Tasks

Introduction to structural equation modeling and mixed models in R

Day 4: SEM

Oksana Buzhdygan

oksana.buzh@fu-berlin.de

Covariance-based assumptions & model fit

Day 4 Task 1

Day 4 Task 1

California, USA.

Photos credit: USFS, and Jon Keeley, USGS

doi.org/10.1186/s42408-019-0041-0

doi.org/10.1071/WF07049

Postfire recovery of plant communities in California shrublands

A number of measures were taken, including:

- Vegetation cover **"cover"**
- Age of stands that burned **"age"**
- Fire severity **"firesev"**

```
# Keeley data
library(piecewiseSEM)
data(keeley)
```

Data: Grace, J.B. and Keeley, J.E. 2006. A structural equation model analysis of postfire plant diversity in California shrublands. Ecological Applications 16:503-514

Day 4 Task 1

Postfire recovery of plant communities in California shrublands

Other measurements:

- Vegetation species richness **"richness"**
- Local abiotic conditions (aspect, soils) **"abiotic"**
- Spatial heterogenity **"hetero"**
- Distance from coast **"distance"**

Measurements:

- Vegetation cover **"cover"**
- Age of stands that burned **"age"**
- Fire severity **"firesev"**

```
# Keeley data
library(piecewiseSEM)
data(keeley)
```

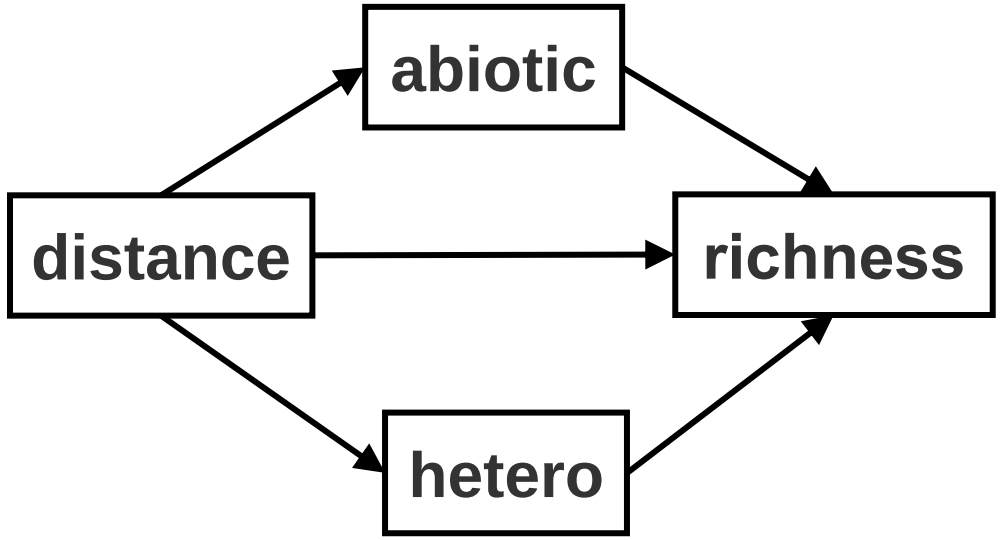
Data: Grace, J.B. and Keeley, J.E. 2006. A structural equation model analysis of postfire plant diversity in California shrublands. Ecological Applications 16:503-514

Day 4 Task 1

```
# Keeley data
library(piecewiseSEM)
data(keeley)
```

Other measurements:

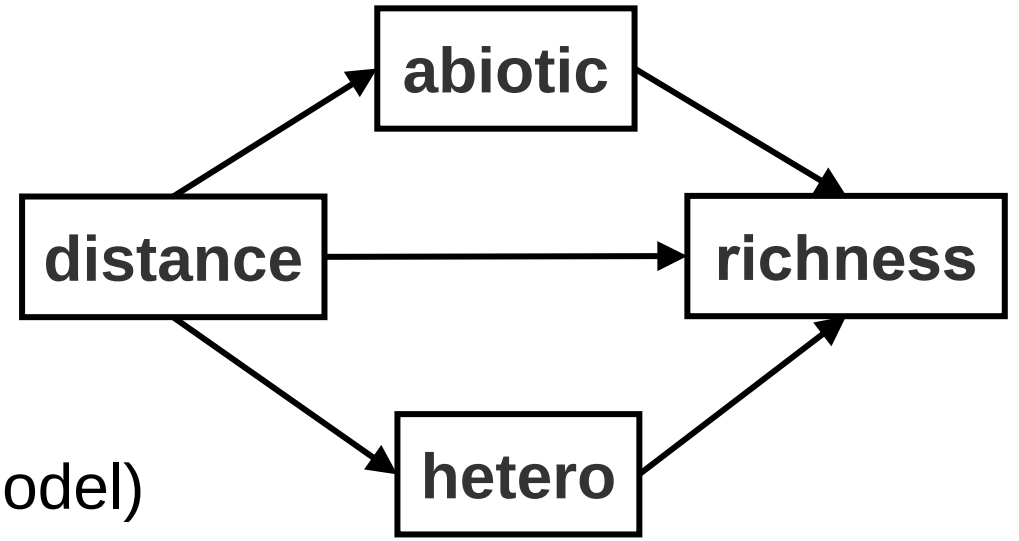
- Vegetation species richness **"richness"**
- Local abiotic conditions (aspect, soils) **"abiotic"**
- Spatial heterogenity **"hetero"**
- Distance from coast **"distance"**



Data: Grace, J.B. and Keeley, J.E. 2006. A structural equation model analysis of postfire plant diversity in California shrublands. Ecological Applications 16:503-514

Day 4 Task 1

1. Specify the following model in lavaan
2. Check assumptions for covariance-based SEM
   - normality of residuals
   - multivariate normality of data
   - multicolinearity
     (function `vif(lm_model)` for each regression model)



3. Fit the model using data `data(keeley)`
4. Get the fit indices
5. Fill in Standardized Coeficients and R^2 for the model
6. Calculate indirect and total effects of distance on plant richness.
   What would you say about direct and indirect effects in this system?

Day 4 Task 1

When you fit the model

```
# Error about data scales
Warning message:
In lav_data_full(data = data, group = group, cluster = cluster,  :
  lavaan WARNING: some observed variances are (at least) a factor 1000 times larger than
others; use varTable(fit) to investigate
```

Day 4 Task 1

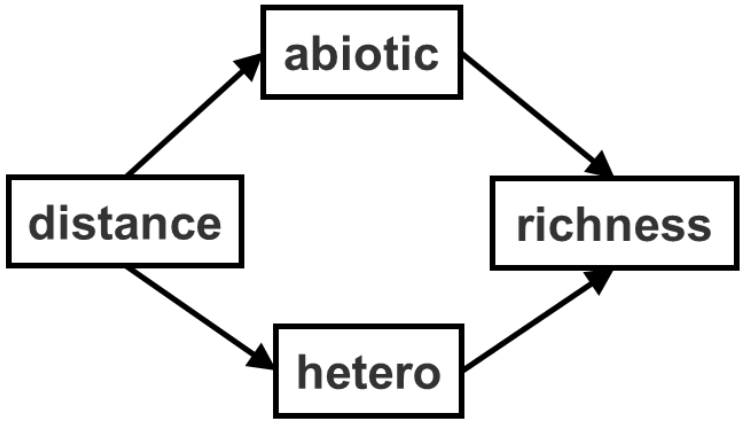
```
# Call the model-implied covariance matrix

lavInspect(SemFit, "obs")$cov



  # Check the data scales

  varTable(SemFit)
```

Model Comparison

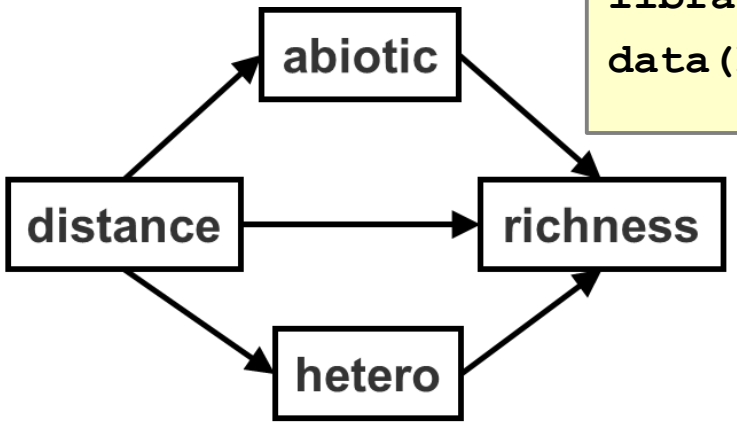
Day 4 Task 2

Day 4 Task 2 Model Comparison

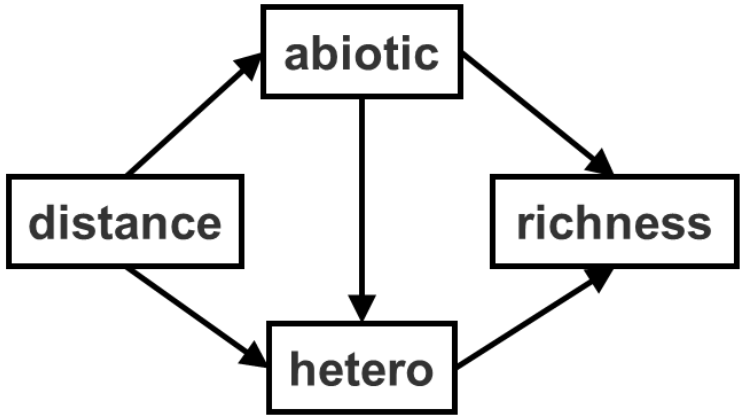
```
# Keeley data
library(piecewiseSEM)
data(keeley)
```



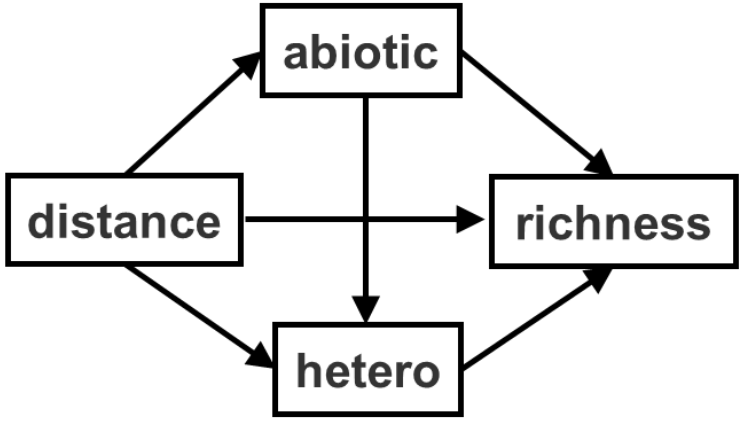
mod1



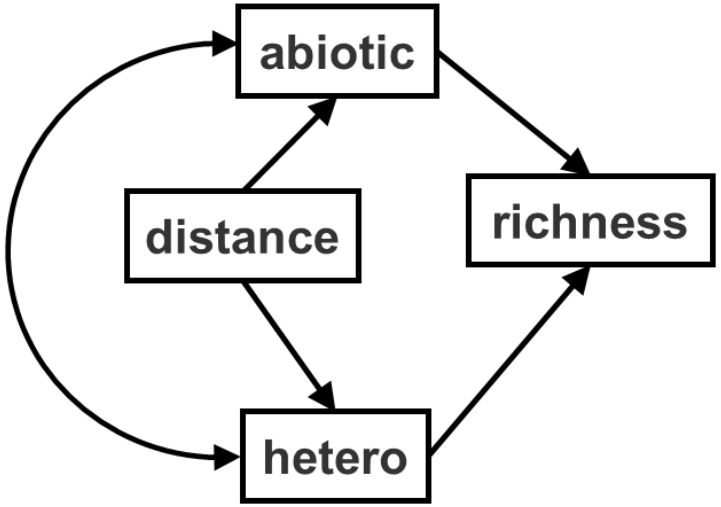
mod2



mod3



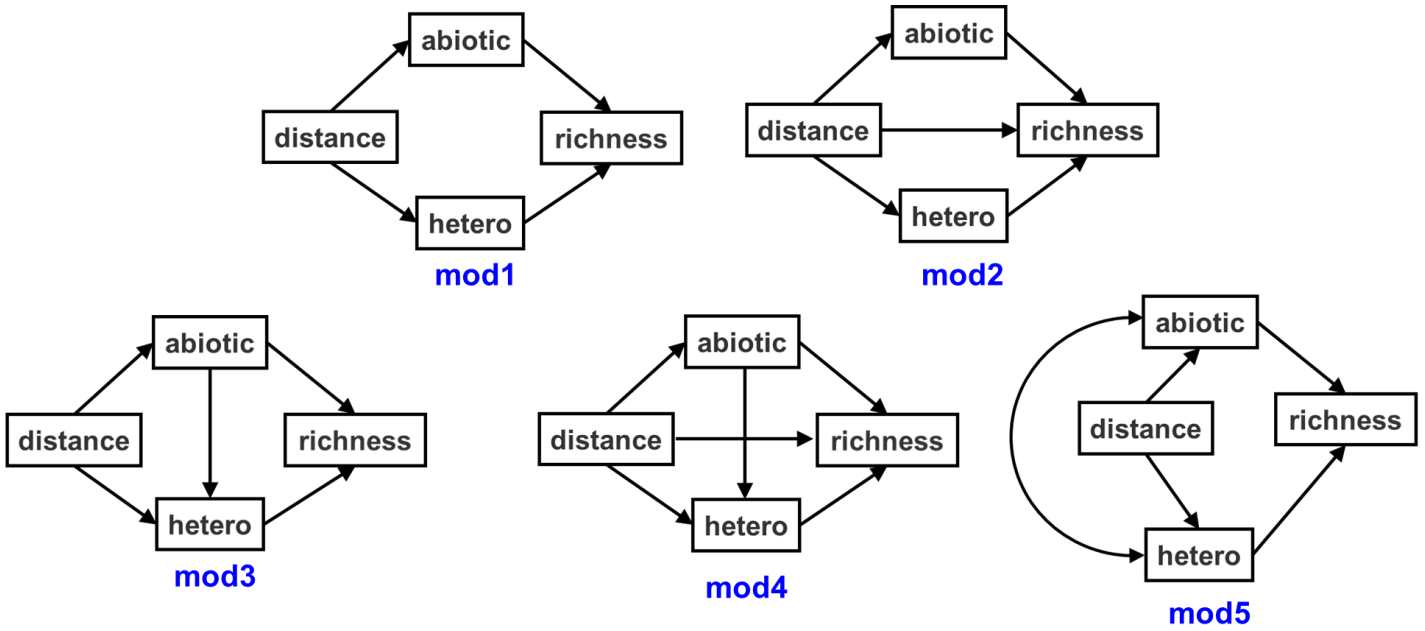
mod4



mod5

Day 4 Task 2

1. Think about the scientific meaning when we include the correlations among “hetero” and “abiotic” (in model 5)
2. Specify each model in lavaan
3. Fit each model using data `data(keeley)`
4. Compare the models using AICc and select the best model

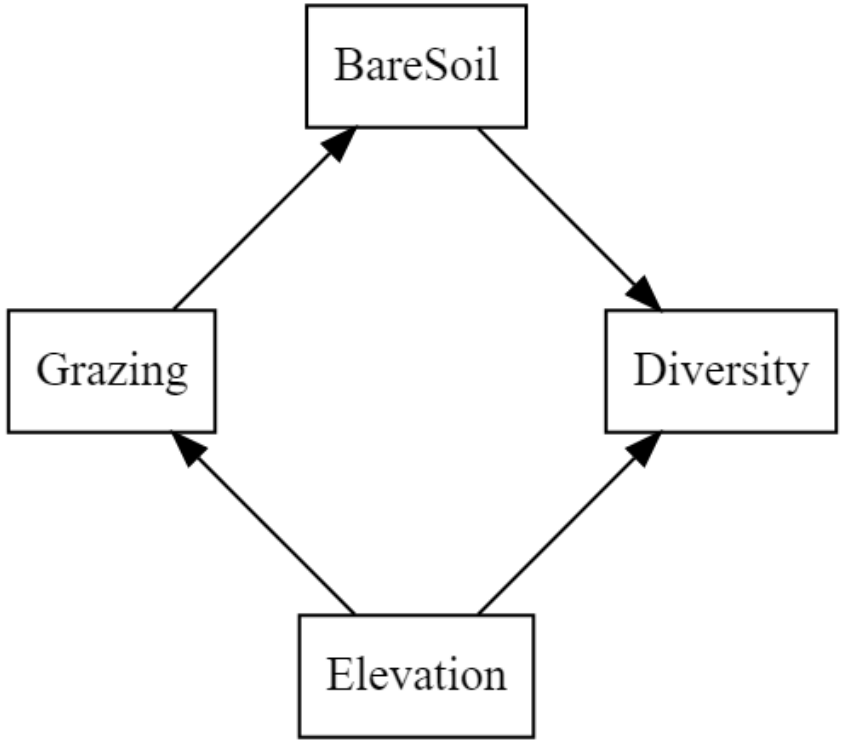


Test of Mediation

Day 4 Task 3

Day 3 Task 3 Test of mediation

Effects of grazing on plant diversity along elevation gradient




```
# data
data <- read.csv("Grassl_data_2.csv")
```

Day 3 Task 3

For the model on Fig. 1:

1. For the variables on Fig 1 and data "Grassl_data_**2**.csv" check the assumptions of the covariance-based SEM
2. If there are any violations of the assumptions use the needed procedures to adjust for these violations (see slide 30 in part 1 Day 4) while you fit the model.
3. Get the fit indices.
4. Follow the procedures in this lecture (part 2) and test the following questions:

- Are we ignoring important links?
- Are all the included links supported by the data?

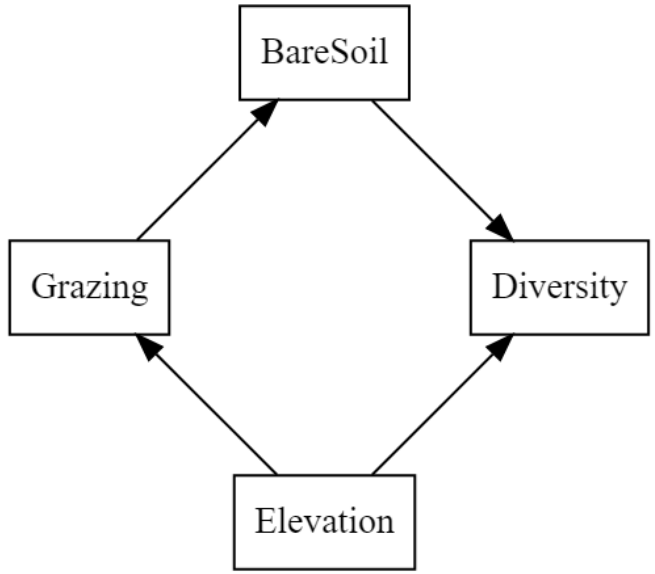


Fig. 1

Day 4 Task 3

For the model on Fig. 1:

1. For the variables on Fig 1 and data "Grassl_data_**2**.csv" check the assumptions of the covariance-based SEM
2. If there are any violations of the assumptions use the needed procedures to adjust for these violations (see slide 30 in part 1 Day 4) while you fit the model.
3. Get the fit indices.
4. Follow the procedures in this lecture (part 2) and test the following questions:

- Are we ignoring important links?
- Are all the included links supported by the data?

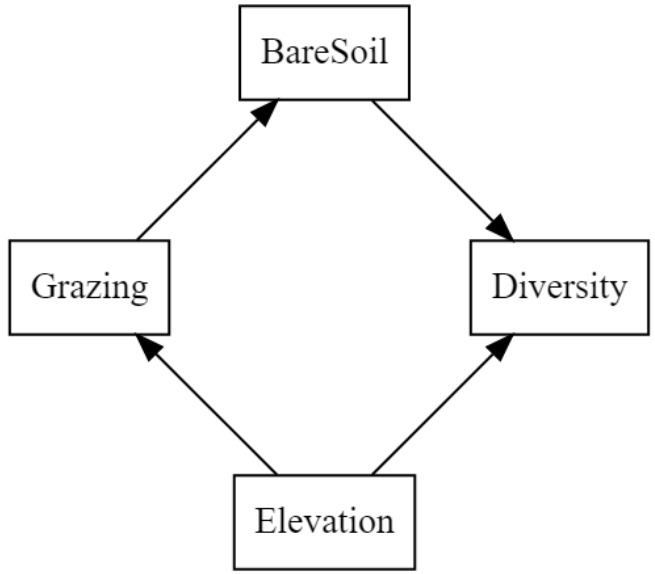


Fig. 1