Advanced Ecology: Communities and Ecosystems PCB5443-U01-1241

Exam 2 - Pulses Time Series

Yuri Souza

March 13, 2024

**Part I: Develop a question or questions that motivate the time-series analysis of the data you chose (10 points).

A:

- Does the defaunation of medium and large herbivores affect plant abundance?

**Part II: Use a time-series model to analyze the data (pick one of the three models, or you may chose to use more than one model). Include the data, the R scripts, figures, and model output (50 points).

A:

- All the data and code are available **_here_** and included in the last pages of this file.

**Part III: Interpret the results of your data (in 1-2 paragraphs) using descriptions of the model output (qualitative and quantitative) and characteristics of pulse dynamics described in Jentsch and White (2019) (40 points).

A:

- The data used in this exploratory analysis was collected over a period of ten years as a part of an exclusion experiment conducted from 2009 to 2018. The experiment involved 46 paired plots, i.e., closed and open, each measuring 5x3 meters. These plots were located across four protected areas in the Atlantic Forest in São Paulo, Brazil. The samples were collected twice a year - one in January and one in June - to capture the variations between the wet and dry seasons. The data pertains to the abundance of plant saplings.

- I have used the "Peak pulse" time series model to identify peaks of plant abundance data in two types of plots: closed (fenced to medium and large herbivores) and open (open to any sort of organism). My hypothesis was that the absence of herbivores would affect the dynamics of plant abundance, with the prediction that abundance would increase in closed plots due to the lack of their predator. For both treatments a threshold of **5** was selected since it seemed to capture and represent better the present peaks. To make this decision, I tested all the values ranging from 1 to 45 (the maximum each model would accept to predict peaks). Values ranging from 1 to 8 showed 4 peaks, while all the other values would show 3.

- According to the models, both treatments had the same number of peaks - four in total - which occurred in the same three years. Overall, the peaks for closed (*Peaks per year: 0.4, average peak magnitude: 1981.75, peak standard deviation: 95.17, peak CV: 0.04*) plot highlight a increase in plant abundance when compared to the open plot (*Peaks per year: 0.4, average peak magnitude: 1290, peak standard deviation: 76.75, peak CV: 0.05* - Figure 1 C and D). However, when we examined the periodicity of peak abundance events over time, we found that they were not statistically significant for either of the treatments (Open: Slope: -0.64, *P-value* = 0.35 - Figure 1 B and C; Closed: Slope: -0.10, *P-value: 0.89* - Figure 1 D and E). Even though the models were not significant, open plots show a decrease in plant abundance after pulse rate events, given the slightly decreased trend of the pulse magnitude (Figure 1C), while in the closed plots, it is constant (figure 1E).

- Both treatments had the same peak patterns within the same year period, with the exception of 2014, which was present only in closed plots, while 2015 was absent. The data structure may not be suitable enough for this method to capture the peaks precisely, given that there were only 10 years of data samples comprising only 2 measurements per year.

- All of the patterns presented here contradict my initial prediction that plant abundance would increase in closed plots exponentially toward a trend of a lower pulse rate and recovery in the last years, which is supported by previous studies conducted within the same experiment (Villar et al., 2019; Souza et al., 2022). Both treatments had an initial pulse rate followed by a recovery period. Visually, closed plots had an increment in the plant's abundance after the start of the experiment, with both lines beginning at almost the same point (Figure 1A). Although the first few years did not exhibit significant variation, there are visible peaks and valleys after 2014 that were not fully captured by the model. Therefore, for future analysis, it is recommended to include the valley *quantmod* method, along with the rate of plant mortality and recruitment. Additionally, exploring the time series at the site level might be more appropriate as each location exhibits a significant variability in plant dynamics.
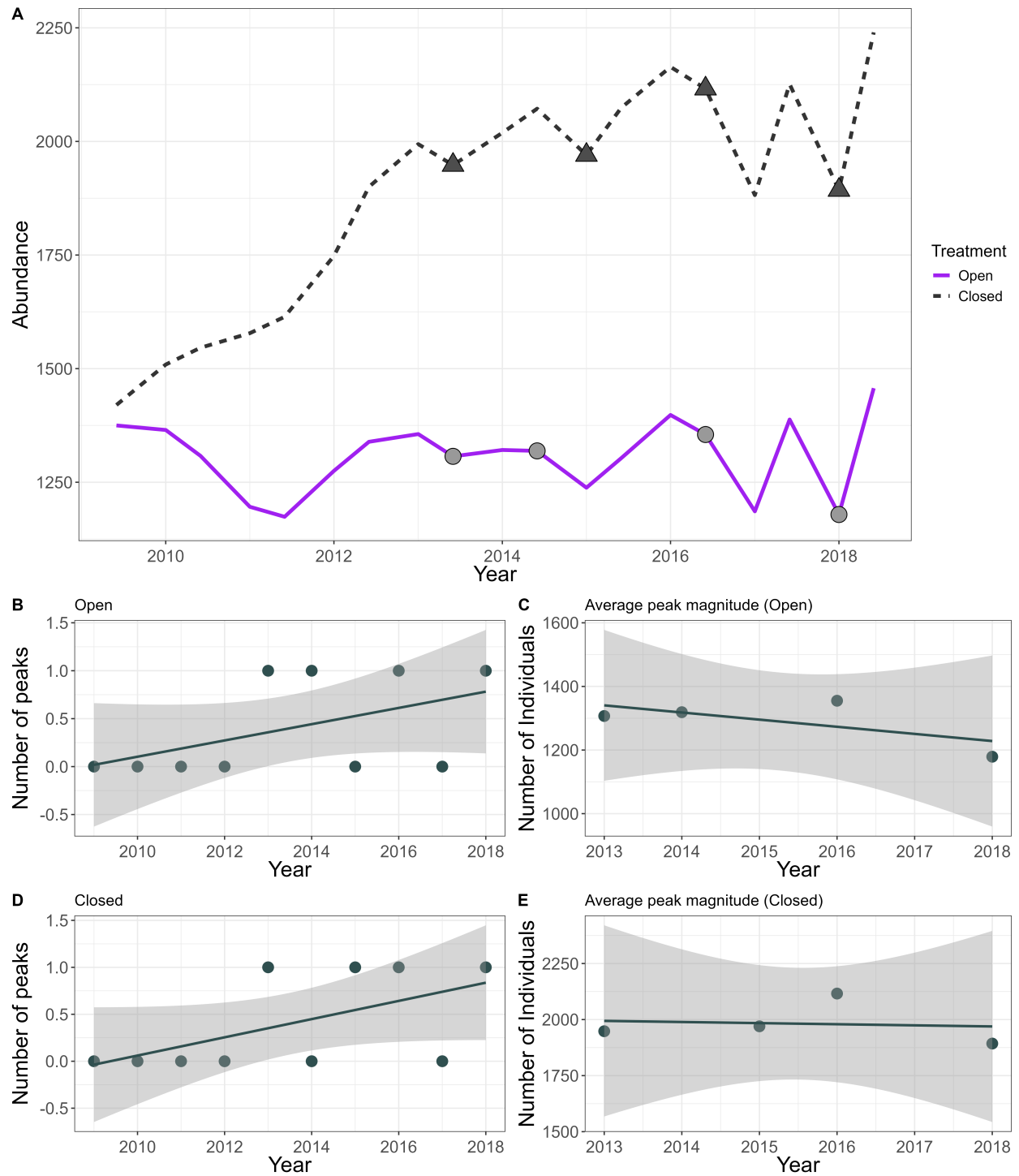
Figure 1. Graph A displays the plant abundance over time for both open (circles) and closed treatments (triangles). Graphs B and D showcase the effect of time on peak occurrence, while C and E represent the trend in magnitudes.

Literature

Villar, N., Siqueira, T., Zipparro, V., Farah, F., Schmaedecke, G., Hortenci, L., . . . & Galetti, M. (2020). The cryptic regulation of diversity by functionally complementary large tropical forest herbivores. Journal of Ecology, 108(1), 279-290.

Souza, Y., Villar, N., Zipparro, V., Nazareth, S., & Galetti, M. (2022). Large mammalian herbivores modulate plant growth form diversity in a tropical rainforest. Journal of Ecology, 110(4), 845-859.

Script

```r
# packages --------------------------------------------------------------

rm(list=ls())

if(!require(tidyverse)) install.packages("tidyverse", dependencies = TRUE)
if(!require(quantmod)) install.packages("quantmod", dependencies = TRUE)
if(!require(xts)) install.packages("xts", dependencies = TRUE)
if(!require(emmeans)) install.packages("emmeans", dependencies = TRUE)
if(!require(ggthemes)) install.packages("ggthemes", dependencies = TRUE)
if(!require(timetk)) remotes::install_github("business-science/timetk")
if(!require(ggpubr)) install.packages("ggpubr", dependencies = TRUE)

# Abundance #  ----------------------------------------------------------


# Data ------------------------------------------------------------------

data_biota <- readr::read_csv("00_data/life_form_2020_for_abudance_recruitment_mortality.csv") |>
  dplyr::select(c(1:3,5:23)) |>
  tidyr::pivot_longer(cols = 4:22, names_to = "Date", values_to = "abnd") |>
  dplyr::mutate(Date = recode(Date,
                              "T0" = "2009-06-01", "T006" = "2010-01-01",
                              "T012" = "2010-06-01", "T018" = "2011-01-01",
                              "T024" = "2011-06-01", "T030" = "2012-01-01",
                              "T036" = "2012-06-01", "T042" = "2013-01-01",
                              "T048" = "2013-06-01", "T054" = "2014-01-01",
                              "T060" = "2014-06-01", "T066" = "2015-01-01",
                              "T072" = "2015-06-01", "T078" = "2016-01-01",
                              "T084" = "2016-06-01", "T090" = "2017-01-01",
                              "T096" = "2017-06-01", "T102" = "2018-01-01",
                              "T108" = "2018-06-01")) |>
  dplyr::group_by(Treatment, Date) |>
  dplyr::summarise(abnd = sum(abnd)) |>
  dplyr::ungroup() |>
  dplyr::mutate(Date = as.POSIXct(paste(Date, "-01", "%Y-%m-%d", tz = "UTC")))


# Open ------------------------------------------------------------------


all_open <- data_biota |>
  dplyr::filter(Treatment == "open")
all_open

all_open_ts <- all_open |>
  timetk::tk_xts()
all_open_ts

str(all_open_ts)
tzone(all_open_ts)
nmonths(all_open_ts)
```

```r
nyears(all_open_ts)

chartSeries(all_open_ts,theme="white")

chartSeries(all_open_ts,up.col='blue',dn.col='red',
            subset='2010::2017-06',theme="white")

# didn't change from 1 to 10
all_open_pks.1 <- findPeaks(all_open_ts, thresh=1)
plot(all_open_ts[all_open_pks.1])

all_open_peaks1 <- as.data.frame(all_open_ts[all_open_pks.1]) |>
  rownames_to_column("Dates") |>
  dplyr::mutate(Treatment = "Open")
all_open_peaks1

#how many peaks per year?
peak_per_y_open <- length(all_open_peaks1$abnd)/nyears(all_open_ts)
peak_per_y_open

#average peak magnitude (in number of individuals per year)
all_peak_mean_open <- mean(all_open_peaks1$abnd)
all_peak_mean_open

#peak standard deviation
all_peak_sd_open <- sd(all_open_peaks1$abnd)
all_peak_sd_open

#peak CV
all_peak_CV_open <- all_peak_sd_open/all_peak_mean_open
all_peak_CV_open

all_open_summary <- cbind("Peaks per year:",peak_per_y_open,
                "average peak magnitude:", all_peak_mean_open,
                "peak standard deviation:", all_peak_sd_open,
                "peak CV:", all_peak_CV_open)
all_open_summary


all_open_peaks1$year<-as.numeric(format(as.POSIXct(all_open_peaks1$Dates),"%Y"))
all_open_peaks1

all_year_open <- min(as.numeric(
  format(as.POSIXct(all_open$Date),"%Y"))):max(as.numeric(
    format(as.POSIXct(all_open$Date),"%Y")))
all_year_open
all_years_open <- as.data.frame(all_year_open)
all_years_open

all_peak.sum_open <- all_open_peaks1 |>
  group_by(year) |>
  summarise(mean.peak=mean(abnd), count=n())
all_peak.sum_open
```

```r
all_peak.number_open <- merge(all_years_open, all_peak.sum_open,
                              by.x="all_year_open", by.y="year", all.x=TRUE)
all_peak.number_open
all_peak.number_open[is.na(all_peak.number_open)] <- 0
all_peak.number_open

all_peak.number.lm_open <- lm(scale(count)~scale(all_year_open),
                              data = all_peak.number_open)
all_lmsum.number_open <- summary(all_peak.number.lm_open)
all_lmsum.number_open
all_peak.number.slope_open <- all_peak.number.lm_open$coefficients[2]
all_peak.number.slope_open
all_peak.number.p_open <- all_lmsum.number_open$coefficients[2,4]
all_peak.number.p_open

all_peak.magnitude.lm_open <- lm(scale(abnd)~ scale(year), data = all_open_peaks1)
all_lmsum.mag_open <- summary(all_peak.magnitude.lm_open)
all_lmsum.mag_open
all_peak.magnitude.slope_open <- all_peak.magnitude.lm_open$coefficients[2]
all_peak.magnitude.slope_open
all_peak.magnitude.p_open <- all_lmsum.mag_open$coefficients[2,4]
all_peak.magnitude.p_open



# Closed -------------------------------------------------------------------


all_closed <- data_biota |>
  dplyr::filter(Treatment == "closed")
all_closed

all_closed_ts <- all_closed |>
  timetk::tk_xts()
all_closed_ts

str(all_closed_ts)
tzone(all_closed_ts)
nmonths(all_closed_ts)
nyears(all_closed_ts)

chartSeries(all_closed_ts,theme="white")

chartSeries(all_closed_ts,up.col='blue',dn.col='red',
            subset='2010::2017-06',theme="white")

# didnt change from 1 to 45
all_closed_pks.1 <- findPeaks(all_closed_ts, thresh=1)
plot(all_closed_ts[all_closed_pks.1])

all_closed_peaks1 <- as.data.frame(all_closed_ts[all_closed_pks.1]) |>
  rownames_to_column("Dates") |>
  dplyr::mutate(Treatment = "Closed")
```

```
all_closed_peaks1

#how many peaks per year?
peak_per_y_closed <- length(all_closed_peaks1$abnd)/nyears(all_closed_ts)
peak_per_y_closed

#average peak magnitude (in number of individuals per year)
all_peak_mean_closed <- mean(all_closed_peaks1$abnd)
all_peak_mean_closed

#peak standard deviation
all_peak_sd_closed <- sd(all_closed_peaks1$abnd)
all_peak_sd_closed

#peak CV
all_peak_CV_closed <- all_peak_sd_closed/all_peak_mean_closed
all_peak_CV_closed

all_closed_summary <- cbind("Peaks per year:",peak_per_y_closed,
                            "average peak magnitude:", all_peak_mean_closed,
                            "peak standard deviation:", all_peak_sd_closed,
                            "peak CV:", all_peak_CV_closed)
all_closed_summary


all_closed_peaks1$year<-as.numeric(format(as.POSIXct(all_closed_peaks1$Dates),"%Y"))
all_closed_peaks1

all_year_closed <- min(as.numeric(
  format(as.POSIXct(all_closed$Date),"%Y"))):max(as.numeric(
    format(as.POSIXct(all_closed$Date),"%Y")))
all_year_closed
all_years_closed <- as.data.frame(all_year_closed)
all_years_closed

all_peak.sum_closed <- all_closed_peaks1 |>
  group_by(year) |>
  summarise(mean.peak=mean(abnd), count=n())
all_peak.sum_closed
all_peak.number_closed <- merge(all_years_closed, all_peak.sum_closed,
                                by.x="all_year_closed",
                                by.y="year",
                                all.x=TRUE)
all_peak.number_closed
all_peak.number_closed[is.na(all_peak.number_closed)] <- 0
all_peak.number_closed

all_peak.number.lm_closed <- lm(scale(count)~scale(all_year_closed),
                                data = all_peak.number_closed)
all_lmsum.number_closed <- summary(all_peak.number.lm_closed)
all_lmsum.number_closed
all_peak.number.slope_closed <- all_peak.number.lm_closed$coefficients[2]
all_peak.number.slope_closed
```

```r
all_peak.number.p_closed <- all_lmsum.number_closed$coefficients[2,4]
all_peak.number.p_closed

all_peak.magnitude.lm_closed <- lm(scale(abnd)~ scale(year),
                                    data = all_closed_peaks1)
all_lmsum.mag_closed <- summary(all_peak.magnitude.lm_closed)
all_lmsum.mag_closed
all_peak.magnitude.slope_closed <- all_peak.magnitude.lm_closed$coefficients[2]
all_peak.magnitude.slope_closed
all_peak.magnitude.p_closed <- all_lmsum.mag_closed$coefficients[2,4]
all_peak.magnitude.p_closed



# Graph -----------------------------------------------------------------

dev.off()

all_biota <- data_biota |>
  dplyr::mutate(Treatment = stringr::str_to_title(Treatment),
               Treatment = fct_relevel(Treatment, c("Open", "Closed")))
all_biota

all_biota_graph <- ggplot(all_biota, aes(x = Date,
                          y = abnd,
                          linetype = Treatment,
                          color = Treatment,
                          fill = Treatment)) +
  geom_line(size = 1.5) +
  scale_color_manual(values=c("purple", "grey20")) +
  geom_point(data = all_open_peaks1, aes(x = as.POSIXct(Dates), y = abnd),
             colour='black', fill = "grey60", size=6, shape = 21) +
  geom_point(data = all_closed_peaks1, aes(x = as.POSIXct(Dates), y = abnd),
             colour='black', fill = "grey30", size=6, shape = 24) +
  labs(x = "Year", y = "Abundance") +
  theme_bw() +
  theme(axis.text.x= element_text(angle= 0, size = 14, vjust = 0.2),
        axis.text.y= element_text(size = 14),
        axis.title.x = element_text(size = 18),
        axis.title.y = element_text(size = 18),
        legend.title = element_text(size = 14),
        legend.text = element_text(size = 12),
        strip.text = element_text(size = 12))
all_biota_graph



# Open
all_peak.number.graph_open <- ggplot(all_peak.number_open,
                                     aes(x = all_year_open,
                                         y = count)) +
  geom_point(color='darkslategrey',size=4) +
  geom_smooth(method = "lm",  se = TRUE, color="darkslategrey") +
```

```r
  labs(x = "Year", y = "Number of peaks", title = "Open") +
  theme_bw() +
  theme(axis.text.x= element_text(angle= 0, size = 14, vjust = 0.2),
        axis.text.y= element_text(size = 14),
        axis.title.x = element_text(size = 18),
        axis.title.y = element_text(size = 18),
        legend.title = element_text(size = 14),
        legend.text = element_text(size = 12),
        strip.text = element_text(size = 12))
all_peak.number.graph_open


all_peak.mag.graph_open <- ggplot(data = all_open_peaks1,
                                  aes(x = year,
                                      y = abnd)) +
  geom_point(color = 'darkslategrey', size=4) +
  geom_smooth(method = "lm",  se = TRUE, color="darkslategrey")+
  labs(x = "Year", y = "Number of Individuals", title = "Average peak magnitude (Open)") +
  theme_bw() +
  theme(axis.text.x= element_text(angle= 0, size = 14, vjust = 0.2),
        axis.text.y= element_text(size = 14),
        axis.title.x = element_text(size = 18),
        axis.title.y = element_text(size = 18),
        legend.title = element_text(size = 18),
        legend.text = element_text(size = 16),
        strip.text = element_text(size = 12))
all_peak.mag.graph_open


# Closed

all_peak.number.graph_closed <- ggplot(all_peak.number_closed,
                                       aes(x = all_year_closed,
                                           y = count)) +
  geom_point(color='darkslategrey',size=4) +
  geom_smooth(method = "lm",  se = TRUE, color="darkslategrey") +
  labs(x = "Year", y = "Number of peaks", title = "Closed") +
  theme_bw() +
  theme(axis.text.x= element_text(angle= 0, size = 14, vjust = 0.2),
        axis.text.y= element_text(size = 14),
        axis.title.x = element_text(size = 18),
        axis.title.y = element_text(size = 18),
        legend.title = element_text(size = 14),
        legend.text = element_text(size = 12),
        strip.text = element_text(size = 12))
all_peak.number.graph_closed


all_peak.mag.graph_closed <- ggplot(data = all_closed_peaks1,
                                    aes(x = year,
                                        y = abnd)) +
  geom_point(color = 'darkslategrey', size=4) +
```

```r
  geom_smooth(method = "lm",  se = TRUE, color="darkslategrey")+
  labs(x = "Year", y = "Number of Individuals", title = "Average peak magnitude (Closed)") +
  theme_bw() +
  theme(axis.text.x= element_text(angle= 0, size = 14, vjust = 0.2),
        axis.text.y= element_text(size = 14),
        axis.title.x = element_text(size = 18),
        axis.title.y = element_text(size = 18),
        legend.title = element_text(size = 14),
        legend.text = element_text(size = 12),
        strip.text = element_text(size = 12))
all_peak.mag.graph_closed

all_graphs <- ggpubr::ggarrange(all_biota_graph,
                                ggarrange(
                                all_peak.number.graph_open,
                                all_peak.mag.graph_open,
                                all_peak.number.graph_closed,
                                all_peak.mag.graph_closed,
                                ncol = 2,
                                labels = c("B", "C", "D", "E")),
                                nrow = 3,
                                labels = "A")
all_graphs


all_biota_peaks <- ggarrange(all_peak.number.graph_open,
                        all_peak.mag.graph_open,
                        all_peak.number.graph_closed,
                        all_peak.mag.graph_closed,
                        ncol=2,
                        nrow=2,
                        labels = c("B", "C","D","E"))

ggarrange(all_biota_graph, all_biota_peaks, ncol=1, nrow=2, labels = "A")

#ggsave("02_figures/20240313_YSS_biota_abnd.png", h = 35, w = 30, units = "cm", dpi = 500)
```