

Ancient Animals' Role Analysis Final Report

Keliang Xu, Shuting Li, Tao He

3/21/2022

Abstract

This study is an analysis of the role of animals in the ancient economy at Teotihuacan Mexico. The purpose of this report is to provide statistical analysis to this study with the main goal of analyzing whether there exist significant differences in animals' diets from different sites. We did some exploratory data analysis to see the distribution of the isotope levels and used two-way ANOVA Test to compare how sites effect the isotope levels. Finally, we concluded there is a statistically-significant difference (p-values < 0.001) in average isotope level among different species and there is no significant difference in diets between two places, except the protein composition.

Introduction

The isotopes we focus on are stable carbon and nitrogen isotopes, which can reflect the level of human-grown plants and protein in the diet respectively. Since isotopes can enter animals' bones by eating, by comparing the average levels of carbon and nitrogen isotopes from ancient animals' bones between two sites of ancient Mexico, we can get insight into whether animals from these two sites had similar diets.

Data Cleaning and Processing

Firstly, to clean this big data set, and also keep the same as the client's data, we dropped 4 useless observations client mentioned, which IDs are MC381, MC389, MC393, MC76.

Secondly, to help focus on the main question, we removed all irrelevant variables, the final variables of interest are "Taxa", "Site", "d13C_VPDB", "ap_13C_VPDB" and "d15N_air". Then we combined "d13C_VPDB", "ap_13C_VPDB" and "d15N_air" into one variable which named "isotope level".

Data Description

Table1: Data Explanation

Variable	Explanation	Description
Taxa	Animal Species	turkey, cottontails, deer, hare
site	Ancient Mexico Cities	BAOX, PATT
isotope level	Carbon and Nitrogen level	d13C_VPDB, ap_13C_VPDB, d15N_air

EDA

Summary of Data:

To double check our client's work, we calculated the statistics of data:

Table 2: Summary of Data

Taxa	Site	variable	min	median	mean	max	sd	count
Cottontail	BAOX	d13C_VPDB	-20.49	-15.69	-16.26	-11.53	3.14	9
Cottontail	BAOX	d15N_air	3.59	6.58	6.99	13.60	2.86	9
Cottontail	BAOX	ap_13C_VPDB	-12.80	-7.69	-8.71	-5.50	2.54	9
Cottontail	Tlajinga	d13C_VPDB	-21.14	-16.92	-16.66	-9.90	3.80	8
Cottontail	Tlajinga	d15N_air	2.77	3.88	4.30	7.00	1.53	8
Cottontail	Tlajinga	ap_13C_VPDB	-13.26	-8.41	-9.25	-4.58	3.03	8
Deer	BAOX	d13C_VPDB	-20.63	-19.38	-18.83	-16.58	1.37	9
Deer	BAOX	d15N_air	3.50	4.97	5.27	6.91	1.22	9
Deer	BAOX	ap_13C_VPDB	-12.84	-10.43	-10.12	-6.51	1.78	9
Deer	Tlajinga	d13C_VPDB	-22.08	-18.92	-19.25	-15.79	2.47	9
Deer	Tlajinga	d15N_air	2.93	4.71	4.67	6.48	1.20	9
Deer	Tlajinga	ap_13C_VPDB	-15.67	-10.94	-10.96	-8.29	2.57	8
Hare	BAOX	d13C_VPDB	-19.51	-16.97	-16.16	-11.19	2.64	11
Hare	BAOX	d15N_air	3.40	6.17	6.28	8.44	1.58	11
Hare	BAOX	ap_13C_VPDB	-11.74	-8.38	-7.98	-4.52	2.38	11
Hare	Tlajinga	d13C_VPDB	-20.92	-14.02	-15.06	-11.85	2.89	8
Hare	Tlajinga	d15N_air	1.69	5.64	5.48	8.29	2.43	8
Hare	Tlajinga	ap_13C_VPDB	-13.05	-7.39	-8.16	-5.58	2.52	8
Turkey	BAOX	d13C_VPDB	-15.26	-12.68	-12.90	-10.35	1.86	8
Turkey	BAOX	d15N_air	5.50	8.26	8.32	10.75	1.63	8
Turkey	BAOX	ap_13C_VPDB	-7.13	-4.14	-4.49	-2.66	1.36	8
Turkey	Tlajinga	d13C_VPDB	-14.21	-11.95	-11.91	-9.81	1.73	6
Turkey	Tlajinga	d15N_air	5.27	7.61	7.31	8.53	1.26	6
Turkey	Tlajinga	ap_13C_VPDB	-7.14	-4.96	-5.39	-4.05	1.21	6

In this form, we can see the mean, median, maximum, minimum, and standard deviation values of isotopes for each kind of animal in each site. The result is exactly same with client's result.

We can see almost all average isotope levels are similar between different sites, except Nitrogen values in cottontail.

Boxplot

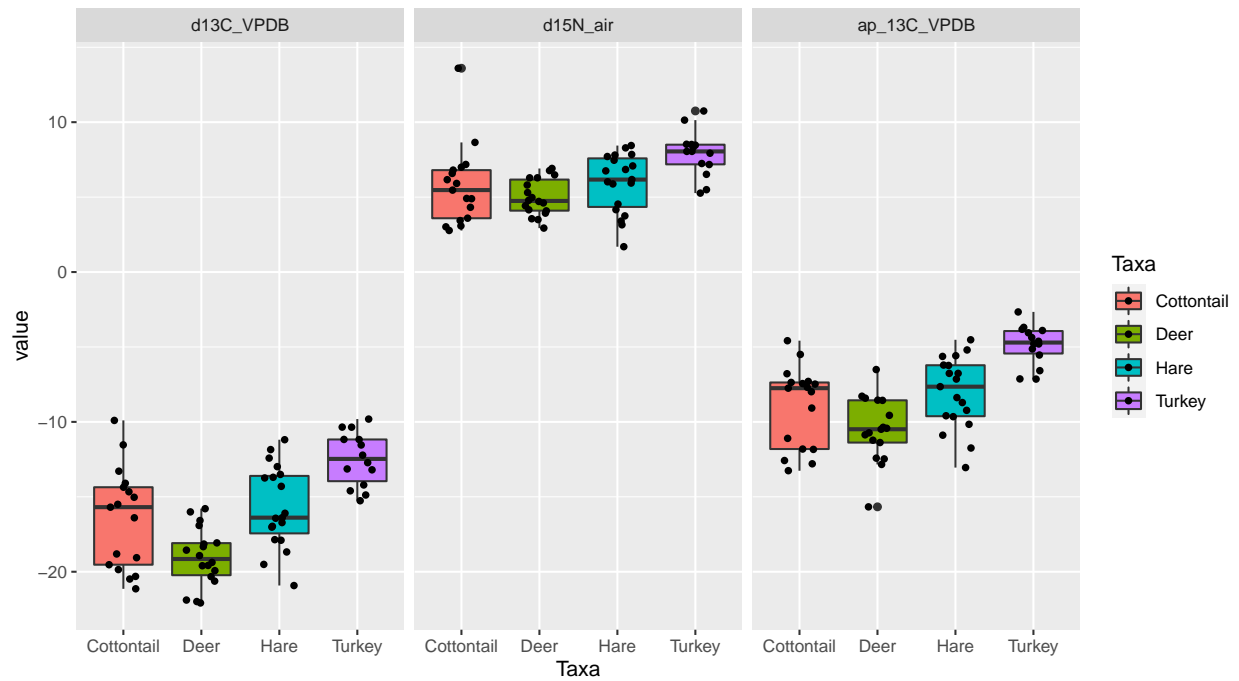


Figure 1: Isotope Levels vs. Taxa

In figure 1, we compared isotopes levels in 4 taxa. It showed the level of d15N_air is less diverse than the other 2 isotopes.

Then we compared isotopes in two sites for each kind of animal. Figure 2 and figure 3 showed that the median level of d13C_VPDB and d15N_air in BAOX is a little higher than Tlajinga of cottontail. Similarly, the median level of d13C_VPDB and d15N_air of hare also show differences in the boxplot.

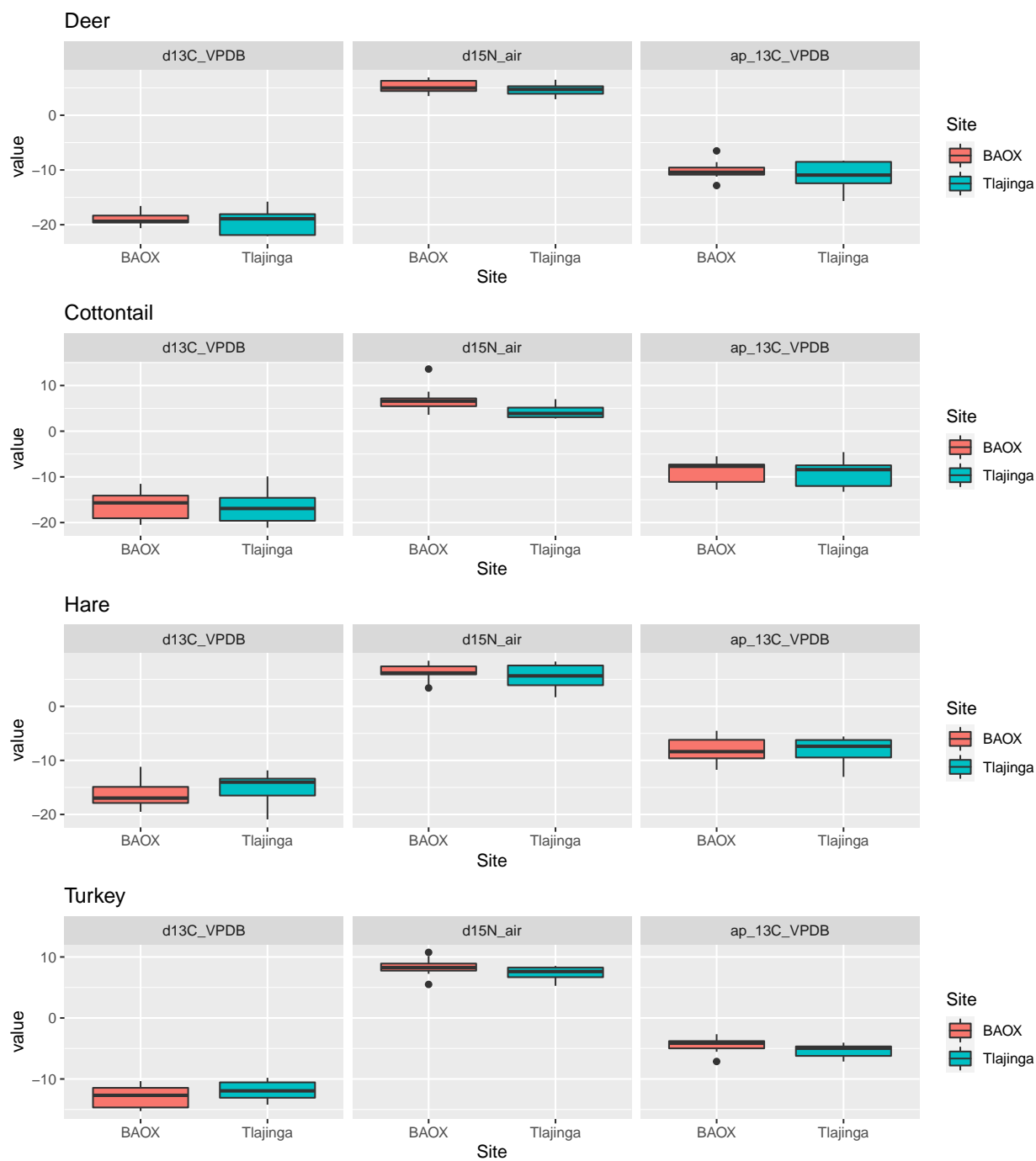


Figure 2: Isotope Levels vs. Site

Testing

Then, we plan to try some statistical tests to compare the levels of the isotope in the two sites more precisely. We use two-way ANOVA tests to figure out how three isotope levels averages vary according to the level of species and location variables.

Firstly, we checked the normality.

Checking Normality

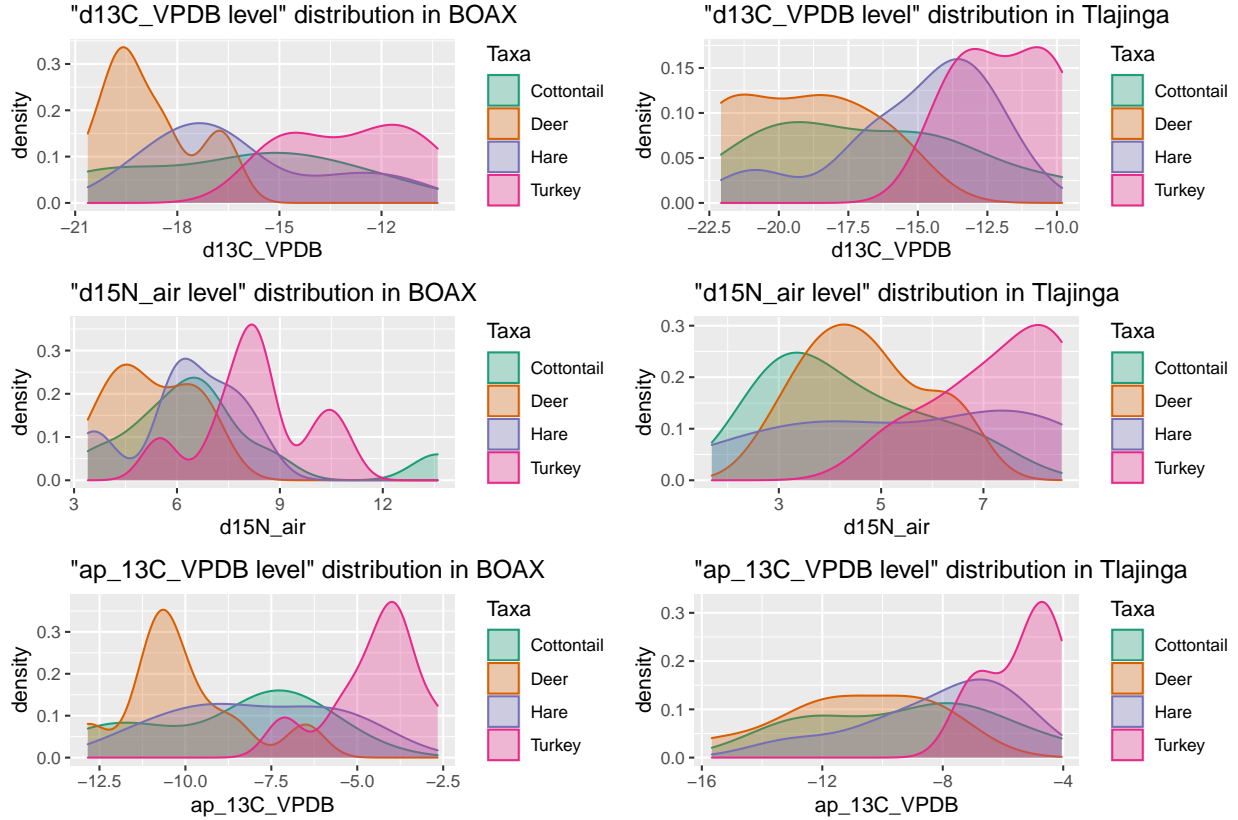


Figure 3: Distribution of Three Isotope Levels For Each Species In Each Place

It is difficult to observe from the figures whether they follow the normal distribution. But comparing the distributions of the two places, most of them follow different distributions. It is reasonable since our sample size is too small.

Therefore, we also plot the Q-Q plots for checking the normality. The purpose of the quantile-quantile (Q-Q) plot is to show if two data sets come from the same distribution. Plotting the first data set's quantiles along the x-axis and plotting the second data set's quantiles along the y-axis is how the plot is constructed.

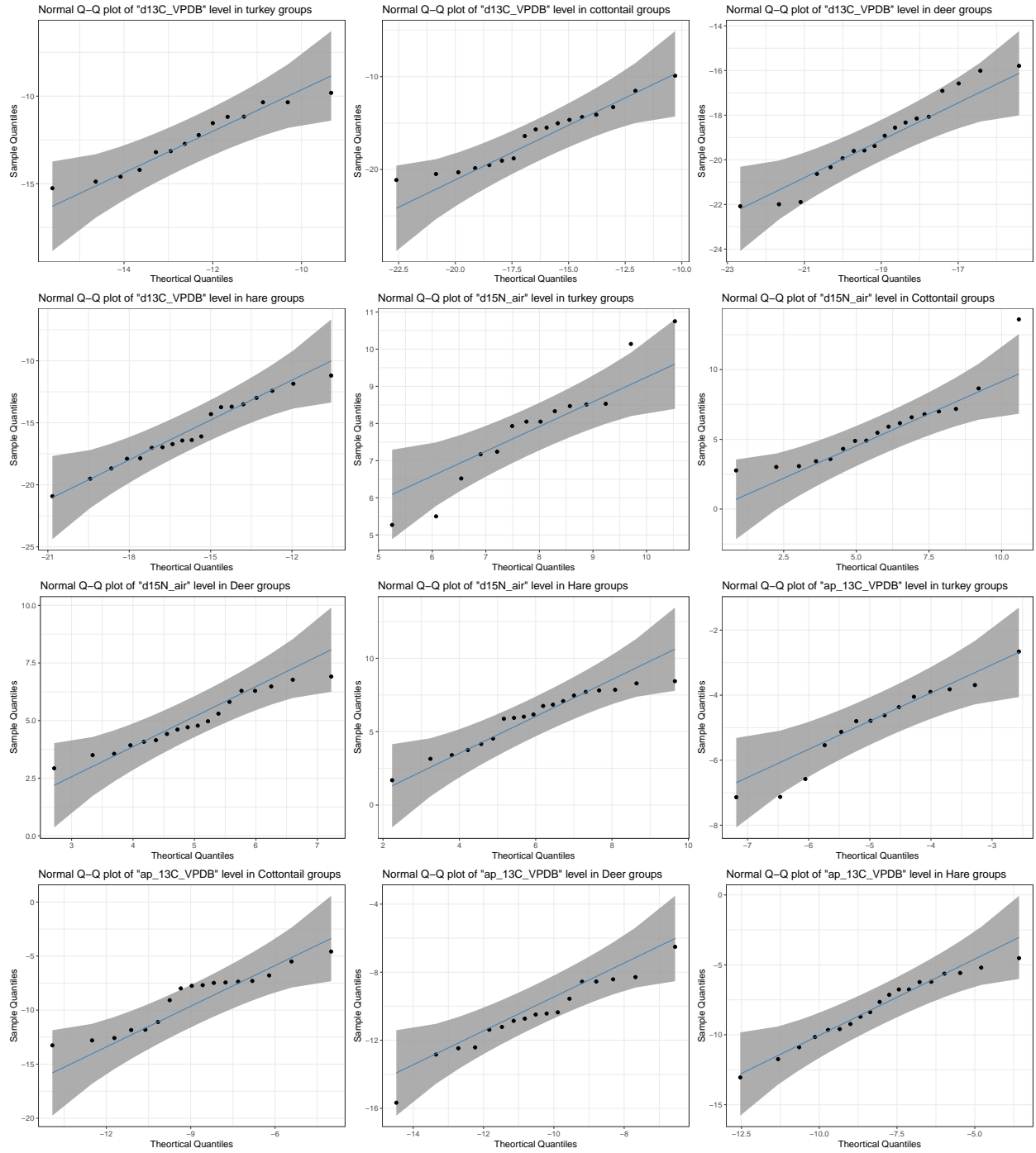


Figure 4: Normal Q-Q Plots in Each Species Group

From the normal Q-Q plots, except the “d15N_air” level in turkey groups, as almost all the points fall approximately along this reference line, we can assume normality to some extent.

Two-Way ANOVA Tests

```
## [1] "'d13C_VPDB' Level ~ animal type + site"

##           Df Sum Sq Mean Sq F value    Pr(>F)
## Taxa       3  345.0   114.99   17.262 2.72e-08 ***
## Site       1    1.4     1.45    0.217   0.643
## Residuals  63  419.7     6.66
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## [1] "'d15N_air' Level ~ animal type + site"

##           Df Sum Sq Mean Sq F value    Pr(>F)
## Taxa       3   70.30   23.433    7.017 0.000381 ***
## Site       1   26.75   26.751    8.011 0.006231 **
## Residuals  63 210.37    3.339
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## [1] "'ap_13C_VPDB' Level ~ animal type + site"

##           Df Sum Sq Mean Sq F value    Pr(>F)
## Taxa       3 256.24    85.41   17.043 3.48e-08 ***
## Site       1   5.72     5.72    1.142   0.289
## Residuals  62 310.72     5.01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 1 observation deleted due to missingness
```

From this output, we found a statistically-significant difference in average isotope level among different *Taxa* (p-values < 0.001). Different animals have different isotope levels.

For the “d13C_VPDB” Level and “ap_13C_VPDB” Level, *Site* is not significant (p-values > 0.001), which means there is no significant difference between two places. However, for the “d15N_air” Level, *Site* factor effects the isotope significantly, means there are some differences in protein composition of animals’ diets from BAOX and PATT.

To check the normality assumption, we also did Q-Q Plots with the residuals that client could refer to in the appendix.

Conclusion

We checked all client’s works, including the summary of data and box plots. Our results are consistent with hers. Then, after we did two-way ANOVA tests to statistically analyze the main problem we want to solve, we concluded there is a statistically-significant difference in average isotope level among different species and there is no significant difference in diets between two places, except the protein composition.

Citation

Rebecca Bevens, “An Introduction to the Two-Way ANOVA.”, Scribbr, 7 Jan. 2021, <https://www.scribbr.com/statistics/two-way-anova/>.

“Two-Way ANOVA Test in R.” STHDA, <http://www.sthda.com/english/wiki/two-way-anova-test-in-r#infos>.

Appendix

Test code

Two-way ANOVA tests

```
# model 1
res.aovdc<-aov(value ~ Taxa + Site, data = data_dc)
# model 2
interaction.aovdc<-aov(value ~ Taxa * Site, data = data_dc)

model.set<-list(res.aovdc, interaction.aovdc)
model.names <- c("two.way", "interaction")
print("'d13C_VPDB' Level ~ animal type + site")
```

```
## [1] "'d13C_VPDB' Level ~ animal type + site"
```

```
aictab(model.set, modnames = model.names)
```

```
##
## Model selection based on AICc:
##
##           K    AICc Delta_AICc AICcWt Cum.Wt      LL
## two.way      6 330.11      0.00   0.96   0.96 -158.37
## interaction  9 336.37      6.26   0.04   1.00 -157.63
```

```
# model 1
res.aovn<-aov(value ~ Taxa + Site, data = data_n)
# model 2
interaction.aovn<-aov(value ~ Taxa * Site, data = data_n)

model.set<-list(res.aovn, interaction.aovn)
model.names <- c("two.way", "interaction")
print("'d15N_air' Level ~ animal type + site")
```

```
## [1] "'d15N_air' Level ~ animal type + site"
```

```
aictab(model.set, modnames = model.names)
```

```
##
## Model selection based on AICc:
##
##           K    AICc Delta_AICc AICcWt Cum.Wt      LL
## two.way      6 283.15      0.00   0.87   0.87 -134.89
## interaction  9 286.94      3.79   0.13   1.00 -132.92
```

```

# model 1
res.aovapc<-aov(value ~ Taxa + Site, data = data_apc)
# model 2
interaction.aovapc<-aov(value ~ Taxa * Site, data = data_apc)

model.set<-list(res.aovapc, interaction.aovapc)
model.names <- c("two.way", "interaction")
print("'ap_13C_VPDB' Level ~ animal type + site")

## [1] "'ap_13C_VPDB' Level ~ animal type + site"

aictab(model.set, modnames = model.names)

```

```

##
## Model selection based on AICc:
##
##           K   AICc Delta_AICc AICcWt Cum.Wt    LL
## two.way    6 306.33      0.00   0.98  0.98 -146.46
## interaction 9 313.78      7.45   0.02  1.00 -146.31

```

Check the homogeneity of variance assumption

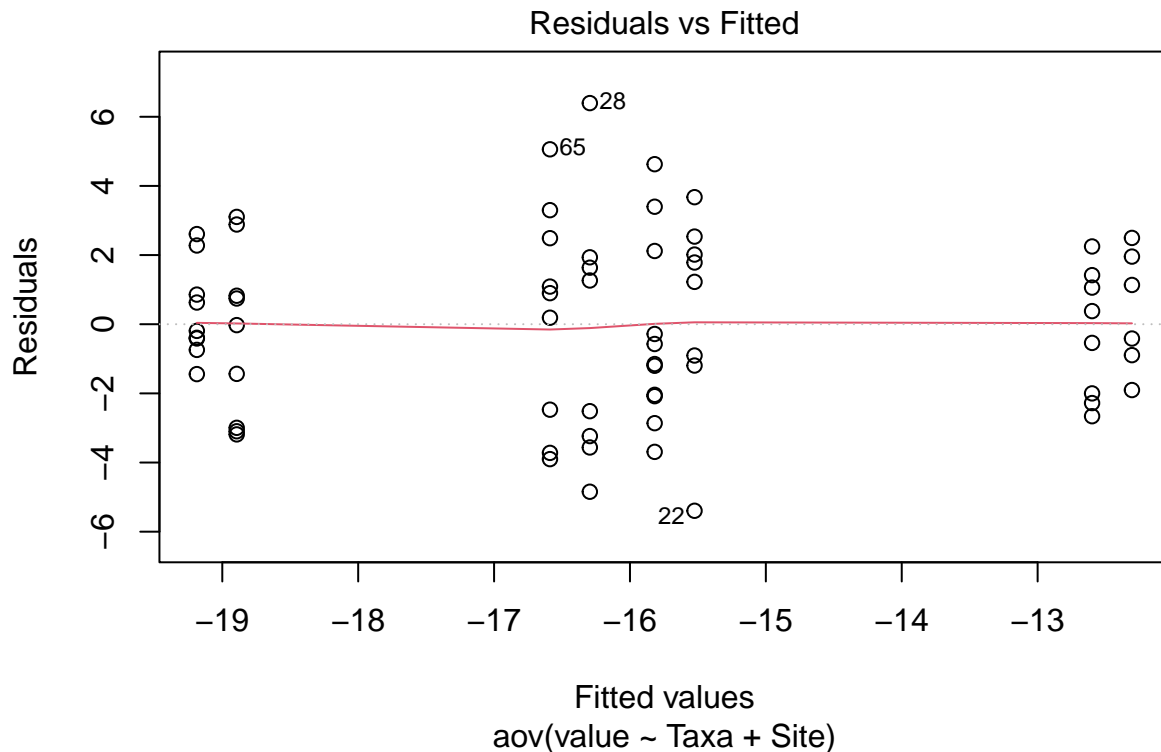


Figure 5: “Residuals versus Fits plot of”d13C_VPDB” Level”

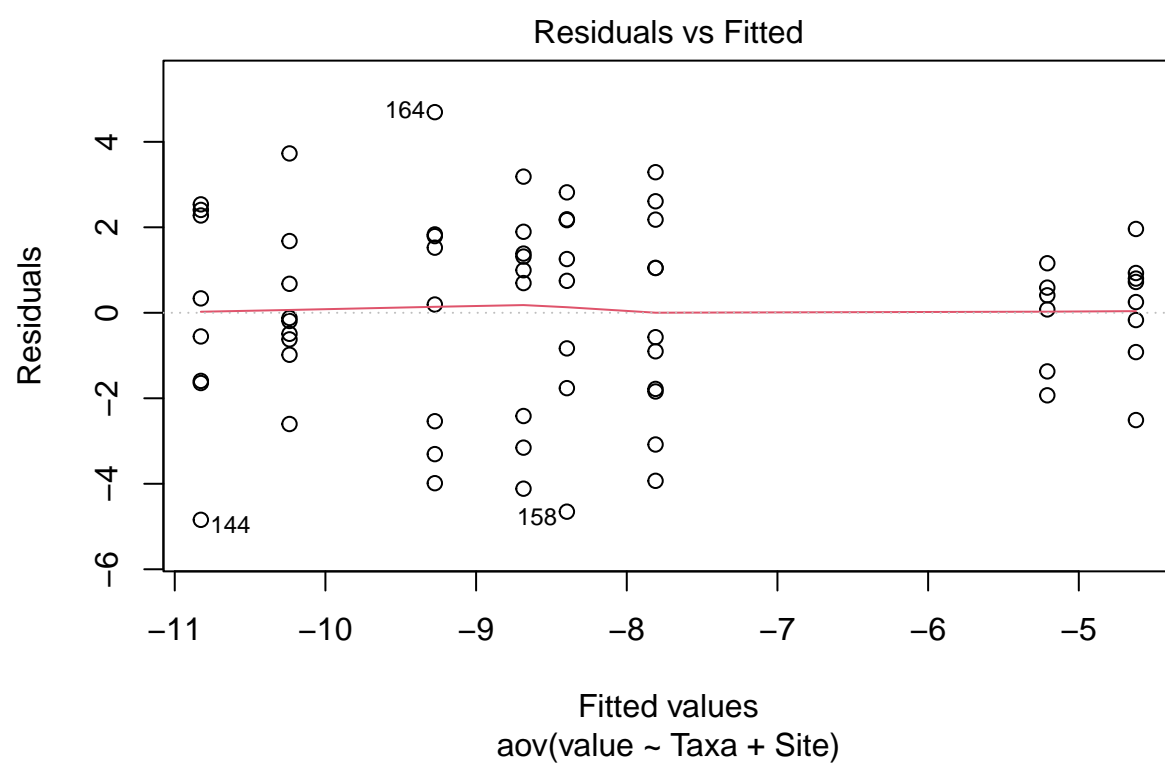


Figure 6: “Residuals versus Fits plot of”d15N_air” Level”

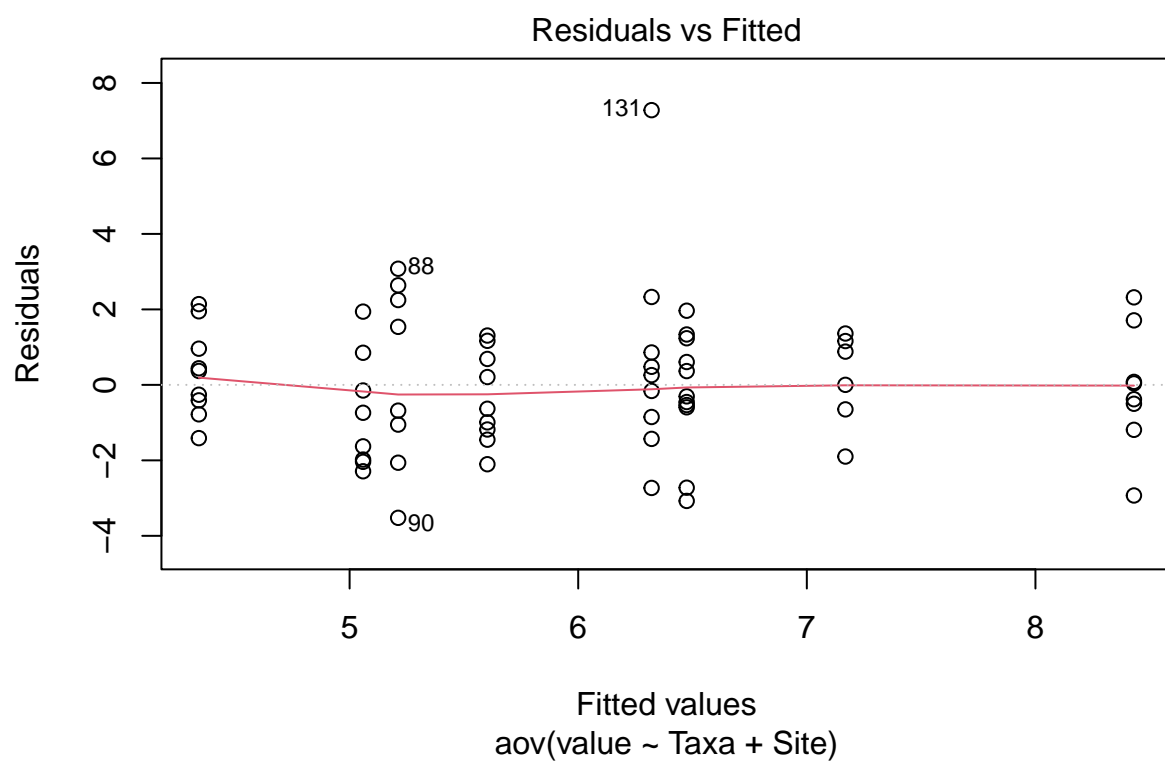


Figure 7: “Residuals versus Fits plot of”ap_13C_VPDB” Level”

```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group 7  1.4051 0.2201
##      60
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group 7  1.1254 0.3595
##      60
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group 7  1.1312 0.3562
##      59
```

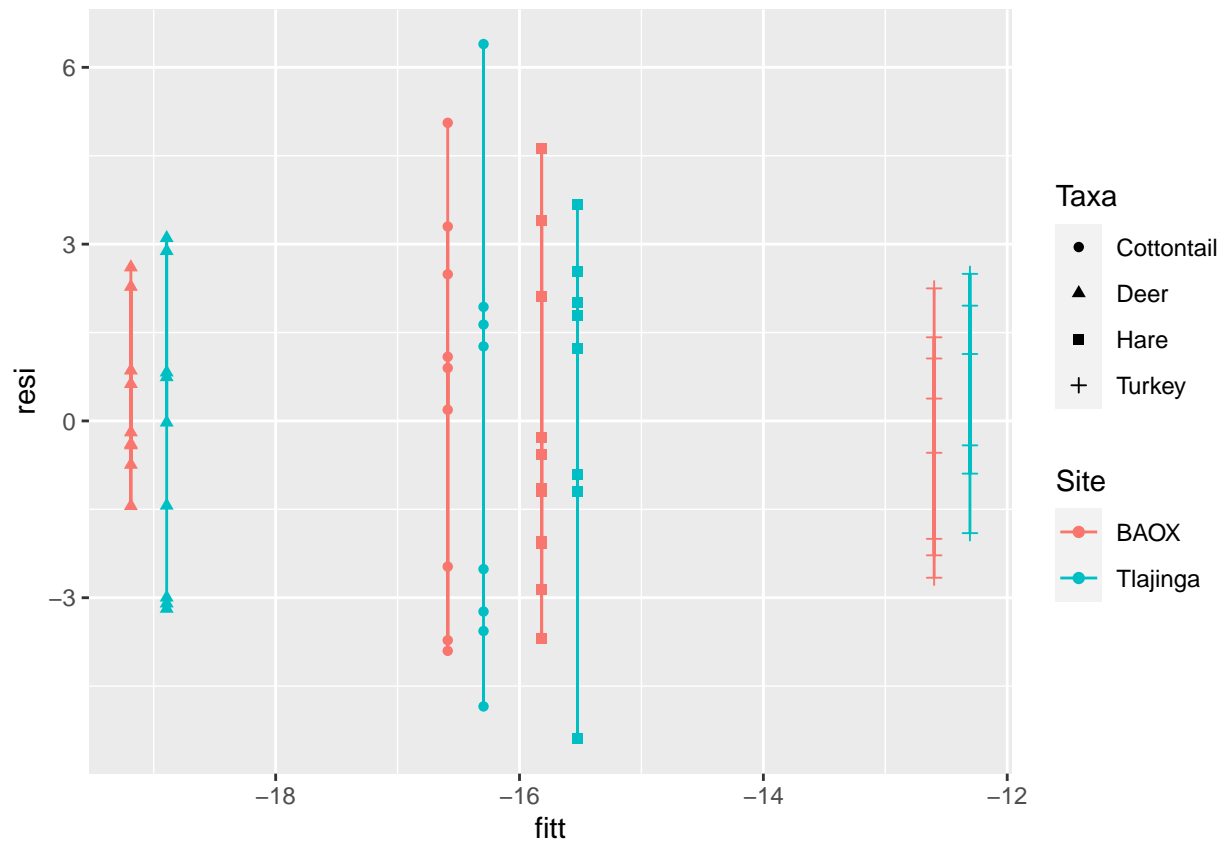


Figure 8: “Residual plot of”d13C_VPDB” Level based on groups”

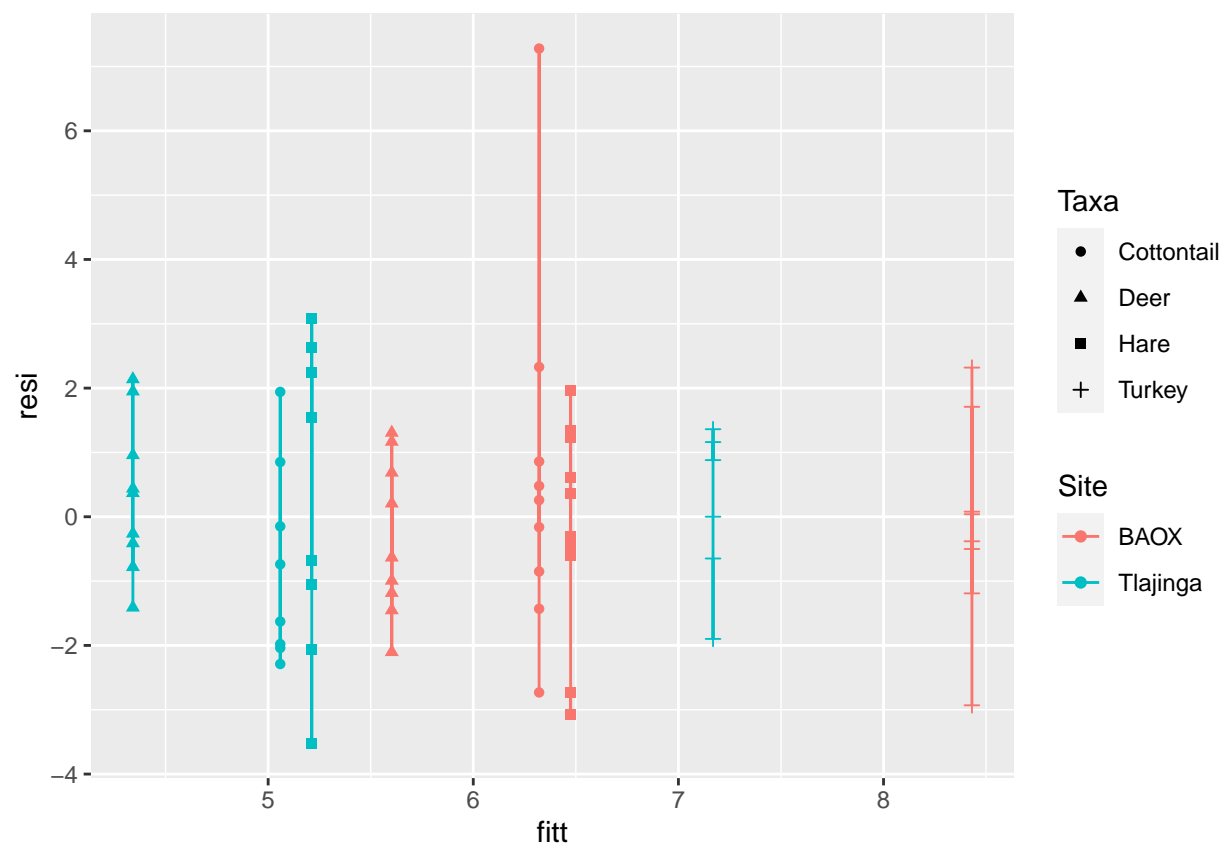


Figure 9: “Residual plot of”d15N_air” Level based on groups”

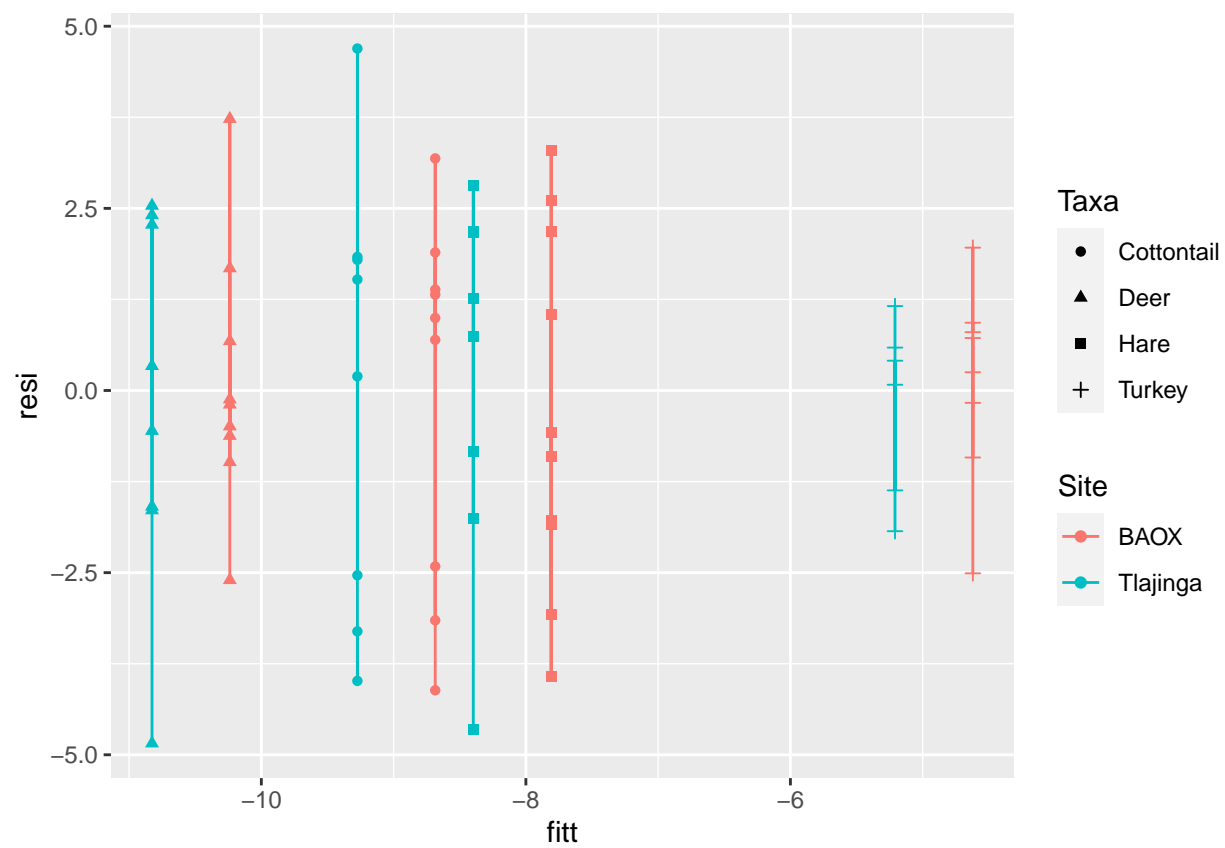


Figure 10: “Residual plot of”ap_13C_VPDB” Level based on groups”

Check the normality assumption

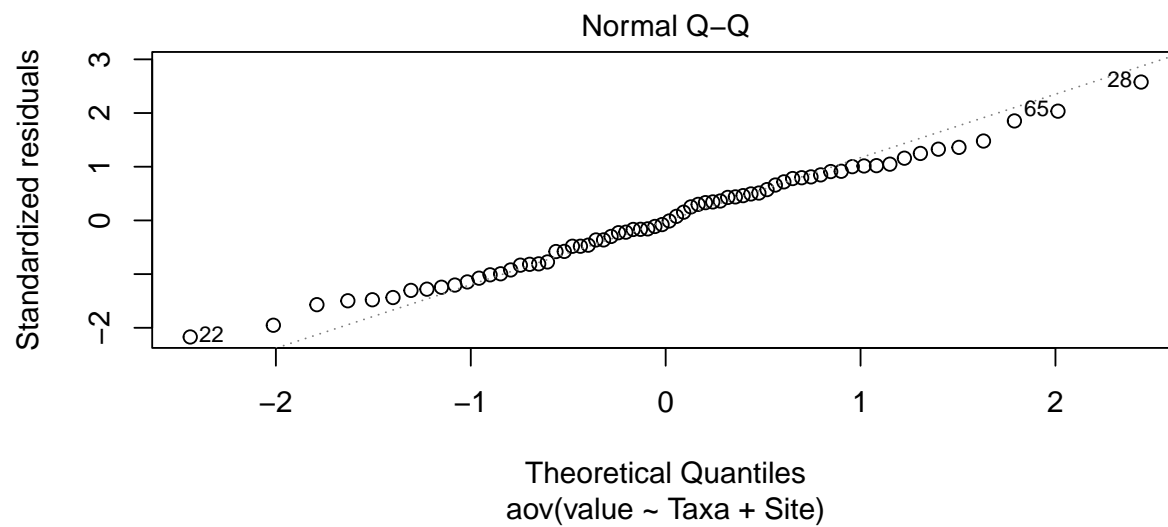


Figure 11: Normality plot of the residuals of “d13C_VPDB” Level

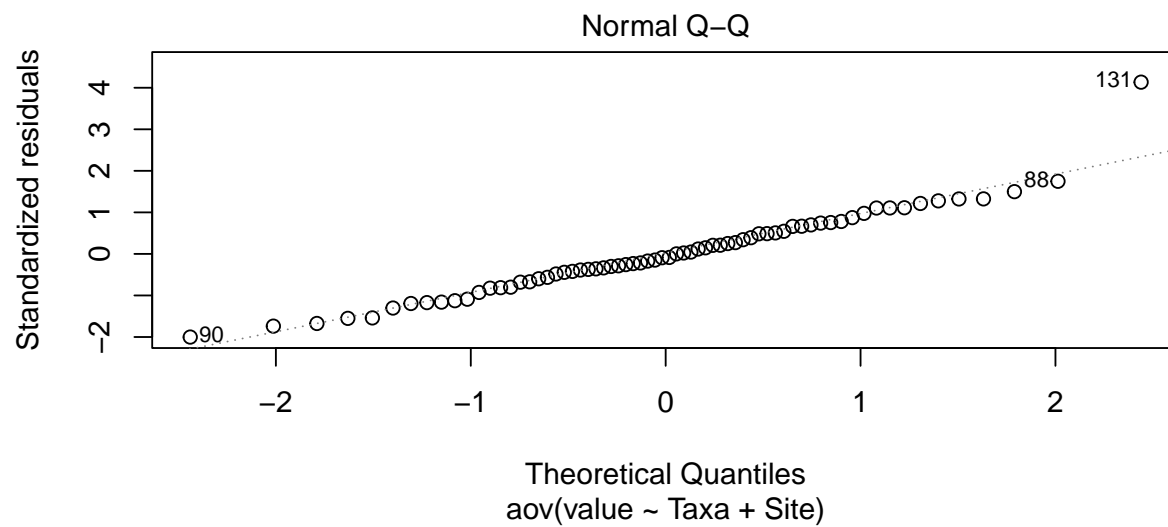


Figure 12: Normality plot of the residuals of “d15N_air” Level

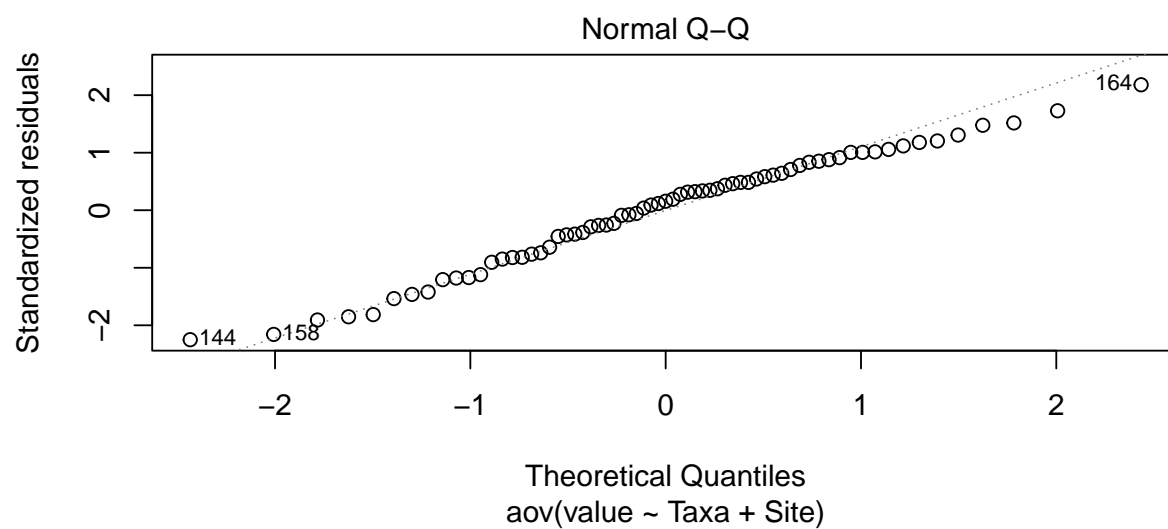


Figure 13: Normality plot of the residuals of “ap_13C_VPDB” Level

Reference

Somerville, A.D., Sugiyama, N., Manzanilla, L.R. et al. Leporid management and specialized food production at Teotihuacan: stable isotope data from cottontail and jackrabbit bone collagen. *Archaeol Anthropol Sci* 9, 83–97 (2017). <https://doi.org/10.1007/s12520-016-0420-2>