Stem density

Tim White

June 2024

Packages and data

```
library(car)
library(tidyverse)
library(lme4)
```

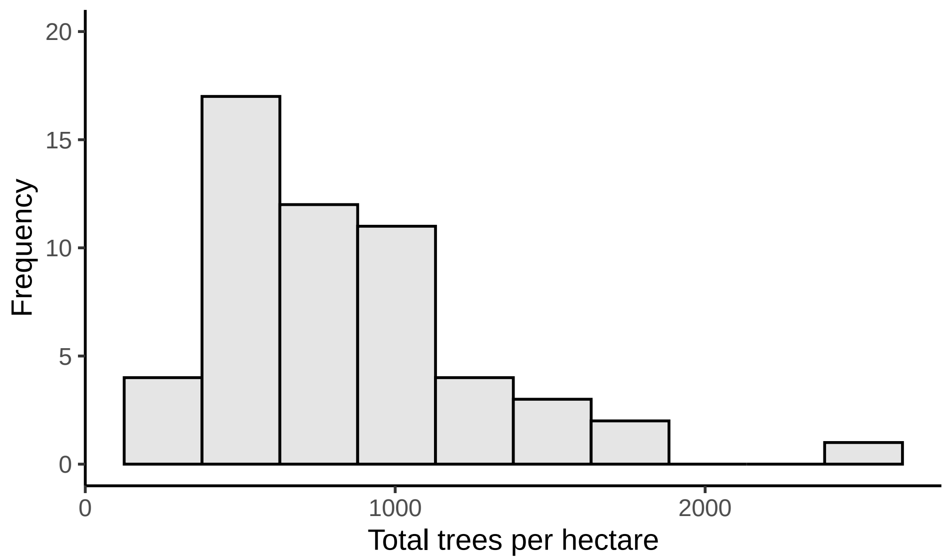
```
data_stems <- read_csv("data/data_stems.csv") %>%
                mutate(plot_id = as.factor(plot_id),
                       size_class = as.factor(size_class),
                       harvested = as.factor(harvested),
                       vegetation_type = as.factor(vegetation_type),
                       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)) %>%
                mutate(seedling_plots_per_ha = 20 / 10000,
                       plots_per_ha = 100 * pi / 10000) %>%
                mutate(stemden_totaltrees_count = stemden_totaltrees * plots_per_ha,
                       stemden_seedlings_count = stemden_seedlings * seedling_plots_per_ha,
                       stemden_saplings_count = stemden_saplings * plots_per_ha,
                       stemden_trees05to09_count = stemden_trees05to09 * plots_per_ha,
                       stemden_trees10to14_count = stemden_trees10to14 * plots_per_ha,
                       stemden_trees15plus_count = stemden_trees15plus * plots_per_ha)

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"))
```

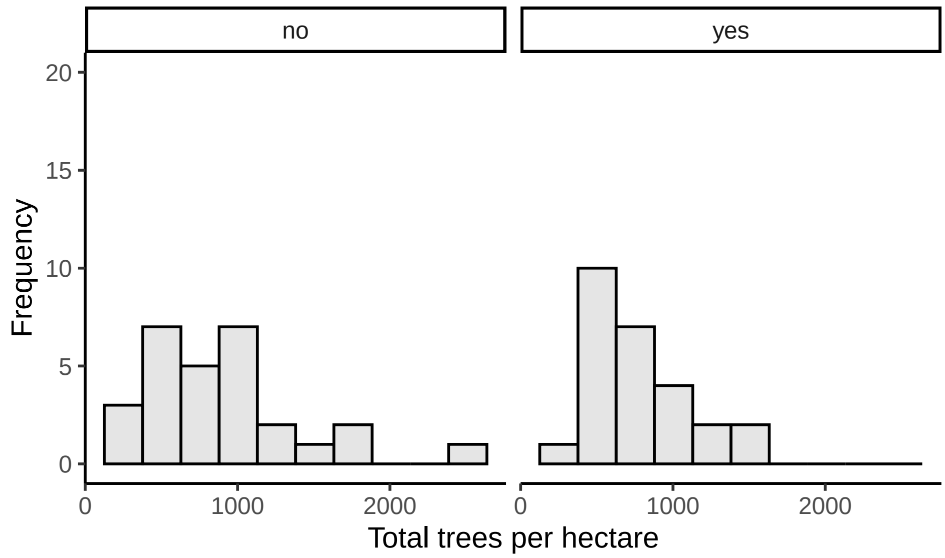
Figures

Histograms of total trees per hectare

```
# All plots
data_plots %>%
  ggplot() +
  geom_histogram(aes(x = stemden_totaltrees), bins = 10, fill = "gray90", col = "black") +
  theme_classic() + ylim(c(0,20)) + labs(x = "Total trees per hectare", y = "Frequency")
```




```
# Harvested vs. unharvested plots
data_plots %>%
  ggplot() +
  geom_histogram(aes(x = stemden_totaltrees), bins = 10, fill = "gray90", col = "black") +
  facet_wrap(~harvested) +
  theme_classic() + ylim(c(0,20)) + labs(x = "Total trees per hectare", y = "Frequency")
```



Five-number summaries of total trees per hectare

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

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   254.6   557.0   779.9   848.8  1018.6  2514.6
```

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

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   254.6   517.3   716.2   778.6   923.1  1496.1
```

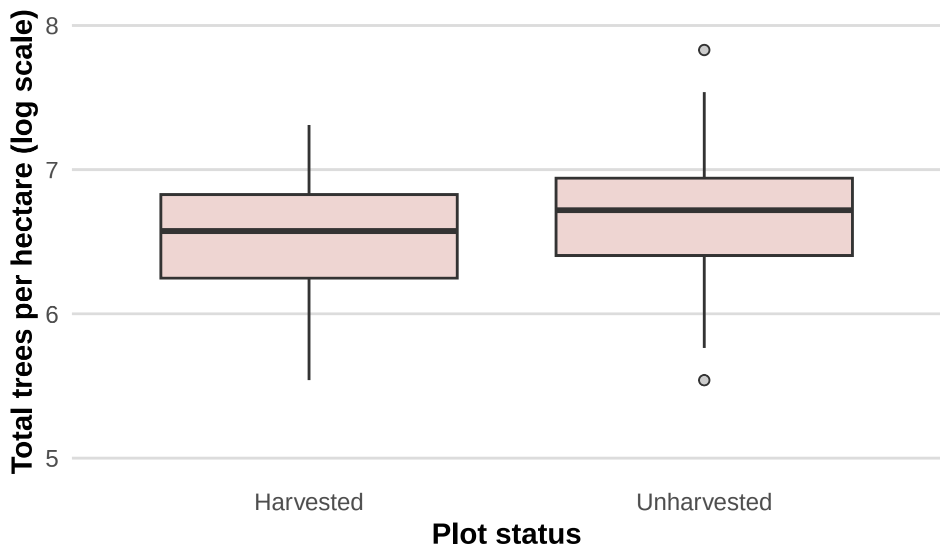
```
data_plots %>% filter(harvested == "no") %>% pull(stemden_totaltrees) %>% summary()
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   254.6   604.8   827.6   914.0  1034.5  2514.6
```

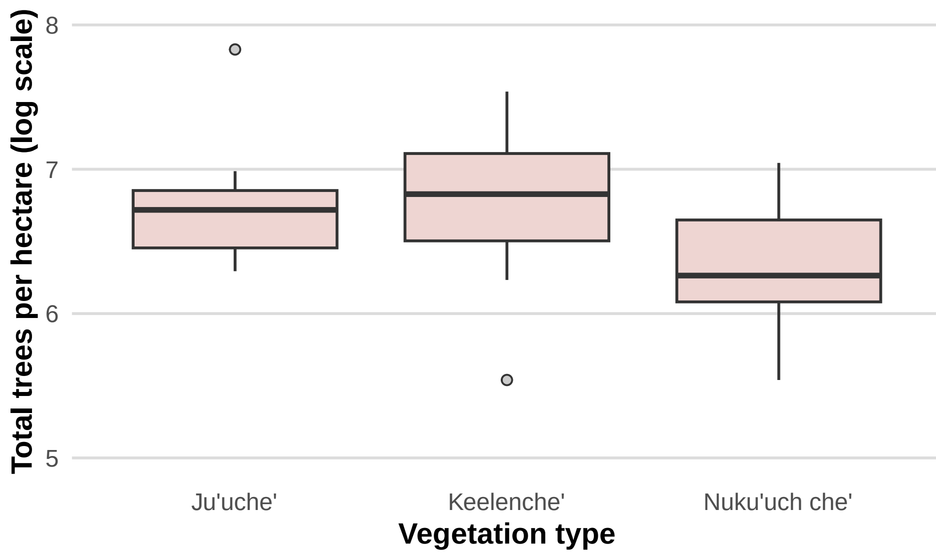
Box plots of total trees per hectare (log scale)

```
# By harvesting status
stemden_fig1 <- data_plots %>%
  mutate(harvested = fct_recode(harvested,
                                "Harvested" = "yes", "Unharvested" = "no")) %>%
  ggplot() +
  geom_boxplot(aes(x = fct_relevel(harvested, "Harvested", "Unharvested"),
                   y = log(stemden_totaltrees)),
               fill = "mistyrose2", outlier.shape = 21,
               outlier.color = "gray20", outlier.fill = "gray80") +
```

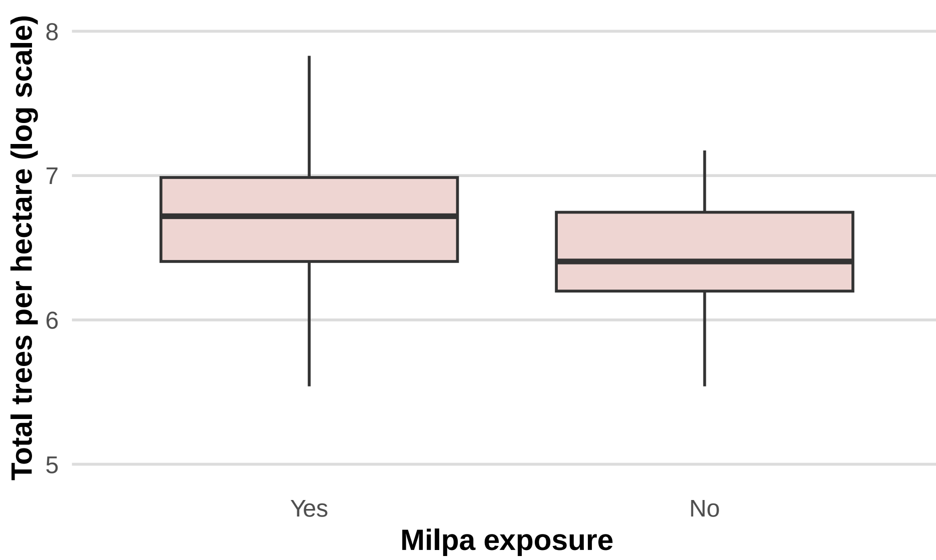
```
  plot_theme +
  coord_cartesian(ylim = c(5, 8)) +
  labs(x = "Plot status", y = "Total trees per hectare (log scale)")
ggsave("figures/stemden_fig1.png", height = 4, width = 6)
stemden_fig1
```




```
# By vegetation type
stemden_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 = log(stemden_totaltrees)),
               fill = "mistyrose2", outlier.shape = 21,
               outlier.color = "gray20", outlier.fill = "gray80") +
  plot_theme +
  coord_cartesian(ylim = c(5, 8)) +
  labs(x = "Vegetation type", y = "Total trees per hectare (log scale)")
ggsave("figures/stemden_fig2.png", height = 4, width = 6)
stemden_fig2
```

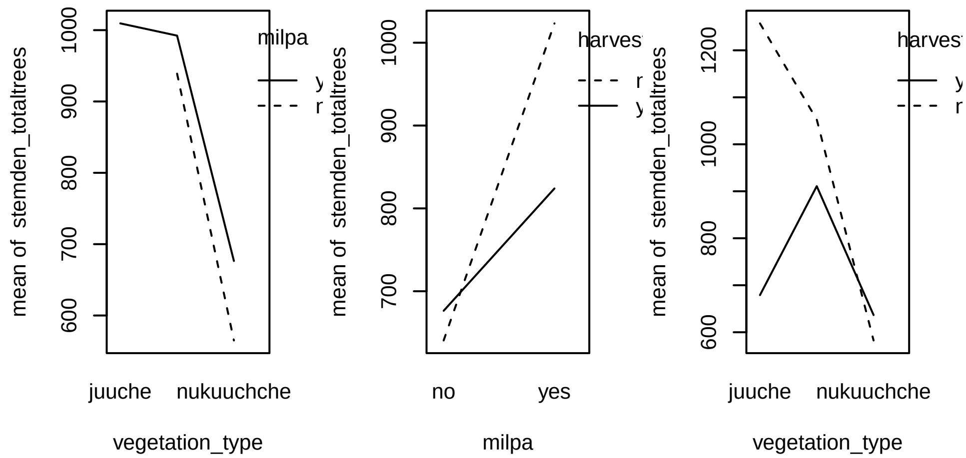



```
# By milpa exposure
stemden_fig3 <- data_plots %>%
  mutate(milpa = fct_recode(milpa,
                            "Yes" = "yes", "No" = "no")) %>%
  ggplot() +
  geom_boxplot(aes(x = fct_relevel(milpa, "Yes", "No"), y = log(stemden_totaltrees)),
               fill = "mistyrose2", outlier.shape = 21,
               outlier.color = "gray20", outlier.fill = "gray80") +
  plot_theme +
  coord_cartesian(ylim = c(5, 8)) +
  labs(x = "Milpa exposure", y = "Total trees per hectare (log scale)")
ggsave("figures/stemden_fig3.png", height = 4, width = 6)
stemden_fig3
```



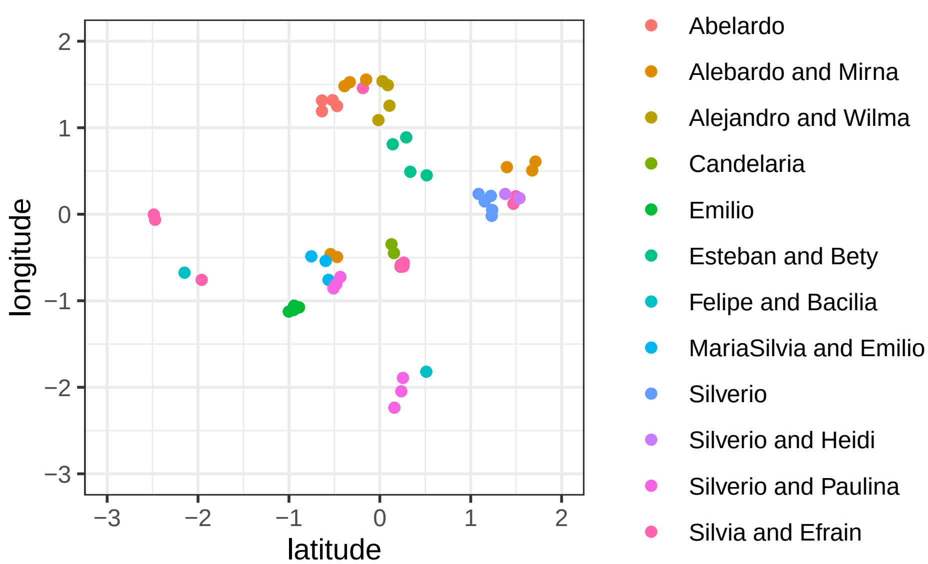
Interaction plots

```
par(mfrow = c(1,3))
with(data_plots, {interaction.plot(vegetation_type, milpa, stemden_totaltrees)
                  interaction.plot(milpa, harvested, stemden_totaltrees)
                  interaction.plot(vegetation_type, harvested, stemden_totaltrees)})
```



Latitude and longitude by artisan

```
data_plots %>%
  ggplot(aes(x = latitude, y = longitude)) +
    geom_point(aes(color = artisan), shape = 19) +
    xlim(-3, 2) +
    ylim(-3, 2) +
    theme_bw()
```



Models

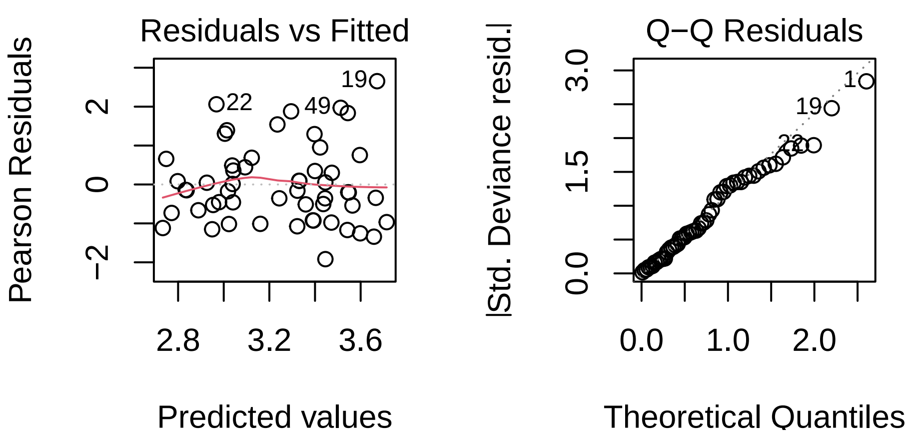
We have fewer than five observations for each artisan, so we elect not to include artisan as a random effect. Also, some individuals appear in more than one level of artisan, so we could not disentangle variability due to each individual.

We decide to fit a negative binomial model with all other covariates, including latitude and longitude. We found that the variance of the residuals increases quadratically with the mean, so we believe a negative binomial model is more appropriate than a quasipoisson GLM.

Fit negative binomial model and check residual plots

```
mod <- MASS::glm.nb(stemden_totaltrees_count ~ harvested + vegetation_type + milpa +
                         harvested:vegetation_type + harvested:milpa + latitude + longitude +
                         offset(plots_per_ha), data = data_plots)
```

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



Summarize the model

```
summary(mod)
```

```
##
## Call:
## MASS::glm.nb(formula = stemden_totaltrees_count ~ harvested +
##     vegetation_type + milpa + harvested:vegetation_type + harvested:milpa +
##     latitude + longitude + offset(plots_per_ha), data = data_plots,
##     init.theta = 8.413800108, link = log)
##
## Coefficients:
##                                         Estimate Std. Error z value Pr(>|z|)
## (Intercept)                             3.406376   0.298190  11.424   <2e-16
## harvestedyes                           -0.555248   0.414112  -1.341   0.1800
## vegetation_typekeelenche               -0.179506   0.227138  -0.790   0.4294
```

```
## vegetation_typenukuuchche              -0.595615   0.285526  -2.086   0.0370
## milpayes                                0.238074   0.214022   1.112   0.2660
## latitude                               -0.004282   0.059109  -0.072   0.9423
## longitude                              -0.086717   0.071057  -1.220   0.2223
## harvestedyes:vegetation_typekeelenche   0.463946   0.339983   1.365   0.1724
## harvestedyes:vegetation_typenukuuchche  0.680179   0.392460   1.733   0.0831
## harvestedyes:milpayes                  -0.056634   0.282870  -0.200   0.8413
##
## (Intercept)                            ***
## harvestedyes
## vegetation_typekeelenche
## vegetation_typenukuuchche              *
## milpayes
## latitude
## longitude
## harvestedyes:vegetation_typekeelenche
## harvestedyes:vegetation_typenukuuchche .
## harvestedyes:milpayes
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(8.4138) family taken to be 1)
##
##     Null deviance: 78.876  on 53  degrees of freedom
## Residual deviance: 53.599  on 44  degrees of freedom
## AIC: 420.64
##
## Number of Fisher Scoring iterations: 1
##
##
##               Theta:  8.41
##           Std. Err.:  2.10
##
##  2 x log-likelihood:  -398.639
```

Construct analysis of deviance table

```
# Use type 2 SS since data are unbalanced)
Anova(mod, type = 2, test.statistic = "LR")
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: stemden_totaltrees_count
##                          LR Chisq Df Pr(>Chisq)
## harvested                  1.1297  1     0.2878
## vegetation_type            3.4255  2     0.1804
## milpa                      1.6884  1     0.1938
## latitude                   0.0056  1     0.9406
## longitude                  1.4938  1     0.2216
## harvested:vegetation_type  3.0312  2     0.2197
## harvested:milpa            0.0414  1     0.8388
```