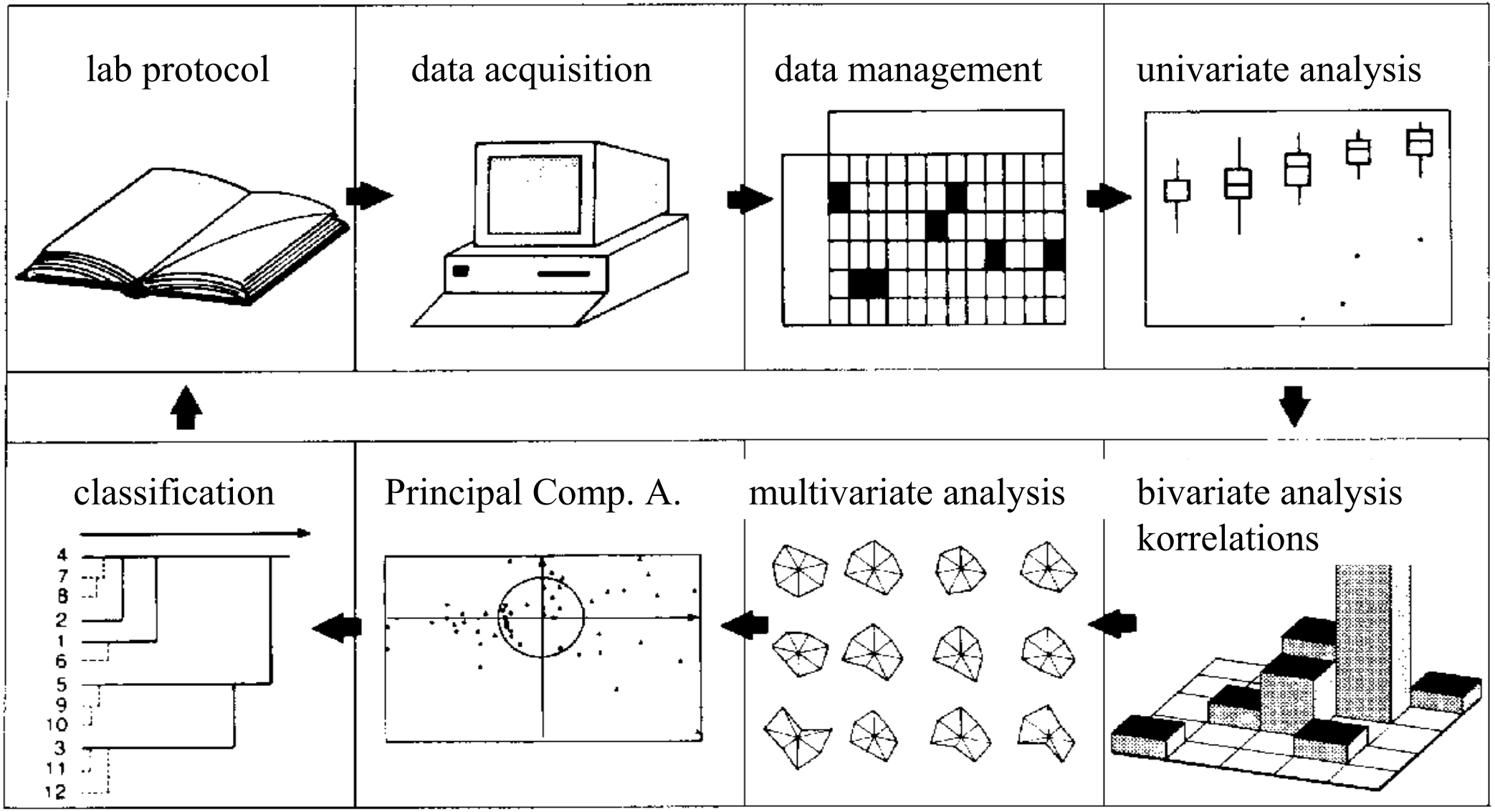
Evaluation of analytical Data - Aims and Purpose
- from Measurements to conceptual Models -



What is the purpose of STATISTICS?

…….to clarify and quantify:

1. Was the experimental design successful or needing upscaling?

What kind of experiment?

Treatment, n Factors? → Differences? → **ANOVA**

T- test, post hoc tests,...

Observation, n Variables? → Relationship? / Clustering? → **ORDINATON**

Corr.A, PCA, Cluster A, NMDS, Discr.A, ...

2. What are your main results, what sound conclusions may be drawn?

Results substancial, how big? → % Δ, $\%R^{2,}$,% Variance explained

Results reliable, how reproducible? → p value / % sig.Level

3. Which presentation graph is best?
4. What are the key parameters for a mathematical model?

A. Data Aquisition: Requirements and Premises

laboratory protocol

- Sample preparation
  All amounts,
  (Weights, Volumina, Tara, FG, TG, Aliquotes, Dilutions)

- Method dokumentation
  All Steps and procedures (Setups, Dilutions, times)

- Primary data tables
  All primary unmodified measurement results

********** Check of Plausibility **********

recapitulate and take down aims of work, approaches, hypotheses !!!
Include this to the primary data tables and in your lab protokol !

B. Preparing the Data Evaluation

Protocol of Evaluation

save primary results, check their comleteness thoroughly
reliability of calibrations, check calibration curves
what is the scattering range of my methods, and which differences are therefore to be taken seriously?

Documentation of the calculation formula
which steps of analysis where carried out in which order (lab protocol)
(at first, do NOT simplify calculations for parallel treatment of samples and blancs)
know the meaning, the values and DIMENSIONS of the primary data e.g. :

Absorbance stands for concentrations	Mol /l	solution
peak areas (quotients) --> amount (weight)	µmol,mg	
ml titer volume stand for Equialents	µVal	

be precise with AMOUNTS and CONCENTRATIONS
(one of the most common errors: mistaking mg for µg and the other way around, avoid calculculating with many zeros behind the comma, eg 0,0001 mg. Use 0.1 µg instead)

C. Calculate and relate the results

systematic compilation of the calculation procedure

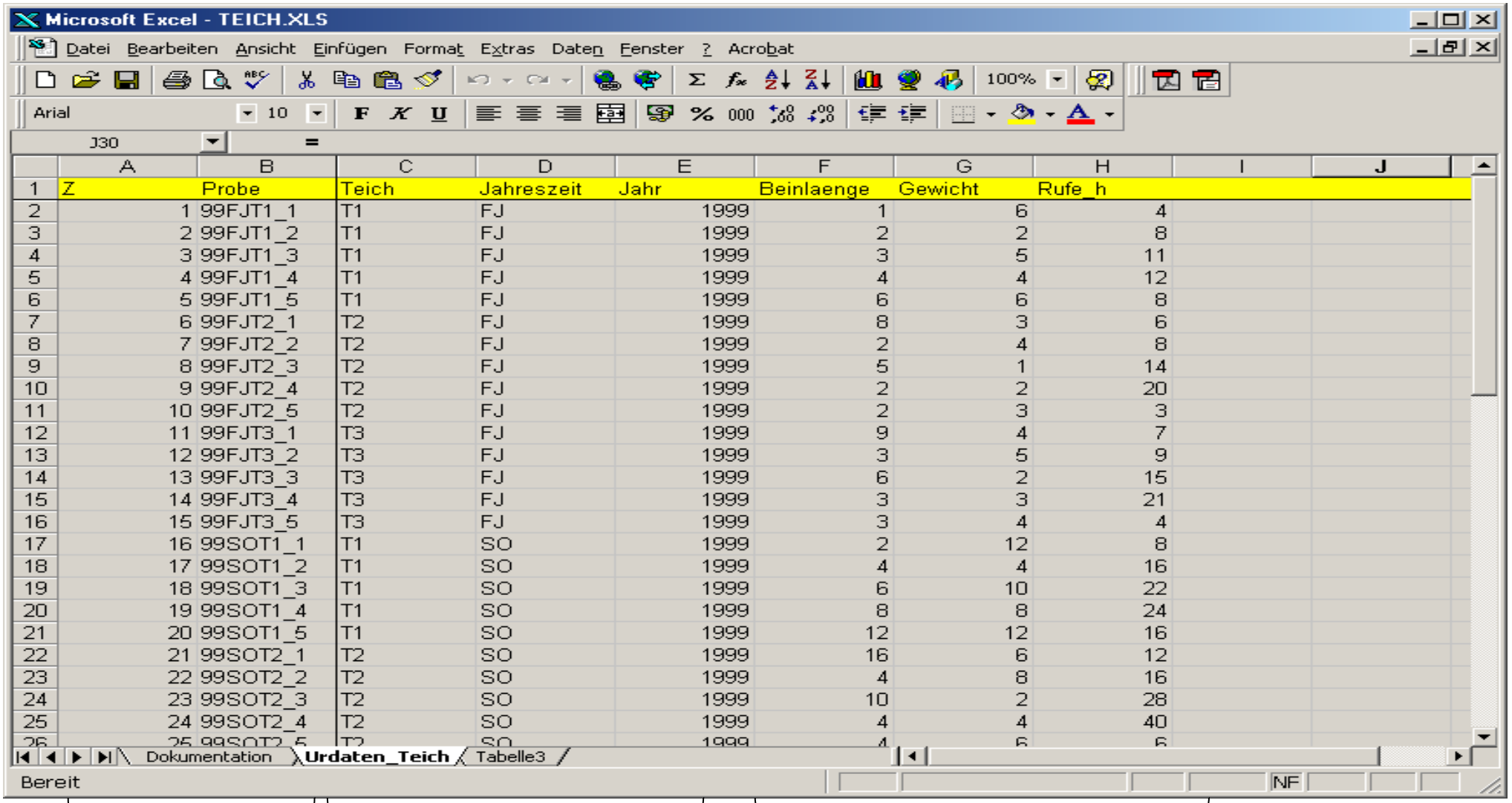
primary data and their dimension
method- dependent calculation
sample preparation dependant recalculations

integrate calculations to formulas with variables and fixed values

********* check plausibility *********

simplification and economisation of calculation methods
(reducing fractions,in a worksheet software- Excel)

- calculation of all Samples
- relating to required dimensions
- adjust and format your tables for easy transfer to statistics programs
- simple statistics: Mean, **Median**, Standard deviation, SE, CI, CV (coeff.of var)



Microsoft Excel - TEICH.XLS

Datei Bearbeiten Ansicht Einfügen Format Extras Daten Fenster ? Acrobat

Arial 10

J30 =

	A	B	C	D	E	F	G	H	I	J
1	Z	Probe	Teich	Jahreszeit	Jahr	Beinlaenge	Gewicht	Rufe_h		
2	1	99FJT1_1	T1	FJ	1999	1	6	4		
3	2	99FJT1_2	T1	FJ	1999	2	2	8		
4	3	99FJT1_3	T1	FJ	1999	3	5	11		
5	4	99FJT1_4	T1	FJ	1999	4	4	12		
6	5	99FJT1_5	T1	FJ	1999	6	6	8		
7	6	99FJT2_1	T2	FJ	1999	8	3	6		
8	7	99FJT2_2	T2	FJ	1999	2	4	8		
9	8	99FJT2_3	T2	FJ	1999	5	1	14		
10	9	99FJT2_4	T2	FJ	1999	2	2	20		
11	10	99FJT2_5	T2	FJ	1999	2	3	3		
12	11	99FJT3_1	T3	FJ	1999	9	4	7		
13	12	99FJT3_2	T3	FJ	1999	3	5	9		
14	13	99FJT3_3	T3	FJ	1999	6	2	15		
15	14	99FJT3_4	T3	FJ	1999	3	3	21		
16	15	99FJT3_5	T3	FJ	1999	3	4	4		
17	16	99SOT1_1	T1	SO	1999	2	12	8		
18	17	99SOT1_2	T1	SO	1999	4	4	16		
19	18	99SOT1_3	T1	SO	1999	6	10	22		
20	19	99SOT1_4	T1	SO	1999	8	8	24		
21	20	99SOT1_5	T1	SO	1999	12	12	16		
22	21	99SOT2_1	T2	SO	1999	16	6	12		
23	22	99SOT2_2	T2	SO	1999	4	8	16		
24	23	99SOT2_3	T2	SO	1999	10	2	28		
25	24	99SOT2_4	T2	SO	1999	4	4	40		
26	25	99SOT2_5	T2	SO	1999	4	6			

Dokumentation | Urdaten_Teich | Tabelle3

Bereit NF

Single samples (cases)

Sampling dates: group variables (factors)

Measuring parameter: measuring variable (variables)

D. Graphical Plotting, Statistics

first::

simple line or barcharts (never waste time and efforts for hasty formatting!)

Excel/Sigmaplot: MEDIAN and ERROR INDICATOR:

Standard deviation STDEV: $\sqrt{\frac{nSx^2-\left(Sx\right)^2}{n\left(n-1\right)}}$

Standard error (error of the mean) SE: $s_x=\frac{s}{\sqrt{n}}$

or Confidence Interval CI: $x\pm t\cdot s_x$ (**t** n=1000...1.96)

for table only: Coefficient of Variance CV: $\frac{s}{x}100$

next:

define the aim of your data evaluation

Differences (sig) Similarities temporal Sequence

define the need for statistics

sign.Differences, Korrelation (Ordination) or Classification?

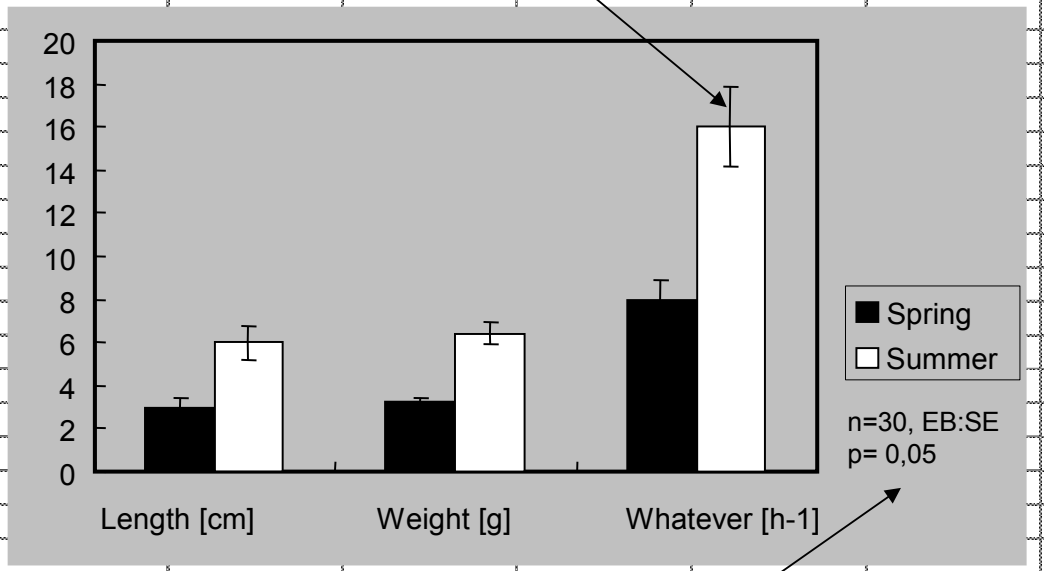
define the aim of your presentation (Poster/Publikation)

Choice of suitable Software -->Formating of the plots

quick plot with error bars prepared with MS Excel / Libre Office Calc

	Length [cm]	Weight [g]	Whatever [h-1]
Spring	3	3,2	8
Summer	6	6,4	16
SE_FJ	0,39259233	0,2530549	0,91515968
SE_SO	0,78518466	0,5061098	1,83031936
CI_FJ	0,80088835	0,51623199	1,86692575
CI_SO	1,6017767	1,03246399	3,7338515

Test for significant Differences
between spring and summer
.
Mediane, Errorbar: SE d. Mittelwertes
Mediane, Errorbar: CI (SE+t(5%), n=30)
t for n=30 at p=0,05: 2,04



Essential informations for Evaluation of plotted Data !

Students T-Test: Δ mean(1,2) < CI1+CI2

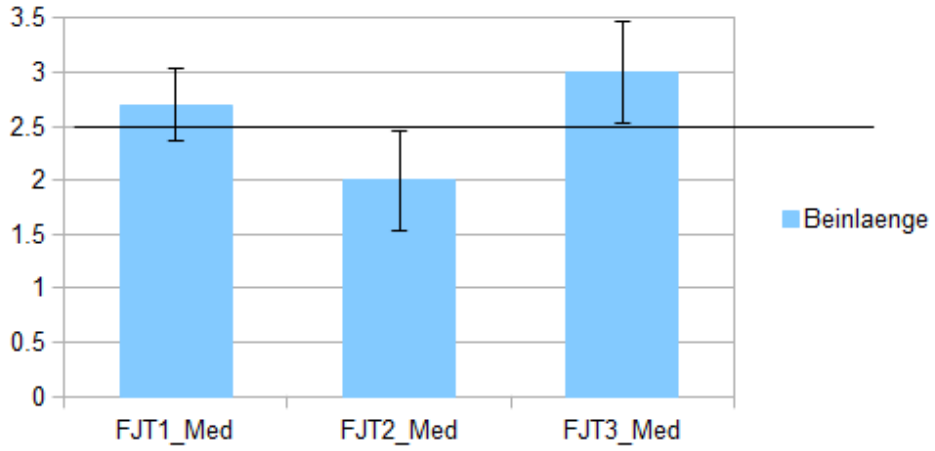
http://en.wikipedia.org/wiki/Post-hoc_analysis

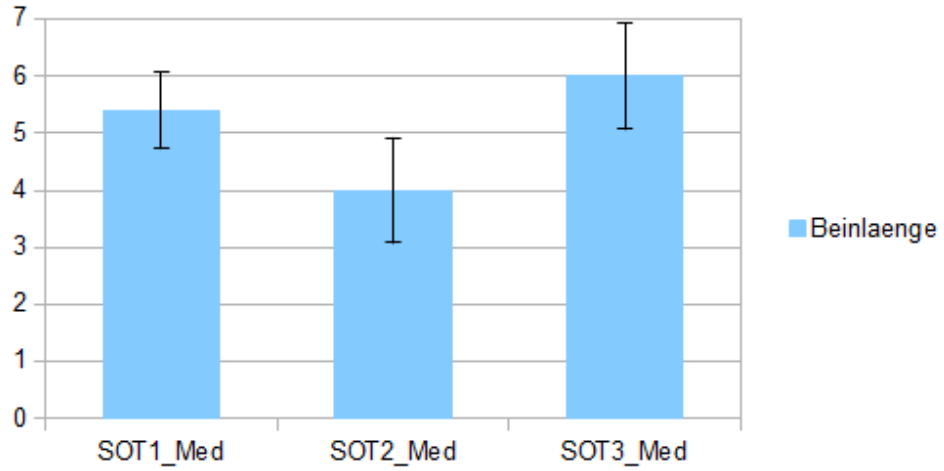
Testing for adequate range of cases
upscaling of undersized experiments

Jahr	Beinlaenge	Gewicht	Rufe_h
FJT1_Med	2.7	4.4	8
FJT2_Med	2	2.4	7.2
FJT3_Med	3	3.2	8.1
SOT1_Med	5.4	8.8	16
SOT2_Med	4	4.8	14.4
SOT3_Med	6	6.4	16.2
FJT1_CI	1.199215608	1.079142333	2.018749636
FJT2_CI	1.663522913	0.723368596	4.221565132
FJT3_CI	1.678154003	0.747335345	4.235950258
SOT1_CI	2.398431215	2.158284666	4.037499273
SOT2_CI	3.327045826	1.446737193	8.443130263
SOT3_CI	3.356308006	1.494670691	8.471900515

		Adäquate n für Beinlänge
'n=10	t Wert für FG 9 = 2,262	
'n=50	t Wert für FG 49 = 2,021	
'n=100	t Wert für FG 99 = 1,987	
'n=150	t Wert für FG 149 = 1,980	
0.333121404		
0.462097962		
0.466162227		
0.666242809		
0.924195925		
0.932324454		

→ =B8/2.262*1.987*SQRT(10)/SQRT(100)





Bezüglich der Beinlänge muss man einen Test für die adequate Stichprobengröße machen.
Bei gleichbleibender Varianz, braucht man ein N von mindestens 100.

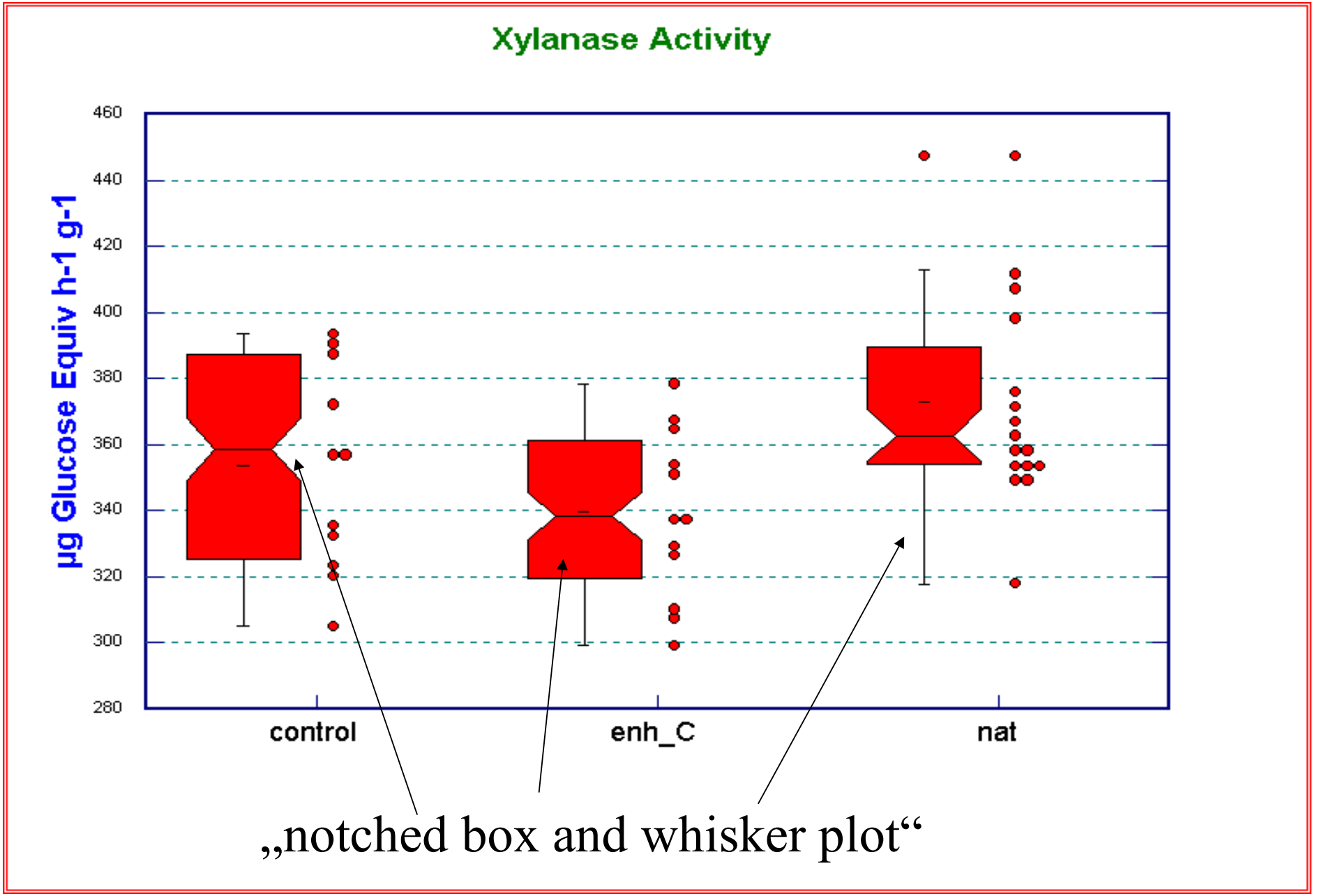
The Significance Level, the Reproducability and Expressability of an Experiment

significant differences (univariate Tests, ANOVA -Analysis Of Variance):

Significance level: P + CI = 100 %
P.... Error probability
CI ...Confidence Interval

highly significant	P < 0,1% (0,001)	***
significant	P < 1% (0,01)	**
weakly significant	P < 5% (0,05)	*
not(?) significant	P > 5%	n.s.

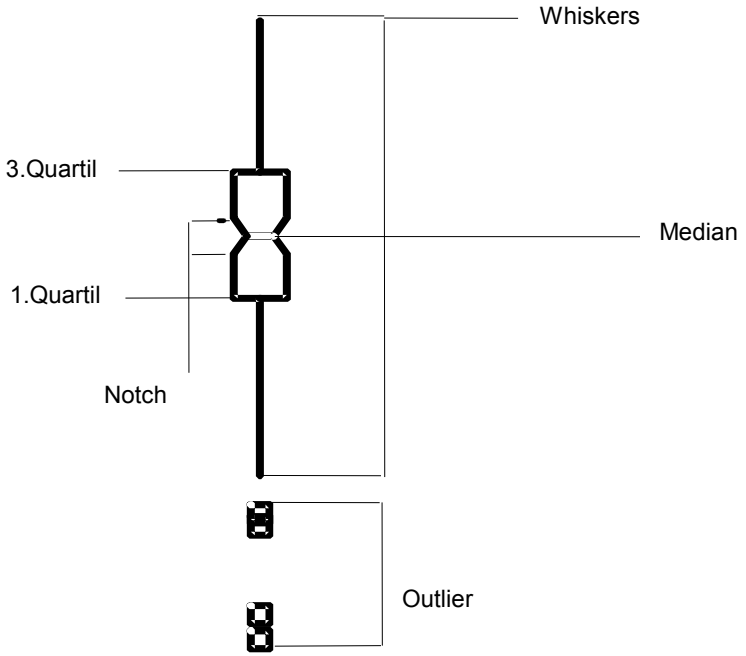
In any case: do provide the actual significance level !
It is not only of interst ***if***, but first of all ***how*** signifikant your results really are.



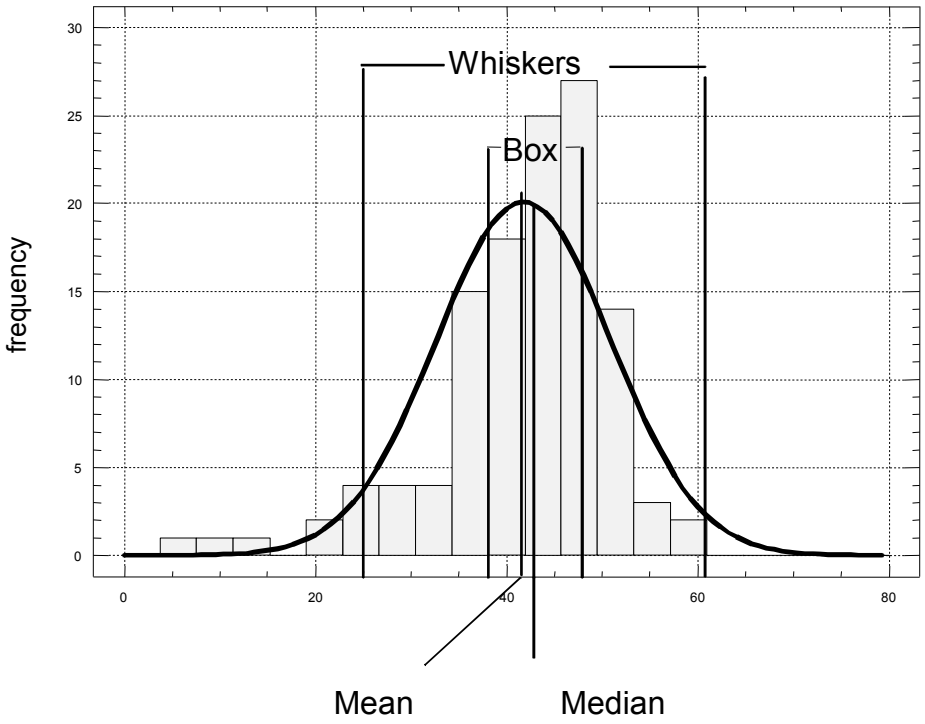
„notched box and whisker plot“

Explanation of notched Box and Whisker Plots

Notched Box and Whisker Plot

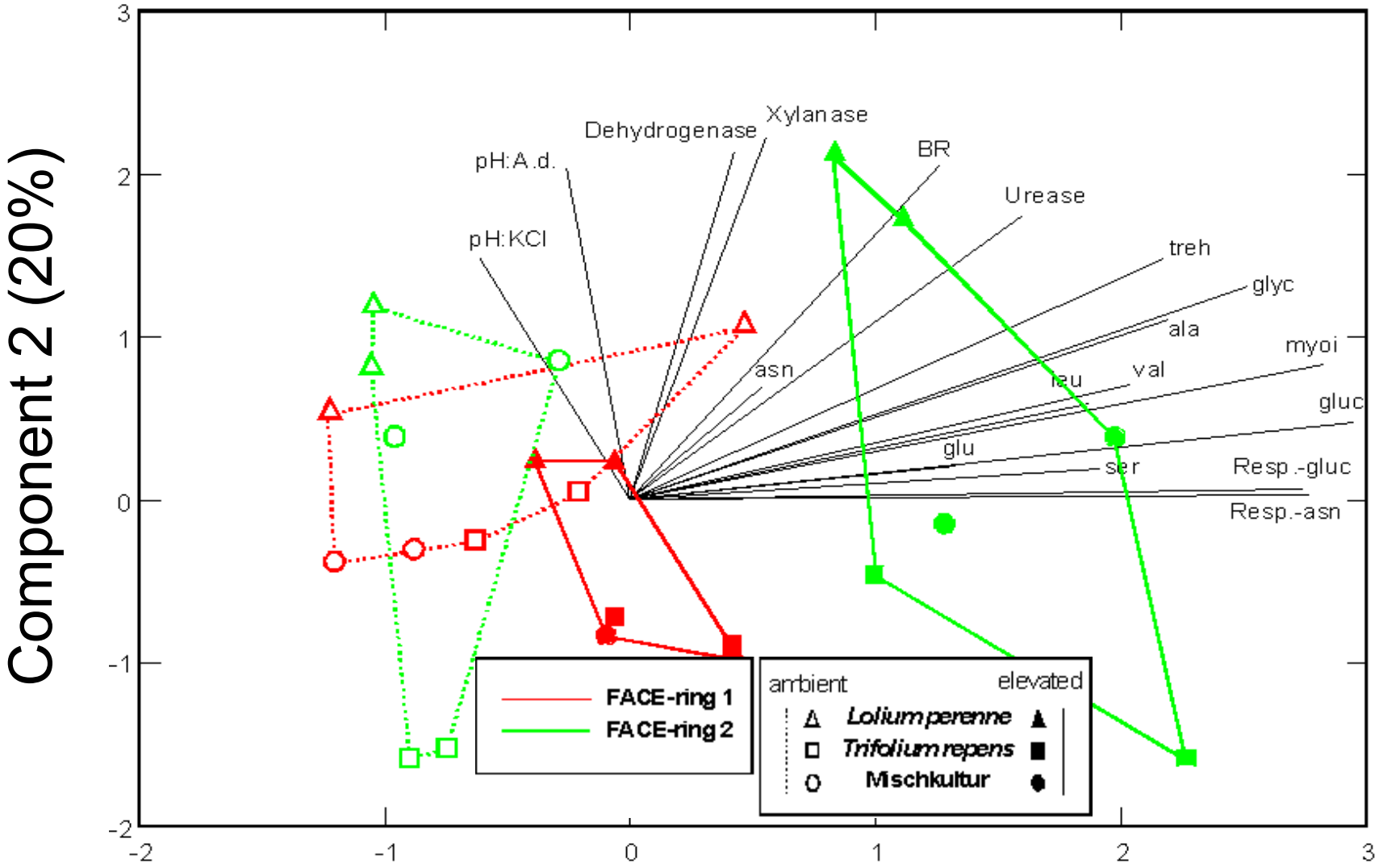


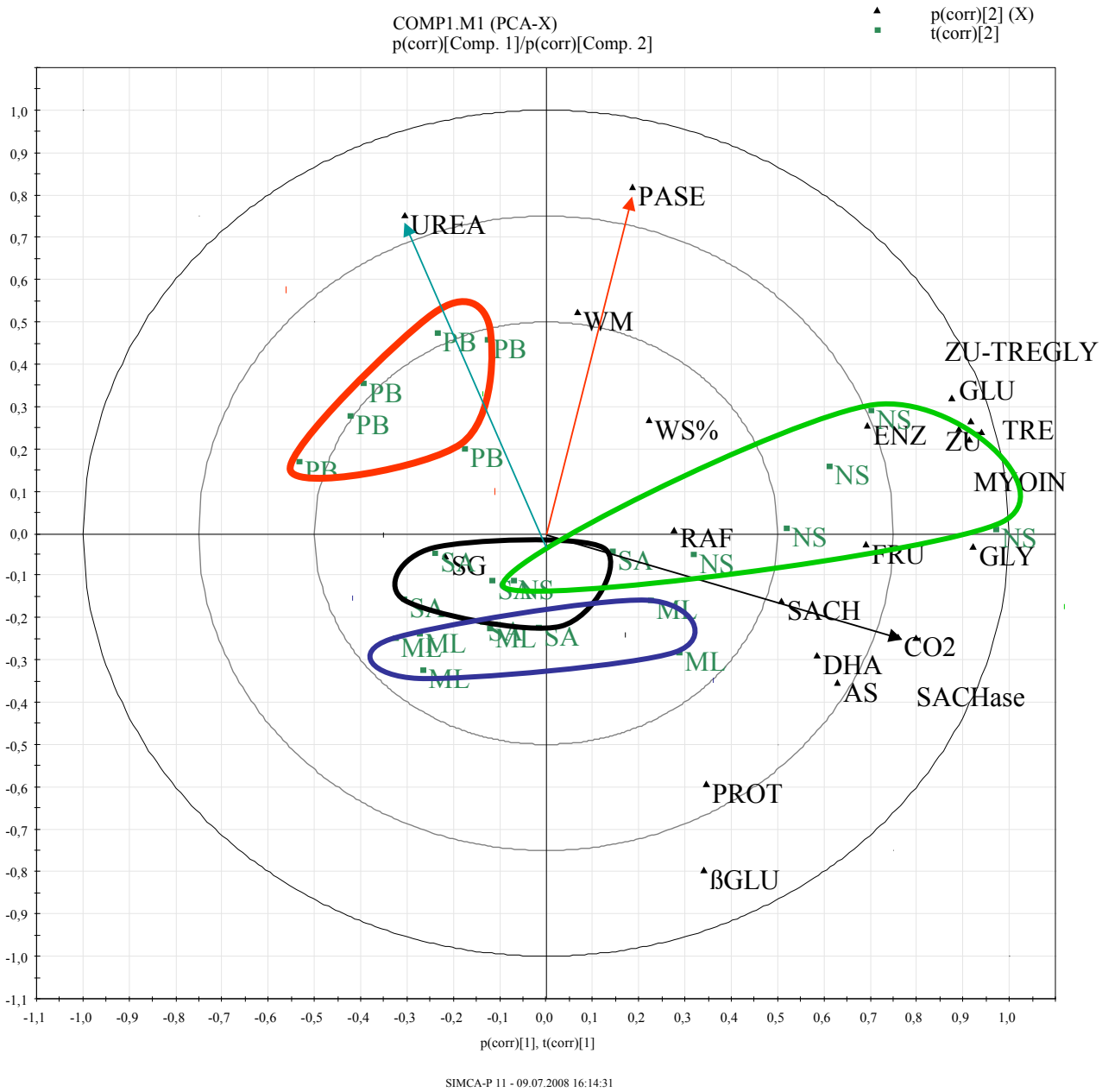
Frequency Histogram



Notch: indicates 95% confidence interval for the Median.
If Notches do not overlap, a significant difference between medians is present

Main Component or Factor analysis



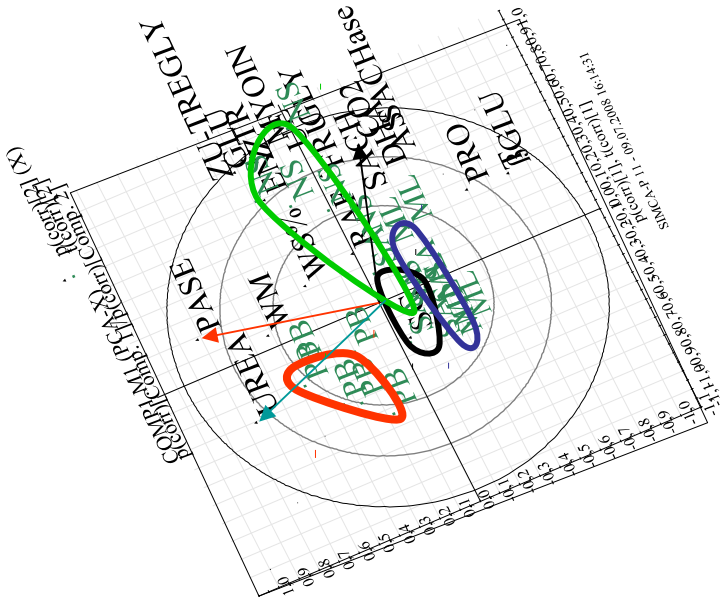


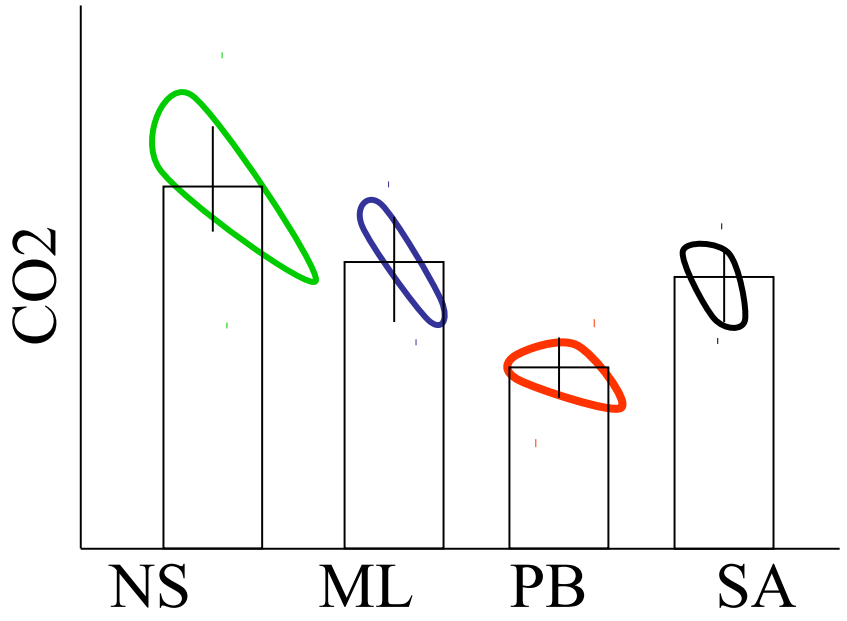
Simca-P
Analysis Loadings Bi-plot
right click, copy
„enh. Metafile“

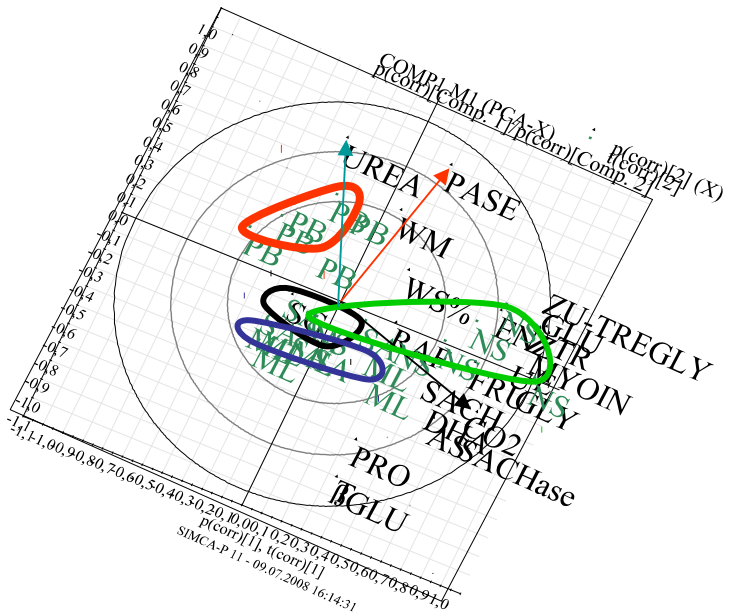
Powerpoint / Impress
reduce size
ungroup
set fonts to 12p
group

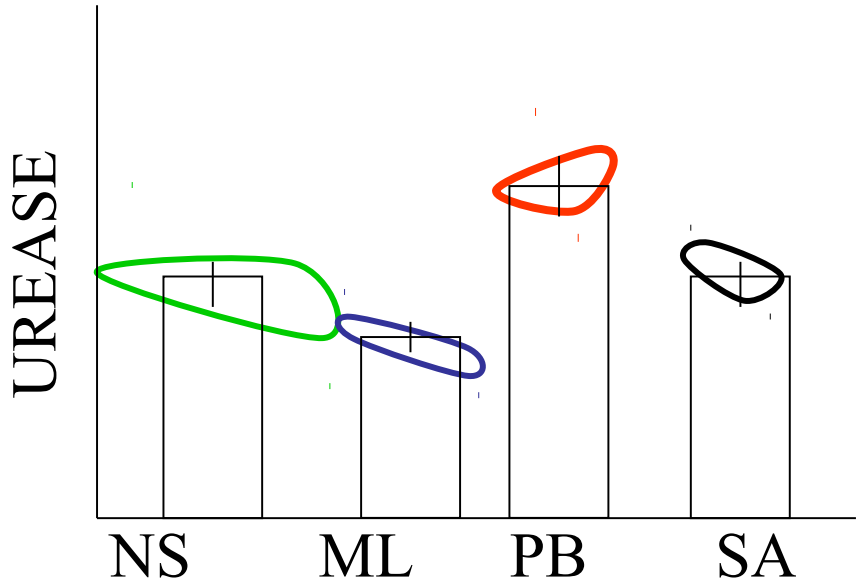
add main vectors
surround cluster by bezier lines

Find the bargraphs (Anova plots) in PCA…

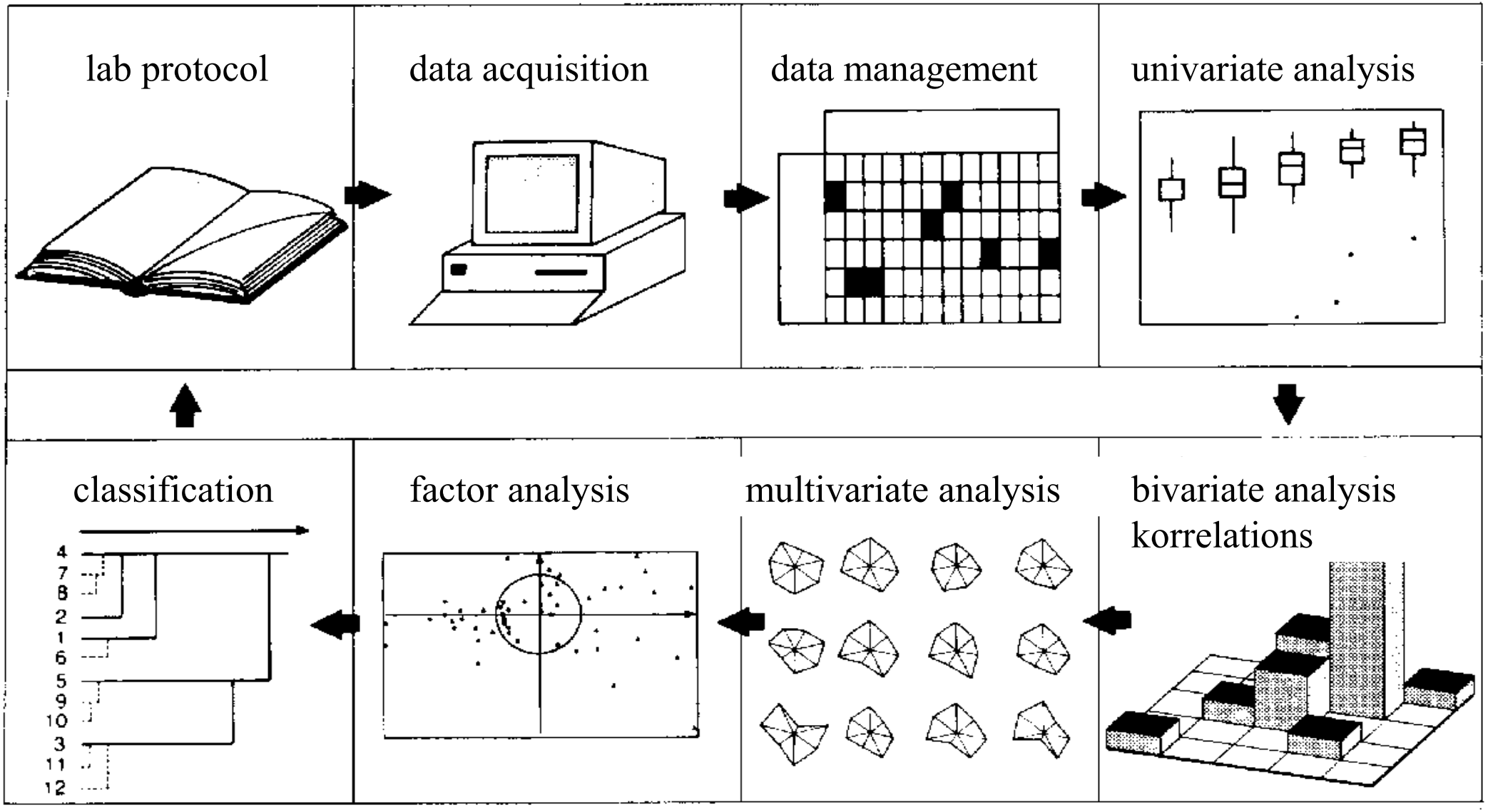




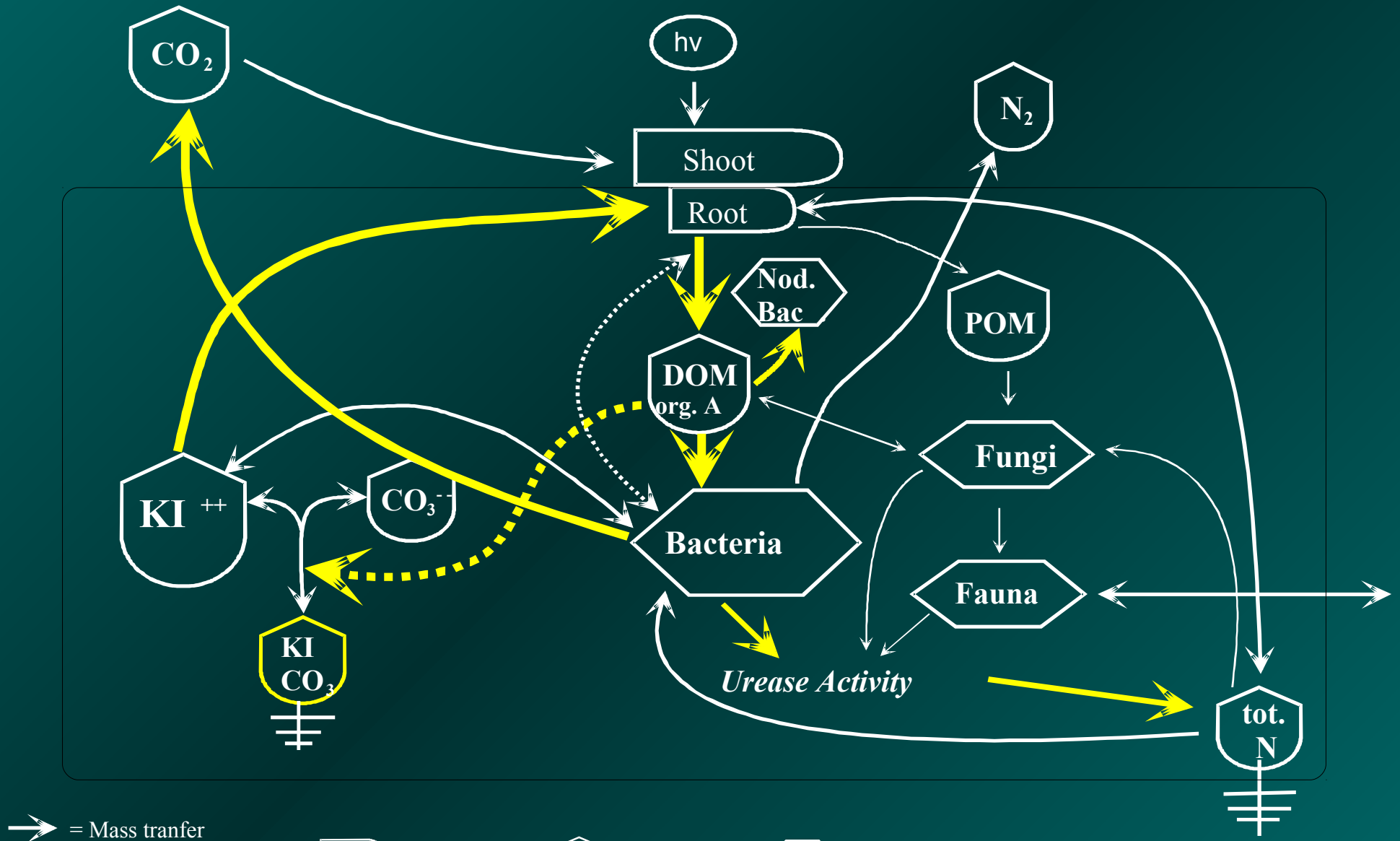




Evaluation of analytical Data - Aims and Purpose
- from Measurements to conceptual Models -



hypothetical model for *Trifolium r.* on alkaline soil



The questions adressed, hypothesis, and the conceptual model determine the statistiks- and plotting necessities !

1.

Questions adressed and investigated::

1 Is there a change in plant productivity due to drougt and fertilization?
2 Is there a combination effect between drought and fertilisation?
3 Is there an Impact of these stresses on soil properties?
4 Are these effects within the bounderies of system dynamics or do they leed to system changes?
5 What are the limiting factors in this soil, how can they be buffered or worked around?