

Análisis de datos exploratorio

Paz_Caballero

27 November 2021

Se evaluaron, en condiciones de laboratorio, los efectos del incremento del pCO₂ sobre las condiciones fisiológicas y metabólicas de ambas especies, y sobre las respuestas fenotípicas a nivel de población. Para esto, *Scurria zebrina* y *Scurria viridula*, fueron expuestos a 500 µatm (valores actuales) y a 800 y 1500 µatm, niveles de pCO₂ que se pronostican para el año 2050 y 2100, respectivamente.

```
## # A tibble: 6 x 12
##   Especie Localidad Etiqueta Tratamiento `Dias Trat.` Largo Ancho `Peso Boyante`
##   <fct>   <fct>       <dbl> <fct>          <dbl> <dbl> <dbl>          <dbl>
## 1 Scurri~ Talcaruca      63 500          15 13.7 11.6          0.113
## 2 Scurri~ Talcaruca      66 500          15 20.6 17.2          0.407
## 3 Scurri~ Talcaruca      67 500          15 15.0 12.9          0.164
## 4 Scurri~ Talcaruca      15 500          15 33.4 27.5          1.86
## 5 Scurri~ Talcaruca      55 500          15 19.1 15.3          0.370
## 6 Scurri~ Talcaruca      26 500          15 15.4 14.8          0.195
## # ... with 4 more variables: Pendiente <dbl>, Tiempo <dtm>, ml_g_h 14C <dbl>,
## #   mgO2_h_1_g_1 <dbl>

## # A tibble: 6 x 12
##   Especie Localidad Etiqueta Tratamiento `Dias Trat.` Largo Ancho `Peso Boyante`
##   <fct>   <fct>       <dbl> <fct>          <dbl> <dbl> <dbl>          <dbl>
## 1 Scurri~ Los Moll~      25 1500          45 23.1 19.7          0.207
## 2 Scurri~ Los Moll~      57 1500          45 25.0 21.4          1.18
## 3 Scurri~ Los Moll~       1 1500          45 22.7 20.2          0.670
## 4 Scurri~ Los Moll~      57 1500          45 14.0 12.3          0.186
## 5 Scurri~ Los Moll~      39 1500          45 15.3 13.4          0.641
## 6 Scurri~ Los Moll~      66 1500          45 22.4 20.0          0.397
## # ... with 4 more variables: Pendiente <dbl>, Tiempo <dtm>, ml_g_h 14C <dbl>,
## #   mgO2_h_1_g_1 <dbl>

##           Especie      Localidad      Etiqueta      Tratamiento
## Scurria viridula:52  Los Molles:49  Min.    : 1.00  500 :36
## Scurria zebrina :58  Talcaruca :61  1st Qu.: 25.00  800 :36
##                               Median : 48.00  1500:38
##                               Mean   : 49.58
##                               3rd Qu.: 63.75
##                               Max.   :351.00
##   Dias Trat.      Largo      Ancho      Peso Boyante
## Min.    :15.00  Min.    :12.87  Min.    :10.74  Min.    :0.1129
## 1st Qu.:15.00  1st Qu.:18.39  1st Qu.:15.99  1st Qu.:0.3608
## Median :30.00  Median :23.06  Median :20.20  Median :0.7285
## Mean   :25.64  Mean   :23.27  Mean   :20.37  Mean   :0.7567
## 3rd Qu.:30.00  3rd Qu.:27.74  3rd Qu.:24.16  3rd Qu.:1.0638
## Max.   :45.00  Max.   :35.23  Max.   :32.74  Max.   :1.8560
##   Pendiente      Tiempo      ml_g_h 14C
```

```

## Min.      :0.01040   Min.      :1899-12-31 00:13:00   Min.      :0.08199
## 1st Qu.:0.01760   1st Qu.:1899-12-31 00:16:15   1st Qu.:0.13876
## Median :0.02090   Median :1899-12-31 00:19:07   Median :0.16478
## Mean    :0.02485   Mean    :1899-12-31 00:20:23   Mean    :0.19590
## 3rd Qu.:0.02838   3rd Qu.:1899-12-31 00:24:07   3rd Qu.:0.22371
## Max.    :0.08550   Max.    :1899-12-31 00:33:15   Max.    :0.67408
## mg02_h_1_g_1
## Min.      :0.01307
## 1st Qu.:0.14230
## Median :0.27614
## Mean    :0.42262
## 3rd Qu.:0.57602
## Max.    :3.24390

## tibble [110 x 12] (S3: tbl_df/tbl/data.frame)
## $ Especie      : Factor w/ 2 levels "Scurria viridula",...: 2 2 2 2 2 2 1 1 1 1 ...
## $ Localidad    : Factor w/ 2 levels "Los Molles","Talcaruca": 2 2 2 2 2 2 2 2 2 2 ...
## $ Etiqueta     : num [1:110] 63 66 67 15 55 26 68 15 69 74 ...
## $ Tratamiento  : Factor w/ 3 levels "500","800","1500": 1 1 1 1 1 1 1 1 1 1 ...
## $ Dias Trat.   : num [1:110] 15 15 15 15 15 15 15 15 15 15 ...
## $ Largo        : num [1:110] 13.7 20.6 15 33.4 19.1 ...
## $ Ancho        : num [1:110] 11.7 17.2 12.9 27.5 15.3 ...
## $ Peso Boyante: num [1:110] 0.113 0.407 0.164 1.856 0.37 ...
## $ Pendiente    : num [1:110] 0.018 0.0284 0.0175 0.0184 0.031 0.0189 0.0302 0.0184 0.0855 0.0609 ...
## $ Tiempo       : POSIXct[1:110], format: "1899-12-31 00:15:00" "1899-12-31 00:17:15" ...
## $ ml_g_h 14C   : num [1:110] 0.142 0.224 0.138 0.145 0.244 ...
## $ mg02_h_1_g_1: num [1:110] 1.257 0.5507 0.8397 0.0782 0.6611 ...

```

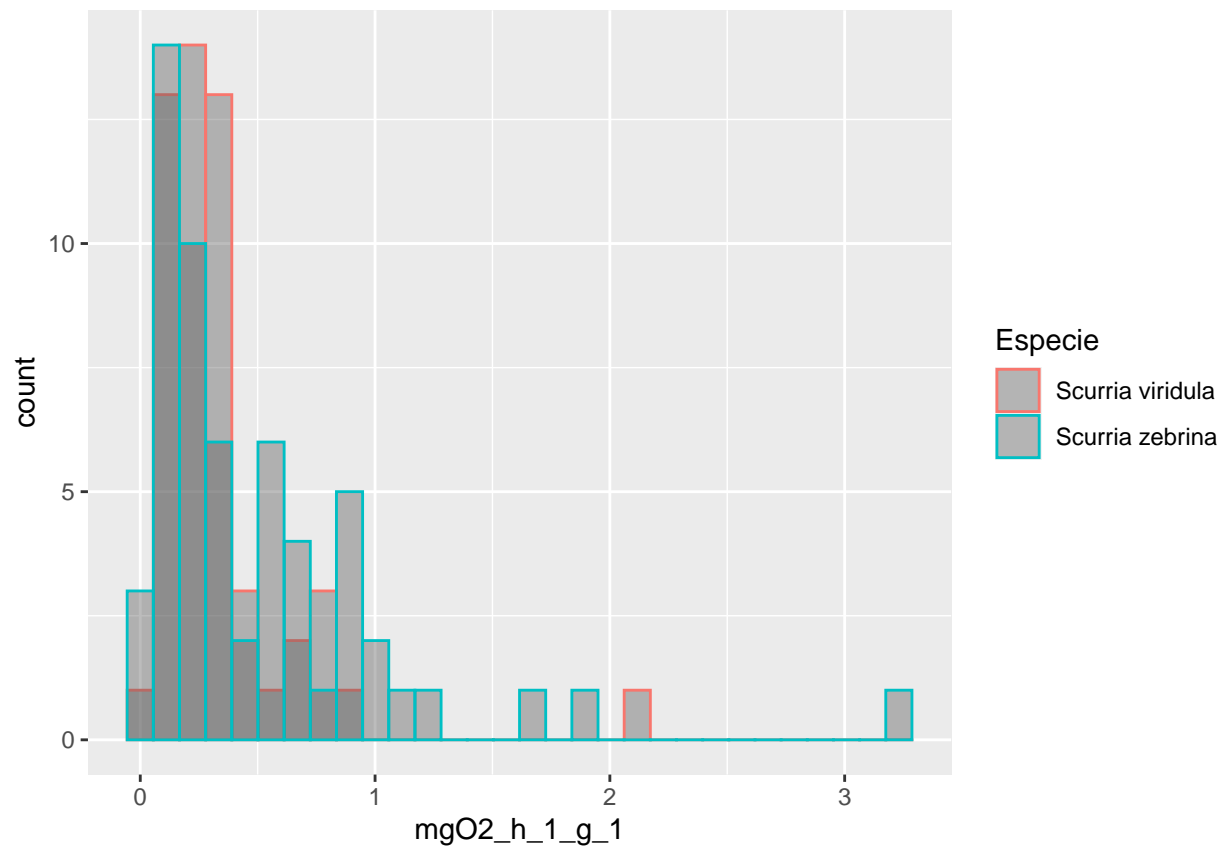


Figura 1. Distribución de metabolismo de ambas especies.

Se observa que no existe una distribución normal, además los datos se encuentran hacia la izquierda de la media, (el metabolismo (variable respuesta) es una variable cuantitativa continua).

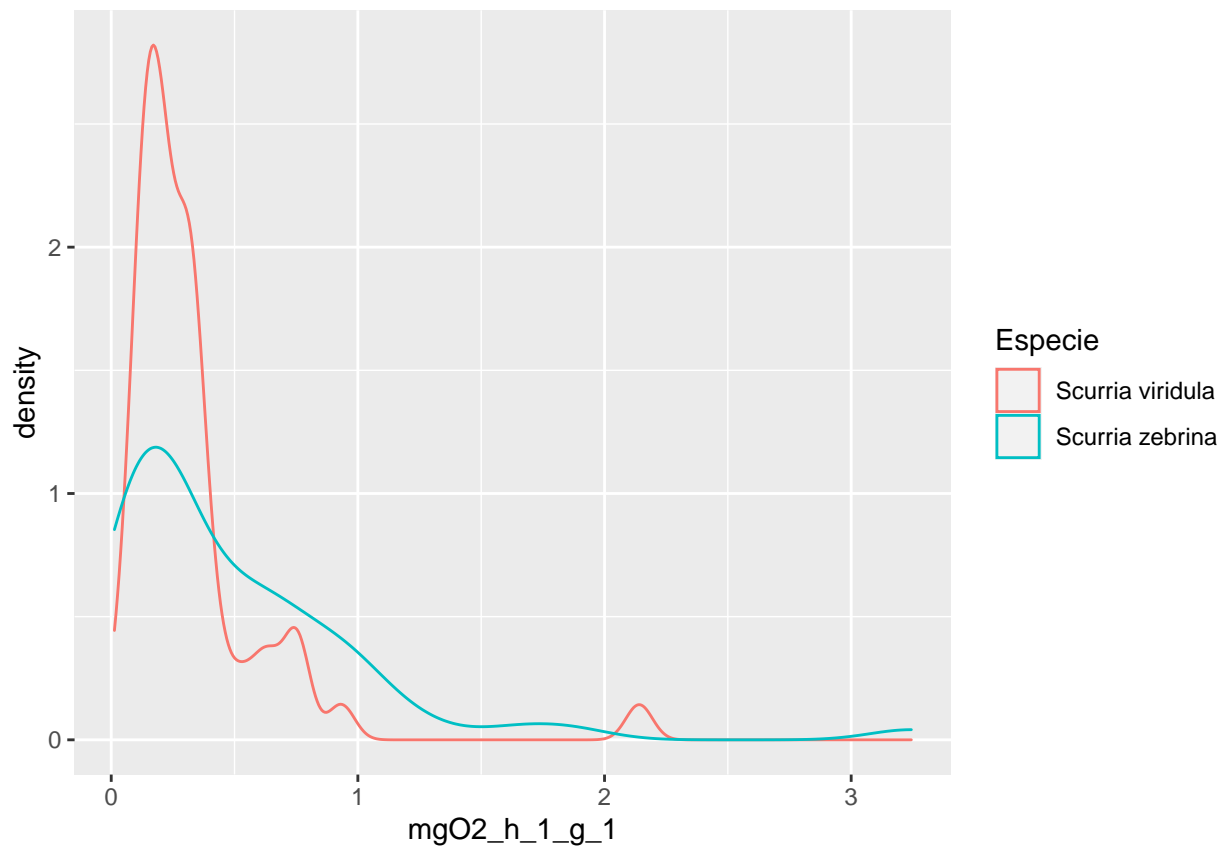


Figura 2. Densidad de probabilidad de cada especie.

```
## [1] 0.4226234
```

```
## [1] 0.4615397
```

```
## [1] 0.01306658 3.24389798
```

```
##          0%          25%          50%          75%          100%
## 0.01306658 0.14230309 0.27613837 0.57601902 3.24389798
```

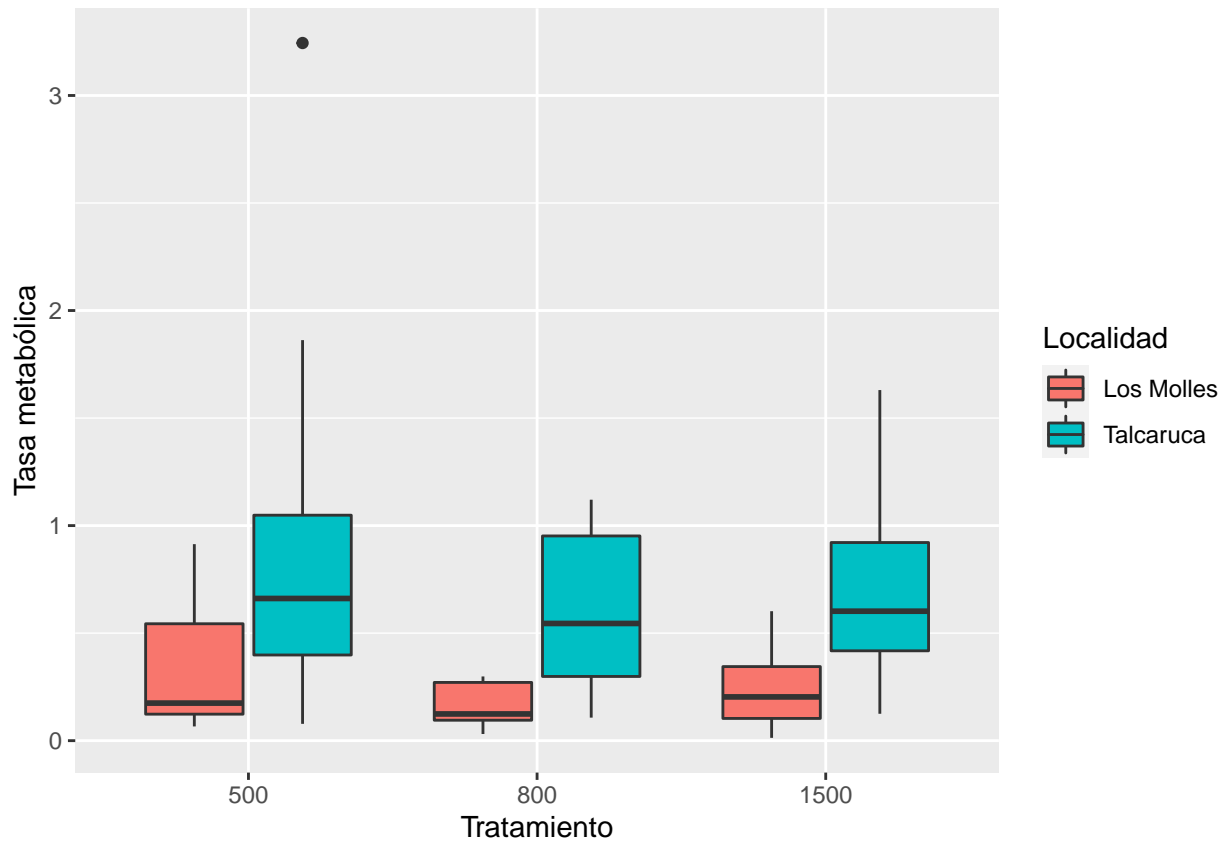


Figura 3. Tasa metabólica de *Scurria zebrina*, correspondiente a las localidades de Talcaruca y Los molles, bajo tres condiciones de acidificación.

La tasa metabólica correspondiente a *Scurria zebrina* de la localidad de Talcaruca presenta mayor variabilidad ya que la longitud de la “caja” es más grande en comparación a la localidad de Los Molles. Además la posición de la mediana en Talcaruca de los tres tratamientos indica que tienen una distribución simétrica en comparación a Los Molles. Además, la población de Los Molles presenta una asimetría negativa en los tratamientos de 500 y 800 microatmósferas.

También se observa 1 dato extremo en la localidad de Talcaruca correspondiente al tratamiento de 500 microatmósferas, y que no existe interacción de la tasa metabólica de ambas localidades cuando se sometieron en el tratamiento de 1500 microatmósferas.

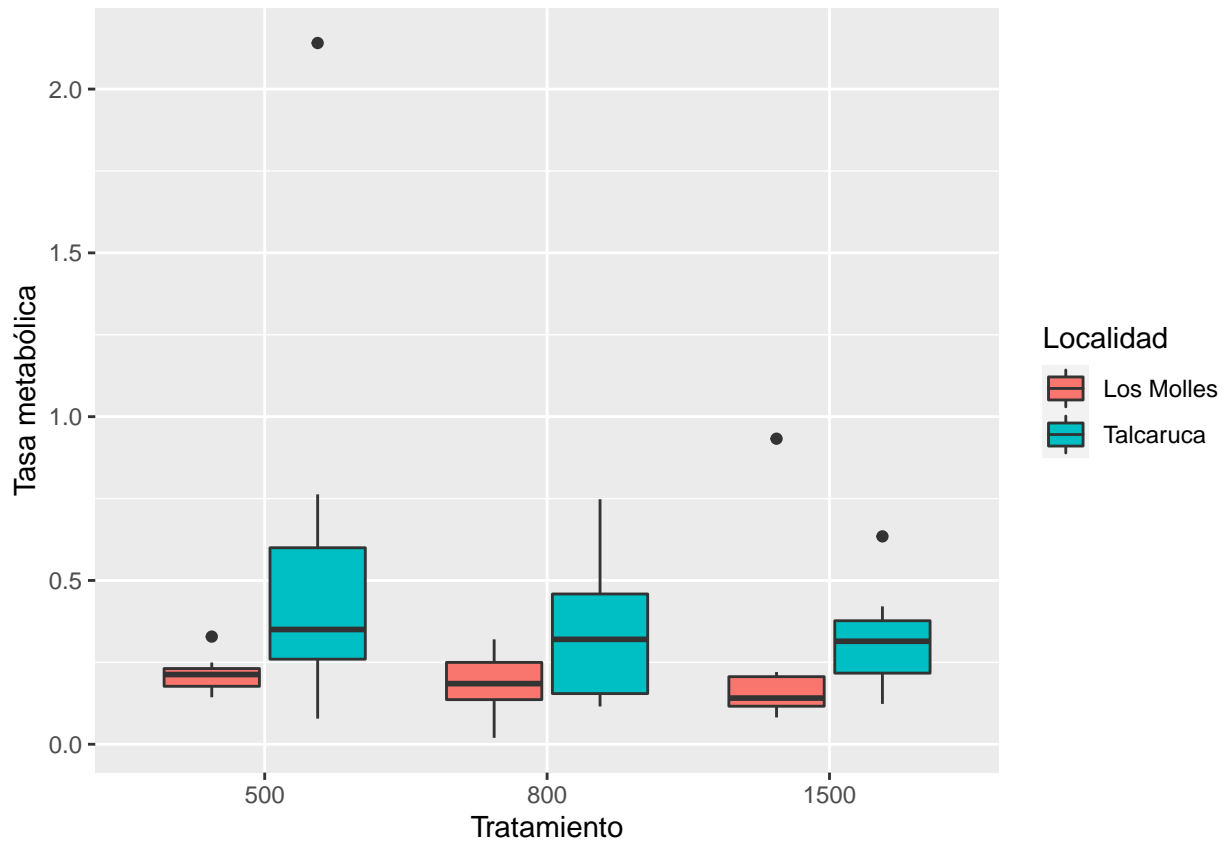
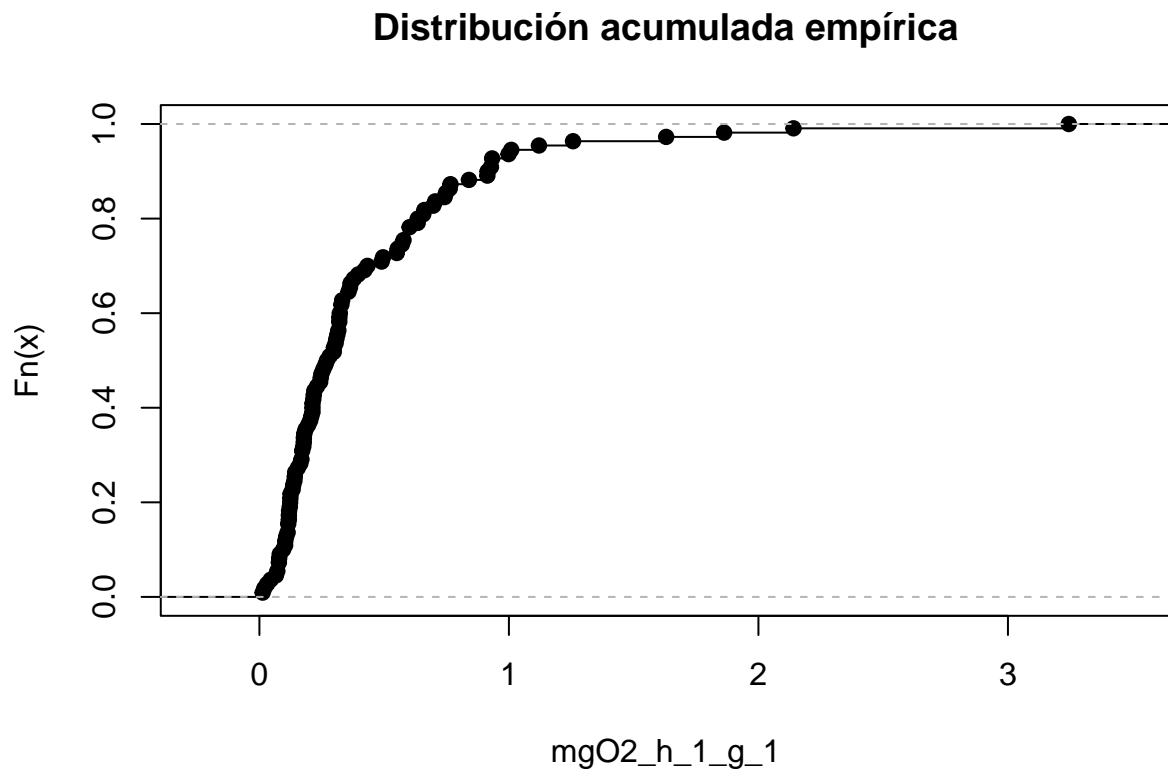


Figura 4. Tasa metabólica de *Scurria viridula* de ambas localidades, bajo tres condiciones de acidificación.

La tasa metabólica correspondiente a *Scurria viridula* de la localidad de Talcaruca presenta mayor variabilidad ya que la longitud de la “caja” es más grande en comparación a la localidad de Los Molles. Además la posición de la mediana en Talcaruca de los tres tratamientos indica que tienen una distribución simétrica en comparación a Los Molles. Se destaca que en la localidad de Los Molles tratamiento de 500, la mediana esta casi pegado al Q3, lo que se traduce en una menor variabilidad y una asimetría de los datos.

También se observa más datos extremos (4) en ambas localidades, y que solo existe correlacion de ambas localidades en el tratamiento de 800 microatmósferas (No se solapan las “cajas”) .



Figura

5. Distribución acumulada empírica de la variable respuesta, metabolismo.

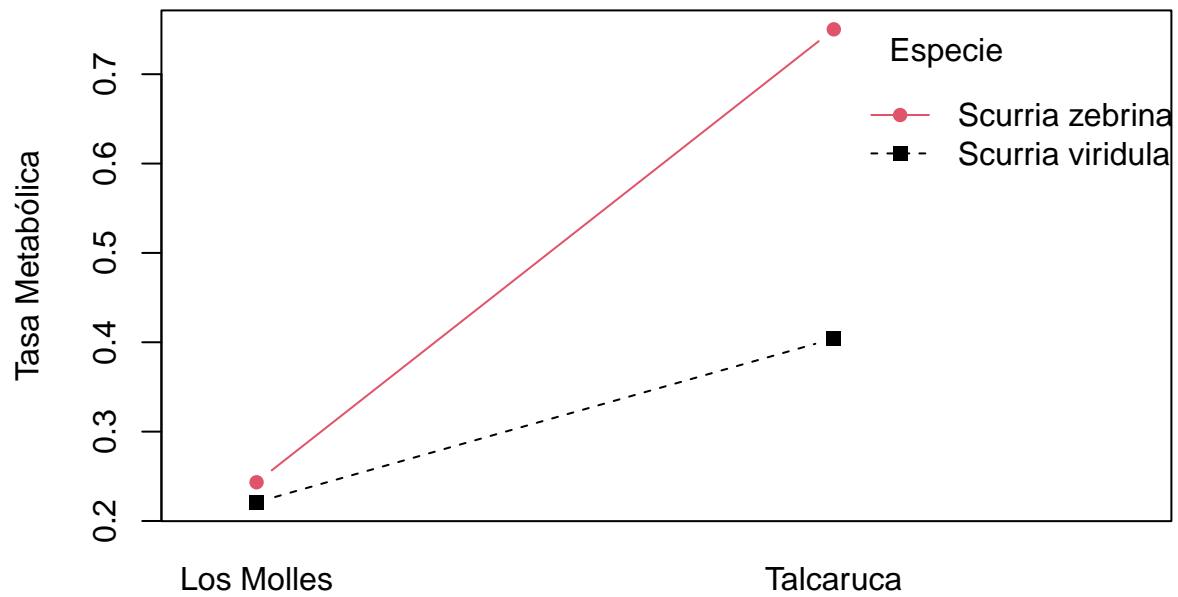
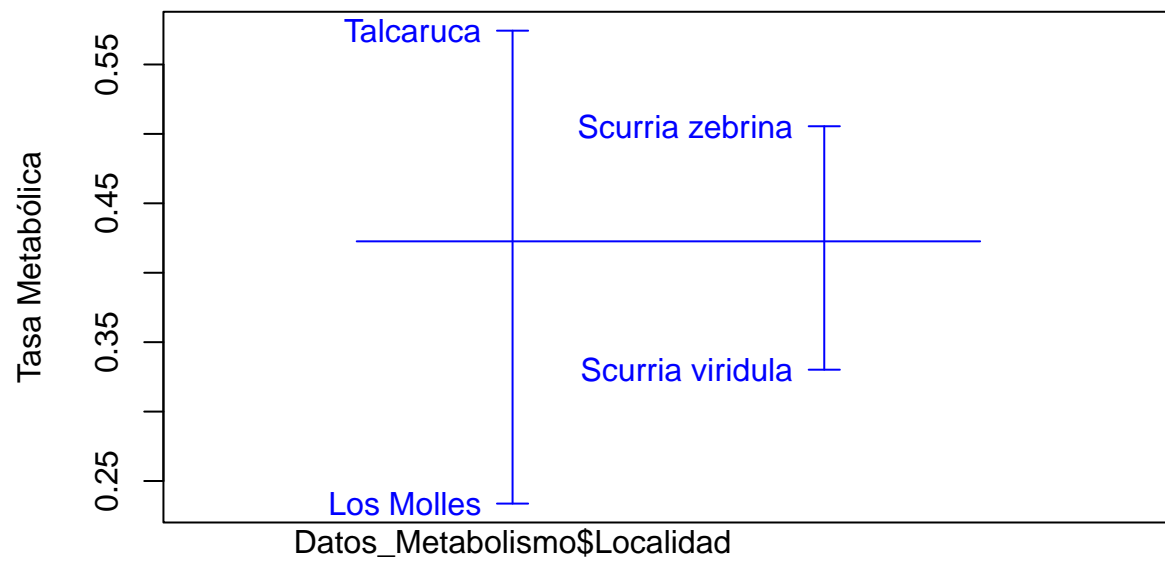


Figura 6. Correlación de ambas especies de la localidad de Talcaruca y Los Molles.

#Ambas especies muestran interacción en la localidad de los Molles



Tratamientos

Figura 7. Tasa metabólica bajo tres tratamientos de acidificación en ambas localidades.

Se observa que el factor localidad tiene mayor efecto sobre la variable respuesta.

Table 1: Media y varianza de Tasa metabólica

Localidad	Especie	media	varianza
Los Molles	Scurria viridula	0.2209761	0.0326127
Los Molles	Scurria zebrina	0.2433236	0.0458745
Talcaruca	Scurria viridula	0.4041385	0.1422985
Talcaruca	Scurria zebrina	0.7502240	0.4163771

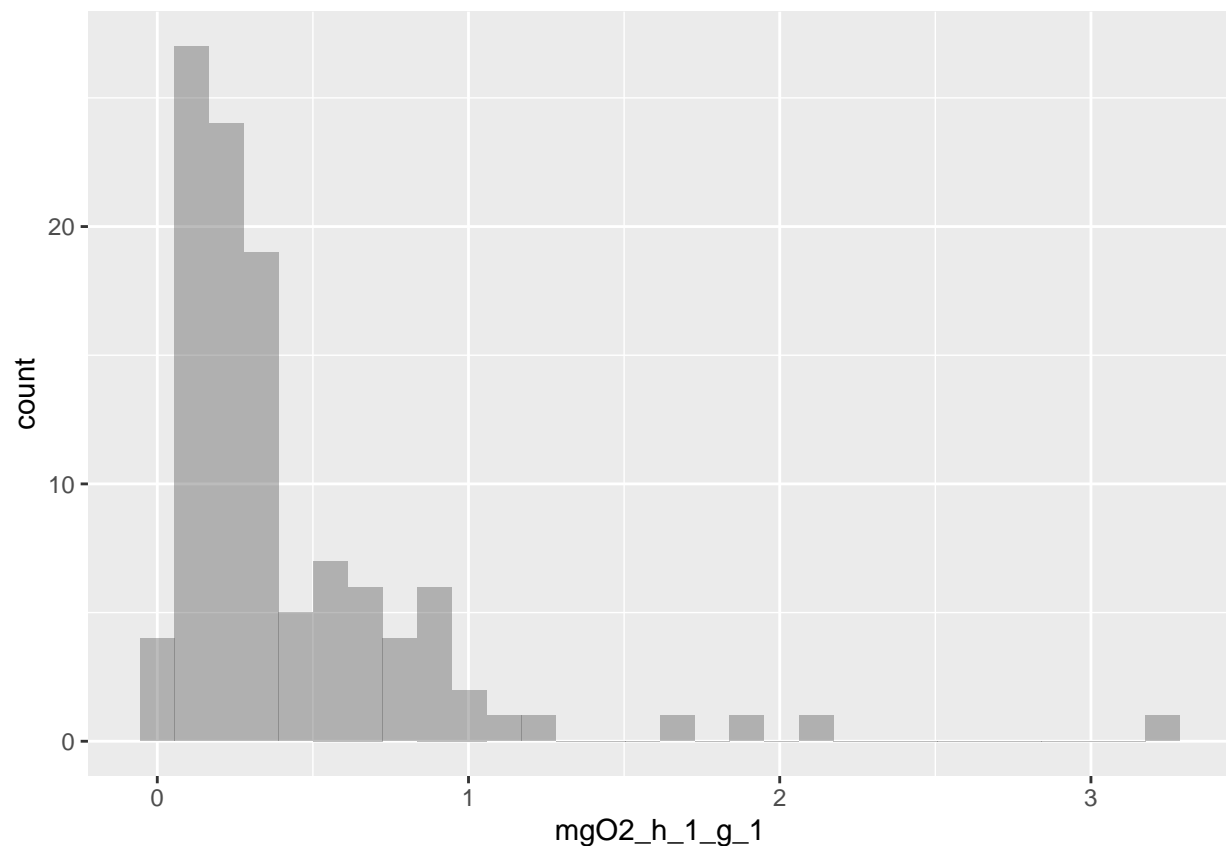
Se observa una media más alta del metabolismo en Scurria zebrina correspondiente a la localidad de Talcaruca. Sin embargo, en la localidad de Los Molles no se observan grandes diferencias.

Análisis de resultados

Debido a que los datos según el histograma no tienen una distribución normal, se debe hacer el análisis con métodos no paramétricos.

H0 : Existen diferencias en la respuesta de rasgos fisiológicos y morfológicos entre poblaciones de Scurria zebrina y scurria viridula bajo condiciones de acidificación.

H1 : No existen diferencias en la respuesta de rasgos fisiológicos y morfológicos entre poblaciones de Scurria zebrina y scurria viridula bajo condiciones de acidificación.

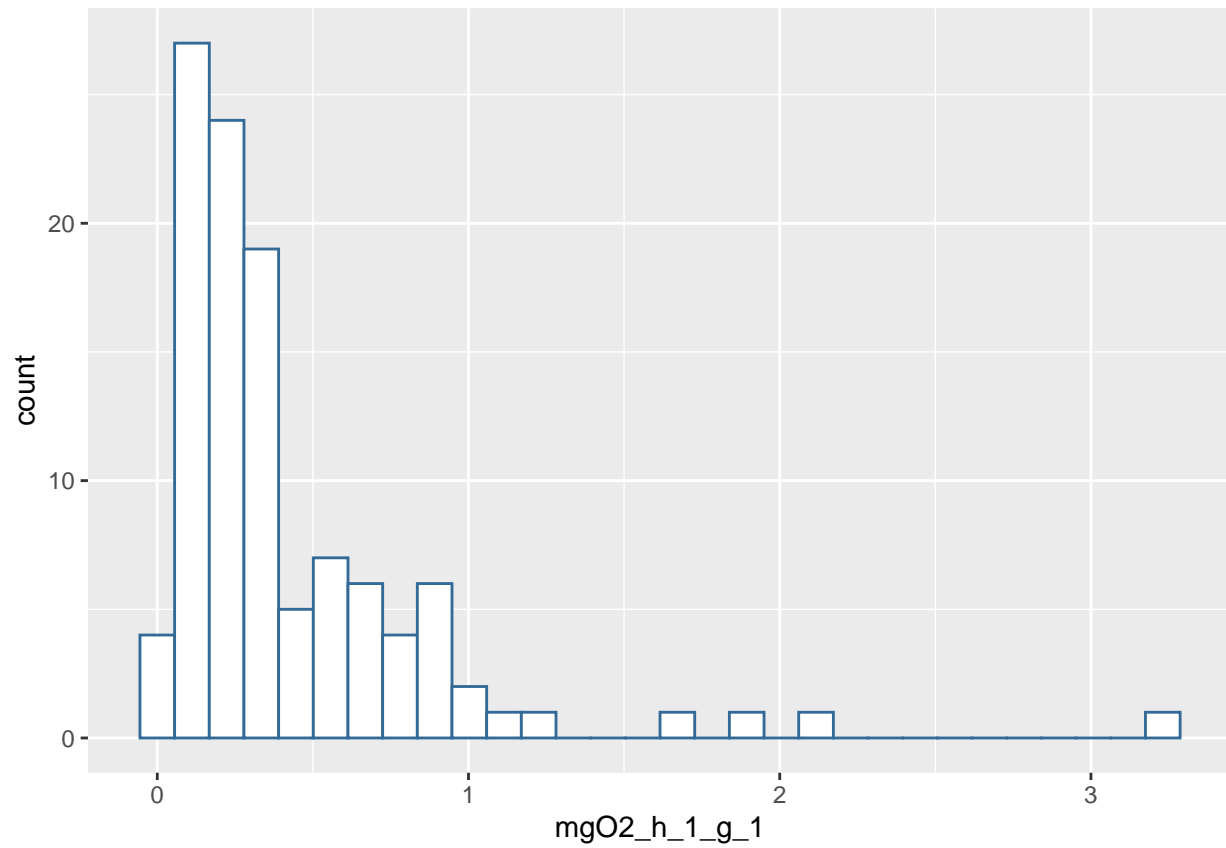


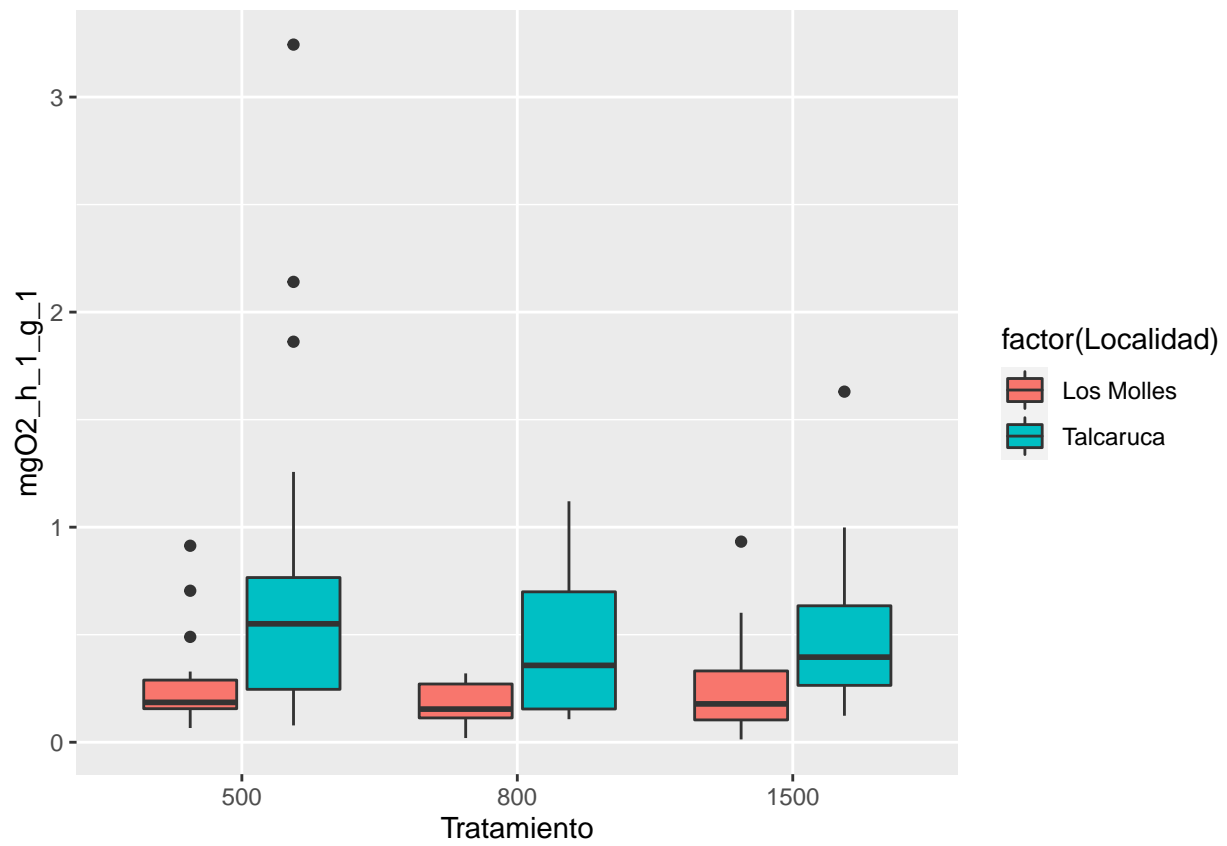
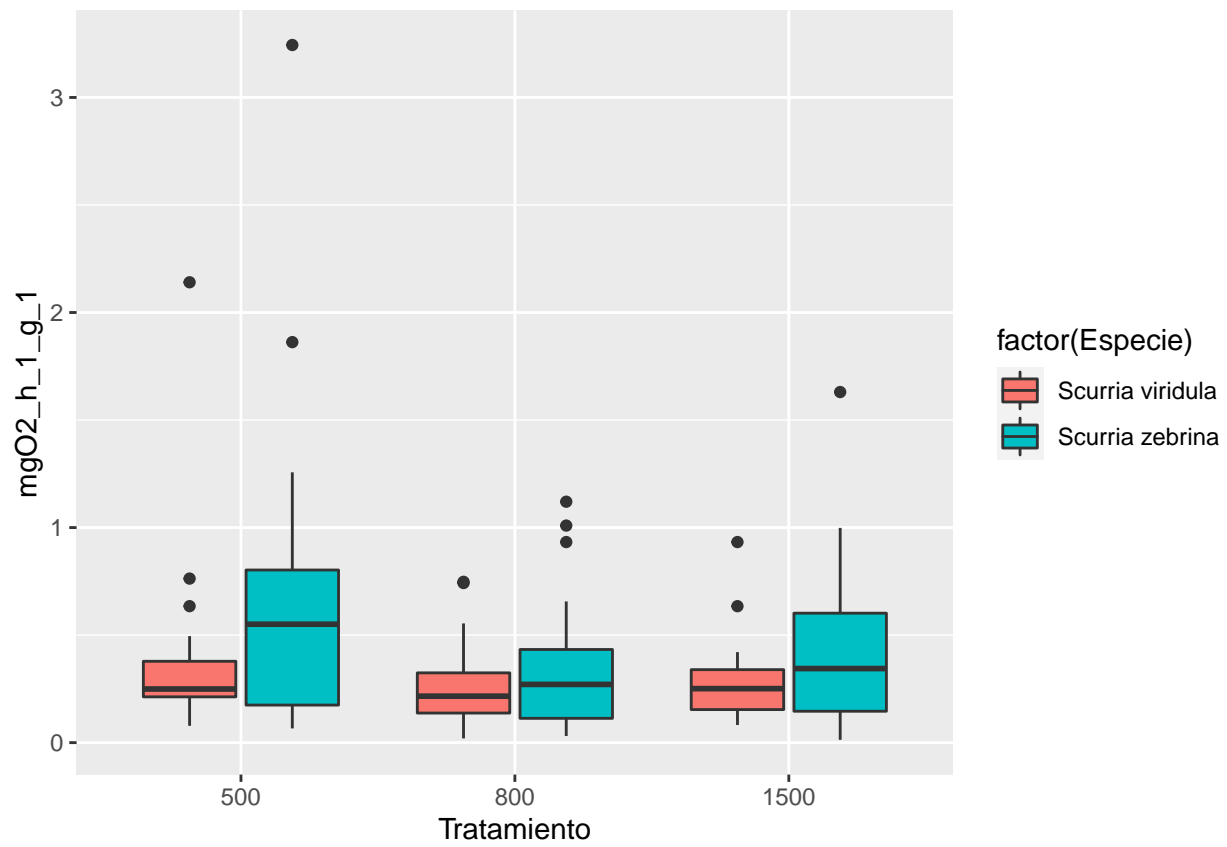
Se realizó Histograma solo con la variable respuesta (mgO2_h_1_g_1) para ver su comportamiento. Dado que su distribución no es Normal, se procede entonces, a hacer un análisis con método no paramétrico.

Modelo lineal simple

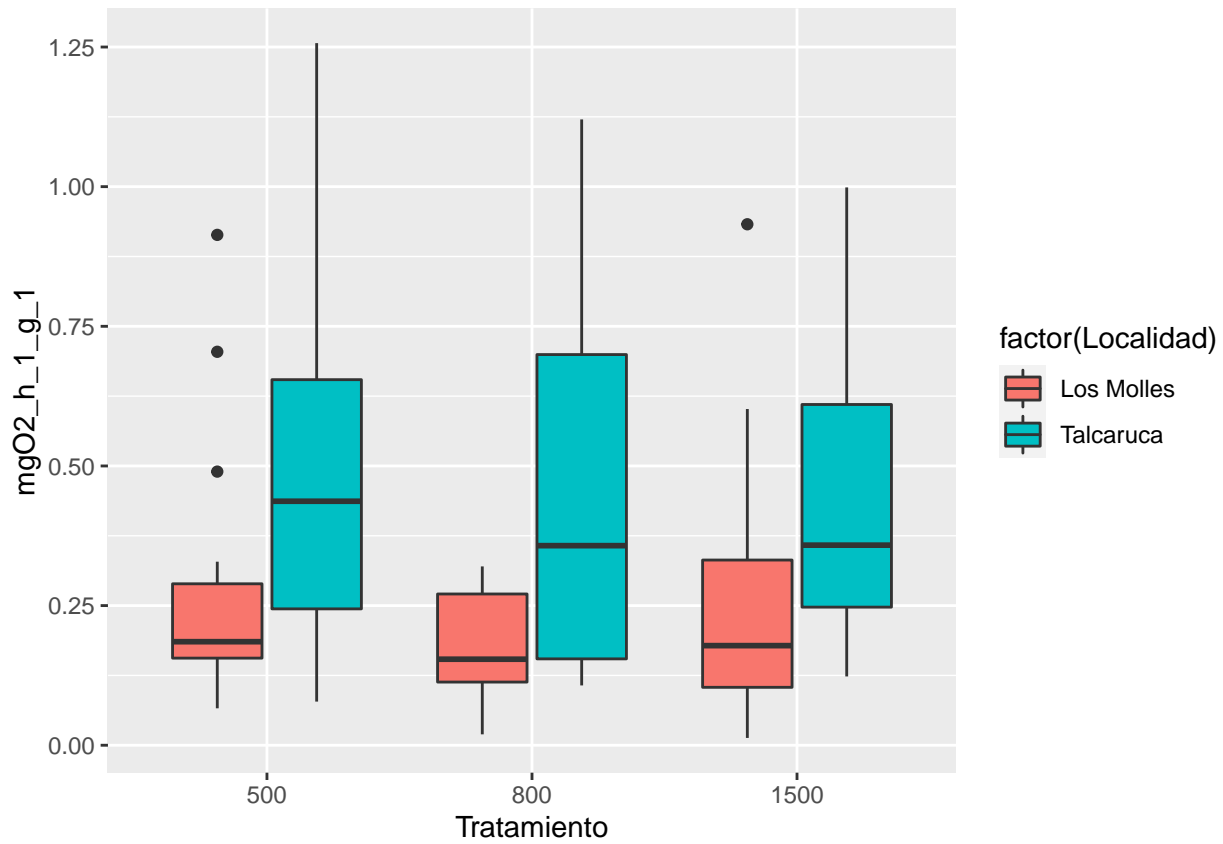
```
##
## Call:
## lm(formula = mgO2_h_1_g_1 ~ Especie + Tratamiento + Localidad,
##     data = Datos_Metabolismo)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.72092 -0.23036 -0.04495  0.08411  2.44482
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.25047    0.09508   2.634   0.0097 **
## EspecieScurria zebrina 0.19992    0.07976   2.507   0.0137 *
```

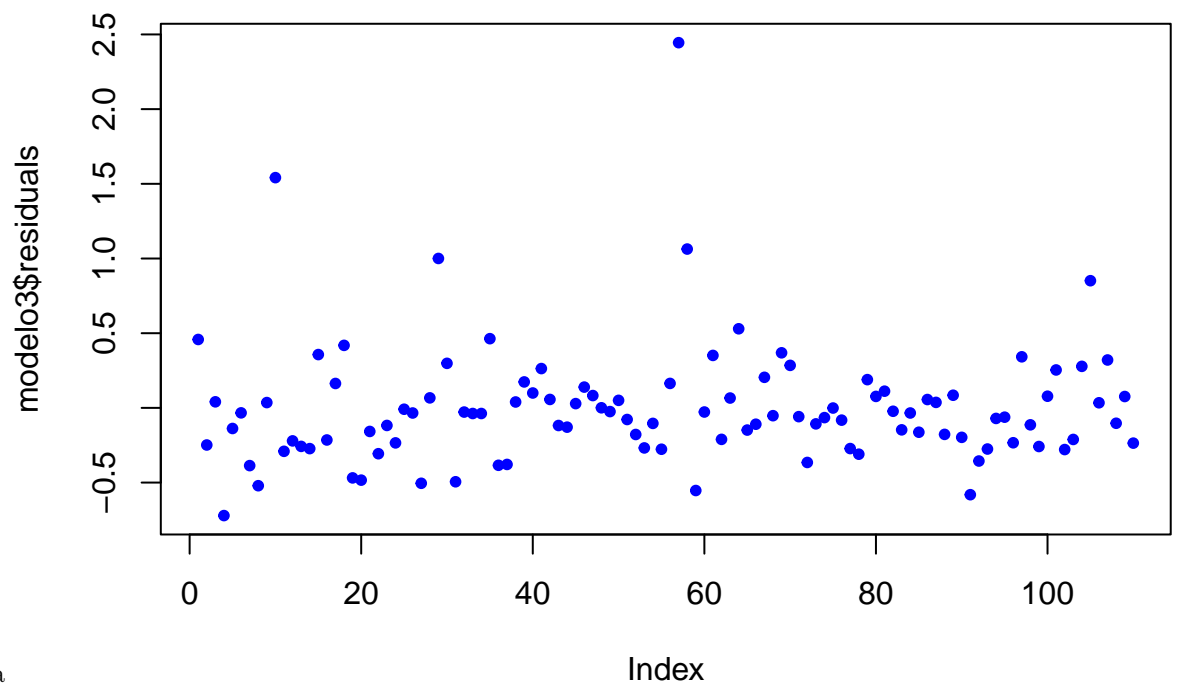
```
## Tratamiento800      -0.20828    0.09795   -2.126    0.0358 *
## Tratamiento1500     -0.16923    0.09655   -1.753    0.0826 .
## LocalidadTalcaruca   0.34868    0.07990    4.364    3e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4147 on 105 degrees of freedom
## Multiple R-squared:  0.2224, Adjusted R-squared:  0.1928
## F-statistic: 7.509 on 4 and 105 DF,  p-value: 2.327e-05
```





```
## # A tibble: 4 x 12
##   Especie Localidad Etiqueta Tratamiento `Dias Trat.` Largo Ancho `Peso Boyante`
##   <fct>   <fct>      <dbl> <fct>          <dbl> <dbl> <dbl> <fct>
## 1 Scurri~ Talcaruca      74 500             15 16.9 16.4 0.2243
## 2 Scurri~ Talcaruca      22 1500            15 15.4 11.7 0.1388
## 3 Scurri~ Talcaruca      25 500             30 15.7 14.0 0.2078
## 4 Scurri~ Talcaruca      21 500             30 14.9 14.8 0.196
## # ... with 4 more variables: Pendiente <dbl>, Tiempo <dtm>, ml_g_h 14C <dbl>,
## #   mgO2_h_1_g_1 <dbl>
```



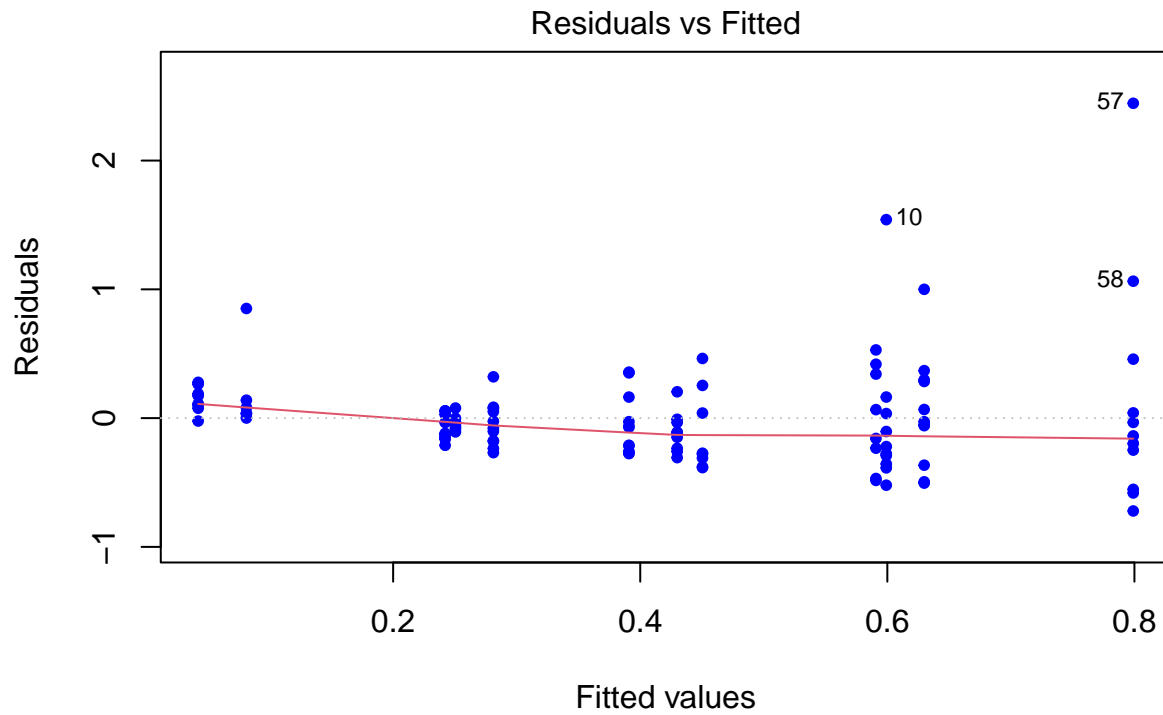


Independencia

```
##
## Durbin-Watson test
##
## data:  mg02_h_1_g_1 ~ Especie + Tratamiento + Localidad
## DW = 2.0296, p-value = 0.8873
## alternative hypothesis: true autocorrelation is not 0
```

No hay independencia

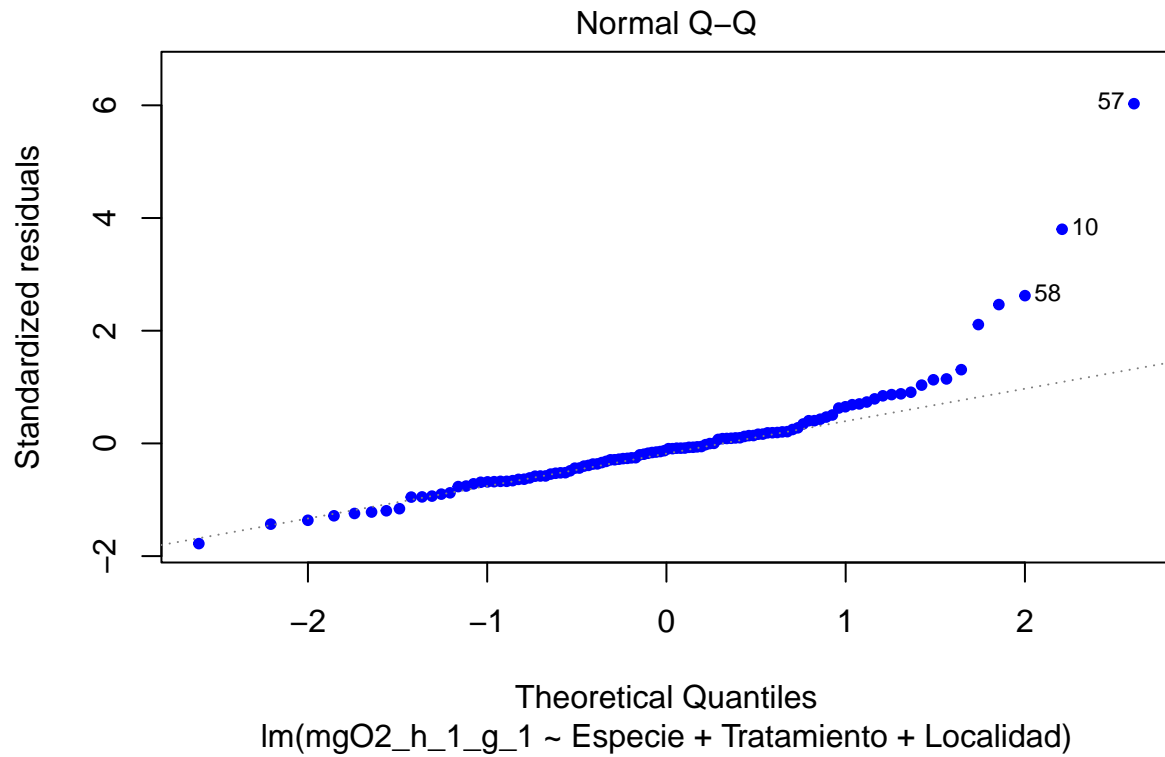
Homogeneidad de varianzas



El

modelo indica que no hay homogeneidad de varianzas, existe correlación.

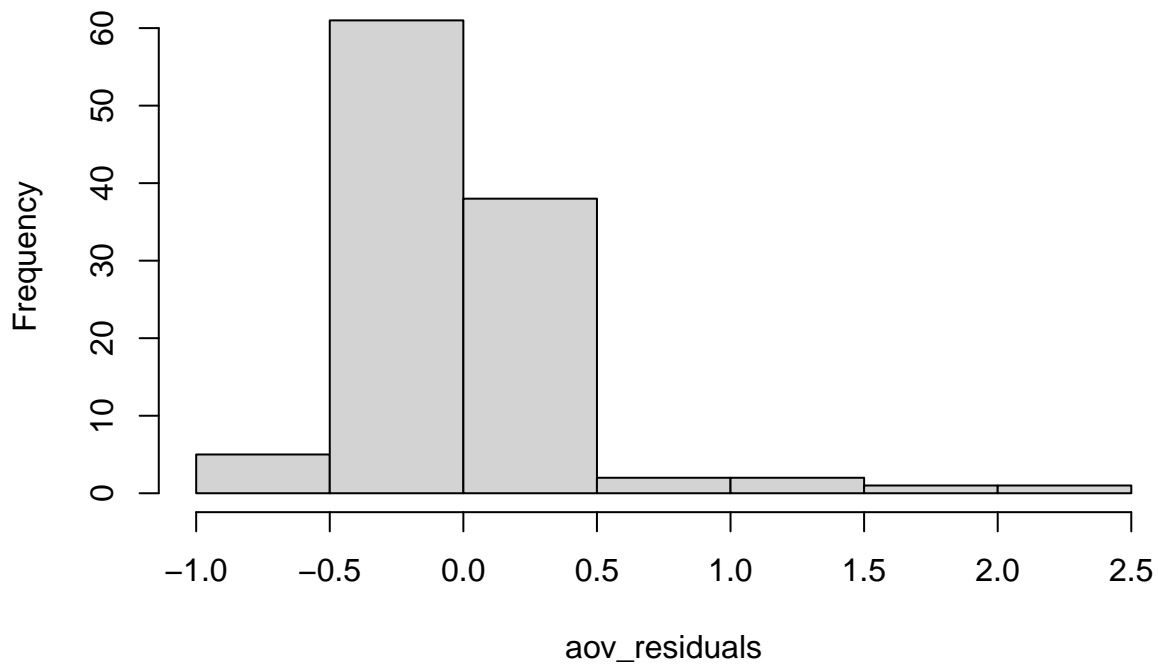
```
## Levene's Test for Homogeneity of Variance (center = "median")
##      Df F value  Pr(>F)
## group 11  2.3448 0.01339 *
##      94
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



El

modelo indica que no se cumple el modelo de normalidad

Histograma de residuales



```
##
## Shapiro-Wilk normality test
##
## data:  aov_residuales
```



```
## W = 0.78715, p-value = 2.518e-11
```

La hipótesis de rechaza debido a que el p valor es menor a 0.05.

Debido a que no se cumplen los supuestos de normalidad, homogeneidad e independencia en los residuos, se propone entonces realizar análisis usando el enfoque no paramétrico.

Enfoque No paramétrico

```
##  
## Kruskal-Wallis rank sum test  
##  
## data: mg02_h_1_g_1 by Especie  
## Kruskal-Wallis chi-squared = 0.93599, df = 1, p-value = 0.3333  
  
##  
## Kruskal-Wallis rank sum test  
##  
## data: mg02_h_1_g_1 by Localidad  
## Kruskal-Wallis chi-squared = 18.87, df = 1, p-value = 1.4e-05  
  
##  
## Kruskal-Wallis rank sum test  
##  
## data: mg02_h_1_g_1 by Tratamiento  
## Kruskal-Wallis chi-squared = 1.4956, df = 2, p-value = 0.4734
```

Se eliminaron datos extremos del análisis, sin embargo, siguen influenciando en gran magnitud sobre el análisis de enfoque no paramétrico.

Sólo la interacción entre mg de O₂ y localidad tiene significancia, ya que el valor p valor es menor a 0.05. Por lo tanto el efecto Localidad si podría ser predictivo para el análisis . Efectivamente, se esperaría que la variable localidad tuviera un efecto significativo sobre la población, ya que se está comparando un sitio (Talcaruca) que presenta condiciones de acidificación constantes con otra localidad (Los Molles) donde sus condiciones son más estables. Por otra parte, el efecto tratamiento y especie no generan diferencias significativas en este estudio.