Basal area

Tim White

June 2024

Packages and data

```r
library(car)
library(tidyverse)

data_stems <- read_csv("data/data_stems.csv") %>%
                mutate(plot_id = as.factor(plot_id),
                        size_class = as.factor(size_class),
                        harvested = as.factor(harvested),
                        vegetation_type = as.factor(vegetation_type),
                        milpa = as.factor(milpa))
data_plots <- read_csv("data/data_plots.csv") %>%
                mutate(plot_id = as.factor(plot_id),
                        harvested = as.factor(harvested),
                        vegetation_type = as.factor(vegetation_type),
                        milpa = as.factor(milpa))

plot_theme <- theme(plot.background = element_rect(fill = "white"),
                        plot.title = element_blank(),
                        plot.subtitle = element_text(family="sans", face="plain"),
                        axis.title.x = element_text(family="sans", face="bold"),
                        axis.title.y = element_text(family="sans", face="bold"),
                        axis.text.x = element_text(family="sans", face="plain"),
                        axis.text.y = element_text(family="sans", face="plain"),
                        panel.background = element_rect(fill="white"),
                        panel.grid.major.x = element_blank(),
                        panel.grid.major.y = element_line(color="gainsboro"),
                        panel.grid.minor = element_blank(),
                        axis.ticks = element_blank(),
                        legend.background = element_rect(color="black", fill = "white"),
                        legend.position = c(0.9, 0.85),
                        legend.title = element_text(color = "black", face = "bold", hjust = 0.5),
                        legend.text = element_text(color = "black"))
```
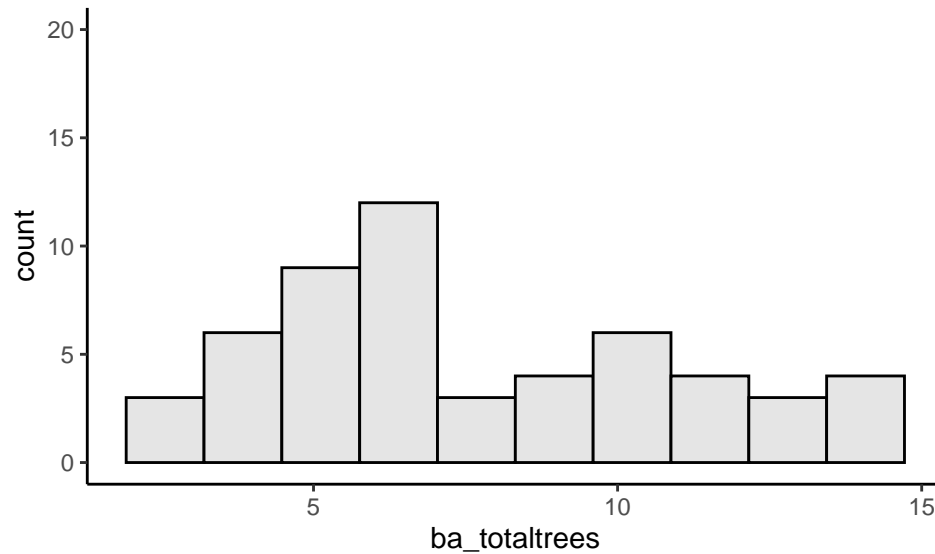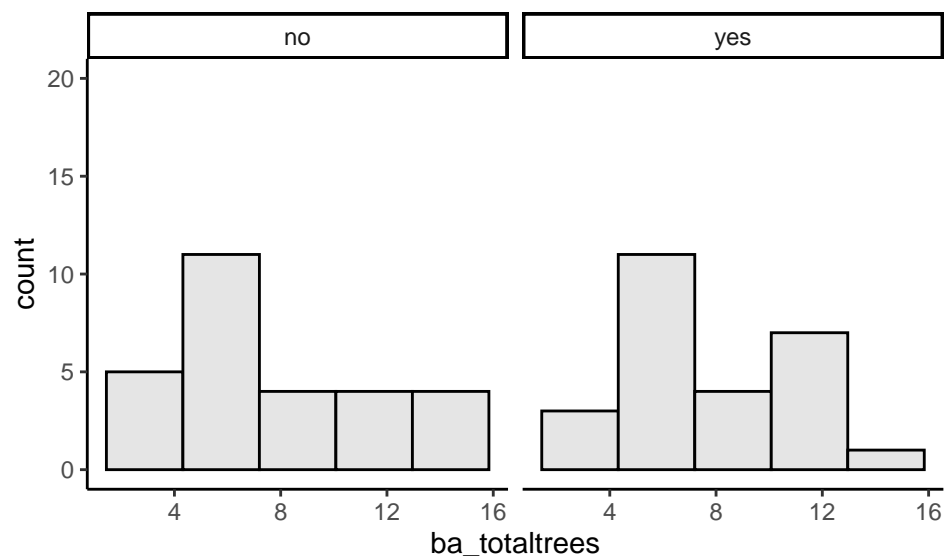
Figures

Histograms of total basal area of trees

```
# All plots
data_plots %>%
  ggplot() +
    geom_histogram(aes(x = ba_totaltrees), bins = 10, fill = "gray90", col = "black") +
    theme_classic() + ylim(c(0,20))
```



```
# Harvested vs. unharvested plots
data_plots %>%
  ggplot() +
  geom_histogram(aes(x = ba_totaltrees), bins = 5, fill = "gray90", col = "black") +
  facet_wrap(~harvested) +
  theme_classic() + ylim(c(0,20))
```

Five-number summaries of total basal area of trees

```
# All plots
summary(data_plots$ba_totaltrees)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   2.498   5.161   6.639   7.660  10.316  14.015
```

```
# Harvested vs. unharvested plots
data_plots %>% filter(harvested == "yes") %>% pull(ba_totaltrees) %>% summary()
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   2.788   4.860   6.506   7.550  10.248  13.785
```

```
data_plots %>% filter(harvested == "no") %>% pull(ba_totaltrees) %>% summary()
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   2.498   5.450   6.765   7.763  10.568  14.015
```

Box plots of total basal area of trees

```
# By harvesting status
ba_fig1 <- data_plots %>%
  mutate(harvested = fct_recode(harvested,
                                "Harvested" = "yes", "Unharvested" = "no")) %>%
  ggplot() +
    geom_boxplot(aes(x = fct_relevel(harvested, "Harvested", "Unharvested"),
                     y = ba_totaltrees),
                 fill = "mistyrose2", outlier.shape = 21,
                 outlier.color = "gray20", outlier.fill = "gray80") +
    plot_theme +
    coord_cartesian(ylim = c(0, 15)) +
    labs(x = "Plot status", y = "Total basal area of trees (m2 per hectare)")
ggsave("figures/ba_fig1.png", height = 4, width = 6)
ba_fig1
```
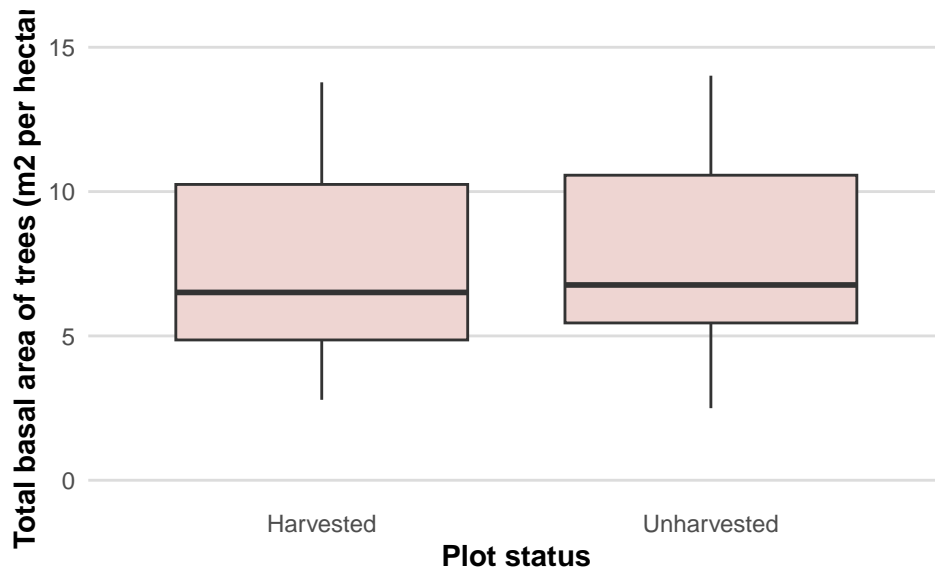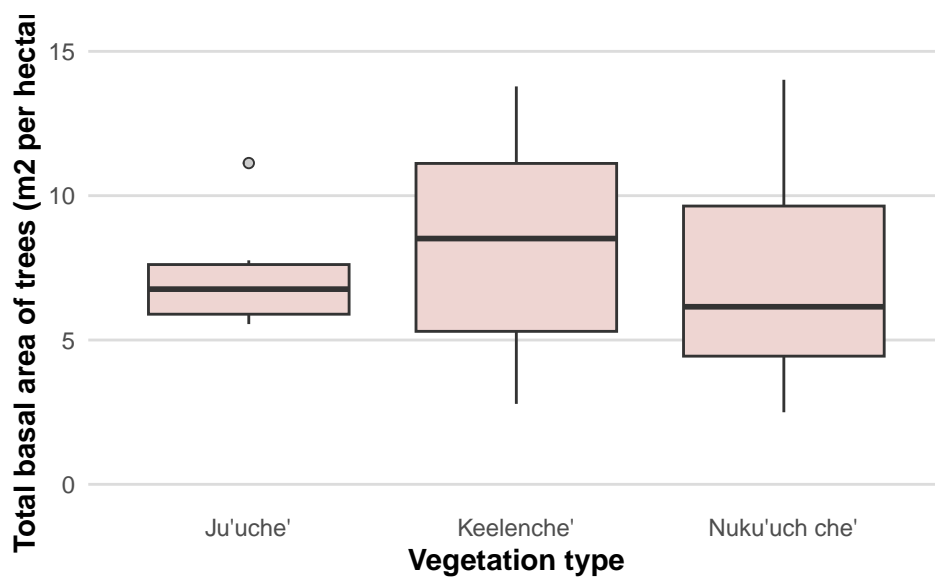
```
# By vegetation type
ba_fig2 <- data_plots %>%
  mutate(vegetation_type = fct_recode(vegetation_type,
                                      "Ju'uche'" = "juuche",
                                      "Keelenche'" = "keelenche",
                                      "Nuku'uch che'" = "nukuuchche")) %>%
  ggplot() +
  geom_boxplot(aes(x = vegetation_type, y = ba_totaltrees),
               fill = "mistyrose2", outlier.shape = 21,
               outlier.color = "gray20", outlier.fill = "gray80") +
  plot_theme +
  coord_cartesian(ylim = c(0, 15)) +
  labs(x = "Vegetation type", y = "Total basal area of trees (m2 per hectare)")
ggsave("figures/ba_fig2.png", height = 4, width = 6)
ba_fig2
```
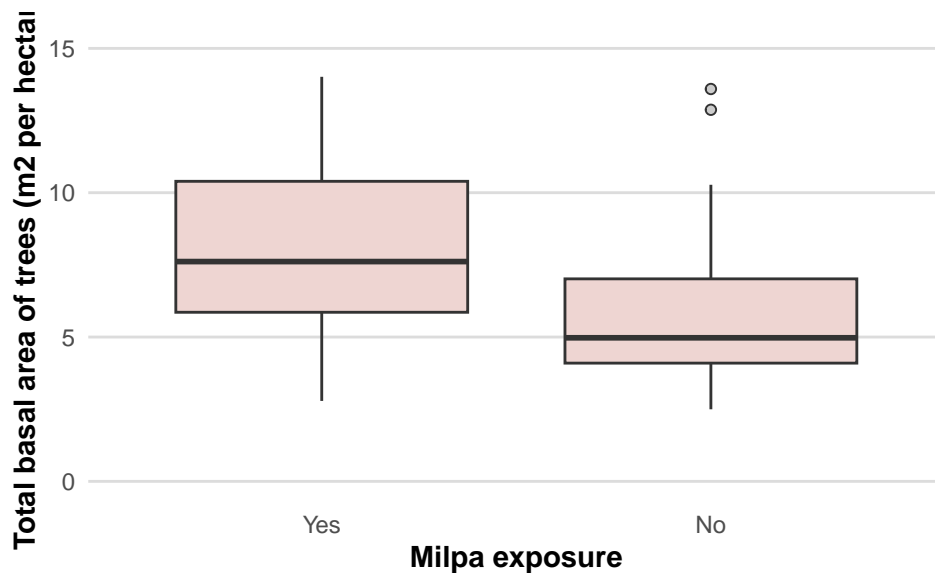
```
# By milpa
ba_fig3 <- data_plots %>%
  mutate(milpa = fct_recode(milpa,
                            "Yes" = "yes", "No" = "no")) %>%
  ggplot() +
  geom_boxplot(aes(x = fct_relevel(milpa, "Yes", "No"), y = ba_totaltrees),
               fill = "mistyrose2", outlier.shape = 21,
               outlier.color = "gray20", outlier.fill = "gray80") +
  plot_theme +
  coord_cartesian(ylim = c(0, 15)) +
  labs(x = "Milpa exposure", y = "Total basal area of trees (m2 per hectare)")
ggsave("figures/ba_fig3.png", height = 4, width = 6)
ba_fig3
```


Models

We found that the residuals of a standard linear model had heteroscedastic variance. We considered a Gaussian GLM and a gamma GLM, both with log link. The residuals of the former also had heteroscedastic variance, so we decide to proceed with the latter.

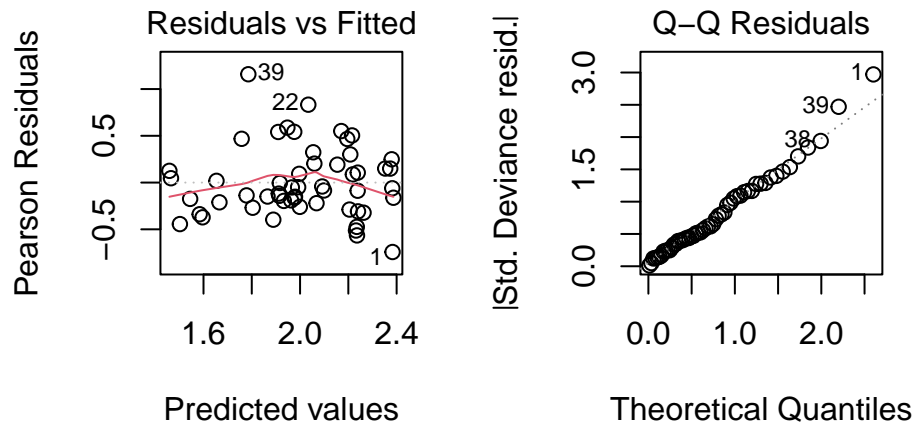### Fit gamma model and check residual plots

```
mod <- glm(ba_totaltrees ~ harvested + vegetation_type + milpa + latitude + longitude,
           family = Gamma(link="log"), data = data_plots)
```

```
par(mfrow = c(1,2))
plot(mod, which = 1:2)
```

Summarize the model

```
summary(mod)
```

```
##
## Call:
## glm(formula = ba_totaltrees ~ harvested + vegetation_type + milpa +
##     latitude + longitude, family = Gamma(link = "log"), data = data_plots)
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)               1.49958    0.23192   6.466 5.29e-08 ***
## harvestedyes              0.01319    0.10639   0.124  0.90184
## vegetation_typekeelenche  0.07888    0.17891   0.441  0.66133
## vegetation_typenukuuchche 0.27827    0.22012   1.264  0.21241
## milpayes                  0.51144    0.15402   3.321  0.00174 **
## latitude                 -0.05475    0.05890  -0.929  0.35740
## longitude                -0.21507    0.06816  -3.155  0.00280 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Gamma family taken to be 0.1521501)
##
##     Null deviance: 10.3415  on 53  degrees of freedom
## Residual deviance:  7.3464  on 47  degrees of freedom
## AIC: 272.3
##
## Number of Fisher Scoring iterations: 7
```

Construct analysis of deviance table

```
Anova(mod, type = 2, test.statistic = "LR")
```

```
## Analysis of Deviance Table (Type II tests)
##
```

```
## Response: ba_totaltrees
##                 LR Chisq Df Pr(>Chisq)
## harvested         0.0152  1  0.9019590
## vegetation_type   2.1595  2  0.3396839
## milpa            11.1473  1  0.0008416 ***
## latitude          0.8519  1  0.3560094
## longitude        10.3153  1  0.0013193 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```