

# Interacción entre factores

Taller GAB, Corrientes 2025

María Gabriela Cendoya  
María Gloria Monterubbianesi

Jueves 2 de octubre de 2025

## Introducción

### ¿Qué es una estructura factorial de los tratamientos?

Cuando el conjunto de tratamientos a estudiar se define como todas las posibles combinaciones de los niveles de varios factores.

Ejemplo, los tratamientos son todas las combinaciones de

- 3 variedades de trigo con
- 4 niveles de fertilización nitrogenada,
- 3 fechas de siembra

en total 36 tratamientos.

### Nos focalizamos ahora en solo estructuras de dos factores

### ¿Qué es la interacción entre factores?

Dos factores interactúan si la diferencia en la respuesta promedio para dos niveles de un factor **NO es constante** a través de los niveles del otro factor.

El efecto de un factor en la variable respuesta **depende** de en qué nivel del otro factor se lo mida.

### Sobre el análisis estadístico

El análisis estadístico adecuado es aquel que responda la pregunta del investigador.

Si estamos en una estructura factorial de los tratamientos

¿Nos interesa estudiar la interacción?, o

¿Estamos en la búsqueda del mejor tratamiento?

### Desarrollo de ejemplos

```
# setwd("~/Gaby/GAB2025/Taller")
```

## Carga de paquetes

```
# library(MASS)
# library(ggplot2)
# library(dplyr)
# library(tidyr)
# library(nlme)
# library(emmeans)
# library(multcomp)
```

```
data(cabbages)
# help(cabbages) # Los tratamientos ¿Tienen estructura factorial?
summary(cabbages)
```

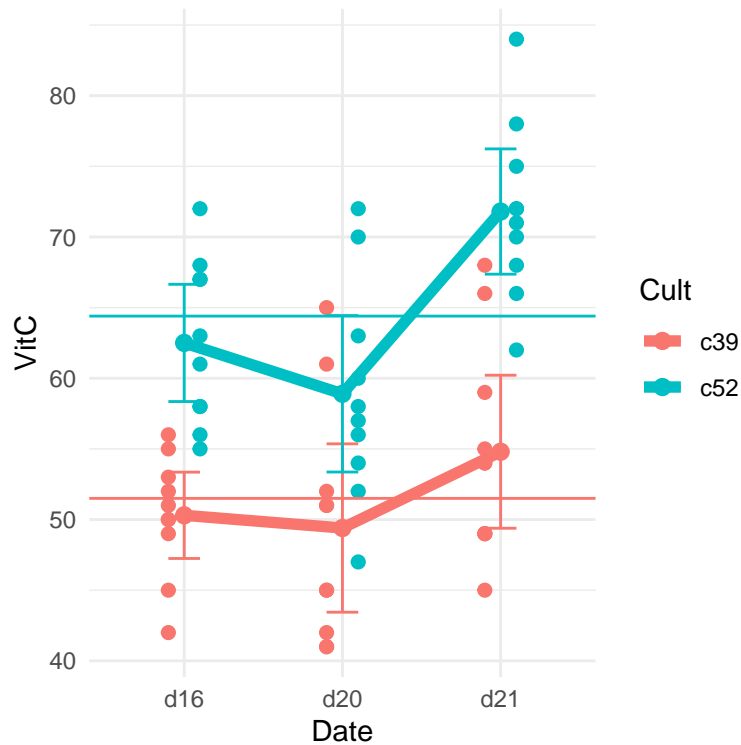
```
##   Cult   Date      HeadWt      VitC
## c39:30 d16:20  Min.    :1.000  Min.    :41.00
## c52:30 d20:20  1st Qu.:1.875  1st Qu.:50.75
##           d21:20  Median :2.550  Median :56.00
##           Mean   :2.593  Mean   :57.95
##           3rd Qu.:3.125  3rd Qu.:66.25
##           Max.   :4.300  Max.   :84.00
```

```
#####
```

## Análisis de Vitamina C

Veamos gráficamente los datos

```
{
promediosxC <- cabbages %>%
  group_by(Cult) %>%
  summarise(VitC_medio = mean(VitC))
promediosxD <- cabbages %>%
  group_by(Date) %>%
  summarise(VitC_medio = mean(VitC))
# windows()
ggplot(cabbages, aes(x=Date, y=VitC, colour=Cult, group=Cult))+
  geom_point(position = position_dodge(width = 0.4), size = 2)+
  stat_summary(fun = mean, geom = "point", shape = 20, size = 4) +
  stat_summary(fun = mean, geom = "line", linewidth = 2) +
  stat_summary(fun.data = mean_cl_normal, geom = "errorbar", width = 0.2)+
  geom_hline(data = promediosxC, aes(yintercept = VitC_medio, colour = Cult), linetype = "solid") +
  theme_minimal()
}
```



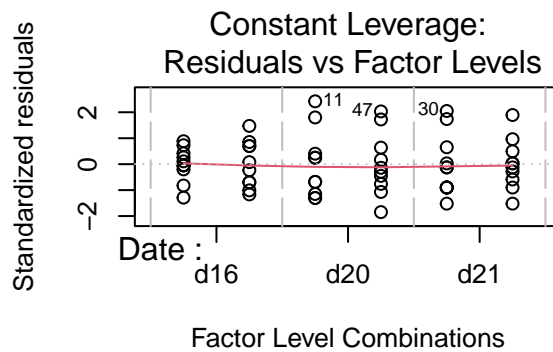
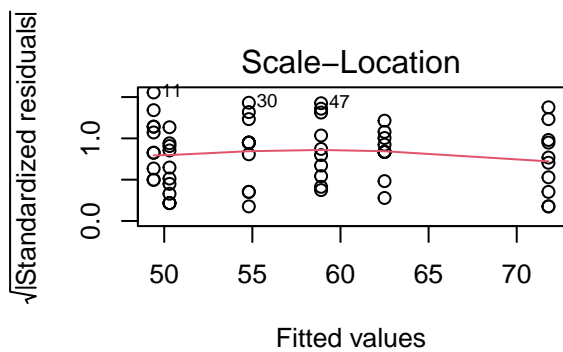
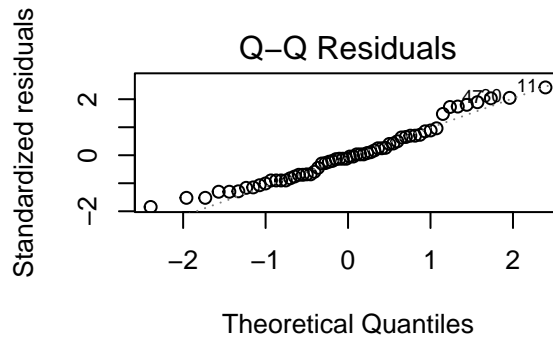
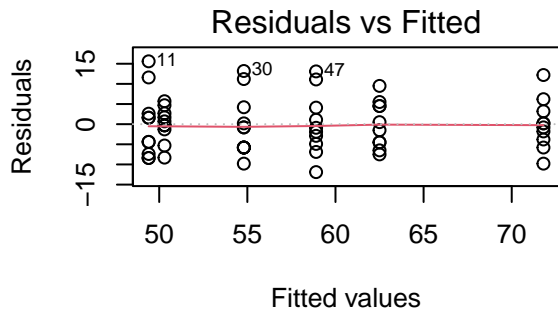
Veamos numéricamente los datos

```
{
tabla_promedios <- cabbages %>%
  group_by(Cult, Date) %>%
  summarise(promedio = mean(VitC), .groups = "drop") %>%
  pivot_wider(names_from = Date, values_from = promedio)
tabla_promedios <- tabla_promedios %>%
  mutate(Promedio_Cult = rowMeans(select(., -Cult), na.rm = TRUE))
tabla_promedios <- tabla_promedios %>%
  bind_rows(
    summarise(.,
      Cult = "Promedio_Date",
      across(where(is.numeric), ~ mean(.x, na.rm = TRUE)))
  )
as.data.frame(tabla_promedios)
}
```

```
##           Cult d16  d20 d21 Promedio_Cult
## 1          c39 50.3 49.40 54.8           51.50
## 2          c52 62.5 58.90 71.8           64.40
## 3 Promedio_Date 56.4 54.15 63.3           57.95
```

Estudio de la interacción

```
MVit0 <- lm(VitC~Date*Cult,data=cabbages)
# windows()
oldpar <- par()
par(mfrow=c(2,2))
plot(MVit0)
```



```
anova(MVit0 ) # Volver al gráfico de interacción
```

```
## Analysis of Variance Table
##
## Response: VitC
##          Df Sum Sq Mean Sq F value    Pr(>F)
## Date      2  909.3   454.65   9.8555 0.0002245 ***
## Cult      1 2496.2  2496.15  54.1095 1.089e-09 ***
## Date:Cult  2  144.3    72.15   1.5640 0.2186275
## Residuals 54 2491.1    46.13
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

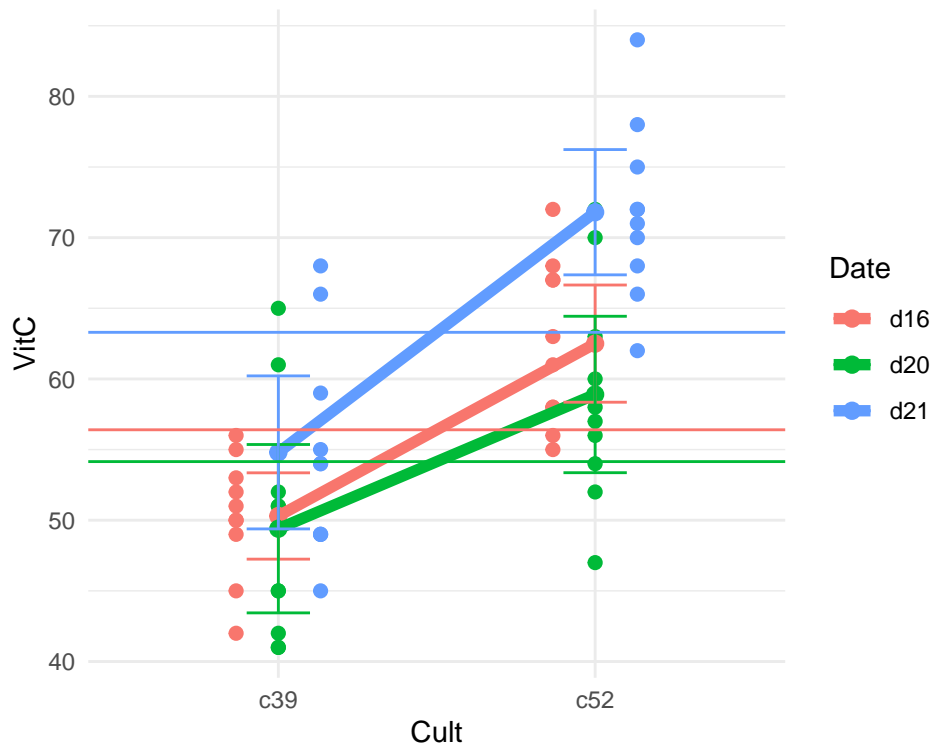
```
par <- oldpar
```

```
# cld(emmeans(MVit0,~Cult),Letters=letters,decreasing=TRUE) #no lo necesito hay sólo dos niveles
# si nos interesa ver la interacción del otro punto de vista
{
# windows()
ggplot(cabbages,aes(x=Cult,y=VitC,colour=Date,group=Date))+
```

```

geom_point(position = position_dodge(width = 0.4), size = 2)+
stat_summary(fun = mean, geom = "point", shape = 20, size = 4) +
stat_summary(fun = mean,geom = "line", linewidth = 2) +
stat_summary(fun.data = mean_cl_normal, geom = "errorbar", width = 0.2)+
geom_hline(data = promediosxD, aes(yintercept = VitC_medio, colour = Date), linetype = "solid") +
theme_minimal()
}

```



```

cld(emmeans(MVit0, ~Date), Letters=letters, decreasing=TRUE)

```

## NOTE: Results may be misleading due to involvement in interactions

```

## Date emmean SE df lower.CL upper.CL .group
## d21 63.3 1.52 54 60.3 66.3 a
## d16 56.4 1.52 54 53.4 59.4 b
## d20 54.1 1.52 54 51.1 57.2 b
##
## Results are averaged over the levels of: Cult
## Confidence level used: 0.95
## P value adjustment: tukey method for comparing a family of 3 estimates
## significance level used: alpha = 0.05
## NOTE: If two or more means share the same grouping symbol,
## then we cannot show them to be different.
## But we also did not show them to be the same.

```

**Conclusión**

```
# El cultivar C52 tiene más VitC que C39, en cualquiera de las 3 fechas de plantación.  
# En la fecha d21 se obtiene más VitC que en las otras 2 fechas.  
# Basados en este experimento no podemos asegurar que la VitC difiera entre las fechas d16 y d20.
```

Si no interesa la interacción (o su ausencia), sólo busco el mejor tratamiento

```
cld(emmeans(MVit0,~Date:Cult),Letters=letters,decreasing=TRUE,adjust="tukey")
```

```
## Note: adjust = "tukey" was changed to "sidak"  
## because "tukey" is only appropriate for one set of pairwise comparisons
```

```
## Date Cult emmean SE df lower.CL upper.CL .group  
## d21 c52 71.8 2.15 54 65.9 77.7 a  
## d16 c52 62.5 2.15 54 56.6 68.4 b  
## d20 c52 58.9 2.15 54 53.0 64.8 bc  
## d21 c39 54.8 2.15 54 48.9 60.7 bcd  
## d16 c39 50.3 2.15 54 44.4 56.2 cd  
## d20 c39 49.4 2.15 54 43.5 55.3 d  
##  
## Confidence level used: 0.95  
## Conf-level adjustment: sidak method for 6 estimates  
## P value adjustment: tukey method for comparing a family of 6 estimates  
## significance level used: alpha = 0.05  
## NOTE: If two or more means share the same grouping symbol,  
## then we cannot show them to be different.  
## But we also did not show them to be the same.
```

```
# Recomendamos el Cultivar c52 en la fecha d21 igual que antes, pero  
# no podemos recomendar c52 en la fecha 20,  
# no podemos recomendar la fecha d21 para C39
```

¿Qué perdimos? ¿Qué ganamos?

¿Por qué ganamos? Potencia!

```
# Veamos intervalos de confianza simultánea para las estimaciones de medias  
emmeans(MVit0,~Cult,adjust="bonferroni")
```

```
## NOTE: Results may be misleading due to involvement in interactions
```

```
## Cult emmean SE df lower.CL upper.CL  
## c39 51.5 1.24 54 48.6 54.4  
## c52 64.4 1.24 54 61.5 67.3  
##  
## Results are averaged over the levels of: Date  
## Confidence level used: 0.95  
## Conf-level adjustment: bonferroni method for 2 estimates
```

```
emmeans(MVit0,~Date,adjust="bonferroni")
```

## NOTE: Results may be misleading due to involvement in interactions

```
## Date emmean SE df lower.CL upper.CL
## d16 56.4 1.52 54 52.6 60.2
## d20 54.1 1.52 54 50.4 57.9
## d21 63.3 1.52 54 59.5 67.1
##
## Results are averaged over the levels of: Cult
## Confidence level used: 0.95
## Conf-level adjustment: bonferroni method for 3 estimates
```

```
emmeans(MVit0, ~Date:Cult, adjust="bonferroni")
```

```
## Date Cult emmean SE df lower.CL upper.CL
## d16 c39 50.3 2.15 54 44.4 56.2
## d20 c39 49.4 2.15 54 43.5 55.3
## d21 c39 54.8 2.15 54 48.9 60.7
## d16 c52 62.5 2.15 54 56.6 68.4
## d20 c52 58.9 2.15 54 53.0 64.8
## d21 c52 71.8 2.15 54 65.9 77.7
##
## Confidence level used: 0.95
## Conf-level adjustment: bonferroni method for 6 estimates
```

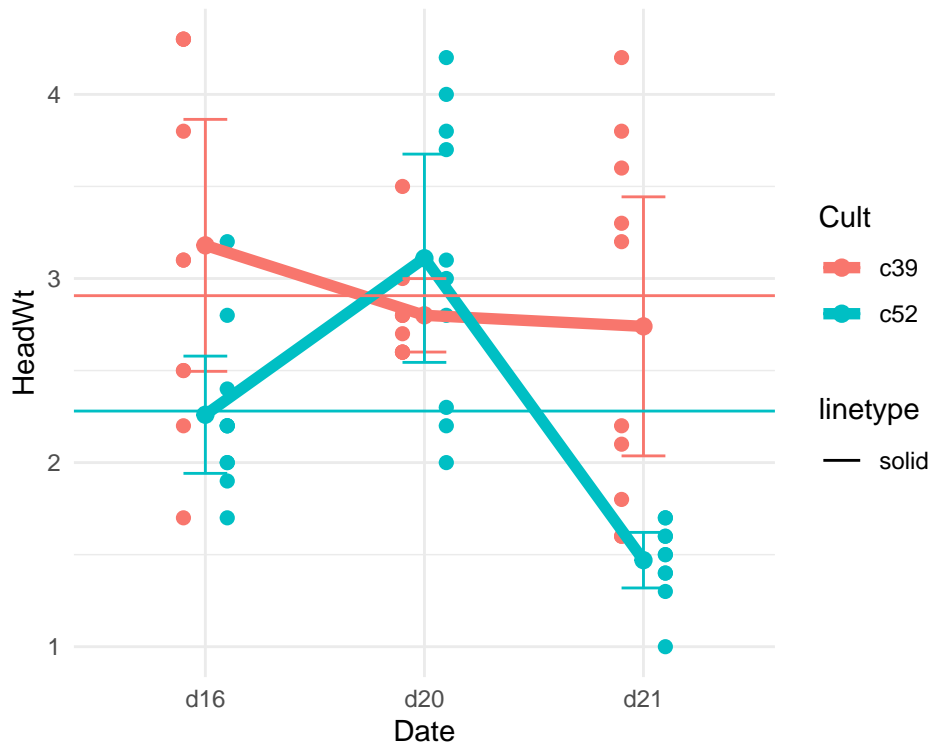
En ausencia de interacción, las comparaciones de efectos principales son más potentes que las comparaciones entre tratamientos, pues comparamos promedios basados en más repeticiones

#####

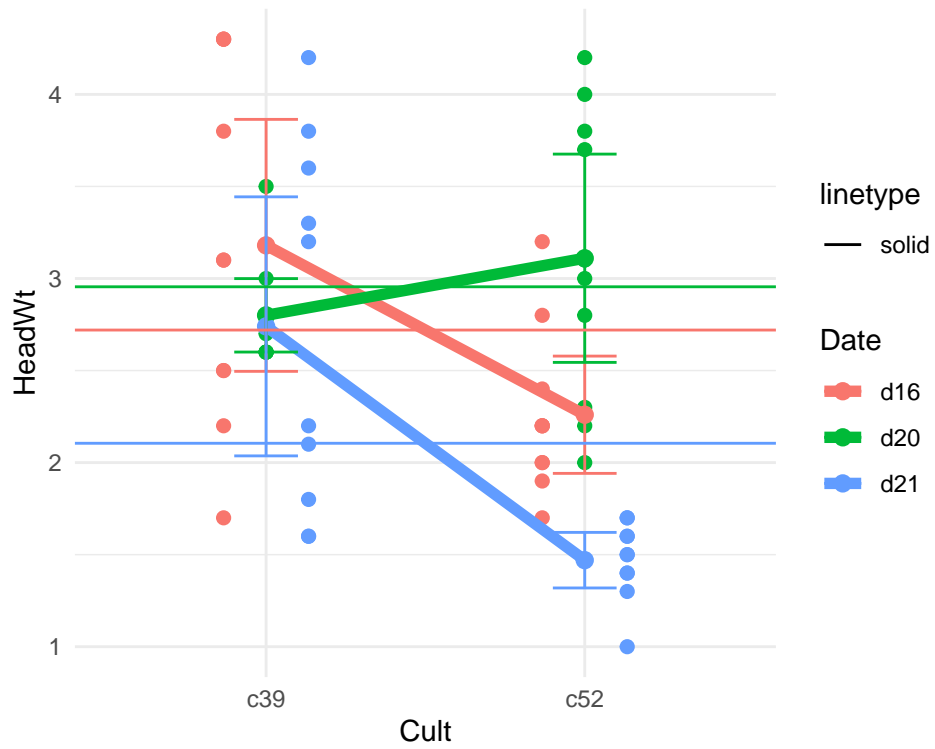
## Análisis del peso de las cabezas de repollo

Veamos gráficamente los datos

```
{
promediosHeadWtxC <- cabbages %>%
  group_by(Cult) %>%
  summarise(HeadWt_medio = mean(HeadWt))
promediosHeadWtxD <- cabbages %>%
  group_by(Date) %>%
  summarise(HeadWt_medio = mean(HeadWt))
# windows()
ggplot(cabbages, aes(x=Date, y=HeadWt, colour=Cult, group=Cult)) +
  geom_point(position = position_dodge(width = 0.4), size = 2) +
  stat_summary(fun = mean, geom = "point", shape = 20, size = 4) +
  stat_summary(fun = mean, geom = "line", linewidth = 2) +
  stat_summary(fun.data = mean_cl_normal, geom = "errorbar", width = 0.2) +
  geom_hline(data = promediosHeadWtxC,
             aes(yintercept = HeadWt_medio, colour = Cult, linetype = "solid")) +
  theme_minimal()
}
```

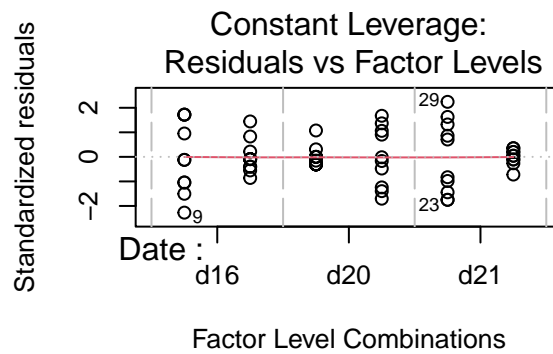
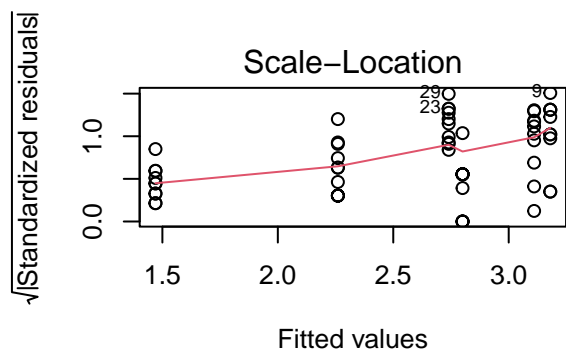
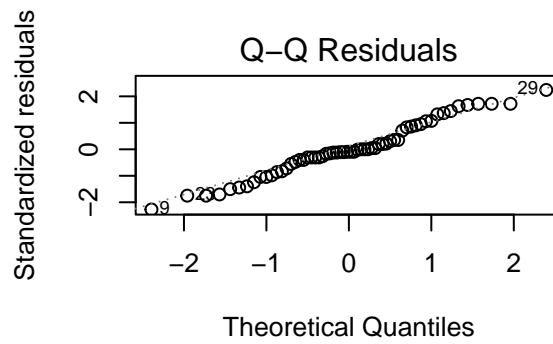
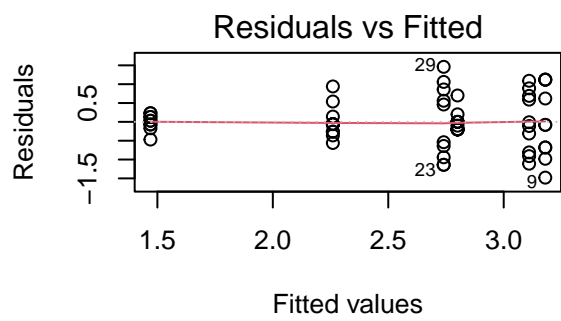


```
# Si necesitamos verlos en el otro sentido
# windows()
ggplot(cabbages, aes(x=Cult, y=HeadWt, colour=Date, group=Date)) +
  geom_point(position = position_dodge(width = 0.4), size = 2) +
  stat_summary(fun = mean, geom = "point", shape = 20, size = 4) +
  stat_summary(fun = mean, geom = "line", linewidth = 2) +
  stat_summary(fun.data = mean_cl_normal, geom = "errorbar", width = 0.2) +
  geom_hline(data = promediosHeadWtxD,
            aes(yintercept = HeadWt_medio, colour = Date, linetype = "solid")) +
  theme_minimal()
```



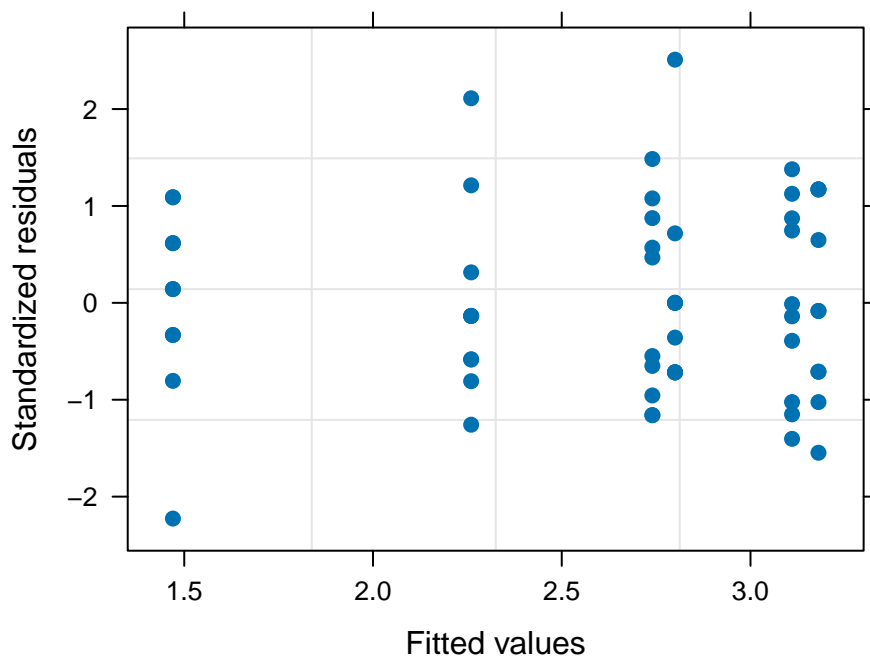
## Comenzando el análisis

```
MHead0 <- lm(HeadWt~Date*Cult,data=cabbages)
# windows()
oldpar <- par()
par(mfrow=c(2,2))
plot(MHead0)
```



```
par <- oldpar
```

```
MHead1 <- gls(HeadWt~Date*Cult,data=cabbages,weights=varIdent(form=~1|Date*Cult))
plot(MHead1,pch=16)
```



```
anova(MHead1 ,MHead0)
```

```
##          Model df      AIC      BIC    logLik  Test  L.Ratio p-value
## MHead1      1 12 120.1883 144.0561 -48.09416
## MHead0      2  7 140.4469 154.3698 -63.22344 1 vs 2 30.25856 <.0001
```

```
anova(MHead1) #Hay interacción entre los Cultivares y las Fechas de siembra
```

```
## Denom. DF: 54
##          numDF  F-value p-value
## (Intercept)    1 1906.9345 <.0001
## Date           2   80.7871 <.0001
## Cult           1    8.1374 0.0061
## Date:Cult      2    8.3873 0.0007
```

```
# Error muy frecuente
```

```
emmeans(MHead1, pairwise ~ Date)
```

```
## NOTE: Results may be misleading due to involvement in interactions
```

```
## $emmeans
## Date emmean  SE    df lower.CL upper.CL
## d16    2.72 0.167 12.46    2.36    3.08
## d20    2.96 0.133 11.20    2.66    3.25
## d21    2.10 0.159  9.82    1.75    2.46
##
```

```
## Results are averaged over the levels of: Cult
## Degrees-of-freedom method: satterthwaite
## Confidence level used: 0.95
```

```
##
## $contrasts
## contrast estimate SE df t.ratio p.value
## d16 - d20 -0.235 0.213 23.1 -1.103 0.5220
## d16 - d21 0.615 0.230 22.4 2.668 0.0357
## d20 - d21 0.850 0.207 19.7 4.105 0.0016
##
## Results are averaged over the levels of: Cult
## Degrees-of-freedom method: satterthwaite
## P value adjustment: tukey method for comparing a family of 3 estimates
```

```
emmeans(MHead1, pairwise ~ Cult)
```

```
## NOTE: Results may be misleading due to involvement in interactions
```

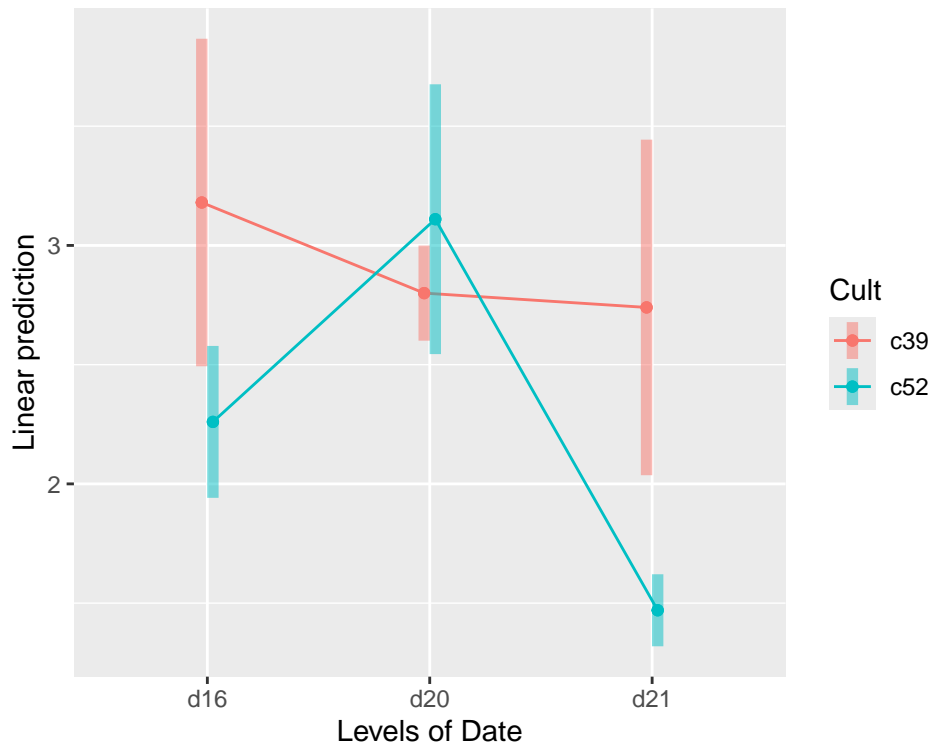
```
## $emmeans
## Cult emmean SE df lower.CL upper.CL
## c39 2.91 0.1476 19.4 2.60 3.22
## c52 2.28 0.0982 15.7 2.07 2.49
##
## Results are averaged over the levels of: Date
## Degrees-of-freedom method: satterthwaite
## Confidence level used: 0.95
##
## $contrasts
## contrast estimate SE df t.ratio p.value
## c39 - c52 0.627 0.177 32.5 3.535 0.0012
##
## Results are averaged over the levels of: Date
## Degrees-of-freedom method: satterthwaite
```

```
#####
# Si estamos en la búsqueda del mejor tratamiento, comparamos todos contra todos.
# Muchas comparaciones con control de Error de tipo I por experimento
```

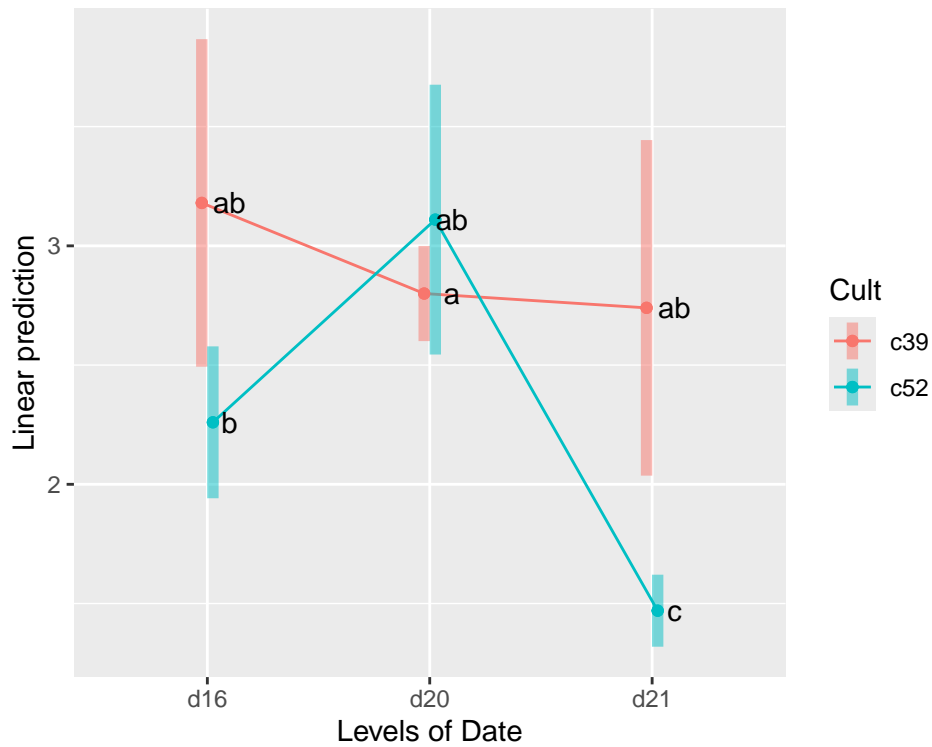
```
cld(emmeans(MHead1, ~Date:Cult), decreasing = TRUE, Letters=letters)
```

```
## Date Cult emmean SE df lower.CL upper.CL .group
## d16 c39 3.18 0.3025 8.80 2.49 3.87 ab
## d20 c52 3.11 0.2501 9.00 2.54 3.68 ab
## d20 c39 2.80 0.0882 9.00 2.60 3.00 a
## d21 c39 2.74 0.3110 8.99 2.04 3.44 ab
## d16 c52 2.26 0.1408 9.00 1.94 2.58 b
## d21 c52 1.47 0.0668 9.00 1.32 1.62 c
##
## Degrees-of-freedom method: satterthwaite
## Confidence level used: 0.95
## P value adjustment: tukey method for comparing a family of 6 estimates
## significance level used: alpha = 0.05
## NOTE: If two or more means share the same grouping symbol,
## then we cannot show them to be different.
## But we also did not show them to be the same.
```

```
TcT <- emmeans(MHead1, ~Date:Cult)
TcT.cld <- data.frame(cld(TcT, decreasing = TRUE, Letters=letters) )
# windows()
emmip(TcT, Cult~Date, CIs=TRUE)
```



```
emmi.TcT <- emmip(TcT, Cult~Date, CIs=TRUE)
emmi.TcT+
  annotate("text", label=trimws(as.character(TcT.cld$.group)),
          x=as.numeric(TcT.cld$Date)+0.1,
          y=TcT.cld$emmean)
```



```
#####
# Análisis de interacción con más potencia,
# con control parcial del error tipo I por experimento
#####
joint_tests(MHead1,by="Cult")
```

```
## Cult = c39:
## model term df1 df2 F.ratio p.value
## Date      2 10.29 0.766 0.4897
##
## Cult = c52:
## model term df1 df2 F.ratio p.value
## Date      2 12.85 29.739 <.0001
```

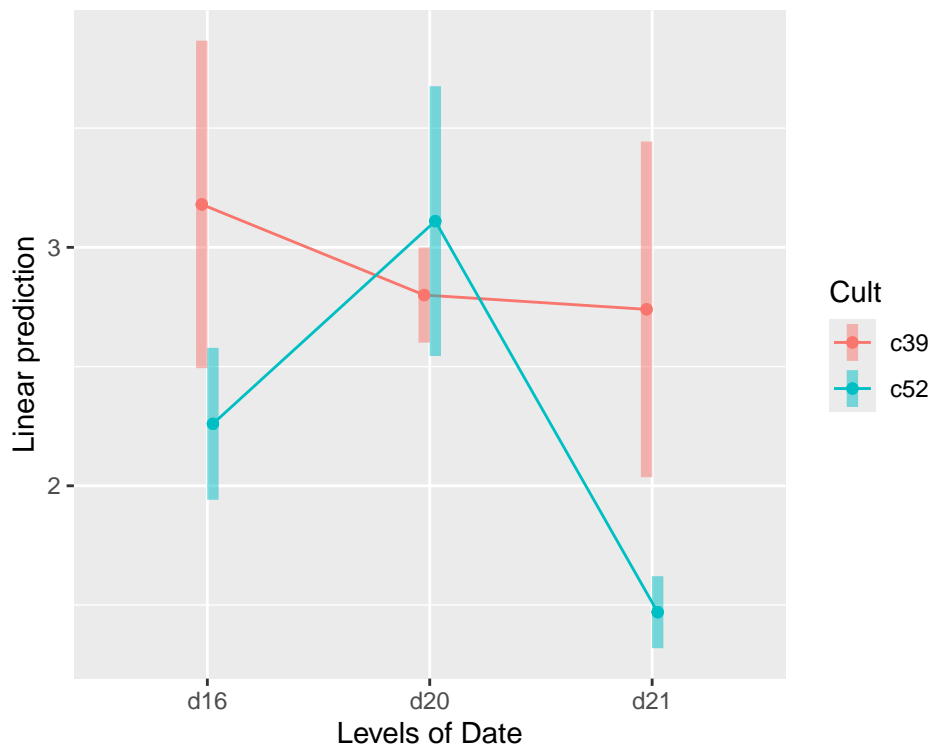
```
joint_tests(MHead1,by="Date")
```

```
## Date = d16:
## model term df1 df2 F.ratio p.value
## Cult      1 12.46 7.602 0.0169
##
## Date = d20:
## model term df1 df2 F.ratio p.value
## Cult      1 11.20 1.367 0.2667
##
## Date = d21:
## model term df1 df2 F.ratio p.value
## Cult      1 9.82 15.943 0.0026
```

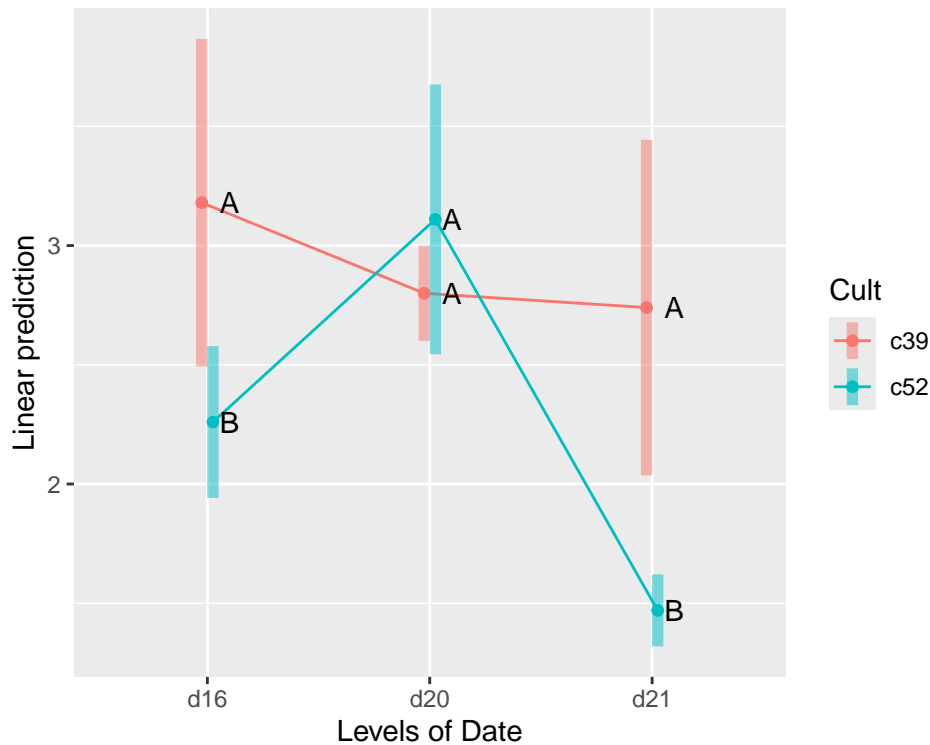
```
cld(CinD <- emmeans(MHead1, ~Cult|Date), decreasing = TRUE, Letters=LETTERS)
```

```
## Date = d16:
## Cult emmean SE df lower.CL upper.CL .group
## c39 3.18 0.3025 8.80 2.49 3.87 A
## c52 2.26 0.1408 9.00 1.94 2.58 B
##
## Date = d20:
## Cult emmean SE df lower.CL upper.CL .group
## c52 3.11 0.2501 9.00 2.54 3.68 A
## c39 2.80 0.0882 9.00 2.60 3.00 A
##
## Date = d21:
## Cult emmean SE df lower.CL upper.CL .group
## c39 2.74 0.3110 8.99 2.04 3.44 A
## c52 1.47 0.0668 9.00 1.32 1.62 B
##
## Degrees-of-freedom method: satterthwaite
## Confidence level used: 0.95
## significance level used: alpha = 0.05
## NOTE: If two or more means share the same grouping symbol,
## then we cannot show them to be different.
## But we also did not show them to be the same.
```

```
CinD.cld <- data.frame(cld(CinD, decreasing = TRUE, Letters=LETTERS) )
# windows()
emmip(CinD, Cult~Date, CIs=TRUE)
```



```
emmi.CinD <-emmi(CinD, Cult~Date,CIs=TRUE)
emmi.CinD+
  annotate("text",label=trimws(as.character(CinD.cld$.group)),x=as.numeric(CinD.cld$Date)+0.1,
          y=CinD.cld$emmean)
```

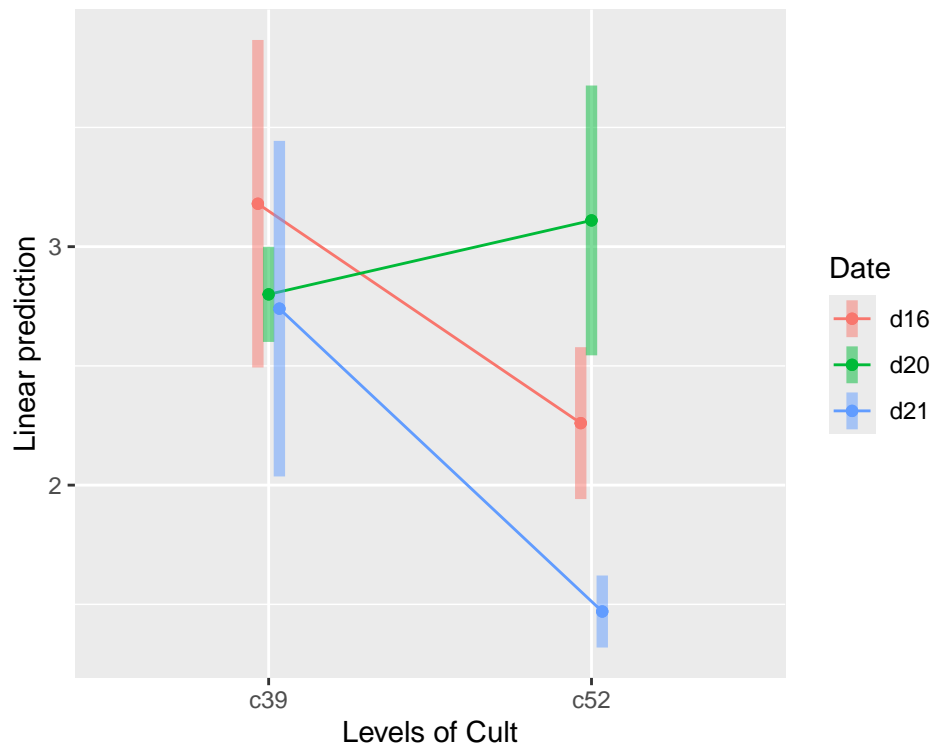


```
cld(DinC <- emmeans(MHead1, ~Date|Cult),decreasing = TRUE,Letters=letters)
```

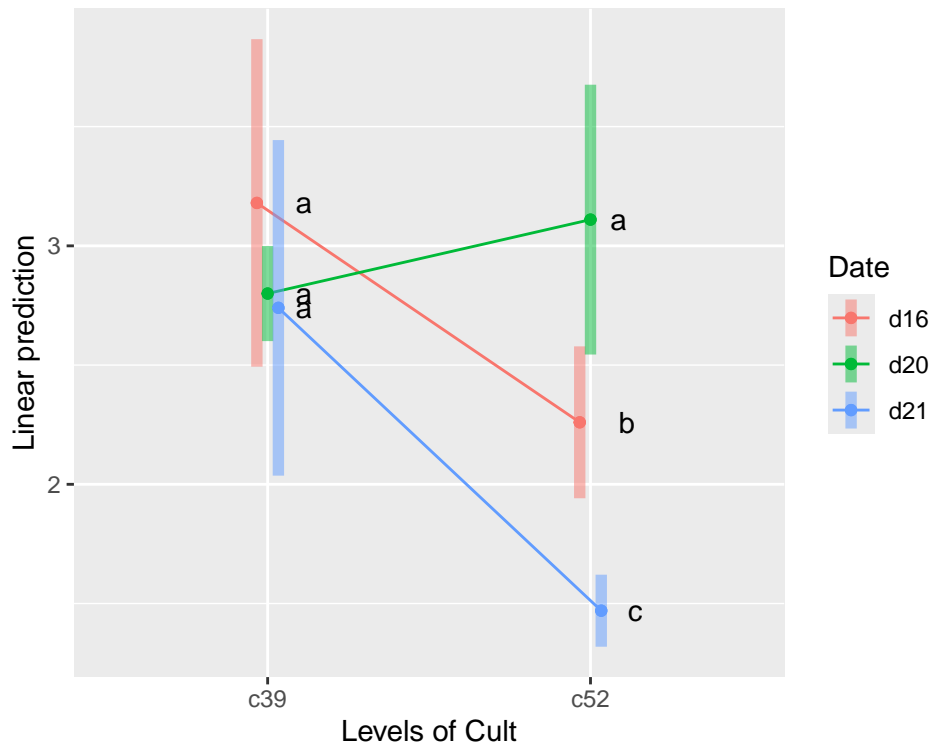
```
## Cult = c39:
## Date emmean SE df lower.CL upper.CL .group
## d16 3.18 0.3025 8.80 2.49 3.87 a
## d20 2.80 0.0882 9.00 2.60 3.00 a
## d21 2.74 0.3110 8.99 2.04 3.44 a
##
## Cult = c52:
## Date emmean SE df lower.CL upper.CL .group
## d20 3.11 0.2501 9.00 2.54 3.68 a
## d16 2.26 0.1408 9.00 1.94 2.58 b
## d21 1.47 0.0668 9.00 1.32 1.62 c
##
## Degrees-of-freedom method: satterthwaite
## Confidence level used: 0.95
## P value adjustment: tukey method for comparing a family of 3 estimates
## significance level used: alpha = 0.05
## NOTE: If two or more means share the same grouping symbol,
## then we cannot show them to be different.
## But we also did not show them to be the same.
```

```
DinC.cld <- as.data.frame(cld(DinC ,decreasing = TRUE,Letters=letters))
```

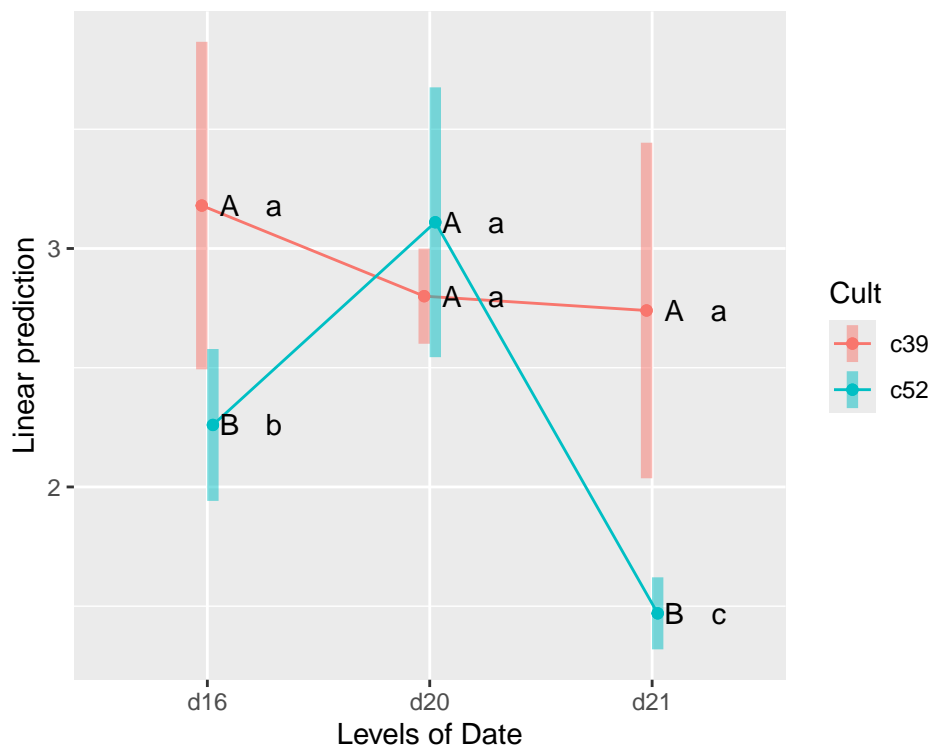
```
# windows()
emmip(DinC, Date~Cult, CIs=TRUE)
```



```
emi.DinC <- emmip(DinC, Date~Cult, CIs=TRUE)
emi.DinC + annotate("text", label=as.character(DinC.cld$.group), x=as.numeric(DinC.cld$Cult)+0.1,
  y=DinC.cld$emmean)
```



```
# windows()
emmi.CinD+
  annotate("text", label=trimws(as.character(CinD.cld$.group)), x=as.numeric(CinD.cld$Date)+0.1,
           y=CinD.cld$emmean)+
  annotate("text", label=trimws(as.character(DinC.cld$.group)), x=as.numeric(DinC.cld$Date)+0.3,
           y=DinC.cld$emmean)
```



## ¿Qué perdimos? ¿Qué ganamos?

¿Por qué ganamos? más potencia, Menos control de error tipo I por experimento!

```
# Análisis de interacción con más potencia, y control de error tipo I por experimento
```

```
ee1 <- emmeans(MHead1, ~Date*Cult)
bb4 <- contrast(ee1, "pairwise", simple="each", combine=TRUE)
bb4
```

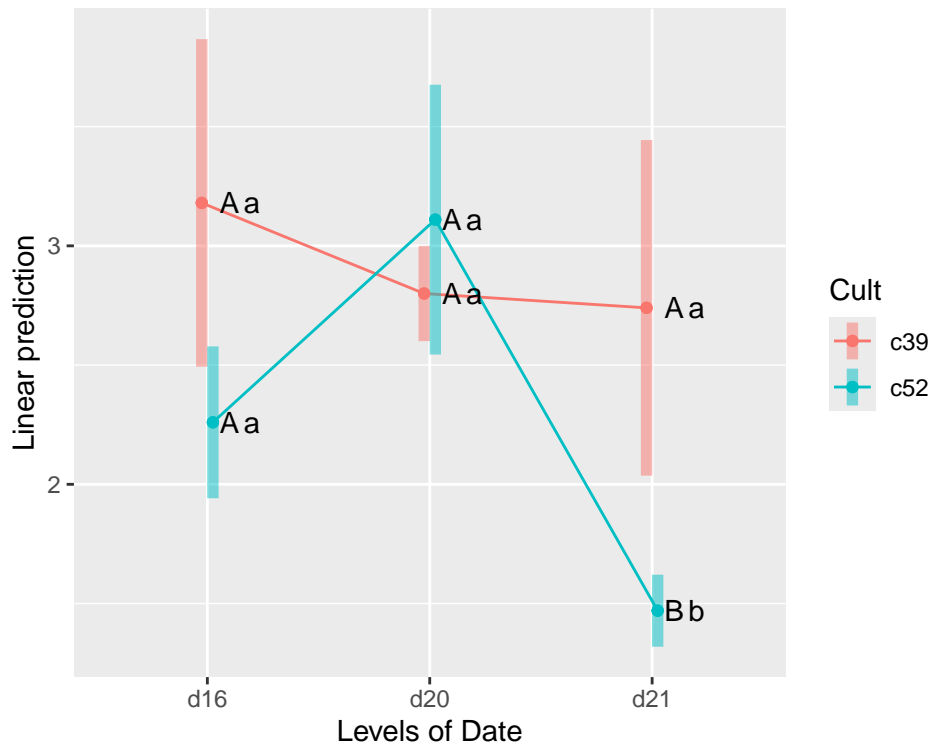
```
## Cult Date contrast estimate SE df t.ratio p.value
## c39 . d16 - d20 0.38 0.315 10.29 1.206 1.0000
## c39 . d16 - d21 0.44 0.434 17.96 1.014 1.0000
## c39 . d20 - d21 0.06 0.323 10.43 0.186 1.0000
## c52 . d16 - d20 -0.85 0.287 14.18 -2.962 0.0917
## c52 . d16 - d21 0.79 0.156 12.85 5.070 0.0020
## c52 . d20 - d21 1.64 0.259 10.27 6.336 0.0007
## . d16 c39 - c52 0.92 0.334 12.46 2.757 0.1517
## . d20 c39 - c52 -0.31 0.265 11.20 -1.169 1.0000
## . d21 c39 - c52 1.27 0.318 9.82 3.993 0.0238
##
## Degrees-of-freedom method: satterthwaite
## P value adjustment: bonferroni method for 9 tests
```

```
# Para el gráfico
```

```
ee2 <- as.data.frame(ee1)
ee2$DinC <- c("a", "a", "a", "a", "a", "b")
ee2$CinD <- c("A", "A", "A", "A", "A", "B")
# ee2
# Date Cult emmean SE df lower.CL upper.CL DinC CinD
# d16 c39 3.18 0.30250881 8.80 2.493274 3.866726 a A
# d20 c39 2.80 0.08819239 9.00 2.600494 2.999506 a A
# d21 c39 2.74 0.31098450 8.99 2.036383 3.443617 a A
# d16 c52 2.26 0.14078974 9.00 1.941502 2.578498 a A
# d20 c52 3.11 0.25008636 9.00 2.544256 3.675744 a A
# d21 c52 1.47 0.06675064 9.00 1.318999 1.621001 b B
```

```
# windows()
```

```
emmi.CinD+
  annotate("text", label=as.character(ee2$DinC), x=as.numeric(ee2$Date)+0.2,
  y=ee2$emmean)+
  annotate("text", label=as.character(ee2$CinD), x=as.numeric(ee2$Date)+0.1,
  y=ee2$emmean)
```



## Resumen

```
contrast(ee1,"pairwise",simple="each",combine=TRUE)
```

```
## Cult Date contrast estimate SE df t.ratio p.value
## c39 . d16 - d20 0.38 0.315 10.29 1.206 1.0000
## c39 . d16 - d21 0.44 0.434 17.96 1.014 1.0000
## c39 . d20 - d21 0.06 0.323 10.43 0.186 1.0000
## c52 . d16 - d20 -0.85 0.287 14.18 -2.962 0.0917
## c52 . d16 - d21 0.79 0.156 12.85 5.070 0.0020
## c52 . d20 - d21 1.64 0.259 10.27 6.336 0.0007
## . d16 c39 - c52 0.92 0.334 12.46 2.757 0.1517
## . d20 c39 - c52 -0.31 0.265 11.20 -1.169 1.0000
## . d21 c39 - c52 1.27 0.318 9.82 3.993 0.0238
##
## Degrees-of-freedom method: satterthwaite
## P value adjustment: bonferroni method for 9 tests
```

```
contrast(ee1,"pairwise",simple="each",adjust="bonferroni")
```

```
## $'simple contrasts for Date'
## Cult = c39:
## contrast estimate SE df t.ratio p.value
## d16 - d20 0.38 0.315 10.3 1.206 0.7645
## d16 - d21 0.44 0.434 18.0 1.014 0.9719
## d20 - d21 0.06 0.323 10.4 0.186 1.0000
##
## Cult = c52:
```

```
## contrast estimate SE df t.ratio p.value
## d16 - d20 -0.85 0.287 14.2 -2.962 0.0306
## d16 - d21 0.79 0.156 12.8 5.070 0.0007
## d20 - d21 1.64 0.259 10.3 6.336 0.0002
##
## Degrees-of-freedom method: satterthwaite
## P value adjustment: bonferroni method for 3 tests
##
## $'simple contrasts for Cult'
## Date = d16:
## contrast estimate SE df t.ratio p.value
## c39 - c52 0.92 0.334 12.46 2.757 0.0169
##
## Date = d20:
## contrast estimate SE df t.ratio p.value
## c39 - c52 -0.31 0.265 11.20 -1.169 0.2667
##
## Date = d21:
## contrast estimate SE df t.ratio p.value
## c39 - c52 1.27 0.318 9.82 3.993 0.0026
##
## Degrees-of-freedom method: satterthwaite
```

```
contrast(ee1,"pairwise",adjust="bonferroni")
```

```
## contrast estimate SE df t.ratio p.value
## d16 c39 - d20 c39 0.38 0.315 10.29 1.206 1.0000
## d16 c39 - d21 c39 0.44 0.434 17.96 1.014 1.0000
## d16 c39 - d16 c52 0.92 0.334 12.46 2.757 0.2529
## d16 c39 - d20 c52 0.07 0.392 17.20 0.178 1.0000
## d16 c39 - d21 c52 1.71 0.310 9.65 5.520 0.0043
## d20 c39 - d21 c39 0.06 0.323 10.43 0.186 1.0000
## d20 c39 - d16 c52 0.54 0.166 15.12 3.250 0.0800
## d20 c39 - d20 c52 -0.31 0.265 11.20 -1.169 1.0000
## d20 c39 - d21 c52 1.33 0.111 16.76 12.025 <.0001
## d21 c39 - d16 c52 0.48 0.341 12.52 1.406 1.0000
## d21 c39 - d20 c52 -0.37 0.399 17.09 -0.927 1.0000
## d21 c39 - d21 c52 1.27 0.318 9.82 3.993 0.0397
## d16 c52 - d20 c52 -0.85 0.287 14.18 -2.962 0.1528
## d16 c52 - d21 c52 0.79 0.156 12.85 5.070 0.0033
## d20 c52 - d21 c52 1.64 0.259 10.27 6.336 0.0011
##
## Degrees-of-freedom method: satterthwaite
## P value adjustment: bonferroni method for 15 tests
```

```
# Si los datos hubiesen sido:
```

```
cabbages$vv<-with(cabbages,ifelse(Date=="d16"&Cult=="c52",HeadWt-0.1,HeadWt))
MHead2 <- gls(vv~Date*Cult,data=cabbages,weights=varIdent(form=~1|Date*Cult))
anova(MHead2)
```

```
## Denom. DF: 54
## numDF F-value p-value
## (Intercept) 1 1886.0039 <.0001
## Date 2 79.2318 <.0001
## Cult 1 9.0526 4e-03
## Date:Cult 2 8.8009 5e-04
```

```
ee1.vv <- emmeans(MHead2, ~ Date*Cult)
contrast(ee1.vv, "pairwise", simple="each", combine=TRUE, adjust="bonferroni")
```

```
## Cult Date contrast estimate SE df t.ratio p.value
## c39 . d16 - d20 0.38 0.315 10.58 1.206 1.0000
## c39 . d16 - d21 0.44 0.434 17.96 1.014 1.0000
## c39 . d20 - d21 0.06 0.323 10.43 0.186 1.0000
## c52 . d16 - d20 -0.95 0.287 14.19 -3.310 0.0457
## c52 . d16 - d21 0.69 0.156 12.85 4.428 0.0063
## c52 . d20 - d21 1.64 0.259 10.28 6.336 0.0007
## . d16 c39 - c52 1.02 0.334 12.78 3.057 0.0840
## . d20 c39 - c52 -0.31 0.265 11.20 -1.169 1.0000
## . d21 c39 - c52 1.27 0.318 9.82 3.993 0.0238
##
## Degrees-of-freedom method: satterthwaite
## P value adjustment: bonferroni method for 9 tests
```

```
contrast(ee1.vv, "pairwise", simple="each", adjust="bonferroni")
```

```
## $'simple contrasts for Date'
## Cult = c39:
## contrast estimate SE df t.ratio p.value
## d16 - d20 0.38 0.315 10.6 1.206 0.7623
## d16 - d21 0.44 0.434 18.0 1.014 0.9719
## d20 - d21 0.06 0.323 10.4 0.186 1.0000
##
## Cult = c52:
## contrast estimate SE df t.ratio p.value
## d16 - d20 -0.95 0.287 14.2 -3.310 0.0152
## d16 - d21 0.69 0.156 12.8 4.428 0.0021
## d20 - d21 1.64 0.259 10.3 6.336 0.0002
##
## Degrees-of-freedom method: satterthwaite
## P value adjustment: bonferroni method for 3 tests
##
## $'simple contrasts for Cult'
## Date = d16:
## contrast estimate SE df t.ratio p.value
## c39 - c52 1.02 0.334 12.78 3.057 0.0093
##
## Date = d20:
## contrast estimate SE df t.ratio p.value
## c39 - c52 -0.31 0.265 11.20 -1.169 0.2667
##
## Date = d21:
## contrast estimate SE df t.ratio p.value
## c39 - c52 1.27 0.318 9.82 3.993 0.0026
##
## Degrees-of-freedom method: satterthwaite
```

```
contrast(ee1.vv, "pairwise", adjust="bonferroni")
```

```
## contrast estimate SE df t.ratio p.value
## d16 c39 - d20 c39 0.38 0.315 10.58 1.206 1.0000
```

```
## d16 c39 - d21 c39      0.44 0.434 17.96    1.014  1.0000
## d16 c39 - d16 c52      1.02 0.334 12.78    3.057  0.1401
## d16 c39 - d20 c52      0.07 0.392 17.45    0.178  1.0000
## d16 c39 - d21 c52      1.71 0.310  9.93    5.520  0.0039
## d20 c39 - d21 c39      0.06 0.323 10.43    0.186  1.0000
## d20 c39 - d16 c52      0.64 0.166 15.12    3.852  0.0232
## d20 c39 - d20 c52     -0.31 0.265 11.20   -1.169  1.0000
## d20 c39 - d21 c52      1.33 0.111 16.76   12.025 <.0001
## d21 c39 - d16 c52      0.58 0.341 12.54    1.699  1.0000
## d21 c39 - d20 c52     -0.37 0.399 17.23   -0.927  1.0000
## d21 c39 - d21 c52      1.27 0.318  9.82    3.993  0.0396
## d16 c52 - d20 c52     -0.95 0.287 14.19   -3.310  0.0762
## d16 c52 - d21 c52      0.69 0.156 12.85    4.428  0.0105
## d20 c52 - d21 c52      1.64 0.259 10.28    6.336  0.0011
##
## Degrees-of-freedom method: satterthwaite
## P value adjustment: bonferroni method for 15 tests
```

```
#####
```

## Ejemplo: problema de CJB 2024

### Contexto biológico

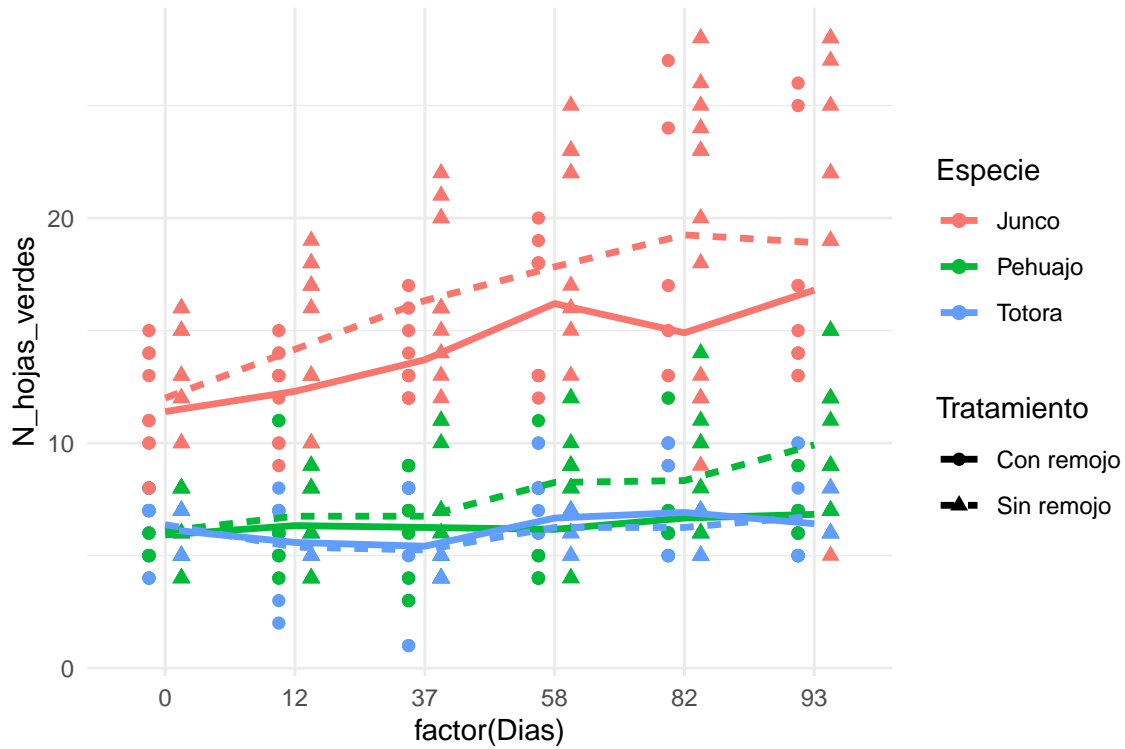
```
# Recuperación de arroyos bonaerenses contaminados mediante la siembra de plantas acuáticas.
# Diseño del ensayo
# Se utilizaron 66 biorrollos de aserrín.
# Algunos recibieron un tratamiento de remojo previo, mientras que otros no.
# En cada biorrollo se sembró una de tres especies de plantas acuáticas,
# y se monitoreó su crecimiento a lo largo del tiempo.
# Preguntas de investigación
# - ¿El efecto del tratamiento varía entre especies?
# - ¿El efecto del tratamiento cambia a lo largo del tiempo?
# - ¿El tratamiento de remojo previo favorece el crecimiento de las plantas?
```

### Los datos

```
load("DatosCJB.RData")
str(datos)
```

```
## 'data.frame':      396 obs. of  7 variables:
## $ ID_biorrollo   : Factor w/ 66 levels "1","2","3","4",...: 1 2 3 4 5 6 7 8 9 10 ...
## $ Especie        : Factor w/ 3 levels "Junco","Pehuajo",...: 1 2 3 1 3 2 2 3 1 2 ...
## $ Tratamiento    : Factor w/ 2 levels "Con remojo","Sin remojo": 1 1 1 1 1 1 1 1 1 1 ...
## $ Raices_debajo  : Factor w/ 2 levels "0","1": 1 1 1 1 1 1 1 1 1 1 ...
## $ Raiz_mas_larga: num  0 0 0 0 0 0 0 0 0 0 ...
## $ N_hojas_verdes: int  14 7 7 11 6 5 5 4 13 8 ...
## $ Dias           : int  0 0 0 0 0 0 0 0 0 0 ...
```

```
# windows()
ggplot(datos, aes(x = factor(Dias), y = N_hojas_verdes,group=Tratamiento))+
  geom_point(aes(color = Especie,shape=Tratamiento),
             position=position_dodge(width=0.5), size = 2) + # Todas las observaciones
  stat_summary(aes(color = Especie, linetype =Tratamiento ,
group = interaction(Especie, Tratamiento)),
              fun = mean, geom = "line", linewidth = 1.2
) + # Línea de medias
  theme_minimal()
```



```
#####
# Notar que en cero no hay, por diseño, efecto de Tratamiento
# En principio es un control del estado inicial de las plantas.
# Una posibilidad para cada especie y tratamiento estudiar
# el "Cambio en número de hojas" respecto al momento inicial
#####
table(datos$Tratamiento,datos$Especie,datos$Dias)
```

```
## , , = 0
##
##
##           Junco Pehuajo Totora
## Con remojo    10     12     12
## Sin remojo    12     12     8
##
## , , = 12
##
##
##           Junco Pehuajo Totora
## Con remojo    10     12     12
```

```

## Sin remajo      12      12      8
##
## , , = 37
##
##
##           Junco Pehuajo Totora
## Con remajo     10      12      12
## Sin remajo     12      12      8
##
## , , = 58
##
##
##           Junco Pehuajo Totora
## Con remajo     10      12      12
## Sin remajo     12      12      8
##
## , , = 82
##
##
##           Junco Pehuajo Totora
## Con remajo     10      12      12
## Sin remajo     12      12      8
##
## , , = 93
##
##
##           Junco Pehuajo Totora
## Con remajo     10      12      12
## Sin remajo     12      12      8

```

```
# Un poco de desbalance pero no extremo
```

```
# Calculamos el cambio en número de hojas verdes
```

```

datos_dif <- datos %>%
  group_by(ID_biorollo) %>%
  mutate(
    N_hojas_verdes_dif = N_hojas_verdes - N_hojas_verdes[Dias == 0]
  ) %>%
  ungroup()
datos_dif <- datos_dif[datos_dif$Dias!=0,]
datos_dif$DiasC <- datos_dif$Dias
datos_dif$Dias <- factor(datos_dif$Dias)
str(datos_dif)

```

```

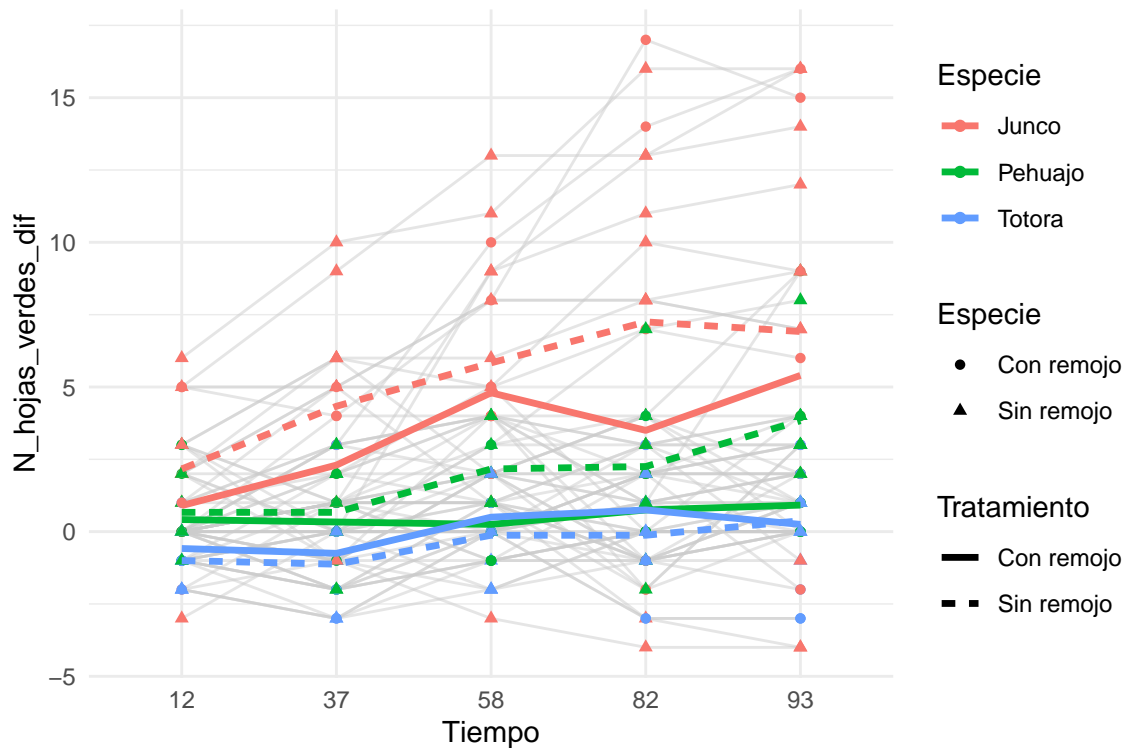
## tibble [330 x 9] (S3: tbl_df/tbl/data.frame)
## $ ID_biorollo      : Factor w/ 66 levels "1","2","3","4",...: 1 2 3 4 5 6 7 8 9 10 ...
## $ Especie         : Factor w/ 3 levels "Junco","Pehuajo",...: 1 2 3 1 3 2 2 3 1 2 ...
## $ Tratamiento     : Factor w/ 2 levels "Con remajo","Sin remajo": 1 1 1 1 1 1 1 1 1 1 ...
## $ Raices_debajo   : Factor w/ 2 levels "0","1": 1 1 1 1 1 1 1 1 1 1 ...
## $ Raiz_mas_larga  : num [1:330] 0 0 0 0 0 0 0 0 0 0 ...
## $ N_hojas_verdes  : int [1:330] 14 6 6 9 6 5 4 3 13 11 ...
## $ Dias            : Factor w/ 5 levels "12","37","58",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ N_hojas_verdes_dif: int [1:330] 0 -1 -1 -2 0 0 -1 -1 0 3 ...
## $ DiasC           : int [1:330] 12 12 12 12 12 12 12 12 12 12 ...

```

```

# windows()
ggplot(datos_dif, aes(x = Dias, y = N_hojas_verdes_dif))+
  geom_line(aes(group = ID_biorollo), color = "grey80", linewidth = 0.5, alpha = 0.5) +
  geom_point(aes(color = Especie, shape=Tratamiento), size = 1.5) + # Todas las observaciones
  stat_summary(aes(color = Especie, linetype =Tratamiento ,
                  group = interaction(Especie, Tratamiento)),
              fun = mean, geom = "line", linewidth = 1.2) + # Línea de medias
  theme_minimal()+
  labs(x = "Tiempo",
       y = "N_hojas_verdes_dif",
       color = "Especie",
       shape = "Especie",
       linetype = "Tratamiento")

```



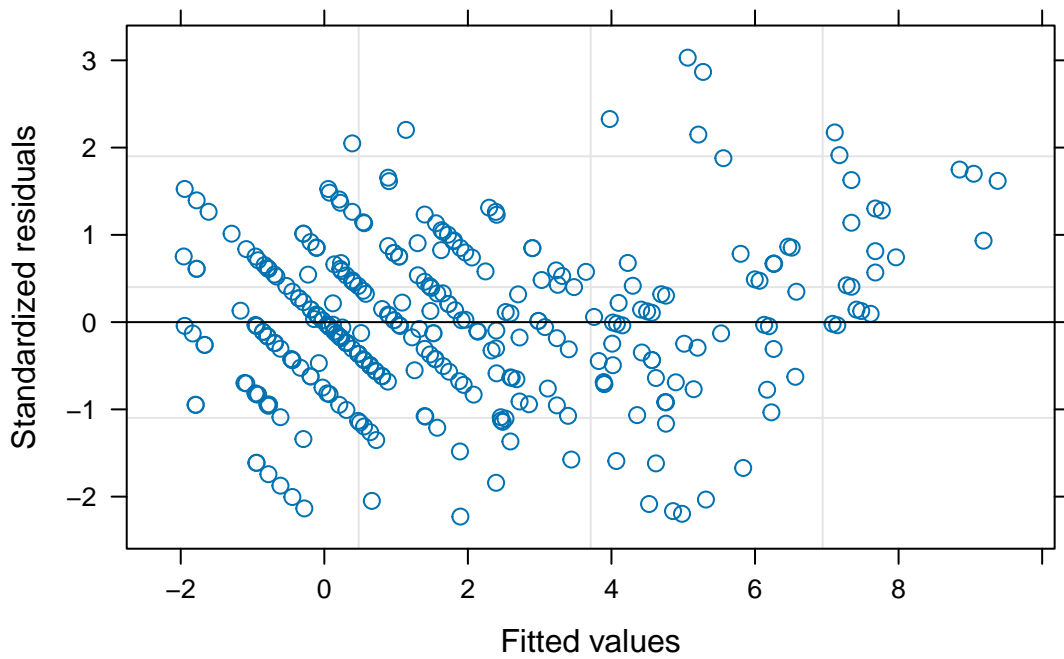
## Análisis estadístico

```

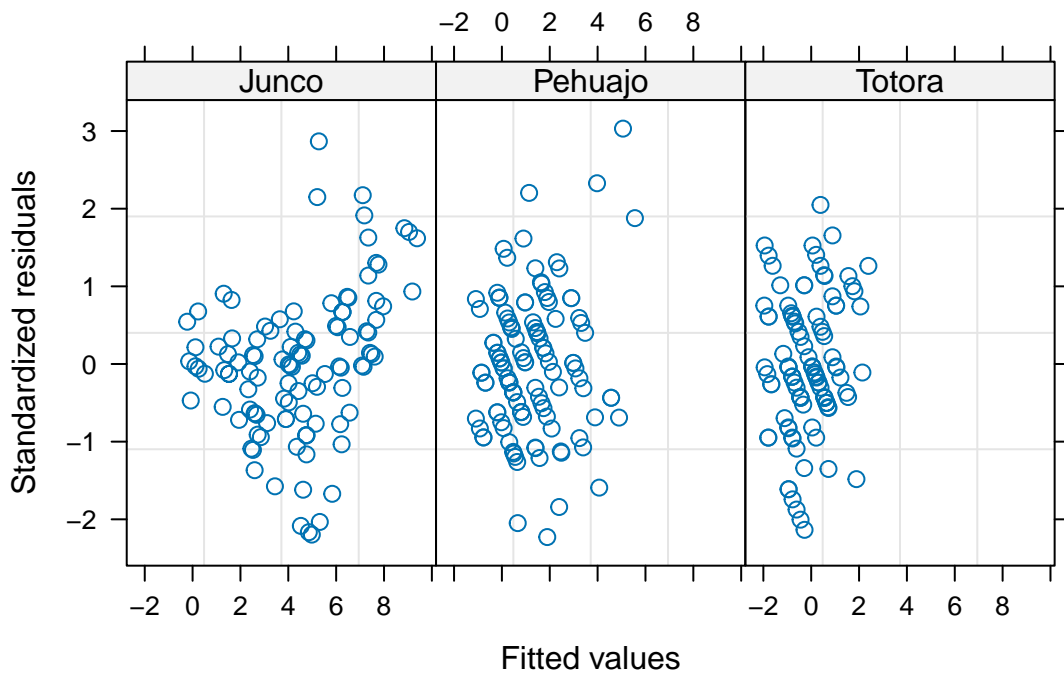
# modelo discreto en el tiempo
modelo_aumhojas <- lme(N_hojas_verdes_dif ~
                      Especie*Tratamiento*Dias,
                      random=~1|ID_biorollo,
                      weights=varIdent(form=~1|Especie),
                      data = datos_dif)

plot(modelo_aumhojas)

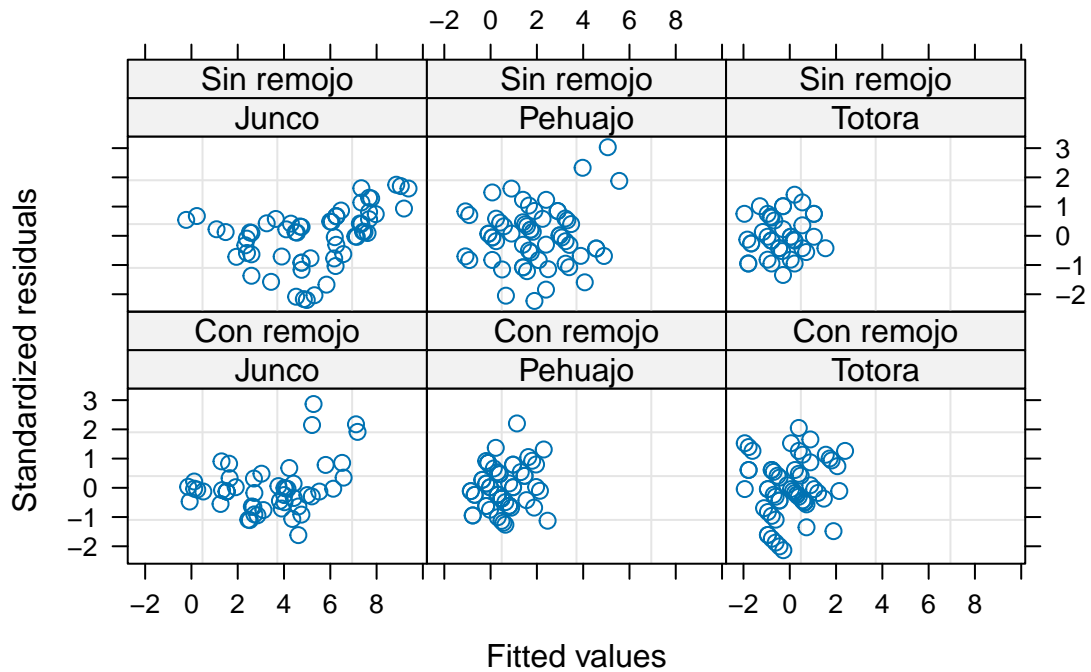
```



```
plot(modelo_aumhojas, resid(., type="normalized")~fitted(.)|Especie)
```



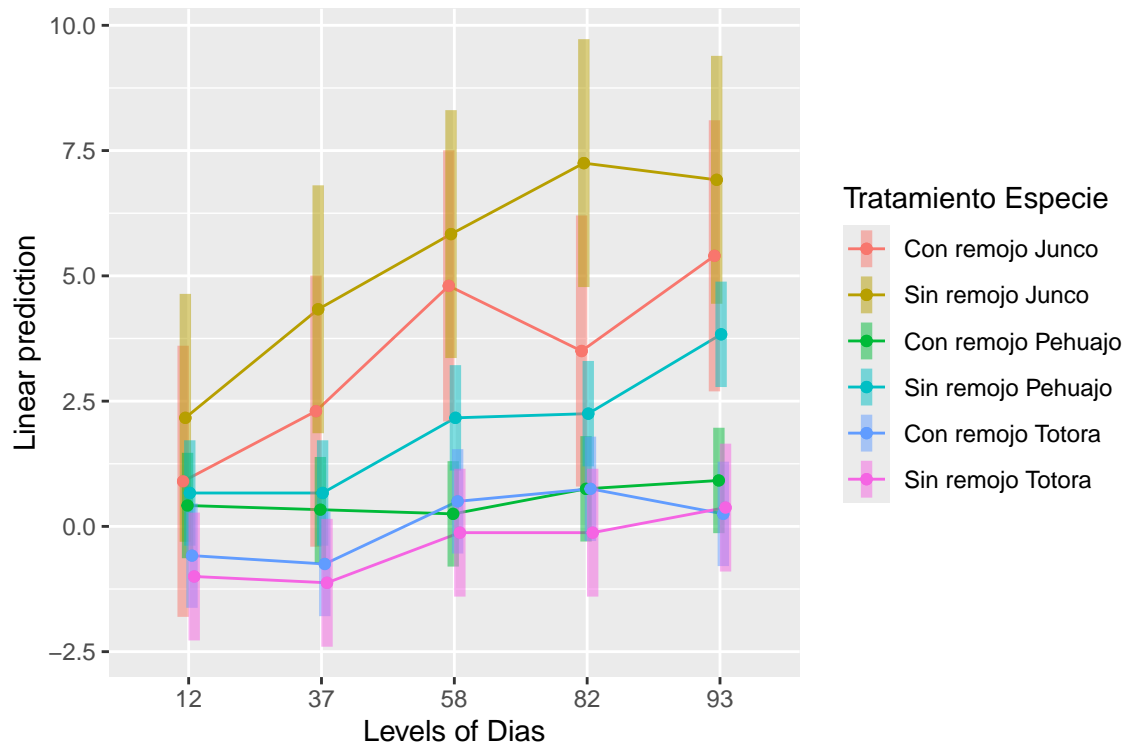
```
plot(modelo_aumhojas, resid(., type="normalized")~fitted(.)|Especie*Tratamiento)
```



```
anova(modelo_aumhojas)
```

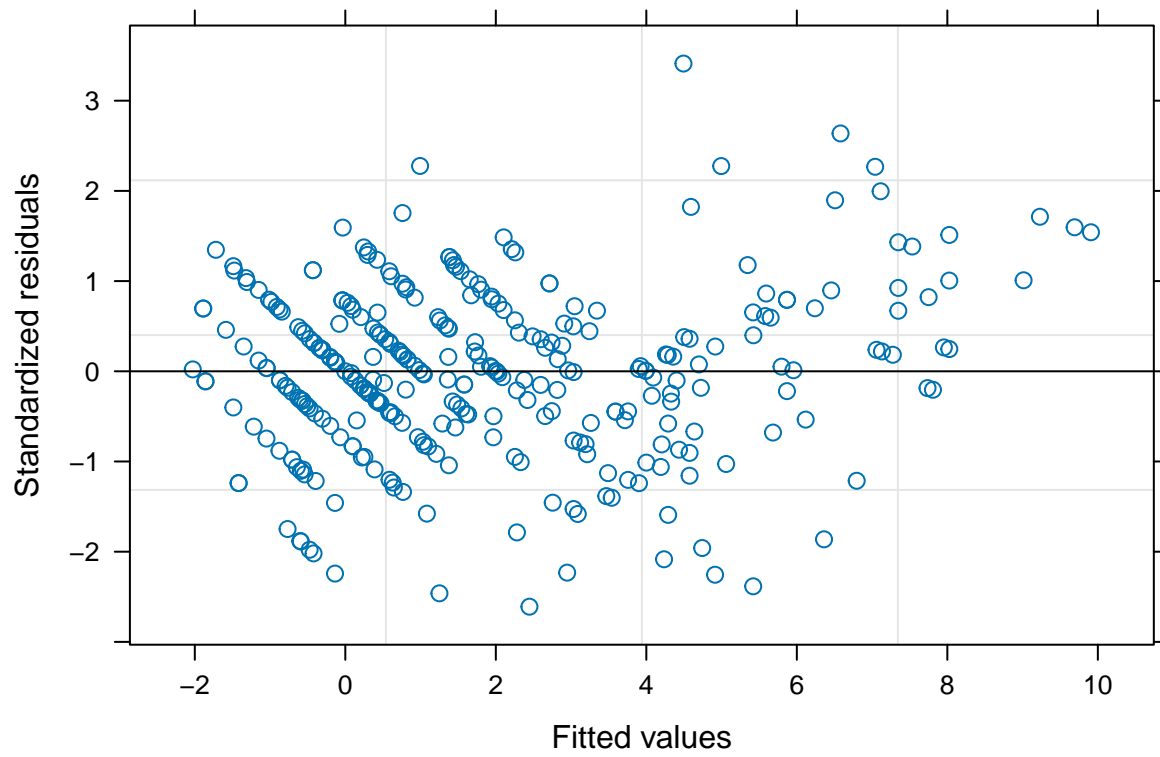
	numDF	denDF	F-value	p-value
## (Intercept)	1	240	40.77750	<.0001
## Especie	2	60	32.20388	<.0001
## Tratamiento	1	60	4.19035	0.0450
## Dias	4	240	14.73520	<.0001
## Especie:Tratamiento	2	60	3.07351	0.0536
## Especie:Dias	8	240	1.70016	0.0990
## Tratamiento:Dias	4	240	2.91536	0.0221
## Especie:Tratamiento:Dias	8	240	1.17516	0.3147

```
emmip(modelo_aumhojas, Tratamiento:Especie~Dias, CIs=TRUE)
```

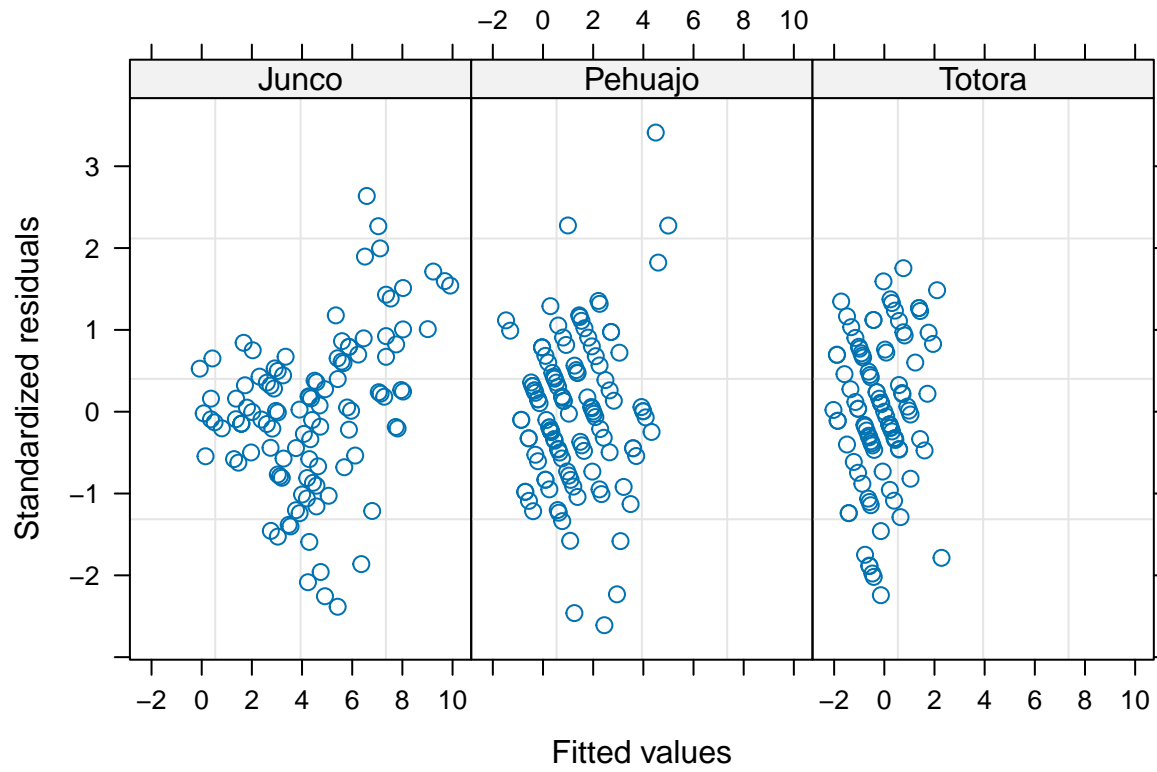


```
# modelo continuo en los dias
modelo_reg <- lme(N_hojas_verdes_dif ~
  Especie * Tratamiento * DiasC,
  random=~1|ID_biorollo,
  weights=varIdent(form=~1|Especie),
  data = datos_dif)

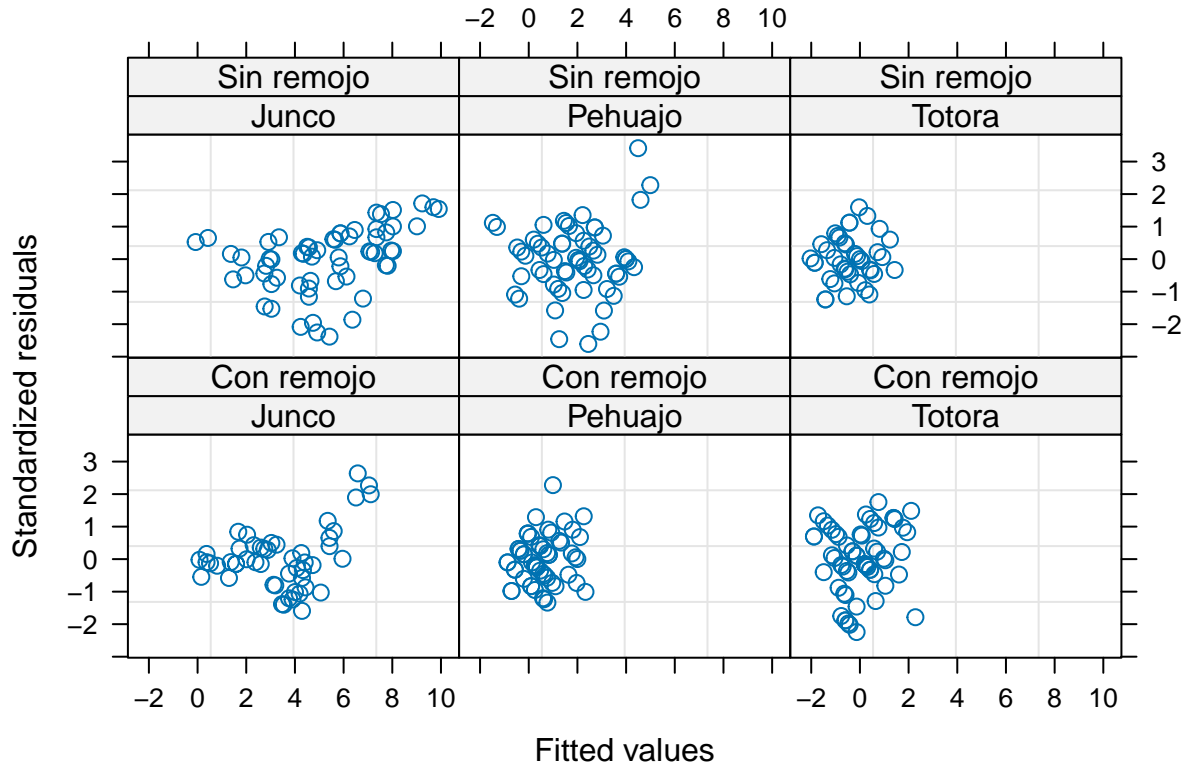
plot(modelo_reg)
```



```
plot(modelo_reg, resid(., type="normalized")~fitted(.)|Especie)
```



```
plot(modelo_reg, resid(., type="normalized")~fitted(.)|Especie*Tratamiento)
```

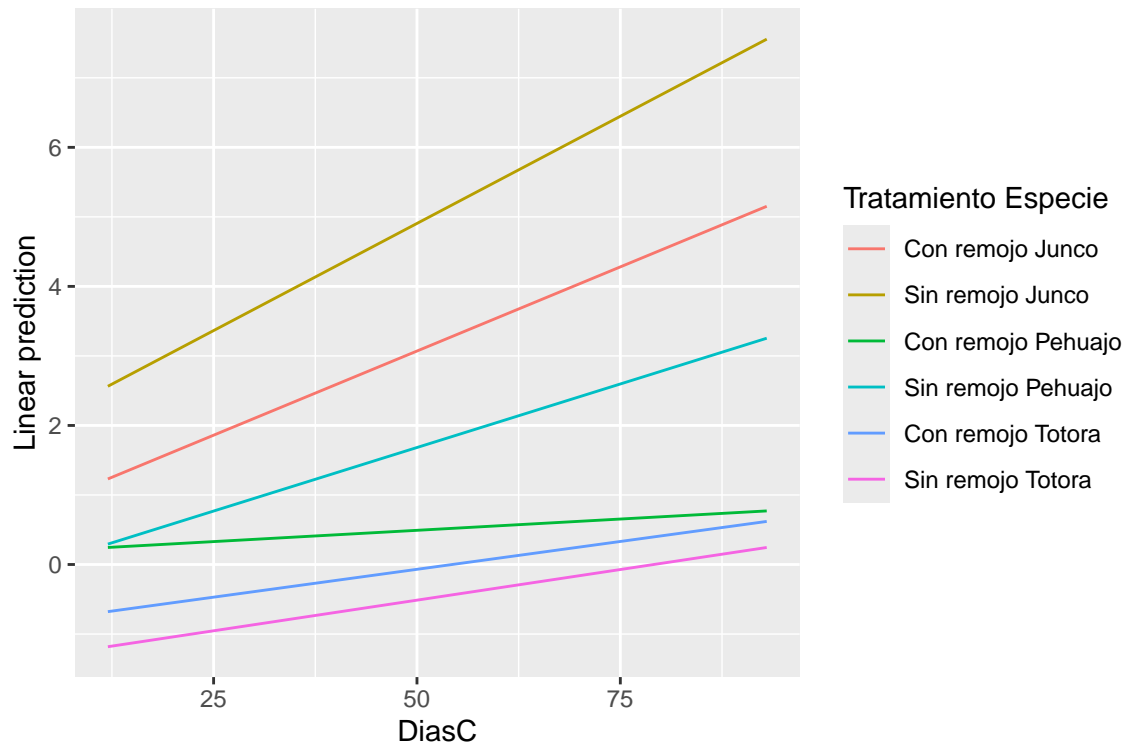


```
anova(modelo_reg)
```

	numDF	denDF	F-value	p-value
## (Intercept)	1	258	41.48911	<.0001
## Especie	2	60	31.89067	<.0001
## Tratamiento	1	60	4.11184	0.0470
## DiasC	1	258	53.05977	<.0001
## Especie:Tratamiento	2	60	2.92029	0.0616
## Especie:DiasC	2	258	4.17939	0.0164
## Tratamiento:DiasC	1	258	8.20330	0.0045
## Especie:Tratamiento:DiasC	2	258	2.80330	0.0625

```
#####
# Responde parcialmente a las preguntas 1 y 2.
# El efecto del tratamiento no es el mismo la largo del tiempo ni entre las especies.
#####
```

```
emmip(modelo_reg, Tratamiento | Especie ~ DiasC, cov.reduce=range)
```



```
emtrends(modelo_reg, pairwise ~ Tratamiento | Especie, var = "DiasC")
```

```
## $emtrends
## Especie = Junco:
## Tratamiento DiasC.trend SE df lower.CL upper.CL
## Con remojo 0.04841 0.01896 258 0.01108 0.0857
## Sin remojo 0.06164 0.01731 258 0.02756 0.0957
##
## Especie = Pehuajo:
## Tratamiento DiasC.trend SE df lower.CL upper.CL
## Con remojo 0.00649 0.00578 258 -0.00490 0.0179
## Sin remojo 0.03655 0.00578 258 0.02516 0.0479
##
## Especie = Totora:
## Tratamiento DiasC.trend SE df lower.CL upper.CL
## Con remojo 0.01602 0.00559 258 0.00501 0.0270
## Sin remojo 0.01762 0.00685 258 0.00413 0.0311
##
## Degrees-of-freedom method: containment
## Confidence level used: 0.95
##
## $contrasts
## Especie = Junco:
## contrast estimate SE df t.ratio p.value
## Con remojo - Sin remojo -0.0132 0.02567 258 -0.515 0.6068
##
## Especie = Pehuajo:
## contrast estimate SE df t.ratio p.value
## Con remojo - Sin remojo -0.0301 0.00818 258 -3.676 0.0003
##
## Especie = Totora:
```

```
## contrast          estimate      SE  df t.ratio p.value
## Con remojo - Sin remojo -0.0016 0.00884 258 -0.181 0.8566
##
## Degrees-of-freedom method: containment
```

```
# corrección de Bonferroni, trabajar cada comparación con alfa=
0.05/3
```

```
## [1] 0.01666667
```

```
emtrends(modelo_reg, pairwise~ Especie|Tratamiento, var="DiasC", adjust="tukey")
```

```
## $emtrends
## Tratamiento = Con remojo:
## Especie DiasC.trend      SE  df lower.CL upper.CL
## Junco      0.04841 0.01896 258 0.01108 0.0857
## Pehuajo    0.00649 0.00578 258 -0.00490 0.0179
## Tatora     0.01602 0.00559 258 0.00501 0.0270
##
```

```
## Tratamiento = Sin remojo:
## Especie DiasC.trend      SE  df lower.CL upper.CL
## Junco      0.06164 0.01731 258 0.02756 0.0957
## Pehuajo    0.03655 0.00578 258 0.02516 0.0479
## Tatora     0.01762 0.00685 258 0.00413 0.0311
##
```

```
## Degrees-of-freedom method: containment
```

```
## Confidence level used: 0.95
```

```
##
```

```
## $contrasts
```

```
## Tratamiento = Con remojo:
## contrast          estimate      SE  df t.ratio p.value
## Junco - Pehuajo    0.04192 0.01982 258 2.115 0.0888
## Junco - Tatora     0.03239 0.01976 258 1.639 0.2313
## Pehuajo - Tatora  -0.00953 0.00804 258 -1.185 0.4633
##
```

```
## Tratamiento = Sin remojo:
## contrast          estimate      SE  df t.ratio p.value
## Junco - Pehuajo    0.02508 0.01825 258 1.375 0.3558
## Junco - Tatora     0.04402 0.01861 258 2.365 0.0490
## Pehuajo - Tatora   0.01894 0.00896 258 2.113 0.0893
##
```

```
## Degrees-of-freedom method: containment
```

```
## P value adjustment: tukey method for comparing a family of 3 estimates
```

```
#####
# Responde parcialmente a las preguntas 2 y 3.
# El efecto de tratamiento cambia a lo largo del tiempo solo
# para la especie "Pehuajó".
# Para las especies "Junco" y "Tatora", si existe efecto del tratamiento no cambia a lo largo
# de los días
#####
```

```
# El efecto del tratamiento cambia en el tiempo, solo para la especie Pehuajó.
#####
mediasR <- emmeans(modelo_reg, ~Tratamiento*Especie*DiasC)
pairs(mediasR, simple=list("Tratamiento"), combine=TRUE) # una comparación no adecuada
```

```
## Especie DiasC contrast          estimate    SE df t.ratio p.value
## Junco  56.4  Con remojo - Sin remojo  -1.920 0.947 60  -2.027  0.1414
## Pehuajo 56.4  Con remojo - Sin remojo  -1.383 0.594 60  -2.327  0.0701
## Tatora 56.4  Con remojo - Sin remojo   0.433 0.661 60   0.656  1.0000
##
## Degrees-of-freedom method: containment
## P value adjustment: bonferroni method for 3 tests
```

```
mediasR_filtrado <- subset(mediasR, Especie != "Pehuajo")
pairs(mediasR_filtrado, simple=list("Tratamiento"), combine=TRUE)
```

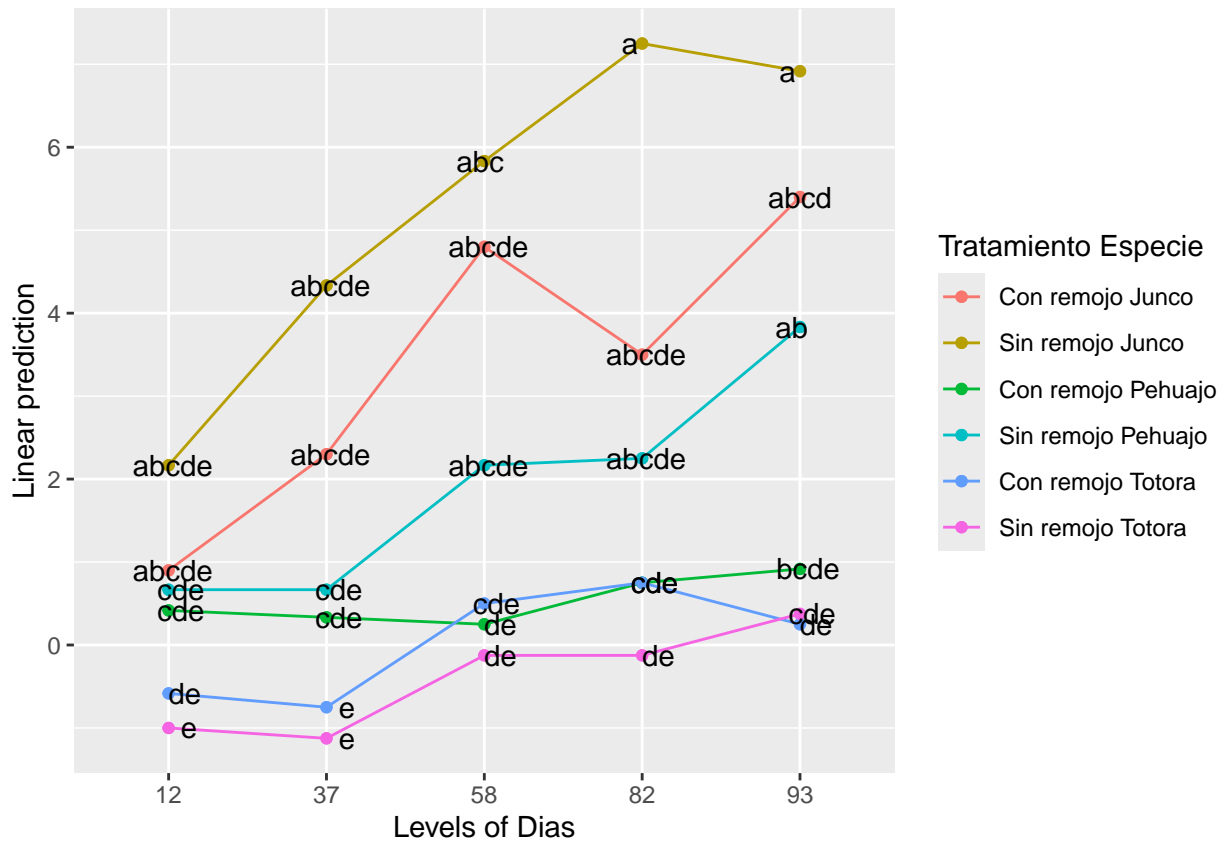
```
## Especie DiasC contrast          estimate    SE df t.ratio p.value
## Junco  56.4  Con remojo - Sin remojo  -1.920 0.947 60  -2.027  0.0943
## Tatora 56.4  Con remojo - Sin remojo   0.433 0.661 60   0.656  1.0000
##
## Degrees-of-freedom method: containment
## P value adjustment: bonferroni method for 2 tests
```

```
#####
# Respondiendo a las preguntas,
# tanto en "Junco" como en "Tatora, el tratamiento no afecta de forma diferencial al
# crecimiento de las plantas,
# en "Pehuajó" la tendencia en crecimiento del número de hojas en plantas con el tratamiento
# "sin remojo" fue mayor que en el tratamiento "con remojo".
# El remojo no mejoró el crecimiento de las plantas de ninguna especie.
#####
```

```
#####
# Nos faltaría probar si la pendiente en "Junco" y la pendiente en "Tatora" difieren de cero,
# para ver si las plantas están creciendo en los biorrollos, pero no era una de las
# preguntas específicas.
#####
```

```
#####
# Una vez detectada la interacción una comparación de todos contra todos
#####
```

```
TcT <- emmeans(modelo_aumhojas, ~Tratamiento*Especie*Dias)
TcT.g <- cld(TcT, adjust="bonferroni", Letters=letters, decreasing = TRUE)
TcT.df <- as.data.frame(TcT.g)
Gtct <- emmip(TcT, Tratamiento:Especie ~Dias)
Gtct+
  annotate("text", label=as.character(TcT.df$.group), x=as.numeric(TcT.df$Dias),
          y=TcT.df$emmean)
```



# Muy difícil sacar respuestas claras.