CHEMOMETRIC STRATEGIES FOR ANALYZING MULTIVARIATE DATA COMING FROM DESIGNED EXPERIMENTS: AN OVERVIEW

Federico Marini

Dept. Chemistry, University of Rome "La Sapienza"

SAPIENZA
UNIVERSITÀ DI ROMA

“If the design of an experiment is faulty, any method of interpretation which makes it out to be decisive, must be faulty too. It is true that there are a great many experimental procedures which are well designed, in that they may lead to decisive conclusions, but in other occasion may fail to do so; in such cases, if decisive conclusions are in fact drawn when they are unjustified, we may say that the fault is wholly in the interpretation, not in the design. But the fault in the interpretation, even in these cases, lies in overlooking the characteristic features of the design, which lead to the result being sometimes inconclusive, or conclusive on some questions but not on all. To understand correctly the one aspect of the problem is to understand the other. ***Statistical procedure and experimental design are only two different aspects of the same whole, and that whole comprises all the logical requirements of the complete process of adding to natural knowledge by experimentation****.”*

R.A. Fisher (1936)

Interpreting univariate results of ED: ANOVA

- Assumes the additivity of the main effects of the controlled factors and their interactions.
- For 2 factors:

$$x_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

- where:
  - i=1:I levels of the first factors
  - j=1:J levels of the second factor
  - k=1:K replicates of each design point

- If only a single replicate (k=1):

$$(\alpha\beta)_{ij} \equiv \varepsilon_{ij}$$

More on univariate ANOVA

- Solution is not unique and there's need of constraints:

$$\sum_{i=1}^{I} \alpha_i = 0$$

$$\sum_{j=1}^{J} \beta_j = 0$$

$$\sum_{i=1}^{I} (\alpha\beta)_{ij} = 0 \quad \forall j$$

$$\sum_{j=1}^{J} (\alpha\beta)_{ij} = 0 \quad \forall i$$

- Accordingly:

$$\mu \to x_{..}$$

$$\alpha_i \to x_{i.} - x_{..}$$

$$\beta_j \to x_{.j} - x_{..}$$

$$(\alpha\beta)_{ij} \to x_{ij} - x_{i.} - x_{.j} + x_{..}$$

ANOVA: Hypothesis testing

- The null hypothesis H_0:

$$\alpha_1 = \alpha_2 = \ldots = \alpha_I$$

- The total sum of squares is decomposed in the sum of individual terms:

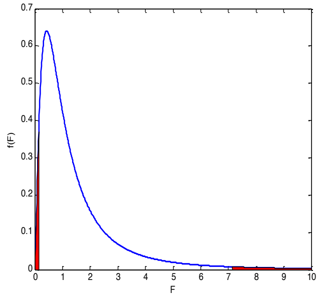
$$\|x-\mu\|^2 = \|\alpha\|^2 + \|\beta\|^2 + \|\alpha\beta\|^2$$

$$\underbrace{\sum_{i,j}\left(x_{ij}-x_{..}\right)^2}_{TSS} = \underbrace{\sum_{i,j}\left(x_{i.}-x_{..}\right)^2}_{SS_\alpha} + \underbrace{\sum_{i,j}\left(x_{.j}-x_{..}\right)^2}_{SS_\beta} + \underbrace{\sum_{i,j}\left(x_{ij}-x_{i.}-x_{.j}+x_{..}\right)^2}_{SS_{\alpha\beta}\equiv SS_R}$$

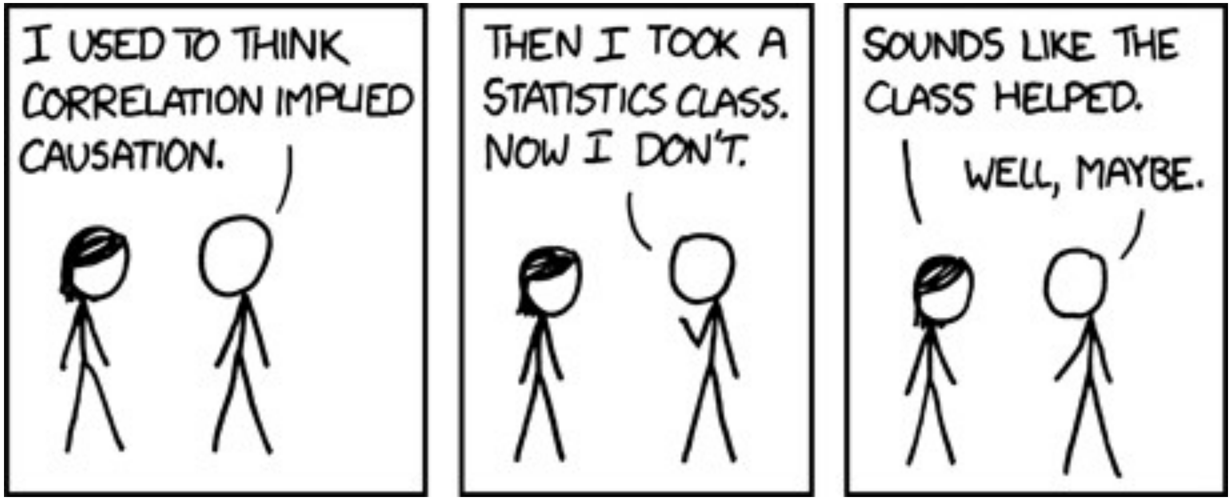
- Significance is (usually) estimated by means of F test:

$$F_\alpha = \frac{SS_\alpha}{SS_R}\frac{N-I-J+1}{I-1}$$

$$F_\beta = \frac{SS_\beta}{SS_R}\frac{N-I-J+1}{J-1}$$



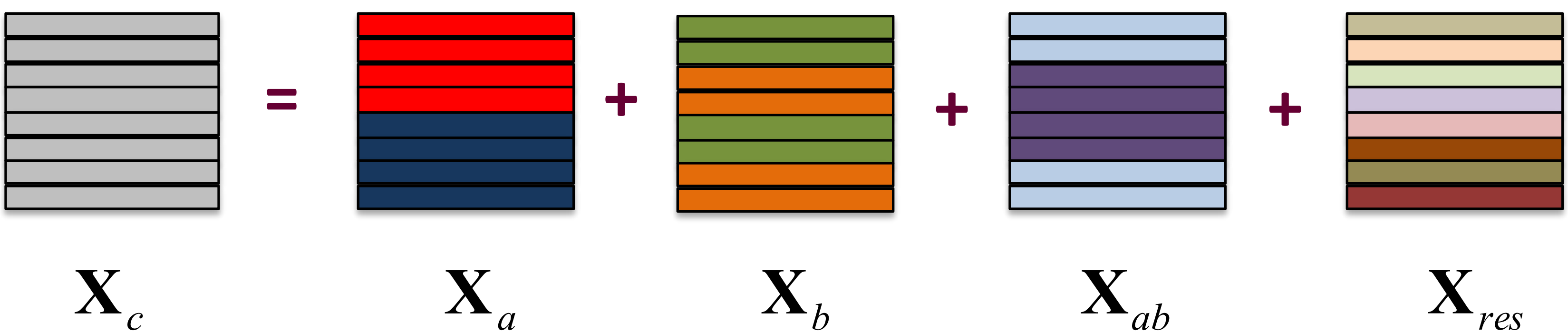
Going multivariate…



Multivariate ANOVA decomposition

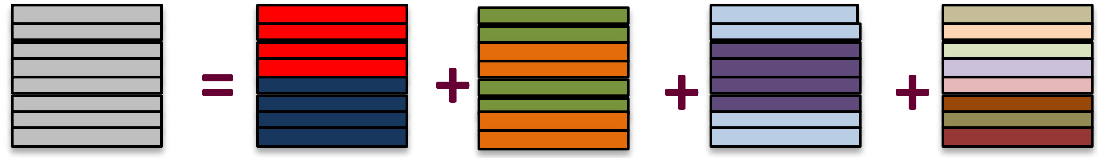
Estimation of the effects

Experiment	Factor a	Factor b
1	+	+
2	+	+
3	+	-
4	+	-
5	-	+
6	-	+
7	-	-
8	-	-

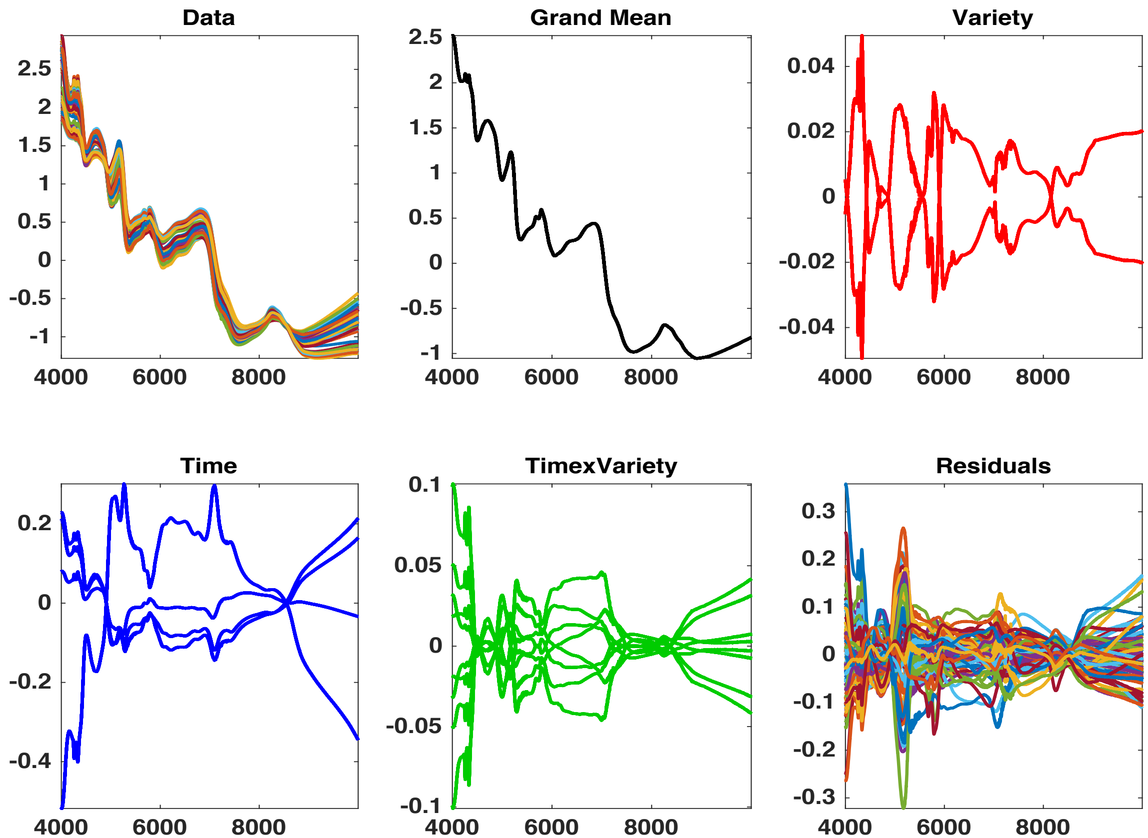


Multivariate ANOVA decomposition

Estimation of the effects



$\mathbf{X}_c \quad \mathbf{X}_a \quad \mathbf{X}_b \quad \mathbf{X}_{ab} \quad \mathbf{X}_{res}$



Two way MANOVA as (multivariate multiple) regression model

- The MANOVA decomposition $y_{ijk} = \boldsymbol{\mu} + \boldsymbol{\alpha}_i + \boldsymbol{\beta}_j + \boldsymbol{\alpha\beta}_{ij} + \boldsymbol{\epsilon}_{ijk}$ can be also expressed as regression model: $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{B} + \boldsymbol{E}$

$$\boldsymbol{X} = [\mathbf{1}_{IJK} \quad \boldsymbol{X}_A \quad \boldsymbol{X}_B \quad \boldsymbol{X}_{AB}] = [\mathbf{1}_{IJK} \quad \boldsymbol{I}_I \otimes \mathbf{1}_{JK} \quad \mathbf{1}_I \otimes \boldsymbol{I}_J \otimes \mathbf{1}_K \quad \boldsymbol{I}_{IJ} \otimes \mathbf{1}_K]$$

$$\boldsymbol{X} = \left[\begin{array}{c|cc|ccc|cccccc}
1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right]$$

S.S Wilks, *Biometrika* **24** (1932) 471-494
P.L. Hsu, *Biometrika* **31** (1940) 221-237

MANOVA

Structure of the model

$$\mathbf{x}_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}$$

$$\varepsilon_{ijk} \sim NID(0, \Sigma)$$

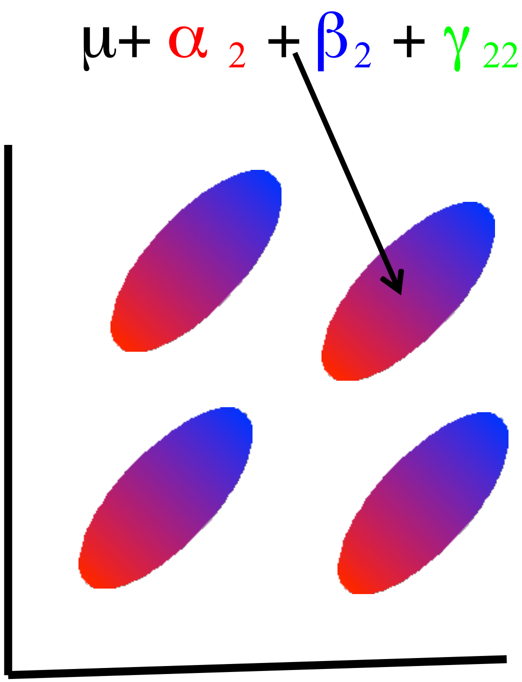
SSQ PARTITIONING

$$\mathbf{T} = \mathbf{B}_\alpha + \mathbf{B}_\beta + \mathbf{B}_\gamma + \mathbf{W}$$

TESTING

$$\Lambda_f = \frac{\mathbf{W}}{\mathbf{B}_f + \mathbf{W}}$$

$$-\left[IJ(n-1) - \frac{v+1-(f-1)}{2}\right]\Lambda_f \geq \chi^2_{v(f-1)}$$



Two way MANOVA as (multivariate multiple) regression model

- The MANOVA decomposition $\boldsymbol{y}_{ijk} = \boldsymbol{\mu} + \boldsymbol{\alpha}_i + \boldsymbol{\beta}_j + \boldsymbol{\alpha\beta}_{ij} + \boldsymbol{\epsilon}_{ijk}$ can be also expressed as regression model: $\boldsymbol{Y} = \boldsymbol{XB} + \boldsymbol{E}$
- Accordingly, the model parameters $\boldsymbol{B}$, i.e. the estimates of factor and interaction levels, can be calculated as: $\hat{\boldsymbol{B}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y}$
- With this formulation, it is also possible to calculate the Hypothesis and Error SSCP as:
  - $\boldsymbol{H}_A = \boldsymbol{Y}^T\boldsymbol{X}_A\left(\boldsymbol{X}_A^T\boldsymbol{X}_A\right)^{-1}\boldsymbol{X}_A^T\boldsymbol{Y}$
  - $\boldsymbol{H}_B = \boldsymbol{Y}^T\boldsymbol{X}_B(\boldsymbol{X}_B^T\boldsymbol{X}_B)^{-1}\boldsymbol{X}_B^T\boldsymbol{Y}$
  - $\boldsymbol{H}_{AB} = \boldsymbol{Y}^T\boldsymbol{X}_{AB}\left(\boldsymbol{X}_{AB}^T\boldsymbol{X}_{AB}\right)^{-1}\boldsymbol{X}_{AB}^T\boldsymbol{Y}$
  - $\boldsymbol{E} = \boldsymbol{Y}^T\left(\boldsymbol{I}_{IJK} - \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\right)\boldsymbol{Y}$
- Unfortunately, due to matrix inversion problems, in order to be able to carry out the tests, inverting the matrix **E**, it is necessary that $L \leq IJK - rank(\boldsymbol{X})$
- The method is unsuitable for highly multivariate responses → **Different available options**

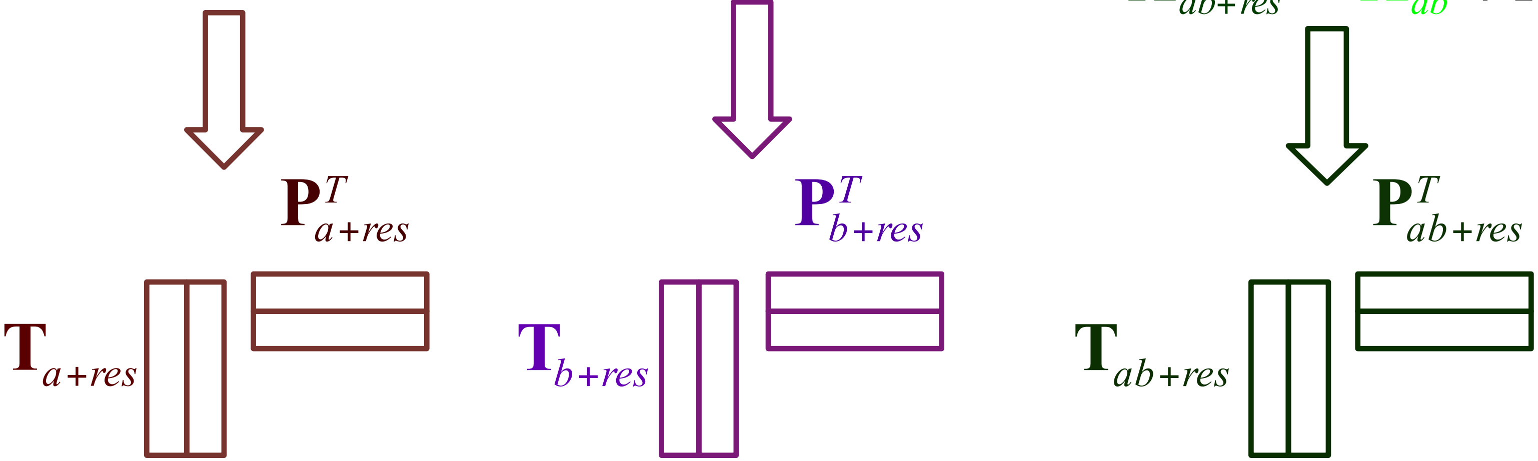
ANOVA-PCA

Structure of the model

$$\mathbf{X} = \mathbf{1}\mathbf{m}^T + \mathbf{X}_a + \mathbf{X}_b + \mathbf{X}_{ab} + \mathbf{X}_{re}$$

EFFECT MATRICES + RESIDUALS

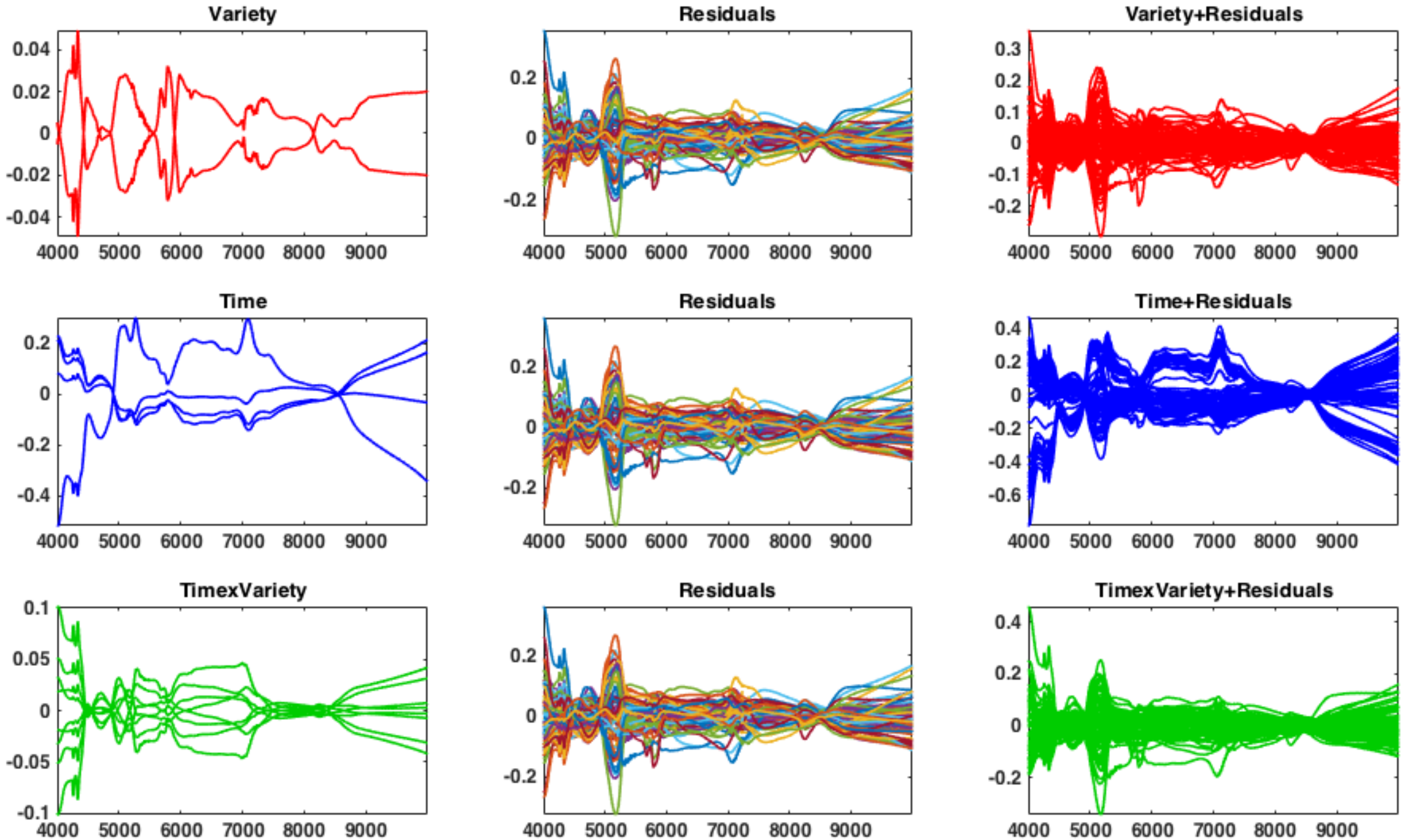
$$\mathbf{X}_{a+res} = \mathbf{X}_a + \mathbf{X}_{res} \qquad \mathbf{X}_{b+res} = \mathbf{X}_b + \mathbf{X}_{res} \qquad \mathbf{X}_{ab+res} = \mathbf{X}_{ab} + \mathbf{X}_{res}$$



P. Harrington *et al.*, *Anal. Chim. Acta* **544** (2005) 118-127

ANOVA-PCA

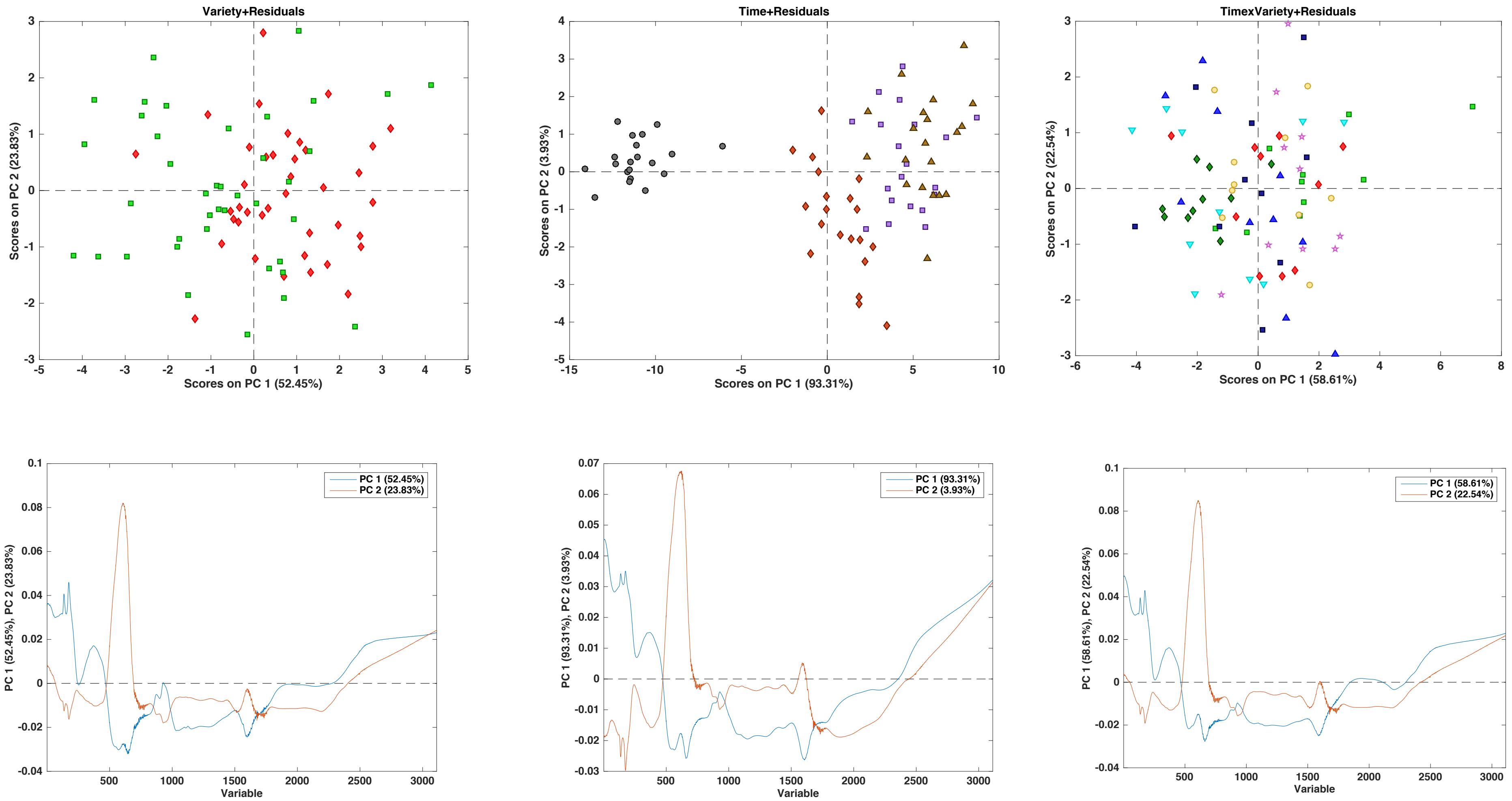
Structure of the model



P. Harrington *et al.*, *Anal. Chim. Acta* **544** (2005) 118-127

ANOVA-PCA

Structure of the model



AComDim

D.J.-R. Bouveresse *et al.*, *Chemolab* **106** (2009) 173-182

Structure of the model

$$\mathbf{W}_{a+res} = \mathbf{X}_{a+res}\mathbf{X}_{a+res}^T \quad \mathbf{W}_{b+res} = \mathbf{X}_{b+res}\mathbf{X}_{b+res}^T \quad \mathbf{W}_{ab+res} = \mathbf{X}_{ab+res}\mathbf{X}_{ab+res}^T \quad \mathbf{W}_{res} = \mathbf{X}_{res}\mathbf{X}_{res}^T$$

MULTIBLOCK ANALYSIS (ComDim) on the ANOVA-PCA MATRICES

$$\mathbf{W}_{tot} = \lambda_{a+res}\mathbf{W}_{a+res} + \lambda_{b+res}\mathbf{W}_{b+res} + \lambda_{ab+res}\mathbf{W}_{ab+res} + \lambda_{res}\mathbf{W}_{res}$$

$\mathbf{q}_1$ first common component

$$\mathbf{W}_{tot} \approx \lambda^1_{a+res}\mathbf{q}_1\mathbf{q}_1^T + \lambda^1_{b+res}\mathbf{q}_1\mathbf{q}_1^T + \lambda^1_{ab+res}\mathbf{q}_1\mathbf{q}_1^T + \lambda^1\ \mathbf{q}_1\mathbf{q}_1^T$$

$$\lambda\ = \mathbf{q}_1^T\mathbf{W}\,\mathbf{q}_1$$

SIGNIFICANCE TESTING

$\mathbf{q}_1$ associated mainly to Noise

$$F_j = \frac{\lambda_{res}}{\lambda_j^1} = \frac{\mathbf{q}_1^T\mathbf{W}_{res}\mathbf{q}_1}{\mathbf{q}_1^T\mathbf{W}_j\mathbf{q}_1}$$

ASCA

(ANOVA - Simultaneous Component Analysis)

Structure of the model and estimation of the effects

A.K. Smilde *et al.*, *Bioinformatics* **21** (2005) 3043-3048
J.J. Jansen *et al. J. Chemometr.* **19** (*2005*) 469-481

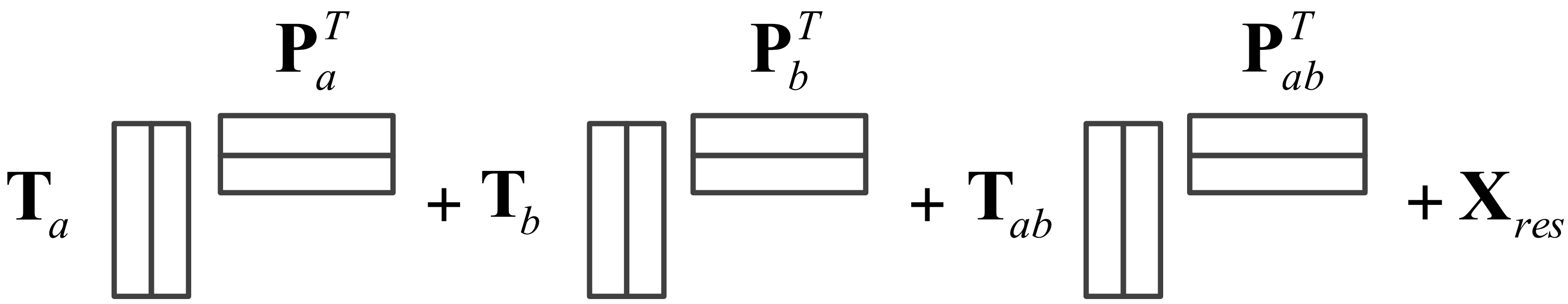
ASCA (ANOVA - Simultaneous Component Analysis)

Structure of the model

$$\mathbf{X} = \mathbf{1}\mathbf{m}^T + \mathbf{X}_a + \mathbf{X}_b + \mathbf{X}_{ab} + \mathbf{X}_{res}$$

SCA

$$\mathbf{X}_c = \mathbf{X} - \mathbf{1}\mathbf{m}^T =$$



J.J. Jansen *et al. J. Chemometr.* **19** (*2005*) 469-481

ASCA (ANOVA - Simultaneous Component Analysis)

Estimation of the effects

$$\mathbf{X}_c \quad = \quad \mathbf{X}_a \quad + \quad \mathbf{X}_b \quad + \quad \mathbf{X}_{ab} \quad + \quad \mathbf{X}_{res}$$

$$\left\|\mathbf{X}_c\right\|^2 = \left\|\mathbf{X} - \mathbf{1}\mathbf{m}^T\right\|^2 = \left\|\mathbf{X}_a\right\|^2 + \left\|\mathbf{X}_b\right\|^2 + \left\|\mathbf{X}_{ab}\right\|^2 + \left\|\mathbf{X}_{res}\right\|^2$$

SUM OF SQUARES

Validation of ASCA

Significance of the effects

Validation of effect significance by permutation tests

- Originally, a non-parametric approach based on permutation tests was proposed for the validation of the multivariate effects in ASCA.
- Exact permutation tests can be built only for main effects
- Approximate tests can be used for the interactions.
  - Permutation of the residuals under the reduced model (similar to Type III sum of squares)

Permutation tests

6	-3
7	-4
8	-2
9	-1
-6	4
-7	2
-8	3
-9	1

X

7.5	-2.5
7.5	-2.5
7.5	-2.5
7.5	-2.5
-7.5	2.5
-7.5	2.5
-7.5	2.5
-7.5	2.5

$\mathbf{X}_A$ ssq=500

6	-3
-6	4
8	-2
9	-1
7	-4
-7	2
-8	3
-9	1

$\mathbf{X}_{perm1}$

4.25	-0.5
4.25	-0.5
4.25	-0.5
4.25	-0.5
-4.25	0.5
-4.25	0.5
-4.25	0.5
-4.25	0.5

$\mathbf{X}_{A,perm1}$ ssq=146.5

6	-3
-6	4
8	-2
-8	3
7	-4
-7	2
9	-1
-9	1

$\mathbf{X}_{perm2}$

0	0.5
0	0.5
0	0.5
0	0.5
0	-0.5
0	-0.5
0	-0.5
0	-0.5

$\mathbf{X}_{A,perm2}$ ssq=2

-8	3
7	-4
8	-2
-9	1
-6	4
-7	2
6	-3
9	-1

$\mathbf{X}_{permn}$

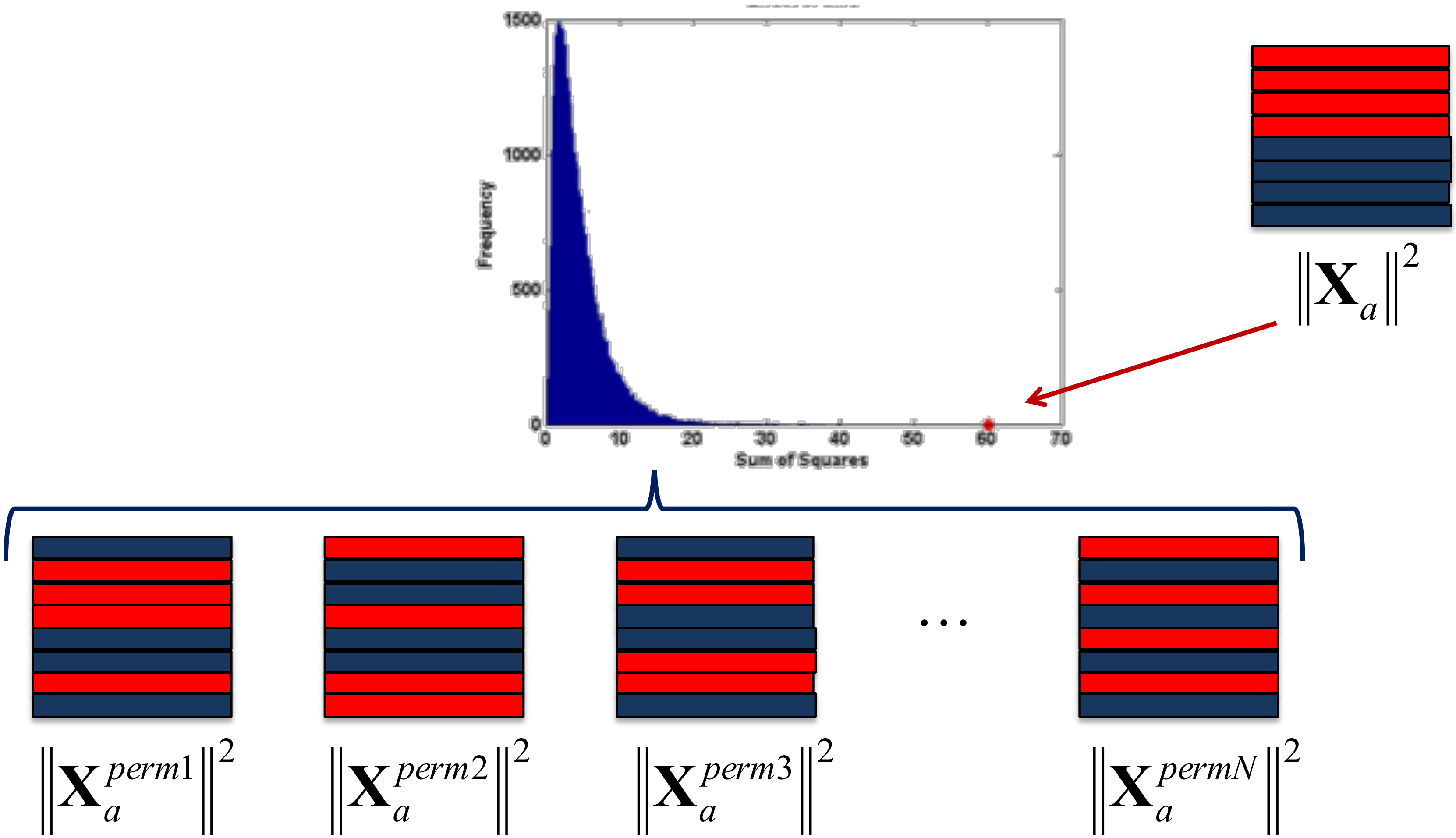
-0.5	-0.5
-0.5	-0.5
-0.5	-0.5
-0.5	-0.5
0.5	0.5
0.5	0.5
0.5	0.5
0.5	0.5

$\mathbf{X}_{A,permn}$ ssq=4

Permutation tests

M.J. Anderson, C.J.F. ter Braak. *J. Stat. Comp. Simul.* **73** (*2003*) 85-113

- Results of permutations are used to estimate the distribution of the Effect size under the null hypothesis



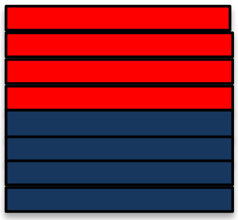
ASCA

(ANOVA - Simultaneous Component Analysis)

Interpretation of the effects

ASCA

Interpretation of the effects

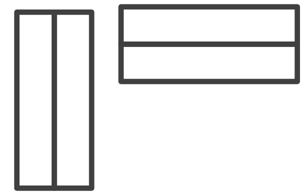


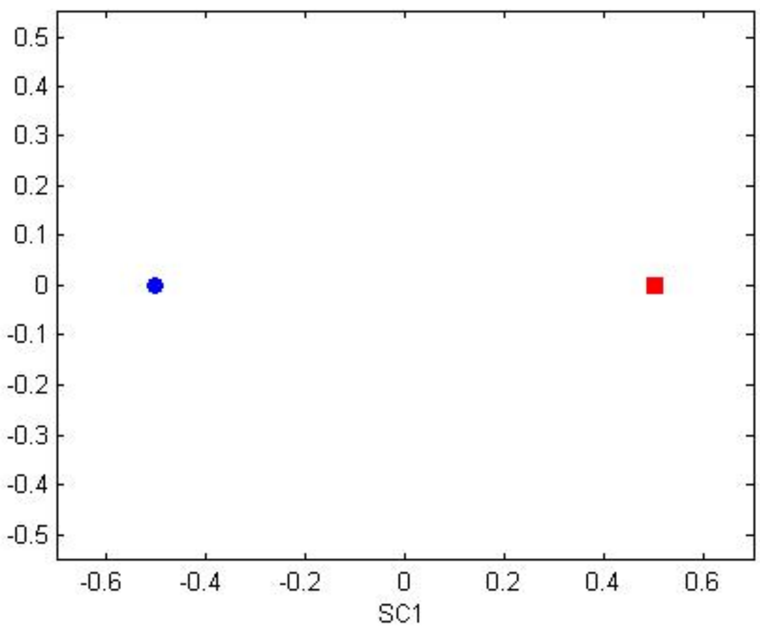
$\mathbf{X}_a$

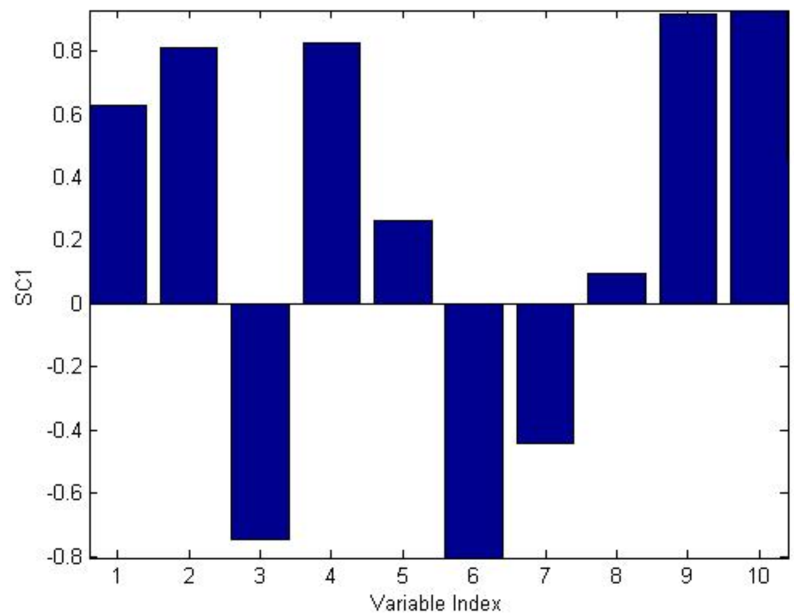
SCA

$\mathbf{P}_a^T$

$$\mathbf{X}_a = \mathbf{T}_a\mathbf{P}_a^T = \mathbf{T}_a$$







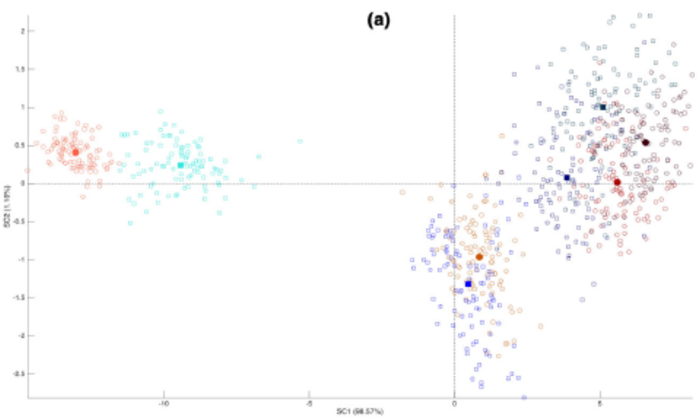
Validation of ASCA

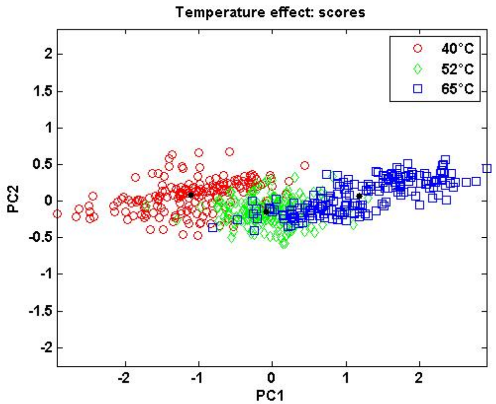
Significance of the effects:
Graphical evaluation

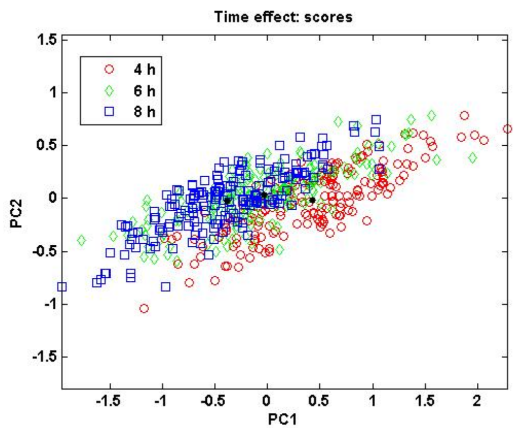
Graphical evaluation of effect significance

- Residuals of ANOVA decomposition are projected back onto the SCA model of the Effect matrix

$$\widehat{\mathbf{X}}_{\mathrm{A}} = \mathbf{T}_{\mathrm{A}}\mathbf{P}_{\mathrm{A}}^{T} \qquad \mathbf{T}'_{\mathrm{A}} = (\mathbf{X}_{\mathrm{A}} + \mathbf{X}_{\mathrm{Res}})\mathbf{P}_{\mathrm{A}}$$





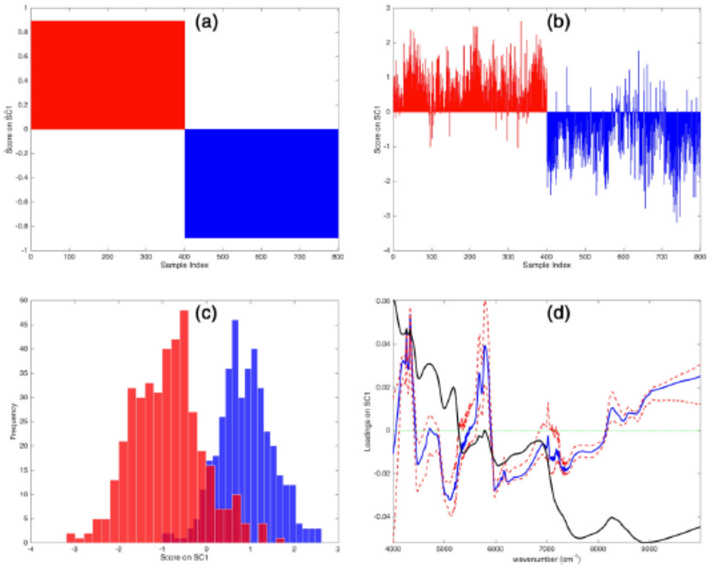


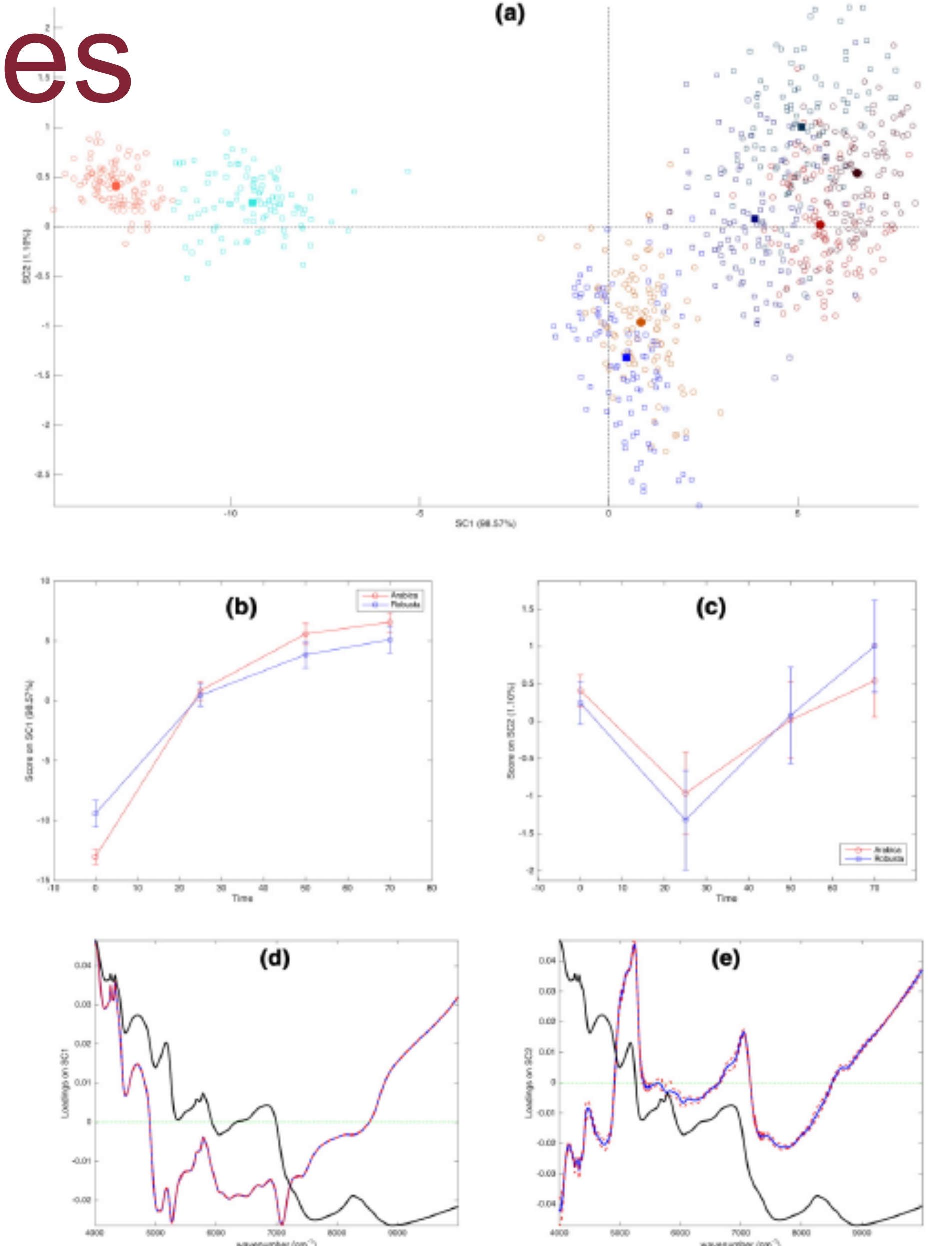
ASCA

(ANOVA - Simultaneous Component Analysis)

Applications

Recent researches





Contents lists available at ScienceDirect

Microchemical Journal

journal homepage: www.elsevier.com/locate/microc

Characterization of the effects of different roasting conditions on coffee samples of different geographical origins by HPLC-DAD, NIR and chemometrics

Silvia De Luca, Martina De Filippis, Remo Bucci, Andrea D. Magrì, Antonio L. Magrì, Federico Marini *

Department of Chemistry, University of Rome "La Sapienza", P.le Aldo Moro 5, I-00185 Rome, Italy

ANOVA-Target Projection (ANOVA-TP)

Journal of Chromatography A, 1405 (2015) 94–102

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Analysis of variance of designed chromatographic data sets: The analysis of variance-target projection approach

Federico Marini[a], Dalene de Beer[b], Elizabeth Joubert[b,c], Beata Walczak[d,*]

[a] Department of Chemistry, University of Rome "La Sapienza", P.le Aldo Moro 5, I-00185 Rome, Italy
[b] Post-Harvest and Wine Technology Division, Agricultural Research Council (ARC), Infruitec-Nietvoorbij, Private Bag X5026, Stellenbosch 7599, South Africa
[c] Department of Food Science, Stellenbosch University, Private Bag X1, Matieland, Stellenbosch, South Africa
[d] Institute of Chemistry, The University of Silesia, Szkolna Street 9, 40-006 Katowice, Poland

Journal of Chromatography A, xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

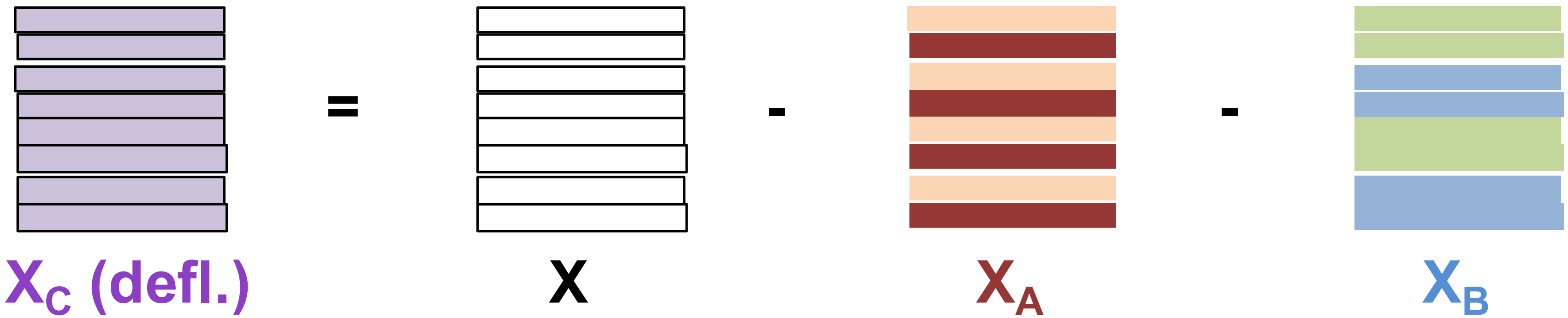
Multivariate analysis of variance of designed chromatographic data. A case study involving fermentation of rooibos tea

Federico Marini[a], Dalene de Beer[b,c], Nico A. Walters[b,c], André de Villiers[d], Elizabeth Joubert[b,c], Beata Walczak[e,*]

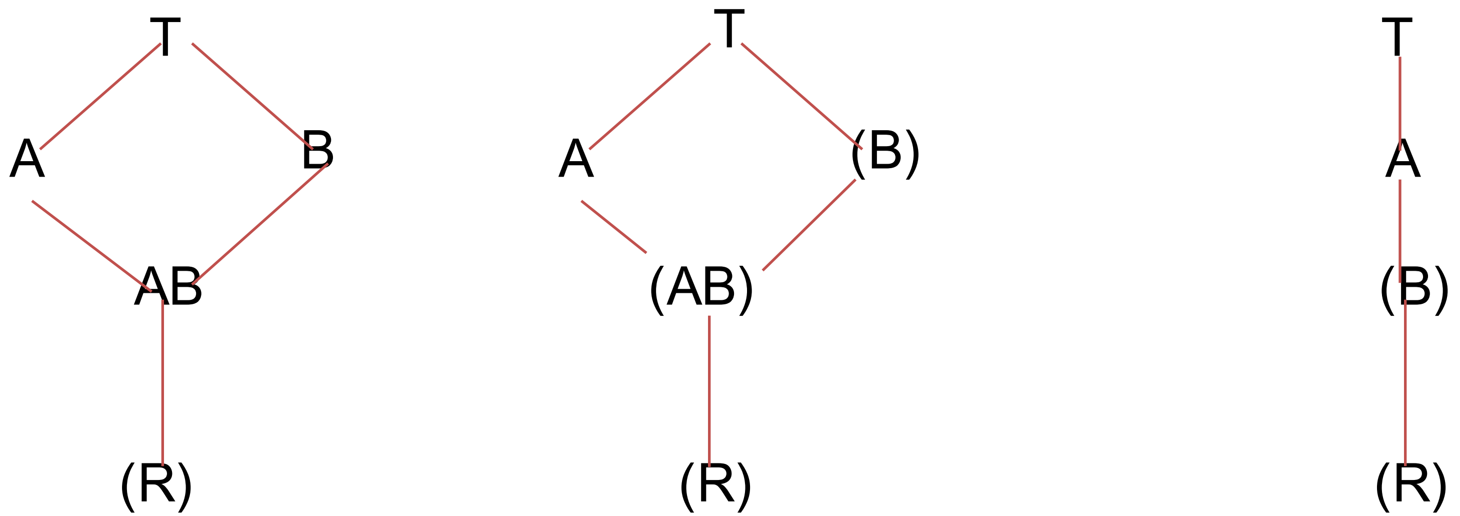
[a] Department of Chemistry, University of Rome "La Sapienza", P.le Aldo Moro 5, I-00185 Rome, Italy
[b] Post-Harvest and Wine Technology Division, Agricultural Research Council (ARC), Infruitec-Nietvoorbij, Private Bag X5026, 7599 Stellenbosch, South Africa
[c] Department of Food Science, Stellenbosch University, Private Bag X1, Matieland, Stellenbosch, South Africa
[d] Department of Chemistry and Polymer Science, Stellenbosch University, Private Bag X1, Matieland, Stellenbosch, South Africa
[e] University of Silesia, Institute of Chemistry, Szkolna 9, 400-006 Katowice, Poland

ANOVA-TP: the algorithm

- Deflate the matrix **X** according to the reduced model:



Hasse diagrams



Crossed design
A,B fixed

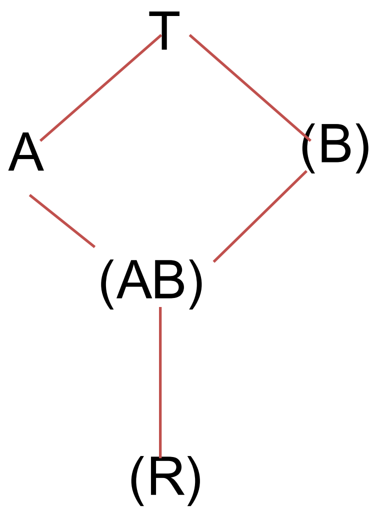
Crossed design
A fixed, B random

Nested design
A fixed, B random

- Hasse diagrams are a powerful tool to define the proper way of deflating **X** to obtain the “reduced” ANOVA model

Hasse diagrams - 2

- The expected MS for a term U is the sum of the contribution of the term + all eligible random terms below it in the diagram
- All random effects not containing a fixed factor not present in T are eligible



Crossed design
A fixed, B random

$MS_A = A + (AB) + (R)$

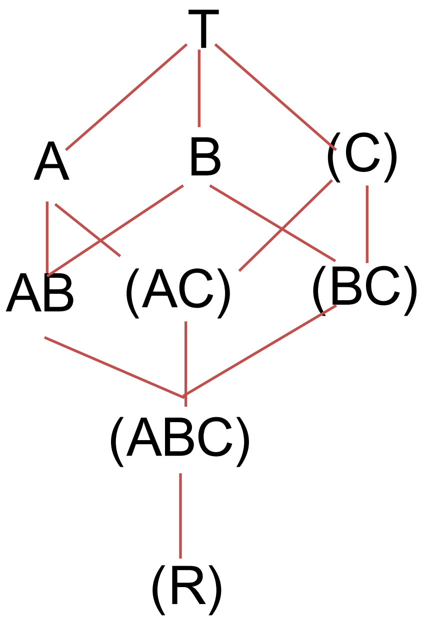
$MS_B = (B) + (R)$

$MS_{AB} = (AB) + (R)$

Reduced model is then calculated by deflating from **X** all other contributions

Hasse diagrams – The reduced ANOVA model

Deflation involves all terms up to the row containing the first eligible random contribution (denominator in a pseudo-F test)

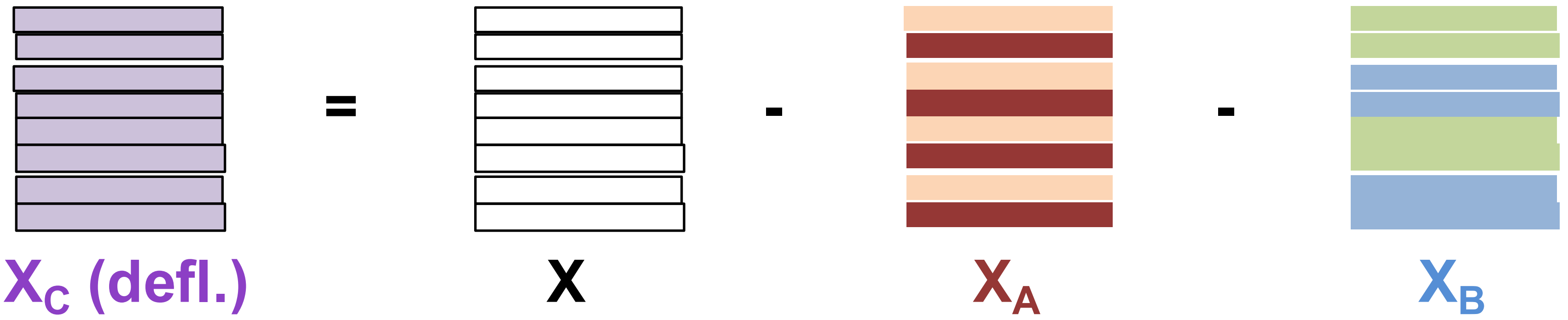


Crossed design
A,B fixed, C random

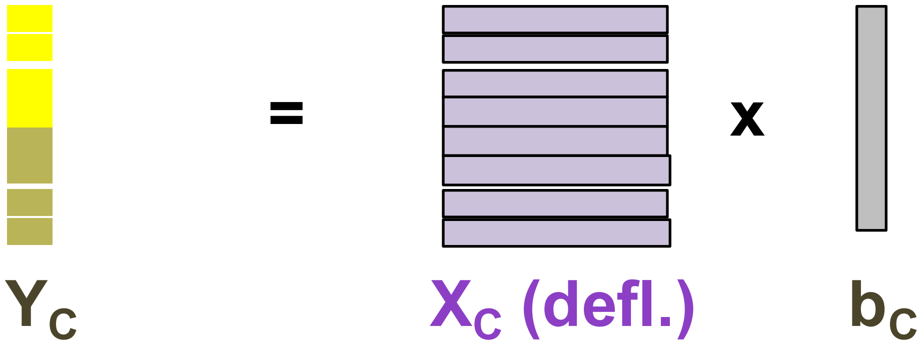
To calculate	Remove (deflate)
A	B, C, AB, BC
B	A, C, AB, AC
C	A, B, AB, AC, BC, ABC
AB	A, B, C, AC, BC
AC	A, B, C, AB, BC, ABC
BC	A, B, C, AB, AC, ABC
ABC	A, B, C, AB, AC; BC

ANOVA-TP: the algorithm

- Deflate the matrix **X** according to the reduced model:

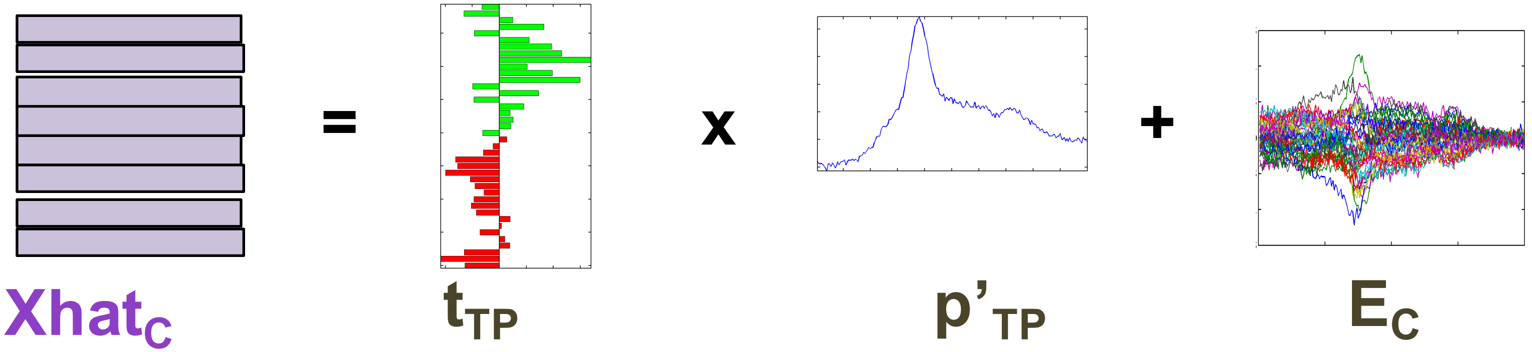


- Calculate a D-PLS model between the deflated matrix and the dummy **Y** coding for the design



ANOVA-TP: the algorithm

- Perform Target projection of the D-PLS solution



PLS and Target projection

- TP reveals the y-relevant variation in the x-variables captured by a multicomponent PLS model on a single latent variable.

$$\mathbf{w}_{TP} = \frac{\mathbf{b}}{\|\mathbf{b}\|} \Leftrightarrow \mathbf{t}_{TP} = \frac{\hat{\mathbf{y}}}{\|\mathbf{b}\|}$$

- TP loadings may be used for interpretation or to build an estimate of the “predictive” part of **X**

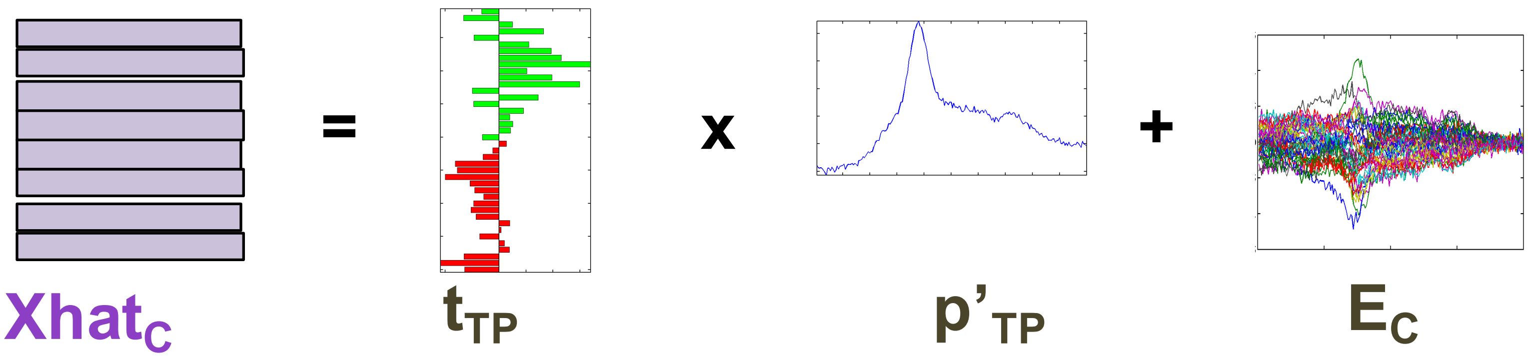
$$\mathbf{p}_{TP} = \frac{\mathbf{X}^T \mathbf{t}_{TP}}{\mathbf{t}_{TP}^T \mathbf{t}_{TP}}$$

$$\mathbf{X} = \mathbf{X}_{TP} + \mathbf{E}_{TP} = \mathbf{t}_{TP} \mathbf{p}_{TP}^T + \mathbf{E}_{TP}$$

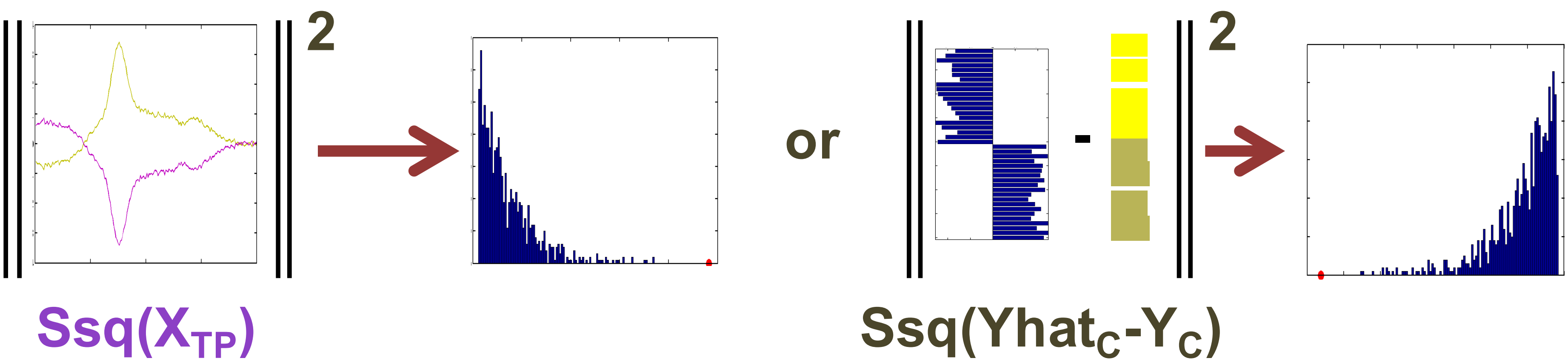
Kvalheim OM, Karstang TV. Chemom. Intell. Lab. Syst. **7** (1989) 39–51

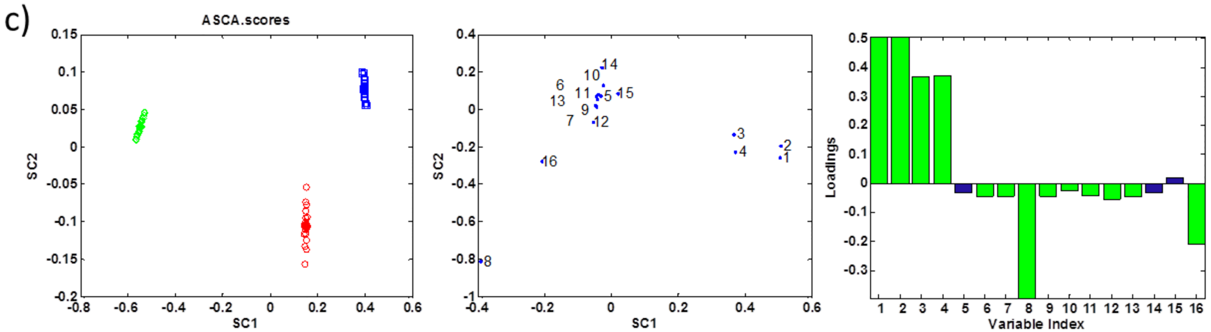
ANOVA-TP: the algorithm

- Perform Target projection of the D-PLS solution

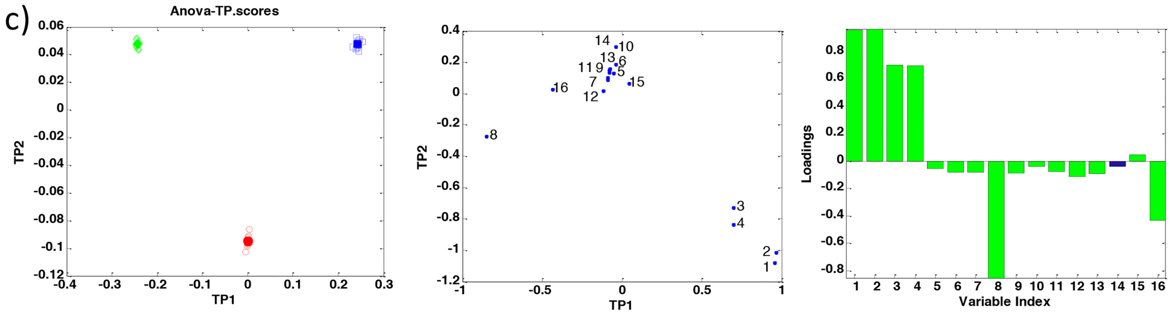


- Evaluate the effect of the factor/interaction and its significance





ASCA



ANOVA-TP

Validation of ASCA revisited

A novel approach based on confidence ellipsoids

Received: 19 May 2017 | Revised: 19 October 2017 | Accepted: 17 November 2017
DOI: 10.1002/cem.2990

RESEARCH ARTICLE

WILEY Journal of CHEMOMETRICS

Confidence ellipsoids for ASCA models based on multivariate regression theory

Kristian Hovde Liland[1,2] | Age Smilde[3] | Federico Marini[4] | Tormod Næs[1]

Confidence ellipsoids for LS means

- Based on the theory of multivariate multiple regression, it is possible to calculate the $1-\alpha$ confidence interval around the LS mean according to:

$$\frac{(n-q)}{na_i}\left(\mathbf{x}_{i,*}^t\widehat{\mathbf{B}}-\mathbf{x}_{i,*}^t\mathbf{B}\right)\hat{\Sigma}^{-1}\left(\mathbf{x}_{i,*}^t\widehat{\mathbf{B}}-\mathbf{x}_{i,*}^t\mathbf{B}\right)^t=T_{1-\alpha,p,n-q}^2$$

- Where:
  - p is the number of variables in **Y**
  - n is the number of observations
  - q is the number of colums in **X**
  - $a_i=\mathbf{x}_{i,*}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_{i,*}$
  - $Cov(\mathbf{e}_{i,*})=\mathbf{\Sigma}$

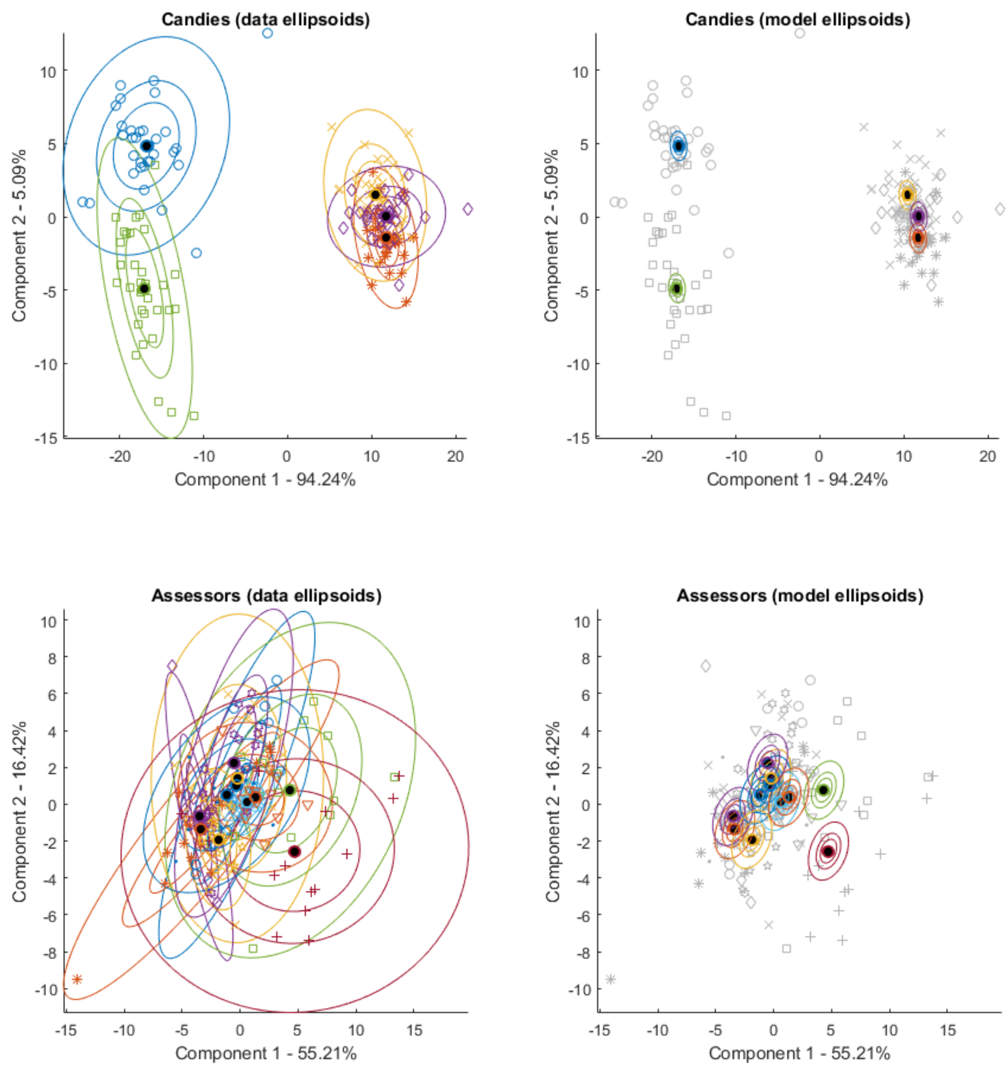
Confidence ellipsoids for single design terms (factors or interactions) in projected space

- These equations can be applied as well to the sub-matrices corresponding to individual effects $\mathbf{X}^{(J)}\mathbf{B}^{(J)}$

$$\frac{(n-q_J)}{na_i^{(\mathrm{J})}}\left(\mathbf{x}_{i,*}^{t(J)}\,\widehat{\mathbf{B}}^{(J)}\mathbf{L}^t-\mathbf{x}_{i,*}^{t(J)}\,\mathbf{B}^{(\mathrm{J})}\mathbf{L}^t\right)\left(\mathbf{L}\widehat{\boldsymbol{\Sigma}}\mathbf{L}^t\right)^{-1}\left(\mathbf{L}\left(\mathbf{x}_{i,*}^{t(J)}\,\widehat{\mathbf{B}}^{(J)}\right)^t-\mathbf{L}(\mathbf{x}_{i,*}^{t(J)}\,\mathbf{B}^{(\mathrm{J})})^t\right)$$
$$=T^2_{1-\alpha,d,n-q_J}.$$

- Where:
  - **L** is the loading matrix
  - *d* is the number of rows in **L** (dimensions of projected space)
  - *n* is the number of observations
  - q_J is the number of colums in $\mathbf{X}^{(J)}$
  - $a_i^{(\mathrm{J})}=\mathbf{x}_{i,*}^{t(J)}\left(\mathbf{X}^{t(J)}\mathbf{X}^{(J)}\right)^{-1}\mathbf{x}_{i,*}^{(J)}$
  - $Cov\left(\mathbf{e}_{\boldsymbol{i},*}\right)=\boldsymbol{\Sigma}$

Examples:



Candies (data ellipsoids)
Candies (model ellipsoids)
Component 2 - 5.09%
Component 1 - 94.24%
Assessors (data ellipsoids)
Assessors (model ellipsoids)
Component 2 - 16.42%
Component 1 - 55.21%

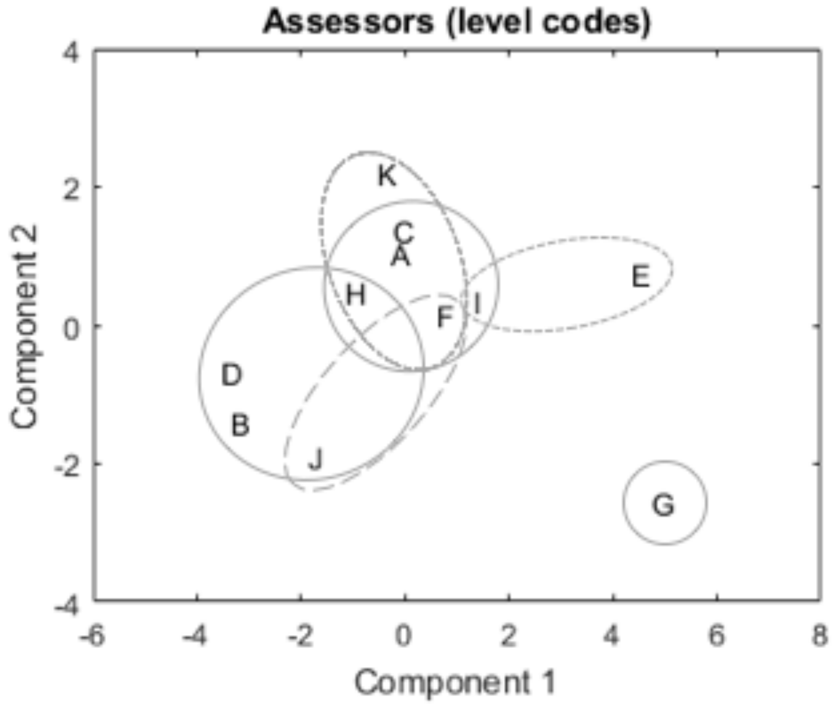
Pairwise comparison of factor levels:

- Based on the theory it is possible to perform pairwise comparisons of effect levels by replacing $\mathbf{x}_{i,*}^{t(J)}\,\widehat{\mathbf{B}}^{(J)}$ with:

$$(\mathbf{x}_{r,*}^{(J)} - \mathbf{x}_{s,*}^{(J)})^t\,\widehat{\mathbf{B}}^{(J)}$$

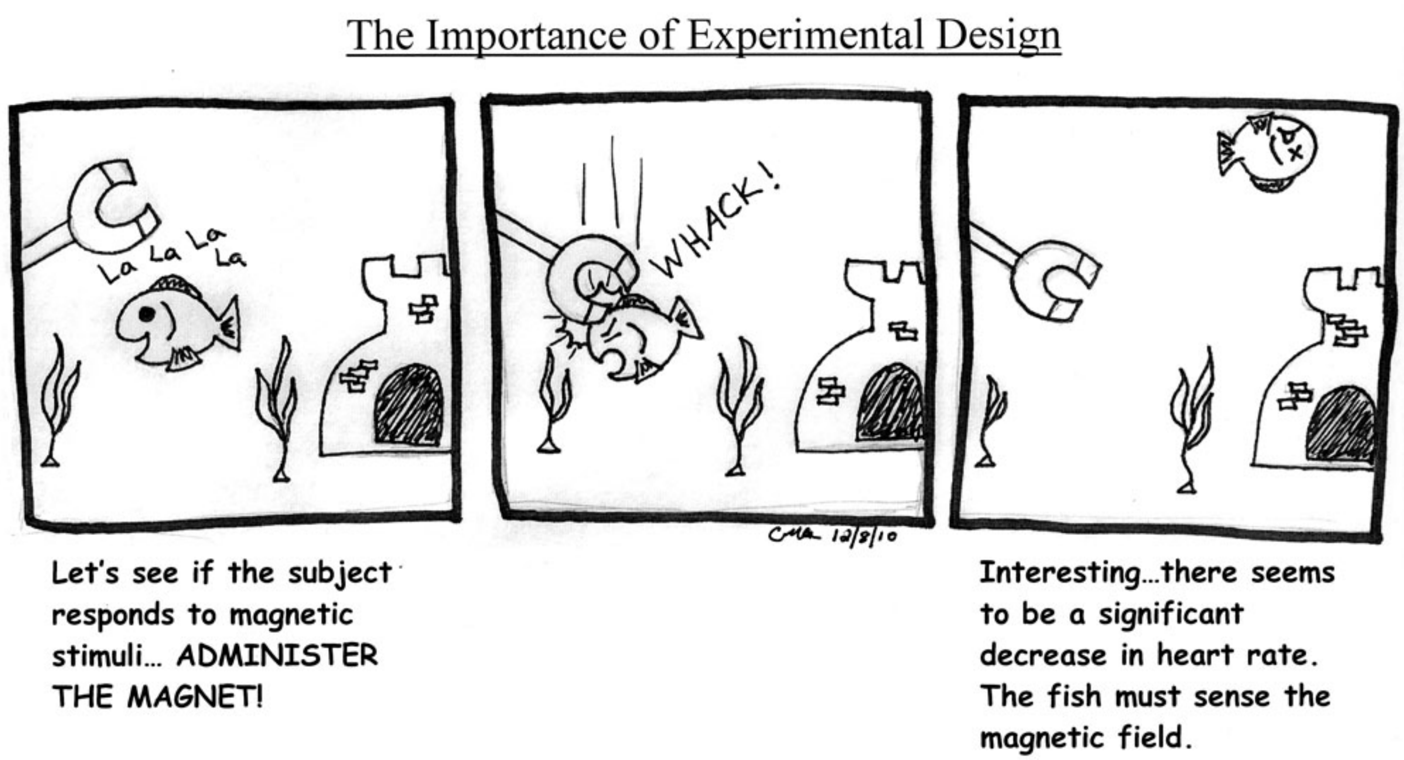
- And verifying if the ellipsoid contains the origin

Assessor	Uncorrected	Bonferroni
E	a	a
I	b	ab
K	c	c
A	cde	bc
C	cd	bc
F	b e	bc e
H	d	bcd
J	f	de
B	g	d
D	g	d
G	h	f



CONCLUSIONS AND PERSPECTIVES

- Designed experiments call for appropriate methods
  - Partitioning of overall variance
  - Significance testing
  - Interpretation
- Possible alternatives:
  - rMANOVA
  - AMOPLS
  - …



Thanks for your attention