

Ecography

E7078

Bunnefeld, N. and Phillimore, A. B. 2011. Island, archipelago and taxon effects: mixed models as a means of dealing with the imperfect design of nature's experiments. – *Ecography* 34: xxx–xxx.

Supplementary material

Appendix 1 – R code

#Load packages lme4 and MuMIn

require('lme4')

require('MuMIn')

*#Test for random effects by varying the random effects added to the model and compare
#models with AICc. Use REML=TRUE and all fixed effects added*

#Fixed effects

#Geol.Age: geological age of the island (continuous)

#Geol.Age2: geological age of the island as a quadratic (continuous)

#logTOTALArea: the area of the island (log-transformed)

#DistMainland: distance to the mainland (continuous)

#Archipel: the archipelago as categorical variable with four levels (Azores, Canaries,

#Galapagos, Hawaii)

*# * denotes the interaction between two fixed effects*

#Random effects

#TAXA: the four different taxa (arthropods, beetles, plants, snails)

#ISLAND: the 39 different islands

#TA: a new categorical variable of TAXA on each archipelago with 14 levels

#(1|ISLAND) defines the random intercept model for islands and (Geol.Age|OA) the random

#slope of geological age within OA

r1<-lmer(log(SIE+1)~ Geol.Age

+Geol.Age2

+DistMainland

+NearestIsland

+logTOTALArea

+(1|ARCHIPEL)

+(1|TAXA)

+(1|ISLAND)

+(1|TA)

,data=data,REML=TRUE)

#with geological age as random slope within the taxa:archipelago random intercept

r2<-lmer(log(SIE+1)~ Geol.Age

+Geol.Age2

+DistMainland

+NearestIsland

+logTOTALArea

+(1|ARCHIPEL)

+(1|TAXA)

+(1|ISLAND)

+(Geol.Age|TA)

,data=data,REML=TRUE)

```
#repeat model fitting with different random effects and give them different names, for
#example r1,r2 etc. Then extract the AICc.
```

```
aicc_r<- AICc(r1,r2,r3,r4,r5,r6,r7,r8)
```

```
aicc_r[order(aicc_r$AICc),]
```

```
#Use the dredge() function in the MuMIn package to compare fixed effects model and
#compare with AICc and AICcweights. Use REML=FALSE.
```

```
f1<-lmer(log(SIE+1)~ Geol.Age
+Geol.Age2
+logTOTALArea
+ARCHIPEL
+ logTOTALArea*ARCHIPEL
+ Geol.Age*ARCHIPEL
+ Geol.Age2*ARCHIPEL
+NearestIsland
+(1|TAXA)
+(1|ISLAND)
+(1|TA)
,data=data,REML=FALSE,
)
d1<-dredge(f1)
```

```
#plot the final model "first1"
```

```
#set the margins
margins<-c(1,1,1,1)
```

```
#set the number of grid cells for the 3d plot
grid<-7
pdf('BAplot.pdf',width=16,height=10.5)
```

```
#set the layout of the plot, so that it is a 4x4 grid with 14 cells (2 empty cells)
```

```
nf <- layout(matrix(c(1,2,0,3,
4,5,6,7,
8,9,10,0,
11,12,13,14), 4, 4, byrow=TRUE),
widths = rep(1,4),
heights =rep(1,4),
respect=FALSE)
```

```

#set outer margins for plotting the axes labels
par(oma=c(5,5,2,2))

#plot Azores Arthropods
s3d <-scatterplot3d(Geol.Age[TA== rownames(coef(first1)$TA)[1]],
                    logTOTALArea[TA== rownames(coef(first1)$TA)[1]],
                    logSIE[TA== rownames(coef(first1)$TA)[1]],
                    pch=16, box=FALSE,color="red",type="h",
                    ylab="",xlab="",zlab="",mar=margins,cex.axis=0.7,
                    xlim=c(min(Geol.Age),max(Geol.Age)),
                    ylim=c(min(logTOTALArea),max(logTOTALArea)),
                    zlim=c(min(logSIE),max(logSIE)))

#plot the grid
for(age in 1:grid){
  agval<-seq(min(Geol.Age[TA==
rownames(coef(first1)$TA)[1]]),max(Geol.Age[TA==rownames(coef(first1)$TA)[1]]),le
ngth.out=le)[age]
  areval<seq(min(logTOTALArea[TA==rownames(coef(first1)$TA)[1]]),
max(logTOTALArea[TA==rownames(coef(first1)$TA)[1]]),length.out=le)
  respvals<-coef(first1)$TAXA[[1]][1]+
  coef(first1)$TAXA[[2]][1]*agval+
  coef(first1)$TAXA[[3]][1]*(agval^2)+
  coef(first1)$TAXA[[4]][1]*areval
  s3d$points3d(rep(agval,le),areval,respvals,type="l",col="blue")
}

for(are in 1:grid){
  agval<-seq(min(Geol.Age[TA==rownames(coef(first1)$TA)[1]]),
max(Geol.Age[TA==rownames(coef(first1)$TA)[1]]),length.out=le)
  areval<-seq(min(logTOTALArea[TA==rownames(coef(first1)$TA)[1]]),
max(logTOTALArea[TA==rownames(coef(first1)$TA)[1]]),length.out=le)[are]
  respvals<-coef(first1)$TAXA[[1]][1]+
  coef(first1)$TAXA[[2]][1]*agval+
  coef(first1)$TAXA[[3]][1]*(agval^2)+
  coef(first1)$TAXA[[4]][1]*areval
  s3d$points3d(agval,rep(areval,le),respvals,type="l",col=ourcolour)
}

#repeat this step for all 14 models

#create axes labels
mtext("Age of the Island (million years)", 1,outer=TRUE,padj=3)
mtext("Number of single island endemics (log(SIE+1))", 2,outer=TRUE,padj=-3)

#create row and column labels
mtext(text=c("Arthropods","Beetles","Plants","Snails"), 3,
      outer=TRUE,at=c(0.13,0.38,0.63,0.88),padj=-0.2)

```

```

mtext(text=c("Hawaii","Galapagos","Canaries","Azores"),4,
      outer=TRUE, at=c(0.11,0.35,0.62,0.85),padj=0.2)
#save the pdf and close the graphics device
dev.off()

```

Appendix 2 – parameter estimates from the most parsimonious model

The estimates of the fixed and random effects of the most parsimonious model. The treatment effects are presented (default in R), where the intercept gives the intercept for the Azores and all other intercepts are the differences between the Azores and Canaries, Galapagos and Hawaii. The slopes for geological age for the different archipelagos can be calculated in a similar way. The random effect variances are the estimates of the variation explained by the random effects. By dividing each variance by the sum of all four variances (including the residual) the variance components can be calculated as proportions.

Fixed effects	Estimate	SE	t-value
Intercept	-1.66	0.47	-3.53
Geol.Age	0.21	0.06	3.41
Geol.Ag2	0.01	0.01	0.65
logArea	0.45	0.04	12.24
Arch.Cana	2.45	0.35	7.01
Arch.Galap	0.35	0.31	1.13
Arch.Hawa	2.68	0.31	8.57
Geol.Age:Arch.Cana	-0.02	0.01	-2.41
Geol.Age:Arch.Galap	-0.02	0.03	-0.92
Geol.Age:Arch.Hawa	-0.02	0.01	-2.08

Random effects	Variance
Island	0.12
Taxa:Archipelago	0.07
Taxa	0.56
Residual	0.25

Appendix 3 Observed and predicted values from the most parsimonious mixed effects model incorporating fixed and random effects (Table 4). Predicted values are the larger symbols in grey.

