainable crop production. *Advances in agronomy*, *128*, 221-275.

February, E. C., Coetsee, C., Cook, G. D., Ratnam, J., & Wigley, B. (2019). Physiological traits of savanna woody species: Adaptations to resource availability. *Savanna woody plants and large herbivores*, 309-329.

Ferreira-Júnior, W. G., Schaefer, C. E., Cunha, C. N., Duarte, T. G., Chieregatto, L. C., & Carmo, F. (2016). Flood regime and water table determines tree distribution in a forest-savanna gradient in the Brazilian Pantanal. *Anais da Academia Brasileira de Ciências*, *88*, 719-731.

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International journal of climatology*, 37(12), 4302-4315.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., ... & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific data*, 2(1), 1-21.

Good, S. P., & Caylor, K. K. (2011). Climatological determinants of woody cover in Africa. *Proceedings of the National Academy of Sciences*, *108*(12), 4902-4907.

Giglio, L., Boschetti, L., Roy, D. P., Humber, M. L., & Justice, C. O. (2018). The Collection 6 MODIS burned area mapping algorithm and product. *Remote sensing of environment*, *217*, 72-85.

Guan, K., Good, S. P., Caylor, K. K., Medvigy, D., Pan, M., Wood, E. F., ... & Xu, X. (2018). Simulated sensitivity of African terrestrial ecosystem photosynthesis to rainfall frequency, intensity, and rainy season length. *Environmental Research Letters*, *13*(2), 025013.

Hempson, G. P., Archibald, S., & Bond, W. J. (2015). A continent-wide assessment of the form and intensity of large mammal herbivory in Africa. *Science*, *350*(6264), 1056-1061.

Hengl, T., Mendes de Jesus, J., Heuvelink, G. B., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., ... & Guevara, M. A. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLoS one*, 12(2), e0169748.

Higgins, S. I., & Scheiter, S. (2012). Atmospheric CO2 forces abrupt vegetation shifts locally, but not globally. *Nature*, *488*(7410), 209-212.

Hijmans, R. J. (2020). Geographic Data Analysis and Modeling [R package raster version 3.4-5].

Hirota, M., Holmgren, M., Van Nes, E. H., & Scheffer, M. (2011). Global resilience of tropical forest and savanna to critical transitions. *Science*, *334*(6053), 232-235.

Hoffmann, W. A., Geiger, E. L., Gotsch, S. G., Rossatto, D. R., Silva, L. C., Lau, O. L., ... & Franco, A. C. (2012). Ecological thresholds at the savanna-forest boundary: how plant

traits, resources and fire govern the distribution of tropical biomes. *Ecology letters*, *15*(7), 759-768.

Hoyos, N., Correa-Metrio, A., Jaramillo, C., Villegas, J. C., & Escobar, J. (2021). Effects of consecutive dry and wet days on the forest–savanna boundary in north-west South America. *Global Ecology and Biogeography*.

Huber, O., de Stefano, R. D., Aymard, G., & Riina, R. (2006). Flora and vegetation of the Venezuelan Llanos: a review. *Neotropical Savannas and Seasonally dry forests*, 95-120.

IPCC, 2013: Climate change 2013: The physical science basis. In T. F. Stocker, D. Qin, G.K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, & P. M. Midgley (Eds.), *Contribution of working group I to the fifth assessment report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY:IPCC

IPCC, 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press. In Press

Jaramillo, C. (2019). 140 million years of tropical biome evolution. *The Geology of Colombia (ed. Gomez, J.). Colombian Geological Survey, Bogota, Colombia*.

Jarvis, A., Reuter, H. I., Nelson, A., & Guevara, E. (2008). Hole-filled SRTM for the globe Version 4. available from the CGIAR-CSI SRTM 90m Database (http://srtm. csi. cgiar. org), 15, 25-54.

Killick, R., & Eckley, I. (2014). changepoint: An R package for changepoint analysis. *Journal of statistical software*, 58(3), 1-19.

Knapp, A. K., Beier, C., Briske, D. D., Classen, A. T., Luo, Y., Reichstein, M., ... & Weng, E. (2008). Consequences of more extreme precipitation regimes for terrestrial ecosystems. *Bioscience*, *58*(9), 811-821.

Knapp, A. K., Hoover, D. L., Wilcox, K. R., Avolio, M. L., Koerner, S. E., La Pierre, K. J., ... & Smith, M. D. (2015). Characterizing differences in precipitation regimes of extreme wet and dry years: implications for climate change experiments. *Global change biology*, *21*(7), 2624-2633.

Kulmatiski, A., & Beard, K. H. (2013). Woody plant encroachment facilitated by increased precipitation intensity. *Nature Climate Change*, 3(9), 833-837.

Lehmann, C. E., Archibald, S. A., Hoffmann, W. A., & Bond, W. J. (2011). Deciphering the distribution of the savanna biome. *New Phytologist*, *191*(1), 197-209.

Lehmann, C. E., Anderson, T. M., Sankaran, M., Higgins, S. I., Archibald, S., Hoffmann, W. A., ... & Hutley, L. B. (2014). Savanna vegetation-fire-climate relationships differ among continents. *Science*, *343*(6170), 548-552.

Liu, J., Ma, X., Duan, Z., Jiang, J., Reichstein, M., & Jung, M. (2020). Impact of temporal precipitation variability on ecosystem productivity. *Wiley Interdisciplinary Reviews: Water*, *7*(6), e1481.

Lloyd, J., Goulden, M. L., Ometto, J. P., Patiño, S., Fyllas, N. M., & Quesada, C. A. (2013). Ecophysiology of forest and savanna vegetation. *Amazonia and global change*, (a), 463-484.

Lloyd, J., Domingues, T. F., Schrodt, F., Ishida, F. Y., Feldpausch, T. R., Saiz, G., ... & Marimon, B. S. (2015). Edaphic, structural and physiological contrasts across Amazon Basin forest–savanna ecotones suggest a role for potassium as a key modulator of tropical woody vegetation structure and function. *Biogeosciences*, 12(22), 6529-6571.

Lipsett-Moore, G. J., Wolff, N. H., & Game, E. T. (2018). Emissions mitigation opportunities for savanna countries from early dry season fire management. *Nature communications*, *9*(1), 1-8.

Lizundia-Loiola, J., Pettinari, M. L., & Chuvieco, E. (2020). Temporal anomalies in burned area trends: satellite estimations of the Amazonian 2019 fire crisis. *Remote Sensing*, *12*(1), 151.

Malhi, Y., Aragão, L. E., Galbraith, D., Huntingford, C., Fisher, R., Zelazowski, P., ... & Meir, P. (2009). Exploring the likelihood and mechanism of a climate-change-induced dieback of the Amazon rainforest. *Proceedings of the National Academy of Sciences*, *106*(49), 20610-20615.

Marengo, J. A., Tomasella, J., Alves, L. M., Soares, W. R., & Rodriguez, D. A. (2011). The drought of 2010 in the context of historical droughts in the Amazon region. *Geophysical Research Letters*, 38(12).

Meeussen, C., Govaert, S., Vanneste, T., Calders, K., Bollmann, K., Brunet, J., ... & De Frenne, P. (2020). Structural variation of forest edges across Europe. *Forest Ecology and Management*, 462, 117929.

Moncrieff, G. R., Scheiter, S., Langan, L., Trabucco, A., & Higgins, S. I. (2016). The future distribution of the savannah biome: model-based and biogeographic contingency. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1703), 20150311.

Murphy, B. P., & Bowman, D. M. (2012). What controls the distribution of tropical forest and savanna?. *Ecology letters*, *15*(7), 748-758.

McCullagh, P., & Nelder, J.A. (1983). Generalized Linear Models (2nd ed.). Routledge. https://doi.org/10.1201/9780203753736

Newberry, B. M., Power, C. R., Abreu, R. C., Durigan, G., Rossatto, D. R., & Hoffmann, W. A. (2020). Flammability thresholds or flammability gradients? Determinants of fire across savanna–forest transitions. *New Phytologist*, *228*(3), 910-921.

Paredes-Trejo, F. J., Alves Barbosa, H., Peñaloza-Murillo, M. A., Moreno, M. A., & Farias, A. (2016). Intercomparison of improved satellite rainfall estimation with CHIRPS gridded product and rain gauge data over Venezuela. Atmósfera, 29(4), 323-342.

Pellegrini, A. F. (2016). Nutrient limitation in tropical savannas across multiple scales and mechanisms. *Ecology*, *97*(2), 313-324.

Pellegrini, A. F., Staver, A. C., Hedin, L. O., Charles-Dominique, T., & Tourgee, A. (2016). Aridity, not fire, favors nitrogen-fixing plants across tropical savanna and forest biomes. *Ecology*, *97*(9), 2177-2183.

Oliveras, I., & Malhi, Y. (2016). Many shades of green: the dynamic tropical forest–savannah transition zones. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1703), 20150308.

Otón, G., Lizundia-Loiola, J., Pettinari, M. L., & Chuvieco, E. (2021). Development of a consistent global long-term burned area product (1982–2018) based on AVHRR-LTDR data. *International Journal of Applied Earth Observation and Geoinformation*, *103*, 102473.

Ratnam, J., Bond, W. J., Fensham, R. J., Hoffmann, W. A., Archibald, S., Lehmann, C. E., ... & Sankaran, M. (2011). When is a 'forest' a savanna, and why does it matter?. *Global Ecology and Biogeography*, *20*(5), 653-660.

Ribeiro, J. W., Pilon, N. A., Rossatto, D. R., Durigan, G., & Kolb, R. M. (2021). The distinct roles of water table depth and soil properties in controlling alternative woodland-grassland states in the Cerrado. *Oecologia*, *195*(3), 641-653.

Rodrigues, M. A., Garcia, S. R., Kayano, M. T., Calheiros, A. J., & Andreoli, R. V. Onset and demise dates of the rainy season in the South American Monsoon region: a cluster analysis result. *International Journal of Climatology*.

Rodríguez-Iturbe, I., & Porporato, A. (2007). *Ecohydrology of water-controlled ecosystems: soil moisture and plant dynamics*. Cambridge University Press.

Romero-Ruiz, M., Etter, A., Sarmiento, A., & Tansey, K. (2010). Spatial and temporal variability of fires in relation to ecosystems, land tenure and rainfall in savannas of northern South America. *Global Change Biology*, *16*(7), 2013-2023.

Romero-Ruiz, M. H., Flantua, S. G. A., Tansey, K., & Berrio, J. C. (2012). Landscape transformations in savannas of northern South America: Land use/cover changes since 1987 in the Llanos Orientales of Colombia. *Applied Geography*, *32*(2), 766-776.

Rosan, T. M., Aragão, L. E., Oliveras, I., Phillips, O. L., Malhi, Y., Gloor, E., & Wagner, F. H. (2019). Extensive 21st-century woody encroachment in South America's savanna. *Geophysical Research Letters*, *46*(12), 6594-6603.

Sánchez Ojeda, F., & Armenteras Pascual, D. (2017). Changes in soil organic carbon after burning in a forest-savanna edge. *Acta Agronómica*, *66*(4), 519-524.

Sankaran, M., Hanan, N. P., Scholes, R. J., Ratnam, J., Augustine, D. J., Cade, B. S., ... & Zambatis, N. (2005). Determinants of woody cover in African savannas. *Nature*, *438*(7069), 846-849.

Sato, H., Kelley, D. I., Mayor, S. J., Martin Calvo, M., Cowling, S. A., & Prentice, I. C. (2021). Dry corridors opened by fire and low CO2 in Amazonian rainforest during the Last Glacial Maximum. *Nature Geoscience*, *14*(8), 578-585.

Schwartz, N. B., Lintner, B. R., Feng, X., & Powers, J. S. (2020). Beyond MAP: A guide to dimensions of rainfall variability for tropical ecology. *Biotropica*, *52*(6), 1319-1332.

Singh, C., van der Ent, R., Wang-Erlandsson, L., & Fetzer, I. (2022). Hydroclimatic adaptation critical to the resilience of tropical forests. *Global Change Biology*.

Staal, A., & Flores, B. M. (2015). Sharp ecotones spark sharp ideas: comment on" Structural, physiognomic and above-ground biomass variation in savanna–forest transition zones on three continents–how different are co-occurring savanna and forest formations?" by Veenendaal et al.(2015). *Biogeosciences*, *12*(18), 5563-5566.

Staal, A., van Nes, E. H., Hantson, S., Holmgren, M., Dekker, S. C., Pueyo, S., ... & Scheffer, M. (2018). Resilience of tropical tree cover: The roles of climate, fire, and herbivory. *Global Change Biology*, *24*(11), 5096-5109.

Staal, A., Fetzer, I., Wang-Erlandsson, L., Bosmans, J. H., Dekker, S. C., van Nes, E. H., ... & Tuinenburg, O. A. (2020). Hysteresis of tropical forests in the 21st century. *Nature communications*, *11*(1), 1-8.

Stark, S. C., Breshears, D. D., Aragón, S., Villegas, J. C., Law, D. J., Smith, M. N., ... & Saleska, S. R. (2020). Reframing tropical savannization: linking changes in canopy structure to energy balance alterations that impact climate. *Ecosphere*, *11*(9), e03231.

Staver, A. C., Archibald, S., & Levin, S. A. (2011). The global extent and determinants of savanna and forest as alternative biome states. *Science*, *334*(6053), 230-232.

Staver, A. C., Botha, J., & Hedin, L. (2017). Soils and fire jointly determine vegetation structure in an African savanna. *New Phytologist*, *216*(4), 1151-1160.

Stevens, N., Lehmann, C. E., Murphy, B. P., & Durigan, G. (2017). Savanna woody encroachment is widespread across three continents. *Global change biology*, *23*(1), 235-244.

Tang, H., Armston, J., & Dubayah, R. (2019). Algorithm Theoretical Basis Document (ATBD) for GEDI L2B Footprint Canopy Cover and Vertical Profile Metrics. *Goddard Space Flight Center: Greenbelt, MD, USA*.

Uribe, M., & Dukes, J. S. (2021). Land cover change alters seasonal photosynthetic activity and transpiration of Amazon forest and Cerrado. *Environmental Research Letters*, *16*(5), 054013.

Veenendaal, E. M., Torello-Raventos, M., Miranda, H. S., Sato, N. M., Oliveras, I., van Langevelde, F., ... & Lloyd, J. (2018). On the relationship between fire regime and vegetation structure in the tropics. *New Phytologist*, *218*(1), 153-166.

Veenendaal, E. M., Torello-Raventos, M., Feldpausch, T. R., Domingues, T. F., Gerard, F., Schrodt, F., ... & Lloyd, J. (2015). Structural, physiognomic and above-ground biomass variation in savanna–forest transition zones on three continents–how different are co-occurring savanna and forest formations?. *Biogeosciences*, *12*(10), 2927-2951.

Viglizzo, E. F., Nosetto, M. D., Jobbágy, E. G., Ricard, M. F., & Frank, F. C. (2015). The ecohydrology of ecosystem transitions: a meta-analysis. *Ecohydrology*, *8*(5), 911-921.

West, L. T., Beinroth, F. H., Sumner, M. E., & Kang, B. T. (1997). Ultisols: Characteristics and impacts on society. *Advances in Agronomy*, *63*, 179-236.

Xu, X., Medvigy, D., Trugman, A. T., Guan, K., Good, S. P., & Rodriguez-Iturbe, I. (2018). Tree cover shows strong sensitivity to precipitation variability across the global tropics. *Global Ecology and Biogeography*, *27*(4), 450-460.

Yang, Y., Donohue, R. J., & McVicar, T. R. (2016). Global estimation of effective plant rooting depth: Implications for hydrological modeling. *Water Resources Research*, 52(10), 8260-8276.

Zeng, Z., Chen, A., Piao, S., Rabin, S., & Shen, Z. (2014). Environmental determinants of tropical forest and savanna distribution: A quantitative model evaluation and its implication. *Journal of Geophysical Research: Biogeosciences*, *119*(7), 1432-1445.

Zeng, Z., Piao, S., Chen, A., Lin, X., Nan, H., Li, J., & Ciais, P. (2013). Committed changes in tropical tree cover under the projected 21 st century climate change. *Scientific reports*, *3*(1), 1-6.

Zheng, B., Ciais, P., Chevallier, F., Chuvieco, E., Chen, Y., & Yang, H. (2021). Increasing forest fire emissions despite the decline in global burned area. *Science advances*, *7*(39), eabh2646.

Zomer, R. J., Trabucco, A., Bossio, D. A., & Verchot, L. V. (2008). Climate change mitigation: A spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. *Agriculture, ecosystems & environment*, *126*(1-2), 67-80.

Supporting information for

"Fire frequency alone does not explain forest - savanna transition: the role of dry season precipitation variability in Northern South America"

Santiago Valencia[1*], Juan Camilo Villegas[2], Juan F. Salazar[1]

[1] Grupo GIGA, Escuela Ambiental, Facultad de Ingeniería, Universidad de Antioquía, Medellín-Colombia
[2] Grupo de investigación en Ecología Aplicada, Escuela Ambiental, Facultad de Ingeniería, Universidad de Antioquía, Medellín-Colombia

***Correspondence:**
Santiago Valencia, Escuela Ambiental, Facultad de Ingeniería, Universidad de Antioquia, Medellín, Colombia
Email: santiago.valencia8@udea.edu.co

Contents of this file:
Supporting Information S1 to S4
Tables S1 to S3
Figures S1 to S12

Supporting information S1

The Global Ecosystem Dynamics Investigation (GEDI) is a new spaceborne LiDAR instrument aboard the International Space Station (ISS) and collecting vegetation structure data at an average footprint resolution of 25 m since April 2019 for a nominal two-years mission (Dubayah et al., 2020). GEDI consists of 3 lasers that reach 25 m diameter on the ground (footprints), resulting in 8 tracks of data, spaced at 60 m along track and 600 m across track within a ~4.2 km swath (Supporting Information Figure S1a). A detailed description of the GEDI products can be found in Tang et al., (2019) and Dubayah et al., (2020). Gridded products based solely on GEDI sample data will have a spatial resolution of at least 1 km. However, those products are still in development (Luthcke et al., 2021). GEDI L2B product data was collected between April 19, 2019 and september 02, 2020. For each footprint, we extracted the percentage canopy cover and Plant Area Volume Density (PAVD) profile using R package rGEDI (Silva et al., 2020). In GEDI, canopy cover represents the percent of the ground covered by the vertical projection of canopy material (i.e., leaves, branches and stems only) considering the gaps between and within the canopy (Tang et al., 2019). For each footprint PAVD profile, we calculated the maximum PAVD ($PAVD_{max}$) as a proxy of vertical vegetation structure (Calders et al., 2014; Decuyper et al., 2017; Meeussen et al., 2020). We excluded low-quality footprints (i.e., poor geolocation performance, waveforms of poor signal quality, and waveforms affected by cloud and other land surface conditions) according to the quality flags of GEDI L2B (Tang et al., 2019).

To obtain gridded GEDI data for the study area, we averaged the footprints within each 0.01° x 0.01° pixel as suggested by Rishmawi et al., (2021). Given the number of GEDI observations within each pixel ranged between 1 and 60 (see Supporting Information Figure S1b), we excluded pixels calculated with less than 20 footprints to minimize the effect of under-sampling

observations within pixels while maintaining a large number of available pixels (Supporting Information Figure S1c). Additionally, the results do not show substantial changes in mean ($\underline{x}$) or variance, for example, of the maximum PAVD ($PAVD_{max}$, m^2/m^3) when considering a threshold of 10 ($\underline{x}_{PAVD_{max}}$ = 0.362 and $s^2_{PAVD_{max}}$ = 0.057) or 30 ($\underline{x}_{PAVD_{max}}$ = 0.384 and $s^2_{PAVD_{max}}$ = 0.059) instead of 20 ($\underline{x}_{PAVD_{max}}$ = 0.371 and $s^2_{PAVD_{max}}$ = 0.058) GEDI observations by 0.01° pixel. The percentage tree cover from MODIS has been used in a wide variety of ecological research including vegetation modelling, estimating restoration potential, and identifying forest-savanna bimodality (Adzhar et al., 2021). Percent tree cover describes the percent of each pixel covered by canopy cover of trees above 5 m in height (DiMiceli et al., 2015). In this case, we used the percent tree cover available for the year 2019 based on MODIS inputs dating from 6 March 2019 to 6 March 2020 at a 250 m resolution.

Pixels were defined as forest or savanna based on each vegetation structure descriptor: canopy cover, tree cover, and $PAVD_{max}$. We used a threshold of 60% of tree cover to define savanna (< 60%) and forest (≥60%) as suggested by Hirota et al., (2011). Given the difference between tree cover (DiMiceli et al., 2015) and canopy cover (Tang et al., 2019) as well as between MODIS and GEDI (e.g., MODIS tree cover saturates at ~80%, while the GEDI canopy cover data maxes out at 100%), we used a threshold of 40% of canopy cover to classificate pixels as savanna (< 40%) or forest (≥60%). Finally, for $PAVD_{max}$, we defined forest and savanna using a threshold of 0.10 m^2/m^3.

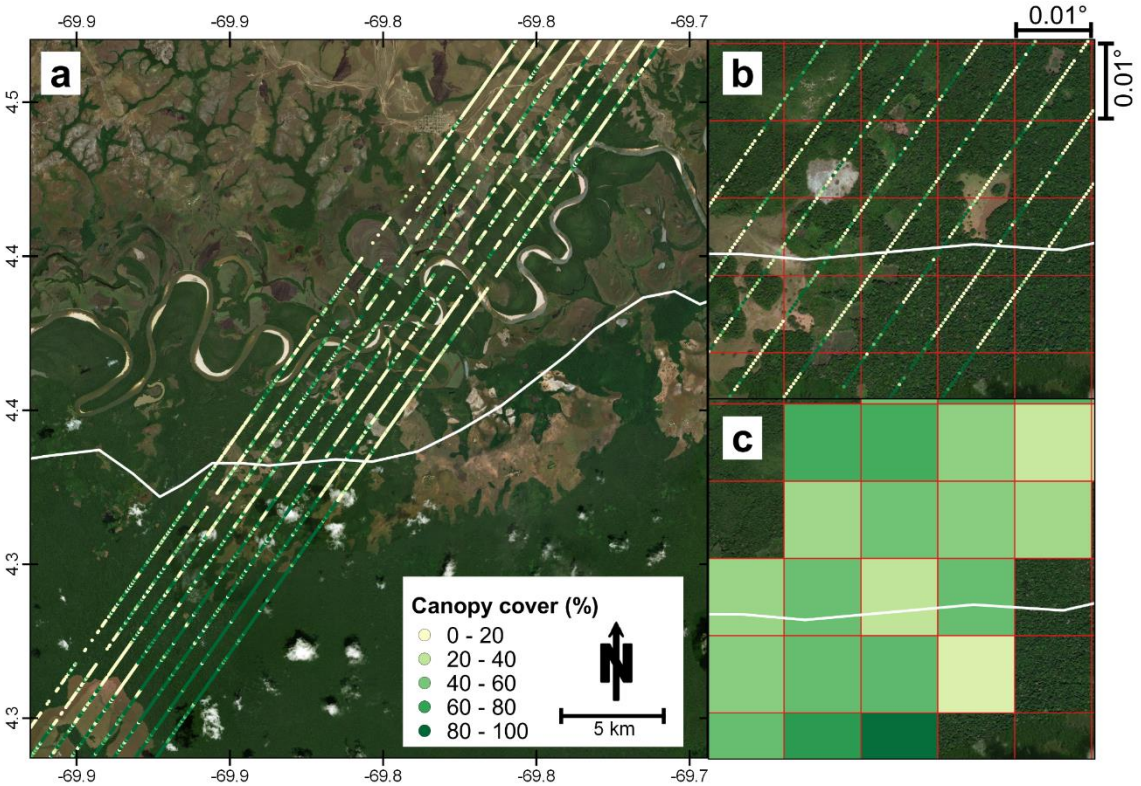

Figure S1. a) Sample GEDI beam ground transects along forest-savanna transition and b) the number of GEDI footprints available in a pixel at 0.01° x 0.01° resolution (red lines). c) Result of average GEDI footprints in each pixel, excluding pixels with less than 20 footprints. The white line represents the limits of the Llanos ecoregion. The Google Earth Image as the background.

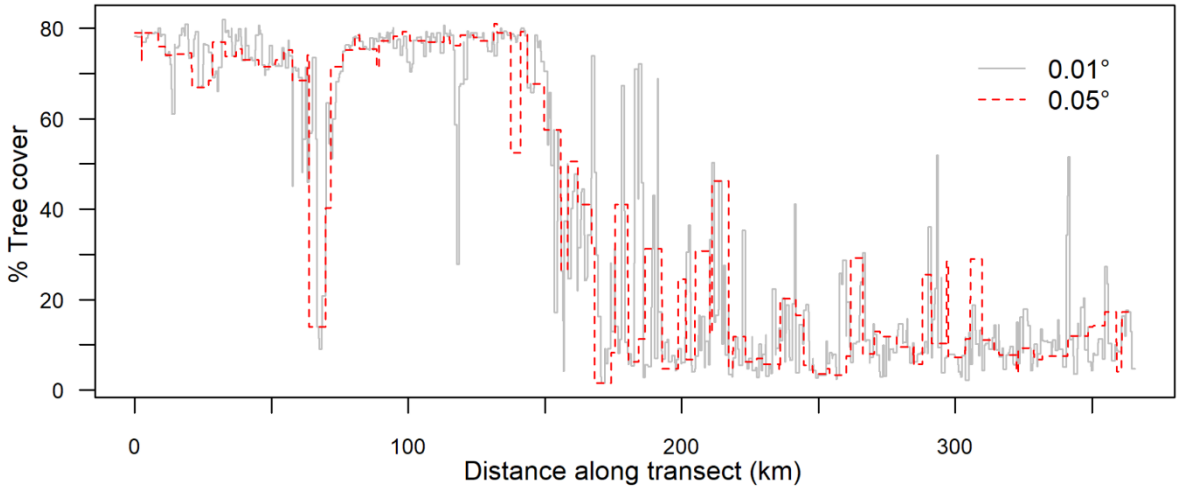

Figure S2. Example of differences in percentage tree cover from MODIS along a forest-savanna transect at 0.01° (solid gray line) and 0.05° (dotted red line) resolution.

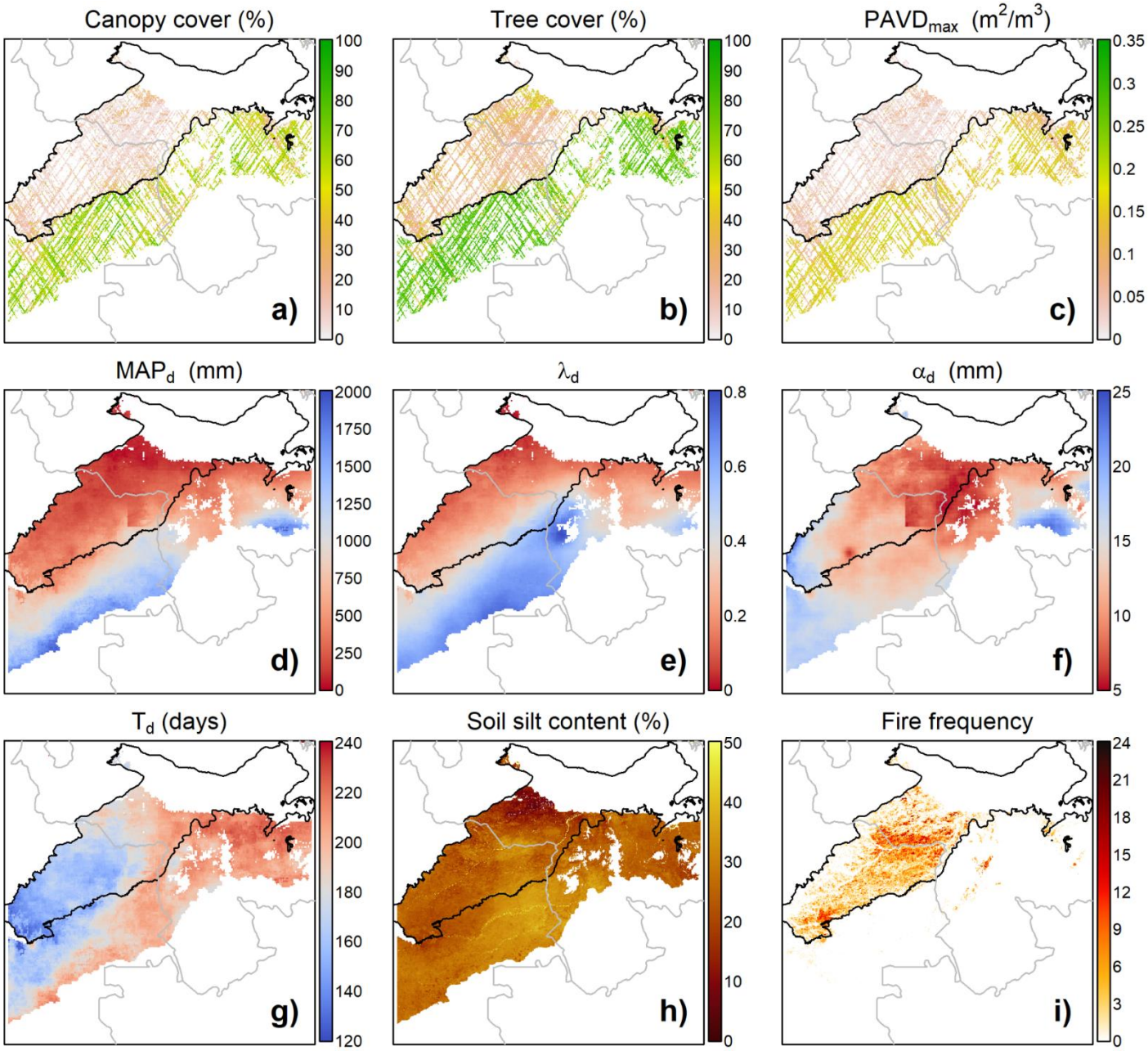

Figure S3. Spatial maps of predictor variables. White area was excluded in our analysis.

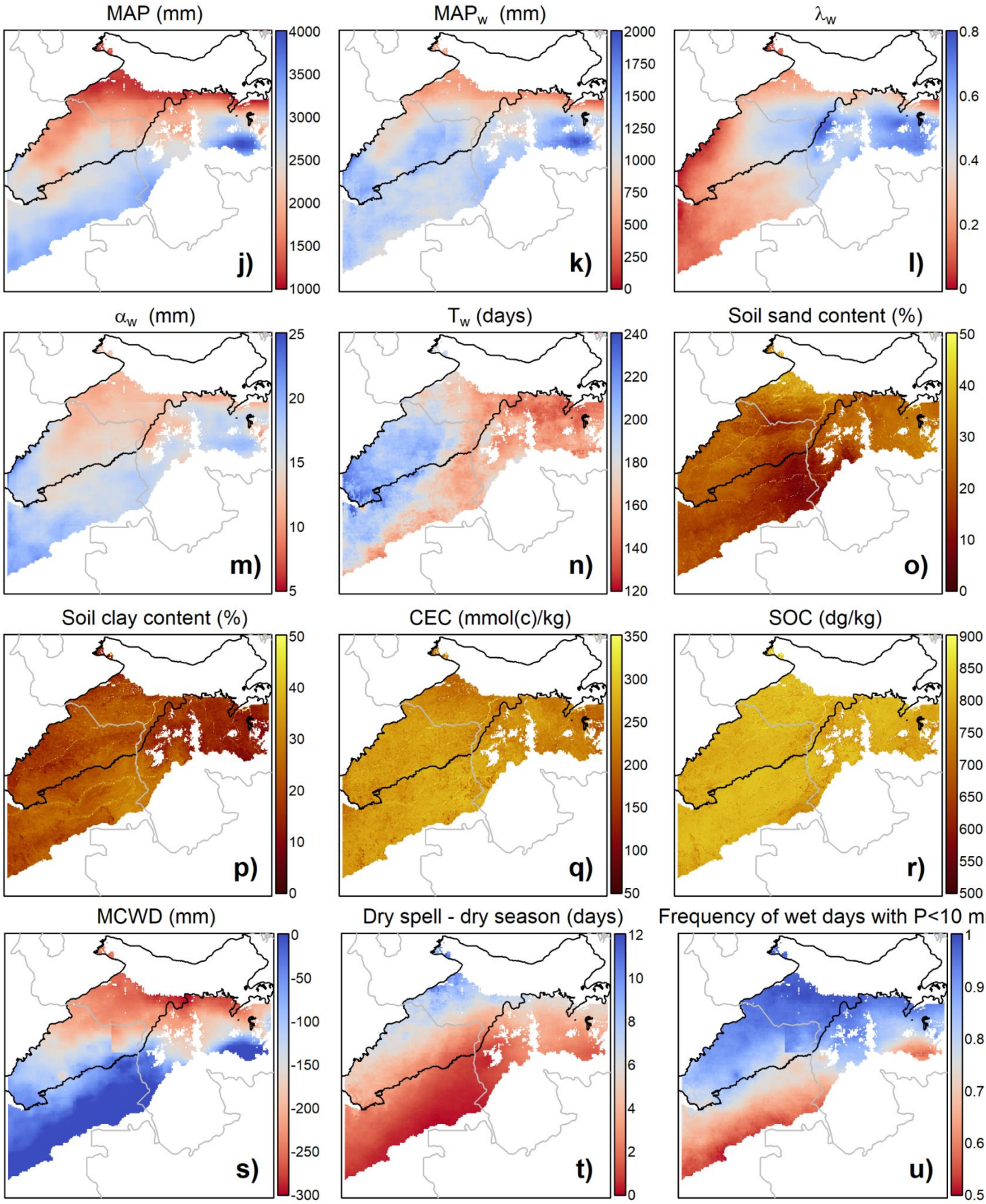

Figure S3. Spatial maps of predictor variables. (Continued)

Supporting information S2

To analyze all the transects, pixels were numbered sequentially, with pixel 1 in the forest and 580 in the savanna. Numbered pixels were aligned to calculate the median as well as 10th, 25th, 75th, and 90th percentiles of each pixel (Supporting Information Figure S4b). Given the length of transects varied depending on the limits of the study area (Figure 1) and GEDI available pixels (Supporting Information Figure S3a), we excluded median and percentiles values calculated with less than 375 pixels (i.e., 50% of GEDI available pixels along all transects analyzed).

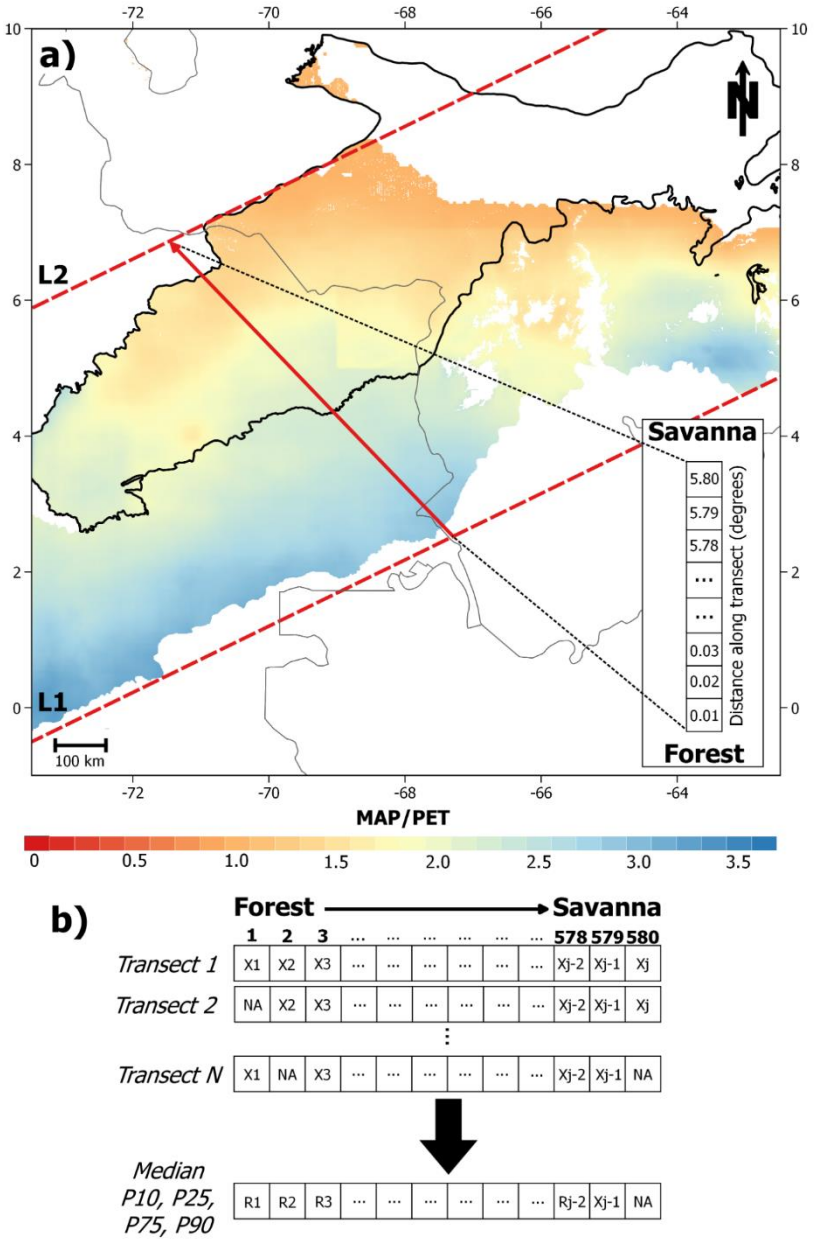

Figure S4. a) Diagram of transect analyzed along forest-savanna transition. Forest and savanna were defined based on limits of ecoregions. Transects were 0.01° (~ 1 km) wide, of variable length, and were sampled in 0.01° increments (0.01° x 0.01° blocks). b) Showing how transects were lined up first on their forest to savanna edge.

Table S1. Median Spearman rank correlation analysis for all predictor variables analyzed. "cover" = % canopy cover; "tree" = % tree cover; "PAVD" = maximum Plant Area Volume Density; "sand" = topsoil sand content; "clay" = topsoil clay content; "silt" = topsoil silt content; "SOC" = topsoil organic carbon; "CEC" = topsoil cation exchange capacity; "MAP" = mean annual precipitation; "$\boldsymbol{\lambda}$" = frequency of rainy days (p > 0); "$\boldsymbol{\alpha}$" = intensity of precipitation; "T" = length of wet or dry seasons; "fire" = fire frequency. "Ds_d" = the length of dry spells in the dry season; "$\boldsymbol{\lambda}_{10d}$" = frequency of wet days with precipitation < 10 mm/day within the dry season. $\boldsymbol{d}$ and $\boldsymbol{w}$ represent the dry and the wet season, respectively. Values are calculated using the selected 1000 random samples of 5% (N=10074) of data. The standard deviation of correlation values is lower than 0.05 for all variables.

	cover	tree	PAVD	sand	clay	silt	SOC	CEC	MAP	$\boldsymbol{\lambda}_w$	$\boldsymbol{\lambda}_d$	$\boldsymbol{\alpha}_w$	$\boldsymbol{\alpha}_d$	MAP_w	MAP_d	T_w	T_d	fire	Ds_d	$\boldsymbol{\lambda}_{10d}$
cover	1																			
tree	0.82	1																		
PAVD	0.99	0.82	1																	
sand	0.21	0.20	0.22	1																
clay	0.01	-0.01	0.01	-0.70	1															
silt	**-0.30**	**-0.29**	**-0.30**	-0.83	0.38	1														
SOC	0.01	0.01	0.01	-0.14	0.26	0.03	1													
CEC	-0.13	-0.11	-0.14	-0.59	0.49	0.51	0.33	1												
MAP	0.56	0.56	0.58	0.49	-0.24	-0.49	0.16	-0.33	1											
$\boldsymbol{\lambda}_w$	-0.01	-0.04	-0.04	0.02	0.09	-0.16	0.12	0.10	0.02	1										
$\boldsymbol{\lambda}_d$	**0.57**	**0.54**	**0.58**	0.64	-0.39	-0.60	0.08	-0.43	0.86	0.07	1									
$\boldsymbol{\alpha}_w$	0.53	0.52	0.54	0.30	-0.06	-0.34	0.19	-0.22	0.86	-0.24	0.65	1								
$\boldsymbol{\alpha}_d$	**0.42**	**0.44**	**0.43**	0.09	0.05	-0.11	0.18	-0.13	0.71	-0.44	0.38	0.81	1							
MAP_w	0.23	0.18	0.25	0.30	-0.09	-0.30	0.15	-0.26	0.74	0.13	0.47	0.64	0.58	1						
MAP_d	0.67	0.66	0.68	0.48	-0.21	-0.51	0.16	-0.30	0.94	0.01	0.92	0.82	0.64	0.56	1					
T_w	-0.26	-0.27	-0.26	0.15	-0.22	0.06	-0.14	-0.32	0.01	-0.52	-0.08	0.02	0.27	0.36	-0.18	1				
T_d	**0.26**	**0.27**	**0.26**	-0.15	0.22	-0.06	0.14	0.32	-0.01	0.52	0.08	-0.02	-0.27	-0.36	0.18	-1*	1			
fire	**-0.64**	**-0.64**	**-0.64**	-0.01	-0.07	0.15	-0.01	0.05	-0.38	0.11	-0.35	-0.42	-0.37	-0.08	-0.49	0.25	-0.25	1		
Ds_d	-0.61	-0.58	-0.62	-0.63	0.38	0.60	-0.09	0.41	-0.88	0.03	-0.99	-0.70	-0.45	-0.48	-0.95	0.10	-0.10	0.40	1	
$\boldsymbol{\lambda}_{10d}$	-0.62	-0.62	-0.63	-0.50	0.26	0.48	-0.12	0.37	-0.95	0.16	-0.88	-0.85	-0.76	-0.61	-0.96	-0.07	0.07	0.44	0.91	1

The bold values indicate the correlations between response variables and predictor variables selected for statistical model analysis. The shaded values indicate low (<|0.2|) and non-statistically significant (p > 0.01) correlation values. *T_w and T_d show a perfect negative correlation because the length of the wet or dry season is linearly dependent on each other (i.e., $T_w = 365 - T_d$).

Table S2. Median Pearson correlation coefficient for all predictor variables analyzed. "cover" = % canopy cover; "tree" = % tree cover; "PAVD" = maximum Plant Area Volume Density; "sand" = topsoil sand content; "clay" = topsoil clay content; "silt" = topsoil silt content; "SOC" = topsoil organic carbon; "CEC" = topsoil cation exchange capacity; "MAP" = mean annual precipitation; "$\boldsymbol{\lambda}$" = frequency of rainy days (p > 0); "$\boldsymbol{\alpha}$" = intensity of precipitation; "T" = length of wet or dry seasons; "fire" = fire frequency. "Ds_d" = the length of dry spells in the dry season; "$\boldsymbol{\lambda}_{10d}$" = frequency of wet days with precipitation < 10 mm/day within the dry season. ***d*** and ***w*** represent the dry and the wet season, respectively. Values are calculated using the selected 1000 random samples of 5% (N=10074) of data. The standard deviation of correlation values is lower than 0.05 for all variables.

	cover	tree	PAVD	sand	clay	silt	SOC	CEC	MAP	$\boldsymbol{\lambda}_w$	$\boldsymbol{\lambda}_d$	$\boldsymbol{\alpha}_w$	$\boldsymbol{\alpha}_d$	MAP_w	MAP_d	T_w	T_d	fire	Ds_d	$\boldsymbol{\lambda}_{10d}$
cover	1																			
tree	0.87	1																		
PAVD	0.99	0.86	1																	
sand	0.24	0.22	0.25	1																
clay	0.03	0.04	0.02	-0.56	1															
silt	**-0.31**	**-0.28**	**-0.31**	-0.77	0.38	1														
SOC	0.02	0.03	0.02	-0.02	0.25	0.01	1													
CEC	-0.10	-0.06	-0.11	-0.51	0.53	0.54	0.35	1												
MAP	0.53	0.54	0.55	0.49	-0.18	-0.52	0.19	-0.32	1											
$\boldsymbol{\lambda}_w$	0.03	0.03	0.03	0.04	0.15	-0.16	0.16	0.11	0.10	1										
$\boldsymbol{\lambda}_d$	**0.57**	**0.58**	**0.58**	0.64	-0.38	-0.60	0.10	-0.40	0.82	0.04	1									
$\boldsymbol{\alpha}_w$	0.51	0.52	0.52	0.30	0.03	-0.36	0.16	-0.20	0.84	-0.25	0.62	1								
$\boldsymbol{\alpha}_d$	**0.39**	**0.42**	**0.40**	0.07	0.09	-0.15	0.15	-0.08	0.69	-0.37	0.34	0.80	1							
MAP_w	0.25	0.24	0.27	0.37	-0.08	-0.41	0.15	-0.30	0.84	0.21	0.49	0.66	0.58	1						
MAP_d	0.64	0.67	0.65	0.47	-0.21	-0.48	0.17	-0.25	0.89	-0.01	0.90	0.78	0.62	0.50	1					
T_w	-0.27	-0.31	-0.26	0.09	-0.22	0.03	-0.14	-0.30	0.06	-0.53	-0.10	0.05	0.28	0.35	-0.20	1				
T_d	**0.27**	**0.31**	**0.26**	-0.09	0.22	0.03	0.14	0.30	-0.06	0.53	0.10	-0.05	-0.28	-0.35	0.20	-1*	1			
fire	**-0.51**	**-0.52**	**-0.51**	0.02	-0.08	0.06	-0.03	0.03	-0.26	0.12	-0.28	-0.32	-0.33	-0.04	-0.37	0.17	-0.17	1		
Ds_d	-0.60	-0.58	-0.61	-0.63	0.30	0.63	-0.13	0.39	-0.87	-0.08	-0.96	-0.70	-0.42	-0.61	-0.88	0.07	-0.07	0.29	1	
$\boldsymbol{\lambda}_{10d}$	-0.61	-0.64	-0.62	-0.47	0.25	0.46	-0.12	0.31	-0.90	0.18	-0.87	-0.82	-0.72	-0.57	-0.96	-0.03	0.03	0.36	0.86	1

The bold values indicate the correlations between response variables and predictor variables selected for statistical model analysis. The shaded values indicate low (<|0.2|) and non-statistically significant (p > 0.01) correlation values. *T_w and T_d show a perfect negative correlation because the length of the wet or dry season is linearly dependent on each other (i.e., $T_w = 365 - T_d$).

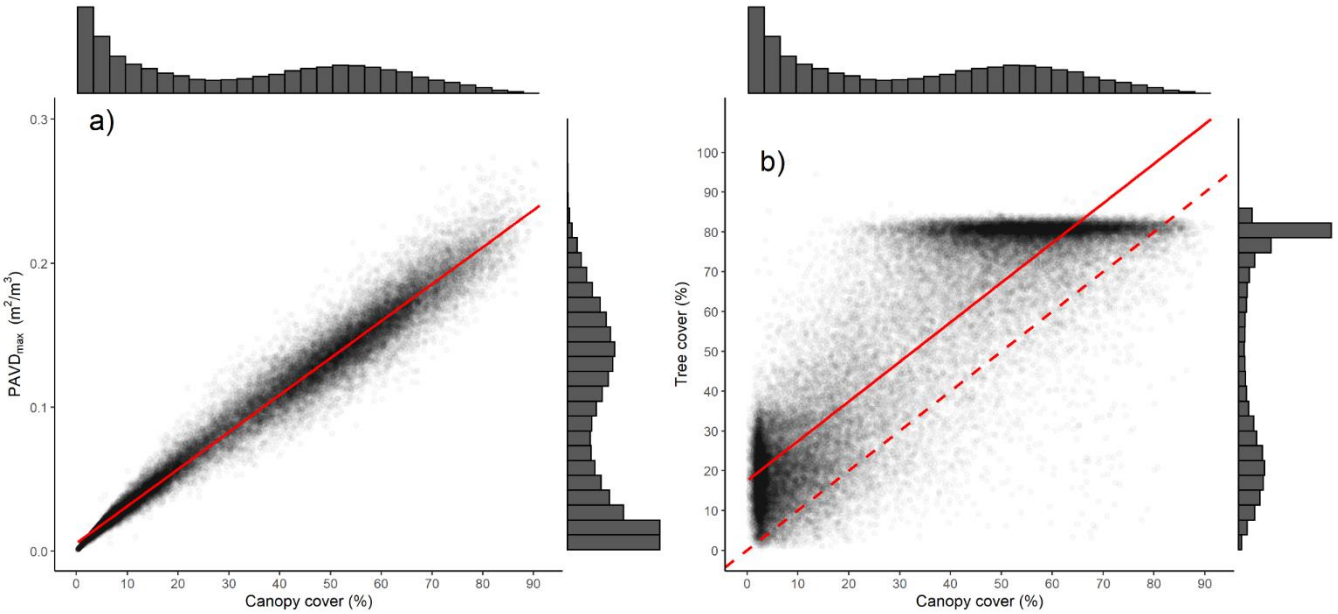

Figure S5. Comparison of a) maximum PAVD ($PAVD_{max}$= 0.559 + 0.257*canopy_cover; the regressionR^2 = 0.97), and b) tree cover (tree cover = 17.463 + 0.996*canopy cover; the regression R^2 = 0.76, Root mean squared difference = 22.10) with canopy cover. The solid and dashed red line represents the linear regression line fit and the 1:1 line, respectively. Black points show a subset (20%, N = ~40000) of observations for purposes of visual clarity.

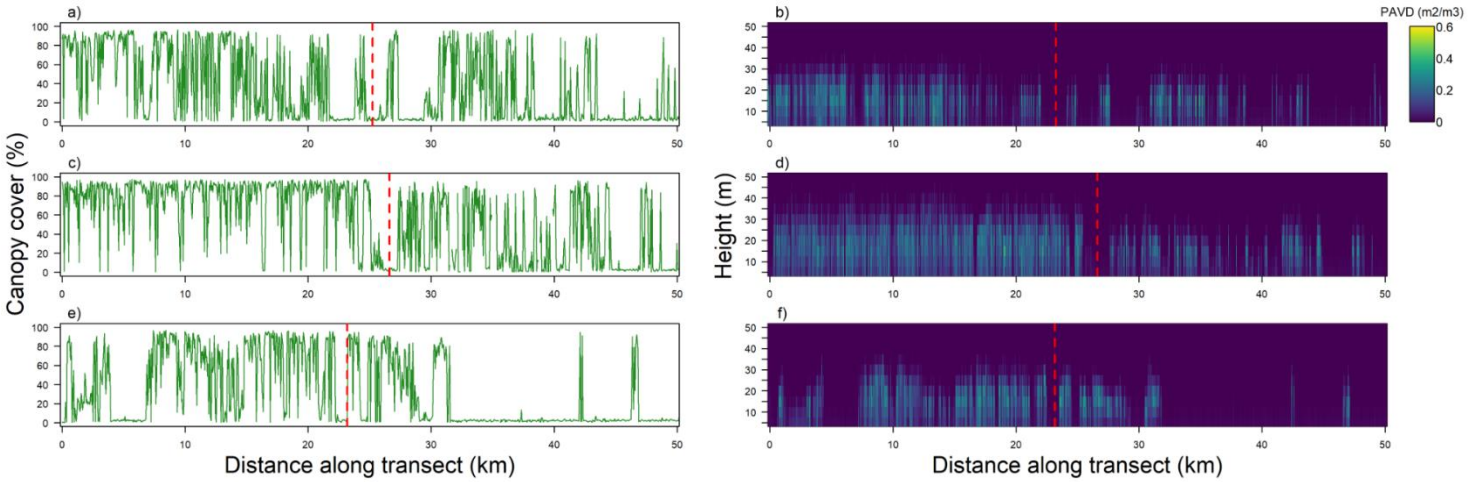

Figure S6. Example of a) % canopy cover and b) PAVD profile variations across three forest-savanna transects. The vertical red lines represent the transition between forest and savanna according to limits of Llanos ecoregion for each transect.

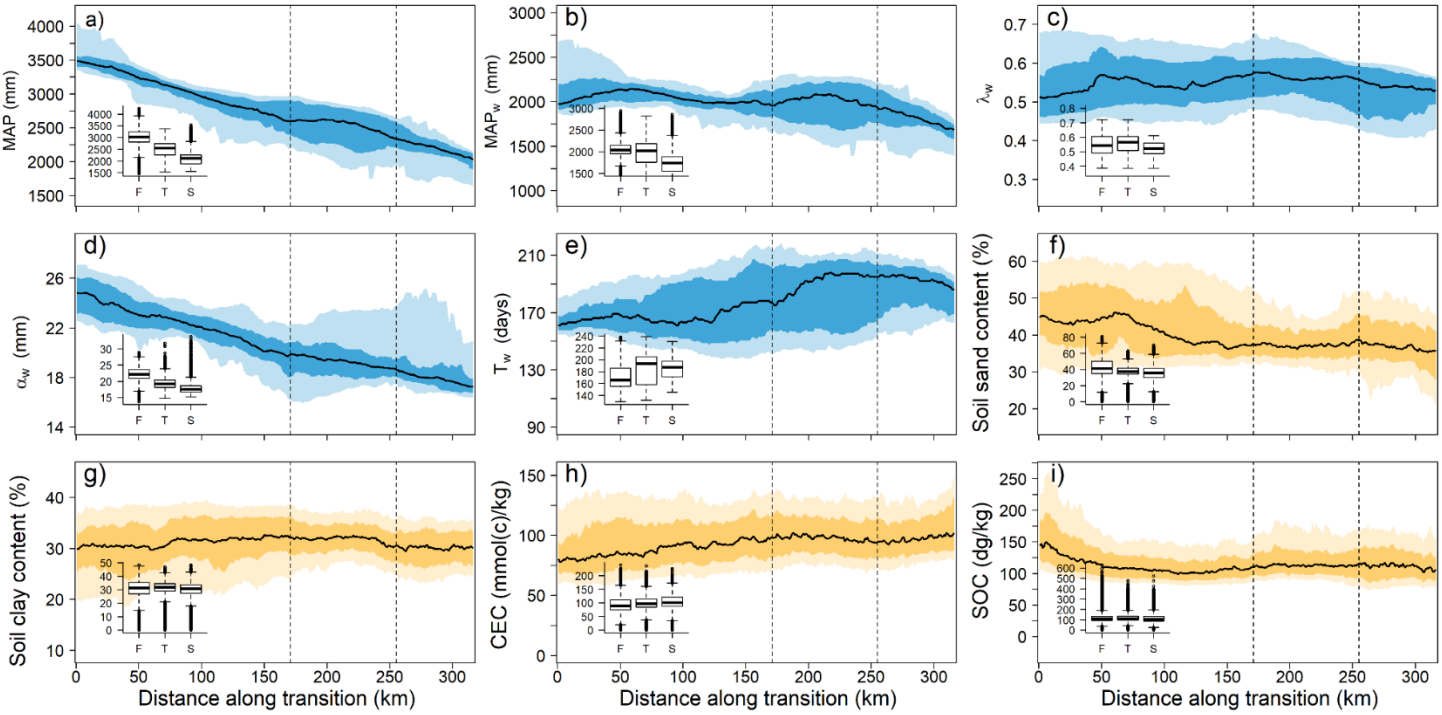

Figure S7. Transects across forest-savanna transition. a) mean annual precipitation (MAP, mm), b) mean total wet-season precipitation (MAP_w), c) frequency of wet days (precipitation > 0) within the wet season ($\boldsymbol{\lambda_w}$), d) intensity of wet days within the wet season ($\boldsymbol{\alpha_w}$), e) length of the wet season (T_w), f) topsoil (0-30 cm) sand content, g) topsoil (0-30 cm) clay content, f) topsoil (0-30 cm) cation exchange capacity (CEC), and i) topsoil (0-30 cm) soil organic carbon (SOC). Black In all panels, the lightest shade is 10^{th} - 90^{th} percentile, the darker shape is 25^{th} - 75^{th} percentile and the black line represents the median (50^{th} percentile) for all transects analyzed (n=1835). Vertical dashed lines show the schematic limits of forest (F), transition (T), and savanna (S) regions quantified using a changepoint analysis for canopy cover as a function of distance along the transition. The inset boxplots show the values of each variable in the F, T, and S regions.

Supporting information S4

Following D'onofrio et al. (2018, 2019), we assumed binomial error distributions with a logit in GLMs, because vegetation descriptors are between 0 and 1 (i.e., canopy cover and tree cover were considered as fractions). In the case of $PAVD_{max}$, which potentially have values greater than 1, was remapped to the range [0:1] before using a logit function $(PAVD_{max})' = \frac{PAVD_{max} - min(PAVD_{max})}{max(PAVD_{max}) - min(PAVD_{max})}$. Due to PV components analyzed are not independent of each other and potential collinearity issues among the other predictor variables, we removed the selected predictor variable(s) with correlation (r or r_s) higher than $\pm$0.70 (Dormann, 2013). In this sense, MAP_d and λ_d were the only selected predictor variables highly correlated (r and r_s > 0.90; Supporting Information Tables S1 and S2). For this reason, we left out MAP_d to maintain variable independence in GLMs. Although MAP_d was higher correlated with vegetation descriptors than λ_d (Supporting Information Tables S1 and S2), we selected λ_d instead MAP_d because precipitation frequency provided more information about precipitation variability during the dry season than total precipitation. Additionally, considering λ_d, α_d, and T_d, we indirectly considered the effect of MAP_d. We also did not include the length of dry spells

(Ds_d) and frequency of wet days with precipitation (λ_{10d}) due to these variables were highly correlated with λ_d and α_d (> |-0.72|) as well as each other (> 0.86; Supporting Information Tables S1 and S2).

Previous studies have also used specific samples of the data for statistical analysis when input data had different spatial resolutions. In this case, we used stratified--by deciles of canopy cover, tree cover, or $PAVD_{max}$-- random samples of 5% instead of 0.1% (Hirota et al., 2011) or 1% (Staal et al., 2016; Bernardino et al., 2019) of all pixels. For each of 1000 samples by response variable and corresponding GLMs, we extracted the estimate (β) and standard error (σ) of each predictor variable and calculated a normal distribution with mean β and standard deviation σ as suggested by Compagnoni et al. (2021). We used the resulting 1000 normal distributions of each predictor variable to estimate its median effect (i.e., direction and strength) and the 95% confidence interval (i.e., variability and significance) on each response variable. Finally, we also evaluated the deviance explained by each predictor variable, including MAP to compare its explanatory power with PV components.

Table S3. Generalized linear models (GLMs) estimates for standardized parameters and 95% confidence interval. The values of deviance explained represent the lowest and highest values from corresponding 1000 GLMs by each vegetation descriptor variable. Fire frequency (fire), mean daily precipitation frequency in dry season ($\boldsymbol{\lambda_d}$), mean daily precipitation intensity in dry season ($\boldsymbol{\alpha_d}$), mean length of dry season (T_d), soil silt content (silt). * Predictor variables that have non-statistically significant effects.

Response variable	Predictors	Estimate	95 % Confidence interval	% Deviance explained (min and max)	% Deviance explained by each variable (min and max)
Canopy cover	fire	-0.79	[-1.00, -0.59]	52.21 - 55.40	32.61 - 35.83
	λ_d	0.40	[0.28, 0.53]		27.53 - 31.46
	α_d	0.25	[0.14, 0.35]		12.24 - 15.58
	T_d	0.26	[0.16, 0.36]		5.56 - 8.55
	silt	-0.04*	[-0.16, 0.08]		7.33 - 10.28
Tree cover	fire	-0.42	[-0.54, -0.30]	54.15 - 57.97	26.21 - 30.57
	λ_d	0.48	[0.36, 0.60]		29.74 - 33.74
	α_d	0.37	[0.26, 0.48]		14.53 - 18. 99
	T_d	0.37	[0.28, 0.47]		6.79 - 11.03
	silt	0.04*	[-0.07, 0.15]		5.46 - 9.52
$PAVD_{max}$	fire	-0.68	[-0.89, -0.47]	52.75 - 56.39	32.14 - 37.17
	λ_d	0.36	[0.23, 0.48]		27.87 - 33.26
	α_d	0.22	[0.11, 0.33]		11.98 - 17.27
	T_d	0.21	[0.10, 0.31]		3.34 - 8.83
	silt	-0.03*	[-0.16, 0.09]		6.81 - 11.62

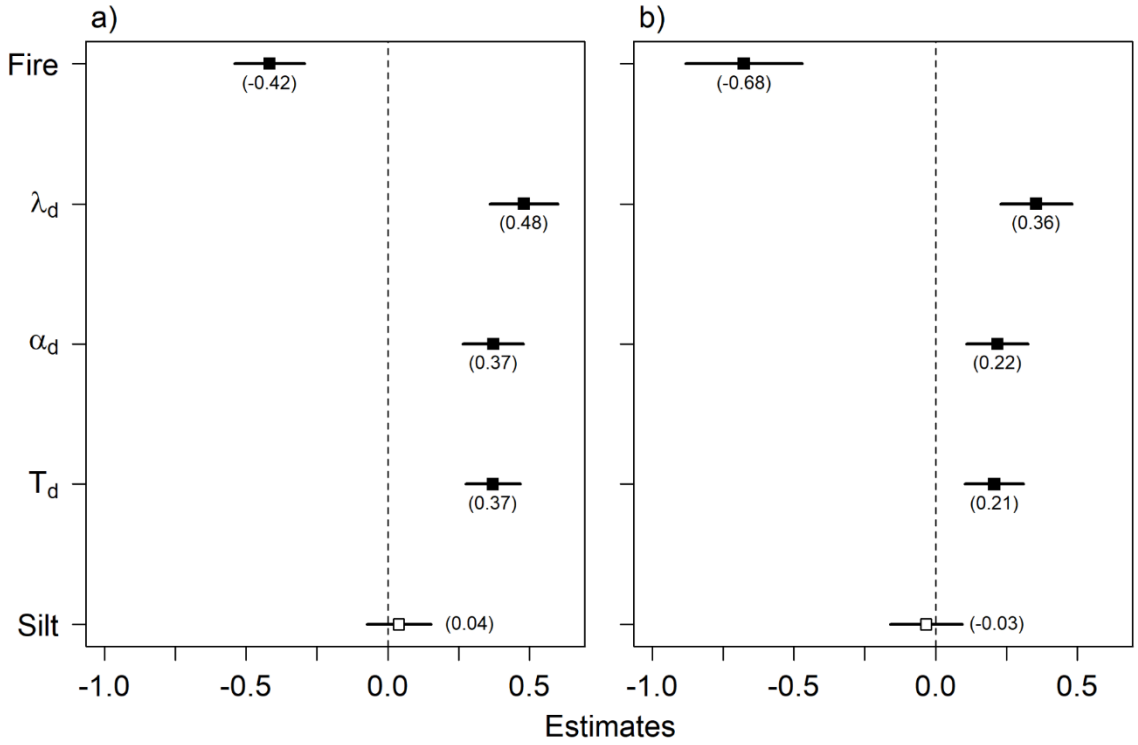

Figure S8. Effect of each predictor variable on a) tree cover and b) $PAVD_{max}$. Standardized slope estimates represent the effect of each variable on vegetation structure. The median estimate (points) and 95% confidence interval (error bars) are based on 1000 generalized linear models (GLM). Terms are not significant (open symbol) when the confidence interval includes zero (dashed vertical line). Details in Supporting information Table S3.

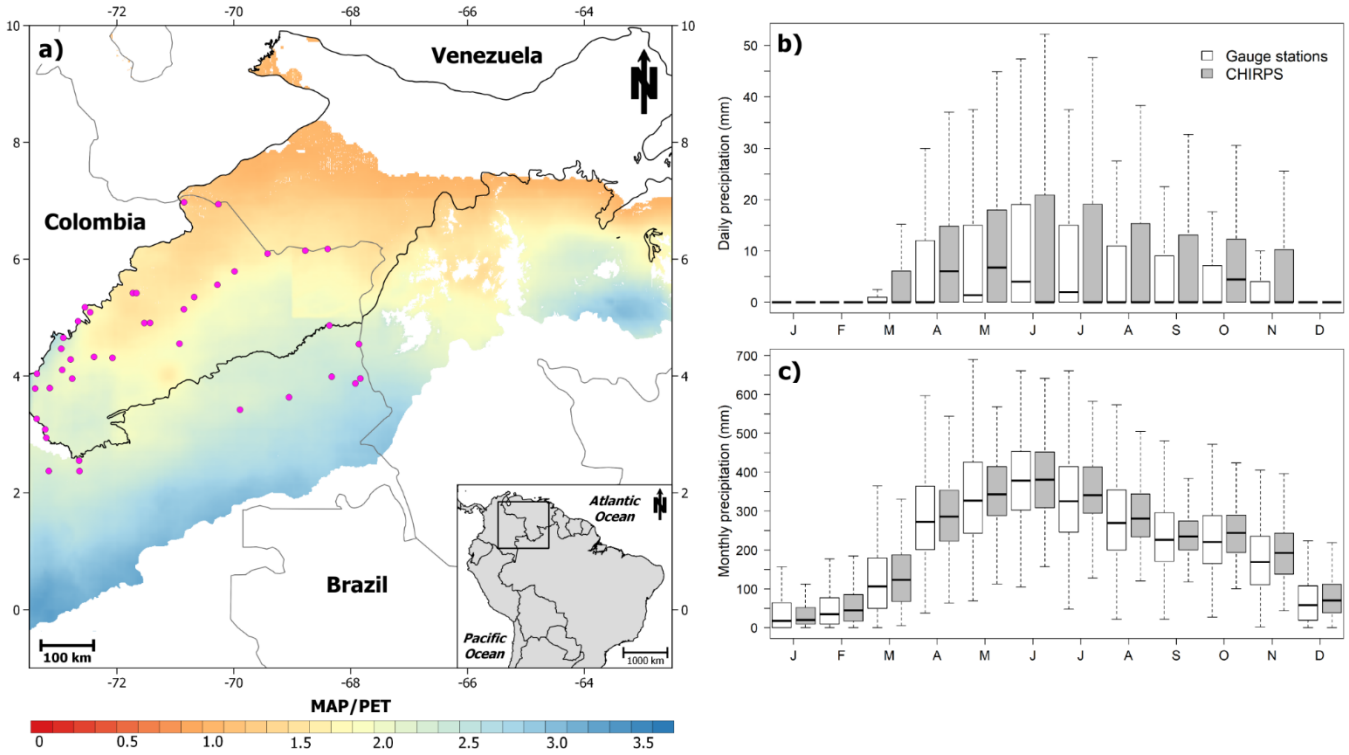

Figure S9. a) Precipitation gauge stations (magenta points, N=40) from the Colombian Institute of Hydrology, Meteorology and Environmental Studies (IDEAM) within the study area with ≤10% of daily missing data for the period 2004-2015 (Valencia et al., in preparation). Daily b) and c) monthly precipitation from gauge stations and the corresponding CHIRPS pixels for the period 2004-2015; months with more than 5 days of missing values in gauge stations were excluded from the analysis. In b) and c), outlier values were excluded for visualization purposes.

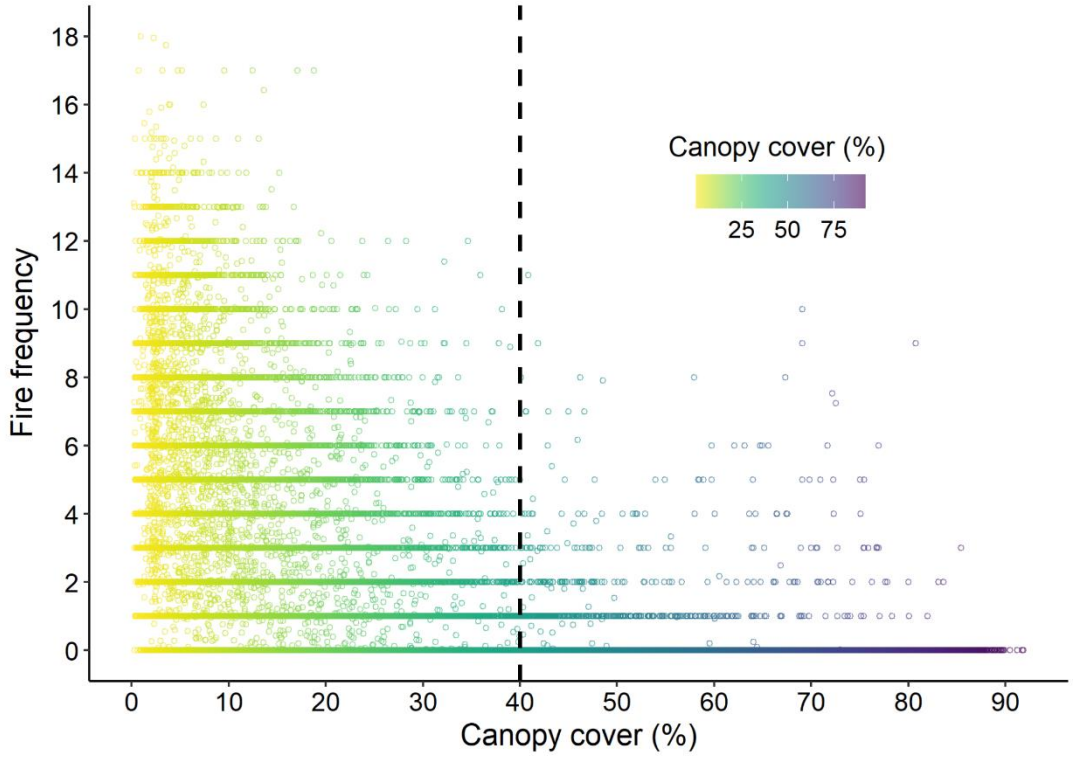

Figure S10. Fire frequency in the forest and savanna pixels that occur in the same MAP (2360 to 3070 mm) and $\boldsymbol{\lambda_d}$ (0.21 to 0.43) range (Figure 4b) for the period 2001-2019. The vertical dashed line indicates the canopy cover threshold used to define savanna (< 40 %) and forest (≥ 40%) pixels.

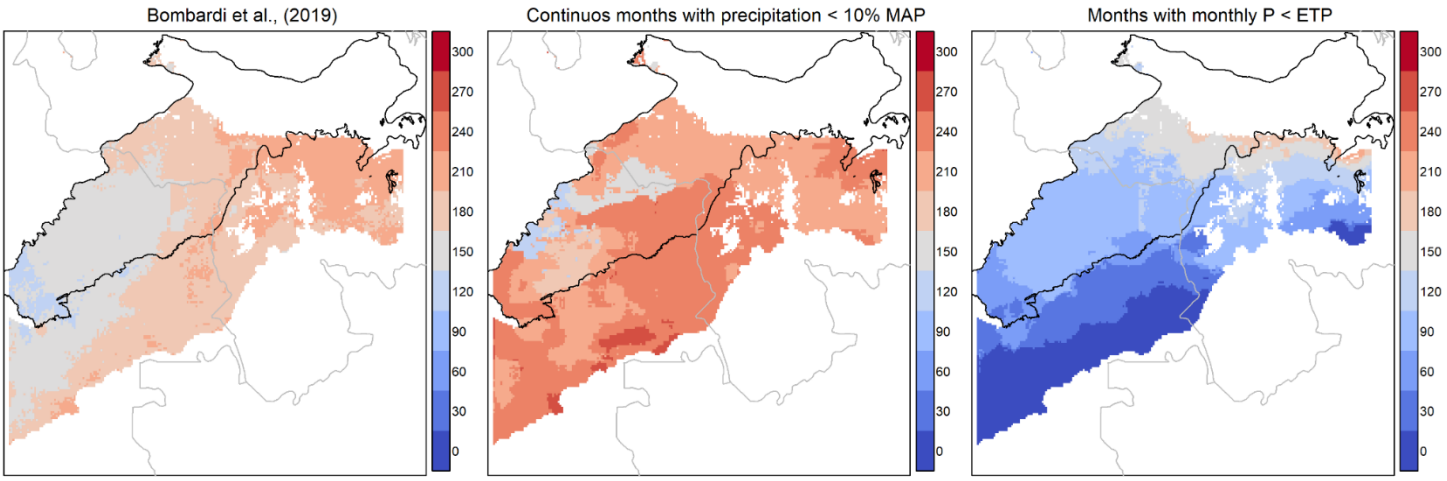

Figure S11. Length of dry season (days) based on a) Bombardi et al. (2019; used in this study), b) continuous months with precipitation (P) lower than 10% of MAP (e.g., Staver et al., 2011), and c) months with monthly precipitation lower than potential evapotranspiration (PET; e.g., Yang et al., 2016).

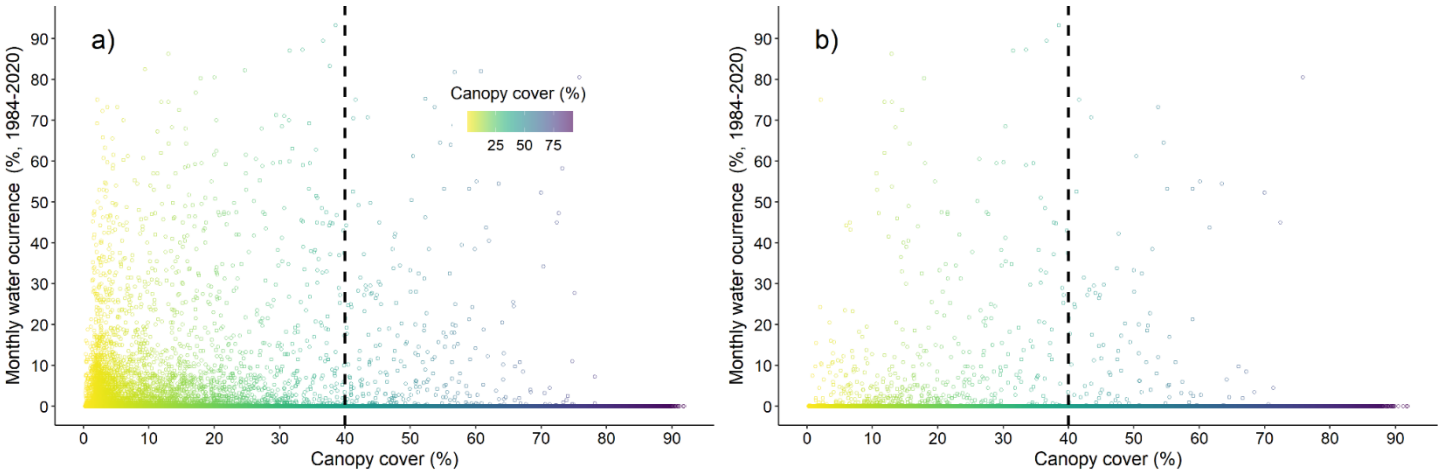

Figure S12. Percentage of monthly water occurrence during the period 1984-2020 obtained from Pekel et al. (2016) for a) all pixels within our study area (N = 201474) and b) forest and savanna pixels (N=68677) in the same climatic space based on MAP (2360 to 3070 mm) and $\boldsymbol{\lambda_d}$ (0.21 to 0.42). Monthly water occurrence was resampled from 30 m to 0.01° (~1 km).

References

Adzhar, R., Kelley, D. I., Dong, N., Torello Raventos, M., Veenendaal, E., Feldpausch, T. R., ... & Gerard, F. (2021). Assessing MODIS Vegetation Continuous Fields tree cover product (collection 6): performance and applicability in tropical forests and savannas. *Biogeosciences Discussions*, 1-20.

Bernardi, R. E., Staal, A., Xu, C., Scheffer, M., & Holmgren, M. (2019). Livestock herbivory shapes fire regimes and vegetation structure across the global tropics. *Ecosystems*, *22*(7), 1457-1465.

Bombardi, R. J., Kinter III, J. L., & Frauenfeld, O. W. (2019). A global gridded dataset of the characteristics of the rainy and dry seasons. *Bulletin of the American Meteorological Society*, *100*(7), 1315-1328.

Calders, K., Armston, J., Newnham, G., Herold, M., & Goodwin, N. (2014). Implications of sensor configuration and topography on vertical plant profiles derived from terrestrial LiDAR. *Agricultural and Forest Meteorology*, *194*, 104-117.

Compagnoni, A., Levin, S., Childs, D. Z., Harpole, S., Paniw, M., Römer, G., ... & Knight, T. M. (2021). Herbaceous perennial plants with short generation time have stronger responses to climate anomalies than those with longer generation time. *Nature communications*, 12(1), 1-8.

D'Onofrio, D., Sweeney, L., von Hardenberg, J., & Baudena, M. (2019). Grass and tree cover responses to intra-seasonal rainfall variability vary along a rainfall gradient in African tropical grassy biomes. *Scientific reports*, *9*(1), 1-10.

D'Onofrio, D., von Hardenberg, J., & Baudena, M. (2018). Not only trees: Grasses determine African tropical biome distributions via water limitation and fire. *Global ecology and biogeography*, *27*(6), 714-725.

Decuyper, M., Mulatu, K. A., Brede, B., Calders, K., Armston, J., Rozendaal, D. M., ... & Bongers, F. (2018). Assessing the structural differences between tropical forest types using terrestrial laser scanning. *Forest Ecology and Management*, 429, 327-335.

Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N. D., Wikramanayake, E., ... & Hansen, M. (2017). An ecoregion-based approach to protecting half the terrestrial realm. *BioScience*, *67*(6), 534-545.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., ... & Lautenbach, S. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36(1), 27-46.

Dubayah, R., Blair, J. B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., ... & Armston, J. (2020). The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of Remote Sensing*, 1, 100002.

Hirota, M., Holmgren, M., Van Nes, E. H., & Scheffer, M. (2011). Global resilience of tropical forest and savanna to critical transitions. *Science*, *334*(6053), 232-235.

Luthcke, S., Sabaka, T., Nicholas, J., Preaux, S., & Hofton, M. (2021). GEDI L3 Gridded Land Surface Metrics, Version 1. *ORNL DAAC*.

Meeussen, C., Govaert, S., Vanneste, T., Calders, K., Bollmann, K., Brunet, J., ... & De Frenne, P. (2020). Structural variation of forest edges across Europe. *Forest Ecology and Management*, 462, 117929.

Pekel, J. F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, *540*(7633), 418-422.

Rishmawi, K., Huang, C., & Zhan, X. (2021). Monitoring Key Forest Structure Attributes across the Conterminous United States by Integrating GEDI LiDAR Measurements and VIIRS Data. *Remote Sensing*, *13*(3), 442.

Staal, A., Dekker, S. C., Xu, C., & van Nes, E. H. (2016). Bistability, spatial interaction, and the distribution of tropical forests and savannas. Ecosystems, 19(6), 1080-1091.

Staver, A. C., Archibald, S., & Levin, S. A. (2011). The global extent and determinants of savanna and forest as alternative biome states. *Science*, *334*(6053), 230-232.

Tang, H., Armston, J., & Dubayah, R. (2019). Algorithm Theoretical Basis Document (ATBD) for GEDI L2B Footprint Canopy Cover and Vertical Profile Metrics. *Goddard Space Flight Center: Greenbelt, MD, USA*.

Yang, Y., Donohue, R. J., & McVicar, T. R. (2016). Global estimation of effective plant rooting depth: Implications for hydrological modeling. *Water Resources Research*, 52(10), 8260-8276.