SMA Regressions

Background: Productivity (P: g dry mass / m2 / year) and biomass (B: g dry mass / m2) data have been collected and compared among different ecosystems for many years. In principle, P should be predicted from B. Annual plants of grasslands make harvest-based measures of P and B relatively easy; thus grasslands have been prominent in productivity-biodiversity research and we focus on those here. *Our mission*: find the best P:B relationship among reasonable alternatives.

So far this semester you have used ordinary least squares (OLS) regressions. Today we compare those to standardized major axis (SMA) regressions for P/B data of grasslands. We need to evaluate log-transformations of P and of B data. Four combinations exist: P/B, P/logB, logP/B, and logP/logB. We evaluate alternative models in two ways:

- Regression assumptions – we should doubt models that do not meet assumptions. We use QQ normality plots and residual plots as before.
- AICc of OLS and SMA models for the same response variable – it is unfair to compare simple P (e.g., 10 $g/m^2/y$) to logP (e.g., log(10 = 1).

Our Mission:

A) Can a strong relationship be found between P and B? What is the "best" equation?
B) Does the logP:logB relationship fit the ¾ slope prediction of Metabolic Theory? (for more detail read Brown et al. 2004)
C) And what is the effect, if any, of grazing on this relationship? Grazing may increase slope if plants respond with more growth, but may also reduce slope by removing growth before it is measured.

Preliminary Steps

1) Install and load the `smatr` package in R. Here we use lm to make OLS models and sma in smatr to make SMA.
2) Load the `bbmle` package
3) Load the `ggplot2` package for graphs
4) Import and attach the `PB.txt` data set from the course web page.
5) First let's just look at the data pattern. Plot P as a function of B for each subset:

```
qplot(B,P,colour=Grazed,pch=Grazed)
```

Three questions:

- Does a strong linear relationship seem apparent?
- Is there a wedge-shaped scatter that may be reduced with log-transforms?
- Does it look like grazed and ungrazed P:B relationships differ? [we revisit this later]

6) Make log-10 transforms (log10 vs. loge is arbitrary – log10 is easier to interpret) and re-plot – more linear? This is one reason why allometry defaults to log-transforms.

7) Now compare residuals of the untransformed and transformed versions (the way you have done this in prior labs).

   - Can we expect log-log assumption tests and models to work better below?

Model Comparison: We compare models with AICc: *we cannot fairly compare a model of logP to a model of simple P*. But we can change predictor variables for any one response variable. So we compare OLS and SMA models in two sets:

A) simple P: [P ~ B] vs. [P ~ logB], each with OLS and SMA models (4 models total)
B) logP: [logP ~ B] vs. [logP ~ logB], each with OLS and SMA models (4 models total)

Here is the code to compute one OLS and one SMA regression for one model [P ~ B].

```
pbols <- lm(P~B) #run lm on P as function of B
par(mfrow=c(2,2)) # make a 2x2 grid for plots
plot(pbols)       #run diagnostic plots
summary(pbols)    #lm output

pbsma <- sma(P~B, method="SMA", nobs=length(B))#now run sma
model, where nobs specifies the number of observations
rpbsma <- residuals(pbsma)    #report resids of sma model
plot(rpbsma~B) #this and next line plot resids
abline(a=0,b=0, lty=2)  # because sma doesn't make the 4 plots
summary(pbsma)    #sma output
```

8. Use the above as a template to also run similar models of the following:
   B) P ~ logB
   C) logP ~ B
   D) logP ~ logB

At the end you should have two sets (OLS and SMA) of each of the four models (P~B, P~logB, logP~B, and logP~logB). Total should = 8 models.

Which model best shows residuals that are

- most evenly distributed along the length,
- normal (i.e., balanced and centered around regression line)
- without some other pattern (e.g., an arch)?

10. You could use bbmle (if you can specify details correctly, and for some !@#$ reason it's not working for me), or you can simply type

    AIC(model) # to get AIC scores to compare among models

- Which model “wins”?
- Of the two “winners” which model would you use to estimate P, based on assumptions and R^2?

Metabolic Theory: Now we test to see if the slope of our best log-log model = ¾ as predicted by Metabolic Theory (Brown et al. 2004). Short version: fractal metabolic networks of plants should cause this slope. We specify a slope.test=0.75 in the logP~logB model:

summary(sma(logP~logB, method="SMA", slope.test=0.75))

- Do the data support the Metabolic Theory of Ecology?

Grazing Effects: Now we use our “best” model to address the effect of grazing: Do grazed and ungrazed P:B relationships differ?

1. repeat your best model (assume here to be SMA of log-log) for just the ungrazed subset by specifying *inside the model parentheses*:

```
subset = PB$Grazed=="n"
```

2. Now replace the “place-holder” values for slope (1) and elevation (intercept; 0.1) below with the actual values for the ungrazed model. This will compare the model for grazed grasslands (Grazed==”y”) to the slope and intercept of the ungrazed model.

```
summary(sma(subset = PB$Grazed=="y", logP~logB, method="SMA",
slope.test=1, elev.test=0.1))
```

- Does grazing change P:B scaling in grasslands?