Title: Phenology and diversity in Zambia

Authors: Godlee, J. L.[1], Abel Siampale[2], Kyle G. Dexter[1,3]

[1]: School of GeoSciences, University of Edinburgh, Edinburgh, United Kingdom

[2]: Forestry Department Headquarters - Ministry of Lands and Natural Resources, Cairo Road, Lusaka, Zambia

[3]: Royal Botanic Garden Edinburgh, Edinburgh, EH3 5LR, United Kingdom

Corresponding author:

John L. Godlee

johngodlee@gmail.com

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

Acknowledgements

Author constribution statement

JLG conceived the study, conducted the analysis, and wrote the first draft of the manuscript. AS coordinated plot data collection in Zambia, and initial data management. All authors contributed to manuscript revisions.

Data accessibility statement

The data that support the findings of this study are held by the Zambian Integrated Land Use Assessment Project (ILUA-II), provided to SEOSAW (A Socio-ecological Observatory for Southern African Woodlands) for research purposes. The data are not publicly available at the time of submission due to privacy concerns surrounding plot location, but can be requested from the corresponding author. An anonymised version will be made available in a data repository following review.

Abstract

1 Introduction

The seasonal timing of tree leaf production in dry deciduous savannas directly influences ecosystem processes and structure (). Leaf Area Index (LAI), leaf area per unit ground area, is tightly coupled with photosynthetic activity and therefore Gross Primary Productivity (GPP) (). Directional shifts in GPP influence the accumulation rate of woody biomass, and affect the delicate balance between tree and grass co-occurrence in these ecosystems (), with potential consequences for transition between closed-canopy forest and open savanna. From a conservation perspective, deciduous savannas with a longer growth period support a greater diversity and abundance of wildlife, particularly bird species but also browsing mammals (). Extreme weather patterns as a result of climate change are leading to shorter but more intense leaf production cycles in these ecosystems which exist at the precipice of their climatic envelope, with severe negative consequences for biodiversity (). Understanding the determinants of seasonal patterns of tree leaf production (land-surface phenology) in dry deciduous savannas can provide valuable information on spatial variation in vulnerability to climate change, and help to model their contribution to land surface models under climate change.

Previous studies have shown that diurnal temperature variation and precipitation are the primary determinants of tree phenological activity in water-limited savannas (). At regional spatial scales, savanna phenological activity can be predicted well using only climatic factors and light environment (Adole, Dash, and Atkinson, 2018), but local variation exists in leaf production cycles which cannot be attributed solely to abiotic environment (). It has been repeatedly suggested that information on biotic environment play a larger role in predicting land-surface phenology (), but implementation is most often limited to coarse ecoregions or functional vegetation types (), which lack the fine-scale resolution which can now be paired with state-of-the-art earth observation data.

Tree species vary in their life history strategy regarding the timing of leaf production (). More conservative species (i.e. slower growing, robust leaves, denser wood) tend to initiate leaf production (green-up) before rainfall has commenced, and persist after the rainy season has finished, despite having lower overall GPP, while more resource acquisative species and juvenile individuals tend to green-up during the rainy season, and create a dense leaf-flush during the mid-season peak of growth (). It has been suggested that this variation in leaf phenological activity between species is one aspect by which increased tree species richness causes an increase in ecosystem-level productivity in deciduous savannas (). Building on research linking biodiversity and ecosystem function, one might expect that an ecosystem with a greater diversity of tree species might be better able to maintain consistent leaf coverage for a longer period over the year, as species vary in their optimal growing conditions due to niche complementarity, whereby coexisting species vary in their occupation of niche space due to competitive exclusion ().

In the water-limited savannas such as those found in large areas of southern Africa (), the ability of conservative tree species to maintain consistent leaf coverage in the upper canopy strata over the growing season, but particularly at the start and end of the growing season, may provide facilitative effects to other tree species and juveniles occupying lower canopy strata that are less

well-adapted to moisture-limiting conditions, but are more productive, by providing shade and influencing below ground water availability through hydraulic lift ().

Variation in tree species composition, as well as species richness, is also expected to have an effect on savanna phenology in southern Africa. Savannas of a number of different types (species composition and structure) are found across southern Africa, but these are often poorly differentiated in regional-scale phenological studies (), resulting in a dearth of information on the phenological behaviour of different woodlands. As our ability to remotely sense tree species composition improves, it allows us to create more tailored models of the carbon cycle which incorporate not only climatic factors, but also biotic factors which govern productivity. We therefore need to understand how species composition and biodiversity metrics affect land-surface phenology.

In the deciduous woodlands of Zambia, a highly pronounced single wet-dry season annual oscillation is observed across the majority of land area, with local exceptions in some mountainous areas (). Variation in leaf phenological activity across the country has a large influence on annual gross primary productivity. Using Zambia as a case study, we can expect similar response from deciduous woodlands across southern Africa, with important consequences for the global carbon cycle ().

While cumulative leaf production across the growing season may be the most important aspect of leaf phenology for GPP, other phenological metrics may be more important for ecosystem function and habitat provision for wildlife. Periods of green-up and senescence which bookend the growing season are key times for invertebrate reproduction (), soil biotic activity () and herbivore browsing activity (). Pre-rainy season green-up in water-limited savannas provides a valuable source of moisture and nutrients before the rainy season, and can moderate the understorey microclimate, increasing humidity, reducing UV exposure, and moderating diurnal oscillations in temperature, reducing ecophysiological stress which can lead to mortality during the dry season. An increase in the time between leading tree growth and the onset of seasonal rains provides a buffer to stressful dry season climatic conditions and wildlife activity. A slower rate of green-up caused by tree species greening at different times provides an extended period of bud-burst, thus maintaining the important food source of nutrient rich young leaves for longer ().

In this study we contend that, across Zambian deciduous savannas, tree species diversity and composition influence three key measurable aspects of the tree phenological cycle: (1) the rates of greening and senescence at the start and end of the seasonal growth phase, (2) the overall length of the growth period, and (3) the lag time between green-up/senescence and the start/end of the rainy season. It is hypothesised that: (H_1) due to variation among species in minimum viable water availability for growth, plots with greater tree species richness will exhibit slower rates of greening and senescence as different species green-up and senesce at different times. We expect that: (H_2) in plots with greater species richness the start of the growing season will occur earlier in respect to the onset of rain due to an increased likelihood of containing a species which can green-up early, facilitating other species to initiate the growing season. We hypothesise that: (H_3) plots with greater species richness will exhibit a longer growth period and greater cumulative green-ness 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: (H_4) irrespective of species diversity, variation in tree species composition and vegetation type will

cause variation in the phenological metrics outlined above.

2 Materials and methods

2.1 Data collection

We used plot-level data on tree species diversity across 704 sites from the Zambian Integrated Land Use Assessment Phase II (ILUA-II), conducted in 2014 (Mukosha and Siampale, 2009; Pelletier et al., 2018). Each site consisted of four 20x50 m (0.2 ha) plots positioned in a square around a central point, with a distance of 500 m between each plot (Figure 2). The original census contained 993 sites, which was filtered in order to define study bounds and to ensure data quality. Only sites with $\geq$50 stems ha^{-1} $\geq$10 cm DBH (Diameter at Breast Height) were included in the analysis, to ensure all sites represented woody savanna rather than 'grassy savanna', which is considered a separate biome with very different species composition and ecosystem processes governing phenology (Parr et al., 2014). Sites in Mopane woodland were removed by filtering sites with greater than 50% of individuals belonging to *Colophospermum mopane*, preserving only plots with Zambesian tree savanna / woodland. To eliminate compositional outliers, plots with fewer than five species with more than one individual were excluded. Plots dominated by non-native tree species ($\geq$50% of individuals), e.g. *Pinus* spp. and *Eucalyptus* spp. were also excluded, as these species may exhibit non-seasonal patterns of leaf production ().

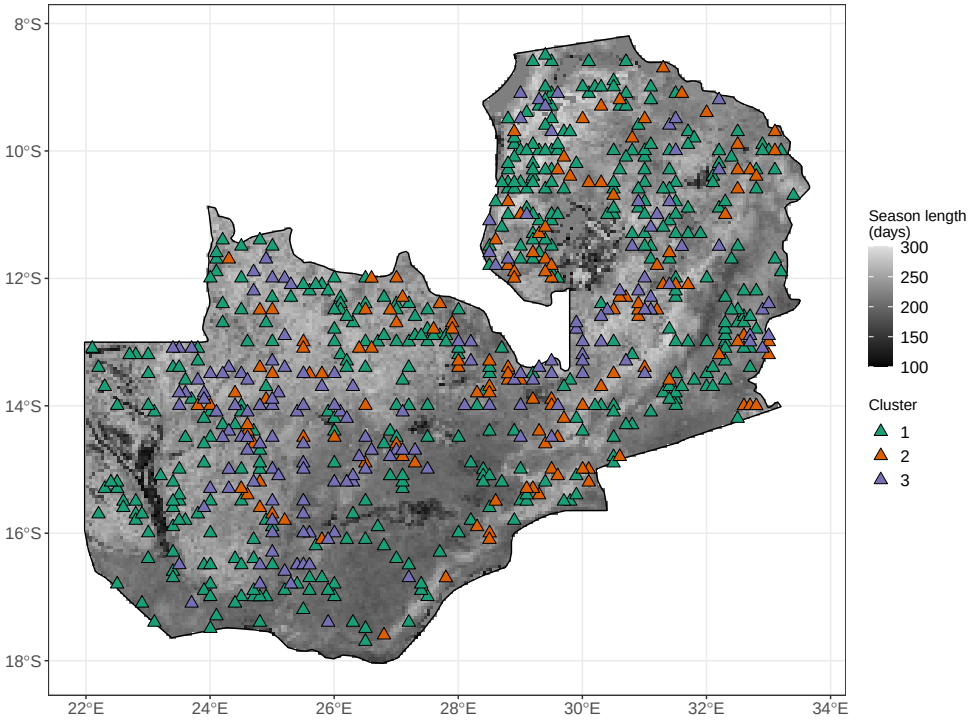

Figure 1: Distribution of study sites within Zambia as triangles, each consisting of four plots. Sites are oloured according to vegetation compositional cluster as identifed by Ward's clustering algorithm on euclidean distance of plots in the first two axes of NSCA ordination space. Zambia is shaded according to growing season length as estimated by the MODIS VIPPHEN-EVI2 product, at 0.05˚spatial resolution (Didan and Barreto, 2016).

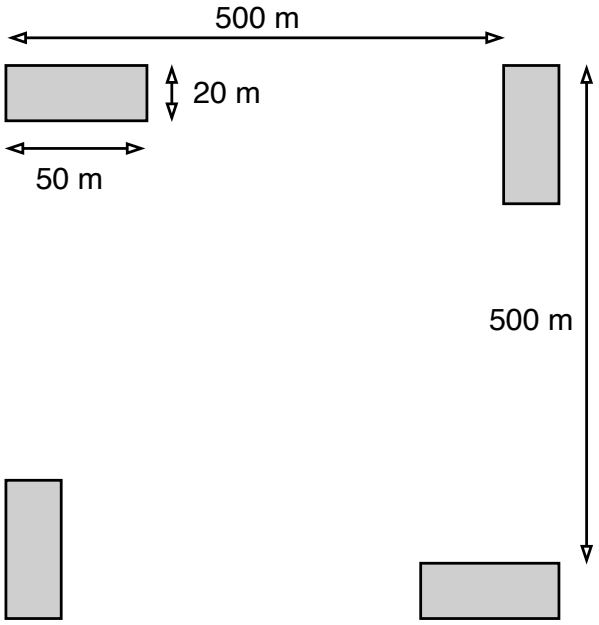

Figure 2: Schematic diagram of plot layout within a site. Each 20x50 m (0.2 ha) plot is shaded grey. The site centre is denoted by a circle. Note that the plot dimensions are not to scale.

Within each plot, the species of all trees with at least one stem ≥10 cm DBH were recorded. Plot data was aggregated to the site level for analyses to avoid pseudo-replication caused by the more spatially coarse phenology data. Tree species composition varied little among the four plots within a site, and were treated as representative of the woodland in the local area. Using the Bray-Curtis dissimilarity index of species abundance data, we calculated that the mean pairwise compositional distance between plots within a site was lower than the mean compositional distance across all pairs of plots in 93.8% of cases.

To quantify phenology at each site, we used the MODIS MOD13Q1 satellite data product at 250 m resolution (Didan, 2015). The MOD13Q1 product provides an Enhanced Vegetation Index (EVI) time series at 16 day intervals. EVI is widely used as a measure of vegetation growth, as an improvement to NDVI (Normalised Differential Vegetation Index), which tends to saturate at higher values. EVI is well-correlated with gross primary productivity and so can act as a suitable proxy (). We used all scenes from January 2015 to August 2020 with less than 20% cloud cover covering the study area. All sites were determined to have a single annual growth season according to the MODIS VIPPHEN product (), which assigns pixels (0.05°, 5.55 km at equator) up to three growth seasons per year. We stacked yearly data between 2015 and 2020 and fit a General Additive Model (GAM) to produce an average EVI curve. We estimated the start and end of the growing season using first derivatives of the GAM. Start of the growing season was identified as the first day where the model slope exceeds half of the maximum positive model slope for a continuous period of 20 or more days, following **White2009**. Similarly, we defined the end of the growing season as the final day of the latest 20 period where the GAM slope meets or exceeds half of the maximum negative slope. We estimated the length of the growing season as the number of days between the start and end of the growing season. We estimated the green-up rate as the slope of a linear model across EVI values between the start of the growing season and the point at which the slope of reduces below half of the maximum positive slope. Similarly the senescence rate was estimated as the slope of a linear model between the latest point where the slope of decrease fell below half of the maximum negative slope and the end of the growing season Figure 3. We validated our calculations of cumulative EVI, mean annual EVI, growing season length, season start date, season end date, green-up rate and senescence rate with calculations made by the MODIS VIPPHEN product with linear models comparing the two datasets across our study sites (Figure S1, Table S1). Sites where our calculation of a phenological metric was drastically different to the MODIS VIPPHEN estimate were excluded, under the assumption that our algorithm had failed to capture the true value or some site specific factor precluded precise estimation. This removed 8 sites.

Precipitation data was gathered using the "GPM IMERG Final Precipitation L3 1 day V06" dataset, which has a pixel size of 0.1°(11.1 km at the equator) (**GPM**), between 2015 and 2020. Daily total precipitation was separated into two periods: precipitation during the growing season (growing season precipitation), and precipitation in the 90 day period before the onset of the growing season (dry season precipitation). Rainy season limits were defined as for the EVI data, using the first derivative of a GAM to create a curve for each site using stacked yearly precipitation data, from which we estimated the half-max positive and negative slope to identify where the GAM model exceeded these slope thresholds for a consistent period of 20 days or more. Mean diurnal

temperature range (Diurnal δT) was calculated as the mean of monthly temperature range from the WorldClim database, using the BioClim variables, with a pixel size of 30 arc seconds (926 m at the equator) (Fick and Hijmans, 2017). averaged across all years of available data (1970-2000).

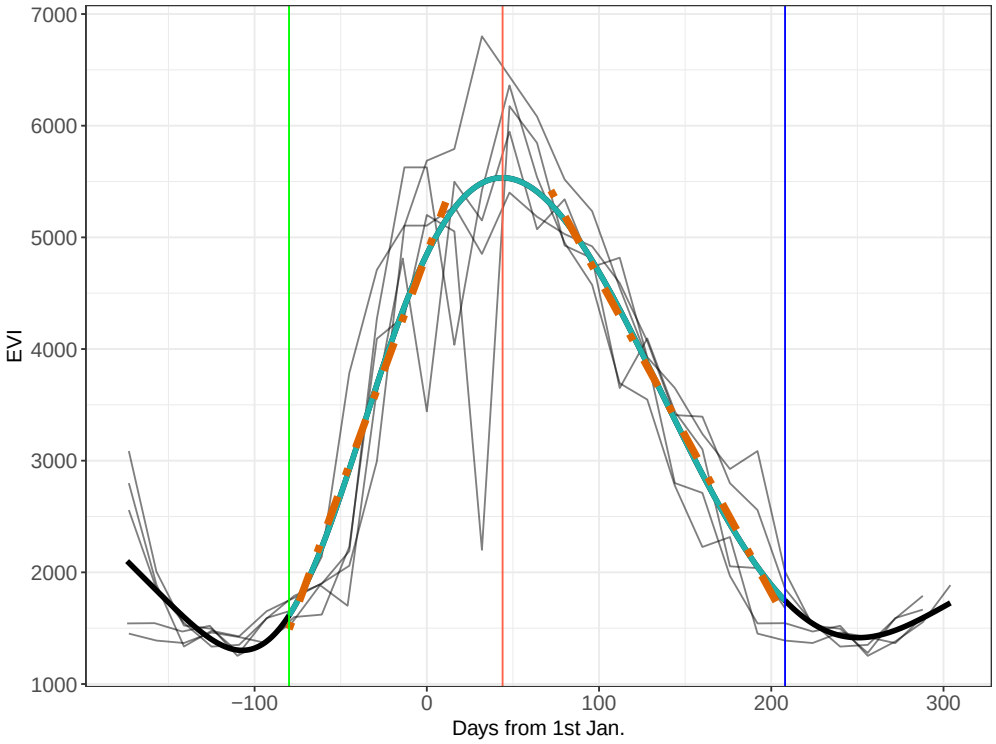

Figure 3: Example EVI time series, demonstrating the metrics derived from it. Thin black lines show the raw EVI time series, with one line for each annual growth season. The thick black line shows the GAM fit. The thin blue lines show the minima which bound the growing season. The red line shows the maximum EVI value reached within the growing season. The shaded cyan area of the GAM fit shows the growing season, as defined by the first derivative of the GAM curve. The two orange dashed lines are linear regressions predicting the green-up rate and senescence rate at the start and end of the growing season, respectively. Note that while the raw EVI time series fluctuate greatly around the middle of the growing season, mostly due to cloud cover, the GAM fit effectively smooths this variation to estimate the average EVI during the mid-season period.

2.2 Data analysis

To measure variation in tree species composition we a used combination of Non-symmetric Correspondence Analysis (NSCA) and agglomerative hirerarchical clustering on species abundance data (**Kreft2010**; **Fayolle2014**). NSCA was performed using the `ade4` R package (Dray and Dufour, 2007). Scree plot analysis demonstrated that 2 axes was optimal to describe our data. These axes accounted for 17.4% of the variance in species composition according to eigenvalue decay. To guard against sensitivity to rare individuals, which can preclude meaningful cluster delineation across such a large species compositional range, we restricted the NSCA to species with greater than five records, and to sites with fewer than five species (). We used Ward's algorithm to define clusters (**Murtagh2014**), based on the euclidean distance of sites in NSCA ordination space. We

determined the optimal number of clusters by maximising the mean silhouette width among clusters (**Rousseeuw1987**) Figure S2. Vegetation type clusters were used later as interaction terms in linear models. We described the vegetation types represented by each of the clusters using a Dufrene-Legendre indicator species analysis (**Dufrene1997**).

We specified multivariate linear models to assess the role of tree species diversity on each of the chosen phenological metrics. We defined tree species diversity using both species richness and abundance evenness as separate independent variables. Abundance evenness was calculated as the Shannon Equitability index ($E_{H'}$) (**Smith1996**) was calculated as the ratio of the Shannon diversity index to the natural log of species richness. We defined a maximal model structure including tree species richness, abundance evenness, the interaction of species richness and vegetation type, and climatic variables shown by previous studies to strongly influence phenology. The quality of the maximal model was compared to models with different subsets of independent variables using the model log likelihood, AIC (Akaike Information Criteria), BIC (Bayesian Information Criteria), and adjusted R^2 values for each model. For each phenological metric, the best model according to the model quality statistics is reported in the results. Where two similar models were within 2 AIC points of each other, the model with fewer terms was chosen as the best model, to maximise model parsimony. All models were fitted using Maximum Likelihood (ML) to allow comparison of models (). The best model was subsequently re-fitted using Restricted Maximum Likelihood for model effect estimation (REML). Independent variables in each model were transformed to achieve normality where necessary and standardised to Z-scores prior to modelling to allow comparison of slope coefficients within a given model.

We used the `ggeffects` package to estimate the marginal means of the interaction effect of species richness and vegetation type, to investigate vegetation type specific effects on each phenological metric (**ggefects**). Estimated marginal means entails generating model predictions across values of a focal variable, in this case species richness, while holding non-focal variables constant. All statistical analyses were conducted in R version 4.0.2 (R Core Team, 2020).

3 Results

Model selection showed that richness and evenness are important determinants of each of the chosen phenological metrics, across vegetation types. The effect of species richness featured and significant in all best models except for senescence lag, while evenness was a significant effect in models for cumulative EVI, season length and senescence lag only Figure 4.

3 vegetation type clusters were identified during hierarchical clustering. Cluster 1, which contains the most sites (416), consists of small stature Zambesian woodlands, as defined by **Olson**, and is not dominated by a particular large canopy tree species. Abundance evenness is high across sites in Cluster 1. Cluster 2 is dominated heavily by *Brachystegia boehmii*, while Cluster 3 is dominated by *Julbenardia paniculata*, both large canopy forming trees. These two clusters likely represent variation among miombo woodland types in dominant canopy tree species. Both Clusters 2 and 3 have a similar composition of non-dominant smaller shrubby species, such as *Pseudolachnostylis maprouneifolia* (Table 1).

As expected (H_3), species richness and wet season precipitation both had positive significant effects on cumulative EVI and season length. In contrast, abundance evenness, the other aspect of tree species diversity in our models, had a significant negative effect on both cumulative EVI and season length (Figure 4).

Species richness caused a significant increase in the lag time between date of green-up and date of rainy season onset (H_2). This effect was comparable to the effects of pre-season precipitation and diurnal temperature range, which also caused an increase in green-up lag. In contrast, senescence lag was poorly defined by our models, suggesting that some unmeasured factor remains the key driver of this phenological metric. The effects of diurnal δT and abundance evenness had wide confidence interval. The best model explained only 1% of the variance in senescence lag, though was still better quality than a climate-only model.

All best models including tree species diversity variables were of better quality than models which included only climatic variables Table 2. The phenological metrics best predicted were green-up lag and cumulative EVI, where models explained 26% and 34% of the variance in these variables, respectively. Senescence rate and senescence lag were the least well predicted phenological metrics, with the best model explaining 3% and 2% of their variance, respectively.

While species richness had a significant negative effect on green-up rate, as predicted by H_1, the best model, which also included pre green-up precipitation and diurnal temperature range, only explained 10% of the variance in this metric.

The slope of the relationship between species richness and phenological metrics varied among vegetation types, but maintained the same direction in all cases Figure 5. Clusters 1 and 2 tended to show similar responses to variation in species richness across phenological metrics. Across all models however, none of the vegetation types were significantly different, according to post-hoc Tukeys's tests on marginal effects (Table S8). Clusters were largely similar in their density distribution of the six phenological metrics Figure 7. The most striking differences are the non-significant () increase in mean green-up lag in Cluster 1 compared to the other 2 clusters, and the presence of some sites in Cluster 1 with particularly high green-up rates. The hierarchical clustering analysis demonstrated that there was little spatial structure to the vegetation clusters identified. The key emergent trend was that Cluster 2 was absent from the southwest of the country (Figure 1) possibly due to the low levels of precipitation in this region, which could preclude many miombo tree species.

Cluster	N sites	Richness	MAP	Diurnal δT	Species	Indicator value
1	416	15(7)	1040(199.8)	14(1.6)	*Pterocarpus angolensis*	0.294
					Diplorhynchus condylocarpon	0.265
					Brachystegia spiciformis	0.252
2	135	16(5)	1051(165.1)	13(1.5)	*Brachystegia boehmii*	0.795
					Psuedolachnostylis maprouneifolia	0.240
					Uapaca kirkiana	0.224
3	153	15(7)	989(153.1)	14(1.4)	*Julbernardia paniculata*	0.717
					Psuedolachnostylis maprouneifolia	0.272
					Diplorhynchus condylocarpon	0.228

Table 1: Climatic information and Dufrene-Legendre indicator species analysis for the vegetation type clusters identified by the PAM algorithm. 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 Diurnal δT are reported as the mean and 1 standard deviation in parentheses. Species richness is reported as the median and the interquartile range in parentheses.

Response	δAIC	δBIC	R^2_{adj}	δlogLik
Cumulative EVI	43.1	34.0	0.34	-23.54
Season length	25.9	21.4	0.17	-13.97
Green-up rate	8.1	3.6	0.10	-5.07
Senescence rate	10.1	1.0	0.03	-7.04
Green-up lag	67.7	63.2	0.26	-34.87
Senescence lag	7.8	7.8	0.02	-3.88

Table 2: Model fit statistics for each phenological metric.

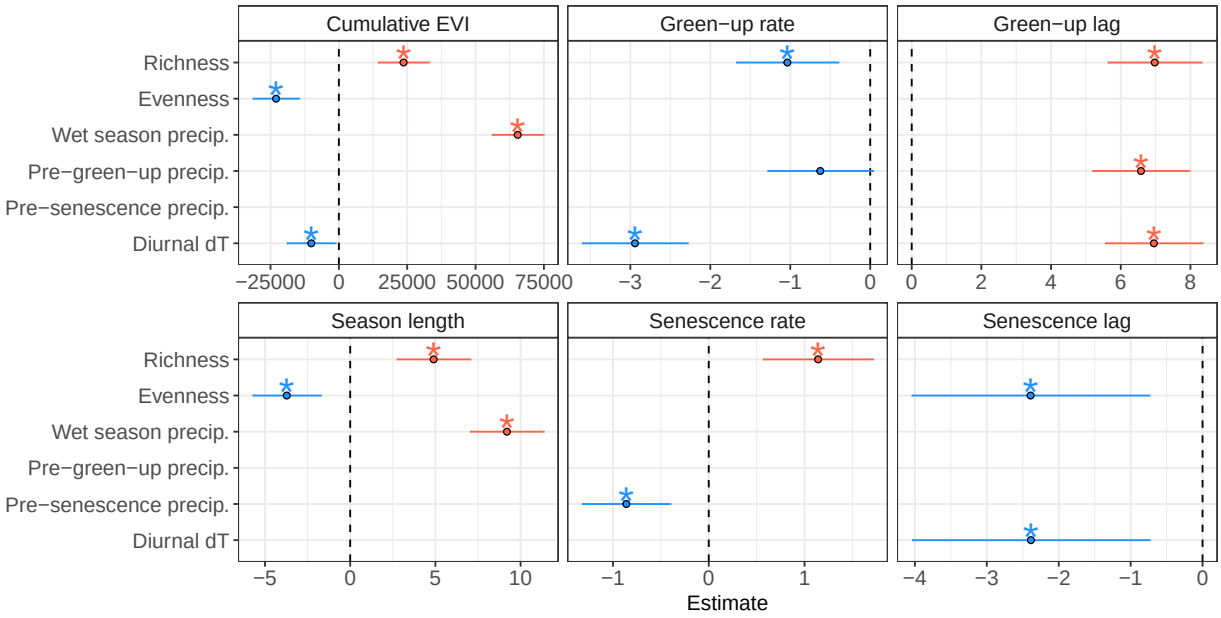

Figure 4: Standardized slope coefficients for each best model of a phenological metric. Slope estimates are ±1 standard error. Slope estimates where the interval (standard error) does not overlap zero are considered to be significant effects.

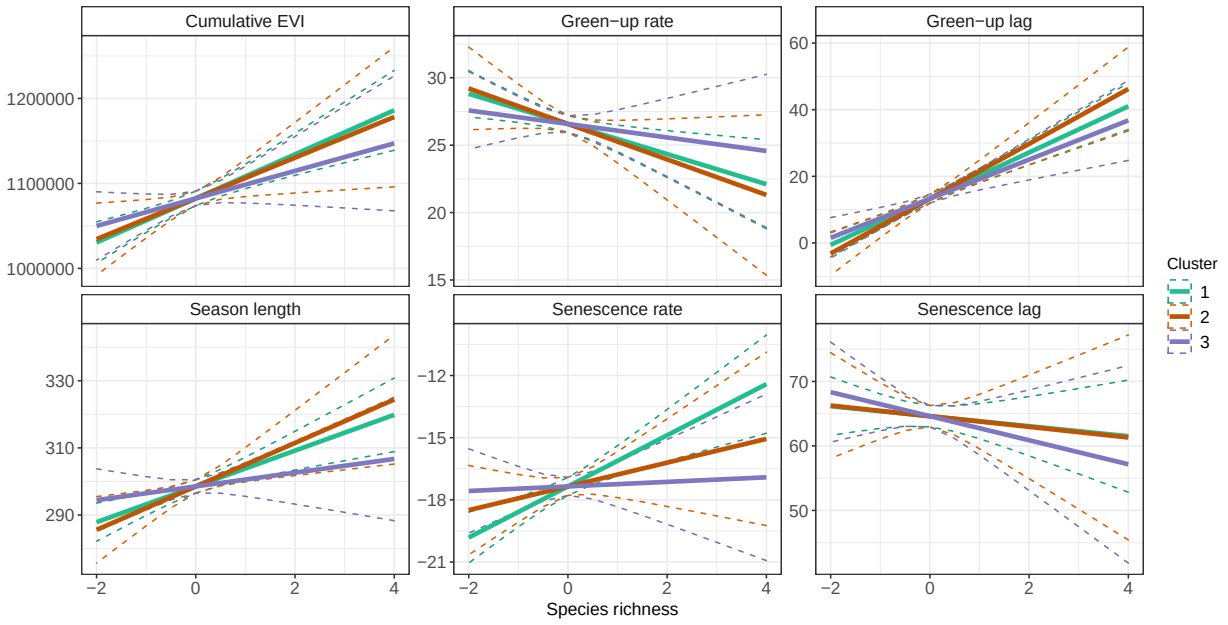

Figure 5: Marginal effects of tree species richness on each of the phenological metrics, for each vegetation type, using the best model including the interaction of species richness and vegetation cluster, for each phenological metric.

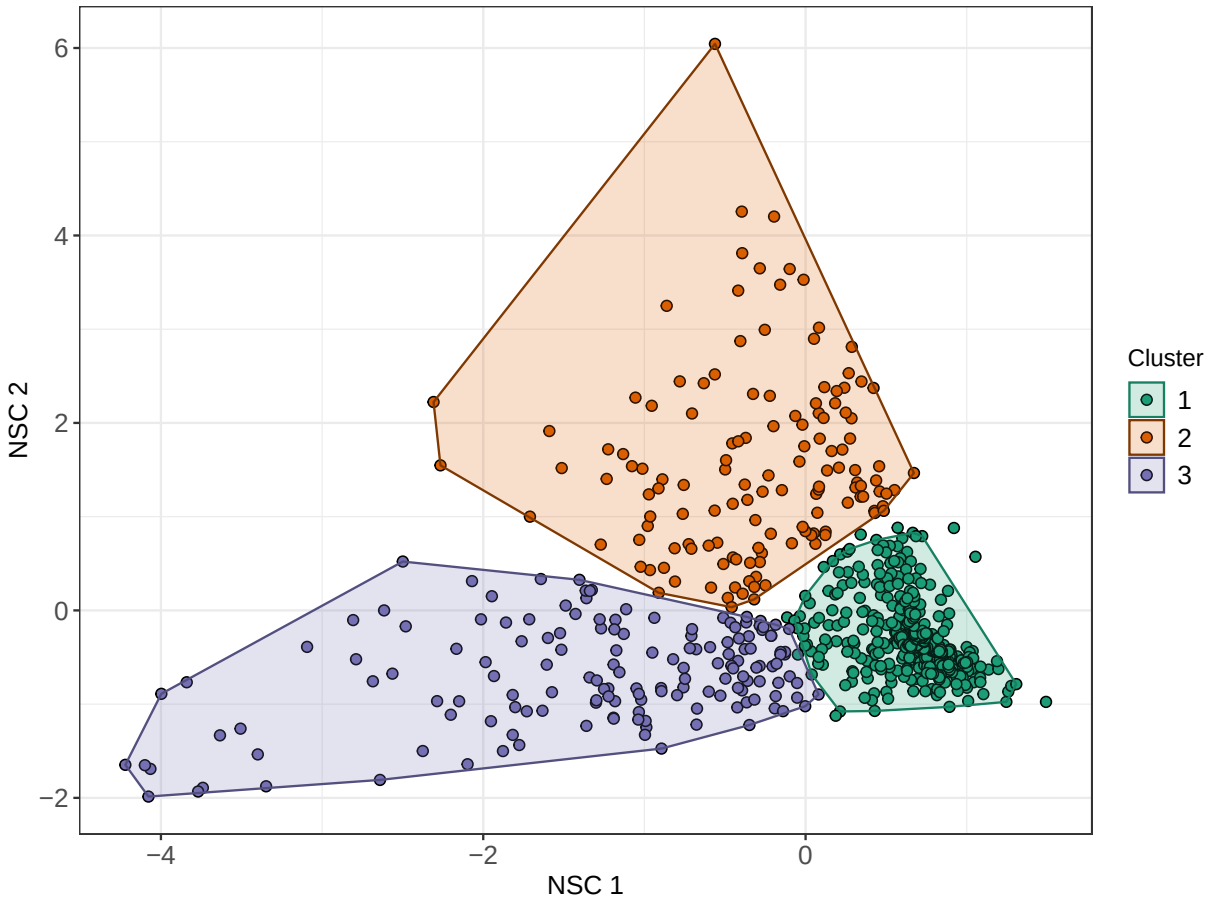

Figure 6: Plot scores of the (A) first and second, and (B) third and fourth axes of the Non-Symmetric Correspondence Analysis of tree species composition. Points are coloured according to clusters defined by Ward's algorithm on euclidean distances of the NSCA ordination axes, along with a convex hull encompassing 95% of the points in each cluster.

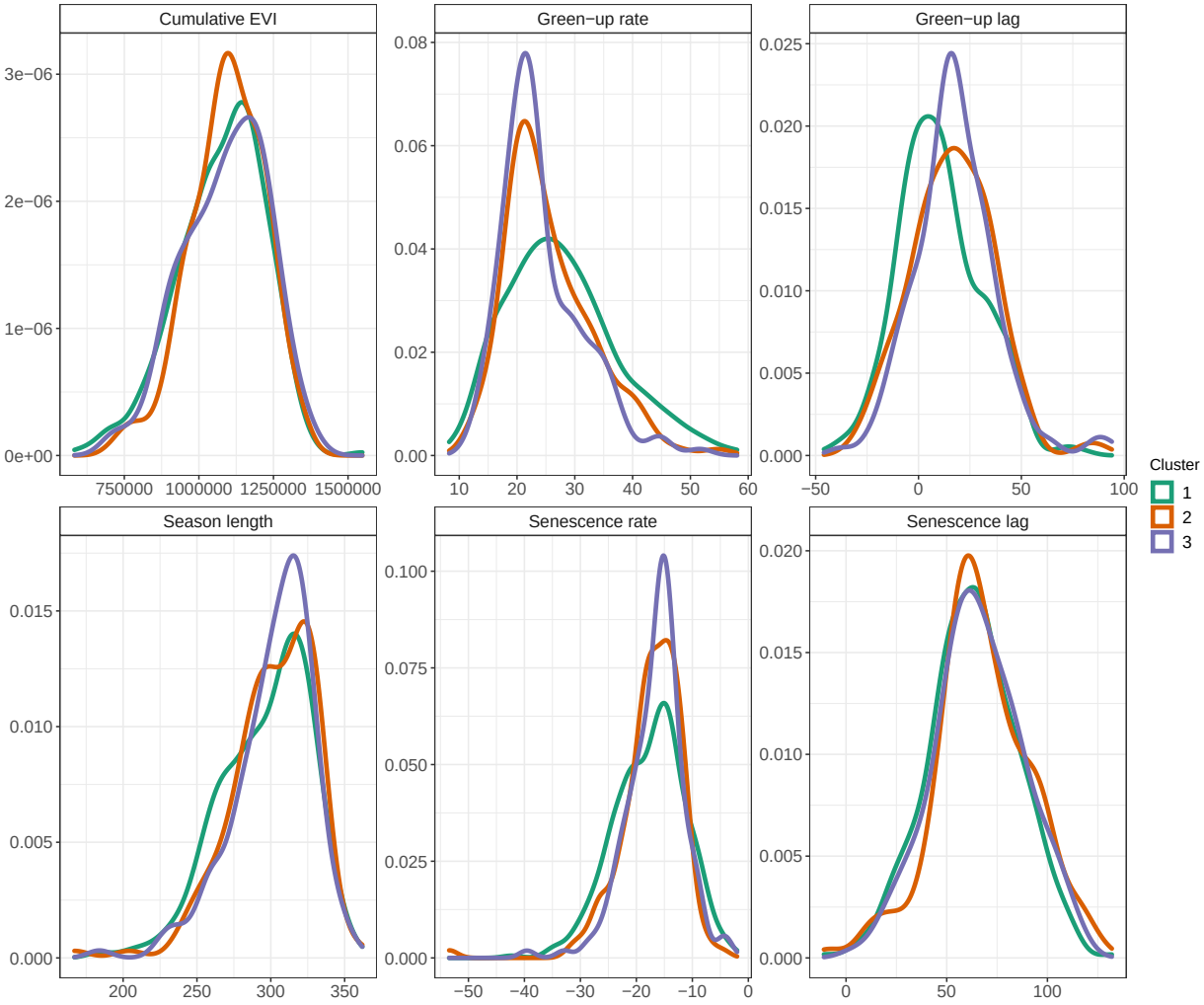

Figure 7

4 Discussion

In this study we have demonstrated a clear and measurable effect of tree species richness across various aspects of land-surface phenology in Zambian deciduous savannas. We showed that tree species richness led to an increase in cumulative EVI and season length. Additionally, species richness led to a slower rate of greening and caused the onset of greening to occur earlier with respect to the start of the rainy season. Our study lends support for a positive biodiversity - ecosystem function relationship in our chosen study area, operating through its influence on phenology. Our results exemplify the key role of tree species biodiversity in driving key ecosystem processes, which affect ecosystem structure, the wildlife provisioning role, and the gross primary productivity of ecosystems.

As methods for the remote sensing of tree species diversity mature, our finding that species richness strongly affects patterns of land-surface phenology should prompt earth system modellers to take advantage of remotely sensed tree species diversity data to create tailored models of the carbon cycle which incorporate not only climatic factors, but also biotic factors which govern productivity in deciduous savannas.

Patterns of senescence were poorly predicted by species richenss and evenness in our models. **Cho2017** found that tree cover, measured by MODIS LAI data, had a significant effect on senescence rates in savannas in South Africa, which have similar climatic conditions to the sites in our study. Other studies both global and within southern African savannas have largely ignored patterns of senescence, instead focussing patterns of green-up (). Most commonly, these studies simply correlate the decline of rainfall with senescence, but the lack of precipitation as a term in our best model suggests that other unmeasured factors are at play. While leaf senescence is not as important for the survival of browsing herbivores as green-up, the timing of senescence with respect to temperature and precipitation has important consequences for the savanna understorey microclimate. The longer leaf material remains in the canopy after the end of the rainy season, the greater the microclimatic buffer for herbaceous understorey plants and animals, which require water and protection from high levels of insolation and dry air which can prevail rapidly after the end of the rainy season (). Our study merely exemplifies that more work needs to be done to properly characterise the drivers of senescence in this biome.

While species richness is a common measure of biodiversity, abundance evenness constitutes a second key axis (**Wilsey2005**; **Hillebrand2008**). While traditionally species richness and evenness were assumed to be highly positively correlated, recent work has demonstrated that in many systems, richness and evenness may be nearly orthogonal (). In this study, we found contrasting effects of richness and evenness on both cumulative EVI and season length. Evenness caused a decrease in these phenological metrics, which we did not expect. It is possible that the negative effect of abundance evenness occurred because an increase in evenness is associated with a reduction in the canopy cover of a few highly dominant large canopy tree species (e.g. *Brachystegia boehmii* and *Julbenardia paniculata*), as part of the transition from woody savanna to thicket vegetation, or following a major disturbance event. Large canopy tree species have access to ground water for a longer part of the year, due to their deep root systems and conservative growth patterns. A future study may choose to explore the differential effects of species diversity in different size classes and in different physiognomic groups defined by functional form.

Our coverage of very short season lengths in Zambia, as estimated by the VIPPHEN product, was restricted, with notable absences of plot data in the northeast of the country around 30.5°E, 11.5°S, and 23.0°E, 15.0°S. Upon further inspection of true colour satellite imagery, these regions are largely seasonally water-logged floodplain and swampland, and were likely ignored by the ILUA-II assessment for this reason. This also explains their divergent phenological patterns.

It is important to note that the remotely sensed EVI measurements used here don't refer only to trees, they represent the landscape as a singe unit. Nevertheless, seasonal patterns of tree leaf phenology in southern African deciduous woodlands, particularly the pre-rainy season green-up phenomenon, is driven almost exclusively by trees, while grasses tend to follow patterns of precipitation more closely (). Grasses contribute to gross primary productivity, and it was therefore in our interests to include their response in our analysis as we seek to demonstrate how tree species richness can affect cycles of carbon exchange. Additionally, the micro-climatic effects of tree leaf canopy coverage and hydraulic lift through tree deep root systems will benefit the productivity of grasses as well as understorey tree individuals.

5 Conclusion

Here we explored the role of tree species diversity on land surface phenology across Zambia. We showed that species richness clearly affects rate of green-up, the lag time between rainy season onset and growth, and the length of the growing season. Our results have a range of consequences for earth system mdoellers and conservation managers, and lend further support to an already well established corpus of the positive effect of species diversity on ecosystem function.

References

Adole, Tracy, Jadunandan Dash, and Peter M. Atkinson (2018). "Large-scale prerain vegetation green-up across Africa". In: *Global Change Biology* 24.9, pp. 4054–4068. DOI: 10.1111/gcb.14310.

Didan, L. (2015). *MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006 [Data set]*. NASA EOSDIS Land Processes DAAC. DOI: 10.5067/MODIS/MOD13Q1.006. (Visited on 08/05/2020).

Didan, L. and A. Barreto (2016). *NASA MEaSUREs Vegetation Index and Phenology (VIP) Phenology EVI2 Yearly Global 0.05Deg CMG [Data set]*. NASA EOSDIS Land Processes DAAC. DOI: 10.5067/MEaSUREs/VIP/VIPPHEN_EVI2.004. (Visited on 08/05/2020).

Dray, Stéphane and Anne–Béatrice Dufour (2007). "The ade4 Package: Implementing the Duality Diagram for Ecologists". In: *Journal of Statistical Software* 22.4, pp. 1–20. DOI: 10.18637/jss.v022.i04.

Fick, S. E. and R. J. Hijmans (2017). "WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas". In: *International Journal of Climatology* 37.12, pp. 4302–4315. DOI: http://dx.doi.org/10.1002/joc.5086.

Mukosha, J and A Siampale (2009). *Integrated land use assessment Zambia 2005–2008*. Lusaka, Zambia: Ministry of Tourism, Environment et al.

Parr, C. L. et al. (2014). "Tropical grassy biomes: misunderstood, neglected, and under threat". In: *Trends in Ecology and Evolution* 29, pp. 205–213. DOI: 10.1016/j.tree.2014.02.004Âă.

Pelletier, J. et al. (2018). "Carbon sink despite large deforestation in African tropical dry forests (miombo woodlands)". In: *Environmental Research Letters* 13, p. 094017. DOI: 10.1088/1748-9326/aadc9a.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

6 Supplementary Material

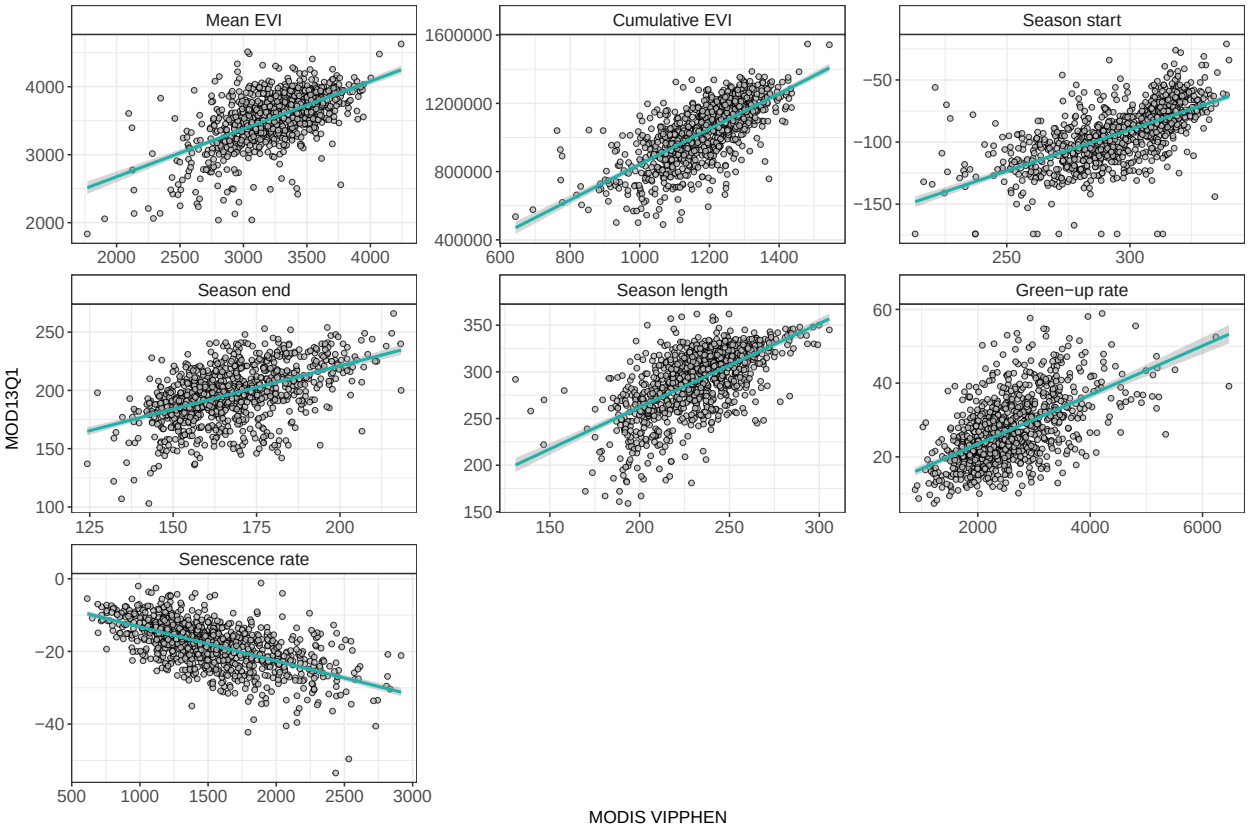

Figure S1

Response	DoF	F	Prob.	R^2
Mean EVI	960	521.3	p<0.05	0.35
Cumulative EVI	960	966.1	p<0.05	0.50
Season start	956	679.6	p<0.05	0.42
Season end	960	337.5	p<0.05	0.26
Season length	960	586.4	p<0.05	0.38
Green-up rate	960	432.4	p<0.05	0.31
Senescence rate	960	541.3	p<0.05	0.36

Table S1: Model fit statistics for comparison of MODIS VIPPHEN and MOD13Q1 products across each of our study sites.

Rank	Precipitation	Diurnal dT	Evenness	Richness	Richness:Cluster	logLik	AIC	ΔIC	W_i
1	✓	✓	✓	✓		**-9198**	**18410**	**0**	**0.681**
2	✓		✓	✓		-9200	18413	3	0.157
3	✓	✓	✓	✓	✓	-9198	18413	3	0.135
4	✓		✓	✓	✓	-9200	18416	6	0.027
5	✓	✓	✓			-9210	18433	23	0.000
6	✓	✓		✓		-9212	18435	25	0.000
7	✓			✓		-9214	18437	27	0.000
8	✓	✓		✓	✓	-9211	18438	28	0.000
9	✓		✓			-9214	18438	28	0.000
10	✓			✓	✓	-9213	18440	30	0.000

Table S2: Cumulative EVI model selection candidate models, with fit statistics. The overall best model is marked by bold text, while the best model with a richness:cluster interaction term is marked by underlined text

Rank	Precipitation	Diurnal dT	Evenness	Richness	Richness:Cluster	logLik	AIC	ΔIC	W_i
1	✓	✓	✓	✓		-3312	6639	0	0.436
2	✓		✓	✓		**-3314**	**6640**	**1**	**0.310**
3	✓	✓	✓	✓	✓	-3312	6641	2	0.130
4	✓		✓	✓	✓	-3313	6641	3	0.120
5	✓	✓		✓		-3319	6650	11	0.002
6	✓			✓		-3320	6651	12	0.001
7	✓	✓		✓	✓	-3317	6651	12	0.001
8	✓			✓	✓	-3319	6651	12	0.001
9	✓		✓			-3324	6657	18	0.000
10	✓	✓	✓			-3323	6658	19	0.000

Table S3: Season length model selection candidate models, with fit statistics. The overall best model is marked by bold text, while the best model with a richness:cluster interaction term is marked by underlined text

Rank	Precipitation	Diurnal dT	Evenness	Richness	Richness:Cluster	logLik	AIC	ΔIC	W_i
1	✓	✓	✓	✓		-2490	4994	0	0.297
2	✓	✓		✓		**-2491**	**4995**	**0**	**0.261**
3		✓		✓		-2493	4996	2	0.127
4		✓	✓	✓		-2492	4996	2	0.124
5	✓	✓	✓	✓	✓	-2490	4997	3	0.069
6	✓	✓		✓	✓	-2491	4998	4	0.051
7		✓	✓	✓	✓	-2492	4999	5	0.029
8		✓		✓	✓	-2493	4999	5	0.026
9	✓	✓	✓			-2495	5002	7	0.009
10	✓	✓				-2496	5003	8	0.004

Table S4: Green-up rate model selection candidate models, with fit statistics. The overall best model is marked by bold text, while the best model with a richness:cluster interaction term is marked by underlined text

Rank	Precipitation	Diurnal dT	Evenness	Richness	Richness:Cluster	logLik	AIC	ΔIC	W_i
1	✓	✓		✓	✓	**-2240**	**4497**	**0**	**0.277**
2	✓	✓		✓		-2242	4497	0	0.229
3	✓	✓	✓	✓	✓	-2240	4498	1	0.133
4	✓	✓	✓	✓		-2242	4498	2	0.122
5	✓			✓	✓	-2242	4499	2	0.090
6	✓			✓		-2245	4500	3	0.062
7	✓		✓	✓	✓	-2242	4500	4	0.045
8	✓		✓	✓		-2244	4501	4	0.036
9		✓		✓		-2249	4507	11	0.001
10		✓		✓	✓	-2247	4507	11	0.001

Table S5: Senescence rate model selection candidate models, with fit statistics. The overall best model is marked by bold text, while the best model with a richness:cluster interaction term is marked by underlined text

Rank	Precipitation	Diurnal dT	Evenness	Richness	Richness:Cluster	logLik	AIC	ΔIC	W_i
1	✓	✓		✓		**-3018**	**6049**	**0**	**0.428**
2	✓	✓	✓	✓		-3017	6049	0	0.388
3	✓	✓		✓	✓	-3018	6051	3	0.103
4	✓	✓	✓	✓	✓	-3017	6052	3	0.081
5		✓	✓	✓		-3057	6127	78	0.000
6		✓		✓		-3059	6128	79	0.000
7		✓	✓	✓	✓	-3057	6130	81	0.000
8		✓		✓	✓	-3058	6131	82	0.000
9	✓		✓	✓		-3062	6135	87	0.000
10	✓		✓	✓	✓	-3060	6135	87	0.000

Table S6: Green-up lag model selection candidate models, with fit statistics. The overall best model is marked by bold text, while the best model with a richness:cluster interaction term is marked by underlined text

Rank	Precipitation	Diurnal dT	Evenness	Richness	Richness:Cluster	logLik	AIC	ΔIC	W_i
1		✓	✓			**-3184**	**6377**	**0**	**0.318**
2		✓	✓	✓		-3183	6378	1	0.226
3	✓	✓	✓	✓		-3182	6379	1	0.153
4	✓	✓	✓			-3183	6379	2	0.149
5		✓	✓	✓	✓	-3183	6382	4	0.035
6	✓	✓	✓	✓	✓	-3182	6382	5	0.023
7			✓			-3188	6383	6	0.016
8		✓		✓		-3187	6383	6	0.016
9		✓				-3188	6383	6	0.016
10	✓	✓		✓		-3186	6384	7	0.010

Table S7: Senescence lag model selection candidate models, with fit statistics. The overall best model is marked by bold text, while the best model with a richness:cluster interaction term is marked by underlined text

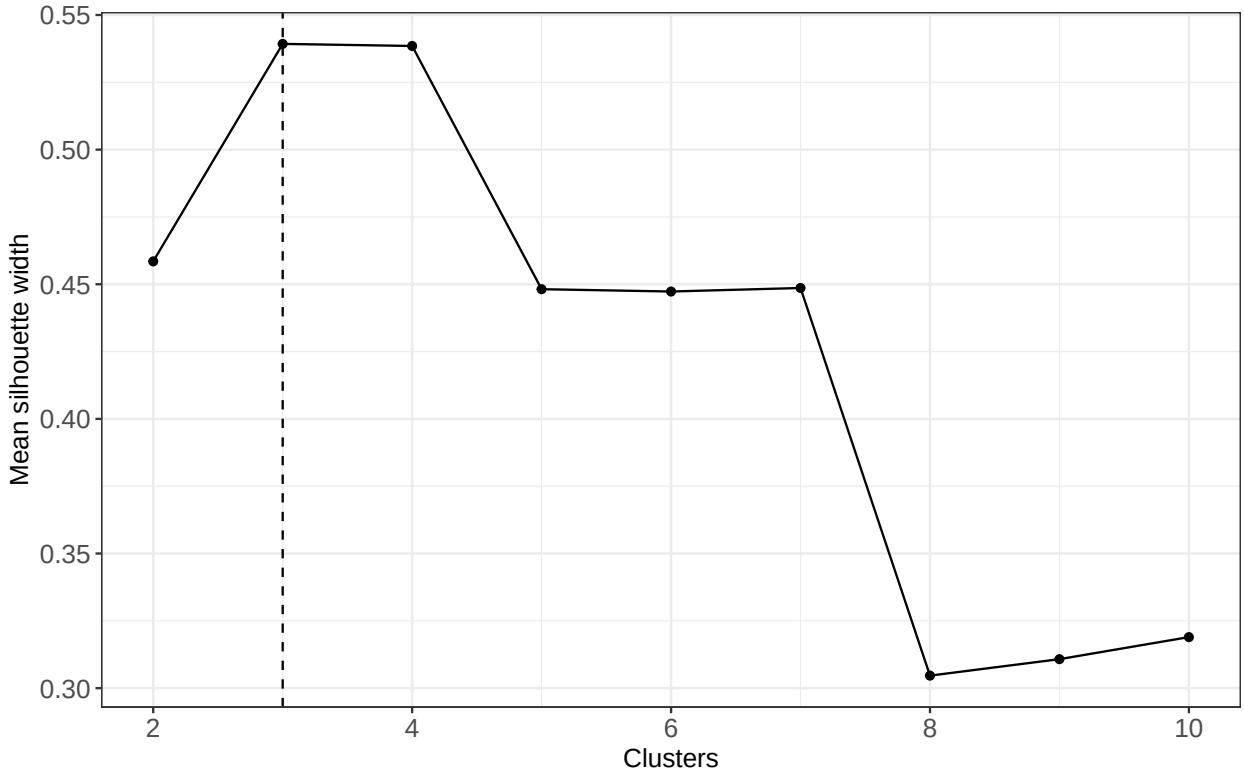

Figure S2

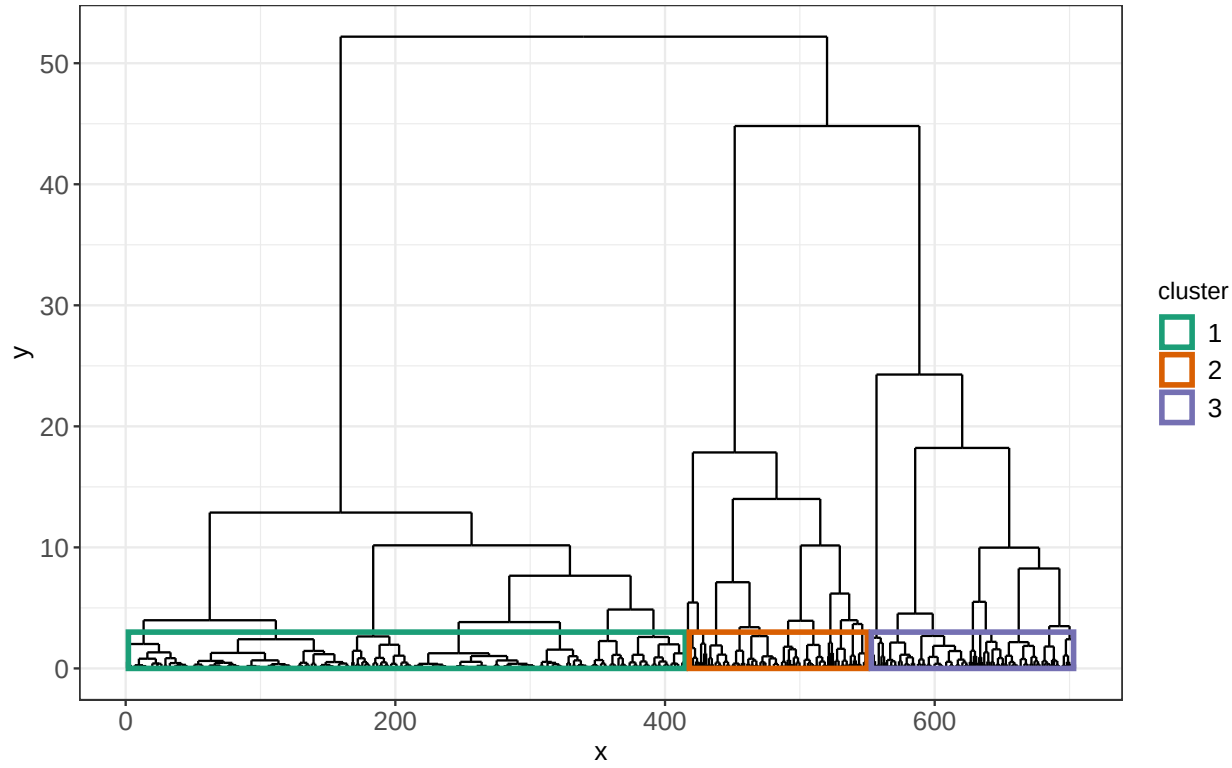

Figure S3

Response	Clusters	Estimate	SE	DoF	T ratio	Prob.
Cumulative EVI	1-2	1.1E-14	6.68E-14	697	0.17	0.98
	1-3	5.5E-14	6.33E-14	697	0.87	0.66
	2-3	4.4E-14	8.16E-14	697	0.54	0.85
Season length	1-2	-6.4E-18	1.56E-17	698	-0.41	0.91
	1-3	1.9E-17	1.48E-17	698	1.26	0.42
	2-3	2.5E-17	1.89E-17	698	1.32	0.38
Green-up rate	1-2	1.1E-18	4.89E-18	698	0.23	0.97
	1-3	-3.5E-18	4.59E-18	698	-0.76	0.73
	2-3	-4.6E-18	5.91E-18	698	-0.78	0.72
Senescence rate	1-2	3.7E-18	3.41E-18	698	1.09	0.52
	1-3	6.3E-18	3.21E-18	698	1.97	0.12
	2-3	2.6E-18	4.14E-18	698	0.63	0.80
Green-up lag	1-2	-7.3E-18	1.03E-17	698	-0.71	0.76
	1-3	6.0E-18	9.71E-18	698	0.62	0.81
	2-3	1.3E-17	1.25E-17	698	1.07	0.54
Senescence lag	1-2	2.9E-19	1.30E-17	698	0.02	1.00
	1-3	6.1E-18	1.23E-17	698	0.50	0.87
	2-3	5.9E-18	1.59E-17	698	0.37	0.93

Table S8: Comparisons of interaction marginal effects using post-hoc Tukey's tests.