Annotated R-script

**Roll, right, repeat: short-term repeatability in the self-righting behaviour of a cold-water sea cucumber**

Jeff C. Clements 1\*, Ellen Schagerström 2,3, Sam Dupont 2, Fredrik Jutfelt 1, Kirti Ramesh 2

1  Department of Biology, Norwegian University of Science and Technology, Høgskoleringen 5, 7491 Trondheim, Norway

2 Department of Biological and Environmental Sciences, University of Gothenburg, Sven Lovén Centre for Marine Infrastructure ‐ Kristineberg, Kristineberg 566, Sweden 45178 Fiskebäckskil, Sweden

3 Swedish Mariculture Research Centre (SWEMARC), University of Gothenburg, P.O Box 463, 405 30 Gothenburg, Sweden

R version 3.5.3 (2019-03-11) -- "Great Truth"

Copyright (C) 2019 The R Foundation for Statistical Computing

Platform: x86\_64-apple-darwin15.6.0 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.

You are welcome to redistribute it under certain conditions.

Type 'license()' or 'licence()' for distribution details.

 Natural language support but running in an English locale

R is a collaborative project with many contributors.

Type 'contributors()' for more information and

'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or

'help.start()' for an HTML browser interface to help.

Type 'q()' to quit R.

[Workspace loaded from ~/.RData]

Loading required package: lmerTest

Loading required package: lme4

Loading required package: Matrix

Attaching package: ‘lmerTest’

The following object is masked from ‘package:lme4’:

 lmer

The following object is masked from ‘package:stats’:

 step

#####test for fixed effects of trial and body length on righting time#####

#upload data

> rightlme<-read.table("/Users/JeffClements/Desktop/cukes.txt",header=T)

> attach(rightlme)

> rightlme$trial<-as.factor(rightlme$trial)

> rightlme$location<-as.factor(rightlme$location)

> rightlme$id<-as.factor(rightlme$id)

> summary(rightlme)

 id location length trial rt

 1 : 5 H:10 Min. :14.50 1:17 Min. : 25.01

 2 : 5 K:75 1st Qu.:17.70 2:17 1st Qu.: 63.77

 3 : 5 Median :19.20 3:17 Median : 86.84

 4 : 5 Mean :20.06 4:17 Mean : 96.72

 5 : 5 3rd Qu.:19.80 5:17 3rd Qu.:120.90

 6 : 5 Max. :29.80 Max. :300.00

 (Other):55

 ln.rt

 Min. :3.219

 1st Qu.:4.155

 Median :4.464

 Mean :4.450

 3rd Qu.:4.795

 Max. :5.704

#build linear mixed effects model

> lmemod1<-lmer(rt~trial\*location+length+(1|id),data=rightlme)

#test assumptions

> plot(lmemod1)

> qqnorm(resid(lmemod1))

> hist(resid(lmemod1))

#assumptions violated

#build new model with ln-transformed righting times

> lmemod2<-lmer(ln.rt~trial\*location+length+(1|id),data=rightlme)

#test assumptions for new model

> plot(lmemod2)

> qqnorm(resid(lmemod2))

> hist(resid(lmemod2))

#assumptions met

#summary of model

> summary(lmemod2)

Linear mixed model fit by REML. t-tests use Satterthwaite's method [

lmerModLmerTest]

Formula: ln.rt ~ trial \* location + length + (1 | id)

 Data: rightlme

REML criterion at convergence: 74.6

Scaled residuals:

 Min 1Q Median 3Q Max

-2.71225 -0.56657 0.03359 0.49814 2.72840

Random effects:

 Groups Name Variance Std.Dev.

 id (Intercept) 0.16683 0.4085

 Residual 0.07153 0.2674

Number of obs: 85, groups: id, 17

Fixed effects:

 Estimate Std. Error df t value Pr(>|t|)

(Intercept) 3.24891 0.69626 15.79677 4.666 0.000267 \*\*\*

trial2 0.15530 0.26745 60.00000 0.581 0.563634

trial3 0.05255 0.26745 60.00000 0.196 0.844893

trial4 0.20420 0.26745 60.00000 0.764 0.448148

trial5 0.20085 0.26745 60.00000 0.751 0.455594

locationK 0.11975 0.37905 22.89635 0.316 0.754924

length 0.05303 0.02606 14.00000 2.035 0.061291 .

trial2:locationK -0.09868 0.28472 60.00000 -0.347 0.730114

trial3:locationK -0.00669 0.28472 60.00000 -0.023 0.981332

trial4:locationK -0.22565 0.28472 60.00000 -0.793 0.431164

trial5:locationK -0.18506 0.28472 60.00000 -0.650 0.518182

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:

 (Intr) trial2 trial3 trial4 trial5 loctnK length trl2:K trl3:K

trial2 -0.192

trial3 -0.192 0.500

trial4 -0.192 0.500 0.500

trial5 -0.192 0.500 0.500 0.500

locationK -0.664 0.353 0.353 0.353 0.353

length -0.868 0.000 0.000 0.000 0.000 0.245

tril2:lctnK 0.180 -0.939 -0.470 -0.470 -0.470 -0.376 0.000

tril3:lctnK 0.180 -0.470 -0.939 -0.470 -0.470 -0.376 0.000 0.500

tril4:lctnK 0.180 -0.470 -0.470 -0.939 -0.470 -0.376 0.000 0.500 0.500

tril5:lctnK 0.180 -0.470 -0.470 -0.470 -0.939 -0.376 0.000 0.500 0.500

 trl4:K

trial2

trial3

trial4

trial5

locationK

length

tril2:lctnK

tril3:lctnK

tril4:lctnK

tril5:lctnK 0.500

#get fixed effect results

> anova(lmemod2)

Type III Analysis of Variance Table with Satterthwaite's method

 Sum Sq Mean Sq NumDF DenDF F value Pr(>F)

trial 0.060324 0.015081 4 60 0.2108 0.93144

location 0.000176 0.000176 1 14 0.0025 0.96117

length 0.296086 0.296086 1 14 4.1394 0.06129 .

trial:location 0.073555 0.018389 4 60 0.2571 0.90422

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

#####test relationship between size and mean righting time#####

#upload data

> rightreg<-read.table("/Users/JeffClements/Desktop/cukeregression.txt",header=T)

> attach(rightreg)

The following object is masked from rightlme:

 length

> summary(rightreg)

 length mean.rt

 Min. :14.50 Min. : 34.76

 1st Qu.:17.70 1st Qu.: 70.40

 Median :19.20 Median : 91.97

 Mean :20.06 Mean : 96.72

 3rd Qu.:19.80 3rd Qu.:118.99

 Max. :29.80 Max. :217.00

> as.data.frame(rightreg)

 length mean.rt

1 18.7 133.77

2 17.4 51.38

3 29.8 217.00

4 14.5 70.40

5 16.4 106.86

6 19.8 155.45

7 15.8 81.96

8 19.0 70.50

9 19.3 48.69

10 19.2 34.76

11 19.2 61.30

12 20.1 91.97

13 17.7 120.72

14 28.3 118.99

15 19.4 78.29

16 26.8 100.41

17 19.6 101.83

#build linear model and test for effect of size on mean righting time

> mod1<-lm(mean.rt~length,data=rightreg)

> summary(mod1)

Call:

lm(formula = mean.rt ~ length, data = rightreg)

Residuals:

 Min 1Q Median 3Q Max

-56.887 -29.631 -4.996 31.758 62.716

Coefficients:

 Estimate Std. Error t value Pr(>|t|)

(Intercept) -21.808 45.948 -0.475 0.6419

length 5.909 2.244 2.634 0.0188 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 38.15 on 15 degrees of freedom

Multiple R-squared: 0.3162, Adjusted R-squared: 0.2706

F-statistic: 6.936 on 1 and 15 DF, p-value: 0.0188

#####determine repeatability of righting time#####

#install MCMCglmm package

> library(MCMCglmm)

Loading required package: coda

Loading required package: ape

#upload & format data

> right<-read.table("/Users/JeffClements/Desktop/cukes.txt",header=T)

> right$id<-factor(right$id)

> right$trial<-factor(right$trial)

> right$location<-factor(right$location)

#obtain summary statistics

> summary(right)

 id location length trial rt

 1 : 5 H:10 Min. :14.50 1:17 Min. : 25.01

 2 : 5 K:75 1st Qu.:17.70 2:17 1st Qu.: 63.77

 3 : 5 Median :19.20 3:17 Median : 86.84

 4 : 5 Mean :20.06 4:17 Mean : 96.72

 5 : 5 3rd Qu.:19.80 5:17 3rd Qu.:120.90

 6 : 5 Max. :29.80 Max. :300.00

 (Other):55

 ln.rt

 Min. :3.219

 1st Qu.:4.155

 Median :4.464

 Mean :4.450

 3rd Qu.:4.795

 Max. :5.704

#run MCMCglmm model without fixed effects (unadjusted model)

> RT<-MCMCglmm(rt~1,random=~id,data=right)

 MCMC iteration = 0

 MCMC iteration = 1000

 MCMC iteration = 2000

 MCMC iteration = 3000

 MCMC iteration = 4000

 MCMC iteration = 5000

 MCMC iteration = 6000

 MCMC iteration = 7000

 MCMC iteration = 8000

 MCMC iteration = 9000

 MCMC iteration = 10000

 MCMC iteration = 11000

 MCMC iteration = 12000

 MCMC iteration = 13000

#derive output for unadjusted model

> summary(RT)

 Iterations = 3001:12991

 Thinning interval = 10

 Sample size = 1000

 DIC: 818.3051

 G-structure: ~id

 post.mean l-95% CI u-95% CI eff.samp

id 2121 846.5 3884 1265

 R-structure: ~units

 post.mean l-95% CI u-95% CI eff.samp

units 743.1 505.9 1005 1000

 Location effects: rt ~ 1

 post.mean l-95% CI u-95% CI eff.samp pMCMC

(Intercept) 97.02 74.25 118.88 1204 <0.001 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

#plot chain trace + posterior density estimations for intercept, Vind, and Ved for unadjusted model

> plot(RT)

Hit <Return> to see next plot:

#calculate agreement repeatability using posterior distributions for unadjusted model

> R.RT<-RT$VCV[,"id"]/(RT$VCV[,"id"]+RT$VCV[,"units"])

> plot(R.RT)

#derive repeatability estimate and 95% CI interval for unadjusted model

#repeatability estimate

> posterior.mode(R.RT)

 var1

0. 7366644

#95% CI

> HPDinterval(R.RT)

 lower upper

var1 0.5482938 0.8762332

attr(,"Probability")

[1] 0.95

#run MCMCglmm model with fixed effects (trial) & fit slopes for body length in each trial (adjusted model)

> adj.RT <-MCMCglmm(rt ~trial\*location+at.level(trial,"1"):length+at.level(trial,"2"):length+at.level(trial,"3"):length+at.level(trial,"4"):length+at.level(trial,"5"):length,random=~id,data=right)

 MCMC iteration = 0

 MCMC iteration = 1000

 MCMC iteration = 2000

 MCMC iteration = 3000

 MCMC iteration = 4000

 MCMC iteration = 5000

 MCMC iteration = 6000

 MCMC iteration = 7000

 MCMC iteration = 8000

 MCMC iteration = 9000

 MCMC iteration = 10000

 MCMC iteration = 11000

 MCMC iteration = 12000

 MCMC iteration = 13000

#derive output for adjusted model

> summary(adj.RT)

Iterations = 3001:12991

 Thinning interval = 10

 Sample size = 1000

 DIC: 815.9711

 G-structure: ~id

 post.mean l-95% CI u-95% CI eff.samp

id 1637 532.6 3185 1000

 R-structure: ~units

 post.mean l-95% CI u-95% CI eff.samp

units 632.5 420.7 903.8 1096

 Location effects: rt ~ trial \* location + at.level(trial, "1"):length + at.level(trial, "2"):length + at.level(trial, "3"):length + at.level(trial, "4"):length + at.level(trial, "5"):length

 post.mean l-95% CI u-95% CI eff.samp pMCMC

(Intercept) -177.6386 -319.7638 -31.6873 1403 0.018 \*

trial2 168.7704 60.5012 277.5268 1000 0.002 \*\*

trial3 90.1022 -18.5984 209.2632 1000 0.110

trial4 181.4686 68.9004 280.0969 1000 0.002 \*\*

trial5 231.2653 124.4876 340.8695 1000 <0.001 \*\*\*

locationK 51.3149 -15.8514 122.1197 1000 0.152

trial2:locationK -40.8564 -94.7612 9.9550 1000 0.134

trial3:locationK -19.6769 -77.1495 34.4124 1000 0.476

trial4:locationK -54.1360 -109.5885 -5.3520 1000 0.046 \*

trial5:locationK -56.6494 -105.6721 0.3963 1000 0.046 \*

at.level(trial, "1"):length 11.5579 6.0420 17.2005 1371 0.002 \*\*

length:at.level(trial, "2") 4.9056 -0.4895 10.8054 1000 0.086 .

length:at.level(trial, "3") 7.8551 2.4258 13.6648 1000 0.012 \*

length:at.level(trial, "4") 4.5745 -0.9739 10.1705 1000 0.104

length:at.level(trial, "5") 2.3979 -3.2882 8.0105 1000 0.416

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

#plot chain trace + posterior density estimations for intercept, Vind, and Ved for adjusted model

> plot(adj.RT)

Hit <Return> to see next plot:

Hit <Return> to see next plot:

Hit <Return> to see next plot:

Hit <Return> to see next plot:

#calculate adjusted repeatability using posterior distribution

> R.adj.RT<-adj.RT$VCV[,"id"]/(adj.RT$VCV[,"id"]+adj.RT$VCV[,"units"])

> plot(R.adj.RT)

#derive repeatability estimate and 95% CI interval for adjusted model

#repeatability estimate

> posterior.mode(R.adj.RT)

 var1

0.7497716

#95% CI

> HPDinterval(R.adj.RT)

 lower upper

var1 0.5196375 0.874569

attr(,"Probability")

[1] 0.95

#create results table

> R.estimates<-rbind(cbind(posterior.mode(R.RT),HPDinterval(R.RT),posterior.mode(R.adj.RT),HPDinterval(R.adj.RT)))

> row.names(R.estimates)<-c("Righting time")

> colnames(R.estimates)<-c("R.agreement","low95","high95","R.adjusted","low95","high95")

> R.estimates

 R.agreement low95 high95 R.adjusted low95

Righting time 0.7366644 0.5482938 0.8762332 0.7497716 0.5196375

 high95

Righting time 0.874569

#derive variance components of RT for adjusted model

#Vind for RT

> posterior.mode(adj.RT$VCV[,"id"])

 var1

1056.014

#Vind for RT

> posterior.mode(adj.RT$VCV[,"units"])

 var1

607.8258