Vegetation drivers of moth assemblages within a tropical agricultural landscape

Aura Alonso-Rodriguez, Lily Shapiro and Ben Camber

12/09/2019

Abstract

Introduction

Land use change, and specifically agricultural expansion, is recognized globally as one of the major drivers of biodiversity loss. The expansion of African oil palm is specifically threatening, since the tropical areas suitable for oil palm are particularly rich in biodiversity and endemism. There is vast evidence of the severe impacts of oil palm expansion in Southeast Asia, where more than 85% of oil palm is cultivated, mostly by means of primary and secondary forest deforestation. Yet due to increasing food and biodiesel demands, oil palm is becoming a major threat in the Neotropics as well. Countries like Costa Rica, Colombia, Guatemala and Brazil, are experiencing increasing land conversion to oil palm, turning previously heterogeneous agricultural matrices into extensive monocultures that support significantly less biodiversity.

The expansion of African oil palm plantations in the Neotropics represents an immense challenge to the conservation of tropical biodiversity. In particular, insects in and around these plantations are easily disturbed by associated deforestation and increased use of agrochemicals. Their complex life cycles and trophic specificity are often ill-suited for large spatial-scale changes in canopy cover, plant species diversity, and forest structure. Conservation efforts in these threatened ecosystems can be more efficient and effective by understanding the complex ecological interactions that occur within them. Therefore, isolating important variables such as the most vulnerable life stage of an endangered species, the most ecotoxic pollutants contaminating a wetland, or the least resilient taxon in a disturbed ecosystem is critical to making useful change.

The South Pacific region of Costa Rica has grown in oil palm area during the last two decades. Evidence has already shown significant decreases in amphibians, reptiles, bats and insect diversity within these plantations, when compared to adjacent primary forests. In Alonso-Rodriguez et al. (2017), moth assemblages were seen to severely degrade in oil palm plantations when compared to primary and young secondary forests. Two moth taxa (Geometridae and Erebidae-Arctiinae, Lepidoptera) showed similar distributions along the disturbance gradient, which was partially explained by plant diversity and vegetation structure. However, which vegetation parameters had a bigger influence on observed distribution patterns and how they differ between moth taxa was not further explored by the authors.

Geometridae and Arctiinae share much of their ranges worldwide, but have very different life history strategies. Geometridae larval fitness is largely derived from their highly effective mimicry of leaves and twigs. On the other hand, Arctiinae larvae are often conspicuous or aposematic and have formidable spines or urticating hairs for defense against a wide variety of predators. The adults of these moth taxa have life histories that differ in a similar way; Geometridae are poor fliers and mimic tree bark, while Arctiinae are strong fliers and mimic wasps or possess threatening coloration. Geometridae are also known for preferring woody plant hosts, while there is a wide variety of host plant preferences within Arctiinae. Taking into account the differences in life history between geometrid and arctiine moths, their responses to structural and floristic vegetation features could help us better explore the potential of predicting diversity and turnover rates of these groups in human-dominated landscapes.

In this study, we ask the following questions: 1) Are structural or floristic features superior predictors of moth species diversity within a habitat disturbance gradient? and 2) Are responses similar in the two

moth groups, despite some major differences in life-history traits of these moth taxa? We predict that modelling will reveal that vegetation variables will have drastically different effects on the success of these taxa. Establishing that superficially similar taxa have inherently different responses to disturbance is crucial for informing conservation policy not only in the case of Geometridae and Arctiinae, but in many other cryptic or non-charismatic taxa worldwide.

Methods

Study Site

This study was conducted in the Pacific lowlands of Costa Rica, in the vicinity of La Gamba Tropical Station and the Piedras Blancas National Park. The area is classified as Tropical Wet Forest Life Zone according to Holdridge et al. (1971). The landscape is dominated by primary forests, surrounded by intensively-used agricultural lands mainly comprised of pastures and oil palm plantations. Study sites for data collection were selected in 4 of the dominant habitat types of the region - Forest Interior (FI), Forest Margin (FM), Young Secondary Forest (YSF) and Oil Palm Plantations (OPP) - with 5 replicates per habitat type. For more details on each of the selected habitat types and local climatic conditions, please refer to Alonso-Rodriguez et al. (2017).

Data Collection

From February to July 2013, Alonso-Rodriguez et al. (2017) surveyed moth assemblages once a month in 20 sites, for a total of 6 trap nights per site spanning 6 months of sampling. Moths were collected with automatic funnel traps, as described in Brehm and Axmacher (2006). The traps were set up at dusk in the understory (1-2 meters above ground) of each site and then collected at sun rise, so they were left running for the entire night. Sampling was restricted to ± 5 days from a New Moon night, to reduce the effects of moonlight on moth catchments. Back at the laboratory, Geometridae and Erebidae-Arctiinae moths were spread, sorted to morphospecies and identified as far as possible using available online taxonomic keys and referencing local entomological museums.

In each site, the authors also measured a variety of forest structural and vegetation parameters. The following represents a list of the measured parameters, which we considered as potential independent variables for our models.

- Tree species composition (abundance matrix)
- Plant species richness
- Vegetation density
- Plant species diversity (Fisher's alpha index)
- Understory density
- Understory complexity
- Undergrowth cover
- Vertical complexity
- Canopy cover
- Tree basal area

Data Analysis and Models

The moth abundance matrix per site and night of collection contained too many ceros. So, we aggregated all the collections per site to obtain an abundance matrix per site for each moth group. Using the vegan package in R, we calculated Fisher's Alpha diversity index for the collected Geometrid and Erebidae-Arctiinae moths at each site.

Since the tree species composition matrix could not be feasibly included in our models, we used non-metric multidimensional scaling (NMDS) with Bray-Curtis distance to capture variation of plant species composition in each site. We then extracted the coordinates of each site from the ordination plot (NMDS axes 1 and 2) to be used as a proxy for plant species composition. These analyses were conducted with the vegan package in R.

We grouped the independent variables mentioned above as either structural or floristic parameters. In order to select the final parameters to keep in our structural and floristic models, we first used the cor and corrplot functions in R to evaluate which variables were autocorrelated (higher than 0.6 correlation coefficient). From correlated variables, we decided which ones to keep based on published literature on moth biology and life history. Our final models were as follows:

Structural Models

Geometrid alpha diversity ~ understory complexity + canopy cover + vertical complexity

Arctiine alpha diversity ~ understory complexity + canopy cover + vertical complexity

Floristic Models

Geometrid alpha diversity ~ NMDS1 + NMDS2 + plant species diversity + vegetation density

Arctiine alpha diversity ~ NMDS1 + NMDS2 + plant species diversity + vegetation density

We examined the distribution of the moth diversity using the Shapiro-Wilk normality test and the R descdist function (fitdistrplus package, Comprehensive R Archive Network). Due to the distribution of the data, we used linear mixed (LMER) models for Arctiinae and generalized mixed (GLMER) models for Geometridae. LMER and GLMER revealed which components of the models were significant predictors of Geometridae and Arctiinae diversity. LMER and GLMER were sourced from the lme4 R package (CRAN), while p values for the predictiveness of the vegetation parameters were generated by the lmeTest package (CRAN).

In our models, we considered the Habitat category as a random effect, since we were not interested in evaluating how this affected moth diversity. Rather, our main interest was in evaluating how the variation in moth diversity is reflected by the chosen floristic and structural vegetation parameters. Since moonlight has shown to affect light trap catchments of flying insects, we also added average moonlight (fraction of the moon illuminated) as a random factor. Moonlight data was extracted from the United States Naval Observatory (http://aa.usno.navy.mil/data/docs/MoonFraction.php) per sampling date, and then averaged for each site. Finally, we used AIC (Akaike's Information Criterion) to compare the strength of the structural and floristic models for geometrids and arctiines separately.

Results

Calculating Fisher's alpha Diversity Index for Geometridae and Arctiinae

```
## Loading required package: permute
```

```
## Loading required package: lattice
```
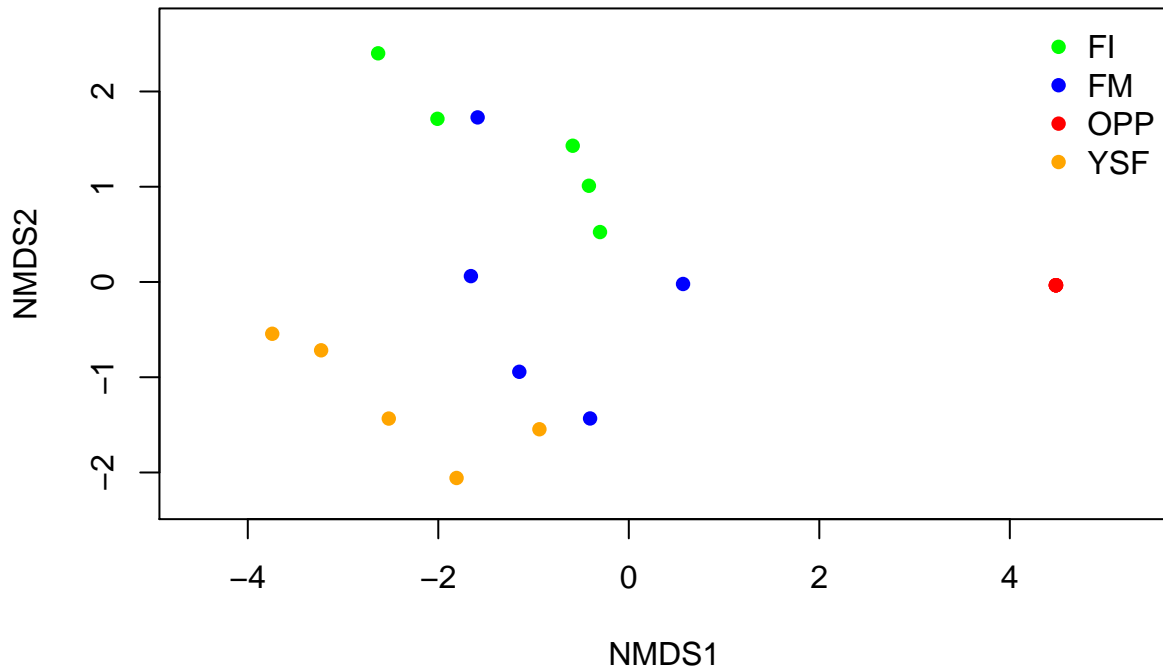
```
## This is vegan 2.5-6
```

Figure 1: Non-metric multidimensional scaling (NMDS) plot derived from matrix of plant species composition across multiple sites. X and Y Axes represent similarity in the composition among sites, with points colored by site type (FI=forest interior, FM=forest margin, OPP=oil palm plantation, YSF= young secondary forest).

Extracting the NMDS axis coordinates for the plant species matrix, to be used as proxy for plant species composition

```
cols <- c("green", "blue", "red", "orange")
plot(PlantsOrd, type = "n")
points(PlantsOrd, col = cols[habitat$Habitat], pch = 16)
legend("topright", legend=levels(habitat$Habitat), bty = "n", col=cols, pch = 16)
```

```
#Extract NMDS axis scores
nms_axis <- scores(PlantsOrd, choices=c(1,2))
#write.table(nms_axis, file="Plant_nms_axis.csv",sep=",",row.names=F)
```

Pearson correlation coefficients between all independent variables

```
## corrplot 0.84 loaded
```

We used the Pearson coefficients to determine correlation between the vegetation parameters. We determined that a floristic model could safely include the following parameters since they lack strong correlations: plant
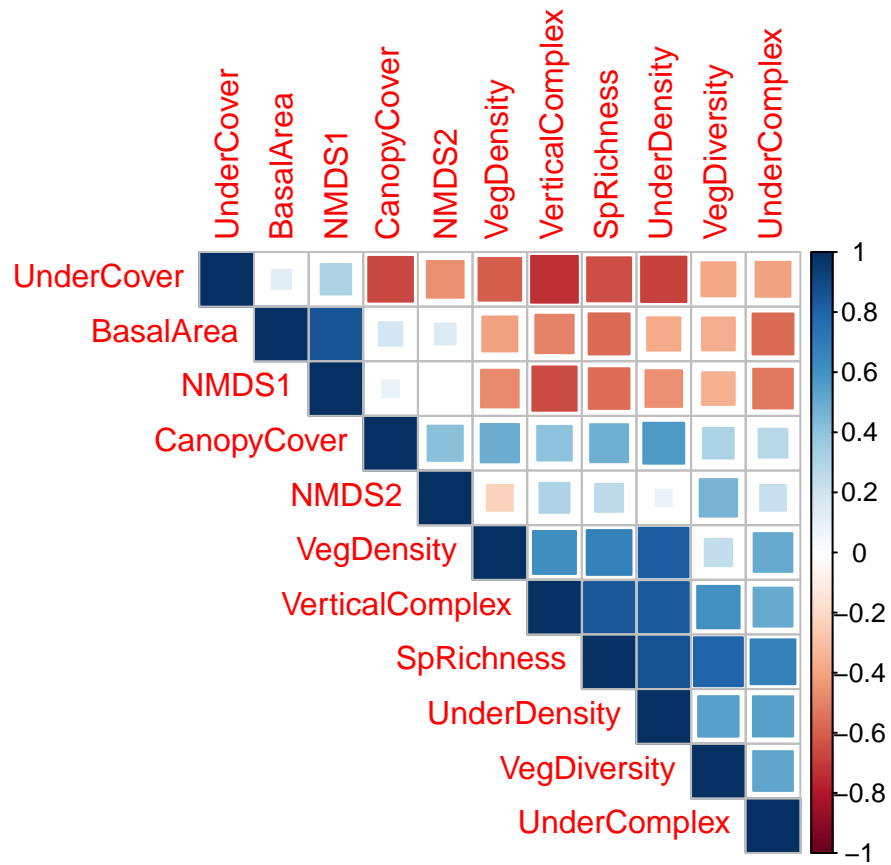
Figure 2: Correlation plot created using Pearson coefficients describing relationship between all listed floristic and structural parameters measured across sites.

species diversity, vegetation density, and a plant species composition matrix (NMDS1 and NMDS2). We also found that a structural model could include the following parameters: understory complexity, canopy cover, and vertical complexity.

Modeling alpha diversity of Geometridae and Arctiinae against structural and floristic variables

Results from our generalized mixed-effects model (Geometridae) and linear mixed-effects model (Arctiinae) showed similar responses of both moth groups to floristic and structural parameters. All models and graphs of the relationship between moth diversity and each independent variable included in the models are shown below.

Evaluating distribution of moth diversity data

```
##
##  Shapiro-Wilk normality test
##
## data:  covar_moth$Geometrid_FisherIndex
## W = 0.87764, p-value = 0.01605


##
##  Shapiro-Wilk normality test
##
## data:  covar_moth$Arctiine_FisherIndex
## W = 0.9299, p-value = 0.1537


## Loading required package: survival


## Loading required package: npsurv


## Loading required package: lsei



## summary statistics
## ------
## min:  6.11273   max:  73.17881
## median:  24.63236
## mean:  24.66262
## estimated sd:  17.7614
## estimated skewness:  1.088706
## estimated kurtosis:  4.46317
```

The Shapiro Wilk Test revealed that the Arctiinae diversity data is normally distributed. So, we used linear mixed models (lmer function in R) for all Arctiinae models, which assume that the data is normally distributed (Gaussian distribution). For Geometridae however, Shapiro-Wilks test revealed that the data is not normally distributed, and the descdist function identified it as having a gamma distribution. For this reason, we used generalized mixed models (glmer function in R) for all Geometridae models, which allow for specifying the distribution of data with non-Gaussian distributions.
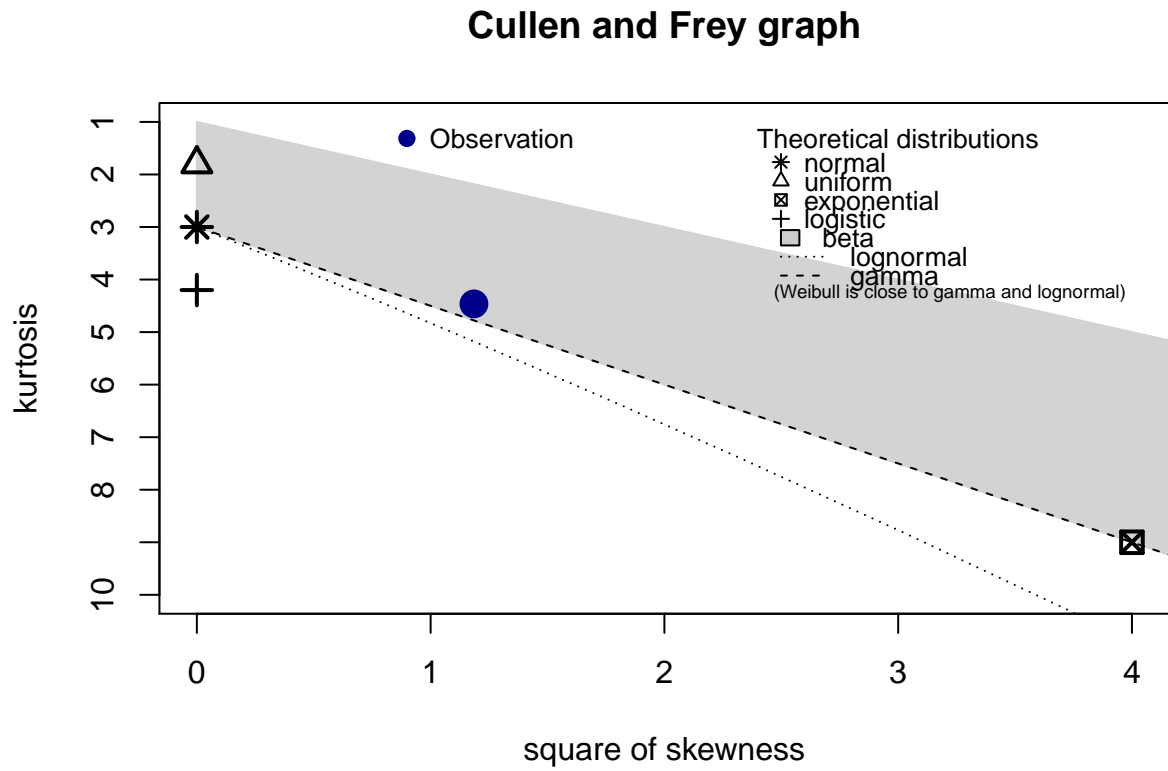
Figure 3: Cullen and Frey graph depicting the appropriateness of theoretical distributions for model application given an input observation. Graph was interpreted according to the methodology of Cullen and Frey (1999) and Delignette-Muller & Dutang (2015).

Geometridae diversity against structural parameters

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: Gamma  ( log )
## Formula:
## Geometrid_FisherIndex ~ UnderComplex + CanopyCover + VerticalComplex +
##     (1 | Moonlight) + (1 | Habitat)
##    Data: covar_moth
##
##      AIC      BIC   logLik deviance df.resid
##    165.0    172.0    -75.5    151.0       13
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.5388 -0.4993 -0.1074  0.5415  1.6999
##
## Random effects:
##  Groups    Name        Variance Std.Dev.
##  Moonlight (Intercept) 0.05608  0.2368
##  Habitat   (Intercept) 0.23188  0.4815
##  Residual              0.20354  0.4512
## Number of obs: 20, groups:  Moonlight, 8; Habitat, 4
##
## Fixed effects:
##                Estimate Std. Error t value Pr(>|z|)
## (Intercept)    3.189614   0.934892   3.412 0.000645 ***
## UnderComplex   0.114630   0.560124   0.205 0.837845
## CanopyCover   -0.003863   0.014314  -0.270 0.787277
## VerticalComplex -0.080419  0.899519  -0.089 0.928762
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##            (Intr) UndrCm CnpyCv
## UnderComplx -0.420
## CanopyCover -0.108 -0.662
## VertclCmplx -0.468  0.742 -0.742
```

The geometrid structural model displayed no significant correlation between the listed structural parameters (understory complexity, canopy cover, and vertical complexity), showing that geometrid diversity is likely not a function of structural forest variables in these study sites.

Arctiinae diversity against structural parameters

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula:
## Arctiine_FisherIndex ~ UnderComplex + CanopyCover + VerticalComplex +
##     (1 | Moonlight) + (1 | Habitat)
##    Data: covar_moth
##
## REML criterion at convergence: 119
```
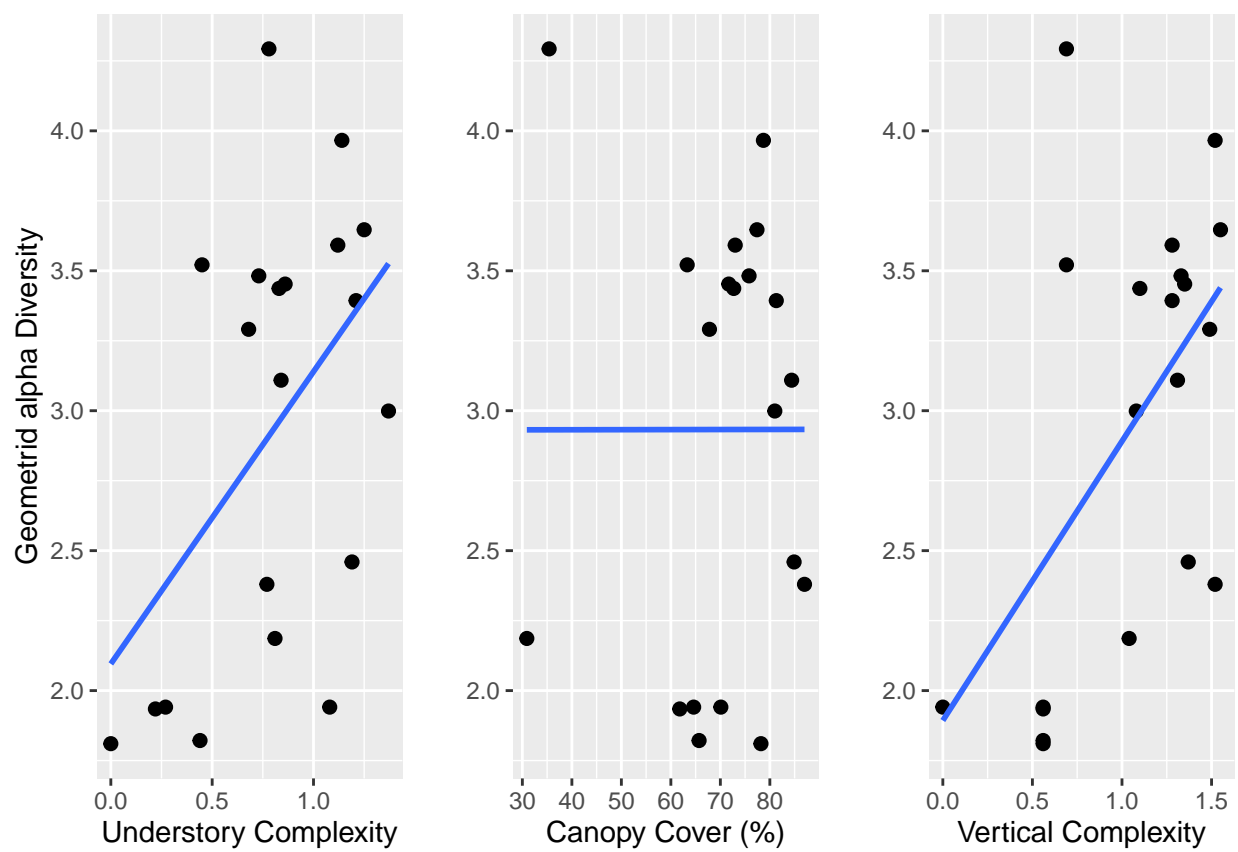
Figure 4: Geometridae structural model, showing the correlations between geometrid Fisher's alpha and understory complexity, canopy cover (%) and vertical complexity.

```
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.0052 -0.5947 -0.2638  0.4222  1.6753
##
## Random effects:
##  Groups    Name        Variance Std.Dev.
##  Moonlight (Intercept)  5.941   2.437
##  Habitat   (Intercept) 97.289   9.864
##  Residual              24.388   4.938
## Number of obs: 20, groups:  Moonlight, 8; Habitat, 4
##
## Fixed effects:
##               Estimate Std. Error     df t value Pr(>|t|)
## (Intercept)     31.985     10.426  8.940   3.068   0.0135 *
## UnderComplex     2.600      4.580 12.263   0.568   0.5805
## CanopyCover     -0.150      0.107  8.499  -1.401   0.1966
## VerticalComplex -7.502      6.208 13.264  -1.208   0.2480
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##            (Intr) UndrCm CnpyCv
## UnderComplx -0.398
## CanopyCover -0.432 -0.226
## VertclCmplx -0.494  0.340 -0.341
```

The arctiine structural model showed little to no relationship between the aforementioned structural variables and alpha diversity within this clade. However, the large output value for habitat as a controlled random effect (97.289) suggests this component plays a significant role in dictating patterns of arctiine diversity.

Geometridae diversity against floristic parameters

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: Gamma  ( log )
## Formula:
## Geometrid_FisherIndex ~ VegDiversity + VegDensity + NMDS1 + NMDS2 +
##     (1 | Moonlight) + (1 | Habitat)
##    Data: covar_moth
##
##      AIC      BIC   logLik deviance df.resid
##    158.2    166.1    -71.1    142.2       12
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.4570 -0.6296 -0.1251  0.5783  1.6631
##
## Random effects:
##  Groups    Name        Variance Std.Dev.
##  Moonlight (Intercept) 0.10413  0.3227
##  Habitat   (Intercept) 0.02786  0.1669
##  Residual              0.14302  0.3782
```
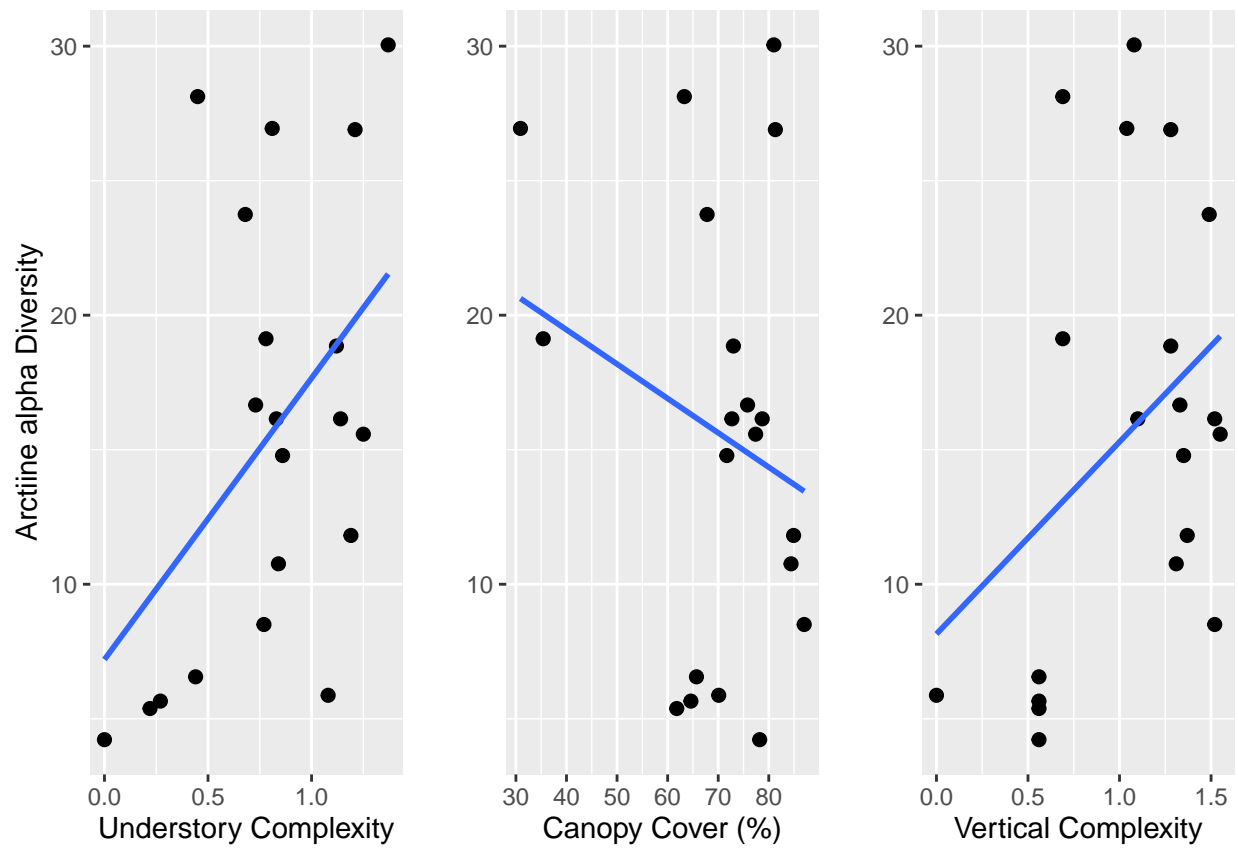
Figure 5: Arctiinae structural model, showing the correlations between arctiine Fisher's alpha and understory complexity, canopy cover (%) and vertical complexity.
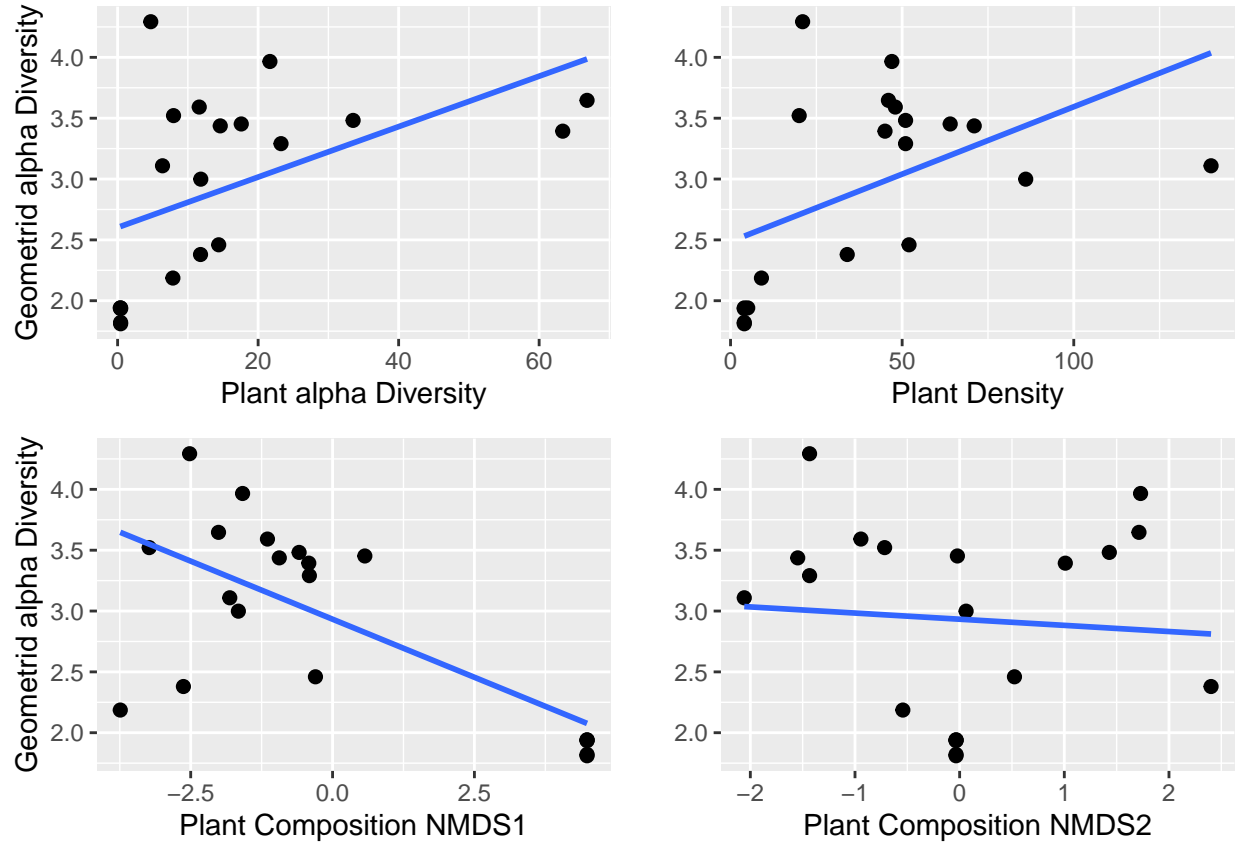
Figure 6: Geometridae floristic model, showing the correlations between geometrid Fisher's alpha and plant alpha diversity, plant density, and plant composition non-metric multidimensional scaling axes 1 and 2.

```
## Number of obs: 20, groups:  Moonlight, 8; Habitat, 4
##
## Fixed effects:
##              Estimate Std. Error t value Pr(>|z|)
## (Intercept)  2.419119   0.308423   7.844 4.38e-15 ***
## VegDiversity 0.025024   0.007390   3.386 0.000709 ***
## VegDensity   0.002470   0.003797   0.650 0.515433
## NMDS1       -0.117604   0.059781  -1.967 0.049154 *
## NMDS2       -0.278201   0.123361  -2.255 0.024122 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##            (Intr) VgDvrs VgDnst NMDS1
## VegDiversty -0.361
## VegDensity  -0.464 -0.087
## NMDS1       -0.148  0.218  0.098
## NMDS2        0.204 -0.323 -0.100  0.242
```

Conversely to the output of the structural models, the floristic models returned significant correlations between select parameters (vegetation diversity, vegetation density, NMDS 1 + 2) and diversity measures. Variance of random effects (moonlight and habitat) were inconsequential in the geometrid floristic model,

while vegetation diversity appears to have a very significant effect on moth alpha diversity (the greatest correlation between diversity and parameters out of all tested models). This result is paralleled with similar studies that discovered a significant relationship between geometrid diversity and vegetation diversity (particularly understory vegetation) *citations*. Results of this model also support a relationship between each of the NMDS axes and geometrid diversity. Negative correlations between each of the NMDS axes and moth diversity are consistent with the NMDS output, since habitats with higher plant diversity are clustered at lower values along the x axis.

Arctiinae diversity against floristic parameters

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula:
## Arctiine_FisherIndex ~ VegDiversity + VegDensity + NMDS1 + NMDS2 +
##     (1 | Moonlight) + (1 | Habitat)
##    Data: covar_moth
##
## REML criterion at convergence: 121.9
##
## Scaled residuals:
##     Min      1Q   Median      3Q      Max
## -1.40414 -0.44870 -0.08145  0.24619  2.45010
##
## Random effects:
##  Groups    Name         Variance Std.Dev.
##  Moonlight (Intercept)  0.00     0.000
##  Habitat   (Intercept) 10.56     3.250
##  Residual              25.19     5.019
## Number of obs: 20, groups:  Moonlight, 8; Habitat, 4
##
## Fixed effects:
##              Estimate Std. Error       df t value Pr(>|t|)
## (Intercept) 15.64545    2.79629  4.74247   5.595  0.00298 **
## VegDiversity  0.17847    0.08285 14.98347   2.154  0.04791 *
## VegDensity   -0.07210    0.04304 14.93753  -1.675  0.11471
## NMDS1        -2.14775    0.68629  3.61846  -3.130  0.04037 *
## NMDS2        -2.67626    1.43069 10.22278  -1.871  0.09027 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##            (Intr) VgDvrs VgDnst NMDS1
## VegDiversty -0.360
## VegDensity  -0.534 -0.183
## NMDS1       -0.180  0.132  0.190
## NMDS2       -0.093 -0.276  0.360  0.166
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

The arctiine floristic model showed an increased value for habitat as a random effect, similarly to the arctiine structural model. Alpha diversity of arctiines was again correlated with vegetation diversity (as with the geometrid floristic model), although not as strongly. NMDS1, but not NMDS2, also displayed a significant negative relationship with diversity. As only oil palm plantations are located at higher values along NMDS1 (which are the least diverse habitats), this correlation is also sensible.
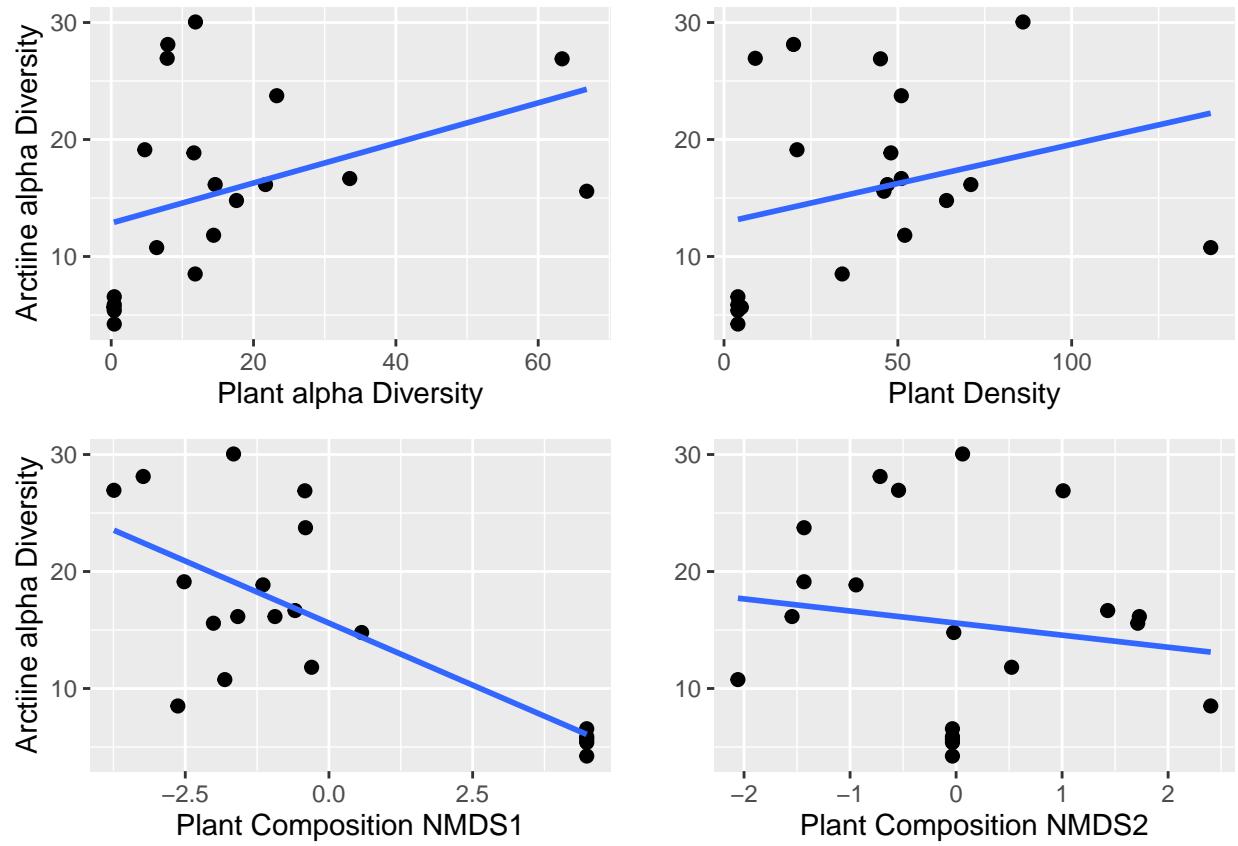
Figure 7: Arctiinae floristic model, showing the correlations between arctiine Fisher's alpha and plant alpha diversity, plant density, and plant composition non-metric multidimensional scaling axes 1 and 2.

Using AIC to compare the floristic and structural models for each moth group

```
##                     df      AIC
## Geometrid.structural  7 165.0320
## Geometrid.floristic   8 158.1526


##                     df      AIC
## Arctiid.structural   7 132.9644
## Arctiid.floristic    8 137.8816
```

AIC was used to compare the strength of floristic and structural models for each clade, although this measurement is likely skewed due to the presence of random effects (affects definition of degrees of freedom (Jiang et al., 2008). AIC ranked the geometrid floristic model higher than the geometrid structural model, in correspondence with the apparent high significance of plant diversity and composition, compared to the failure of structural parameters to adequately describe geometrid diversity patterns. By contrast, the arctiine structural model was ranked higher than the corresponding floristic model, but this may be due to the large effect of habitat acting in the background.

Discussion

coming soon

Easton comments

- You don't need this for the class project, but eventually you should add more citations in the introduction
- Good job in the introduction setting up why this work is important and how you plan on building on Alonso-Rodriguez et al. (2017). I also like how you ended the section with a list of questions and predictions.
- I would break up your methods into subsections
- You don't need to do this for the class project, but for a final manuscript, you will need to do some residual analyses to make sure you are meeting your modeling assumptions
- I split one of your R code chunks in order to suppress the model output but to retain a plot

What I need from you for the final version:

- Clarify what data you averaged together. I believe you averaged all the data within a site for all the different months. Is this true?
- add appropriate figure captions
- Break up your methods into subsections (e.g. study site, data collection, models)
- No need to add to the discussion section at this point

References

Alonso-Rodríguez, A. M., Finegan, B., & Fiedler, K. (2017). Neotropical moth assemblages degrade due to oil palm expansion. Biodiversity and Conservation, 26(10), 2295-2326.

Axmacher, J. C., Brehm, G., Hemp, A., Tünte, H., Lyaruu, H. V., Müller-Hohenstein, K., & Fiedler, K. (2009). Determinants of diversity in afrotropical herbivorous insects (Lepidoptera: Geometridae): plant diversity, vegetation structure or abiotic factors?. Journal of Biogeography, 36(2), 337-349.

Axmacher, J. C., Liu, Y., Wang, C., Li, L., & Yu, Z. (2011). Spatial a-diversity patterns of diverse insect taxa in Northern China: Lessons for biodiversity conservation. Biological Conservation, 144(9), 2362–2368.

Beck, J., Brehm, G., & Fiedler, K. (2011). Links between the environment, abundance and diversity of Andean moths. Biotropica, 43(2), 208-217.

Beck, J. A. N., Schulze, C. H., Linsenmair, K. E., & Fiedler, K. (2002). From forest to farmland: diversity of geometrid moths along two habitat gradients on Borneo. Journal of tropical ecology, 18(1), 33-51.

Brehm, G., & Axmacher, J. C. (2006). A comparison of manual and automatic moth sampling methods (Lepidoptera: Arctiidae, Geometridae) in a rain forest in Costa Rica. Environmental Entomology, 35(3), 757-764.

Brehm, G., Homeier, J., & Fiedler, K. (2003). Beta diversity of geometrid moths (Lepidoptera: Geometridae) in an Andean montane rainforest. Diversity and Distributions, 9(5), 351–366.

Cullen, A. C., Frey, H. C., & Frey, C. H. (1999). Probabilistic techniques in exposure assessment: a handbook for dealing with variability and uncertainty in models and inputs. Springer Science & Business Media.

Delignette-Muller M.L. and Dutang C. (2015). fitdistrplus: An R Package for Fitting Distributions. Journal of Statistical Software, 64(4), 1-34.

Fuentes-Montemayor, E., Goulson, D., Cavin, L., Wallace, J. M., & Park, K. J. (2012). Factors influencing moth assemblages in woodland fragments on farmland: Implications for woodland management and creation schemes. Biological Conservation, 153, 265-275.

Gray, R. E., Slade, E. M., Chung, A. Y., & Lewis, O. T. (2019). Movement of moths through riparian reserves within oil palm plantations. bioRxiv, 204990.

Hilt, N., & Fiedler, K. (2005). Diversity and composition of Arctiidae moth ensembles along a successional gradient in the Ecuadorian Andes. Diversity and Distributions, 11(5), 387-398.

Holdridge, L. R., & Grenke, W. C. (1971). Forest environments in tropical life zones: a pilot study. Forest environments in tropical life zones: a pilot study. Pergamon Press, Oxford.

Jiang, J., Rao, J., Gu, Z., & Nguyen, T. (2008). FENCE METHODS FOR MIXED MODEL SELECTION. The Annals of Statistics, 36(4), 1669–1692.

Ober, H., & Hayes, J. (2010). Determinants of nocturnal Lepidopteran diversity and community structure in a conifer-dominated forest. Biodiversity and Conservation, 19(3), 761–774.