

Cleaner shrimp remove parasite eggs on fish cages

David B. Vaughan*, Alexandra S. Grutter, Kate S. Hutson

*Corresponding author: david.vaughan@jcu.edu.au

Aquaculture Environment Interactions 10: 429–436 (2018)

Statistical approach

```
>### Run mixed effects random intercept model with tanks as the random effect and count data log-transformed:
```

```
>
> library(nlme)
>
Model1=lme(log(N_girellaeCH4)~TreatmentCH4*DayCH4,random=~1|TankCH4,data=CH4.data)
>
> summary(Model1)
Linear mixed-effects model fit by REML
Data: CH4.data
      AIC      BIC    logLik
951.3817 984.6714 -467.6909
```

Random effects:

```
Formula: ~1 | TankCH4
      (Intercept) Residual
StdDev:  0.1917087 0.6237589
```

Fixed effects: log(N_girellaeCH4) ~ TreatmentCH4 * DayCH4

	Value	Std.Error	DF	t-value	p-value
(Intercept)	6.875135	0.11853900	468	57.99893	0.0000
TreatmentCH4T	-2.114827	0.16763946	6	-12.61533	0.0000
DayCH4Day2	-0.617147	0.09862495	468	-6.25751	0.0000
DayCH4Day3	-0.601230	0.09862495	468	-6.09613	0.0000
TreatmentCH4T:DayCH4Day2	0.750258	0.13947674	468	5.37909	0.0000
TreatmentCH4T:DayCH4Day3	0.325959	0.13947674	468	2.33701	0.0199

Correlation:

	(Intr)	TrCH4T	DCH4D2	DCH4D3	TCH4T:DCH4D2	
TreatmentCH4T		-0.707				
DayCH4Day2		-0.416	0.294			
DayCH4Day3		-0.416	0.294	0.500		
TreatmentCH4T:DayCH4Day2		0.294	-0.416	-0.707	-0.354	
TreatmentCH4T:DayCH4Day3		0.294	-0.416	-0.354	-0.707	0.500

Standardized Within-Group Residuals:

	Min	Q1	Med	Q3	Max
	-3.37876326	-0.56688351	-0.03197522	0.52067114	4.91476087

```
Number of Observations: 480
```

```
Number of Groups: 8
```

```
>
>### Diagnostic plot for residuals:
```

```
>
> qqnorm(residuals(Model1))
>
```

```

>### Clear difference in the variance in treatment and control groups = biological
effect of density dependent predation
>
>### Run second lme model but with correlation of variance structure to account
for different variation by treatment groups:
>
>
Model2=lme(log(N_girellaeCH4)~TreatmentCH4*DayCH4,random=~1|TankCH4,weights=varIde
nt(form=~1|TreatmentCH4),data=CH4.data)
>
> summary(Model2)
Linear mixed-effects model fit by REML
Data: CH4.data
      AIC      BIC    logLik
 878.7914 916.2423 -430.3957

Random effects:
Formula: ~1 | TankCH4
      (Intercept) Residual
StdDev:   0.1838733 0.7699959

Variance function:
Structure: Different standard deviations per stratum
Formula: ~1 | TreatmentCH4
Parameter estimates:
      C      T
1.0000000 0.5600797
Fixed effects: log(N_girellaeCH4) ~ TreatmentCH4 * DayCH4
              Value Std.Error DF   t-value p-value
(Intercept)    6.875135 0.1259505 468   54.58601 0.0000
TreatmentCH4T  -2.114827 0.1632197   6  -12.95693 0.0000
DayCH4Day2     -0.617147 0.1217470 468   -5.06909 0.0000
DayCH4Day3     -0.601230 0.1217470 468   -4.93836 0.0000
TreatmentCH4T:DayCH4Day2  0.750258 0.1395419 468    5.37658 0.0000
TreatmentCH4T:DayCH4Day3  0.325959 0.1395419 468    2.33592 0.0199
Correlation:
              (Intr) TrCH4T DCH4D2 DCH4D3 TCH4T:DCH4D2
TreatmentCH4T  -0.772
DayCH4Day2     -0.483  0.373
DayCH4Day3     -0.483  0.373  0.500
TreatmentCH4T:DayCH4Day2  0.422 -0.427 -0.872 -0.436
TreatmentCH4T:DayCH4Day3  0.422 -0.427 -0.436 -0.872  0.500

Standardized Within-Group Residuals:
      Min      Q1      Med      Q3      Max
-4.85671049 -0.55309498 -0.04219193  0.65219878  3.96456705

Number of Observations: 480
Number of Groups: 8
>
#### Allow for variance to differ between each treatment-day combination, rather
than just treatments (there are differences in the spread of variance between
days, not just between treatments; see below:)

# See differences in variance between days:

> with(CH4.data, tapply(log(N_girellaeCH4),list(TreatmentCH4,DayCH4),var ))
      Day1      Day2      Day3
C 0.5424765 0.7718764 0.5857066

```

```

T 0.3437419 0.1113239 0.1468044
>
# See differences in means between days:
>
> with(CH4.data, tapply(log(N_girellaeCH4),list(TreatmentCH4,DayCH4),mean ))
      Day1      Day2      Day3
C 6.875135 6.257988 6.273905
T 4.760308 4.893419 4.485036
>
> library(lattice)
> bwplot(log(N_girellaeCH4)~DayCH4|TreatmentCH4,data=CH4.data)

# create vargroup (new variable) whose levels are each treatment-day combination,
and lets each one have its own variance:

> CH4.data$vargp=with(CH4.data,factor(paste(DayCH4,TreatmentCH4)))

> str(CH4.data)
'data.frame':  480 obs. of  6 variables:
 $ LineCH4      : num  1 2 3 4 5 6 7 8 9 10 ...
 $ DayCH4       : Factor w/  3 levels "Day1","Day2",...: 1 1 1 1 1 1 1 1 1 1 ...
 $ TreatmentCH4 : Factor w/  2 levels "C","T": 1 1 1 1 1 1 1 1 1 1 ...
 $ N_girellaeCH4: num  559 639 2024 428 1669 ...
 $ TankCH4      : Factor w/  8 levels "C1","C2","C3",...: 1 1 1 1 1 1 1 1 1 1 ...
 $ vargp        : Factor w/  6 levels "Day1 C","Day1 T",...: 1 1 1 1 1 1 1 1 1 1 ...
>
# Run third model to include the vargroup:
>
Model3=lme(log(N_girellaeCH4)~TreatmentCH4*DayCH4,random=~1|TankCH4,weights=varIdent(form=~1|vargp),data=CH4.data)
>
> anova(Model1,Model2,Model3)
      Model df      AIC      BIC    logLik   Test  L.Ratio p-value
Model1     1  8 951.3817 984.6714 -467.6909
Model2     2  9 878.7914 916.2423 -430.3957 1 vs 2 74.59028 <.0001
Model3     3 13 858.3051 912.4008 -416.1526 2 vs 3 28.48630 <.0001
>
# The AIC (lowest) suggests this model is even better (the best).
>
> summary(Model3)
Linear mixed-effects model fit by REML
Data: CH4.data
      AIC      BIC    logLik
858.3051 912.4008 -416.1526

Random effects:
Formula: ~1 | TankCH4
      (Intercept) Residual
StdDev:  0.1732571 0.699756

Variance function:
Structure: Different standard deviations per stratum
Formula: ~1 | vargp
Parameter estimates:
      Day1 C      Day2 C      Day3 C      Day1 T      Day2 T      Day3 T
1.0000000 1.2214233 1.0679785 0.7989372 0.4482660 0.5573069
Fixed effects: log(N_girellaeCH4) ~ TreatmentCH4 * DayCH4
      Value Std.Error DF   t-value p-value
(Intercept)                6.875135 0.1167272 468  58.89917 0.0000

```

```
TreatmentCH4T      -2.114827  0.1582296   6 -13.36556  0.0000
DayCH4Day2         -0.617147  0.1234994 468  -4.99717  0.0000
DayCH4Day3         -0.601230  0.1144635 468  -5.25259  0.0000
TreatmentCH4T:DayCH4Day2  0.750258  0.1427896 468   5.25429  0.0000
TreatmentCH4T:DayCH4Day3  0.325959  0.1375129 468   2.37039  0.0182
```

Correlation:

```
(Intr) TrCH4T DCH4D2 DCH4D3 TCH4T:DCH4D2
TreatmentCH4T      -0.738
DayCH4Day2         -0.425  0.313
DayCH4Day3         -0.458  0.338  0.433
TreatmentCH4T:DayCH4Day2  0.367 -0.444 -0.865 -0.374
TreatmentCH4T:DayCH4Day3  0.381 -0.461 -0.360 -0.832  0.511
```

Standardized Within-Group Residuals:

```
      Min      Q1      Med      Q3      Max
-3.77002063 -0.58999501 -0.01999284  0.67764778  4.07818424
```

Number of Observations: 480

Number of Groups: 8

>

> anova(Model3)

```
              numDF denDF  F-value p-value
(Intercept)         1   468 6858.621 <.0001
TreatmentCH4         1     6  173.357 <.0001
DayCH4               2   468   31.201 <.0001
TreatmentCH4:DayCH4  2   468   13.870 <.0001
```

>

Confidence intervals for Model3 (these are used in table 1 of the manuscript):

>

> intervals(Model3)

Approximate 95% confidence intervals

Fixed effects:

```
              lower      est.      upper
(Intercept)  6.64576045  6.8751347  7.1045090
TreatmentCH4T -2.50200097 -2.1148271 -1.7276531
DayCH4Day2   -0.85982872 -0.6171468 -0.3744649
DayCH4Day3   -0.82615633 -0.6012302 -0.3763041
TreatmentCH4T:DayCH4Day2  0.46966962  0.7502578  1.0308459
TreatmentCH4T:DayCH4Day3  0.05573963  0.3259589  0.5961781
```

attr(,"label")

[1] "Fixed effects:"

Random Effects:

Level: TankCH4

```
              lower      est.      upper
sd((Intercept)) 0.08489948 0.1732571 0.3535714
```

Variance function:

```
              lower      est.      upper
Day2 C 0.9809402 1.2214233 1.5208621
Day3 C 0.8582701 1.0679785 1.3289268
Day1 T 0.6400008 0.7989372 0.9973436
Day2 T 0.3585310 0.4482660 0.5604604
Day3 T 0.4461444 0.5573069 0.6961670
```

attr(,"label")

[1] "Variance function:"

Within-group standard error:

```
      lower      est.      upper
0.5986807 0.6997560 0.8178958
>
> hist(residuals(Model3))
> plot(Model3)
```