EPFL

Multivariate statistics in R

Hannes PETER
Martin BOUTROUX

This lecture benefits from teaching experience at different universities by Alexandre Buttler, François Gillet and Daniel Borcard.

Aims of the course

- Overview of principles of multidimensional data analysis
- Choice of statistical tools
- Learn how to use these tools
- Interpretation of the results
- Efficient communication with other experts
- Use

Course structure

- Wednesday 9:15 – ~13:00
  - ~1h theoretical lecture
  - ~1h introduction to practical (commented R script)
    - Doubs river fish dataset
  - 1-2h hands-on practical work
    - oribatide mites datase
    - before you leave: briefly present your results to a teacher (~2-3 minutes)
- short paper discussions
  - read paper for discussion the following week
- group work

Tentative schedule

- 11.09. **session 1** Introduction to multivariate statistics using R
- 18.09. **session 2** Similarity and distance
- 25.09. **group work - define research topic**
- 02.10. **session 3** Cluster analysis
- 09.10. ***«modern R» with Martin*** (tidyverse)
- 16.10. **group work - cluster analysis**
- 23.10. *break*
- 30.10. **session 4** Ordination
- 06.11. **session 5** Constrained ordination
- 13.11. **mid-term exam**, **group work - ordination techniques**
- 20.11. **session 6** Variance partitioning
- 27.11. **session 7** Auxilliary multivariate tools
- 04.12./11.12. **group work**
- 18.12. hand in report
- 08.01. group presentations 1
- 15.01. group presentations 2

EPFL

Evaluation

- **mid-term exam** (40%) (individual)
  - ca. 5 multiple-choice questions
  - 1-2 questions to develop
- **written report** (30%) (group)
  - introduction, background, hypotheses
  - data source, types
  - results, figures, tables
  - discussion, interpretation, conclusions
  - 8 pages <u>maximum</u>, 3 different multivariate analyses, ~5 Figures, deadline 18.12.2024
- **oral presentation** (30%) (group)
  - in January 2025, 10 + 5 minutes

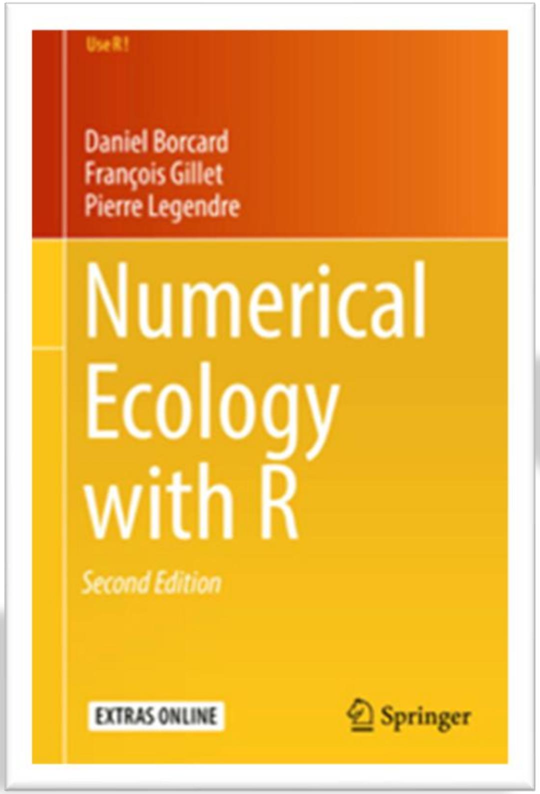
Group projects

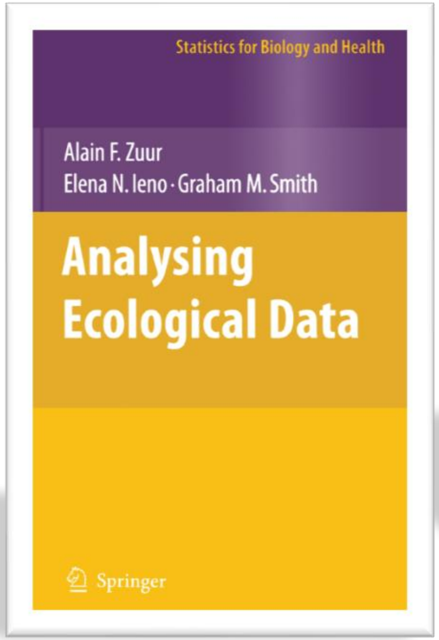
- 7 groups of 5 students
- define a research question
- collect/find data
- present research idea (1 page)
- perform multivariate analyses to address research question
- write report
- present results to the class

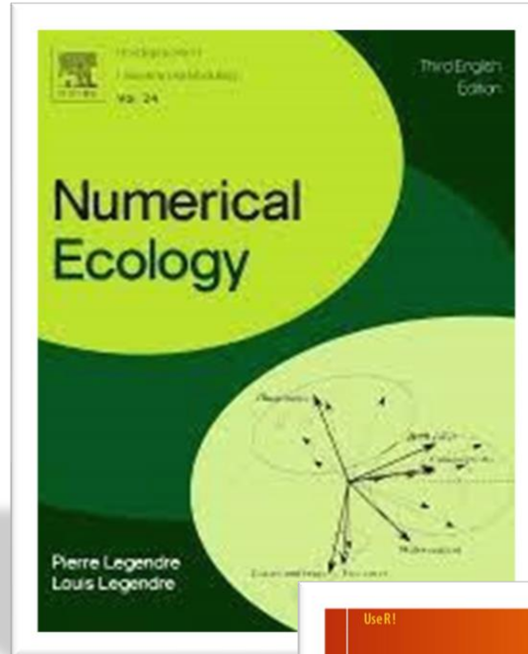
> Check moodle for data sources

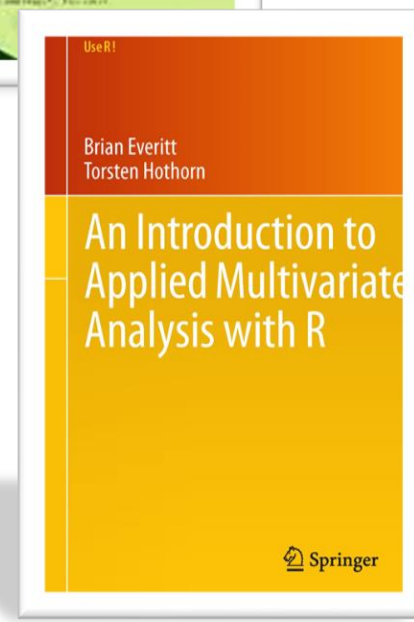
EPFL

material









http://www.numericalecology.com/numecolR/index.html

available as pdf

Online resources

- https://ordination.okstate.edu/
- https://www.davidzeleny.net/anadat-r/doku.php/en:start
- https://environmentalcomputing.net/statistics/mvabund/
- https://sites.google.com/site/mb3gustame/

Multivariate statistics (in Ecology)

- « *Domain of quantitative ecology dealing with the numerical analysis of complex data* » (Legendre & Legendre 1998)

- Origin in the ecology of biological communities (synecology, community ecology)

- Original numerical methods, often developed by ecologists (e.g. diversity/similarity measures)

Why is it important?

You will try to answer difficult questions about the complex world we live in.

Which test should I apply?

There is not a single «test», but rather a series of analyses - data exploration, model formulation, evaluation and interpretation is required.

EPFL

Practical Example

Indicator Species Analysis (ISA)

Indicator Species

Monitoring the occurrence (or abundance) of indicator species is often used in long-term environmental monitoring for conservation or ecological management.

Indicator species:

1. reflect the environment (abiotic or biotic)
2. respond to environmental change
3. representative of other species (i.e. can be used to predict the diversity of other species, taxa or communities)

EPFL

Lawton & Gaston, Encyclopedia of Biodiversity, 2001

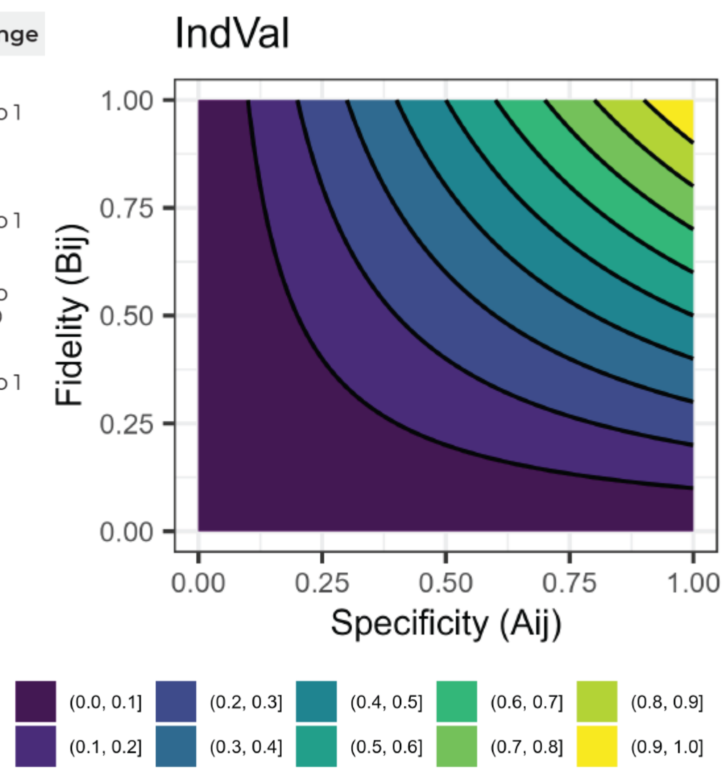
Indicator Species Analysis (ISA)

ISA requires classification of sample units into groups.

ISA involves calculating the **specificity** (*A*ij; relative abundance) and **fidelity** (*B*ij; relative frequency) of species *i* in group *j*.
These values are then multiplied to yield the test statistic, the **Indicator Value (*IV*ij)**.

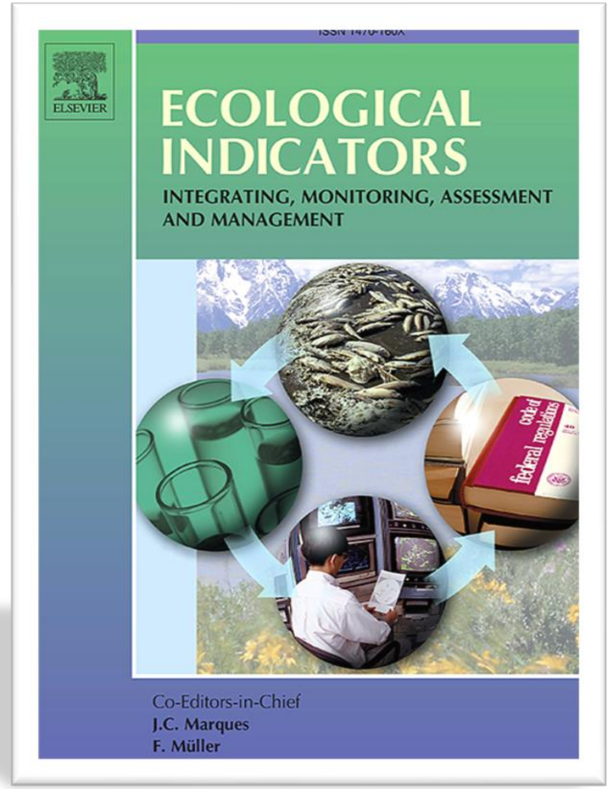
	Formula	Verbal Interpretation	Range
Specificity / Relative Abundance:	$A_{ij} = \frac{\bar{x}_{ij}}{\sum_j \bar{x}_{i\bullet}}$	Mean cover of species *i* in group *j* as a proportion of its mean cover in all groups	0 to 1
Fidelity / Relative Frequency:	$B_{ij} = \frac{n_{ij}}{n_{\bullet j}}$	Proportion of plots in group *j* on which species *i* occurs	0 to 1
Indicator Value:	$IV_{ij} = A_{ij} \times B_{ij} \times 100$	As proposed by Dufrêne & Legendre (1997)	0 to 100
	$IV_{ij} = \sqrt{A_{ij} \times B_{ij}}$	As reported in `indicspecies::multipatt()`	0 to 1

- $\bar{x}_{ij}$ is the mean cover of species *i* within group *j*
- $\sum_j \bar{x}_{i\bullet}$ is the sum of the mean cover of species *i* in all groups
- n_{ij} is the number of plots in group *j* occupied by species *i*
- $n_{\bullet j}$ is the total number of plots in group *j*.



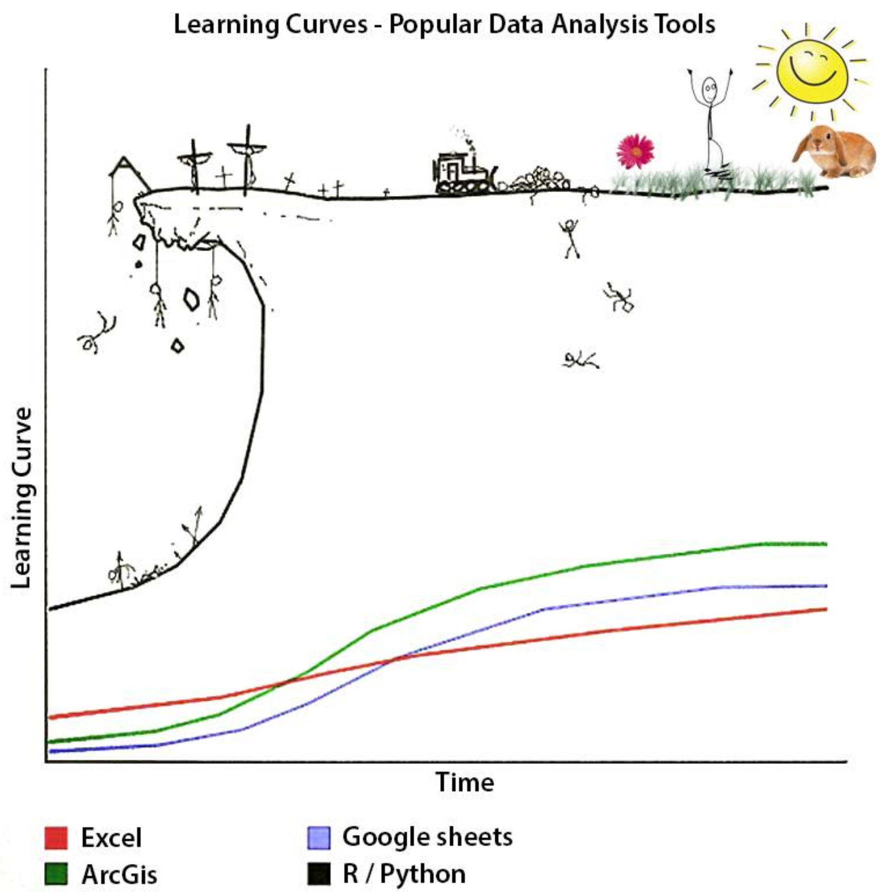
Heft 85, 2019
WSL Berichte
ISSN 2296-3456
Zustand und Entwicklung der Biotope von nationaler Bedeutung: Resultate 2011–2017 der Wirkungskontrolle Biotopschutz Schweiz
Ariel Bergamini, Christian Ginzler, Benedikt R. Sch
Angéline Bedolla, Steffen Boch, Klaus Ecker, Ulric
Helen Küchler, Meinrad Küchler, Oliver Dosch,
Rolf Holderegger
Eidg. Forschungsanstalt für Wald, Schnee und
CH-8903 Birmensdorf
info fauna karch, Bellevaux 51, 2000 Neuchâte

2023 | Umwelt-Zustand
Biodiversität in der Schweiz
Zustand und Entwicklung
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra
Bundesamt für Umwelt BAFU



ELSEVIER
ECOLOGICAL INDICATORS
INTEGRATING, MONITORING, ASSESSMENT AND MANAGEMENT
Co-Editors-in-Chief
J.C. Marques
F. Müller

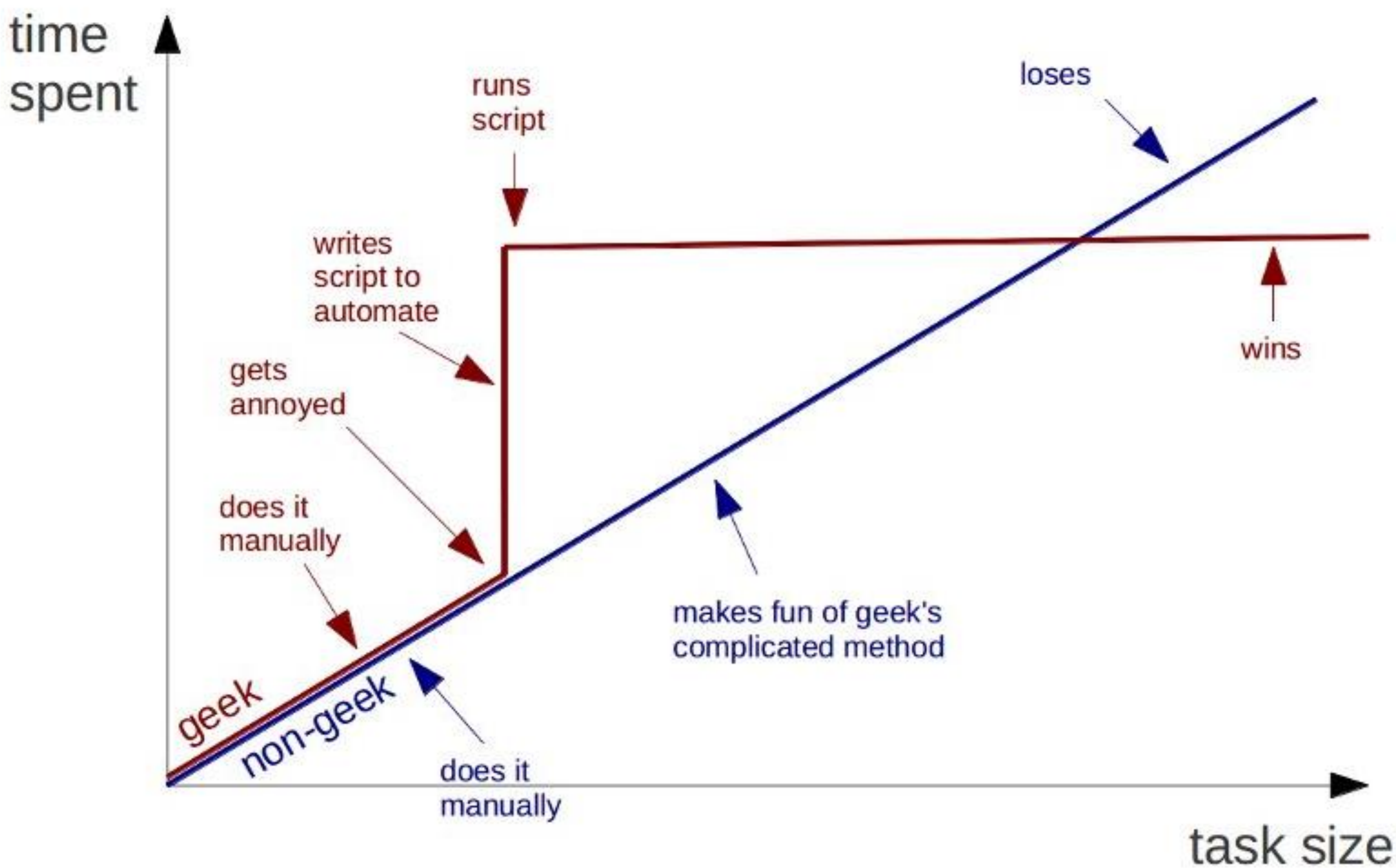
Why R?



https://nceas.github.io/oss-lessons/

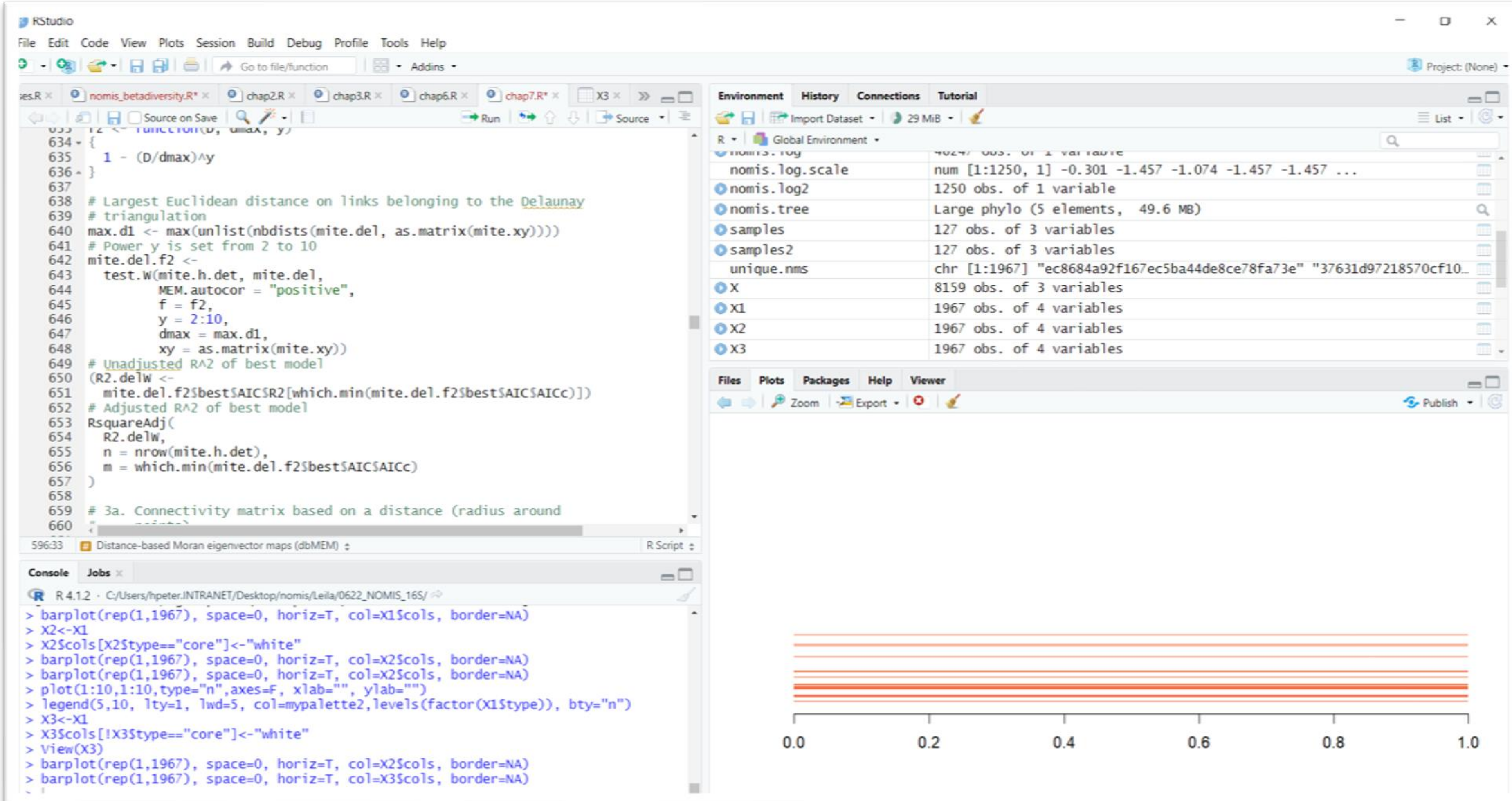
- freely available
- constant development
- professional visualization
- documentation and help
- large community (support)
- reproducability (code sharing)

Geeks and repetitive tasks



https://nceas.github.io/oss-lessons/

RStudio


RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Project: (None)
nomis_betadiversity.R*
chap2.R
chap3.R
chap6.R
chap7.R*
X3
Source on Save
Run
Source
{
  1 - (D/dmax)^y
}

# Largest Euclidean distance on links belonging to the Delaunay
# triangulation
max.d1 <- max(unlist(nbdists(mite.del, as.matrix(mite.xy))))
# Power y is set from 2 to 10
mite.del.f2 <-
  test.W(mite.h.det, mite.del,
         MEM.autocor = "positive",
         f = f2,
         y = 2:10,
         dmax = max.d1,
         xy = as.matrix(mite.xy))
# Unadjusted R^2 of best model
(R2.delW <-
  mite.del.f2$best$AIC$R2[which.min(mite.del.f2$bestAICAICc)])
Adjusted R^2 of best model
RsquareAdj(
  R2.delW,
  n = nrow(mite.h.det),
  m = which.min(mite.del.f2$best$AIC$AICc)
)

3a. Connectivity matrix based on a distance (radius around
596:33
Distance-based Moran eigenvector maps (dbMEM)
R Script
Environment History Connections Tutorial
Import Dataset
29 MiB
List
R
Global Environment
nomis.log.scale num [1:1250, 1] -0.301 -1.457 -1.074 -1.457 -1.457 ...
nomis.log2 1250 obs. of 1 variable
nomis.tree Large phylo (5 elements, 49.6 MB)
samples 127 obs. of 3 variables
samples2 127 obs. of 3 variables
unique.nms chr [1:1967] "ec8684a92f167ec5ba44de8ce78fa73e" "37631d97218570cf10...
X 8159 obs. of 3 variables
X1 1967 obs. of 4 variables
X2 1967 obs. of 4 variables
X3 1967 obs. of 4 variables
Files Plots Packages Help Viewer
Zoom
Export
Publish
Console Jobs
R 4.1.2 · C:/Users/hpeter.INTRANET/Desktop/nomis/Leila/0622_NOMIS_16S/
> barplot(rep(1,1967), space=0, horiz=T, col=X1$cols, border=NA)
> X2<-X1
> X2$cols[X2$type=="core"]<-"white"
> barplot(rep(1,1967), space=0, horiz=T, col=X2$cols, border=NA)
> barplot(rep(1,1967), space=0, horiz=T, col=X2$cols, border=NA)
> plot(1:10,1:10,type="n",axes=F, xlab="", ylab="")
> legend(5,10, lty=1, lwd=5, col=mypalette2,levels(factor(X1$type)), bty="n")
> X3<-X1
> X3$cols[!X3$type=="core"]<-"white"
> View(X3)
> barplot(rep(1,1967), space=0, horiz=T, col=X2$cols, border=NA)
> barplot(rep(1,1967), space=0, horiz=T, col=X3$cols, border=NA)
0.0
0.2
0.4
0.6
0.8
1.0

Datasets used in this class

built-in

iris - Edgar Anderson's Iris Data

mtcars - Motor Trend Car Road Tests

doubs – Verneaux doubs fish

dune - Vegetation and Environment in Dutch Dune Meadows.

mite - Oribatid Mite Data with Explanatory Variables

BCI - Barro Colorado Island Tree Counts

varechem/varespec - Vegetation and environment in lichen pastures

own (group) datasets

R libraries used in this class

- vegan
- ade4
- adespatial

…

CRAN –The Comprehensive R Archive Network

https://cran.r-project.org/

vegan: Community Ecology Package

Ordination methods, diversity analysis and other functions for community and vegetation ecologists.

Version:	2.6-2
Depends:	permute (≥ 0.9-0), lattice, R (≥ 3.4.0)
Imports:	MASS, cluster, mgcv
Suggests:	parallel, tcltk, knitr, markdown
Published:	2022-04-17
Author:	Jari Oksanen [aut, cre], Gavin L. Simpson [aut], F. Guillaume Blanchet [aut], Roeland Kindt [aut], Pierre Legendre [aut], Peter R. Minchin [aut], R.B. O'Hara [aut], Peter Solymos [aut], M. Henry H. Stevens [aut], Eduard Szoecs [aut], Helene Wagner [aut], Matt Barbour [aut], Michael Bedward [aut], Ben Bolker [aut], Daniel Borcard [aut], Gustavo Carvalho [aut], Michael Chirico [aut], Miquel De Caceres [aut], Sebastien Durand [aut], Heloisa Beatriz Antoniazi Evangelista [aut], Rich FitzJohn [aut], Michael Friendly [aut], Brendan Furneaux [aut], Geoffrey Hannigan [aut], Mark O. Hill [aut], Leo Lahti [aut], Dan McGlinn [aut], Marie-Helene Ouellette [aut], Eduardo Ribeiro Cunha [aut], Tyler Smith [aut], Adrian Stier [aut], Cajo J.F. Ter Braak [aut], James Weedon [aut]
Maintainer:	Jari Oksanen <jhoksane at gmail.com>
BugReports:	https://github.com/vegandevs/vegan/issues
License:	GPL-2
URL:	https://github.com/vegandevs/vegan
NeedsCompilation:	yes
Materials:	NEWS
In views:	Environmetrics, Psychometrics, Spatial
CRAN checks:	vegan results

Documentation:

Reference manual:	vegan.pdf
Vignettes:	Design decisions and implementation Diversity analysis in vegan Introduction to ordination in vegan Partition of Variation vegan FAQ

Task views

reference manual

vignettes!

CRAN Task Views

CRAN Task View: Analysis of Ecological and Environmental Data

Maintainer:	Gavin Simpson
Contact:	ucfagls at gmail.com
Version:	2022-03-10
URL:	https://CRAN.R-project.org/view=Environmetrics
Source:	https://github.com/cran-task-views/Environmetrics/
Contributions:	Suggestions and improvements for this task view are very welcome and can be made through issues or pull requests on GitHub or via e-mail to the maintainer address. For further details see the Contributing guide.
Citation:	Gavin Simpson (2022). CRAN Task View: Analysis of Ecological and Environmental Data. Version 2022-03-10. URL https://CRAN.R-project.org/view=Environmetrics.
Installation:	The packages from this task view can be installed automatically using the ctv package. For example, `ctv::install.views("Environmetrics", coreOnly = TRUE)` installs all the core packages or `ctv::update.views("Environmetrics")` installs all packages that are not yet installed and up-to-date. See the CRAN Task View Initiative for more details.

Introduction

This Task View contains information about using R to analyse ecological and environmental data.

The base version of R ships with a wide range of functions for use within the field of environmetrics. This functionality is complemented by a plethora of packages available via CRAN, which provide specialist methods such as ordination & cluster analysis techniques. A brief overview of the available packages is provided in this Task View, grouped by topic or type of analysis. As a testament to the popularity of R for the analysis of environmental and ecological data, a special volume of the *Journal of Statistical Software* was produced in 2007.

Those interested in environmetrics should consult the Spatial view. Complementary information is also available in the Cluster, and SpatioTemporal task views.

If you have any comments or suggestions for additions or improvements, then please contact the maintainer or submit an issue or pull request in the GitHub repository linked above.

A list of available packages and functions is presented below, grouped by analysis type.

General packages

These packages are general, having wide applicability to the environmetrics field.

- Package EnvStats is the successor to the S-PLUS module *EnvironmentalStats*, both by Steven Millard. A user guide in the form of a book has recently been released.

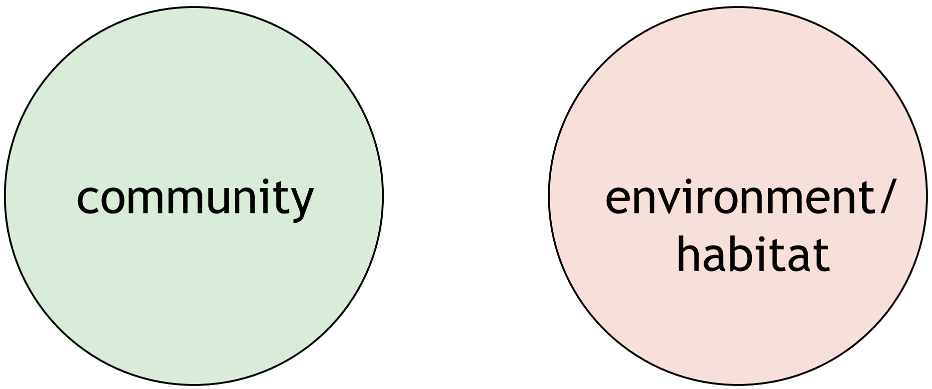
Modelling species responses and other data

Analysing species response curves or modelling other data often involves the fitting of standard statistical models to ecological data and includes simple (multiple) regression, Generalized Linear Models (GLM), extended regression (e.g. Generalized Least Squares [GLS]), Generalized Additive Models (GAM), and mixed effects models, amongst others.

- The base installation of R provides `lm()` and `glm()` for fitting linear and generalized linear models, respectively.

https://cran.r-project.org/web/views/

Possible approaches to problems in Ecology



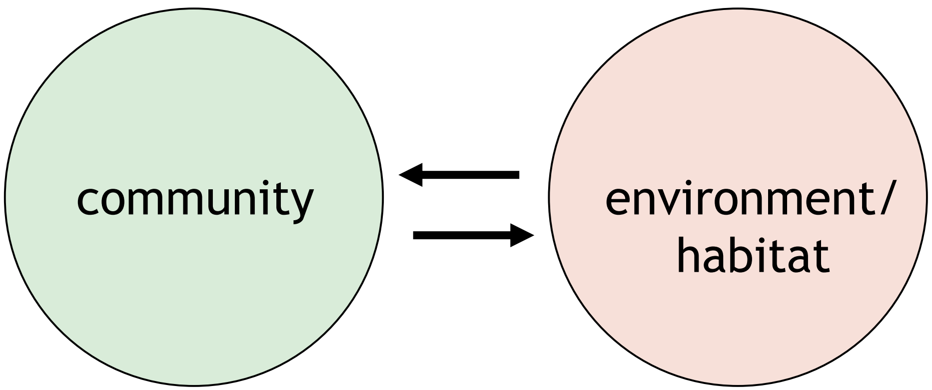
1. Independent description of community and environment

- Community structure, species composition, diversity, abundance distribution
- Spatial structure
- Temporal variation

⇒ Observational and inductive approach, no hypotheses

⇒ Qualitative or intuitive links e.g. bioindicator

Possible approaches to problems in Ecology



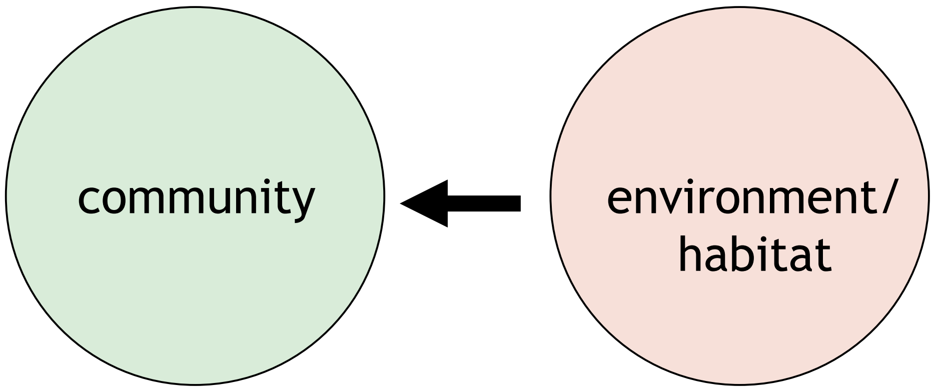
2. Simultaneous description of community and environment

- Correlation between biological and environmental data
- Quantitative empirical links

⇒ No explanatory model, no hypotheses

⇒ Observational approach

Possible approaches to problems in Ecology



« Response variables » « Explanatory variables »

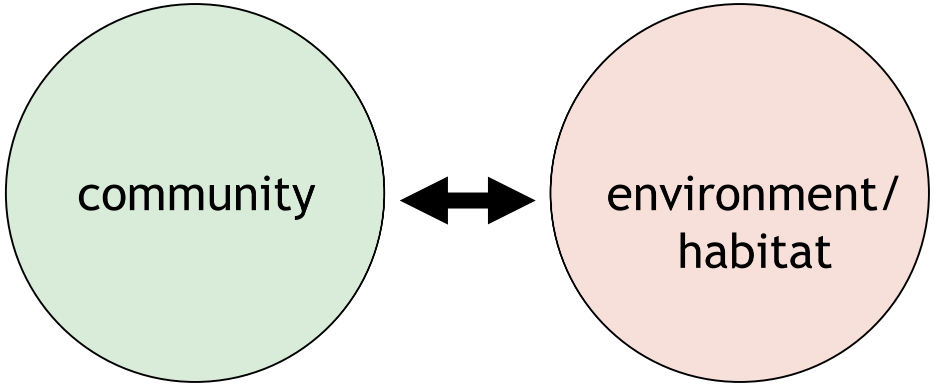
3. Unilateral influence of the habitat (explanatory variables) on the community (biological response)

- Regression: biological *versus* environmental data
- Empirical explanatory model
  - Observational (implicit causality), static
  - Experimental, manipulative (treatments), predictive, (explicit causality), kinetic

⇒ Deductive approach, hypotheses

EPFL

Possible approaches to problems in Ecology



4. Reciprocal interactions between habitat and community

- Systemic approach
- Predictive, theoretical model (explicit causality with feedback effects), dynamic

⇒ Ecological simulations

Multidimensional ecological/environmental data

- **Univariate analysis -** one variable
- **Bivariate analysis -** two variables
- **Multivariate analysis** - more than two variables

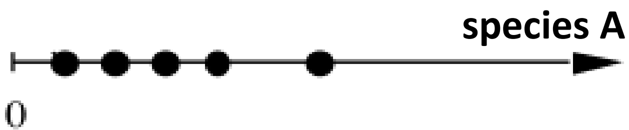
  Every object (sample) is characterized by several descriptors

  Direct graphical representation is impossible beyond 3 dimensions

Unidimensional data

Objects: samples
Descriptors

	1	2	3	4	5	6
species A	1	4	2	1	6	3



Multidimensional data

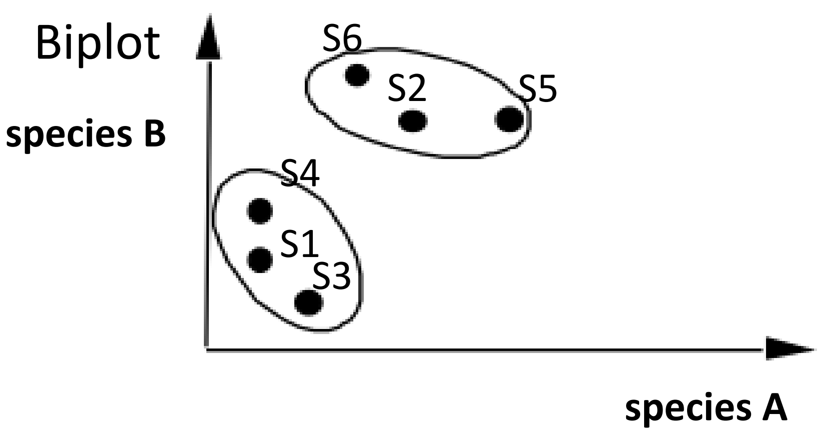
samples

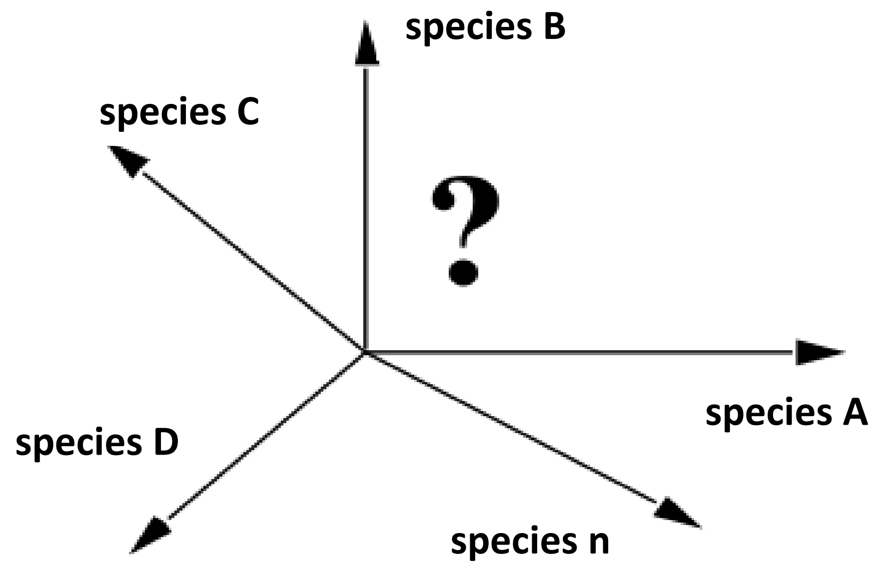
	1	2	3	4	5	6
species A	1	4	2	1	6	3
species B	2	5	1	3	5	6
species C	1	4	3	1	2	2
species D	3	1	6	5	6	2
:						
species n	1	6	3	2	2	4

Bidimensional data

samples

	1	2	3	4	5	6
species A	1	4	2	1	6	3
species B	2	5	1	3	5	6





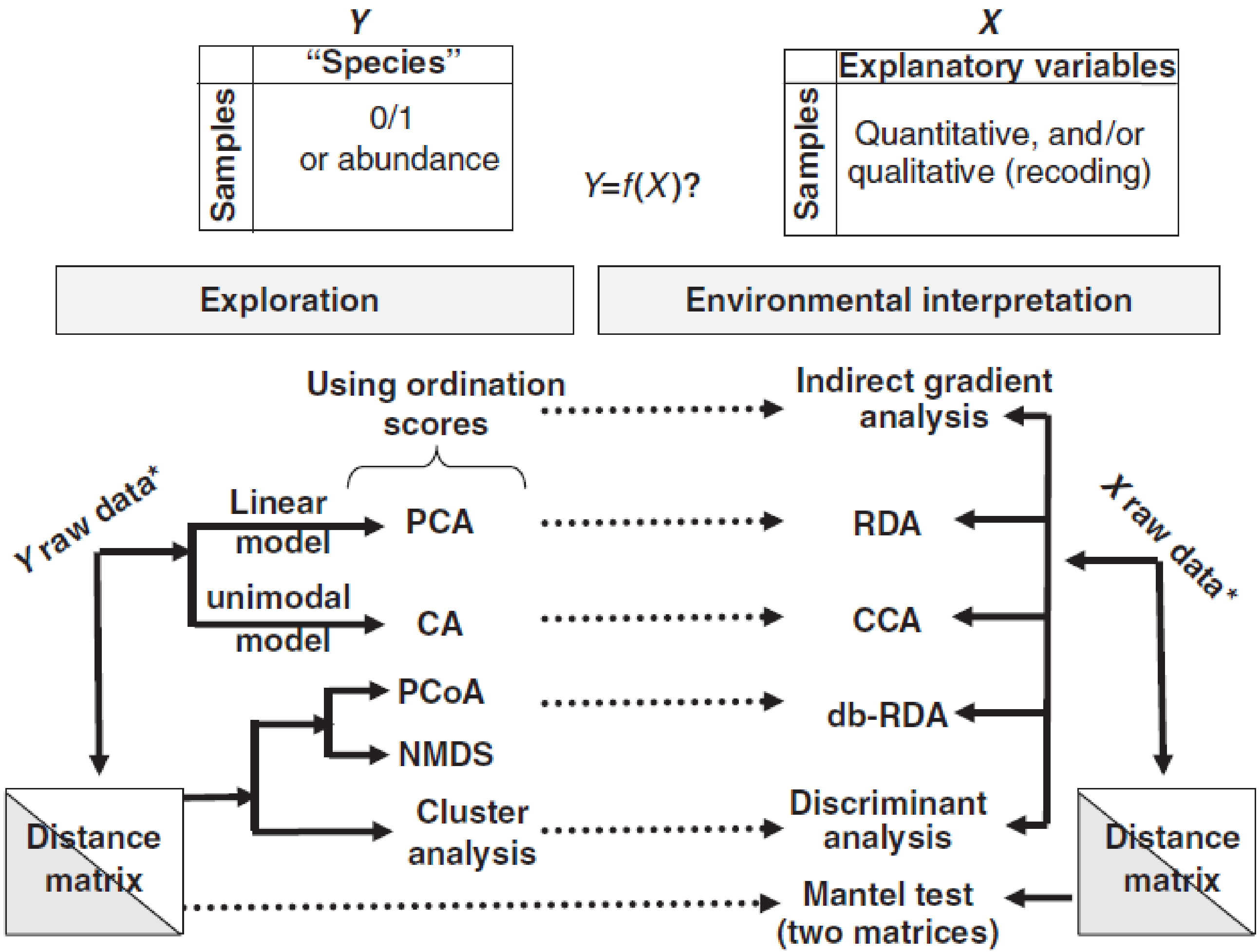
EPFL

Types of scientific tasks

most suited to the application of multivariate methods

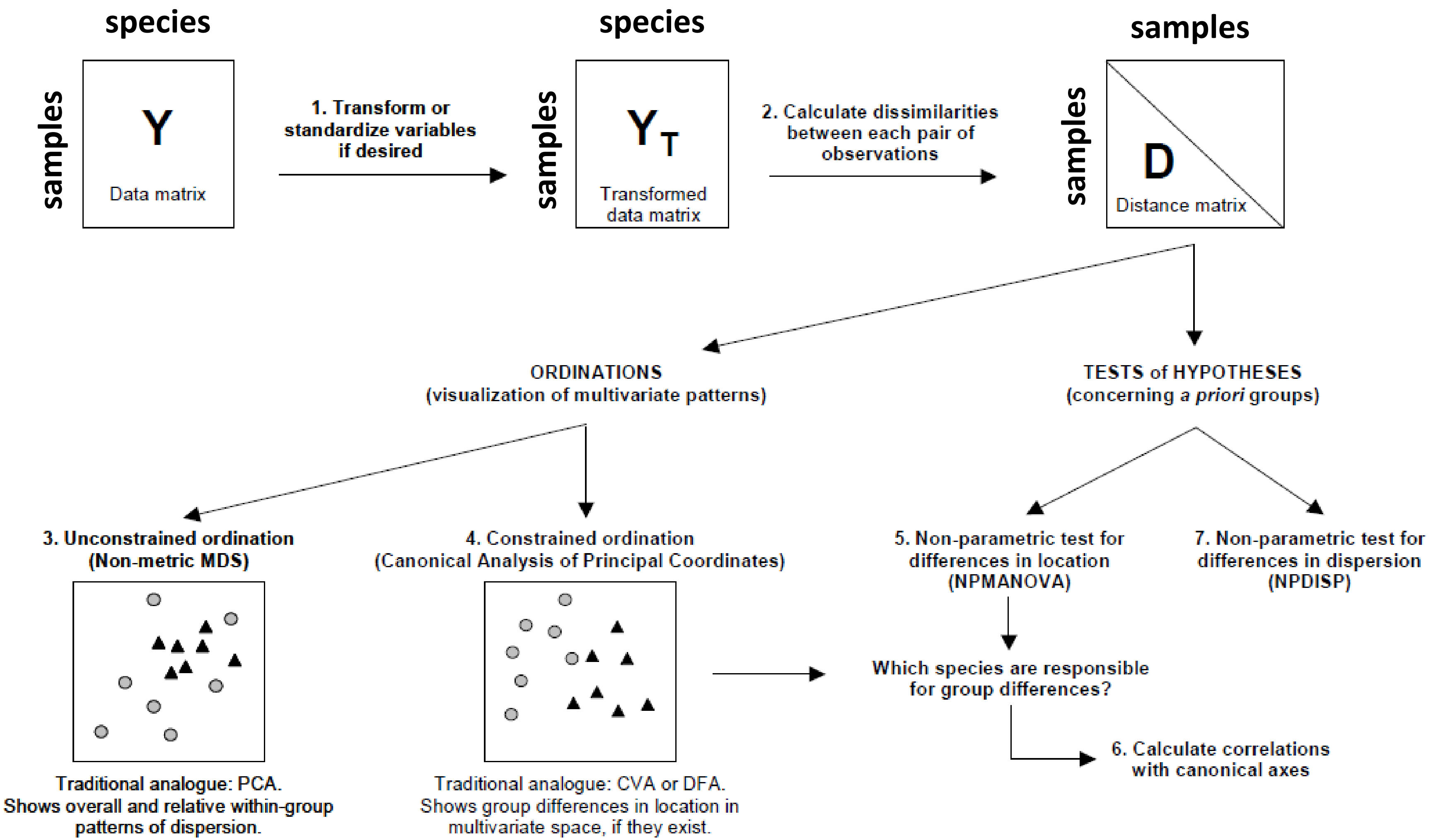
- **Data reduction and simplification**
  - the summary of multiple variables via a small set of (synthetic) variables. High-dimensional patterns are presented in a lower-dimensional space, aiding interpretation.
  - *Principal Component Analysis*
- **Sorting and grouping**
  - Tasks concerned with the similarity of samples and their assignment to groups.
  - *Cluster analysis*
- **Investigation of dependence among variables**
  - Methods that detect dependence among variables are valuable in detecting influence or covariation.
  - *Redundancy Analysis*
- **Hypothesis construction and testing**
  - Exploratory techniques can reveal patterns in data from which hypotheses may be constructed.
  - *Mantel test, PERMANOVA*

Accurate choice of methods…



Ramette 2007

Accurate choice of methods…



EPFL

Objects and descriptors

- Objects = observations
  - samples, field surveys, sites, experimental units

- Descriptors = measured variables
  - **Biological variables**
    - species, presence/absence, abundance (or their attributes)
    - traits
    - evolutionary/phylogenetic relationships
    - activity measurements
  - **Environmental variables**
    - temperature, pH, soil type, nutrients…
  - **Spatial variables**
    - Geographical coordinates (x, y), island size, latitude, …

Descriptor types

- Binary (boolean, qualitative with two modalities)
  - *Ex. : presence (1) or absence (0) of a species, terrestrial vs aquatic*
- Multiple
  - Unsorted (nominal, qualitative multiclass, categorial)
    - *Ex. : soil type, group affiliation (for instance following cluster analysis)*
  - Sorted
    - Semi-quantitative (ordinal, rank)
      - *Ex. : weak – medium – strong (coded 1 2 3)*
      - *Ex. : dominance code of a species (r + 1 2 3 4 5)*
    - Quantitative (cardinal)
      - Discreet (integer)
        - *Ex. : number of individuals of a species (abundance s.s.)*
      - Continuous (numerical)
        - *Ex. : biomass, altitude, activity rate measurement*
- Synthetic (complex)
  - *Ex. : relative abundance of a species*
  - *Ex. : C/N ratios of organic matter*

Doubs

abundance of 27 fish species along 30 sites in the river Doubs

	Cogo	Satr	Phph	Neba	Thth	Teso	Chna	Chto	Lele	Lece	Baba	Spbi	Gogo	Eslu	Pefl
1	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	5	4	3	0	0	0	0	0	0	0	0	0	0	0
3	0	5	5	5	0	0	0	0	0	0	0	0	0	1	0
4	0	4	5	5	0	0	0	0	0	1	0	0	1	2	2
5	0	2	3	2	0	0	0	0	5	2	0	0	2	4	4
6	0	3	4	5	0	0	0	0	1	2	0	0	1	1	1
7	0	5	4	5	0	0	0	0	1	1	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	1	3	0	0	0	0	0	5	0	0	0	0	0
10	0	1	4	4	0	0	0	0	2	2	0	0	1	0	0
11	1	3	4	1	1	0	0	0	0	1	0	0	0	0	0
12	2	5	4	4	2	0	0	0	0	1	0	0	0	0	0
13	2	5	5	2	3	2	0	0	0	0	0	0	0	0	0
14	3	5	5	4	4	3	0	0	0	1	1	0	1	1	0
15	3	4	4	5	2	4	0	0	3	3	2	0	2	0	0
16	2	3	3	5	0	5	0	4	5	2	2	1	2	1	1
17	1	2	4	4	1	2	1	4	3	2	3	4	1	1	2
18	1	1	3	3	1	1	1	3	2	3	3	3	2	1	3
19	0	0	3	5	0	1	2	3	2	1	2	2	4	1	1
20	0	0	1	2	0	0	2	2	2	3	4	3	4	2	2
21	0	0	1	1	0	0	2	2	2	2	4	2	5	3	3
22	0	0	0	1	0	0	3	2	3	4	5	1	5	3	4

environmental parameter

Variable	Code	Units
Distance from the source	dfs	km
Elevation	ele	m a.s.l.
Slope	slo	‰
Mean minimum discharge	dis	$m^3 \cdot s^{-1}$
pH of water	pH	–
Hardness (Ca concentration)	har	$mg \cdot L^{-1}$
Phosphate concentration	pho	$mg \cdot L^{-1}$
Nitrate concentration	nit	$mg \cdot L^{-1}$
Ammonium concentration	amm	$mg \cdot L^{-1}$
Dissolved oxygen	oxy	$mg \cdot L^{-1}$
Biological oxygen demand	bod	$mg \cdot L^{-1}$

+ spatial coordinates

traits

	sort	LatinName	Family	EnglishName	FrenchName	BodyLength	BodyLengthMax	ShapeFactor	TrophicLevel	omni
Cogo	1	Cottus gobio	Cottidae	Bullhead	Chabot commun	80	120	3.13	3.10	
Satr	2	Salmo trutta fario	Salmonidae	Brown trout	Truite fario	280	800	5.03	4.00	
Phph	3	Phoxinus phoxinus	Cyprinidae	Eurasian minnow	Vairon	60	100	2.95	3.10	
Babl	4	Barbatula barbatula	Nemacheilidae	Stone loach	Loche franche	80	130	4.50	3.10	
Thth	5	Thymallus thymallus	Salmonidae	Grayling	Ombre commun	300	450	4.33	3.10	
eso	6	Telestes souffia	Cyprinidae	Vairone	Blageon	90	180	5.26	3.40	
na	7	Chondrostoma nasus	Cyprinidae	Common nase	Hotu	280	480	6.50	2.00	
Pato	8	Parachondrostoma toxostoma	Cyprinidae	South-west European nase	Toxostome	160	250	5.00	2.00	
Lele	9	Leuciscus leuciscus	Cyprinidae	Common dace	Vandoise	180	360	2.81	2.57	
Sqce	10	Squalius cephalus	Cyprinidae	European chub	Chevaine	240	500	3.97	3.50	

R> library(vegan)

Data(

Verneaux J. 1973. - Cours d'eau de Franche-Comté (Massif du Jura).
Recherches écologiques sur le réseau hydrographique du Doubs. Essai de biotypologie.

Oribatide mites

abundance of 35 species of mites in mosses collected in 70 sites in Montreal, Canada

	Brachy	PHTH	HPAV	RARD	SSTR	Protopl	MEGR	MPRO	TVIE	HMIN	HMIN2	NPRA	TVEL	ONOV	SUCT	LCIL	Oribatl1	Cera
1	17	5	5	3	2	1	4	2	2	1	4	1	17	4	9	50	3	
2	2	7	16	0	6	0	4	2	0	0	1	3	21	27	12	138	6	
3	4	3	1	1	2	0	3	0	0	0	6	3	20	17	10	89	3	
4	23	7	10	2	2	0	4	0	1	2	10	0	18	47	17	108	10	
5	5	8	13	9	0	13	0	0	0	3	14	3	32	43	27	5	1	
6	19	7	5	9	3	2	3	0	0	20	16	2	13	38	39	3	5	
7	17	3	8	2	3	0	3	0	0	19	3	0	22	27	37	0	4	
8	5	4	8	2	1	2	3	0	0	1	4	10	12	25	26	0	6	
9	3	3	2	2	1	1	12	0	0	0	3	9	20	8	19	30	7	
10	22	4	5	3	0	0	0	0	0	11	4	5	16	33	29	0	8	
11	36	7	35	9	0	2	5	0	0	0	6	1	42	17	15	0	2	
12	28	2	12	13	0	0	1	0	0	11	1	2	27	34	59	0	3	
13	3	2	4	12	0	0	2	0	0	1	1	2	11	57	63	0	10	
14	41	5	12	0	2	0	2	0	0	24	20	3	19	61	20	0	17	
15	6	0	6	0	0	0	1	0	0	0	5	1	21	7	7	29	8	
16	7	2	3	2	0	0	3	0	0	1	0	6	19	8	30	1	3	
17	9	0	1	2	0	0	2	0	0	2	1	3	24	9	31	0	0	
18	19	3	7	0	0	0	2	0	0	10	4	0	9	32	16	3	1	
19	12	2	10	1	0	0	4	1	0	16	5	2	5	14	18	7	3	
20	3	1	7	1	0	0	17	0	0	13	7	7	25	51	12	0	14	
21	5	2	8	2	0	0	3	0	0	24	0	1	11	20	31	0	1	

environmental parameter

	SubsDens	WatrCont	Substrate	Shrub	Topo
1	39.18	350.15	Sphagn1	Few	Hummock
2	54.99	434.81	Litter	Few	Hummock
3	46.07	371.72	Interface	Few	Hummock
4	48.19	360.50	Sphagn1	Few	Hummock
5	23.55	204.13	Sphagn1	Few	Hummock
6	57.32	311.55	Sphagn1	Few	Hummock
7	36.95	378.93	Sphagn1	Few	Hummock
8	80.59	266.78	Interface	Many	Blanket
9	61.43	310.70	Litter	Many	Blanket
10	32.14	220.73	Sphagn1	Many	Hummock
11	35.59	134.13	Sphagn3	Many	Blanket
12	46.80	405.91	Sphagn1	Few	Hummock
13	27.97	243.70	Sphagn1	Many	Hummock
14	37.25	239.51	Interface	Many	Blanket
15	59.93	350.64	Interface	Many	Blanket
16	35.41	321.87	Interface	Few	Hummock
17	29.56	296.95	Interface	Many	Hummock
18	25.84	276.44	Sphagn1	Many	Hummock
19	44.10	383.83	Interface	Many	Blanket
20	38.61	145.68	Interface	Many	Hummock
21	25.93	184.04	Sphagn4	Few	Hummock
22	32.74	185.89	Sphagn1	Many	Hummock

+ spatial coordinates

Borcard, D., Legendre, P. & Drapeau, P. (1992) Partialling out the spatial component of ecological variation. *Ecology* 73: 1045-1055.