Vocational Aqualabs

# Vocational Aqualabs - Vocational Generic Skills for Researchers

## Experimental Design

### Introduction to multivariate analysis

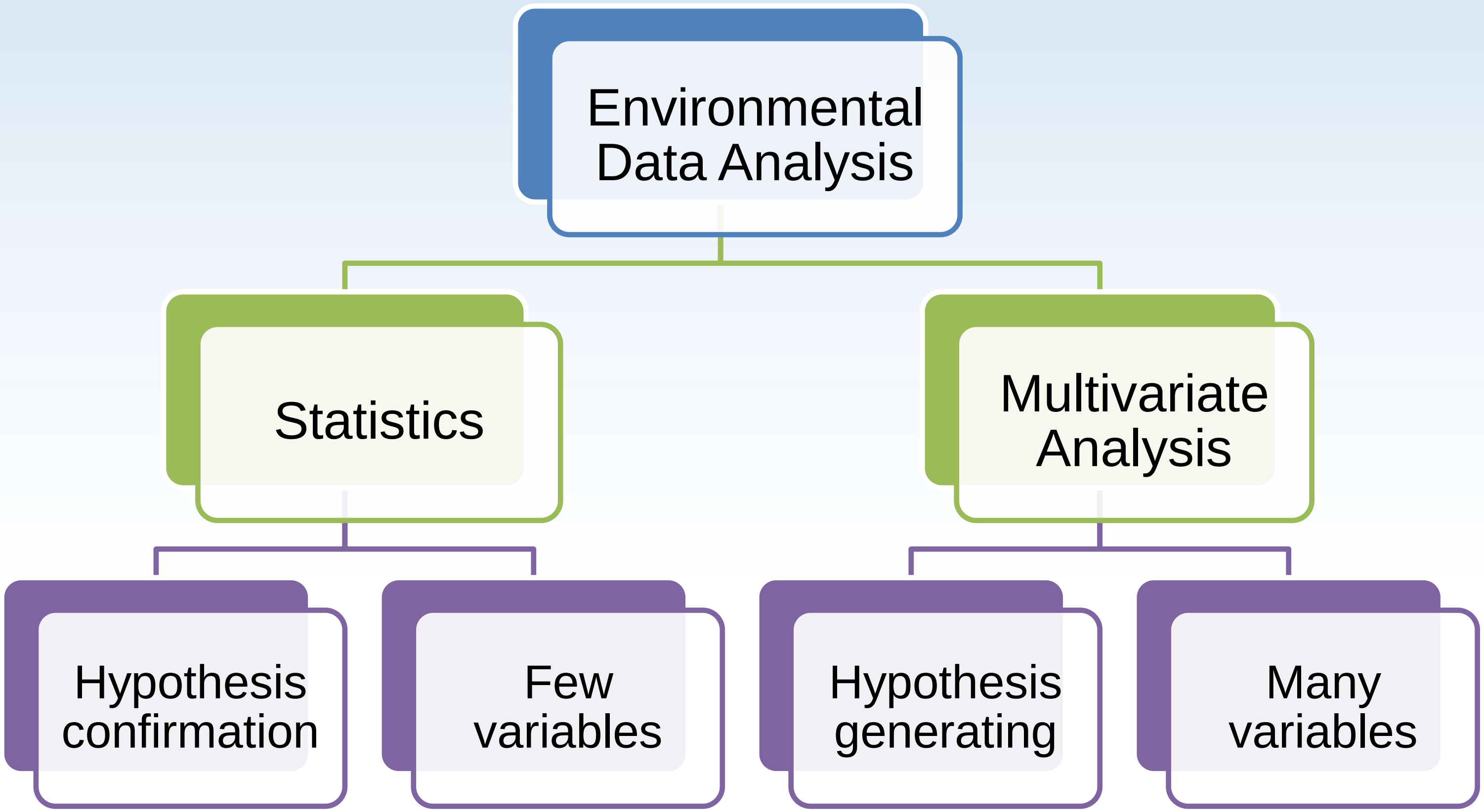
Trevor Telfer/James Bron
Senior Lecturer
University of Stirling

Education and Culture DG
Lifelong Learning Programme

# Multivariate analysis

- Community level data is important for interpretation of ecological effects from properly constructed field studies (see unit 4), though can be used for a variety of lab-derived multivariate datasets
- Community level data is multivariate (many variables) as each sample site is described by the abundances of a number of species or environmental factors
- **What is Multivariate Analysis?**
- **Why use it for analysing environmental data?**
- **What are the main techniques of MVA?**

# Multivariate analysis




Vocational Aqualabs

# Multivariate analysis

**There are three main advantages of using MVA**

1. user appreciates the problem and foresees data type, sampling design and analysis method
2. there is a consistency of implementation
3. detects smaller differences and more subtle changes in community structure

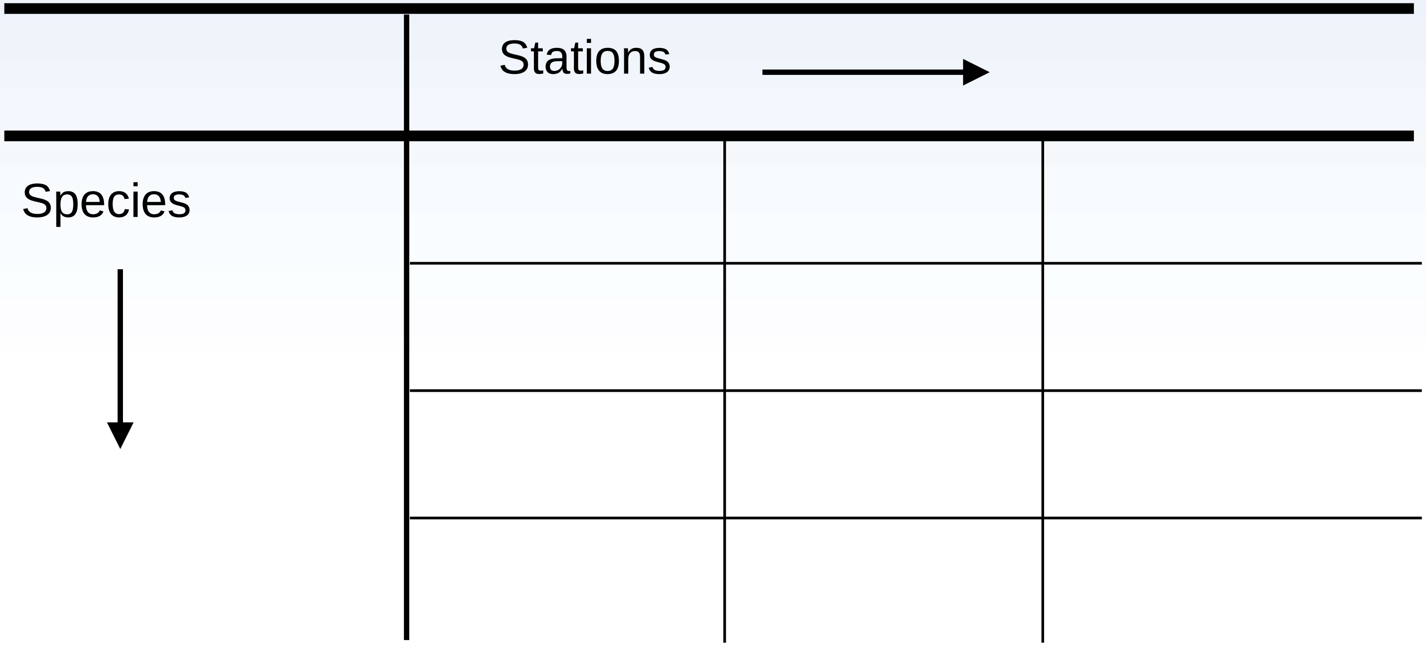
# Multivariate analysis – main types

- **Classification**
- **Ordination**

- **<u>Data input</u>**

# Multivariate analysis – data input

## Data matrix (two way table):



# Multivariate analysis - transformations

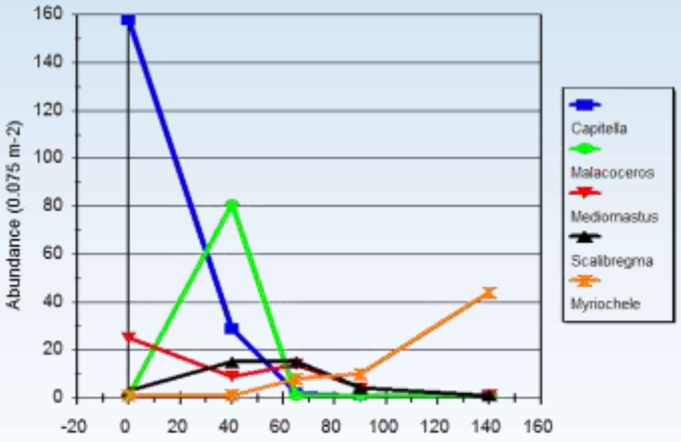
- **for MVA is primarily used to weight the data in order to even out the contribution made by all variables, e.g. species, environmental factors**
- **Commonly used transforms:**
  - square root (sq rt)
  - log transform (lg)
  - fourth root (4rt)
  - binary or presence/absence
- **Strength: sq rt < lg < 4rt < +/-**

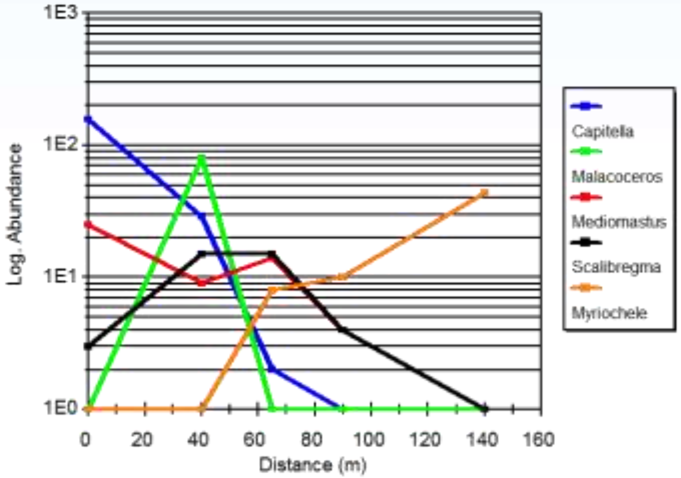


Vocational Aqualabs

# Multivariate analysis - transformations

- **log transformation**
  - downweights abundant species
  - increases importance of rare species
- **log(x + 1)**
  - 1000 ind  > 3.0004
  - 100 ind  > 2.0043
  - 10 ind  > 1.0413
  - 1 ind > 0.3010






Vocational Aqualabs

# Multivariate analysis – main types

- <u>Classification</u>
- Ordination



Vocational Aqualabs

# Multivariate analysis - classification

- **Groups similar entities together in clusters or groups**

- **Two basic kinds:**
  - Non-hierarchical classification, e.g. twoway table
  - Hierarchical classification, e.g. cluster analysis

# Multivariate analysis - classification

- simple ordering or species and stations so their occurrences give a patterned structure along the gradient
- good for large data sets
- done “by eye” or by computer

Table 1.3. *German meadow samples arranged by the Braun-Blanquet method*

| | Sample* | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | B | B | B | B | B | B | B | B | G | G | G | G | G | G | C | O | C | C | C | C | C | C | C | C | C |
| Species | 4 | 10 | 1 | 9 | 15 | 3 | 24 | 2 | 16 | 18 | 12 | 20 | 5 | 22 | 23 | 19 | 13 | 17 | 6 | 11 | 8 | 21 | 25 | 14 | 7 |
| *Ia* | | | | | | | | | | | | | | | | | | | | | | | | | |
| *Bromus erectus* | 9 | 7 | 8 | 8 | 8 | 8 | 6 | | | | | | | | | | | | | | | | | | |
| *Scabiosa columbaria* | 2 | 1 | 1 | | 1 | | | 3 | | | | | | | | | | | | | | | | | |
| *Thymus serpyllum* | 1 | 3 | 2 | | | 1 | | 1 | | | | | | | | | | | | | | | | | |
| *Salvia pratensis* | | 5 | | 3 | 2 | 4 | 4 | | | | | | | | | | | | | | | | | | |
| *Koelria pyramidata* | 3 | 4 | 4 | 4 | | | | | | | | | | | | | | | | | | | | | |
| *Festuca ovina* | | 3 | 3 | 2 | | | | 2 | | | | | | | | | | | | | | | | | |
| *Ib* | | | | | | | | | | | | | | | | | | | | | | | | | |
| *Campanula glomerata* | 2 | | 1 | 2 | | 2 | 1 | | 1 | | 2 | | | | | | | | | | | | | | |
| *Viola hirta* | | 4 | | 3 | 1 | 1 | 1 | | | 1 | | | | | | | | | | | | | | | |
| *Briza media* | | 2 | 2 | 3 | | | | 2 | 3 | | | | | | | | | | | | | | | | |
| *Linum catharticum* | 1 | | 1 | 1 | | 1 | | | 1 | | | | | | | | | | | | | | | | |
| *IIb* | | | | | | | | | | | | | | | | | | | | | | | | | |
| *Geum rivale* | | | | | | | | | 2 | 1 | 1 | 1 | | 1 | | 2 | 1 | 1 | 3 | 4 | 5 | 2 | 1 | | 2 |
| *Holcus lanatus* | | | | | | | | | | | | | 2 | | 1 | 1 | 2 | 3 | 2 | 3 | 3 | 3 | 6 | 3 | 2 |
| *Melandrium diurnum* | | | | | | | | | 2 | | 1 | | | 1 | | 1 | 1 | 1 | 3 | | | 1 | 4 | 2 | 2 |
| *Alopecurus pratensis* | | | | | | | | | | 3 | | | 3 | 6 | | | 4 | | 5 | 5 | 6 | 6 | 2 | | |
| *Lysimachia nummularia* | | | | | | | | | | | 1 | 1 | | | | | | 1 | 1 | 1 | 2 | 1 | 1 | | |
| *Lychnis flos-cuculi* | | | | | | | | | 2 | | | | | | | 1 | | 1 | | 1 | 1 | 1 | 1 | | |
| *Glechoma hederacea* | | | | | | | | | | | 1 | | 1 | 1 | | | | | | 1 | 1 | | | | 1 |
| *IIa* | | | | | | | | | | | | | | | | | | | | | | | | | |
| *Cirsium oleraceum* | | | | | | | | | | | | 1 | 1 | 1 | 2 | 1 | 1 | 3 | 6 | 4 | 7 | 3 | 4 | 7 | 7 |
| *Deschampsia caespitosa* | | | | | | | | | | | | | | | | 3 | | 6 | | 7 | 3 | 5 | | 2 | 6 |
| *Angelica sylvestris* | | | | | | | | | | | | | | | 2 | | 1 | | 2 | | | | 1 | 2 | 3 |
| *Carex acutiformis* | | | | | | | | | | | | | | | | 3 | | | | 4 | | 2 | 3 | | 6 |
| *Filipendula ulmaria* | | | | | | | | | | | | | | | | | 1 | | 4 | | 3 | | | 1 | |
| *Pimpinella magna* | | | | | | | | | | | | | | | | | 2 | 1 | | | | 2 | | | 1 |
| *Polygonum bistorta* | | | | | | | | | | | 1 | | | | | 2 | | | | | | | 4 | | 2 |
| *Arrhenatherum elatius* | 3 | 5 | 1 | 5 | 6 | 3 | 7 | 2 | 6 | 7 | 6 | 7 | 4 | 7 | 7 | 7 | 6 | 6 | 4 | 6 | 6 | 7 | 8 | 6 | 4 |
| *Poa pratensis* | 5 | 5 | 4 | 6 | 6 | 6 | 7 | 9 | 6 | 6 | 7 | 5 | 4 | 6 | 6 | 2 | 3 | 5 | 2 | 5 | 4 | 2 | 6 | 5 | 5 |
| *Crepis biennis* | 1 | 5 | 2 | 3 | 2 | 4 | 5 | 1 | 2 | | 4 | 4 | 1 | 2 | 2 | 1 | 2 | 2 | 4 | 1 | 2 | 1 | | 3 | 3 |



Vocational Aqualabs

# Multivariate analysis - classification

- groups similar entities together into classes in a hierarchical manner
- single analysis may be viewed on several levels, e.g. general / detailed
- relationships are expressed among the entities classified
- poor for large datasets, good for smaller datasets
- two types:
    - **Monothetic divisive** *(all samples in a single cluster then progressively divide into smaller ones)*
    - **Polythetic agglomerative** *(each sample in a single cluster then agglomerate these into larger and larger clusters)*

# Multivariate analysis – assoc. matrix

- comparison of sample pairs or species pairs
- many different techniques, e.g. percentage similarity

PERCENT SIMILARITY

| | MID | A0 | A15 | A30 | A50 | A75 |
|---|---|---|---|---|---|---|
| MID | 100 | 75.794 | 52.241 | 17.121 | 2.006 | 4.839 |
| A0 | 75.794 | 100 | 60.171 | 20.701 | 0 | 0 |
| A15 | 52.241 | 60.171 | 100 | 43.164 | 17.844 | 24.758 |
| A30 | 17.121 | 20.701 | 43.164 | 100.000 | 54.561 | 57.640 |
| A50 | 2.006 | 0 | 17.844 | 54.561 | 100 | 49.979 |
| A75 | 4.839 | 0 | 24.758 | 57.640 | 49.979 | 100 |
| A100 | 1.956 | 0 | 20.994 | 49.447 | 65.520 | 56.377 |
| B0 | 73.289 | 73.947 | 49.904 | 20.060 | 0 | 4.901 |
| B15 | 16.673 | 14.291 | 36.491 | 68.575 | 53.796 | 56.250 |
| B30 | 2.139 | 0 | 21.585 | 53.373 | 61.295 | 55.675 |
| C0 | 75.601 | 74.378 | 46.425 | 17.878 | 1.993 | 0 |
| C15 | 40.886 | 48.576 | 59.823 | 45.219 | 21.957 | 27.091 |
| C30 | 27.396 | 32.810 | 53.496 | 64.443 | 37.842 | 45.030 |
| C50 | 16.729 | 15.864 | 35.951 | 69.550 | 56.097 | 59.998 |
| C75 | 11.721 | 11.107 | 28.022 | 67.388 | 58.387 | 48.990 |
| C100 | 1.545 | 0 | 17.747 | 56.802 | 63.073 | 52.641 |
| D0 | 79.475 | 73.622 | 47.078 | 21.461 | 1.979 | 7.139 |
| D15 | 17.582 | 14.976 | 37.532 | 62.970 | 7.407 | 59.444 |
| D30 | 2.052 | 2.024 | 26.919 | 62.486 | 62.275 | 61.908 |
| REF2 | 4.729 | 2.328 | 27.201 | 52.518 | 56.907 | 51.905 |
| REF3 | 4.417 | 0 | 23.599 | 53.562 | 54.597 | 64.218 |

# Multivariate analysis – sorting methods

- **Agglomeration procedures**

- **e.g.**

  **Nearest neighbour**

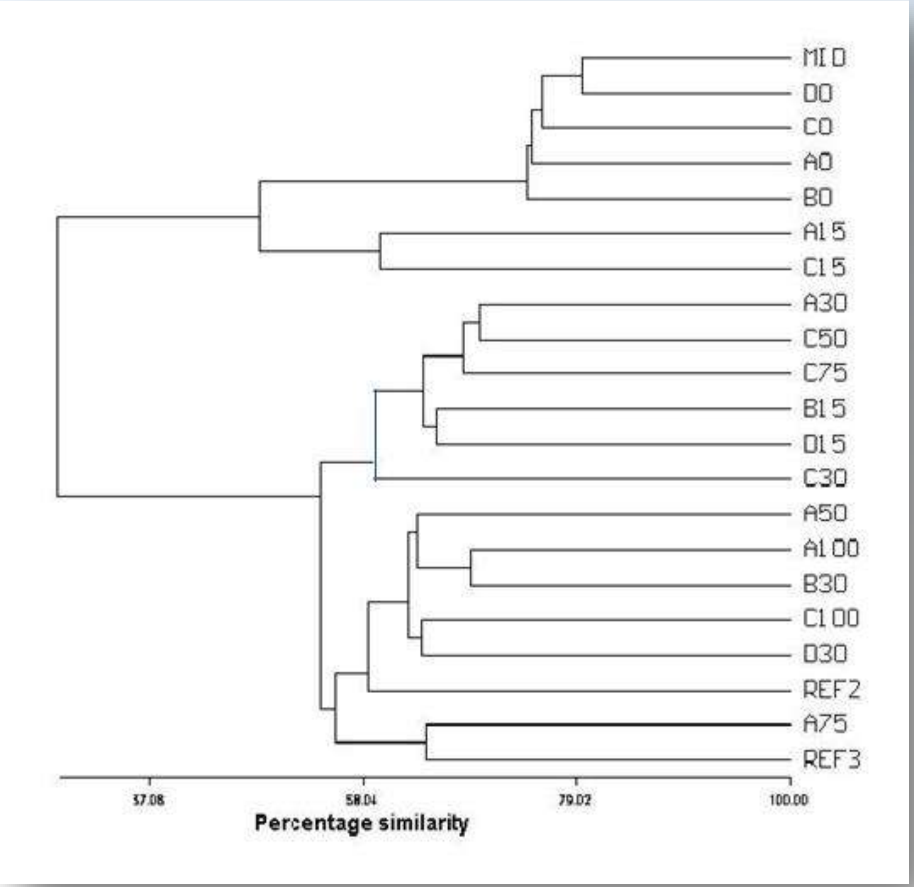
  (lowest similarity measure)

  **Furthest neighbour**

  (highest similarity measure)

  **UPGMA**

  (average similarity measure)

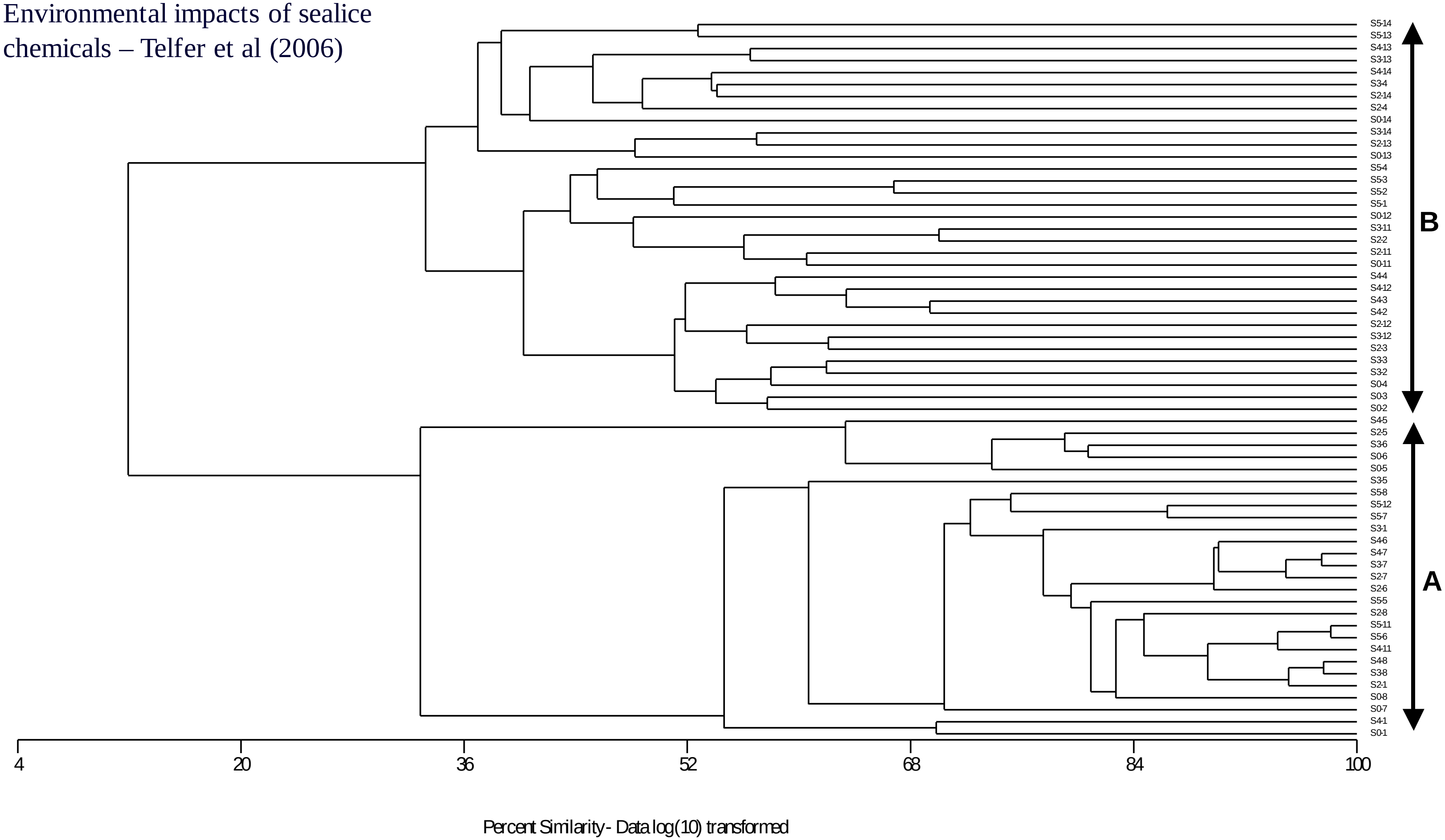





Vocational Aqualabs

# Multivariate analysis – dendrogram

Environmental impacts of sealice chemicals – Telfer et al (2006)





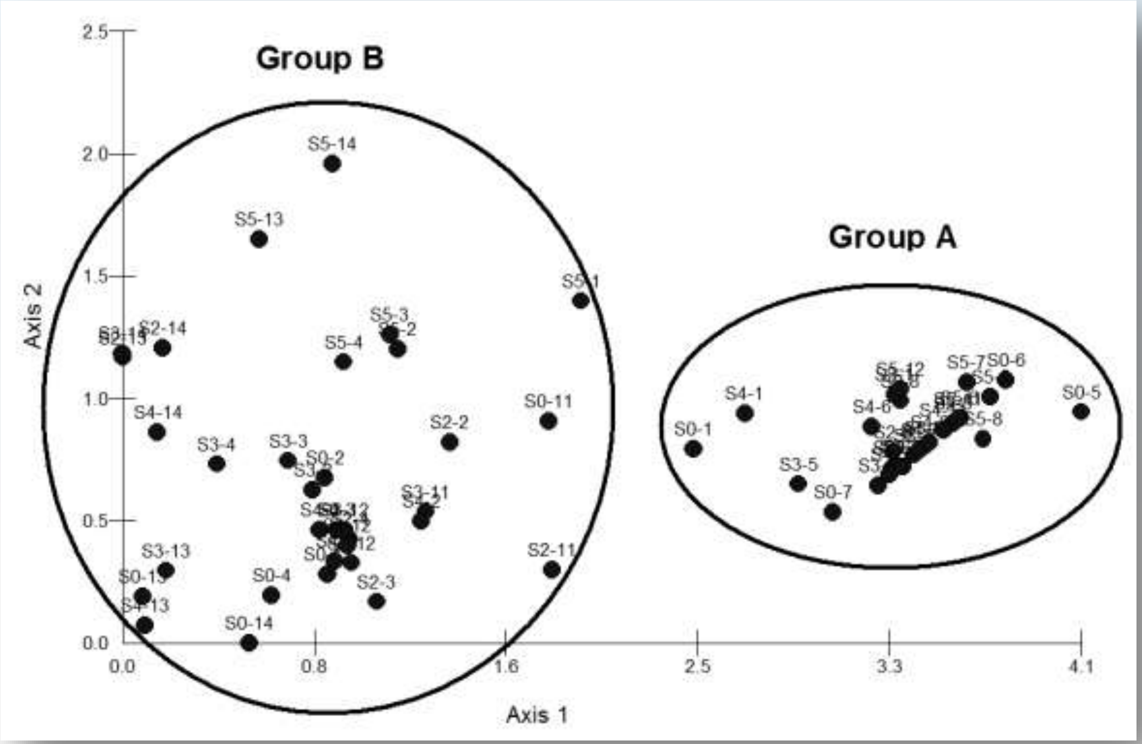

Vocational Aqualabs

# Multivariate analysis – main types

- **Classification**
- **Ordination**

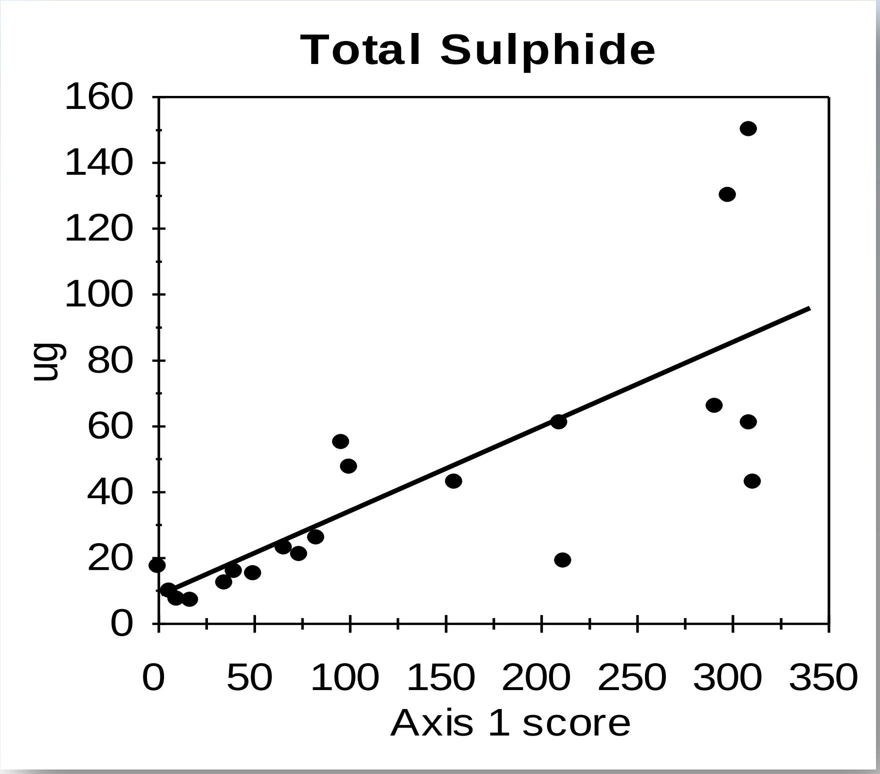
# Multivariate analysis - ordination

- represents samples and species relationships in a low dimensional space
- plot where similar samples/species situated near to each other
- arrangement of samples or species related to environmental factors
  - graphically
  - statistically




Environmental impacts of sealice chemicals – Telfer et al (2006)

# Multivariate analysis - ordination




*Graph - Regression analysis*

Regression analysis

# Multivariate analysis - ordination

Spearman Rank Correlation

| Variable | Axis 1 | Axis 2 | n |
|---|---|---|---|
| Carbon | 0.513* | 0.428 | 20 * |
| Total S | 0.847** | 0.288 | 20 |
| Free S | 0.764** | 0.185 | 19 |

* correlation at $p<0.05$

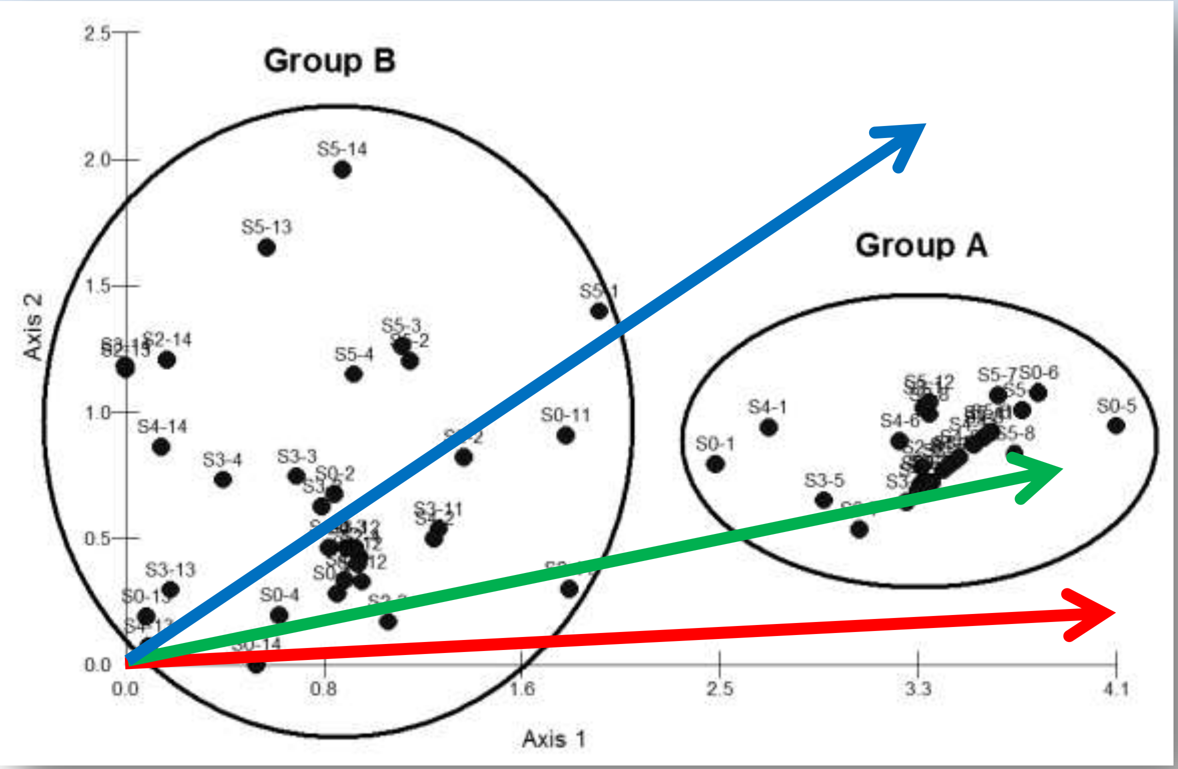
** correlation at $p<0.01$

n represents the number of stations used for the correlation as the chemical parameters were not done at all stations

$$r_s = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

# Multivariate analysis - ordination

- The real relationship between the environmental factors and the positions of the sampling stations would be:



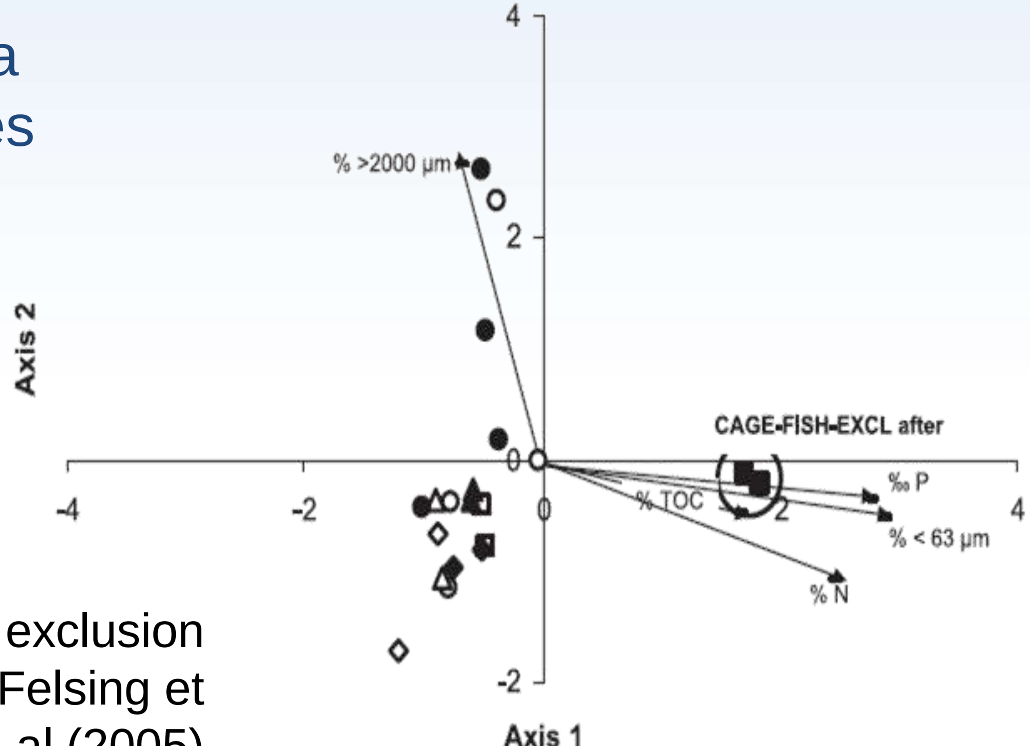

Free Sulphide

Total Sulphide

Total Carbon

# Multivariate analysis - ordination

- **Canonical Correspondence Analysis**
- ordinates both species data and environmental variables
- based on DECORANA
- gives the actual axis directions




Fish exclusion experiments – Felsing et al (2005)

# Multivariate analysis – useful refs

**Greig-Smith, P.** (1982) *Quantitative plant ecology. Third edition.* Studies in Ecology Volume 9. Blackwell Scientific Publications, Oxford.

**Gauch, H.G.** (1982) *Multivariate analysis in community ecology.* Cambridge Studies in Ecology. Cambridge University Press. Cambridge.

**Kent, M & Coker, P** (1994) *Vegetation description and analysis. A practical approach*. John Wiley & Sons, Chichester. 363pp.



Vocational Aqualabs

Vocational Aqualabs

# Thank you

**Trevor Telfer/James Bron**
**Institute of Aquaculture**
**University of Stirling**

ANKARA ÜNİVERSİTESİ 1946

AQUATT

1 4 9 5 UNIVERSITY OF ABERDEEN

UNIVERSITY OF STIRLING

INNOVAMAR

Aquark

Education and Culture DG

Lifelong Learning Programme

This project has been funded with support from the European Commission. This publication reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained herein.