Final models

Tim White

June 2024

Packages, data, and functions

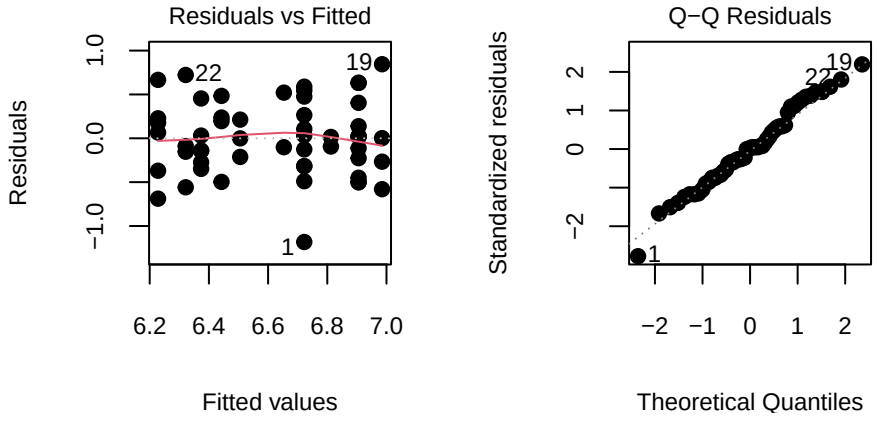
```
library(car)
library(tidyverse)
library(knitr)

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))

summarize_lm <- function(mod) {
  summary <- summary(mod)
  table <- tibble(var = c("Intercept", "harvestedYes", "vegetationKeelenche",
                          "vegetationNukuuchche", "milpaYes", "harvestedYes:vegetationKeelenche",
                          "harvestedYes:vegetationNukuuchche", "harvestedYes:milpaYes"),
                  coef = summary$coefficients[,1], se = summary$coefficients[,2],
                  t = summary$coefficients[,3], p = summary$coefficients[,4]) %>%
              mutate_if(is.numeric, function(col) {round(col, 2)}) %>%
              kable(col.names = c("", "Coefficient", "SE", "t-statistic", "p-value"), align = "c")
  return(table)
}

summarize_anova <- function(mod) {
  summary <- Anova(mod)
  table <- tibble(var = c("harvested", "vegetation", "milpa", "harvested:vegetation",
                          "harvested:milpa", "residuals"),
                  ss = summary$`Sum Sq`, df = summary$Df,
                  f = summary$`F value`, p = summary$`Pr(>F)`) %>%
              mutate_if(is.numeric, function(col) {round(col, 2)}) %>%
              kable(col.names = c("", "SS", "df", "F-statistic", "p-value"), align = "c")
  return(table)
}
```

Stem density

Model

$$\log(\text{total trees per ha}) \sim \beta_0 + \beta_1 \cdot \text{harvestedYes} + \beta_2 \cdot \text{vegetationKeelenche} + \beta_3 \cdot \text{vegetationNukuuchche} + \beta_4 \cdot \text{milpaYes} + \beta_5 \cdot \text{harvestedYes:vegetationKeelenche} + \beta_6 \cdot \text{harvestedYes:vegetationNukuuchche} + \beta_7 \cdot \text{harvestedYes:milpaYes} + \varepsilon$$

Diagnostic plots



Linear regression output

This model has an R^2 of 0.278.

	Coefficient	SE	t-statistic	p-value
Intercept	6.89	0.31	22.08	0.00
harvestedYes	-0.45	0.46	-0.99	0.33
vegetationKeelenche	-0.08	0.25	-0.31	0.76
vegetationNukuuchche	-0.66	0.29	-2.26	0.03
milpaYes	0.09	0.22	0.43	0.67
harvestedYes:vegetationKeelenche	0.30	0.38	0.77	0.44
harvestedYes:vegetationNukuuchche	0.60	0.44	1.38	0.17
harvestedYes:milpaYes	-0.02	0.31	-0.08	0.94

ANOVA output

	SS	df	F-statistic	p-value
harvested	0.14	1	0.69	0.41
vegetation	1.69	2	4.28	0.02
milpa	0.05	1	0.27	0.60
harvested:vegetation	0.40	2	1.02	0.37
harvested:milpa	0.00	1	0.01	0.94
residuals	9.09	46		

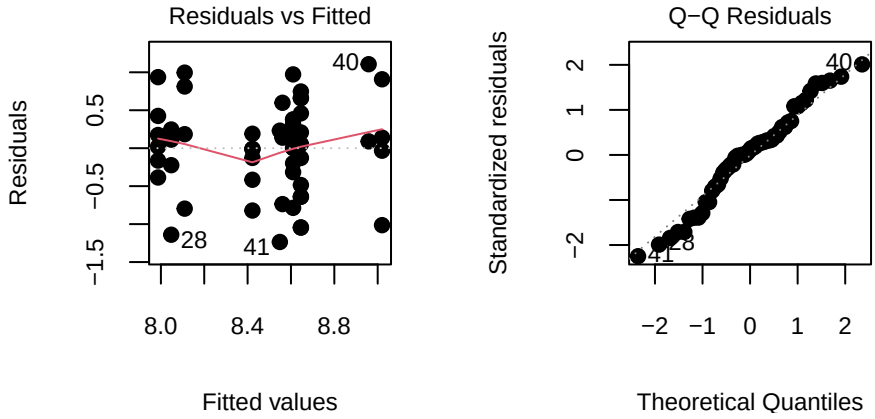
Stem density by size class

Seedlings

Model

$$\log(\text{seedlings per ha}) \sim \beta_0 + \beta_1 \cdot \text{harvestedYes} + \beta_2 \cdot \text{vegetationKeelenche} + \beta_3 \cdot \text{vegetationNukuuchche} + \beta_4 \cdot \text{milpaYes} + \beta_5 \cdot \text{harvestedYes:vegetationKeelenche} + \beta_6 \cdot \text{harvestedYes:vegetationNukuuchche} + \beta_7 \cdot \text{harvestedYes:milpaYes} + \varepsilon$$

Diagnostic plots



Linear regression output

This model has an R^2 of 0.21.

	Coefficient	SE	t-statistic	p-value
Intercept	9.33	0.45	20.92	0.00
harvestedYes	-0.83	0.66	-1.27	0.21
vegetationKeelenche	-0.37	0.36	-1.03	0.31
vegetationNukuuchche	-0.91	0.42	-2.17	0.04
milpaYes	-0.31	0.31	-1.00	0.32
harvestedYes:vegetationKeelenche	0.42	0.55	0.77	0.44
harvestedYes:vegetationNukuuchche	0.40	0.62	0.64	0.53
harvestedYes:milpaYes	0.37	0.44	0.84	0.40

ANOVA output

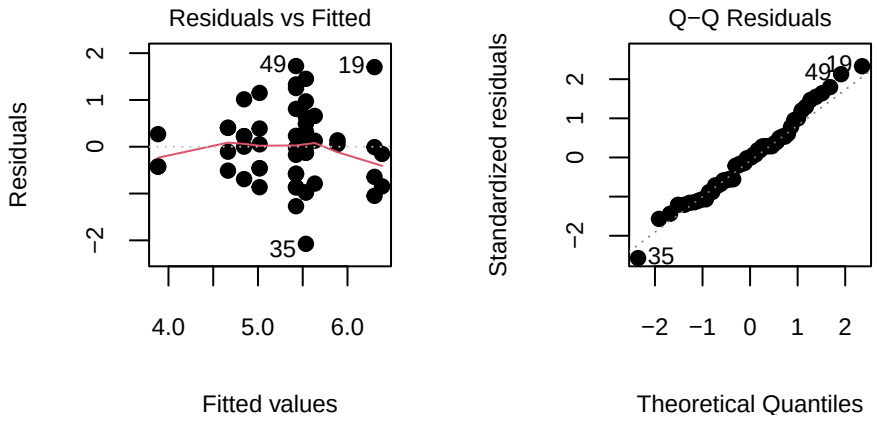
	SS	df	F-statistic	p-value
harvested	0.60	1	1.48	0.23
vegetation	3.30	2	4.09	0.02
milpa	0.13	1	0.32	0.57
harvested:vegetation	0.25	2	0.30	0.74
harvested:milpa	0.29	1	0.71	0.40
residuals	18.57	46		

Saplings (0-4 cm DBH)

Model

$$\log(\text{saplings per ha}) \sim \beta_0 + \beta_1 \cdot \text{harvestedYes} + \beta_2 \cdot \text{vegetationKeelenche} + \beta_3 \cdot \text{vegetationNukuuchche} + \beta_4 \cdot \text{milpaYes} + \beta_5 \cdot \text{harvestedYes:vegetationKeelenche} + \beta_6 \cdot \text{harvestedYes:vegetationNukuuchche} + \beta_7 \cdot \text{harvestedYes:milpaYes} + \varepsilon$$

Diagnostic plots



Linear regression output

This model has an R^2 of 0.374.

	Coefficient	SE	t-statistic	p-value
Intercept	7.26	0.59	12.27	0.00
harvestedYes	-1.28	0.87	-1.46	0.15
vegetationKeelenche	-0.88	0.48	-1.82	0.08
vegetationNukuuchche	-2.42	0.56	-4.34	0.00
milpaYes	-0.96	0.42	-2.31	0.03
harvestedYes:vegetationKeelenche	0.78	0.73	1.07	0.29
harvestedYes:vegetationNukuuchche	1.45	0.83	1.75	0.09
harvestedYes:milpaYes	0.61	0.59	1.03	0.31

ANOVA output

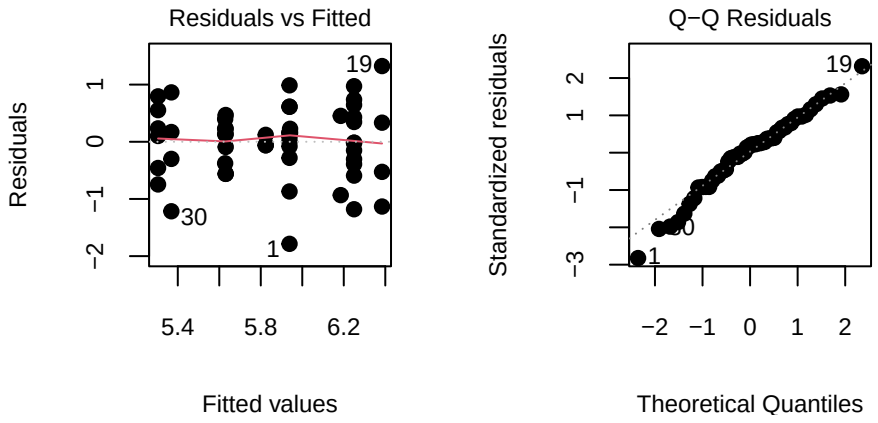
	SS	df	F-statistic	p-value
harvested	0.09	1	0.12	0.73
vegetation	17.15	2	12.07	0.00
milpa	3.55	1	5.00	0.03
harvested:vegetation	2.27	2	1.60	0.21
harvested:milpa	0.76	1	1.07	0.31
residuals	32.69	46		

Trees (5-9 cm DBH)

Model

$$\log(\text{5-9 cm trees per ha}) \sim \beta_0 + \beta_1 \cdot \text{harvestedYes} + \beta_2 \cdot \text{vegetationKeelenche} + \beta_3 \cdot \text{vegetationNukuuchche} + \beta_4 \cdot \text{milpaYes} + \beta_5 \cdot \text{harvestedYes:vegetationKeelenche} + \beta_6 \cdot \text{harvestedYes:vegetationNukuuchche} + \beta_7 \cdot \text{harvestedYes:milpaYes} + \varepsilon$$

Diagnostic plots



Linear regression output

This model has an R^2 of 0.246.

	Coefficient	SE	t-statistic	p-value
Intercept	6.32	0.46	13.64	0.00
harvestedYes	-0.49	0.68	-0.72	0.47
vegetationKeelenche	-0.13	0.38	-0.36	0.72
vegetationNukuuchche	-1.02	0.44	-2.33	0.02
milpaYes	0.06	0.33	0.20	0.84
harvestedYes:vegetationKeelenche	0.25	0.57	0.44	0.66
harvestedYes:vegetationNukuuchche	0.82	0.65	1.27	0.21
harvestedYes:milpaYes	-0.07	0.46	-0.15	0.89

ANOVA output

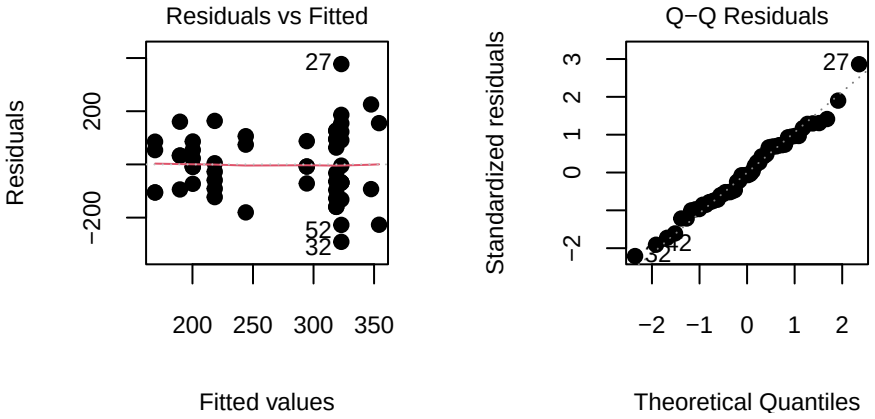
	SS	df	F-statistic	p-value
harvested	0.17	1	0.38	0.54
vegetation	3.37	2	3.87	0.03
milpa	0.01	1	0.02	0.89
harvested:vegetation	0.95	2	1.09	0.35
harvested:milpa	0.01	1	0.02	0.89
residuals	20.04	46		

Trees (10-14 cm DBH)

Model

$$\begin{aligned}\text{10-14 cm trees per ha} \sim\ & \beta_0 + \beta_1 \cdot \text{harvestedYes} + \beta_2 \cdot \text{vegetationKeelenche} + \beta_3 \cdot \text{vegetationNukuuchche} \\ & + \beta_4 \cdot \text{milpaYes} + \beta_5 \cdot \text{harvestedYes:vegetationKeelenche} \\ & + \beta_6 \cdot \text{harvestedYes:vegetationNukuuchche} + \beta_7 \cdot \text{harvestedYes:milpaYes} + \varepsilon\end{aligned}$$

Diagnostic plots



Linear regression output

This model has an R^2 of 0.189.

	Coefficient	SE	t-statistic	p-value
Intercept	325.60	96.31	3.38	0.00
harvestedYes	-52.85	141.93	-0.37	0.71
vegetationKeelenche	28.52	78.36	0.36	0.72
vegetationNukuuchche	-125.33	90.72	-1.38	0.17
milpaYes	-31.17	67.62	-0.46	0.65
harvestedYes:vegetationKeelenche	46.24	118.18	0.39	0.70
harvestedYes:vegetationNukuuchche	70.97	134.30	0.53	0.60
harvestedYes:milpaYes	2.45	95.76	0.03	0.98

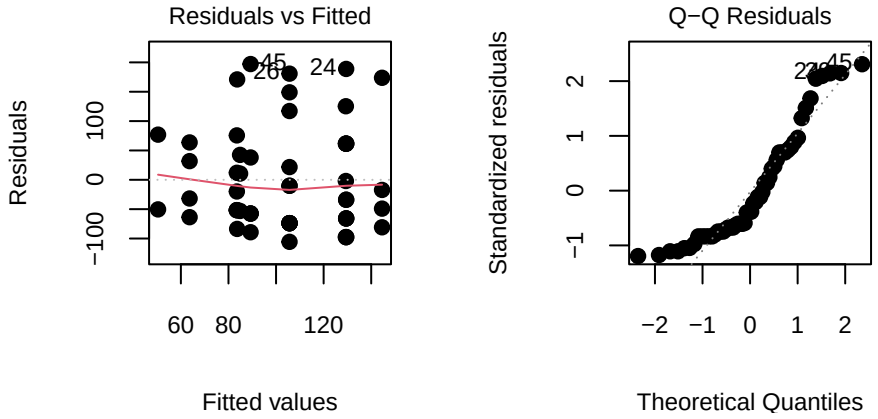
ANOVA output

	SS	df	F-statistic	p-value
harvested	34.09	1	0.00	0.97
vegetation	179553.12	2	4.77	0.01
milpa	7359.88	1	0.39	0.53
harvested:vegetation	5252.71	2	0.14	0.87
harvested:milpa	12.30	1	0.00	0.98
residuals	865374.17	46		

Trees (15-19 cm DBH)

$$\begin{aligned}\text{15-19 cm trees per ha} \sim\ & \beta_0 + \beta_1 \cdot \text{harvestedYes} + \beta_2 \cdot \text{vegetationKeelenche} + \beta_3 \cdot \text{vegetationNukuuchche} \\ & + \beta_4 \cdot \text{milpaYes} + \beta_5 \cdot \text{harvestedYes:vegetationKeelenche} \\ & + \beta_6 \cdot \text{harvestedYes:vegetationNukuuchche} + \beta_7 \cdot \text{harvestedYes:milpaYes} + \varepsilon\end{aligned}$$

Diagnostic plots



Linear regression output

This model has an R^2 of 0.081.

	Coefficient	SE	t-statistic	p-value
Intercept	8.40	64.62	0.13	0.90
harvestedYes	30.65	95.23	0.32	0.75
vegetationKeelenche	42.00	52.58	0.80	0.43
vegetationNukuuchche	80.90	60.87	1.33	0.19
milpaYes	55.26	45.37	1.22	0.23
harvestedYes:vegetationKeelenche	2.60	79.30	0.03	0.97
harvestedYes:vegetationNukuuchche	-36.43	90.11	-0.40	0.69
harvestedYes:milpaYes	-9.43	64.25	-0.15	0.88

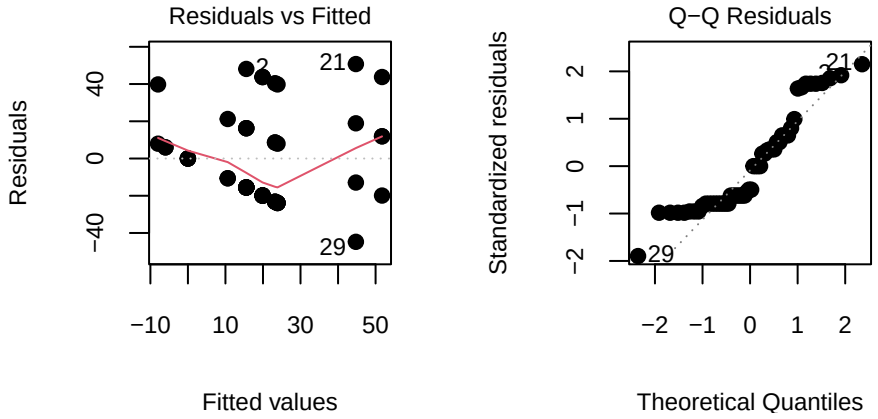
ANOVA output

	SS	df	F-statistic	p-value
harvested	1869.31	1	0.22	0.64
vegetation	16857.86	2	1.00	0.38
milpa	20978.30	1	2.48	0.12
harvested:vegetation	3479.61	2	0.21	0.82
harvested:milpa	182.45	1	0.02	0.88
residuals	389582.70	46		

Trees (20+ cm DBH)

$$\begin{aligned}\text{20+ cm trees per ha} \sim{}& \beta_0 + \beta_1 \cdot \text{harvestedYes} + \beta_2 \cdot \text{vegetationKeelenche} + \beta_3 \cdot \text{vegetationNukuuchche} \\ &+ \beta_4 \cdot \text{milpaYes} + \beta_5 \cdot \text{harvestedYes:vegetationKeelenche} \\ &+ \beta_6 \cdot \text{harvestedYes:vegetationNukuuchche} + \beta_7 \cdot \text{harvestedYes:milpaYes} + \varepsilon\end{aligned}$$

Diagnostic plots



Linear regression output

This model has an R^2 of 0.269.

	Coefficient	SE	t-statistic	p-value
Intercept	-27.85	18.38	-1.52	0.14
harvestedYes	16.92	27.09	0.62	0.54
vegetationKeelenche	19.89	14.96	1.33	0.19
vegetationNukuuchche	51.73	17.32	2.99	0.00
milpaYes	27.85	12.91	2.16	0.04
harvestedYes:vegetationKeelenche	-14.95	22.56	-0.66	0.51
harvestedYes:vegetationNukuuchche	-17.58	25.64	-0.69	0.50
harvestedYes:milpaYes	-6.31	18.28	-0.35	0.73

ANOVA output

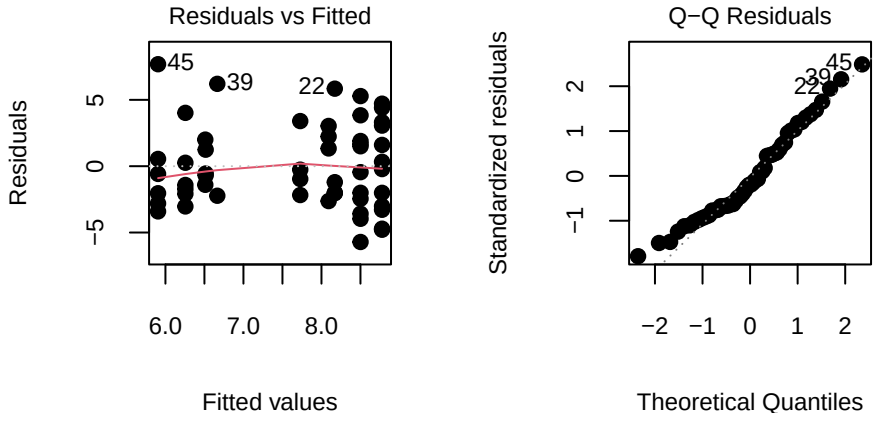
	SS	df	F-statistic	p-value
harvested	31.70	1	0.05	0.83
vegetation	10868.45	2	7.93	0.00
milpa	5008.54	1	7.31	0.01
harvested:vegetation	358.27	2	0.26	0.77
harvested:milpa	81.75	1	0.12	0.73
residuals	31530.82	46		

Basal area

Model

$$\text{total tree BA (m2/ha)} \sim \beta_0 + \beta_1 \cdot \text{harvestedYes} + \beta_2 \cdot \text{vegetationKeelenche} + \beta_3 \cdot \text{vegetationNukuuchche} + \beta_4 \cdot \text{milpaYes} + \beta_5 \cdot \text{harvestedYes:vegetationKeelenche} + \beta_6 \cdot \text{harvestedYes:vegetationNukuuchche} + \beta_7 \cdot \text{harvestedYes:milpaYes} + \varepsilon$$

Diagnostic plots



Linear regression output

This model has an R^2 of 0.111.

	Coefficient	SE	t-statistic	p-value
Intercept	5.46	2.34	2.34	0.02
harvestedYes	-0.78	3.45	-0.23	0.82
vegetationKeelenche	1.05	1.90	0.55	0.58
vegetationNukuuchche	0.44	2.20	0.20	0.84
milpaYes	2.27	1.64	1.38	0.17
harvestedYes:vegetationKeelenche	0.94	2.87	0.33	0.75
harvestedYes:vegetationNukuuchche	1.13	3.26	0.35	0.73
harvestedYes:milpaYes	-0.43	2.33	-0.18	0.85

ANOVA output

	SS	df	F-statistic	p-value
harvested	0.51	1	0.05	0.83
vegetation	12.31	2	0.55	0.58
milpa	34.54	1	3.11	0.08
harvested:vegetation	1.45	2	0.07	0.94
harvested:milpa	0.38	1	0.03	0.85
residuals	510.25	46		

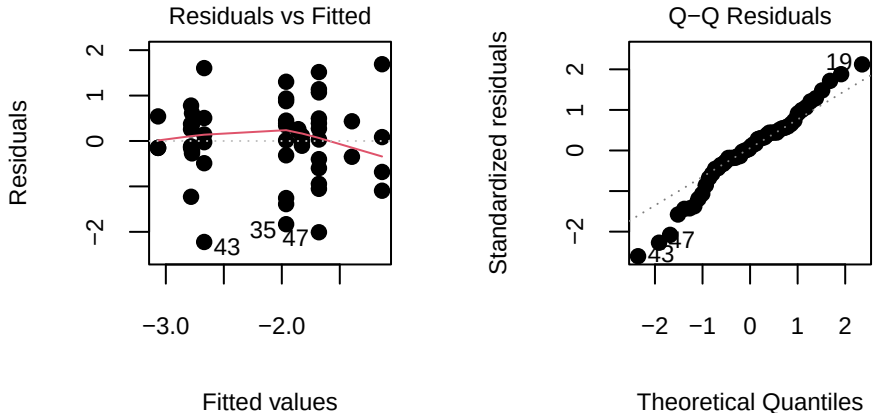
Basal area by size class

Saplings (0-4 cm DBH)

Model

$$\log(\text{sapling BA (m2/ha)}) \sim \beta_0 + \beta_1 \cdot \text{harvestedYes} + \beta_2 \cdot \text{vegetationKeelenche} + \beta_3 \cdot \text{vegetationNukuuchche} + \beta_4 \cdot \text{milpaYes} + \beta_5 \cdot \text{harvestedYes:vegetationKeelenche} + \beta_6 \cdot \text{harvestedYes:vegetationNukuuchche} + \beta_7 \cdot \text{harvestedYes:milpaYes} + \varepsilon$$

Diagnostic plots



Linear regression output

This model has an R^2 of 0.312.

	Coefficient	SE	t-statistic	p-value
Intercept	-0.85	0.65	-1.32	0.19
harvestedYes	-0.87	0.95	-0.91	0.37
vegetationKeelenche	-0.54	0.53	-1.04	0.31
vegetationNukuuchche	-1.93	0.61	-3.18	0.00
milpaYes	-0.28	0.45	-0.63	0.53
harvestedYes:vegetationKeelenche	0.41	0.79	0.51	0.61
harvestedYes:vegetationNukuuchche	0.98	0.90	1.09	0.28
harvestedYes:milpaYes	0.18	0.64	0.28	0.78

ANOVA output

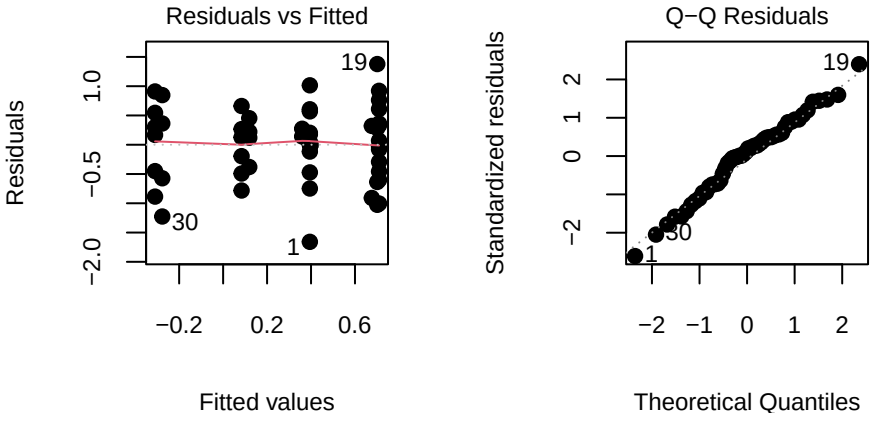
	SS	df	F-statistic	p-value
harvested	0.41	1	0.48	0.49
vegetation	13.26	2	7.84	0.00
milpa	0.31	1	0.37	0.55
harvested:vegetation	1.18	2	0.69	0.50
harvested:milpa	0.07	1	0.08	0.78
residuals	38.92	46		

Trees (5-9 cm DBH)

Model

$$\log(\text{5-9 cm tree BA (m2/ha)}) \sim \beta_0 + \beta_1 \cdot \text{harvestedYes} + \beta_2 \cdot \text{vegetationKeelenche} + \beta_3 \cdot \text{vegetationNukuuchche} + \beta_4 \cdot \text{milpaYes} + \beta_5 \cdot \text{harvestedYes:vegetationKeelenche} + \beta_6 \cdot \text{harvestedYes:vegetationNukuuchche} + \beta_7 \cdot \text{harvestedYes:milpaYes} + \varepsilon$$

Diagnostic plots



Linear regression output

This model has an R^2 of 0.258.

	Coefficient	SE	t-statistic	p-value
Intercept	0.67	0.47	1.43	0.16
harvestedYes	-0.31	0.69	-0.45	0.66
vegetationKeelenche	0.01	0.38	0.02	0.98
vegetationNukuuchche	-0.98	0.44	-2.23	0.03
milpaYes	0.03	0.33	0.10	0.92
harvestedYes:vegetationKeelenche	-0.01	0.57	-0.02	0.99
harvestedYes:vegetationNukuuchche	0.70	0.65	1.08	0.29
harvestedYes:milpaYes	0.00	0.46	0.00	1.00

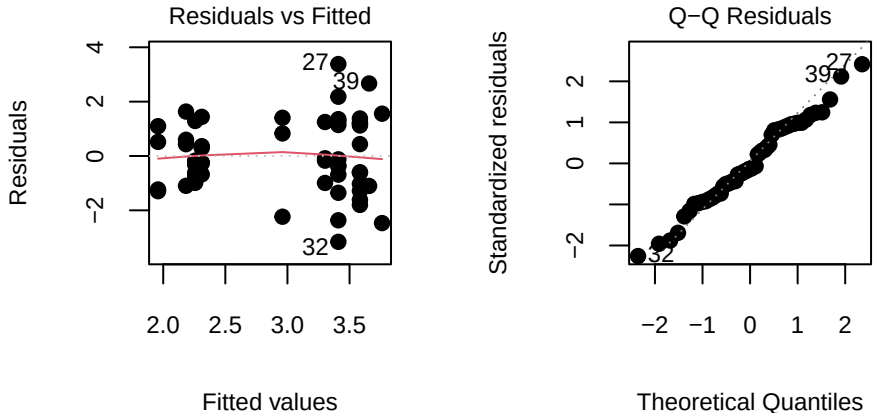
ANOVA output

	SS	df	F-statistic	p-value
harvested	0.03	1	0.08	0.78
vegetation	3.74	2	4.25	0.02
milpa	0.01	1	0.02	0.88
harvested:vegetation	1.17	2	1.33	0.27
harvested:milpa	0.00	1	0.00	1.00
residuals	20.27	46		

Trees (10-14 cm DBH)

$$\begin{aligned}\text{10-14 cm tree BA (m2/ha)} \sim\ & \beta_0 + \beta_1 \cdot \text{harvestedYes} + \beta_2 \cdot \text{vegetationKeelenche} + \beta_3 \cdot \text{vegetationNukuuchche} \\ & + \beta_4 \cdot \text{milpaYes} + \beta_5 \cdot \text{harvestedYes:vegetationKeelenche} \\ & + \beta_6 \cdot \text{harvestedYes:vegetationNukuuchche} + \beta_7 \cdot \text{harvestedYes:milpaYes} + \varepsilon\end{aligned}$$

Diagnostic plots



Linear regression output

This model has an R^2 of 0.179.

	Coefficient	SE	t-statistic	p-value
Intercept	3.65	1.02	3.57	0.00
harvestedYes	-0.62	1.51	-0.41	0.68
vegetationKeelenche	0.11	0.83	0.13	0.90
vegetationNukuuchche	-1.34	0.96	-1.40	0.17
milpaYes	-0.35	0.72	-0.49	0.63
harvestedYes:vegetationKeelenche	0.52	1.25	0.41	0.68
harvestedYes:vegetationNukuuchche	0.57	1.43	0.40	0.69
harvestedYes:milpaYes	0.28	1.02	0.27	0.79

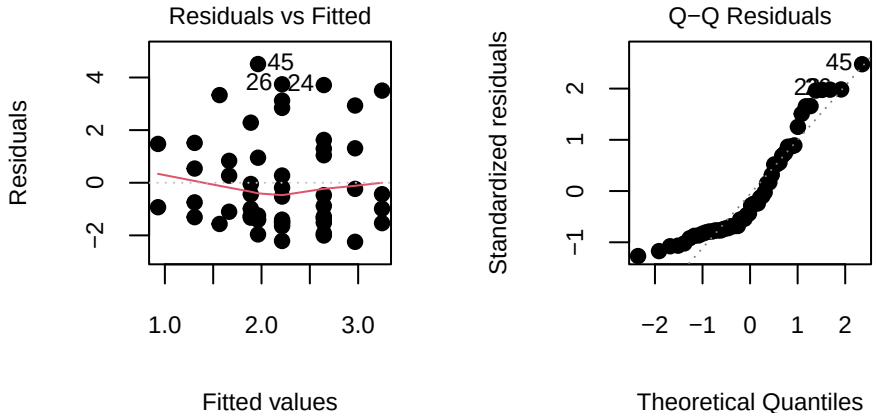
ANOVA output

	SS	df	F-statistic	p-value
harvested	0.03	1	0.01	0.91
vegetation	18.14	2	4.28	0.02
milpa	0.37	1	0.18	0.68
harvested:vegetation	0.40	2	0.09	0.91
harvested:milpa	0.16	1	0.07	0.79
residuals	97.56	46		

Trees (15-19 cm DBH)

$$\text{15-19 cm tree BA (m2/ha)} \sim \beta_0 + \beta_1 \cdot \text{harvestedYes} + \beta_2 \cdot \text{vegetationKeelenche} + \beta_3 \cdot \text{vegetationNukuuchche} + \beta_4 \cdot \text{milpaYes} + \beta_5 \cdot \text{harvestedYes:vegetationKeelenche} + \beta_6 \cdot \text{harvestedYes:vegetationNukuuchche} + \beta_7 \cdot \text{harvestedYes:milpaYes} + \varepsilon$$

Diagnostic plots



Linear regression output

This model has an R^2 of 0.092.

	Coefficient	SE	t-statistic	p-value
Intercept	0.02	1.38	0.02	0.99
harvestedYes	0.56	2.03	0.28	0.78
vegetationKeelenche	0.90	1.12	0.81	0.42
vegetationNukuuchche	1.94	1.30	1.49	0.14
milpaYes	1.28	0.97	1.33	0.19
harvestedYes:vegetationKeelenche	0.08	1.69	0.04	0.96
harvestedYes:vegetationNukuuchche	-0.64	1.92	-0.33	0.74
harvestedYes:milpaYes	-0.20	1.37	-0.15	0.88

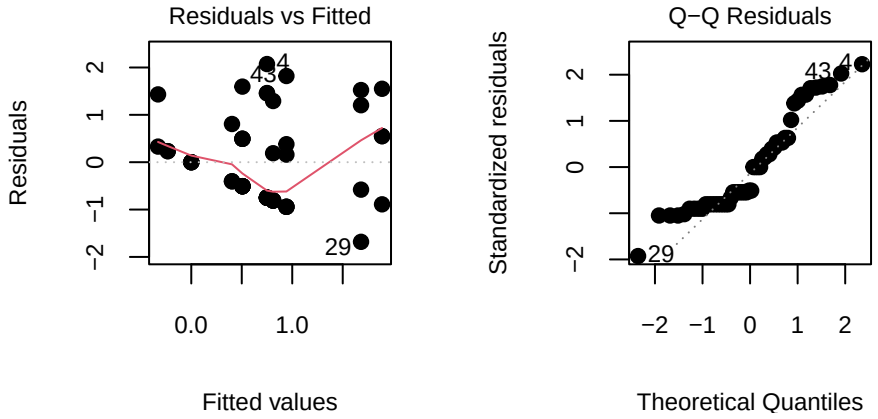
ANOVA output

	SS	df	F-statistic	p-value
harvested	0.65	1	0.17	0.68
vegetation	11.30	2	1.47	0.24
milpa	11.45	1	2.97	0.09
harvested:vegetation	1.15	2	0.15	0.86
harvested:milpa	0.09	1	0.02	0.88
residuals	177.36	46		

Trees (20+ cm DBH)

$$\begin{aligned}\text{20+ cm tree BA (m2/ha)} \sim\ & \beta_0 + \beta_1 \cdot \text{harvestedYes} + \beta_2 \cdot \text{vegetationKeelenche} + \beta_3 \cdot \text{vegetationNukuuchche} \\ & + \beta_4 \cdot \text{milpaYes} + \beta_5 \cdot \text{harvestedYes:vegetationKeelenche} \\ & + \beta_6 \cdot \text{harvestedYes:vegetationNukuuchche} + \beta_7 \cdot \text{harvestedYes:milpaYes} + \varepsilon\end{aligned}$$

Diagnostic plots



Linear regression output

This model has an R^2 of 0.276.

	Coefficient	SE	t-statistic	p-value
Intercept	-1.08	0.68	-1.59	0.12
harvestedYes	0.74	1.00	0.74	0.46
vegetationKeelenche	0.75	0.55	1.36	0.18
vegetationNukuuchche	1.89	0.64	2.95	0.00
milpaYes	1.08	0.48	2.26	0.03
harvestedYes:vegetationKeelenche	-0.65	0.83	-0.78	0.44
harvestedYes:vegetationNukuuchche	-0.61	0.95	-0.64	0.52
harvestedYes:milpaYes	-0.34	0.68	-0.50	0.62

ANOVA output

	SS	df	F-statistic	p-value
harvested	0.03	1	0.03	0.86
vegetation	15.26	2	8.14	0.00
milpa	6.79	1	7.24	0.01
harvested:vegetation	0.57	2	0.31	0.74
harvested:milpa	0.24	1	0.25	0.62
residuals	43.14	46		