UNIVERSITÀ DEGLI STUDI DI PADOVA

DEPARTMENT OF INDUSTRIAL ENGINEERING

Design of Experiments
Lesson #10

Academic year 2025-2026

Prof. Pierantonio Facco

CAPE-Lab, Computer-Aided Process Engineering Laboratory

Email: pierantonio.facco@unipd.it

URL: https://research.dii.unipd.it/capelab/

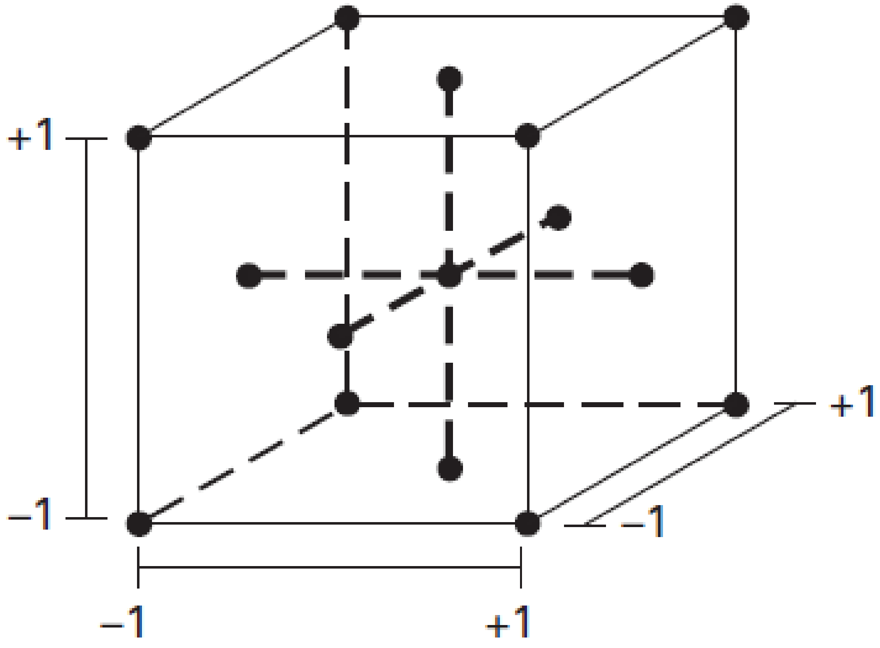
Recap on DoE

task	number of factors	needed resources	selected design	experimental burden	obtained information	response surface model
initial screening/ discovery	very high	few	**fractional factorial**	low	important factors are identified, negligible removed	$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \varepsilon$
in-depth analysis	intermediate	intermediate	**full factorial**	intermediate	interactions are identified	$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_{12} x_1 x_2 + \varepsilon$
process/product optimization	limited	considerable	**central composite design**	high	optimal points are found	$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_{12} x_1 x_2 + \cdots + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \cdots + \varepsilon$

Product/process optimization: central composite designs

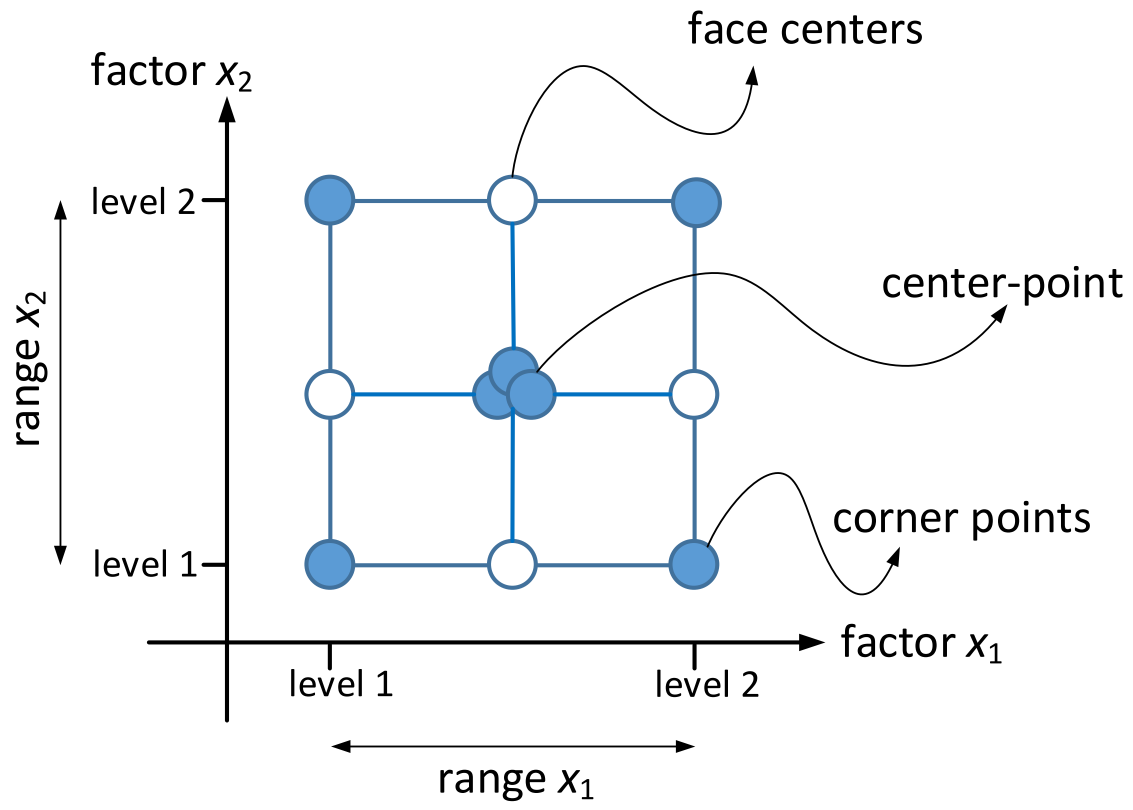
Central composite designs

- **Central composite designs** are suitable designs for **optimization purposes**
  - a larger number of experiments allows larger number of estimated parameters and a more complex version of the regression model, i.e., quadratic terms
  - cubic design with the center of the cube and of the faces (a total of $2^K + 2K + n_C$ experiments)
    - circumscribed in a spherical domain
    - faced in a cubic domain
  - the best method to estimate the nonlinear effects
  - in Matlab®: `ccdesign`



Face-centered central composite design

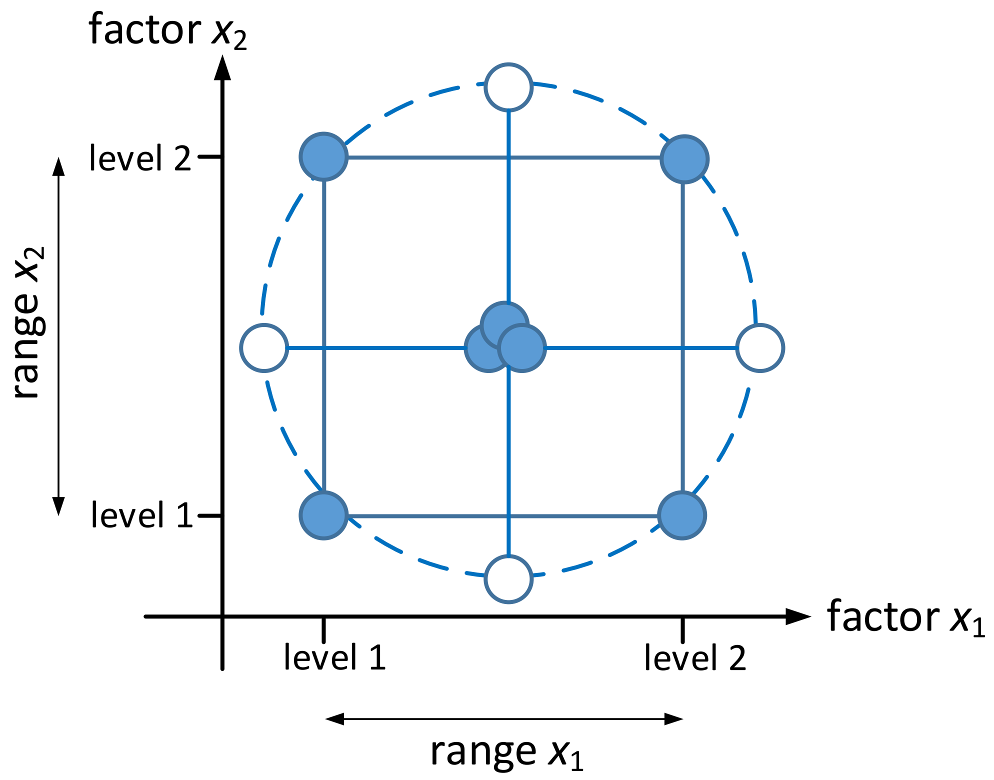
- **Face-centered central composite design** considers the corners of the domain, the center point and the centers of the faces of the experimental domain



factor x_1	factor x_2
-1	-1
-1	1
1	-1
1	1
-1	0
1	0
0	-1
0	1
0	0
0	0
0	0

Circumscribed central composite design

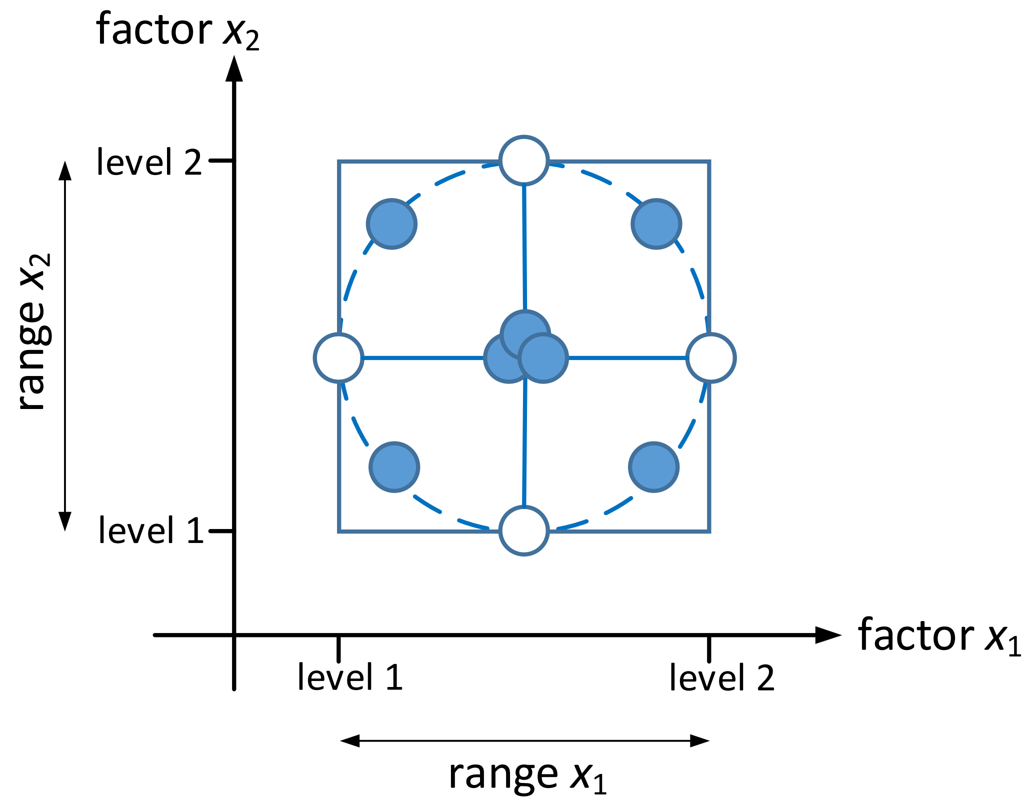
- **Circumscribed central composite design** considers the corners of the domain, the center point and the projection of the face centers of the experimental domain in a circle which is circumscribed to the experimental domain



factor x_1	factor x_2
-1	-1
-1	1
1	-1
1	1
-1.41421	0
1.41421	0
0	-1.41421
0	1.41421
0	0
0	0
0	0

Inscribed central composite design

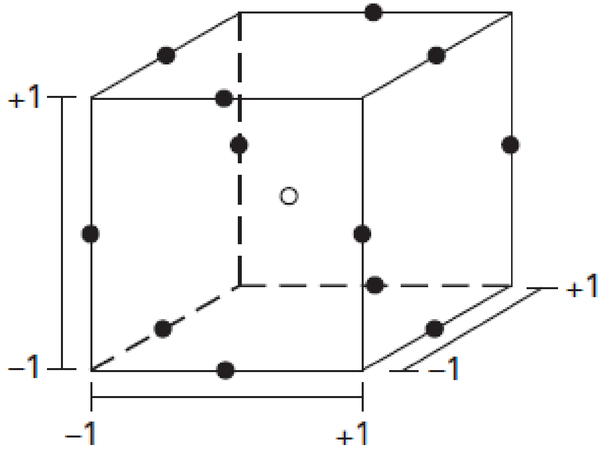
- **Inscribed central composite design** considers the center point and the face centers of the experimental domain and the projections of the corner points in a circle which is inscribed to the experimental domain



factor x_1	factor x_2
-0.707107	-0.707107
-0.707107	0.707107
0.707107	-0.707107
0.707107	0.707107
-1	0
1	0
0	-1
0	1
0	0
0	0
0	0

Efficient alternative for optimization: Box–Behnken design

- **Box–Behnken design** (BBD) is a three-level design for fitting response surfaces for optimization purposes:
  - a sort of complimentary design with respect to face-centered central composite design
- BBDs
  - are usually very **efficient** in terms of the **number of required runs**
    - have fewer design points with respect to CCD
    - can be less expensive to perform than CCD with the same number of factors
  - allow efficient estimation of the first- and second-order coefficients
  - they are not suited for sequential experiments
    - they do not have an embedded factorial design
  - prove useful if you know the safe operating zone for your process, while CCD has axial points outside the design space
  - not all factors are not set at their high levels at the same time
    - convenient if the corners are critical points
    - not convenient if it exclude part of the interesting design space



Run	x_1	x_2	x_3
1	−1	−1	0
2	−1	1	0
3	1	−1	0
4	1	1	0
5	−1	0	−1
6	−1	0	1
7	1	0	−1
8	1	0	1
9	0	−1	−1
10	0	−1	1
11	0	1	−1
12	0	1	1
13	0	0	0
14	0	0	0
15	0	0	0

Finding optimal points

- From the regression model $y = f(\mathbf{x})$ the optimal point can be found, namely the **optimal combination of factors** in vector $\mathbf{x}$ to **maximize/minimize the response** y:
  - **minimum** (of waste, scraps, reworks, etc…)

$$\min_{\mathbf{x}} f(\mathbf{x})$$

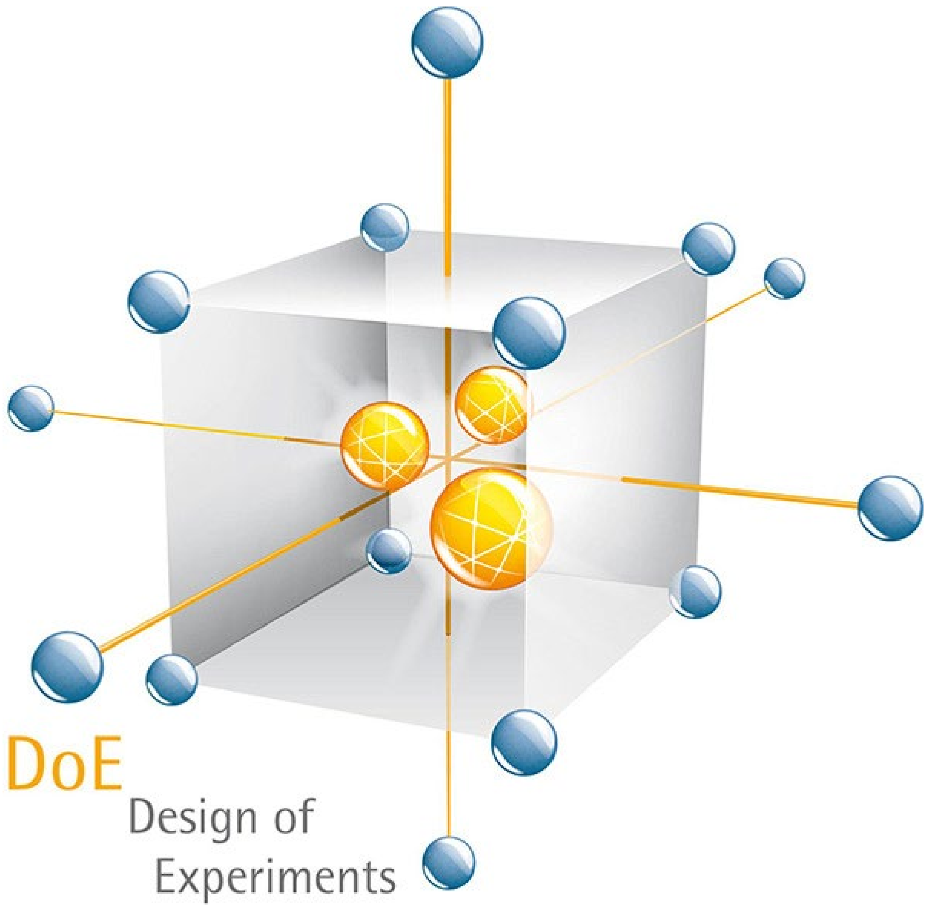
  - **maximum** (of yield, productivity, stability, etc…)

$$\min_{\mathbf{x}}[-f(\mathbf{x})]$$

- Constraints may be utilized for both the factors and the response
- In Matlab® use command: `fmincon`

Concluding remarks

- A lot of types of design exist:
  - Plackett-Burman designs
    - removing non-relevant variables
  - Dohelert design
    - hexagonal domains
  - Optimal designs
  - Multi-criteria decision making
  - Mixture designs
  - etc.

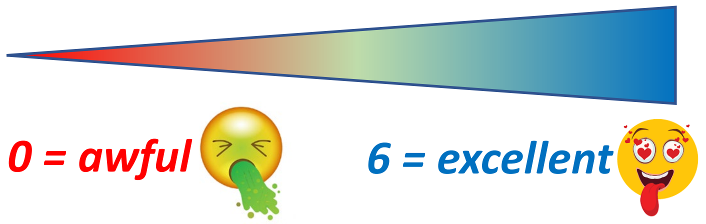


Exercise:
cake recipe and taste

Industrial production of cakes

- The case study is related to a food industry which produces cakes
- Objective: to map a process producing a cake mix to be sold in a box at a supermarket or shopping malls

- Factors:
  - flour
  - shortening
  - egg powder
- Response:
  - taste
    (from a panel)



flour [g]	shortening [g]	egg powder [g]
200	50	50
400	50	50
200	100	50
400	100	50
200	50	100
400	50	100
200	100	100
400	100	100
300	75	75
300	75	75
300	75	75

taste
3.52
3.66
4.74
5.20
5.38
5.90
4.36
4.86
4.73
4.61
4.68

What design is this?

flour [g]	shortening [g]	egg powder [g]
200	50	50
400	50	50
200	100	50
400	100	50
200	50	100
400	50	100
200	100	100
400	100	100
300	75	75
300	75	75
300	75	75

taste
3.52
3.66
4.74
5.20
5.38
5.90
4.36
4.86
4.73
4.61
4.68

Full factorial design with 3 replicates of the center point!

N-way ANOVA

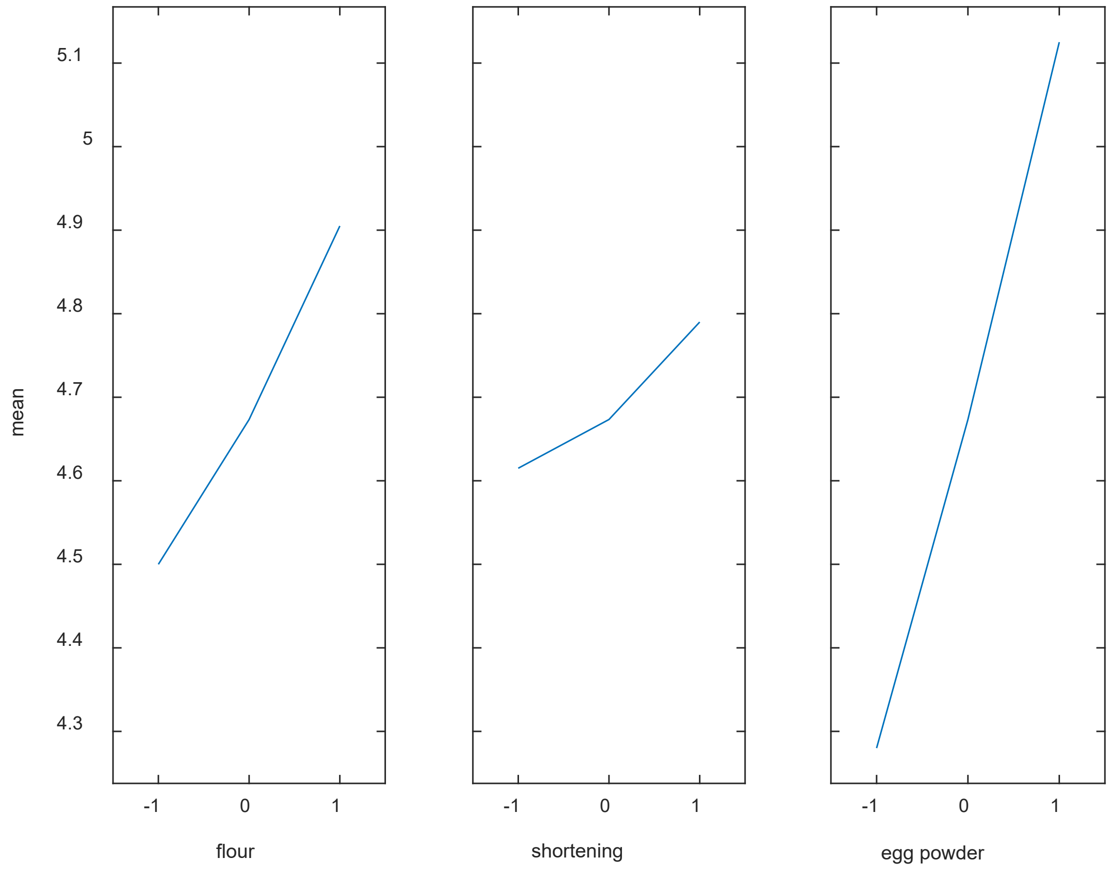
- The ANOVA model shows that:
  - the main effects are related to:
    - **egg powder**
    - **interaction among shortening and egg powder**
  - the reference for the error could be the triplicate of the center point whose
  $$SS_E = 0.0036$$
    - the flour and shortening have valuable effects on the taste, as well

Analysis of Variance

Source	Sum Sq.	d.f.	Mean Sq.	F	Prob>F
X1	0.32805	1	0.32805	Inf	NaN
X2	0.06125	1	0.06125	Inf	NaN
X3	1.42805	1	1.42805	Inf	NaN
X1*X2	0.01125	1	0.01125	Inf	NaN
X1*X3	0.02205	1	0.02205	Inf	NaN
X2*X3	2.90405	1	2.90405	Inf	NaN
X1*X2*X3	0.01445	1	0.01445	Inf	NaN
Error	-0	0	-0		
Total	4.76915	7			

Constrained (Type III) sums of squares.

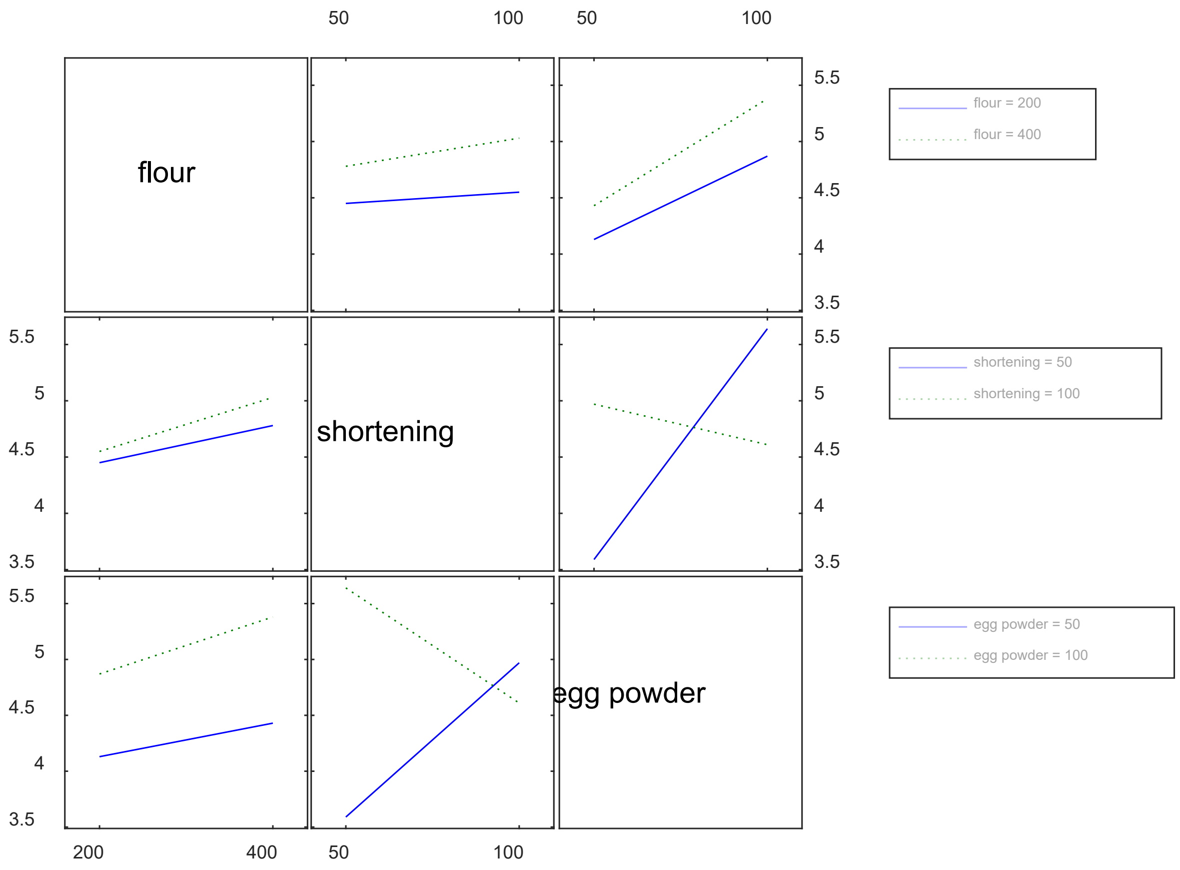
Main effects



- The main effect plot confirms that:
  - the main effect is the one of **egg powder**
  - the effect of shortening is low

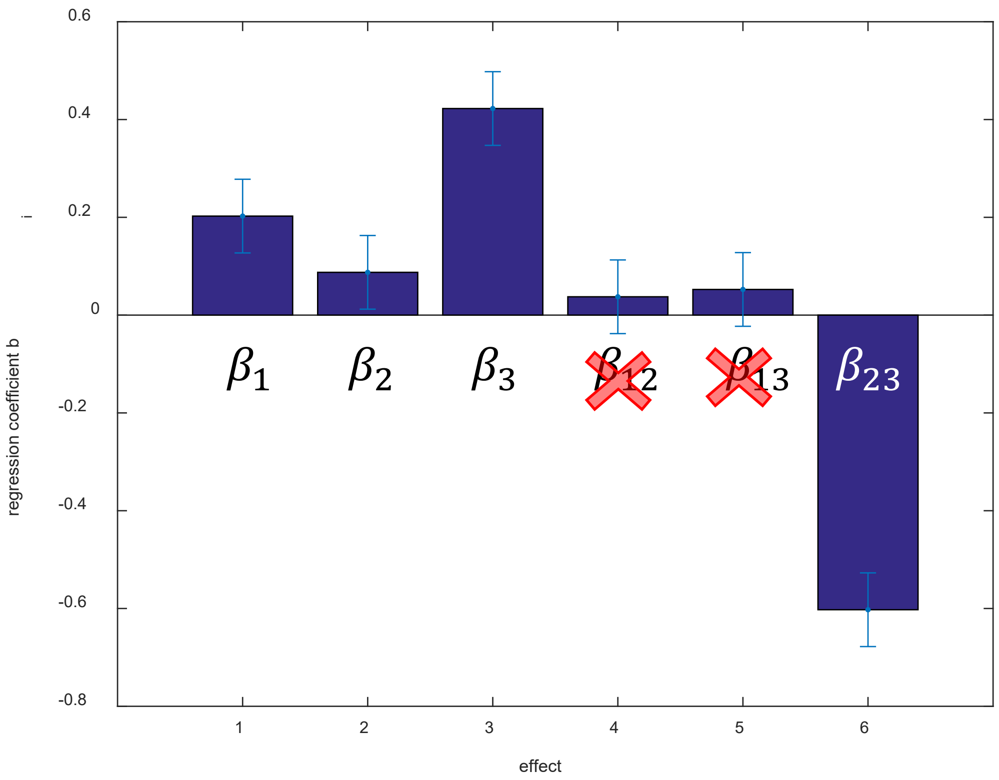
Interactions

- The interaction plot confirms that there is a strong **interaction among shortening and egg powder**



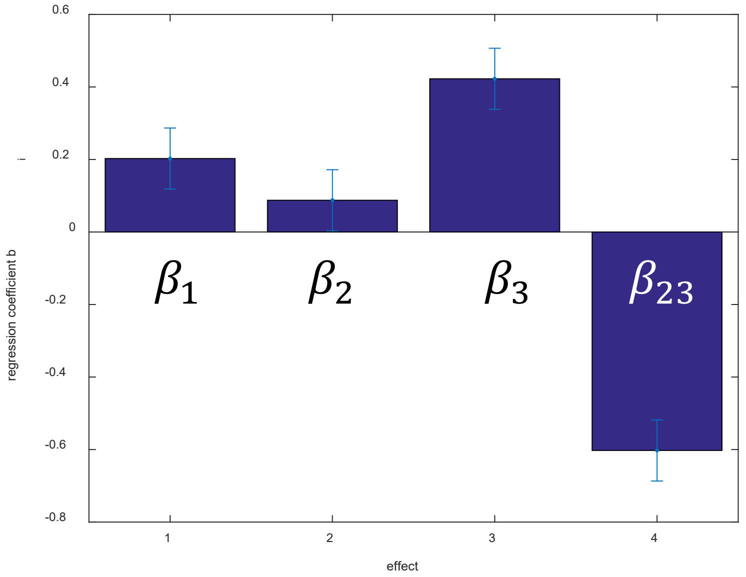
Response surface modelling

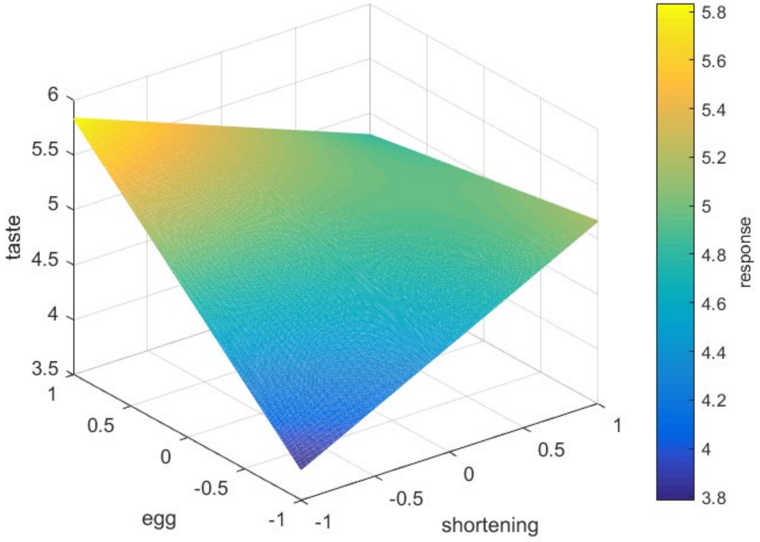
- The bar plot of the regression coefficients confirms the abovementioned results
  - this result suggests to refine the model excluding the interactions between the flour and the other two factors

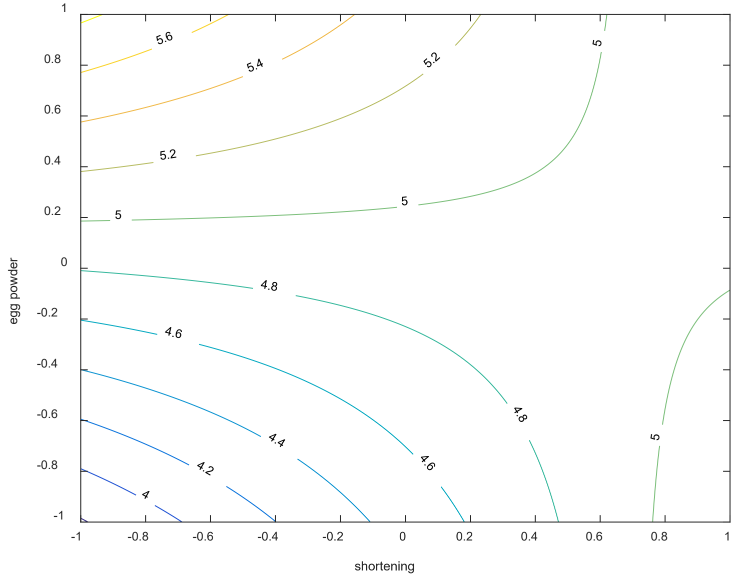


Updated response surface model

- **To improve the cake taste** the main actions could be:
  - **increase the flour content**
  - **increase the egg powder content**
  - **decrease the shortening content**







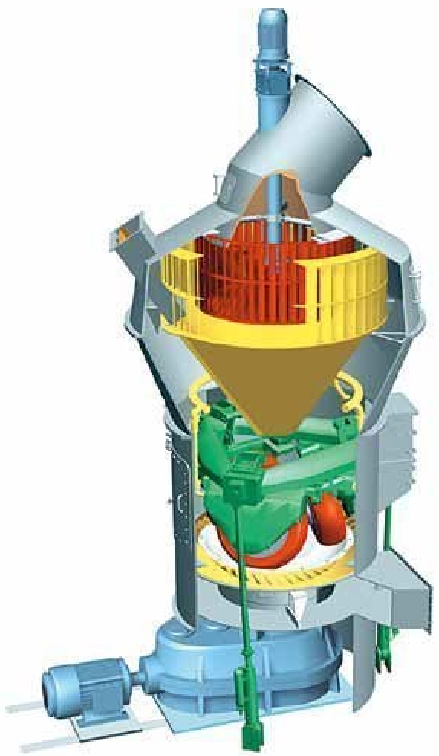
Example on multiple responses: coal mill

Coal milling

- Process: **coal milling**

- Objective: obtaining optimal operating conditions ensuring that
  - small particles >70%
  - large particles <1%

- Factors:
  - load of the mill [tons/h]
  - position of the classifier (curtain partially closing the mill)
- Responses:
  - y_1 = % of small particles (<200 mesh)
  - y_2 = % of large particles (>50 mesh)



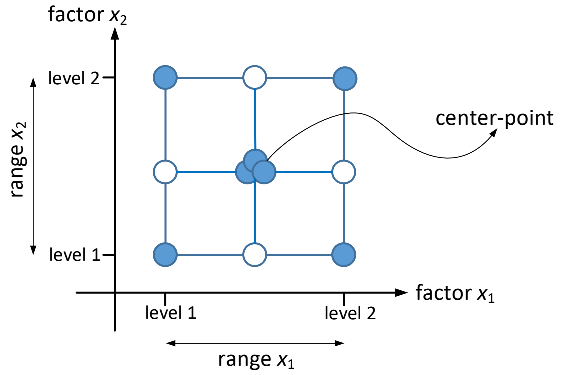
Experimental plan

- The experimental plan is summarized in the following table

load [tons/h]	classifier position
5	1
5	3
15	1
15	3
5	2
15	2
10	1
10	3
10	2

3^2 full factorial design

= face-centered central composite design



Preliminary ANOVA model on small particles

- The ANOVA model shows that:
  - the classifier position seems to be an important variable
  - interactions seem to be limited

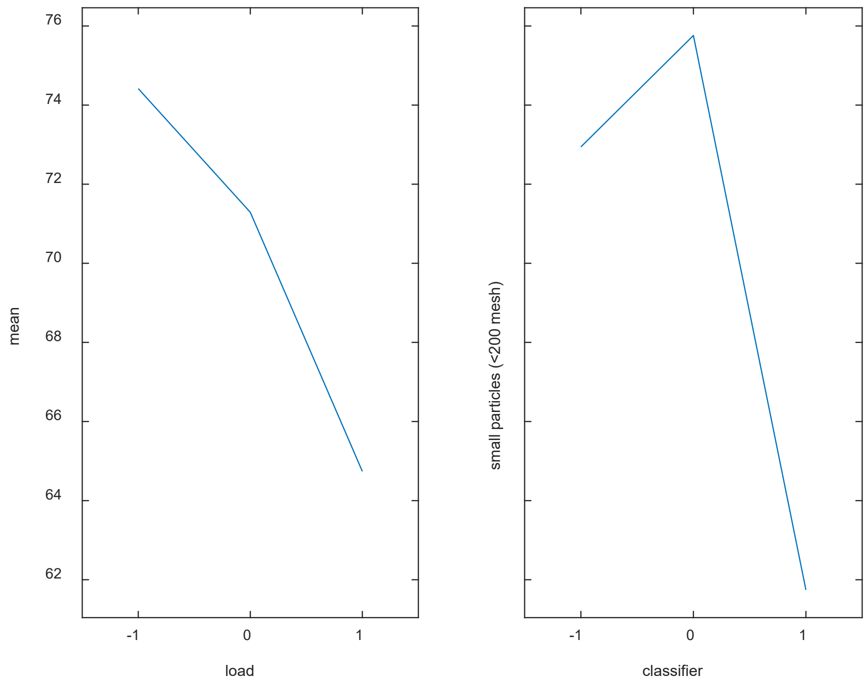
Analysis of Variance

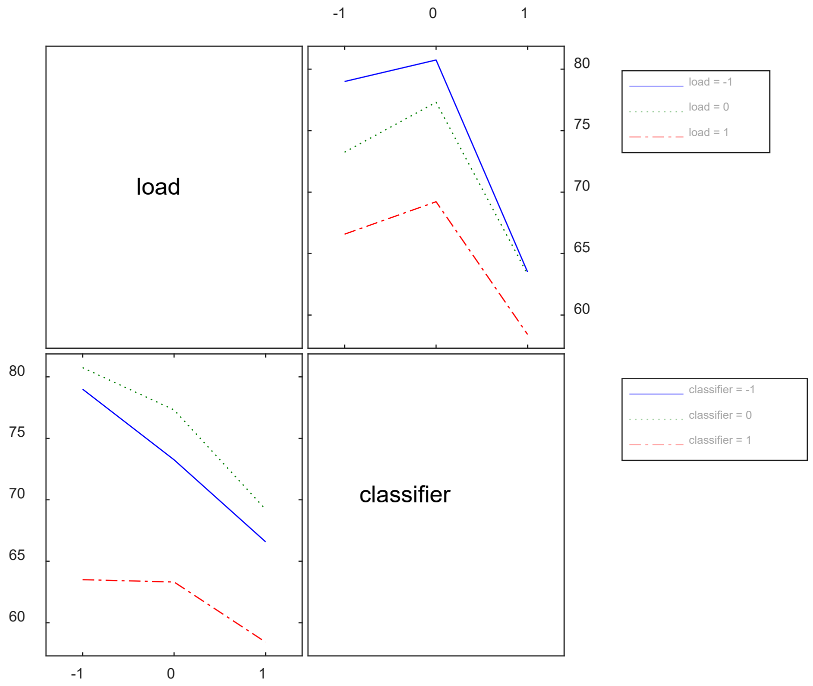
Source	Sum Sq.	d.f.	Mean Sq.	F	Prob>F
X1	146.316	2	73.158	Inf	NaN
X2	329.841	2	164.92	Inf	NaN
X1*X2	17.612	4	4.403	Inf	NaN
Error	-0	0	-0		
Total	493.769	8			

Constrained (Type III) sums of squares.

Main effect and interactions for small particles

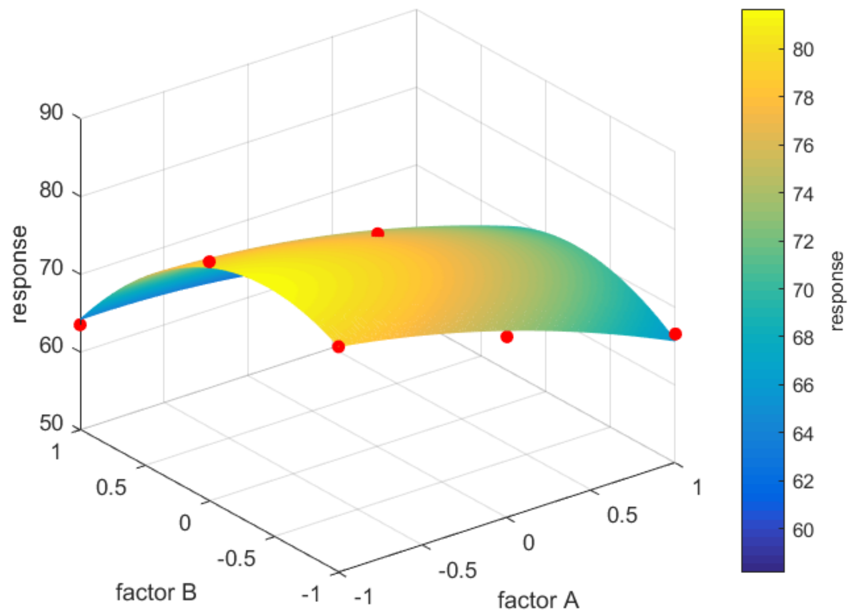
- The previous indications on effect and interactions are confirmed in the main effect and the interaction plots
- The lowest load and the first position of the classifier seem to be related to a good product (high content of small particles)

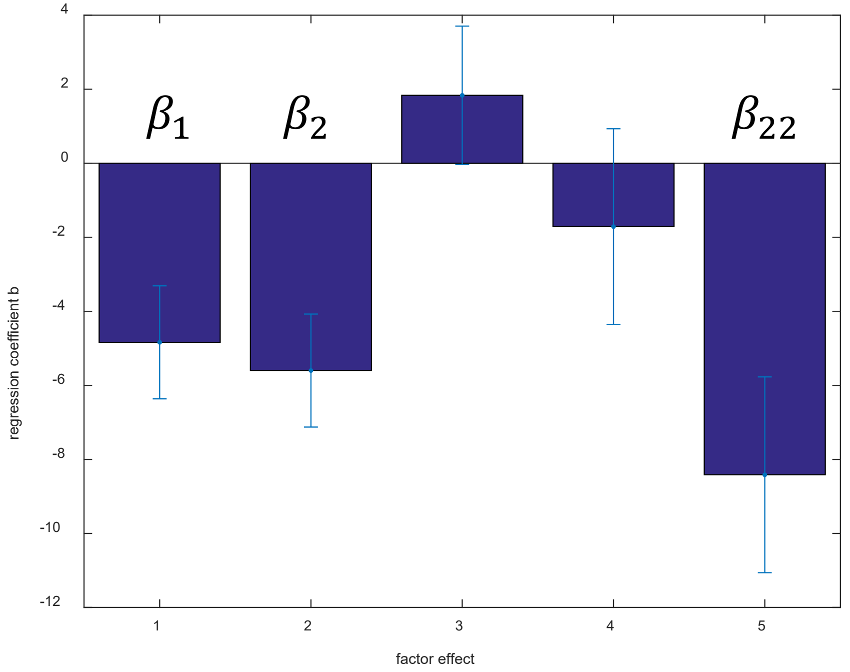




Regression model for small particles

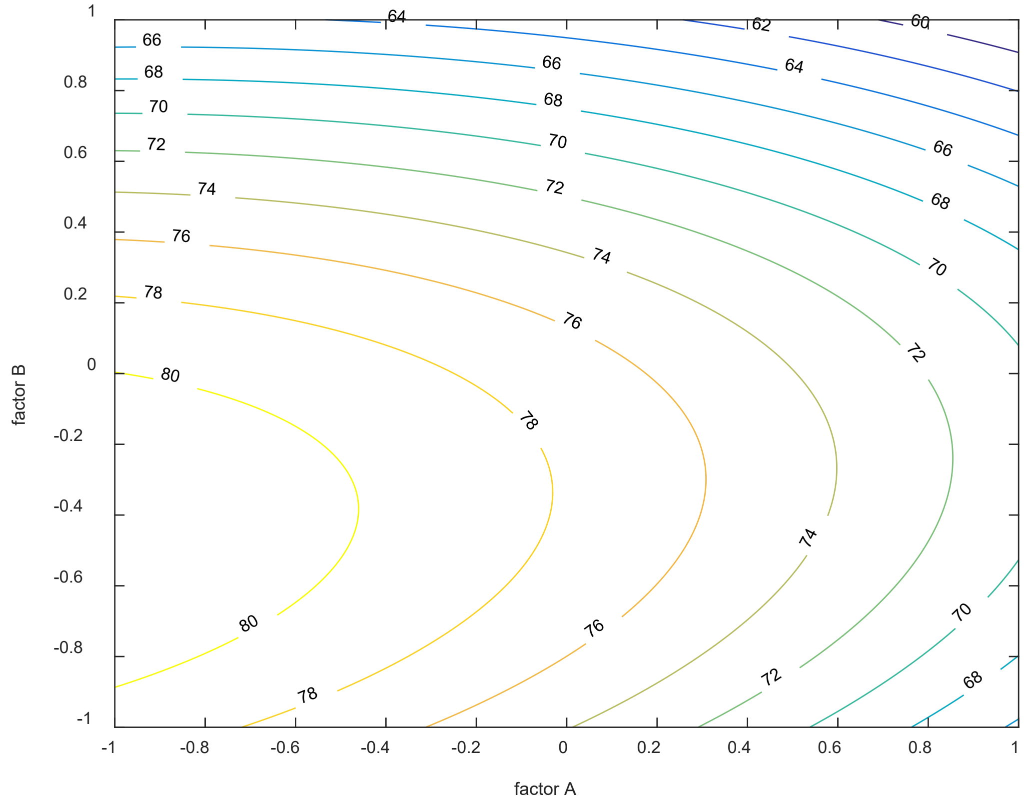
- The selected regression model structure for optimization is: $y_1 = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2$
- The response surface shows that:
  - the fitting provided by the model is very good (response surface very close to the real measurements, i.e., the red points)
  - the regression coefficients confirm that:
    - the **load and the classifier position are the most important effects**
      - lower loads and the first position of the classifier are related to the highest content of small particles in the product
    - the **interaction is not important for small particles**
    - there is a significant effect of the **squared value of the classifier position**





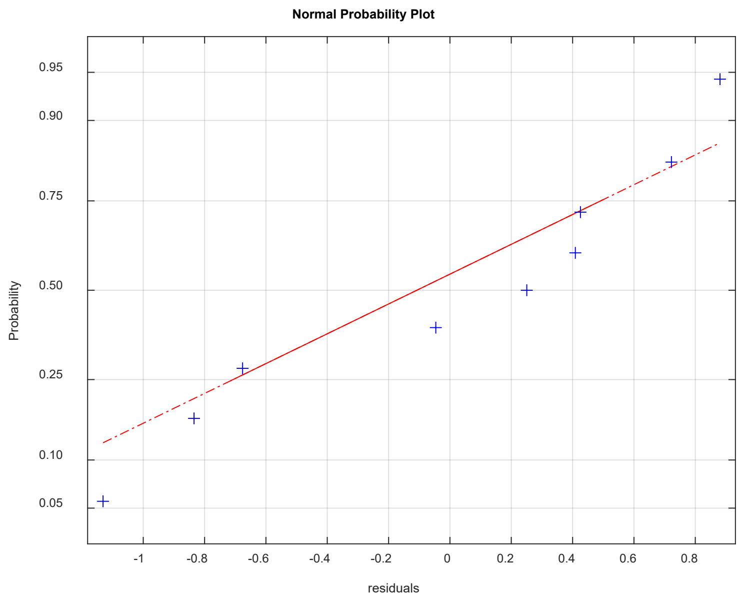
Optimal point for small particles

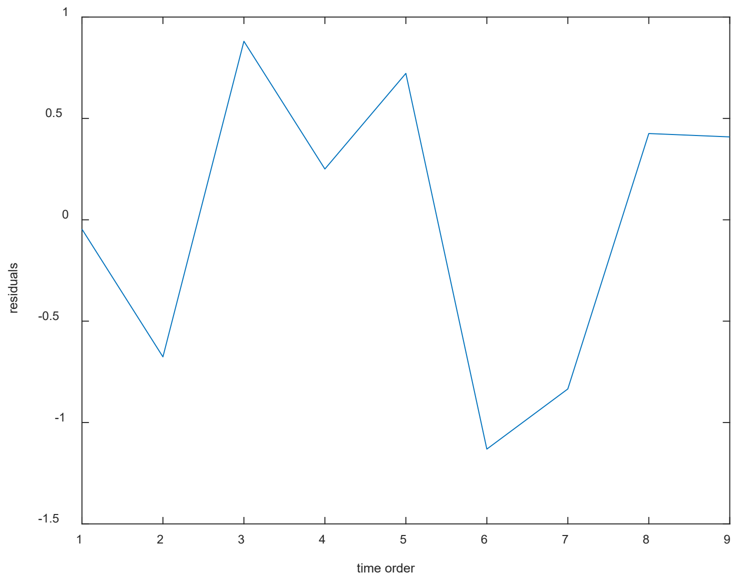
- The optimal point of operability to guarantee a product with small particles is the one with the lowest loads and the classifier in an intermediate position between the first and the second

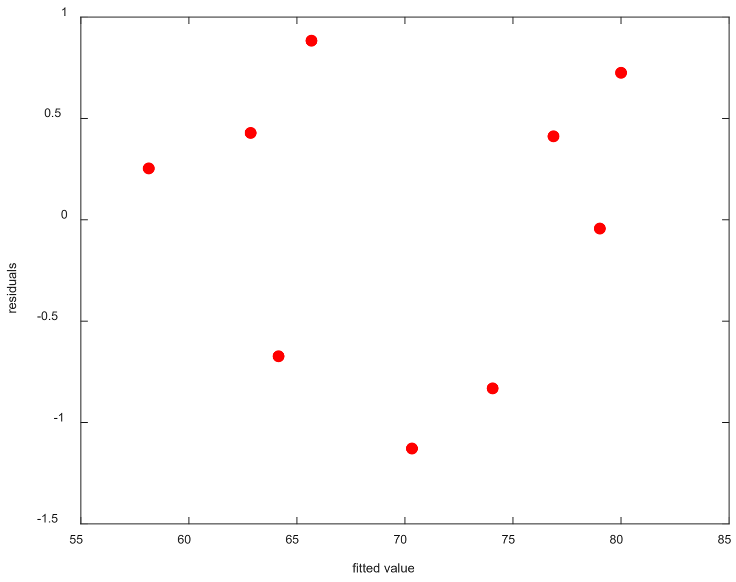


Analysis of the residuals

- The model residuals do not show any pattern







Preliminary ANOVA model on large particles

- The main effect on large particles seems to be related to the load of the feeding material

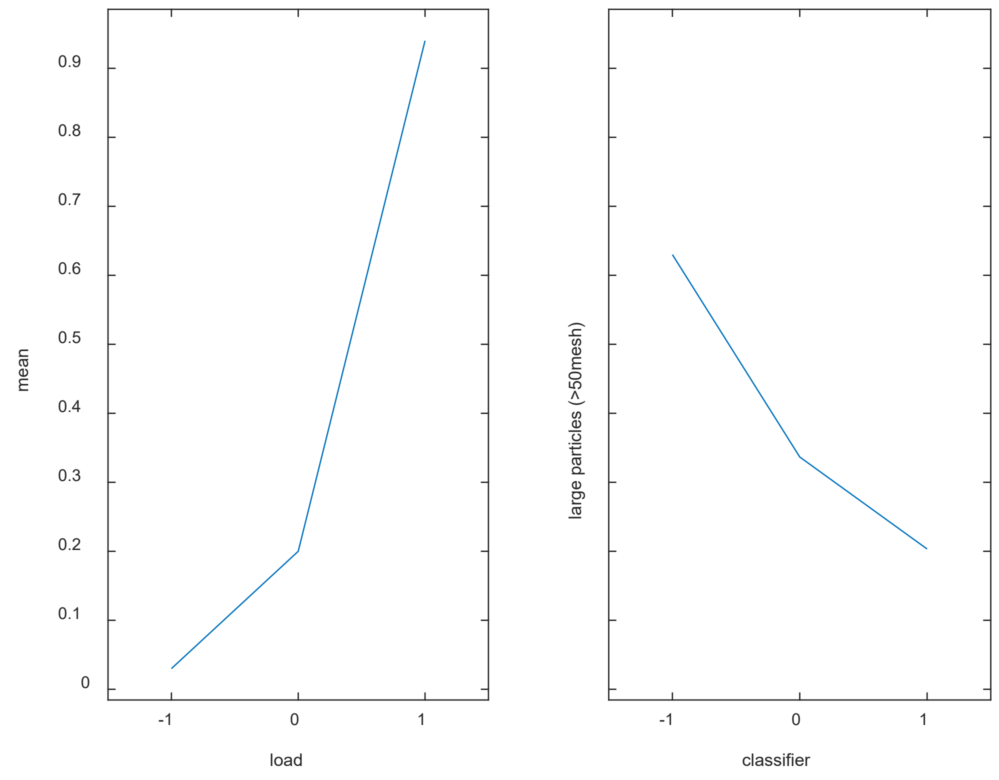
Analysis of Variance

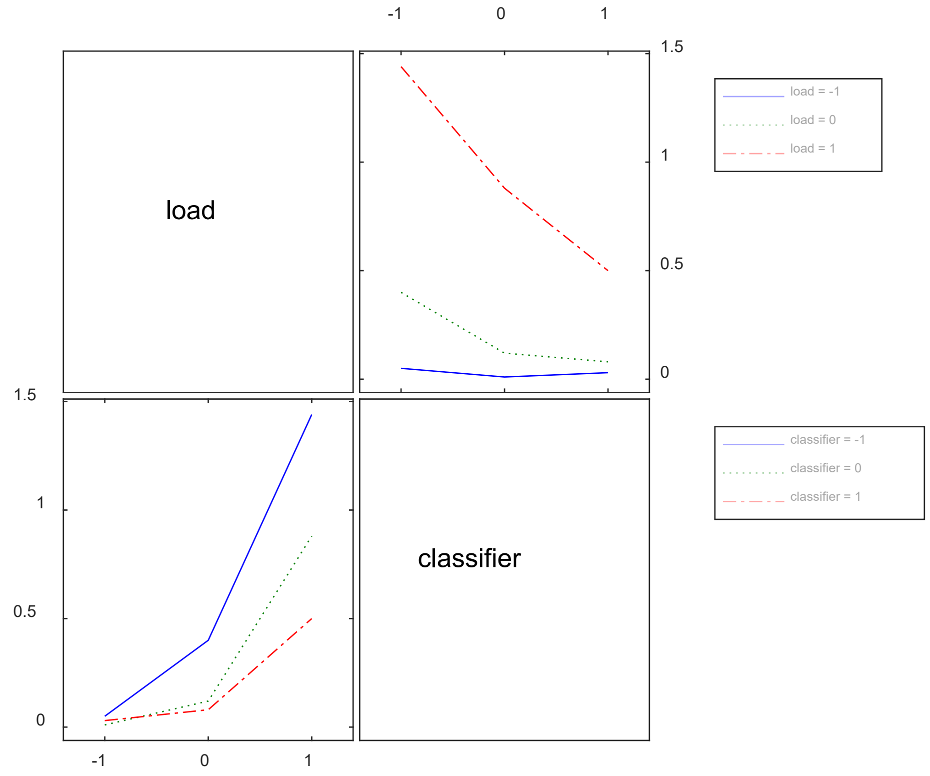
Source	Sum Sq.	d.f.	Mean Sq.	F	Prob>F
X1	1.4046	2	0.7023	Inf	NaN
X2	0.28587	2	0.14293	Inf	NaN
X1*X2	0.22293	4	0.05573	Inf	NaN
Error	-0	0	-0		
Total	1.9134	8			

Constrained (Type III) sums of squares.

Main effect and interactions for large particles

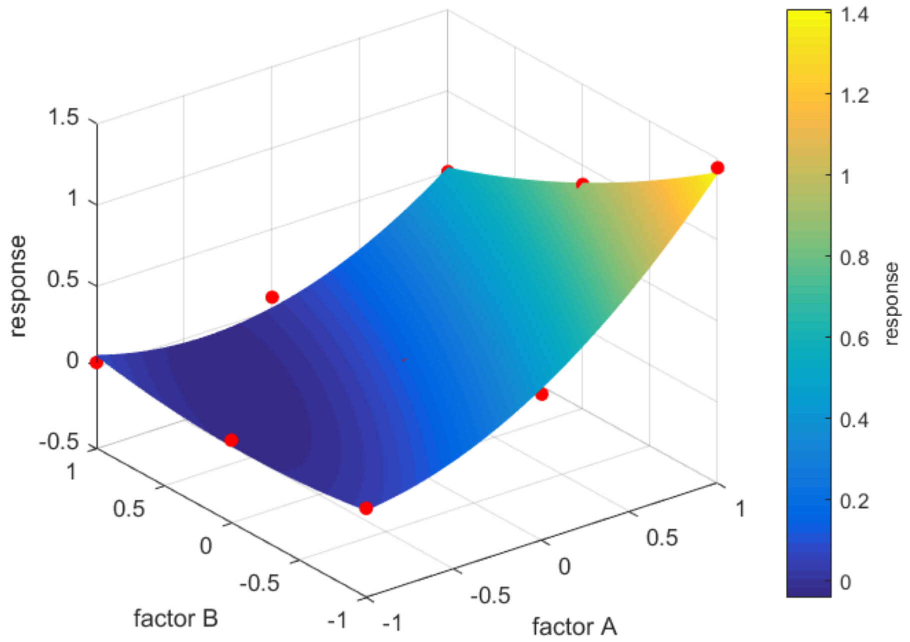
- The results of the main plots confirm the ANOVA model outcomes
- Higher loads seem to determine higher content of large particles in the product

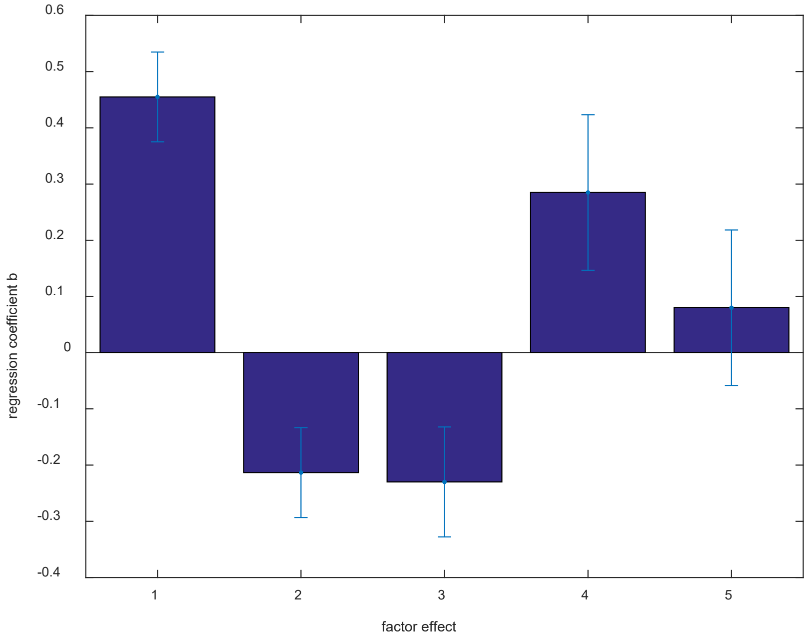




Regression model for large particles

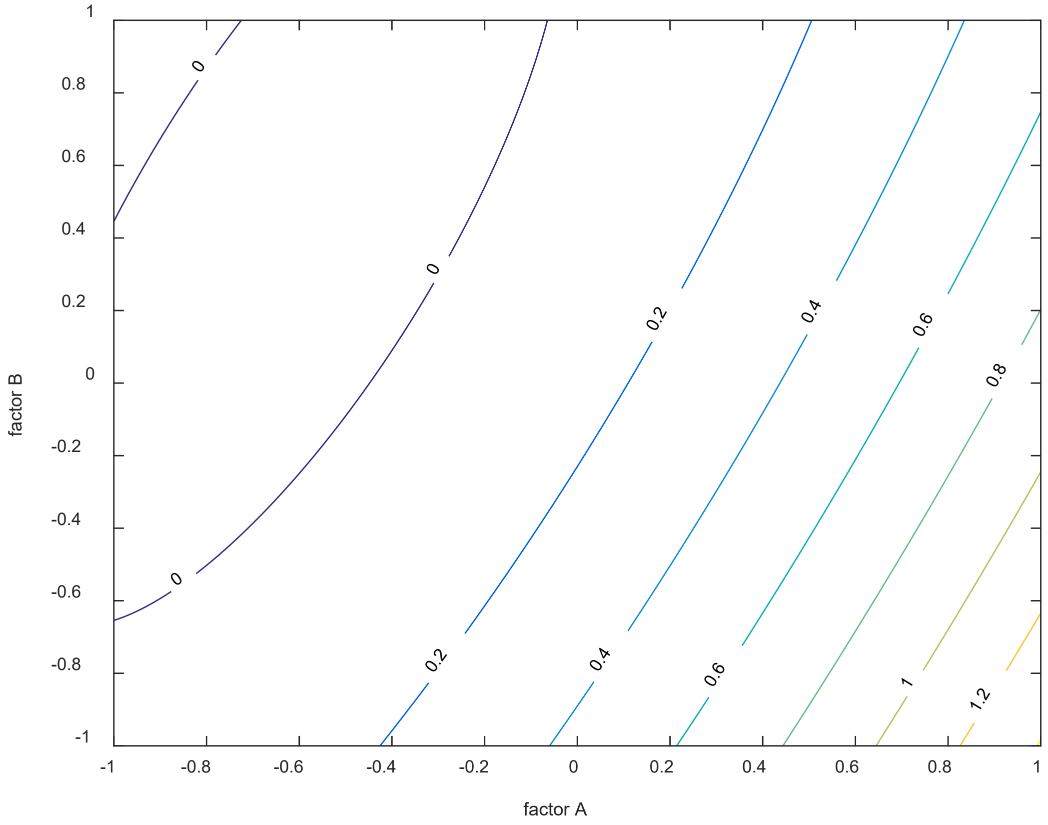
- The selected regression model structure for optimization is: $y_2 = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2$
- The response surface shows that:
  - the fitting of the model is very good
  - the regression coefficients confirm that:
    - the load is the most important effects on large particles
      - the higher the load, the higher the content of large particles in the product
    - the interaction and the classifier position seem to be not so important on large particles
    - a significant effect of the squared value of the load is present





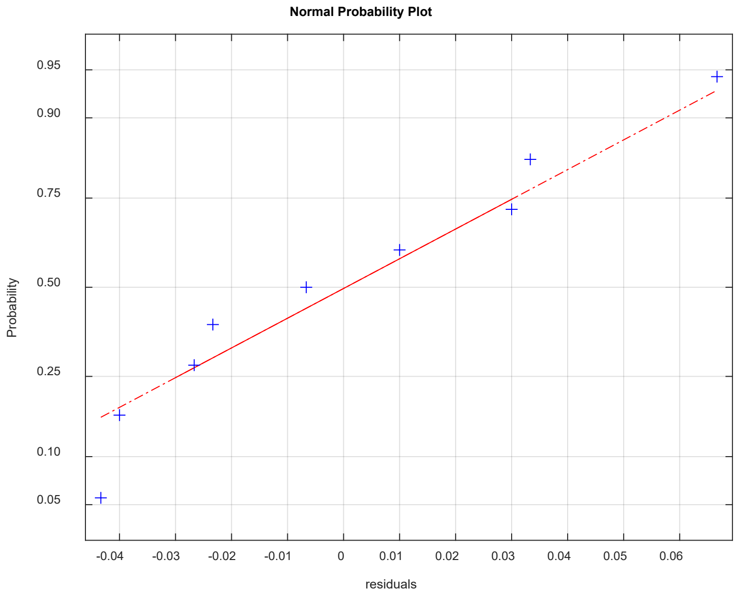
Optimization

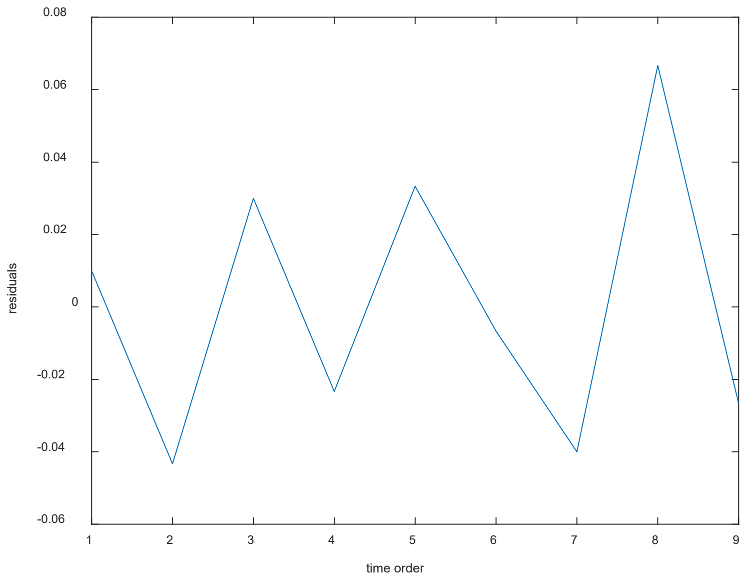
- If lower content of large particles is desired it is suggested to keep low the load

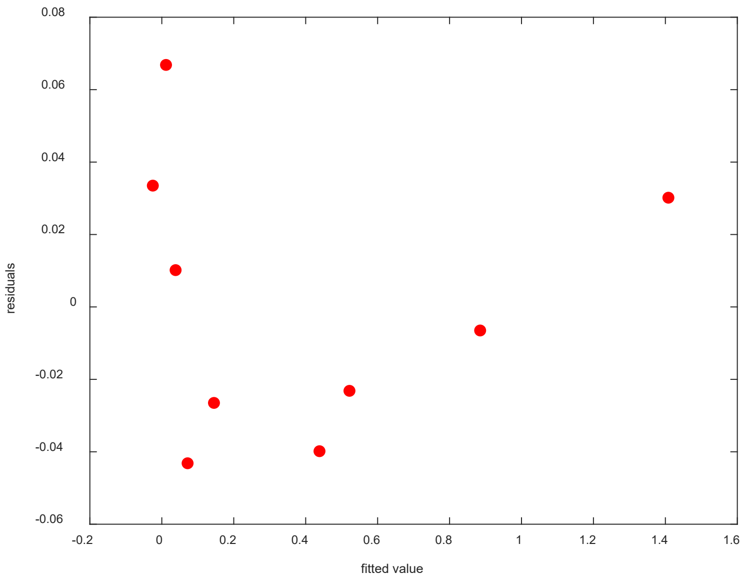


Analysis of the residuals

- The model residuals do not show any pattern



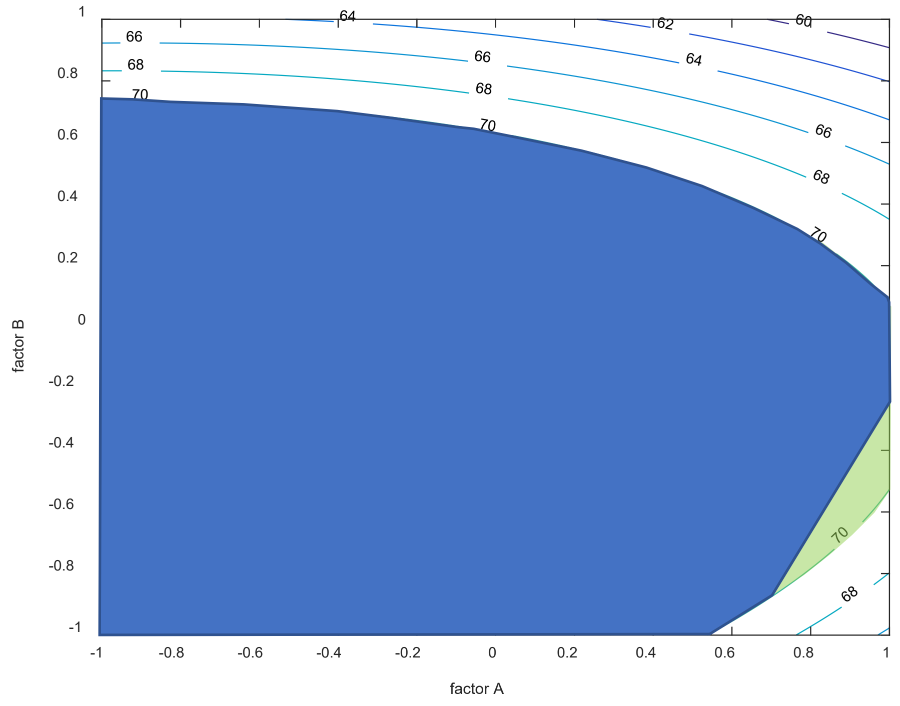


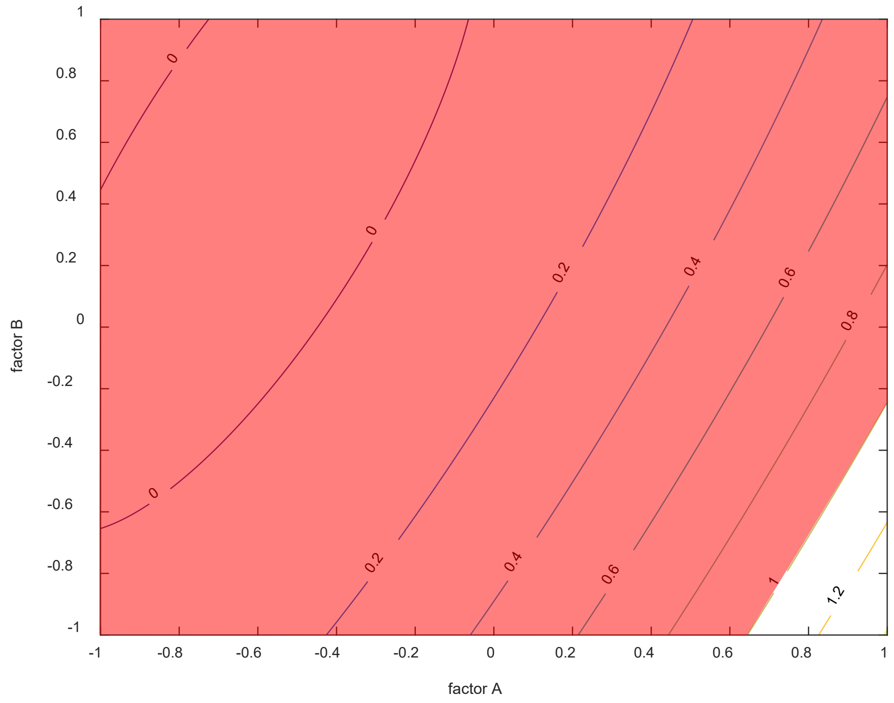


Optimal operating region

- If the small particles should be >70% and the large particles should be <1%

the optimal experimental domain is the intersection of the two domains in the abovementioned response surfaces





Homework #2

- You will find **Homework #2** in Moodle:
  - this is a typical exercise on DoE that you can find in the final exam
  - deadline: 1 week before the final exam
    - however, the sooner you complete it, the sooner you can have a feedback on your preparation for the final test

…and thank you very much for having attended the course!
I am pretty sure that it will be useful in your professional life! ☺

Good luck for your exams!!!