Appendix

for

'Examining Fouquieria splendens in an environmental and ecological context: Effect of topography and interspecific neighbors on ocotillo morphology and distribution';

Anastasia Bernat, Acacia Tsz So Tang, Allegra Steenson, Eric Larsen, Emma Greig

Contents

1 Details of the Analyses

This document was generated by R Markdown on 2022-12-03 using R version 4.2.1 (2022-06-23). The document provides the step-by-step analytical methods used in the manuscript by Anastasia Bernat (AVB), Acacia Tsz So Tang (ATST), Allegra Steenson (AS), Eric Larsen (EL), and Emma Greig (EG). Draft scripts were written by AVB and ATST between 2019-06-01 and 2021-01-01 until being distilled and complied by AVB at the University of Chicago into this comprehensive script. All draft scripts can be viewed in the GitHub repository, ocotillo-research (https://github.com/avbernat/ocotillo-research).

All code and output from the statistical analyses are shown. Code for data cleaning and the generation of plots is not displayed, but can be viewed in the **appendix.Rmd** file and its accompanying sourced scripts. To repeat analyses and the generation of plots, all data files and sourced scripts should follow the directory structure presented in the ocotillo-research repository.

1.1 Description of the Data

Ocotillos, *Fouquieria splendens*, were measured in Summer 2019 in the Sonoran Desert at Organ Pipe Cactus National Monument

1.2 Abbreviations Used in the Data and Code

1.3 Data Transformations

- `_b` - a column name that ends in `_b` is a column that has been recodified into binary data (0's and 1's). Example columns:
- `_c` - a column name that ends in `_c` is a column that has been centered. Example columns:
- `_s` - a column name that ends in `_s` is a column that has been standardized. Example columns:
- `log` - a column name that starts with `log` is a column that has been log transformed. Example columns:
- `_baj` - a dataset that ends in `_baj` is a dataset that only contains ocotillo measurements from ocotillos across a bajada. Example datasets: ocos_baj, segs_baj

1.4 Read in Libraries

```
library(dplyr)
library(outliers) # dixon.test
# library(ggplot2)
# library(olsrr)
# library(lme4)
#
# library(tidyverse)
# library(ggpubr)
# library(rstatix)
#
# library(FactoMineR)
# library(factoextra)
# library(readr)
# library(corrplot)
```

1.5 Read Source Scripts

```
source("src/cleaning_data.R")
source("src/regression_output.R")
source("src/diagnostics.R")
source("src/pretty_reg.R")
```

1.6 Read the Data

```
ocotillo_data = read.csv("data/General_Oco_Data3.csv",
                         fileEncoding="UTF-8-BOM",  stringsAsFactors=TRUE)
segment_data = read.csv("data/Terminal_5Segs.csv",
                         fileEncoding="UTF-8-BOM",  stringsAsFactors=TRUE)

ocos = clean_ocos_data(ocotillo_data)               # all ocotillos
ocos_baj = ocos[1:20,]                              # only ocotillos across the bajada
segs = clean_segs_data(segment_data, ocotillo_data) # all ocotillos
segs_baj = segs[1:1000,]                            # only ocotillos across the bajada
```

1.7 Normality

All measurements followed log-normal distributions except for circumference, median branch length, number of nodes, and distance to the nearest arroyo, which were normally distributed. Ocotillo data were log-transformed before analyses to meet assumptions of normality, linear regressions, and homogeneity for parametric analyses.

1.8 Outliers

Terminal segment length IQR has an outlier - Ocotillo 1 (see graphs.Rmd). In turn, the outlier was removed and logHeight is predicted with a smaller dataset containing all ocotillos except Ocotillo 1.

```
dixon.test(ocos$Terminal_SegIQR)
```

```
##
##  Dixon test for outliers
##
## data:  ocos$Terminal_SegIQR
## Q = 0.67857, p-value < 2.2e-16
## alternative hypothesis: highest value 19 is an outlier
```

```
ocos_data = ocotillo_data[ocotillo_data$Tree != 1,]
segs_data = segment_data[segment_data$Tree != 1,]

# rerun data cleaning to generate newly transformed columns
ocos = clean_ocos_data(ocos_data)
ocos_baj = ocos[1:20,]
segs = clean_segs_data(segs_data, ocos_data)
segs_baj = segs[1:1000,]
```

2 Ocotillo Morphology

Analyses below are multiple variate models of Fouquieria splendens morphology for ocotillos located on both a bajada and a plain in Organ Pipe National Monument, Arizona. All models were grouped by their response variable and ordered by their ascending AIC values. Dataset "ocos" indicates all individuals measured on the bajada and plain while "ocos_baj" indicates only the individuals measured on the bajada. Ocotillos located on the bajada were encoded with site = 0 while ocotillos on the plain were encoded with site = 1. Interspecific neighbor group is split between two types – shrub and cactus – where cactus = 0 and shrub = 1.

2.1 Principal Component Analysis

2.2 Multiple Variate Modeling

2.2.1 Testing Covariates

Only significant covariates are listed below:

```
output_color = FALSE
```

```
## lm logHeight ~ Elevation_c ocos
## AIC:  -8.591236
## (Intercept)  coeff:  1.0939148   Pr(>|t|):  2.62618e-20 *
## Elevation_c  coeff:  -0.0052053  Pr(>|t|):  0.001364881 *

## lm Height_c ~ Site ocos
## AIC:  41.70617
## (Intercept)  coeff:  -0.3335223  Pr(>|t|):  0.007768289 *
## Site         coeff:  1.238797    Pr(>|t|):  9.21175e-06 *

## lm logHeight ~ Site ocos
## AIC:  -16.71477
## (Intercept)  coeff:  0.9937884   Pr(>|t|):  2.496203e-19 *
## Site         coeff:  0.3718982   Pr(>|t|):  2.708724e-05 *

## lm Height ~ Median_BL_c ocos
## AIC:  46.26668
## (Intercept)  coeff:  3.0661538   Pr(>|t|):  4.632554e-20 *
## Median_BL_c  coeff:  0.0110914   Pr(>|t|):  8.101066e-05 *

## lm logHeight ~ Median_BL_c ocos
## AIC:  -15.83864
## (Intercept)  coeff:  1.0939148   Pr(>|t|):  9.99643e-22 *
## Median_BL_c  coeff:  0.0035489   Pr(>|t|):  4.114291e-05 *

## lm Height ~ NumNodes_c ocos
## AIC:  52.55392
## (Intercept)  coeff:  3.0661538   Pr(>|t|):  7.712155e-19 *
## NumNodes_c   coeff:  0.0106211   Pr(>|t|):  0.001707761 *

## lm logHeight ~ NumNodes_c ocos
## AIC:  -8.057626
## (Intercept)  coeff:  1.0939148   Pr(>|t|):  3.337813e-20 *
## NumNodes_c   coeff:  0.0032965   Pr(>|t|):  0.001774288 *
```