

Diversity and carbon across the tropical forest biome

Martin Sullivan et al.

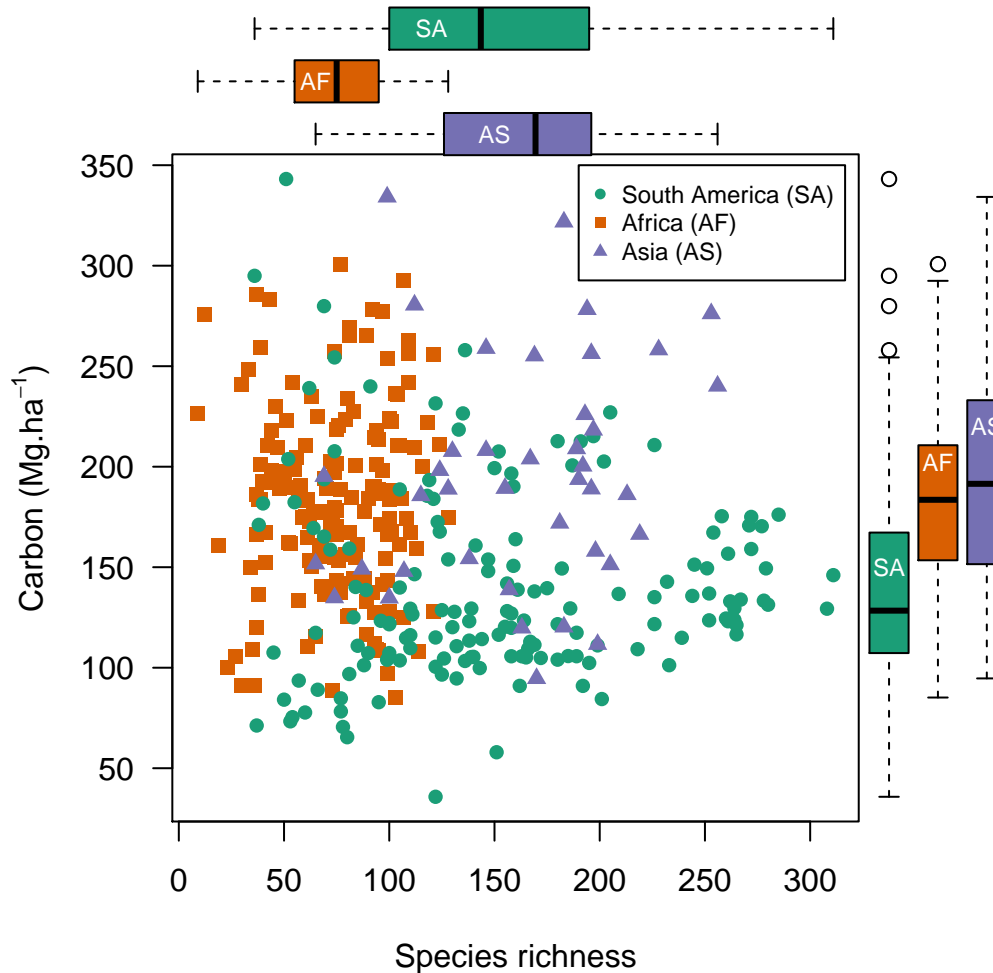
25 November 2016

Diversity carbon relationship among 1 ha plots

We will load the data, and then plot the bivariate relationship between tree diversity and carbon storage across the tropical forest biome

```
# Set working directory
setwd("~/Desktop/Talbot paper/Data_package")
# Load 1 ha data
plots <- read.csv("Data_1ha.csv", header = T, as.is = T, na.strings = "#N/A")

# Convert AGBChave to carbon
plots$Carbon <- 0.471 * plots$AGBChave
```



We will now test whether this relationship is statistically significant, first pan-tropically and then in each continent

```

# Look at pan-tropical bivariate diversity-carbon
# relationship Weight by forest cover - from Saatchi et al
# 2011 PNAS, using 30% cover values (AF=26.65%, SA = 53.25%,
# AS = 20.04%)
wv <- ifelse(plots$continent == "South America", 191.7/158, ifelse(plots$continent ==
  "Africa", 95.96/162, 72.13/40))
mod <- lm(log(Carbon) ~ SpeciesRich, data = plots, weights = wv)
summary(mod)

##
## Call:
## lm(formula = log(Carbon) ~ SpeciesRich, data = plots, weights = wv)
##
## Weighted Residuals:
##      Min       1Q   Median       3Q      Max
## -1.60964 -0.22140  0.03601  0.23679  1.03082
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.0669952  0.0408399 124.070  <2e-16 ***
## SpeciesRich -0.0002306  0.0002736  -0.843    0.4
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3464 on 358 degrees of freedom
## Multiple R-squared:  0.00198,    Adjusted R-squared:  -0.0008075
## F-statistic: 0.7104 on 1 and 358 DF,  p-value: 0.3999

# Kendall's tau correlations
c1 <- with(plots[plots$continent == "South America", ], cor.test(Carbon,
  SpeciesRich, method = "k"))
c2 <- with(plots[plots$continent == "Africa", ], cor.test(Carbon,
  SpeciesRich, method = "k"))
c3 <- with(plots[plots$continent == "Asia", ], cor.test(Carbon,
  SpeciesRich, method = "k"))

k.cor <- data.frame(Continent = c("South America", "Africa",
  "Asia"), Tau = round(c(c1$estimate, c2$estimate, c3$estimate),
  3), P = round(c(c1$p.value, c2$p.value, c3$p.value), 3))

```

Table 1. Correlation between species richness and carbon

Continent	Tau	P
South America	0.084	0.120
Africa	0.014	0.788
Asia	0.132	0.230

Variation in diversity and carbon among continents

```

# Variation in diversity and carbon among continents Split
# data by continent so can run functions on each continent
dat <- split(plots, plots$continent)

```

```

# Define function to randomly sample data with replacement
boot.fun <- function(x, n = 10000) {
  m.vec <- c()
  for (i in 1:n) {
    b.samp <- sample(x, length(x), replace = T)
    m.vec[i] <- mean(b.samp)
  }
  LCL <- quantile(m.vec, 0.025)
  UCL <- quantile(m.vec, 0.975)
  res <- paste(round(LCL, 0), round(UCL, 0), sep = " - ")
  return(res)
}

# Define wrapper function to do bootstrapping and return mean
# and 95% CIs
boot.wrap <- function(x) {
  bs <- boot.fun(x)
  bs <- paste("(", bs, ")", sep = "")
  mv <- round(mean(x), 0)
  res <- paste(mv, bs, sep = " ")
  return(res)
}

# Run function for each variable
cont.sum <- rbind(c("Carbon (Mg/ha)", sapply(dat, function(x) boot.wrap(x$Carbon))),
  c("Fisher's alpha", sapply(dat, function(x) boot.wrap(x$FishersAlpha))),
  c("Species richness / ha", sapply(dat, function(x) boot.wrap(x$SpeciesRich))),
  c("SpRch / 300 stems", sapply(dat, function(x) boot.wrap(na.omit(as.numeric(x$Speciesper300stems)))),
  c("Genus richness / ha", sapply(dat, function(x) boot.wrap(x$GenusRich))),
  c("GnRch / 300 stems", sapply(dat, function(x) boot.wrap(na.omit(as.numeric(x$Generaper300stems)))),
  c("Family richness / ha", sapply(dat, function(x) boot.wrap(x$FamilyRich))),
  c("FamRch / 300 stems", sapply(dat, function(x) boot.wrap(na.omit(as.numeric(x$Familiesper300stems))))))

# Test whether variables differ significantly among
# continents
plots2 <- plots
plots2$continent <- as.factor(plots2$continent)
mf1 <- lm(log(Carbon) ~ continent, data = plots2)
mf2 <- lm(sqrt(FishersAlpha) ~ continent, data = plots2)
mf3 <- glm.nb(SpeciesRich ~ continent, data = plots2)
mf4 <- lm(as.numeric(Speciesper300stems) ~ continent, data = plots2)
mf5 <- glm.nb(GenusRich ~ continent, data = plots2)
mf6 <- lm(as.numeric(Generaper300stems) ~ continent, data = plots2)
mf7 <- glm.nb(FamilyRich ~ continent, data = plots2)
mf8 <- lm(as.numeric(Familiesper300stems) ~ continent, data = plots2)

# Different letters indicate significant differences
mc <- rbind(cld(glht(mf1, linfct = mcp(continent = "Tukey")))$mcletters$Letters,
  cld(glht(mf2, linfct = mcp(continent = "Tukey")))$mcletters$Letters,
  cld(glht(mf3, linfct = mcp(continent = "Tukey")))$mcletters$Letters,
  cld(glht(mf4, linfct = mcp(continent = "Tukey")))$mcletters$Letters,
  cld(glht(mf5, linfct = mcp(continent = "Tukey")))$mcletters$Letters,
  cld(glht(mf6, linfct = mcp(continent = "Tukey")))$mcletters$Letters,

```

```

cld(glht(mf7, linfct = mcp(continent = "Tukey")))$mcletters$Letters,
cld(glht(mf8, linfct = mcp(continent = "Tukey")))$mcletters$Letters)

cont.sum <- as.data.frame(cbind(cont.sum, mc))
names(cont.sum) <- c("Variable", "Africa", "Asia", "South America",
  "Africa.sig", "Asia.sig", "South America.sig")

```

Table 2. Variation in diversity and carbon among continents. Different letters indicate statistically significant differences.

Variable	Africa	Asia	South America	Africa.sig	Asia.sig	South America.sig
Carbon (Mg/ha)	183 (176 - 190)	197 (179 - 215)	140 (133 - 148)	b	b	a
Fisher's alpha	28 (26 - 30)	84 (73 - 96)	80 (71 - 89)	a	b	b
Species richness / ha	74 (70 - 78)	162 (147 - 177)	152 (142 - 163)	a	b	b
SpRch / 300 stems	65 (62 - 68)	120 (111 - 130)	109 (102 - 116)	a	b	b
Genus richness / ha	59 (56 - 62)	87 (81 - 93)	91 (87 - 96)	a	b	b
GnRch / 300 stems	54 (51 - 56)	71 (66 - 75)	72 (68 - 75)	a	b	b
Family richness / ha	27 (27 - 28)	40 (38 - 42)	38 (37 - 40)	a	b	b
FamRch / 300 stems	26 (25 - 27)	35 (34 - 37)	33 (32 - 34)	a	b	b

Examine diversity-carbon relationship controlling for climate and soil

```

# Examine diversity-carbon relationship controlling for other
# variables NOTE: Lat long only to 2 dp in this file, so
# results of SAR will differ slightly to those in paper Need
# to allow for zero distances between plots
set.ZeroPolicyOption(TRUE)

## [1] FALSE

# Collapse variation in soil texture into two axes
soil.pca <- prcomp(plots[, 32:34])
plots$PCA1 <- soil.pca$x[, 1]
plots$PCA2 <- soil.pca$x[, 2]

# Create coords object
coords <- as.matrix(cbind(plots$Long, plots$Lat))
# Scale predictor variables
plots[, c(10:12, 14:23, 30, 31, 35:37, 39, 40)] <- scale(plots[,
  c(10:12, 14:23, 30, 31, 35:37, 39, 40)])

# Set diversity variable
resp <- "SpeciesRich"

# Set continent
cont = "South America"

# Create formula
form <- as.formula(paste("log(Carbon)~", resp, "+MAT+MAP+CWD+PCA1+PCA2+TEB.HWSDbest+C.N.HWSDbest"))

# Fit OLS model
glob.mod <- lm(form, data = plots[plots$continent == cont, ],

```

```

    na.action = "na.fail")
adj.r <- summary(glob.mod)$adj.r.squared
coef <- summary(glob.mod)$coefficients

# Select neighbourhood distance that gives lowest AIC value
# for SAR model
aics <- c()
dist.seq <- seq(20, 1000, by = 20)
for (i in 1:length(dist.seq)) {
  nb <- dnearneigh(coords[plots$continent == cont, ], 0, dist.seq[i],
    longlat = T)
  # Spatial weights with coding style 'W' (row standardized)
  nbw <- nb2listw(nb, style = "W", zero.policy = TRUE)
  # SAR error model
  sp.mod <- errorsarlm(glob.mod, data = plots[plots$continent ==
    cont, ], listw = nbw, zero.policy = TRUE, na.action = na.fail)
  aics[i] <- AIC(sp.mod)
}
best.neigh <- dist.seq[grep(min(aics), aics)]
# Run with best neighbourhood size
nb <- dnearneigh(coords[plots$continent == cont, ], 0, best.neigh,
  longlat = T)
nbw <- nb2listw(nb, style = "W", zero.policy = TRUE)
sp.mod <- errorsarlm(glob.mod, data = plots[plots$continent ==
  cont, ], listw = nbw, zero.policy = TRUE, na.action = na.fail)
sp.sum <- summary(sp.mod, Nagelkerke = TRUE)
sp.coef <- sp.sum$Coef

# Run all simplifications of model
dd1 <- dredge(glob.mod)

## Fixed term is "(Intercept)"
dd2 <- dredge(sp.mod)

## Fixed term is "(Intercept)"

# Model averaging
avgmod1 <- model.avg(dd1, cumsum(weight) <= 0.95)
avgmod2 <- model.avg(dd2, cumsum(weight) <= 0.95)

# Collate OLS results
a1 <- summary(avgmod1)$coefmat.subset
b <- summary(avgmod1)$importance
ols <- cbind(a1, b[match(dimnames(a1)[[1]], names(b))])
ols <- as.data.frame(ols)
ols$Continent <- cont
ols$Response <- resp
ols$Mod <- "OLS"
ols <- ols[, c(1, 2, 5, 6, 7, 8, 9)]
names(ols) <- c("Beta", "SE", "P", "AIC wi", "Cont", "Resp",
  "Mod")

# Collate SAR results
a2 <- summary(avgmod2)$coefmat.subset

```

```

b <- summary(avgmod2)$importance
sar <- cbind(a2, b[match(dimnames(a2)[[1]], names(b))])
sar <- as.data.frame(sar)
sar$Continent <- cont
sar$Response <- resp
sar$Mod <- "SAR"
sar <- sar[, c(1, 2, 4, 5, 6, 7, 8)]
names(sar) <- c("Beta", "SE", "P", "AIC wi", "Cont", "Resp",
               "Mod")
ols

```

```

##              Beta      SE      P      AIC wi      Cont
## (Intercept)  4.93595094 0.02890617 0.0000000      NA South America
## C.N.HWSDbest  0.04852711 0.02470883 0.0512592 0.7280081 South America
## CWD          0.16843498 0.03131219 0.0000001 1.0000000 South America
## PCA1         0.08019778 0.02729920 0.0035216 0.9708428 South America
## PCA2        -0.04417959 0.01964551 0.0255596 0.8517373 South America
## SpeciesRich -0.04660722 0.03285775 0.1588869 0.5020595 South America
## MAT          0.02970411 0.02487022 0.2360389 0.4041722 South America
## MAP         -0.01789122 0.03399817 0.6015953 0.2634435 South America
## TEB.HWSDbest -0.01100952 0.02114794 0.6053089 0.2761468 South America
##              Resp Mod
## (Intercept) SpeciesRich OLS
## C.N.HWSDbest SpeciesRich OLS
## CWD          SpeciesRich OLS
## PCA1         SpeciesRich OLS
## PCA2         SpeciesRich OLS
## SpeciesRich  SpeciesRich OLS
## MAT          SpeciesRich OLS
## MAP          SpeciesRich OLS
## TEB.HWSDbest SpeciesRich OLS
sar

```

```

##              Beta      SE      P      AIC wi      Cont
## lambda       0.848165818 0.05476865 0.0000000      NA South America
## (Intercept)  4.842160693 0.10927324 0.0000000      NA South America
## CWD          0.111601402 0.03286838 0.0006853 0.9983029 South America
## PCA2        -0.026593674 0.01706670 0.1191810 0.5223580 South America
## TEB.HWSDbest 0.032724776 0.01531342 0.0325981 0.7750329 South America
## MAT         -0.011297196 0.02130746 0.5959739 0.2622629 South America
## SpeciesRich  0.008356920 0.02462011 0.7342823 0.2454242 South America
## MAP          0.004592115 0.03261451 0.8880281 0.2715761 South America
## C.N.HWSDbest 0.008155481 0.01772305 0.6453997 0.2612701 South America
## PCA1         0.005975005 0.02019077 0.7672854 0.2453732 South America
##              Resp Mod
## lambda       SpeciesRich SAR
## (Intercept) SpeciesRich SAR
## CWD          SpeciesRich SAR
## PCA2         SpeciesRich SAR
## TEB.HWSDbest SpeciesRich SAR
## MAT          SpeciesRich SAR
## SpeciesRich  SpeciesRich SAR
## MAP          SpeciesRich SAR
## C.N.HWSDbest SpeciesRich SAR

```

```
## PCA1 SpeciesRich SAR
```

Within plot analysis

```
# Within plot analysis Load subplot data
subplot <- read.csv("Data_subplot.csv", header = T, as.is = T,
  na.string = "NA")
# Convert AGB to carbon
subplot$Carbon <- 0.471 * subplot$AGBChave

# Split by plot so can run functions on each plot
s1 <- split(subplot, subplot$plotviewid)

# Run regressions controlling for stem density in each plot
beta.vec <- sapply(s1, function(x) as.numeric(coef(lm(log(AGBChave) ~
  log(SpeciesRich) + log(stems) + I(log(stems)^2), data = x))[2]))
# Get continent ID for each plot
cont.list <- sapply(s1, function(x) unique(x$continent))

# Mean across plots
m.all <- mean(beta.vec)
# Bootstrap to find 95% CIs
m.samp <- c()
for (i in 1:10000) {
  b2 <- beta.vec[sample(1:length(beta.vec), length(beta.vec),
    replace = T)]
  m.samp[i] <- mean(b2)
}
LCL <- quantile(m.samp, 0.025)
UCL <- quantile(m.samp, 0.975)

# Define as a function so can run on all continents
b.func <- function(x) {
  m.all2 <- mean(x)
  m.samp <- c()
  for (i in 1:10000) {
    b2 <- x[sample(1:length(x), length(x), replace = T)]
    m.samp[i] <- mean(b2)
  }
  LCL2 <- quantile(m.samp, 0.025)
  UCL2 <- quantile(m.samp, 0.975)
  return(c(m.all2, LCL2, UCL2))
}

# Run on each continent
Af <- b.func(beta.vec[cont.list == "Africa"])
As <- b.func(beta.vec[cont.list == "Asia"])
Sa <- b.func(beta.vec[cont.list == "South America"])
```

