Analysis Stand Features

AJ Perez-Luque ((???))

2018 Jan

Prepare Data

- Two datasets: focal tree and competence

```
# Compute diameter (mm)
tree <- tree %>%
  mutate(dn_mm = (perim_mm / pi))

### - old code 4 sites
# Set levels of eleveation
# sj_lowcode  <- paste0('A', str_pad(1:10, 2, pad='0'))
# sj_highcode <- paste0('A', 11:20)
# ca_lowcode <- c(paste0('B', str_pad(1:10, 2, pad='0')),
#              paste0('B', 26:30))
# ca_highcode <- paste0('B', 11:25)
#
# tree <- tree %>%
#   mutate(elevF = ifelse(id_focal %in% sj_lowcode, 'Low',
#                       ifelse(id_focal %in% sj_highcode, 'High',
#                              ifelse(id_focal %in% ca_lowcode, 'Low', 'High')))) %>%
#   mutate(site = paste0(loc, '_', elevF))
#
#


sj_code  <- paste0('A', str_pad(1:20, 2, pad='0'))
ca_lowcode <- c(paste0('B', str_pad(1:10, 2, pad='0')), paste0('B', 26:30))
ca_highcode <- paste0('B', 11:25)

tree <- tree %>% mutate(
  site = as.factor(case_when(
    id_focal %in% sj_code ~ "sj",
    id_focal %in% ca_highcode ~ "caH",
    id_focal %in% ca_lowcode ~ "caL")))

# Get only focal trees
ft <- tree %>%
  filter(sp=='Focal') %>%
  filter(id_focal!='Fresno')

# Get only no focal trees
nft <- tree %>%
  filter(sp!='Focal')
```

General Variables

Numbers of focal trees by site

```
general_var <- ft %>% group_by(site) %>% count()

general_var %>% kable
```

site	n
caH	15
caL	15
sj	20

Spatial Info

Coordinates of the centroid for each site

```
## Get coordinates of spatial data
sp_ca <- as.data.frame(field_work_ca) %>%
  dplyr::select(ele, name, lat = coords.x2, long = coords.x1)

sp_sj <- as.data.frame(field_work_sj) %>%
  dplyr::select(ele, name, lat = coords.x2, long = coords.x1)

sp_sites <- rbind(sp_sj, sp_ca) %>%
  mutate(site = as.factor(case_when(
    name %in% sj_code ~ "sj",
    name %in% ca_highcode ~ "caH",
    name %in% ca_lowcode ~ "caL")))


coord_sites <- sp_sites %>%
  group_by(site) %>%
  dplyr::summarise(lat_m = mean(lat), long_m = mean(long))

coord_sites %>% kable()
```

site	lat_m	long_m
caH	36.96613	-3.420703
caL	36.95645	-3.424107
sj	37.13121	-3.374908

```
# old code
#
#
#
# ## Add code site and get centroid (see # http://rspatial.org/analysis/rst/8-pointpat.html)
# sp_ca <- sp_ca %>%
#   mutate(loc = 'CA',
#
```

```
#
#           elevF = ifelse(name %in% ca_lowcode, 'Low', 'High'),
#           site = paste0(loc, '_', elevF))
#
# coord_ca <- sp_ca %>%
#   group_by(site) %>%
#   summarise(lat_m = mean(lat),
#             long_m = mean(long))
#
# # plot(sp_ca$long, sp_ca$lat, pch=19, col='gray')
# # points(coord_ca$long_m, coord_ca$lat_m, pch=19, col='blue')
#
# sp_sj <- sp_sj %>%
#   mutate(loc = 'SJ',
#           elevF = ifelse(name %in% sj_lowcode, 'Low', 'High'),
#           site = paste0(loc, '_', elevF))
# coord_sj <- sp_sj %>%
#   group_by(site) %>%
#   summarise(lat_m = mean(lat),
#             long_m = mean(long))
#
# coords_sites <- coord_sj %>% rbind(coord_ca)
#
# #plot(sp_sj$long, sp_sj$lat, pch=19, col='gray')
# #points(coord_sj$long_m, coord_sj$lat_m, pch=19, col='blue')
#
# coords_sites %>% kable()
```

Competition data

- Read data
- Create a custom function to compare between sites (aov & post hoc)
- Export data into text files (see `/out/anovas_competition/`)

Distance-Independet Indices

Basal Area

Table 3: Mean values (ba)

site	mean	sd	se	min	max	tukey	variable
caH	1.229	0.7637	0.1972	0.0118	2.817	b	ba
caL	0.5661	0.2235	0.0577	0.2512	0.8633	a	ba
sj	0.3658	0.1717	0.0384	0.1205	0.7658	a	ba

Table 4: ANOVA table (ba)

term	df	sumsq	meansq	statistic	p.value
site	2	6.696	3.348	16.7	**0**
Residuals	47	9.424	0.2005		

Stand Density

Table 5: Mean values (std)

site	mean	sd	se	min	max	tukey	variable
caH	348	147.1	37.98	63.66	541.1	a	std
caL	409.6	226	58.35	159.2	1050	a	std
sj	339	130.3	29.14	127.3	636.6	a	std

Table 6: ANOVA table (std)

term	df	sumsq	meansq	statistic	p.value
site	2	47401	23701	0.8309	0.442
Residuals	47	1340699	28526		

Plot Density

Table 7: Mean values (pd)

site	mean	sd	se	min	max	tukey	variable
caH	0.0348	0.0147	0.0038	0.0064	0.0541	a	pd
caL	0.041	0.0226	0.0058	0.0159	0.105	a	pd
sj	0.0339	0.013	0.0029	0.0127	0.0637	a	pd

Table 8: ANOVA table (pd)

term	df	sumsq	meansq	statistic	p.value
site	2	0.00047	0.00024	0.8309	0.442
Residuals	47	0.01341	0.00029		

Number of competitors within *r* meters (10 m)

Table 9: Mean values (n_competitors)

site	mean	sd	se	min	max	tukey	variable
caH	10.93	4.621	1.193	2	17	a	n_competitors
caL	12.87	7.1	1.833	5	33	a	n_competitors
sj	10.65	4.095	0.9156	4	20	a	n_competitors

Table 10: ANOVA table (n_competitors)

term	df	sumsq	meansq	statistic	p.value
site	2	46.78	23.39	0.8309	0.442
Residuals	47	1323	28.15		

Number of competitors within *r* meters (10 m) such that $ dbh_j > dbh_i $

Table 11: Mean values (n_competitors_higher)

site	mean	sd	se	min	max	tukey	variable
caH	1.2	1.474	0.3805	0	4	a	n_competitors_higher
caL	0.6667	0.8165	0.2108	0	2	a	n_competitors_higher
sj	0.5	0.6882	0.1539	0	2	a	n_competitors_higher

Table 12: ANOVA table (n_competitors_higher)

term	df	sumsq	meansq	statistic	p.value
site	2	4.387	2.193	2.115	0.1319
Residuals	47	48.73	1.037		

Sum of size of trees within *r* meters (10 m)

Table 13: Mean values (sum_sizes)

site	mean	sd	se	min	max	tukey	variable
caH	3.374	1.444	0.373	0.1735	6.201	b	sum_sizes
caL	2.549	1.073	0.2771	1.275	5.28	ab	sum_sizes
sj	2.077	0.8657	0.1936	0.8324	4.269	a	sum_sizes

Table 14: ANOVA table (sum_sizes)

term	df	sumsq	meansq	statistic	p.value
site	2	14.5	7.249	5.718	**0.00599**
Residuals	47	59.58	1.268		

Size ratio

Table 15: Mean values (sr)

site	mean	sd	se	min	max	tukey	variable
caH	0.2153	0.1957	0.0505	0.0794	0.876	a	sr
caL	0.1705	0.0682	0.0176	0.0648	0.3309	a	sr
sj	0.1512	0.0643	0.0144	0.0662	0.2961	a	sr

Table 16: ANOVA table (sr)

term	df	sumsq	meansq	statistic	p.value
site	2	0.03592	0.01796	1.242	0.2982
Residuals	47	0.6797	0.01446		

Distance-Dependet Indices

Distance to nearest tree

Table 17: Mean values (dnn)

site	mean	sd	se	min	max	tukey	variable
caH	3.412	1.859	0.4801	0.88	6.75	a	dnn
caL	3.123	1.308	0.3377	1.44	5.53	a	dnn
sj	2.4	1.328	0.2969	0.67	4.99	a	dnn

Table 18: ANOVA table (dnn)

term	df	sumsq	meansq	statistic	p.value
site	2	9.665	4.833	2.146	0.1283
Residuals	47	105.8	2.252		

Crowding

Table 19: Mean values (crowding)

site	mean	sd	se	min	max	tukey	variable
caH	0.7169	0.4719	0.1219	0.0215	1.64	a	crowding
caL	0.4818	0.1844	0.0476	0.2535	0.9069	a	crowding
sj	0.4681	0.2521	0.0564	0.1922	1.026	a	crowding

Table 20: ANOVA table (crowding)

term	df	sumsq	meansq	statistic	p.value
site	2	0.6212	0.3106	3.04	**0.05732**
Residuals	47	4.802	0.1022		

Lorimer

Table 21: Mean values (lorimer)

site	mean	sd	se	min	max	tukey	variable
caH	7.664	4.711	1.216	0.1569	18.62	a	lorimer

site	mean	sd	se	min	max	tukey	variable
caL	7.779	3.937	1.017	2.749	18.51	a	lorimer
sj	9.345	4.069	0.91	3.796	19.29	a	lorimer

Table 22: ANOVA table (lorimer)

term	df	sumsq	meansq	statistic	p.value
site	2	31.72	15.86	0.8849	0.4195
Residuals	47	842.3	17.92		

Negative Exponential size ratio

Table 23: Mean values (nesr)

site	mean	sd	se	min	max	tukey	variable
caH	0.0586	0.085	0.0219	0	0.2509	a	nesr
caL	0.0348	0.0325	0.0084	0.0013	0.1035	a	nesr
sj	0.0771	0.0946	0.0211	0.0034	0.3422	a	nesr

Table 24: ANOVA table (nesr)

term	df	sumsq	meansq	statistic	p.value
site	2	0.01534	0.00767	1.261	0.2927
Residuals	47	0.2858	0.00608		

Negative Exponential Weighted size ratio

Table 25: Mean values (newsr)

site	mean	sd	se	min	max	tukey	variable
caH	0.4838	0.6224	0.1607	9e-04	1.681	a	newsr
caL	0.1792	0.2233	0.0576	2e-04	0.7879	a	newsr
sj	0.3965	0.6886	0.154	5e-04	2.57	a	newsr

Table 26: ANOVA table (newsr)

term	df	sumsq	meansq	statistic	p.value
site	2	0.7463	0.3731	1.159	0.3226
Residuals	47	15.13	0.3219		

Size ratio proportional to distance

Table 27: Mean values (srd)

site	mean	sd	se	min	max	tukey	variable
caH	0.9085	0.629	0.1624	0.0156	2.379	a	srd
caL	0.8896	0.4385	0.1132	0.3213	2.06	a	srd
sj	1.111	0.5199	0.1163	0.4794	2.247	a	srd

Table 28: ANOVA table (srd)

term	df	sumsq	meansq	statistic	p.value
site	2	0.5441	0.272	0.9565	0.3916
Residuals	47	13.37	0.2844		

Size difference proportional to distance

Table 29: Mean values (sdd)

site	mean	sd	se	min	max	tukey	variable
caH	-0.4931	0.4018	0.1037	-1.226	-0.0625	a	sdd
caL	-0.4143	0.2291	0.0592	-1.022	-0.0853	ab	sdd
sj	-0.2091	0.0972	0.0217	-0.4064	-0.0826	b	sdd

Table 30: ANOVA table (sdd)

term	df	sumsq	meansq	statistic	p.value
site	2	0.7649	0.3825	5.662	**0.00626**
Residuals	47	3.175	0.06755		

Topographic data

- Read data

```
# Read topo data
topo <- read.csv(file=here::here("/data/topo", "topo.csv"), header=TRUE, sep=',')

# topo <- topo %>%
#   mutate(loc = ifelse(str_detect(name, "A"), 'SJ', 'CA'),
#          elevF = ifelse(name %in% sj_lowcode, 'Low',
#                         ifelse(name %in% sj_highcode, 'High',
#                                ifelse(name %in% ca_lowcode, 'Low', 'High')))) %>%
#   mutate(site = paste0(loc, '_', elevF)) %>%
#   mutate(site = as.factor(site),
#          loc = as.factor(loc),
#          elevF = as.factor(elevF))

topo <- topo %>%
  mutate(site = as.factor(case_when(
```

```
    name %in% sj_code ~ "sj",
    name %in% ca_highcode ~ "caH",
    name %in% ca_lowcode ~ "caL")))
```

- Compare data and export data into text files (see `/out/anovas_topo/`)

Elevation

Table 31: Mean values (mde)

site	mean	sd	se	min	max	tukey	variable
caH	1865	12.14	3.135	1846	1884	c	mde
caL	1719	21.9	5.655	1691	1751	b	mde
sj	1395	59.68	13.35	1322	1474	a	mde

Table 32: ANOVA table (mde)

term	df	sumsq	meansq	statistic	p.value
site	2	2044877	1022439	628.5	**0**
Residuals	47	76459	1627		

Slope

Table 33: Mean values (slope)

site	mean	sd	se	min	max	tukey	variable
caH	12.11	3.275	0.8457	6.853	18.23	a	slope
caL	12.86	2.984	0.7705	8.669	18.04	a	slope
sj	27.33	5.593	1.251	16.88	34.27	b	slope

Table 34: ANOVA table (slope)

term	df	sumsq	meansq	statistic	p.value
site	2	2647	1323	71.57	**0**
Residuals	47	869.1	18.49		

Focal tree summary

```
## Comparison
# Select only variables to compare
ft_sel <- ft %>%
  mutate(dn = dn_mm / 10,
         h = height_cm / 100) %>%
  dplyr::select(dn, h, site)
```

```r
# Get vector with variables
variables <- c('dn','h')

for (i in variables){

  # apply comparison
  out_compara <- compara(df=ft_sel, mivariable = i)

  out_name  <- paste0('aov_', i)
  assign(out_name, out_compara)

}


# Loop to export into txt files (see ./out/anovas_ft ... )
for (i in variables){

  out <- get(paste0('aov_', i))

  file_out <- file(here::here("out/anovas_ft", paste0("aov_", i, ".txt")), "w")
  sink(file_out)

  cat("MODEL \n")
  print(out$mymodel)
  cat("\n")

  cat("MODEL pretty \n")
  print(out$tm)
  cat("\n")

  cat("POST HOC \n")
  print(out$mymult)
  cat("\n")

  cat("SUMMARY VALUES \n")
  print(as.data.frame(out$summ_comparison))
  cat("\n")

  while (sink.number()>0) sink()
  # close(file_out)

}

while (sink.number()>0) sink()
```

dn Focal tree

Table 35: Mean values (dn)

site	mean	sd	se	min	max	tukey	variable
caH	69.75	20.51	5.297	44.56	122.5	c	dn

site	mean	sd	se	min	max	tukey	variable
caL	45.9	8.6	2.221	35.97	63.66	b	dn
sj	31.86	3.728	0.8335	26.42	39.79	a	dn

Table 36: ANOVA table (dn)

term	df	sumsq	meansq	statistic	p.value
site	2	12356	6178	40.38	**0**
Residuals	47	7191	153		

height Focal tree

Table 37: Mean values (h)

site	mean	sd	se	min	max	tukey	variable
caH	15.42	1.784	0.4606	11.9	18	b	h
caL	12.61	1.575	0.4065	9.5	14.7	a	h
sj	11.81	2.304	0.5153	5.7	16.2	a	h

Table 38: ANOVA table (h)

term	df	sumsq	meansq	statistic	p.value
site	2	117.4	58.68	15.31	**1e-05**
Residuals	47	180.1	3.833		

```
ft_summary <- ft_sel %>% group_by(site) %>%
  dplyr::summarise(h_mean = mean(h),
                   h_sd = sd(h),
                   dn_mean = mean(dn),
                   dn_sd = sd(dn)) %>%
  as.data.frame()

compe_summary <- compe %>%
  mutate(ba_ha = (ba * 10000)/(pi*100)) %>%
  group_by(site) %>%
  dplyr::select(ba_ha, std, srd, site) %>%
  dplyr::summarise(ba_mean = mean(ba_ha),
                   ba_sd = sd(ba_ha),
                   stand_density_mean = mean(std),
                   stand_density_sd = sd(std),
                   srd_mean = mean(srd),
                   srd_sd = sd(srd)) %>%
  as.data.frame()


# DBH and H of all trees (no focal tree)
```

```r
all <- nft %>% mutate(dn = dn_mm / 10,
                      h = height_cm / 100) %>%
  dplyr::select(dn, h, site) %>%
  bind_rows(ft_sel) %>%
  group_by(site) %>%
  dplyr::summarise(h_mean_all = mean(h),
                   h_sd_all = sd(h),
                   dn_mean_all = mean(dn),
                   dn_sd_all = sd(dn)) %>%
  as.data.frame()






stand_biotic <- ft_summary %>%
  inner_join(compe_summary, by = 'site') %>%
  inner_join(all, by='site')


# Export data
write.csv(stand_biotic, file=here::here("data/competence", "biotic_stand_features.csv"), row.names = FA

pander(stand_biotic)
```

Table 39: Table continues below

site	h_mean	h_sd	dn_mean	dn_sd	ba_mean	ba_sd	stand_density_mean
caH	15.42	1.784	69.75	20.51	39.13	24.31	348
caL	12.61	1.575	45.9	8.6	18.02	7.113	409.6
sj	11.81	2.304	31.86	3.728	11.64	5.466	339

Table 40: Table continues below

stand_density_sd	srd_mean	srd_sd	h_mean_all	h_sd_all	dn_mean_all
147.1	0.9085	0.629	10.8	4.352	34.12
226	0.8896	0.4385	9.02	2.847	21.69
130.3	1.111	0.5199	9.7	3.64	20.56

dn_sd_all
24.32
14.39
8.09

```r
# Join all data and export as table

# Get name of variables
variables <- compe %>% dplyr::select(ba:dnn) %>% names() %>% rbind()
```

```r
# Add topo and focal tree variables
nombres <- c(variables, 'mde','slope', 'dn', 'h')

out <- c()

for (i in nombres){
  # name of aov object
  aov_name <- paste0('aov_',i)
  # get aov_object
  aux <- get(aov_name)
  # get summ comparison dataframe
  aux2 <- aux$summ_comparison
  out <- rbind(out, aux2)
}

out2 <- out %>% dplyr::select(site, mean, sd, tukey, mivar = variable)


# See this solution
# http://stackoverflow.com/questions/39066811/long-to-wide-with-dplyr
# http://garrettgman.github.io/tidying/
out3 <- gather(out2, variable, value, mean, sd, tukey) %>%
  unite(var, variable, site) %>%
  spread(var, value, convert=TRUE)


# out4 <- out3 %>%
#   transmute(variables = mivar,
#             CA_High_m = round(mean_CA_High, 3),
#             CA_High_sd = round(sd_CA_High, 3),
#             CA_High_l = tukey_CA_High,
#             CA_Low_m = round(mean_CA_Low, 3),
#             CA_Low_sd = round(sd_CA_Low, 3),
#             CA_Low_l = tukey_CA_Low,
#             SJ_High_m = round(mean_SJ_High, 3),
#             SJ_High_sd = round(sd_SJ_High, 3),
#             SJ_High_l = tukey_SJ_High,
#             SJ_Low_m = round(mean_SJ_Low, 3),
#             SJ_Low_sd = round(sd_SJ_Low, 3),
#             SJ_Low_l = tukey_SJ_Low)

out4 <- out3 %>%
  transmute(variables = mivar,
            caH_m = round(mean_caH, 3),
            caH_sd = round(sd_caH, 3),
            caH_l = tukey_caH,
            caL_m = round(mean_caL, 3),
            caL_sd = round(sd_caL, 3),
            caL_l = tukey_caL,
            sj_m = round(mean_sj, 3),
            sj_sd = round(sd_sj, 3),
            sj_l = tukey_sj)
```

```r
# Export data
# write.csv(out4, file=paste(di, "data/proto_tables/site_features.csv", sep=""), row.names = FALSE)
write.csv(out4, file=here::here("data/competence","resumen_medias_indices_competencia.csv"), row.names =

pander(out4, caption='Summary Values')
```

Table 42: Summary Values (continued below)

variables	caH_m	caH_sd	caH_l	caL_m	caL_sd	caL_l	sj_m	sj_sd	sj_l
ba	1.229	0.764	b	0.566	0.223	a	0.366	0.172	a
crowding	0.717	0.472	a	0.482	0.184	a	0.468	0.252	a
dn	69.75	20.51	c	45.9	8.6	b	31.86	3.728	a
dnn	3.412	1.859	a	3.123	1.308	a	2.4	1.328	a
h	15.42	1.784	b	12.61	1.575	a	11.81	2.304	a
lorimer	7.664	4.711	a	7.779	3.937	a	9.345	4.07	a
mde	1865	12.14	c	1719	21.9	b	1395	59.68	a
n_competitors	10.93	4.621	a	12.87	7.1	a	10.65	4.095	a
n_competitors_higher	1.2	1.474	a	0.667	0.816	a	0.5	0.688	a
nesr	0.059	0.085	a	0.035	0.032	a	0.077	0.095	a
newsr	0.484	0.622	a	0.179	0.223	a	0.397	0.689	a
pd	0.035	0.015	a	0.041	0.023	a	0.034	0.013	a
sdd	-0.493	0.402	a	-0.414	0.229	ab	-0.209	0.097	b
slope	12.11	3.275	a	12.86	2.984	a	27.32	5.593	b
sr	0.215	0.196	a	0.171	0.068	a	0.151	0.064	a
srd	0.908	0.629	a	0.89	0.438	a	1.111	0.52	a
std	348	147.1	a	409.6	226	a	339	130.3	a
sum_sizes	3.374	1.444	b	2.549	1.073	ab	2.077	0.866	a

General topo of the sites

```r
topo_site <- topo %>% group_by(site) %>% summarise(
  elev_mean = mean(mde),
  elev_min = min(mde),
  elev_max = max(mde),
  elev_sd = sd(mde),
  elev_se = sd(mde)/sqrt(n()),
  slope_mean = mean(slope),
  slope_sd = sd(slope),
  slope_se = sd(slope)/sqrt(n()),
)


site_general <- general_var %>%
  left_join(coord_sites, by='site') %>%
  left_join(topo_site, by='site') %>%
  as.data.frame()

pander(site_general)
```

Table 44: Table continues below

site	n	lat_m	long_m	elev_mean	elev_min	elev_max	elev_sd
caH	15	36.97	-3.421	1865	1846	1884	12.14
caL	15	36.96	-3.424	1719	1691	1751	21.9
sj	20	37.13	-3.375	1395	1322	1474	59.68

elev_se	slope_mean	slope_sd	slope_se
3.135	12.11	3.275	0.8457
5.655	12.86	2.984	0.7705
13.35	27.33	5.593	1.251

```r
write.csv(site_general, file=here::here("data/sites_summary", "site_general.csv"), row.names = FALSE)
```