Università degli Studi di Padova

Department of Industrial Engineering

Design of Experiments
Lesson #9

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/

Fractional Factorial Design

Fractional factorial designs

- In a 2^K factorial design, the number of runs required to complete a full factorial experimentation rapidly outgrows the resources of most experimenters as the number of factors K increases
  - a complete run of a 2^6 design requires 64 runs
  - only 6 of the 63 degrees of freedom correspond to main effects
  - only 15 degrees of freedom correspond to two-factor interactions
    - as a total of the previous ones only 21 degrees of freedom associated with effects that are likely to be of major interest
  - the remaining 42 degrees of freedom are associated with three-factor and higher interactions
- **Fractional factorial designs**: if the experimenter can reasonably assume that *certain high-order interactions are negligible,* information on the main effects and low-order interactions may be obtained by running a **limited *fraction of the complete factorial*** experiment
  - fractional factorial designs are among the most widely used types of designs
  - applications:
    - product and process design
    - process improvement
    - industrial/business experimentation
    - etc…

Screening experimentation

- A major use of fractional factorials is in **screening experiments**
  - *many factors* are considered
  - objective is to identify those factors (if any) that have large effects
  - usually performed in the *early stages of a project*
  - the factors identified as important are then investigated more thoroughly in subsequent experiments
- The successful use of fractional factorial designs is based on three key ideas:
  1. **sparsity of effects principle**
     - when there are several variables, the system or process is likely to be **driven primarily by some of the main effects and low-order interactions**
  2. **projection property**
     - fractional factorial designs can be **projected into stronger larger designs in the subset of significant factors**
  3. **sequential experimentation**
     - it is possible **to combine the runs of two (or more) fractional factorials to construct sequentially a larger design** to estimate the factor effects and interactions of interest

Basic principles of the fractional factorial

- Consider a situation in which three factors $K = 3$ are selected, each at two levels
- Suppose that the experimenters cannot afford to run all the $2^3 = 8$ treatment combinations, but can afford 4 runs
  - this suggests a **one-half fraction** of a 2^3 design. Because the design contains $2^{3-1} = 4$ treatment combinations, a one-half fraction of the 2^3 design is often called a $\mathbf{2^{3-1}}$ **design**
- The four treatment are selected as a combinations *a, b, c* and *abc* as our one-half fraction
  - the 2^{3-1} design is formed by selecting only those treatment combinations that have a plus in the *ABC* column
  - *ABC* is called the **generator** of this particular fraction, and ABC **words**
  - the identity column *I = ABC* is always plus, so we call the **defining relation** for our design
    - the defining relation for a fractional factorial will always be the set of all columns that are equal to the identity column *I*
- **The treatment combinations in this design yield three degrees of freedom that we may use to estimate the main effects**

Treatment Combination	Factorial Effect							
	I	*A*	*B*	*C*	*AB*	*AC*	*BC*	*ABC*
a	+	+	−	−	−	−	+	+
b	+	−	+	−	−	+	−	+
c	+	−	−	+	+	−	−	+
abc	+	+	+	+	+	+	+	+
ab	+	+	+	−	+	−	−	−
ac	+	+	−	+	−	+	−	−
bc	+	−	+	+	−	−	+	−
(1)	+	−	−	−	+	+	+	−

Main effects and alias

- In the effects it is impossible to differentiate among:
  - *A* and *BC*
  - *B* and *AC*
  - *C* and *AB*
- These are called **aliases**:
  - when estimating A, we are really estimating A+BC
  - when estimating B, we are really estimating B+AC
  - when estimating C, we are really estimating C+AB
- The **alternate** (and **complementary**) one-half fraction can be chosen through the *other* one-half fraction I = -ABC
  - the treatment combinations associated with minus in the *ABC* column
  - in principle, it does not matter which fraction is actually used, because the fractions belong to the same **family**

$$[A] = \frac{1}{2}(a - b - c + abc)$$
$$[B] = \frac{1}{2}(-a + b - c + abc)$$
$$[C] = \frac{1}{2}(-a - b + c + abc)$$

$$[BC] = \frac{1}{2}(a - b - c + abc)$$
$$[AC] = \frac{1}{2}(-a + b - c + abc)$$
$$[AB] = \frac{1}{2}(-a - b + c + abc)$$

Fractional factorial design resolution

- In general, a design is of **resolution** R if no p-factor effect is aliased with another effect containing less than $R - p$ factors
- Designs of resolution III, IV, and V are particularly important. The definitions of these designs and an example of each follow:
  - **resolution III designs**
    - designs in which no main effects are aliased with any other main effect
    - main effects are aliased with two-factor interactions
    - some two-factor interactions may be aliased with each other
    - the 2^{3-1} design is of resolution III
  - **resolution IV designs**
    - no main effect is aliased with any other main effect or with any two-factor interaction
    - two-factor interactions are aliased with each other
    - a 2^{4-1} design with $I = ABCD$ is a resolution IV design
  - **resolution V designs**
    - no main effect or two-factor interaction is aliased with any other main effect or two-factor interaction
    - two factor interactions are aliased with three-factor interactions
    - a 2^{5-1} design with $I = ABCDE$ is a resolution V design
- It is desirable to employ fractional designs that have the **highest possible resolution** consistent with the degree of fractionation required
  - *the higher the resolution, the less restrictive the assumptions that are required regarding which interactions are negligible to obtain a unique interpretation of the results*

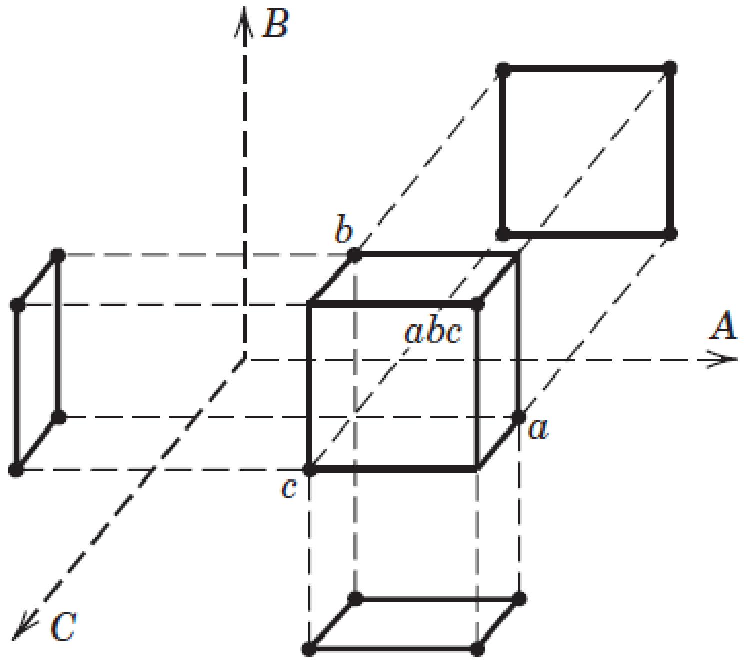
Building a one-half fractional factorial design

- A one-half fraction of the 2^K design of the highest resolution may be constructed by:
  - writing down a basic design consisting of the runs for a *full* 2^{K-1} factorial
  - adding the K-th factor by identifying its plus and minus levels with the plus and minus signs of the highest order interaction
  - for example, a one-half 2^3 fractional factorial is obtained by:
    - writing down the full 2^2 factorial as the basic design
    - then equating factor *C* to the *AB* interaction
    - the alternate fraction would be obtained by equating factor *C* to the *AB* interaction
- In Matlab® use: `fracfact` and `fracfactgen`

Full 2^2 Factorial (Basic Design)		2_{III}^{3-1}, $I = ABC$		
A	B	A	B	$C = AB$
−	−	−	−	+
+	−	+	−	−
−	+	−	+	−
+	+	+	+	+

Projections of fractions into factorials

- *Any fractional factorial design of resolution R contains complete factorial designs (possibly replicated factorials) in any subset of $(R-1)$ factors*:
  - if an experimenter has several factors of potential interest, but believes that only $(R-1)$ of them have important effects, then a fractional factorial design of resolution R is the appropriate choice of design
    - if the experimenter is correct, the fractional factorial design of resolution R will project into a full factorial in the $(R-1)$ significant factors
  - for example, a 2^3 design projects into every subset of two factors
- Because the maximum possible resolution of a one-half fraction of the 2^K design is $(R-K)$, every 2^{K-1} design will project into a full factorial in any $(K-1)$ of the original K factors
  - furthermore, a 2^{K-1} design may be projected into two replicates of a full factorial in any subset of $K-2$ factors, four replicates of a full factorial in any subset of $K-3$ factors, and so on…



Confirmation experiments

- **Confirmation Experiments**:
  - adding the alternate fraction to the principal fraction may be thought of as a type of **confirmation experiment**
    - it provides information that allow strengthening the initial conclusions about the two-factor interaction effects
  - a very simple confirmation experiment is to use the model equation to **predict the response at a point of interest** in the design space (this should not be one of the runs in the current design) and then actually run that treatment combination (perhaps several times), comparing the predicted and observed responses:
    - reasonably close agreement indicates that the interpretation of the fractional factorial was correct
    - serious discrepancies mean that the interpretation was problematic
    - if the predicted and observed values in a confirmation experiment are not this close, it will be necessary to answer the question of whether the two values are sufficiently close
      - one way to answer this question is to construct a **prediction interval** on the future observation for the confirmation run and then see if the actual observation falls inside the prediction interval

Let’s recall the example of the filtration rate

- See slide 26 in the previous lesson (DoE Lesson #8)

Example: filtration rate

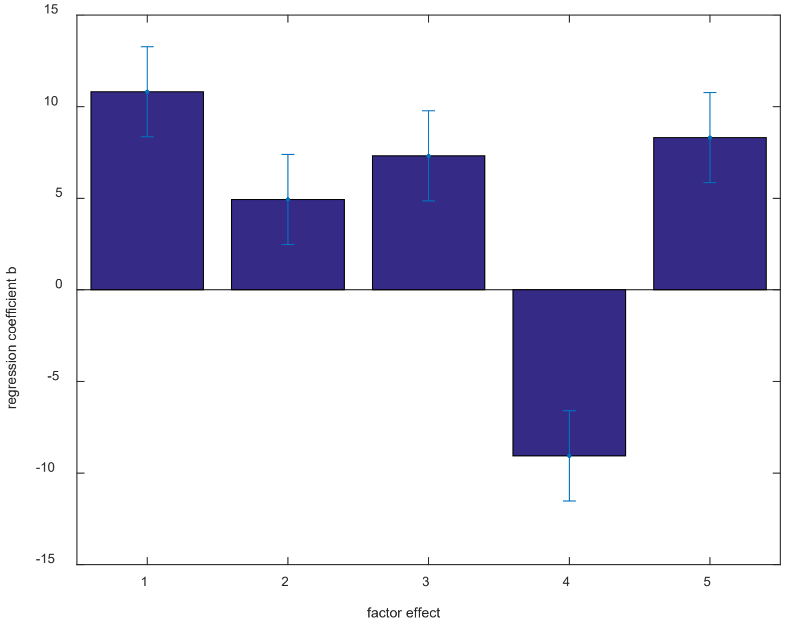
- A chemical product is produced in a **pressure vessel**
- <u>Objective</u>: experiments are carried out in the pilot plant to study how the factors influence the filtration rate of the product
- Experimentation:
  - 4 factors:
    - temperature T (factor A)
    - pressure P (factor B)
    - concentration of formaldehyde $C_{formaldehyde}$ (factor C)
    - stirring rate SR (factor D)
  - 2^4 factorial factors are
    - each factor is present at two levels
    - single replicate experiments

17/05/2024 Prof. Pierantonio Facco - Machine learning for process engineering 26

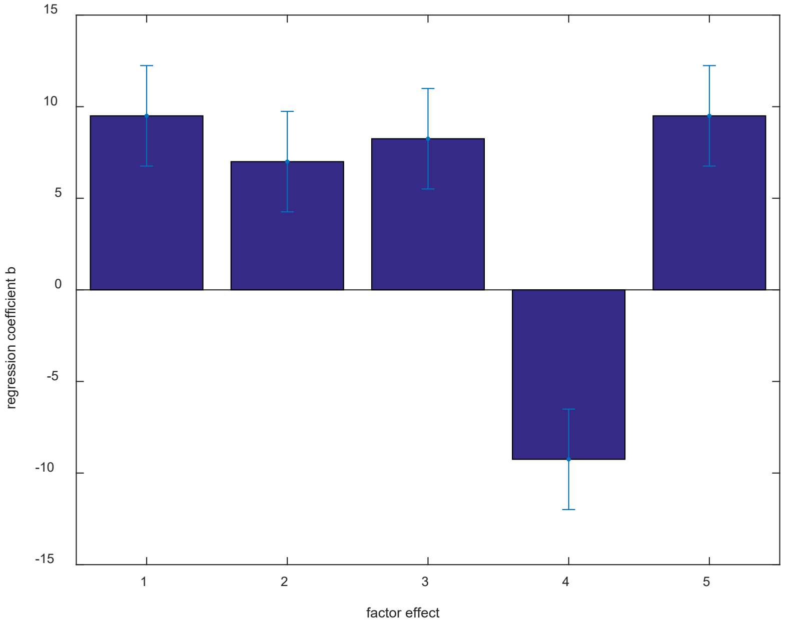
Example of Filtration rate: comparison (1/3)

- Limited differences are found in the regression model parameter estimation among the full factorial design and the one half fractional factorial design
  - see: `filtration_rate.m`

full factorial



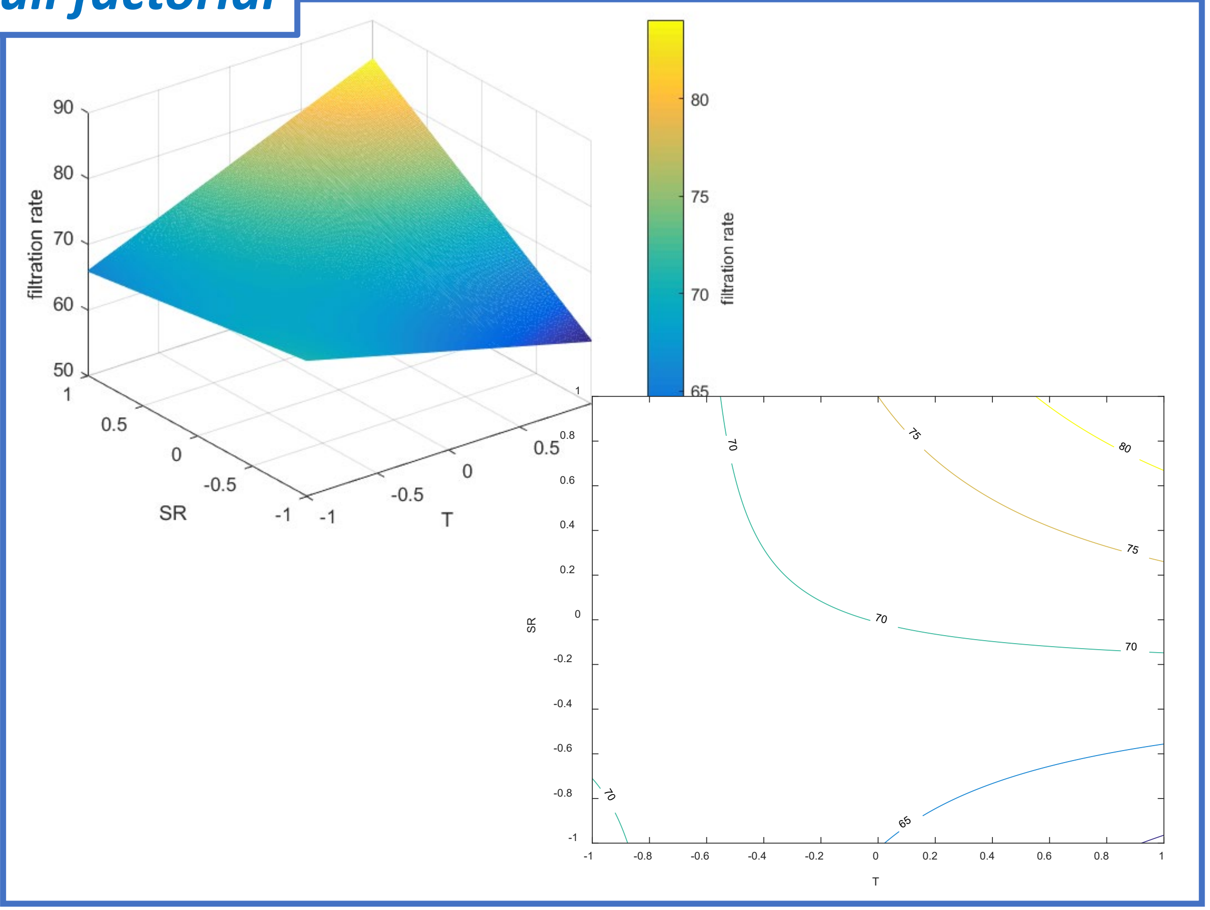
fractional factorial



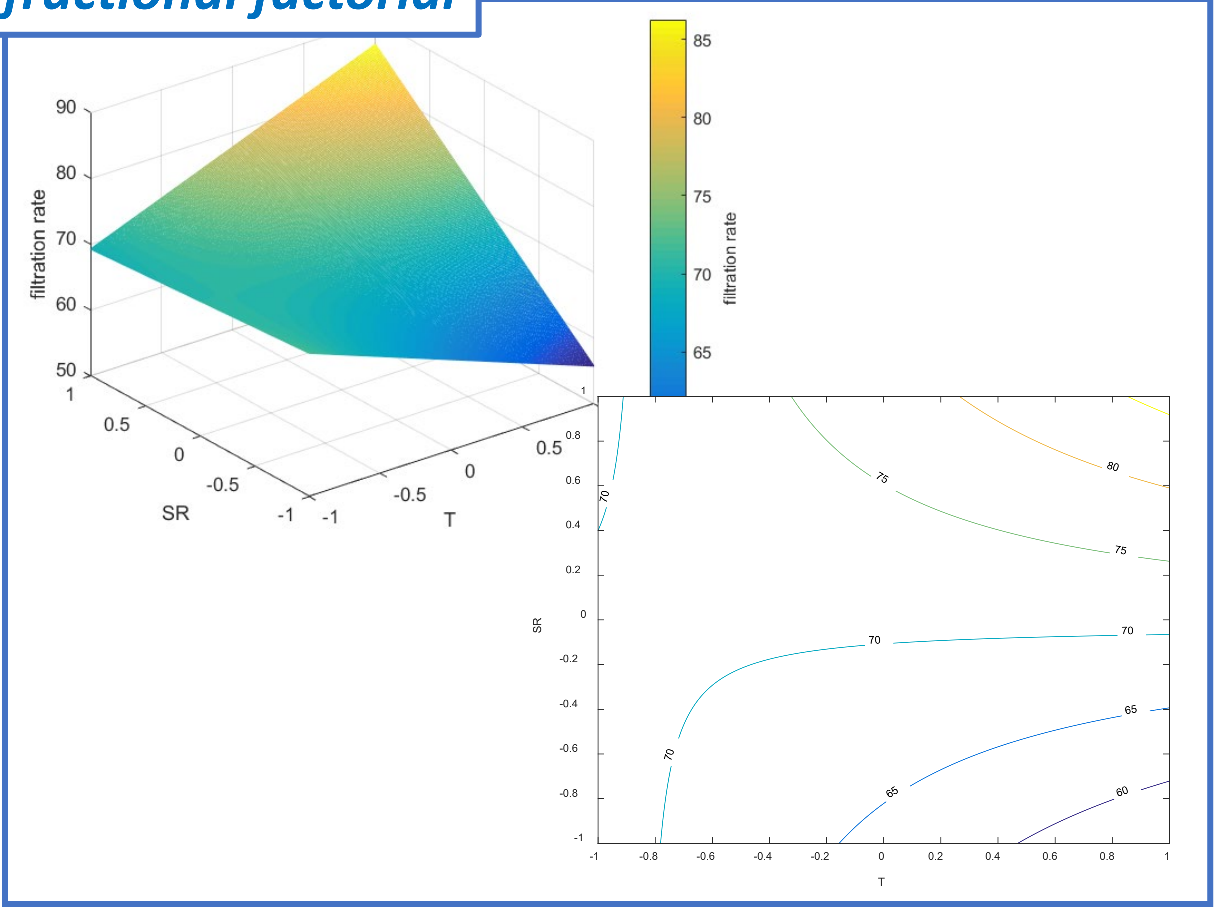
Example of Filtration rate: comparison (2/3)

- Only slight differences can be seen in the response surface at constant temperature level

full factorial



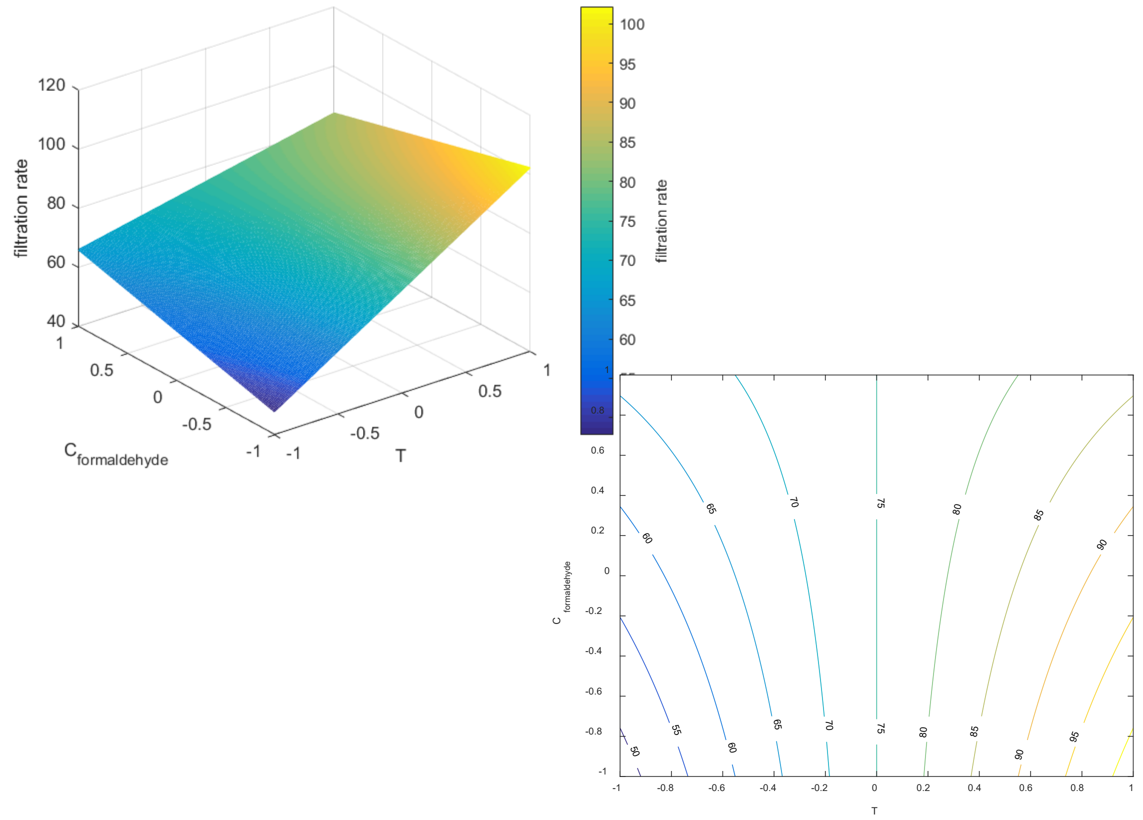
fractional factorial



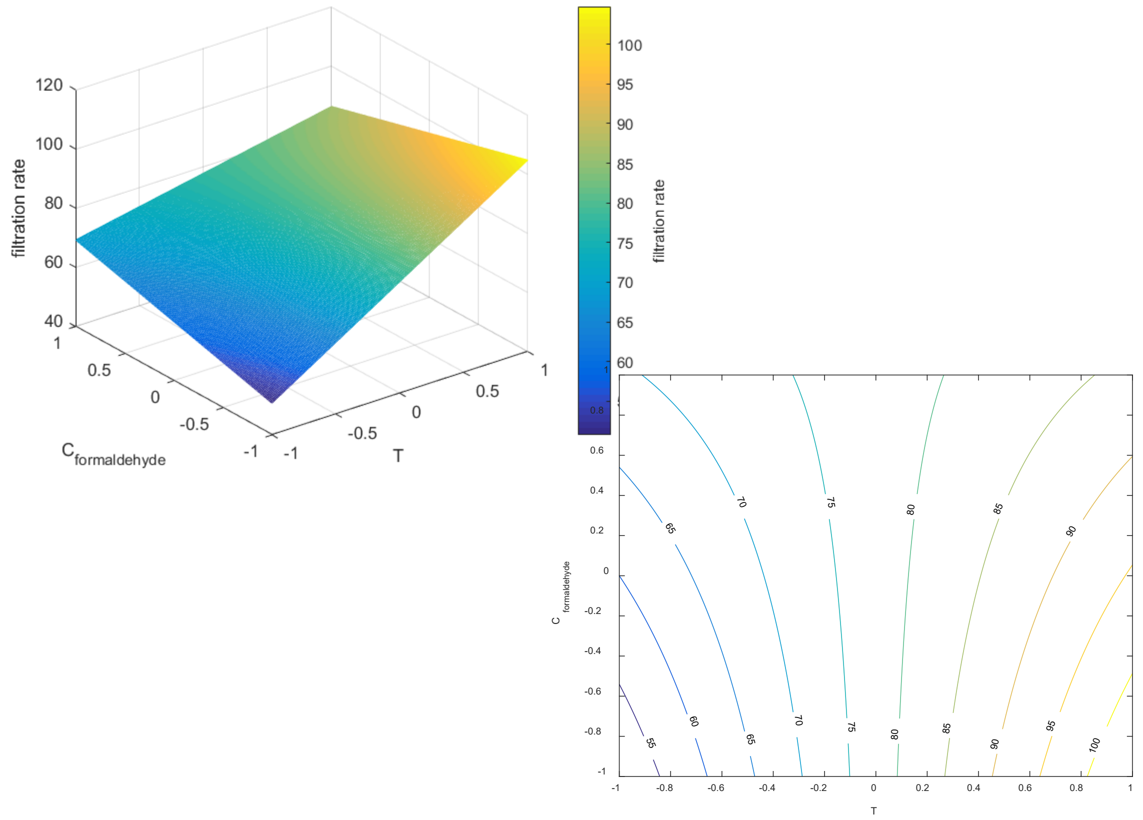
Example of Filtration rate: comparison (3/3)

- Only slight differences can be seen in the response surface at constant SR level

full factorial

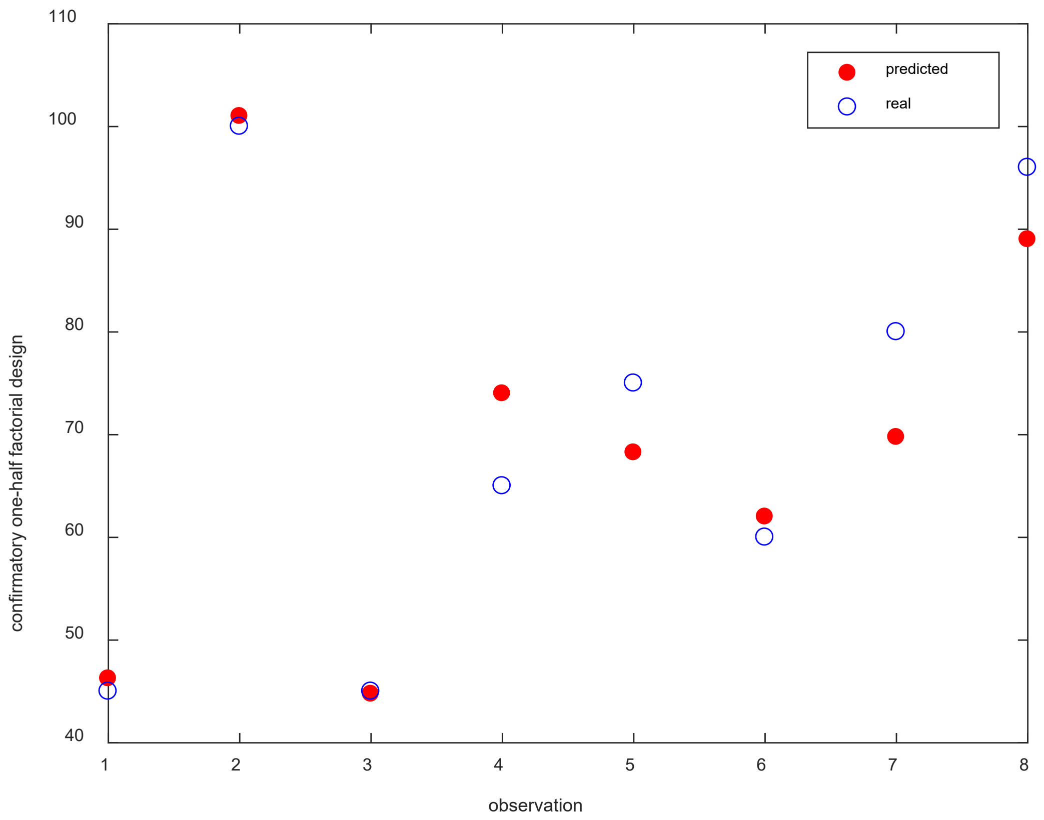


fractional factorial



Confirmatory experimentation

- Prediction of the samples of the one-half fractional factorial design with the model built on the complementary design:
  - predictions are absolutely satisfactory (namely, accurate)
  - compare the prediction errors with the expected variability of the predictions (namely, the prediction uncertainty)!



Verification of the confirmatory experiment

- Pay attention that prediction errors are found
- Are these errors acceptable?
- Let's consider the prediction uncertainty

$$\hat{y}(\mathbf{x}_n) - t_{\alpha,N-V}\sqrt{\hat{\sigma}^2\mathbf{x}_n^{\mathrm{T}}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{x}_n} \leq \boldsymbol{\mu}_y(\mathbf{x}_n) \leq \hat{y}(\mathbf{x}_n) + t_{\alpha,N-V}\sqrt{\hat{\sigma}^2\mathbf{x}_n^{\mathrm{T}}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{x}_n}$$

- Verify if the error $y - \hat{y}$ is lower than the uncertainty:

$$y \in \left[\hat{y}(\mathbf{x}_n) - t_{\alpha,N-V}\sqrt{\hat{\sigma}^2\mathbf{x}_n^{\mathrm{T}}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{x}_n}, \hat{y}(\mathbf{x}_n) + t_{\alpha,N-V}\sqrt{\hat{\sigma}^2\mathbf{x}_n^{\mathrm{T}}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{x}_n}\right]$$

2^{K-P} fractional factorial designs

- A 2^K fractional factorial design containing 2^{K-P} runs is called a $1/2^P$ fraction of the 2^K design or, more simply, a 2^{K-P} **fractional factorial design**
  - these designs require the selection of P **independent generators**
  - the defining relation for the design consists of:
    - the P generators initially chosen
    - their $(2^P - P - 1)$ generalized interactions
- The alias structure may be found by multiplying each effect column by the defining relation
- Care should be used in choosing the generators so that effects of potential interest are not aliased with each other
  - each effect has $(2P - 1)$ aliases
  - for moderately large values of K, we usually assume higher order interactions (say, third- or fourth-order and higher) to be negligible ➡ the alias structure is greatly simplified
- It is important to select the P generators for a 2^{K-P} fractional factorial design in such a way as to obtain the **best possible alias relationships**
  - a reasonable criterion is to select the generators such that the resulting 2^{K-P} design has the **highest possible resolution**

Example of maximum resolution design

- Consider the following design

	Basic Design					
Run	A	B	C	D	$E = ABC$	$F = BCD$
1	−	−	−	−	−	−
2	+	−	−	−	+	−
3	−	+	−	−	+	+
4	+	+	−	−	−	+
5	−	−	+	−	+	+
6	+	−	+	−	−	+
7	−	+	+	−	−	−
8	+	+	+	−	+	−
9	−	−	−	+	−	+
10	+	−	−	+	+	+
11	−	+	−	+	+	−
12	+	+	−	+	−	−
13	−	−	+	+	+	−
14	+	−	+	+	−	−
15	−	+	+	+	−	+
16	+	+	+	+	+	+

- The generators $E = ABC$ and $F = BCD$ are used thereby producing a design of resolution IV
- This is the maximum resolution design
  - if we had selected $E = ABC$ and $F = ABCD$, the complete defining relation would have been $I = ABCE = ABCDF = DEF$, and the design would be of resolution III
    - this is an inferior choice because it needlessly sacrifices information about interactions

Minimum aberration designs

- Sometimes resolution alone is not sufficient to distinguish between designs
- Consider the three designs in the following Table

Design A Generators: $F = ABC,\ G = BCD$ $I = ABCF = BCDG = ADFG$	**Design B Generators:** $F = ABC,\ G = ADE$ $I = ABCF = ADEG = BCDEFG$	**Design C Generators:** $F = ABCD,\ G = ABDE$ $I = ABCDF = ABDEG = CEFG$
Aliases (two-factor interactions)	Aliases (two-factor interactions)	Aliases (two-factor interactions)
$AB = CF$	$AB = CF$	$CE = FG$
$AC = BF$	$AC = BF$	$CF = EG$
$AD = FG$	$AD = EG$	$CG = EF$
$AG = DF$	$AE = DG$	
$BD = CG$	$AF = BC$	
$BG = CD$	$AG = DE$	
$AF = BC = DG$		

- All of these designs are of resolution IV, but they have rather different alias structures
  - assumed that three-factor and higher interactions are negligible
  - design A has more extensive aliasing
  - design C has the least aliasing, so it would be the best choice
  - C minimizes the number of words in the defining relation that are of minimum length: **minimum aberration design**
- Minimizing aberration in a design of resolution R ensures that the design has the minimum number of main effects aliased with interactions of order $R-1$, the minimum number of two-factor interactions aliased with interactions of order $R-2$, and so forth...

Blocking and confounding

- Occasionally, a fractional factorial design requires so many runs that all of them cannot be made under homogeneous conditions
- In many problems it is also impossible to perform a complete replicate of a factorial design in one block:
  - the price we have to pay when reducing the number of experiments is that our factor effects can no longer be computed completely free of one another
  - the effects are said to be *confounded*, that is, to a certain degree mixed up with each other
- Fractional factorials may be separated in **blocks**:
  - tables are available which contain recommended blocking arrangements for many fractional factorial designs
  - the minimum block size for these designs is eight runs
- **Confounding** is a design technique for arranging a complete factorial experiment in blocks, where the block size is smaller than the number of treatment combinations in one replicate
  - the technique causes information about certain treatment effects (usually high-order interactions) to be **indistinguishable from**, or **confounded with**, **blocks**
  - even though the designs presented are **incomplete block designs** because each block does not contain all the treatments or treatment combinations, the special structure of the factorial system allows a simplified method of analysis

Example: jet engine blades

Example: jet turbine engine

- A five-axis CNC (Computer Numerical Control) machine is used to produce an impeller for a **jet turbine engine**
  - the blade profiles are an important quality characteristic
  - deviations of the blade profile from the one specified on the engineering drawing is of interest
- <u>Objective</u>: to determine which machine parameters affect profile deviation
- Experimentation:
  - 8 factors are selected
  - response variable:
    - standard deviation of the difference between the actual profile and the specified profile
    - the machine has four spindles and the process engineers feel that the spindles should be treated as blocks

Factor	Low Level (−)	High Level (+)
A = x-Axis shift (0.001 in.)	0	15
B = y-Axis shift (0.001 in.)	0	15
C = z-Axis shift (0.001 in.)	0	15
D = Tool supplier	1	2
E = a-Axis shift (0.001 deg)	0	30
F = Spindle speed (%)	90	110
G = Fixture height (0.001 in.)	0	15
H = Feed rate (%)	90	110

Choice of the fractional factorial

- Three-factor and higher interactions do not seem to be important, but *ignoring the two-factor interactions is not fully appropriate*
- Two designs initially appear appropriate: the design with 16 runs and the design with 32 runs
  - if the *16-run design* is used, there will be fairly extensive aliasing of two-factor interactions
    - this design cannot be run in four blocks without confounding four two-factor interactions with blocks
  - it is convenient to use the design *in four blocks*
    - this confounds one three-factor interaction alias chain and one two-factor interaction (*EH*) and its three-factor interaction aliases with blocks
    - the *EH* interaction is the one between the *a*-axis shift and the feed rate, which is considered to be fairly unlikely

Designs with 8 Factors

(l) 2^{8-4}; 1/16 fraction of 8 factors in 16 runs **Resolution IV**

Design Generators

$E = BCD \quad F = ACD \quad G = ABC \quad H = ABD$

Defining relation: $I = BCDE = ACDF = ABEF = ABCG = ADEG = BDFG = CEFG = ABDH = ACEH = BCFH = DEFH = CDGH = BEGH = AFGH = ABCDEFGH$

Aliases

$A = CDF = BEF = BCG = DEG = BDH = CEH = FGH$	$AB = EF = CG = DH$
$B = CDE = AEF = ACG = DFG = ADH = CFH = EGH$	$AC = DF = BG = EH$
$C = BDE = ADF = ABG = EFG = AEH = BFH = DGH$	$AD = CF = EG = BH$
$D = BCE = ACF = AEG = BFG = ABH = EFH = CGH$	$AE = BF = DG = CH$
$E = BCD = ABF = ADG = CFG = ACH = DFH = BGH$	$AF = CD = BE = GH$
$F = ACD = ABE = BDG = CEG = BCH = DEH = AGH$	$AG = BC = DE = FH$
$G = ABC = ADE = BDF = CEF = CDH = BEH = AFH$	$AH = BD = CE = FG$
$H = ABD = ACE = BCF = DEF = CDG = BEG = AFG$	

2 blocks of 8: $AB = EF = CG = DH$

(m) 2^{8-3}; 1/8 fraction of 8 factors in 32 runs **Resolution IV**

Design Generators

$F = ABC \quad G = ABD \quad H = BCDE$

Defining relation: $I = ABCF = ABDG = CDFG = BCDEH = ADEFH = ACEGH = BEFGH$

Aliases

$A = BCF = BDG$	$AE = DFH = CGH$	$DE = BCH = AFH$
$B = ACF = ADG$	$AF = BC = DEH$	$DH = BCE = AEF$
$C = ABF = DFG$	$AG = BD = CEH$	$EF = ADH = BGH$
$D = ABG = CFG$	$AH = DEF = CEG$	$EG = ACH = BFH$
$E =$	$BE = CDH = FGH$	$EH = BCD = ADF = ACG = BFG$
$F = ABC = CDG$	$BH = CDE = EFG$	$FH = ADE = BEG$
$G = ABD = CDF$	$CD = FG = BEH$	$GH = ACE = BEF$
$H =$	$CE = BDH = AGH$	$ABE = CEF = DEG$
$AB = CF = DG$	$CG = DF = AEH$	$ABH = CFH = DGH$
$AC = BF = EGH$	$CH = BDE = AEG$	$ACD = BDF = BCG = AFG$
$AD = BG = EFH$		

2 blocks of 16: $ABE = CEF = DEG$

4 blocks of 8: $ABE = CEF = DEG$, $ABH = CFH = DGH$, $EH = BCD = ADF = ACG = BFG$

Generation of the fractional factorial design

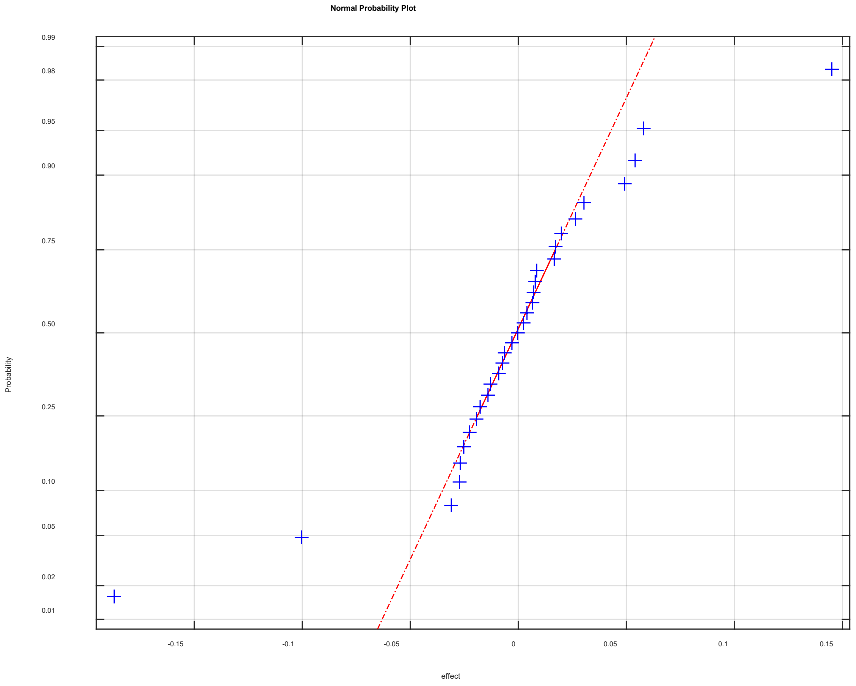
- In Matlab®:
  - `generators=('a b c d e abc abd bcde');`
  - `[dfF,confounding]=fracfact(generators,'MaxInt',3,'FactorNames',{'A','B','C','D','E','F','G','H'});`
- The outcome is:

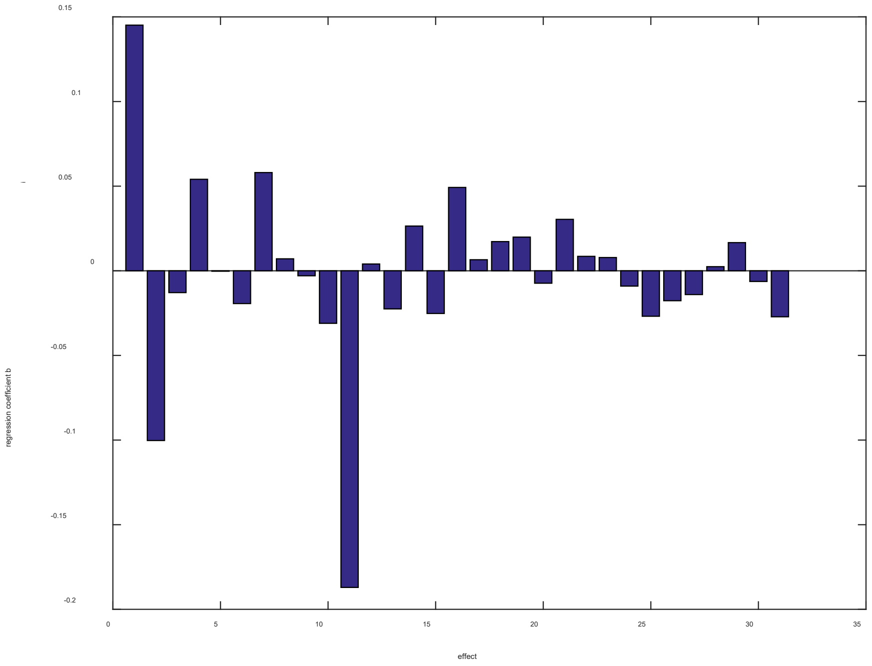
A	B	C	D	E	F=ABC	G=ABD	H=BCDE
-1	-1	-1	-1	-1	-1	-1	1
-1	-1	-1	-1	1	-1	-1	-1
-1	-1	-1	1	-1	-1	1	-1
-1	-1	-1	1	1	-1	1	1
-1	-1	1	-1	-1	1	-1	-1
-1	-1	1	-1	1	1	-1	1
-1	-1	1	1	-1	1	1	1
-1	-1	1	1	1	1	1	-1
-1	1	-1	-1	-1	1	1	-1
-1	1	-1	-1	1	1	1	1
-1	1	-1	1	-1	1	-1	1
-1	1	-1	1	1	1	-1	-1
-1	1	1	-1	-1	-1	1	1
-1	1	1	-1	1	-1	1	-1
-1	1	1	1	-1	-1	-1	-1
-1	1	1	1	1	-1	-1	1
1	-1	-1	-1	-1	1	1	1
1	-1	-1	-1	1	1	1	-1
1	-1	-1	1	-1	1	-1	-1
1	-1	-1	1	1	1	-1	1
1	-1	1	-1	-1	-1	1	-1
1	-1	1	-1	1	-1	1	1
1	-1	1	1	-1	-1	-1	1
1	-1	1	1	1	-1	-1	-1
1	1	-1	-1	-1	-1	-1	-1
1	1	-1	-1	1	-1	-1	1
1	1	-1	1	-1	-1	1	1
1	1	-1	1	1	-1	1	-1
1	1	1	-1	-1	1	-1	1
1	1	1	-1	1	1	-1	-1
1	1	1	1	-1	1	1	-1
1	1	1	1	1	1	1	1

The experimental campaign is carried out based on this strategy

ANOVA and regression models

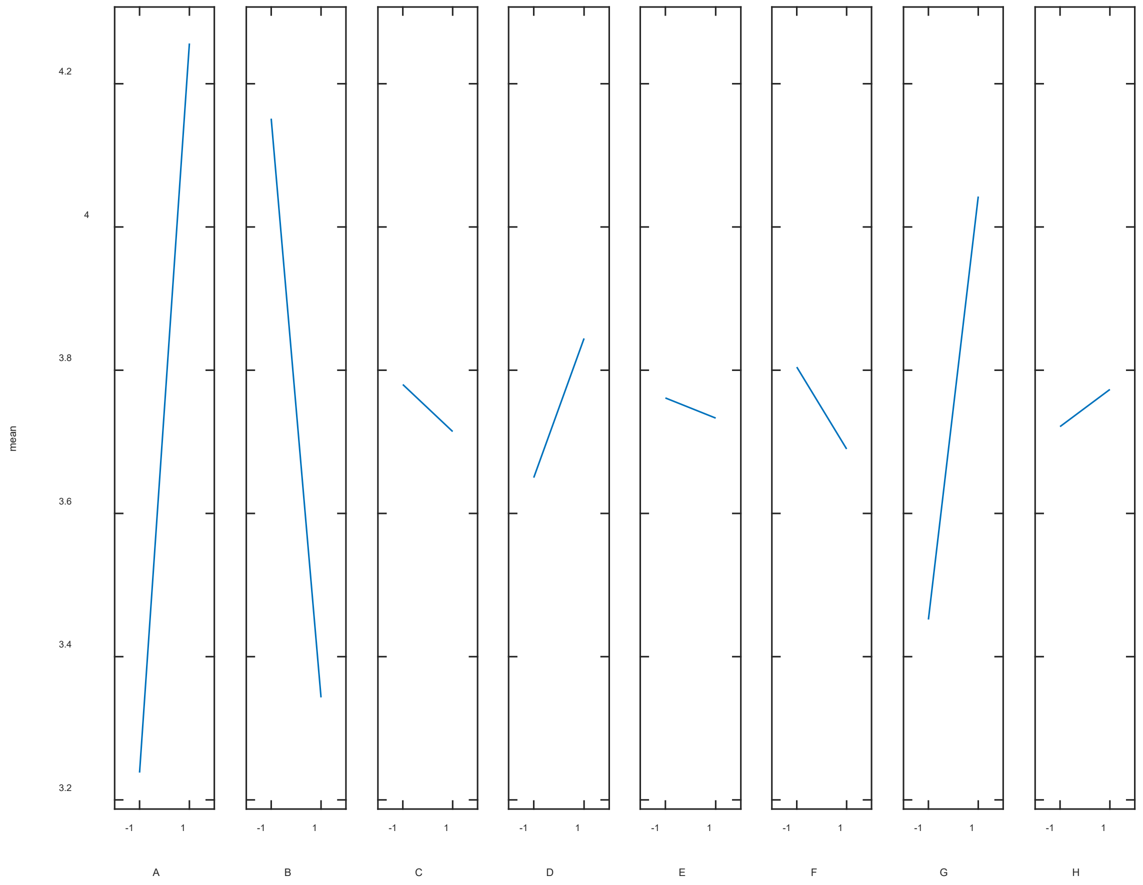
- Only one term per alias is included into the model
- The main effects are:
  - A, x-axis shift
  - B, y-axis shift
  - AD (interaction x-axis shift and tool supplier) or the alias BG (interaction y-axis shift and fixture height)
    - not the alias CF because C and F are not in the main effects





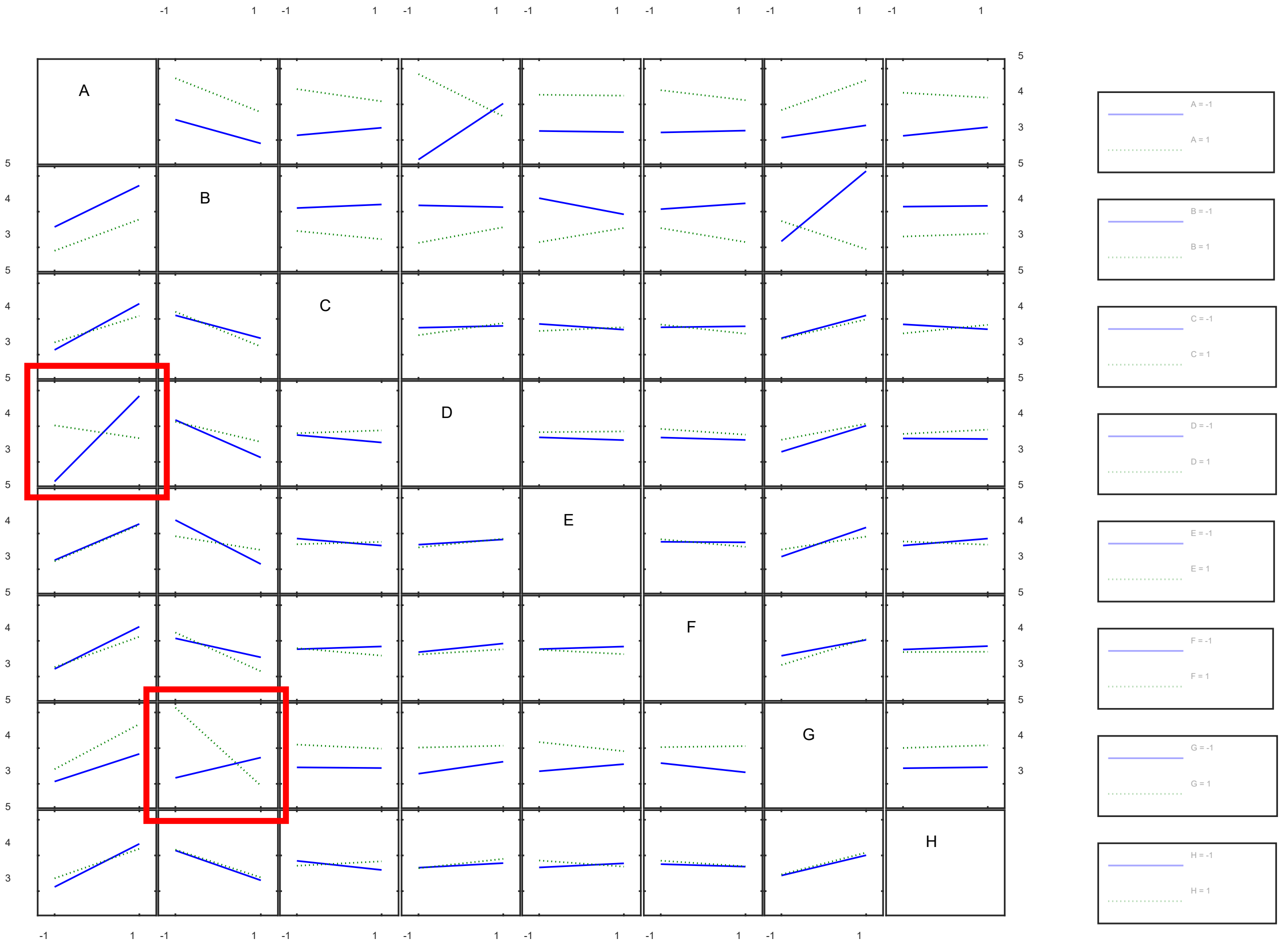
Main effects

- Main effects of A and B are confirmed in the main effect plot



Aliasing and interactions

- The aliasing can be seen in the interaction plot, as well



… per sempre a fianco a me!