

1 Title: Phenology and species diversity in Zambian woodlands

2 Authors: Godlee, J. L.¹, Ryan, C. M.¹, Dexter, K. G.¹

3 ¹: School of GeoSciences, University of Edinburgh, Edinburgh, United Kingdom

4 ²: Some other address

5 Corresponding author:

6 John L. Godlee

7 johngodlee@gmail.com

8 School of GeoSciences, University of Edinburgh, Edinburgh, United Kingdom

9 **1 Acknowledgements**

Blinded Main Text File

Title: Phenology and species diversity in Zambian woodlands

Running title: Phenology and diversity in Zambia

2 Abstract

3 Introduction

The seasonal timing of tree leaf production in deciduous woodlands directly influences ecosystem-level productivity ().

Previous studies have shown that diurnal temperature variation and precipitation are the primary determinants of tree phenological activity in water-limited savannas, but uncertainty remains in the prediction of leaf production cycles in these ecosystems ().

It is important to control for the effects of climate in regional analyses of phenology if we wish to look at the effects of tree species diversity.

It is important to separate the signal of tree growth from grass growth. Grass tends to green-up after the rainy season starts, while trees can often green-up just before the rainy season. **MORE**

In this study we contend that tree species composition and tree species diversity influence two key measurable aspects of the leaf phenology cycle: (1) the rate of greening at the start of the seasonal growth phase, and the overall length of the growth period. It is hypothesised that: (H₁) due to variation among species in minimum viable water availability for growth, plots with greater species richness will exhibit a slower rate of greening. Additionally, we hypothesise that: (H₂) plots with greater species richness will exhibit a longer growth period and greater cumulative greenness over the course of the growth period, due to a higher resilience to variation in water availability, acting as a buffer to ecosystem-level productivity. Finally, we hypothesise that: (H₃) irrespective of species diversity, variation in tree species composition will cause variation in growth season length.

4 Materials and methods

4.1 Data collection

We used plot-level data on tree species diversity across sites from the **ILU Aii Zambian Forestry Commission national census** (). Each site consisted of four 20x50 m (0.2 ha) plots positioned north, east, south and west of a central point, with a distance of **20 m** from the central point to the long axis of each plot. Only sites with $> 1 \text{ ha}^{-1}$ were included in the analysis, to ensure all sites represented woodland rather than “grassy savanna”, which is considered a separate biome with very different species composition and ecosystem processes governing phenology (Parr et al., 2014). Sites in Mopane woodland were removed by filtering sites with greater than % of trees belonging to *Colophospermum mopane*, preserving only plots with Zambesian tree savanna / woodland. Mopane woodlands **have different processes governing their phenology, so it’s not sensible to include them.**

Within each of these plots all tree stems > 5 cm diameter at breast height (DBH) were inventoried. The following data was available for each stem: species, trunk diameter (DBH), and tree identity. Plots were measured in **2011**. Plot data was aggregated to site data for analyses to avoid pseudo-replication. Tree species composition of the four plots at each site was assumed to be representative of the larger area.

Climatic variables were derived from the WorldClim database, using the BioClim variables, with a

pixel size of 30 arc seconds (926 m at the equator) (?). Mean Annual Precipitation was calculated as the yearly sum of daily precipitation, averaged across all years of available data (1970-2000). Mean diurnal temperature range was calculated as the mean of monthly temperature range. **WHY ARE THESE USED? INTRO OR HERE?**

To quantify phenology at each site, we used the MODIS VIPPHEN satellite data product at 0.05° resolution (~5600 m at equator) (). The VIPPHEN product uses the modified 2-band Enhanced Vegetation Index (EVI2) and provides a number of scientific datasets (SDS) including phenological metrics such as the start, peak and length of the growing season as well as cumulative EVI2. EVI2 is well-correlated with gross primary productivity and so can act as a suitable proxy ().

We used data from all 18 available years (2000-2018) in the VIPPHEN product. A single value of each variable for each pixel was calculated by taking the mean over the available years.

All sites in the study occurred within discrete image pixels. All sites were determined to have a single annual growth season according to the VIPPHEN product, which can assign a pixel up to three growth seasons per year. For each site, we matched pixel values from the VIPPHEN SDS, extracting length of growth season, day of start of growth season, cumulative EVI2 across the growth season, and rate of greening at the start of the growth season. Growth season start and end are estimated **MORE**.

4.2 Data analysis

To quantify variation in tree species composition we computed a Principle Coordinate Analysis (PCoA), with Cailliez correction for potential negative eigenvalues (?), on a Bray-Curtis dissimilarity matrix calculated from a tree species abundance matrix per plot, using the **ape** R package (?). The first three axes of this PCoA explained of the variation in species composition among plots according to eigenvalue analysis. These three axes were used in further statistical modelling.

We used simple multivariate linear models to assess the role of tree species diversity on our three chosen phenological metrics. We defined a maximal model structure including tree species richness, the first 4 principle coordinate analysis axes of tree species composition, and climatic variables shown by previous studies to strongly influence phenology. A model selection process was used to determine if tree species diversity markedly improved model fit. Models were compared using AIC and the R^2 of the given phenological metric used as a response variable. Explanatory variables in each model were transformed to achieve normality where necessary and standardised to Z-scores prior to modelling. Standardisation put each explanatory variable on the same scale, with a mean of zero and a standard deviation of one. Standardisation allows regression coefficients to be easily compared within the same model.

Hierarchical partitioning was used to assess the independent and joint effects of each independent variable in the optimal model for each phenological metric (??), using the **hier.part** R package (?). Hierarchical partitioning calculates goodness-of-fit across all combinations of independent variables

All statistical analyses were conducted in R version 4.0.2 (?).

5 Results

6 Discussion

7 Conclusion