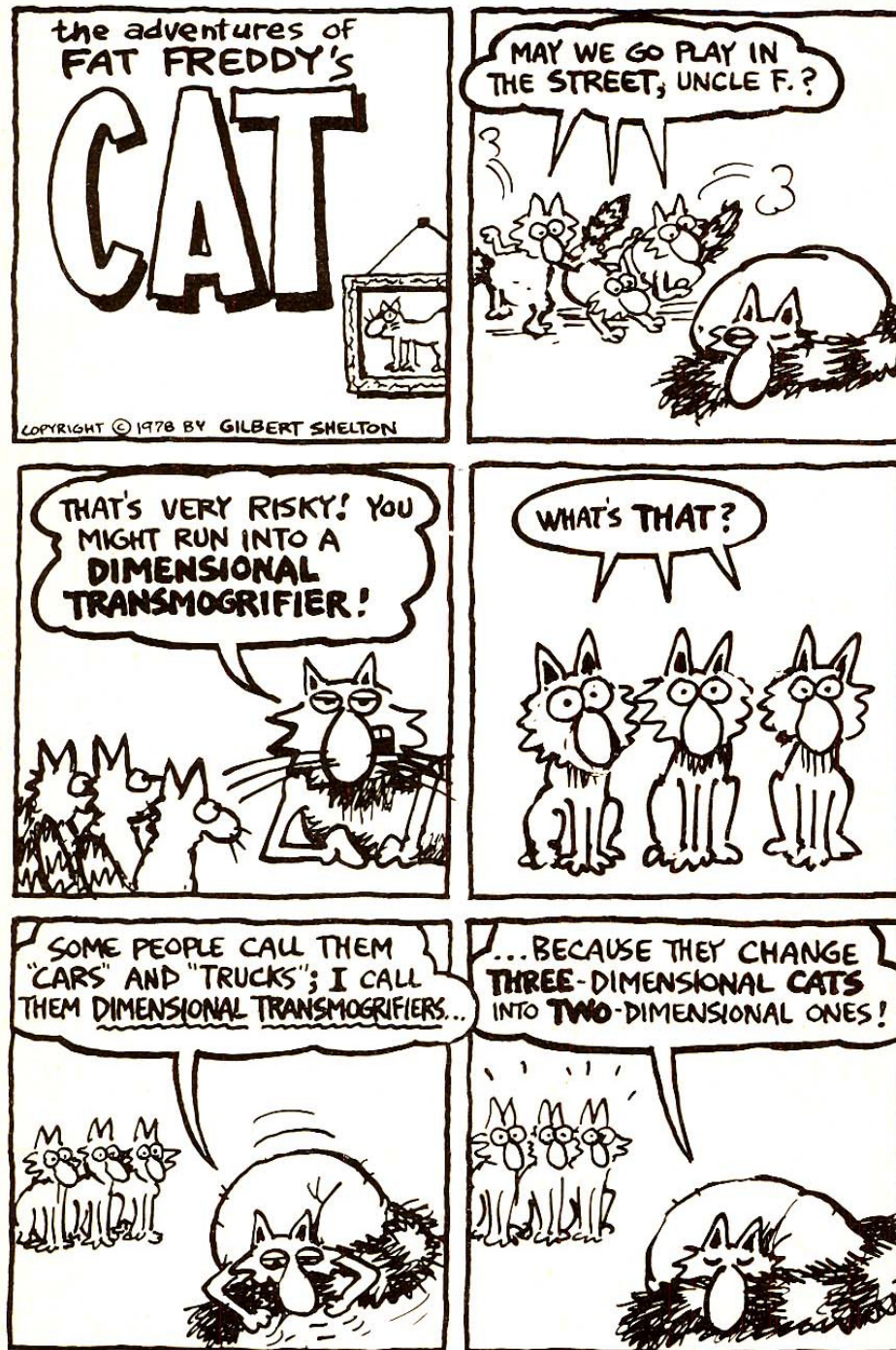


Introduction to ordination

Gary Bradfield
Botany Dept.

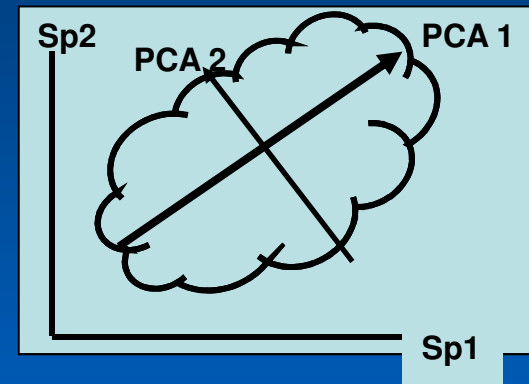


Ordination

“...there appears to be no word in English which one can use as an antonym to “classification”; I would like to propose the term “ordination.” (Goodall, D. W. 1954. Amer. J. Bot. 2: p.323)

MAIN USES:

- Data reduction and graphical display
- Detection of main structure and relationships
- Hypothesis generation
- Data transformation for further analysis



Ordination info & software

<http://ordination.okstate.edu/>



OSU
OSU ECOLOGY

Ordination Methods for Ecologists

JUMP TO: | [OVERVIEW](#) |
[Ordination Topics](#) | [Ordination
Software Links](#) | [Ordination
Glossary](#) | [Other Ordination
Links](#) | [Ordination Listserv](#) |
[OSU Botany](#)

Ordination Topics

Ordination is a widely-used family of methods which attempts to reveal the relationships between ecological communities. For definitions, go [HERE](#).

This ordination web page is designed to address some of the most frequently asked questions about ordination. It is my intention to gear this page towards the student and the practitioner rather than the ordination specialist, so please contact me if the jargon is unintelligible!

The ecological literature is filled with papers describing, contrasting, and modifying existing ordination techniques. Then why is an ordination web page needed? My main motivation is based upon the following observation: many of us, when we start to use ordination methods, make the same simple mistakes. If we are good scientists, we will learn from our own mistakes. But wouldn't it save a lot of time if we could also learn from other people's mistakes?


It turns out that there are a number of frequently asked questions concerning ordination, as well as a number of "tricks of the trade" and "rules of thumb". It is probably not worthwhile filling the pages of our scientific journals with such rules; many of them are quite trivial. However, the Web, which is easy to update, modify, and rearrange, is an ideal forum for presenting such ideas.

As in any scientific endeavor, the field of ordination methodology is filled with conflicting opinions and world views. While I try to be objective, it is difficult to remain completely detached. I would like this to be an open forum! Please send in your comments, reactions, flames, questions, etc. if you would like to see them included in the web page! I am also very eager to

<p>General and Reference</p> <ul style="list-style-type: none"> • Overview of ordination methods • A Glossary for terms used in Ordination • Milestones in the history of Ordination • Ordination terminology: some confusions • The ideal ordination method • Recommendations for ordination: a key • Suggested references for self-education • Hypothesis-driven and Exploratory Analysis • Ordination links 	<p>Indirect Gradient Analysis</p> <ul style="list-style-type: none"> • Distance-based ordination methods • Eigenanalysis-based ordination methods • Principal Components Analysis • Correspondence Analysis • Detrended Correspondence Analysis
<p>Statistics and Background</p> <ul style="list-style-type: none"> • Basic statistical concepts • Multiple Regression • Randomization tests • Centroids and Inertia • Similarity, Distance and 	<p>Direct Gradient Analysis</p> <ul style="list-style-type: none"> • Environmental Variables in Constrained Ordination (CCA, RDA) • Reducing the number of variables • The robustness of CCA • Partial Ordination

<http://home.centurytel.net/~mjm/index.htm>

MjM
Software
Design



[PC-ORD Information](#)
[Analysis of Ecological Communities](#)
[HyperNiche Information](#)
[PC-ORD 5 Fixes](#)
[PC-ORD 4 Fixes](#)
[HyperNiche 2 Fixes](#)
[HyperNiche 1 Fixes](#)
[Send E-mail](#)

PC-ORD™ Version 5
 Multivariate Analysis of Ecological Data
 Software for multivariate statistical analysis of ecological communities. Includes cluster analysis, ordination, 3D graphics, and species diversity.

Analysis of Ecological Communities
 Methods for Analyzing Multivariate Data in Community Ecology
 Book on the various ecological analysis methods, many of which are available in **PC-ORD**. This book should help you to decide which methods to use with your data. It will also help you to understand the basis for those decisions.

HyperNiche™ Version 2
 Multiplicative Habitat Modeling
 Software for habitat modeling and nonparametric regression where the predictors are combined multiplicatively rather than additively. Includes cross-validation, 3D graphics, and GIS input/output.

For technical questions either send e-mail to mjm@centurytel.net or fax to 1-541-764-3935

<http://cc.oulu.fi/~jarioksa/softhelp/vegan.html>

Vegan: R functions for vegetation ecologists

R is a free, open source statistical environment which is not unlike S-plus. It can be downloaded freely for your operating system from the [R archive site](#). More information, such as guides and FAQs, can be found on the same site or through the [R home page](#). Those reading Finnish can get my [R opas ekologille](#). R is so similar to S-plus that you can use S-plus documentation as well, so that perhaps one of the best book form introductions to R is [Modern Applied Statistics with S-plus](#) by Venables & Ripley (Springer, 1999).

Vegan package is intended to help vegetation ecologists and other community ecologists to use R. It contains all major ordination methods, ecologically meaningful dissimilarity indices, tools to analysis of diversity, species richness and abundance models, plus numerous support functions.

Public Release Version

- [link to vegan homepage at CRAN](#), latest official release with binaries for Windows and MacOS X.
- [Vegan tutorial \(pdf file\)](#): First version of a tutorial of community ordination with R. This is a [Sweave](#) document and all output (including graphics) were generated by R when processing the Sweave source into LaTeX so that you should be able to re-generate all analyses.
- [FAQ](#): First version of Frequently Asked Questions. This is intended to be non-technical, and really to answer to questions asked.

Development versions

Vegan development happens now at [R Forge](#).

- Vegan home page is <http://vegan.r-forge.r-project.org/>.
- Install Windows or MacOS X binaries or source files through R-Forge from an R prompt: `install.packages("vegan", repos="http://r-forge.r-project.org/").`
- [ChangeLog](#) at R-Forge. More detailed and up-to-date info can be found in [vegan R-Forge](#).
- Vegan help [web pages](#) (links work only between vegan functions; links to other packages are broken).

Vegetation Science in R

Dave Roberts has [excellent tutorials](#) on using R/S-plus in vegetation and community analysis.

My brief [lecture notes](#) discuss vegetation analysis in R. Unfortunately they are not quite finished: The lectures are over for 2003, and I may not have motivation to correct and complete the notes before the next season.

Ordination background:

Community-unit hypothesis:
“classification” of discrete variation

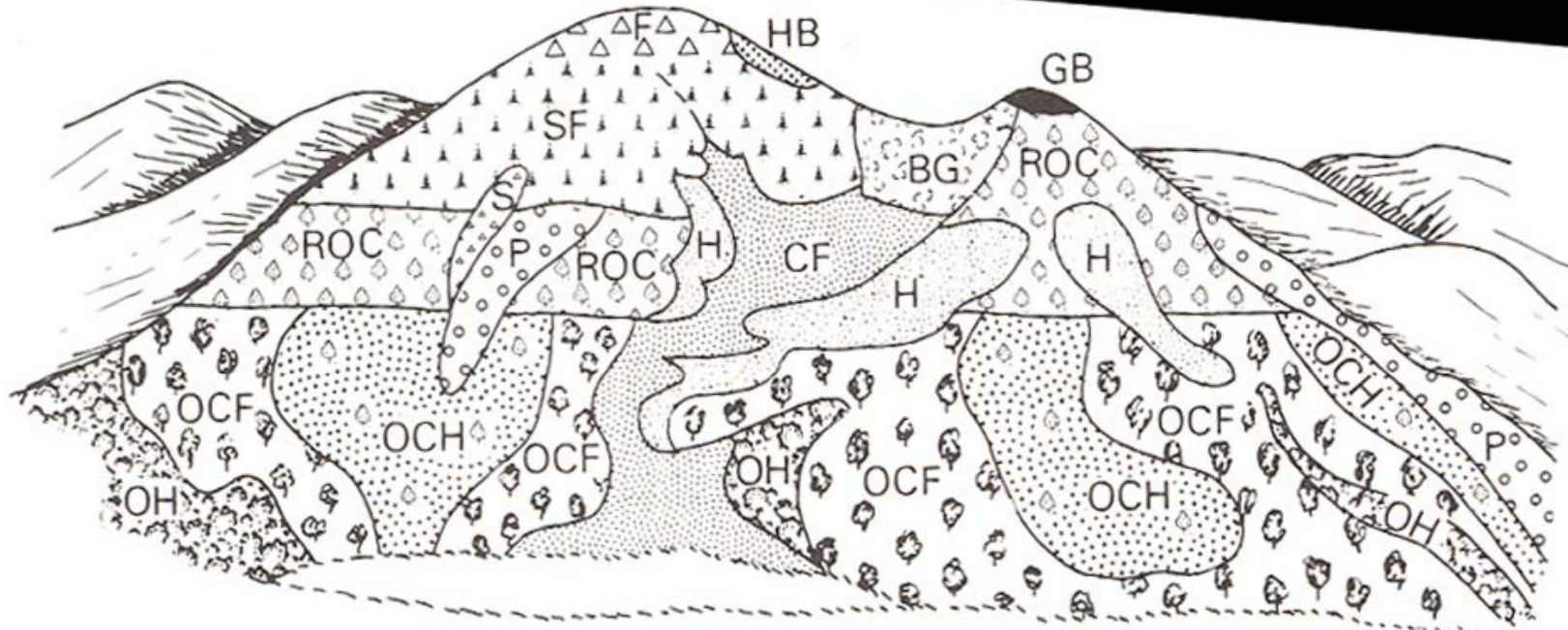
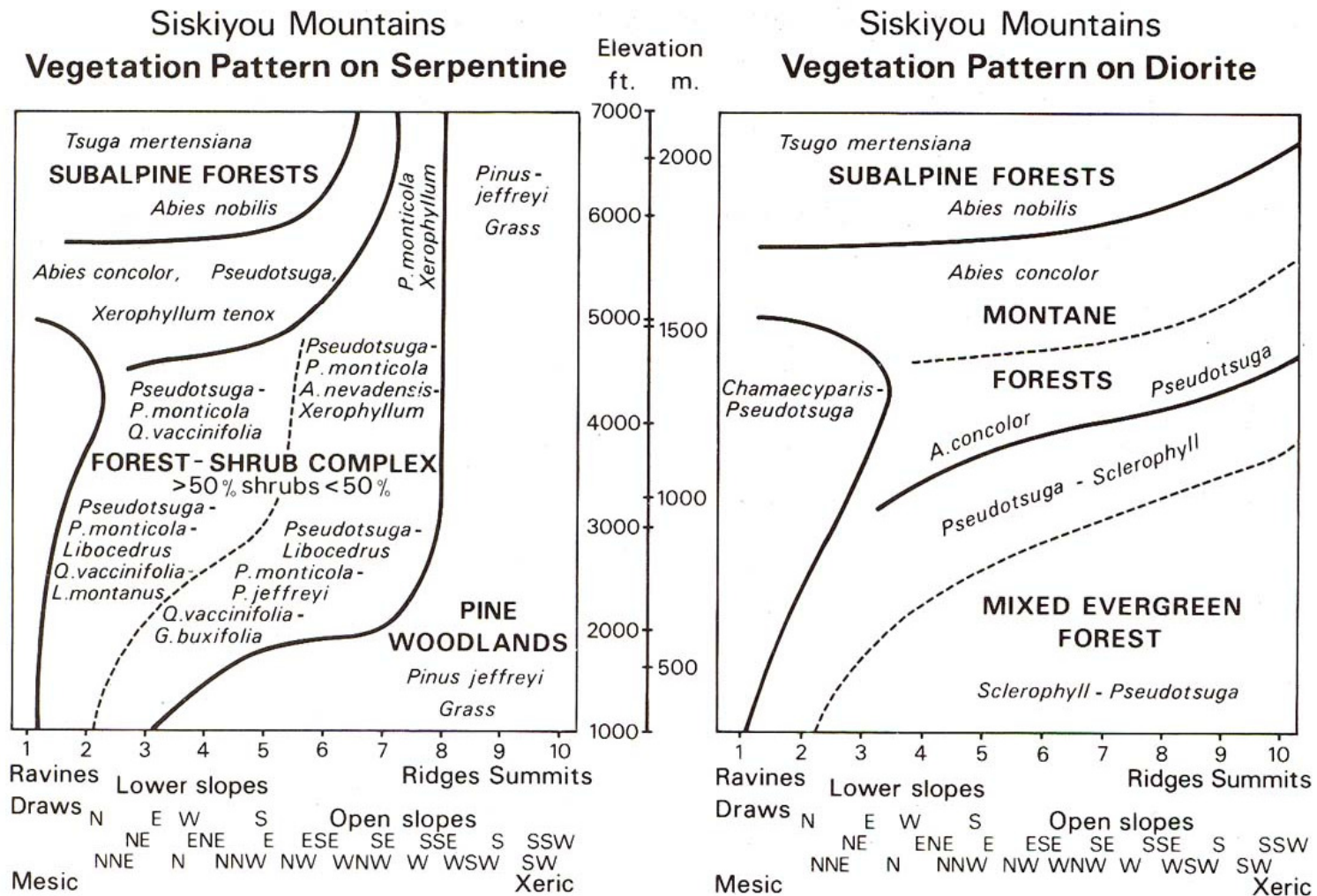


Figure 1.9 Topographic distributions of vegetation types on an idealised west-facing mountain and valley in the Great Smokey Mountains, USA. Vegetation types: BG, beech gap; CF, cove forest; F, Fraser fir forest; GB, grassy bald; H, hemlock forest; HB, heath bald; OCF, chestnut oak–chestnut forest; OCH, chestnut oak–chestnut heath; OH, oak–hickory forest; P, pine forest and pine heath; ROC, red oak–chestnut forest; S, spruce forest; SF, spruce–fir forest. (Redrawn from Whittaker, 1956; with kind permission of *Ecological Monographs*)

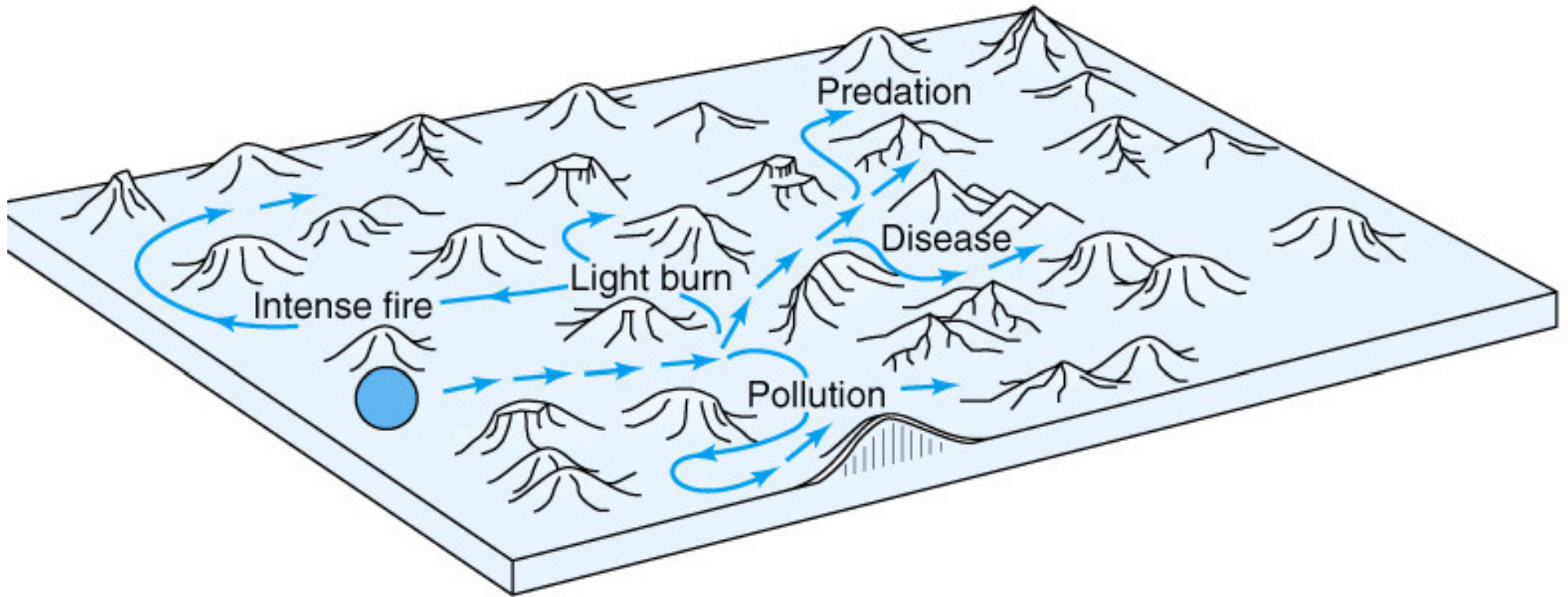
Ordination background:

Individualistic hypothesis: “ordination” of continuous variation



Ordination background:

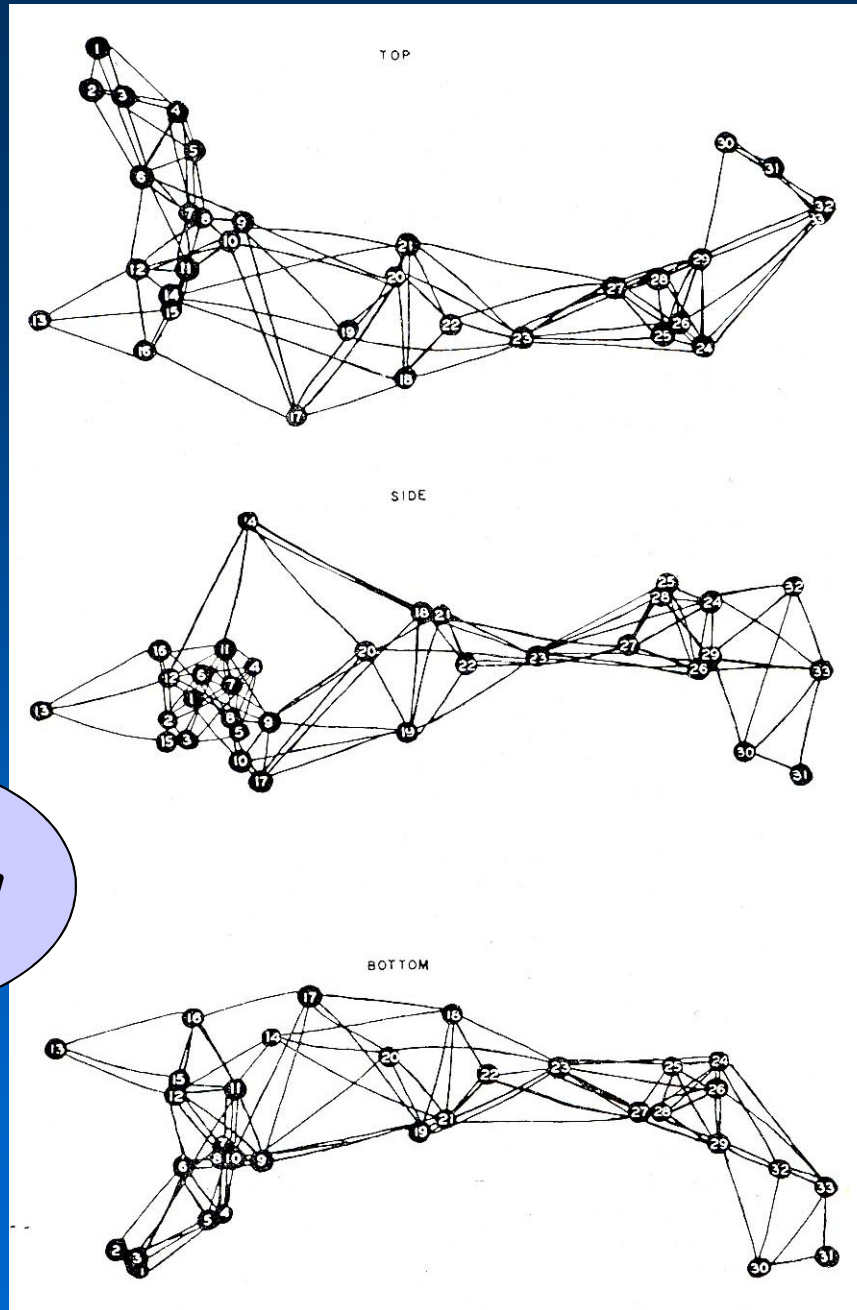
Nonequilibrium landscape model



- continuous interplay of spatial & temporal processes
- consistent with ordination approach to analysis

Early ordinations:

Plexus diagram
of plant species
in Saskatchewan
(Looman 1963)

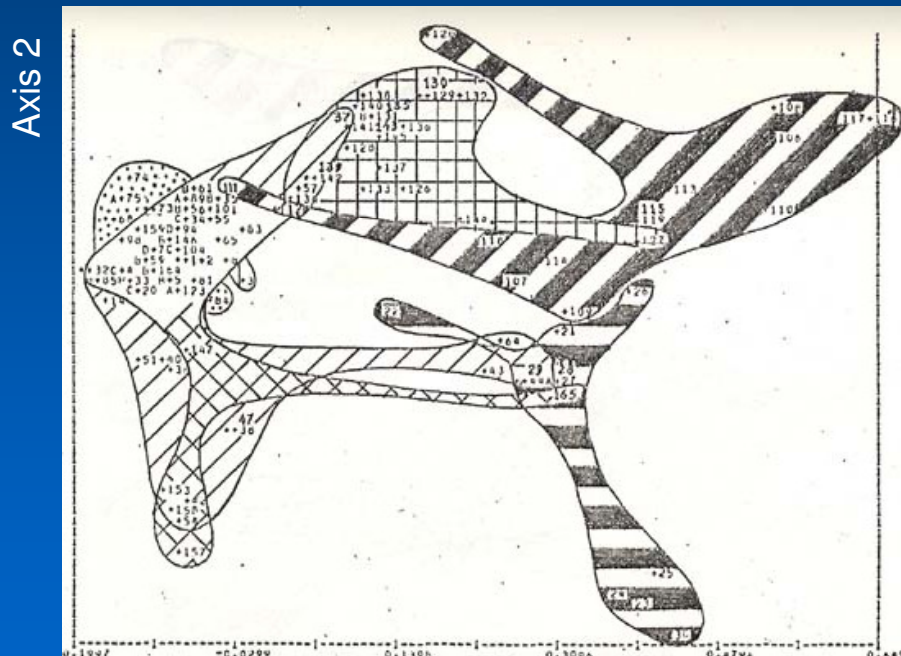


Bow-wow

Early ordinations:

PCA of *Eucalyptus* forest localities after fire in S.E. Australia (Bradfield 1977)

Species covariance

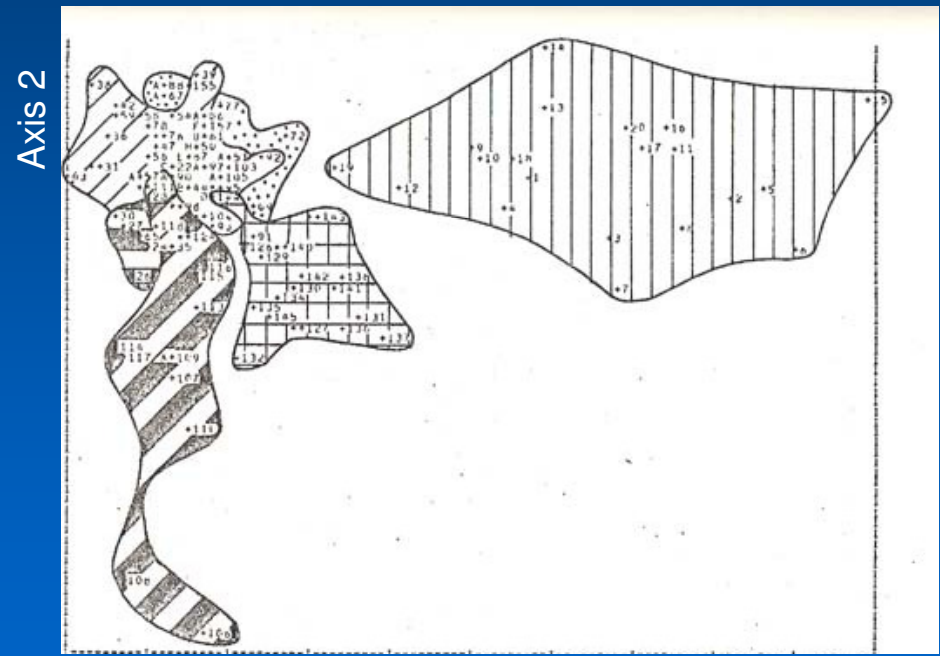


shrub cover



Axis 1

Species correlation



rare species



Axis 1

Early ordinations:

NMS ordination of Scottish cities (Coxon 1982)

Matrix of ranked distances between cities

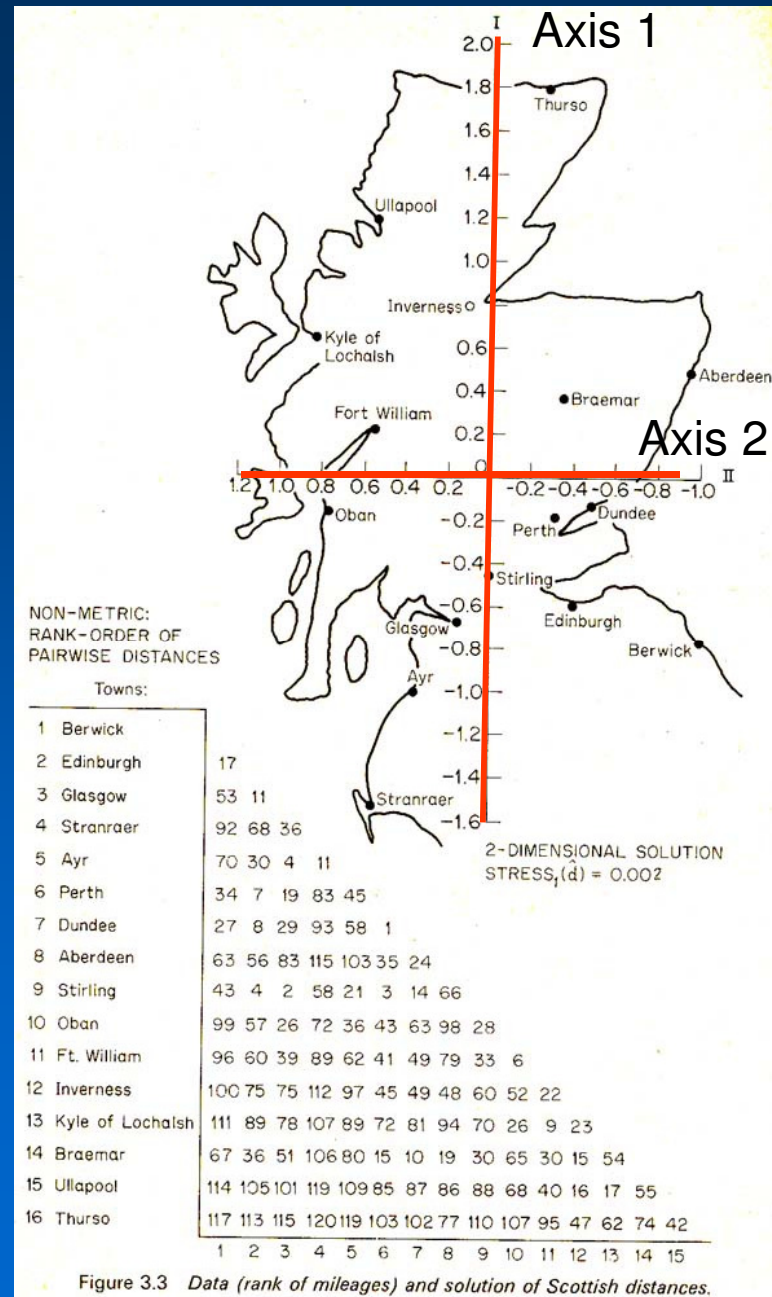
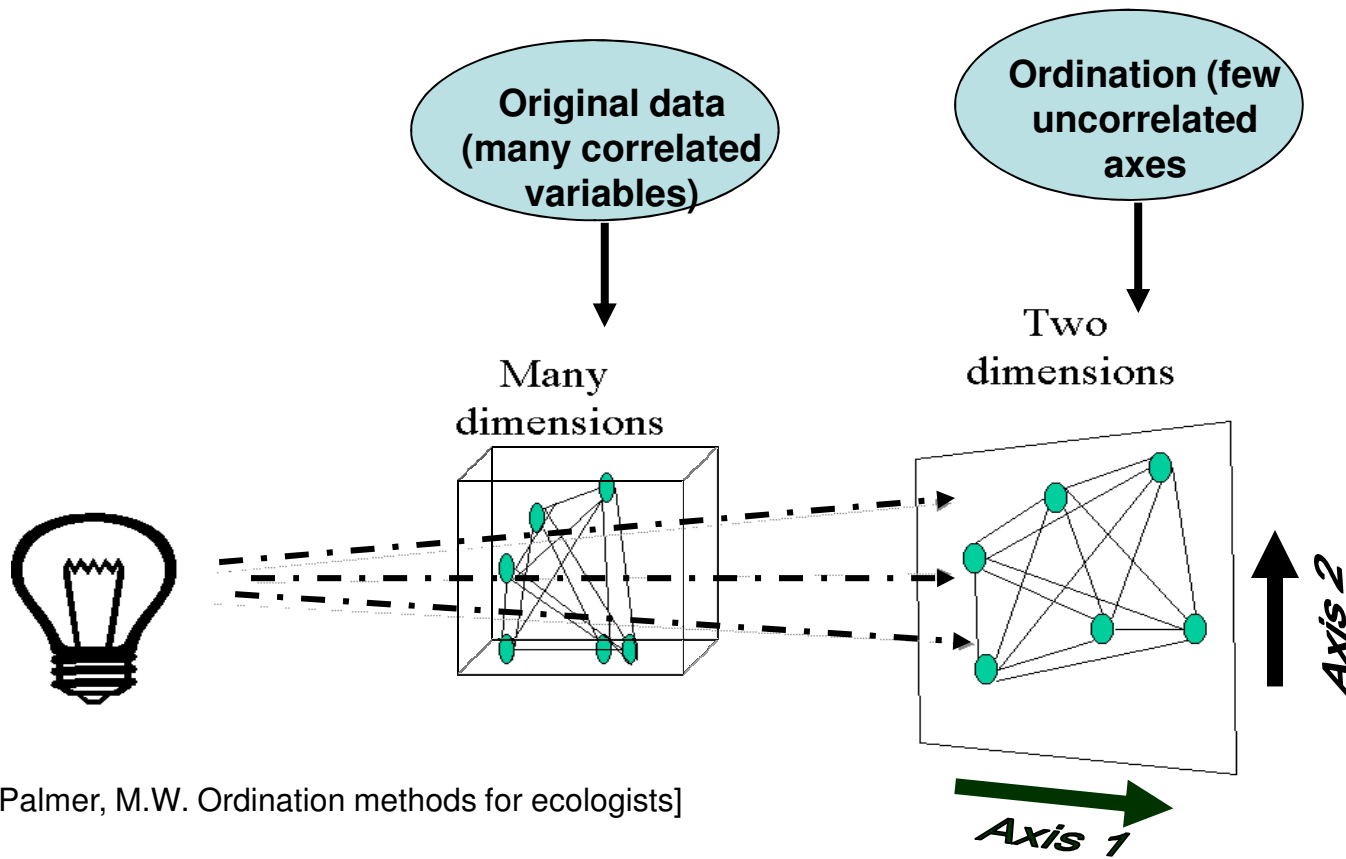


Figure 3.3 Data (rank of mileages) and solution of Scottish distances.

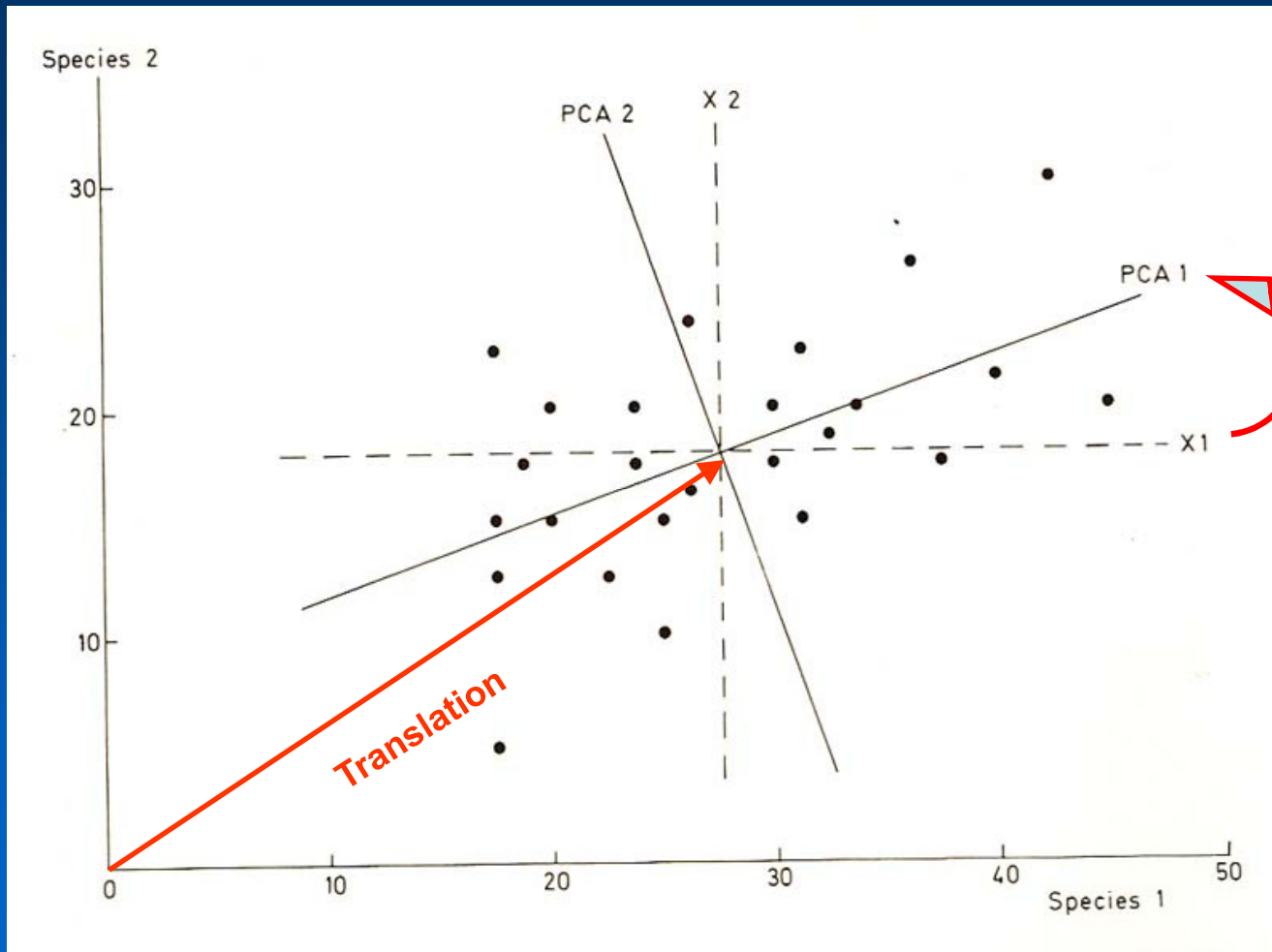
Basic idea of ordination:



[Source: Palmer, M.W. Ordination methods for ecologists]

<http://ordination.okstate.edu/>

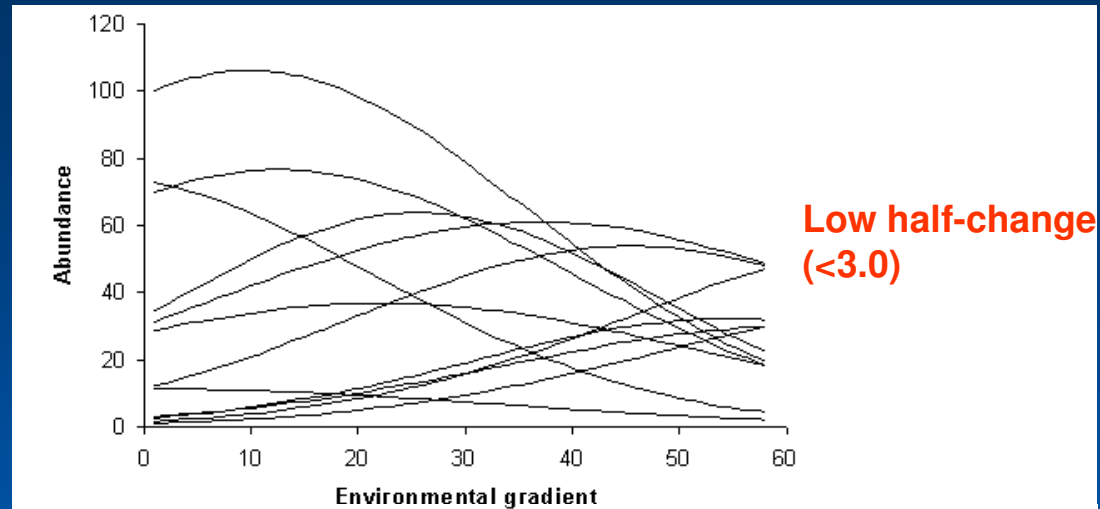
Geometric model of PCA



Rotation
"eigenanalysis"

PCA assumes linear relations among species

Linear



Non-linear

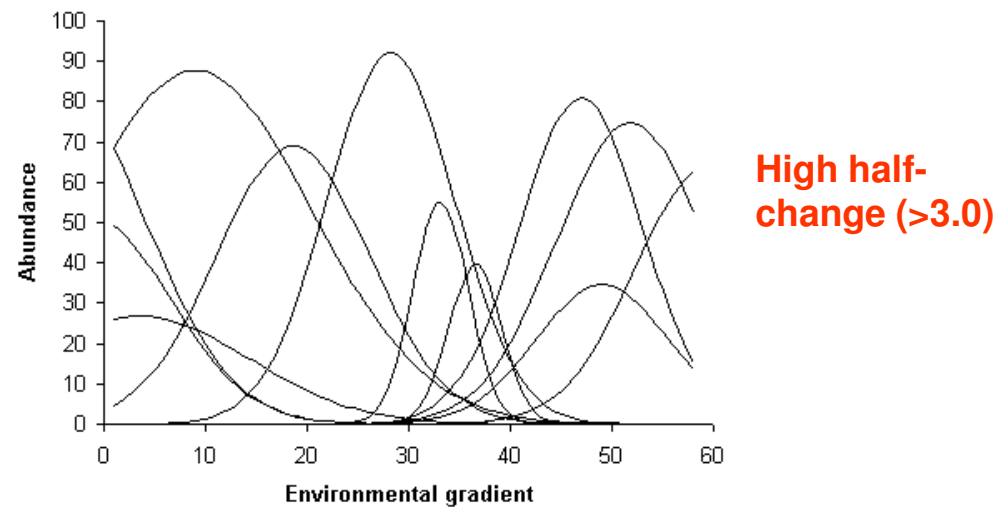


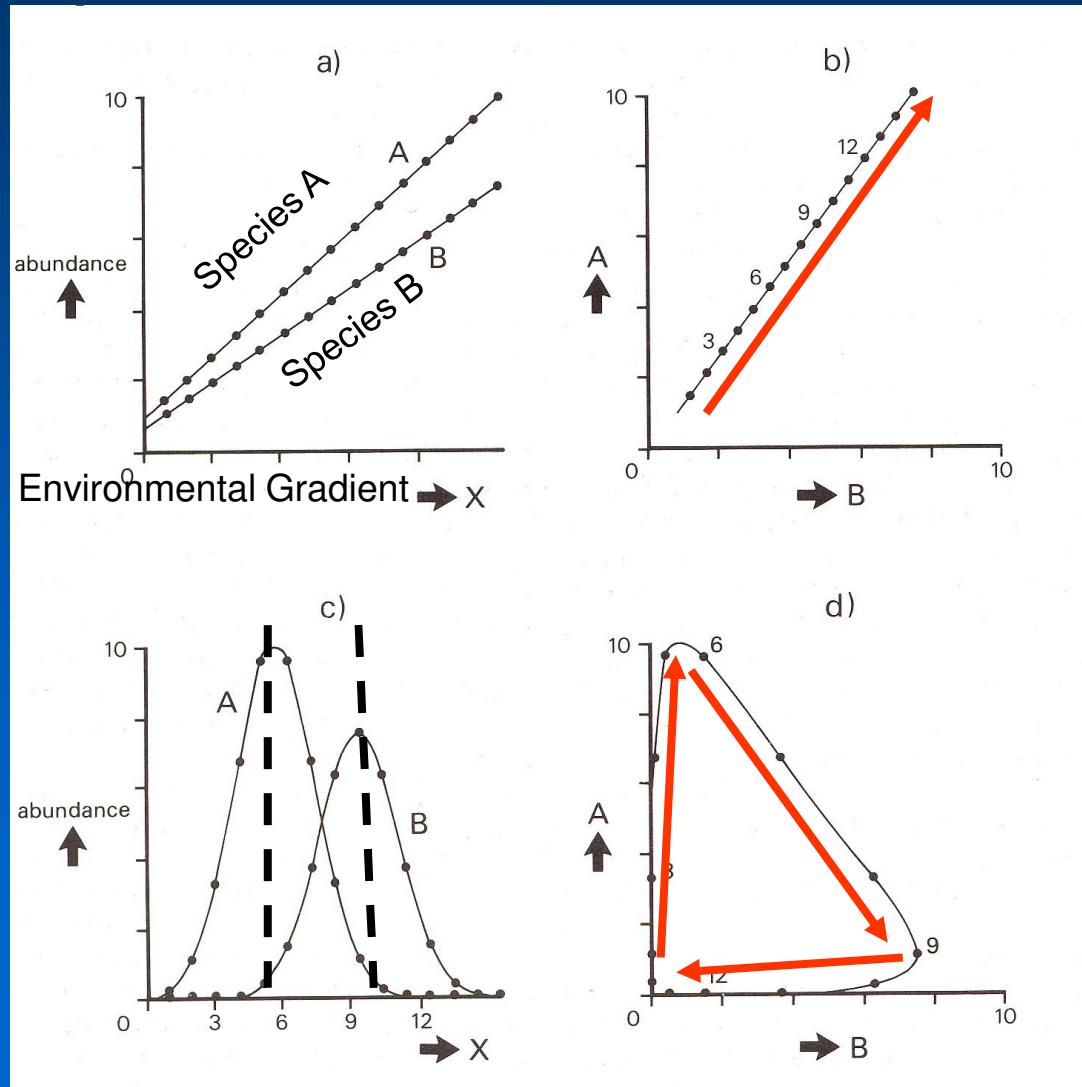
Figure 2: Two coenoclines. Note these are hypothetical examples; real examples would have much noise. example.

PCA assumes linear relations among species

Environment space

Species space

Linear



Non-linear

CHOOSING AN ORDINATION METHOD

Unconstrained methods	Constrained methods
<p>Methods to describe the structure in a single data set:</p> <ul style="list-style-type: none">• PCA (principal component analysis on a covariance matrix or a correlation matrix)• CA (correspondence analysis, also known as reciprocal averaging)• DCA (detrended correspondence analysis)• NMS (nonmetric multidimensional scaling, also known as NMDS)	<p>Methods to explain one data set by another data set (ordinations constrained by explanatory variables):</p> <ul style="list-style-type: none">• RDA (redundancy analysis, the canonical form of PCA)• CCA (canonical correspondence analysis, the canonical form of CA)• CANCOR (canonical correlation analysis)• “Partial” analysis (methods to describe the structure in a data set after accounting for variation explained by a second data set i.e. covariable data)

NMS (Nonmetric multidimensional scaling)

- Goal of NMS is to position objects in a space of reduced dimensionality while preserving rank-order relationships as well as possible (i.e. make a nice picture)
- Wide flexibility in choice of distance coefficients
- Makes no assumptions about data distributions
- Often gives "better" 2 or 3 dimensional solution than PCA (but NMS axes are arbitrary)
- Success measured as that configuration with lowest "stress"

$$\sqrt{\frac{\sum_{i < j} (d_{ij} - \hat{d}_{ij})^2}{\sum_{i < j} d_{ij}^2}}$$

<u>Stress</u>	<u>Fit</u>
0.40	Poor
0.30	Fair
0.20	Good
0.10	Excellent
0.00	Perfect

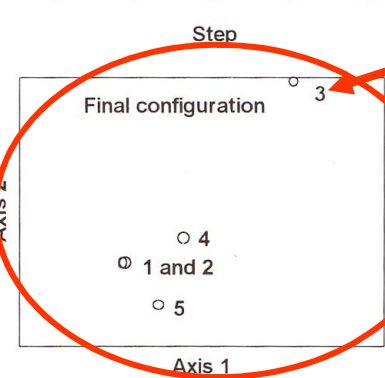
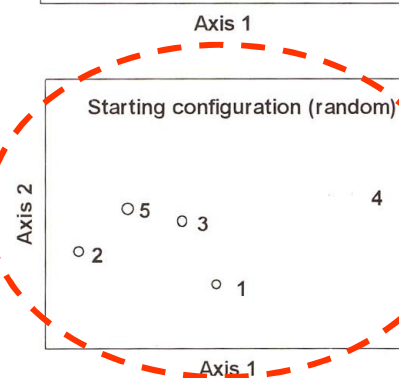
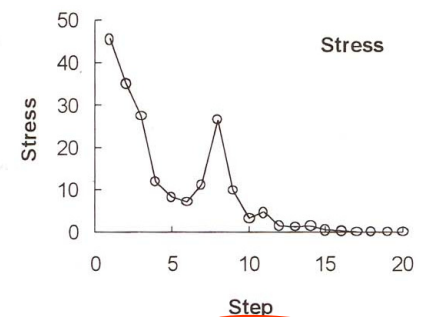
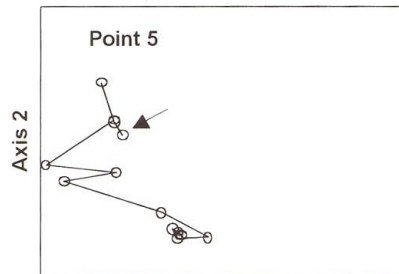
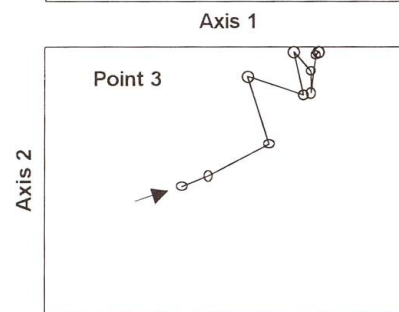
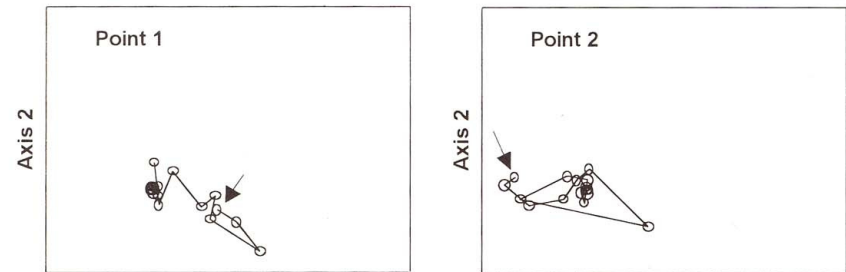
NMS illustration (McCune & Grace 2002)

Table 16.5. Abundance of six species in each of five sample units.

SU	Species					
	sp1	sp2	sp3	sp4	sp5	sp6
1	1	2	3	4	5	5
2	1	3	2	4	6	6
3	0	3	0	1	0	1
4	1	2	2	2	3	4
5	5	2	1	3	5	6

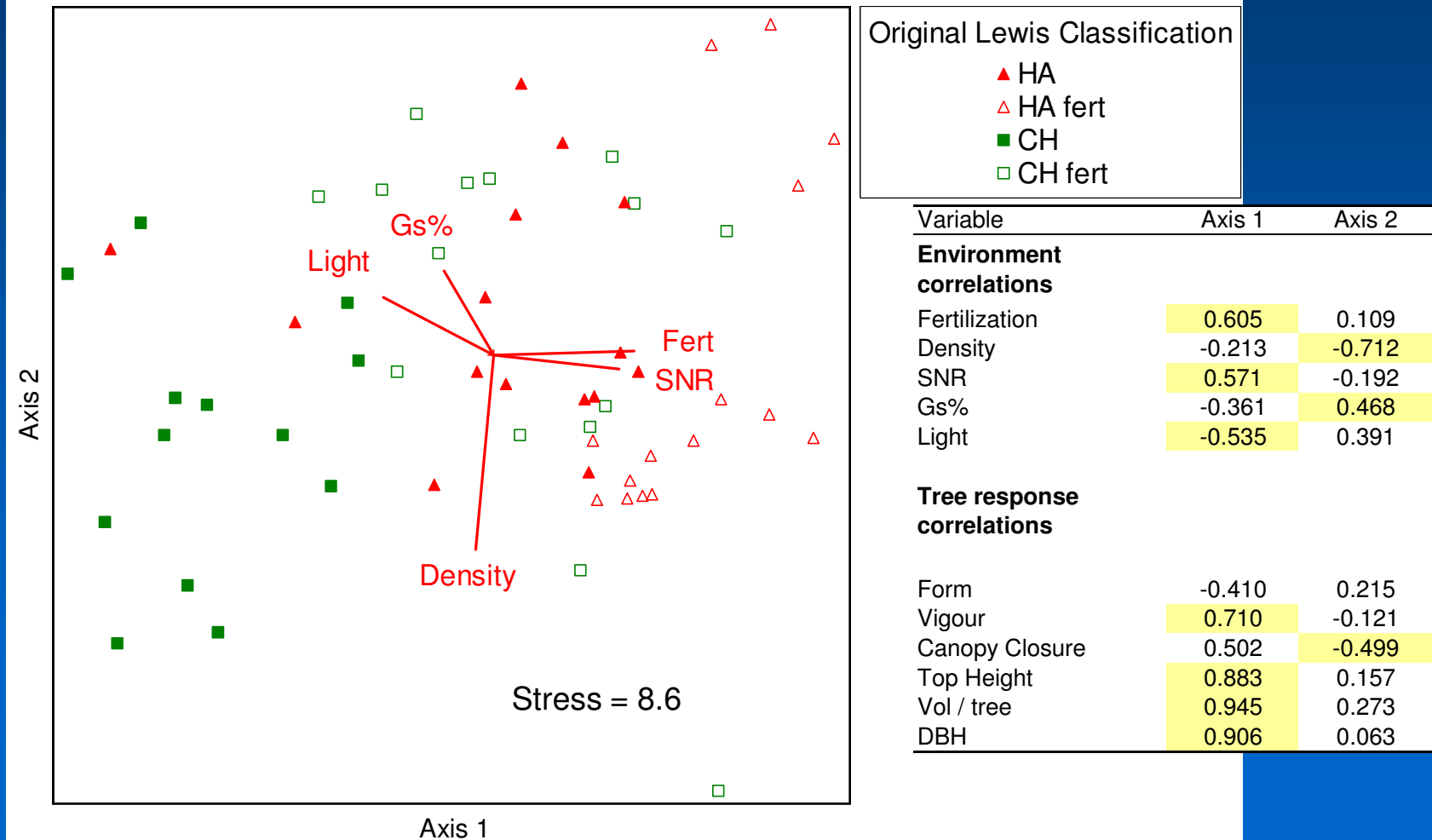
Table 16.6. Sørensen distances among the five sample units from Table 16.5.

SU	SU				
	1	2	3	4	5
1	0				
2	0.0952	0			
3	0.6800	0.6296	0		
4	0.1765	0.2222	0.5789	0	
5	0.1905	0.1818	0.7037	0.2778	0



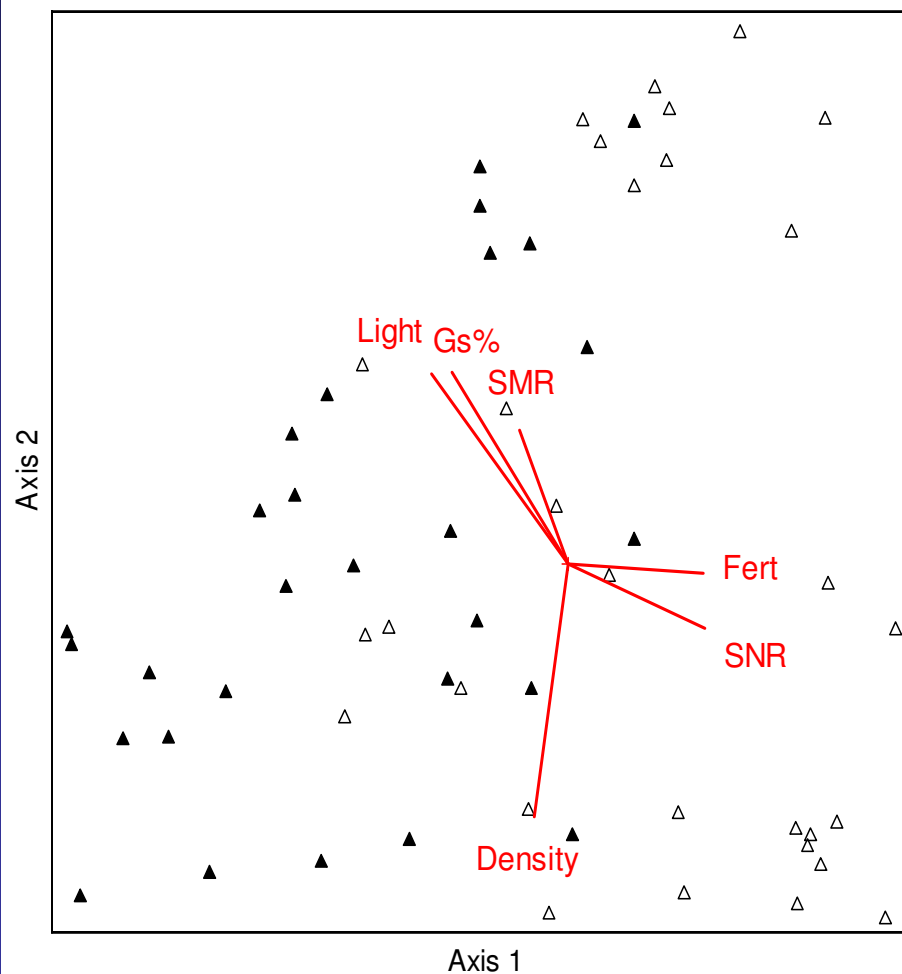
Example: Planted hemlock trees – northern Vancouver Island (Shannon Wright MSc thesis)

NMS (Nonmetric Multidimensional Scaling)



Example: Planted hemlock trees – northern Vancouver Island (Shannon Wright MSc thesis)

CCA (Canonical Correspondence Analysis)

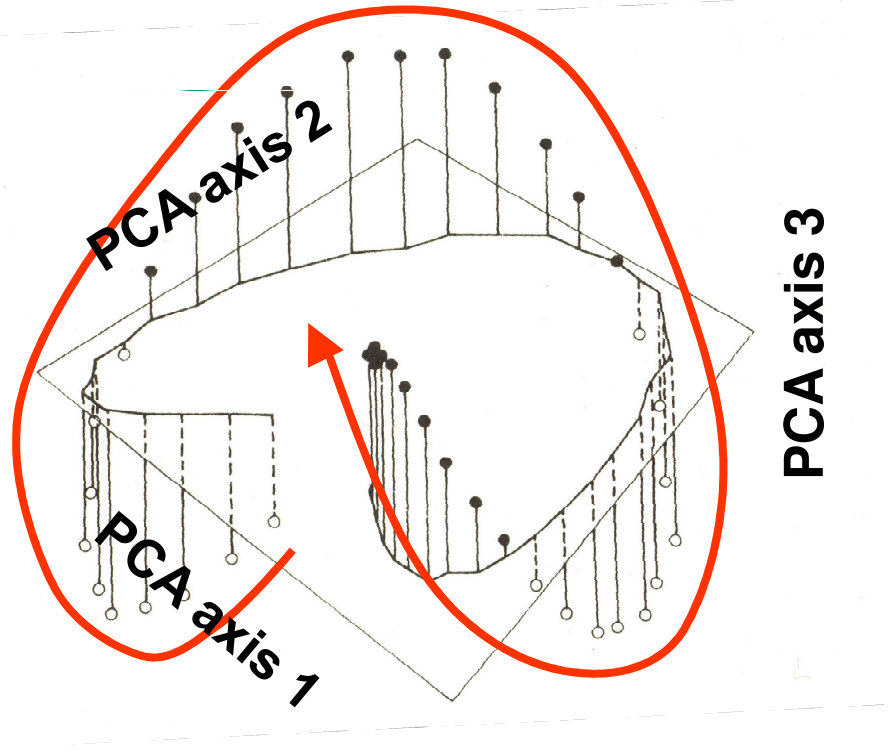
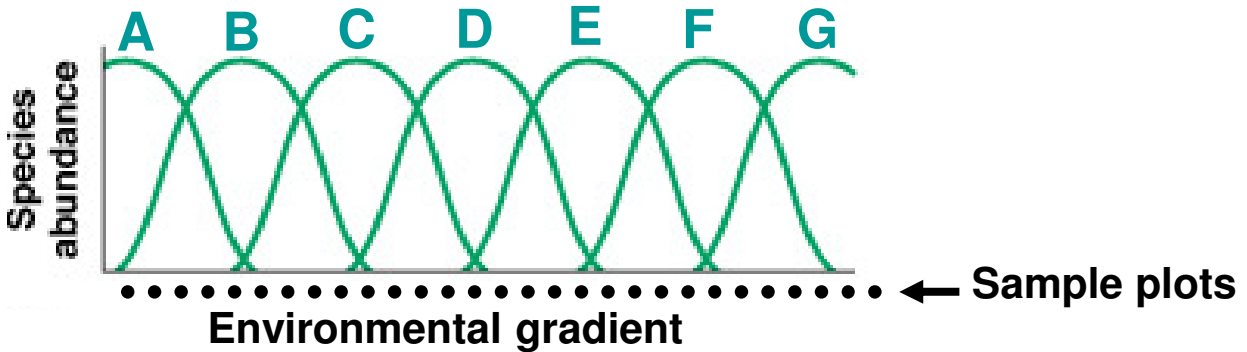


Variable	Axis 1	Axis 2
Environment intraset correlations		
Fertilization	0.620	-0.125
Scarification	-0.163	-0.433
Density	-0.314	-0.892
SMR	-0.094	0.531
SNR	0.588	-0.327
FFcm	-0.236	0.267
Gs%	-0.496	0.757
Rs%	0.123	0.009
For Flr	0.279	-0.320
Light	-0.532	0.776
Tree response correlations		
Form	0.040	0.080
Vigour	0.744	-0.158
Canopy Closure	0.469	-0.721
Top Height	0.834	-0.132
Vol / tree	0.812	-0.050
DBH	0.870	0.056

Evaluating an ordination method:

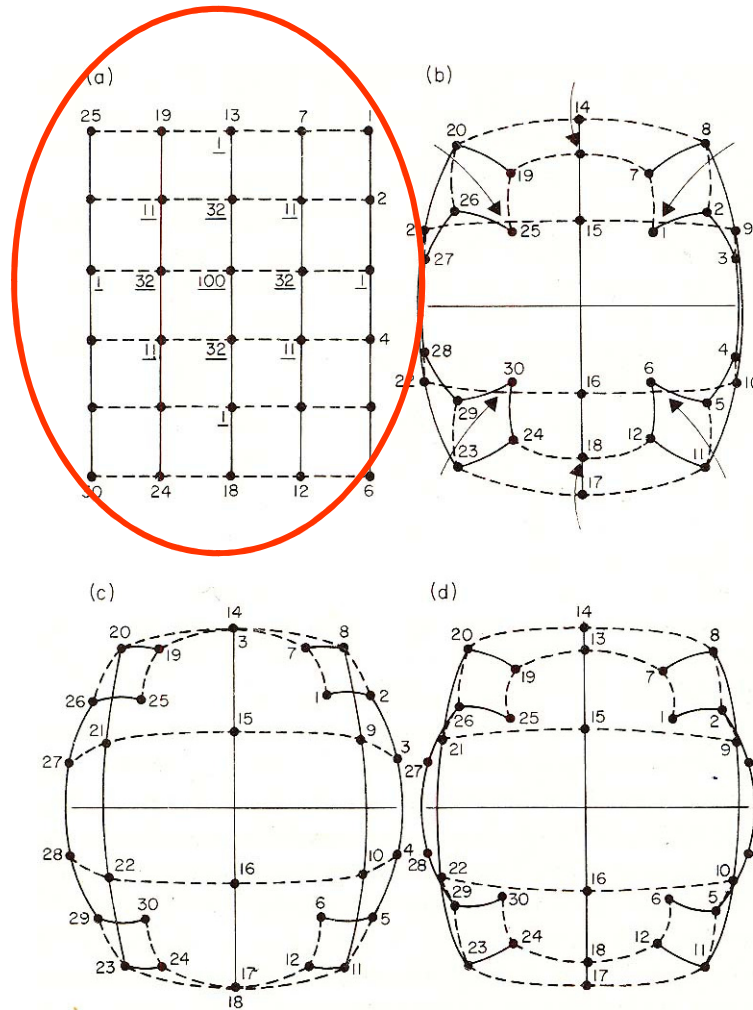
- “Eyeballing” – Does it make sense?
- Summary stats:
 - variance explained (PCA) $(\lambda_i / \sum \lambda_i) * 100\%$
 - correlations with axes (all methods)
 - stress (NMS)
$$\sqrt{\frac{\sum_{i < j} (d_{ij} - \hat{d}_{ij})^2}{\sum_{i < j} d_{ij}^2}}$$
- Performance with simulated data:
 - coenocline: single dominant gradient
 - coenoplane: two (orthogonal) gradients

Simulated data: 1-D coenocline (>2 species, 1 gradient)

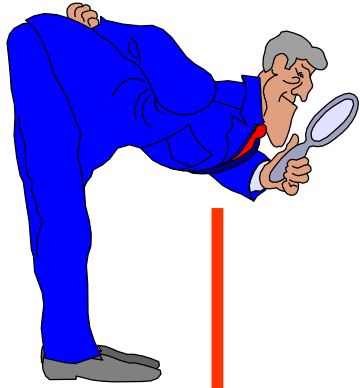


Simulated data: 2-D coenoplane (>2 species, 2 gradients)

Sampling
grid (30
plots x 30
species)



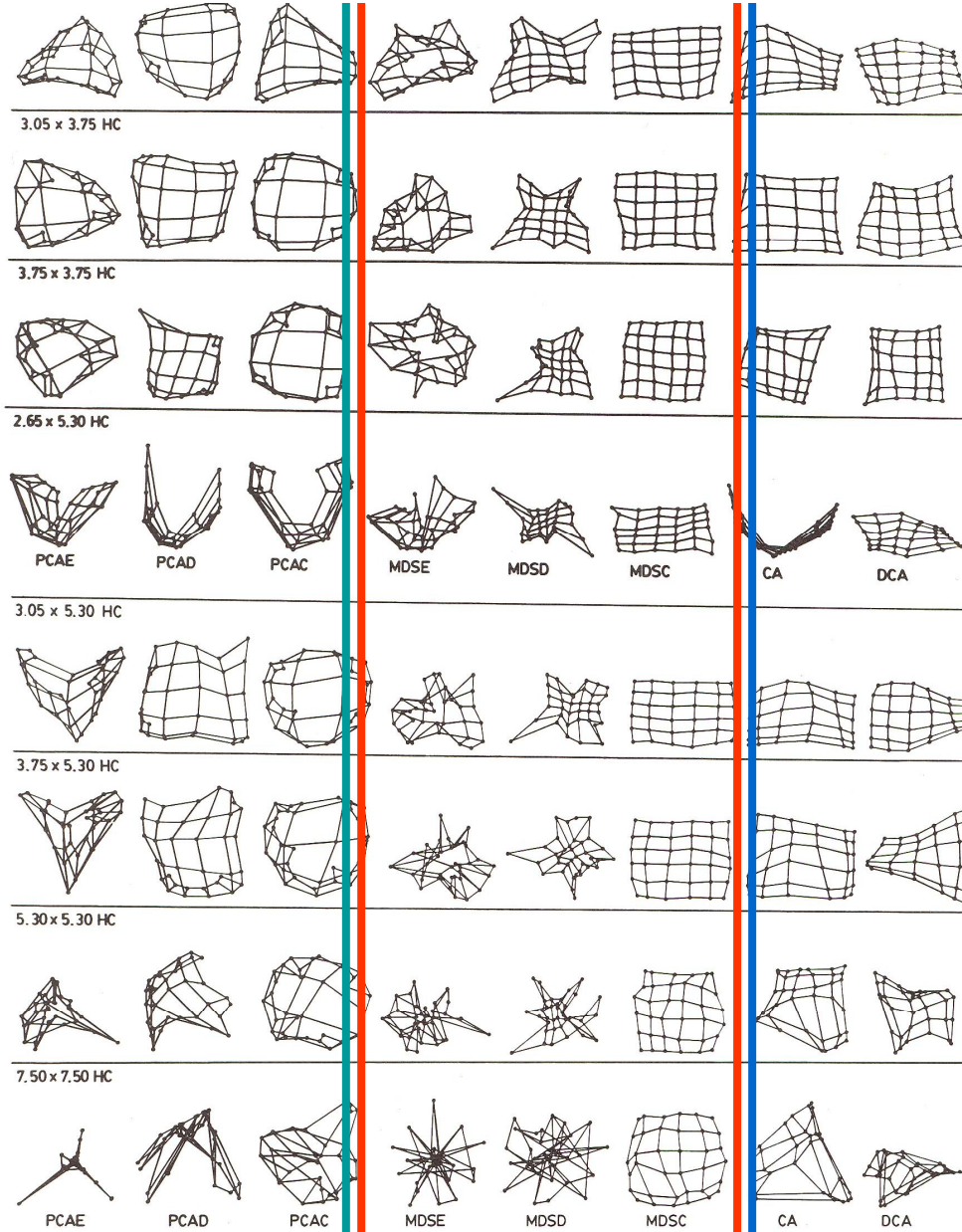
PCA
ordinations
(various
data
standardiza
tions)



Applying Metric and Nonmetric Multidimensional Scaling to Ecological Studies: Some New Results

N. C. Kenkel and L. Orlóci

Ecology, Vol. 67, No. 4 (Aug., 1986), pp. 919-928



PCA

MDS

CA &
DCA

Increasing
half-changes

SUMMARY : ORDINATION STRATEGY

1. Data transformation.
2. Standardization of variables and/or sampling units.
3. Selection of ordination method.

CHOICES AT STEPS 1 and 2 ARE AS IMPORTANT AS CHOICE AT STEP 3.

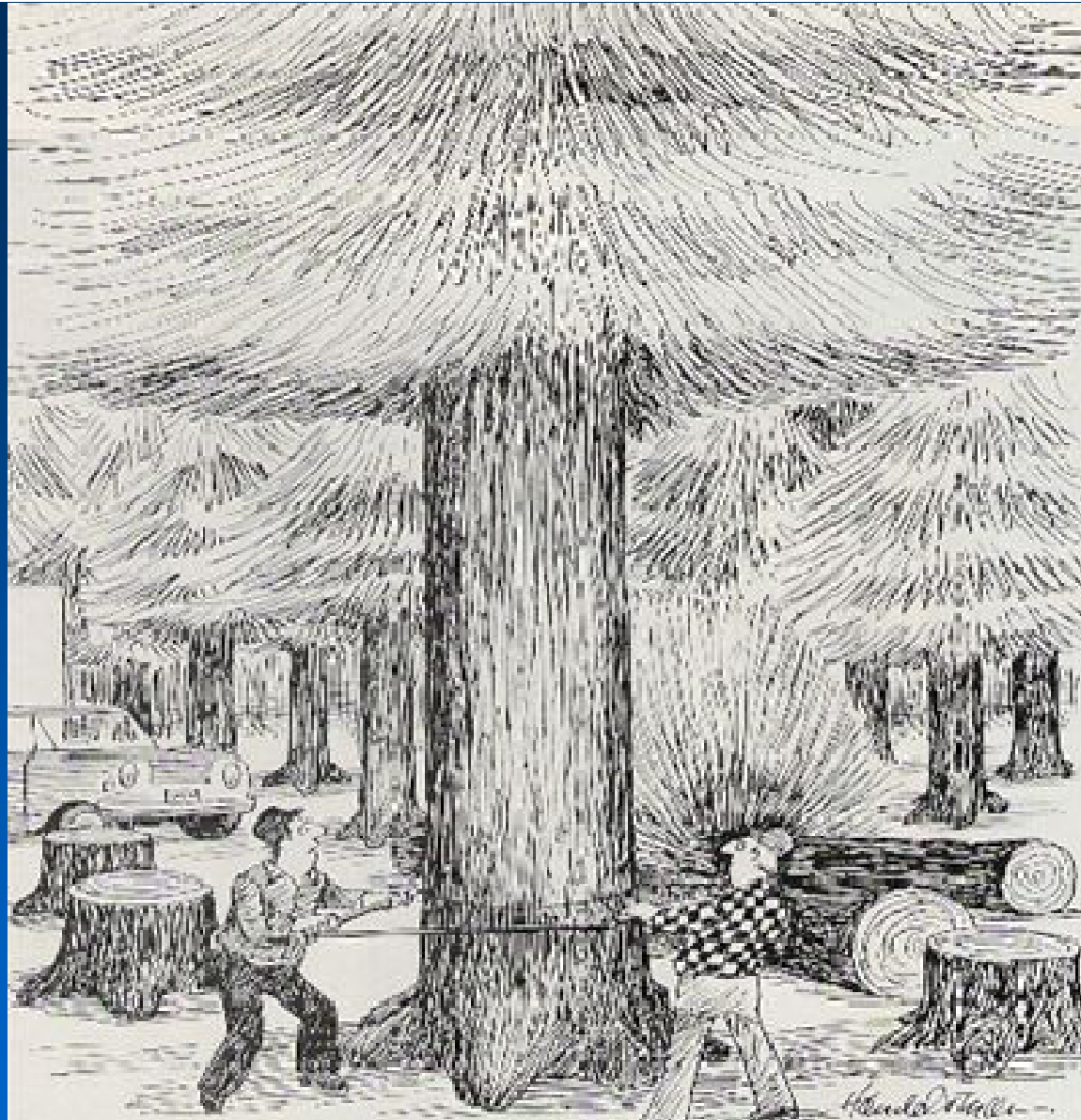
SUMMARY: ORDINATION RECOMMENDATIONS

- Abiotic (environment) survey data:
 - Principal Component Analysis.
 - Standardize variables to “z-scores” (correlation).
 - Log-transform data (continuous variables).
- Biotic (species) survey data:
 - Principal Component Analysis.
 - Do not standardize variables.
 - Log-transform data (continuous variables).
 - Examine results carefully for evidence of unimodal species responses. If so, try correspondence analysis (CA) but be aware that infrequent species may dominate.

NON-METRIC MULTIDIMENSIONAL SCALING

also good but...

- Limitations:
 - Iterative method: solution is not unique and may be sub-optimal or degenerate.
 - Ordination axes merely define a coordinate system: order and direction are meaningless concepts.
 - Variable weights (biplot scores) are not produced.
 - Ordination configuration is based only on ranks, not absolutes.
 - User must choose distance measure, and solution is highly dependent on measure chosen.



"That's life. You stand straight and tall and proud for a thousand years and the next thing you know, you're junk mail."