

Valeria Chiriboga Escobar.

Curso Redbioma, Lenguaje de R.

Docente: Jonathan V. Solórzano.

Proyecto Final “Distribución Altitudinal de Tábanos del Género Boldodimyia en Siete Provincias del Ecuador”

1. Objetivo

El objetivo de este análisis es estudiar la distribución altitudinal (elevación en metros) del género Boldodimyia de la familia Tabanidae (Diptera) a lo largo de diferentes meses del año con el fin de evaluar si existen diferencias significativas en la distribución de elevaciones entre los meses analizados.

2. Metodología

Se utilizó un conjunto de datos que incluye registros de ejemplares del género Boldodimyia de la familia Tabanidae, recolectados en Ecuador en las provincias de Cañar, El Oro, Esmeraldas, Guayas, Manabí, Morona Santiago y Napo. Las variables analizadas fueron el mes del año en el que se realizaron las colectas y la elevación en metros.

Análisis estadísticos

Se utilizó ANOVA para evaluar si las elevaciones difieren significativamente entre los meses. Posteriormente se realizó una prueba de normalidad donde se evaluaron los residuos del modelo usando la prueba de Shapiro-Wilk. Se verificó la homogeneidad de varianzas con la prueba de Levene. Finalmente se hizo el análisis de prueba no paramétrica de Kruskal-Wallis.

Los gráficos fueron realizados con (*geom_point*) como gráfico de puntos para observar la distribución de elevaciones por mes, y (*geom_boxplot*) como gráfico de cajas para visualizar las distribuciones y compararlas entre meses.

3. Resultados

Se realizó un análisis ANOVA para evaluar si existen diferencias significativas en la elevación entre los meses, obteniéndose un valor de $F = 3.693$ y un p-valor de 0.0636 . Aunque el valor F indica cierta variabilidad entre los meses, el valor de p sugiere que las diferencias no son estadísticamente significativas al nivel de confianza establecido. Sin embargo, la Prueba de Shapiro-Wilk mostró que los residuos no siguen una distribución normal ($p = 0.04982$), y la Prueba de Levene reveló que las varianzas no son homogéneas ($p = 0.02542$).

Dado que no se cumplen los supuestos de normalidad y homogeneidad de varianzas para el ANOVA, se utilizó la prueba no paramétrica de Kruskal-Wallis, que arrojó un p-valor de 0.08818 . Esto indica que no hay evidencia suficiente para afirmar la existencia de diferencias estadísticamente significativas en la elevación entre los meses.

Figure 1 Gráfico de puntos de distribución de la elevación por mes

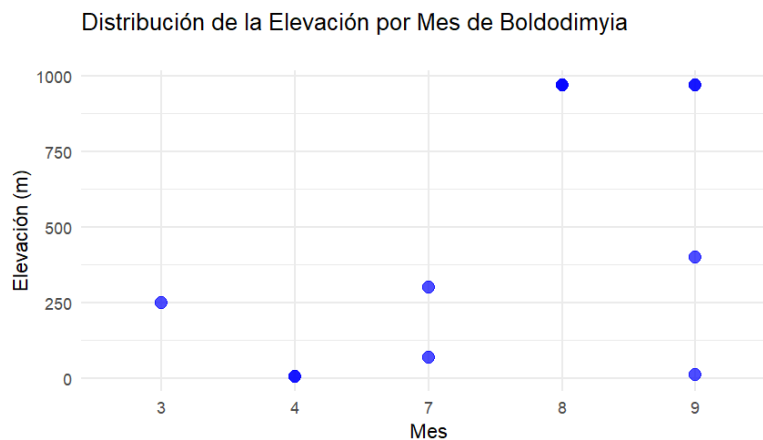
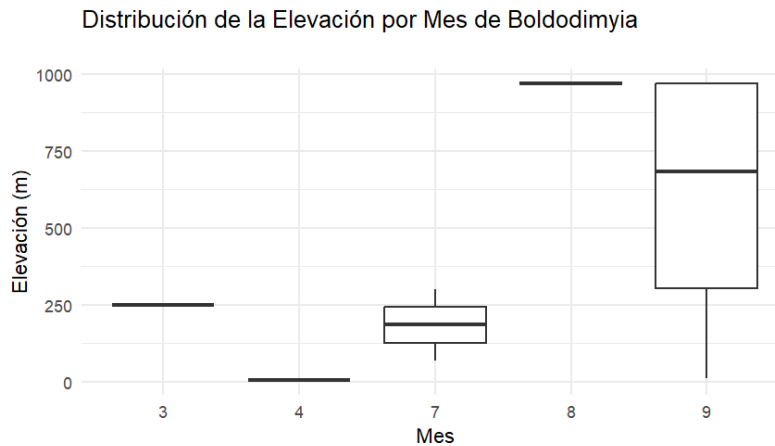


Figure 2 Gráfico de cajas de distribución de la elevación por mes



Ambos gráficos resaltan que, en la época seca del año, en los meses Agosto y septiembre presentan elevaciones mayores en comparación con otros meses, mientras que, en el mes Marzo (época lluviosa) y Julio (época seca) tienen registros a elevaciones más bajas.

4. Discusión

Los resultados obtenidos muestran el género Boldodimyia puede tener patrones amplios en la distribución en elevación a lo largo de diferentes meses, que van desde los 5msnm hasta los 1000msnm. Se observa que las elevaciones mayores tienden a registrarse durante la época seca del año, en los meses Agosto y septiembre, mientras que las menores se concentran en marzo (época lluviosa) y Julio (época seca), lo que podría reflejar adaptaciones ecológicas del género a las condiciones ambientales estacionales.

Las provincias que presentan registros de elevaciones más altas incluyen Cañar y Napo, mientras que las provincias con registros a menores elevaciones, como Manabí y El Oro, podrían estar influenciadas por la topografía y clima locales. Aunque no se encontraron diferencias estadísticamente significativas entre los meses, estos patrones sugieren posibles tendencias que podrían explorarse con un mayor tamaño de muestra o análisis más detallados.

5. Anexos

Script de R

```
> library(readxl)
Aviso:
package 'readxl' was built under R version 4.4.2
> ruta_excel <- "C://Users//valec//Documents//Tabanide Data.xlsx"
> (ruta_excel)
[1] "C://Users//valec//Documents//Tabanide Data.xlsx"
> Tabanidae <- read_excel(ruta_excel)
> View(Tabanidae)
> str(Tabanidae)
tibble [12 × 29] (S3: tbl_df/tbl/data.frame)
 $ kingdom      : chr [1:12] "Animalia" "Animalia" "Animalia" "Animalia" ...
 $ phylum     : chr [1:12] "Arthropoda" "Arthropoda" "Arthropoda" "Arthropoda" ...
 $ class        : chr [1:12] "Insecta" "Insecta" "Insecta" "Insecta" ...
 $ order        : chr [1:12] "Diptera" "Diptera" "Diptera" "Diptera" ...
 $ family       : chr [1:12] "Tabanidae" "Tabanidae" "Tabanidae" "Tabanidae" ...
 $ scientificName : chr [1:12] "Boldodimyia bicolor" "Boldodimyia bicolor"
"Boldodimyia celeroides" "Boldodimyia celeroides" ...
 $ taxonID      : logi [1:12] NA NA NA NA NA NA NA ...
 $ scientificNameAuthorship: chr [1:12] "Bigot 1892" "Bigot 1892" "Stone 1954" "Stone
1954" ...
 $ genus        : chr [1:12] "Boldodimyia" "Boldodimyia" "Boldodimyia" "Boldodimyia"
...
 $ subgenus     : chr [1:12] "bicolor" "bicolor" "celeroides" "celeroides" ...
 $ taxonRank    : chr [1:12] "especie" "especie" "especie" "especie" ...
 $ identifiedBy : chr [1:12] "Jaime Buestan" "Jaime Buestan" "Jaime Buestan" "Jaime
Buestan" ...
 $ recordedBy   : chr [1:12] "J. Buestan" "J. Buestan" "J. Buestan" "J. Buestan" ...
```

```

$ eventDate      : chr [1:12] "20/04/2006" "20/04/2006" "01/09/12" "01/09/16" ...
$ year          : num [1:12] 2006 2006 2012 2016 2012 ...
$ month         : num [1:12] 4 4 9 9 8 3 7 9 9 8 ...
$ day          : num [1:12] 20 20 1 1 1 29 30 1 1 1 ...
$ lifeStage     : chr [1:12] "Adulto" "Adulto" "Adulto" "Adulto" ...
$ individualCount : chr [1:12] "Trampa Malaise" "Trampa Malaise" "Trampa Malaise"
"Trampa Malaise" ...
$ preparations  : chr [1:12] "Montado en alfiler" "Montado en alfiler" "Montado en
alfiler" "Montado en alfiler" ...
$ country       : chr [1:12] "Ecuador" "Ecuador" "Ecuador" "Ecuador" ...
$ stateProvince : chr [1:12] "Manabi" "Manabi" "Cañar" "Morona Santiago" ...
$ municipality  : chr [1:12] NA NA NA NA NA ...
$ locality      : chr [1:12] "Rio Mache" "Rio Mache" "Sorocoche" "Milagro" ...
$ locationRemarks : logi [1:12] NA NA NA NA NA NA NA ...
$ decimalLatitude : num [1:12] 0.215 0.215 NA -2.131 NA ...
$ decimalLongitude : num [1:12] -79.9 -79.9 NA -79.6 NA ...
$ geodeticDatum  : chr [1:12] "WGS84" "WGS84" "WGS84" "WGS84" ...
$ verbatimElevation : num [1:12] 5 5 970 13 970 250 70 400 970 970 ...

```

```
> datos$month <- as.factor(datos$month)
```

Error in `\$<-`:

! Assigned data `as.factor(datos\$month)` must be compatible with existing data.

✘ Existing data has 6 rows.

✘ Assigned data has 0 rows.

i Only vectors of size 1 are recycled.

Caused by error in `vectbl_recycle_rhs_rows()`:

! Can't recycle input of size 0 to size 6.

Run `rlang::last_trace()` to see where the error occurred.

Aviso:

Unknown or uninitialised column: `month`.

```
> Tabanidae$month <- as.factor(Tabanidae$month)
```

```
> modelo_anova <- aov(elevation ~ month, Tabanidae = Tabanidae)
```

Error en eval(predvars, data, env): objeto 'elevation' no encontrado

```
>
```

```
> modelo_anova <- aov(verbatimElevation ~ month, Tabanidae = Tabanidae)
```

Error en eval(predvars, data, env):

objeto 'verbatimElevation' no encontrado

```
> modelo_anova <- aov(verbatimElevation ~ month, Tabanidae = Tabanidae)
```

Error en eval(predvars, data, env):

objeto 'verbatimElevation' no encontrado

```
> modelo_anova <- aov(verbatimElevation ~ month, data = Tabanidae)
```

```
> summary(modelo_anova)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
month	4	1443898	360975	3.693	0.0636 .
Residuals	7	684267	97752		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```
>
```

```
> shapiro.test(residuos)
```

Error: objeto 'residuos' no encontrado

```
> residuos <- residuals(modelo_anova)
```

```
> shapiro.test(residuos)
```

Shapiro-Wilk normality test

data: residuos

W = 0.86065, p-value = 0.04982

```
> library(car)
```

```
Cargando paquete requerido: carData
```

```
Avisos:
```

```
1: package 'car' was built under R version 4.4.2
```

```
2: package 'carData' was built under R version 4.4.2
```

```
> leveneTest(verbatimElevation ~ month, data = Tabanidae)
```

```
Levene's Test for Homogeneity of Variance (center = median)
```

```
    Df F value Pr(>F)
```

```
group 4  5.486 0.02542 *
```

```
    7
```

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> kruskal.test(elevation ~ month, data = Tabanidae)
```

```
Error en eval(predvars, data, env): objeto 'elevation' no encontrado
```

```
> kruskal.test(verbatimElevation ~ month, data = Tabanidae)
```

```
Kruskal-Wallis rank sum test
```

```
data: verbatimElevation by month
```

```
Kruskal-Wallis chi-squared = 8.0943, df = 4, p-value = 0.08818
```

```
> library(ggplot2)
```

```
> ggplot(Tabanidae, aes(x = month, y = verbatimElevation)) +
```

```
+ geom_boxplot() +
```

```
+ labs(title = "Distribución de la Elevación por Mes",
```

```
+   x = "Mes",
```

```
+   y = "Elevación (m)") +
```

```
+ theme_minimal()
```

```
> ggplot(Tabanidae, aes(x = month, y = verbatimElevation)) +
```

```
+ geom_point(color = "blue", size = 3, alpha = 0.7) +
+ labs(title = "Distribución de la Elevación por Mes",
+       x = "Mes",
+       y = "Elevación (m)") +
+ theme_minimal()
> ggplot(Tabanidae, aes(x = month, y = verbatimElevation)) +
+ geom_boxplot() +
+ labs(title = "Distribución de la Elevación por Mes de Boldodimya
+ ",
+       x = "Mes",
+       y = "Elevación (m)") +
+ theme_minimal()
> ggplot(Tabanidae, aes(x = month, y = verbatimElevation)) +
+ geom_point(color = "blue", size = 3, alpha = 0.7) +
+ labs(title = "Distribución de la Elevación por Mes de Boldodimya
+ ",
+       x = "Mes",
+       y = "Elevación (m)") +
+ theme_minimal()
>
```