DOE
DESIGN OF EXPERIMENT
EXPERIMENTAL DESIGN

Lesson 10

Universitat Pompeu Fabra
Barcelona

Framework of lesson 10:

- Response Surface Methodology
- Research of the ascending direction (Steepest ascent method)
- Local explorative phase
- Examples
- Some complete studies

RESPONSE SURFACE

- Until now we have often spoken generically about "factors". It has been seen that these can be "category" or "continuous type“
- Among the first we can mention the battery brand, the instrument on which it is used, the medicine tested on patients, ... We have seen how the ANOVA results can be deepened with the concepts of orthogonality.
- Among the factors "continuous type" you can have the selling price, the space dedicated on the shelf, the welding temperature, ... In the case of continuous factors the ANOVA results can be completed with a linearity analysis, evaluating if there are significant non-linearity components.
- For continuous factors, usually called "variables", there is a further type of analysis, that of the study of the surface of response.

RESPONSE SURFACE

- It is usual to divide the experimentation into three phases:

1. Screening phase: this is a preliminary analysis (usually performed with complete or fractional factorial plans, Taguchi method), which considers all the factors evaluated at a few levels (often 2), oriented to understand which factors are significant and which ones can be discarded.
2. Project phase of the actual experiment: it is a subsequent phase, in which the effects of the previously revealed factors are investigated, with analysis of the variance on the few factors at many levels (orthogonal decompositions, linearity evaluations).

RESPONSE SURFACE

3. Study phase of the response surface: for continuous variables it is aimed at determining the mathematical function that links the response to the independent variables, secondly it seeks the search for the optimal combination of variables, which maximizes the response itself.

1^{st} Phase (Screening)	Complete factorial plans (2k) or fractional plans (2k-p), Taguchi method
2^{nd} Phase (Effects)	ANOVA with 1 or 2 factors or more Latin and Greek-Latin quadrature factors, variance decomposition, linearity verification
3^{rd} Phase	Search for the response surface and its function: optimization of treatments to maximize the result

RESPONSE SURFACE: EXAMPLE

- Consider the following question. We want to optimize the thickness of a casting by acting on the values of two factors, namely continuous variables (called X_1 and X_2). The factors are the mixture ratio of the two components of the epoxy resin and the position of the pour point in the mold.
- The independent variable is represented by the thickness to be optimized.

Variabili indipendenti	**Variabile dipendente**
Variable X_1: resin	Answer Y: Thickness of the piece made by casting
Variable X_2: position of the mold in the cast	

RESPONSE SURFACE: INTERPOLATION

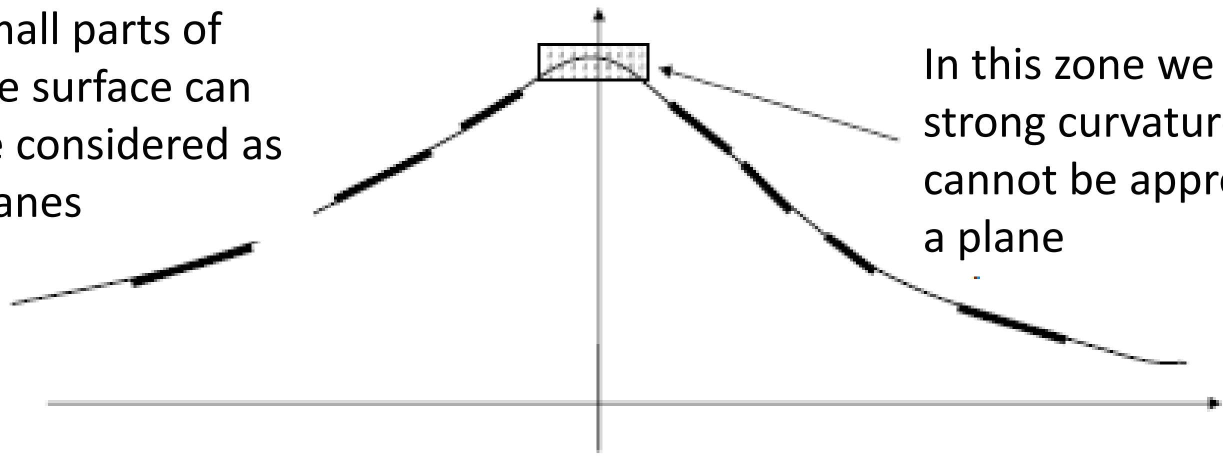
- The aim is to identify the relationship that links X_1 and X_2 (independent variables) to the dependent variable Y. This relationship will later allow us to identify the combination of X1 and X2 that maximizes the response. In the event that it is minimized (for example, reduction of product defects, pollutant emissions, ...), just consider the maximization of (-Y).
- The relation is of the type:
- $Y=f(X_1,X_2) + \varepsilon$
- $f(X_1,X_2)$ represents an interpolating polynomial relationship, while ε is an error term (deviation between real value and value obtained by interpolation).

RESPONSE SURFACE: INTERPOLATION

- The interpolating function can be of the first or second order.
- Approximation according to a plan (first order):
- $f(X_1, X_2) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$
- Approximation according to a bilinear form of the second order:
- $f(X1, X2) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{11} X_1^2 + \beta_{22} X_2^2 + \beta_{12} X_1 X_2$
- The term $\beta_{12} X_1 X_2$ is closely linked to the interaction phenomena between the two variables.
- Initially, in the attempt to map the dependent variable with respect to the inputs, the linear approximation can be used, later, near the maximum, the quadratic terms must also be considered.

RESPONSE SURFACE: INTERPOLATION

Small parts of the surface can be considered as planes



In this zone we have a strong curvature which cannot be approximated to a plane

- The planar approximation is usually acceptable in an area of the response function sufficiently far from the point of maximum and of limited extension.
- The actual adequacy of the planar model can however be verified with appropriate tests.
- By mapping the surface, you can locate the maximum and therefore the optimal combination.

PHASE ONE: EXPLORATION OF THE FIELD

- The question of optimizing the answer is usually answered, mapping the answer itself according to the inputs. Initially we rely on a linear model and do a series of experiments on two levels with two purposes: to understand how far you are from the maximum and identify the direction to move to the maximum.
- This method is called "method of the steepest ascent“.
- Assuming a planar type model is equivalent to considering only linear terms, in the absence of quadratic terms and terms related to interactions. With only linear terms at stake, only two levels can be considered for each factor, using complete or fractional factorial plans.

PHASE TWO: LOCAL EXPLORATION

- Only when, at the end of the experiments of wide-ranging exploration, we get close to the maximum and the planar model becomes inadequate, we must use a quadratic model. A particular ANOVA test is used to ascertain the inadequacy of the planar approximation.
- This second phase usually involves only one experiment, but on several levels (therefore with more combinations of treatments).
- The method adopted in this phase is called "local exploration".
- Given a problem, for example the one aimed at optimizing the thickness of a casting, we proceed according to the two phases.

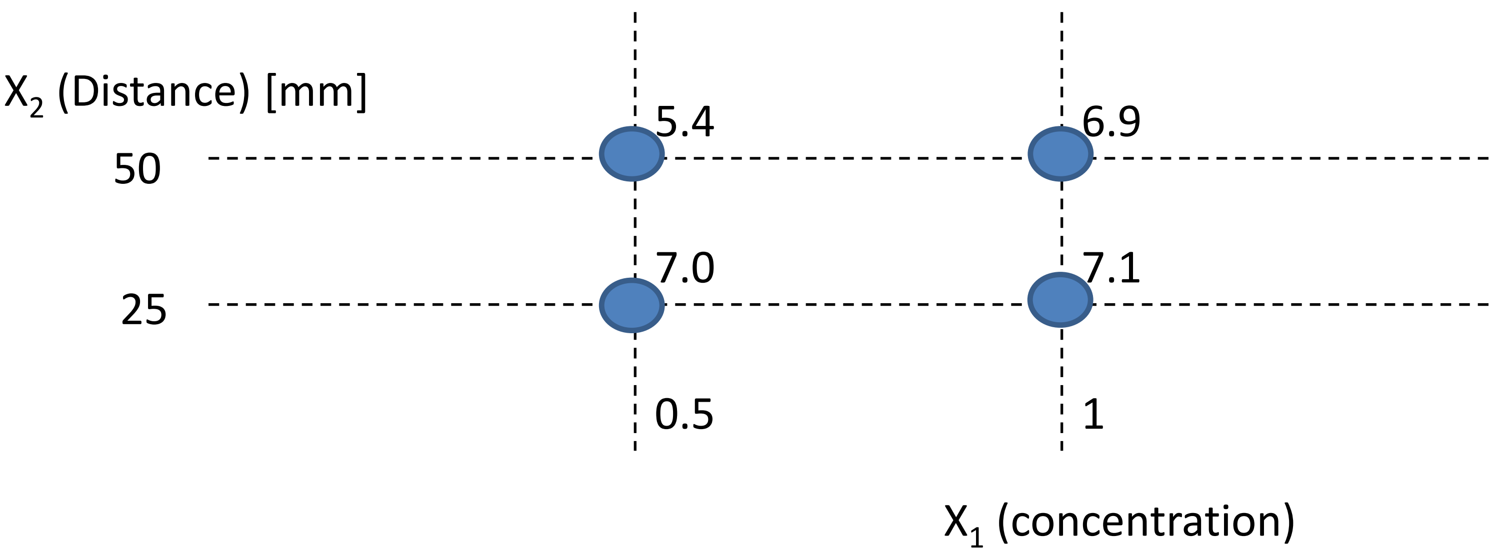
STEEPEST ASCENT METHOD

- You start by identifying two levels for each variable. These may be two typical values for the variables under examination, which, depending on the state of the art or previous knowledge, may not be too far from the optimal values.
- In correspondence with these values a 2^k plane is set.
- This is the first of the series of experiments in the first phase of analysis.

	Low Level	**High Level**
Variable X_1: Concentration of resin	1:2 (0.5)	1:1 (1)
Variable X_2: pouring position in the mold	25 mm from the side of the mold	50 mm from the side of the mold

STEEPEST ASCENT METHOD

- Returning to the plan 2^2:
- Here, too, an optimization is carried out according to the two phases. It starts from the first experiment of the first phase performed under the conditions mentioned (in the absence of repetitions).

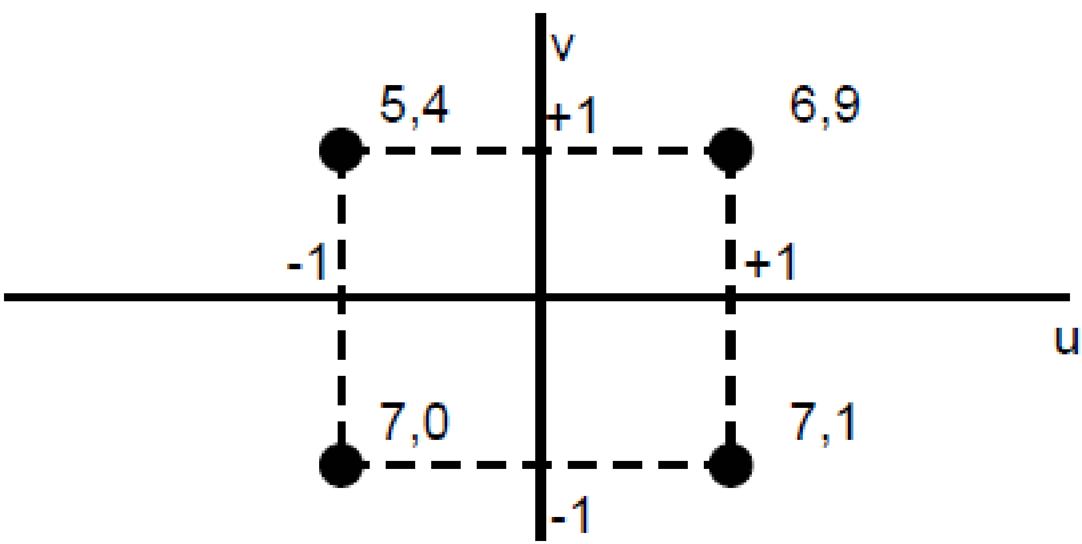


STEEPEST ASCENT METHOD

Usually we refer to quantities (here indicated with u and v) defined on [-1; 1]:

Therefore:

- $X_1 = [0,5; 1] \rightarrow u = [-1; 1]$;
- $X_2 = [25; 50] \rightarrow v = [-1; 1]$
- $u = 4 \cdot X_1 - 3$; $v = 2/25 \cdot X_2 - 3$



STEEPEST ASCENT METHOD

Usually we refer to quantities (here indicated with u and v) defined on [-1; 1]:
Therefore:

- X_1 = [0,5; 1] → u = [-1; 1];
- X_2 = [25; 50] → v = [-1; 1]
- $u = 4 \cdot X_1 - 3$; $v = 2/25 \cdot X_2 - 3$ (note that $X_1 \equiv A$, $X_2 \equiv B$)
- The answer function is of the type:
- $Y = \delta_0 + \delta_1 \cdot u + \delta_2 \cdot v + \varepsilon$, this represents the relationship between the inputs and the dependent variable, unless an error ε.

Combinations	Results	(1)	(2)	Esteems of…	
(-1,-1) →1	7.0	14.1	26.4	26.4/4=6.6	μ (mean)
(1,-1) →a	7.1	12.3	1.6	1.6/2=0.8	A
(-1,1) →b	5.4	0.1	-1.8	-1.8/2=-0.9	B
(-1,-1) →ab	6.9	1.5	1.4	1.4/2=0.7	AB

STEEPEST ASCENT METHOD

- δ_0 = Average
- $\delta_1 = A/2$
- $\delta_2 = B/2$
- $Y = \delta_0 + \delta_1 \cdot u + \delta_2 \cdot v + \varepsilon = \mu + A/2 \cdot u + B/2 \cdot v + \varepsilon$
- So the estimate of the response plan is:
- $f(u, v) = \mu + A/2 \cdot u + B/2 \cdot v = 6{,}60 + 0{,}4 \cdot u - 0{,}45 \cdot v$
- Both u and v vary from -1 to +1, that is, two "drawing units" (from -1 to 0 and from 0 to +1).
- The divisions for two take into account the fact that the effects represent the increase in the response variable due to changes in two drawing units.

STEEPEST ASCENT METHOD

- $f(u, v) = \mu + A/2 \cdot u + B/2 \cdot v = 6{,}60 + 0{,}4 \cdot u - 0{,}45 \cdot v$
- The goal is to maximize the response, ie the thickness of casting. This is greater, the higher the value of u (higher concentration ratio) and the lower the value of v (the pour point closest to the edge of the mold).
- One can now ask whether the terms μ, A / 2 and B / 2 are acceptable estimates of δ0, δ1, δ2 respectively. To make sure of this, an ANOVA test is performed, to evaluate the significance or not of the A and B effects, referable to u and v. If they were significant, then this would indicate that the result is significantly increased with at least one factor changing, indicating that we might be close to the point of maximum.

STEEPEST ASCENT METHOD

- To evaluate the significance, we consider the usual parallelism between the considered 2^2 plane and a monofactorial ANOVA with 2^2, i.e. 4 levels.

1	a	b	ab	Grand Mean
7.0	7.1	5.4	6.9	6.6

- TSS = (7,0 - 6,6)2 + (7,1- 6,6)2 + (5,4 - 6,6)2 + (6,9- 6,6)2 = 0,16 + 0,25 + 1,44 + 0,09 = 1,94
- The problem is that in the absence of replicas it is not possible to evaluate the error term, so SSW = 0 with zero degrees of freedom. It also has that SSB_C = TSS.

STEEPEST ASCENT METHOD

- Assuming negligible interactions (as in the planar model), we use the term interaction AB to estimate the error: we say that this term is not null only due to the effect of uncertainty.
- The individual decompositions are of the type:
- $(\text{effect})^2 \cdot \text{n° ripetitions} \cdot 2^{k-2}$
- $SSQ_U = A^2 \cdot 1 \cdot 2^{2-2} = A^2 = (0,8)^2 = 0,64$
- $SSQ_V = B^2 \cdot 1 \cdot 2^{2-2} = B^2 = (-0,9)^2 = 0,81$
- $SSQ_{Errore} = SSW = AB^2 \cdot 1 \cdot 2^{2-2} = A^2 = (0,7)^2 = 0,49$
- Effettivamente $TSS = SSB_C = SSQ_U + SSQ_V + SSW =$
- $= 0,64 + 0,81 + 0,49 = 1,94$

STEEPEST ASCENT METHOD

	SSQ	DoF	MSQ	Fcalc.	p-value
SSB_C	1.94	3			
$A(u,\delta_1)$	0.64	1	0.64	1.3	46%
$B(v,\delta_1)$	0.81	1	0.81	1.7	42%
Errore (SSW)	0.49	1	0.49		
Total	1.94	3			

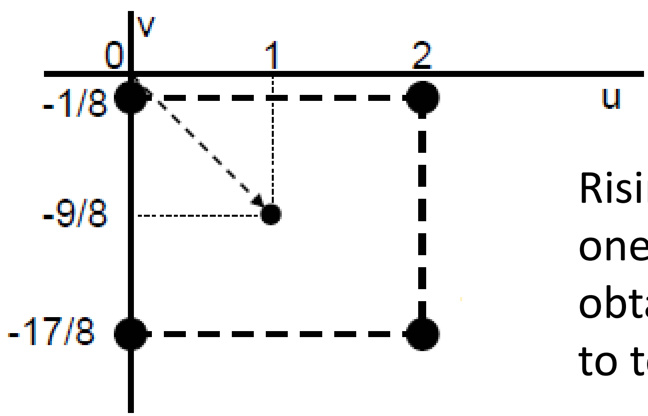
It is therefore obtained that there is no significance, neither for A (u, δ1), nor for B (v, δ2). This indicates that the response surface is flat and suggests that it is still far from the maximum point.

STEEPEST ASCENT METHOD

- The next experiment is then formulated, moving in the direction of the "steepest climb".
- This is the direction given by the vector of the type [δ1; δ2], in practice. In this case: [0.40; -0.45] = [1; -9/8]. The searched direction has the equation: $v = -9 / 8 \cdot u$. To center the new experiment below, it will be necessary to move in this direction.
- Therefore: the first experiment is centered in [0; 0], to center the second one, it is necessary to move along the line $v = -9/8 \cdot u$.
- The problem is related to how much to move from the center: if you move too little, there is the risk of having to use many steps, to get to the point of maximum, if too much, there is a risk of overcoming the maximum point.

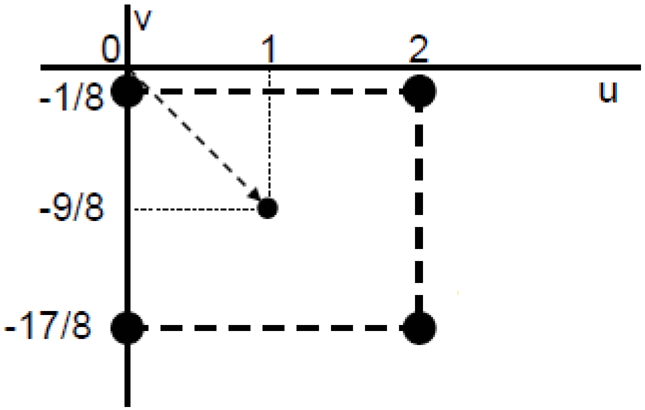
STEEPEST ASCENT METHOD

- There is no fixed rule: one goes by trial, also relying on experience. It would be important to know what is the maximum value to be expected.
- Assume that you have defined the following new center: u = +1, v = -9/8



Rising and descending of one unit in the plot we obtain the combinations to test in the new 2^2

STEEPEST ASCENT METHOD

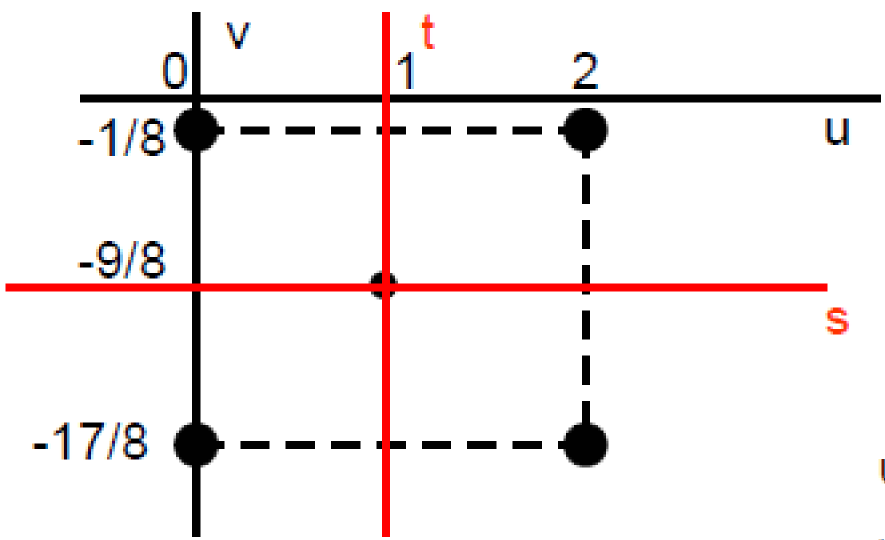


u	V	X_1	X_2
0	-1/8	0.75	35.9
0	-17/8	0.75	10.9
2	-1/8	1.25	35.9
2	-17/8	1.25	10.9

$$\begin{cases} u = 4X_1 - 3 \\ v = \dfrac{2}{25}X_2 - 3 \end{cases} \Leftrightarrow \begin{cases} X_1 = \dfrac{u+3}{4} \\ X_2 = \dfrac{25}{2}\cdot(v+3) \end{cases}$$

STEEPEST ASCENT METHOD

Now a reference system change is being made: the new 2^2 plan is centered with respect to an s-t system.



$$\begin{cases} s = u - 1 \\ t = v + \dfrac{9}{8} \end{cases}$$

u = [0; 2] → s = [-1; 1];

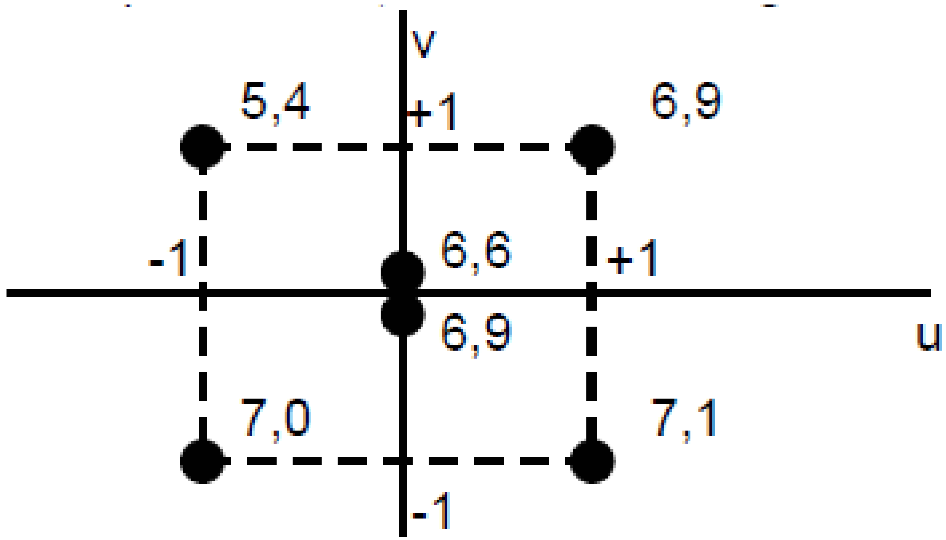
v = [-17/8; -1/8] → t = [-1; 1]

STEEPEST ASCENT METHOD

- At this point we continue, performing a certain number of experiments. At the end of each experiment, the effects ascribable to the two variables are assessed and their significance is tested. In the absence of significance, identifies the maximum increase direction and chooses a new center, to undertake another experiment. If, on the other hand, there is a significance of at least one factor, we stop and proceed with a further check: that of the adequacy of the planar approximation. This verification, satisfied far from the maximum point, should no longer be near the peak.
- To carry out this verification, which is however well to be carried out at the end of each experiment, it is necessary to add some central points.

STEEPEST ASCENT METHOD

- For convenience we take back the values of the previous example, adding 2 central points, we found the following results:



- There are two additional points at the center, which will be used to estimate the error and verify planarity.

STEEPEST ASCENT METHOD

Characteristics of the central points:

- They have no influence on the estimations of the coefficient (of slope) δ1 and δ2.
- It allows an estimate of the error, evaluating the differences in the results of the replicas.
- Allows the execution of the planarity check.

Grand mean = (7,0 + 7,1 + 5,4 + 6,9 + 6,6 + 6,9)/6 = 6,65

TSS = $(7{,}0 – 6{,}65)^2 + (7{,}1 – 6{,}65)^2 + (5{,}4 – 6{,}65)^2 +$

$+ (6{,}9 – 6{,}65)^2 + (6{,}6 – 6{,}65)^2 + (6{,}9 – 6{,}65)^2 = 2{,}015$

STEEPEST ASCENT METHOD

$SSQ_u = 0{,}64$

$SSQ_v = 0{,}81$

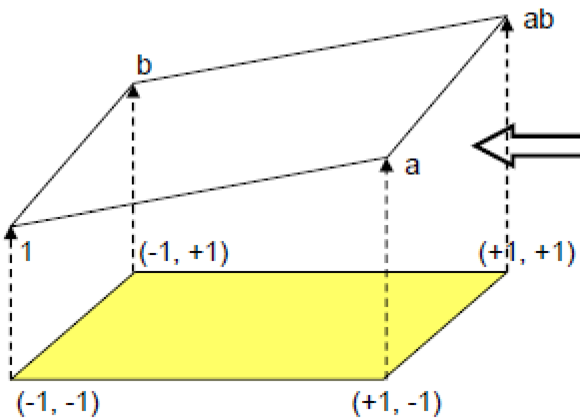
$SSQ_{error} = SSW = (6{,}6 – 6{,}75)2 + (6{,}9 – 6{,}75)2 = 0{,}045$

$SSQ_{deviation\ from\ planarity} = TSS – SSQ_u – SSQ_v – SSW =$

$=2{,}015 – 0{,}64 – 0{,}81 – 0{,}045 = 0{,}52$

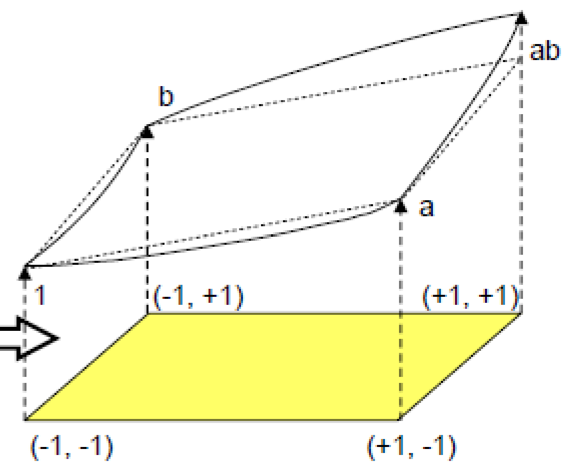
- The term $SSQ_{deviation\ from\ planarity}$ takes into account all those causes, which can determine non-planar behavior: essentially non-linearity and interaction effects.
- The figures in the slide show how the surface of response loses the planarity characteristics, if there is interaction.

STEEPEST ASCENT METHOD



Response surface
without interactions

Response surface
with interactions



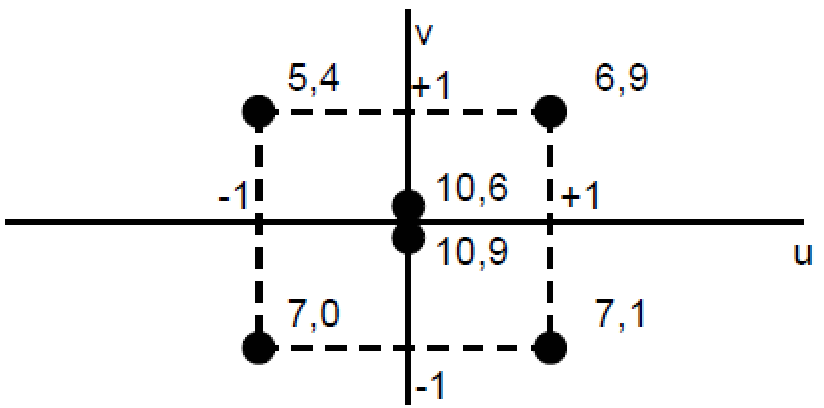
STEEPEST ASCENT METHOD

Effect	SSQ	DoF	MSQ	Fcalc.	p-value
A(u,δ_1)	0.64	1	0.64	1.3	46%
B(v,δ_1)	0.81	1	0.81	1.7	42%
Deviation from the plan e configuration	0.52	2	0.26	5.78	28.2%
Error (SSW)	0.045	1	0.045		
Total	2.015	5			

It must therefore be concluded that no significant effects are observed (at the 15% threshold), both as regards the effects directly attributable to the two variables and as regards non-planarity. The planar model is therefore still adequate, it is still necessary to proceed in the direction of the steepest ascent to find the maximum.

STEEPEST ASCENT METHOD

- If instead it was found, as regards the points in the center:



- We repeat the analysis of the variance:

STEEPEST ASCENT METHOD

Grand mean = (7,0 + 7,1 + 5,4 + 6,9 + 10,6 + 10,9)/6 = 7,98

TSS = (7,0 – 7,98)2 + … + (10,9 – 7,98)2 = 24,95

SSQ_u = 0,64

SSQ_v = 0,81

SSQ_{error} = SSW = (10,6 – 10,75)2 + (10,9 – 10,75)2 = 0,045

$SSQ_{deviation\ from\ planarity}$ = TSS – SSQ_u – SSQ_v – SSW = 24,95 – 0,64 – 0,81 – 0,045 = 23,45

Effect	SSQ	DoF	MSQ	Fcalc.	p-value
$A(u,\delta_1)$	0.64	1	0.64	1.3	46%
$B(v,\delta_1)$	0.81	1	0.81	1.7	42%
Deviation from the planar configuration	23.45	2	11.73	260.6	4.4%
Error (SSW)	0.045	1	0.045		
Total	24.95	5			

STEEPEST ASCENT METHOD

- In this case it is observed that the deviation from planarity is significant, while the effects of the two variables are always below the threshold of significance.
- Indeed, the much greater value of the two central points is interpreted as the presence of a peak inside the four lateral points.
- Therefore, being sure to be near the maximum point, in order to find this more precisely, it is necessary to refine the model, adding the quadratic terms and those linked to the interaction. The goal is to determine the combination of treatments, which maximizes the response and maximized response value.
- The local exploration method is then applied.

METHOD OF THE LOCAL EXPLORATION

We can apply two kinds of plans:

- Cetral composite plan
- Box Plan- Behnken

The central composite plan provides for the evaluation of experimental points inserted according to different criteria.

a) Factorial plan complete or fractional for the estimation of the main effects and interaction and related coefficients
b) Points with a star arrangement, for the estimation of the quadratic components
c) Central points, for the estimation of the error and the evaluation of the curvature of the surface of response.

CENTRAL COMPOSIT PLAN

a) Complete factorial plan or fractional factorial

- The most important requirement is that this plan must be able to estimate the main effects, in addition to two-factor interactions. In other words the resolution (minimum number of letters of the confused effects) must be 5 or higher. If it is 5, the main effects are in combination with 4-factor interactions, the interactions between 2 factors are interactions with 3 factors.
- For example, if you have 4 factors, a fractional plan can have at most 4 (main effects in conjunction with 3-factor interactions, 2-factor interactions in combination with each other, so you can not resort to a 2^{4-1}) must keep the plan complete 2^4. If the factors are 5, however, you can use a plan 2^{5-1}.

CENTRAL COMPOSIT PLAN

b) Components with a star arrangment

- Points placed axially at $\pm P_s$ distances from the center are added in positive and negative directions. The coordinates of these points for a generic number of factors are substantially the following.

X_1	$+P_s$	0	0	...	0	0
	$-P_s$	0	0	...	0	0
X_2	0	$+P_s$	0	...	0	0
	0	$-P_s$	0	...	0	0
	...					
X_k	0	0	0	...	0	$+P_s$
	0	0	0	...	0	$-P_s$

Ps=f(k,k-p)

For instance in a complete 2^4,

Ps=2,

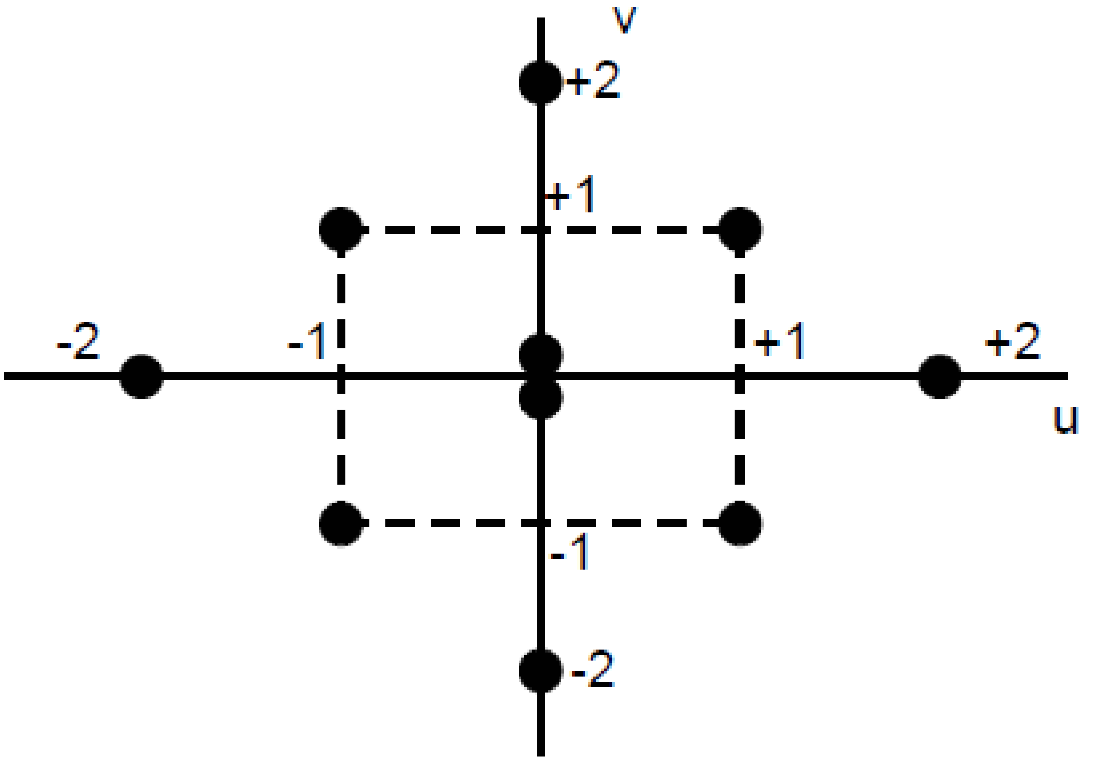
Typical value: 2

CENTRAL COMPOSIT PLAN

c) Central Points

- Finally, there are some central points, which, as seen, allow the estimation of the error, as well as evaluations on the curvature of the surface, allowing a more reliable calculation of the quadratic terms. There is an optimal number of points depending on k and (k-p), typically it is between 2 and 4.
- In the case of two factors, a plan like the following should be used:
- a) Points in the four corners → 2^2 plan
- b) Points over the axes → star points
- c) Points in the center → central points

CENTRAL COMPOSIT PLAN



PIANO BOX-BEHNKEN

- It is a project strategy that requires that the factors be evaluated at only 3 levels (unlike the composite central plan which provides for 5-level evaluations: -Ps, -1, 0, +1, + Ps).
- It is made up of the combination of all the possible plans 2^2 (complete factorial with 2 factors, each with 2 levels), to involve all the factors.
- Then, central points are added, in a similar way to the composite central plane, to carry out an estimate of the uncertainty and to evaluate the curvature of the response surface. Typically, 2 to 4 center points are added.

PIANO BOX-BEHNKEN

- In the presence of 4 factors, X_1, X_2, X_3, X_4, to evaluate all the possible combinations, 4 floors 2^2 must be considered. The table shows the coordinates of the points.

	N.	X_1	X_2	X_3	X_4
2^2 fra X_1 e X_2	1	-1	-1	0	0
	2	+1	-1	0	0
	3	-1	+1	0	0
	4	+1	+1	0	0
2^2 fra X_1 e X_3	5	-1	0	-1	0
	6	+1	0	-1	0
	7	-1	0	+1	0
	8	+1	0	+1	0

	N.	X_1	X_2	X_3	X_4
2^2 fra X_1 e X_4	9	-1	0	0	-1
	10	+1	0	0	-1
	11	-1	0	0	+1
	12	+1	0	0	+1
2^2 fra X_2 e X_3	13	0	-1	-1	0
	14	0	+1	-1	0
	15	0	-1	+1	0
	16	0	+1	+1	0

9 → 12: 2^2 fra X_1 e X_4

13 → 16: 2^2 fra X_2 e X_3

PIANO BOX-BEHNKEN

N.	X_1	X_2	X_3	X_4
17	0	-1	0	-1
18	0	+1	0	-1
19	0	-1	0	+1
20	0	+1	0	+1
21	0	0	-1	-1
22	0	0	+1	-1
23	0	0	-1	+1
24	0	0	+1	+1

2^2 fra X_2 e X_4 (prove 17–20)

2^2 fra X_3 e X_4 (prove 21–24)

Punti centrali

N.	X_1	X_2	X_3	X_4
25	0	0	0	0
26	0	0	0	0
...	0	0	0	0

CENTRAL COMPOSIT VS. BOX-BEHNKEN

In the case of 4 factors, the composite central plane requires:

- A complete factorial plan 2^4 for a total of 16 points
- $2 \cdot 4 = 8$ star points along the axes
- 2 to 4 central points

In the case of 4 factors, the Box-Behnken plane requires:

- 6 complete factorial plans 2^2, for a total of $6 \cdot 4 = 24$ points
- 2 to 4 central points

In conclusion both plans require the execution of experimental tests at 24 points, with the addition of 2 or 4 central points. For 4 factors the number of combinations is the same in the two cases. The cases are now evaluated by number of factor k minor or major 4.

CENTRAL COMPOSIT VS. BOX-BEHNKEN

In the case of 3 factors, the composite central plane requires:

- A complete factorial plan 2^3 for a total of 8 points
- $2 \cdot 3 = 6$ star points along the axes
- 2 to 4 central points

In the case of 3 factors, the Box-Behnken plan requires:

- 3 complete factorial plans 2^2, for a total of $3 \cdot 4 = 12$ points
- 2 to 4 central points

So for $k = 3$ is the Box-Behnken plan to require less combinations (12 versus 14, apart from the central points), this is generally valid for $k < 4$ too.

CENTRAL COMPOSIT VS. BOX-BEHNKEN

In the case of 5 factors, the composite central plane requires:

- A complete factorial plan 2^{5-1} for a total of 16 points
- $2 \cdot 5 = 10$ star points along the axes
- 2 to 4 central points

In the case of 5 factors, the Box-Behnken plan requires:

- 10 complete factorial plans 2^2, for a total of $10 \cdot 4 = 40$ points
- 2 to 4 central points

So for k = 5 is the composite central plane to require less combinations (26 versus 40, apart from the central points), this is generally valid for k> 4. In general the composite central plane is preferable, unless it is very expensive to perform tests on 5 levels.

METHOD OF THE LOCAL EXPLORATION

At this point, based on the detected experimental data, a mathematical model of the second order can be outlined, which links the dependent variable to the independent k variables (X_1, ..., X_k). This model considers, in addition to the linear terms, also the quadratic and interaction ones that are neglected in the first phase of the experimentation. All terms of the third order or of higher orders are neglected. The model is presented in the form:

- $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 +\ldots+ \beta_k X_k +$ Linear Terms
- $+ \beta_{11} X_1 2 + \beta_{22} X_2 2 +\ldots+ \beta_{kk} X_k 2 +$ Quadratic Terms
- $+ \beta_{12} X_1 X_2 + \beta_{13} X_1 X_3 +\ldots+ \beta_1 kX1Xk +\ldots+ \beta_{23} X_2 X_3 + \beta_{24} X_2 X_4 +\ldots+ \beta_{2k} X_2 X_k +\ldots+ \beta_{(k-1)k} X_{k-1} X_k$ Interaction Terms

METHOD OF THE LOCAL EXPLORATION

This mathematical relation can be used to make numerical predictions on the output value for different combinations and above all to define among them the optimal one, which maximizes the response itself.

To find the maximum point, calculate the Y gradient and set it to zero.

$$\begin{cases} \dfrac{\partial Y}{\partial X_1} = 0 \\ \dfrac{\partial Y}{\partial X_2} = 0 \\ \dots \\ \dfrac{\partial Y}{\partial X_k} = 0 \end{cases}$$

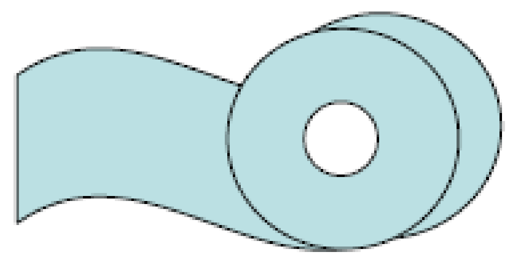
The points thus found can be points of maximum, minimum or saddle points (i.e. maximum for some directions, minimum for others). It is therefore necessary to calculate the function in a neighborhood of the points found, to understand if they are actually points of maximum.

CASE STUDY 1: FRICTION

A leading company in the field of automatic packaging machines commissioned a study aimed at investigating the behavior of plastic film friction. We want to evaluate how the dynamic friction coefficient is influenced by 5 operating factors. In particular, one wants to know from which factors depends and with what kind of dependence.

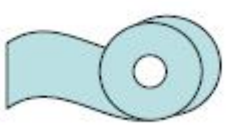
Packet:
dimensions:
80 mm x 60 mm x 20 mm

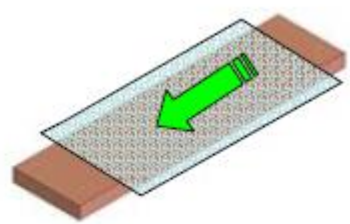
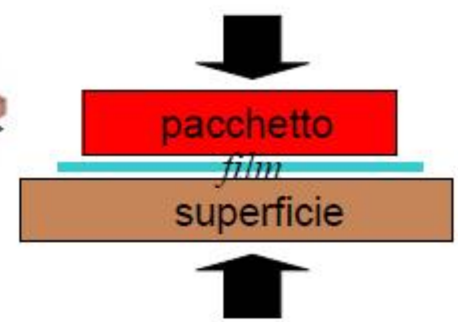
BOPP
polipropilene film a double orientation



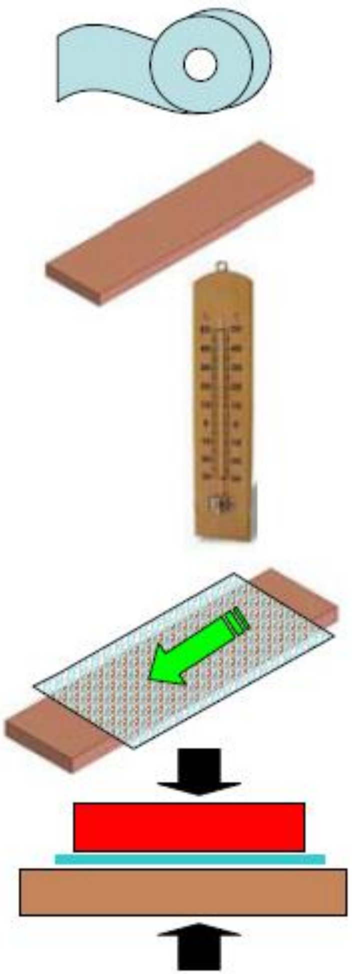
FACTORS

- Research factors having a potential impact over the process.
  a) Physical and chemical properties of the film
  b) Characteristics of the sliding surface
  c) Temperature
  d) Sliding speed
  e) Pressure at the packet-film interface



CASE STUDY 1: SELECTION OF LEVELS



Categorical variables:

- a) 15 films from different vendors
- b) 4 different surfaces

For continuous variables, to get the eventual non-linearity

- c) at least 3 levels of temperature (22°C ÷ 150°C)
- d) at least 3 levels of sliding speeds (150 ÷ 1200 mm/min)
- e) at least 3 levels of interface pressure

CASE STUDY 1: WORK PLANIFICATION

The execution of tests in all possible combinations would have involved the performance of 15·4·3·3·3 = 1620 experiments. Considering the performance of 3 replicas, to allow the estimation of uncertainty and to give statistical relief to the data, the number of trials would have risen to 1620·3 = 4860. This amount of evidence was incompatible with the available resources. We proceeded therefore considering two phases of the work.

1st Phase (screening)	**Using a $2^{k-p} = 2^{5-1}$ plan to understand which factors are significant and which of these are the most influential.**
2nd Phase (deepening)	2-factor ANOVA with decomposition of the quadratic sums and linearity analysis.

CASE STUDY 1: FIRST PHASE SCREENING

- The first phase basically serves to reduce the number of factors. Since we wanted to do three repetitions and still keep the resources for the second phase, we worked with a 2^{5-1} plan.
- The 32 combinations are split by subdivision into two blocks, of which only one is tested. The 5-factor plan is then tested at the price of one to 4, with 16 combinations and $16 \cdot 3 = 48$ experiments.
- Of the 31 (2^5-1) effects, one is lost (which can be chosen arbitrarily) and there remain 30. These 30 are divided into 15 pairs aliases (2^4-1) by two effects each.
- It is chosen to confuse the interaction between all the factors (it is the effect that interests less, probably negligible).

CASE STUDY 1: FIRST PHASE SCREENING

	Low level (0)	High Level (1)
Factor A: film kind	C72	Al 5083 alloy
Factor B: surface material	GLS 20	RC 25
Factor C: temperature	22º C	50º C
Factor D: scroll speed	150 mm/min	1200 mm/min
Factor E: interface pressure	Reference weight	4 (reference weight)

CASE STUDY 1: FIRST PHASE SCREENING

- I = ABCDE
- Alias pairs:
- A = BCDE; B = ACDE; C = ABDE; D = ABCE; E = ABCD;
- AB = CDE; AC = BDE; AD = BCE; AE = BCD; BC = ADE;
- BD = ACE; BE = ACD; CD = ABE; CE = ABD; DE = ABC
- Then we will see if the (=) should be replaced with a (+) or a (-).
- Resolution = 5, it is possible to evaluate main effects and interactions between two factors
- Once this picture of alias pairs was accepted, the two blocks were determined, starting from the main one. In this case (I = ABCDE) all the combinations that have an even number of letters in Yates notation are part of the main block.

CASE STUDY 1: FIRST PHASE SCREENING

Principal Block		Secondary Block	
1	cd	a	acd
ab	ce	b	ace
ac	de	c	ade
ad	abcd	d	bcd
ae	abce	e	bce
bc	abde	abc	bde
bd	acde	abd	cde
be	bcde	abe	abcde

The combinations of the main block were chosen and the order of the trials was randomized. Once the tests were carried out, the effects were evaluated using the modified Yates algorithm.

CASE STUDY 1: FIRST PHASE SCREENING

Yates Comb.	Tested Comb.	Results	(1)	(2)	(3)	(4)	Effects	
1	1	0.1892	0.40	0.74	1.65	3.30	0.2064	Average
a	a(d)	0.2136	0.33	0.92	1.65	-0.22	-0.0272	A-BCDE
b	b(d)	0.1717	0.61	0.74	-0.07	-0.82	-0.1024	B-ACDE
ab	ab	0.1616	0.30	0.91	-0.15	0.42	0.0522	AB-CDE
c	c(d)	0.3670	0.43	0.01	-0.38	0.35	0.0444	C-ABDE
ac	ac	0.2443	0.30	-0.08	-0.44	-0.24	-0.0296	AC-BDE
bc	bc	0.1317	0.61	0.00	0.13	-0.42	-0.0524	BC-ADE
abc	abc(d)	0.1729	0.30	-0.15	0.29	0.23	0.0283	ABC-DE
e	(d)e	0.2507	0.02	-0.07	0.18	0.00	-0.0002	E-ABCD
ae	ae	0.1835	-0.01	-0.31	0.17	-0.08	-0.0104	AE-BCD
be	be	0.1204	-0.12	-0.13	-0.10	-0.07	-0.0084	BE-ACD
abe	ab(d)e	0.1831	0.04	-0.31	-0.14	0.16	0.0198	ABE-CD
ce	ce	0.3823	-0.07	-0.03	-0.24	0.00	-0.0006	CE-ABD
ace	ac(d)e	0.2304	0.06	0.16	-0.18	-0.05	-0.0057	ACE-BD
bce	bc(d)e	0.1470	-0.15	0.13	0.20	0.06	0.0069	BCE-AD
abce	abce	0.1530	0.01	0.16	0.03	-0.17	-0.0213	ABCE-D

CASE STUDY 1: FIRST PHASE SCREENING

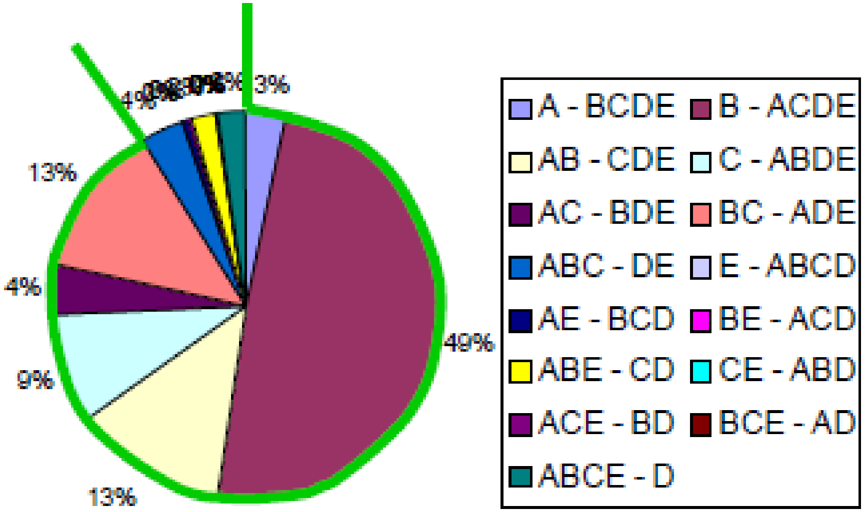
	SSQ	DoF	MSQ	F_{calc}	p-value
SSBC	0.2575	15	0.0172	168.27	4.62E-24
A - BCDE	0.0089	1	0.0089	87.02	1.21E-08
B - ACDE	0.1259	1	0.1259	1234.32	3.92E-25
AB - CDE	0.0327	1	0.0327	320.20	3.16E-16
C - ABDE	0.0236	1	0.0236	231.42	3.35E-14
AC - BDE	0.0105	1	0.0105	103.35	1.51E-09
BC - ADE	0.0330	1	0.0330	323.28	2.74E-16
ABC - DE	0.0096	1	0.0096	94.39	4.57E-09
E - ABCD	0.0000	1	0.0000	0.01	94.02
AE - BCD	0.0013	1	0.0013	12.71	0.12
BE - ACD	0.0008	1	0.0008	8.28	0.71
ABE - CD	0.0047	1	0.0047	46.17	1.11E-05
CE - ABD	0.0000	1	0.0000	0.04	83.59
ACE - BD	0.0004	1	0.0004	3.83	5.90
BCE - AD	0.0006	1	0.0006	5.59	2.43
ABCE - D	0.0055	1	0.0055	53.48	2.59E-06
SSW	0.0033	32	0.0001		
Total	0.2608	47			

CASE STUDY 1: FIRST PHASE SCREENING

- A and B, C and D> 0, (friction increases with temperature and speed), E <0 (the coefficient of friction decreases, very weakly, with pressure).
- To evaluate significance, the usual parallelism was considered with a monofactorial ANOVA of 16 (2^{5-1}) levels. Subsequently, the decomposition of the SSBC is carried out. Every single SSQ_i is of the type: (estimate) 2 · number of repetitions · 2^{k-p-2} = $(\text{estimate})^2 \cdot 3 \cdot 22 = (\text{estimate})^2 \cdot 12$.
- Of the factors, B (surface) is undoubtedly the most significant, then we have C (temperature) and A (type of film) and their interactions are also strong. Also the factor D (velocity) is above the significance threshold, while E (interface pressure) has a very high p-value.

CASE STUDY 1: FIRST PHASE SCREENING

The pie chart shows that the difference between the results of the various treatments is essentially attributable (for 88%, sector highlighted in the figure) of the total to the factors A, B and C and their interactions. D is also significant, but has far less influence.



CASE STUDY 1: SECOND PHASE

- In the next phase we focused on factors A, B and C. We made two-factor ANOVA analyzes (temperature and type of film), taking into account 4 different sliding surfaces.
- Identify the surfaces that guarantee the lowest coefficient of friction and a stable behavior in temperature.
- Evaluate the non-linearity with respect to temperature by performing tests on 3 levels (22, 50 and 80 ° C).

	Kind of Film (15 levels)
Temperature (3 levels)	45 Combinations 45 * 3=135 total tests

CASE STUDY 1: SECOND PHASE

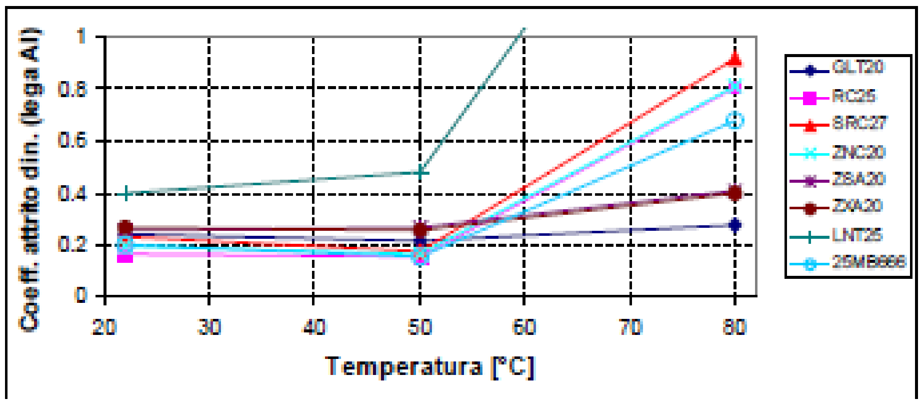
- In the case of Alluminum alloy Al 5083:

	SSQ	DoF	MSQ	Fcalc.	p-value
SSB_R	2.6416	2	1.3208	2284	7.45E-76
SSB_C	5.4938	14	0.3924	679	9.78E-83
Interaction	6.5276	28	0.2331	403	7.48E-81
Error	0.0520	90	0.0006		
Total	14.7151	134			

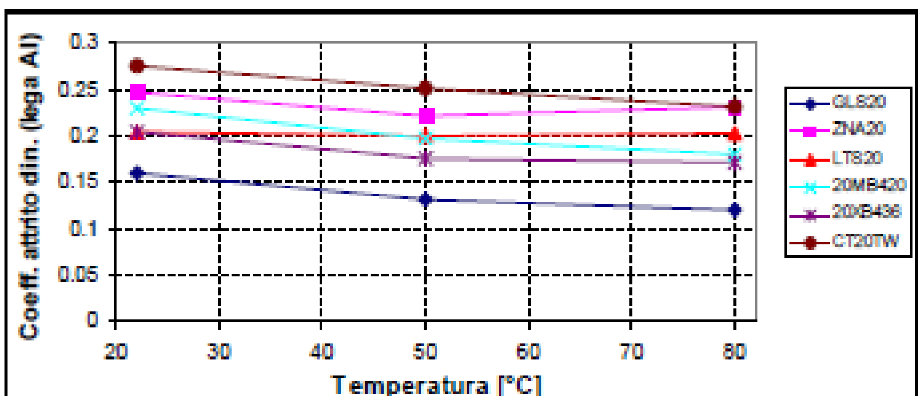
- Strong significance for main effects and interaction.

CASE STUDY 1: SECOND PHASE

- In the case of Alluminum alloy Al 5083:



Non-linear Coeff. Increase per temp. increase



Linear Coeff. Decreasing per temp. decreasing

CASE STUDY 1: SECOND PHASE

- In the case of aluminum alloy Al 5083:
- Evaluation of linearity: the SSB_R (the only numerical variable) is broken into two linear and quadratic components: A_L and A_Q. A in this phase refers to the row factor, the temperature, B to that of column, the type of film. The orthonormal matrix has 2 rows (as many as the DoF) and 45 columns (as many as the combinations).
- Orthogonal:

a_1b_1	a_1b_2	...	a_1b_{15}	a_2b_1	a_2b_2	...	a_2b_{15}	a_3b_1	a_3b_2	...	a_3b_{15}
-28.7	-28.7	...	-28.7	-0,67	-0,67	...	-0,67	29.3	29.3	...	29.3
30	30	...	30	-58	-58	...	-58	28	28	...	28

- Orthonormal:

a_1b_1	a_1b_2	...	a_1b_{15}	a_2b_1	a_2b_2	...	a_2b_{15}	a_3b_1	a_3b_2	...	a_3b_{15}
-0,18	-0,18	...	-0,18	0,00	0,00	...	0,00	0,18	0,18	...	0,18
0,11	0,11	...	0,11	-0,21	-0,21	...	-0,21	0,10	0,10	...	0,10

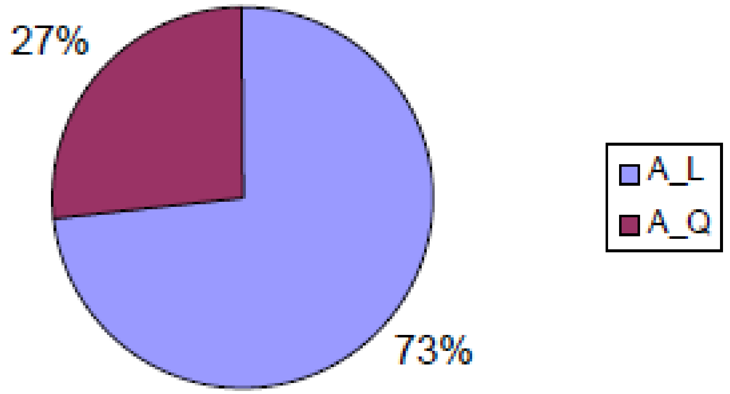
CASE STUDY 1: SECOND PHASE

- By applying the orthonormal matrix to the vector (45 x 1) of the average results in the various combinations, we obtain:
- A_L = 0,80; A_Q = 0,49, from which
- $SSQ_{AL} = (0,80)2 \cdot 3 = 1.9269$; $SSQ_{AQ} = (0,49)2 \cdot 3 = 0.7147$

	SSQ	DoF	MSQ	Fcalc.	p-value
SSB_R	2.6416	2	1.3208	2284	7.45E-76
A_L	1.9269	1	1.9269	3332	6.72E-71
A_Q	0.7147	1	0.7147	1236	2.33E-52
SSB_C	5.4938	14	0.3924	679	9.78E-83
Interaction	6.5276	28	0.2331	403	7.48E-81
Error	0.0520	90	0.0006		
Total	14.7151	134			

CASE STUDY 1: SECOND PHASE

Both the linear and the quadratic effects are significant, the decomposition of the SSB_R is shown below from a graphical point of view. It is noted that the linear component is still prevalent on the quadratic one. On the other hand, however, a non-linearity component is undoubtedly present, as is also clear from the graphs.



CASE STUDY 1: SECOND PHASE

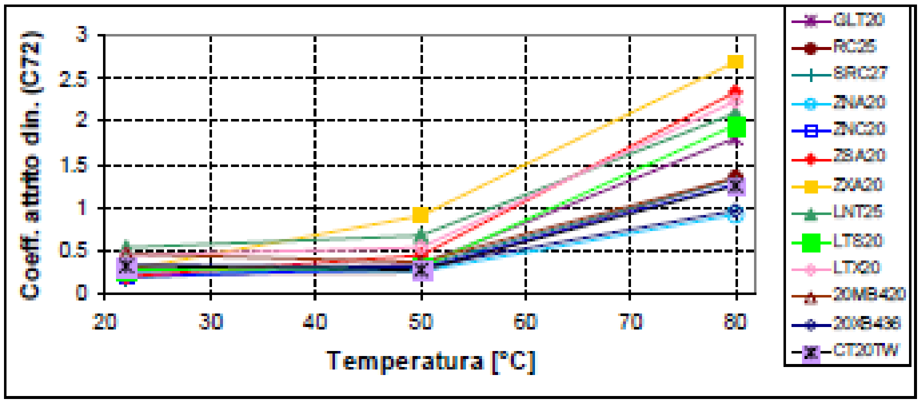
- In the case of steel C72:

	SSQ	DoF	MSQ	Fcalc.	p-value
SSB_R	47.0056	2	23.5028	1181	2.57E-63
SSB_C	8.0103	14	0.5722	29	3.34E-25
Interaction	6.8976	28	0.2463	12	2.31E-18
Error	1.7909	90	0.0199		
Total	63.7044	134			

- Strong significance for main effects and interaction.

CASE STUDY 1: SECOND PHASE

- In the case of steel C72:



- Non linearity
- Trend at high values of the coefficient of friction

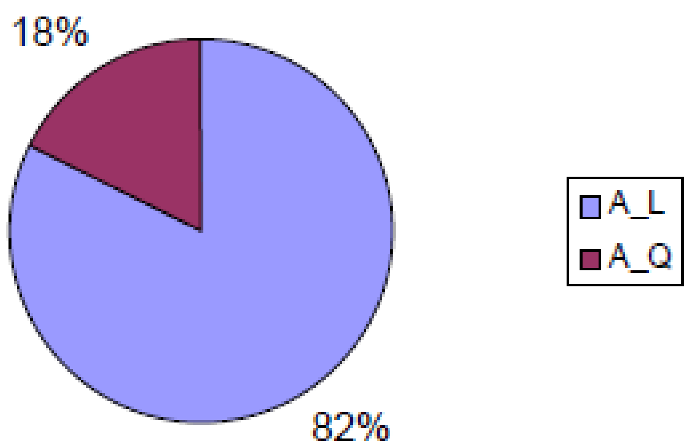
CASE STUDY 1: SECOND PHASE

- By applying the orthonormal matrix (same as before) to the vector of the average results in the various combinations, we obtain:
- A_L = 0,80; AQ = 0,49, from which
- $SSQ_{AL} = (3,59)2{\cdot}3 = 38,5978$; $SSQ_{AQ} = (1,67)2{\cdot}3 = 8,4079$

	SSQ	DoF	MSQ	Fcalc.	p-value
SSB_R	47.0056	2	23.50	1181	2.57E-63
A_L	38.5978	1	38.60	1940	1.10E-60
A_Q	8.4079	1	8.41	423	9.29E-34
SSB_C	8.0103	14	0.57	29	3.34E-25
Interaction	6.8976	28	0.25	12	2.31E-18
Error	1.7909	90	0.02		
Total	63.7044	134			

CASE STUDY 1: SECOND PHASE

Both the linear and the quadratic effects are significant, the decomposition of the SSB_R is shown below from a graphical point of view. The same considerations as previously seen regarding the prevalence of the linear contribution on the quadratic one, however present, are valid.



CASE STUDY 1: SECOND PHASE

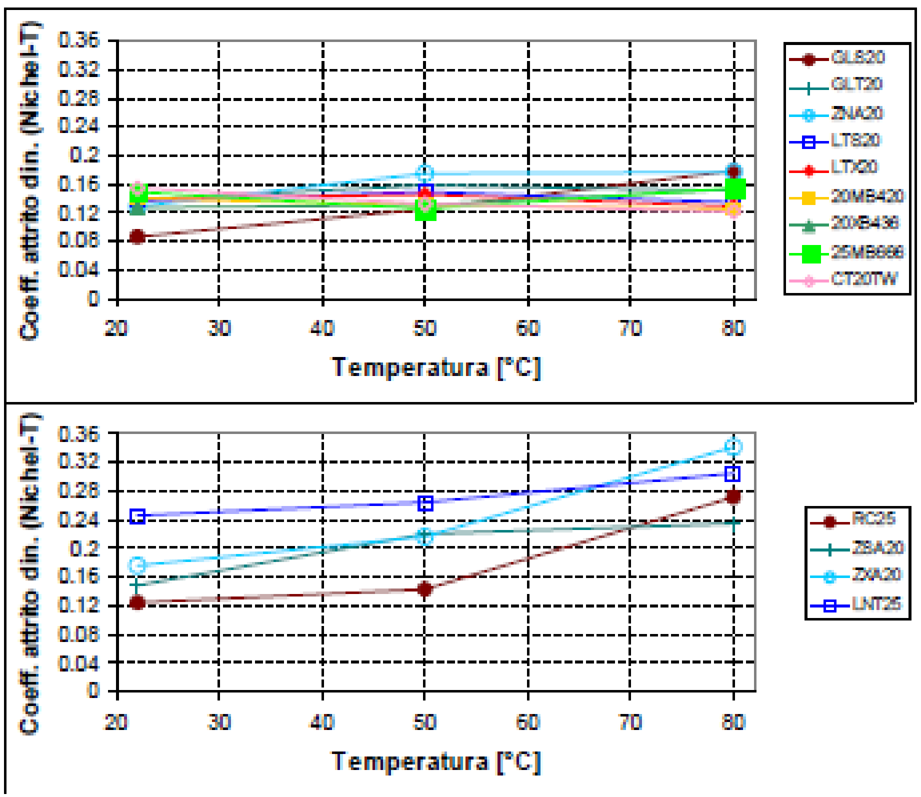
- In the case of the Nickel-T coating :

	SSQ	DoF	MSQ	Fcalc.	p-value
SSB_R	0.0547	2	0.0273	680	4.71E-53
SSB_C	0.2339	14	0.0167	416	2.71E-73
Interaction	0.1048	28	0.0037	93	7.16E-53
Error	0.0036	90	0.0000		
Total	0.3971	134			

- Strong significance for main effects and interaction

CASE STUDY 1: SECOND PHASE

- In the case of the Nickel-T coating :



Stable behaviour

Values increasing with good linearity

CASE STUDY 1: SECOND PHASE

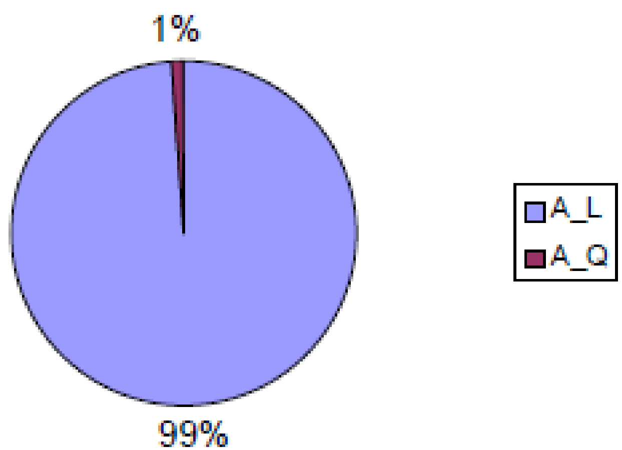
By applying the orthonormal matrix (same as before) to the vector of the average results in the various combinations, we obtain:

- $A_L = 0{,}134$; $A_Q = 0{,}013$, da cui
- $SSQ_{AL} = (0{,}134)2 \cdot 3 = 0{,}0541$; $SSQ_{AQ} = (0{,}013)2 \cdot 3 = 6E\text{-}4$

	SSQ	DoF	MSQ	Fcalc.	p-value
SSB_R	0.0547	2	0.0273	680	4.71E-53
A_L	0.0541	1	0.0541	1347	6.25E-54
A_Q	6E-4	1	6E-4	14	3.69E-02
SSB_C	0.2339	14	0.0167	416	2.71E-73
Interaction	0.1048	28	0.0037	93	7.16E-53
Error	0.0036	90	4.02E-5		
Total	0.3971	134			

CASE STUDY 1: SECOND PHASE

Both the linear and the quadratic effects are significant, the decomposition of the SSB_R is shown below from a graphical point of view. We note that the quadratic component is almost zero, just 1% compared to the linear one.



CASE STUDY 1: SECOND PHASE

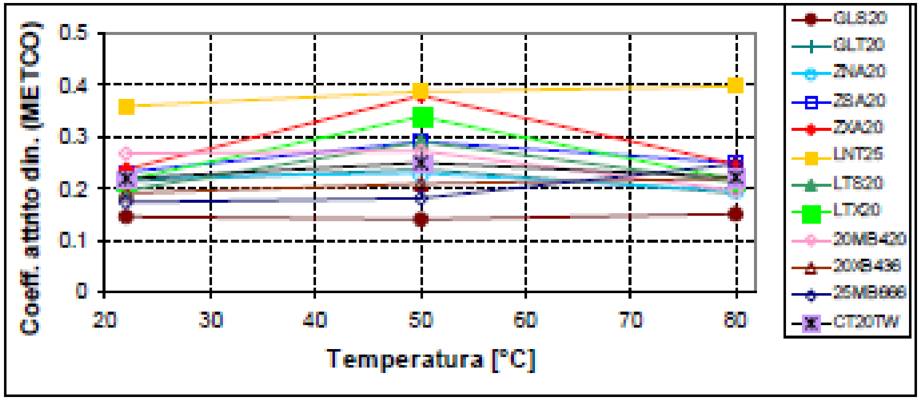
- In the case of the METCO coating:

	SSQ	DoF	MSQ	Fcalc.	p-value
SSB_R	0.0994	2	0.0497	612	3.92E-51
SSB_C	0.4473	14	0.0319	393	3.11E-72
Interaction	0.5439	28	0.0194	239	8.63E-71
Error	0.0073	90	0.0001		
Total	1.0980	134			

- Strong significance for main effects and interaction.

CASE STUDY 1: SECOND PHASE

- In the case of the METCO coating:



- Stable value of temperature
- Relative maximum at t = 50°C

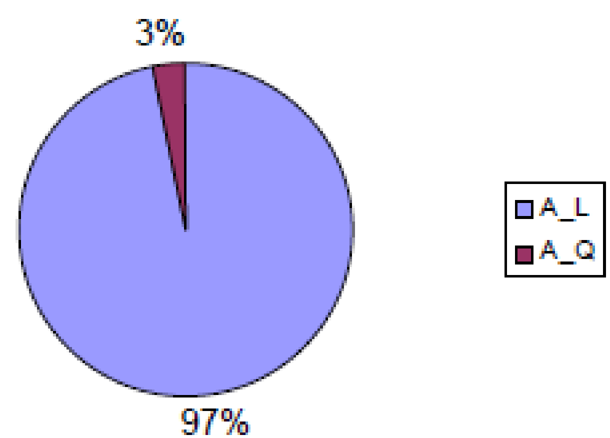
CASE STUDY 1: SECOND PHASE

- By applying the orthonormal matrix (same as before) to the vector of the average results in the various combinations, we obtain:
- $A_L = 0,179$; $A_Q = -0,032$, from which
- $SSQ_{AL} = (0,179)2 \cdot 3 = 0,0963$;
- $SSQ_{AQ} = (-0,032)2 \cdot 3 = 0,0032$

	SSQ	DoF	MSQ	Fcalc.	p-value
SSB_R	0.0994	2	0.0497	612	3.92E-51
A_L	0.0963	1	0.0963	1186	1.32E-51
A_Q	0.0032	1	0.0032	39	1.42E-06
SSB_C	0.4473	14	0.0319	393	3.11E-72
Interaction	0.5439	28	0.0194	239	8.63E-71
Error	0.0073	90	0.0001		
Total	1.0980	134			

CASE STUDY 1: SECOND PHASE

Both the linear and the quadratic effects are significant, the decomposition of the SSB_R is shown below from a graphical point of view. It is noted that the quadratic component is almost zero, just 3% compared to the linear one.



CASE STUDY 1: SECOND PHASE

- Final balance :
- Initially scheduled tests : 4860

48 in Screening phase

- Tests actually implemented : 588
- (reduction of a factor of 9, -88%)

540 in deepening phase

Preferable nickel-T and METCO coatings for values lower than the dynamic coefficient of friction, as well as for better stability in temperature with predominantly linear behavior (even if a non-linearity component is always significant).

Questions