

# Summary Analysis of Fire severity impact on understory vegetation species richness in the Araucaria-Nothofagus forest.

Pham, Peter

\*Preface: The client was interested in exploring the effects of wildfires on recovering areas. Of particular interest was to develop a model that would describe the relationship between burn severity and species richness. As such, the first step in the process is conducting a hypothesis test to gain a broad understanding of the data. The analysis conducted in this document is an excerpt from the full report.

The full R code used to produce this markdown document can be found in the Appendix\*

## Introduction

In 2015, a wildfire occurred in the Araucaria-Nothofagus forest in the La Araucania region in southern Chile. The wildfire was active for 23 days, burning a total area of 6.599 ha, of which around 50% were within conservation areas. The clients presented a dataset that was collected five years after the occurrence. Field technicians collected data from 60 assigned circular plots that were subdivided into  $1 \times 1 \text{ m}^2$ , resulting in a total sample size of 224 subplots. The focus of this study to examine the effect of burn severity on species richness. The current body of research that describes the long-term effects of forest floor species diversity is sparse.

## Exploratory Analysis

Exploratory analysis is conducted to gain a stronger sense in which the data is collected. Mean species richness for each burn severity group are slightly different from each other (unburned (UN):  $\mu = 4.84$ , low-burned (L):  $\mu = 3.43$ , high-burned (H):  $\mu = 2.11$ ). In figure 1, below, we can see this relative distribution of species richness from the burn categories. Not only does mean species richness decrease, but so does the variability of species richness found in each plot decreases as burn severity increases. This can be seen as the “whiskers” of each box plot centralizes around the mean as the categories shift.

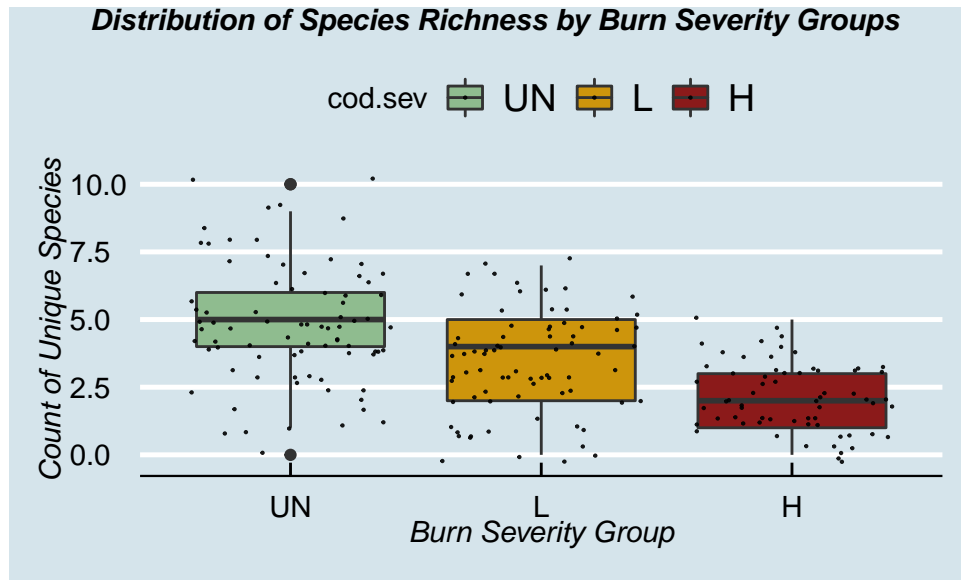


Figure 1: Distributions of Unique Specices by Burn Severity

## Distribution Analysis (Continued)

Table 1 below provides summary statistics of this behavior - as burn severity increases, each summary statistic decreases. This notable as we see a stark decrease in variability as well. This may indicate that only a few species may survive wildfire events.

Additional examination of comparative histograms (figure 2) also reinforces the notion of decrease in spread and mean species richness as fire severity increases. Based on visual analysis, we posit that there may exist a flat decrease in value that is related to burn severity.

Table 1: Summary Statistics for Species Richness by Burn Severity

Burn Severity	Minimum	Q1	Median	Mean	S.Error	S.D.	Q3	Maximum
UN	0	4	5	4.84	0.25	2.20	6	10
L	0	2	4	3.43	0.21	1.86	5	7
H	0	1	2	2.11	0.16	1.30	3	5
All	0	2	4	3.56	0.15	2.15	5	10

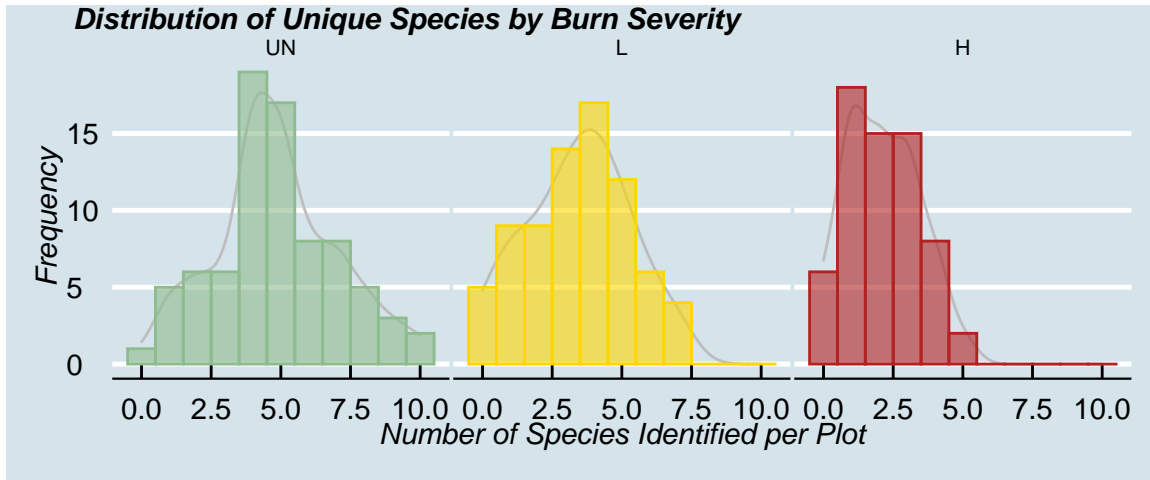


Figure 2: Individual Distribution of Each Burn Category

## Methods

We look to test if the mean species richness of each burn group is significantly different. An analysis of variance (ANOVA) test is preliminarily selected to test the hypothesis:

$H_0$  = there does not exist a difference in means between each burn severity group  $H_1$  = there does exist a difference in means between each burn severity group

Each group each other through hypothesis testing. Distribution normality is to be assessed and an appropriate test is to be selected. Further more, we wish to define the relationship with a regression tree model.

Shapiro-Wilks is a standard test to assess normality of each group. Conducting this test yields p-value of  $1.616 \times 10^{-13}$ ; thus, we conclude that the samples collected are not normally distributed and a non-parametric test is required. We then proceed apply Kruskal-Wallis' ranked mean determine if there exists a significant difference between mean species richness of each burn category.

```
## Kruskal-Wallis rank sum test
##
```

Table 2: Shapiro-Wilks Test for All Groups

Burn Severity	N Observations	Test Stat	P-Value
UN	80	0.97	0.05
L	76	0.96	0.01
H	64	0.93	0.00

```
## data: x and group
## Kruskal-Wallis chi-squared = 58.907, df = 2, p-value = 0
##
##
##              Comparison of x by group
##              (No adjustment)
## Col Mean-|
## Row Mean |          H          L
## -----+-----
##      L | -4.110901
##      |  0.0000*
##      |
##      UN | -7.674090 -3.680515
##      |  0.0000*   0.0001*
##
##
## List of pairwise comparisons: Z statistic (p-value)
## -----
## H - L : -4.110901 (0.0000)*
## H - UN : -7.674090 (0.0000)*
## L - UN : -3.680515 (0.0001)*
##
## alpha = 0.05
## Reject Ho if p <= alpha/2
##
## dunn$comparisons dunn$Z dunn$P.adjusted
## 2      H - UN -7.674090  8.329851e-15
## 1      H - L -4.110902  1.970583e-05
## 3      L - UN -3.680516  1.163813e-04
```

Table 3: Dunn's Pairwise Test for Burn Severity Groups

Group Comparisons	Test Statistic	P-Value
H - L	-4.1109	0e+00
H - UN	-7.6741	0e+00
L - UN	-3.6805	1e-04

```
knitr::opts_chunk$set(echo = TRUE)
knitr::opts_chunk$set(tidy.opts=list(width.cutoff=40), tidy=TRUE)
knitr::opts_chunk$set(fig.pos = '!h')
# Packages Used
library(caret)
library(leaps)
library(formatR)
```

```

library(broom)
library(plyr)
library(ggpubr)
library(ggthemes)
library(hrbrthemes)
library(kableExtra)
library(knitr)
library(magrittr)
library(tidyverse)
library(xtable)
library(dunn.test)
library(pander)

# Import data from .csv
dt = data.frame(read_csv("ChinaMuerta.csv"), header=TRUE, sep=",")

# Subset "dt" and remove categories that are not "native" or "exotic"
spp <-subset(dt, origen=="native" | origen=="exotic")
spp<-droplevels(spp)

# Quick calculation of the total species richness
# dim(tapply(spp$origen, spp$species, length)) #66 spp

# Reorganize "spp" by species richness by treatment
ddply(spp, c("site","cod.sev"), function(df)
  return(c(species=length(unique(df$species)))))

# Reorganize "spp" further by variables of interest.
plot.rich<-ddply(spp, c("site","n.plot" ,"cod.sev", "transect"), function(df)
  return(c(species=length(unique(df$species)))))

# spread function
# avoiding to lose subplot without species, and filling them with zeros
dt1<-spread(plot.rich, transect, species, fill=0)

# gather function & reorganizing subplots (transects) in one column
dt2<-gather(dt1, "east", "north", "south", "west", key="transect", value=species)

# Set cod.sev as a factor and assign a specific order
dt2$cod.sev = factor(dt2$cod.sev, levels = c("UN","L","H"), ordered=TRUE)

# -----Create data frame for multiple regression-----
df = dt %>%
  filter(origen != "NA") %>%
  select(species, cod.sev, slope, aspect, elev.m,
         cov.50N, cov.50E, cov.50S, cov.50W, cov.140N, cov.140E, cov.140S,
         cov.140W) %>%
  rowwise() %>%
  mutate(c50.bar = mean(c(cov.50N, cov.50E, cov.50S, cov.50W), na.rm=TRUE),
         c140.bar = mean(c(cov.140N, cov.140E, cov.140S, cov.140W), na.rm=TRUE),
         cod.sev = recode(cod.sev,"UN" = "1", "L" = "2", "H" = "3")) %>%
  group_by(cod.sev, aspect, slope, elev.m) %>%
  summarize(c50.mean = mean(c50.bar, na.rm = TRUE),

```

```

    c140.mean = mean(c140.bar, na.rm = TRUE),
    species = length(species))

# Creating a boxplot for the groups
bp = ggplot(dt2, aes(y=species, fill=cod.sev, x=cod.sev)) +
  geom_boxplot() +
  geom_jitter(color="black", size=0.1, alpha=0.9) +
  theme_economist() +
  scale_fill_manual(values=c("darkseagreen", "darkgoldenrod3", "firebrick4")) +
  labs(y = expression(italic("Count of Unique Species")),
       x = expression(italic("Burn Severity Group")),
       subtitle = expression(
         bold(
           "")))

# Boxplot Annotation
annotate_figure(bp,
               fig.lab = "Distribution of Species Richness by Burn Severity Groups",
               fig.lab.face="bold.italic",
               fig.lab.pos = "top.left")

# Creating a histogram for the frequency of the species count
hg = ggplot(dt2, aes(x=species, fill = cod.sev, color=cod.sev)) +
  geom_histogram(binwidth=1, alpha=0.5, position="identity")+
  theme_economist() +
  scale_fill_manual(values=c("darkseagreen", "darkgoldenrod3", "firebrick")) +
  scale_color_manual(values=c("darkseagreen", "black", "firebrick")) +
  labs(y = expression(italic("Frequency of Occurrence")),
       x = expression(italic("Count of Unique Species")),
       subtitle = expression(
         bold(
           "")))
  )) +
  theme(legend.position = "top")

# Histogram Annotation
annotate_figure(hg,
               fig.lab = "Overlaid Histograms for Each Burn Severity",
               fig.lab.face="bold.italic",
               fig.lab.pos = "top.left")

# Table 1: Summary Statistics for groups

summary = bind_rows( # We are creating two tables, this function will bind them together

dt2 %>% # Produce summary statistics by burn group
dplyr::group_by(cod.sev) %>%
dplyr::summarise(
  Minimum = min(species),
  Q1 = quantile(species, 0.25),
  Median = median(species),
  Mean = mean(species),
  S.Error = sd(species)/sqrt(length(species)),
  S.D. = sd(species),

```

```

    Q3 = quantile(species, 0.75),
    Maximum = max(species)
  ),

dt2 %>% # Produce summary statistics for all groups combined
  summarise(
    Minimum = min(species),
    Q1 = quantile(species, 0.25),
    Median = median(species),
    Mean = mean(species),
    S.Error = sd(species)/sqrt(length(species)),
    S.D. = sd(species),
    Q3 = quantile(species, 0.75),
    Maximum = max(species)
  ) %>%
  mutate(cod.sev = "All")
)

# Using KableExtra package to clean up table data and to add styling options
kable(summary,
  caption =
    "Summary Statistics for Species Richness by Burn Severity",
  col.names = c("Burn Severity", names(summary)[-1]),
  align="c",
  booktabs = T,
  digits=2) %>%
  kable_styling(position = "center", latex_options="hold_position") %>%
  row_spec(0, bold=TRUE)

hg.all = dt2 %>%
  mutate(cod.sev = fct_reorder(cod.sev, species, .desc=TRUE)) %>%
  ggplot(aes(x=species, fill=cod.sev, color=cod.sev)) +
  geom_density(color="gray", size=0.5, alpha=0, aes(y=..count..)) +
  geom_histogram(binwidth=1, alpha=0.6, position="identity") +
  scale_fill_manual(values=c("darkseagreen", "gold", "firebrick")) +
  scale_color_manual(values=c("darkseagreen", "gold", "firebrick")) +
  theme_economist() +
  theme(
    legend.position="none",
    panel.spacing = unit(0.1, "lines"),
    strip.text.x = element_text(size = 8)
  ) +
  xlab(expression(italic("Number of Species Identified per Plot"))) +
  ylab(expression(italic("Frequency"))) +
  facet_wrap(~cod.sev)

# Annotate Figure
annotate_figure(hg.all,
  fig.lab = "Distribution of Unique Species by Burn Severity",
  fig.lab.face="bold.italic",
  fig.lab.pos = "top.left")

```

```

sw.all = dt2 %>%
  dplyr::group_by(cod.sev) %>%
  dplyr::summarize(
    N = length(species),
    Test.Stat = shapiro.test(species)$statistic,
    P.Value = shapiro.test(species)$p.value)

# Using KableExtra package to clean up table data and to add styling options
kable(sw.all,
  caption =
    "Shapiro-Wilks Test for All Groups",
  col.names = c("Burn Severity", "N Observations", "Test Stat", "P-Value"),
  align="c",
  booktabs = T,
  digits=2) %>%
  kable_styling(position = "center", latex_options="hold_position") %>%
  row_spec(0, bold=TRUE)

# Kruskal-Wallace
kw = kruskal.test(species ~ cod.sev, data = dt2)

# Post Hoc Analysis
dunn = dunn.test(dt2$species, g=dt2$cod.sev, kw=TRUE, list=TRUE)
dunn.table = cbind.data.frame(dunn$comparisons, dunn$Z, dunn$P.adjusted)
dunn.table[order(dunn.table$dunn$P.adjusted),]

# Using KableExtra package to clean up table data and to add styling options
kable(dunn.table,
  caption = "Dunn's Pairwise Test for Burn Severity Groups",
  col.names = c("Group Comparisons", "Test Statistic", "P-Value"),
  align="c",
  booktabs = T,
  digits=4) %>%
  kable_styling(position = "center", latex_options="hold_position") %>%
  row_spec(0, bold=TRUE)

```