

Basal area by size class

Tim White

June 2024

Packages, data, and functions

```
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))

compute_bootstrap_interval <- function(var, statistic = mean, num_iters = 50000,
                                         percentiles = c(0.025, 0.975)) {
  boot_samples <- numeric(num_iters)

  for (i in 1:num_iters) {
    boot_samples[i] <- statistic(sample(var, size = length(var), replace = TRUE))
  }

  return(quantile(boot_samples, percentiles))
}

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

Average basal area by size class

Compute bootstrap intervals and organize data

```

harvested <- data_plots %>%
  filter(harvested == "yes") %>%
  select(ba_saplings, starts_with("ba_trees"))
unharvested <- data_plots %>%
  filter(harvested == "no") %>%
  select(ba_saplings, starts_with("ba_trees"))

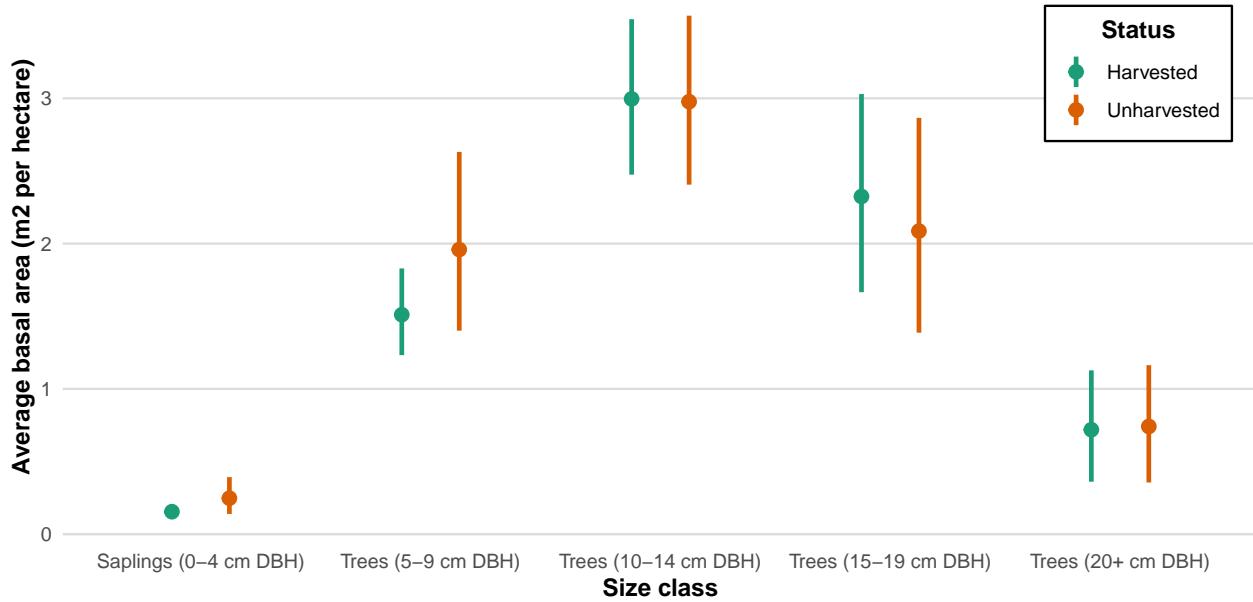
set.seed(1)
boot_intervals <- bind_rows(compute_bootstrap_interval(unharvested$ba_saplings),
                             compute_bootstrap_interval(unharvested$ba_trees05to09),
                             compute_bootstrap_interval(unharvested$ba_trees10to14),
                             compute_bootstrap_interval(unharvested$ba_trees15to19),
                             compute_bootstrap_interval(unharvested$ba_trees20plus),
                             compute_bootstrap_interval(harvested$ba_saplings),
                             compute_bootstrap_interval(harvested$ba_trees05to09),
                             compute_bootstrap_interval(harvested$ba_trees10to14),
                             compute_bootstrap_interval(harvested$ba_trees15to19),
                             compute_bootstrap_interval(harvested$ba_trees20plus))

ba_mean_data <- data_plots %>%
  select(harvested, ba_saplings, starts_with("ba_trees")) %>%
  group_by(harvested) %>%
  summarize_all(mean) %>%
  pivot_longer(cols = starts_with("ba"), names_prefix = "ba_",
               names_to = "size_class", values_to = "mean") %>%
  bind_cols(boot_intervals) %>%
  select(harvested, size_class, lower = `2.5%`, mean, upper = `97.5%`) %>%
  mutate(size_class = as.factor(size_class)) %>%
  mutate(harvested = fct_recode(harvested,
                                "Harvested" = "yes", "Unharvested" = "no")) %>%
  mutate(size_class = fct_recode(size_class,
                                 "Saplings (0-4 cm DBH)" = "saplings",
                                 "Trees (5-9 cm DBH)" = "trees05to09",
                                 "Trees (10-14 cm DBH)" = "trees10to14",
                                 "Trees (15-19 cm DBH)" = "trees15to19",
                                 "Trees (20+ cm DBH)" = "trees20plus"))

```

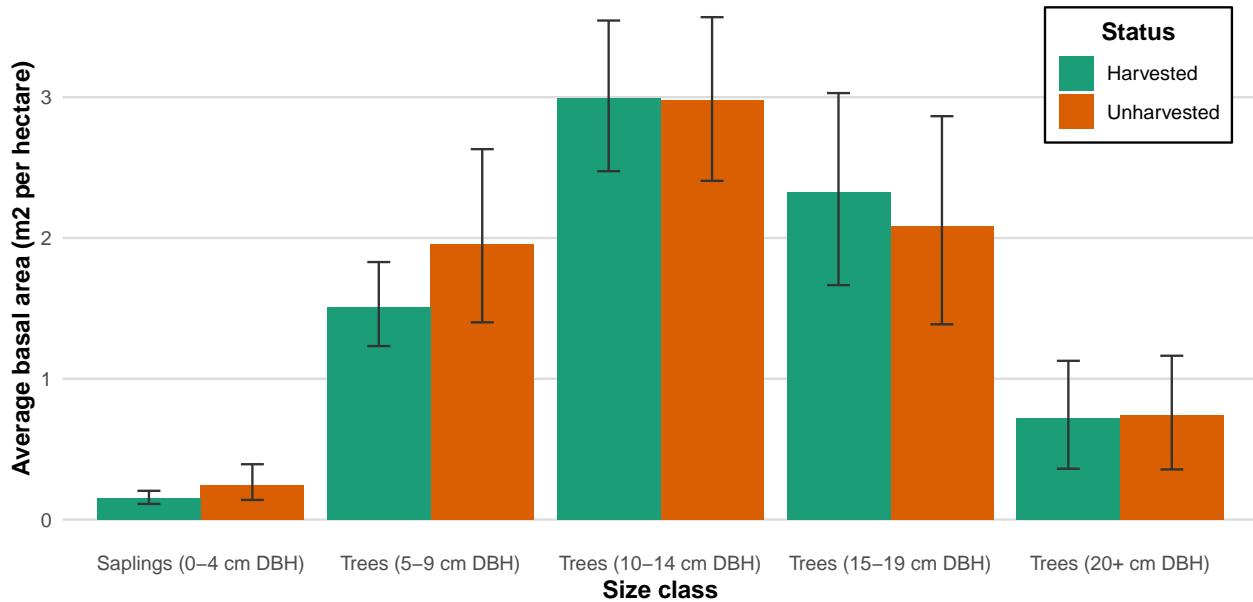
Pointrange plot

```
ba_sizeclass_fig1 <- ba_mean_data %>%
  ggplot() +
  geom_pointrange(aes(x = size_class, col = fct_relevel(harvested, "Harvested", "Unharvested"),
                       ymin = lower, y = mean, ymax = upper),
                  position = position_dodge(width = 0.5), size = 0.5, linewidth = 1) +
  plot_theme +
  scale_color_brewer(palette = "Dark2") +
  labs(x = "Size class", y = "Average basal area (m2 per hectare)") +
  guides(col=guide_legend(title = "Status"))
ggsave("figures/ba_sizeclass_fig1.png", height = 4, width = 8)
ba_sizeclass_fig1
```



Bar plot

```
ba_sizeclass_fig2 <- ba_mean_data %>%
  ggplot(aes(x = size_class,
             fill = fct_relevel(harvested, "Harvested", "Unharvested"))) +
  geom_col(aes(y = mean),
            position = "dodge") +
  geom_errorbar(aes(x = size_class, ymin = lower, ymax = upper), col = "gray20",
                position = position_dodge(width = 0.9), width = 0.2, show.legend = FALSE) +
  plot_theme +
  scale_fill_brewer(palette = "Dark2") +
  labs(x = "Size class", y = "Average basal area (m2 per hectare)") +
  guides(fill=guide_legend(title = "Status"))
ggsave("figures/ba_sizeclass_fig2.png", height = 4, width = 8)
ba_sizeclass_fig2
```



Distribution of basal area by size class

Organize data

```
ba_dist_data <- data_plots %>%
  select(harvested, ba_saplings, starts_with("ba_trees")) %>%
  pivot_longer(cols = starts_with("ba"),
               names_prefix = "ba_",
               names_to = "size_class",
               values_to = "ba") %>%
  mutate(size_class = as.factor(size_class)) %>%
  mutate(harvested = fct_recode(harvested,
                                "Harvested" = "yes",
                                "Unharvested" = "no")) %>%
  mutate(size_class = fct_recode(size_class,
                                 "Saplings (0-4 cm DBH)" = "saplings",
                                 "Trees (5-9 cm DBH)" = "trees05to09",
                                 "Trees (10-14 cm DBH)" = "trees10to14",
                                 "Trees (15-19 cm DBH)" = "trees15to19",
                                 "Trees (20+ cm DBH)" = "trees20plus"))
```

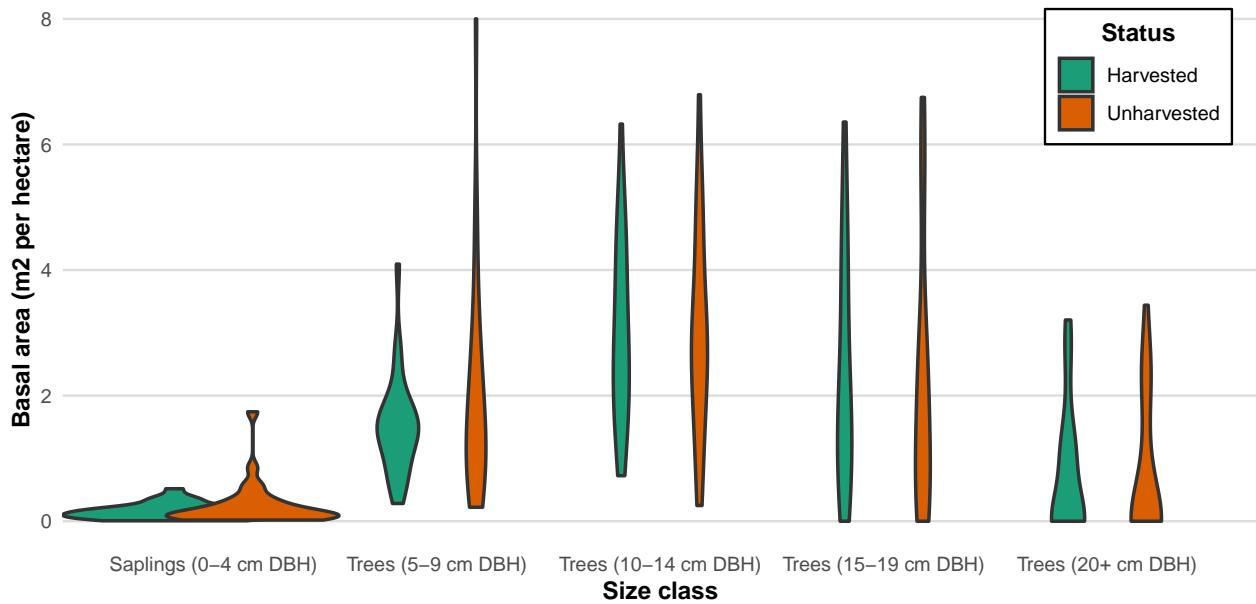
Violin plot

```
ba_dist_data %>%
  ggplot() +
  geom_violin(aes(x = size_class, y = ba,
                  fill = fct_relevel(harvested, "Harvested", "Unharvested")),
              scale = "area", width = 2, position = position_dodge(width = 0.7),
              lwd = 0.75, kernel = "gaussian", adjust = 1.25) +
```

```

plot_theme +
scale_fill_brewer(palette = "Dark2") +
labs(x = "Size class", y = "Basal area (m2 per hectare)") +
guides(fill = guide_legend(title = "Status"))

```

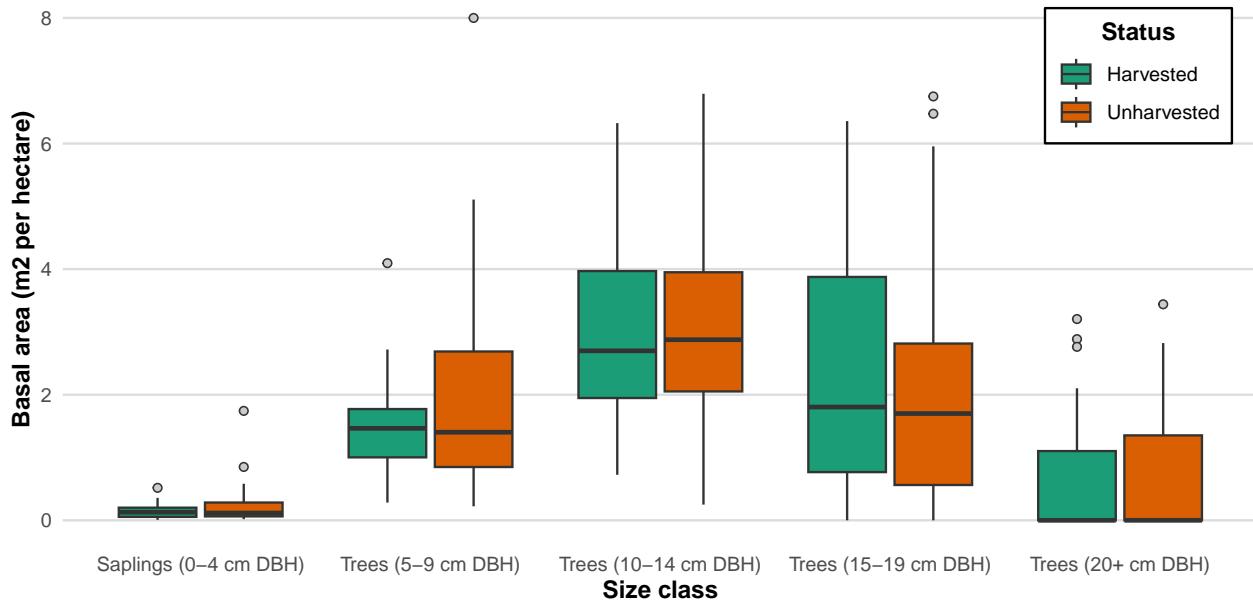


Box plot

```

ba_sizeclass_fig3 <- ba_dist_data %>%
ggplot() +
geom_boxplot(aes(x = size_class, y = ba,
fill = fct_relevel(harvested, "Harvested", "Unharvested")),
outlier.shape = 21, outlier.color = "gray20",
outlier.fill = "gray80") +
plot_theme +
scale_fill_brewer(palette = "Dark2") +
labs(x = "Size class", y = "Basal area (m2 per hectare)") +
guides(fill = guide_legend(title = "Status"))
ggsave("figures/ba_sizeclass_fig3.png", height = 4, width = 8)
ba_sizeclass_fig3

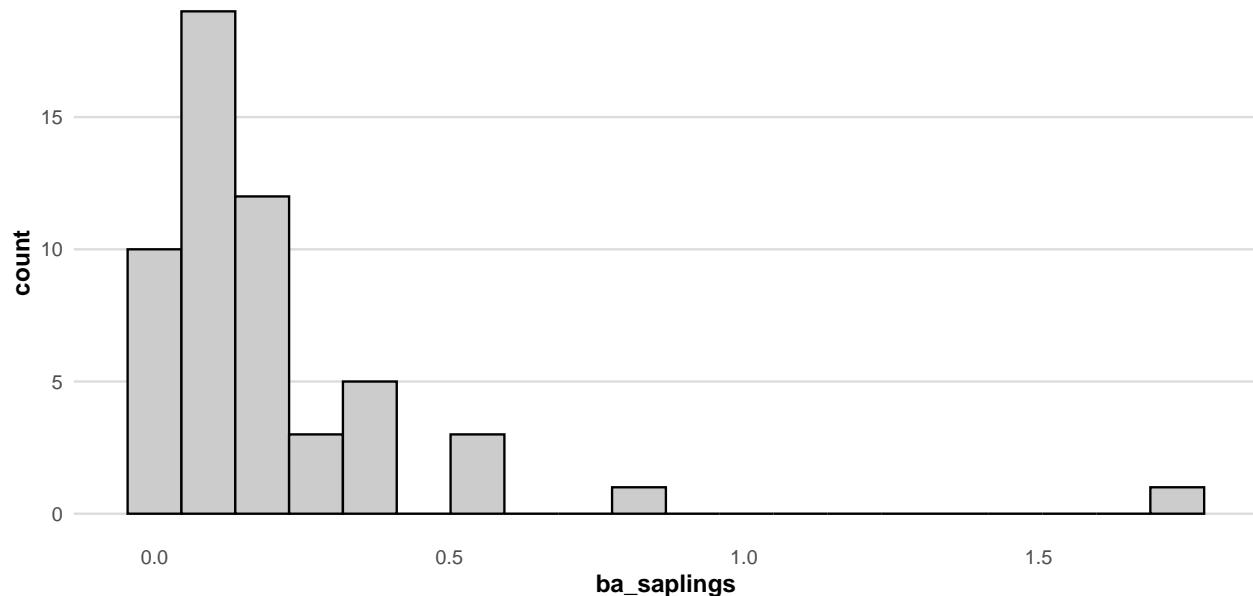
```



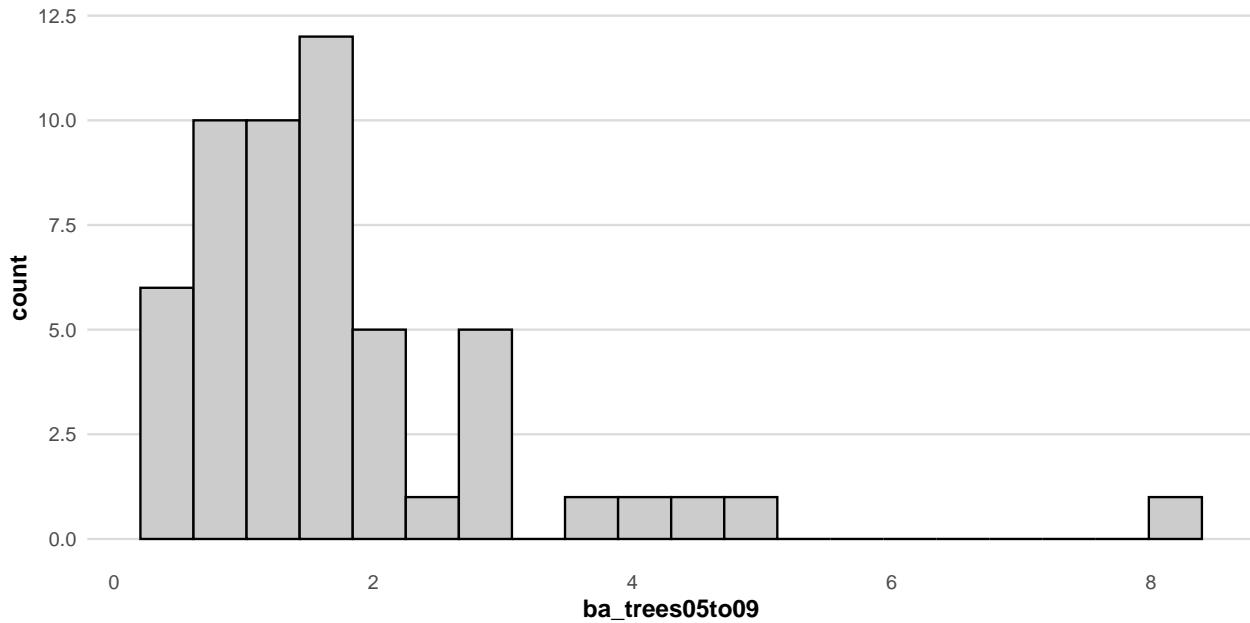
Basal area by size class for different covariates

Histograms, harvesting status

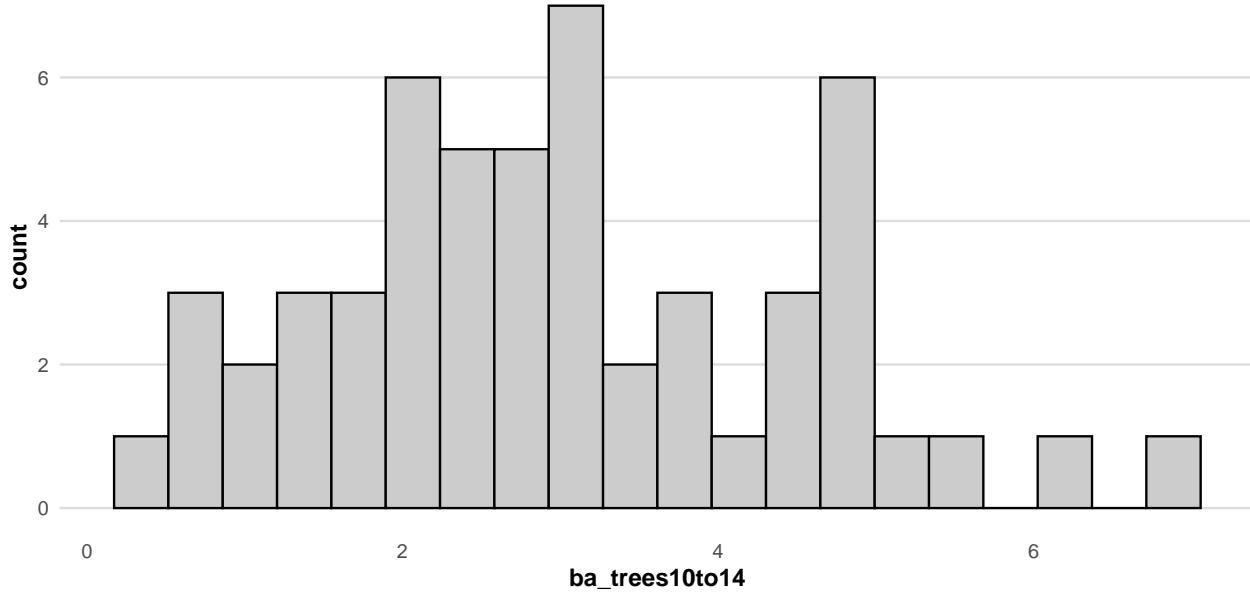
```
data_plots %>%
  ggplot() + geom_histogram(aes(x = ba_saplings), bins = 20,
                            col = "black", fill = "gray80") + plot_theme
```



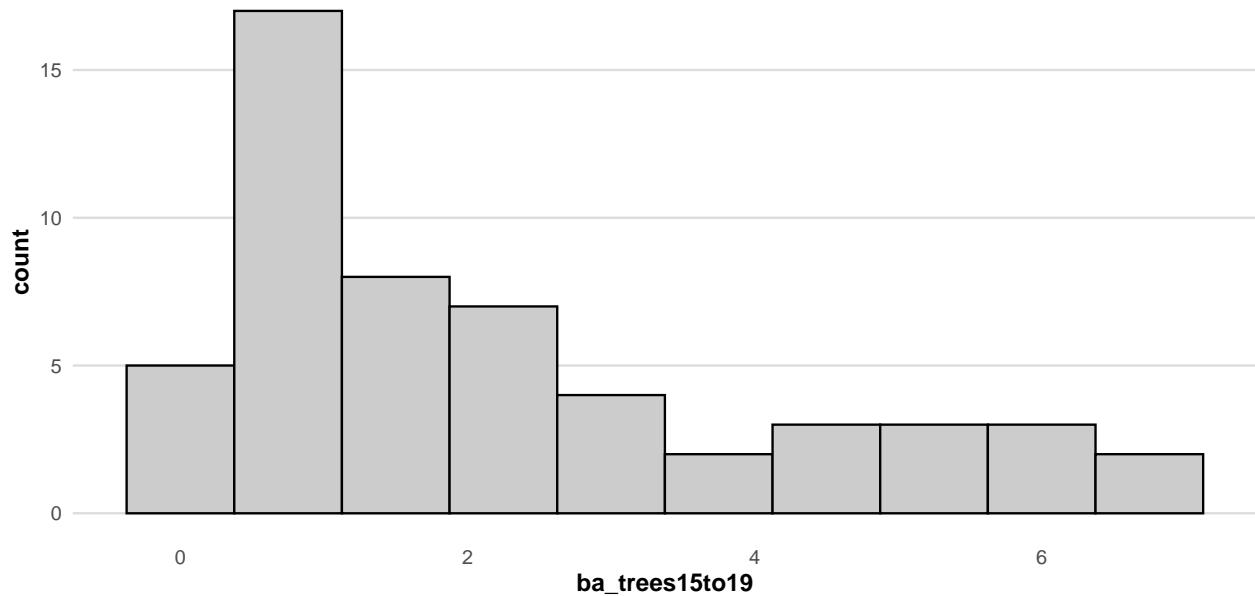
```
data_plots %>%
  ggplot() + geom_histogram(aes(x = ba_trees05to09), bins = 20,
                            col = "black", fill = "gray80") + plot_theme
```



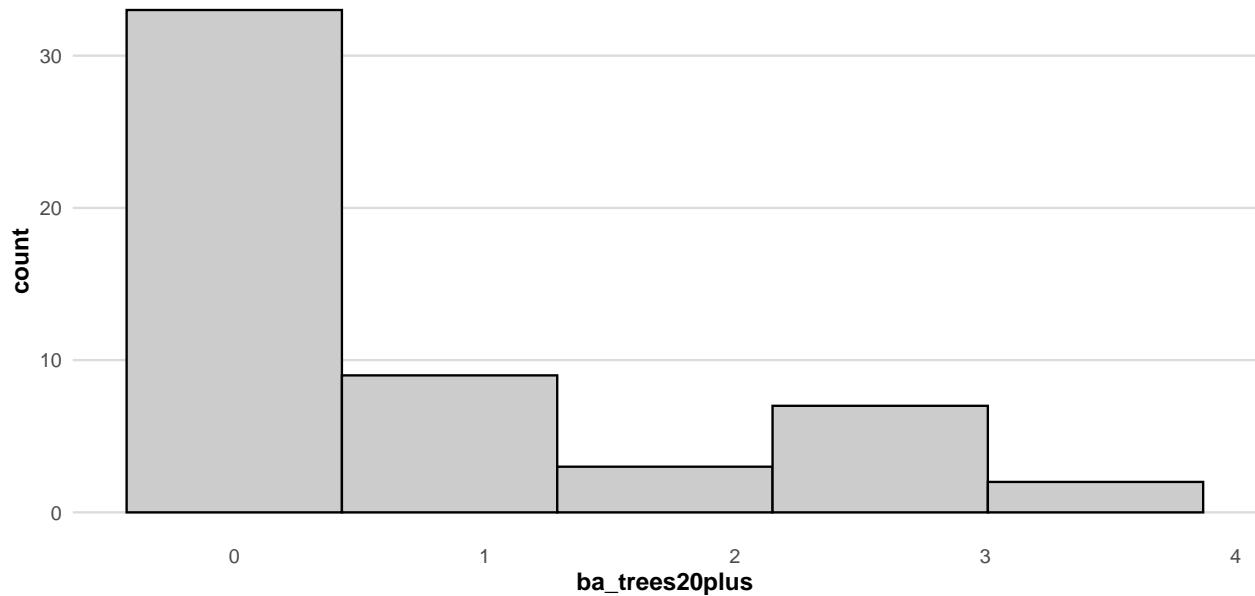
```
data_plots %>%
  ggplot() + geom_histogram(aes(x = ba_trees10to14), bins = 20,
                            col = "black", fill = "gray80") + plot_theme
```



```
data_plots %>%
  ggplot() + geom_histogram(aes(x = ba_trees15to19), bins = 10,
                            col = "black", fill = "gray80") + plot_theme
```



```
data_plots %>%
  ggplot() + geom_histogram(aes(x = ba_trees20plus), bins = 5,
                            col = "black", fill = "gray80") + plot_theme
```



Five-number summaries, harvesting status

```
summary(data_plots$ba_saplings)
```

```
##      Min. 1st Qu. Median      Mean 3rd Qu.      Max.
## 0.0075  0.0575  0.1200  0.2029  0.2169  1.7425
```

```

data_plots %>% filter(harvested == "yes") %>% pull(ba_saplings) %>% summary()

##      Min. 1st Qu. Median     Mean 3rd Qu.     Max.
## 0.00750 0.05437 0.12875 0.15471 0.20125 0.51750

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

##      Min. 1st Qu. Median     Mean 3rd Qu.     Max.
## 0.01812 0.06250 0.11625 0.24770 0.28281 1.74250

summary(data_plots$ba_trees05to09)

##      Min. 1st Qu. Median     Mean 3rd Qu.     Max.
## 0.2225 0.9025 1.4650 1.7430 1.9406 8.0000

data_plots %>% filter(harvested == "yes") %>% pull(ba_trees05to09) %>% summary()

##      Min. 1st Qu. Median     Mean 3rd Qu.     Max.
## 0.2825 1.0025 1.4650 1.5106 1.7719 4.0950

data_plots %>% filter(harvested == "no") %>% pull(ba_trees05to09) %>% summary()

##      Min. 1st Qu. Median     Mean 3rd Qu.     Max.
## 0.2225 0.8488 1.4025 1.9588 2.6881 8.0000

summary(data_plots$ba_trees10to14)

##      Min. 1st Qu. Median     Mean 3rd Qu.     Max.
## 0.250 1.989 2.751 2.986 3.969 6.793

data_plots %>% filter(harvested == "yes") %>% pull(ba_trees10to14) %>% summary()

##      Min. 1st Qu. Median     Mean 3rd Qu.     Max.
## 0.725 1.948 2.699 2.996 3.969 6.325

data_plots %>% filter(harvested == "no") %>% pull(ba_trees10to14) %>% summary()

##      Min. 1st Qu. Median     Mean 3rd Qu.     Max.
## 0.250 2.053 2.876 2.977 3.950 6.793

summary(data_plots$ba_trees15to19)

##      Min. 1st Qu. Median     Mean 3rd Qu.     Max.
## 0.0000 0.6606 1.7388 2.2009 2.8919 6.7500

```

```

data_plots %>% filter(harvested == "yes") %>% pull(ba_trees15to19) %>% summary()

##      Min. 1st Qu. Median     Mean 3rd Qu.    Max.
## 0.0000  0.7675 1.8037  2.3242  3.8756  6.3575

data_plots %>% filter(harvested == "no") %>% pull(ba_trees15to19) %>% summary()

##      Min. 1st Qu. Median     Mean 3rd Qu.    Max.
## 0.0000  0.5625 1.7000  2.0863  2.8150  6.7500

summary(data_plots$ba_trees20plus)

##      Min. 1st Qu. Median     Mean 3rd Qu.    Max.
## 0.0000  0.0000 0.0000  0.7306  1.1025  3.4400

data_plots %>% filter(harvested == "yes") %>% pull(ba_trees20plus) %>% summary()

##      Min. 1st Qu. Median     Mean 3rd Qu.    Max.
## 0.0000  0.0000 0.0000  0.7189  1.1025  3.2050

data_plots %>% filter(harvested == "no") %>% pull(ba_trees20plus) %>% summary()

##      Min. 1st Qu. Median     Mean 3rd Qu.    Max.
## 0.0000  0.0000 0.0000  0.7415  1.3525  3.4400

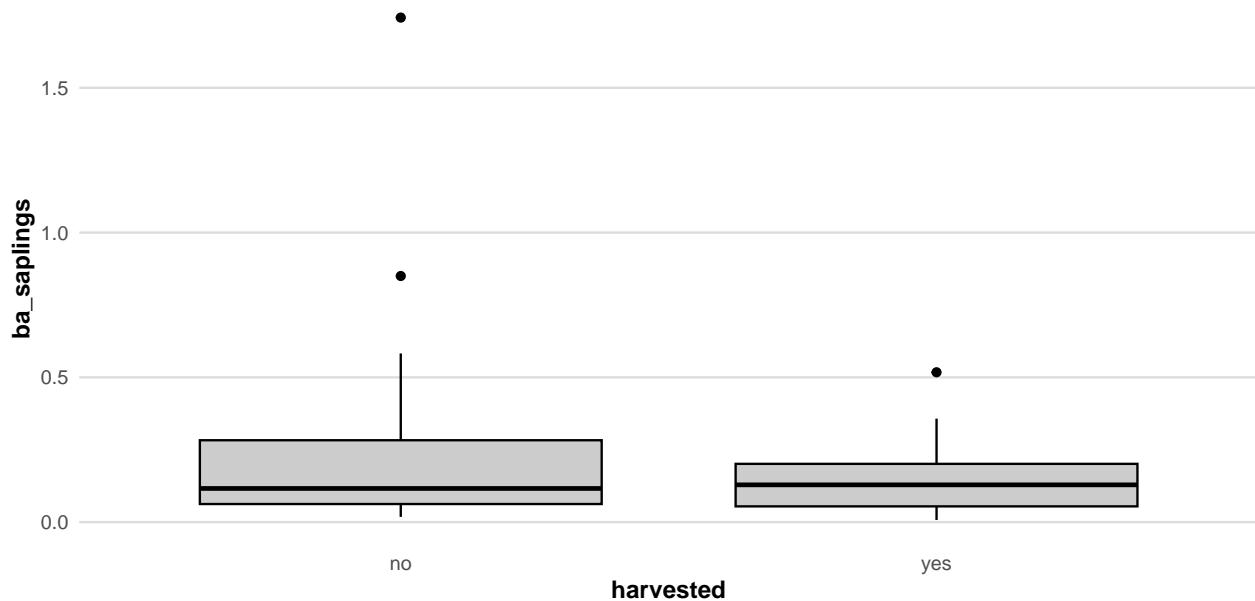
```

Box plots, harvesting status

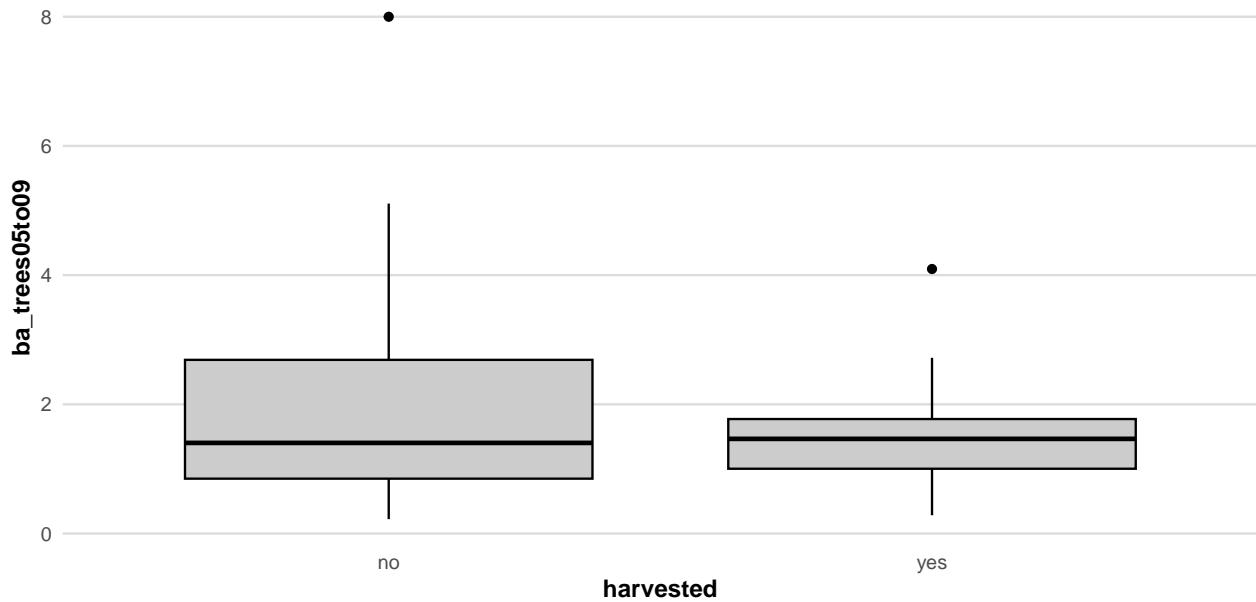
```

data_plots %>% ggplot() +
  geom_boxplot(aes(x = harvested, y = ba_saplings),
                col = "black", fill = "gray80") + plot_theme

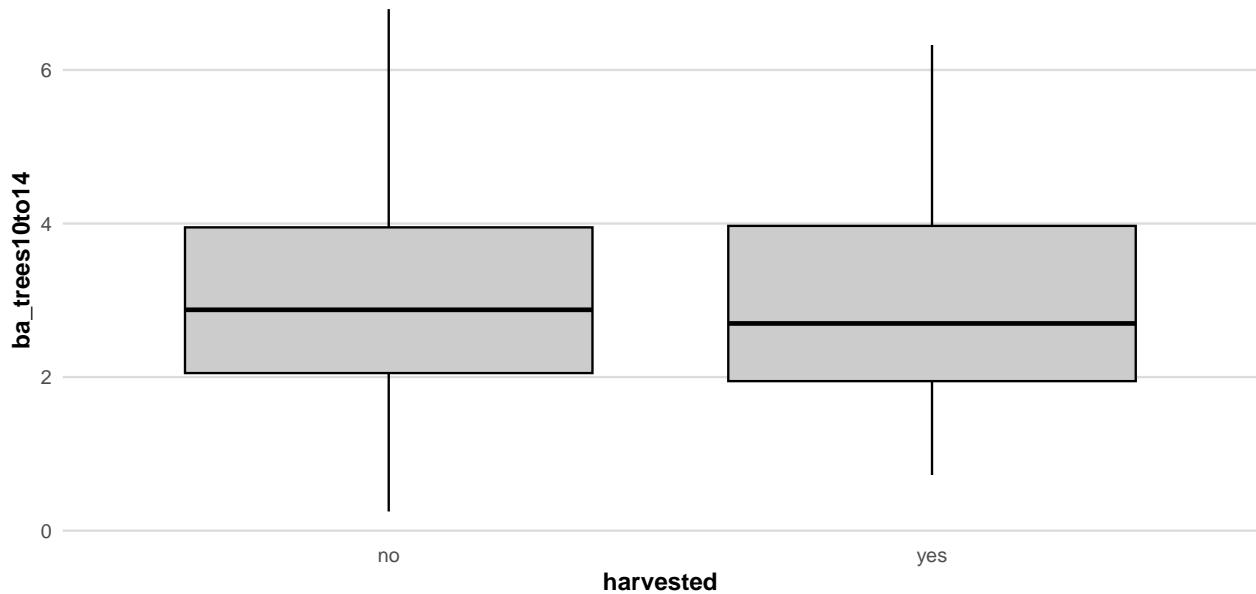
```



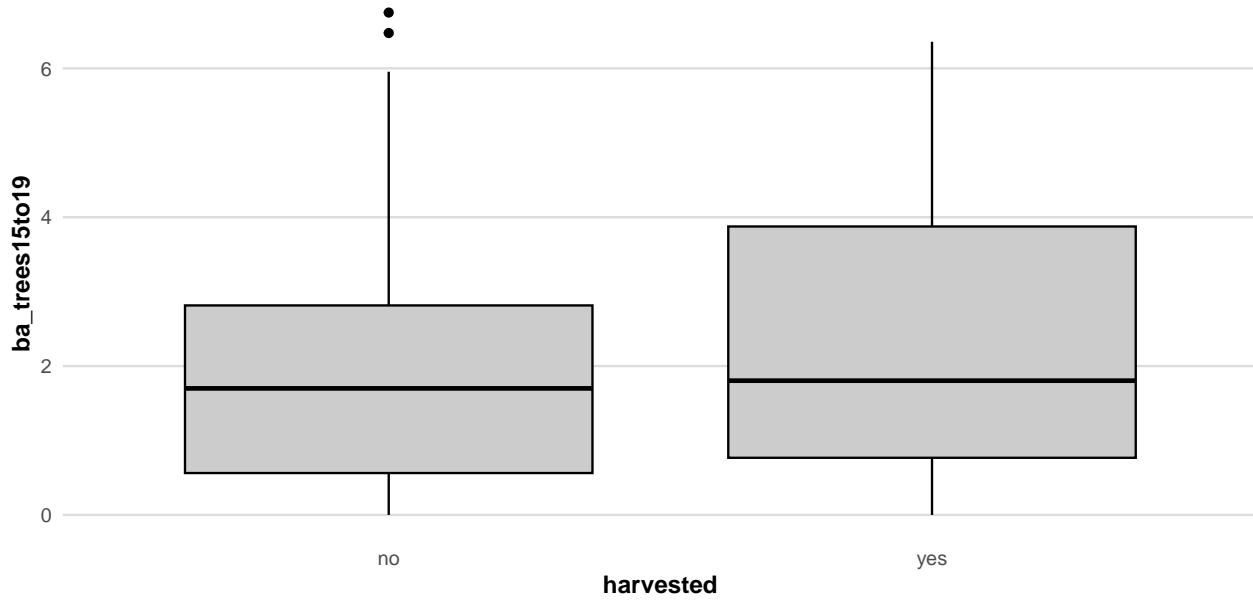
```
data_plots %>% ggplot() +
  geom_boxplot(aes(x = harvested, y = ba_trees05to09),
    col = "black", fill = "gray80") + plot_theme
```



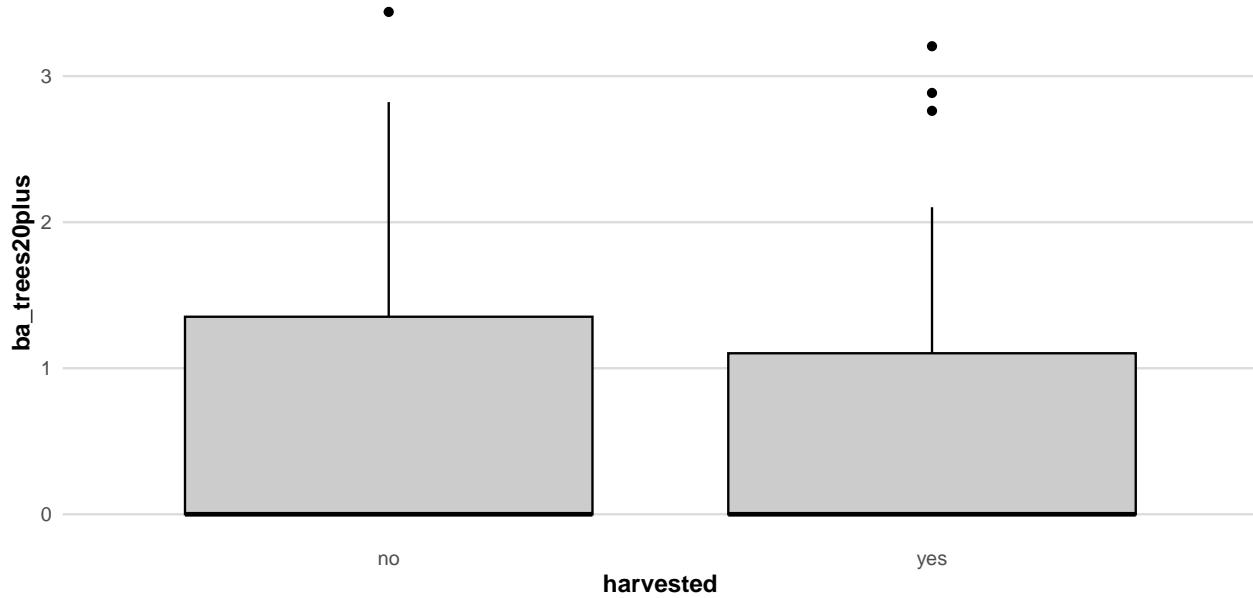
```
data_plots %>% ggplot() +
  geom_boxplot(aes(x = harvested, y = ba_trees10to14),
    col = "black", fill = "gray80") + plot_theme
```



```
data_plots %>% ggplot() +
  geom_boxplot(aes(x = harvested, y = ba_trees15to19),
    col = "black", fill = "gray80") + plot_theme
```

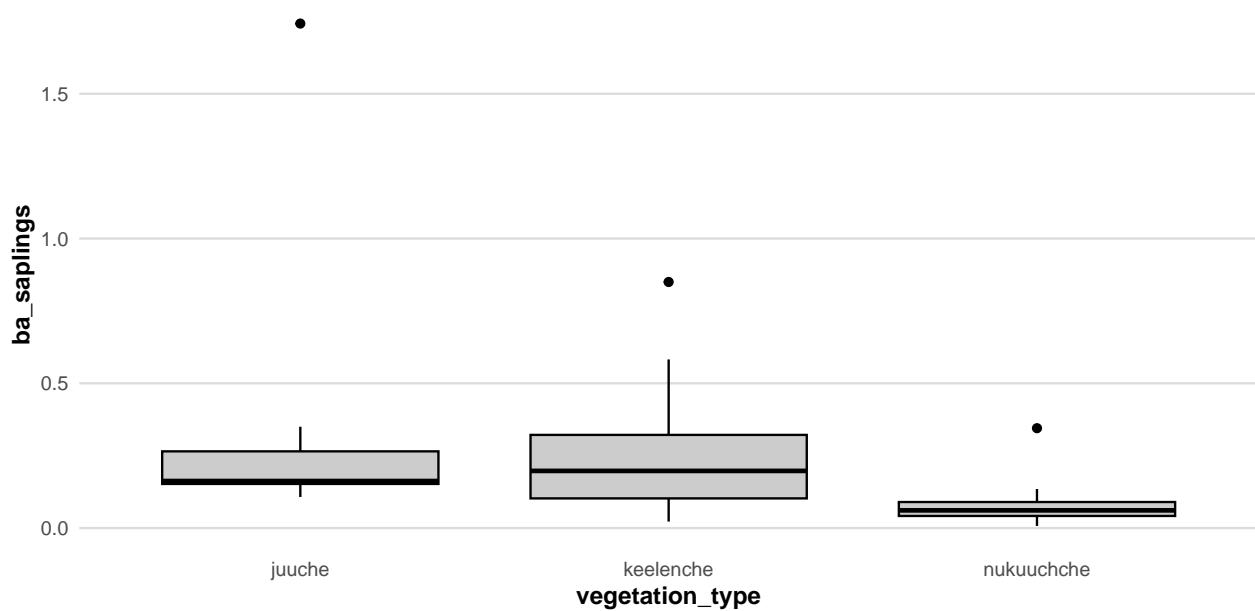


```
data_plots %>% ggplot() +
  geom_boxplot(aes(x = harvested, y = ba_trees20plus),
               col = "black", fill = "gray80") + plot_theme
```

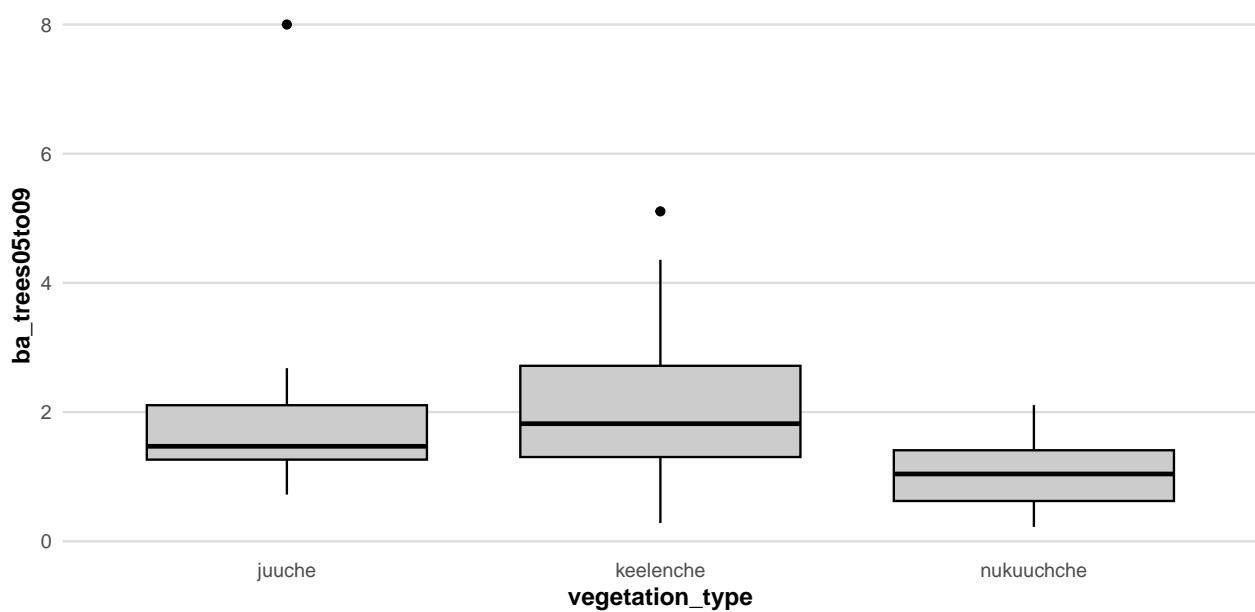


Box plots, vegetation type

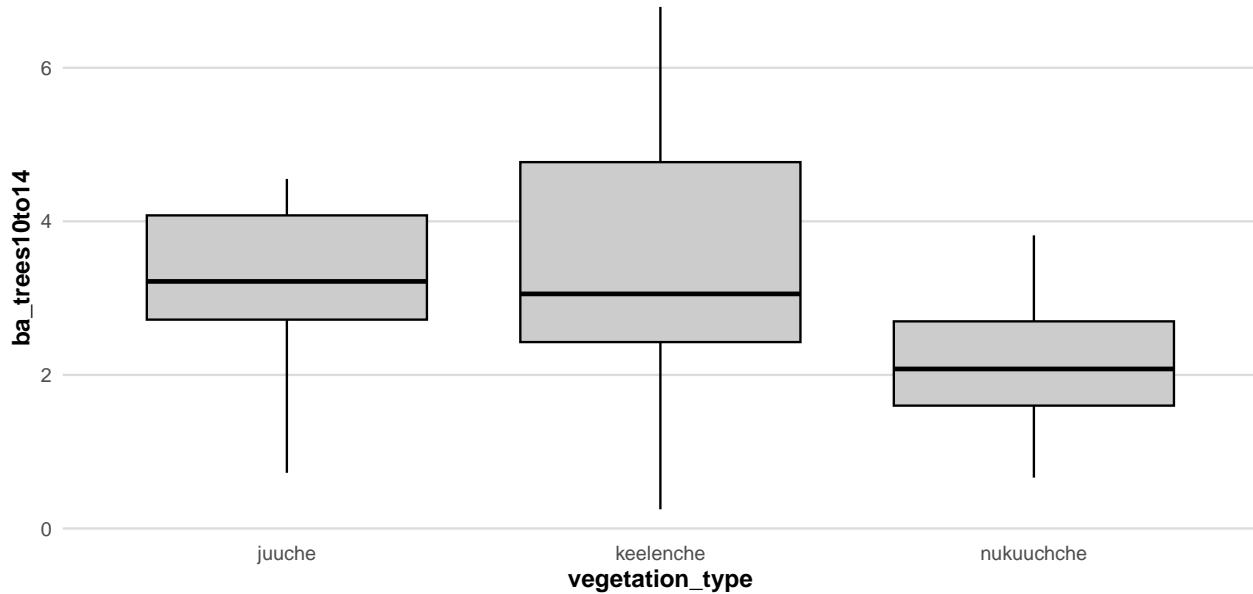
```
data_plots %>% ggplot() +
  geom_boxplot(aes(x = vegetation_type, y = ba_saplings),
               col = "black", fill = "gray80") + plot_theme
```



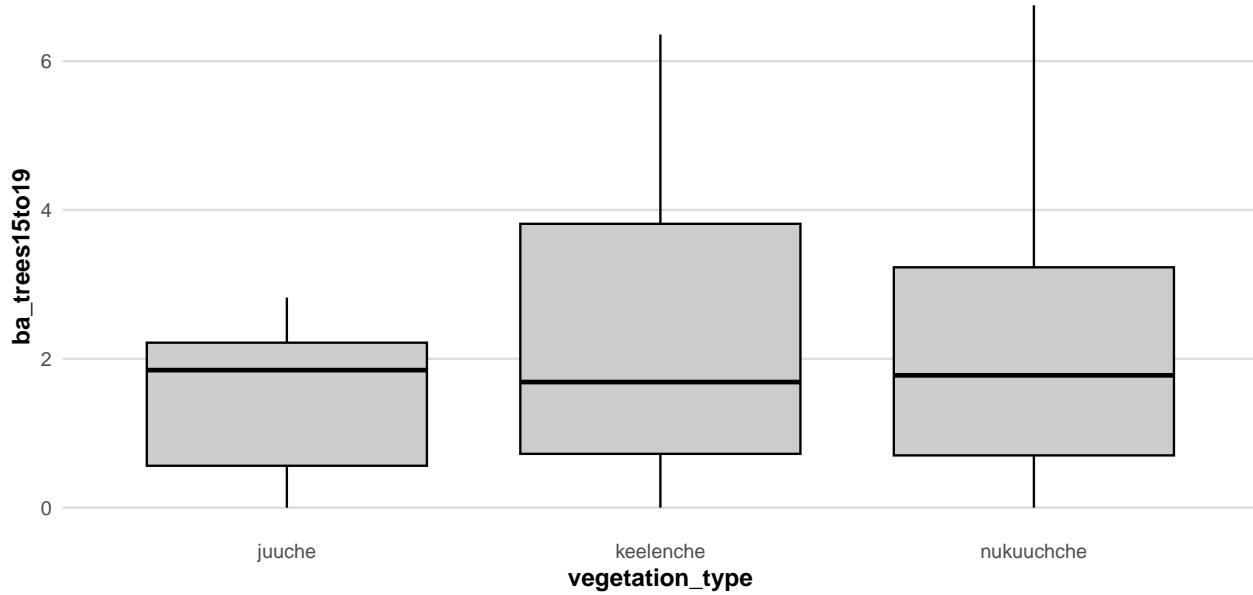
```
data_plots %>% ggplot() +
  geom_boxplot(aes(x = vegetation_type, y = ba_trees05to09),
               col = "black", fill = "gray80") + plot_theme
```



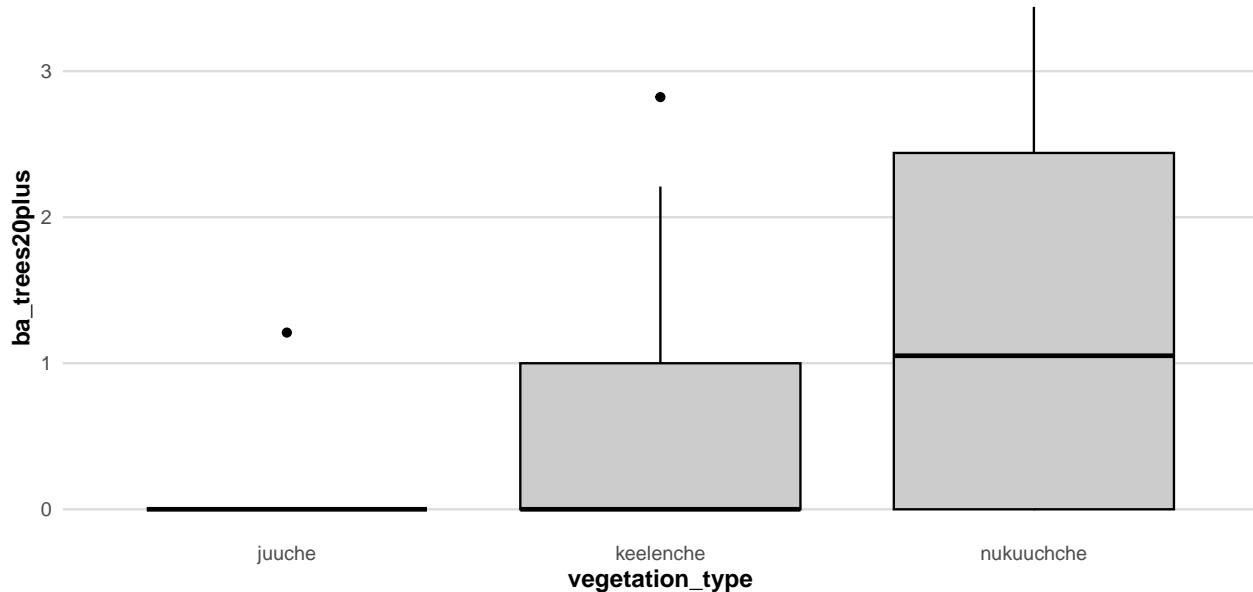
```
data_plots %>% ggplot() +
  geom_boxplot(aes(x = vegetation_type, y = ba_trees10to14),
               col = "black", fill = "gray80") + plot_theme
```



```
data_plots %>% ggplot() +
  geom_boxplot(aes(x = vegetation_type, y = ba_trees15to19),
               col = "black", fill = "gray80") + plot_theme
```

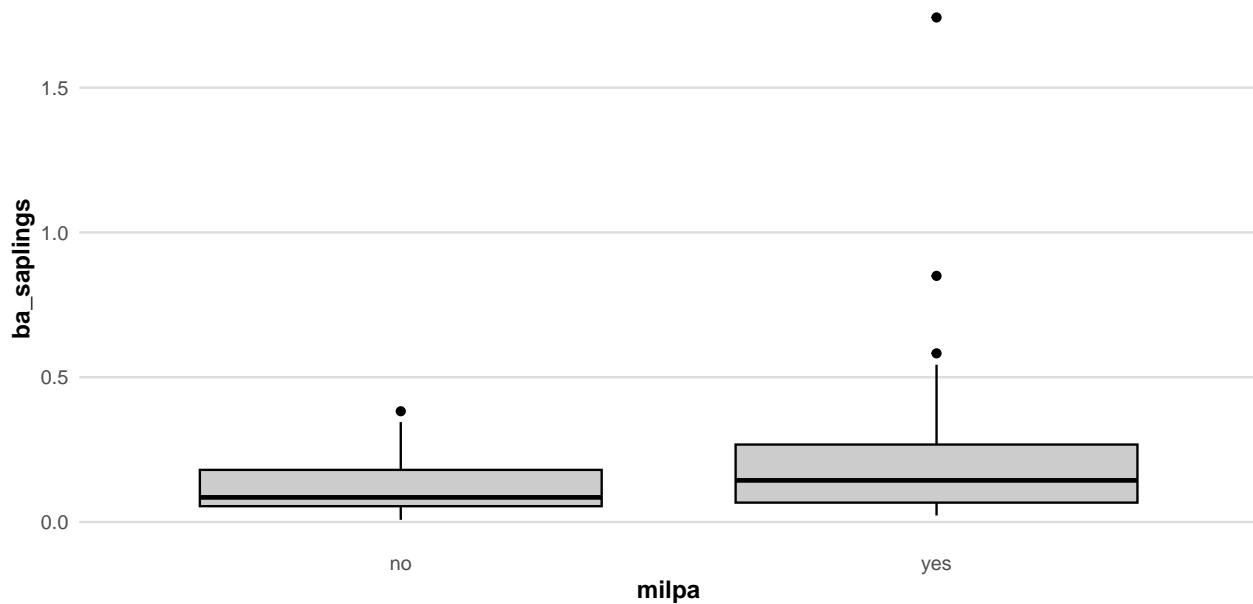


```
data_plots %>% ggplot() +
  geom_boxplot(aes(x = vegetation_type, y = ba_trees20plus),
               col = "black", fill = "gray80") + plot_theme
```

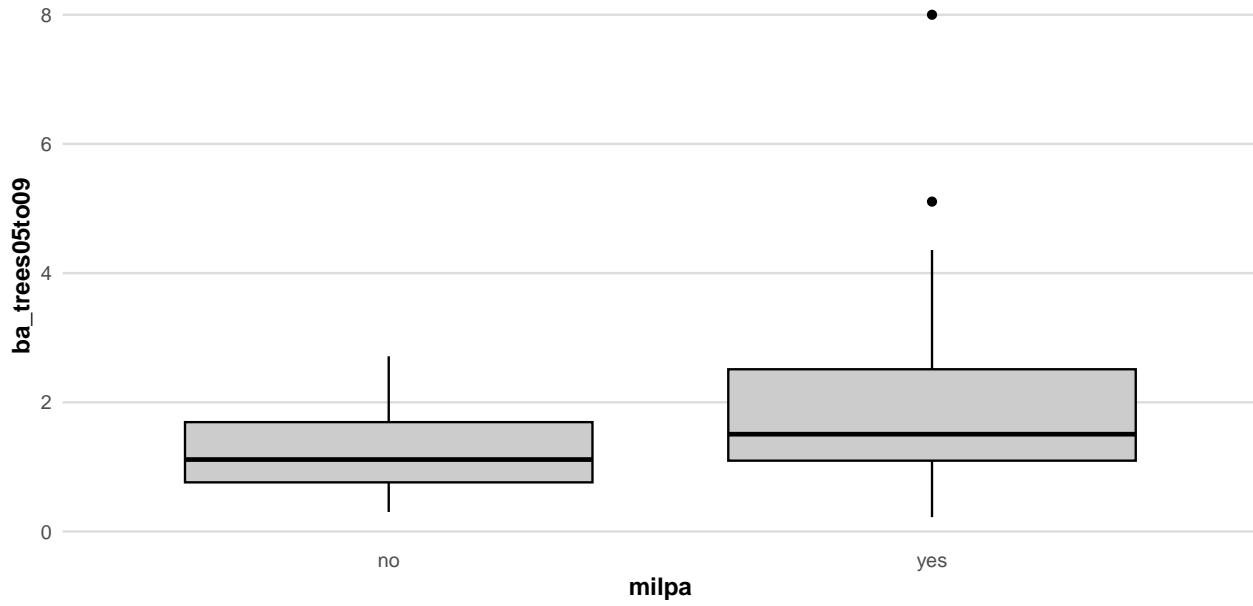


Box plots, milpa exposure

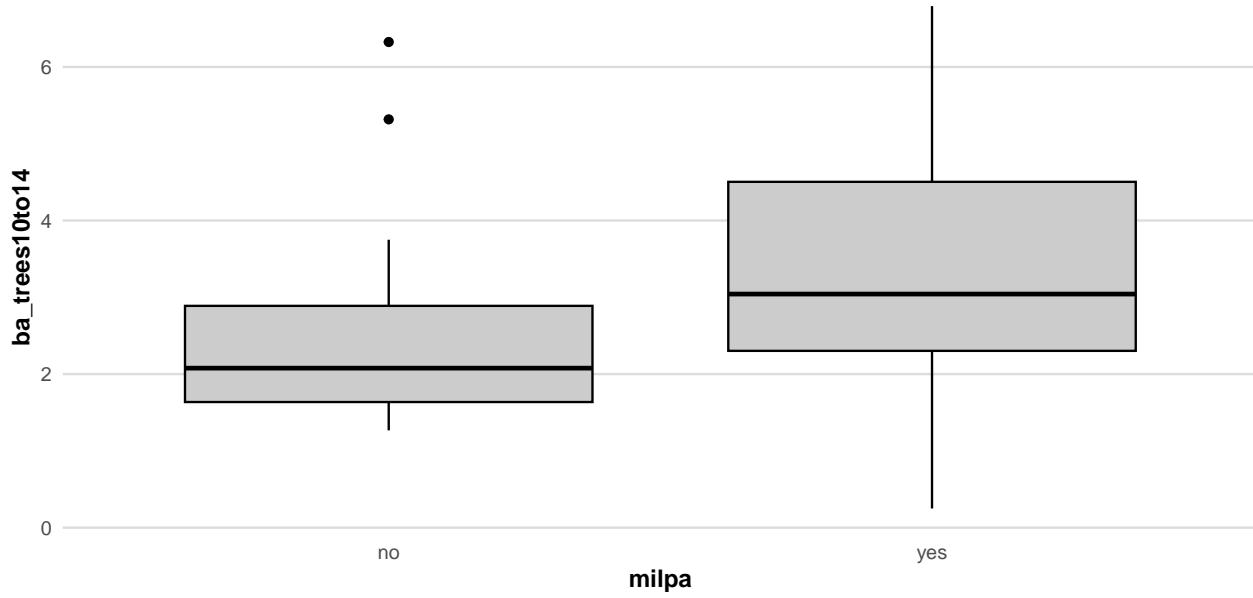
```
data_plots %>% ggplot() +
  geom_boxplot(aes(x = milpa, y = ba_saplings),
               col = "black", fill = "gray80") + plot_theme
```



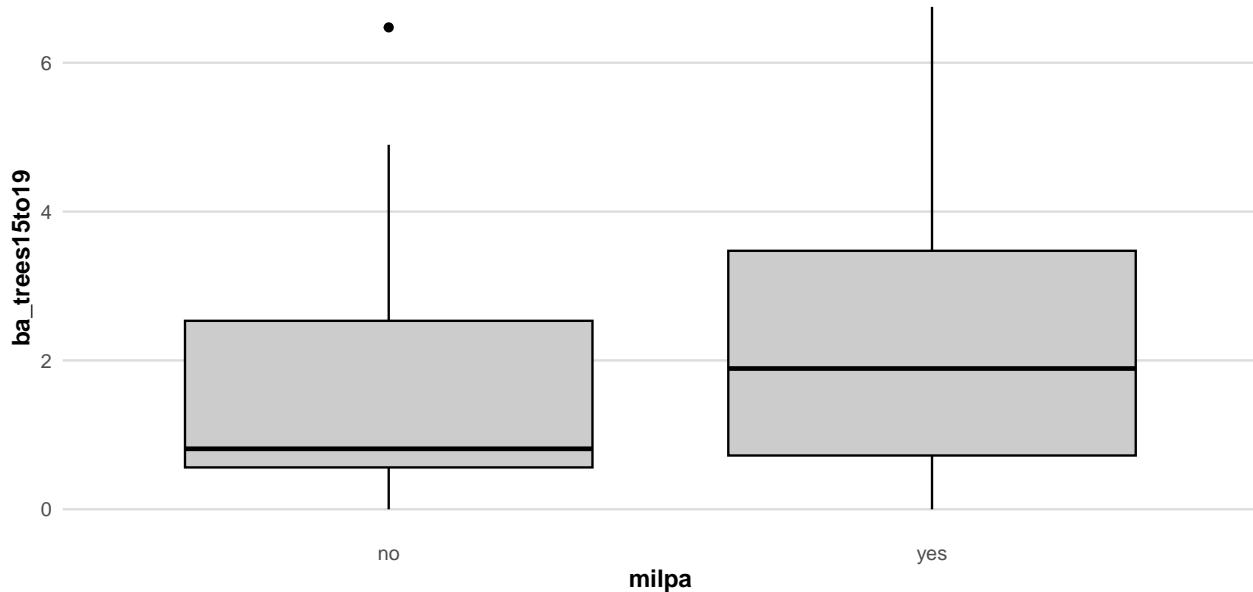
```
data_plots %>% ggplot() +
  geom_boxplot(aes(x = milpa, y = ba_trees05to09),
               col = "black", fill = "gray80") + plot_theme
```



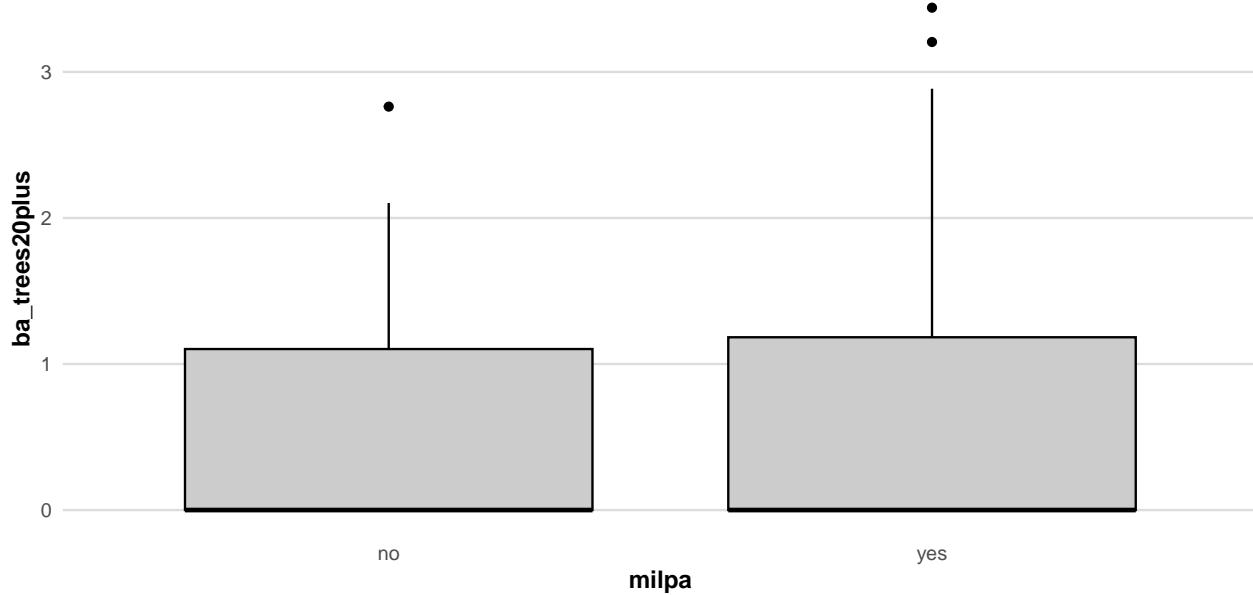
```
data_plots %>% ggplot() +
  geom_boxplot(aes(x = milpa, y = ba_trees10to14),
               col = "black", fill = "gray80") + plot_theme
```



```
data_plots %>% ggplot() +
  geom_boxplot(aes(x = milpa, y = ba_trees15to19),
               col = "black", fill = "gray80") + plot_theme
```

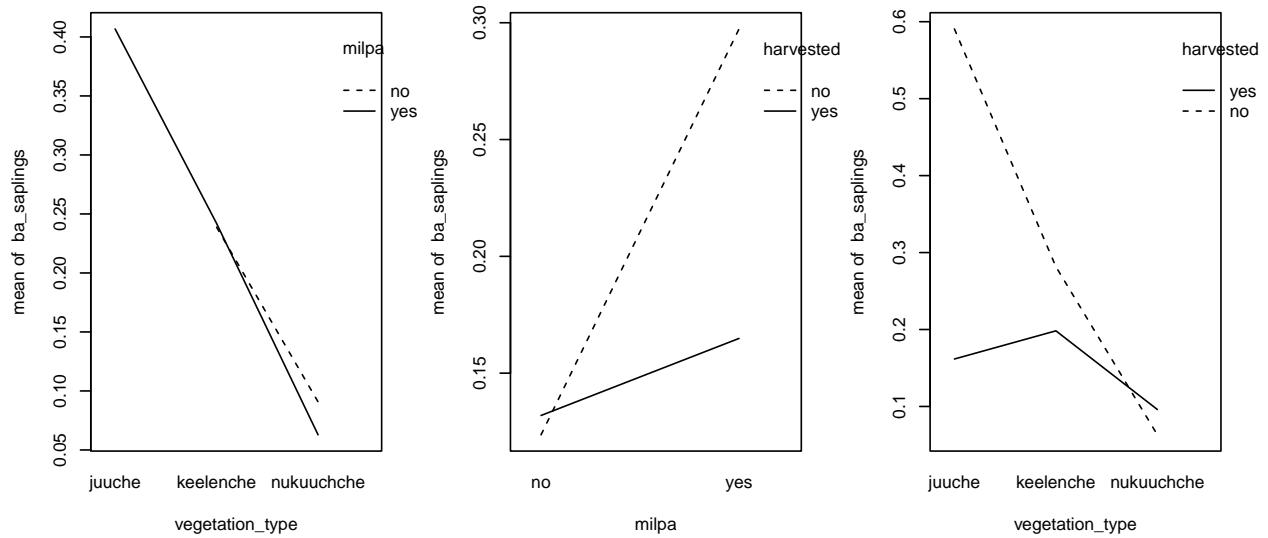


```
data_plots %>% ggplot() +
  geom_boxplot(aes(x = milpa, y = ba_trees20plus),
    col = "black", fill = "gray80") + plot_theme
```

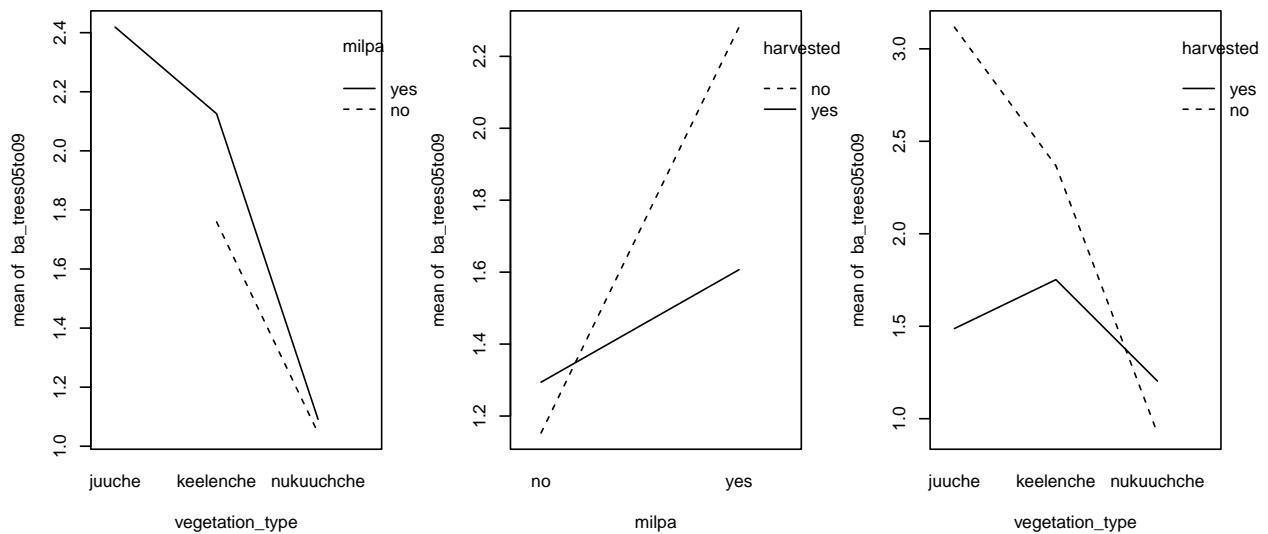


Interaction plots

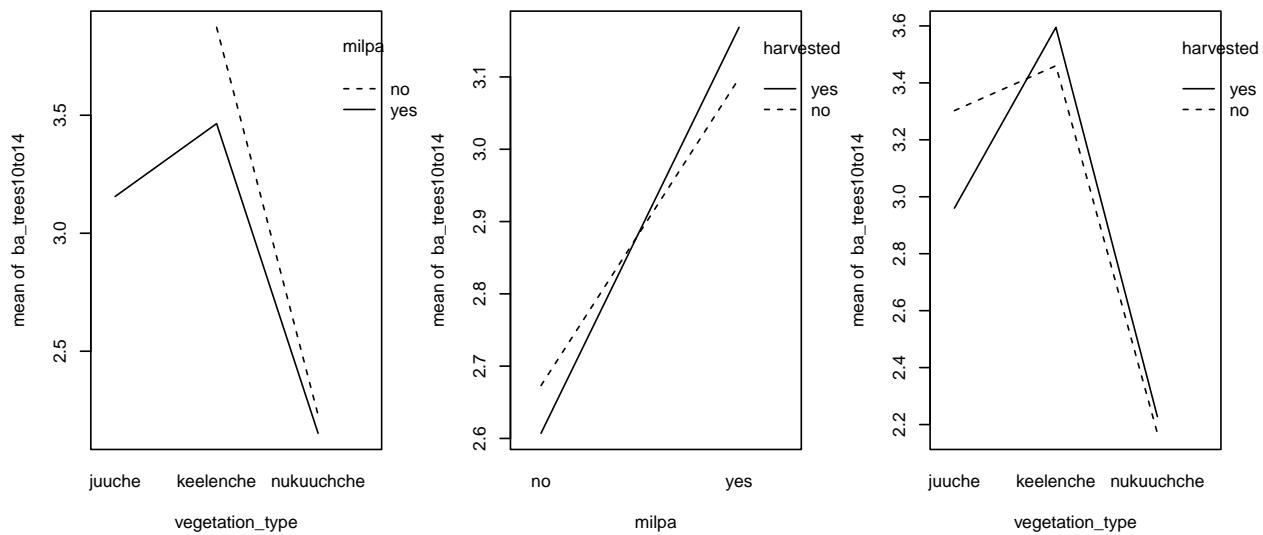
```
par(mfrow = c(1,3))
with(data_plots, {interaction.plot(vegetation_type, milpa, ba_saplings)
  interaction.plot(milpa, harvested, ba_saplings)
  interaction.plot(vegetation_type, harvested, ba_saplings)})
```



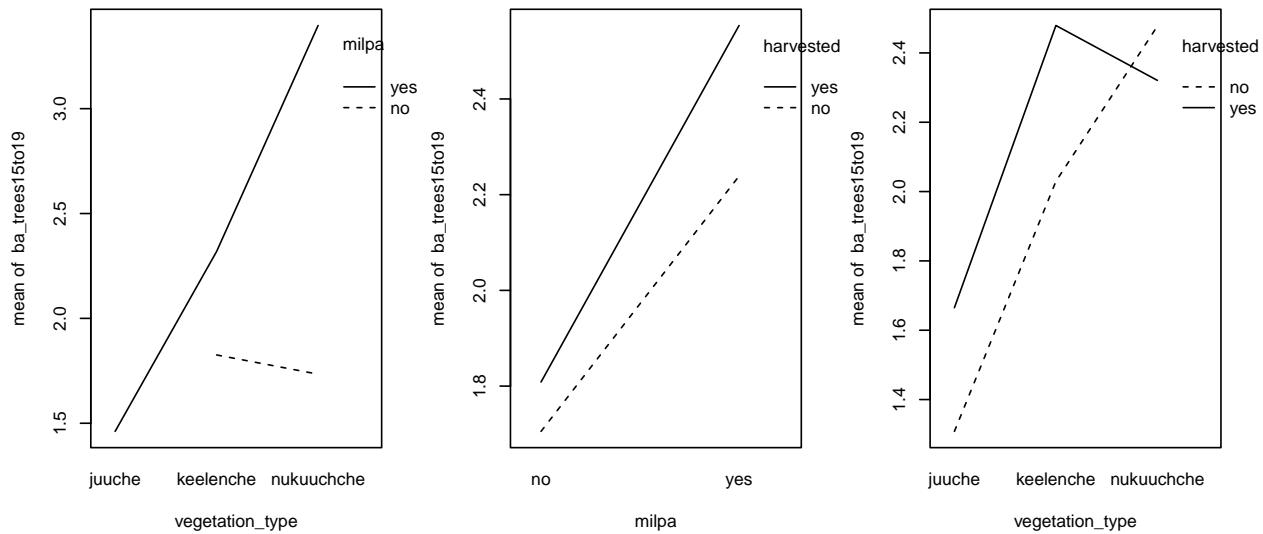
```
with(data_plots, {interaction.plot(vegetation_type, milpa, ba_trees05to09)
                  interaction.plot(milpa, harvested, ba_trees05to09)
                  interaction.plot(vegetation_type, harvested, ba_trees05to09)})
```



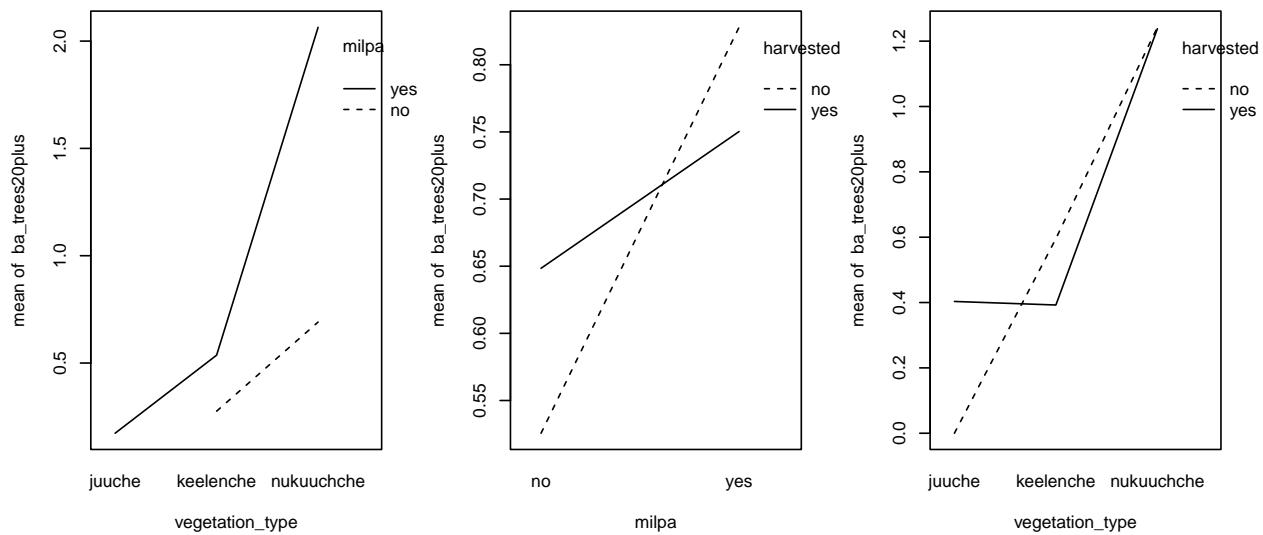
```
with(data_plots, {interaction.plot(vegetation_type, milpa, ba_trees10to14)
                  interaction.plot(milpa, harvested, ba_trees10to14)
                  interaction.plot(vegetation_type, harvested, ba_trees10to14)})
```



```
with(data_plots, {interaction.plot(vegetation_type, milpa, ba_trees15to19)
                 interaction.plot(milpa, harvested, ba_trees15to19)
                 interaction.plot(vegetation_type, harvested, ba_trees15to19)})
```



```
with(data_plots, {interaction.plot(vegetation_type, milpa, ba_trees20plus)
                 interaction.plot(milpa, harvested, ba_trees20plus)
                 interaction.plot(vegetation_type, harvested, ba_trees20plus)})
```



Models

Seedlings

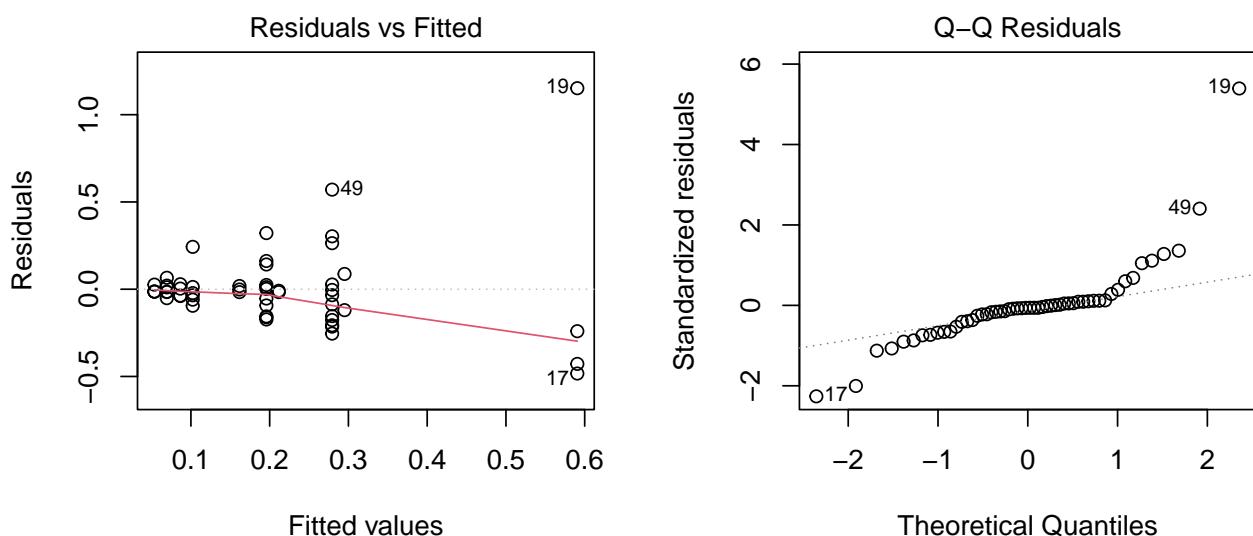
Seedlings have zero basal area by definition, so there is no need to fit a model for seedlings.

Saplings (0-4 cm DBH)

Fit model and check assumptions

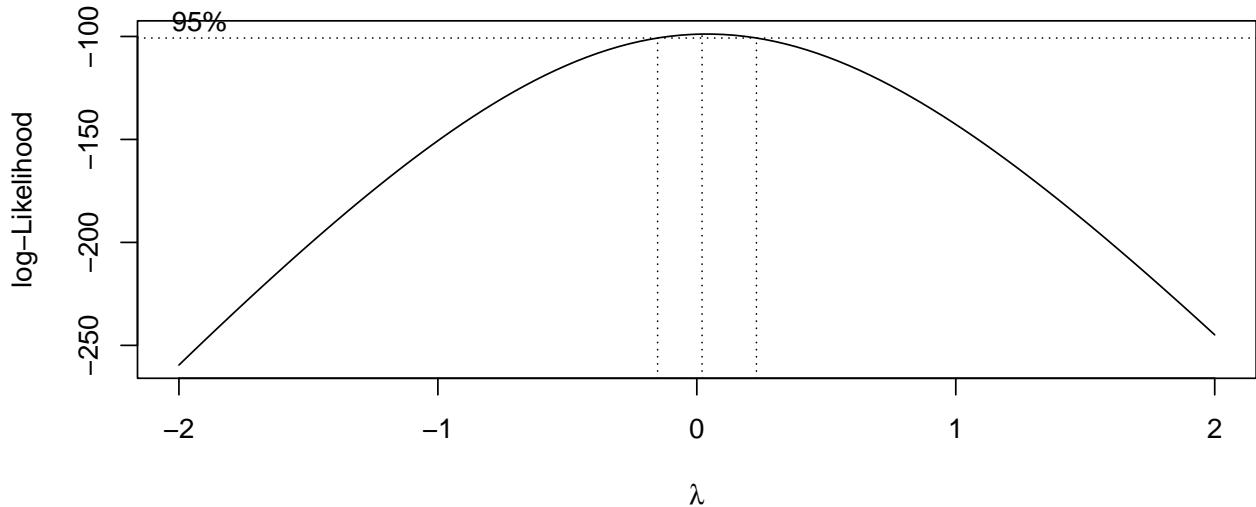
```
mod.saplings <- lm(ba_saplings ~ harvested + vegetation_type + milpa +
                     harvested:vegetation_type + harvested:vegetation_type, data = data_plots)

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



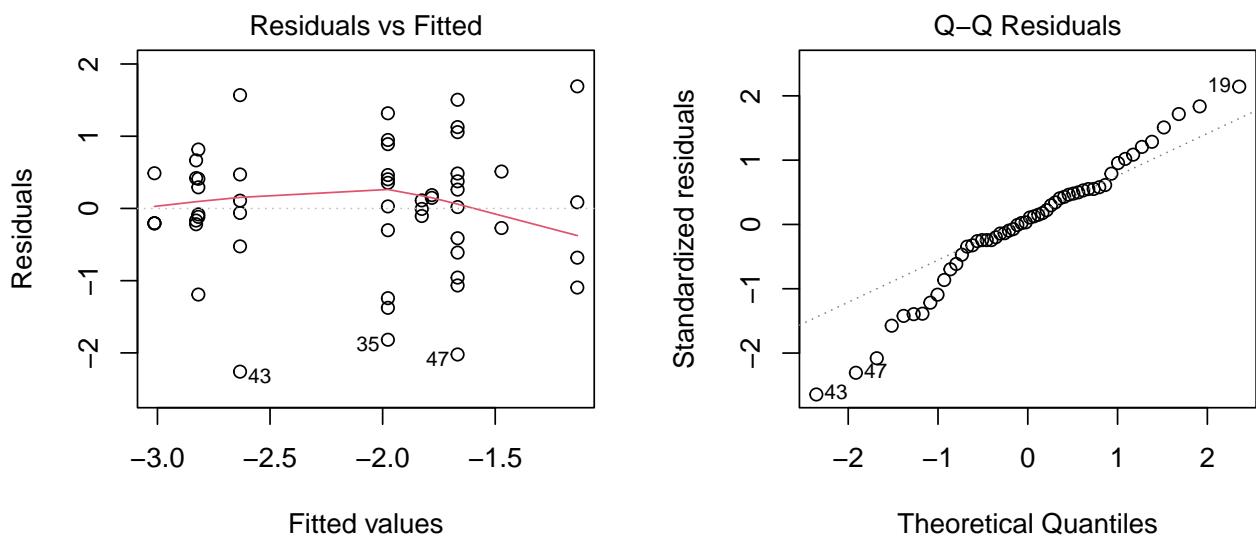
Attempt transformation of response

```
# Very distinct megaphone shape on the residuals vs. fitted; try Box-Cox
MASS::boxcox(mod.saplings)
```



```
# Box-Cox strongly suggests log transformation; re-fit model with logged response
mod.saplings.transform <- lm(log(ba_saplings) ~ harvested + vegetation_type + milpa +
                             harvested:vegetation_type + harvested:vegetation_type, data = data_pl)

# Check constant variance assumption again
par(mfrow = c(1,2))
plot(mod.saplings.transform, which = 1:2)
```



Assumptions look significantly better, so we will proceed with the transformed model.

Summarize selected model and run ANOVA

```
# Check for outliers
outlierTest(mod.saplings.transform)

## No Studentized residuals with Bonferroni p < 0.05
## Largest |rstudent|:
##      rstudent unadjusted p-value Bonferroni p
## 43 -2.835106          0.0067852     0.3664

# No significant outliers; look at model summary
summary(mod.saplings.transform)

##
## Call:
## lm(formula = log(ba_saplings) ~ harvested + vegetation_type +
##     milpa + harvested:vegetation_type + harvested:vegetation_type,
##     data = data_plots)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.25958 -0.29368  0.05692  0.46846  1.69078
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                -0.9404    0.5554 -1.693  0.09702 .
## harvestedyes               -0.6907    0.6956 -0.993  0.32583
## vegetation_typekeelenche  -0.5317    0.5183 -1.026  0.31027
## vegetation_typenukuuchche -1.8780    0.5716 -3.286  0.00193 **
## milpayes                  -0.1951    0.3179 -0.614  0.54245
## harvestedyes:vegetation_typekeelenche  0.3812    0.7790  0.489  0.62691
## harvestedyes:vegetation_typenukuuchche  0.8758    0.8061  1.087  0.28278
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9107 on 47 degrees of freedom
## Multiple R-squared:  0.311, Adjusted R-squared:  0.2231
## F-statistic: 3.536 on 6 and 47 DF,  p-value: 0.005699

# Construct ANOVA table (use type 2 SS since data are unbalanced)
Anova(mod.saplings.transform, type = 2)

## Anova Table (Type II tests)
##
## Response: log(ba_saplings)
##              Sum Sq Df F value    Pr(>F)
## harvested           0.408  1  0.4925  0.486290
## vegetation_type    13.305  2  8.0208 0.001007 **
## milpa              0.312  1  0.3765  0.542449
## harvested:vegetation_type 1.226  2  0.7392  0.482946
## Residuals          38.984 47
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We find that vegetation_type is significant, but harvested, milpa, and the interaction terms are not.

Pairwise comparisons

```
# Pairwise comparison for vegetation_type using Tukey HSD
# Nuku'uch che' has a significantly lower sapling BA than Ju'uche' and Keelenche'
TukeyHSD(aov(mod.saplings.transform), "vegetation_type", conv.level = 0.95)
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = mod.saplings.transform)
##
## $vegetation_type
##              diff      lwr      upr      p adj
## keelenche-juuche -0.3446386 -1.279480 0.5902032 0.6478545
## nukuuchche-juuche -1.3573744 -2.325312 -0.3894365 0.0039532
## nukuuchche-keelenche -1.0127357 -1.662988 -0.3624838 0.0013048
```

```
# Confirm lack of significance of harvested using Tukey HSD
TukeyHSD(aov(mod.saplings.transform), "harvested", conv.level = 0.95)
```

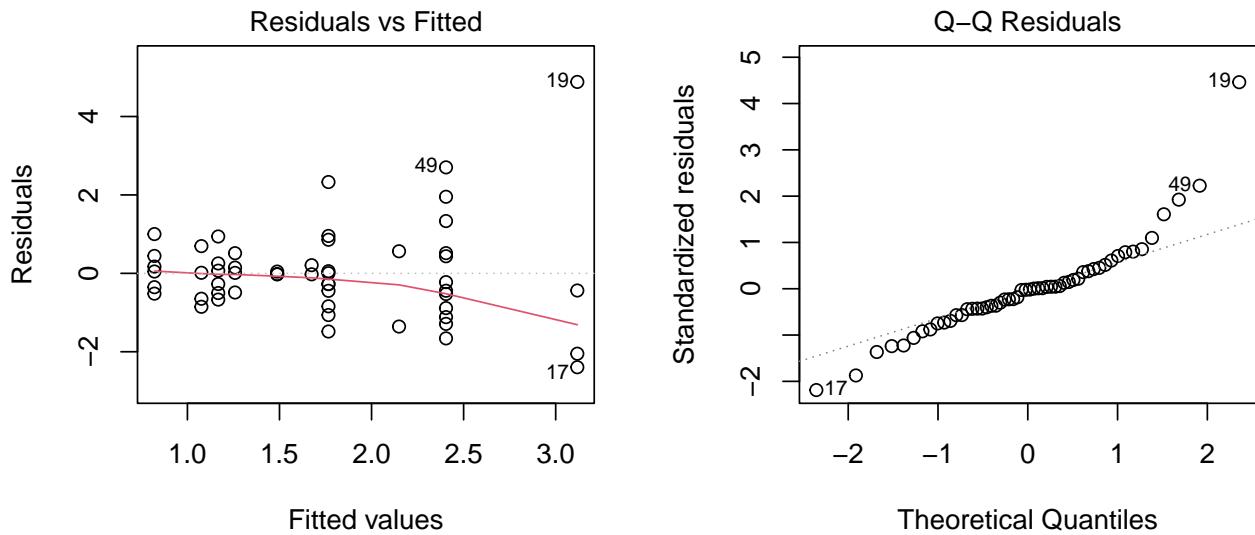
```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = mod.saplings.transform)
##
## $harvested
##              diff      lwr      upr      p adj
## yes-no -0.2105371 -0.7095314 0.2884571 0.4002954
```

Trees (5-9 cm DBH)

Fit model and check assumptions

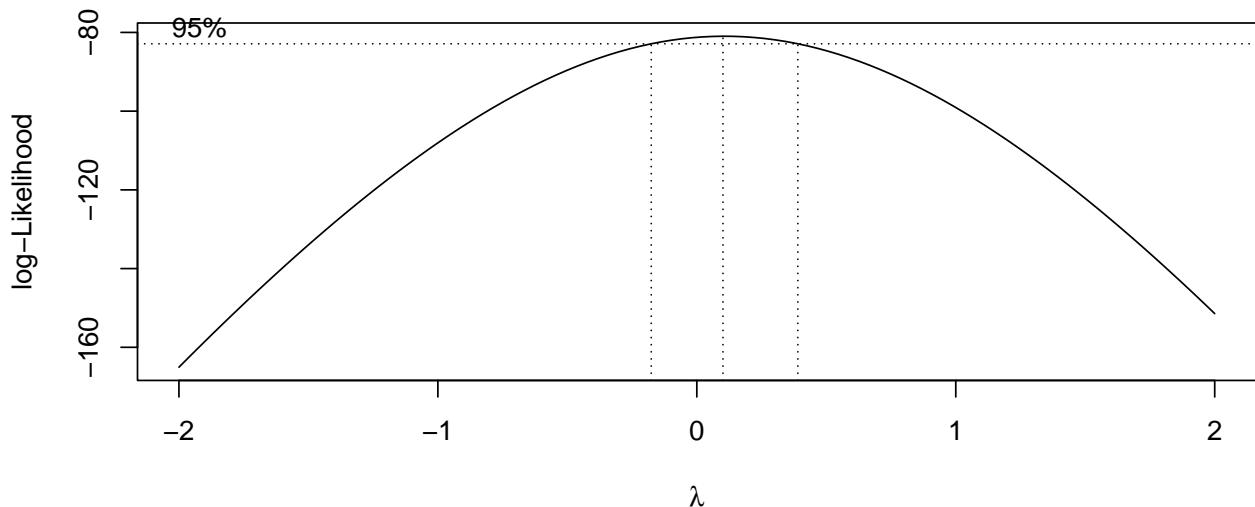
```
mod.05to09 <- lm(ba_trees05to09 ~ harvested + vegetation_type + milpa +
                     harvested:vegetation_type + harvested:milpa, data = data_plots)

par(mfrow = c(1,2))
plot(mod.05to09, which = 1:2)
```



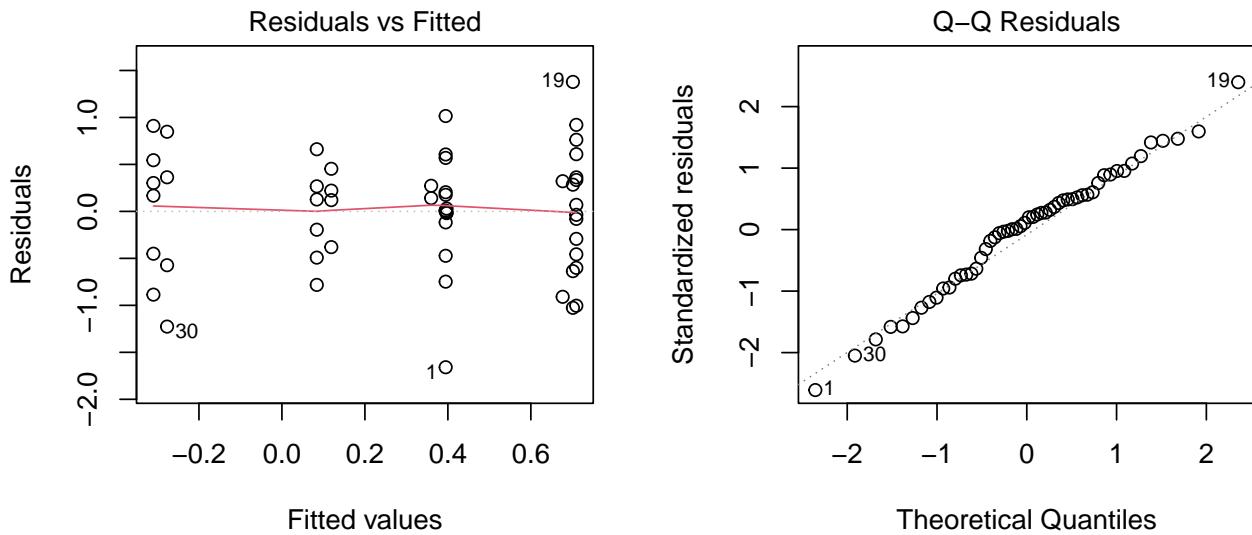
Attempt transformation of response

```
# Very distinct megaphone shape on the residuals vs. fitted; try Box-Cox
MASS:::boxcox(mod.05to09)
```



```
# Box-Cox strongly suggests log transformation; re-fit model with logged response
mod.05to09.transform <- lm(log(ba_trees05to09) ~ harvested + vegetation_type + milpa +
                           harvested:vegetation_type + harvested:milpa, data = data_plots)

# Check constant variance assumption again
par(mfrow = c(1,2))
plot(mod.05to09.transform, which = 1:2)
```



Assumptions look significantly better, so we will proceed with the transformed model.

Summarize selected model and run ANOVA

```
# Check for outliers
outlierTest(mod.05to09.transform)
```

```
## No Studentized residuals with Bonferroni p < 0.05
## Largest |rstudent|:
##      rstUDENT unadjusted p-value Bonferroni p
## 1 -2.795767          0.0075867     0.40968
```

```
# No significant outliers; look at model summary
summary(mod.05to09.transform)
```

```
##
## Call:
## lm(formula = log(ba_trees05to09) ~ harvested + vegetation_type +
##     milpa + harvested:vegetation_type + harvested:milpa, data = data_plots)
##
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -1.65878 -0.45530  0.09487  0.35507  1.37806 
##
## Coefficients:
## (Intercept)            Estimate Std. Error t value Pr(>|t|)    
## harvestedyes           0.668496  0.466107  1.434   0.1583  
## vegetation_typekeelenche 0.008184  0.379233  0.022   0.9829  
## vegetation_typenkuuchche -0.978240  0.439063 -2.228   0.0308 *  
## milpayes                0.032889  0.327258  0.100   0.9204  
## harvestedyes:vegetation_typekeelenche -0.010349  0.571960 -0.018   0.9856  
## harvestedyes:vegetation_typenkuuchche  0.700400  0.649984  1.078   0.2868
```

```

## harvestedyes:milpayes          0.002112   0.463434   0.005   0.9964
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6638 on 46 degrees of freedom
## Multiple R-squared:  0.2577, Adjusted R-squared:  0.1448
## F-statistic: 2.282 on 7 and 46 DF,  p-value: 0.04412

# Construct ANOVA table (use type 2 SS since data are unbalanced)
Anova(mod.05to09.transform, type = 2)

```

```

## Anova Table (Type II tests)
##
## Response: log(ba_trees05to09)
##                               Sum Sq Df F value Pr(>F)
## harvested                  0.0337  1  0.0765 0.78338
## vegetation_type            3.7415  2  4.2456 0.02032 *
## milpa                      0.0095  1  0.0215 0.88418
## harvested:vegetation_type  1.1703  2  1.3280 0.27498
## harvested:milpa             0.0000  1  0.0000 0.99638
## Residuals                  20.2691 46
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

We find that vegetation_type is significant, but harvested, milpa, and the interaction terms are not.

Pairwise comparisons

```

# Pairwise comparison for vegetation_type using Tukey HSD
# Keelenche' has a significantly higher 5-9 cm tree BA than Nuku'uch che'
TukeyHSD(aov(mod.05to09.transform), "vegetation_type", conf.level = 0.95)

```

```

## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = mod.05to09.transform)
##
## $vegetation_type
##               diff      lwr      upr      p adj
## keelenche-juuche -0.01431319 -0.6961652 0.66753878 0.9985761
## nukuuchche-juuche -0.66526202 -1.3712534 0.04072937 0.0684727
## nukuuchche-keelenche -0.65094883 -1.1252275 -0.17667017 0.0048831

```

```

# Confirm lack of significance of harvested using Tukey HSD
TukeyHSD(aov(mod.05to09.transform), "harvested", conf.level = 0.95)

```

```

## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = mod.05to09.transform)

```

```

## 
## $harvested
##           diff      lwr      upr     p adj
## yes-no -0.0685562 -0.4324633 0.2953509 0.7062783

```

Trees (10-14 cm DBH)

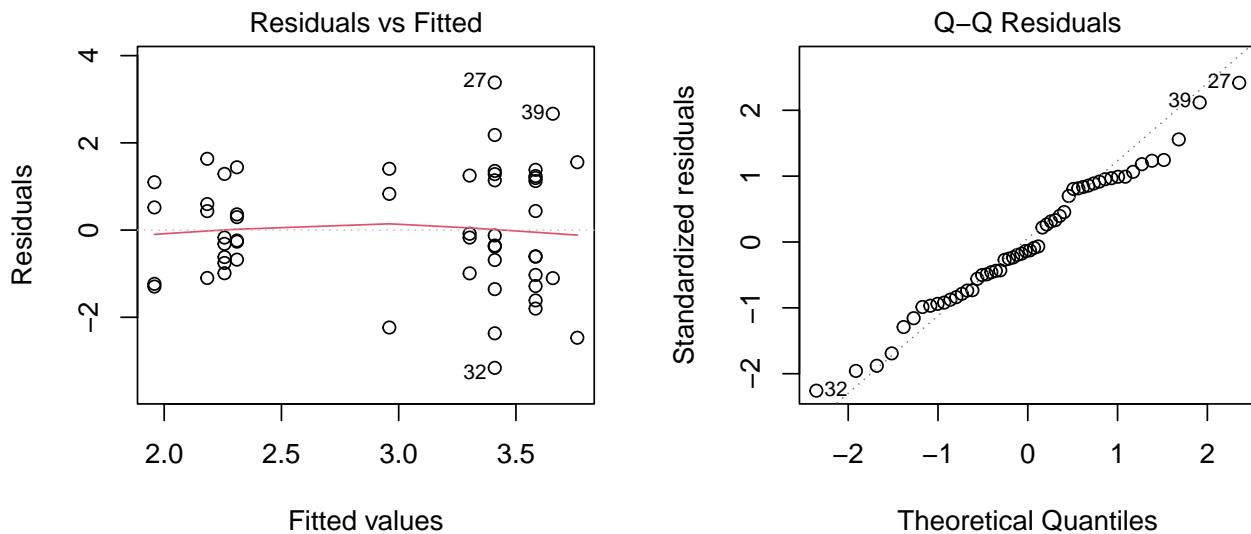
Fit model and check assumptions

```

mod.10to14 <- lm(ba_trees10to14 ~ harvested + vegetation_type + milpa +
                     harvested:vegetation_type + harvested:milpa, data = data_plots)

par(mfrow = c(1,2))
plot(mod.10to14, which = 1:2)

```



Assumptions appear to be satisfied.

Summarize selected model and run ANOVA

```

# Check for outliers
outlierTest(mod.10to14)

## No Studentized residuals with Bonferroni p < 0.05
## Largest |rstudent|:
##   rstudent unadjusted p-value Bonferroni p
## 27 2.558604       0.013944      0.75299

# No significant outliers; look at model summary
summary(mod.10to14)

```

```

## 
## Call:
## lm(formula = ba_trees10to14 ~ harvested + vegetation_type + milpa +
##     harvested:vegetation_type + harvested:milpa, data = data_plots)
##
## Residuals:
##    Min      1Q  Median      3Q      Max 
## -3.1597 -0.9918 -0.1723  1.1804  3.3828 
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)    
## (Intercept)                 3.6544    1.0226   3.574  0.00084 ***
## harvestedyes                -0.6209    1.5069  -0.412  0.68223    
## vegetation_typekeelenche    0.1072    0.8320   0.129  0.89802    
## vegetation_typenukuuchche   -1.3442    0.9632  -1.395  0.16958    
## milpayes                     -0.3519    0.7180  -0.490  0.62632    
## harvestedyes:vegetation_typekeelenche  0.5165    1.2548   0.412  0.68255    
## harvestedyes:vegetation_typenukuuchche  0.5678    1.4260   0.398  0.69234    
## harvestedyes:milpayes            0.2784    1.0167   0.274  0.78545    
## ---                        
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.456 on 46 degrees of freedom
## Multiple R-squared:  0.1793, Adjusted R-squared:  0.05445 
## F-statistic: 1.436 on 7 and 46 DF,  p-value: 0.2143

# Construct ANOVA table (use type 2 SS since data are unbalanced)
Anova(mod.10to14, type = 2)

```

```

## Anova Table (Type II tests)
## 
## Response: ba_trees10to14
##                               Sum Sq Df F value Pr(>F)    
## harvested                      0.027  1  0.0125 0.91132    
## vegetation_type                  18.136  2  4.2757 0.01981 *
## milpa                          0.373  1  0.1758 0.67699    
## harvested:vegetation_type       0.400  2  0.0942 0.91024    
## harvested:milpa                 0.159  1  0.0750 0.78545    
## Residuals                      97.555 46
## ---                        
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

We find that vegetation type is significant, but harvested, milpa, and the interaction terms are not.

Pairwise comparisons

```

# Pairwise comparison for vegetation_type using Tukey HSD
# Keelenche' has a significantly higher 10-14 cm tree BA than Nuku'uch che'
TukeyHSD(aov(mod.10to14), "vegetation_type", conf.level = 0.95)

```

```

## Tukey multiple comparisons of means

```

```

##      95% family-wise confidence level
##
## Fit: aov(formula = mod.10to14)
##
## $vegetation_type
##              diff      lwr      upr     p adj
## keelenche-juuche    0.3682668 -1.127618 1.8641518 0.8228298
## nukuuchche-juuche   -0.9584649 -2.507308 0.5903786 0.3009605
## nukuuchche-keelenche -1.3267316 -2.367231 -0.2862325 0.0093790

```

```

# Confirm lack of significance of harvested using Tukey HSD
TukeyHSD(aov(mod.10to14), "harvested", conf.level = 0.95)

```

```

## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = mod.10to14)
##
## $harvested
##              diff      lwr      upr     p adj
## yes-no 0.01925824 -0.7791016 0.8176181 0.9614837

```

Trees (15-19 cm DBH)

Fit model and check assumptions

```

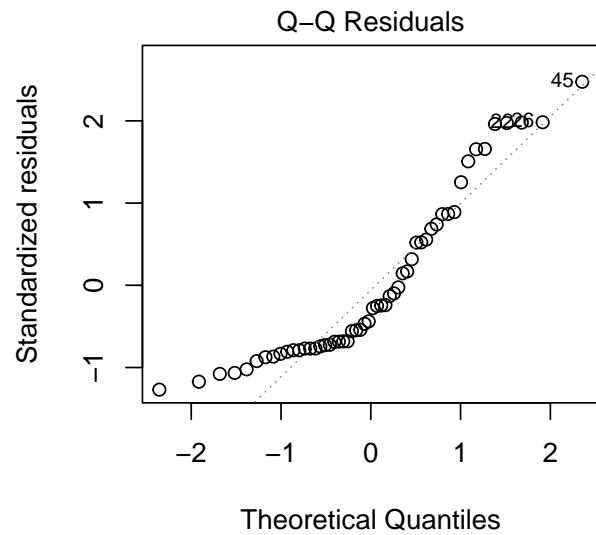
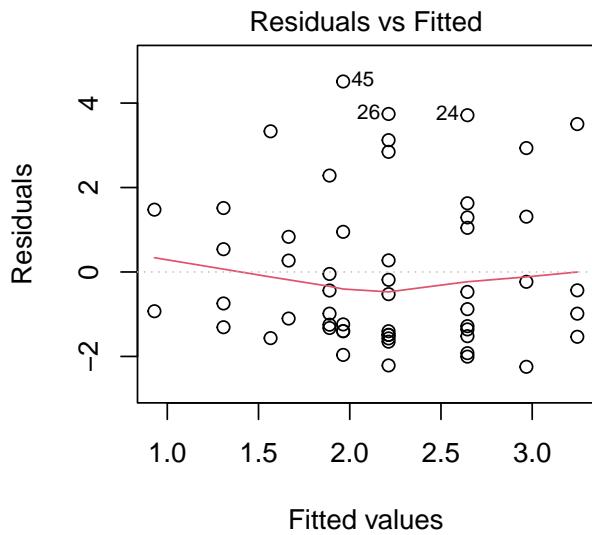
mod.15to19 <- lm(ba_trees15to19 ~ harvested + vegetation_type + milpa +
                     harvested:vegetation_type + harvested:milpa, data = data_plots)

```

```

par(mfrow = c(1,2))
plot(mod.15to19, which = 1:2)

```



Assumptions appear to be satisfied.

Summarize selected model and run ANOVA

```

# Check for outliers
outlierTest(mod.15to19)

## No Studentized residuals with Bonferroni p < 0.05
## Largest |rstudent|:
##   rstudent unadjusted p-value Bonferroni p
## 45 2.630107      0.011643      0.62873

# No significant outliers; look at model summary
summary(mod.15to19)

## 
## Call:
## lm(formula = ba_trees15to19 ~ harvested + vegetation_type + milpa +
##     harvested:vegetation_type + harvested:milpa, data = data_plots)
##
## Residuals:
##    Min     1Q Median     3Q    Max 
## -2.2456 -1.4014 -0.6355  1.2306  4.5111 
## 
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)    
## (Intercept)                 0.02481   1.37881   0.018   0.986    
## harvestedyes                0.56132   2.03188   0.276   0.784    
## vegetation_typekeelenche   0.90467   1.12182   0.806   0.424    
## vegetation_typenkuuchche   1.93911   1.29881   1.493   0.142    
## milpayes                     1.28332   0.96807   1.326   0.192    
## harvestedyes:vegetation_typekeelenche  0.07535   1.69193   0.045   0.965    
## harvestedyes:vegetation_typenkuuchche -0.63604   1.92274  -0.331   0.742    
## harvestedyes:milpayes        -0.20444   1.37090  -0.149   0.882    
## 
## Residual standard error: 1.964 on 46 degrees of freedom
## Multiple R-squared:  0.09173,   Adjusted R-squared:  -0.04648 
## F-statistic: 0.6637 on 7 and 46 DF,  p-value: 0.7013

# Construct ANOVA table (use type 2 SS since data are unbalanced)
Anova(mod.15to19, type = 2)

## Anova Table (Type II tests)
## 
## Response: ba_trees15to19
##              Sum Sq Df F value    Pr(>F)    
## harvested            0.645  1 0.1674  0.68433  
## vegetation_type      11.303  2 1.4658  0.24148  
## milpa                11.453  1 2.9705  0.09151 .  
## harvested:vegetation_type  1.147  2 0.1487  0.86224  
## harvested:milpa       0.086  1 0.0222  0.88210  
## Residuals           177.365 46                        
## ---                
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

We find that no predictors are significant.

Pairwise comparisons

```
# Confirm lack of significance of harvested using Tukey HSD
TukeyHSD(aov(mod.15to19), "harvested", conf.level = 0.95)
```

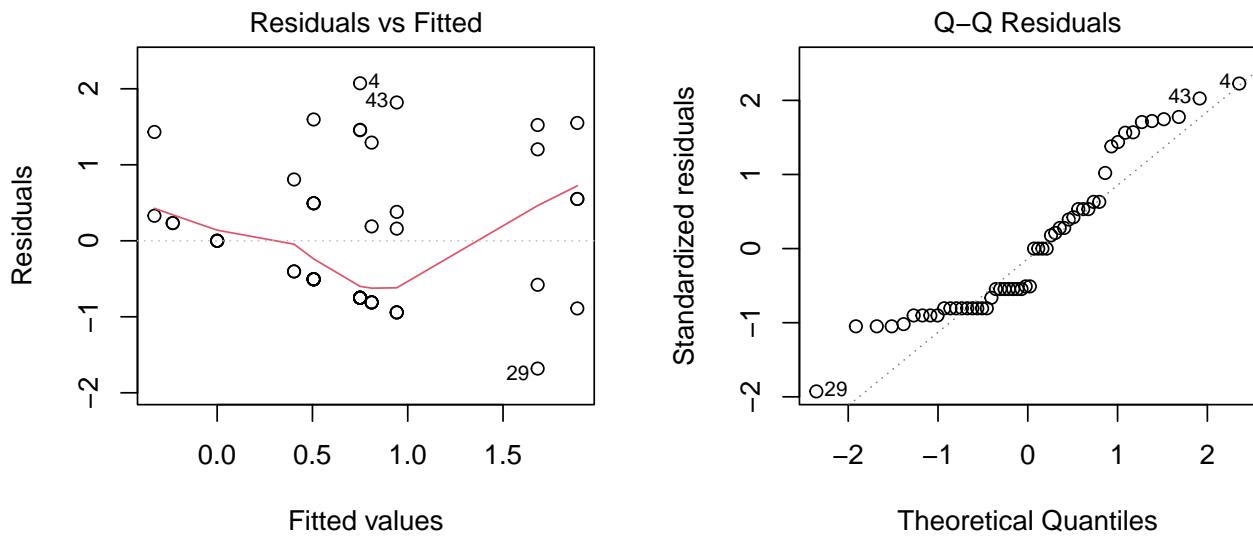
```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = mod.15to19)
##
## $harvested
##          diff      lwr      upr   p adj
## yes-no 0.2378915 -0.8385922 1.314375 0.6585297
```

Trees (20+ cm DBH)

Fit model and check assumptions

```
mod.20plus <- lm(ba_trees20plus ~ harvested + vegetation_type + milpa +
                     harvested:vegetation_type + harvested:milpa, data = data_plots)

par(mfrow = c(1,2))
plot(mod.20plus, which = 1:2)
```



There is a slight megaphone shape on the residuals vs. fitted. We can't do a log transformation since the response has zeros, so we'll try a square root transformation.

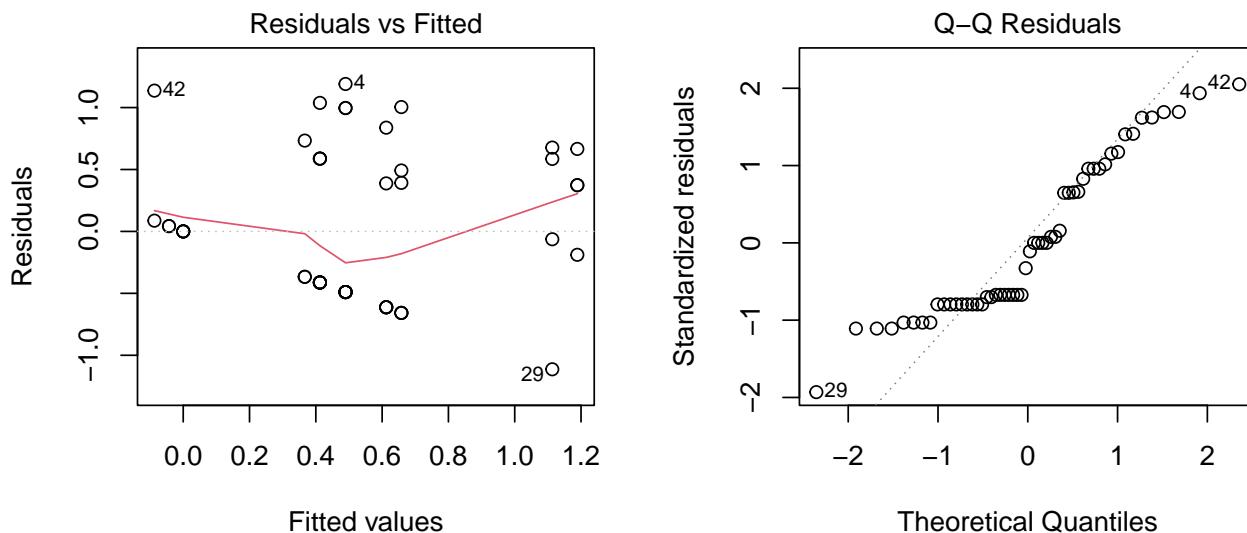
Attempt transformation of response

```

mod.20plus.transform <- lm(sqrt(ba_trees20plus) ~ harvested + vegetation_type + milpa +
                           harvested:vegetation_type + harvested:milpa, data = data_plots)

# Check constant variance assumption again
par(mfrow = c(1,2))
plot(mod.20plus.transform, which = 1:2)

```



No substantial improvement in assumptions, so we will proceed with the untransformed model.

Summarize selected model and run ANOVA

```
# Check for outliers
outlierTest(mod.20plus)

## No Studentized residuals with Bonferroni p < 0.05
## Largest |rstudent|:
##   rstudent unadjusted p-value Bonferroni p
## 4  2.332063           0.024232          NA

# No significant outliers; look at model summary
summary(mod.20plus)
```

```
##  
## Call:  
## lm(formula = ba_trees20plus ~ harvested + vegetation_type + milpa +  
##       harvested:vegetation_type + harvested:milpa, data = data_plots)  
##  
## Residuals:  
##      Min       1Q   Median       3Q      Max  
## -1.6816 -0.7499 -0.4033  0.4938  2.0726  
##  
## Coefficients:
```

```

##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  -1.0791    0.6800 -1.587  0.11935
## harvestedyes                   0.7432    1.0020  0.742  0.46206
## vegetation_typekeelenche      0.7499    0.5532  1.355  0.18189
## vegetation_typenukuuchche     1.8897    0.6405  2.950  0.00498 **
## milpayes                      1.0791    0.4774  2.260  0.02857 *
## harvestedyes:vegetation_typekeelenche -0.6470    0.8344 -0.775  0.44208
## harvestedyes:vegetation_typenukuuchche -0.6115    0.9482 -0.645  0.52220
## harvestedyes:milpayes          -0.3398    0.6761 -0.503  0.61760
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9684 on 46 degrees of freedom
## Multiple R-squared:  0.2757, Adjusted R-squared:  0.1655
## F-statistic: 2.501 on 7 and 46 DF,  p-value: 0.02895

# Construct ANOVA table (use type 2 SS since data are unbalanced)
Anova(mod.20plus, type = 2)

```

```

## Anova Table (Type II tests)
##
## Response: ba_trees20plus
##                               Sum Sq Df F value    Pr(>F)
## harvested                     0.030  1  0.0317 0.8595325
## vegetation_type                15.258  2  8.1355 0.0009438 ***
## milpa                          6.791  1  7.2422 0.0098941 **
## harvested:vegetation_type     0.574  2  0.3059 0.7379379
## harvested:milpa                0.237  1  0.2527 0.6175960
## Residuals                      43.135 46
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

We find that vegetation type and milpa are significant, but harvested and the interaction terms are not.

Pairwise comparisons

```

# Pairwise comparison for vegetation_type using Tukey HSD
# Nuku'uch che' has significantly higher 20+ cm tree BA than ju'uche' and keelenche';
# no significant difference between the latter two
TukeyHSD(aov(mod.20plus), "vegetation_type", conf.level = 0.95)

```

```

## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = mod.20plus)
##
## $vegetation_type
##               diff      lwr      upr   p adj
## keelenche-juuche  0.3262077 -0.66848589 1.320901 0.7084233
## nukuuchche-juuche 1.0688804  0.03897193 2.098789 0.0403830
## nukuuchche-keelenche 0.7426727  0.05078945 1.434556 0.0328677

```

```
# Pairwise comparison for milpa using Tukey HSD
# Plots that have been milpa have significantly higher 20+ cm tree BA
TukeyHSD(aov(mod.20plus), "milpa", conf.level = 0.95)
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = mod.20plus)
##
## $milpa
##      diff      lwr      upr      p adj
## yes-no 0.6640428 0.08313967 1.244946 0.025975
```

```
# Confirm lack of significance of harvested using Tukey HSD
TukeyHSD(aov(mod.20plus), "harvested", conf.level = 0.95)
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = mod.20plus)
##
## $harvested
##      diff      lwr      upr      p adj
## yes-no -0.02257555 -0.5534475 0.5082964 0.9321565
```