Species diversity and stand structure as drivers of canopy complexity in southern African woodlands

John L. Godlee

8th August 2021

1   **Abstract**

2   Atmospheric CO_2 enrichment and human-induced climate change are expected to drive
3   woody encroachment and an increase in tree cover across African savannas, with consequences
4   for ecosystem function, particularly related to carbon dynamics. The patch dynamics of
5   savanna-woodland mosaics are complex however, as woody growth is mediated by seasonal
6   fire that is itself driven by woody canopy structure. It is unclear how variation in existing
7   tree species composition and stand structure in this ecosystem affects canopy structure, and
8   how this might determine future vegetation dynamics. Here, I conducted the first study of
9   canopy structure in miombo savannas using terrestrial LiDAR, at two sites in Angola and
10   Tanzania, to explore relationships between tree species diversity, stand structure, and canopy
11   structure. I found that **WHAT**

1   Introduction

13   Atmospheric CO_2 enrichment, coupled with climate change and changing disturbance regimes,
14   is expected to drive woody encroachment, i.e. proliferation of trees in previously non-wooded
15   areas, and woody thickening, i.e. increased growth of trees in currently wooded areas, across
16   the savanna biome over the coming century (**Criado2020**; **Stevens2016**; **Mitchard2013**). As
17   atmospheric CO_2 concentrations increase, C_3 trees are expected to gain a competitive edge over
18   C_4 grasses due to differences in photosynthetic pathway (**Buitenwerf2012**), with cascading
19   effects on canopy cover, grass growth, and therefore disturbance regime (**Bond2012**). If realised,
20   woody encroachment and thickening will have significant effects on the global carbon cycle, as
21   more CO_2 is stored as woody biomass, as well as myriad other effects on ecosystem structure
22   (**Donohue2013**). Indeed, tropical savannas have been identified as the fastest increasing
23   component of the terrestrial carbon sink (**Sitch2015**). Previous studies however, have reported
24   wide variation in rates of woody encroachment and thickening (**Mitchard2013**), particularly
25   in disturbance-prone savannas such as miombo woodlands in southern Africa (**Lewis2009**),
26   and it is unclear how the fertilisation effect of atmospheric CO_2 enrichment will interact other
27   ecosystem properties to alter vegetation (**Korner2017**; **Reich2014**).

28   Savanna vegetation is defined by the coexistence of trees and grasses (**Scholes1997**). In the
29   tropical mesic savannas of southern Africa, disturbance by fire and herbivory are the main
30   limitations on tree cover, preventing the competitive exclusion of shade-sensitive C_4 grasses where
31   climatic conditions would otherwise allow for closed canopy forest (**Sankaran2005**). C_4 grasses
32   also provide the main fuel source for seasonal fires in these savannas (**Frost1996**), producing a
33   positive feedback where an increase in tree cover reduces grass fuel load, reducing fire frequency
34   and intensity, increasing tree cover, and so on (**Staver2015**). As such, even small perturbations
35   in tree cover can lead to large changes in vegetation structure if critical thresholds of tree cover
36   are crossed (**Hirota2011**). Previous research has sought to identify environmental factors which
37   affect tree cover and its responses to atmospheric CO_2 enrichment, but few have considered the
38   functional role of the existing tree community and its effect on ecosystem processes.

39   Canopy structure describes the spatial distribution of tree canopy foliage (**Lowman2004**).
40   Canopy structural complexity, i.e. the spatial heterogeneity of foliage distribution within the

canopy, has been linked to increased net ecosystem productivity (**Hardiman2011**; **Chen2012**; **Law2001**; **Baldocchi2001**; **Morin2015**), increased resilience of productivity (**Pretzsch2014**), reduced understorey light penetration (**Scheuermann2018**; **Fotis2018**), and greater moderation of understorey micro-climate (**Wright2017**). Furthermore, in temperate and boreal forests, functional differences among coexisting tree species in their vertical and horizontal canopy occupation provide a link between species diversity, canopy structural complexity and canopy density, with canopy structure constituting a mechanism for observed positive biodiversity-ecosystem function effects in wooded ecosystems (**Pretzsch2014**; **Barry2019**). In tropical savannas, tree species diversity might therefore influence ecosystem-level woody thickening in response to elevated atmospheric CO_2, where diverse tree communities are less limited by competition due to niche separation, and can more effectively increase foliage density and reduce understorey light penetration, excluding grass and thus reducing disturbance.

As well as the species diversity of trees, the spatial distribution and relative size of trees, i.e. stand structure, is also expected to affect canopy structural complexity (**Stark2015**). Heterogeneity in tree size, whether a result of species diversity, disturbance history or some other factor, is expected to increase canopy complexity and canopy density as individuals of different sizes occupy different parts of the vertical canopy space (**Panzou2020**), and may differ in light requirements (**CharlesDominique2018**). Additionally, clustering of individuals in space is expected to increase canopy structural heterogeneity across the wider savanna landscape, but ultimately decrease total foliage density due to an increase in competitive interactions (**Dohn2017**). Clustering may occur as a result of disturbance history, or as a result of strong facilitation effects among individuals in stressful environments (**Ratcliffe2017**). More diverse communities may allow more dense clustering, as differences in canopy occupancy among species can reduce competition, meaning that diversity may reduce the negative effect of disturbance on tree cover ().

Functional differences among floristic types of savanna may also drive variation in canopy structure, irrespective of species diversity. Some savanna trees form denser canopies than others, as a result of variation in leaf size and branch architecture. Previous studies have compared the branch architecture of ex-Acacia and miombo archetypal tree species. While ex-Acacia species tend to inhabit drier, heavily grazed or seasonally flooded areas, miombo species tend to inhabit dystrophic wetter areas structured heavily by fire (**Ribeiro2020**). These studies have shown that ex-Acacia species develop sparser canopies, cagey branch architecture, and wider spreading crowns, while miombo species develop thicker canopies and can grow to large trees (**Mugasha2013**; **Archibald2003**; **Privette2004**). Under identical stem densities, miombo woodland species may therefore exclude grass more effectively.

Canopy structure is multi-dimensional and has previously been explained using a plethora of simple metrics that originated in forest and community ecology (**Kershaw2017**). Assessments of canopy structure have most often modelled tree canopies as a series of ellipses (2D), ellipsoids or cones (3D) based on field measurements with measuring tapes (**Jucker2015**). Measurements of this kind are time consuming and yet are an over-simplification of canopy structure. Alternatively, canopy cover is often measured using indirect optical methods which partition sky from canopy material, i.e. with hemispherical photography or the commonly used LAI-2000, providing a 2D representation of the canopy but lacking information on vertical canopy structure (**Jonckheere2004**). In recent years, particularly in temperate and boreal forests, LiDAR (Light Detection And Ranging) has emerged as a suitable technology for rapidly and precisely assessing canopy structure in 3D, conserving information on 3D structure of the calibre that is required to understand it's complexities (**Muir2018**; **Calders2020**). In tropical savannas, very few studies have used terrestrial LiDAR for vegetation analyses, and in southern Africa all existing studies have been located the the Skukuza Flux Tower in Kruger National Park, South Africa (**Muumbe2021**). Pioneering work describing the ecology of southern African savannas placed

large emphasis on canopy structural diversity as a mediator of ecosystem function (**Solbrig1996**), but much of that understanding of savanna vegetation structure was derived from traditional mensuration methods. Using terrestrial LiDAR to measure canopy structure in southern African savannas therefore offers a unique chance to validate accepted theory and describe differences in ecosystem structure among savanna vegetation types in finer detail than previously possible.

In this study I applied terrestrial LiDAR techniques to woodland-savanna mosaics at two sites in southern Africa, with the aim of increasing our understanding of how various metrics of tree canopy structural complexity relate to tree neighbourhood diversity and stand structure. I aim to develop our understanding of how biotic ecosystem properties in savannas might mediate responses to atmospheric CO_2 enrichment and climate change. I hypothesise that neighbourhoods with greater tree diversity and greater structural diversity allow greater canopy complexity, and foliage density. Thus, more diverse savannas might more effectively increase growth under elevated atmospheric CO_2 and are more likely to experience woody thickening through their greater occupation of environmental niche space. I also consider the functional differences among tree species in these communities and assess how combinations of these functional groups affect canopy structure and understorey light environment.

2 Materials and methods

2.1 Study sites

Field measurements were conducted at two sites, the first in Bicuar National Park, southwest Angola (S15.1°, E14.8°), and the second in and around Mtarure Forest Reserve, southeast Tanzania (S9.0°, E39.0°) (Figure 1). At each site, 1 ha (100x100 m) plots were located in areas of savanna-woodland vegetation, across a gradient of stem density and range of savanna floristic archetypes. In Angola, 15 plots were sampled, while in Tanzania, seven were sampled following the curtailment of fieldwork due to COVID-19 travel restrictions. Fieldwork was conducted between February and April at both sites, during the peak growth period of each site in order to capture the greatest foliage volume in the canopy.
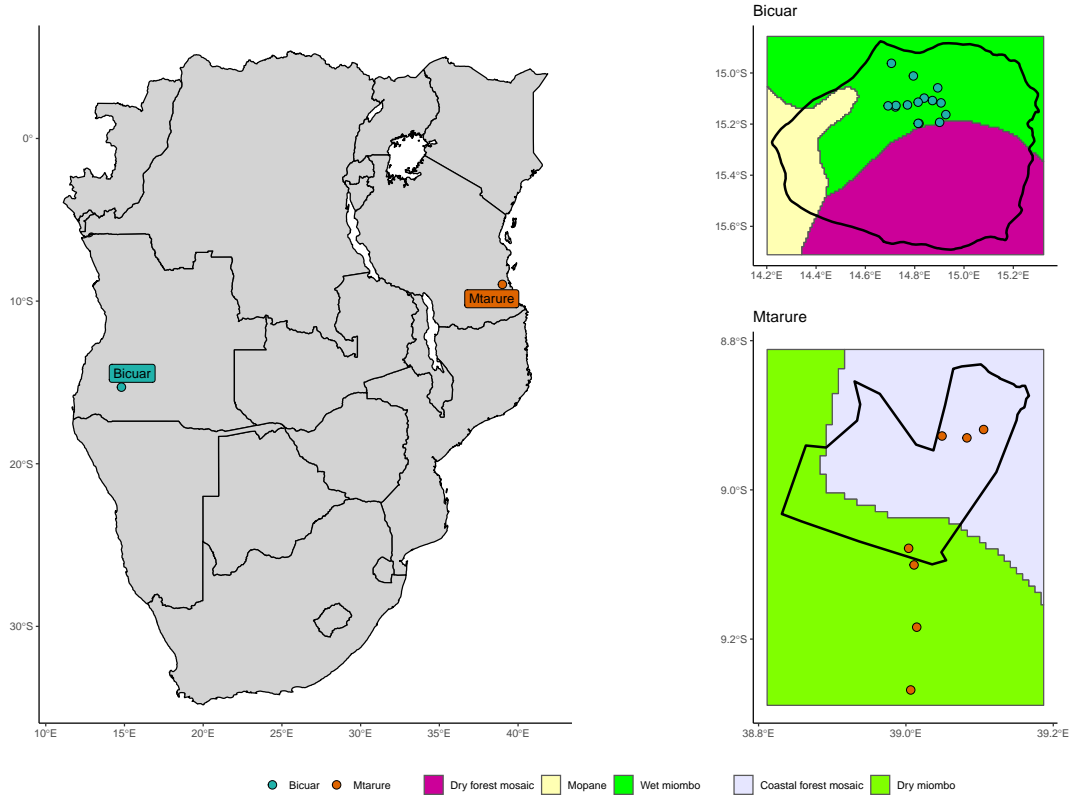
Figure 1: Location of study sites within southern Africa (left), and of 1 ha plots within each site (right). The black outlines in each site map denote the boundaries of protected areas which encompass the majority of study sites, Bicuar National Park in Angola (top), and Mtarure Forest Reserve in Tanzania (bottom). The background of each site map is a re-classified version of White's vegetation map (**White1983**). Note that all maps are on different scales.

2.2 Field measurements

Within each 1 ha plot we identified each woody stem >5 cm stem diameter to species, measured stem diameter (Diameter at Breast Height - 1.3 m) and recorded stem location within the plot using tape measures. Each 1 ha plot was further subdivided into nine 10 m diameter circular subplots arranged in a regular grid, with a 15 m buffer from the plot edge and 35 m between subplots. For each subplot, we recorded the distance and direction from the subplot centre of each stem >5 cm diameter with canopy material inside the subplot. Within each subplot, a variable number of scans were recorded using a Leica HDS6100 phase-shift Terrestrial Laser Scanner (TLS). The number and position of scans within a subplot was determined by the arrangement of canopy material in the subplot, to minimise shadows within the canopy of the subplot, and to maximise canopy penetration. The number of scans per subplot ranged between one and five across both sites.

2.3 Data analysis

2.3.1 TLS processing

Point clouds from scans in each subplot were registered and unified using Leica Cyclone (version 9.1), using five reflective cross targets visible to all scans. Point clouds were voxelised to cubic

4

voxel sizes of different sizes depending on the application of the data. For subplot height profile estimation and gap fraction we used 5 cm^3 voxels, and for whole plot canopy rugosity we used 50 cm^3 voxels. Voxels were classified as 'filled' if they intersected one or more points. Variation in voxel size reflects the spatial scale of each analysis, and is bounded by the beam divergence of the scanner over longer distances (). Choosing voxels that are too small can result in pock-marked representations of surfaces that are especially problematic when calculating larger scale canopy structure metrics such as canopy top roughness, while voxels that are too large can result in an over-estimation of plant volume when estimating canopy foliage density at the subplot scale (**Seidel2012**; **Cifuentes2014**). We used the noise reduction algorithm from **Rusu2008** to discard points based on mean nearest neighbour distances, with a mean number of neighbours of eight, and a standard deviation threshold of 1.96. This effectively removed 'ghost points' produced by partial beam interceptions and also removed many erroneous returns caused by airborne dust particles, which was common at our study sites. Raw points clouds for each subplot had a mean of ~2.9e+08 points, ~4.5e+07 points after voxelisation to 5 cm^3, and ~2.1e+07 points after noise reduction. Ground points were classified using the Progressive Morphological Filter (PMF) from **Zhang2003**. Point cloud height was reclassified height based on this revised ground layer by measuring the vertical distance between the nearest ground point and each point.

At the subplot level, we used ray-tracing to estimate canopy cover at the subplot centre from multiple TLS scans. Hemispherical images were created using the POV-ray ray-tracing software (**Povray2004**). Filled voxels were represented as matt black cubes filling the voxel volume, with a white sky box and no light source. A 'camera' with a 180° fisheye lens was placed at the subplot centre within POV-Ray, at a height of 1.8 m pointing directly upwards. The images produced by POV-Ray were analysed using Hemiphot (**HemiPhot**) to estimate canopy cover as the proportion of pixels filled by canopy material. Canopy cover estimates from the TLS were validated with hemispherical photographs taken at the same location and processed using the same method in Hemiphot, and compared using Pearson's correlation ($r(195) = 0.89$, p<0.001). The Effective Number of Layers (ENL) measures subplot canopy complexity and was calculated as the Shannon entropy of foliage density among 50 cm height bins within each subplot. The uniformity of foliage distribution was calculated by fitting a linear model to the cumulative foliage density profile, then extracting the standard error on the slope estimate of this linear model. Total foliage density was calculated as the area under the curve of the foliage height profile.

At the plot level, canopy complexity was measured by two metrics. Canopy top roughness was measured as the standard deviation of canopy height across the plot. Canopy rugosity was measured according to **Hardiman2011**, as the standard deviation of vertical and horizontal foliage density within 0.5 m cubic bins. We also estimated plot-level canopy cover by calculating the mean of the canopy cover values from each subplot.

2.3.2   Stand structure

For each subplot, we calculated an adapted version of the Hegyi index to estimate crowding, as an alternative to stem density which does not adequately capture crowding at small spatial scales, when only a small number of trees are included in the sample (**Hegyi1974**). To estimate subplot structural diversity we calculated the coefficient of variation of stem diameter as a measure of the heterogeneity of tree size in the neighbourhood.

At the plot level, we estimated the regularity of species spatial distribution using the spatial mingling index (**Gadow2002**). We also measured the uniformity of whole plot stem distribution using the winkelmass, which measures the spatial clustering of stems (**Gadow2002**). Finally, we calculated plot level stem density.

2.3.3   Statistical analysis

To describe variation in species composition among plots, we conducted Non-metric Multi-dimensional Scaling (NMDS) analysis on genus-level basal area weighted abundance in each plot. We excluded stems that could not be identified to genus from this analysis, which accounted for 0.2% of the total basal area recorded. Four distinct vegetation types were identified, two from each site. These vegetation types are summarised in Table 1.

Linear mixed effects models tested the effects of tree species diversity and stand structural diversity on canopy complexity. Mixed models were used to account for the nested sampling design of subplots within plots and plots within vegetation types. Models were conducted at both the subplot and plot level. Subplot models included random effects for plot nested within vegetation type, while plot level models included random effects for vegetation type only. Separate models were fitted for each canopy complexity metric, resulting in four models at the subplot level and four models at the plot level. We compared effect sizes among fixed effects in each model, and also compared AIC values and Akaike weights of models with different combinations of fixed effects to determine which diversity and structural metrics best explained variation in each canopy complexity metric.

To test the hypothesis that tree species diversity may influence canopy complexity indirectly through its effect on stand structure, we conducted a path analysis using the `piecewiseSEM` R package (). The path analysis investigated the direct effect of plot species richness on mean plot canopy cover, as well as the indirect effect of richness on cover via the coefficient of variation of diameter, with random intercept terms for each vegetation type.

Table 1: Climatic information and Dufrene-Legendre indicator species analysis for the vegetation type clusters identified by the PAM algorithm, based on basal area weighted species abundances. The three species per cluster with the highest indicator values are shown along with other key statistics for each cluster. MAP (Mean Annual Precipitation) and δT (Diurnal temperature range) are reported as the mean and 1 standard deviation in parentheses. Species richness is reported as the median and the interquartile range in parentheses.

Cluster	N sites	Richness	Stem dens.	AGB	Species	Ind. value
1	12	17(2)	642(194)	41(8.4)	Strychnos spinosa	0.83
					Combretum collinum	0.74
					Julbernardia paniculata	0.70
2	5	23(4)	411(137)	72(11.9)	Pteleopsis myrtifolia	1.00
					Diplorhynchus condylocarpon	0.89
					Pseudolachnostylis maprouneifolia	0.81
3	3	6(1)	196(55)	77(7.3)	Baikiaea plurijuga	0.94
					Baphia massaiensis	0.83
					Philenoptera nelsii	0.45
4	2	12(2)	288(73)	9(0.2)	Vachellia nilotica	0.99
					Combretum apiculatum	0.70
					Senegalia polyacantha	0.62

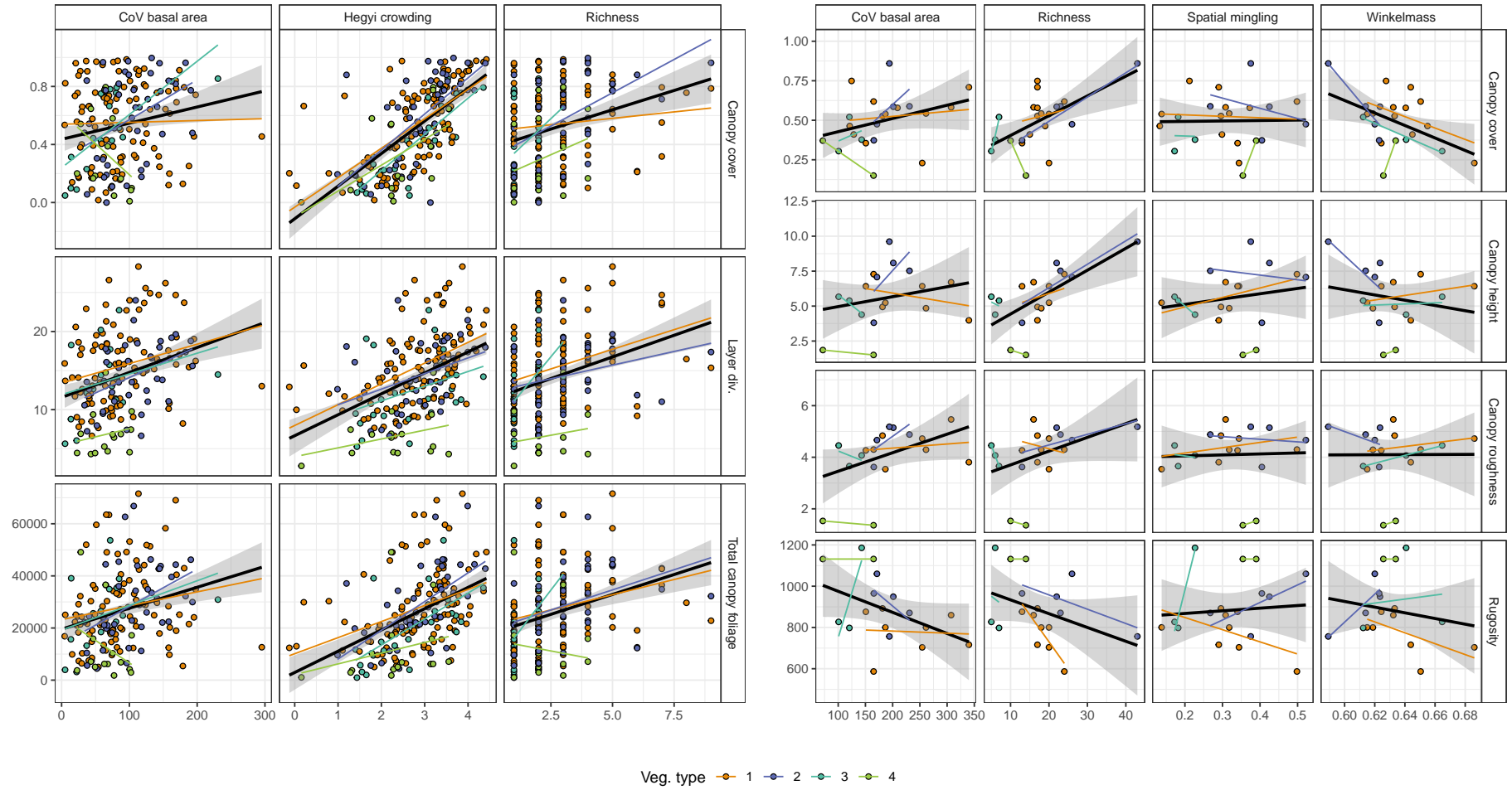# 3 Results

3.1 Vertical canopy complexity

Figure 2: Bivariate relationships between canopy structure metrics (y axis) and diversity/stand structure metrics (x axis), at both the subplot level (left) and the plot level (right). Points and linear model lines of best fit are coloured by vegetation type. The black line of best fit is a linear model including all plots, with a 95% confidence interval. See **supp. material** for a comparison of linear model fits by vegetation type.

Bivariate plots showed that subplot species diversity, measured by species richness of the tree neighbourhood around each 10 m diameter subplot, appeared to have weak but positive effects on canopy layer diversity and total canopy cover (Figure 2). The Hegyi crowding index and both stand structural diversity metrics had strong positive effects on canopy complexity, for all metrics except for uniformity of foliage distribution and height of peak foliage density. The two sites in our study had similar bivariate relationships, with interaction effects of site in the bivariate linear models being non-significant in all cases **(supp. material)**.

Linear mixed effects models showed that species richness of the subplot neighbourhood had variable effects across different measures of canopy structure, but the effect sizes were not significant (slope standard errors not overlapping zero) for any model (Figure 3). One exception being the negative effect of richness on canopy height in Mtarure only. As in the bivariate plots, the Hegyi crowding index had strong positive effects on three of six canopy complexity metrics. Heterogeneity of stem diameter had a positive effect on layer diversity and total foliage density, and a marginally significant positive effect on canopy height. Variation in crown area was only seen to have significant effects in Bicuar plots, where it correlated with a decrease in vertical uniformity of foliage distribution, and total canopy foliage density.
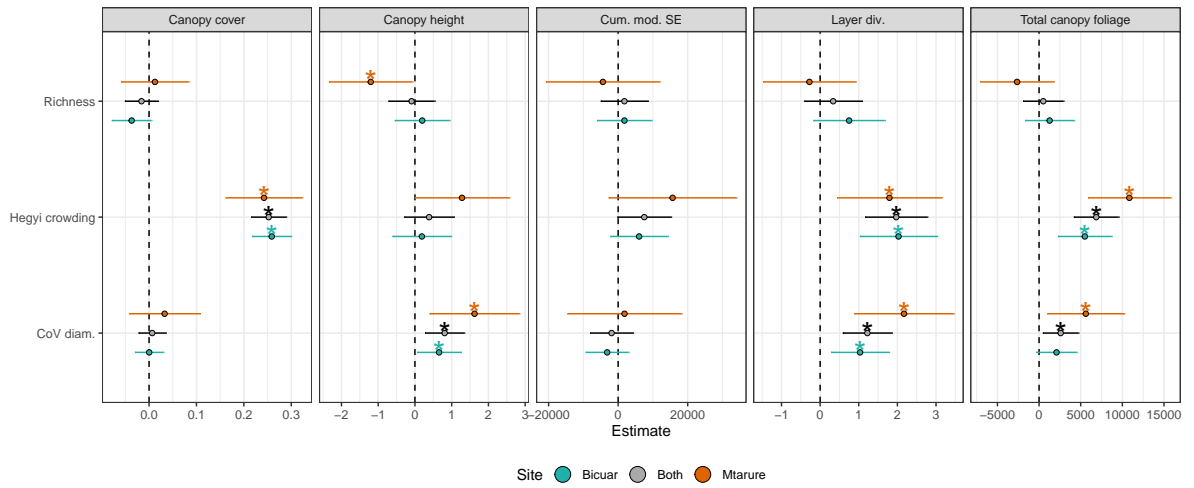

Figure 3: Standardized fixed effect slopes for each model of a canopy structure metric. Slope estimates are ±1 standard error. Slope estimates where the interval (standard error) does not overlap zero are considered to be significant effects. Points are coloured according to site.

The model selection process showed that the best model for layer diversity included species richness. Stand structural diversity metrics were included in the best models for all canopy complexity metrics except for canopy cover, which was predicted solely by the Hegyi crowding index. Models of layer diversity, total foliage density, and canopy cover were predicted well by a combination of crowding and stand structural diversity. Models of height of peak foliage density, canopy height, and uniformity of foliage distribution were poorly constrained by the available fixed effects, with R^2_m of ~5%. The majority of the total model effect on canopy height came from the random effects of site and plot identity.
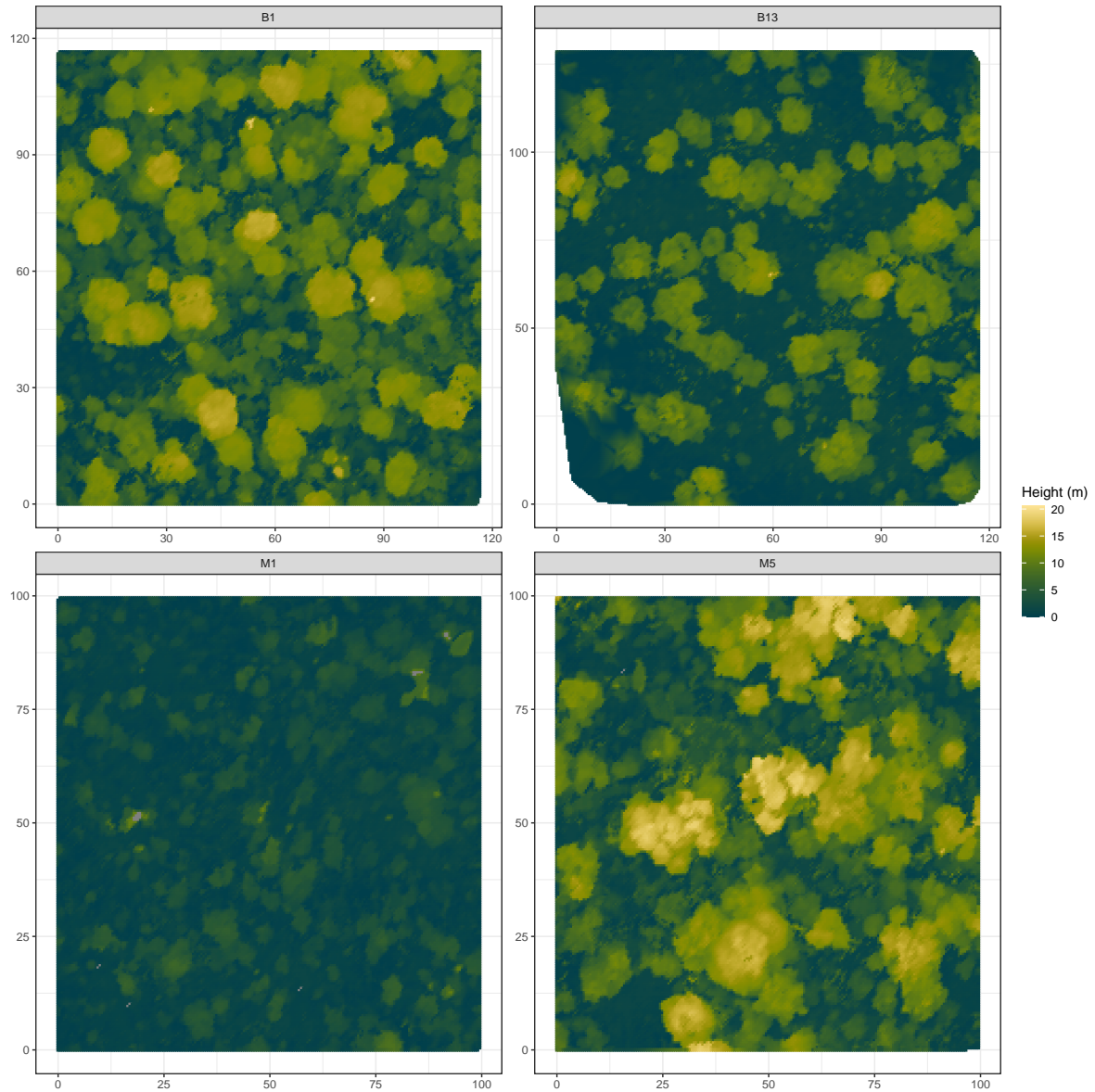
3.2 Canopy rugosity

Figure 4: Representative canopy surface models for each vegetation type identified in the Non-metric Multi-dimensional Scaling (NMDS) analysis. Plot titles show the plot name and the vegetation type.

Similar to the subplot analyses, at the whole-plot scale tree species diversity, measured here by the Shannon index, tended to have weak positive effects on canopy complexity metrics, while stand structural diversity metrics had stronger positive effects (**??**). Strong positive relationships of basal area on canopy complexity are driven mostly by two plots with particularly low basal area in Mtarure, M3 and M4. These plots are sparse thorny savanna, dominated by *Senegalia* spp. (Figure 6). Indeed, linear models using only plots in bicuar show divergent relationships. These two plots also have particularly low canopy cover, canopy height, and canopy top roughness, despite having similar tree species diversity and spatial distribution of trees (winkelmass) as other plots.

Linear mixed effects models show that increased spatial clustering of trees causes a decrease in

canopy cover. Increased spatial mingling of tree species causes an increase in canopy rugosity, while in contrast an increase in overall tree species diversity appears to cause a decrease in canopy rugosity. An increase in the heterogeneity of stem diameter causes an increase in canopy top roughness.
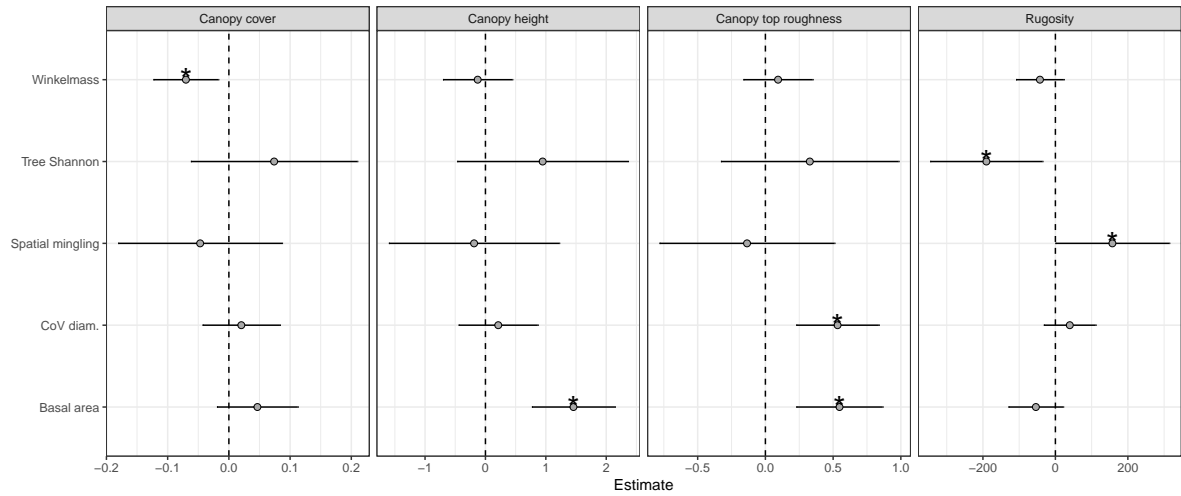

Figure 5: Standardized fixed effect slopes for whole-plot canopy rugosity. Slope estimates are ±1 standard error. Slope estimates where the interval (standard error) does not overlap zero are considered to be significant effects.

Model selection showed that all plot canopy complexity metrics except canopy rugosity were best modelled by a combination of basal area and either species diversity or structural diversity.
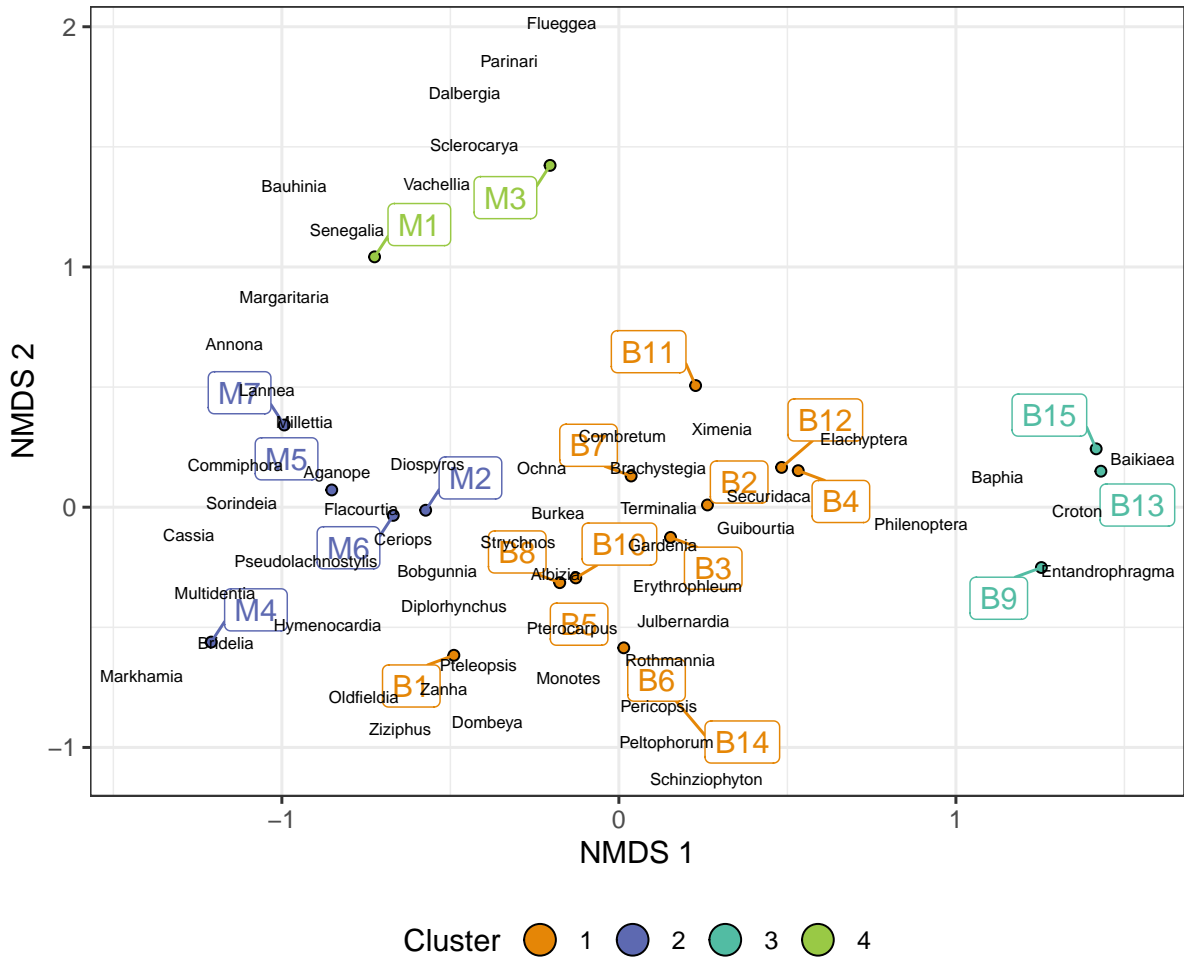
Figure 6: The first two axes of a Non-metric Multi-Dimensional Scaling (NMDS) analysis of tree species diversity in each plot. Species scores are labelled as black text, while plot scores are labelled as coloured points. Plots can be split into four principal groups: 1) B9, B13 and B15, dominated by *Baikiaea plurijuga*; 2) the other Bicuar plots; 3) M2, M5, M6, and M7,dominated by *Julbernardia* spp., *Brachystegia* spp. and *Ochna* spp.; 4) M1, M3, and M4, dominated by *Senegalia* spp. and *Vachellia* spp..

3.3 Comparing subplot and plot measures of canopy structure

Plot-level and subplot-level canopy structure metrics were highly correlated in many cases (Figure 5). Plot canopy height especially, tended to be strongly positively correlated with subplot canopy complexity. Additionally, as canopy top roughness increases, many subplot canopy complexity and density metrics increase. In the majority of cases, both sites had similar correlations of subplot and plot measures of canopy structure, with notable exceptions for plot roughness vs. layer diversity, plot roughnesss vs. canopy cover, and plot canopy height vs. canopy cover.

Variance of plot canopy height and plot roughness was larger in Mtarure than Bicuar. The increase in variance was caused by two particularly sparse thorny savanna plots in Mtarure, M3 and M4, which had very low canopy height and roughness.
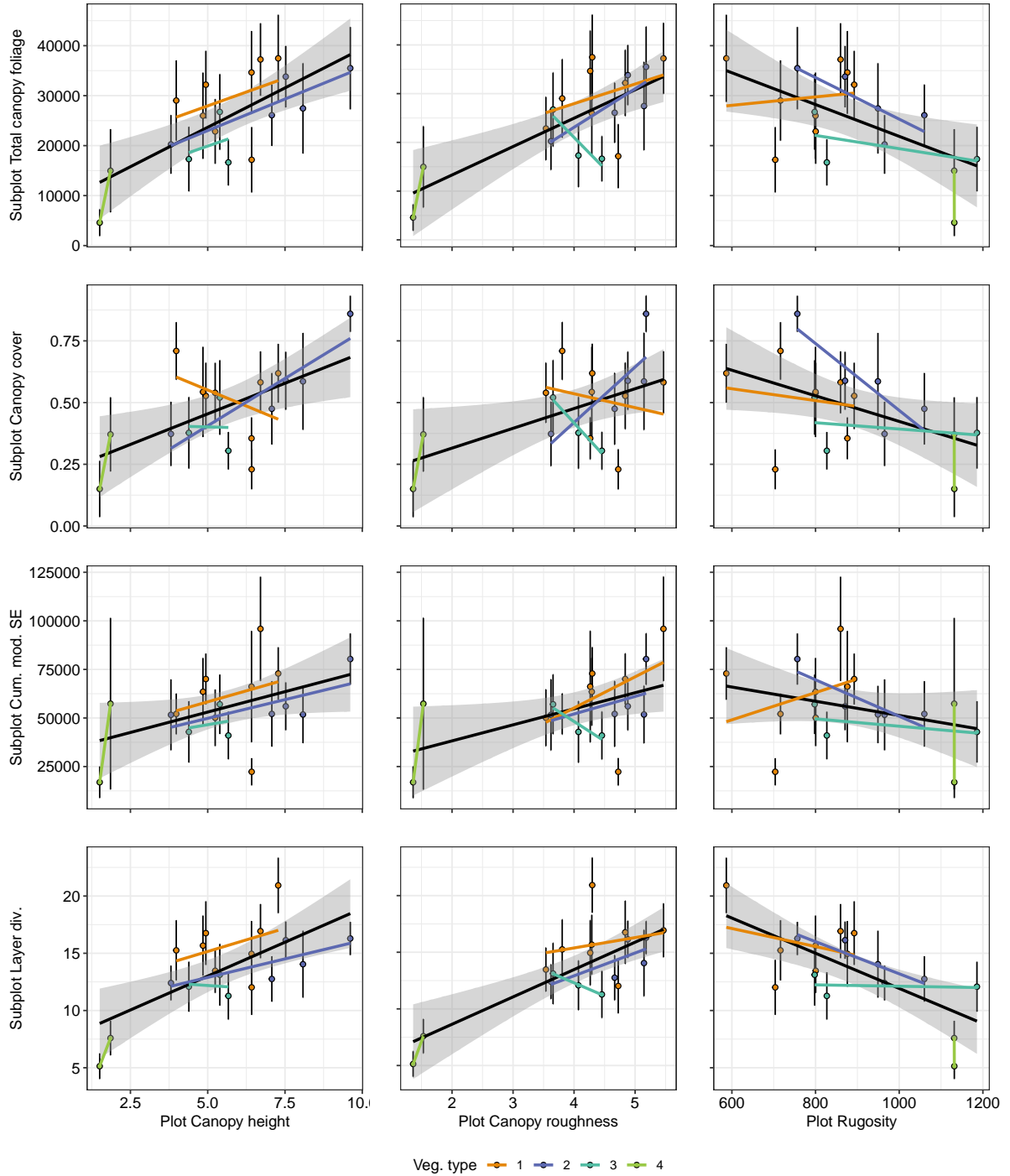
Figure 7: Bivariate plots of canopy structural metrics at the subplot (y axis) and plot level (x axis). Each point represents the mean values of a single plot. Points and linear model fits are coloured according to vegetation type. The black linear model combines all vegetation types. Error bars on points are the standard deviation of mean subplot metrics across the plot.

4   Discussion

We investigated the effects of tree species diversity and structural diversity on several metrics of canopy complexity that were hypothesised to affect plot productivity. Species diversity appeared

to generally have weak positive effects on canopy complexity at both the subplot and plot scales, while stand structural diversity had much stronger effects. The strongest determinant of canopy complexity was stem crowding, as measured by basal area and the Hegyi crowding index.

The positive relationships between species richness and subplot canopy complexity metrics observed in the subplot bivariate models were not seen in the linear mixed effects models. This is likely because the observed species richness effect was itself driven by stand structure. The Hegyi crowding index increases with stem density, i.e. decreased distance of individuals from the subplot centre. Species richness also increases with stem density, as a greater number of individuals is more likely to hold more species simply through sampling effects. **Jucker2015** however, did find that increased species diversity led to greater canopy packing in European forests, with trees in mixed forests having generally larger crowns. Our result that species diversity did not have consistent effects on canopy complexity may be specific to the vegetation type studied here. Southern African open woodlands are much more heavily affected by disturbance from fire and herbivory than temperate forests, meaning the effects of inter-specific competition are weakened as a driver of stand and canopy structure ().

Canopy structure at the plot level was less well predicted by stand structure and species diversity than subplot level canopy structure. Results at the plot level suggest that woodland vegetation type and basal area has the greatest effect on canopy complexity. The two thorny savanna plots in Mtarure produced strong positive effects of basal area and diameter variantion on canopy cover, canopy height, and canopy roughness, but when these plots are removed the remaining points do not produce strong relationships.

Facilitation might be more important in these woodlands than in temperate woodlands. Large canopy trees may cause micro-climate amelioration for understorey saplings, protecting them from drying conditions caused by the sun and wind. Facilitation has been under-played in BEFR research (**Wright2021**).

5   Conclusion