GENERALISED LEAST SQUARES

Alejandro Ordonez

Assistant Professor - Department of Biology

Section for Ecoinformatics & Biodiversity

Center for Biodiversity Dynamics in a Changing World (BIOCHANGE)

MIXED EFFECT MODELS
WHAT WE WILL TALK ABOUT TODAY

The basics of generalises Least Squares – the mathematical principle

Dealing with Heteroscedasticity – Solving non equal variance

- Modelling the model variance– Solving non-equal variances
- Autocorrelation– Solving non-independence

Model selection – Combing ML and RMEL

Any questions?
Ready to start?

A GENERAL LINEAR MODEL

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \boxed{\begin{bmatrix} 1 & x_{11} & x_{12} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}} + \boxed{\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}}$$

Fixed part

Error part

A GENERAL LINEAR MODEL

Random part

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \boxed{\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}}$$

Normally distributed errors

$$\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} : N_{(0,\mathbf{V})}$$

Homogeneity of variances

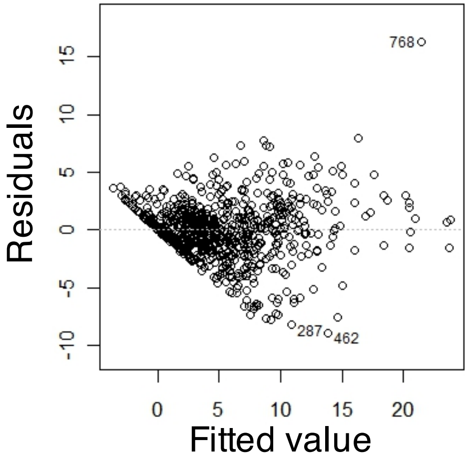
$$\mathbf{V} = \begin{bmatrix} \sigma^2 & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & \sigma^2 \end{bmatrix}$$

Variance-covariance matrix

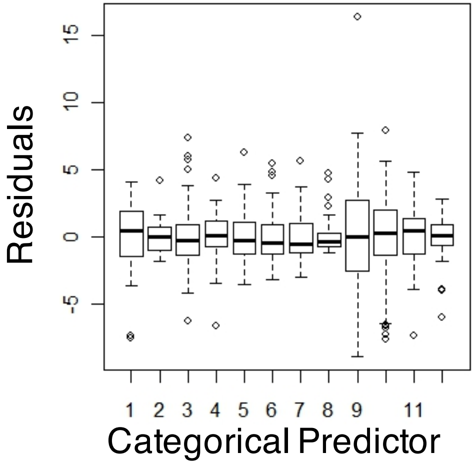
VIOLATING THE ASSUMPTIONS OF A GENERAL LINEAR MODEL

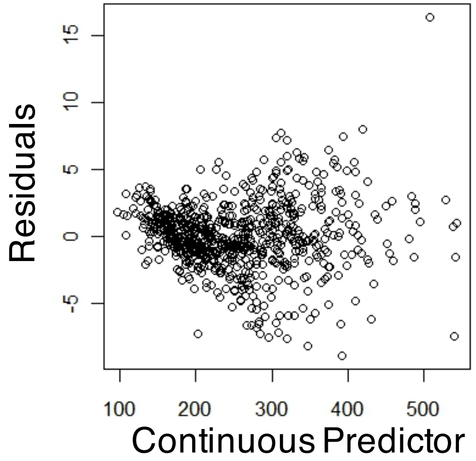
What problem is visualised here?

Homogeneity of variances



Independence





The variance of error terms is not similar across the independent variable(s)

Heteroscedasticity

How can you see that problem in these figures?

VIOLATING THE HOMOGENEITY OF VARIANCES ASSUMPTION

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \boxed{\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}}$$

Random part

Variances are not equal across observations

Normally distributed errors

$$\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} : N_{(0,\mathbf{V})}$$

Variance heterogeneity

$$\mathbf{V} = \begin{bmatrix} \sigma^2_{(1)}(\mathbf{X}) & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & \sigma^2_{(n)}(\mathbf{X}) \end{bmatrix}$$

VIOLATING THE HOMOGENEITY OF VARIANCES ASSUMPTION

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \boxed{\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}}$$

Random part

However observations are **still** independent

Normally distributed errors

$$\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} : N_{(0,\mathbf{V})}$$

Variance heterogeneity

$$\mathbf{V} = \begin{bmatrix} \sigma^2_{(1)}(\mathbf{X}) & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & \sigma^2_{(n)}(\mathbf{X}) \end{bmatrix}$$

MODELING VARIANCES IN R
`GLS` NOT `LM` MODELS

Remember that in **basic** linear models parameters are estimated using **Ordinary least squares**.

But..

As we move to model the variance we a need a method of estimation that:

- Model the fixed effects of the predictors
- Accounts for structure in the error term → how the variance changes!

Generalised Least Squares (GLS) is a method of estimation that allows accounting for structure in the error term

MODELING VARIANCES IN R WITH GLS

```
gls (y ~ x ,
      weights = VAR.Fn
         method = “ML/REML”)
```

Name of the function in R	What does it do?
`VarIdent`	Different variances per stratum
`VarFixed`	Fixed variance
`VarPower`	Power of the variance covariate
`VarExp`	Exponential of the variance covariate
`VarConstPower`	Constant plus power of the variance covariate
`VarComb`	A combination of variance functions

The syntax is almost the same as an `lm` but You can specify a description of the within-group heteroscedasticity structure as **weights**

MODELING VARIANCES IN R WITH GLS

```
gls (y ~ x ,
      weights = VAR.Fn
      method = "ML/REML")
```

The general approach

Gaussian errors

```
MASS::glmPQL (y ~ x ,
      weights = VAR.Fn
      family = "poisson")
```

The generalised approach

Gaussian/binomial/Poisson errors

So far so good?

Any questions?

Ready to move on?

HOW TO DEAL WITH UNEQUAL VARIANCES

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Variance heterogeneity

$$\mathbf{V} = \begin{bmatrix} \sigma^2_{(1)}(\mathbf{X}) & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & \sigma^2_{(n)}(\mathbf{X}) \end{bmatrix}$$

Build a model to describe how the variability changes as a function of the data

R does this by adding a weights to the variance of each observation

HOW TO DEAL WITH UNEQUAL VARIANCES

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Variance heterogeneity

$$\mathbf{V} = \begin{bmatrix} \sigma^2_{(1)}(\mathbf{X}) & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & \sigma^2_{(n)}(\mathbf{X}) \end{bmatrix}$$

Weights are included by defining a function to describe the change in variance between observations

$\sigma^2_{(i)} \sim \sigma^2_{(class_i)}$ → Variance depend on a category

$\sigma^2_{(i)} \sim \dfrac{\sigma^2}{x_i}$ → Variance **proportional** to a variance covariate

$\sigma^2_{(i)} \sim \sigma^2 F_{(x_i)}$ → Variance as a **function** of a variance covariate

DIFFERENT VARIANCES PER STRATUM

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

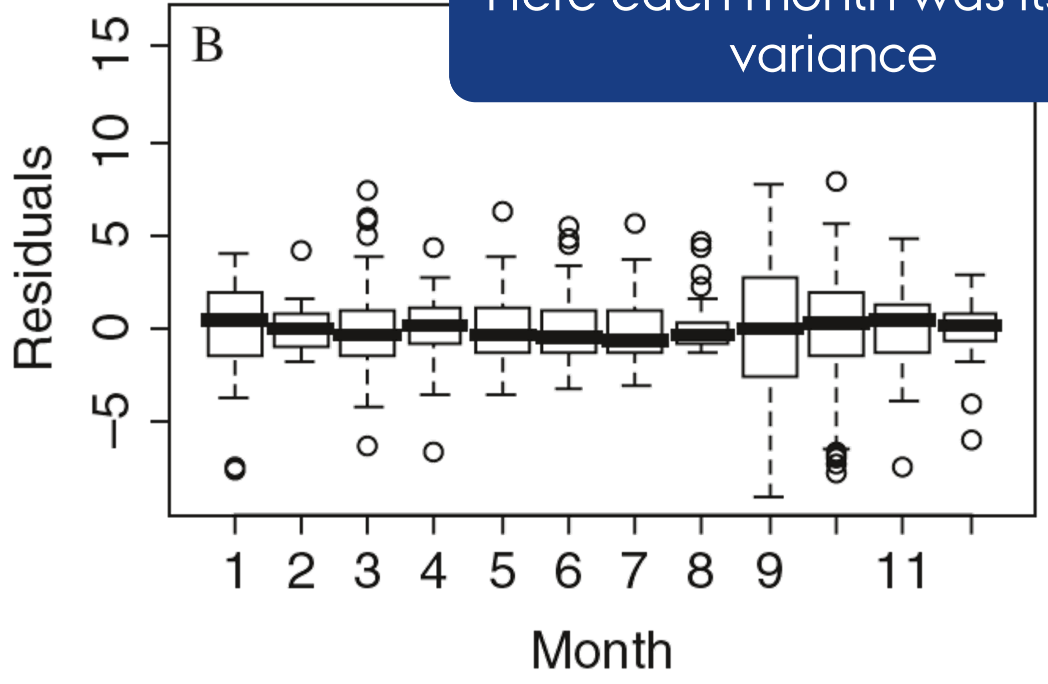
Variance heterogeneity

$$\mathbf{V} = \begin{bmatrix} \sigma^2_{(1)}\boldsymbol{X} & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & \sigma^2_{(n)}\boldsymbol{X} \end{bmatrix}$$

$$\sigma^2_{(i)} \sim \sigma^2_{(class_i)}$$

Variance depend on a category

Here each month was its own variance



WEIGHTED VARIANCES

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

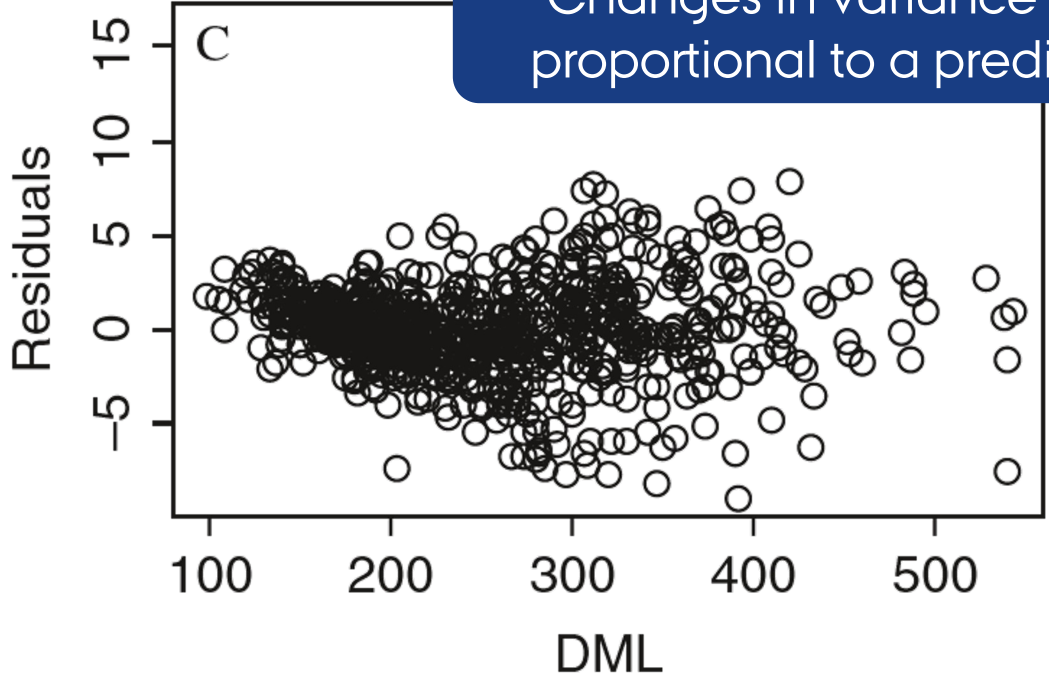
Variance heterogeneity

$$\mathbf{V} = \begin{bmatrix} \sigma^2_{(1)}\boldsymbol{X} & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & \sigma^2_{(n)}\boldsymbol{X} \end{bmatrix}$$

$$\sigma^2_{(i)} \sim \frac{\sigma^2}{x_i} \rightarrow$$

Variance as a function of a predictor(s)

Changes in variance are proportional to a predictor



DIFFERENT VARIANCES PER STRATUM

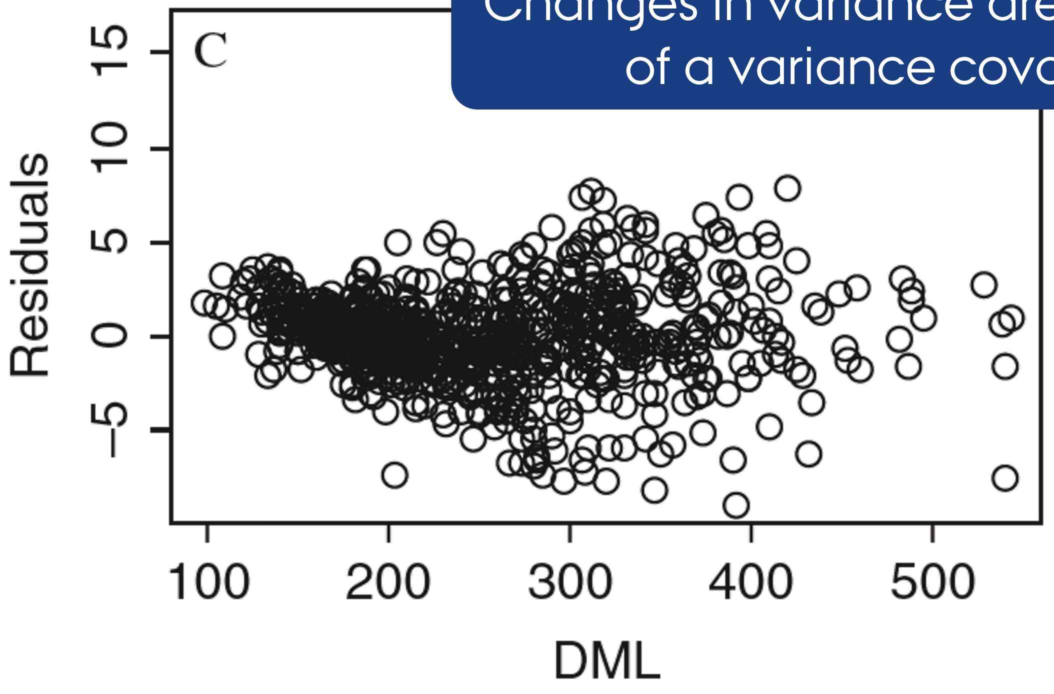
$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Variance heterogeneity

$$\mathbf{V} = \begin{bmatrix} \sigma^2_{(1)}\boldsymbol{X} & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & \sigma^2_{(n)}\boldsymbol{X} \end{bmatrix}$$

$\sigma^2_{(i)} \sim \sigma^2 F_{(x_i)}$ → Variance as a function of a variance covariate

Changes in variance are a function of a variance covariate



VARIANCES AS A FUNCTION

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Variance heterogeneity

$$\mathbf{V} = \begin{bmatrix} \sigma^2_{(1)}\boldsymbol{X} & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & \sigma^2_{(n)}\boldsymbol{X} \end{bmatrix}$$

$$\sigma^2_{(i)} \sim \sigma^2 F_{(x_i)}$$ → Variance as a function of a variance covariate

Functions can be of different types:

$\sigma^2_{(i)} \sim \sigma^2 \times |x_i|^{2\delta}$ Power Function

$\sigma^2_{(i)} \sim \sigma^2 \times e^{2\delta \times x_i}$ Exponential Function

$\sigma^2_{(i)} \sim \sigma^2 \times \left(\delta_1 + |x_i|^{2\delta_2}\right)^2$ Constant + Power Function

VARIANCES AS A FUNCTION

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Variance heterogeneity

$$\mathbf{V} = \begin{bmatrix} \sigma_{(1)}^2 \boldsymbol{X} & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & \sigma_{(n)}^2 \boldsymbol{X} \end{bmatrix}$$

Functions can combine categorical and continuous variance covariate(s)

$\sigma_{(i)}^2 \sim \sigma^2 \times |x_i \times Class_i|^{2\delta}$ Power Function

$\sigma_{(i)}^2 \sim \sigma^2 \times e^{2\delta \times x_i \times Class_i}$ Exponential Function

$\sigma_{(i)}^2 \sim \sigma^2 \times \left(\delta_1 + |x_i \times Class_i|^{2\delta_2}\right)^2$ Constant + Power Function

So far so good?

Any questions?

Ready to move on?

OVERVIEW OF ALL VARIANCE STRUCTURES

Name of the function in R	What does it do?
`VarIdent`	Different variances per stratum
`VarFixed`	Fixed variance
`VarPower`	Power of the variance covariate
`VarExp`	Exponential of the variance covariate
`VarConstPower`	Constant plus power of the variance covariate
`VarComb`	A combination of variance functions

Rather limited – Problematic when Variance-Convariance takes negative Values

OVERVIEW OF ALL VARIANCE STRUCTURES

Name of the function in R	What does it do?
`VarIdent`	Different variances per stratum
`VarFixed`	Fixed variance
`VarPower`	Power of the variance covariate
`VarExp`	Exponential of the variance covariate
`VarConstPower`	Constant plus power of the variance covariate
`VarComb`	A combination of variance functions

Flexible, but **do not** use if the variance covariate takes the value of zero

OVERVIEW OF ALL VARIANCE STRUCTURES

Name of the function in R	What does it do?
`VarIdent`	Different variances per stratum
`VarFixed`	Fixed variance
`VarPower`	Power of the variance covariate
`VarExp`	Exponential of the variance covariate
`VarConstPower`	Constant plus power of the variance covariate
`VarComb`	A combination of variance functions

Flexible, **can use** if the variance covariate takes the value of zero

So far so good?

Any questions?

Ready to move on?

MIXED EFFECTS MODELS
THE PROBLEM OF INDEPENDENCE

Observational/experiments can have:

- Spatial-Temporal-Evolutionary proximity
  - Autocorrelation → removes independence between samples
- A nested observational/experimental design
  - samples taken across different spatial-temporal-taxonomic scales

What does this mean?

Samples have “Pseudoreplication”

MIXED EFFECTS MODELS
(PSEUDO)REPLICATION

What is replication?

Independent observations of same measurement

What is (Pseudo)replication?

Replicates are **not independent** (spatial-temporal), or "treatments are truly **not replicated**"

What is independent?

Independent is when the value of one observation does not affect the value of other observations

INDEPENDENCE OF OBSERVATIONS

What does it mean?

- Objects are sampled only once.
- The occurrence of one event doesn't change the probability for another.
- Sampling of one observation does not affect the choice of the second observation.

How can we ensure independence of observations?

- Sample each object only once.
- Make sure that replicates have no relation → if they do, they are repetitions.
- Use a fully randomized design!!!

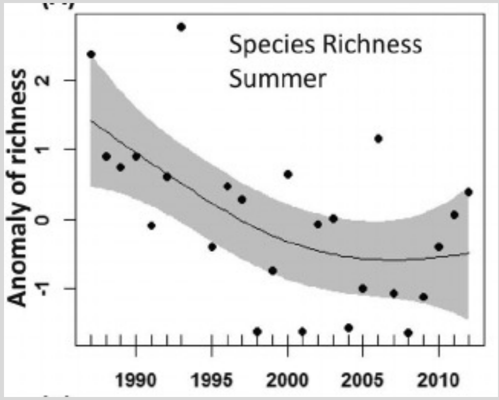
So far so good?

Any questions?

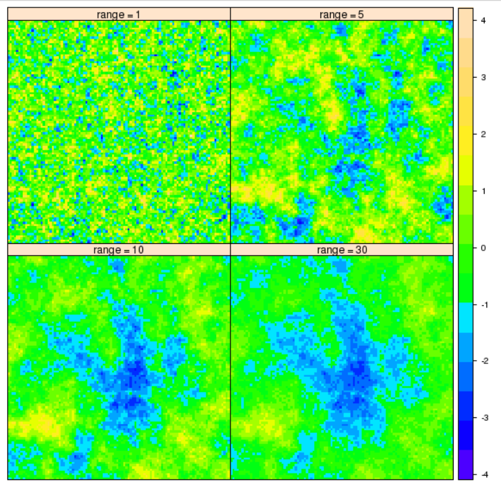
Ready to move on?

AUTOCORRELATION:
A MORE COMPLEX VARIANCE STRUCTURE

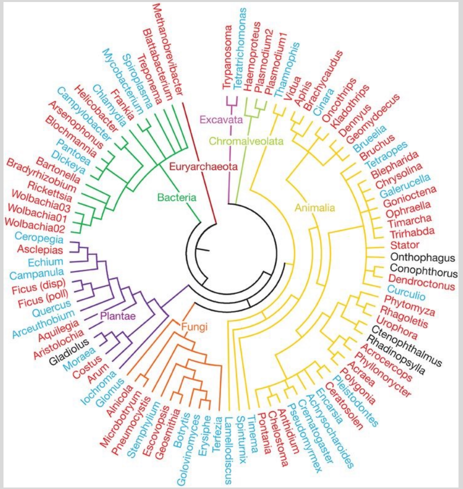
Time



Space



Species



AUTOCORRELATION
THE MATH BEHIND THE MODEL

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \boxed{\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}}$$

Random part

Normally distributed errors

$$\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} : N_{(0,\mathbf{V})}$$

Variance heterogeneity

$$\mathbf{V} = \begin{bmatrix} \sigma^2_{(1,1)} & \cdots & \sigma^2_{(j,1)} \\ \vdots & \vdots & \vdots \\ \sigma^2_{(1,n)} & \cdots & \sigma^2_{(j,n)} \end{bmatrix}$$

Observations are not independent

Variances are not equal across observations

MODELING CORRELATIONS IN R WITH GLS

```
gls (y ~ x ,
    correlation = Cor.Fn,
      method = "ML/REML")
```

```
MASS::glmmPQL (y ~ x ,
        correlation = Cor.Fn,
          family = "poisson")
```

The syntax is almost the same as an `lm` but You can specify a description of the correlation between observations as a **correlation**

Name of the function in R	What does it do?
`corAR1`	autoregressive process of order 1
`corARMA`	autoregressive moving average process
`corExp`	exponential spatial correlation
`corGaus`	Gaussian spatial correlation
`corLin`	linear spatial correlation
`corRatio`	Rational quadratics spatial correlation
`corSpher`	spherical spatial correlation
`corBrownian`	Brownian motion mode
`corPagel`	The covariance matrix defined Freckelton et al. (2002)

So far so good?

Any questions?

Ready to move on?

TEMPORAL AUTOCORRELATION
THE MATH BEHIND THE MODEL

$$\varepsilon_s \sim N(0, \sigma^2)$$

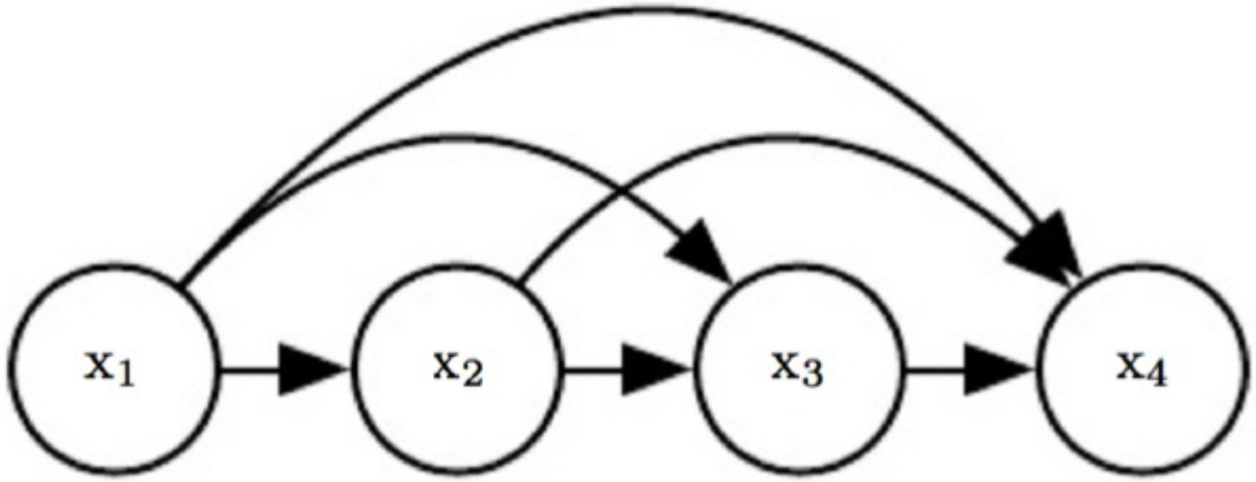
$$cov(\varepsilon_s, \varepsilon_t) = \begin{cases} \sigma^2 \ if \ s = t \\ \boxed{h(\varepsilon_s, \varepsilon_{s+k}, \rho) \ else} \end{cases}$$

Correlation function
function that measures the spatial dependence between two sites

We assume stationarity
Correlation between the residuals ε_s and ε_t only depends on their time difference (s – t)

The goal
Find the best optimal parameterisation of the correlation function

TEMPORAL AUTOCORRELATION
THE TEMPORAL CASE

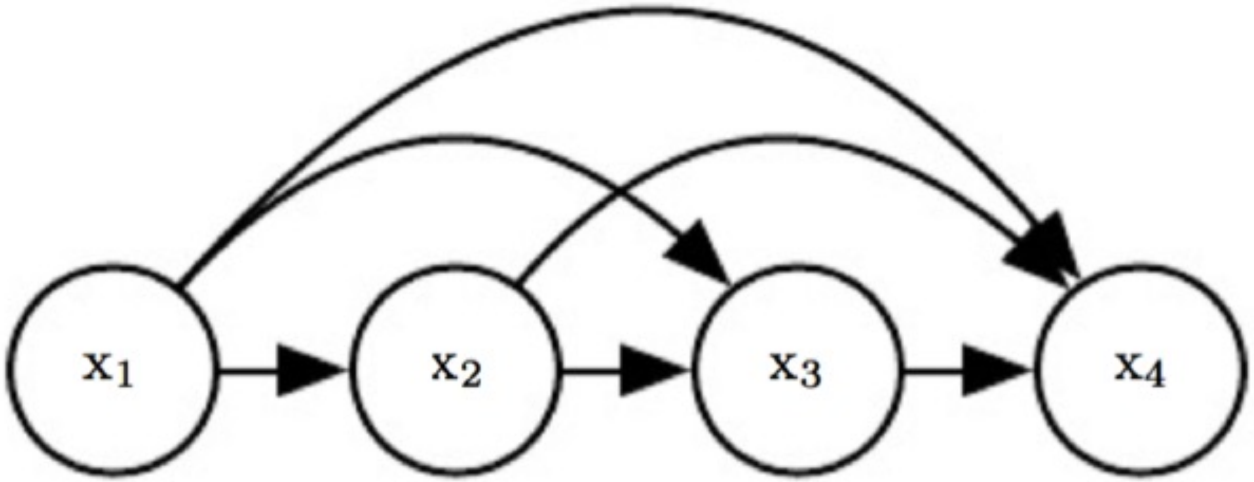


Output variable depends linearly on its own previous values

That means :
To estimate the response at time t_n we need to know the value/errors of the Predictor Xi at that time t_n AND the state of the response at time t_{n-1}

TEMPORAL AUTOCORRELATION

AUTOREGRESSIVE MODEL



Output variable depends linearly on its own previous values

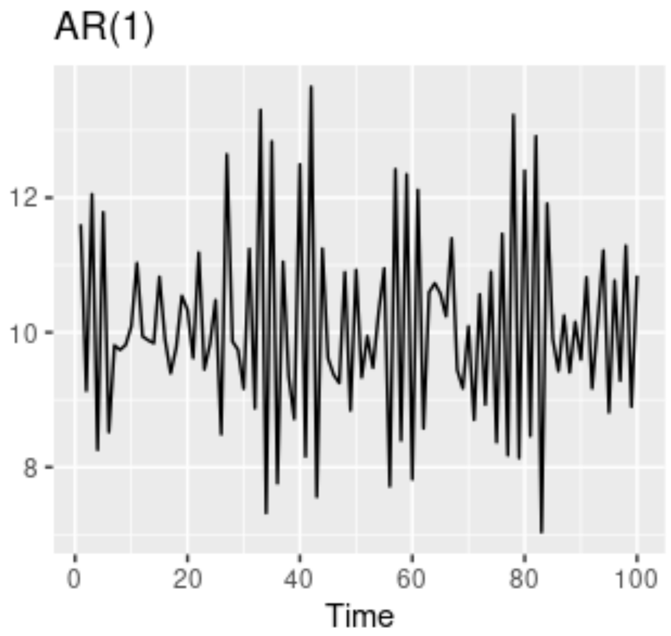
If we **use the preceding value,** we call this an **Autoregressive** model

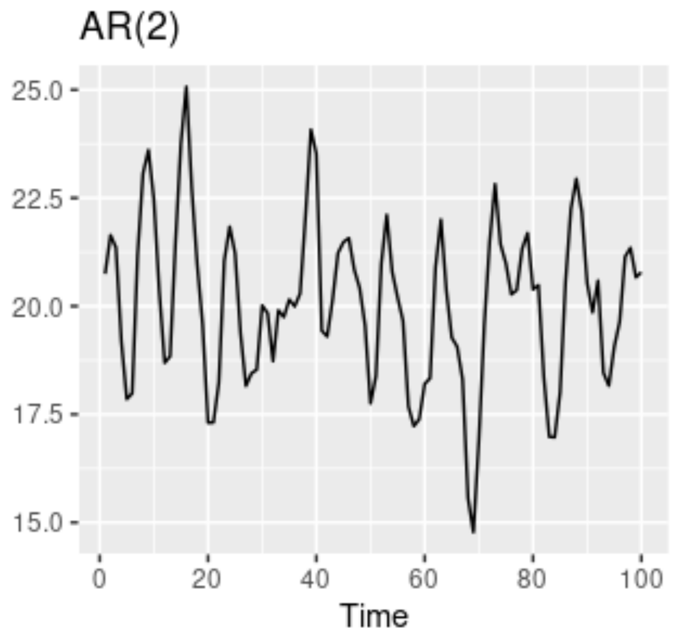
Autoregression = regression of the variable against itself.

$$X_t = \sum_{i=1}^{p} \varphi_i X_{t-1} + \varepsilon_t$$

TEMPORAL AUTOCORRELATION
AUTOREGRESSIVE MODEL

AR(1) process is one in which the current value is based on the immediately preceding value



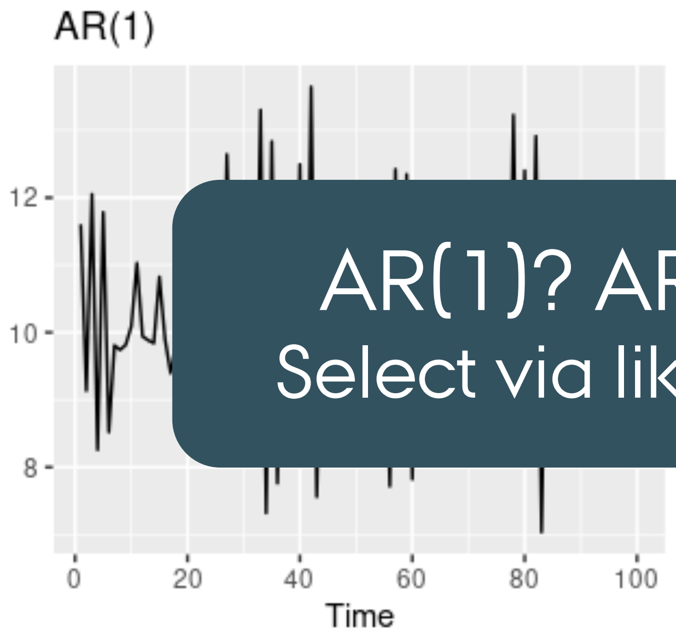


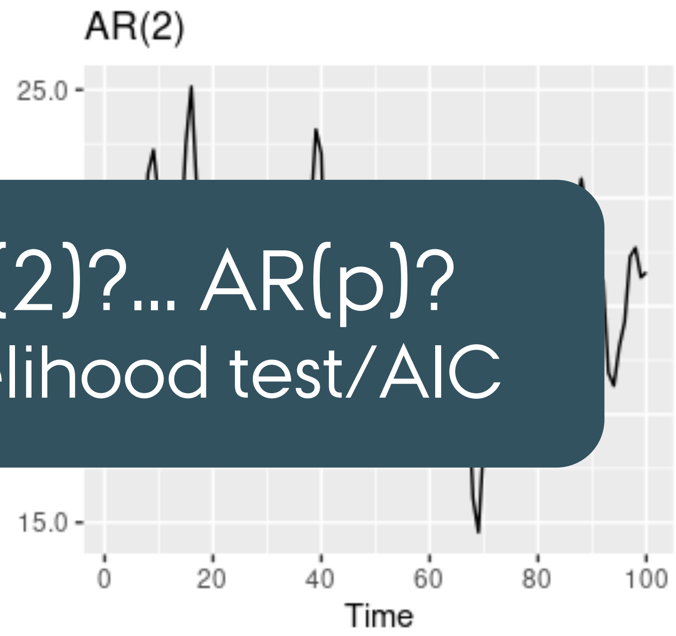
AR(2) process is one in which the current value is based on the previous two values

The correlation is uni-directional as past events afects future ones, not the other way

TEMPORAL AUTOCORRELATION
AUTOREGRESSIVE MODEL

AR(1) process is one in which the current value is based on the immediately preceding value



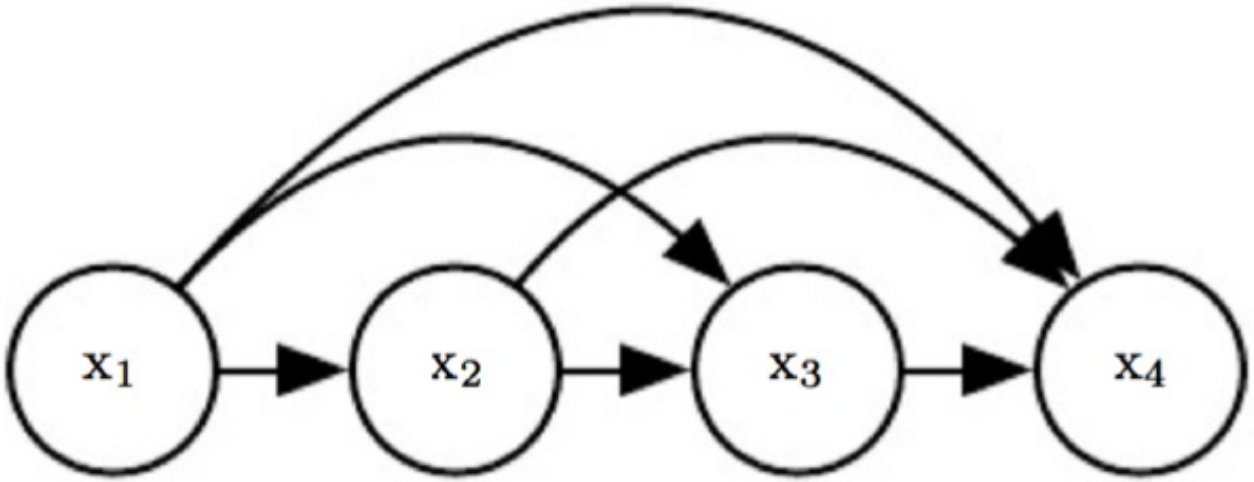


AR(1)? AR(2)?... AR(p)?
Select via likelihood test/AIC

AR(2) process is one in which the current value is based on the previous two values

The correlation is uni-directional as past events afects future ones, not the other way

TEMPORAL AUTOCORRELATION
MOVING AVERAGE MODEL



Output variable depends linearly on its own previous values

If we **use the preceding error(s)**, we call this a **Moving average** model

Error defines the difference form the time series mean.

$$X_t = \mu + \sum_{i=1}^{q} \theta_i \varepsilon_{t-1} + \varepsilon_t$$

TEMPORAL AUTOCORRELATION
AUTOREGRESIVE + MOVIG AVERAGE

Autoregresive
Forcast based on the observations

Movig average
Forcast based on the Errors

ARMA
AutoRegressive
Moving Average

TEMPORAL AUTOCORRELATION

R – IMPLEMENTATION

autoregressive parameters

Covariate with time

```
corAR1(value, form)
```

```
corARMA(value, form, p, q)
```

autoregressive & **moving average** parameters

Covariate with time

autoregressive and moving average order

So far so good?

Any questions?

Ready to move on?

SPATIAL AUTOCORRELATION
THE MATH BEHIND THE MODEL

Random part

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \boxed{\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}}$$

Normally distributed errors

$$\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} : N_{(0,\mathbf{V})}$$

Variance heterogeneity

$$\mathbf{V} = \begin{bmatrix} \sigma^2_{(1,1)} & \cdots & \sigma^2_{(j,1)} \\ \vdots & \vdots & \vdots \\ \sigma^2_{(1,n)} & \cdots & \sigma^2_{(j,n)} \end{bmatrix}$$

Observations are not independent

Variances are not equal across observations

SPATIAL AUTOCORRELATION
FITTING CORRELATION FUNCTIONS

Exponential correlation
corExp

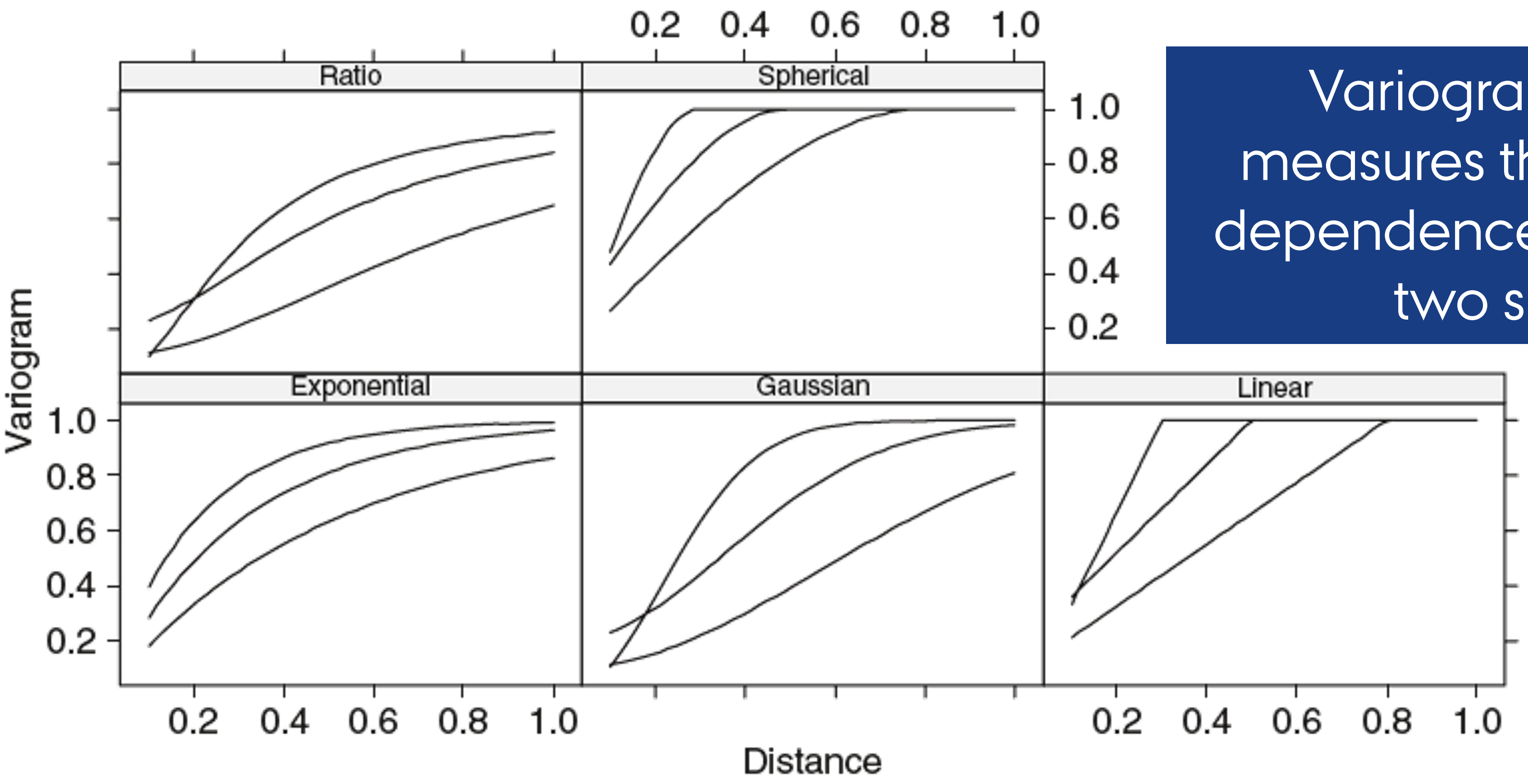
Spherical correlation
corSpher

Linear correlation
corLin

Gaussian correlation
corGaus

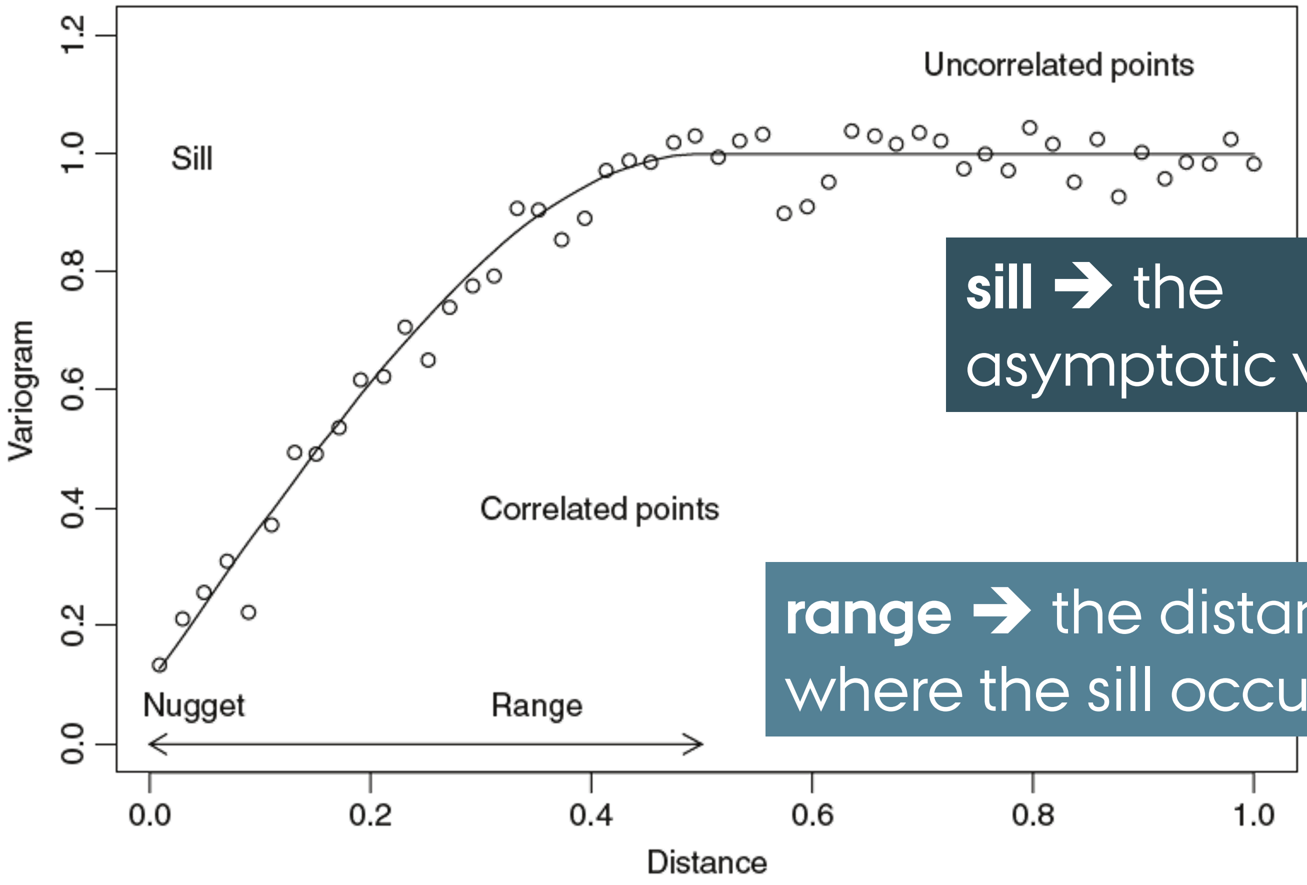
Rational quadratic correlation
corRatio

SPATIAL AUTOCORRELATION
FITTING CORRELATION FUNCTIONS



Variograms are measures the spatial dependence between two sites

SPATIAL AUTOCORRELATION
FITTING CORRELATION FUNCTIONS



sill → the asymptotic value

range → the distance where the sill occurs

nugget → the y-value when the distance is 0.

SPATIAL AUTOCORRELATION
R – IMPLEMENTATION

```
corSpher( form, nugget, metric)
corLin   ( form, nugget , metric)
corRatio( form, nugget , metric)
corGaus  ( form, nugget , metric)
corExp   ( form, nugget , metric)
```

The spatial covarianties	Nugget effect	distance metric
Lat/Long x/y	Cor at lag dist =0	Euclidean? Manhattan?

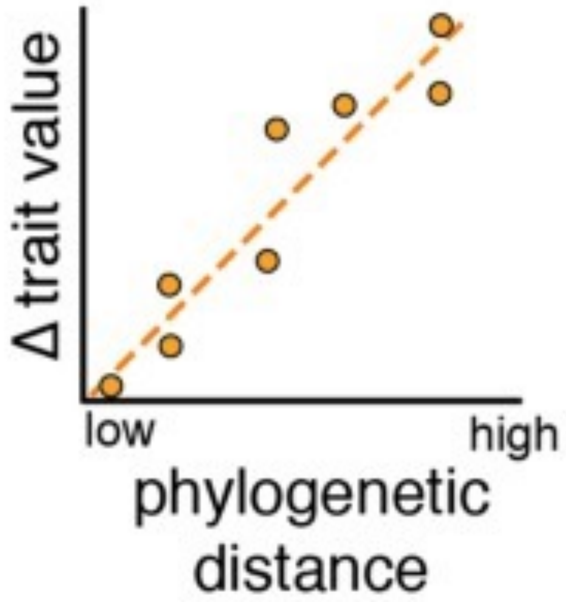
So far so good?

Any questions?

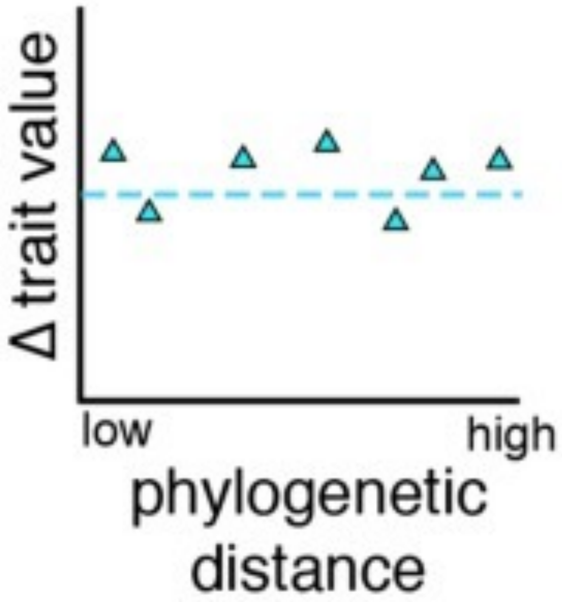
Ready to move on?

PHYLOGENETIC AUTOCORRELATION

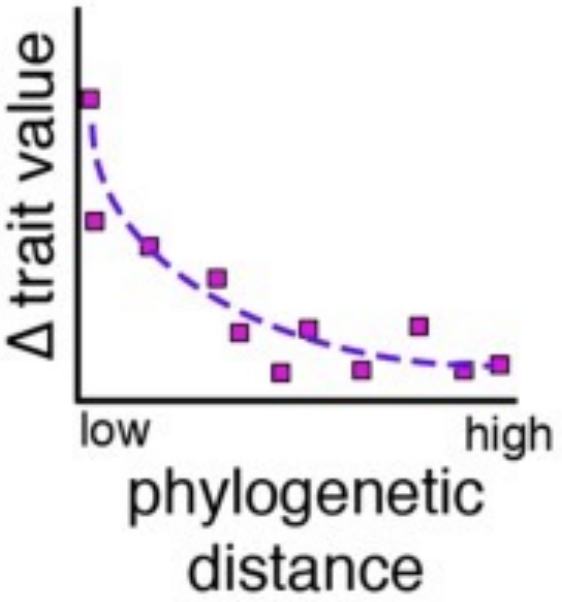
Positive Phylogenetic Autocorrelation
Evolutionary similarity means similar traits



No Phylogenetic Autocorrelation
Evolutionary similarity does not affect trait similarity



Negative Phylogenetic Autocorrelation
Evolutionary Similarity means different traits

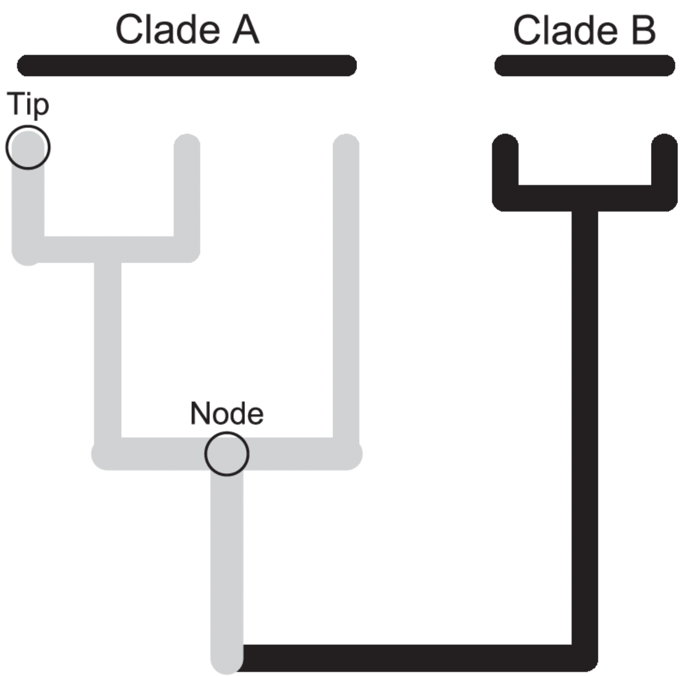


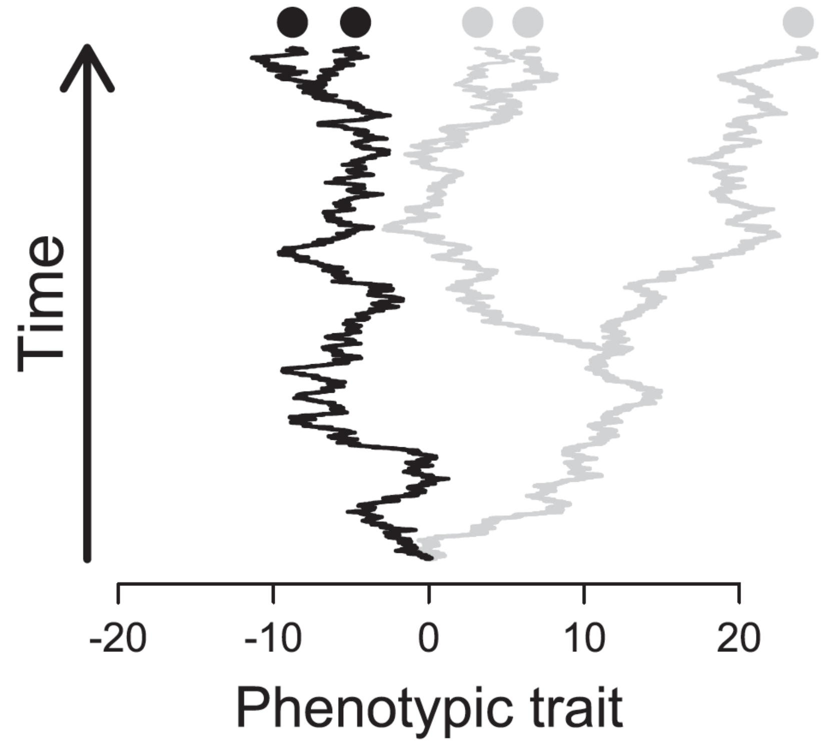
PHYLOGENETIC AUTOCORRELATION

The correlation between species/nodes is determined by evolutionary:

Tempo
Amount of change per unit time

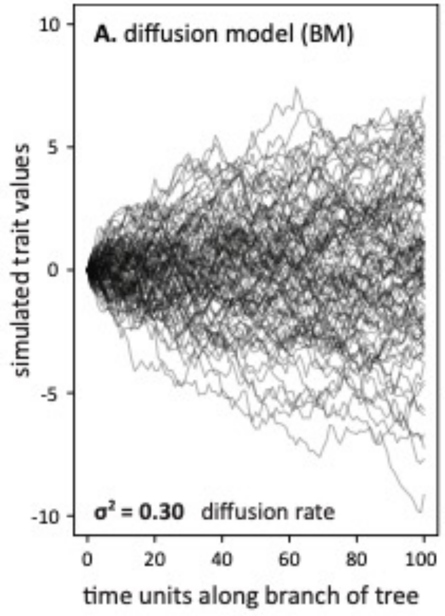
Mode
mechanisms responsible for evolution and their patterns



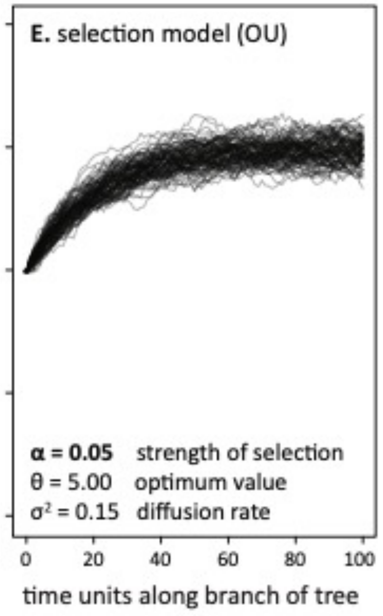


PHYLOGENETIC AUTOCORRELATION
MODELS OF EVOLUTION

Brownian motion
Random drift and constant rate through time and across lineages



Ornstein–Uhlenbeck
Changes got towards a fitness optimum on an adaptive landscape



PHYLOGENETIC AUTOCORRELATION
R – IMPLEMENTATION [BROWNIAN]

Phylogeny

```
corBrownian(phy)
```

```
corPagel(value, phy)
```

value of the parameter λ
λ=Phylo Signal

Phylogeny

PHYLOGENETIC AUTOCORRELATION
R-IMPLEMENTATION [ORNSTEIN-UHLENBECK]

```
corMartins(value, phy)
```

value of the parameter α
α=optima

Phylogeny

So far so good?

Any questions?

Ready to move on?

MODEL SELECTION – TWO STEPS

Model selection is done in two stages:

1. Define the best variance/correlation structure.
   - Comapre a full model in fixed effects to one with a given variance/correlation structure.
   - Full and variance/correlation model are fitted using "REML"

2. Define the best fixed structure.
   - Reduce the pool of fixed effects from the model selected above
   - Models beeing reduced are build using "ML".

MODEL SELECTION –STEP-1

Define the best variance/correlation structure

```
B1A<- gls(A ~ towns + region + Sex, data= eire.Dat,
         correlation=corSpher(form=~x + y, nugget=TRUE),
         method = 'REML')
```

One note of model Selection

If you need to include and the explanation of the (co)variance it is more important to have an expectation of it in you model that finding the "best" description.

```
         method = 'REML')
B1E <- gls(A ~ towns + region + Sex, data= eire.Dat,
         correlation=corExp(form=~x + y, nugget=TRUE) ,
         method = 'REML')
```

MODEL SELECTION –STEP-1

Define the best fixed structure.

```
B1EFull <- gls(A ~ towns + region + Sex, data= eire.Dat,
            correlation=corExp(form=~x + y, nugget=TRUE) ,
             method = 'ML')
B1EMod1 <- gls(A ~ towns + region, data= eire.Dat,
            correlation=corExp(form=~x + y, nugget=TRUE) ,
             method = 'ML')
B1EMod2 <- gls(A ~ towns + Sex, data= eire.Dat,
            correlation=corExp(form=~x + y, nugget=TRUE) ,
             method = 'ML')
B1EMod3 <- gls(A ~ region + Sex, data= eire.Dat,
            correlation=corExp(form=~x + y, nugget=TRUE) ,
             method = 'ML')
```

RESTRICTED MAXIMUM LIKELIHOOD (REML) MAXIMUM LIKELIHOOD (ML) ANALYSIS

Maximum Likelihood (**ML**) Analysis

- Focused on getting the coefficients of the linear model right.
- That is the focus is predicting the mean correctly.

Restricted Maximum Likelihood (**REML**)

- Focused on getting the variances in the model right.
- That is the focus is predicting the variance correctly.

A good explanation of the diff between these is here:
https://towardsdatascience.com/maximum-likelihood-ml-vs-reml-78cf79bef2cf

So far so good?
Any questions?
Ready to finish?

IN SUMMARY

1. Generalized Least Squares is an approach that allows controlling for:
   - Heterogeneous variances: via modelling the variance using multiple strategies.
   - Autocorrelated Variables: Via the inclusion of correlation structures
     - Autoregresing / moving averages
     - Semivarigram/correlogram
     - Phylogenetic relatedness

IN SUMMARY

2. Model selection in mixed models is a two step process:
   - **Part one** – define the best variance/covariance structure
     - Models build using REML
   - **Part two** – define the best fixed effect structure
     - Models build using ML

AARHUS
UNIVERSITY