Alma Mater Studiorum – Università di Bologna
Facoltà di Scienze Matematiche, Fisiche e Naturali

Laurea Magistrale in Biologia Marina

Insegnamento di

RICERCA SCIENTIFICA SUBACQUEA

A.A. 2001/2011

Prof. Marco Abbiati - Dr. Massimo Ponti

Ecology

The science that explore the interaction among organisms and between them and the physical environment… the collection of biotic and abiotic processes that maintain life on Earth (Krebs, 1994)

These processes are sensitive to changes in pattern of distribution, abundance and diversity of species

Heterogeneity and variability

Natural systems are extremely complex, they are regulated by a network of biotic and abiotic relationship, which vary in space and time and at different observation scales

Aims in ecological studies

Description and characterization of species distribution (biodiversity and biogeography)

Test biological and ecological theories (population dynamics, species interaction, etc.) in order to understand natural events

Monitoring and assessment of natural changes and anthropic disturbance in order to preserve biodiversity and guarantee sustainable development

"Community"

All the species living in a habitat, interacting among them and with the environment

"Assemblages"

Group of organism living in a selected area and studied together

Macrobenthic invertebrates… in ecological studies

- Reduced mobility
- Rapid respond to both natural and anthropogenic disturbances
- Long life span
- Different range of sensitivity/tolerance
- Crucial role in nutrient and other chemical cycle between bottom and sea water column
- Important (direct or indirect) commercial or recreational role

Measurable biotic variables and derived indices

- Presence/absence
- Abundance (density, cover area)
- Biomass
- Production (somatic growth)

Measurable biotic variables and derived indices

- Species diversity indices
  - Species richness
  - Evenness
  - Heterogeneity
- Trophic indices
- Size and biomass spectra
- Abundance/Biomass relationship
- Biotic indices based on sensitivity/tolerance
- Similarity indices in order to compare “assemblages”

- **Species diversity indices (how many species and individuals?)**
  - **Species richness = S**
  - **Evenness e.g.: Pielou index (J)**
  - **Heterogeneity or species diversity e.g.: Shannon index (H')**

$$H' = -\sum_{i=1}^{s} p_i \log_2 p_i$$

$$J' = \frac{H'}{H'_{\max}}, \quad J' = \frac{H'}{Log_2 S}$$

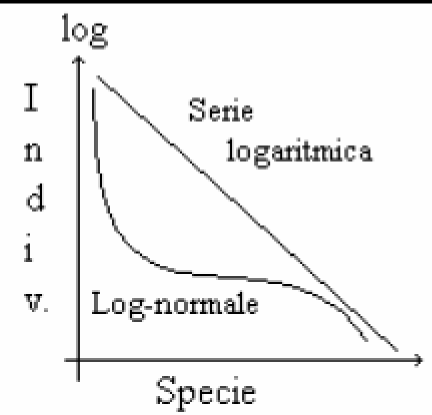


Fig. 2.C - Grafico di Whittaker per vedere il tipo di andamento seguito da una comunità.

Univariate analyses

- **Comparison between means**: t-test, Analysis of Variance **(ANOVA)**
- **Correlation**: relation between two variables
- **Regression**: an independent variable influence a dependent variable

Multivariate analyses

(*based on similarity indices*)

- **Classification techniques (Cluster)**
- **Ordination methods (PCA, nMDS, PCO, CAP)**
- **Test of hypothesis (ANOSIM, PERMANOVA)**
- **Contribution of single species to similarity (SIMPER)**
- **Correlation between biotic and abiotic variables (BIOENV, DISTLM)**

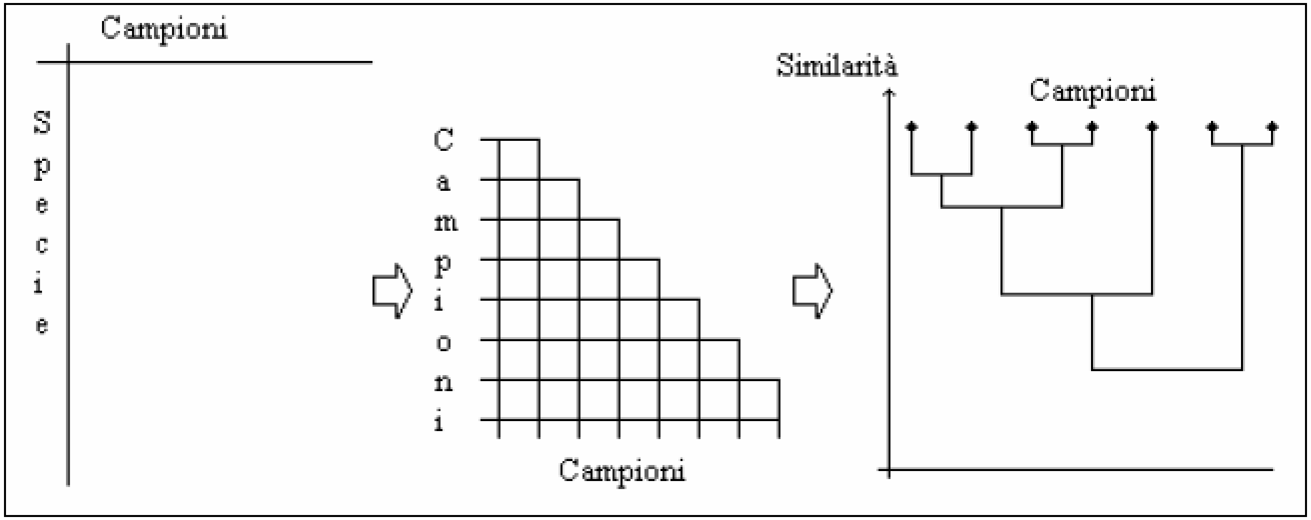


Fig. 2.B - Stadi successivi nell'*analisi dei cluster*: matrice bruta, matrice triangolare di similarità, dendrogramma.

Trasformazione	Formula	Commento
Radice quadrata	$x' = \sqrt{x}$	trasformazione debole
Doppia radice quadrata	$x' = \sqrt[4]{x}$	trasformazione media
Logaritmica	$x' = \ln(x+1)$	trasformazione forte
Presenza / Assenza	0-1	trasformazione estrema
Standardizzazione percentuale	$x' = \frac{x}{\sum x_i} \cdot 100$	per ridurre l'effetto della taglia del campione

Tab. 2.D - Esempi di trasformazioni dei dati: da operarsi quando vi è almeno un ordine di grandezza fra l'abbondanza delle specie più abbondanti e quelle più rare. La standardizzazione invece serve nel caso di campioni di taglia differente.

Indice	Formula	Commento
Jaccard	$S_J = \frac{a}{a+b+c}$	non tiene conto delle doppie assenze
Sorensen	$S_S = \frac{2a}{2a+b+c}$	non tiene conto delle doppie assenze e da molto peso alle comunanze
Incontro semplice	$S_{SM} = \frac{a+d}{a+b+c+d}$	utilizzabile solo se si conoscono i reali assenti
Baroni-Urbani & Buser	$S_B = \frac{\sqrt{ad}+a}{a+b+c+\sqrt{ad}}$	utilizzabile solo se si conoscono i reali assenti

Tab. 2.B - Indici di similarità fra due campioni con dati binari: a: n° specie presenti in entrambi i campioni; b: n° specie presenti solo nel campione A; n° specie presenti solo nel campione B; n° specie assenti in entrambi i campioni.

Indice	Formula	Commento
Distanza euclidea	$\Delta_{JK} = \sqrt{\sum (x_{ij} - x_{ik})^2}$	di scomodo utilizzo perché i valori vanno da 0 a ∞, inoltre aumenta col numero di specie presenti
Distanza euclidea media	$d_{JK} = \sqrt{\frac{\Delta_{JK}^2}{n}}$	come sopra ma non aumenta col numero di specie presenti
Manhattan	$d_M = \sum \lvert x_{ij} - x_{ik} \rvert$	è una delle misure più semplici
Bray - Curtis	$B = \frac{\sum \lvert x_{ij} - x_{ik} \rvert}{\sum (x_{ij} + x_{ik})}$ $S_{BC} = 1 - B$	è una standardizzazione del Manhattan, il valore va da 0 a 1, esalta però l'effetto delle specie abbondanti e non tiene conto delle doppie assenze; si può usare anche come indice di similarità (S_{BC})
Camberra Metric	$C = \frac{1}{n} \sum_{i=1}^{S} \frac{\lvert x_{ij} - x_{ik} \rvert}{x_{ij} + x_{ik}}$ $S_{CM} = 1 - C$	esalta l'effetto delle specie rare, il valore va da 0 a 1, i valori zero vanno sostituiti con valori minimi (es. 0,01); si può usare anche come indice di similarità (S_{CM})

Tab. 2.C - Indici di distanza fra due campioni. x_{ij}: n° di individui della specie i-esima nel campione A; x_{ik}: n° di individui della specie i-esima nel campione B (le sommatorie sono per specie i).

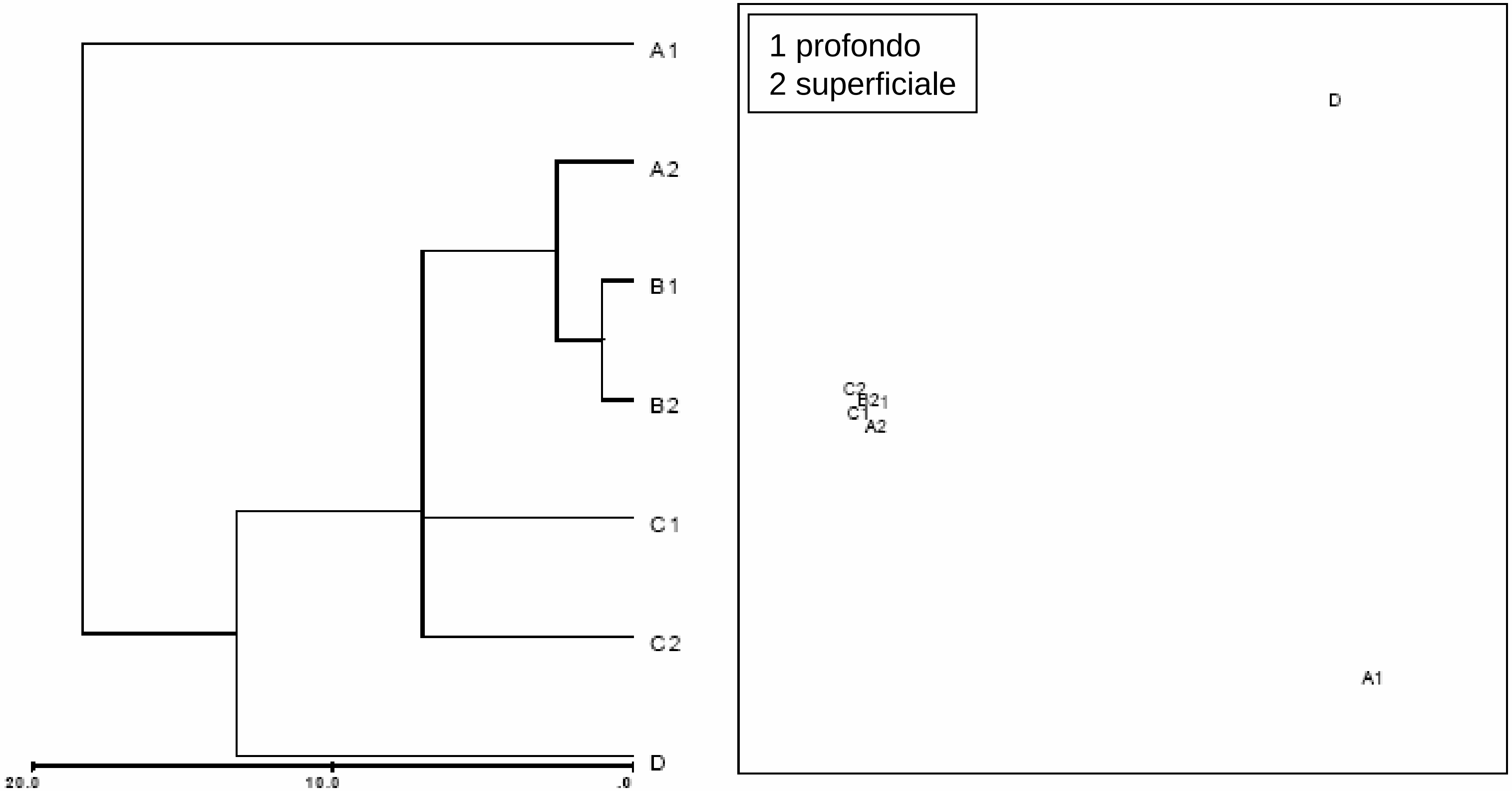


Fig. 3.M - Dendrogramma e MDS non metrico ottenuti dall'analisi dei popolamenti sommati per sito e profondità (standardizzazione percentuale, trasformazione con doppia radice, indice di Bray-Curtis Ranked; stress 0,01).

Stress	interpretazione
< 0,05	rappresentazione eccellente
< 0,10	buon ordinamento, senza false deduzioni
< 0,20	utilizzabile, ma con valori vicini al limite superiore è possibile compiere errori di interpretazione soprattutto dei dettagli
> 0,20	non utilizzabile, possibili errori d'interpretazione

Tab. 2.E - Interpretazione dei valori di stress (Clarke, 1993).

Scientific approach

Hypothetic-deductive approach

Every studies start from the prior knowledge (**literature**) and subjective "field" observation.

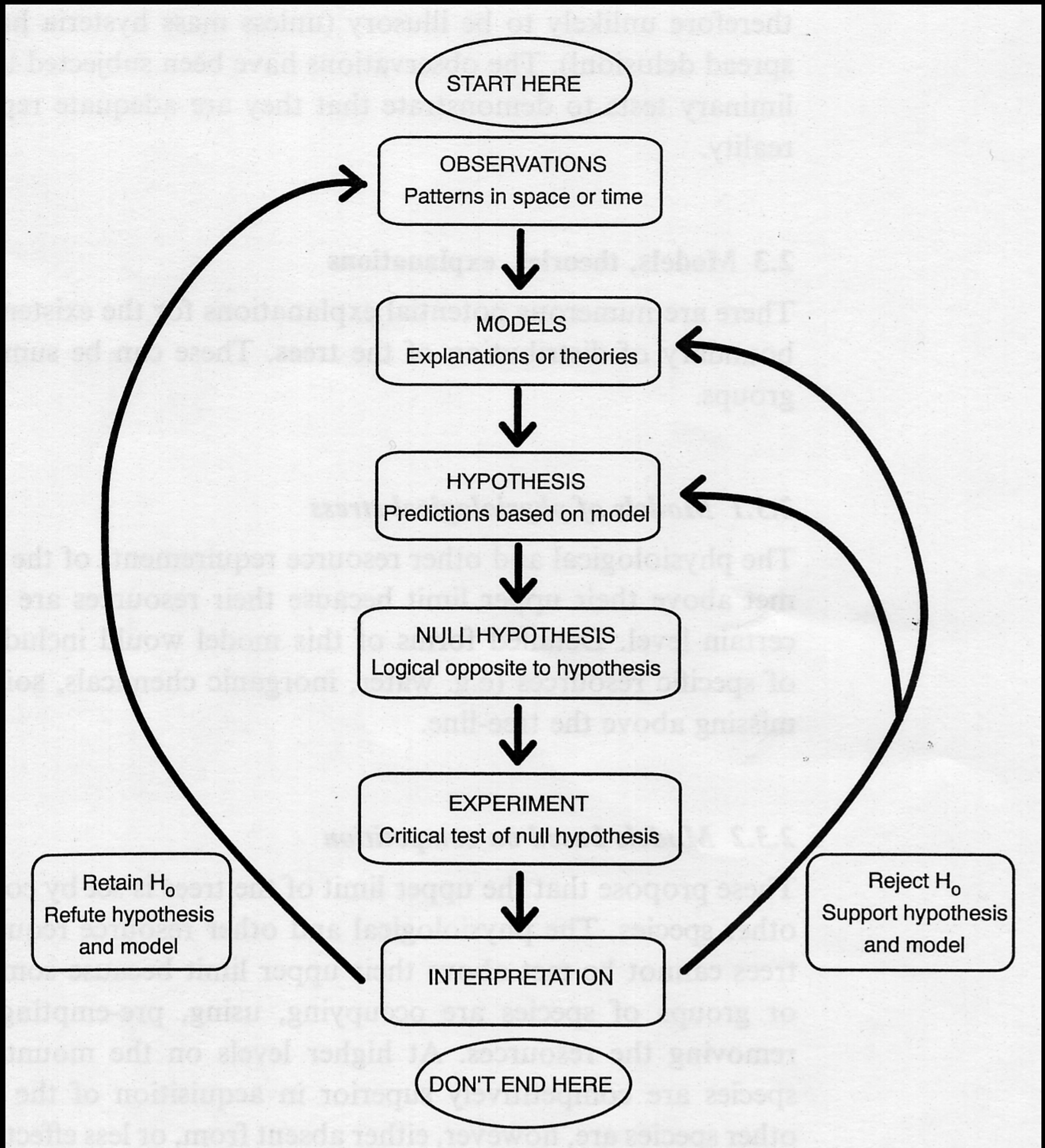
To avoid subjectivity in the interpretation of phenomena and processes under investigation a **formal procedure** is needed!

The formal procedure must to start with a clear and **univocal definition of the specific aims** of the study. Data collected without specific scopes are simple observation and do not allows to analyses any theories or models!

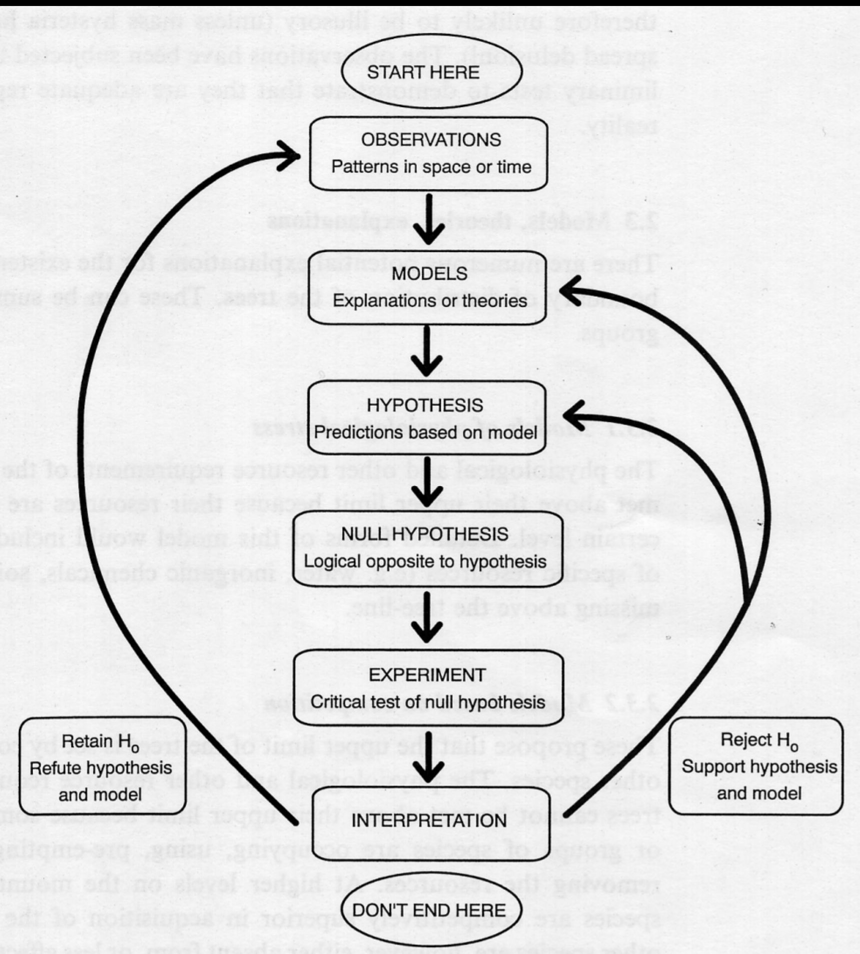
The nature of the data that will lead to acceptance or rejection of a clear defined conjecture must be specified ***before*** and not ***after*** the data have been collected and analyzed.

It is impossible to prove the truth of a theory (unless measuring the whole universe). We can only try to falsify a theory (Popper, 1959).

Hypothetic-deductive approach



Hypothetic-deductive approach



OBSERVATIONS

Every study start from the prior knowledge (**literature**) and <u>subjective</u> “field” observation

MODELS

theory that explain the observation

$H_{1,2,...}$ HYPOTHESIS

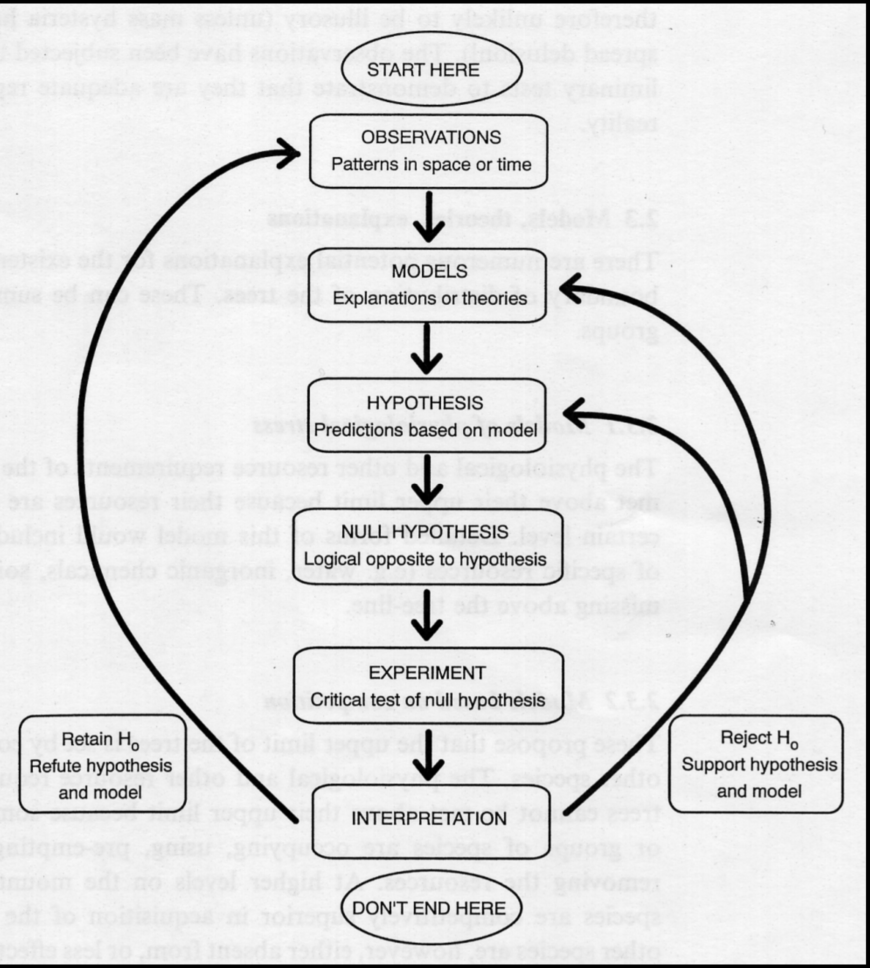
“if what I propose is true, then making this particular change to the system I must observe this specific response …”

… since is impossible to prove the truth, we can try to falsify the opposite

H_0 NULL HYPOTHESIS

Logical opposite… including all the possibilities

Hypothetic-deductive approach



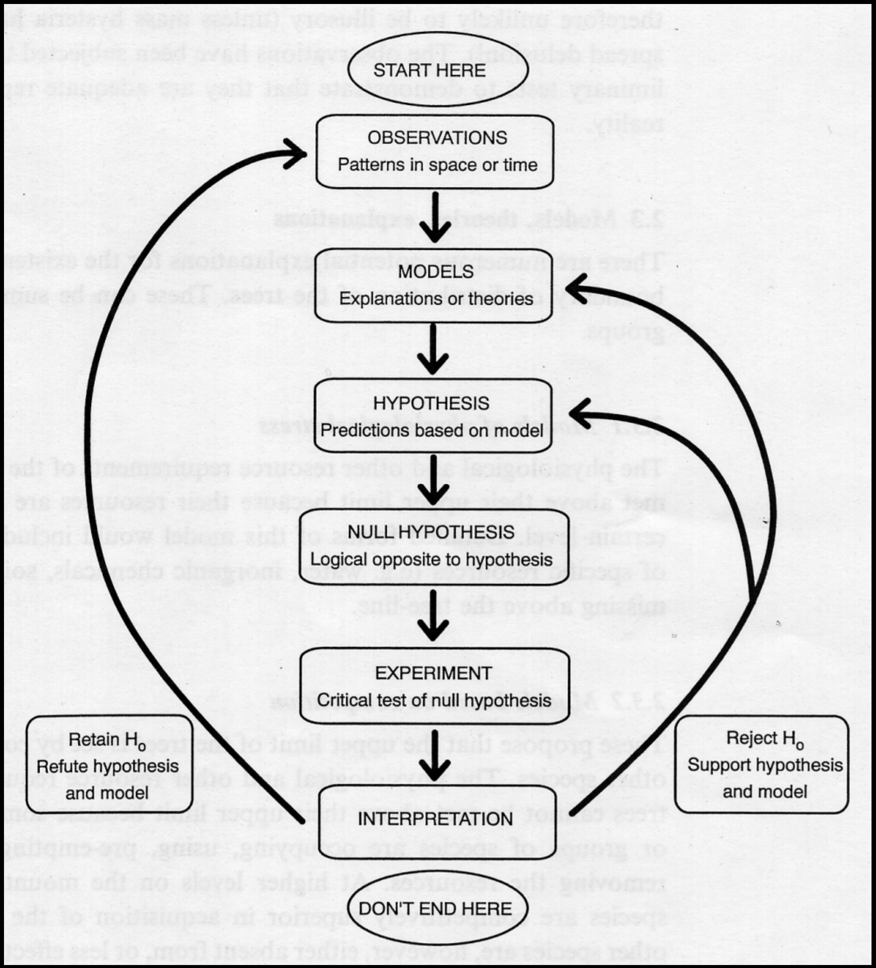
EXPERIMENT

Aimed to try to falsify the null hypothesis…

If we reject the null hypothesis (at *a priori* level of probability, generally 1 or 5%), by definition the original hypothesis remain the only valid alternative. In this case we corroborate the original model… till further evidence…

In fact, rejection of H_0 do not mean that the model is true. It is possible that the model cannot explain the same observation in another spatial and/or temporal context, or that another model explain the same phenomenon more accurately.

Hypothetic-deductive approach



If we retain the null hypothesis (at *a priori* level of probability, generally 1 or 5%), we have to reject the original hypothesis and the model

Don't forget

To retain the logic of any test of an hypothesis requires close attention to issues of replication, independence of data and controls for experimental procedures… these make the design of most experiments necessarily complex

Experimental design

Descriptive approach (*mensurative experiment*)

examine relationship among variables without identifying cause-effect relationship (e.g.: studies on pattern of distribution and abundance of species), furthermore

- **provide the fundamental observation that enable causal inferences**
- **often suggest which processes may be discarded from successive analyses**

Experimental approach (*manipulative experiment*)

examine cause-effect relationship (including controls for artifacts)

Experimental design

Factor: predictive variables (e.g.: depth, exposure, pollution, protection, etc.)

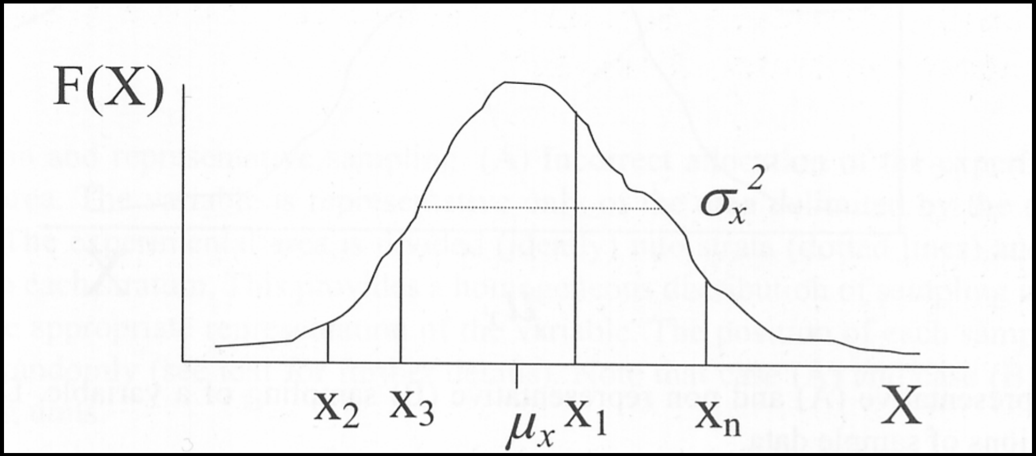
Level (treatment): conditions or intensities of the factor (e.g.: chosen depth, pollution level, etc.)

Experimental design

- the number of factors
- the relationship among factors
- the criteria that determine the choice of levels of a factor
- the number of levels of a factor
- the way of assigning experimental units to factors
- the spatial and temporal distribution of experimental units
- the number of replicates in each factor

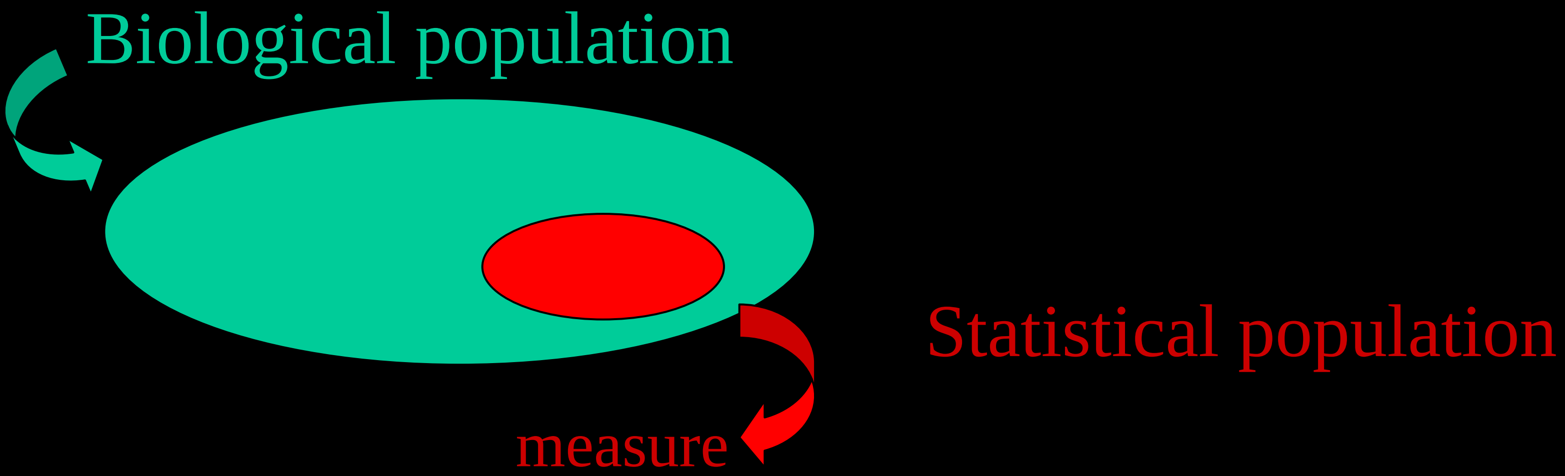
Variable estimation

Since we cannot measure the whole universe of a statistical population (e.g. the number of individual of a species in the study area) we use an estimation of the parameters of the frequency distribution of the study variable:



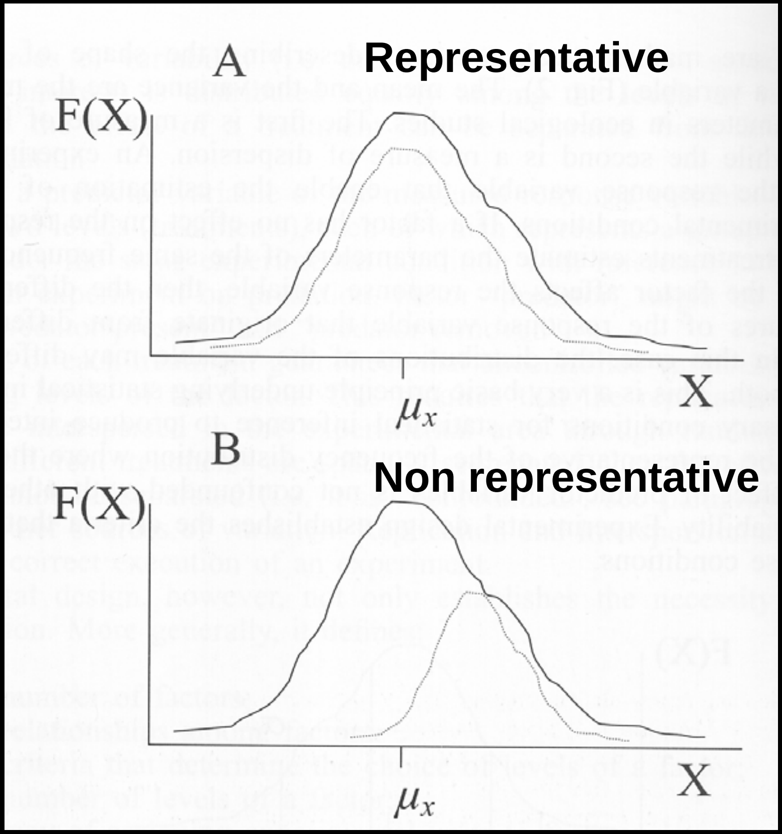
mean (central tendency) (e.g.: mean number of individual per square meter)

variance (dispersion)



Representative sampling

Sampling is representative if the entire "range" of possible values of the variables is reproduced by the sample. This condition is achieved if at the start of the experiment all possible observations have the same probability to be included in the sample.

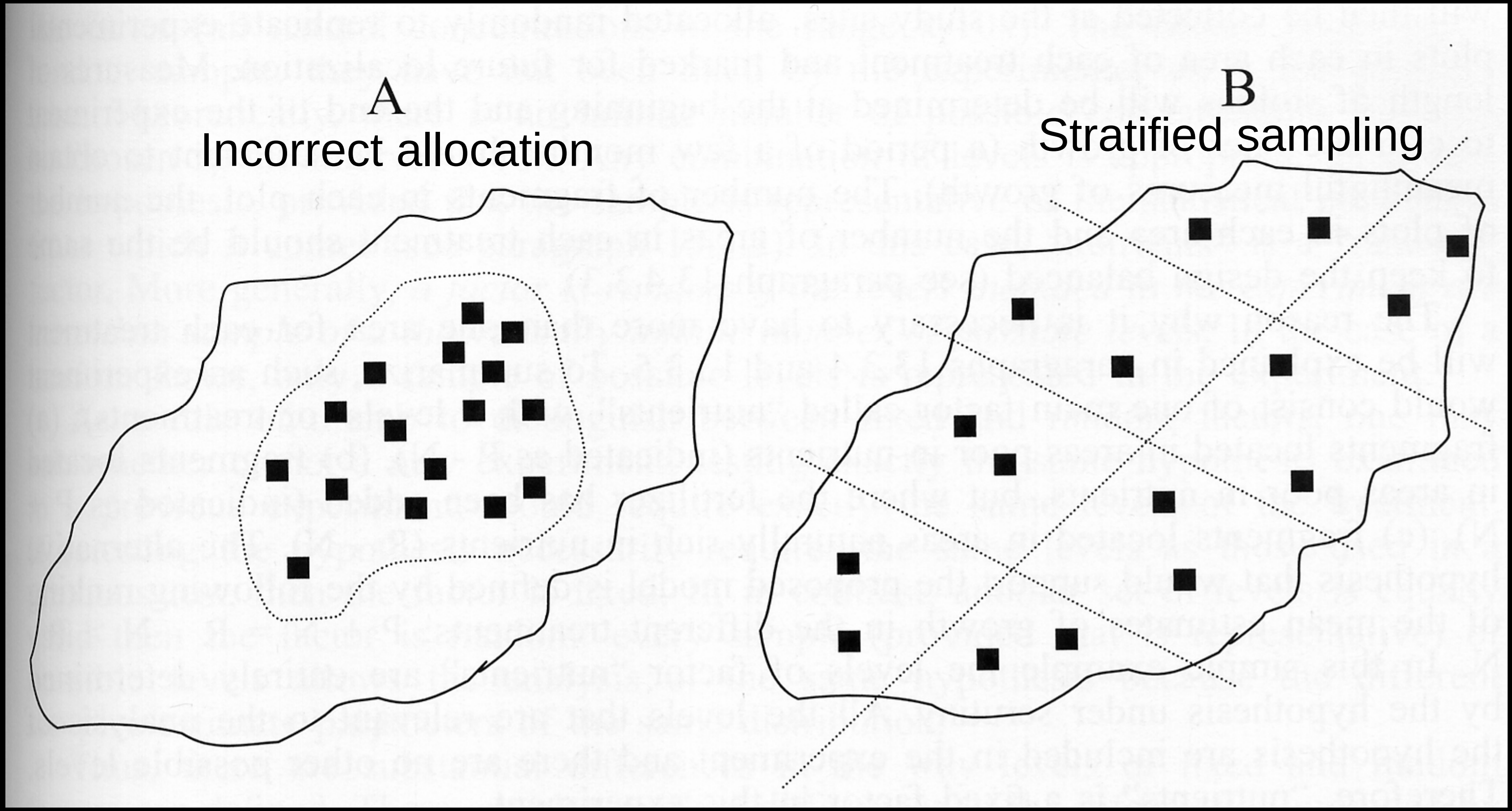


The only way is to collect **random** samples and high number of replicates.

Representative sampling

1. Collect random replicate samples everywhere covering all the area of interest

2. Stratify the samples unit according to an a priori hypothesis about the nature of spatial variation… in order to obtain a homogeneous distribution of the samples in the study area



Fixed and random factors

Whether a factor is fixed or random depends on how the levels of the factor are chosen for the experiment according to the hypothesis to test.

Fixed factors

when all the levels that are relevant for a test of the hypothesis are included in the experiment.

1. eventually differences between levels can be analyzed by "*a posteriori*" statistical test
2. results can not be generalized to different levels (not included)

example: pollution (polluted vs. unpolluted sites)

Random factors

if the levels included in the experiment are a random sample of a theoretically infinite number of possible levels

1. eventually differences between levels can NOT be analyzed by "*a posteriori*" statistical test
2. results can be generalized to all possible levels

example: temporal variation (random sampling date)

Complex hypothesis

Orthogonal factors (crossed or factorial design)

if each level of factor A is represented in every level of B and viceversa

interaction among factors

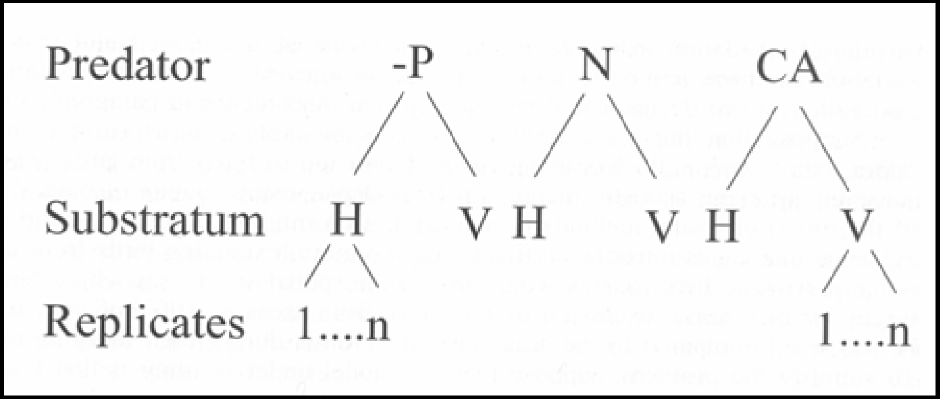
A

B

	Predator +	Predator -
Substratum 1	replicates	replicates
Substratum 2	replicates	replicates

A

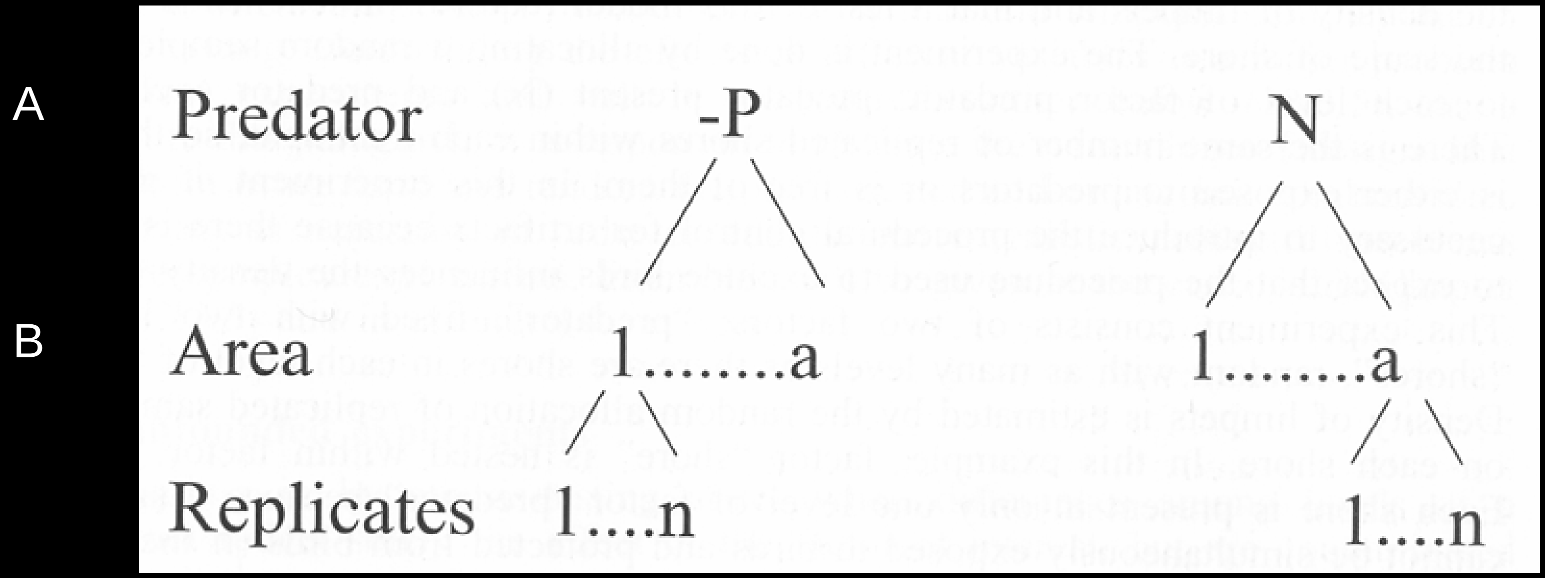
B



Complex hypothesis

Hierarchical (nested) factor

Factor B is nested in A if each level of B is present in a single level of A
Nested factor is always “random”



Confounded experiments

A correct experimental design must to guarantee that all the variability out of control (not associated with the study hypothesis, i.e. not included in the factor) have been equally distributed among the factors. That allows to separate the effect of the factors from the “background noise”

Therefore each level must be randomly replicate and intersperse

Confounded experiments

Spatial confound

e.g.: spatial separation between treatment

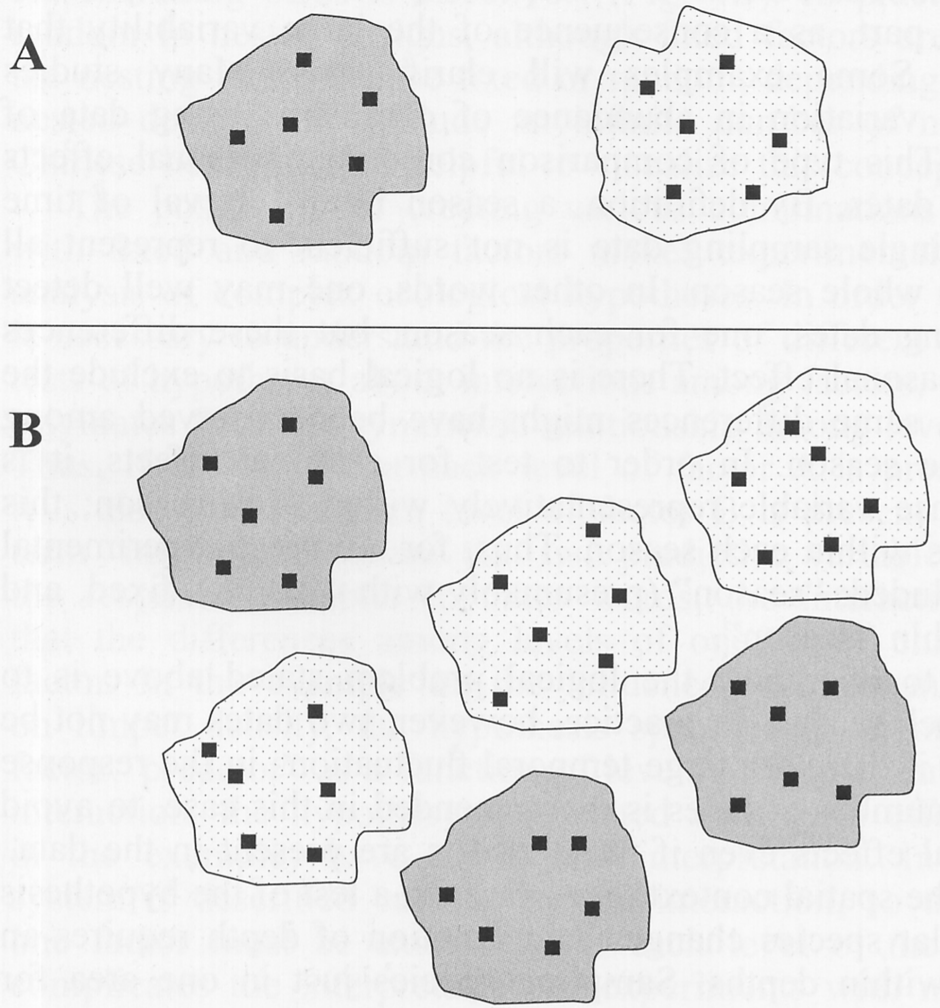


Fig. 6 - Confounded experiments on predation and a solution of the problem.

Predator present;
Predator removed.

(A) Schematic representation of a confounded experiment. The predator is removed from a single area and left at natural density in a second area. The abundance of prey is estimated by sampling each area with *n* replicated sampling units. This design confounds the effect of the predator with the intrinsic variability among areas.

(B) Correct design. There are several areas for each level of predation and *n* sampling units in each area. By replicating areas in each level of predation, it is possible to tease apart the effect of predators from the natural variability among areas. See text for further details.

Confounded experiments

Temporal confound

e.g.: to assess seasonal patterns is needed more sampling data for each season