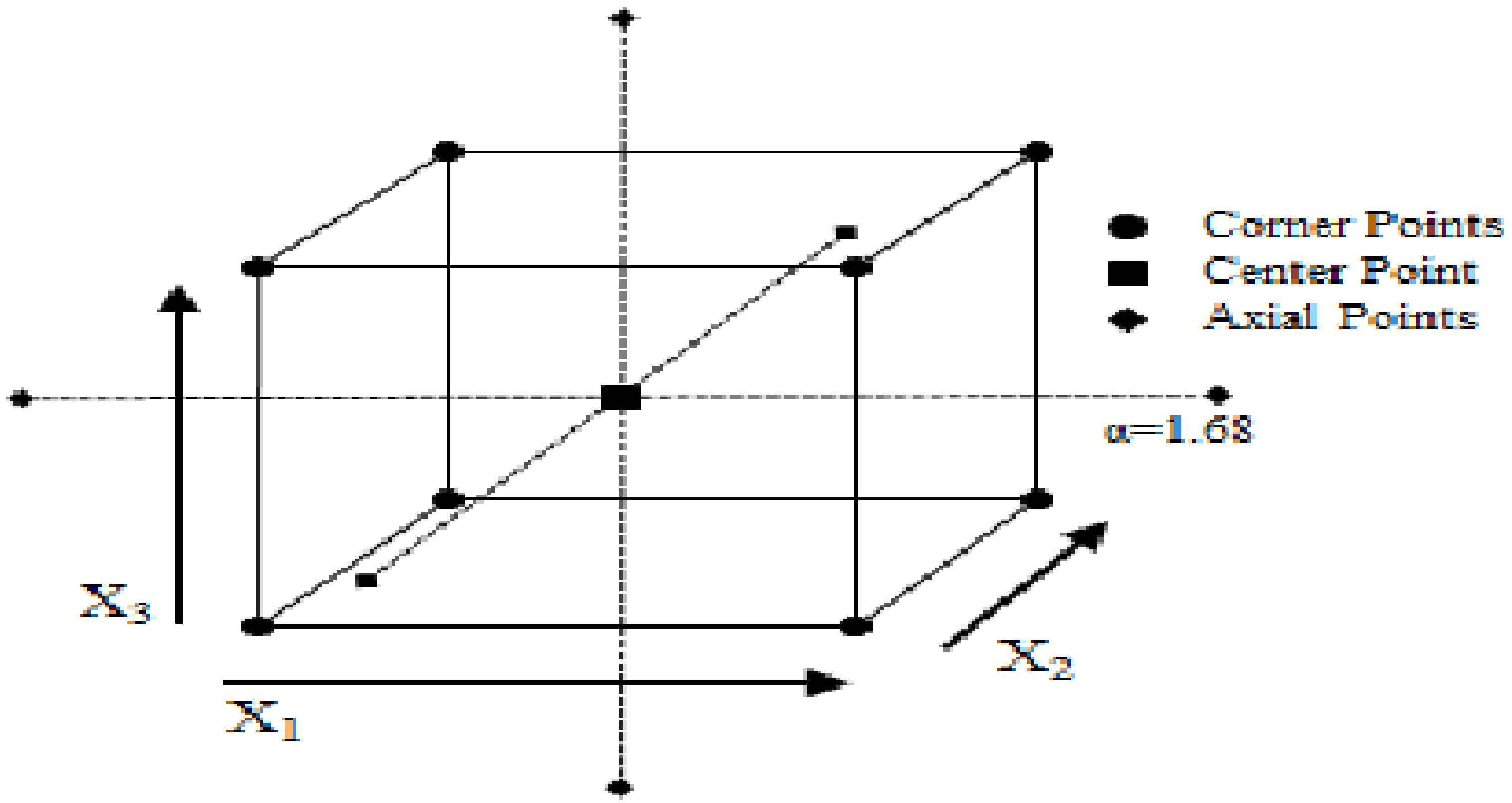
Central composite design

In experimental design strategy, a **central composite design** is an experimental design, useful in creating relatively sophisticated models, usually, a second order (quadratic) model for the output function without needing to use other designs, for instance a complete three-level factorial design (requiring more experimental work).

After the designed experiment is performed, the calculation algorithm follows the regression strategy, sometimes iteratively, to obtain results. Coded variables are often used when constructing this design.

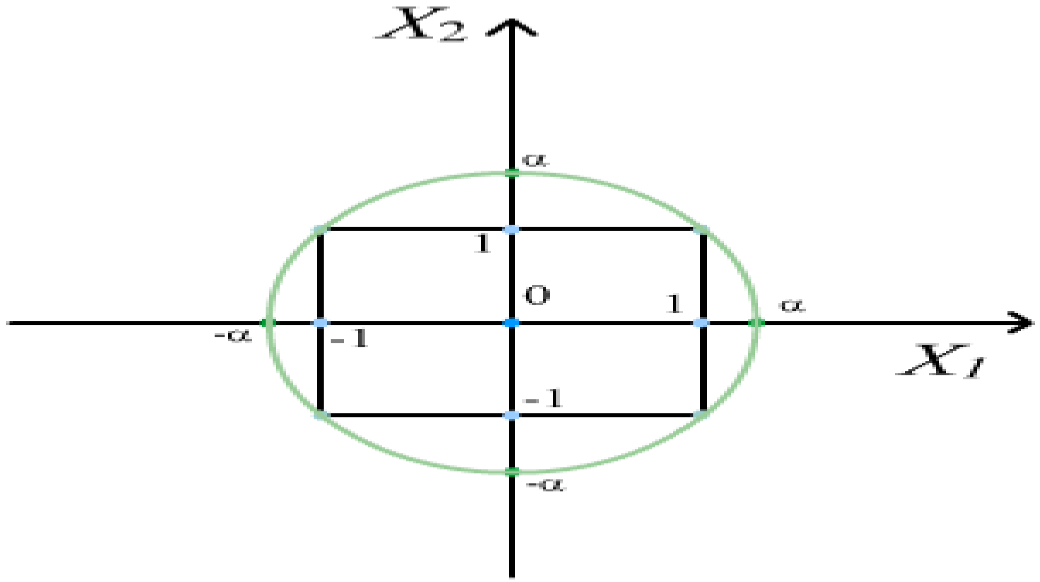
Graphical presentation



Legend to the figure

- The figure presented in the previous slide demonstrates the design for 3 input factors as central composite design;
- The total number of required experiments is 15 as follows:
- 8 experiments according to the scheme of FFD 2^3 (totally 8);
- 6 experiments at so called “star” or “axial” points;
- 1 experiment in the center of the design (center point).

The choice of various “points” for the design



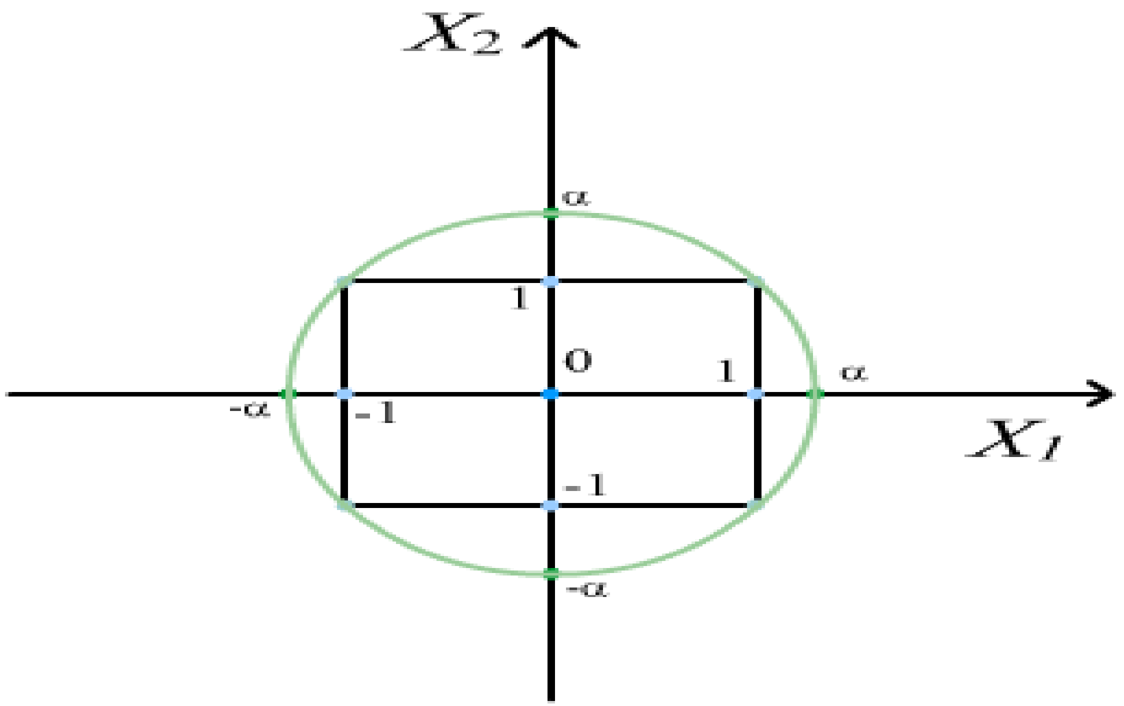
Corner points

- Previous slide shows the simplest central composite design for two input factors – it needs totally 9 experimental points;
- Corner points: these are the experiments according to the design of the type full factorial design at two levels of variation of the factors;
- The choice of the intervals of variation for the factors is an important task for the scientist and, as we already know, depends on preliminary experience and information; the factors are respective coded as “+1” and “-1”;
- These are the experiments at the corners of the quadrat (4 points).

Axial or star points

- These points are usually marked by “α”;
- They are located symmetrically around the design center and the value of α depends on the number of input factors involved (this value is determined *a prioi* and could be found in statistical tables (software products); the experiments at the axial points are carried out with coordinates (0, α) or (0, -α) and for 2 factors the total number of axial experiments is 2x2 = 4; this part of the central composite design is the “superstructure” to the full factorial design.

The central point of the design marked by 0



Some clarifications

- The central experiment is performed when the input factors have as coordinates the middle of the interval of variation (0, 0);
- In the scale of the coded input variables the corner experiments have as coordinates combinations of +1 and -1; the axial experiments – combinations of 0 and α (the α value for two factors is 1.4, so the real value for the experiment could be easily calculated); the center experiment was considered above.
- All necessary steps for carrying out the real experiments are as in case of full factorial experiment – assessment of experimental error, randomization etc.

The statistical model

- It fits a complete quadratic model;
- The model is checked for error homogeneity, significance of the regression coefficients and validity (comparison between calculated by the model and experimentally found value)

$$Y = a_0 + \sum_{i=1}^{n} a_i\, x_i + \sum_{i=1}^{n} a_{ii}\, x_{ii}^2 + \sum_{i,j=1}^{n} a_i\, a_j x_i x_j$$

Optimization experiments

- Very often the final goal of a research (experimental) study is to reach an extreme (minimal or maximal) value of the response (output) function, e.g. the response has to be optimized;
- Possible actions are:
- Mapping experiments (experiments around the maximum);
- Simplex optimization;
- Box – Wilson gradient methods (steepest slope approach).

“Mapping” experiments approach

- After carrying out some kind of experimental design the research will reach a maximal value of the response for certain combination of input factor levels; Is this the real optimum?
- In order to check this assumption single-at-a-time experiments around the condition accepted as “optimal” for the response are organized and performed; it is easy to find out if higher (or lower) values as compared to that by the design are obtained as a result of the mapping.

Graphical example

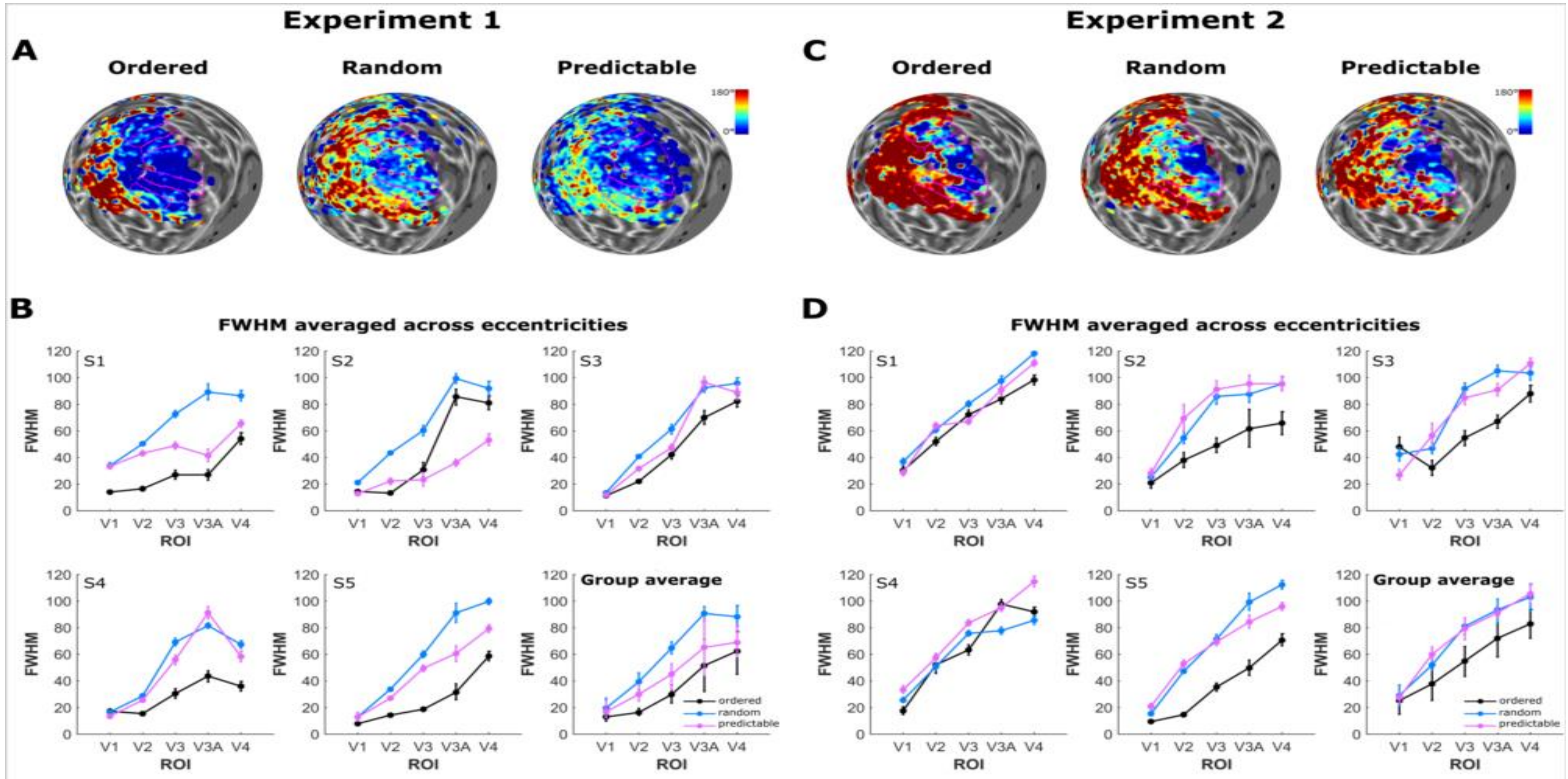


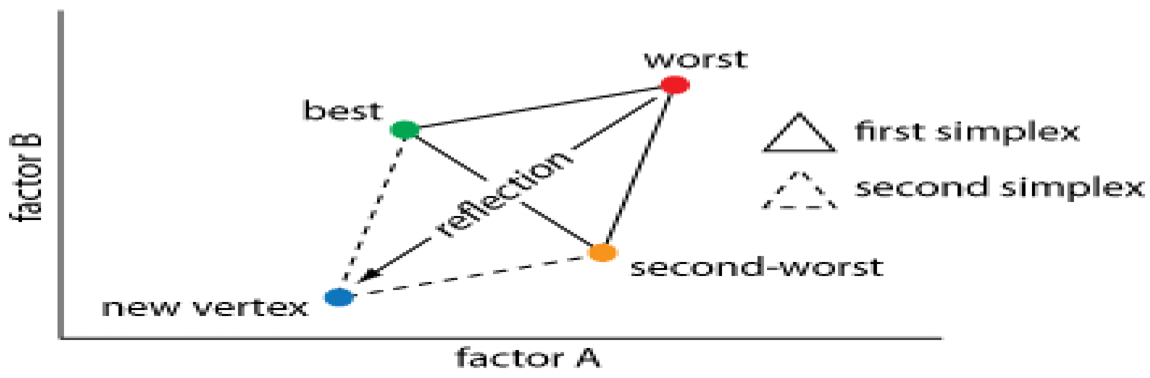
Figure 2. Full-Width Half-Maximum estimates different mapping conditions in Experiment 1 and 2. (A, C) Example FWHM maps for ordered, random and predictable sequences for the same (left) hemisphere of one participant in Experiment 1 (A) and Experiment 2 (C). (B, D) Individual and group average FWHM estimates for different sequence conditions, averaged separately for areas V1-V4 in Experiment 1 (B) and Experiment 2 (D).

Simplex optimization procedure

- What is a **simplex**? Geometric figure with number of apexes equal to the number of input factors + 1; it means that the simplex for 2 factors has three vertexes, i.e. it is triangle, for 3 factors is a pyramid and for more factors is a topological figure which could not be drawn on the plane of the sheet;
- In order to carry out experiments following the algorithm for Simplex optimization coordinates of the initial Simplex are determined (by the researcher according his/her goals and information).

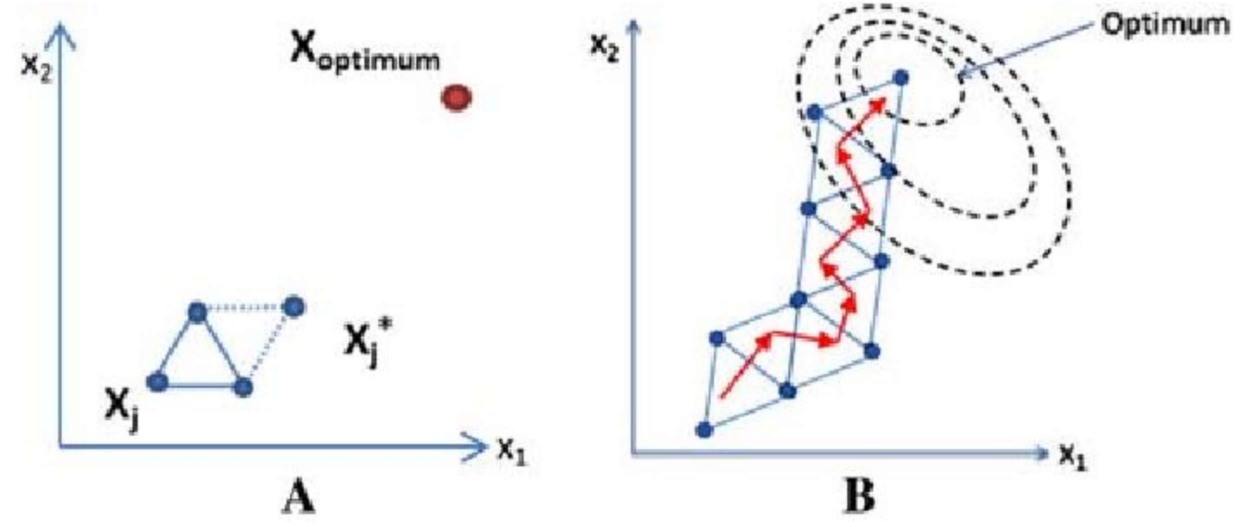
The movement of the Simplex towards optimum

- Let’s consider “triangle” case (2 factors) – three initial experiment at the apexes; three responses – the worst one is eliminated and by reflection a new Simplex is formed (the coordinates of the newly introduced point are easily calculated.



Next steps for Simplex

- The movement continues using the same algorithm for the next simplexes:

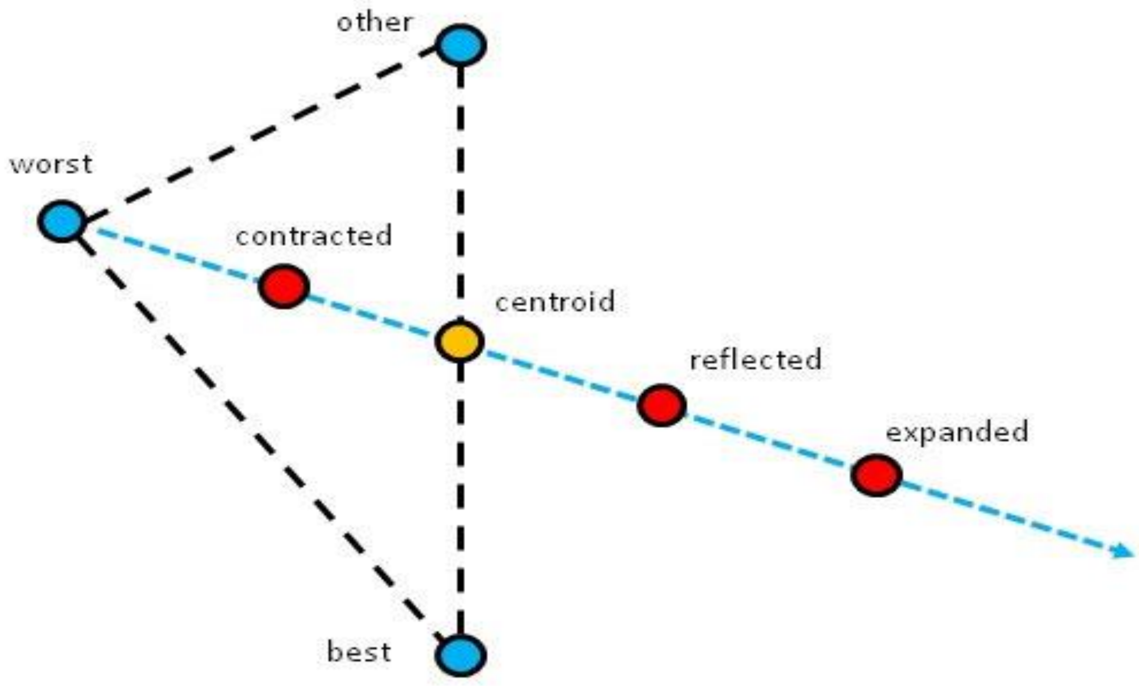


Basic rules

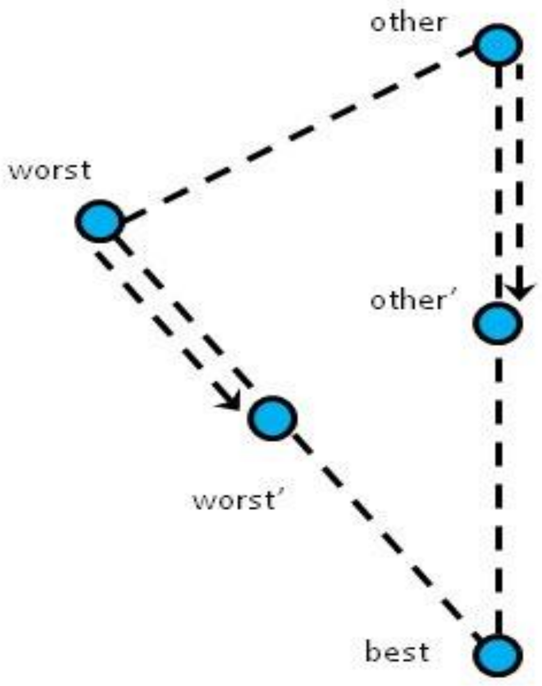
- The movement of the simplex needs some basic rules:

1. Rotation – the simplex could stop at certain point and starts rotation around its centroid; the direction of movement has to be changes by elimination of the second worst response;
2. At a certain point the Simplex could jump over the coordinate system (movement in not allowed space, e.g. negative concentration); then the direction should be changed by using the second worst response;
3. Appearance of one and the same vertex in three consecutive simplexes (probably the region of optimum is achieved)
4. There are options for acceleration of the movement (expansion) or delay (contraction).

Graphical presentation of the rules



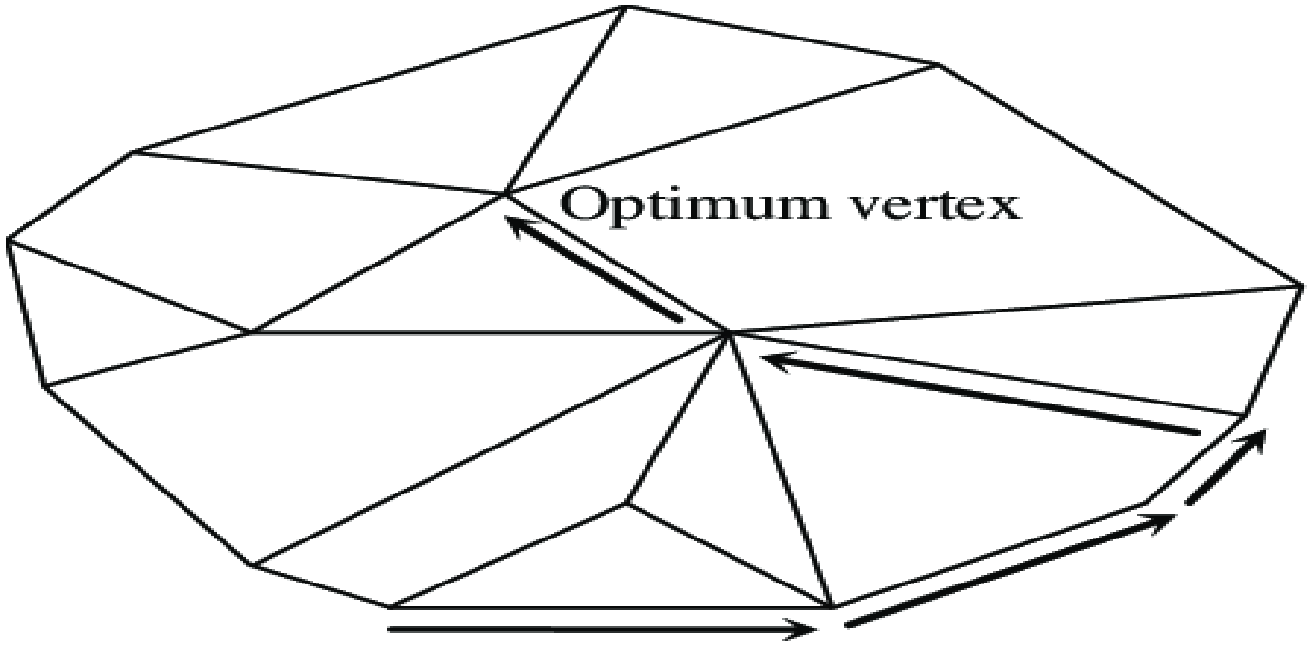
Three candidates to replace worst



Shrinking

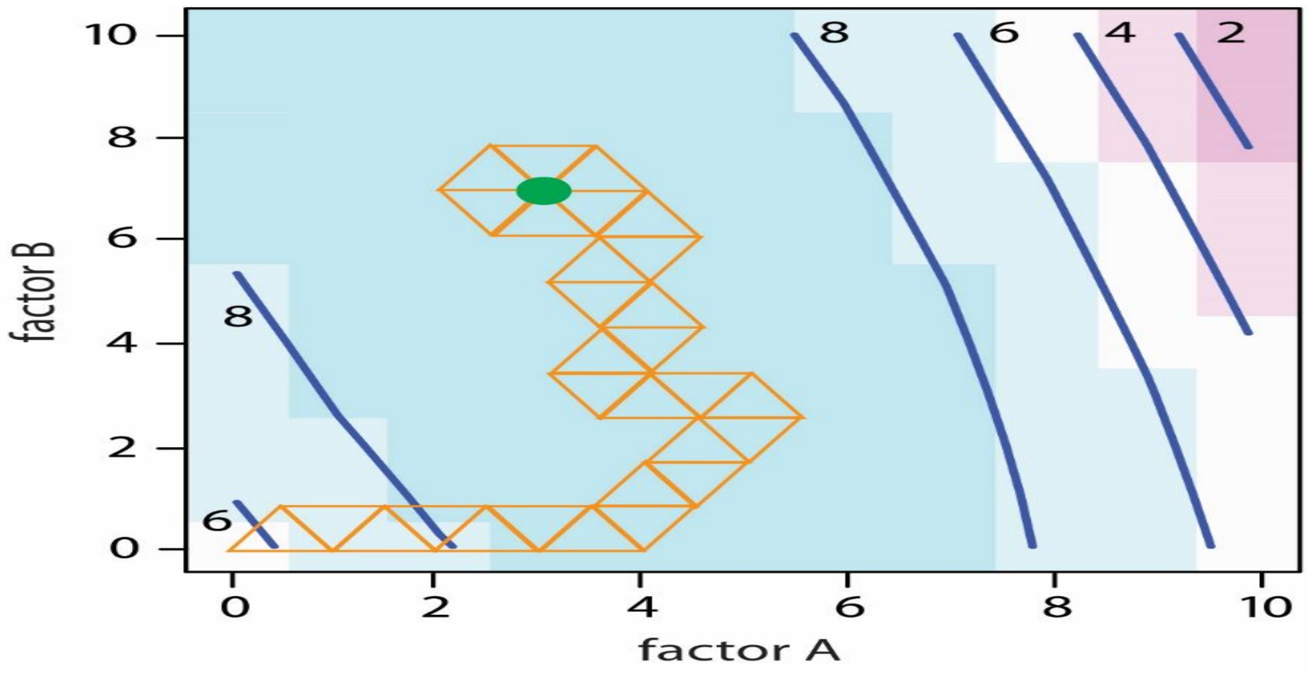
Optimal vertex or area of optimum

- It has to be kept in mind that the Simplex could miss the optimal vertex due to many reasons; that is why a satisfactory outcome is the location of area of optimum, plateau.



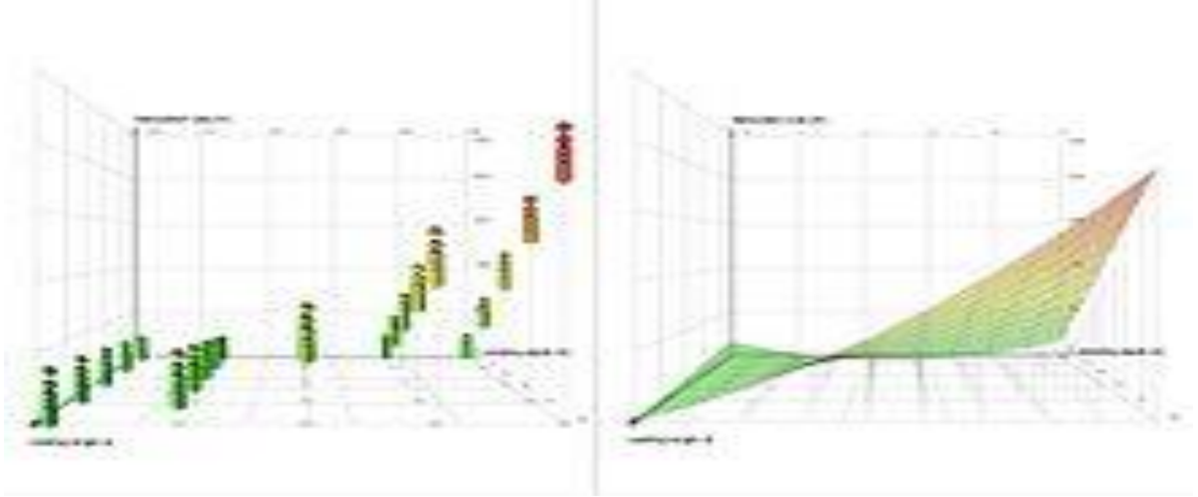
Is this the real optimum?

- One good option to check if the real optimum is achieved is to start a simplex with different initial coordinates. If the same peak or region of optimum is reached, then the final goal is validated.



Box – Wilson gradient method for optimization

- In statistics, **response surface methodology (RSM)** explores the relationships between several input factors and one or more output functions. The main idea of RSM is to use a sequence of designed experiments to obtain an optimal response. The authors acknowledge that this model is only an approximation, but they use it because such a model is easy to estimate and apply, even when little is known about the process.



Steepest descent method

Steepest Descent Method

If minimization is desired then we call this technique the "method of steepest descent".

$$\bar{y} = \beta_o + \sum_{i=1}^{k} \bar{\beta}_i \, x_i$$

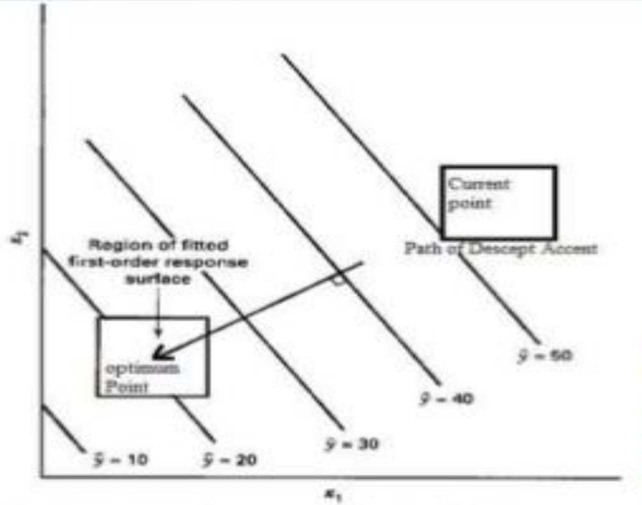


Figure 11-4 First-order response surface and path of steepest ascent.

Steepest ascent method

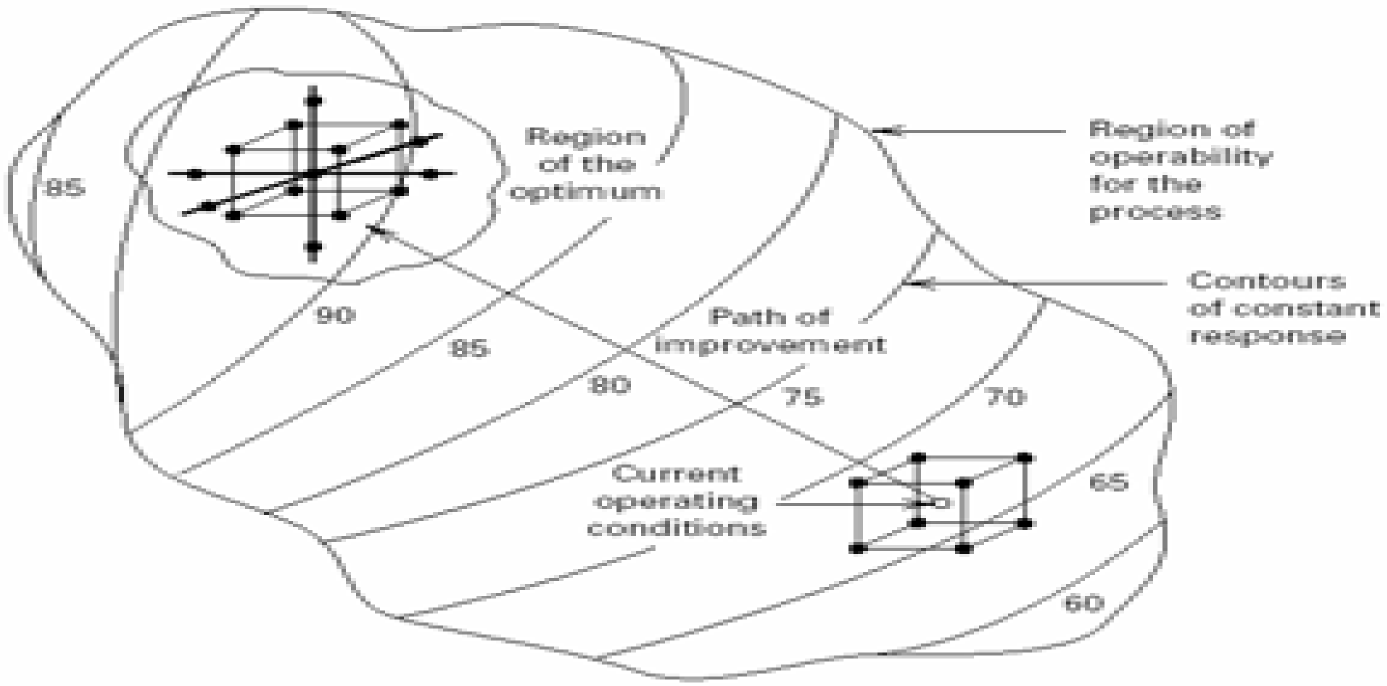
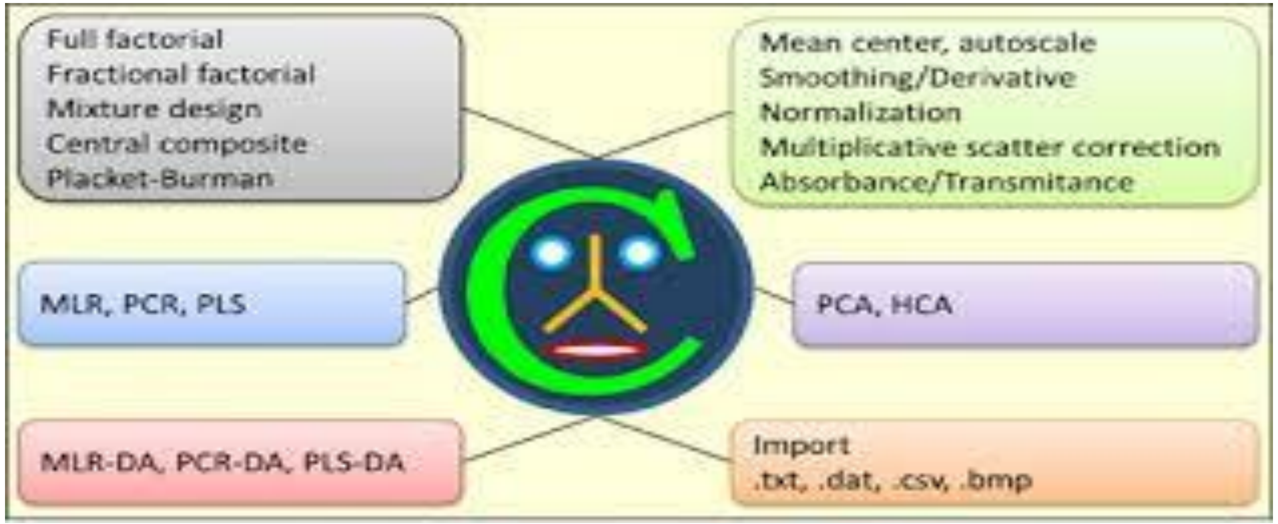


Figure 11-3 The sequential nature of RSM.

General Steps of the method

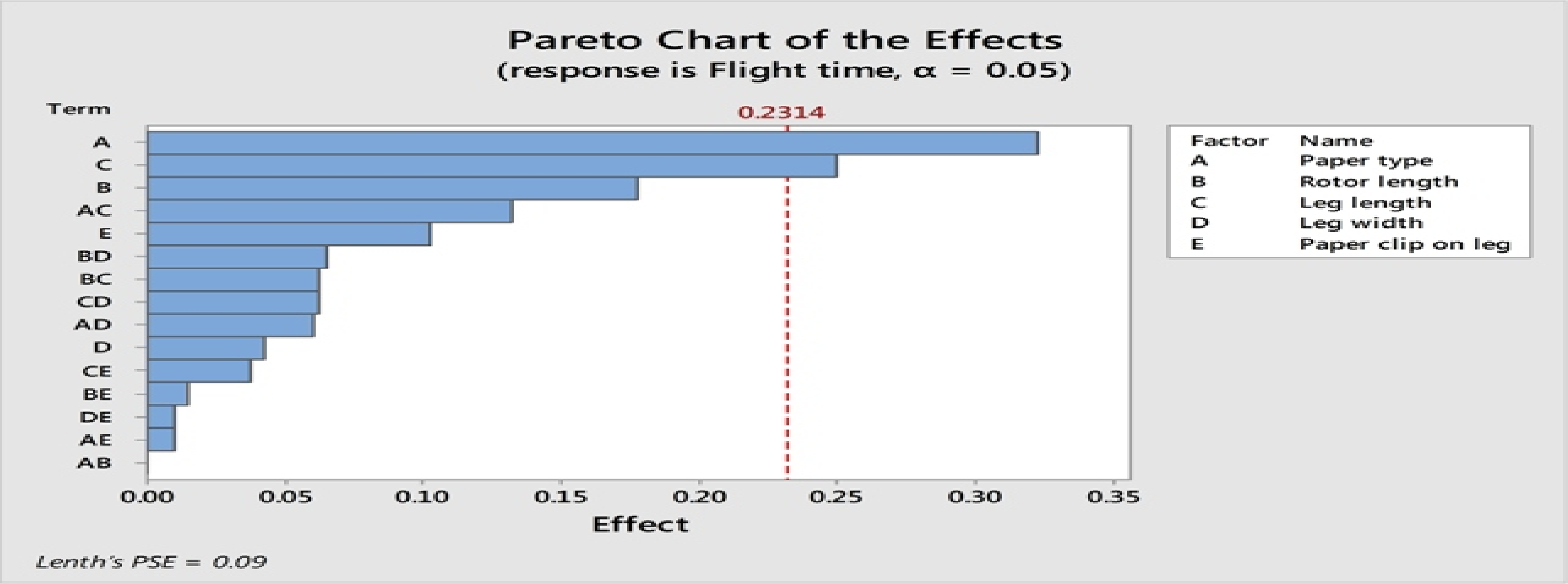
- 1. Full factorial design
- 2. Determination of the steepest ascent – the factor with the most significant weight
- 3. Determination of the new intervals of variation with respect to steepest ascent and carrying out next FFD
- 4. Mental or real experiments toward the trend of optimum
- 5. Change of direction (sign of the regression coefficient), if necessary
- 6. Location of optimal region

Simple software package CHEMOFACE



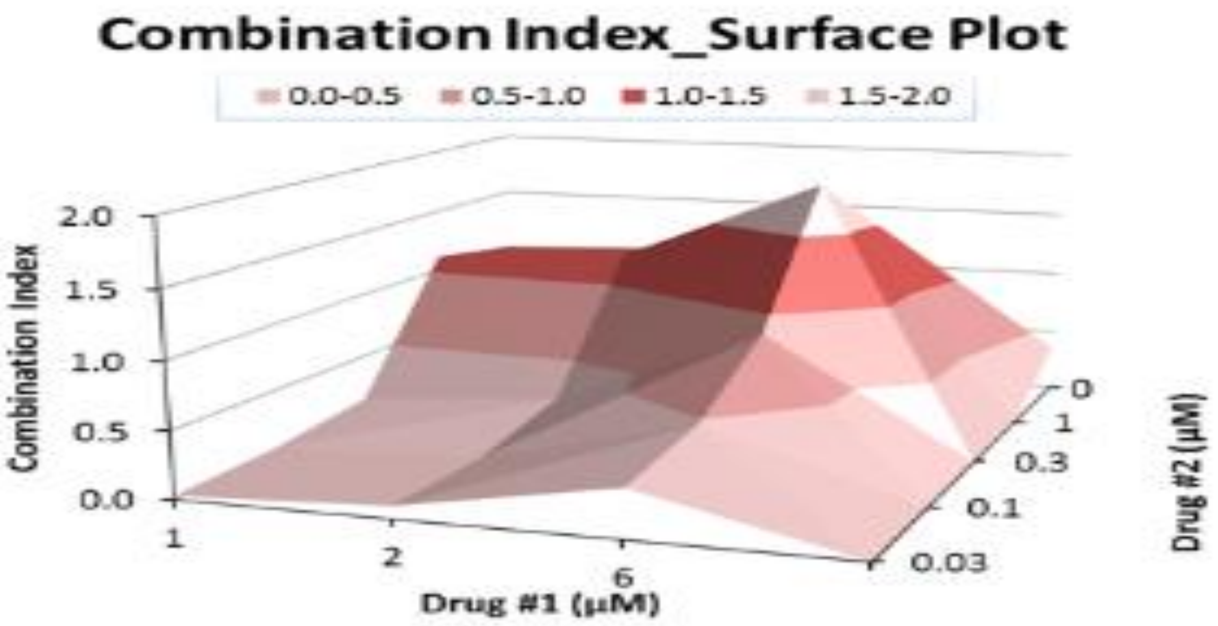
Some interesting outputs from Chemoface

- Pareto charts to indicate the significance of the input factors

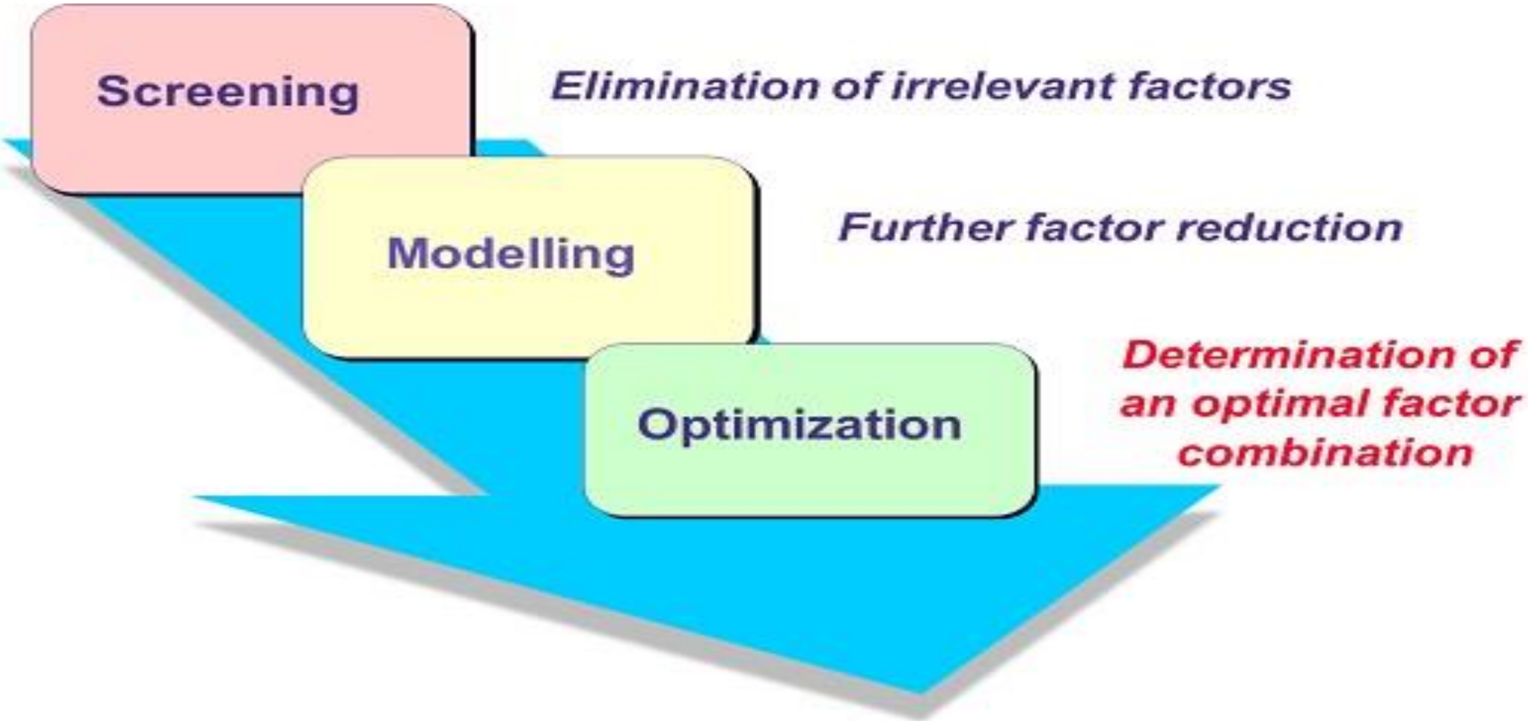


More outputs…

- Response surface



Before starting with designs



The road to multivariate statistics