Title: “Circling back: *Vespula* spp. reinvasion after knockout from volcanic ash fall in Patagonia”

Subtitle: "Complete Statistical Analysis"

Authors:

- Ana Julia Pereira

- Juan C. Corley

- Gustavo N. Gimenez

- Maité Masciocchi

date: "`r format(Sys.time(), '%d %B %Y')`"

output:

 html\_document:

 toc: true

 toc\_depth: 3

 number\_sections: true

 theme: united

 highlight: tango

knitr::opts\_chunk$set(echo = TRUE)

The attached file \*\*Datos\_Vespula.csv\*\* has wasp abundance data for two species:

- *Vespula germanica* (VG)

- *Vespula vulgaris* (VV).

Both are invasive species that are being targeted for eradication/control in the Neuquén and Río Negro mountain ranges. When ash fell from the Puyehue volcano, most of the arthropod fauna was affected, leading to a natural defaunation event in many areas. The following year, we began sampling the same sites annually for three years to study the reinvasion of these species and to determine whether one of the two is better at colonization. Based on the literature, we hypothesized that VV is a better invader. However, the models do not indicate significant differences. You will see that we use the amount of ash at the site (thickness in cm), the distance to the nearest urban center (in km), and the species (VV or VG) as predictor variables.

The data correspond to the study of the species *Vespula germanica* and *Vespula* *vulgaris* during the times of the volcanic eruption. The variable of interest is abundance, and the species are listed in the columns labeled "species."

```{r datos\_vespula}

#vespula = read.csv("Datos\_vespula1.csv", dec = ",")

vespula = read.csv("datos\_anadic12.csv", dec = ".", sep = "\t")

head(vespula)

```

The records for each site include georeferenced coordinates and also measure the distance from the urban center

## Summary of *Vespula* Data

Regarding abundance, there are records ranging from 0 to 142 captures at a single site. The median is 0, indicating that at least 50% of the data are zeros.

summary(vespula)

## Another way to summarize

library(tidyverse)

glimpse(vespula)

The response variable in this case is Abundance, and its behavior is being explored

source("/home/gustavo/Documentos/R/Funciones\_Mias/ggcont\_prop0.R")

Abundance = vespula$Abundance

#ggsave("prop0\_vespula.png")

ggcont\_prop0(Abundance)

#dev.off()

A notable feature is that a very high proportion of the data consists of zeros (more than 50%).

barplot(table(vespula$Abundance))

There is a high frequency of zeros, and overdispersion of counts is evident.

Abundance is presented in a boxplot removing the zeros.

source("/home/gustavo/Documentos/R/Funciones\_Mias/ggboxdot.R")

vespula |> filter(Abundance > 0) |> ggboxdots(Species, Abundance)

There is greater dispersion in the abundance of the VG species when excluding zeros.

Since abundance is a count and therefore an integer value, we can count the number of values the variable takes.

tabla\_vespula = vespula |> group\_by(Abundance) |> count()

#ggsave("barplot\_vespula.png")

vespula |> ggplot(aes(x = Abundance)) + geom\_bar(width = 2) +

 xlab("Abundance") + ylab("Frequency") + theme\_bw()

#dev.off()

The zero is observed with a frequency exceeding 300.

#ggsave("Incidencias2\_vespula.png")

ggplot(vespula, aes(Abundance, fill = Species))+geom\_histogram(binwidth = 1.2)+

 facet\_grid(Year ~ Species, margins = TRUE, scales = "free") + theme\_bw() + xlab("Abundance") + ylab("Frequency") +

 theme(legend.position="none")

#dev.off()

No significant differences are detected in terms of frequency between VG and VV. It appears that there may be differences in captures between the years.

Next, the abundance (total captures) at different sites and by species is explored.

ggplot(vespula, aes(Site, Abundance, fill = Species))+

 geom\_col() + theme\_bw() + xlab("Abundance") + ylab("Frequency")

Descriptively, higher captures are observed at some sites such as C2-7, C2-9, and C2-8. In C2-5, VG predominates over VV.

**Captures by species and site**

Using a Lollipop chart (remember these are total captures per site).

vesp\_tot\_sitl = vespula |> group\_by(Site, Species) |>

 summarise(Total = sum(Abundance)) |> arrange(desc(Total))

 vesp\_tot\_sitlt = vesp\_tot\_sitl |> pivot\_wider(names\_from = Species, values\_from = Total)

 tabfin = vesp\_tot\_sitlt |>

 rowwise() |>

 mutate(suma\_esp = sum(c(`V. germanica`, `V. vulgaris`) )) |>

 arrange(suma\_esp) |>

 mutate(Site = factor(Site, Site)) |> print(n = 23)

library(hrbrthemes)

# Plot

ggplot(tabfin) +

 geom\_segment( aes(x=Site, xend=Site, y = `V. germanica`, yend = `V. vulgaris`), color="grey") +

 geom\_point( aes(x=Site, y=`V. germanica`), color=rgb(0.2,0.7,0.1,0.5), size=3 ) +

 geom\_point( aes(x=Site, y=`V. vulgaris`), color=rgb(0.7,0.2,0.1,0.5), size=3 ) +

 theme\_ipsum() +

 xlab("Site") +

 ylab("Abundance") +

 scale\_fill\_manual(name='Species',

 breaks=c('`V. germanica`', '`V. vulgaris`'),

 values=c('`V. germanica`'=rgb(0.2,0.7,0.1,0.5), '`V. vulgaris`'=rgb(0.7,0.2,0.1,0.5))) +

 theme(legend.title=element\_text(size=20),

 legend.text=element\_text(size=14))+

 theme( legend.position = "top") +

 coord\_flip() +

 annotate("text", x=6, y=150, label= "bolditalic(`V. vulgaris`)",

 col=rgb(0.7,0.2,0.1,0.5), size=3, parse=TRUE)+

 annotate("text", x=6, y=100, label= "bolditalic(`V. germanica`)",

 col=rgb(0.2,0.7,0.1,0.5), size=3, parse=TRUE)

In the graph, sites are observed according to abundance, where red dots indicate captures of *Vespula vulgaris* and green dots indicate captures of *Vespula germanica*. The sites are not ordered geographically but by capture frequency, from highest to lowest. The gray segment shows the difference between both species for each site. In the case of C2-9, it's the site with the highest number of captures, with *Vespula vulgaris* being predominant, while in C2-5, *Vespula germanica* stands out, showing a significant difference compared to the other species.

**Captures by species and site per year using a Lollipop chart**

In the summary, it's noted that these are total captures per site.

library(hrbrthemes)

# Plot

ggplot(tabfin) +

 geom\_segment( aes(x=Site, xend=Site, y=`V. germanica`, yend=`V. vulgaris`), color="cyan3") +

 geom\_point( aes(x=Site, y=`V. germanica`, color="V. germanica"), size = 2.5) +

 geom\_point( aes(x=Site, y=`V. vulgaris`, color="V. vulgaris"), size = 2.5, alpha = 0.6) +

 theme\_bw()+

 xlab("Site") +

 ylab("Abundance") +

 scale\_color\_manual(name='Species',

 guide = guide\_legend(),

 values=c('V. germanica'="#E7B800",

 'V. vulgaris'="gray6")) +

 theme(legend.title=element\_text(size=20),

 legend.text=element\_text(size=14,

 face = "italic"),

 axis.text.y = element\_text(size=10),

 axis.text.x = element\_text(size=10),

 panel.background = element\_rect(fill = 'white', color = 'white'))+

 theme( legend.position = "top",

 legend.title = element\_text(size=12),

 legend.text = element\_text(size=12),

 plot.background = element\_rect(fill = "white", color = "white")) +

 coord\_flip() + facet\_wrap(~Year)

ggsave("Fig\_1b.tiff", device = "tiff",

 dpi = 600)

dev.off()

In the graph, it's evident that there is a greater number of captures in general in 2014.

**Lollipop chart with sites ordered by distance**

In the summary, it's noted that these are total captures per site.

vesp\_tot\_dist = vespula |> group\_by(Site,Dist\_Urb, Species, Year) |>

 summarise(Total\_Captures = sum(Abundance)) |> arrange(desc(Total\_Captures))

vesp\_tot\_dist2 = vesp\_tot\_dist |> pivot\_wider(names\_from = Species, values\_from = Total\_Captures)

 tabfin2 = vesp\_tot\_dist2 |>

 rowwise() |>

 mutate(sum\_species = sum(c(`V. germanica`,`V. vulgaris`) )) |>

 arrange(Dist\_Urb) |>

 mutate(Site = factor(Site, Site)) |> print(n = 23)

library(hrbrthemes)

# Plot

ggplot(tabfin2) +

 geom\_segment( aes(x=Site, xend=Site, y=`V. germanica`, yend=`V. vulgaris`), color="grey") +

 geom\_point( aes(x=Site, y=`V. germanica`), color=rgb(0.2,0.7,0.1,0.5), size=3 ) +

 geom\_point( aes(x=Site, y=`V. vulgaris`), color=rgb(0.7,0.2,0.1,0.5), size=3 ) +

 theme\_ipsum() +

 xlab("Site") +

 ylab("Abundance") +

 scale\_fill\_manual(name='Species',

 breaks=c('`V. germanica`', '`V. vulgaris`'),

 values=c('`V. germanica`'=rgb(0.2,0.7,0.1,0.5), '`V. vulgaris`'=rgb(0.7,0.2,0.1,0.5))) +

 theme(legend.title=element\_text(size=20),

 legend.text=element\_text(size=14))+

 theme(legend.position = "top") +

 coord\_flip() +

 annotate("text", x=17, y=180, label= "bolditalic(`V. vulgaris`)",

 col=rgb(0.7,0.2,0.1,0.5), size=2.5, parse=TRUE)+

 annotate("text", x=17, y=70, label= "bolditalic(`V. germanica`)",

 col=rgb(0.2,0.7,0.1,0.5), size=2.5, parse=TRUE) + facet\_wrap(~Year)

**Total captures per species according to latitude and longitude**

vesp\_tot\_sit = vespula |> group\_by(LAT, LONG, Site, Species) |> summarise(Total\_Captures = sum(Abundance)) |> arrange(desc(Total\_Captures))

ggplot(vesp\_tot\_sit, aes(x = LAT, y = LONG, col = Species)) +

 geom\_point(aes(size = Total\_Captures)) + theme\_bw()

vesp\_tot\_sit\_anio = vespula |>

 group\_by(LAT, LONG, Site, Species, Year) |>

 summarise(Total\_Captures = sum(Abundance)) |>

 arrange(desc(Total\_Captures))

ggplot(vesp\_tot\_sit\_anio, aes(x = LAT, y = LONG, col = Species)) +

 geom\_point(aes(size = Total\_Captures)) +

 facet\_wrap(~Year)

vesp\_tot\_sitdis = vespula |>

 group\_by(Site, Dist\_Urb, Species) |>

 summarise(Total\_Captures = sum(Abundance)) |>

 arrange(desc(Total\_Captures))

**Geographic analysis**

**Working with leaflet**

Worksite

library(leaflet)

lng1 = mean(vespula$LONG)

lat1 = mean(vespula$LAT)

leaflet() |>

 addTiles(group = "StreetMap") |>

 addProviderTiles(providers$Stamen.Terrain) |>

 setView(lng = lng1, lat = lat1, zoom = 10)

vespula\_mapa = vespula |> filter(Year == 2012, FECHA == "12/03/2012") |> rename(long = LONG, lat = LAT)

bees = makeIcon("wasp.png", iconWidth = 19, iconHeight = 19)

# Show first 20 rows from the `quakes` dataset

leaflet(data = vespula\_mapa) |> addTiles() |>

 addMarkers(~long, ~lat, icon = bees)

leaflet(vespula\_mapa) |> addTiles() |> addMarkers(

 clusterOptions = markerClusterOptions()

**Studying the relationship between the number of captures and distance to an urban center**

The relationship between total captures and the distance to an urban center for the two species was studied.

library(hrbrthemes)

 ggplot(vesp\_tot\_sitdis,

 aes(x = Dist\_Urb, y = Total\_Captures, col = Species)

 )+

 geom\_point() +

 theme\_ipsum() +

 geom\_jitter(width = 0.5)

vesp\_tot\_sitdis = vesp\_tot\_sitdis |> mutate(inverse\_captures = 1/(Total\_Captures + 1))

 ggplot(vesp\_tot\_sitdis,

 aes(x = Dist\_Urb, y = inverse\_captures, col = Species)

 )+

 geom\_point() +

 theme\_ipsum()

**Smoothing to find the best model**

ggplot(vesp\_tot\_sitdis,

 aes(x = Dist\_Urb, y = Total\_Captures, col = Species)

 )+

 geom\_point() +

 theme\_ipsum() + geom\_smooth(method = "lm", formula = y ~ poly(x, 2, raw = TRUE), se = FALSE)

ggplot(vesp\_tot\_sitdis,

 aes(x = Dist\_Urb, y = Total\_Captures, col = Species)

 )+

 geom\_point() +

 theme\_ipsum() + geom\_smooth(method = lm,

 formula = y ~ splines::bs(x, degree = 5, knots = 5), se = FALSE)

vesp\_tot\_sitdis = vespula |> group\_by(Site, Dist\_Urb, Species) |> summarise(Total\_Captures = sum(Abundance)) |> arrange(desc(Total\_Captures)) |> print()

library(magrittr)

vespula %<>% mutate(Year = as.factor(Year), Moment = if\_else(FECHA == "12/03/2012", "M1", if\_else(FECHA == "26/03/2012", "M2", "A1")))

 vespula |> select(c(Ash, Year)) |> ggplot(aes(x = Year, y = Ash)) + geom\_boxplot()

vespula |> select(c(Ash, Moment, Year)) |>

 ggplot(aes(x = Moment, y = Ash)) +

 geom\_boxplot() + facet\_wrap(~Year)

vesp\_tot\_sitdisCe = vespula |> group\_by(Site, Dist\_Urb) |> summarise(Total\_Captures = sum(Abundance), Avg\_Ash = mean(Ash)) |> arrange(desc(Total\_Captures))

library(viridis)

 ggplot(vesp\_tot\_sitdisCe,

 aes(x = Dist\_Urb, y = Total\_Captures, size = Avg\_Ash)) +

 geom\_point(alpha=0.5, shape=21, color="black") +

 scale\_size(range = c(.1, 10), name="Average Ash") +

 theme\_ipsum() +

 scale\_fill\_viridis(discrete=TRUE, guide = "none", option="A")

library(fitdistrplus)

fit\_abundance = fitdist(vespula$Abundance, "pois")

summary(fit\_abundance)

par(mfrow = c(2, 2), mar = c(1, 1, 1, 1))

plot(fit\_abundance)

fit\_abundance2 = fitdist(vespula$Abundance, "nbinom")

summary(fit\_abundance2)

par(mfrow = c(2, 2), mar = c(1, 1, 1, 1))

plot(fit\_abundance2)

**Modeling**

**Poisson modeling**

vesp\_glm\_pois <- glm(Abundance ~ Year \* Species \* Ash,

 family = poisson,

 data = vespula)

vesp\_glm\_pois2 <- glm(Abundance ~ Species \* Year,

 family = poisson,

 data = vespula)

summary(vesp\_glm\_pois2)

anova(vesp\_glm\_pois)

car::Anova(vesp\_glm\_pois2)

deviance(vesp\_glm\_pois2)

**Applying a negative binomial model with GLM**

library(MASS)

#### Using the "glm.nb" function to estimate the parameter value

vespula\_glm\_bn <- glm.nb(Abundance ~ Species \* Year \* Dist\_Urb,

 data = vespula)

summary(vespula\_glm\_bn)

car::Anova(vespula\_glm\_bn)

#### Using the "glm.nb" function to estimate the parameter value

vespula\_glm\_bn0 <- glm.nb(Abundance ~ Species \* Dist\_Urb + Year,

 data = vespula)

summary(vespula\_glm\_bn0)

car::Anova(vespula\_glm\_bn0)

library(topmodels)

vesp0\_root\_zip <- rootogram(vespula\_glm\_bn0, style = "hanging", plot = FALSE)

autoplot(vesp0\_root\_zip, xlim = c(0,250))

vesp\_glmbn <- glm(Abundance ~ Species \* Year \* Dist\_Urb,

 data = vespula,

 family = negative.binomial(2.47)

 )

summary(vesp\_glmbn)

car::Anova(vesp\_glmbn)

**Analyzing with GLMM**

Considering the site effect

library(lme4)

vesp\_glmer.bn\_p <- glmer.nb(Abundance ~ Species \* Dist\_Urb + Year +

 (1|Site),

 control=glmerControl(optimizer="bobyqa"),

 data = vespula)

summary(vesp\_glmer.bn\_p)

car::Anova(vesp\_glmer.bn\_p)

df.residual(vesp\_glmer.bn\_p)

deviance(vesp\_glmer.bn\_p)

performance::check\_singularity(vesp\_glmer.bn\_p )

performance::check\_overdispersion(vesp\_glmer.bn\_p)

performance::check\_zeroinflation(vesp\_glmer.bn\_p)

vesp\_glmer.bn\_p1 <- glmer.nb(Abundance ~ Species \* Year + Dist\_Urb +

 (1|Site),

 control=glmerControl(optimizer="bobyqa"),

 data = vespula)

summary(vesp\_glmer.bn\_p1)

car::Anova(vesp\_glmer.bn\_p1)

df.residual(vesp\_glmer.bn\_p1)

deviance(vesp\_glmer.bn\_p1)

performance::check\_singularity(vesp\_glmer.bn\_p1 )

performance::check\_overdispersion(vesp\_glmer.bn\_p1)

performance::check\_zeroinflation(vesp\_glmer.bn\_p1)

vesp\_glmer.bn\_p2 <- glmer.nb(Abundance ~ Species + Dist\_Urb + Year +

 (1|Site),

 control=glmerControl(optimizer="bobyqa"),

 data = vespula)

summary(vesp\_glmer.bn\_p2)

car::Anova(vesp\_glmer.bn\_p2)

df.residual(vesp\_glmer.bn\_p2)

deviance(vesp\_glmer.bn\_p2)

performance::check\_singularity(vesp\_glmer.bn\_p2)

performance::check\_overdispersion(vesp\_glmer.bn\_p2)

performance::check\_zeroinflation(vesp\_glmer.bn\_p2)

anova(vesp\_glmer.bn\_p, vesp\_glmer.bn\_p1, vesp\_glmer.bn\_p2)

library(bbmle)

AICtab(vesp\_glmer.bn\_p, vesp\_glmer.bn\_p1,vesp\_glmer.bn\_p2)

**Checking the fit of the NB model from the deviance residuals**

res <- resid(vesp\_glmer.bn\_p, type = "deviance")

pred <- predict(vesp\_glmer.bn\_p)

par(mfrow=c(1,2))

plot(pred, res, xlab = "linear predictor", ylab = "deviance residuals")

qqnorm(res, ylab = "deviance residuals")

qqline(res)

vespula |> rowid\_to\_column("obs\_id") |> #Create a column with the id number

 mutate( residuos\_pearson = resid(vesp\_glmer.bn\_p, type = "pearson"), predichos =predict(vesp\_glmer.bn\_p,type = "response") ) |>

 ggplot() + aes(x = predichos, y = residuos\_pearson, color = abs(residuos\_pearson) > 3) + geom\_hline(yintercept = c(0, -3, 3)) + geom\_point() +

 geom\_text(aes(label = ifelse(abs(residuos\_pearson) > 3, obs\_id, "")), position = position\_nudge(x = 30)) + scale\_color\_manual(values = c("Black", "Red"))

**Ashes are by site, not by year. See the interaction or analysis**

#install.packages("R2admb")

#install.packages("glmmADMB",

# repos=c("http://glmmadmb.r-forge.r-project.org/repos",

# getOption("repos")),

# type="source")#library("glmmADMB")

vespula = vespula |>

 mutate(FECHA = as.factor(FECHA), Year = as.factor(Year), Site = as.factor(Site))

library(glmmADMB)

M1<- glmmadmb(Abundance~ Ash + Dist\_Urb + Species + (1|FECHA/Year/Site),

 data = vespula, zeroInflation=TRUE, family="nbinom")

summary(M1)

M2<- glmmadmb(Abundance~ Ash + Dist\_Urb + Species + (1|FECHA/Year/Site), data=vespula, zeroInflation=TRUE, family="nbinom1")

summary(M2)

**Using the glmmTMB package**

*Zero-inflated with Poisson*

library("glmmTMB")

vespula\_zipoisTMB <- glmmTMB(Abundance ~ Species \* Year + Dist\_Urb +

 (1|Site),

 data = vespula, ziformula=~1,family=poisson)

summary(vespula\_zipoisTMB)

**Save the used data:**

vespula |> dplyr::select(!c(FECHA,Momento)) |> write\_csv2(file = "vespula\_data.csv")

**Create a plot:**

emmip(vespula\_NB1TMB25ok, ~ Year,

 type = "response",

 CIs = T,

 CIarg = list(lwd = 1, alpha = 1),

 inearg = list(linetype = "Year")) +theme\_bw(base\_size = 18) +

 ylab("Abundance") + xlab("Year") +

 geom\_point(size = 2) + scale\_y\_continuous(breaks = 0:7, limits = c(0,7.05))+

geom\_errorbar(aes(ymin = summary(meanyead)$asymp.LCL, ymax=summary(meanyead)$asymp.UCL), width=.05)

**Another plot:**

fib.rg = ref\_grid(vespula\_NB1TMB25ok,

 at = list(Dist\_Urb= c(0.000000, 5.449074, 14.209447, 30.279438, 47.895142))

)

emmip\_vesp <- emmip(fib.rg, ~ Dist\_Urb |Year, type = "response", CIs = T,

 CIarg = list(lwd = 1, alpha = 1),

 inearg = list(linetype = "Dist\_Urb")) + theme\_bw(base\_size = 18) + ylab("Abundance") +

 geom\_point(size = 2) + scale\_x\_continuous(limits = c(0,50)) +

 theme(legend.title=element\_text(size=22),

 legend.text=element\_text(size=16)

 )

emmip\_vesp + geom\_errorbar(aes(

 ymin = ggplot\_build(emmip\_vesp )$data[[2]]$ymin,

 ymax = ggplot\_build(emmip\_vesp )$data[[2]]$ymax,

 width=.9), colour="black"

 )

guides(col = F, size = F) +

 geom\_point(size = 2) + geom\_errorbar(aes(

 ymin = ggplot\_build(emmip\_vesp )$data[[2]]$ymin,

 ymax = ggplot\_build(emmip\_vesp )$data[[2]]$ymax,

 width=.1)

 )

datosmedias <- ggplot\_build(emmip\_vesp )$data[[2]]

ggplot(datosmedias, aes(x = x, y = y)) + geom\_point(size = 2) + theme\_bw(base\_size = 18) + ylab("Abundance") + xlab("Distance Urbane") + geom\_point(size = 2) + geom\_line()+

 theme(legend.title=element\_text(size=22),

 legend.text=element\_text(size=16)

 ) + geom\_errorbar(aes(

 ymin = ggplot\_build(emmip\_vesp )$data[[2]]$ymin,

 ymax = ggplot\_build(emmip\_vesp )$data[[2]]$ymax,

 width=.9), colour="black"

 ) + scale\_y\_continuous(breaks = 0:7, limits = c(0,7.05))

**Further plot modifications and saving the plot:**

datosmedias1 <- ggplot\_build(emmip\_vesp)$data[[2]]

datosmedias2 <- datosmedias1 |>

 rename(Year = PANEL) |>

 mutate(Year = as.factor(Year), Year = if\_else(Year == 1, 2012, ifelse(Year == 2, 2013, 2014)))

ggplot(datosmedias2, aes(x = x, y = y)) + theme\_classic(base\_size = 10) +

 ylab("Abundance (average catch values)") + xlab("Distance from nearest urban center (km)") + geom\_line() +

 facet\_wrap(~Year) +

 theme(text = element\_text(family = "Arial"),

 axis.line = element\_line(color='black'),

 plot.background = element\_blank(),

 panel.grid.minor = element\_blank(),

 panel.grid.major = element\_blank(),

 panel.border = element\_blank(),

 legend.title=element\_text(size=10),

 legend.text=element\_text(size=10),

 axis.title.x = element\_text(margin = margin(t = 0.25, unit = "in")),

 axis.title.y = element\_text(margin = margin(r = 0.25, unit = "in"))

 ) + geom\_errorbar(aes(

 ymin = ymin,

 ymax = ymax), colour="#E7B800" ,

 linewidth=2, width = 0,

 ) + geom\_point(size = 2.5)

ggsave("grafbar.tiff", height = 5.2, width = 7.7, device = "tiff", dpi = 600)

dev.off()

**Further data analysis:**

vespulas = vespula |> mutate(Ash = as.factor(Ash))

vespula\_NB1TMB27a <- glmmTMB(Abundance ~ Species + Ash + Dist\_Urb + Year +

 (1|Site),

 data = vespulas,

 disp = ~Year + Species,

 family = nbinom1

 )

summary(vespula\_NB1TMB27a)

car::Anova(vespula\_NB1TMB27a)

**New site names:**

Datos\_ANA = read.csv("datos\_anadic12.csv", dec = ".", sep = "\t")

head(Datos\_ANA)

**Further plotting:**

library(hrbrthemes)

###Lollipop by year (NOT ordered by Dist\_Urb)###

vesp\_tot\_sitl = Datos\_ANA |> group\_by(ID, Species, Year) |>

 summarise(Capturas\_tot = sum(Abundance)) |> arrange(desc(Capturas\_tot))

vesp\_tot\_sitlt = vesp\_tot\_sitl |> pivot\_wider(names\_from =Species, values\_from = Capturas\_tot)

tabfin = vesp\_tot\_sitlt |>

 rowwise() |>

 mutate(suma\_esp = sum(c(`V. germanica`,`V. vulgaris`) )) |>

 arrange(suma\_esp) |>

 mutate(ID = factor(ID, ID)) |> print(n = 23)

library(hrbrthemes)

# Plot

ggplot(tabfin) +

 geom\_segment( aes(x=ID, xend=ID, y=`V. germanica`, yend=`V. vulgaris`), color="grey") +

 geom\_point( aes(x=ID, y=suma\_esp), size=3 ) +

#, color="#E7B800" +

#geom\_point( aes(x=ID, y=`V. germanica`),

#insert the image

#install.packages("png")

#install.packages("patchwork")

library(patchwork)

library(png)

library(grid)

#readPNG("C:/Users/ajper/Dropbox/Paper #Ceniza/analisis\_R\_2022/anlisisestadsticovespula/wasp.png", native = TRUE)

xpos<-1:4

ypos<-xpos\*\*2

path <- "wasp2.png"

img <- readPNG(path, native = TRUE)

legend <- rasterGrob(image=img, interpolate = F, height =unit(1,"cm"))

#dist + annotation\_custom(legend, xmin=58, xmax=60, ymin=148, ymax=150)

 # grid.arrange(dist,legend)

wasp<-rasterGrob(img, interpolate = T, height =unit(1.4,"cm"))

img\_graph <- dist + inset\_element(p=wasp, left = 0.9, bottom = 0.6, right = 1, top = 0.8)

img\_graph

print (img\_graph)

ggsave("graf5sqr.tiff", height = 7, width = 6, device = "tiff", dpi = 600)

#ggsave("graf2.tiff", device = "tiff", dpi = 600)

dev.off()