Introduction to structural equation modeling and mixed models in 

Day 9 – Part 3: SEM

Oksana Buzhdygan

oksana.buzh@fu-berlin.de

Day 9 – Part 3 Outline

- Extensions to GLM, LMM, and GLMM

GLM in SEM

Example

wikipedia.org

Serengeti you think of…

Serengeti you may not think of…

GLM in SEM

Example

ECOLOGY
ECOLOGICAL SOCIETY OF AMERICA

Article | Full Access

Landscape-scale analyses suggest both nutrient and antipredator advantages to Serengeti herbivore hotspots

T. Michael Anderson, J. Grant C. Hopcraft, Stephanie Eby, Mark Ritchie, James B. Grace, Han Olff

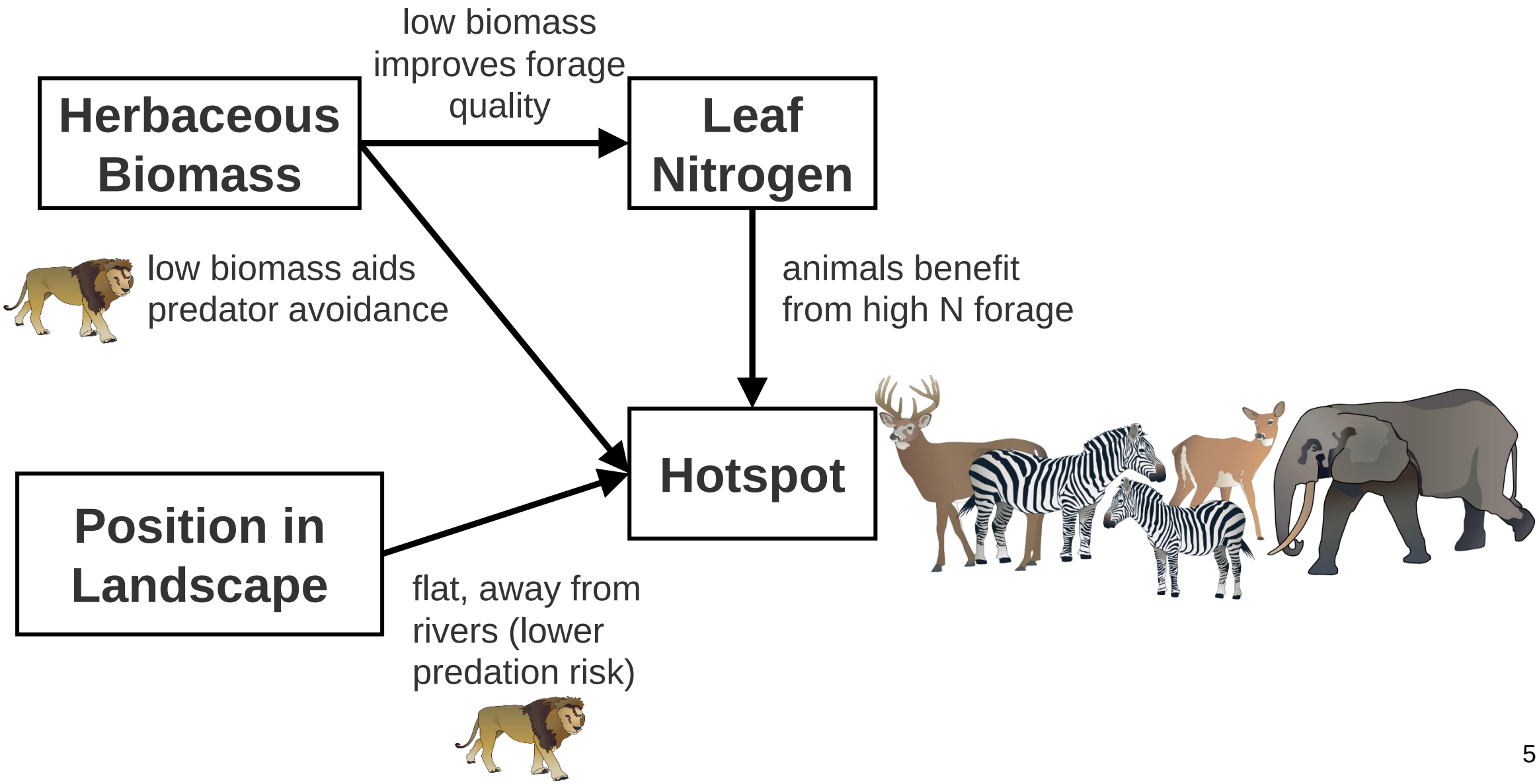
First published: 01 May 2010 | https://doi.org/10.1890/09-0739.1 | Citations: 71

Corresponding Editor: T. P. Young.

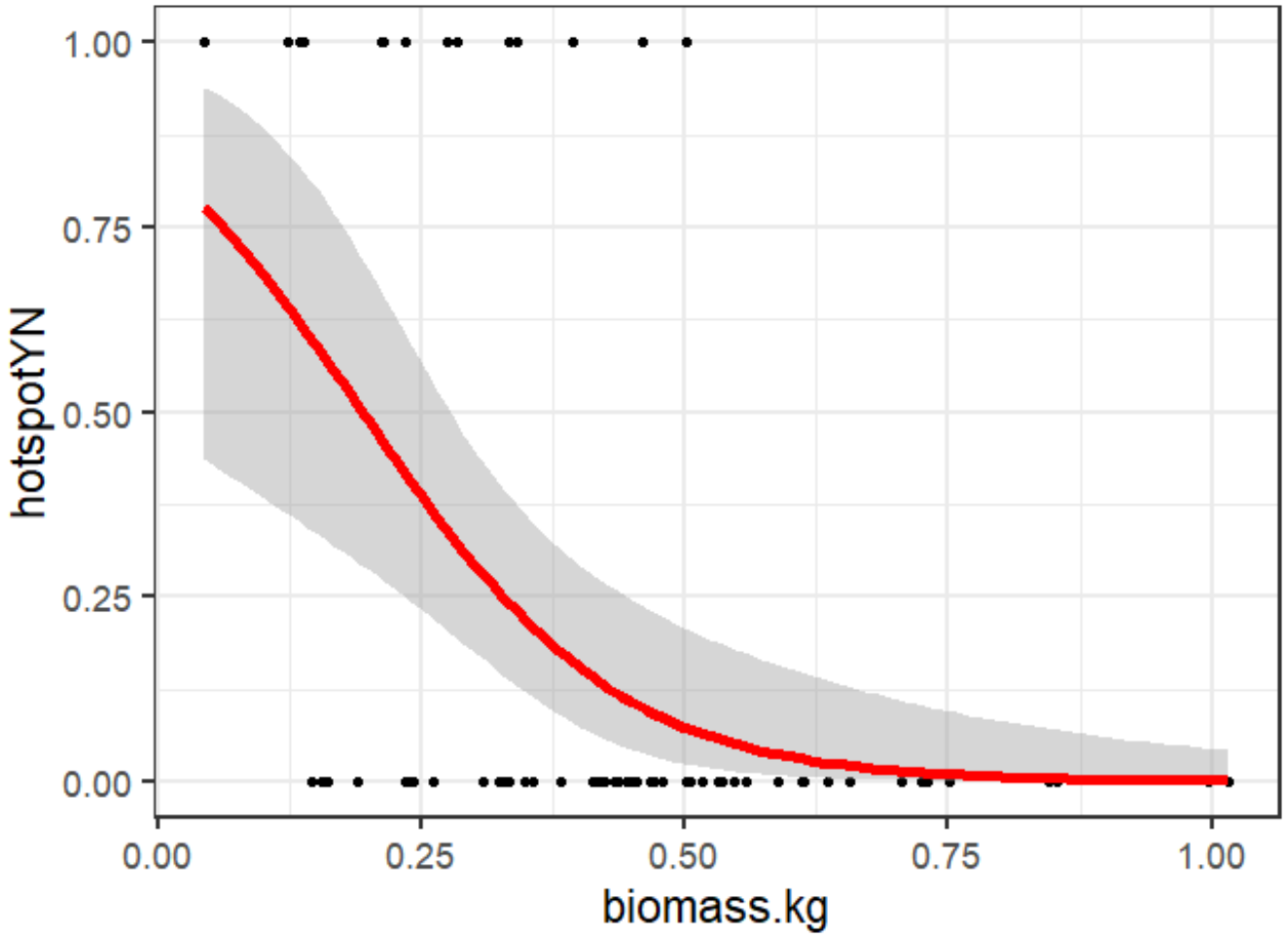
Abstract

Mechanistic explanations of herbivore spatial distribution have focused largely on either resource-related (bottom-up) or predation-related (top-down) factors. We studied direct and indirect influences on the spatial distributions of Serengeti herbivore hotspots, defined as temporally stable areas inhabited by mixed herds of resident grazers. Remote sensing and variation in landscape features were first used to create a map of the spatial distribution of hotspots, which was tested for accuracy against an independent data set of herbivore observations. Subsequently, we applied structural equation modeling to

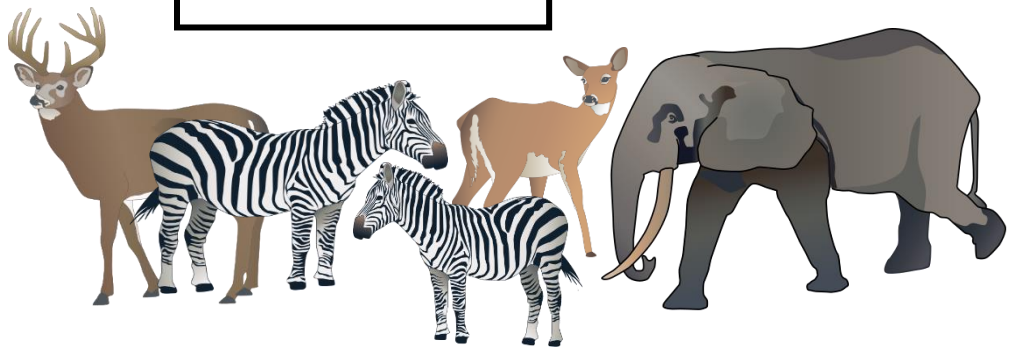
GLM in SEM


low biomass improves forage quality
Herbaceous Biomass
Leaf Nitrogen
low biomass aids predator avoidance
animals benefit from high N forage
Hotspot
Position in Landscape
flat, away from rivers (lower predation risk)

GLM in SEM

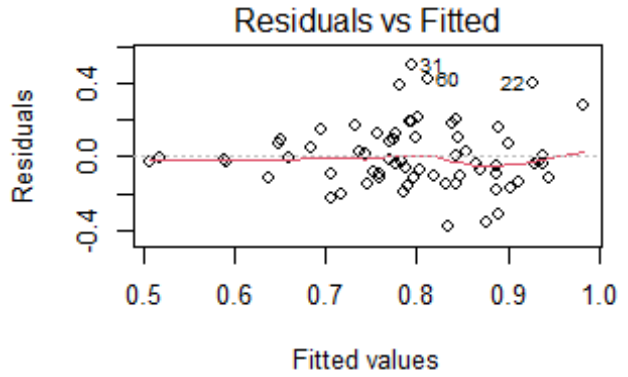


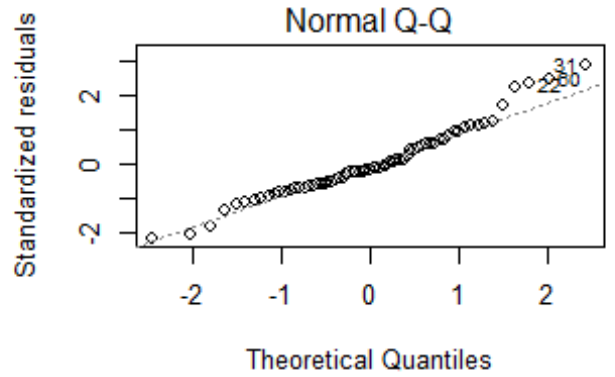
Hotspot

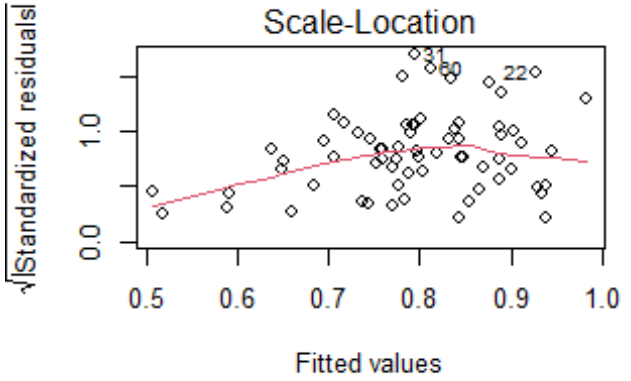


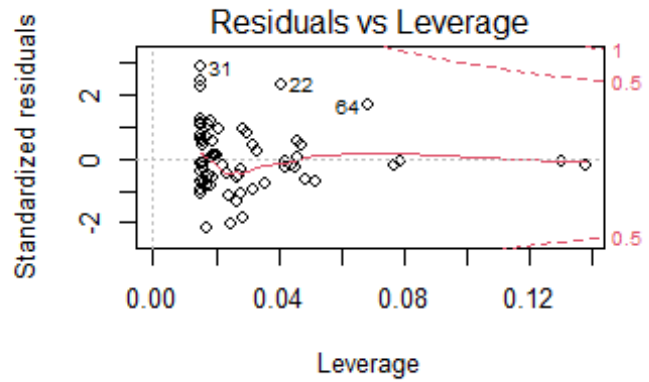
GLM in SEM

```
anderson <- read.csv("Data/Anderson.csv")
str(anderson)
# check assumptions
m1 <- lm(leafN ~ biomass.kg, anderson)
plot(m1)

m2 <- glm(hotspotYN ~ leafN + biomass.kg +
            landscape, family = "binomial",
            data=anderson)
summary(m2) # check overdispersion

anderson.sem <- psem( m1,m2)
summary(anderson.sem)
```



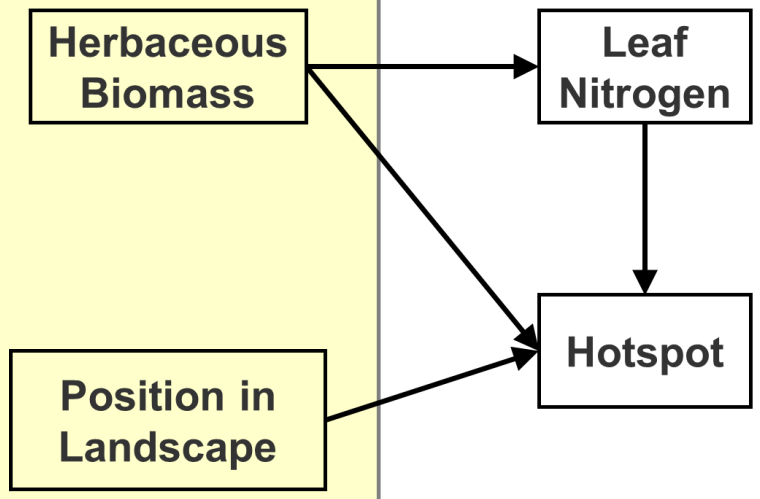






GLM in SEM

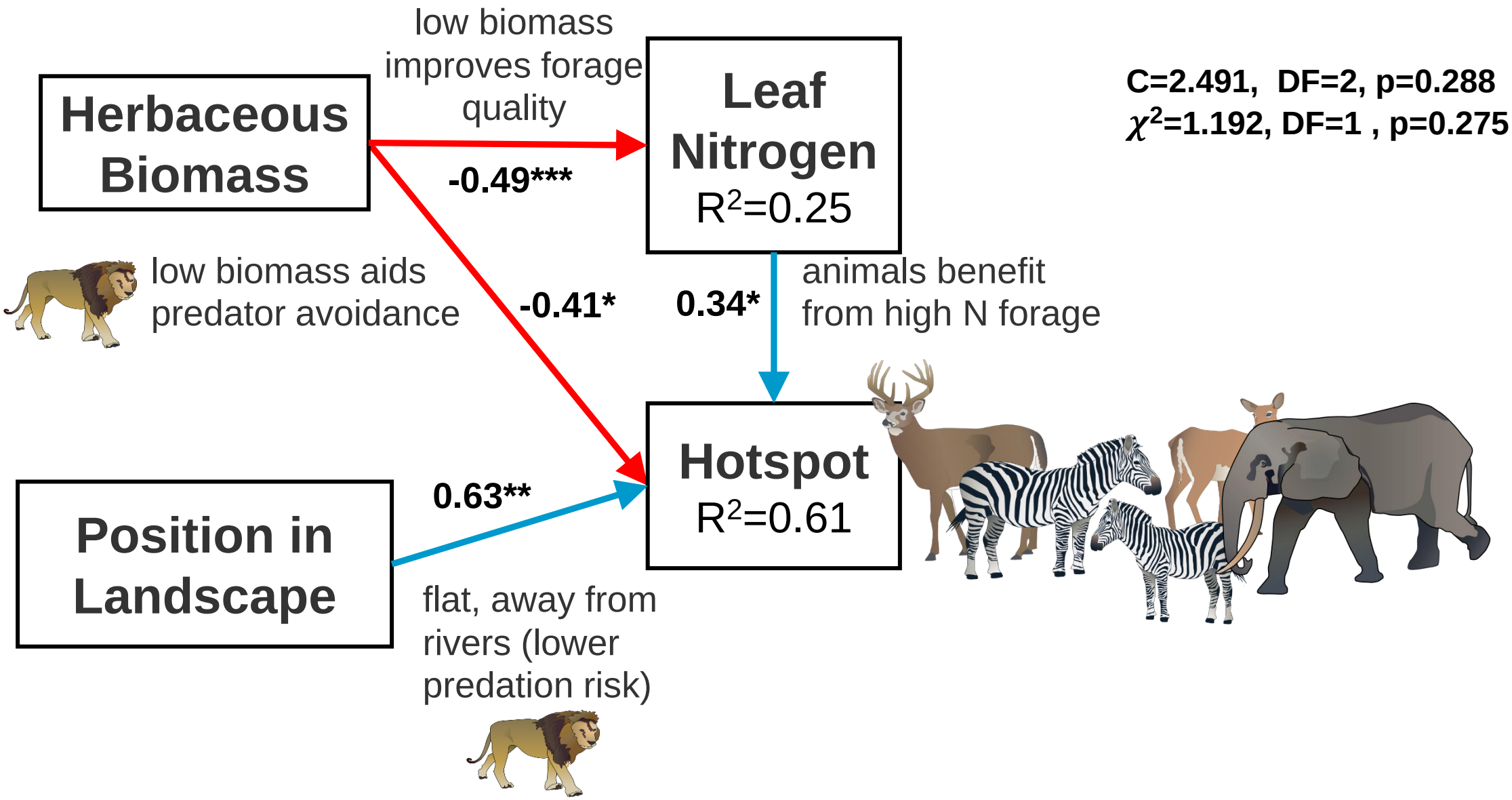
```
> summary(anderson.sem)

Structural Equation Model of anderson.sem
Call:
  leafN ~ biomass.kg
  hotspotYN ~ leafN + biomass.kg + landscape

    AIC
 4.617

---
Tests of directed separation:

           Independ.Claim Test.Type DF Crit.Value P.Value
  leafN ~ landscape + ...      coef 64    -1.0718  0.2878
--
Global goodness-of-fit:

Chi-Squared = 1.192 with P-value = 0.275 and on 1 degrees of freedom
Fisher's C = 2.491 with P-value = 0.288 and on 2 degrees of freedom
```



GLM in SEM

```
> ...
---
Coefficients:

   Response  Predictor Estimate Std.Error DF Crit.Value P.Value Std.Estimate
      leafN biomass.kg  -0.4880    0.1050 65    -4.6486  0.0000      -0.4995 ***
  hotspotYN      leafN   6.6867    2.7818 63     2.4037  0.0162       0.3399   *
  hotspotYN biomass.kg  -7.7838    3.5694 63    -2.1807  0.0292      -0.4050   *
  hotspotYN  landscape   1.3600    0.4955 63     2.7449  0.0061       0.6332  **

  Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05

---
Individual R-squared:

   Response     method R.squared
      leafN       none      0.25
  hotspotYN nagelkerke      0.61
```

GLM in SEM

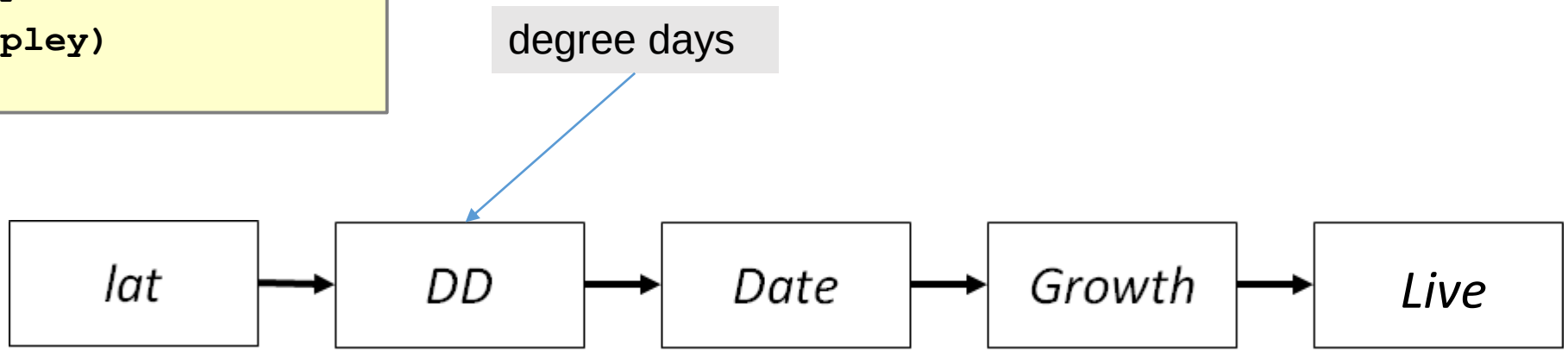


C=2.491, DF=2, p=0.288

χ^2=1.192, DF=1 , p=0.275

Day 9 Task 3


```
# Shipley data
library(piecewiseSEM)
data(shipley)
```



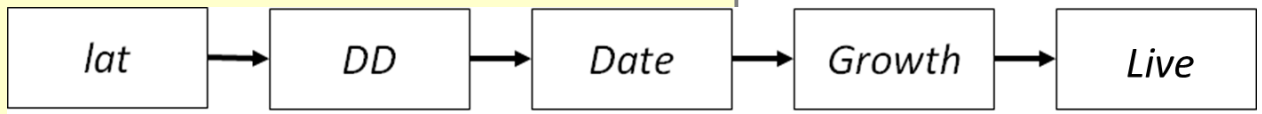
- Dataset: predicting latitude effect on survival of a tree species
- Repeated measures on 5 trees at 20 sites from 1970-2006
- Live (0/1) influenced by phenology (degree days until bud break, Julian days until bud break), size (stem diameter growth)

Day 9 Task 3


```
# Shipley data
library(piecewiseSEM)
data(shipley)
> str(shipley)
'data.frame': 1900 obs. of  9 variables:
 $ site    : int  1 1 1 1 1 1 1 1 1 1 ...
 $ tree    : int  1 2 3 4 5 1 2 3 4 5 ...
 $ lat     : num  40.4 40.4 40.4 40.4 40.4 ...
 $ year    : int  1970 1970 1970 1970 1970 1972 1972 1972 1972
1972 ...
 $ Date    : num  115 118 116 111 121 ...
 $ DD      : num  161 159 160 161 157 ...
 $ Growth  : num  61.4 43.8 44.7 48.2 50 ...
 $ Survival: num  1 0.843 0.944 0.957 0.976 ...
 $ Live    : int  1 1 1 1 1 1 1 1 1 1 ...
```

Day 9 Task 3


```
library(nlme)
library(lme4)

 lme(DD ~ lat, random = ~ 1 | site / tree, na.action = na.omit,
     data = shipley),

  lme(Date ~ DD, random = ~ 1 | site / tree, na.action = na.omit,
     data = shipley),

  lme(Growth ~ Date, random = ~ 1 | site / tree, na.action = na.omit,
     data = shipley),

  glmer(Live ~ Growth + (1 | site) + (1 | tree),
       family = binomial(link = "logit"), data = shipley)
```



Task: Use these sub-models in piecewiseSEM as a part of the SEM model shown above. Think about the study design and explain the results.