ESA\_Recovery\_overview\_doc

Annabelle

6/7/2021

##Introduction We are trying to answer the following Qs (which are the same as in the “Overview” doc) 1. What are the conservation initiatives for a species 2. What are the actions specified in the federal register important for the species? 3. Who are the organizations/players involved in doing these conservation initiatives? 4. What are the actions each organization is doing (optional?)

This document has the summary statistics for questions 1 & 2 (added 06/07/2021) and each section is question specific

Notes - Slight issue - missing at least 10 species data from tyler for initial draft of data (10 out of 38 species) … is this data accessible from the drives??

##Set up file structure

##load in different datasets

##Modify datasets

##Modify datasets Tyler Variables needed: # range size –> area c32 # area weighted footprint –> c33 ?have options # total threat count –> c 26 # type of threat –> c27 # employment (all relevant industries) -> 55:59 [avg\_employment ???] # percent public and percent federal c66, c68 # diversity of landowners - shannons H (land use complextity) # taxa (plants v animals // vert v invert) taxa c16 # percent public private lands now added (6/27)

missing species (presumming Tyler dropped them..) # Euphilotes pallescens arenamontana # Nysius wekiuicola

Species Name changes/differences accross datasets Chorizanthe parryi var fernandina -> Chorizanthe parryi var. fernandina Moxostoma sp 2 -> Moxostoma sp.

## 1. What are the conservation initiatives for a species

#### 1 - Summary stats

* Number of iniatitves per species

# iniatives per species (mean, median, quartiles)

* Count of efforts, agreements, groups per species

#### 1 - Regression Predictor Modification

* Working group mentioned – logistic regression
* Numbers of efforts or agreements – linear regression or negative binomial or poisson regression (since count data)

### Base model

* Note: percent public variable is still in matrix but was removed from model

VIFS

model summary - variations

Robust Standard Errors in R

Trying stargazer

lmg is the R^2 contribution averaged over orderings amoung regressors

## 2. What are the actions specified in the federal register important for the species?

#### 1 - Summary stats

Count of outcomes as proactive vs responsive

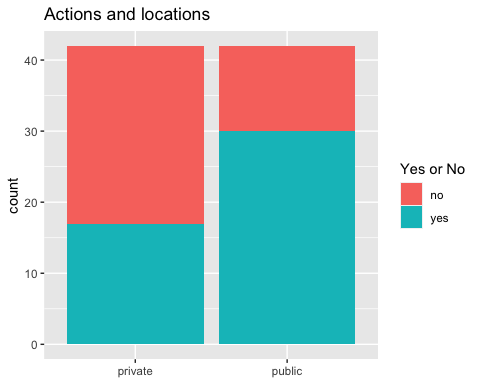
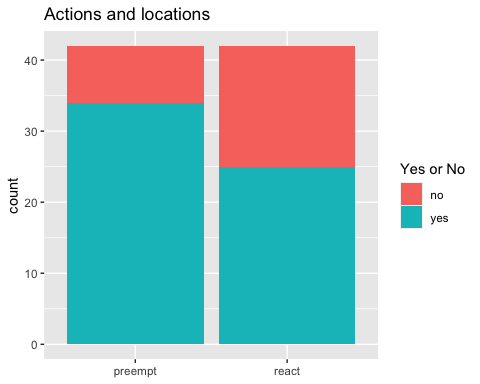
Count of outcomes on public vs private land

|  |  |  |
| --- | --- | --- |
| one | yes | no |
| preempt | 34 | 8 |
| react | 25 | 17 |
| public | 30 | 12 |
| private | 17 | 25 |



##   
## Pearson's Chi-squared test with Yates' continuity correction  
##   
## data: test\_act  
## X-squared = 3.6447, df = 1, p-value = 0.05625

##   
## Fisher's Exact Test for Count Data  
##   
## data: test\_act  
## p-value = 0.0551  
## alternative hypothesis: true odds ratio is not equal to 1  
## 95 percent confidence interval:  
## 0.980416 8.939537  
## sample estimates:  
## odds ratio   
## 2.852948



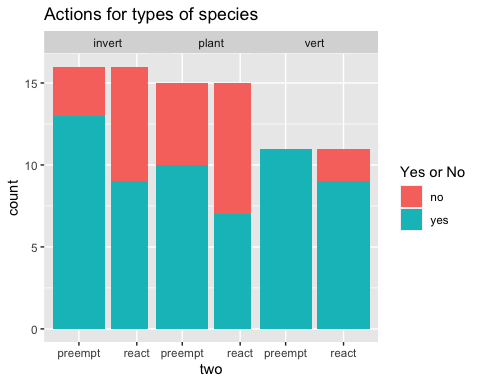
##   
## Pearson's Chi-squared test with Yates' continuity correction  
##   
## data: test\_land  
## X-squared = 6.9557, df = 1, p-value = 0.008355

##   
## Fisher's Exact Test for Count Data  
##   
## data: test\_land  
## p-value = 0.00795  
## alternative hypothesis: true odds ratio is not equal to 1  
## 95 percent confidence interval:  
## 1.35466 10.13523  
## sample estimates:  
## odds ratio   
## 3.615964

#### 

#### Taxa

##   
## Pearson's Chi-squared test  
##   
## data: test\_taxa\_actions  
## X-squared = 11.871, df = 5, p-value = 0.0366



|  |  |  |  |
| --- | --- | --- | --- |
| one | two | yes | no |
| plant | preempt | 10 | 5 |
| plant | react | 7 | 8 |

## # A tibble: 1 x 3  
## n p p.signif  
## \* <dbl> <dbl> <chr>   
## 1 30 0.462 ns

|  |  |  |  |
| --- | --- | --- | --- |
| one | two | yes | no |
| vert | preempt | 11 | 0 |
| vert | react | 9 | 2 |

## # A tibble: 1 x 3  
## n p p.signif  
## \* <dbl> <dbl> <chr>   
## 1 22 0.476 ns

|  |  |  |  |
| --- | --- | --- | --- |
| one | two | yes | no |
| plant | preempt | 10 | 5 |
| vert | preempt | 11 | 0 |
| invert | preempt | 13 | 3 |

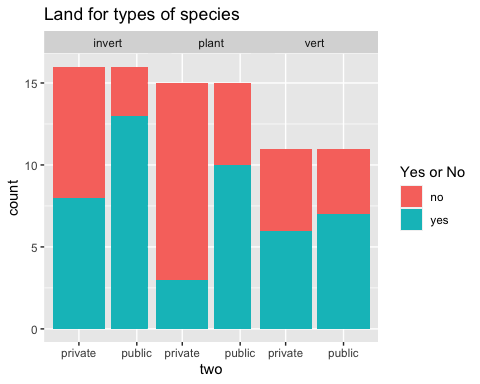
## # A tibble: 3 x 6  
## group1 group2 n p p.adj p.adj.signif  
## \* <chr> <chr> <dbl> <dbl> <dbl> <chr>   
## 1 row1 row2 26 0.0527 0.158 ns   
## 2 row1 row3 31 0.433 1 ns   
## 3 row2 row3 27 0.248 0.744 ns

|  |  |  |  |
| --- | --- | --- | --- |
| one | two | yes | no |
| plant | react | 7 | 8 |
| vert | react | 9 | 2 |
| invert | react | 9 | 7 |

## # A tibble: 3 x 6  
## group1 group2 n p p.adj p.adj.signif  
## \* <chr> <chr> <dbl> <dbl> <dbl> <chr>   
## 1 row1 row2 26 0.109 0.327 ns   
## 2 row1 row3 31 0.724 1 ns   
## 3 row2 row3 27 0.231 0.693 ns

##

## Pearson's Chi-squared test  
##   
## data: test\_taxa\_land  
## X-squared = 13.223, df = 5, p-value = 0.02138



|  |  |  |  |
| --- | --- | --- | --- |
| one | two | yes | no |
| plant | public | 10 | 5 |
| plant | private | 3 | 12 |

## # A tibble: 1 x 3  
## n p p.signif  
## \* <dbl> <dbl> <chr>   
## 1 30 0.0253 \*

|  |  |  |  |
| --- | --- | --- | --- |
| one | two | yes | no |
| vert | public | 7 | 4 |
| vert | private | 6 | 5 |

## # A tibble: 1 x 3  
## n p p.signif  
## \* <dbl> <dbl> <chr>   
## 1 22 1 ns

|  |  |  |  |
| --- | --- | --- | --- |
| one | two | yes | no |
| invert | public | 13 | 3 |
| invert | private | 8 | 8 |

## # A tibble: 1 x 3  
## n p p.signif  
## \* <dbl> <dbl> <chr>   
## 1 32 0.135 ns

|  |  |  |  |
| --- | --- | --- | --- |
| one | two | yes | no |
| plant | public | 10 | 5 |
| vert | public | 7 | 4 |
| invert | public | 13 | 3 |

## # A tibble: 3 x 6  
## group1 group2 n p p.adj p.adj.signif  
## \* <chr> <chr> <dbl> <dbl> <dbl> <chr>   
## 1 row1 row2 26 1 1 ns   
## 2 row1 row3 31 0.433 1 ns   
## 3 row2 row3 27 0.391 1 ns

|  |  |  |  |
| --- | --- | --- | --- |
| one | two | yes | no |
| plant | private | 3 | 12 |
| vert | private | 6 | 5 |
| invert | private | 8 | 8 |

## # A tibble: 3 x 6  
## group1 group2 n p p.adj p.adj.signif  
## \* <chr> <chr> <dbl> <dbl> <dbl> <chr>   
## 1 row1 row2 26 0.103 0.309 ns   
## 2 row1 row3 31 0.135 0.405 ns   
## 3 row2 row3 27 1 1 ns

#### 

#### Land actions

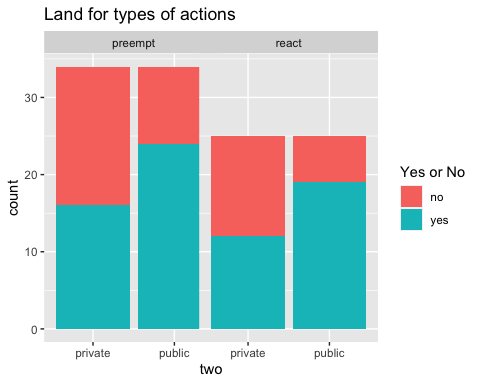
##   
## Pearson's Chi-squared test  
##   
## data: test\_landactions  
## X-squared = 8.1376, df = 3, p-value = 0.04325

|  |  |  |  |
| --- | --- | --- | --- |
| one | two | yes | no |
| preempt | public | 24 | 10 |
| preempt | private | 16 | 18 |

## # A tibble: 1 x 3  
## n p p.signif  
## \* <dbl> <dbl> <chr>   
## 1 68 0.0837 ns

|  |  |  |  |
| --- | --- | --- | --- |
| one | two | yes | no |
| react | public | 19 | 6 |
| react | private | 12 | 13 |

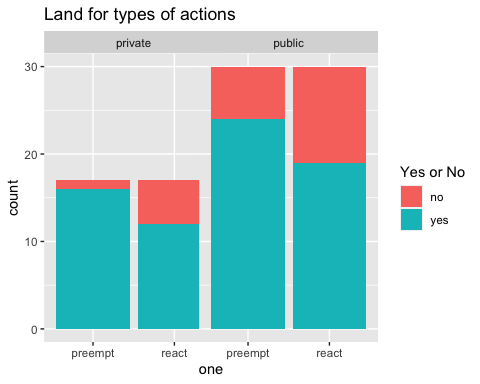
## # A tibble: 1 x 3  
## n p p.signif  
## \* <dbl> <dbl> <chr>   
## 1 50 0.0792 ns



##

## Pearson's Chi-squared test  
##   
## data: test\_landactions  
## X-squared = 6.1418, df = 3, p-value = 0.1049

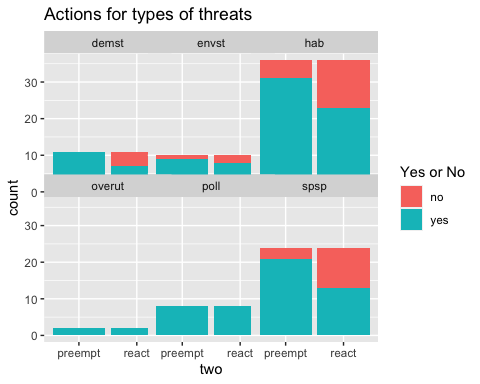
## # A tibble: 4 x 5  
## group n p p.adj p.adj.signif  
## \* <chr> <dbl> <dbl> <dbl> <chr>   
## 1 1 94 0.61 1 ns   
## 2 2 94 0.0622 0.249 ns   
## 3 3 94 0.0742 0.297 ns   
## 4 4 94 0.756 1 ns



####threat types

Note - missing three species because don’t have data

##   
## Pearson's Chi-squared test  
##   
## data: test\_threat\_actions  
## X-squared = 26.151, df = 11, p-value = 0.006163



|  |  |  |  |
| --- | --- | --- | --- |
| one | two | yes | no |
| hab | preempt | 31 | 5 |
| overut | preempt | 2 | 0 |
| poll | preempt | 8 | 0 |
| spsp | preempt | 21 | 3 |
| envst | preempt | 9 | 1 |
| demst | preempt | 11 | 0 |

## # A tibble: 15 x 6  
## group1 group2 n p p.adj p.adj.signif  
## \* <chr> <chr> <dbl> <dbl> <dbl> <chr>   
## 1 row1 row2 38 1 1 ns   
## 2 row1 row3 44 0.566 1 ns   
## 3 row1 row4 60 1 1 ns   
## 4 row1 row5 46 1 1 ns   
## 5 row1 row6 47 0.322 1 ns   
## 6 row2 row3 10 1 1 ns   
## 7 row2 row4 26 1 1 ns   
## 8 row2 row5 12 1 1 ns   
## 9 row2 row6 13 1 1 ns   
## 10 row3 row4 32 0.555 1 ns   
## 11 row3 row5 18 1 1 ns   
## 12 row3 row6 19 1 1 ns   
## 13 row4 row5 34 1 1 ns   
## 14 row4 row6 35 0.536 1 ns   
## 15 row5 row6 21 0.476 1 ns

|  |  |  |  |
| --- | --- | --- | --- |
| one | two | yes | no |
| hab | react | 23 | 13 |
| overut | react | 2 | 0 |
| poll | react | 8 | 0 |
| spsp | react | 13 | 11 |
| envst | react | 8 | 2 |
| demst | react | 7 | 4 |

## # A tibble: 15 x 6  
## group1 group2 n p p.adj p.adj.signif  
## \* <chr> <chr> <dbl> <dbl> <dbl> <chr>   
## 1 row1 row2 38 0.538 1 ns   
## 2 row1 row3 44 0.082 1 ns   
## 3 row1 row4 60 0.592 1 ns   
## 4 row1 row5 46 0.46 1 ns   
## 5 row1 row6 47 1 1 ns   
## 6 row2 row3 10 1 1 ns   
## 7 row2 row4 26 0.492 1 ns   
## 8 row2 row5 12 1 1 ns   
## 9 row2 row6 13 1 1 ns   
## 10 row3 row4 32 0.0292 0.438 ns   
## 11 row3 row5 18 0.477 1 ns   
## 12 row3 row6 19 0.103 1 ns   
## 13 row4 row5 34 0.251 1 ns   
## 14 row4 row6 35 0.721 1 ns   
## 15 row5 row6 21 0.635 1 ns

|  |  |  |  |
| --- | --- | --- | --- |
| one | two | yes | no |
| hab | preempt | 31 | 5 |
| hab | react | 23 | 13 |

## # A tibble: 1 x 3  
## n p p.signif  
## \* <dbl> <dbl> <chr>   
## 1 72 0.055 ns

|  |  |  |  |
| --- | --- | --- | --- |
| one | two | yes | no |
| overut | preempt | 2 | 0 |
| overut | react | 2 | 0 |

## # A tibble: 1 x 3  
## n p p.signif  
## \* <dbl> <dbl> <chr>   
## 1 4 1 ns

|  |  |  |  |
| --- | --- | --- | --- |
| one | two | yes | no |
| poll | preempt | 8 | 0 |
| poll | react | 8 | 0 |

## # A tibble: 1 x 3  
## n p p.signif  
## \* <dbl> <dbl> <chr>   
## 1 16 1 ns

|  |  |  |  |
| --- | --- | --- | --- |
| one | two | yes | no |
| spsp | preempt | 21 | 3 |
| spsp | react | 13 | 11 |

## # A tibble: 1 x 3  
## n p p.signif  
## \* <dbl> <dbl> <chr>   
## 1 48 0.0243 \*

|  |  |  |  |
| --- | --- | --- | --- |
| one | two | yes | no |
| envst | preempt | 9 | 1 |
| envst | react | 8 | 2 |

## # A tibble: 1 x 3  
## n p p.signif  
## \* <dbl> <dbl> <chr>   
## 1 20 1 ns

|  |  |  |  |
| --- | --- | --- | --- |
| one | two | yes | no |
| demst | preempt | 11 | 0 |
| demst | react | 7 | 4 |

## # A tibble: 1 x 3  
## n p p.signif  
## \* <dbl> <dbl> <chr>   
## 1 22 0.0902 ns

Chi squares

## 3. Who are the organizations/players involved in conserving the conservation of the species (eg in an agreement of effort)?

#### 3 - Data wrangling

o Summary stats… histograms/bar charts… (perhaps stacked bar charts where color code stacks for low, medium, high confidence in having collected most of partners)?

o To account for the uncertainty in number of partners, we put this information into bins and then “then use an ordered probit to look at how well variables predict the rough estimate of numbers of partners.”

o We use logit regression models to predict participation by different “types” of partners

#### Set up model

### Base model

* Note: percent public variable is still in matrix but was removed from model

VIFS

Adding stargazer - <https://rpubs.com/omerorsun/week3_stargazer>

lmg is the R^2 contribution averaged over orderings amoung regressors

Testing heteroskedasticity

Robust Standard Errors in R

##   
## t test of coefficients:  
##   
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -5.639863 2.155417 -2.6166 0.013604 \*   
## log(range\_area) 0.267593 0.096884 2.7620 0.009571 \*\*  
## plant -0.421257 0.236643 -1.7801 0.084858 .   
## nserve\_pop\_stability 0.158461 0.103995 1.5237 0.137713   
## private\_land\_proportion 1.294449 0.587351 2.2039 0.035089 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Predicting values - <https://www.youtube.com/watch?v=ahDFXHAdZRU> -

## Paper

###Summary stats

* number of unique partners working on all species (note - using the HCP dataset)

### Table with both models

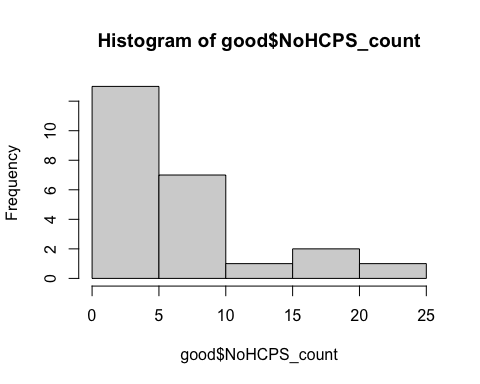
Let’s try with sjPlot

Scrap - testing ECOS csv against our list of species

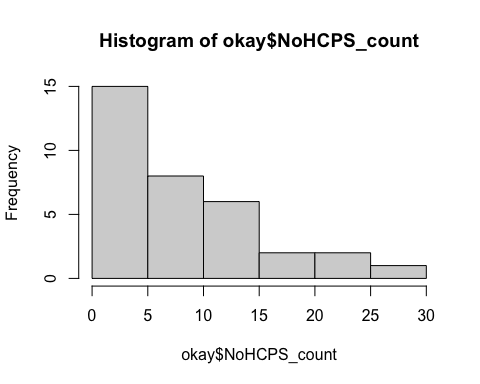
run confidence check

## Joining, by = c("scientific\_name", "common\_name")

## Joining, by = "scientific\_name"



##   
## Call:  
## lm(formula = log(NoHCPS\_count) ~ log(range\_area) + plant + total\_threats +   
## private\_land\_proportion, data = good)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.45667 -0.47158 -0.00002 0.52812 1.24481   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -4.36439 3.10160 -1.407 0.1774   
## log(range\_area) 0.23843 0.12509 1.906 0.0737 .  
## plant -0.54846 0.35426 -1.548 0.1400   
## total\_threats 0.02466 0.15549 0.159 0.8759   
## private\_land\_proportion 1.14626 0.67645 1.695 0.1084   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.7652 on 17 degrees of freedom  
## (2 observations deleted due to missingness)  
## Multiple R-squared: 0.3411, Adjusted R-squared: 0.186   
## F-statistic: 2.2 on 4 and 17 DF, p-value: 0.1125



##   
## Call:  
## lm(formula = log(NoHCPS\_count) ~ log(range\_area) + plant + total\_threats +   
## private\_land\_proportion, data = okay)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.49513 -0.38167 0.07007 0.44253 1.23390   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -6.37636 2.25217 -2.831 0.00883 \*\*  
## log(range\_area) 0.32167 0.09535 3.373 0.00234 \*\*  
## plant -0.45424 0.27660 -1.642 0.11259   
## total\_threats 0.04625 0.12525 0.369 0.71491   
## private\_land\_proportion 1.40696 0.48794 2.883 0.00779 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.7097 on 26 degrees of freedom  
## (3 observations deleted due to missingness)  
## Multiple R-squared: 0.4546, Adjusted R-squared: 0.3707   
## F-statistic: 5.418 on 4 and 26 DF, p-value: 0.002609

test relationship between number of initiatives and number of partners

Q - how many agreement documents are we missing

* Breusch-Pagan test
* <https://www.r-bloggers.com/2016/01/how-to-detect-heteroscedasticity-and-rectify-it/>
* <https://cran.r-project.org/web/packages/olsrr/vignettes/heteroskedasticity.html>

While it doesn’t give us the critical value to compare the test statistic, all you need to look at is the p-value to determine whether or not you should reject the null. If the p-value is less than the level of significance (in this case if the p-value is less than α=0.05), then you reject the null hypothesis. Since 0.006579 < 0.05, we can reject the null hypothesis.

With a p-value of 0.91, we fail to reject the null hypothesis (that variance of residuals is constant) and therefore infer that ther residuals are homoscedastic. Lets check this graphically as well.