

Beispiel 8.4. Einfluss einer Vegetationsentfernung auf die Molchlarven in Auengewässern

Kapitel 8.3.1 aus Henle, K., A. Grimm-Seyfarth & B. Gruber: Erfassung und Analyse von Tierpopulationen. Ulmer Verlag

Annegret Grimm-Seyfarth

2025-04-22

In diesem Beispiel zeigen wir eine einfache BACI-Analyse in R. Für klassische Regressionsanalysen mittels *lm* und *glm* sind keine weiteren Pakete nötig. Allerdings erleichtern die Pakete MuMIn (Bartón 2024) und emmeans (Lenth 2025) den Umgang mit Modellselektion und Interaktionen. Außerdem nutzen wir das Paket ggplot2 für Visualisierungen (Wickham 2016). Als Datensatz nutzen wir ein Larvenmonitoring im Leipziger Auwald, nachdem ein Gewässer von Vegetation freigeschnitten wurde.

```
# check.packages function: install and load multiple R packages.
# Function from: https://gist.github.com/smithdanielle/9913897
check.packages <- function(pkg){
  new.pkg <- pkg[!(pkg %in% installed.packages()[, "Package"])]
  if (length(new.pkg))
    install.packages(new.pkg, dependencies = TRUE, type = "source")
  sapply(pkg, require, character.only = TRUE)
}

# benoetigte R pakete
pakete <- c("MuMIn", "emmeans", "ggplot2")

# Pruefe und installiere
check.packages(pakete)
```

```
## MuMIn emmeans ggplot2
## TRUE TRUE TRUE
```

Daten einladen

Im Leipziger Auwald leben viele Amphibien in den temporären Gewässern (Lachen). Im Zuge eines Renaturierungsprojektes stehen die Larven von Kamm- (*Triturus cristatus*) und Teichmolchen (*Lissotriton vulgaris*) im Fokus. Im Herbst 2023 wurde ein Gewässer im Zuge einer Pflegemaßnahme für eine Oberleitung komplett von Vegetation freigeschnitten. Wir wollen nun die Auswirkungen auf den Reproduktionserfolg der beiden Molcharten testen. Dazu nutzen wir neben dem Einflussgewässer noch ein weiteres in naher Umgebung, das jedoch weit genug weg ist, dass es von der Maßnahme nicht betroffen sein dürfte.

Beide Gewässer wurden jeweils 2023 und 2024 im Juni für fünf aufeinanderfolgende Tage mittels Minnow Traps befangen. Da jeweils die gleiche Fallenzahl genutzt wurde (6 Fallen, jeweils 2 Öffnungen), können

die Daten direkt als relative Häufigkeiten verglichen werden. Andernfalls hätte man einen Index berechnen müssen, der die Anzahl Fallen bzw. Fallenöffnungen und Fangtage berücksichtigt.

Zunächst laden wir die Daten ein:

```
BAdata <- read.csv2("extdata/Burgaue_Larven.csv")
head(BAdata)
```

```
##   Saison      Date      Group Time LV_larvae_Numbers_of_Individual
## 1      3 15.06.2023 Control1 Before                      1
## 2      3 15.06.2023 Treatment Before                     1
## 3      3 16.06.2023 Control1 Before                      1
## 4      3 16.06.2023 Treatment Before                     1
## 5      3 17.06.2023 Control1 Before                      1
## 6      3 17.06.2023 Treatment Before                     2
##   TC_larvae_Numbers_of_Individual
## 1                                2
## 2                                1
## 3                                2
## 4                                1
## 5                                1
## 6                                0
```

Das Kontrollgewässer wurde hier als “Control1”, das Einflussgewässer als “Treatment” benannt. Als Zeit gilt 2023 als “Before” und 2024 als “After”.

Da in R standardmäßig alle Faktoren alphanumerisch sortiert werden, müssen wir für die Zeit die Sortierung ändern.

```
BAdata$Time <- factor(BAdata$Time, levels = c("Before", "After"))
```

Generalisiertes lineares Modell

Da wir hier als Ergebnisvariable Zähldaten verwenden, sind die Residuen in unserem Modell Poissonverteilt. Wir bauen nun die beiden Modelle mit der Interaktion Group und Time, dargestellt durch ein “*“.

```
# Teichmolch
glm.lv <- glm(LV_larvae_Numbers_of_Individual ~ Group * Time,
              data = BAdata, family = "poisson")
# Kammolch
glm.tc <- glm(TC_larvae_Numbers_of_Individual ~ Group * Time,
              data = BAdata, family = "poisson")
```

Schauen wir uns die Arten im Detail an.

Teichmolch

Vergleichen wir zunächst die Modelle mit jeweils linearen Kombinationen der Parameter. Dazu nutzen wir die *dredge* Funktion aus dem R-Paket MuMIn.

```
# Modellvergleich
```

```
glm.lv.mod <- dredge(glm.lv, options(na.action = "na.fail"))
glm.lv.mod
```

```
## Global model call: glm(formula = LV_larvae_Numbers_of_Individual ~ Group * Time,
##      family = "poisson", data = BAdata)
## ---
## Model selection table
##      (Int) Grp Tim Grp:Tim df  logLik  AICc delta weight
## 4 -0.68960  +   +         3 -33.412  74.3  0.00  0.716
## 8 -0.22310  +   +         + 4 -32.756  76.2  1.85  0.283
## 2  0.09531  +           2 -42.316  89.3 15.01  0.000
## 3  0.26240      +         2 -44.962  94.6 20.31  0.000
## 1  1.04700           1 -53.866 110.0 35.63  0.000
## Models ranked by AICc(x)
```

Wir sehen, dass die lineare Kombination aus Group und Time besser abschneidet als die Interaktion. Allerdings ist das Modell mit Interaktion nur 1,85 AICc Punkte hinten - ein Einfluss der Intervention (Freischneidung der Fläche) kann also nicht ausgeschlossen werden. Eine Einführung in Modellvergleiche findet sich in Kapitel 9.3 des Buches.

Schauen wir uns die Schätzwerte der beiden besten Modelle an.

```
get.models(glm.lv.mod, subset = 1:2)
```

```
## $'4'
##
## Call:  glm(formula = LV_larvae_Numbers_of_Individual ~ Group + Time +
##      1, family = "poisson", data = BAdata)
##
## Coefficients:
##      (Intercept)  GroupTreatment      TimeAfter
##          -0.6896         1.4307         1.2192
##
## Degrees of Freedom: 19 Total (i.e. Null);  17 Residual
## Null Deviance:      60.44
## Residual Deviance: 19.54      AIC: 72.82
##
## $'8'
##
## Call:  glm(formula = LV_larvae_Numbers_of_Individual ~ Group + Time +
##      Group:Time + 1, family = "poisson", data = BAdata)
##
## Coefficients:
##      (Intercept)      GroupTreatment      TimeAfter
##          -0.2231         0.8109         0.5596
## GroupTreatment:TimeAfter
##          0.8541
##
## Degrees of Freedom: 19 Total (i.e. Null);  16 Residual
## Null Deviance:      60.44
## Residual Deviance: 18.22      AIC: 73.51
##
```

```
## attr("rank")
## function (x)
## do.call("rank", list(x))
## <environment: 0x0000022a7f3bf808>
## attr("call")
## AICc(x)
## attr("class")
## [1] "function"      "rankFunction"
## attr("beta")
## [1] "none"
```

Im "Treatment" Gewässer gab es mehr Teichmolchlarven, ebenso 2024 ggü. 2023.

Schauen wir noch nach möglicher Signifikanz der Parameter

```
drop1(glm.lv, test="Chi")
```

```
## Single term deletions
##
## Model:
## LV_larvae_Numbers_of_Individual ~ Group * Time
##           Df Deviance   AIC    LRT Pr(>Chi)
## <none>           18.222 73.511
## Group:Time  1    19.535 72.824 1.3128  0.2519
```

```
drop1(glm(LV_larvae_Numbers_of_Individual ~ Group + Time,
          data = BAdata, family = "poisson"),
      test="Chi")
```

```
## Single term deletions
##
## Model:
## LV_larvae_Numbers_of_Individual ~ Group + Time
##           Df Deviance   AIC    LRT  Pr(>Chi)
## <none>           19.535 72.824
## Group    1    42.635 93.924 23.100 1.538e-06 ***
## Time     1    37.344 88.632 17.808 2.443e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Während die Gesamtinteraktion nicht signifikant ist, sind es die Einzelparame-
ter. Wenn es einen Interaktionseffekt (also Einfluss der Intervention) gibt, ist dieser schwächer als die Effekte der einzelnen Parameter.

Interaktionseffekte

Die einzelnen Interaktionseffekte können wir uns auch mithilfe des emmeans Paketes anschauen.

```
l.em <- emmeans(glm.lv, ~ Group * Time)
l.em
```

```
## Group    Time  emmean    SE df asymp.LCL asymp.UCL
## Control1 Before -0.223 0.500 Inf   -1.2031    0.757
```

```
## Treatment Before 0.588 0.333 Inf -0.0655 1.241
## Control1 After 0.336 0.378 Inf -0.4043 1.077
## Treatment After 2.001 0.164 Inf 1.6793 2.324
##
## Results are given on the log (not the response) scale.
## Confidence level used: 0.95
```

Im Treatment Gewässer waren also auch 2023 mehr Larven als 2024, jedoch ist der Unterschied 2024 deutlich stärker zu sehen. In beiden Gewässern gibt es 2024 allerdings mehr Larven als 2023, weshalb der mögliche positive Effekt der Pflegemaßnahme maskiert sein könnte.

Graphische Darstellung

```
ggplot(BAdata, aes(x = Group, y = LV_larvae_Numbers_of_Individual,
                    fill = Time, group = interaction(Group, Time))) +
  geom_boxplot() +
  labs(
    x = "",
    y = "",
    title = "Veränderung der Teichmolchlarven vor und nach dem Eingriff",
    fill = "Zeit"
  ) +
  theme_classic() +
  scale_fill_manual(
    values = c("Before" = "yellowgreen", "After" = "seagreen") # Anpassung der Farben
  ) +
  theme(
    text = element_text(family = "Times"),

    # Achsentitel
    axis.title.x = element_text(size = 18, color = "black"),
    axis.title.y = element_text(size = 18, color = "black"),

    # Achsenbeschriftungen
    axis.text.x = element_text(size = 18, color = "black"),
    axis.text.y = element_text(size = 18, color = "black"),

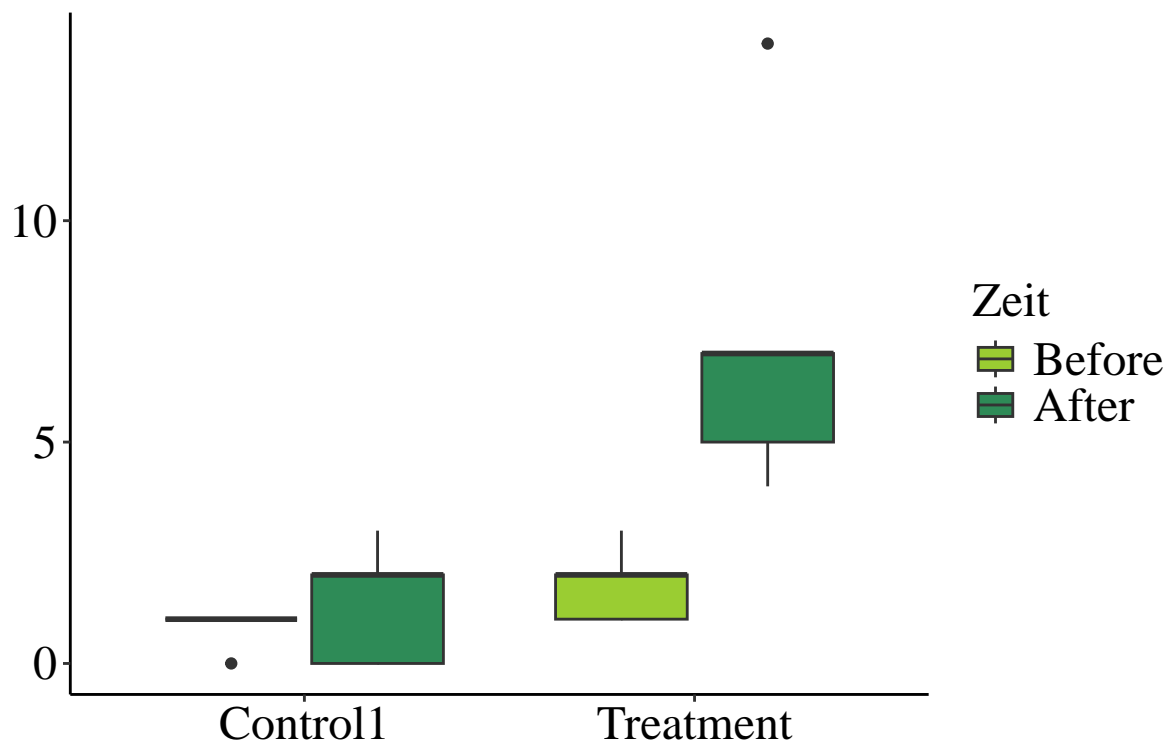
    # Titel
    plot.title = element_text(size = 14, face = "bold", hjust = 0.5, color = "black"),

    #legend.key.size = unit(1.5, "cm"),

    # Ränder
    #plot.margin = margin(t = 50, r = 20, b = 20, l = 20),

    # Legende
    legend.title = element_text(size = 18, color = "black"),
    legend.text = element_text(size = 18, color = "black"),
    legend.position = "right"
  )
```

Veränderung der Teichmolchlarven vor und nach dem Eingriff



Dieses Ergebnis deckt sich mit den Modellergebnissen. Achtung: in einem Boxplot werden keine Mittelwerte, sondern Mediane dargestellt, weshalb der Zahlenwert nicht mit denen aus emmeans überein stimmen kann.

Kammmolch

Vergleichen wir zunächst die Modelle mit jeweils linearen Kombinationen der Parameter. Dazu nutzen wir die *dredge* Funktion aus dem R-Paket MuMIn.

```
# Modellvergleich
glm.tc.mod <- dredge(glm.tc, options(na.action = "na.fail"))
glm.tc.mod

## Global model call: glm(formula = TC_larvae_Numbers_of_Individual ~ Group * Time,
##   family = "poisson", data = BAdata)
## ---
## Model selection table
##      (Int) Grp Tim Grp:Tim df  logLik AICc delta weight
## 1  0.13980      +      1 -32.470 67.2  0.00  0.428
## 2 -0.10540      +      2 -31.922 68.5  1.39  0.214
## 8  0.33650      +      4 -29.122 68.9  1.75  0.178
## 3  0.09531      +      2 -32.448 69.6  2.44  0.126
## 4 -0.14980      +      3 -31.900 71.3  4.14  0.054
## Models ranked by AICc(x)
```

Wir sehen, dass es hier keinen statistischen Unterschied zwischen dem Null-Modell (ohne Parameter), der

Group bzw. deren Interaktion mit Time gibt. Alle Modelle liegen nur wenige AICc Punkte auseinander. Ein Einfluss der Intervention (Freischneidung der Fläche) kann also nicht ausgeschlossen werden.

Schauen wir uns die Schätzwerte der drei besten Modelle an.

```
get.models(glm.tc.mod, subset = c(1:3))
```

```
## $'1'
##
## Call: glm(formula = TC_larvae_Numbers_of_Individual ~ 1, family = "poisson",
## data = BAdata)
##
## Coefficients:
## (Intercept)
## 0.1398
##
## Degrees of Freedom: 19 Total (i.e. Null); 19 Residual
## Null Deviance: 35.16
## Residual Deviance: 35.16 AIC: 66.94
##
## $'2'
##
## Call: glm(formula = TC_larvae_Numbers_of_Individual ~ Group + 1, family = "poisson",
## data = BAdata)
##
## Coefficients:
## (Intercept) GroupTreatment
## -0.1054 0.4418
##
## Degrees of Freedom: 19 Total (i.e. Null); 18 Residual
## Null Deviance: 35.16
## Residual Deviance: 34.06 AIC: 67.84
##
## $'8'
##
## Call: glm(formula = TC_larvae_Numbers_of_Individual ~ Group + Time +
## Group:Time + 1, family = "poisson", data = BAdata)
##
## Coefficients:
## (Intercept) GroupTreatment TimeAfter
## 0.3365 -0.5596 -1.2528
## GroupTreatment:TimeAfter
## 2.1691
##
## Degrees of Freedom: 19 Total (i.e. Null); 16 Residual
## Null Deviance: 35.16
## Residual Deviance: 28.47 AIC: 66.24
##
## attr("rank")
## function (x)
## do.call("rank", list(x))
## <environment: 0x0000022a815a0080>
## attr("call")
## AICc(x)
```

```
## attr("class")
## [1] "function"      "rankFunction"
## attr("beta")
## [1] "none"
```

Im "Treatment" Gewässer gab es insgesamt über beide Jahre mehr Teichmolchlarven. Der Zeiteffekt steckt jedoch in der Interaktion, dies ist schwierig zu interpretieren - zur Sicherheit schaut man sich unten die Interaktionseffekte an.

Schauen wir noch nach möglicher Signifikanz der Parameter

```
drop1(glm.tc, test="Chi")
```

```
## Single term deletions
##
## Model:
## TC_larvae_Numbers_of_Individual ~ Group * Time
##           Df Deviance   AIC   LRT Pr(>Chi)
## <none>           28.466 66.245
## Group:Time  1   34.021 69.800 5.555  0.01843 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(glm(TC_larvae_Numbers_of_Individual ~ Group + Time,
          data = BAdat, family = "poisson"),
      test="Chi")
```

```
## Single term deletions
##
## Model:
## TC_larvae_Numbers_of_Individual ~ Group + Time
##           Df Deviance   AIC   LRT Pr(>Chi)
## <none>           34.021 69.800
## Group    1   35.116 68.896 1.09568  0.2952
## Time     1   34.064 67.843 0.04349  0.8348
```

In diesem Modell ist die Gesamtinteraktion signifikant, während die Einzelparameter (eines Modelles ohne Interaktion) es nicht wären. Es gibt also mit großer Sicherheit einen Interaktionseffekt und damit einen Effekt der Maßnahme.

Interaktionseffekte

Die einzelnen Interaktionseffekte können wir uns auch mithilfe des emmeans Paketes anschauen.

```
l.em <- emmeans(glm.tc, ~ Group * Time)
l.em
```

```
## Group    Time  emmean    SE df asymp.LCL asymp.UCL
## Control1 Before  0.336 0.378 Inf   -0.4043    1.077
## Treatment Before -0.223 0.500 Inf   -1.2031    0.757
## Control1 After -0.916 0.707 Inf   -2.3022    0.470
## Treatment After  0.693 0.316 Inf    0.0734    1.313
```



```
##
## Results are given on the log (not the response) scale.
## Confidence level used: 0.95
```

Während 2023 im Treatment Gewässer weniger Kammolchlarven als im Kontrollgewässer waren (mit nur sehr geringem Unterschied), befanden sich nach der Maßnahme im Treatmentgewässer deutlich mehr Kammolchlarven, im Kontrollgewässer deutlich weniger als im Vorjahr.

Graphische Darstellung

```
ggplot(BAdata, aes(x = Group, y = TC_larvae_Numbers_of_Individual,
                    fill = Time, group = interaction(Group, Time))) +
  geom_boxplot() +
  labs(
    x = "",
    y = "",
    title = "Veränderung der Kammolchlarven vor und nach dem Eingriff",
    fill = "Zeit"
  ) +
  theme_classic() +
  scale_fill_manual(
    values = c("Before" = "yellowgreen", "After" = "seagreen") # Anpassung der Farben
  ) +
  theme(
    text = element_text(family = "Times"),

    # Achsentitel
    axis.title.x = element_text(size = 18,color = "black"),
    axis.title.y = element_text(size = 18,color = "black"),

    # Achsenbeschriftungen
    axis.text.x = element_text(size = 18, color = "black"),
    axis.text.y = element_text(size = 18,color = "black"),

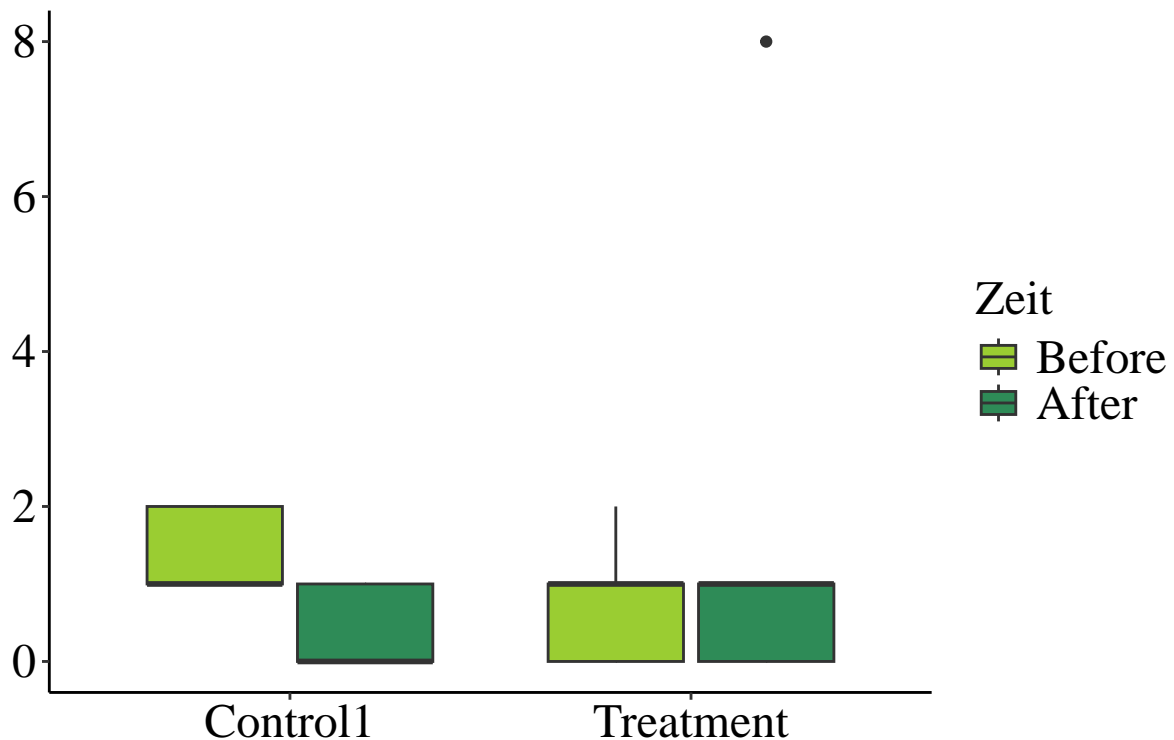
    # Titel
    plot.title = element_text(size = 14, face = "bold", hjust = 0.5,color = "black"),

    #legend.key.size = unit(1.5, "cm"),

    # Ränder
    #plot.margin = margin(t = 50, r = 20, b = 20, l = 20),

    # Legende
    legend.title = element_text(size = 18,color = "black"),
    legend.text = element_text(size = 18,color = "black"),
    legend.position = "right"
  )
```

Veränderung der Kammolchlarven vor und nach dem Eingriff



Schauen wir uns nur die Boxplots an (also die Mediane, nicht die Mittelwerte der Interaktionen), sehen wir, dass die Kammolchlarven von 2023 nach 2024 eigentlich zurückgegangen sind, wie im Kontrollgewässer sichtbar. Im Gewässer mit der Pflegemaßnahme sind sie jedoch nicht zurückgegangen, was darauf schließen lässt, dass die Maßnahme für die Kammolche positiv war.

Insgesamt hatte die Freischneidung also vor allem auf die Kammolche einen positiven Effekt, wobei auch die Teichmolche etwas davon profitiert haben.

Bei Analysen mit mehreren Gebieten vor und nach dem Eingriff mit jeweils mehreren Beobachtungen sollte man ggf. statt eines (generalisierten) linearen Modells ein gemischtes (generalisiertes) lineares Modell nutzen, bei dem das Gebiet als Random Intercept eingegeben wird.

Literaturverzeichnis

Bartoń, K. 2024. MuMIn: Multi-model inference. R package version 1.48.9/r534, <https://R-Forge.R-project.org/projects/mumin/>

Lenth R. 2025. emmeans: Estimated marginal means, aka least-squares means. R package version 1.11.0, <https://CRAN.R-project.org/package=emmeans>

Wickham, H. 2016. ggplot2: Elegant graphics for data analysis. Springer-Verlag New York.