Reliability of Ordination Analyses

- *Objectives:*

Discuss Reliability

Define Consistency and Accuracy

Discuss Validation Methods

Opening Thoughts

- Inference Space:

  What is it ?

  Inference space can be defined in many ways, but is generally described as the limits to how broadly the particular model / test results apply.

  Defines the extent to which generalizations can be made from data.

  The region (of species – habitat association or variable space) that has been investigated and characterized in a statistical model.

  How is it defined ?

  - By sampling: defined by the sampling universe or the population.

  - By the analysis: defined by the largest entity that is described.

Why Ordination ?

- Data Exploration:

The primary goal of ordination considered "exploratory".

Ordination often considered as much an art as a science.

Traditionally, it was the job of the ecologists to use their knowledge and intuition to collect and interpret data;

pure objectivity could potentially interfere with ability to find important gradients – with inadequate sampling design

(Gauch 1982).

- Hypothesis Testing:

Rigorous hypothesis testing requires complete objectivity, and provides result repeatability and falsifiability.

Why Ordination – Exploration

- According to: Gauch, H.G., Jr. 1982. Noise reduction by eigenvalue ordinations. Ecology 63:1643-1649.

"Ordination primarily endeavors to represent sample and species relationships as faithfully as possible in low-dimensional space"

Thus, ordination is a 'noise reduction technique'.

Ecological significance:

high degree of observed variance (pattern) explained

Statistical significance:

strength of observed patterns versus expected patterns

Why Ordination – Exploration

- How can we validate exploratory ordination ?

Criteria: Variance Explained (Sensitive to 3 manipulations)

- Cross-correlations of the variables (transformations)
- Species / Samples Analyzed (outliers)
- Units of variables (scaling / relativizing)

Why Ordination – Exploration

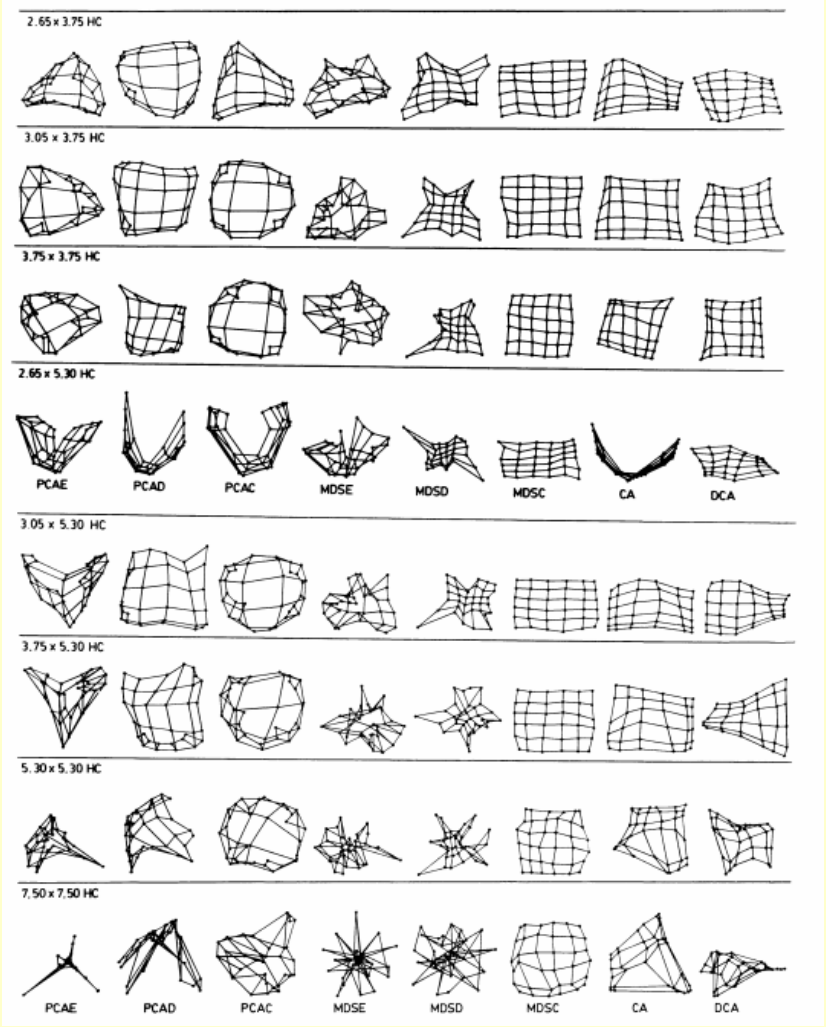
- Validating exploratory ordination

Method	Standardization or version
Principal coordinates analysis (P-Co-A)	unstandardized (PCAE) simultaneous double standardization (PCAD) stand norm standardization (PCAC)
Nonmetric multidimensional scaling (NMDS)	unstandardized (MDSE) simultaneous double standardization (MDSD) stand norm standardization (MDSC)
Correspondence analysis (CA)	unmodified (CA) detrended (DCA)

(Kenkel and Orloci 1986)

Why Ordination – Exploration

- Validating exploratory ordination



(Kenkel and Orloci 1986)

- Save and compare scores quantitatively

Graph - GRAPHROW.GPH - GRAPHCOL.GPH

```
Moss_NMDS_Sorensen
   504 points
111            0.51156      0.13811
112            0.90813      0.17990
113            0.71843      0.13485
114            0.71841      0.13491
121            0.90813      0.17990
122            0.90813      0.17990
123            0.90813      0.17990
124            0.90813      0.17990
```

```
Moss_NMDS_Sorensen
    50 points
Ancu          -0.03153     -0.92260
Clad          -0.18735      0.04084
Cloc          -1.45200     -0.12784
Clcr          -0.79254      0.62354
Deab          -1.24571      1.43251
Difu          -0.09498      0.02934
Disc          -0.29512      0.18033
Dita           0.14181     -1.70019
```

Why Ordination – Comparing Scores

- Distance measure for first and second matrix = Euclidean

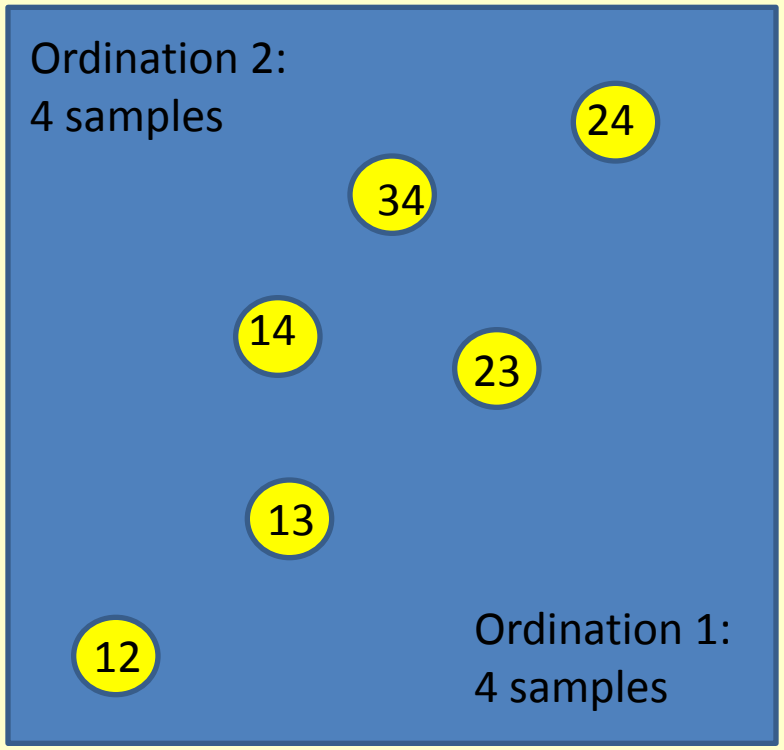
REDUNDANCY OF TWO SETS OF ORDINATION SCORES:

- R-squared:

Squared standardized Mantel statistic (compares ordination distances among all object pairs)

- Percent redundancy:

Shared variance between both matrices (100% * R-squared)

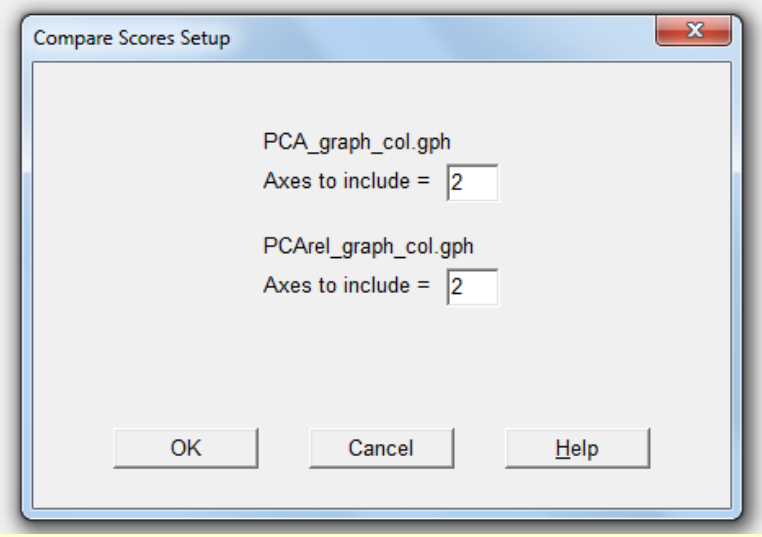


Reconciling Different Ordinations

- Analyze the same dataset (number and identity of objects) is analyzed using multiple approaches, we can compare the resulting ordinations quantitatively:

- Columns Coordinates (species)
- Row Coordinates (samples)

- NOTE: select number of axes for analysis

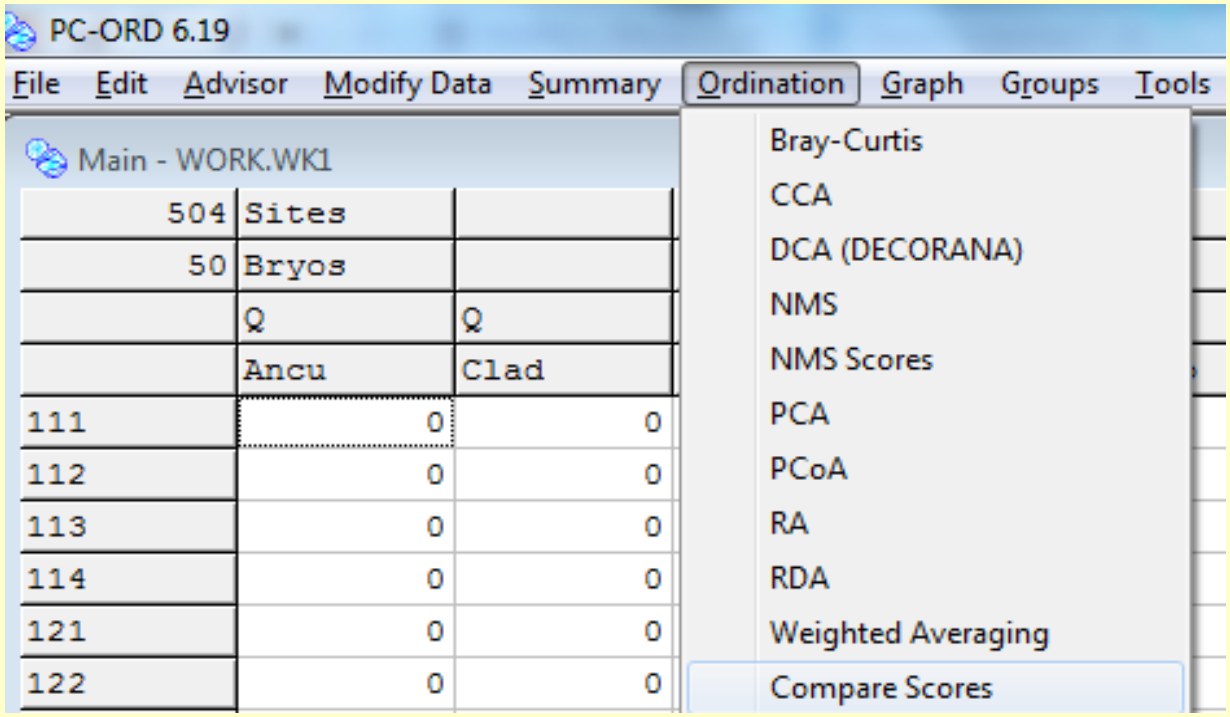


Reconciling Different Ordinations

- Scores can be compared for different objects:

  Samples with samples OR species with species

- Comparing Scores in PC-Ord



Reconciling Different Ordinations

- STATISTICAL TEST OF ORDINATION SCORES:

(Null Hypothesis: no association between matrices)

Mantel`s asymptotic approximation method

+ 0.70 = r = Mantel statistic

Z = (Obs r – Exp r) / (SD of r)

What is the expected value of r ? (given null hypothesis)

Reconciling Different Ordinations

- STATISTICAL TEST OF ORDINATION SCORES:

  - 2.33E+05 = Observed Z
  - 1.83E+05 = Expected Z
  - 1.72E+03 = Standard error of Z
  - 2.89E+01 = t (P < 0.000001)

obs	exp	diff
2.33E+05	1.83E+05	4.98E+04

se		t
1.72E+03		2.89E+01

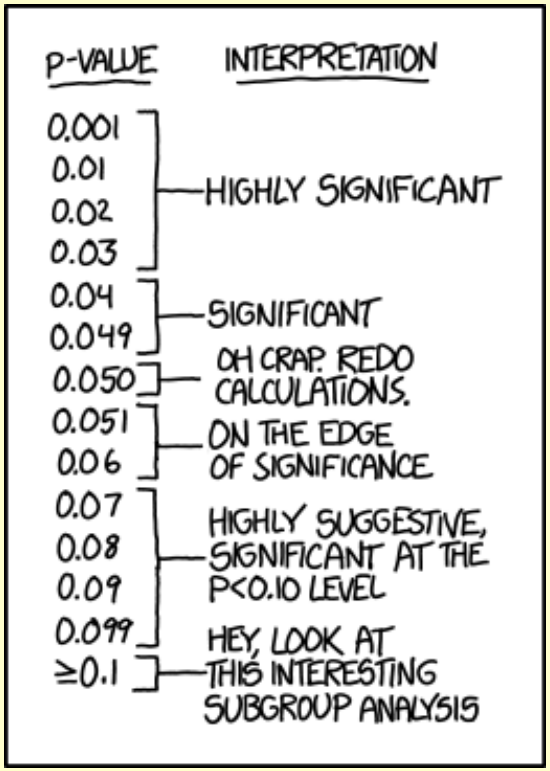
t-distribution with infinite degrees of freedom

- using asymptotic approximation of Mantel (1967).

- If t < 0, then negative association is indicated.
- If t > 0, then positive association is indicated.

Why Ordination – Hypothesis Testing

- Free Ordination: strength of a hypothesized gradient (defined by 2 poles) relative to other empirical gradients
- Constrained Ordination: biological patterns defined by cookie-cutter environmental relationships
- Limitations of Hypothesis testing:
  - Arbitrary nature of p value
  - Testing many patterns at once



Why Ordination – Hypothesis Testing

- How can we validate hypothesis testing ordination ?

Why Ordination – Hypothesis Testing

- Checking type-I Error Rate

Definition:

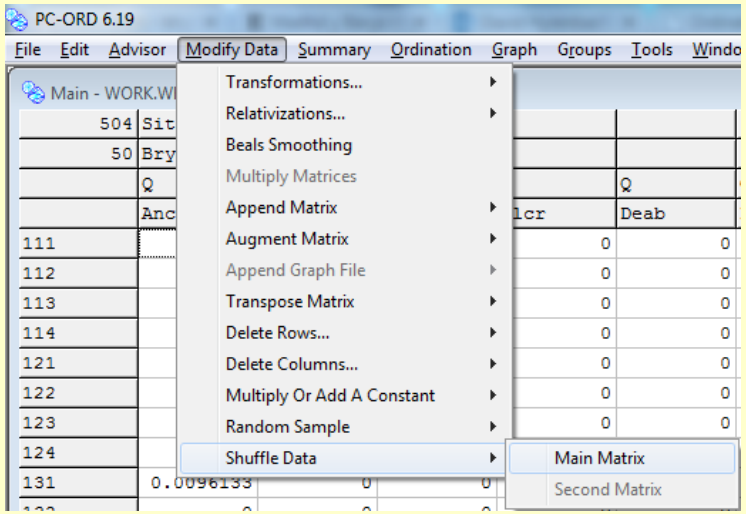
a **type I error** is the incorrect rejection of a true null hypothesis (a "false positive")

Approach:

generate random pattern and determine chance of finding significant results

Remember: Two types of tests (within / across matrices)

- intra-matrix structure (shuffle matrix rows)
- inter-matrix relationship (shuffle one matrix)



Ordination Reliability – Motivation

- When we interpret ordinations, we assume that the configurations define a definite structure – a relationship between samples and species.

- If the study was replicated, the answer should be the same. Obviously, it would not be exactly the same, but how different would it be? Depends on the "multivariate landscape".

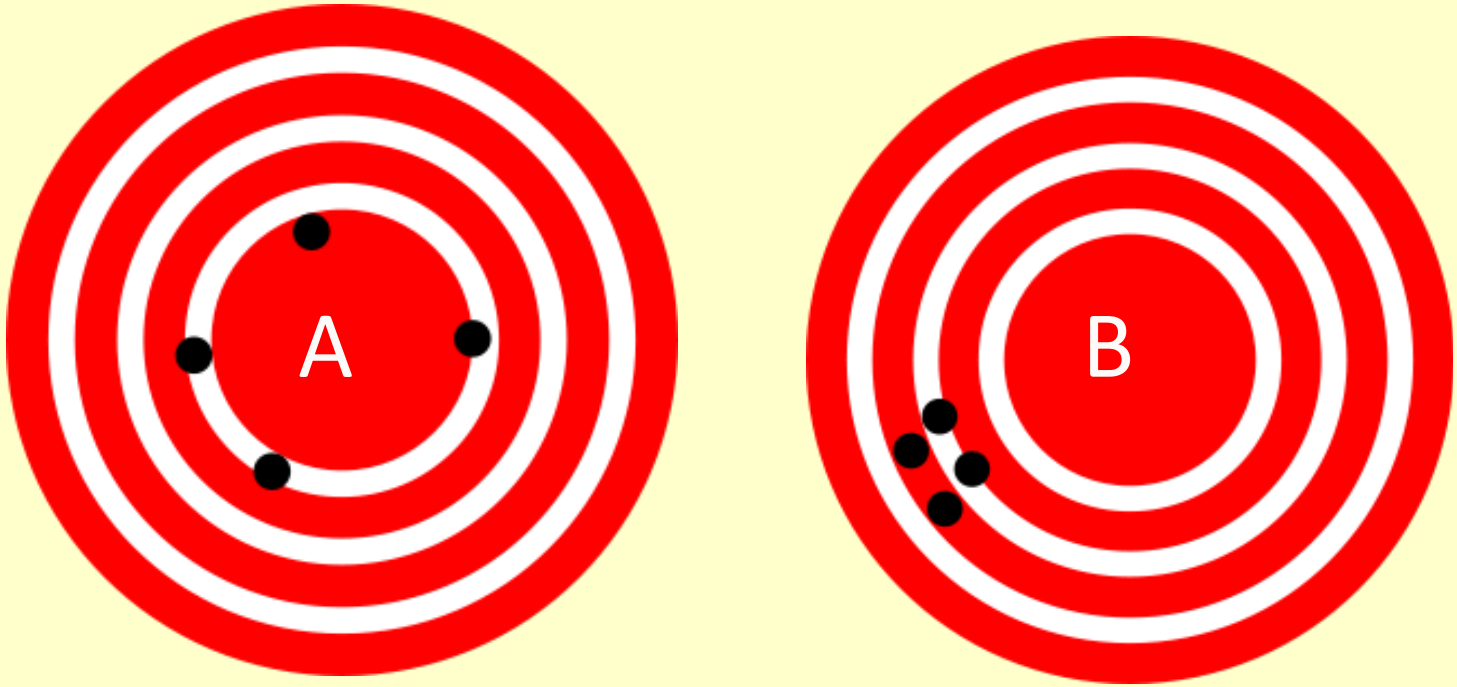
- Methods for evaluating reliability (robustness) of ordinations look for consistency in results across options / methods.

- When same general patterns emerge, user gains confidence.

- However, other more quantitative methods are available.

Consistency / Accuracy

- Consistency: Different methods give same answer
- Accuracy: Results approximate the real pattern



Ordination Reliability – Approaches

- There are several sources of discrepancies:

  Measurement error
  Sampling error
  Methodological artifacts
  Ecological inconsistencies

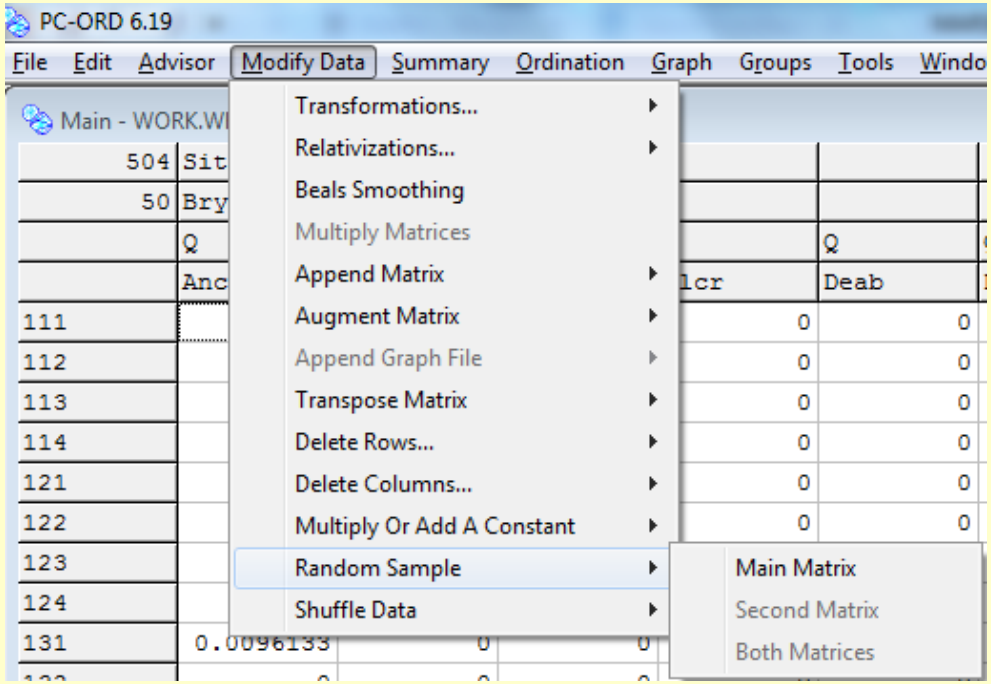
- There are several quantitative methods at our disposal:

  Strength of pattern compared to null model
  Proportion of total variance explained
  Consistency of the pattern across data subsets
  Accuracy of results, with respect to underlying pattern

Validation of Ordination - Approach

- Divide dataset into two: use one to predict the other.

- Develop model with "training dataset"
- Use relationship to predict response
- Compare response with "validation" dataset



Approach – Validation of Ordination

- This approach provides two “independent” analyses of the patterns, using two separate datasets

- What metrics could we use to perform this validation?

- This approach provides a second perspective – how well did one dataset predict the other dataset?

- But, what if the answer depends on the sample size?

Approach – Validation of Ordination

- Partitioning dataset into many subsamples via bootstrapping

- Create many datasets
(smaller than original dataset)

- Calculate statistics of similarity
(across all individual answers)

- Compare against “full” dataset
(how well did the individual bootstrapped answers do compared to the original answer from full dataset?)

Approach - Validation

Journal of Vegetation Science 10: 895-902, 1999

The bootstrapped ordination re-examined

Pillar, Valério DePatta

Abstract. A method is described to determine the number of significant dimensions in metric ordination of a sample. The method is probabilistic, based on bootstrap resampling. An iterative algorithm takes bootstrap samples with replacement from the sample. It finds in each bootstrap sample ordination coordinates and computes, after Procrustean adjustments, the correlation between observed and bootstrap ordination scores. It compares this correlation to the same parameter generated in a parallel bootstrapped ordination of randomly permuted data, which upon many iterations will generate a probability. The method is assessed in principal coordinates analysis of simulated data sets that have varying number of variables and correlation levels, uniform or patterned correlation structure.

Approach - Validation

- Partitioning dataset into many subsamples (Pillar 1999)

1. Save ordination scores for *k* axes from the complete data set ($n \times p$). Call the $n \times k$ scores the original ordination.
2. Draw a bootstrapped sample of size *n*. Ordinate the sample.
3. Perform rotation of the *k* axes from the bootstrapped ordination, maximizing alignment with the original ordination.
4. Calculate correlation coefficient between the original and bootstrapped ordination scores, saving a separate coefficient for each axis. The higher the correlation, the better the agreement between scores for full data set and the bootstrap.

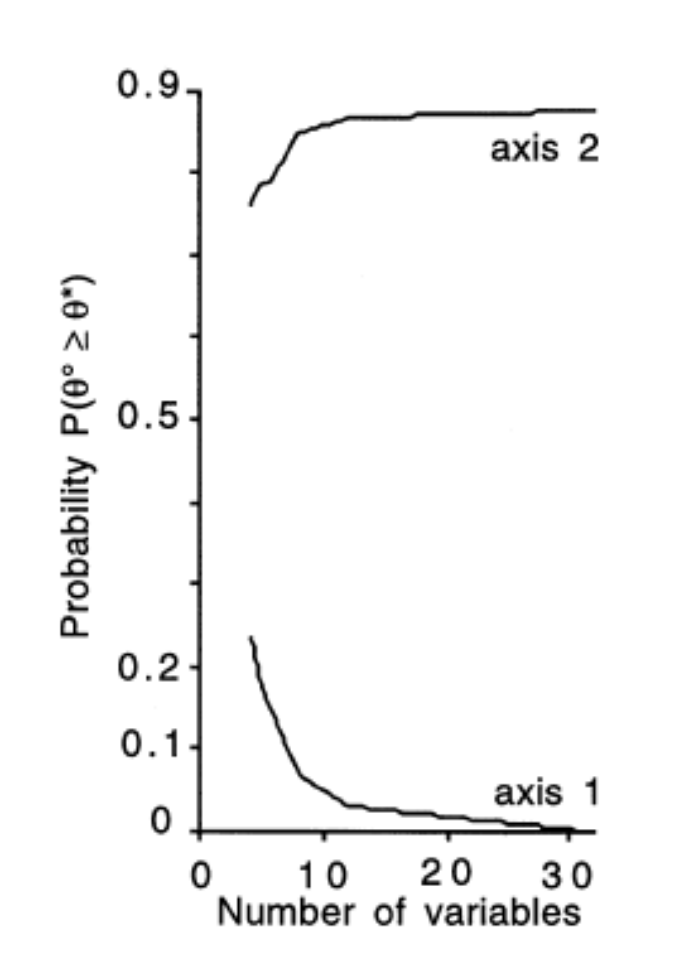
Approach - Validation

- Approach: (Pillar 1999)

5. Repeat steps 1-4 for randomization of original dataset. Elements of complete dataset randomly permuted within columns.
6. For each axis, if correlation coefficient from step 4 for randomized dataset is greater than or equal to the correlation coefficient from the nonrandomized data set, then increment a frequency counter, $F = F + 1$.
7. Repeat steps above many times (B = 40 or more).
8. For null hypothesis that the ordination structure of the dataset is not stronger than expected by chance, calculate a probability of type I error: $p = F/B$

Approach - Validation

- Example: (Pillar 1999)

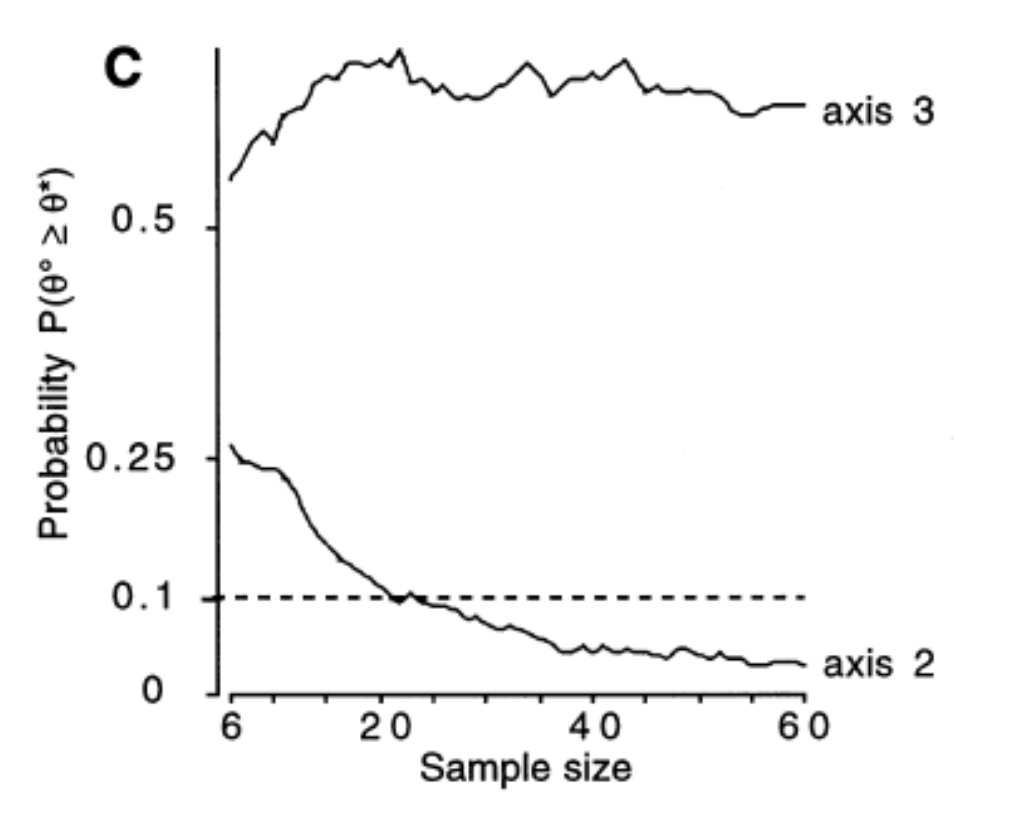


Investigating the influence of the number of variables on the significance of ordination axes.

From bootstrapping of simulated datasets containing 40 sampling units and 32 variables (from 4 to 32)

Approach - Validation

- Example: (Pillar 1999)



Investigating the influence of the number of samples on the significance of ordination axes

From bootstrapping of simulated datasets containing 6 to 60 sampling units and 20 environmental variables

Approach – Validation of Sampling

ON SAMPLE SIZE OPTIMALITY IN ECOSYSTEM SURVEY

L. ORLÓCI and V. De PATTA PILLAR

SUMMARY

Sampling is discussed as a process. In this the sample structure evolves and attains increasing stability as the sample size n increases. The minimum n at which the sample structure begins to attain stability is suggested as a lower bound for optimal sample size in ecosystem survey.

Approach – Validation of Sampling

- Determining the adequacy of sampling (Orloci & Pillar 1989)

- Mapping distance between sample pairs:

  - distance on the basis of species abundances
  - distance on the basis of environmental data

An n x n symmetric matrix **D** of quadrat distances defines vegetation structure based on s species as variables. A second n x n symmetric matrix Δ of quadrat distances defines another structure based on t environmental variables. The similarity of the **D** and Δ configurations is a measure of the two structures' affinity.

Approach – Validation of Sampling

- Stress compares the structure from species / environment:

The relationship of **D** and Δ evolves as sample size increases. We monitored this by a stress function

$$\sigma_{VE} = \sqrt{1 - \rho^2(\mathbf{D};\Delta)} \qquad (2)$$

in which $\rho(\mathbf{D};\Delta)$ is a product moment correlation. Other definitions are possible (e.g., SHEPARD and CARROLL, 1966).

P = correlation of pairwise distances from species OR environment

If p = +1, then sigma = ?

If p = 0, then sigma = ?

If p = -1, then sigma = ?

Approach – Validation of Sampling

- How should the stress behave as we sample more and develop a more stable description of the community structure?

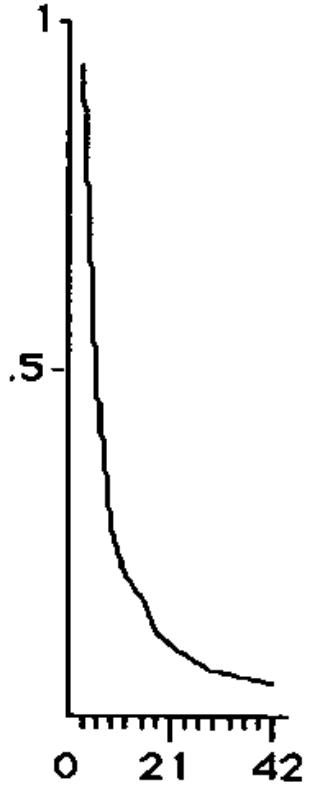


Figure 1: Changing relative stress (vertical scale, Eq. 2), comparing vegetation structure **D** and environmental structure Δ at increasing sample size (n, horizontal scale) in random sampling on the Elk Lake site. Although random sampling continues until 42 quadrats, structural stability is demonstrated already at about n = 18.

Validation - Conclusions

- The number of variables influences the ability to identify significant axes of variability in ordinations.

- The number of sampling units influences the ability to identify significant axes of variability in ordinations.

- Validation of model results using subsampling and bootstrapping provides insights into reliability of the results and their dependency on sample size

Validation - References

- Pillar, V.D. 1999. The bootstrapped ordination reexamined. *Journal of Vegetation Science* 10: 895-902.

- Orlóci, L. & V.D. Pillar. 1989. On sample size optimality in ecosystem survey. *Biometrie-Praximetrie* 29: 173-184.

Overview of Ordination

GOAL: You want to explore gradients (of samples / species)

- Quantify Species / Variable Responses to Gradients
  YES – keep going
- Are there linear relationships in the main matrix ?
  YES – PCA NO – keep going
- If not linear, do you want to test a single gradient ?
  YES – keep going NO – NMDS
- If single gradient, do you want to use a priori weights?
  YES – WA NO – keep going
- Is chi-square distance measure OK?
  YES – CA NO – Polar Ordination

What Else Can We Do ?

GOAL: You want to compare groupings (of samples / species)

- Compare group composition – defined by 1 factor
- Compare group composition – defined by 2 or more factors
- Compare groups – designed experiments (nested / factorial)
- Identify species – indicative of certain groups