Analysis Stand Features

AJ Perez-Luque (@ajpelu)

2017 May

Prepare Data

- Two datasets: focal tree and competence

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


# 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))


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

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

General Variables

Numbers of focal trees by site

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

general_var %>% kable
```

loc	site	n
CA	CA_High	15
CA	CA_Low	15
SJ	SJ_High	10

loc	site	n
SJ	SJ_Low	10

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)

## 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()
```

site	lat_m	long_m
SJ_High	37.12916	-3.365722
SJ_Low	37.13326	-3.384094
CA_High	36.96613	-3.420703
CA_Low	36.95645	-3.424107

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
CA_High	1.2293916	0.7636880	0.19718340	0.0118272	2.8166207	b	ba
CA_Low	0.5661200	0.2234657	0.05769860	0.2512500	0.8633280	a	ba
SJ_High	0.4510437	0.1808913	0.05720286	0.2212393	0.7657899	a	ba
SJ_Low	0.2804943	0.1156196	0.03656213	0.1205221	0.4492367	a	ba

Table 4: ANOVA table (ba)

term	df	sumsq	meansq	statistic	p.value
site	3	6.841	2.28	11.3	**1e-05**
Residuals	46	9.279	0.2017		

Stand Density

Table 5: Mean values (std)

site	mean	sd	se	min	max	tukey	variable
CA_High	348.0188	147.0867	37.97761	63.66198	541.1268	a	std
CA_Low	409.5587	225.9990	58.35268	159.15494	1050.4226	a	std
SJ_High	404.2536	119.1479	37.67788	286.47890	636.6198	a	std
SJ_Low	273.7465	110.4698	34.93361	127.32395	445.6338	a	std

Table 6: ANOVA table (std)

term	df	sumsq	meansq	statistic	p.value
site	3	132562	44187	1.619	0.1979
Residuals	46	1255538	27294		

Plot Density

Table 7: Mean values (pd)

site	mean	sd	se	min	max	tukey	variable
CA_High	0.03480188	0.01470867	0.003797761	0.006366198	0.05411268	a	pd
CA_Low	0.04095587	0.02259990	0.005835268	0.015915494	0.10504226	a	pd
SJ_High	0.04042536	0.01191479	0.003767788	0.028647890	0.06366198	a	pd
SJ_Low	0.02737465	0.01104698	0.003493361	0.012732395	0.04456338	a	pd

Table 8: ANOVA table (pd)

term	df	sumsq	meansq	statistic	p.value
site	3	0.00133	0.00044	1.619	0.1979
Residuals	46	0.01256	0.00027		

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

Table 9: Mean values (n_competitors)

site	mean	sd	se	min	max	tukey	variable
CA_High	10.93333	4.620864	1.193102	2	17	a	n_competitors
CA_Low	12.86667	7.099966	1.833203	5	33	a	n_competitors
SJ_High	12.70000	3.743142	1.183685	9	20	a	n_competitors
SJ_Low	8.60000	3.470511	1.097472	4	14	a	n_competitors

Table 10: ANOVA table (n_competitors)

term	df	sumsq	meansq	statistic	p.value
site	3	130.8	43.61	1.619	0.1979
Residuals	46	1239	26.94		

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
CA_High	1.2000000	1.4735768	0.3804759	0	4	a	n_competitors_higher
CA_Low	0.6666667	0.8164966	0.2108185	0	2	a	n_competitors_higher
SJ_High	0.5000000	0.7071068	0.2236068	0	2	a	n_competitors_higher
SJ_Low	0.5000000	0.7071068	0.2236068	0	2	a	n_competitors_higher

Table 12: ANOVA table (n_competitors_higher)

term	df	sumsq	meansq	statistic	p.value
site	3	4.387	1.462	1.38	0.2607

term	df	sumsq	meansq	statistic	p.value
Residuals	46	48.73	1.059		

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

Table 13: Mean values (sum_sizes)

site	mean	sd	se	min	max	tukey	variable
CA_High	3.374255	1.4444380	0.3729523	0.1734789	6.200677	b	sum_sizes
CA_Low	2.549089	1.0733727	0.2771437	1.2748311	5.280124	ab	sum_sizes
SJ_High	2.524675	0.8782956	0.2777415	1.5931410	4.268536	ab	sum_sizes
SJ_Low	1.629428	0.6046458	0.1912058	0.8323804	2.334803	a	sum_sizes

Table 14: ANOVA table (sum_sizes)

term	df	sumsq	meansq	statistic	p.value
site	3	18.5	6.168	5.106	**0.00392**
Residuals	46	55.57	1.208		

Size ratio

Table 15: Mean values (sr)

site	mean	sd	se	min	max	tukey	variable
CA_High	0.2153327	0.19568030	0.050524437	0.07941010	0.8759954	a	sr
CA_Low	0.1705044	0.06824190	0.017619982	0.06483256	0.3308519	a	sr
SJ_High	0.1184228	0.02789219	0.008820286	0.06615599	0.1521739	a	sr
SJ_Low	0.1840603	0.07446098	0.023546628	0.10928962	0.2960969	a	sr

Table 16: ANOVA table (sr)

term	df	sumsq	meansq	statistic	p.value
site	3	0.05746	0.01915	1.339	0.2735
Residuals	46	0.6582	0.01431		

Distance-Dependet Indices

Distance to nearest tree

Table 17: Mean values (dnn)

site	mean	sd	se	min	max	tukey	variable
CA_High	3.412000	1.859444	0.4801063	0.88	6.75	a	dnn
CA_Low	3.122667	1.307792	0.3376705	1.44	5.53	a	dnn

site	mean	sd	se	min	max	tukey	variable
SJ_High	2.285000	1.467971	0.4642132	0.67	4.90	a	dnn
SJ_Low	2.514000	1.240100	0.3921542	1.03	4.99	a	dnn

Table 18: ANOVA table (dnn)

term	df	sumsq	meansq	statistic	p.value
site	3	9.928	3.309	1.442	0.2429
Residuals	46	105.6	2.295		

Crowding

Table 19: Mean values (crowding)

site	mean	sd	se	min	max	tukey	variable
CA_High	0.7169079	0.4719477	0.12185637	0.02151754	1.6399550	b	crowding
CA_Low	0.4818314	0.1844421	0.04762275	0.25348648	0.9068871	ab	crowding
SJ_High	0.6034893	0.2735479	0.08650345	0.26860548	1.0260638	ab	crowding
SJ_Low	0.3327240	0.1362680	0.04309171	0.19222016	0.6687375	a	crowding

Table 20: ANOVA table (crowding)

term	df	sumsq	meansq	statistic	p.value
site	3	0.9878	0.3293	3.415	**0.02501**
Residuals	46	4.435	0.09642		

Lorimer

Table 21: Mean values (lorimer)

site	mean	sd	se	min	max	tukey	variable
CA_High	7.664087	4.710569	1.2162638	0.1569462	18.61649	a	lorimer
CA_Low	7.778860	3.937159	1.0165702	2.7494835	18.51367	a	lorimer
SJ_High	11.528850	3.811393	1.2052684	7.3519688	19.28918	a	lorimer
SJ_Low	7.160582	3.135968	0.9916801	3.7963327	12.97633	a	lorimer

Table 22: ANOVA table (lorimer)

term	df	sumsq	meansq	statistic	p.value
site	3	127.1	42.38	2.61	**0.06273**
Residuals	46	746.9	16.24		

Negative Exponential size ratio

Table 23: Mean values (nesr)

site	mean	sd	se	min	max	tukey	variable
CA_High	0.05864130	0.08498713	0.021943582	3.101805e-05	0.2508758	a	nesr
CA_Low	0.03478072	0.03247587	0.008385235	1.274032e-03	0.1035436	a	nesr
SJ_High	0.10784464	0.11721176	0.037065613	3.389832e-03	0.3421948	a	nesr
SJ_Low	0.04631867	0.05507588	0.017416522	3.519526e-03	0.1908227	a	nesr

Table 24: ANOVA table (nesr)

term	df	sumsq	meansq	statistic	p.value
site	3	0.03427	0.01142	1.969	0.1318
Residuals	46	0.2668	0.0058		

Negative Exponential Weighted size ratio

Table 25: Mean values (newsr)

site	mean	sd	se	min	max	tukey	variable
CA_High	0.4837733	0.6224122	0.16070614	0.0009384470	1.6811993	a	newsr
CA_Low	0.1792381	0.2232725	0.05764871	0.0002055700	0.7879362	a	newsr
SJ_High	0.6168525	0.9008731	0.28488108	0.0004798334	2.5701475	a	newsr
SJ_Low	0.1762010	0.2855108	0.09028646	0.0022657241	0.8746652	a	newsr

Table 26: ANOVA table (newsr)

term	df	sumsq	meansq	statistic	p.value
site	3	1.717	0.5724	1.86	0.1497
Residuals	46	14.16	0.3078		

Size ratio proportional to distance

Table 27: Mean values (srd)

site	mean	sd	se	min	max	tukey	variable
CA_High	0.9084877	0.6290275	0.1624142	0.01556593	2.379239	a	srd
CA_Low	0.8896226	0.4384576	0.1132093	0.32131106	2.060190	a	srd
SJ_High	1.3883389	0.5049906	0.1596921	0.81700518	2.246604	a	srd
SJ_Low	0.8345799	0.3811389	0.1205267	0.47940177	1.668270	a	srd

Table 28: ANOVA table (srd)

term	df	sumsq	meansq	statistic	p.value
site	3	2.077	0.6924	2.692	**0.05708**
Residuals	46	11.83	0.2572		

Size difference proportional to distance

Table 29: Mean values (sdd)

site	mean	sd	se	min	max	tukey	variable
CA_High	-0.4931479	0.4017918	0.10374219	-1.2261101	-0.06249655	a	sdd
CA_Low	-0.4142691	0.2291280	0.05916059	-1.0221632	-0.08533653	ab	sdd
SJ_High	-0.2353267	0.1100053	0.03478672	-0.4064358	-0.08260647	ab	sdd
SJ_Low	-0.1827803	0.0795404	0.02515288	-0.3214838	-0.08321985	b	sdd

Table 30: ANOVA table (sdd)

term	df	sumsq	meansq	statistic	p.value
site	3	0.7787	0.2596	3.778	**0.01665**
Residuals	46	3.161	0.06872		

Topographic data

- Read data

```
# Read topo data
topo <- read.csv(file=paste(di, "/data/topo/topo.csv", sep=""), 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_summary <- topo %>%
#   group_by(site) %>%
#   summarise(mde_m = mean(mde),
#             mde_sd = sd(mde),
#             mde_min = min(mde),
#             mde_max = max(mde),
#             slope_m = mean(slope),
#             slope_sd = sd(slope))
```

```
#
#
#
# # Another way to obtain summary values
# variables <- c('mde','slope','aspect')
# auxdf <- data.frame()
#
# for (i in variables){
# aux <- topo %>%
#   dplyr::group_by(site) %>%
#   summarise_each_(funs(mean, sd, se=sd(.)/sqrt(n())), i) %>% mutate(variable=i)
#
# auxdf <- rbind(auxdf, aux) }
```

- 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
CA_High	1864.649	12.14112	3.134824	1845.60	1883.88	d	mde
CA_Low	1718.537	21.90183	5.655029	1691.31	1750.80	c	mde
SJ_High	1450.954	22.55849	7.133620	1417.74	1473.53	b	mde
SJ_Low	1339.607	11.10027	3.510213	1322.49	1355.49	a	mde

Table 32: ANOVA table (mde)

term	df	sumsq	meansq	statistic	p.value
site	3	2106868	702289	2233	**0**
Residuals	46	14468	314.5		

Slope

Table 33: Mean values (slope)

site	mean	sd	se	min	max	tukey	variable
CA_High	12.11195	3.275225	0.8456595	6.852460	18.22588	a	slope
CA_Low	12.85756	2.983954	0.7704537	8.668827	18.03590	a	slope
SJ_High	32.27118	1.553920	0.4913926	29.341582	34.27049	c	slope
SJ_Low	22.37899	3.042151	0.9620127	16.879122	26.38996	b	slope

Table 34: ANOVA table (slope)

term	df	sumsq	meansq	statistic	p.value
site	3	3136	1045	126.6	**0**
Residuals	46	379.9	8.258		

Focal tree summary

```
## Comparison
# Select only variables to compare
ft_sel <- ft %>%
  mutate(dn = dn_mm / 1000,
         h = height_cm / 100) %>%
  dplyr::select(dn, h, site)


# 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(paste0(di,'out/anovas_ft/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
CA_High	0.6975231	0.20514102	0.05296718	0.4456338	1.2254931	c	dn
CA_Low	0.4590029	0.08600428	0.02220621	0.3596902	0.6366198	b	dn
SJ_High	0.3157634	0.03819129	0.01207715	0.2641972	0.3978874	a	dn
SJ_Low	0.3214930	0.03816770	0.01206969	0.2705634	0.3724226	a	dn

Table 36: ANOVA table (dn)

term	df	sumsq	meansq	statistic	p.value
site	3	1.236	0.4119	26.36	**0**
Residuals	46	0.7189	0.01563		

height Focal tree

Table 37: Mean values (h)

site	mean	sd	se	min	max	tukey	variable
CA_High	15.42000	1.783736	0.4605587	11.9	18.0	b	h
CA_Low	12.60667	1.574560	0.4065496	9.5	14.7	a	h
SJ_High	10.94000	2.344829	0.7415000	5.7	13.1	a	h
SJ_Low	12.69000	2.002471	0.6332368	10.5	16.2	a	h

Table 38: ANOVA table (h)

term	df	sumsq	meansq	statistic	p.value
site	3	132.7	44.22	12.34	**0**
Residuals	46	164.8	3.583		