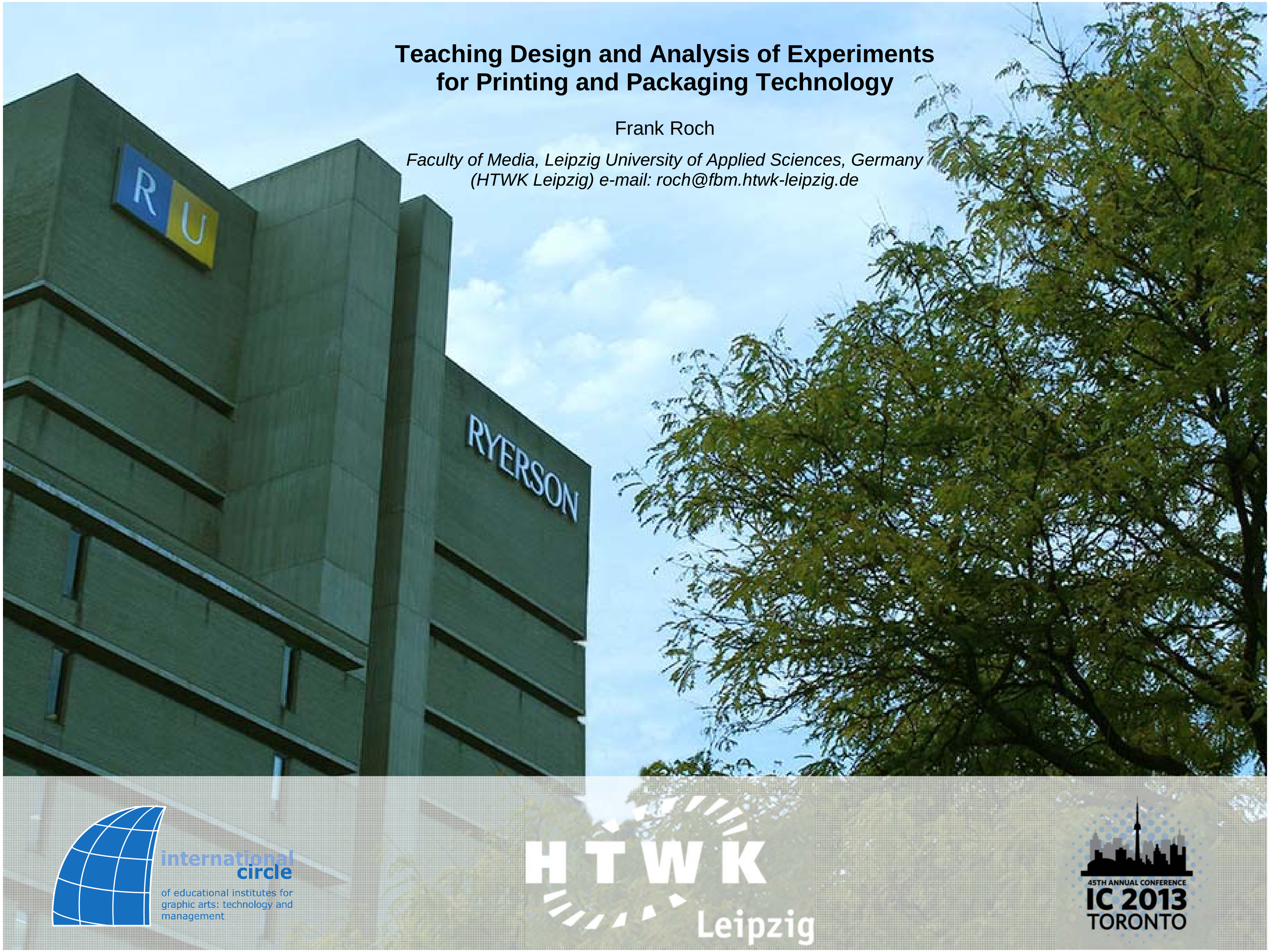
Teaching Design and Analysis of Experiments for Printing and Packaging Technology

Frank Roch

Faculty of Media, Leipzig University of Applied Sciences, Germany (HTWK Leipzig) e-mail: roch@fbm.htwk-leipzig.de



international circle of educational institutes for graphic arts: technology and management

HTWK Leipzig

45TH ANNUAL CONFERENCE IC 2013 TORONTO

Teaching Design and Analysis of Experiments for Printing and Packaging Technology

Frank Roch

Faculty of Media, Leipzig University of Applied Sciences, Germany (HTWK Leipzig) e-mail: roch@fbm.htwk-leipzig.de

DoE Structure of the presentation

- Significance of Experimental Data
- Factorial Design
- Teaching Example of Use
- Guidelines for Designing Experiments
- Conclusion and Recommendation

Significance of Experimental Data

Students in many cases

- on the one hand **overlook effects**,
  that means to <u>ignore</u> them
  (because of measurement uncertainties
  or too less replications),
- on the other hand **construe effects**
  although they <u>don’t exist truly</u>,
  that means to read tea leaves
  (because of useless precision assumed
  and missing correct statistical calculations).

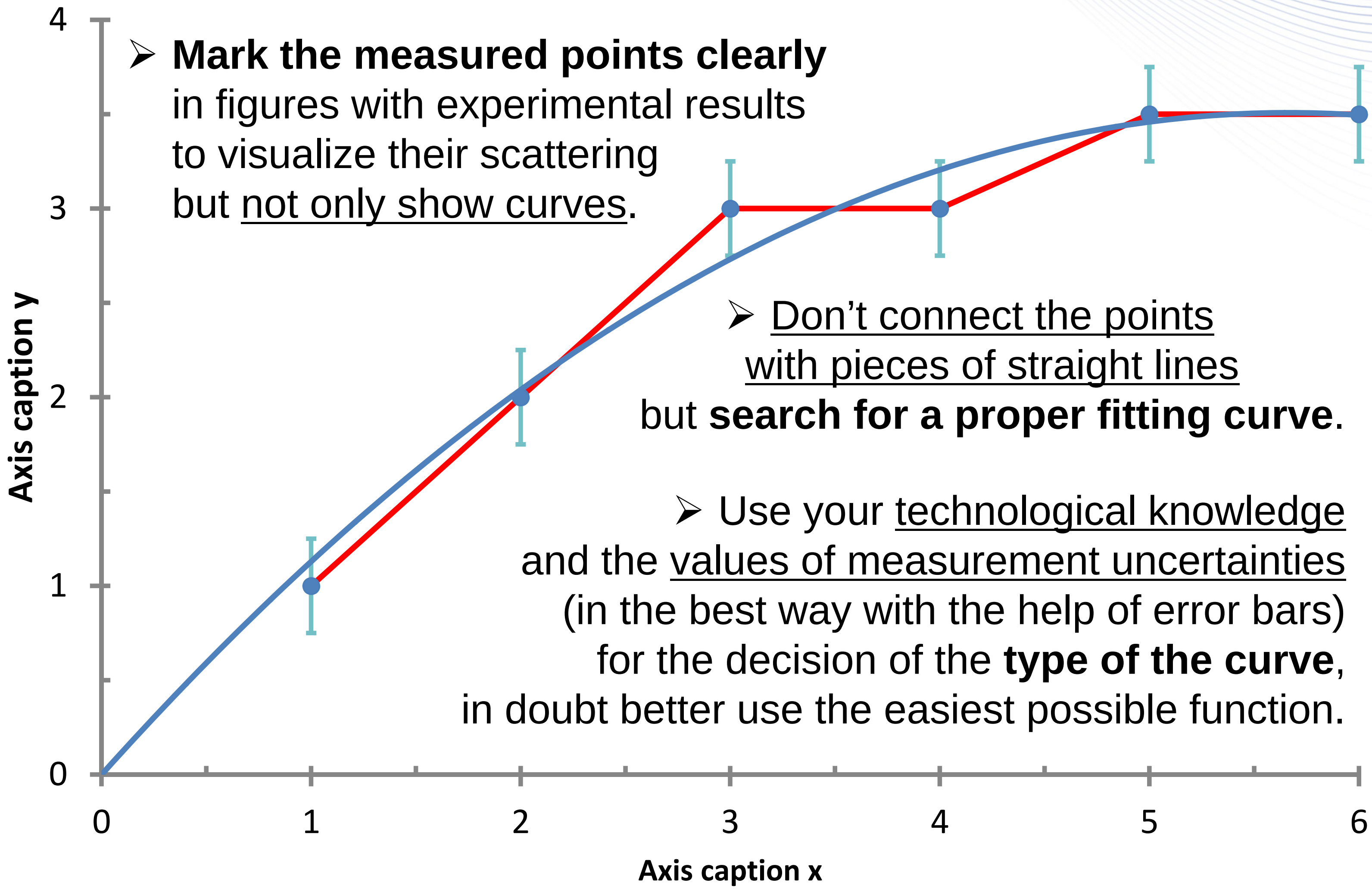
Recommended literature (partly used here)

[1] MONTGOMERY, Douglas C:
Design and Analysis of Experiments,
John Wiley & Sons Inc. (2009)

[2] BOX, George E. P.; HUNTER, J. Stuart; HUNTER, William G.:
Statistics for Experimenters,
John Wiley & Sons Inc. (2009)

[Several References] Bibliography on text books covering DoE
in German language (“Statistische Versuchsplanung”)
to be requested from the author of this presentation

Graphic charts of experimental data with measurement uncertainties

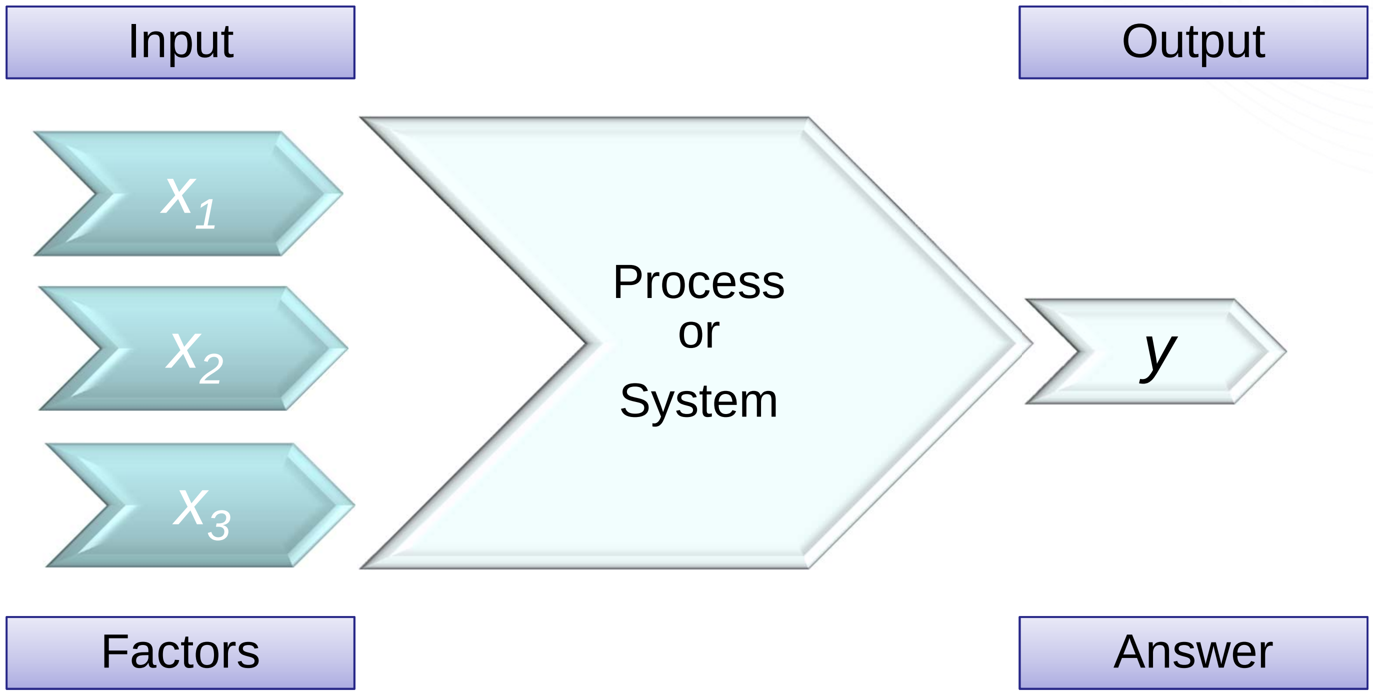


- **Mark the measured points clearly** in figures with experimental results to visualize their scattering but <u>not only show curves</u>.

- <u>Don't connect the points with pieces of straight lines</u> but **search for a proper fitting curve**.

- Use your <u>technological knowledge</u> and the <u>values of measurement uncertainties</u> (in the best way with the help of error bars) for the decision of the **type of the curve**, in doubt better use the easiest possible function.

Factorial Design



Comparison of the approaches OFT and DoE

Treatment	x_1	x_2	x_3	Output
1	0	0	0	
2	1	0	0	
3	0	1	0	
4	0	0	1	

a) One-Factor-at-a-Time (left)

b) Design of Experiments (right)

Treatment	x_1	x_2	x_3	Output
1	-	-	-	
2	+	-	-	
3	-	+	-	
4	+	+	-	
5	-	-	+	
6	+	-	+	
7	-	+	+	
8	+	+	+	

number k of factors $k = 3$

number l of levels $l = 2$

→ l^k factorial design $a = 2^3 = 8$

2^3-design

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 + b_{123}x_1x_2x_3$$

Treatment	I	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$	Output
1	+	-	-	-	+	+	+	-	y_1
2	+	+	-	-	-	-	+	+	y_2
3	+	-	+	-	-	+	-	+	y_3
4	+	+	+	-	+	-	-	-	y_4
5	+	-	-	+	+	-	-	+	y_5
6	+	+	-	+	-	+	-	-	y_6
7	+	-	+	+	-	-	+	-	y_7
8	+	+	+	+	+	+	+	+	y_8
Effects	$\bar{y}$	$ME1$	$ME2$	$ME3$	$IE12$	$IE13$	$IE23$	$IE123$	
Coefficients	b_0	b_1	b_2	b_3	b_{12}	b_{13}	b_{23}	b_{123}	

a is the number of treatments

$$\bar{y} = \sum_{i=1}^{a} y_i$$

$$E_{\dots} = \frac{\sum_{i=1}^{a} x_{i\dots}y_i}{a/2}$$

$$b_{\dots} = E_{\dots}/2$$

$$b_0 = \bar{y}$$

Normally, the output is realised n times.

For factorial designs of first order (linear model) with a treatments which are replicated n times the experimental variance of an effect can be calculated as the sum of the variances of the mean responses at high and low level of the factors

$$var(E) = s_E^2 = \frac{s^2}{an/2} + \frac{s^2}{an/2} = \frac{4s^2}{N}$$

with a total number $N = a \cdot n$ of measured single values.

Choosing the level of significance α (probability to predicate an effect E although it is not true) the propagated measuring error ΔE of an effect is

$$\Delta E = \mathrm{t}(\nu, P) \cdot s_E = t\left(\nu; 1 - \frac{\alpha}{2}\right) \cdot \frac{2s}{\sqrt{N}}$$

with the degree ν of freedom for the estimation s^2 of the single values experimental variance. The values of $\mathrm{t}(\nu, P)$ are quantiles of the t-distribution.

A true effect η will fall into the confidence interval

$$E - \Delta E \leq \eta \leq E + \Delta E$$

in the environment of the calculated effect E. If this interval contains the value zero the effect is not verified with a risk α to be really existent.

Comparison of the approaches OFT and DoE

Treatment	x_1	x_2	x_3	Output
1	0	0	0	
2	1	0	0	
3	0	1	0	
4	0	0	1	

a) One-Factor-at-a-Time (left)

Treatment	x_1	x_2	x_3	Output
1	-	-	-	
2	+	-	-	
3	-	+	-	
4	+	+	-	
5	-	-	+	
6	+	-	+	
7	-	+	+	
8	+	+	+	

b) Design of Experiments (right)

number k of factors $k = 3$

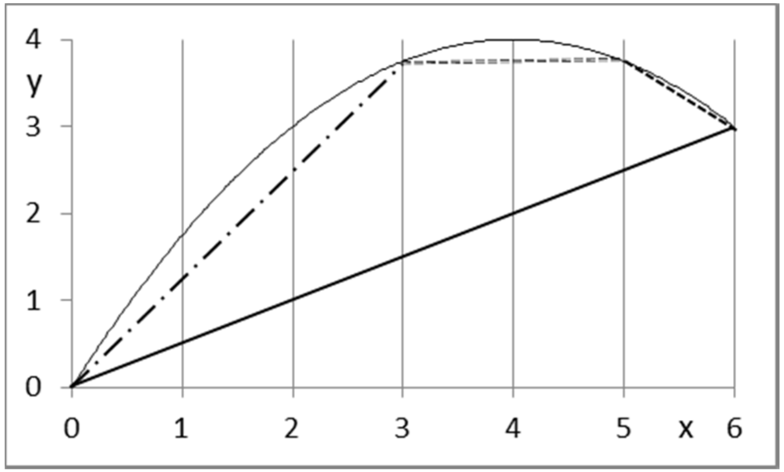
number l of levels $l = 2$

→ l^k factorial design $a = 2^3 = 8$

The use of factorial designs with **DoE**
will truncate the confidence intervals of the effects at the same time
with the decrease of the necessary number of experiments.

Restrictions of the Linear Model

It is very simple, but don't forget the danger of mistake in reasoning shot
if the true dependence of y from (one of the) x_i doesn't follow the linear model.



In this example the increase of the output within the full set range (1… 6) of input is stronger in the range for instance (0… 3), but for example between the values 5 and 6 there is even a decrease.
Moreover, if the input setting is just to 3 and 5 there would be the conclusion that the output doesn't depend from the input value at all.

Second-Order Factorial Designs

k input variables, each of them 3 levels with standardised values (-1; 0; 1)

Quadratic model with 2 variables:

complete 3^2-factorial with 9 treatments
regression equation

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_{11} x_1^2 + b_{22} x_2^2 + b_{12} x_1 x_2$$

only six terms → over-determined

With a mathematical trick it is possible to ascribe the calculation of the regression coefficients of a quadratic model to that of a linear model.

Special design for $l = 3$ levels and $k = 3$ factors: BOX-BEHNKEN
Designs containing the point (0; 0; 0): Central Composite Designs

Response surface methodology
is a collection of techniques to optimise the response.
Many response surface problems involve the analysis of multiple responses.
The task is to find a set of operation conditions of the system
that optimises all responses or at least keep them in desired ranges.

Analysis of Variance

Sources of variability of the system answer:

- Varying treatment, hopefully of course
- Measuring errors, sadly always
- Model defect, checked with over-determined design

Approach:

Partitioning of the total variability into its component parts using sums of squares in the Analysis of Variance (ANOVA)

Realisation:

Comparison of the variance of the effect to be proved with the variance coming from the disturbance

Decision about Significance:

Dedicated standard test function is the F-Distribution**.**
The effect of a variable is significant with the probability α of error, if the calculated value of F is greater than the related tabulated quantile.

Teaching Example of Use

Seal Strength of special polymer foils

The seal is made by a laboratory sealing device.
The *strength* is measured as force in the unit N to break the seal.

Factors:

1. the sealing temperature x_1
   with the levels 165 and 185 in the unit degree Celsius,

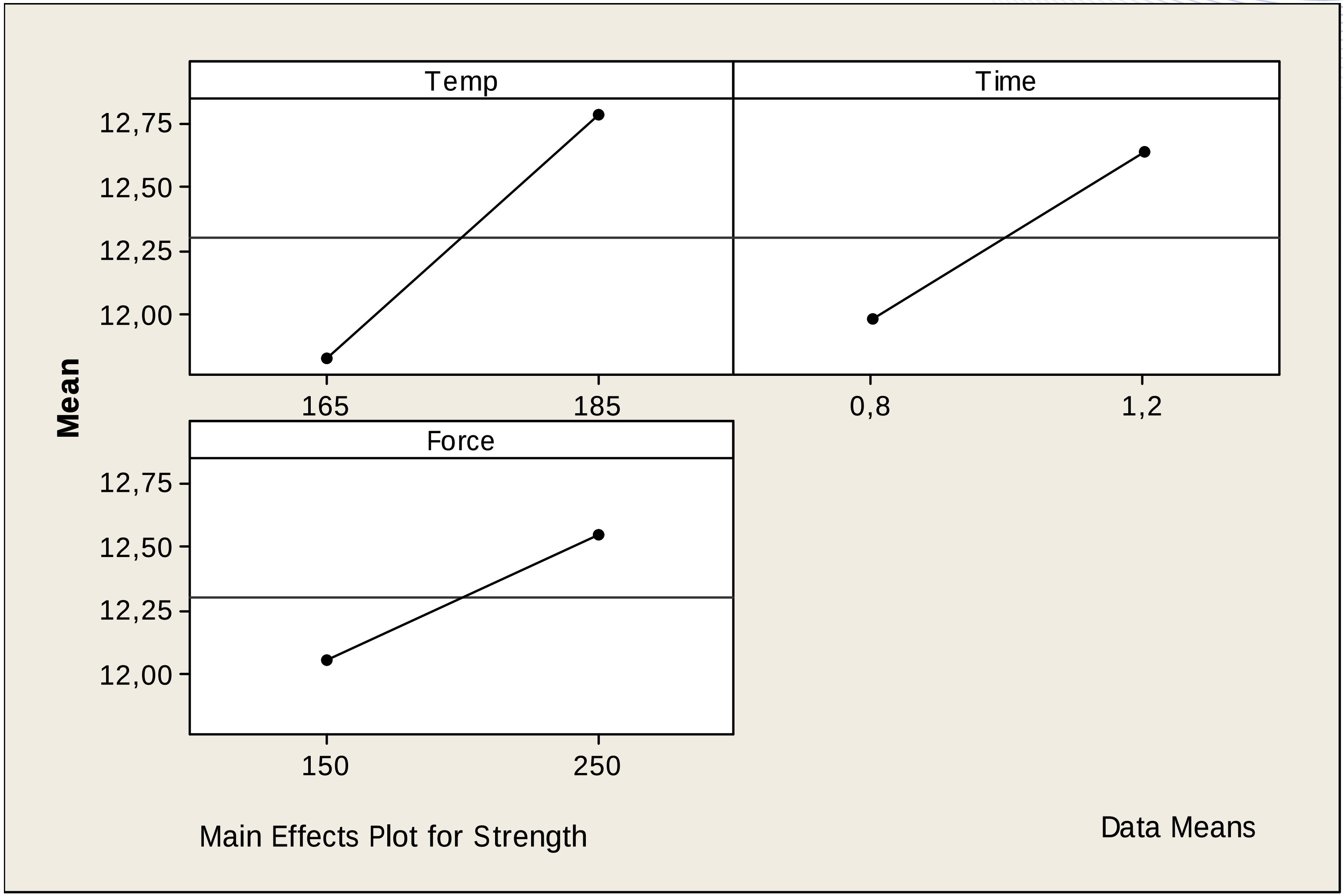
2. the sealing duration ("time") x_2
   with the values 0.8 and 1.2 in the unit seconds

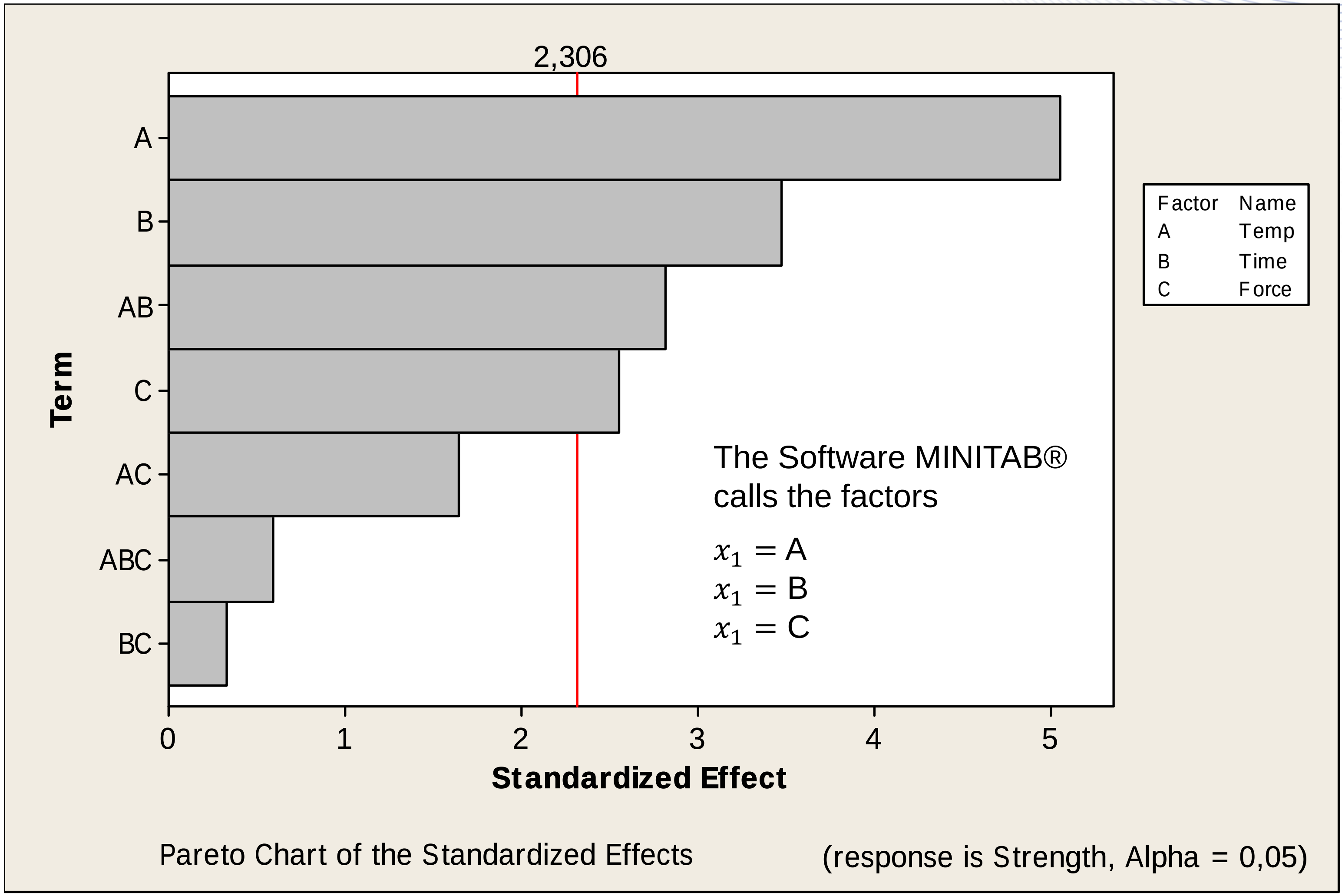
3. and the sealing pressure x_3
   with the level values 150 and 250 in Newton.

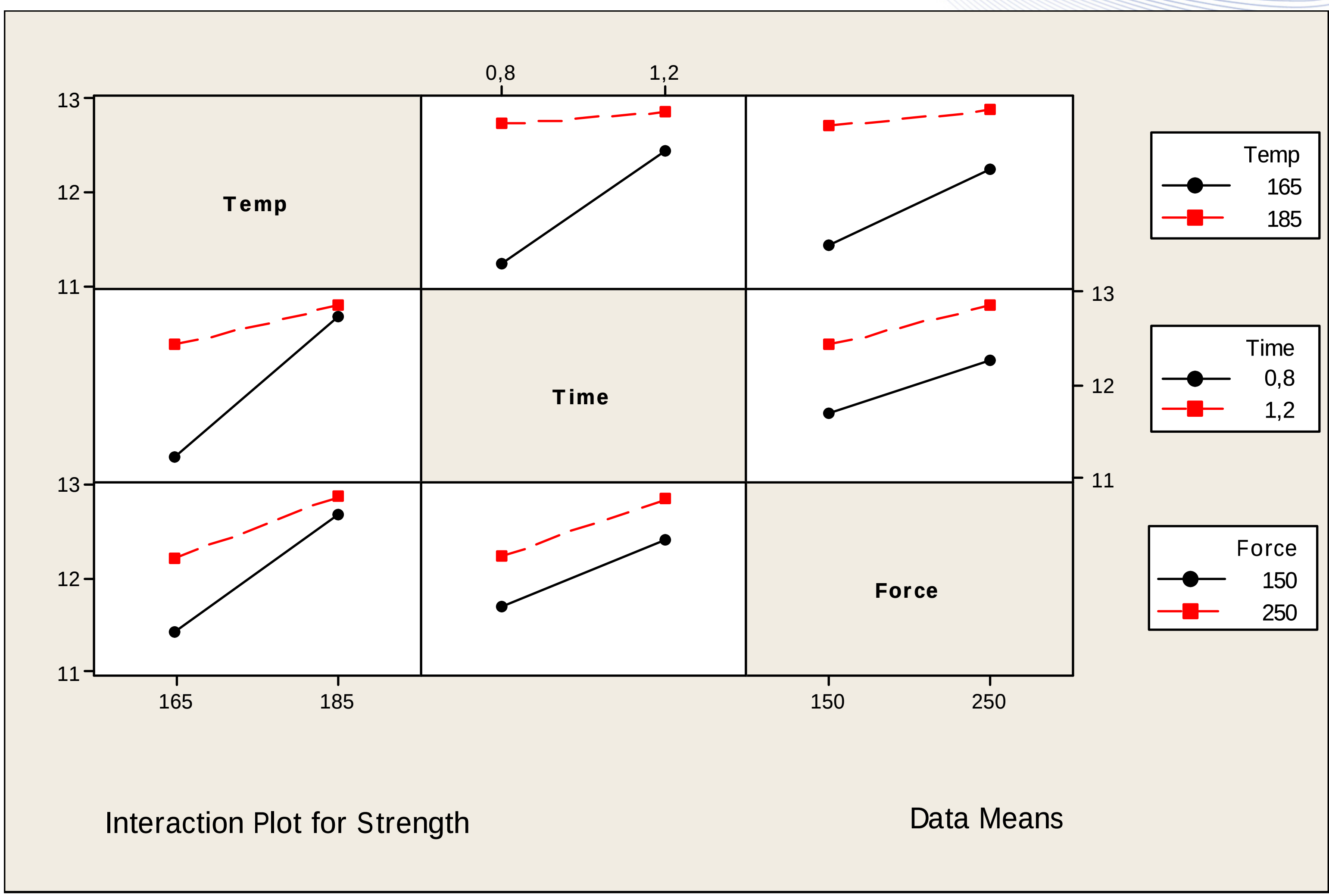
2^3 factorial design with the experimental data after only two replications: Effects with an error $\alpha < 5\%$ are significant whose calculated F-values are greater than the tabulated quantile 5.32 for the related degrees of freedom.

Y = b0 + b1*x1 + b2*x2 + b3*x3 + b12*x1*x2 + b13*x1*x3 + b23*x2*x3 + b123*x1*x2*x3											2432,29	**SS (raw)**
											2423,10	**SS (b0)**
SS (rest)		**Temp.**	**Time**	**Force**					**Seal Strength**		9,19	**SS (total)**
syst. Nr.	**x0**	**x1**	**x2**	**x3**	**x1*x2**	**x1*x3**	**x2*x3**	**x1x2x3**	**y1**	**y2**	**(y1+y2)/2**	SS (error)
1	1	-1	-1	-1	1	1	1	-1	10,6	11,1	10,85	0,13
2	1	1	-1	-1	-1	-1	1	1	12,8	12,3	12,55	0,12
3	1	-1	1	-1	-1	1	-1	1	12,4	11,6	12,00	0,32
4	1	1	1	-1	1	-1	-1	-1	13,0	12,7	12,85	0,04
5	1	-1	-1	1	1	-1	-1	1	11,8	11,4	11,60	0,08
6	1	1	-1	1	-1	1	-1	-1	13,2	12,6	12,90	0,18
7	1	-1	1	1	-1	-1	1	-1	13,2	12,5	12,85	0,25
8	1	1	1	1	1	1	1	1	13,0	12,7	12,85	0,05
	8	8	8	8	8	8	8	8				
	98,45	3,85	2,65	1,95	-2,15	-1,25	-0,25	-0,45	**F(0,95;1;8)=5,32**			
Reg.-Coeff.	12,31	**0,48**	**0,33**	**0,24**	**-0,27**	-0,16	-0,03	-0,06				
SS(b...)	2423,10	3,71	1,76	0,95	1,16	0,39	0,02	0,05			1,17	**1,17**
Fo=SS/s^2	16639	**25,45**	**12,06**	**6,53**	**7,94**	2,68	0,11	0,35			**s^2 =**	**0,15**

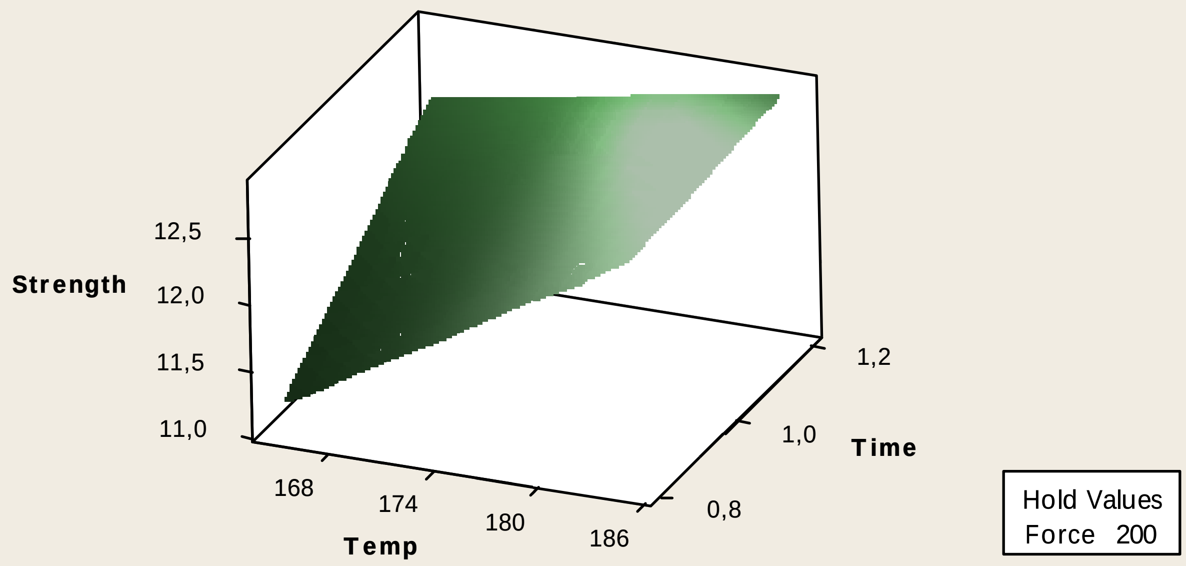
In the example all main effects positive. Largest influence has temperature.
Remarkable: significant negative interaction between Temp. and Time.
All other interaction effects are only a result of random.


Temp
Time
Force
Mean
12,75
12,50
12,25
12,00
165
185
0,8
1,2
150
250
Main Effects Plot for Strength
Data Means

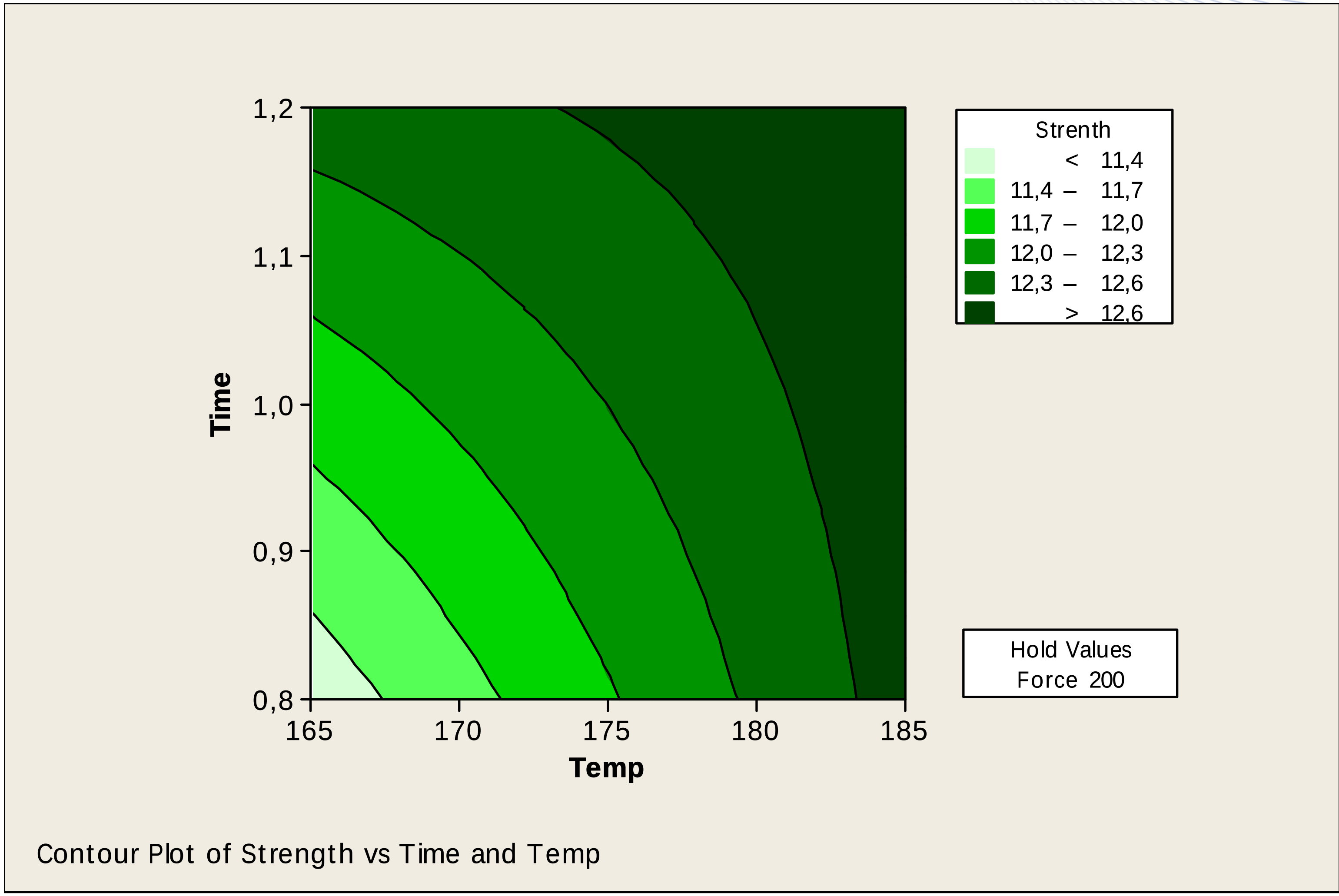

2,306
A
B
AB
C
AC
ABC
BC
Term
0
1
2
3
4
5
Standardized Effect
Factor Name
A Temp
B Time
C Force
The Software MINITAB®
calls the factors
x_1 = A
x_1 = B
x_1 = C
Pareto Chart of the Standardized Effects
(response is Strength, Alpha = 0,05)



0,8
1,2
13
12
11
Temp
Time
Force
165
185
150
250
Temp
165
185
Time
0,8
1,2
Force
150
250
Interaction Plot for Strength
Data Means



Surface Plot of Strength vs Time and Temp


1,2
1,1
1,0
0,9
0,8
Time
165
170
175
180
185
Temp
Strenth
< 11,4
11,4 – 11,7
11,7 – 12,0
12,0 – 12,3
12,3 – 12,6
> 12,6
Hold Values
Force 200
Contour Plot of Strength vs Time and Temp

Guidelines for Designing Experiments

The following steps can give a roadmap (in some dependence on [1]):

1. Recognition and statement of the problem
2. Choice of output quantities (called “response variables”)
3. Selection of input quantities (called “factors”)
4. Determination of the experimental design
5. Specification of the factor levels
6. Performance of the experiments
7. Statistical analysis of the data
8. Conclusions and recommendations

Students should keep in mind the following advices:

a) Use at first your non-statistical knowledge of the problem.

b) Keep the design as simple as possible.

c) Recognise the difference between statistical and practical significance.

d) Don't forget the restrictions in whose frame the results are valid.

Last, but not least,

experimental data must never be accepted uncritical!

Conclusion and Recommendation

DoE

- ❖ greatly increases the efficiency of experiments,
- ❖ is inevitably in the interpretation of effects of interacting factors.

To achieve real skills in utilisation computer software generally three steps are recommended:

- ✓ Set very simple examples to calculate them manually in application of the mathematical basics.
- ✓ Write simple programs yourself (e. g. using spreadsheet) and compare the results with the previous step.
- ✓ Use sophisticated software at first to compute the same simple examples from the previous step and compare the results.