

# Temporal Analysis of Zoobenthos Data

Trang Nguyen; El Park

05 mars, 2025

## Biodiversity Analysis of Zoobenthos in the Baltic Sea

### 1. Introduction

In this analysis, we explore the dynamics of zoobenthos species richness and their relationship with environmental factors in the Baltic Sea over a time span from 1980 to 2005. The analysis incorporates various methods, including species richness calculation, time series analysis, beta diversity, rank shift analysis, and community stability assessment

1. Questions for Analysis In order to guide this project and achieve a deeper understanding of zoobenthos biodiversity in the Baltic Sea, the following questions were formulated:

What is the temporal pattern of zoobenthos species richness across different sites?

How does species richness change over time at different sites in the Baltic Sea? Are there any significant trends, peaks, or declines in species richness at specific times or sites? How does site location affect species richness?

Is there a difference in species richness between different sites (S0, S2, S3) over time? Are some sites more stable in terms of species richness than others? What is the turnover rate of species at each site?

How much does species composition change over time? Are there years where new species appear or existing species disappear significantly? What is the Mean Rank Shift (MRS) of species?

How much do species rankings change from one year to the next? Does the rank order of species remain stable, or are there significant shifts in the species composition over time? How stable are the zoobenthos communities over time?

What is the variability in species richness across years at each site, and how does this affect community stability? Can we identify periods of community instability or resilience?

### 2. Data Preparation

#### 2.1 Setting Up the Environment

```
# Set up working directory (use your own directory path)
rm(list=ls())
getwd()
```

```
## [1] "C:/Users/ttran/OneDrive - Indiana University/SP25 - Quantitative Biodiversity/QB2025-Project"
```

```

# Load the required packages
package.list = c('vegan', 'tidyr', 'dplyr', 'codyn', 'ggplot2', 'nlme', 'forecast', 'emmeans', 'tseries')
for (package in package.list) {
  if (!require(package, character.only = TRUE, quietly = TRUE)) {
    install.packages(package, repos='http://cran.us.r-project.org')
    library(package, character.only = TRUE)
  }
}

```

```
## Warning: le package 'vegan' a été compilé avec la version R 4.4.2
```

```
## Warning: le package 'permute' a été compilé avec la version R 4.4.2
```

```
## This is vegan 2.6-8
```

```
##
```

```
## Attachement du package : 'dplyr'
```

```
## Les objets suivants sont masqués depuis 'package:stats':
```

```
##
```

```
##      filter, lag
```

```
## Les objets suivants sont masqués depuis 'package:base':
```

```
##
```

```
##      intersect, setdiff, setequal, union
```

```
## Warning: le package 'codyn' a été compilé avec la version R 4.4.2
```

```
## Warning: le package 'ggplot2' a été compilé avec la version R 4.4.2
```

```
##
```

```
## Attachement du package : 'nlme'
```

```
## L'objet suivant est masqué depuis 'package:dplyr':
```

```
##
```

```
##      collapse
```

```
## Warning: le package 'forecast' a été compilé avec la version R 4.4.2
```

```
## Registered S3 method overwritten by 'quantmod':
```

```
##      method      from
```

```
##      as.zoo.data.frame zoo
```

```
##
```

```
## Attachement du package : 'forecast'
```

```
## L'objet suivant est masqué depuis 'package:nlme':
```

```
##
```

```
##      getResponse
```

```
## Warning: le package 'emmeans' a été compilé avec la version R 4.4.2
```

```
## Welcome to emmeans.
```

```
## Caution: You lose important information if you filter this package's results.
```

```
## See '? untidy'
```

```
## Warning: le package 'tseries' a été compilé avec la version R 4.4.2
```

## 2.2 Loading the Dataset

We load the dataset containing information on zoobenthos species abundance and environmental factors.

```
# Load the dataset
data = read.csv("data/data_temporal.csv") # Update path to your dataset
```

## 2.3 Converting to a Time-by-Species Matrix

To analyze species richness across years and sites, we create a time-by-species matrix.

```
# Convert to time-by-species matrix: Group by year and site_id, count species
time.by.species = group_by(data, year, site_id) %>%
  count(species) %>%
  spread(key = species, value = n, fill = 0)

# Calculate observed richness per site and year
richness = rowSums(time.by.species[, -c(1:2)]) # Remove year and site_id columns
richness_data = data.frame(time.by.species, richness)

# Convert columns to appropriate formats
richness_data$year = as.factor(richness_data$year)
richness_data$site_id = as.factor(richness_data$site_id)
richness_data$richness = as.numeric(richness_data$richness)

head(richness_data)
```

```
##   year site_id Abra.alba Acanthodoris.pilosa Actinia.equina Alcyonidium.polyoum
## 1 1984      S0          1              0            0              0
## 2 1984      S2          0              0            0              0
## 3 1984      S3          0              0            0              0
## 4 1985      S0          1              0            0              0
## 5 1985      S2          0              0            0              0
## 6 1985      S3          0              0            0              0
##   Amauropsis.islandica Ampharete.acutifrons Ampharete.baltica Arctica.islandica
## 1                    0                    0                    0              1
## 2                    0                    0                    1              0
## 3                    0                    0                    0              1
## 4                    0                    0                    0              1
## 5                    0                    0                    1              0
## 6                    0                    0                    0              1
##   Arenicola.marina Aricidea.cerrutii Aricidea.suecica Astarte.borealis
## 1                    0                    0                    0              0
```

## 2	0	0	0	1
## 3	0	0	0	0
## 4	0	0	0	0
## 5	0	0	0	1
## 6	0	0	0	0
##	Astarte.elliptica	Astarte montagui	Asterias.rubens	Balanus.crenatus
## 1	0	0	0	0
## 2	0	0	0	0
## 3	1	0	0	0
## 4	0	0	0	0
## 5	0	0	0	0
## 6	1	0	0	0
##	Balanus.improvisus	Bathyporeia.pilosa	Bougainvillia.ramosa	Bylgides.sarsi
## 1	0	0	0	1
## 2	0	0	0	0
## 3	0	0	0	0
## 4	0	0	0	1
## 5	0	0	0	1
## 6	0	0	0	0
##	Callopora.lineata	Capitella.capitata	Cephalothrix	Cerastoderma.glaucum
## 1	0	0	0	0
## 2	0	0	0	0
## 3	0	0	0	0
## 4	0	0	0	0
## 5	0	0	0	0
## 6	0	1	0	0
##	Cirrophorus.eliasoni	Clytia.hemisphaerica	Corbula.gibba	Corophium.crassicorne
## 1	0	0	1	0
## 2	0	0	0	0
## 3	0	0	0	0
## 4	0	0	1	0
## 5	0	0	0	0
## 6	0	0	0	0
##	Corophium.insidiosum	Corophium.volutator	Crangon.crangon	Diastylis.rathkei
## 1	0	0	0	1
## 2	1	0	0	1
## 3	0	0	0	0
## 4	0	0	0	1
## 5	1	0	0	1
## 6	0	0	0	1
##	Electra.crustulenta	Electra.pilosa	Eteone.flava	Eteone.longa
## 1	0	0	0	0
## 2	0	0	0	0
## 3	0	0	0	0
## 4	0	0	0	0
## 5	0	0	0	0
## 6	0	0	0	0
##	Euchone.papillosa	Eucratea.loricata	Eulalia.bilineata	Facelina.bostoniensis
## 1	0	0	0	0
## 2	0	0	0	0
## 3	0	0	0	0
## 4	0	0	0	0
## 5	0	0	0	0
## 6	0	0	0	0

##	Gammarellus.homari	Gammarus.oceanicus	Gammarus.salinus	Gammarus.zaddachi
## 1	0	0	0	0
## 2	0	1	0	0
## 3	0	0	0	0
## 4	0	0	0	0
## 5	0	1	0	0
## 6	0	0	0	0
##	Gastrosaccus.spinifer	Gonothyraea.loveni	Halacaridae	Halicryptus.spinulosus
## 1	0	0	0	1
## 2	0	0	0	0
## 3	0	0	0	1
## 4	0	0	0	1
## 5	0	0	0	1
## 6	0	0	0	1
##	Halisarca.dujardinii	Halitholus.yoldiaarcticae	Harmothoe.imbricata	
## 1	0	0	0	
## 2	0	0	0	
## 3	0	0	0	
## 4	0	0	0	
## 5	0	0	0	
## 6	0	0	0	
##	Harmothoe.impar	Hartlaubella.gelatinosa	Hediste.diversicolor	
## 1	0	0	0	
## 2	0	0	0	
## 3	0	0	0	
## 4	0	0	0	
## 5	0	0	0	
## 6	0	0	0	
##	Heteromastus.filiformis	Hydrobia.ulvae	Idotea.balthica	Jaera.albifrons
## 1	1	0	0	0
## 2	0	1	0	0
## 3	1	0	0	0
## 4	1	0	0	0
## 5	0	1	0	0
## 6	1	0	0	0
##	Lafoeina.tenuis	Lagis.koreni	Lamellidoris.muricata	Laomedea.flexuosa
## 1	0	1	0	0
## 2	0	0	0	0
## 3	0	0	0	0
## 4	0	1	0	0
## 5	0	0	0	0
## 6	0	0	0	0
##	Lepidonotus.squamatus	Levinsenia.gracilis	Lineus.ruber	Macoma.balthica
## 1	0	0	0	0
## 2	0	0	1	1
## 3	0	0	0	0
## 4	0	0	0	0
## 5	0	0	1	1
## 6	0	0	0	0
##	Malacobdella.grossa	Marenzelleria.neglecta	Metridium.senile	
## 1	0	0	0	
## 2	0	0	0	
## 3	0	0	0	
## 4	0	0	0	

## 5	0	0	0	
## 6	0	0	0	
##	Microdeutopus.gryllotalpa	Molgula.manhattensis	Musculus.discors	
## 1	0	0	0	
## 2	0	0	0	
## 3	0	0	0	
## 4	0	0	0	
## 5	0	0	0	
## 6	0	0	0	
##	Musculus.marmoratus	Mya.arenaria	Mya.truncata	Mysella.bidentata Mysis.mixta
## 1	0	0	0	0 0
## 2	0	1	0	0 0
## 3	0	0	0	0 0
## 4	0	0	0	0 0
## 5	0	1	0	0 0
## 6	0	0	0	0 0
##	Mytilus.edulis	Nais.elinguis	Nemertina	Neomysis.integer Nephtys Nephtys.caeca
## 1	0	0	0	0 0 0
## 2	1	0	1	0 0 0
## 3	0	0	0	0 0 0
## 4	0	0	0	0 0 0
## 5	1	0	1	0 0 0
## 6	0	0	0	0 0 0
##	Nephtys.ciliata	Nephtys.hombergii	Nereimyra.punctata	Nymphon.brevirostre
## 1	1	0	0	0
## 2	0	0	0	0
## 3	0	0	0	0
## 4	1	1	0	0
## 5	0	0	0	0
## 6	0	0	0	0
##	Odostomia.rissoides	Oligochaeta	Opercularella.lacerata	Opercularella.pumila
## 1	0	0	0	0
## 2	0	0	0	0
## 3	0	0	0	0
## 4	0	0	0	0
## 5	0	0	0	0
## 6	0	0	0	0
##	Ophelia.limacina	Ophiura.albida	Palaemon.elegans	Paraonis.fulgens
## 1	0	0	0	1
## 2	0	0	0	0
## 3	0	0	0	0
## 4	0	0	0	1
## 5	0	0	0	0
## 6	0	0	0	0
##	Parvicardium.ovale	Phaxas.pellucidus	Pherusa.plumosa	Philine.aperta Pholoe
## 1	0	1	0	0 0
## 2	0	0	0	0 0
## 3	0	0	0	0 0
## 4	0	1	0	0 0
## 5	0	0	0	0 0
## 6	0	0	0	0 0
##	Pholoe.assimilis	Pholoe.baltica	Pholoe.inornata	Phoxocephalus.holbolli
## 1	0	0	0	0
## 2	0	0	0	1

## 3	0	0	0	0
## 4	0	0	0	0
## 5	0	0	0	1
## 6	0	0	0	0
##	Phyllodoce	Phyllodoce..Anaitides..maculata	Phyllodoce..Anaitides..mucosa	
## 1	0		0	0
## 2	0		0	0
## 3	0		0	0
## 4	0		0	0
## 5	0		0	0
## 6	0		0	0
##	Polydora	Polydora..Polydora..quadrilobata	Polydora.caulleryi	Polydora.ciliata
## 1	0		1	0
## 2	0		1	0
## 3	0		0	0
## 4	0		1	0
## 5	0		1	0
## 6	0		0	0
##	Pontoporeia.femorata	Priapulus.caudatus	Prostoma.obscura	
## 1		0	1	0
## 2		0	0	0
## 3		0	0	0
## 4		0	1	0
## 5		0	0	0
## 6		0	0	0
##	Pseudopolydora.antennata	Pseudopolydora.pulchra	Pusillina.inconspicua	
## 1		0	0	0
## 2		0	0	0
## 3		0	0	0
## 4		0	0	0
## 5		0	0	0
## 6		0	0	0
##	Pygospio.elegans	Retusa.obtusa	Retusa.truncatula	Saduria.entomon
## 1	0	0	0	0
## 2	0	0	0	0
## 3	0	0	0	0
## 4	0	0	0	0
## 5	1	0	0	0
## 6	0	0	0	0
##	Scalibregma.inflatum	Scoloplos.armiger	Spio.filicornis	Spio.goniocephala
## 1	0	0	0	0
## 2	0	1	0	0
## 3	0	0	0	0
## 4	0	0	0	0
## 5	0	1	1	0
## 6	0	0	0	0
##	Streptosyllis.websteri	Terebellides.stroemii	Travisia.forbesii	
## 1	0	0	0	
## 2	0	0	0	
## 3	0	0	0	
## 4	0	0	0	
## 5	0	0	0	
## 6	0	0	0	
##	Trochochaeta.multisetosa	Tubifex.costatus	Tubificoides.benedeni	Turbellaria

```
## 1          0          0          0          0
## 2          0          0          0          0
## 3          0          0          0          0
## 4          1          0          0          0
## 5          0          0          0          0
## 6          1          0          0          0
## richness
## 1         13
## 2         14
## 3          4
## 4         15
## 5         18
## 6          7
```

### 3. Visualizing Temporal Patterns

#### 3.1 Time Series Analysis of Richness

We convert the data into a time series format for each site and plot the species richness over time.

```
## ===== Visualizing temporal patterns =====

# Convert the data to a time series format for each site
S0.ts = ts(richness_data[richness_data$site_id == "S0", ]$richness, start = 1984, end = 2004, frequency
S2.ts = ts(richness_data[richness_data$site_id == "S2", ]$richness, start = 1984, end = 2004, frequency
S3.ts = ts(richness_data[richness_data$site_id == "S3", ]$richness, start = 1984, end = 2004, frequency

# Plot time series
# Save the plots as separate files
pdf("plots/S0_Richness_Time_Series.pdf")
plot.ts(S0.ts, main = "Richness over Time at Site 0", ylab = "Richness", xlab = "Year")
plot.ts(S2.ts, main = "Richness over Time at Site 2", ylab = "Richness", xlab = "Year")
plot.ts(S3.ts, main = "Richness over Time at Site 3", ylab = "Richness", xlab = "Year")
dev.off()
```

```
## pdf
## 2
```

#### 3.2 Result:

The time series plots for sites S0, S2, and S3 show the richness at each site. There is an upward trend in all the sites, but all of the time series show high volatility.

### 4. RM-ANOVA (Repeated Measures ANOVA)

#### 4.1 Model Fitting

We fit a **linear mixed-effects model** (LME) to analyze the effect of `year` and `site_id` on species richness, with `site_id` as a random effect.



```
## ===== RM-ANOVA =====
# Fit a simpler model without the interaction term
rich.rm_simple = lme(richness ~ year + site_id,
                     random = ~ 1 | site_id,
                     data = richness_data)

# Check summary output of the model
summary(rich.rm_simple)
```

```
## Warning in pt(-abs(tVal), fDF): Production de NaN
```

```
## Linear mixed-effects model fit by REML
## Data: richness_data
##      AIC      BIC    logLik
## 343.2222 385.4442 -146.6111
##
## Random effects:
## Formula: ~1 | site_id
##      (Intercept) Residual
## StdDev:      3.873432 6.656361
##
## Fixed effects: richness ~ year + site_id
##              Value Std.Error DF   t-value p-value
## (Intercept) 10.66667  5.583825 40   1.910280 0.0633
## year1985     3.000000  5.434896 40   0.551988 0.5840
## year1986    -2.333333  5.434896 40  -0.429324 0.6700
## year1987     1.333333  5.434896 40   0.245328 0.8075
## year1988     0.333333  5.434896 40   0.061332 0.9514
## year1989    -0.666667  5.434896 40  -0.122664 0.9030
## year1990     1.666667  5.434896 40   0.306660 0.7607
## year1991     3.000000  5.434896 40   0.551988 0.5840
## year1992     0.000000  5.434896 40   0.000000 1.0000
## year1993     3.000000  5.434896 40   0.551988 0.5840
## year1994     6.666667  5.434896 40   1.226641 0.2271
## year1995     4.666667  5.434896 40   0.858649 0.3957
## year1996    19.333333  5.434896 40   3.557259 0.0010
## year1997    14.000000  5.434896 40   2.575946 0.0138
## year1998    17.333333  5.434896 40   3.189267 0.0028
## year1999    11.333333  5.434896 40   2.085290 0.0435
## year2000    15.000000  5.434896 40   2.759942 0.0087
## year2001    19.666667  5.434896 40   3.618591 0.0008
## year2002    11.000000  5.434896 40   2.023958 0.0497
## year2003    21.333333  5.434896 40   3.925251 0.0003
## year2004    19.333333  5.434896 40   3.557259 0.0010
## site_idS2     6.904762  5.850357  0   1.180229    NaN
## site_idS3    -7.904762  5.850357  0  -1.351159    NaN
## Correlation:
##      (Intr) yr1985 yr1986 yr1987 yr1988 yr1989 yr1990 yr1991 yr1992 yr1993
## year1985  -0.487
## year1986  -0.487  0.500
## year1987  -0.487  0.500  0.500
## year1988  -0.487  0.500  0.500  0.500
## year1989  -0.487  0.500  0.500  0.500  0.500
```

```

## year1990 -0.487 0.500 0.500 0.500 0.500 0.500
## year1991 -0.487 0.500 0.500 0.500 0.500 0.500 0.500
## year1992 -0.487 0.500 0.500 0.500 0.500 0.500 0.500 0.500
## year1993 -0.487 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500
## year1994 -0.487 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500
## year1995 -0.487 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500
## year1996 -0.487 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500
## year1997 -0.487 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500
## year1998 -0.487 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500
## year1999 -0.487 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500
## year2000 -0.487 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500
## year2001 -0.487 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500
## year2002 -0.487 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500
## year2003 -0.487 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500
## year2004 -0.487 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500
## site_idS2 -0.524 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
## site_idS3 -0.524 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
## yr1994 yr1995 yr1996 yr1997 yr1998 yr1999 yr2000 yr2001 yr2002 yr2003
## year1985
## year1986
## year1987
## year1988
## year1989
## year1990
## year1991
## year1992
## year1993
## year1994
## year1995 0.500
## year1996 0.500 0.500
## year1997 0.500 0.500 0.500
## year1998 0.500 0.500 0.500 0.500
## year1999 0.500 0.500 0.500 0.500 0.500
## year2000 0.500 0.500 0.500 0.500 0.500 0.500
## year2001 0.500 0.500 0.500 0.500 0.500 0.500 0.500
## year2002 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500
## year2003 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500
## year2004 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500 0.500
## site_idS2 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
## site_idS3 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
## yr2004 st_dS2
## year1985
## year1986
## year1987
## year1988
## year1989
## year1990
## year1991
## year1992
## year1993
## year1994
## year1995
## year1996
## year1997

```

```
## year1998
## year1999
## year2000
## year2001
## year2002
## year2003
## year2004
## site_idS2 0.000
## site_idS3 0.000 0.500
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -2.20340611 -0.55085153  0.03576958  0.57589023  1.86717206
##
## Number of Observations: 63
## Number of Groups: 3
```

**Random Effects** The random effects here shows how much variability there is across the levels of site\_id. In this case, each site has its own baseline richness level. The random intercept for site\_id has a standard deviation of 3.87, this means that there is variability in species richness across the different sites. The residual standard deviation is 6.66, this shows the variability in richness after accounting for the fixed effects (year and site\_id).

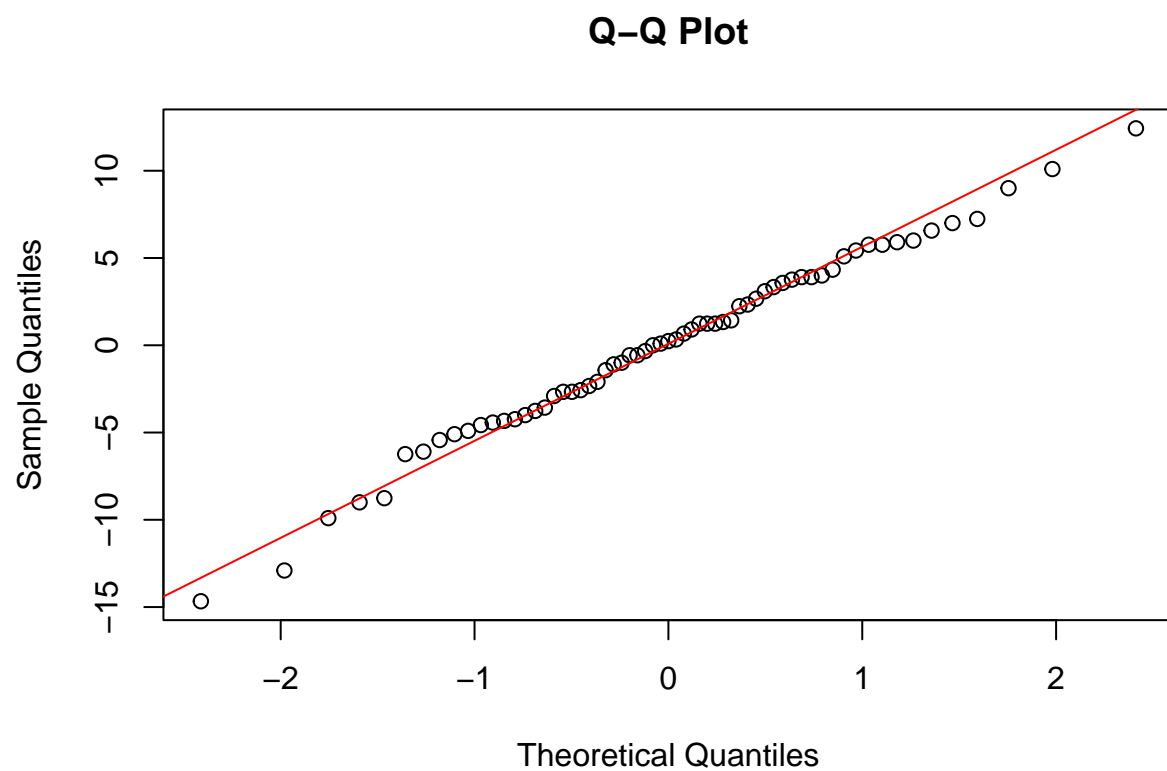
**Fixed Effects** The fixed effects shows how the predictors (in this case, year and site\_id) are associated with the outcome variable richness. It provides the estimated effects (coefficients), standard errors, t-values, and p-values for each of these predictors.

(Intercept): 10.67 (p-value = 0.0633) The intercept represents the expected species richness in the baseline year (1984) and at site S1. The p-value is just above 0.05, suggesting a marginally significant difference from 0, but not quite statistically significant. Year Effects: Coefficients for each year are provided. These represent the difference in species richness relative to the baseline year (1984). For example, year1985 has a coefficient of 3.00, but the p-value is 0.5840, meaning there is no significant difference in species richness between 1985 and 1984. year1996 has a coefficient of 19.33 (p-value = 0.0010), which is statistically significant. This suggests that the species richness in 1996 is significantly higher compared to 1984. year2001 also has a significant positive effect with a coefficient of 19.67 and p-value = 0.0008, indicating a significant increase in richness in this year. Other years (e.g., 1986, 1987, 1990) do not show significant differences from 1984, with p-values greater than 0.05.

Site Effects: Coefficients for the sites (site\_idS2, site\_idS3) show the difference in species richness between these sites and the reference site (S1). site\_id S2: Coefficient is 6.90 (p-value = NaN), indicating some positive difference in richness, though the p-value is not provided, suggesting no statistical test for this site's effect. site\_id S3: Coefficient is -7.90 (p-value = NaN), indicating a negative difference in richness for site S3, but again, the p-value is not available.

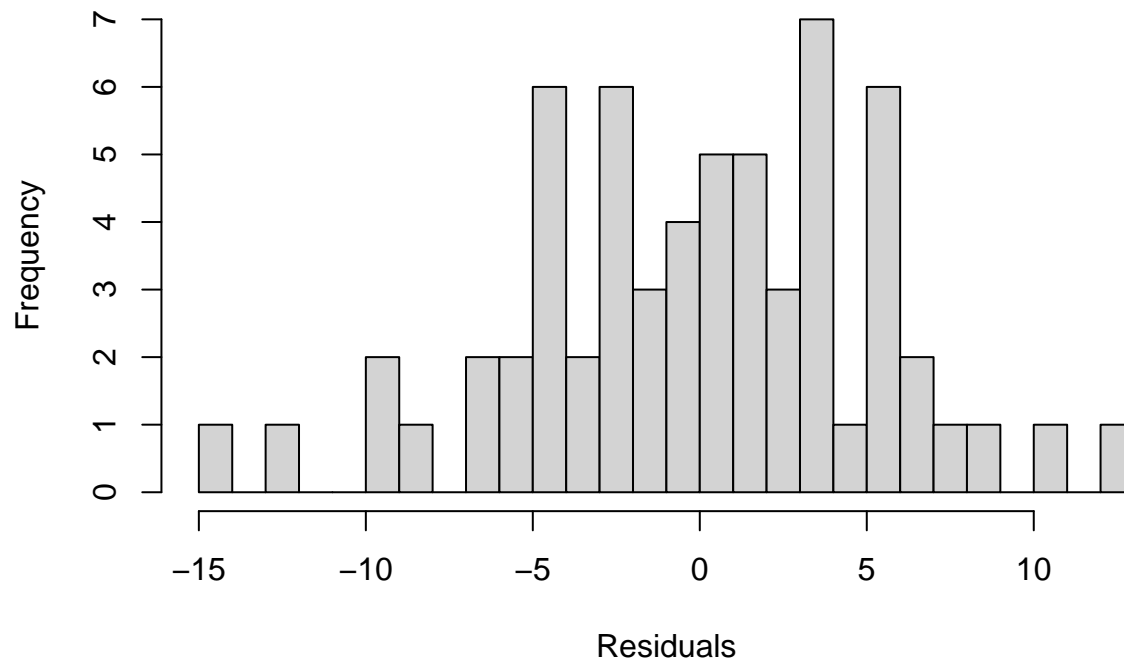
**Correlation of Fixed Effects** The correlation matrix of the fixed effects shows how correlated the predictors are with each other. The correlations here are all moderate (e.g., 0.500 for year-to-year comparisons), meaning there is some relationship between the years, but it is not overly high. This indicates that multicollinearity is not a major issue.

```
# Plots
# Q-Q plot for residuals
qqnorm(resid(rich.rm_simple), main = "Q-Q Plot")
qqline(resid(rich.rm_simple), col = "red")
```



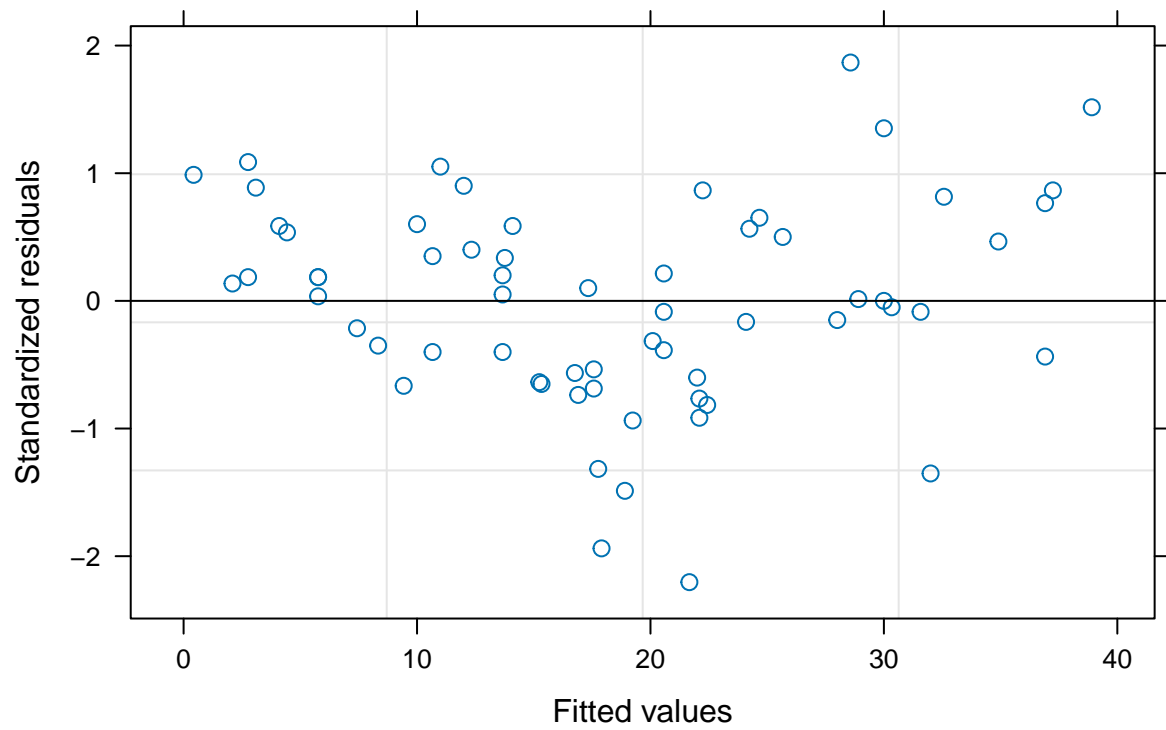
```
# Histogram of residuals  
hist(resid(rich.rm_simple), main = "Histogram of Residuals", xlab = "Residuals", breaks = 20)
```

## Histogram of Residuals

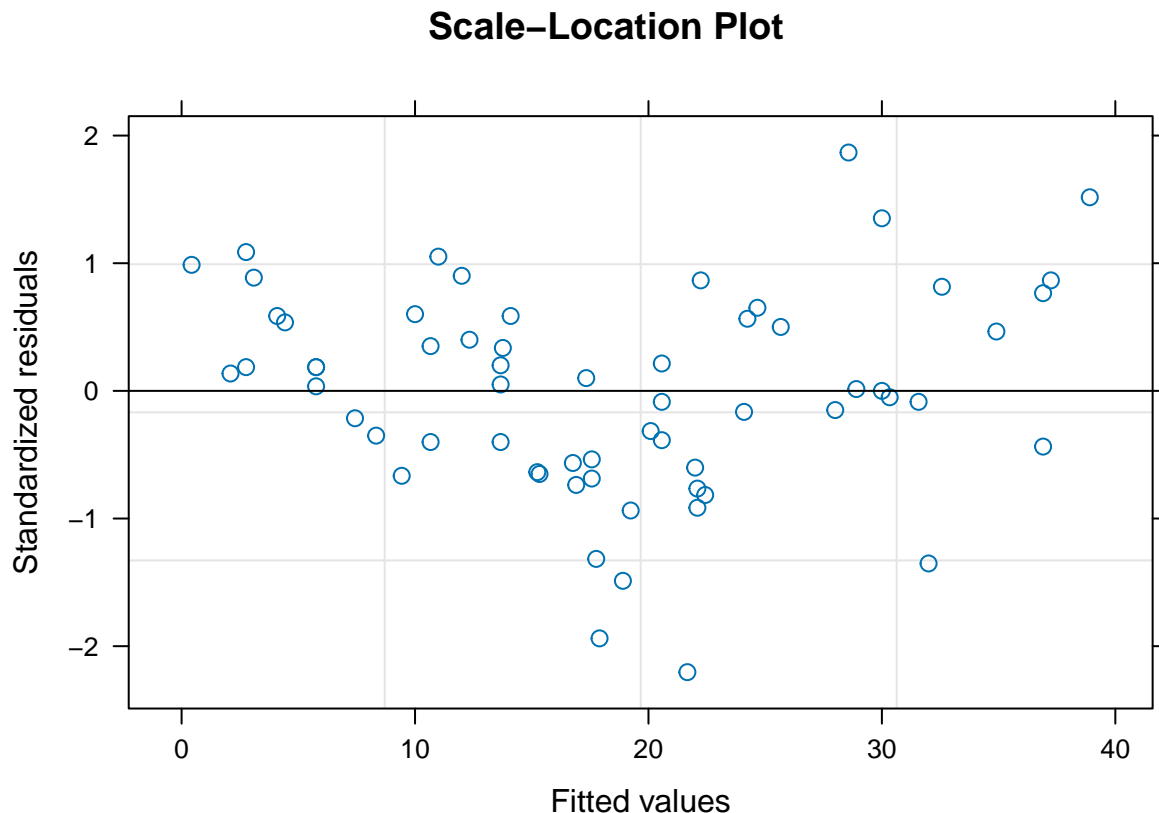


```
# Residuals vs Leverage plot  
plot(rich.rm_simple, which = 5, main = "Residuals vs Leverage")
```

## Residuals vs Leverage



```
# Scale-Location plot  
plot(rich.rm_simple, which = 3, main = "Scale-Location Plot")
```



### ### 4.2 ANOVA and Marginal Means

We then perform the **ANOVA** to assess the statistical significance of the fixed effects and calculate **LSMeans** to estimate marginal means.

```
# Perform the ANOVA
anova_results = anova(rich.rm_simple)
pander(anova_results)

# Calculate marginal means (LSMeans) for year and site_id
library(lsmmeans)
lsmeans_results = lsmeans(rich.rm_simple, ~ year | site_id)
summary(lsmeans_results)
```

**ANOVA Results** Intercept and year are both significant, meaning the model shows meaningful variation in species richness over time. `site_id` has a significant F-value, but the p-value could not be calculated due to a problem with the denominator degrees of freedom.

## 5. Beta Diversity Analysis

### 5.1 Turnover Metrics

We calculate **turnover metrics** (total turnover, appearance, disappearance) to examine how species composition changes across sites and years.

```
# Calculate turnover for the three sites (S0, S2, and S3)
species.abunds = group_by(data, year, site_id) %>% count(species)
```

```
# Calculate turnover metrics
```

```
total_turnover = turnover(df = species.abunds,
  time.var = "year",
  species.var = "species",
  abundance.var = "n",
  replicate.var = "site_id",
  metric = "total")
```

```
appearance = turnover(df = species.abunds,
  time.var = "year",
  species.var = "species",
  abundance.var = "n",
  replicate.var = "site_id",
  metric = "appearance")
```

```
disappearance = turnover(df = species.abunds,
  time.var = "year",
  species.var = "species",
  abundance.var = "n",
  replicate.var = "site_id",
  metric = "disappearance")
```

```
# Combine turnover metrics
```

```
turnover = full_join(total_turnover, disappearance) %>%
  full_join(appearance)
```

```
## Joining with 'by = join_by(year, site_id)'
## Joining with 'by = join_by(year, site_id)'
```

```
# Convert to long format
```

```
turnover = gather(turnover, key = metric, value = turnover,
  total, appearance, disappearance)
```

```
# Visualize turnover
```

```
pdf("plots/Turnover_Metrics.pdf")
ggplot(turnover, aes(x = year, y = turnover, color = metric)) +
  geom_line(size = 1) +
  facet_wrap(~ site_id, ncol = 1) +
  xlim(1984, 2004) +
  xlab("Year") +
  ylab("Turnover") +
  scale_color_grey()
```

```
## Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use 'linewidth' instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```



```
dev.off()
```

```
## pdf  
## 2
```

## 5.2 Result:

The turnover analysis helps us visualize changes in species composition over time. The high points in turnover indicates significant changes in community structure, while low turnover suggests more stability in species composition.

## 7. Community Stability (Coefficient of Variation)

### 7.1 Calculating Community Stability

We calculate the **Coefficient of Variation (CV)** to assess community stability across time.

```
# Community Stability (Coefficient of Variation)  
portal.stab = community_stability(df = species.abunds,  
                                time.var = "year",  
                                abundance.var = "n",  
                                replicate.var = "site_id")  
  
portal.stab
```

### 7.2 Result:

A higher CV means greater variability or fluctuations in species richness across time. This means that at Site S0, species richness fluctuates more over the years, and the community is less stable compared to the other sites. The CV for Site S2 is lower than for Site S0, indicating that the species richness at this site fluctuates less over time and the community is more stable compared to Site S0. Site S3 has the lowest CV, indicating the most stable species richness over time among the three sites.

---