An Introduction to Design of Experiments (DOE) – Classical and more recent applications

Ing-Marie Olsson, PhD

Application specialist, Application manager

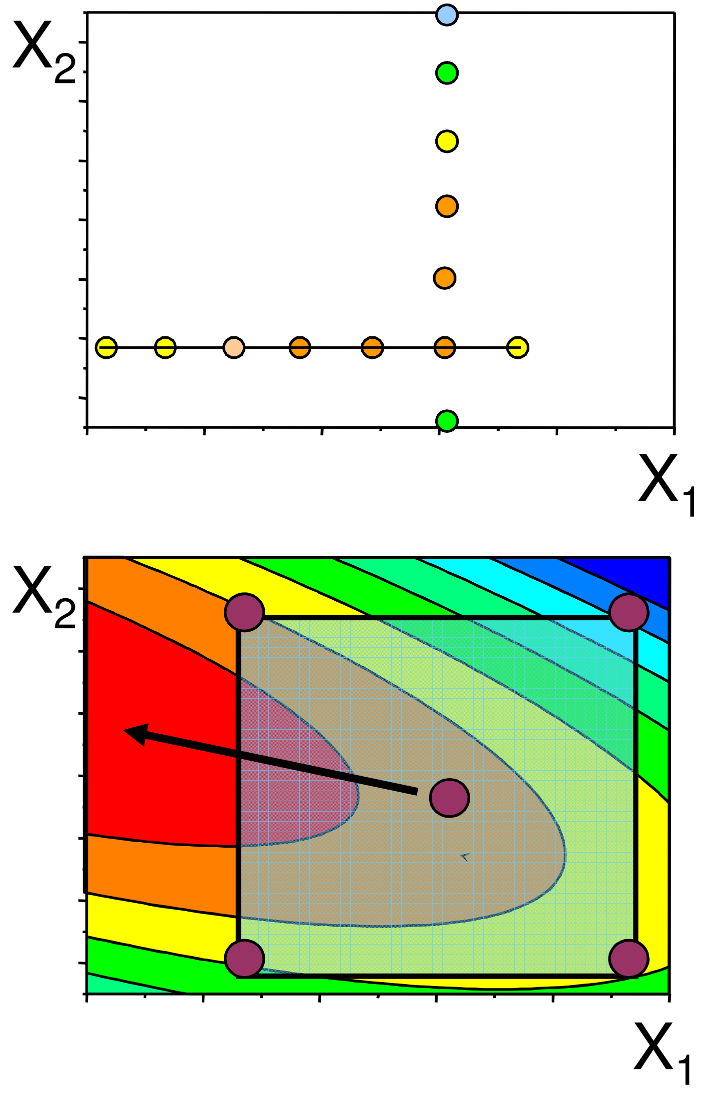
Umetrics AB

Contents

- Introduction to Design of Experiments
  - Why, when, what and how?
- Problem formulation
- Stepwise DoE
- Design types
  - Screening
  - Optimization
  - Robustness testing
- Mixture Design
- D-optimal design
- **LUNCH**
- Applications of DoE
  - Optimization of enzyme activity assay using RED-MUPs
  - Selection of diverse sub set using D-optimal Onion designs
- News in MODDE 9 and M-link

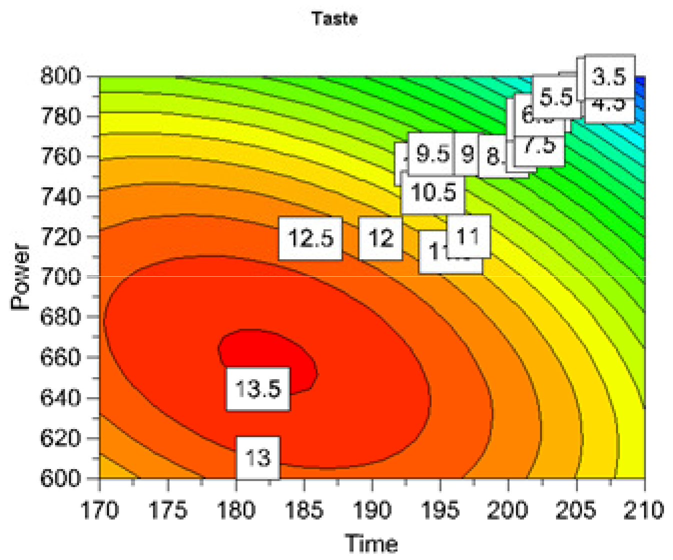
The "intuitive" (COST) approach to experimental work

- Changing one separate factor at a time (COST) does not lead to the real optimum, and gives different implications with different starting points
- Leads to many experiments and little information
- No quantification of interactions !!!



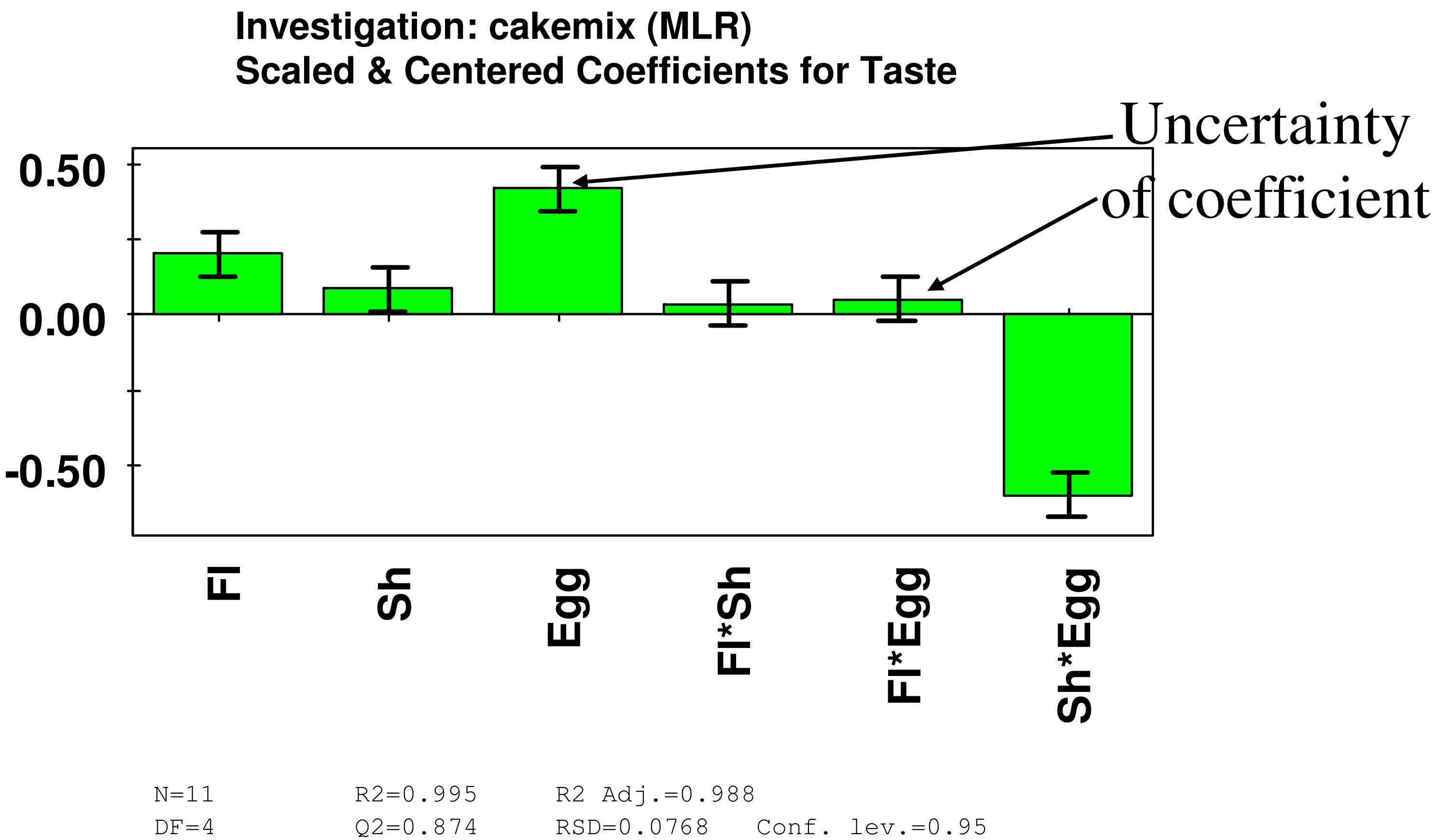
Design of Experiments- the Key to knowledge

- Knowledge is power!
- DoE allows
  - Description of system as mathematical model
    - In a controlled region
  - Separate true effects from noise
  - Separately estimate noise
  - Estimate interactions
- Result: Detailed map of investigated system



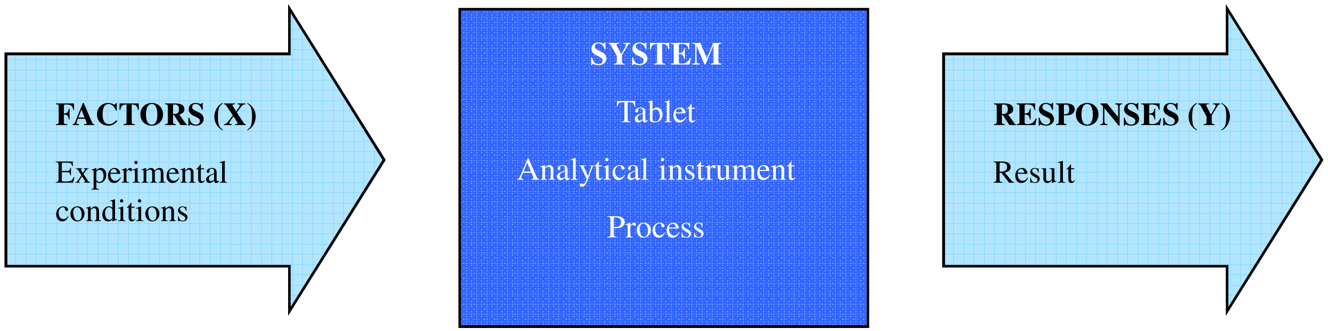
Design of Experiments: Estimate real effects and noise

- Real effects are estimated by the coefficients, and the noise is contained in the confidence intervals



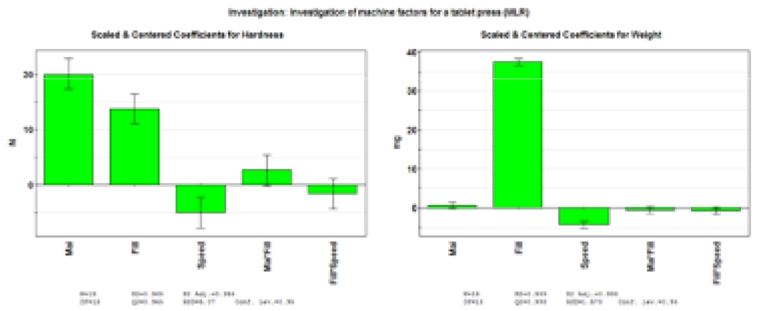
Design of Experiments- Use of knowledge?

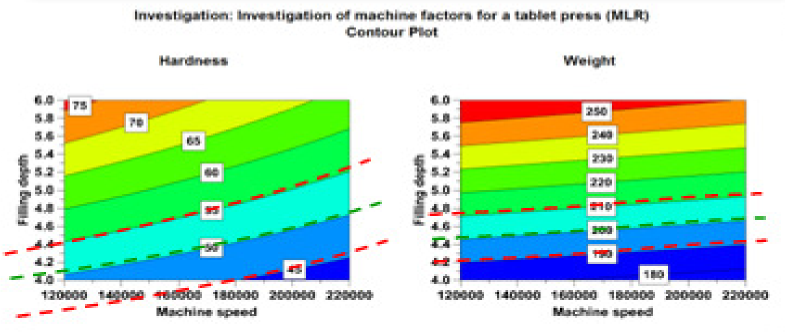
- Root cause analysis for issues in processes
- Optimization of existing products and processes
  - Product properties, quality, efficiency, robustness
- Development of new products and processes
- Minimization of production costs and polluting outlets

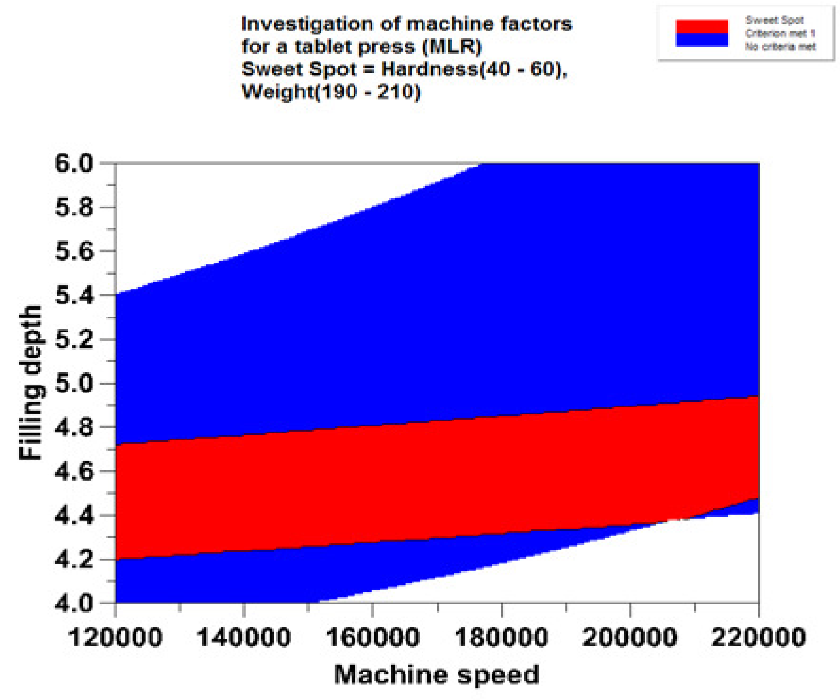


Use of DOE for Design Space

- Key technology in QbD discussions
- Tabletting process optimised using DOE
  - **Factors:** Pre-pressure, Main-pressure, Filling depth, Machine Speed
  - **Responses:** Tablet weight, Tablet hardness







UMETRICS
AN MKS COMPANY

Definition design space according to ICH Q8(R1) Pharmaceutical Development Revision 1

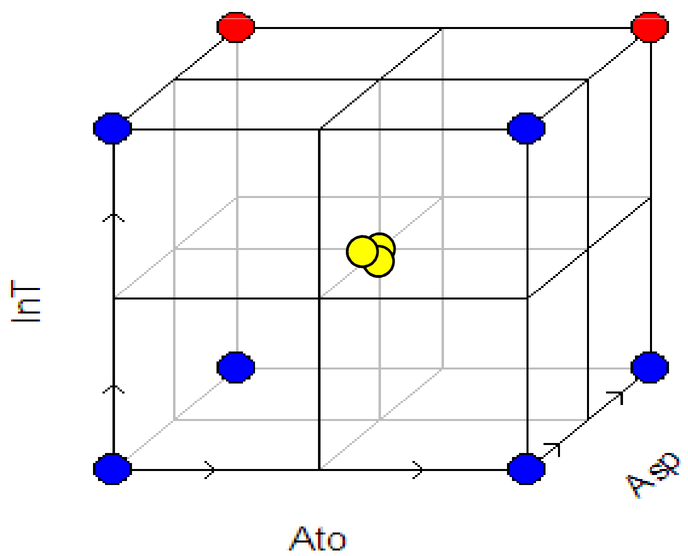
A design space can be defined in terms of ranges of input variables or parameters ***(DOE),*** or through more complex mathematical relationships. It is possible to define a design space as a time dependent function (e.g., temperature and pressure cycle of a lyophilisation cycle), or as a combination of variables such as principal components of a multivariate model ***(MVA).*** factors can also be included if the design space is intended to span multiple operational scales. Analysis of historical data can provide the basis for establishing a design space ***(MVA).*** Regardless of how a design space is developed, it is expected that operation within the design space will result in a product meeting the defined quality attributes*.*

*Blue text added by Umetrics

What is Design of Experiments?

When you vary all investigated factors simultaneously,

-according to a well designed plan

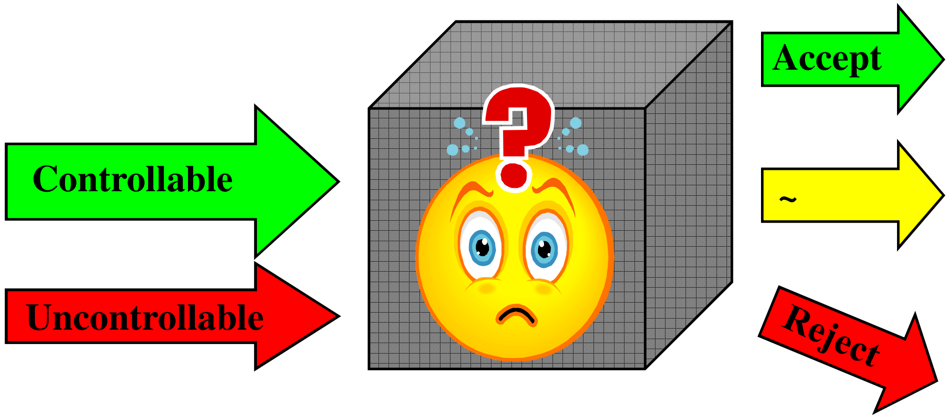


How Design of Experiments?

- Problem formulation
  - Properly performed is the key to success
- Prepare the design
- Evaluate the data
- Perform modeling
- Visualize and interpret model
- Reporting and basis for decision making

Problem formulation

- Identify question
  - Define goal
    - Preferably quantitative
  - How much do we know?
    - Black box? Grayish box?

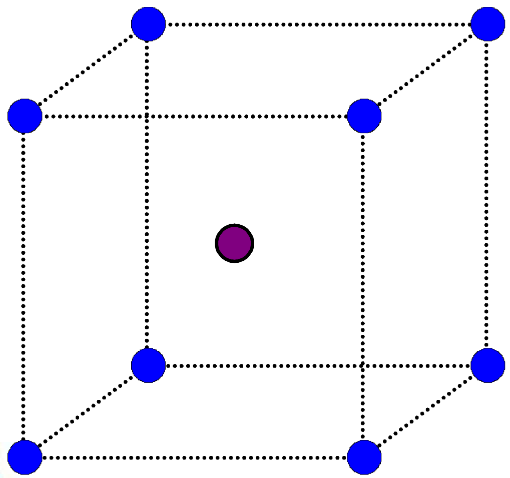


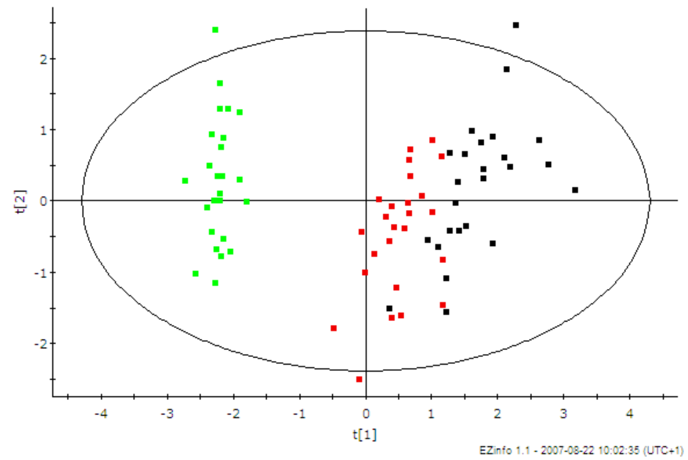
- Possibility to meet goal
  - Is the goal realistic?
  - Resources

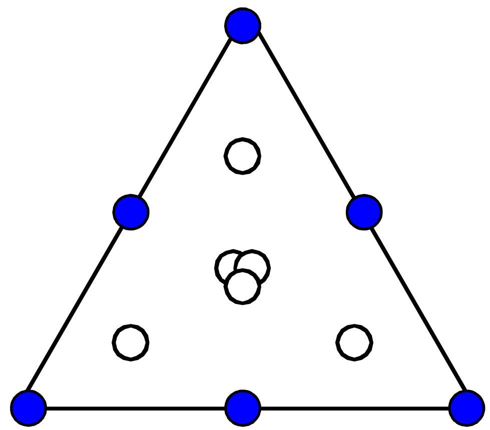
- How reach the goal?
  - Influential parameters
  - Controllable/ uncontrollable?

DOEs for different types of factors

- Continuous: Temp 50-85ºC, 5-10% MeOh
- Discrete: Catalyst A or B, 0 or 1
- Non-continuous quantitative: Spectral data (D-optimal)
- Formulation: Ratios of ingredients (Mixture)





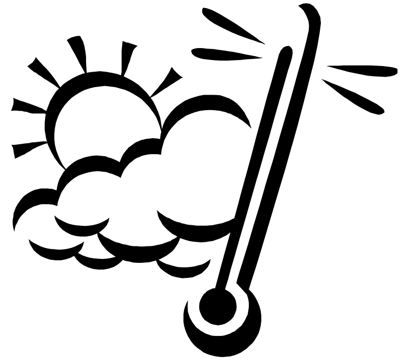


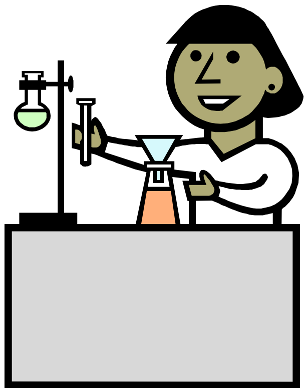
Uncontrollable factors

- Some things can't be controlled….

- Uncontrolled quantitative: Atmospheric pressure, outdoor temp…
- Uncontrolled qualitative: Analysis instrument, reaction vessel, persons…

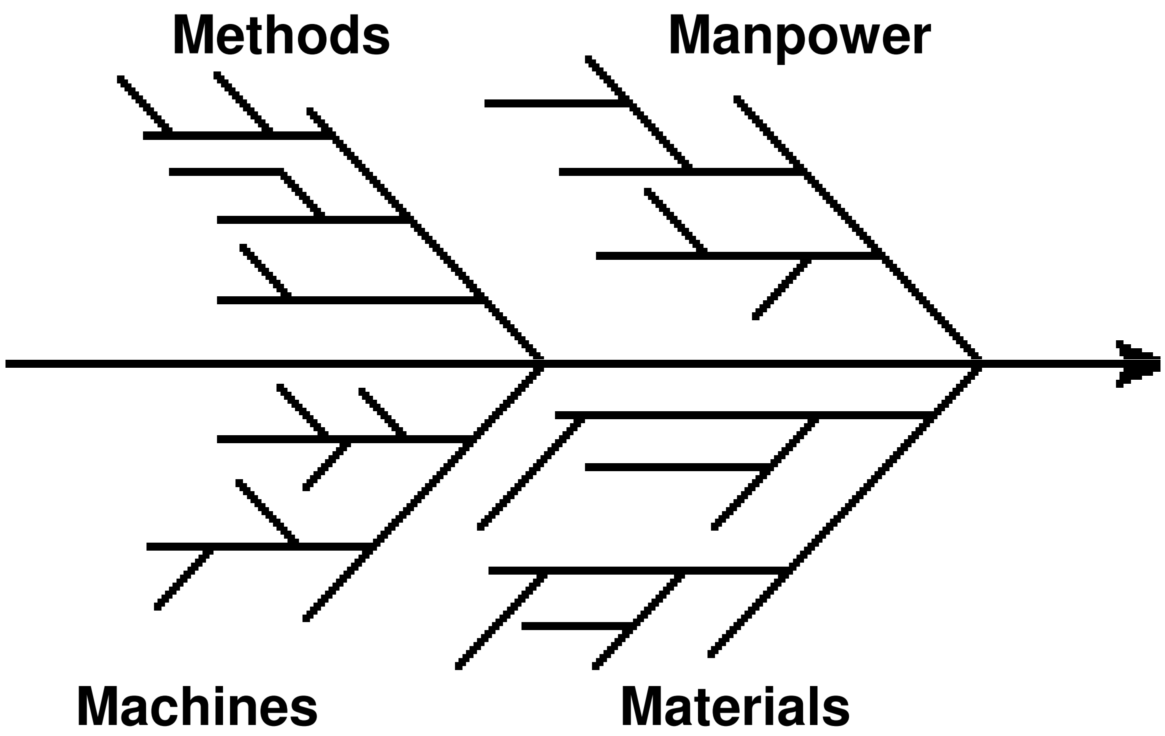
- Uncontrolled but measurable





Ishikawa diagram

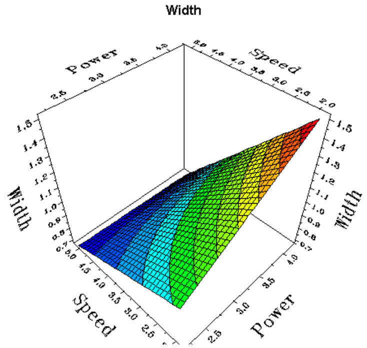
- The Ishikawa, or fishbone, system diagram is a very helpful method to overview all factors
- Reduces the risk of missing a critical factor
- The four M´s
- Practical maximum depth 4-5 levels

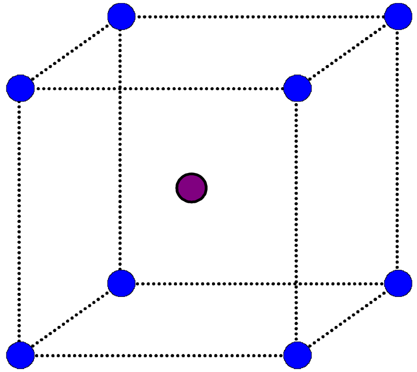


Design of experiments- Three Objectives

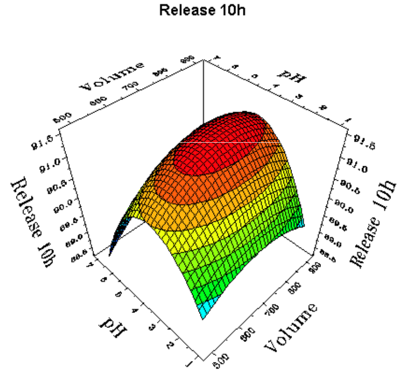
Different objectives require different types of Designs

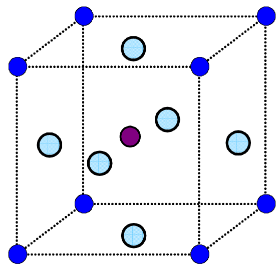
- Screening
  - Many factors, few experiments
  - Find important factors and relevant factor ranges



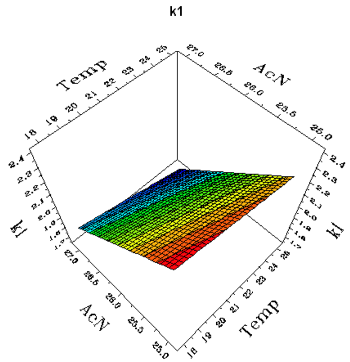


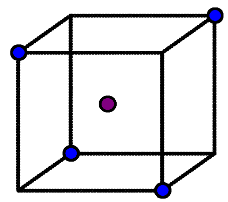
- Optimization
  - Few factors, many experiments
  - Find detailed information about investigated system
  - Find optima





- Robustness testing
  - Few factors, few experiments
  - Is the system robust within “normal” system variation?

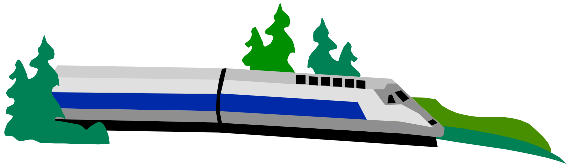


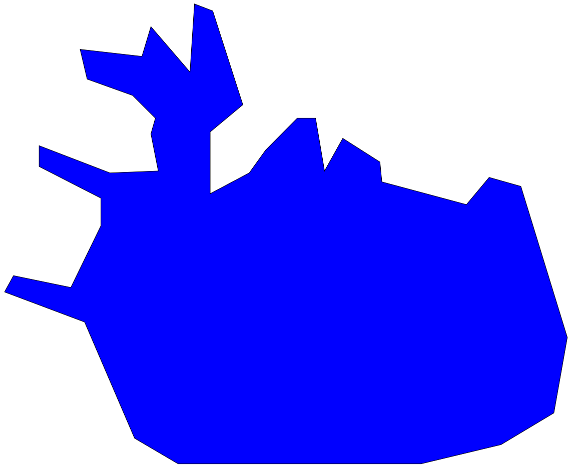


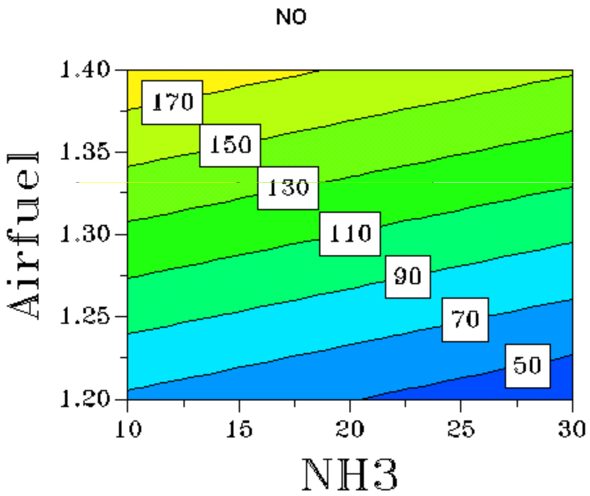
UMETRICS
AN MKS COMPANY

The model concept

- Models are not reality, but approximate representations of some aspects of reality







Toy train **Map of Iceland** **Response contour plot**

GENERAL EXAMPLE

UMETRICS
AN MKS COMPANY

Making DOE understandable

- How do you explain DOE in a fun way?

- Mission Popcorn; carried out during recent summer break

- Root cause (at Legoland, Billund) was well tasting cotton candy but distasteful popcorn (burnt, unpleasent odor)

Problem formulation: Selection of Objective

- *Practical objectives*:
  - Find out how to make good popcorn!
  - To explain to kids what DOE means using an everyday problem (i.e., how to get good popcorn from the microwave) as illustration.
  - Understand what dad is working with

- *Experimental objective*: Optimization

Problem formulation: Definition of factors

- The dataset contains two factors, Time and Power, both adjustable on a continuous scale:
  - Time (seconds), low level 170 seconds, high level 210 seconds.
  - Power (watt), low level 600 watts, high level 800 watts.

Factors

	Name	Abbr.	Units	Type	Use	Settings	Transform	Prec.	MLR Scale	PLS Scale
1	Time	Tim	seconds	Quantitative	Controlled	170 to 210	None	Free	Orthogonal	Unit Variance
2	Power	Pow	W	Quantitative	Controlled	600 to 800	None	Free	Orthogonal	Unit Variance

UMETRICS AN MKS COMPANY

Specification of response(s)

- The dataset also contains two responses, Kernels and Taste.
  - *Kernels*, this is simply the number of unpopped kernels.
  - *Taste*, each person expressed his liking on a five-level scale (1=bad taste, …., 5=optimal taste). The response value is the sum across three persons (we could not use the average as this was too complicated for the little brother).

Responses

	Name	Abbr.	Units	Transform	MLR Scale	PLS Scale	Type
1	Kernels	Ker		None	None	Unit Variance	Regular
2	Taste	Tas		None	None	Unit Variance	Regular

Generation of experimental design

- The design used was a CCF optimization design, by default encoding 8+3 experiments in MODDE 8. One centerpoint was dropped since we bought a ten-pack of microwave popcorn.

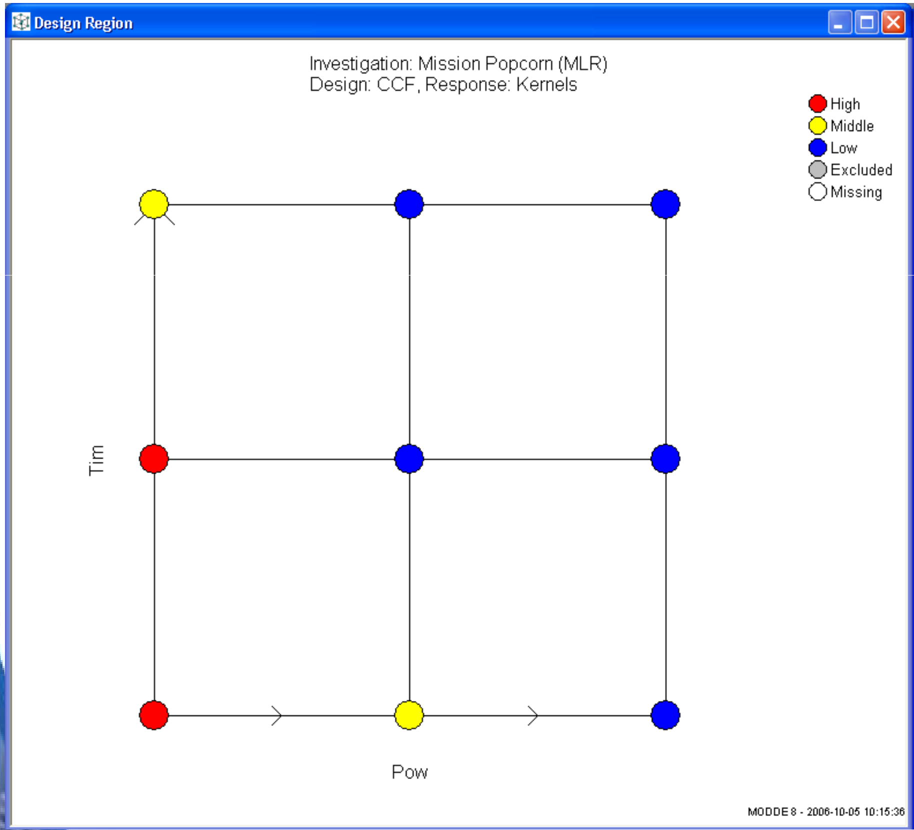
Worksheet

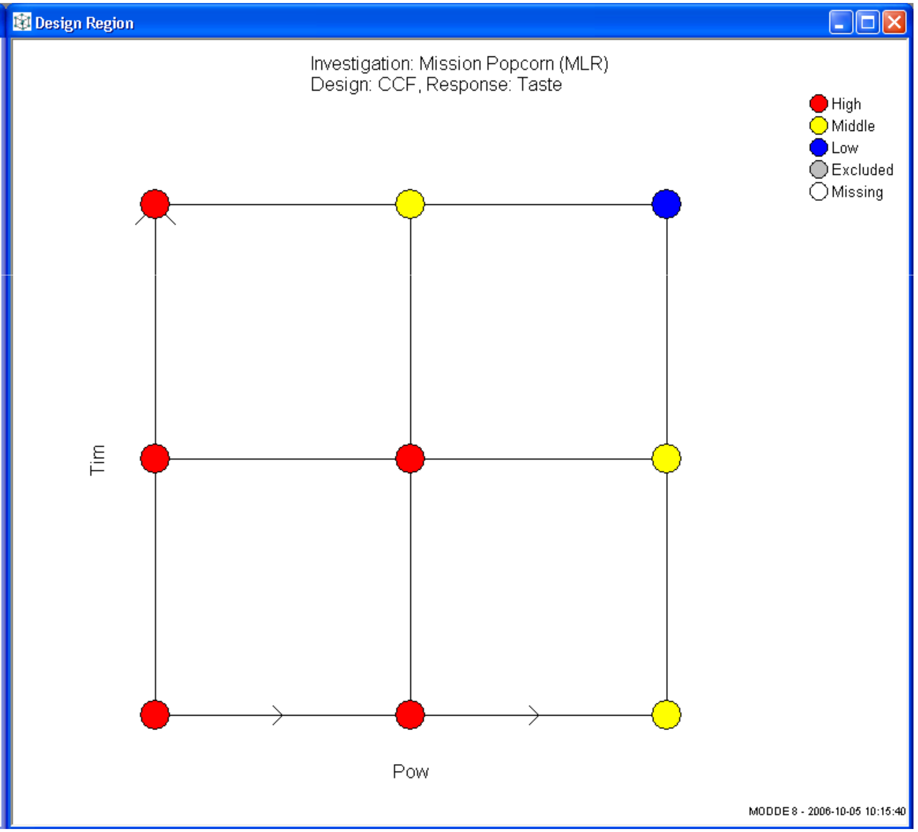
	1	2	3	4	5	6	7	8
1	Exp No	Exp Name	Run Order	Incl/Excl	Time	Power	Kernels	Taste
2	1	N1	9	Incl	170	600	142	12
3	2	N2	2	Incl	210	600	53	11
4	3	N3	10	Incl	170	800	11	10
5	4	N4	3	Incl	210	800	0	3
6	5	N5	1	Incl	170	700	48	12
7	6	N6	6	Incl	210	700	0	8
8	7	N7	8	Incl	190	600	121	12
9	8	N8	4	Incl	190	800	1	7
10	9	N9	5	Incl	190	700	22	14
11	10	N10	7	Incl	190	700	18	13

UMETRICS
AN MKS COMPANY

Visualize geometry of design

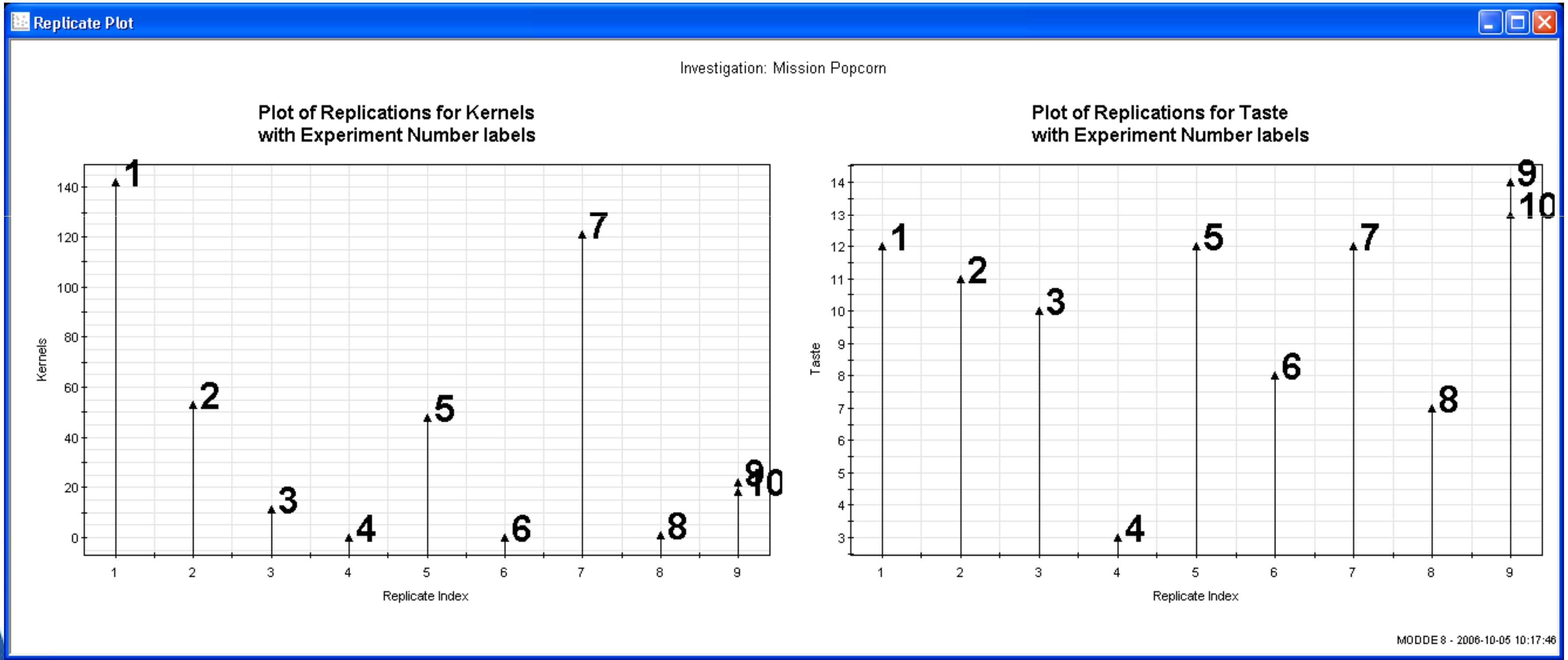
- Colour coding provides an easy-to-understand overview.





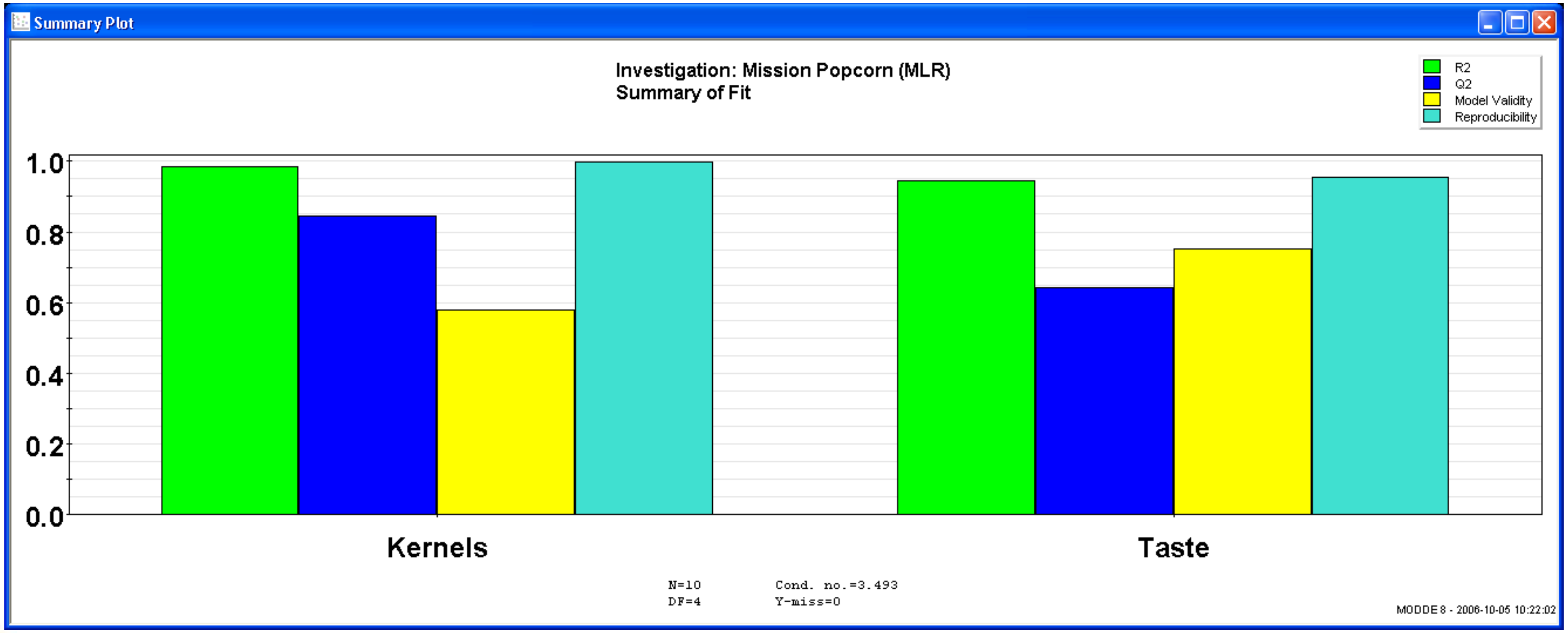
Data analysis: Evaluate raw data

- The replicate plots indicate small variability among the replicates.



Data analysis: Compute regression model

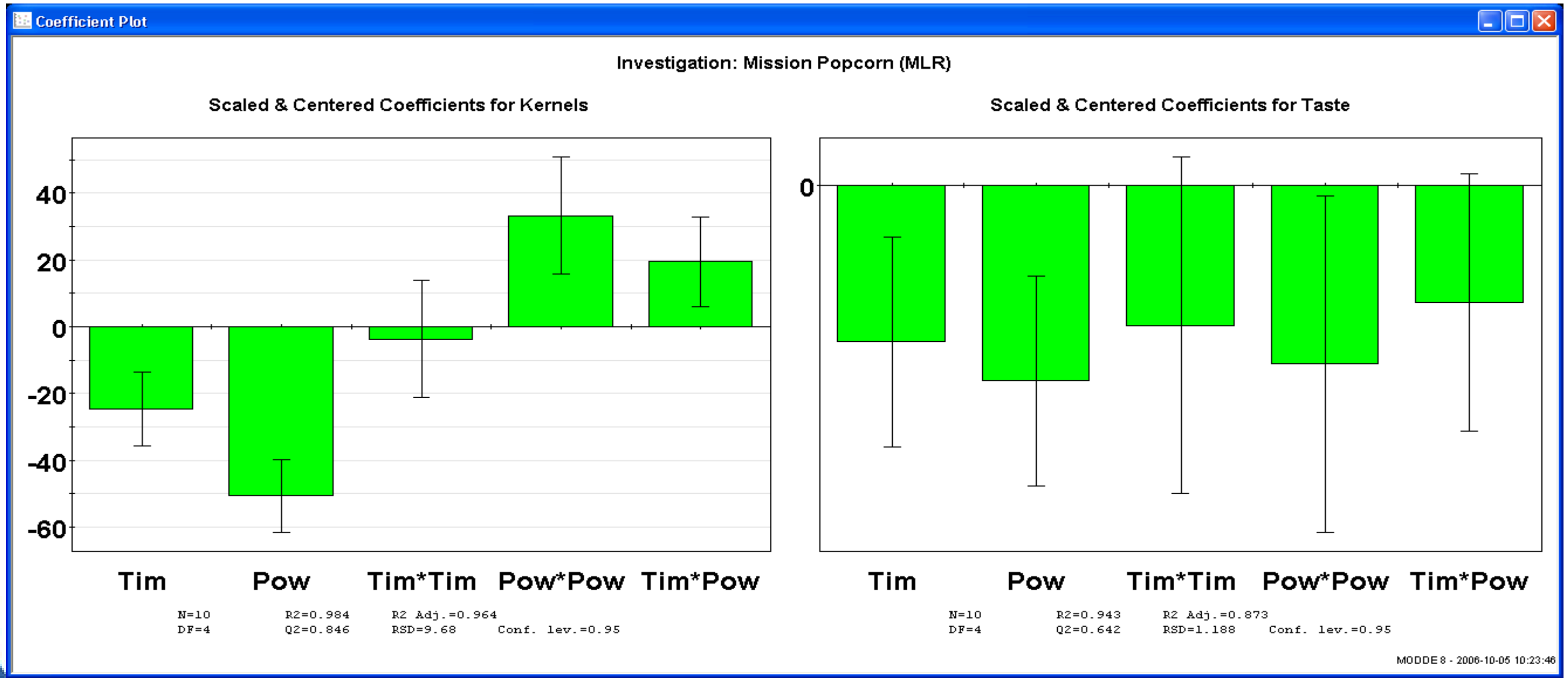
- Model summary plot- All columns high = good!
  - R2: How well does the model fit to the data? Amount of explained variation (range 0-1)
  - Q2: How well can the model predict? (range -∞ - 1)
  - Model validity: Based on statistical test of Lack of Fit (ANOVA), above 0,25 for valid model
  - Reproducibility: Reflects replicate error



UMETRICS
AN MKS COMPANY

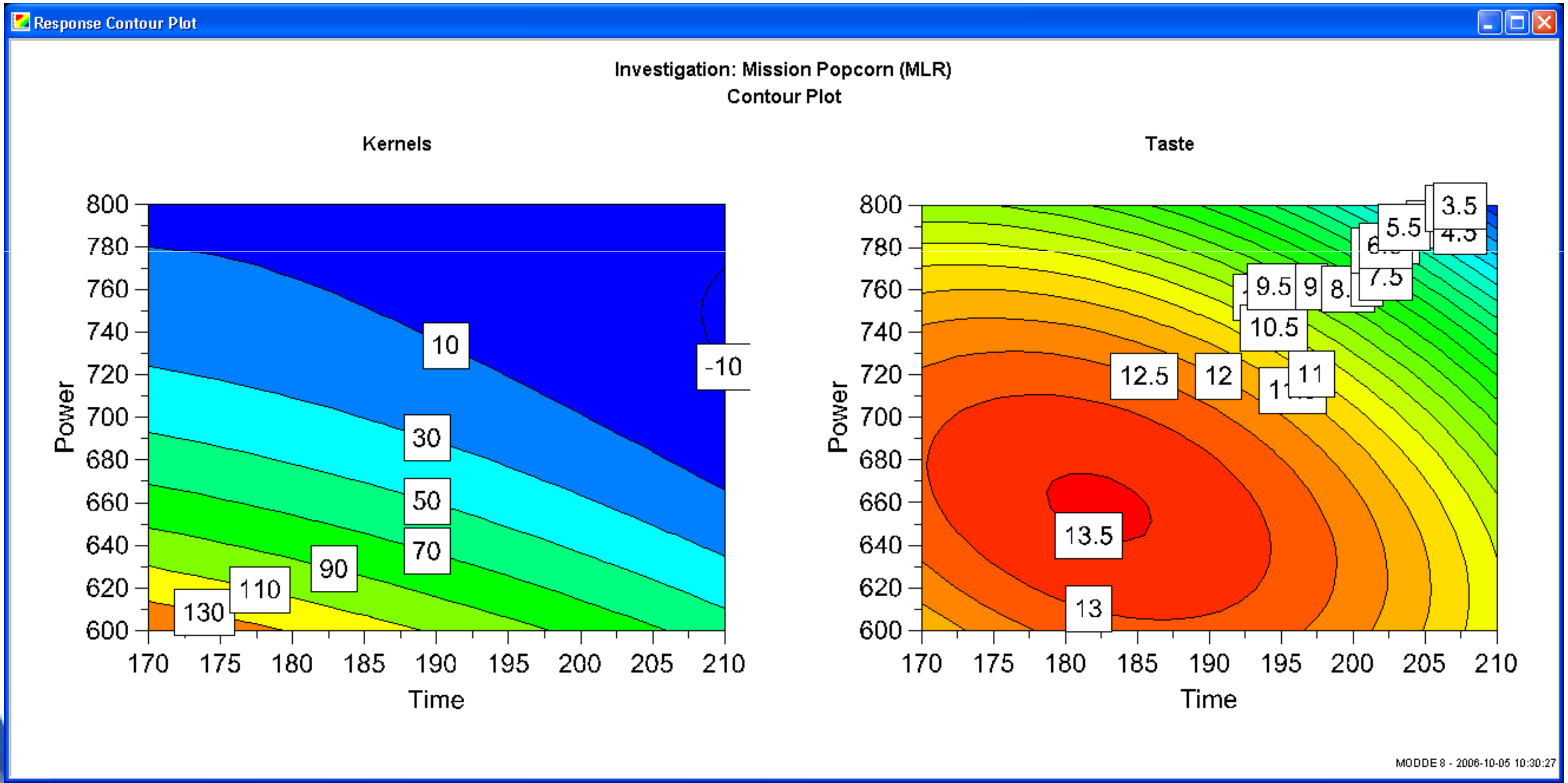
Data analysis: Interpretation of models

- Coefficient plots with confidence interval, size and reliability of a factors (model terms) impact
- Coefficients show:
  - To minimize the number of Kernels both factors should be set high.
  - Time and Power seem to have a similar impact on Taste.
  - Adjusting both factors on a lower value corresponds to increasing the Taste.



Use of model (Contour plots)

- Time ≈ 182 seconds and Power ≈ 657 watts give highest taste. This point does, however, not correspond to the lowest number of Kernels.



Use of model: Finding the optimum

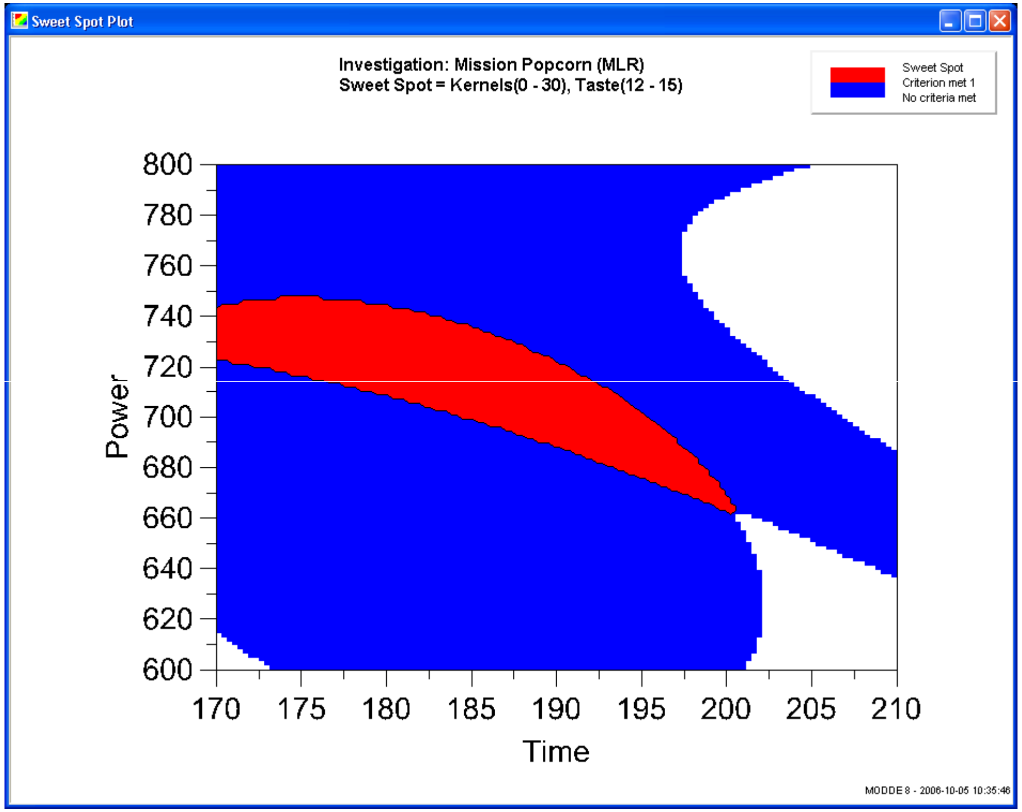
- To arrive at a ‘final’ point to use, we sat down and together specified what we wanted. We agreed that a Taste of 12 or higher would be fully acceptable. Having 10 kernels per bowl was also deemed OK (hence a total of 30).

- Thus, we set up the following response desirabilities:

	Response	Criteria	Weight	Min	Target	Max
1	Kernels	Minimize	1		0	30
2	Taste	Maximize	1	12	15	

Use of model: Graphical display of optimal region

- Software optimizer was run
- Convergence was instantanous
- Sweet Spot plot shows there is a region of optimum inside the searched space (the "Design Space")



Mission Popcorn: End result

- Based on our joint efforts we were able to find out a suitable combination of Time (= 190 secs) and Power (= 700 watts).
- We are currently using this combination with great satisfaction. It produces well tasting popcorn without undesirable side effects such as burning and unpleasant odor. One resulting bag is seen to the right.
- The final result (apart from the popcorn) for the two end users (i.e., the two boys) was better understanding for dad's work plus having a lot of fun together with their father.

Benefits of DOE

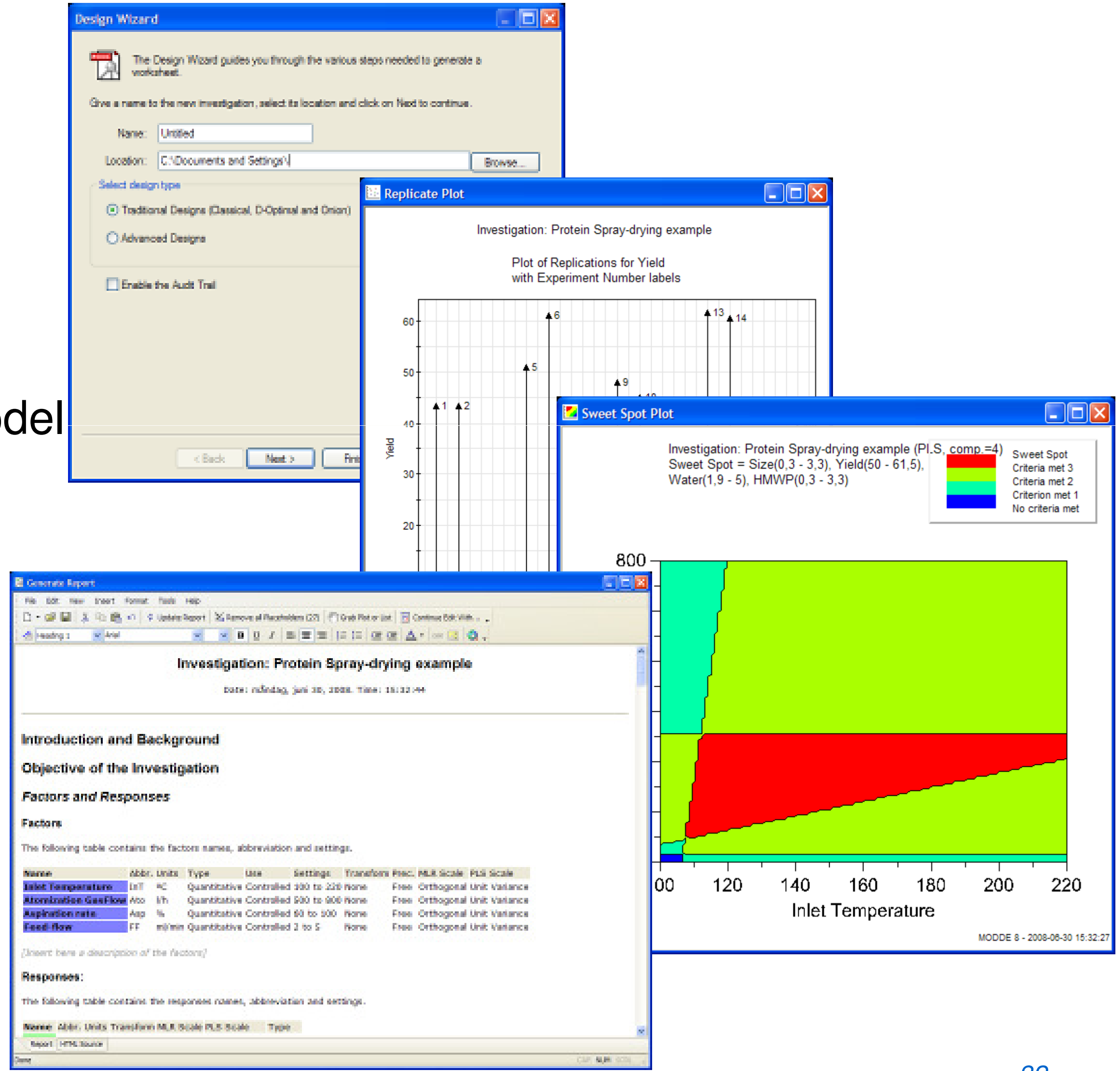
- Organized approach which connects experiments in a rational manner
- More useful information is obtained (the influence of all factors together)
- More precise information is acquired in fewer experiments
- Results are evaluated in the light of variability
- Support for decision-making: Map of the system (response contour plot)

UMETRICS AN MKS COMPANY

DoE in MODDE

MODDE 8 combines:

- User friendliness
  - Design wizard
  - Analysis advisor
- Full battery of functionality
  - For raw data diagnosis, model diagnosis, use of model
  - Report generator
- Dedicated solutions
  - Custom made for specific applications

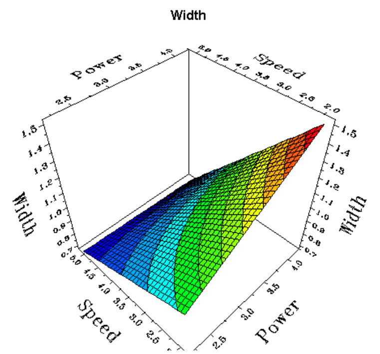


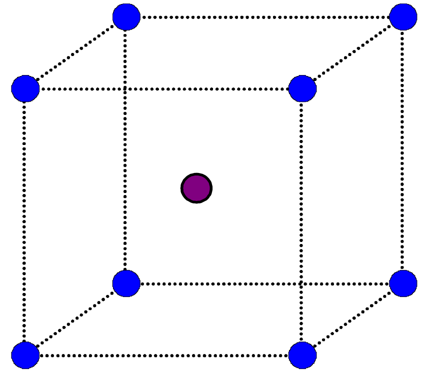
UMETRICS
AN MKS COMPANY

Design of experiments- Three Objectives

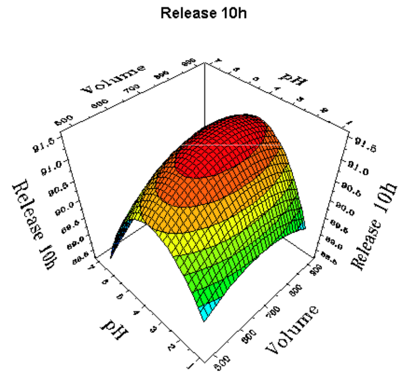
Different objectives require different types of Designs

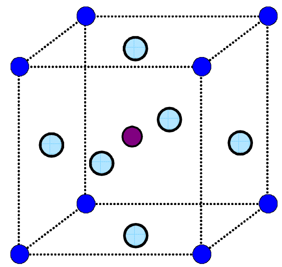
- Screening
  - Many factors, few experiments
  - Find important factors and relevant factor ranges



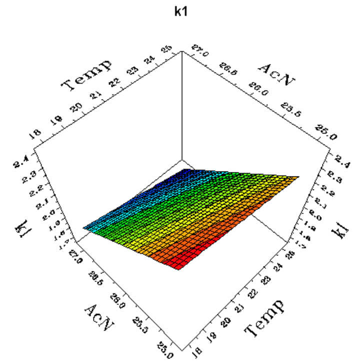


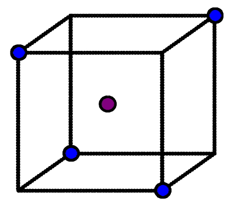
- Optimization
  - Few factors, many experiments
  - Find detailed information about investigated system
  - Find optima





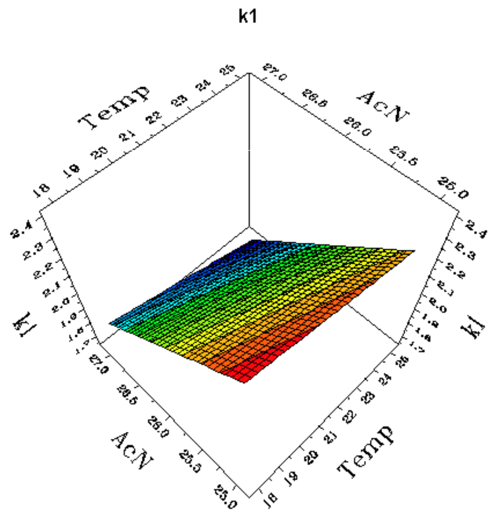
- Robustness testing
  - Few factors, few experiments
  - Is the system robust within “normal” system variation?





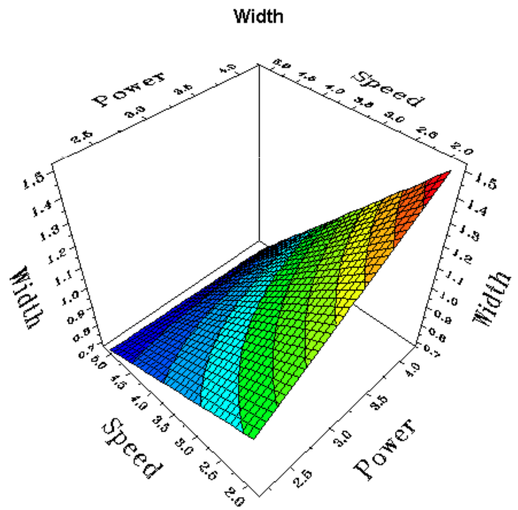
Objective and Model complexity

- We distinguish between three main types of polynomial models
  - linear: $\mathbf{y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \varepsilon}$
  - interaction: $\mathbf{y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \ldots + \varepsilon}$
  - quadratic: $\mathbf{y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \ldots + \varepsilon}$



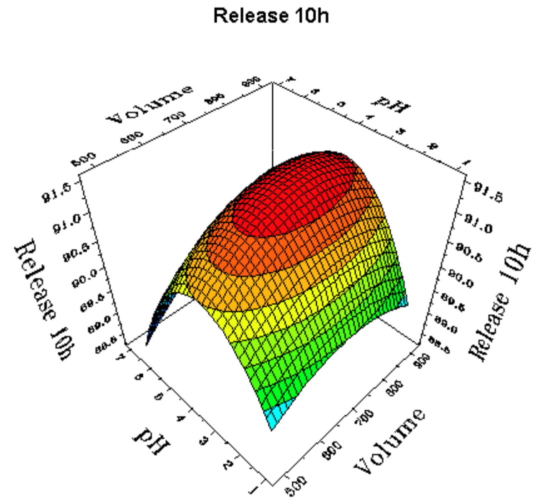
Linear:

Screening & Rob. Test.



Interaction:

Screening



Quadratic:

Optimization

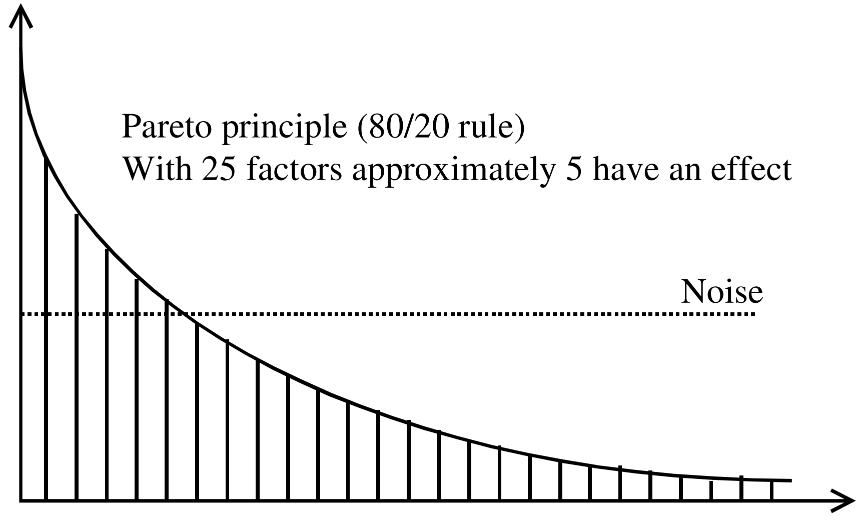
UMETRICS
AN MKS COMPANY

OBJECTIVE SCREENING

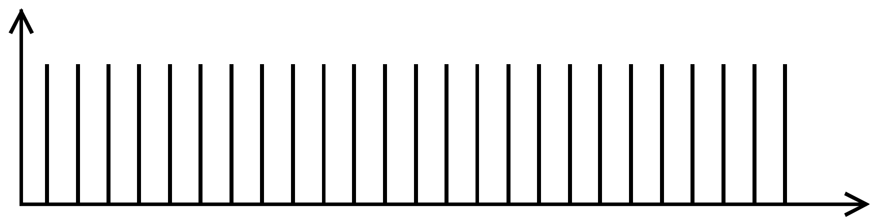
Screening - Introduction

- Useful when one wants to find out a little about many factors

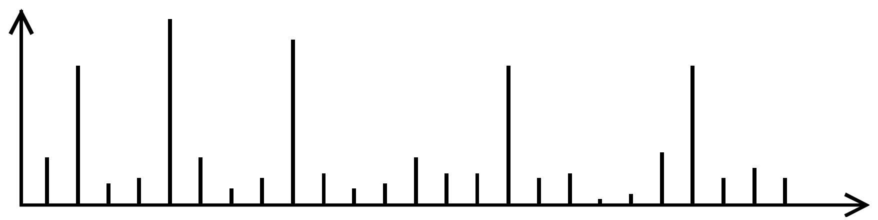
- Goal: To uncover the important factors and their appropriate ranges. Is factor/response relationship linear or non-linear?



- Results before ….



… and after screening



General Example 1: Screening

- Protein spray drying
- Background
  - Study made on a model protein at AstraZeneca AB
- Objective: determine which process parameters influence the quality of the spray-dried product
  - to produce *particles of controlled size*
  - Reduce water content
  - Increase yield, avoid denaturation

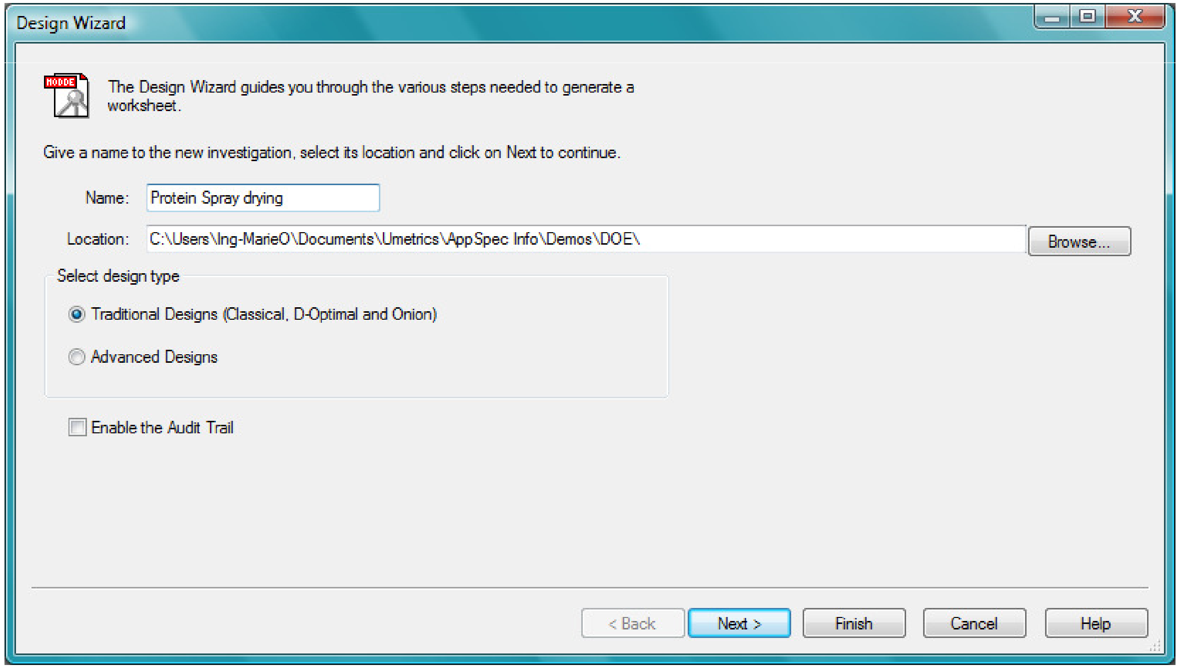
Reference: Cronholm, M., The Effect of Process Variables on a Spray-dried Protein Intended for Inhalation, Undergraduate Research Study, Department of Pharmaceutics, Uppsala University, Uppsala, Sweden, 1998.

DoE example: Protein spray-drying

Background

- Study made on a model protein at AstraZeneca AB

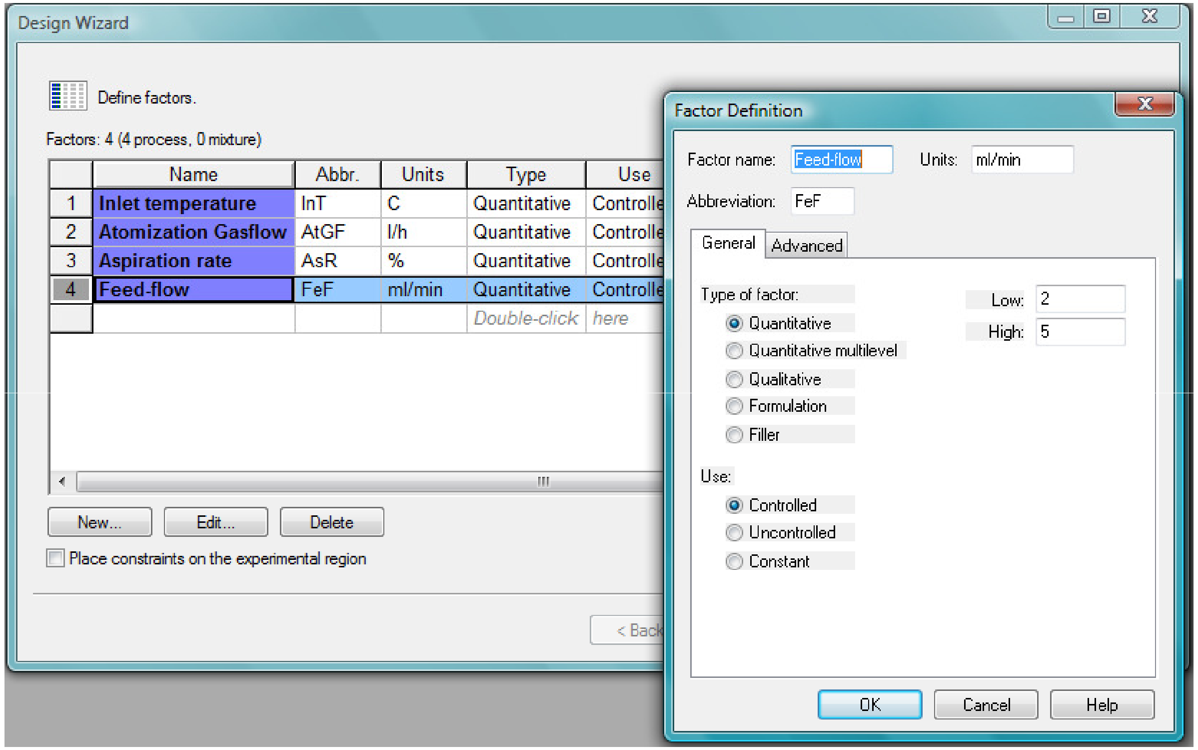
- Objective: determine which process parameters influence the quality of the spray-dried product
  - to produce *particles of controlled size*
  - Reduce water content
  - Increase yield
  - Avoid denaturation



Reference: Cronholm, M., The Effect of Process Variables on a Spray-dried Protein Intended for Inhalation, Undergraduate Research Study, Department of Pharmaceutics, Uppsala University, Uppsala, Sweden, 1998.

Define investigated process factors

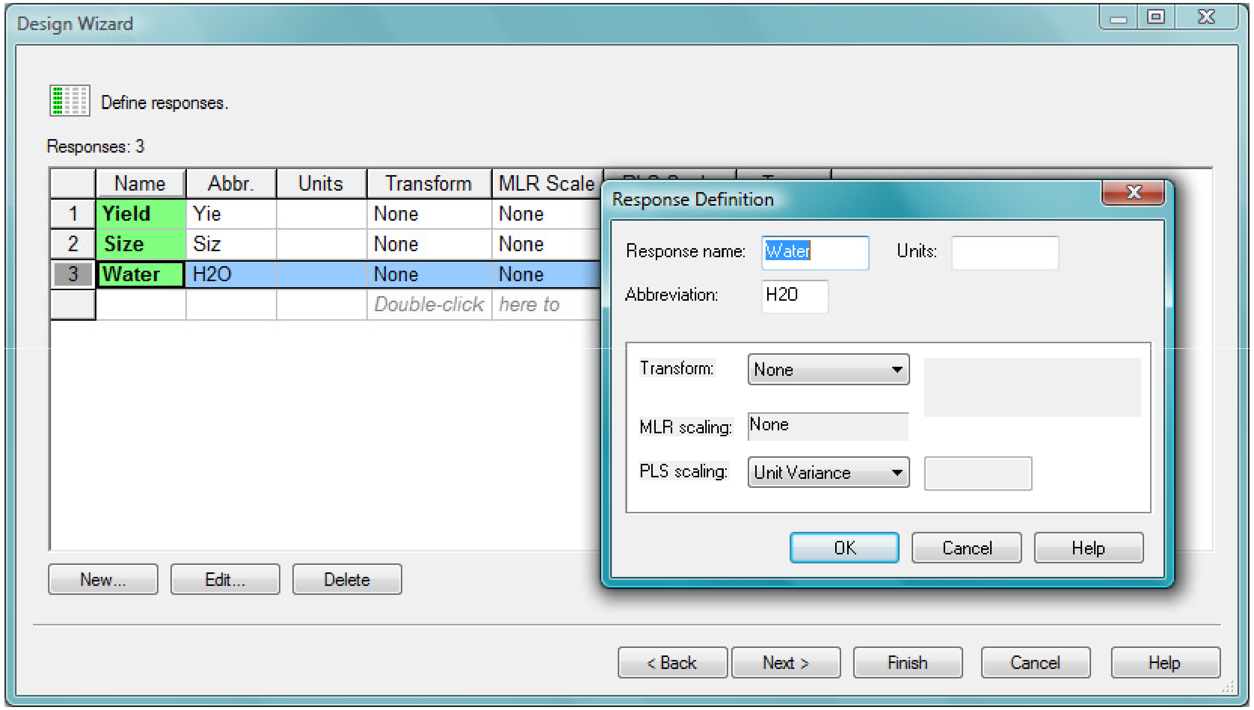
- Inlet Temperature – temperature of drying air at the inlet of the equipment
  - 100ºC to 220ºC denaturation expected at high temperature
- Atomization gas flow
  - 500 l/h to 800 l/h maximum flow with this spray-dryer
- Aspiration rate
  - 60 to 100%
- Feed-flow – indicates the material flow through the equipment
  - 2 to 5ml/min



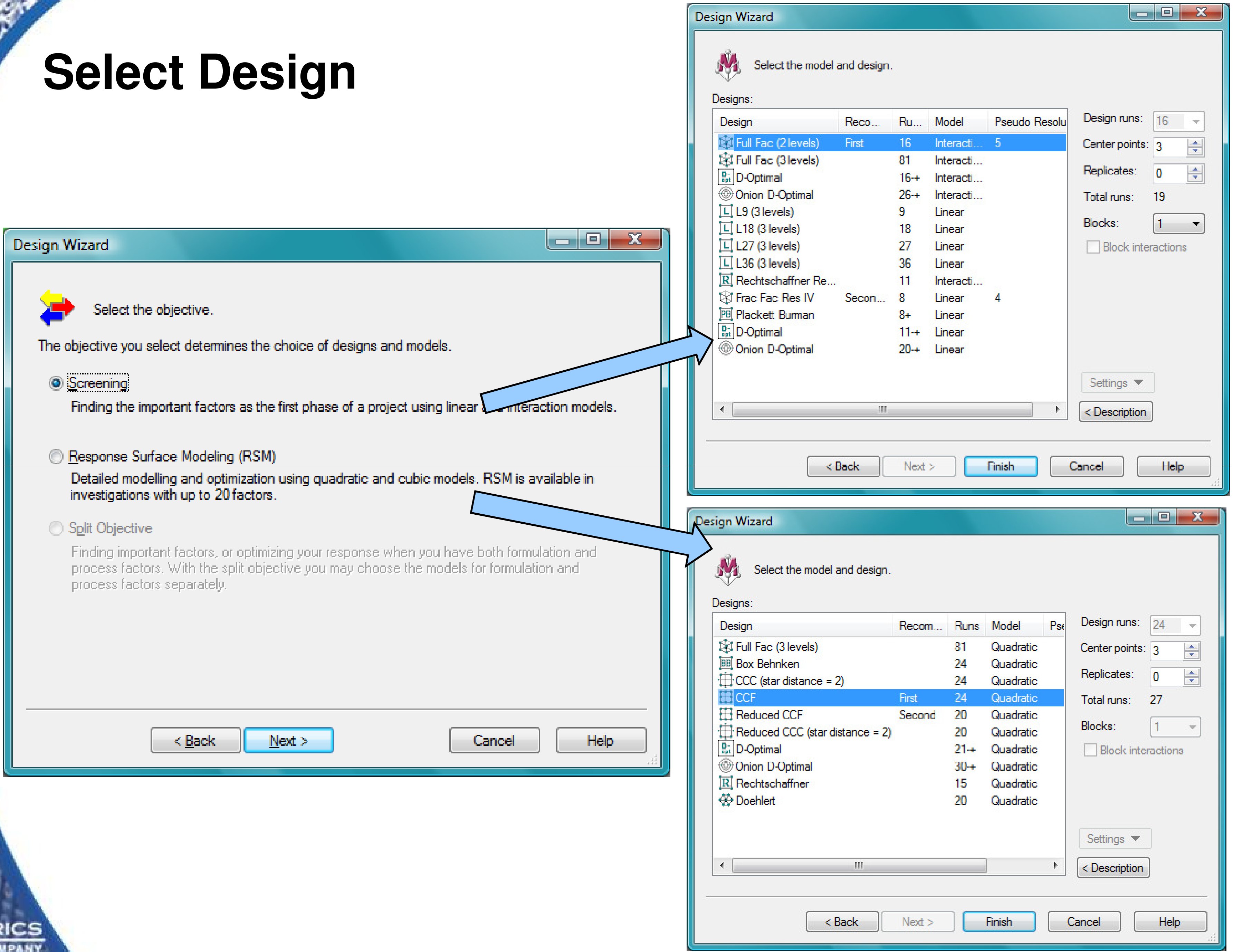
UMETRICS
AN MKS COMPANY

Define Responses

- Size – particle size
  - Specification: 0.5 – 3.3 µm
- Yield – the amount of product produced.
  - To be maximized
- Water – water content in spray-dried protein.
  - To be minimized



Select Design



Overview of data analysis using Screening Example: Protein spray drying

- 2^4 full factorial design

Worksheet

	1	2	3	4	5	6	7	8	9	10	11
1	Exp No	Exp Name	Run Order	Incl/Excl	Inlet temperature	Atomization Gasflow	Aspiration rate	Feed-flow	Yield	Size	Water
2	1	N1	1	Incl	100	500	60	2	43,5	3,1	4,5
3	2	N2	6	Incl	220	500	60	2	43,5	3,8	3,2
4	3	N3	17	Incl	100	800	60	2	8,5	2,2	4,5
5	4	N4	15	Incl	220	800	60	2	9	2	3,8
6	5	N5	19	Incl	100	500	100	2	51	3,4	3,1
7	6	N6	11	Incl	220	500	100	2	61	4,2	1,9
8	7	N7	16	Incl	100	800	100	2	22,5	1,6	3,3
9	8	N8	18	Incl	220	800	100	2	27	1,6	2,8
10	9	N9	12	Incl	100	500	60	5	48	3,2	4,8
11	10	N10	8	Incl	220	500	60	5	45	3,6	2,3
12	11	N11	5	Incl	100	800	60	5	12,5	2,1	5,1
13	12	N12	4	Incl	220	800	60	5	11	2,3	3,3
14	13	N13	9	Incl	100	500	100	5	61,5	3,2	4
15	14	N14	14	Incl	220	500	100	5	60,5	4,4	2,2
16	15	N15	13	Incl	100	800	100	5	35	1,6	4,2
17	16	N16	3	Incl	220	800	100	5	33	1,75	2,7
18	17	N17	2	Incl	160	650	80	3,5	34,5	2,2	3,4
19	18	N18	10	Incl	160	650	80	3,5	39	2,4	3,7
20	19	N19	7	Incl	160	650	80	3,5	38	2,2	3,2

UMETRICS
AN MKS COMPANY

Analysis/ Evaluate

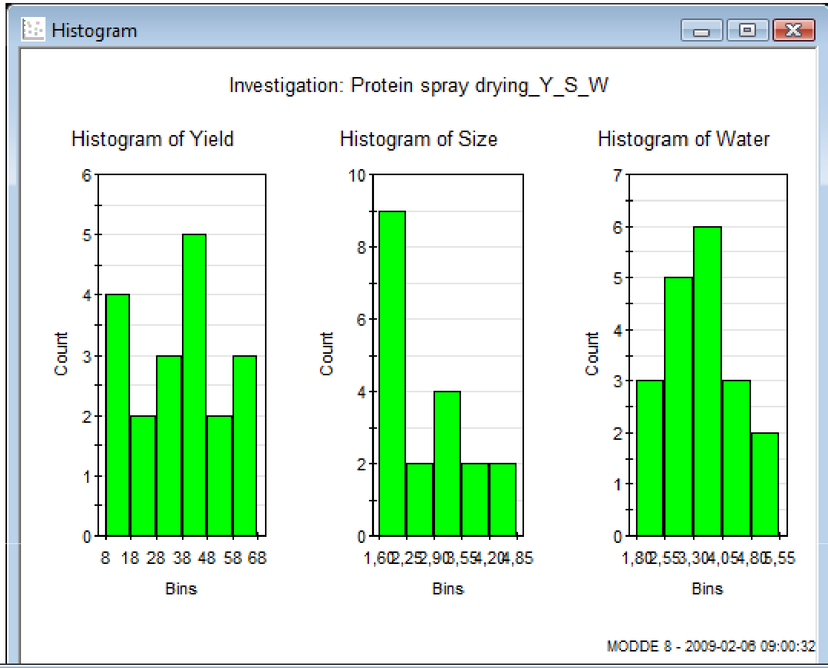
Worksheet/ Histogram

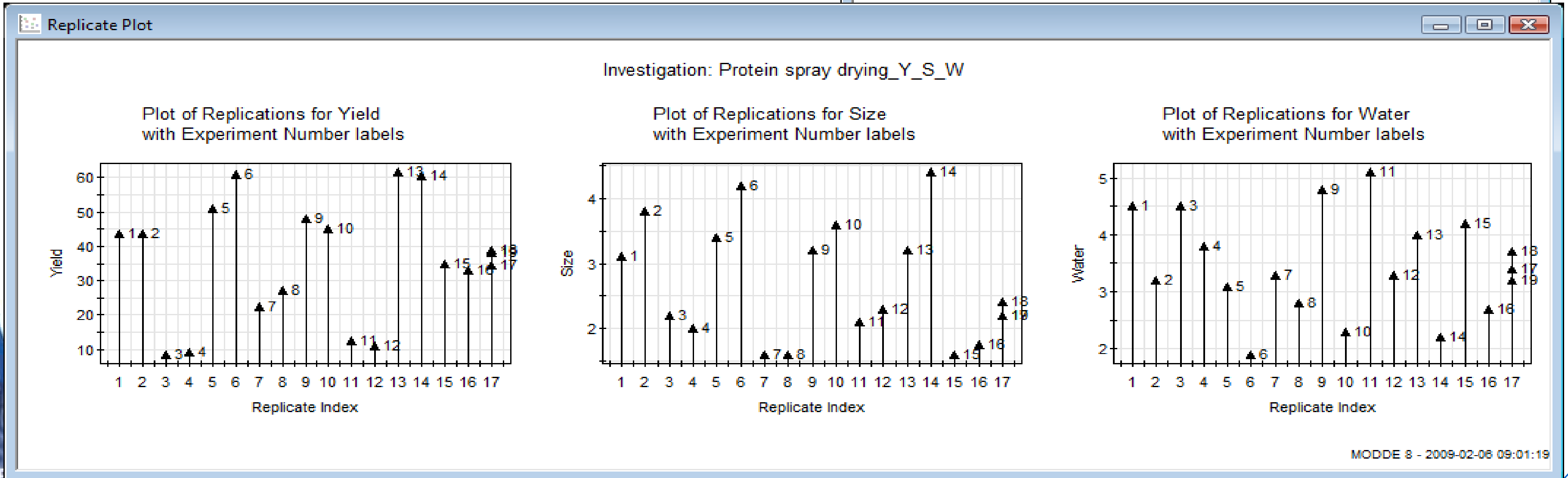
Worksheet/ Replicate plot

Raw data evaluation

Evaluate

	1	2
1		**Evaluation of MLR model**
2		All factors are orthogonally scaled
3	Condition number	1,08972
4	Worksheet runs	19
5	Model terms	11
6	DF residual	8
7	DF lack of fit	6
8	DF pure error (repl. runs)	2





Raw data evaluation

- Correlations between model terms?
- Correlations between factors and responses?

Correlation Matrix

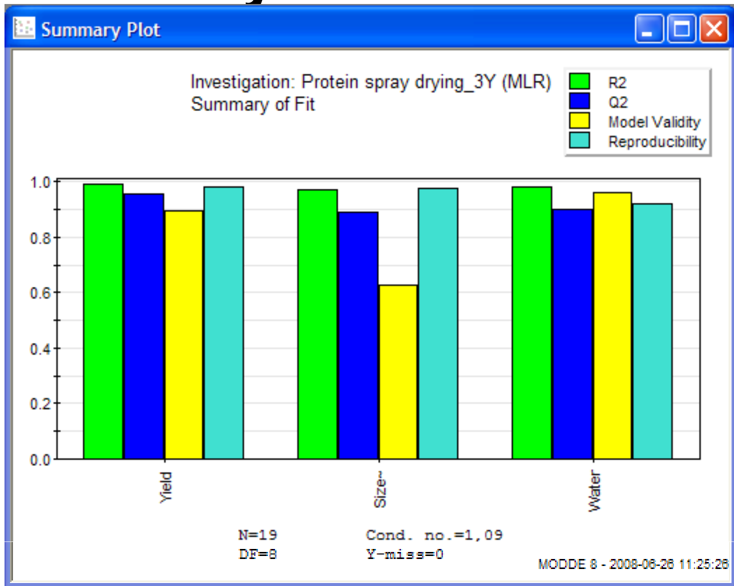
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1		InT	AtGF	AsR	FeF	InT*AtGF	InT*AsR	InT*FeF	AtGF*AsR	AtGF*FeF	AsR*FeF	Yie	Siz	H2O
2	InT	1	0	0	0	0	0	0	0	0	0	0,0255415	0,212047	-0,745571
3	AtGF	0	1	0	0	0	0	0	0	0	0	-0,870115	-0,897124	0,244125
4	AsR	0	0	1	0	0	0	0	0	0	0	0,444423	-0,0358849	-0,481652
5	FeF	0	0	0	1	0	0	0	0	0	0	0,137924	0,0163113	0,0989696
6	InT*AtGF	0	0	0	0	1	0	0	0	0	0	-0,0153249	-0,192474	0,151753
7	InT*AsR	0	0	0	0	0	1	0	0	0	0	0,0527859	0,0685076	0,0857737
8	InT*FeF	0	0	0	0	0	0	1	0	0	0	-0,0766246	0,0424095	-0,257321
9	AtGF*AsR	0	0	0	0	0	0	0	1	0	0	0,0766246	-0,231621	-0,00659796
10	AtGF*FeF	0	0	0	0	0	0	0	0	1	0	0,0289471	0,0293604	0,0197939
11	AsR*FeF	0	0	0	0	0	0	0	0	0	1	0,0561914	0,00326229	0,164949
12	Yie	0,02554	-0,87011	0,44442	0,13792	-0,01532	0,052785	-0,07662	0,076624	0,028947	0,056191	1	0,753866	-0,404094
13	Siz	0,21204	-0,89712	-0,0358	0,01631	-0,19247	0,068507	0,042409	-0,23162	0,029360	0,003262	0,753866	1	-0,377578
14	H2O	-0,7455	0,24412	-0,4816	0,09896	0,151753	0,085773	-0,25732	-0,00659	0,019793	0,164949	-0,404094	-0,377578	1

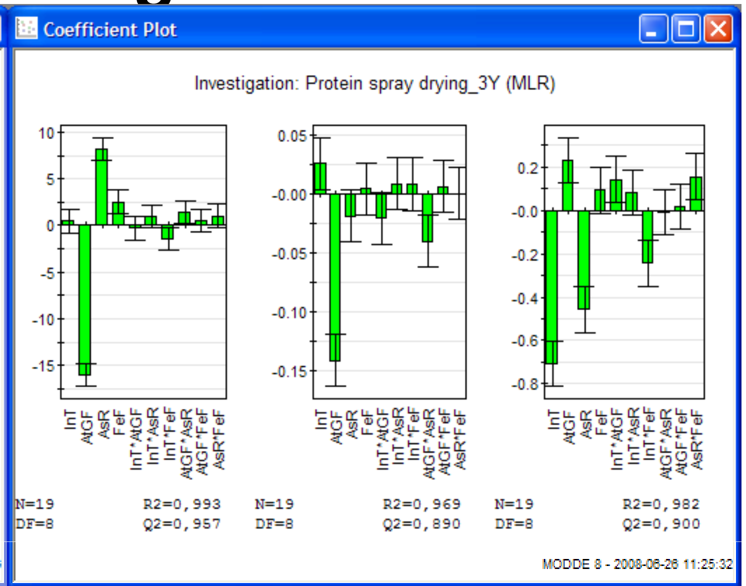
Threshold 0,3

UMETRICS
AN MKS COMPANY

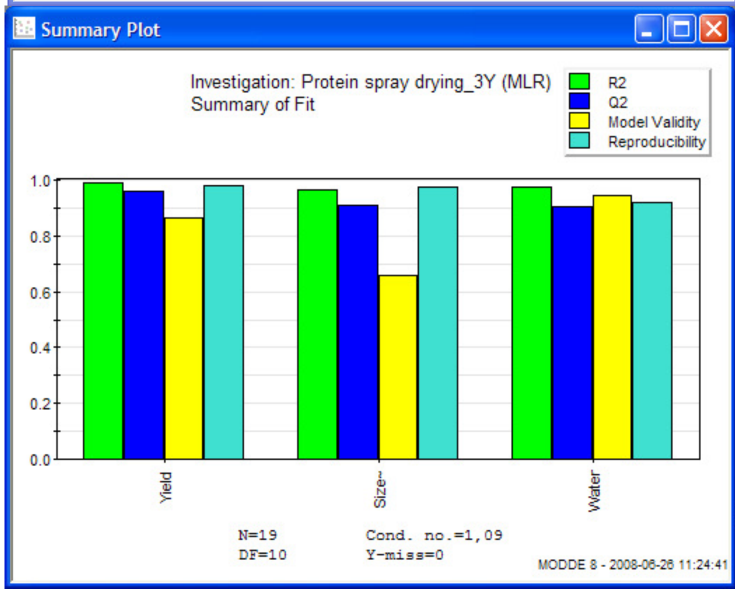
Regression analysis and model diagnostics

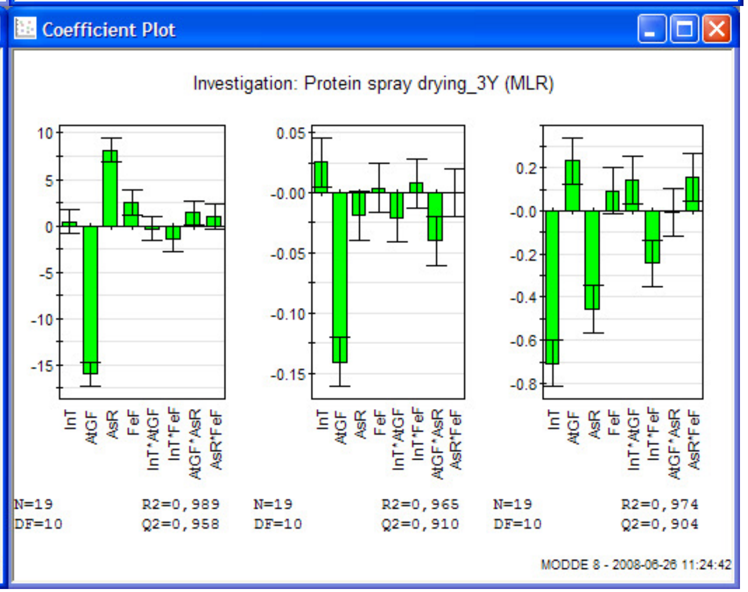
- Original model





- Refined model
- ANOVA OK

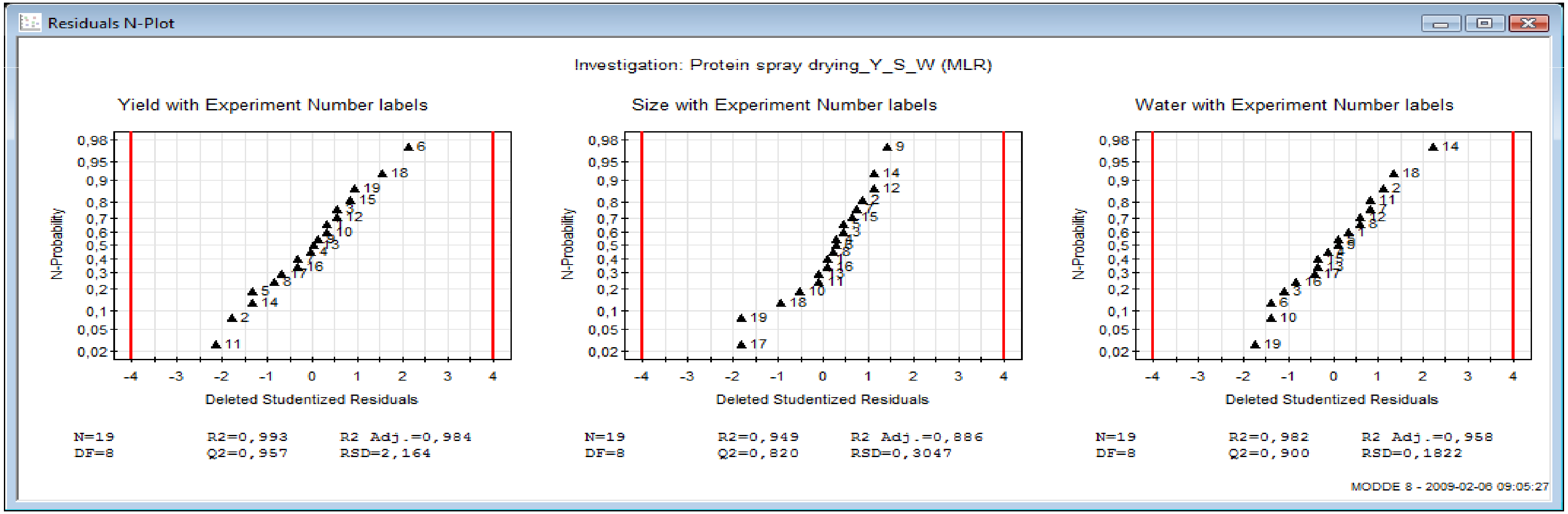




UMETRICS
AN MKS COMPANY

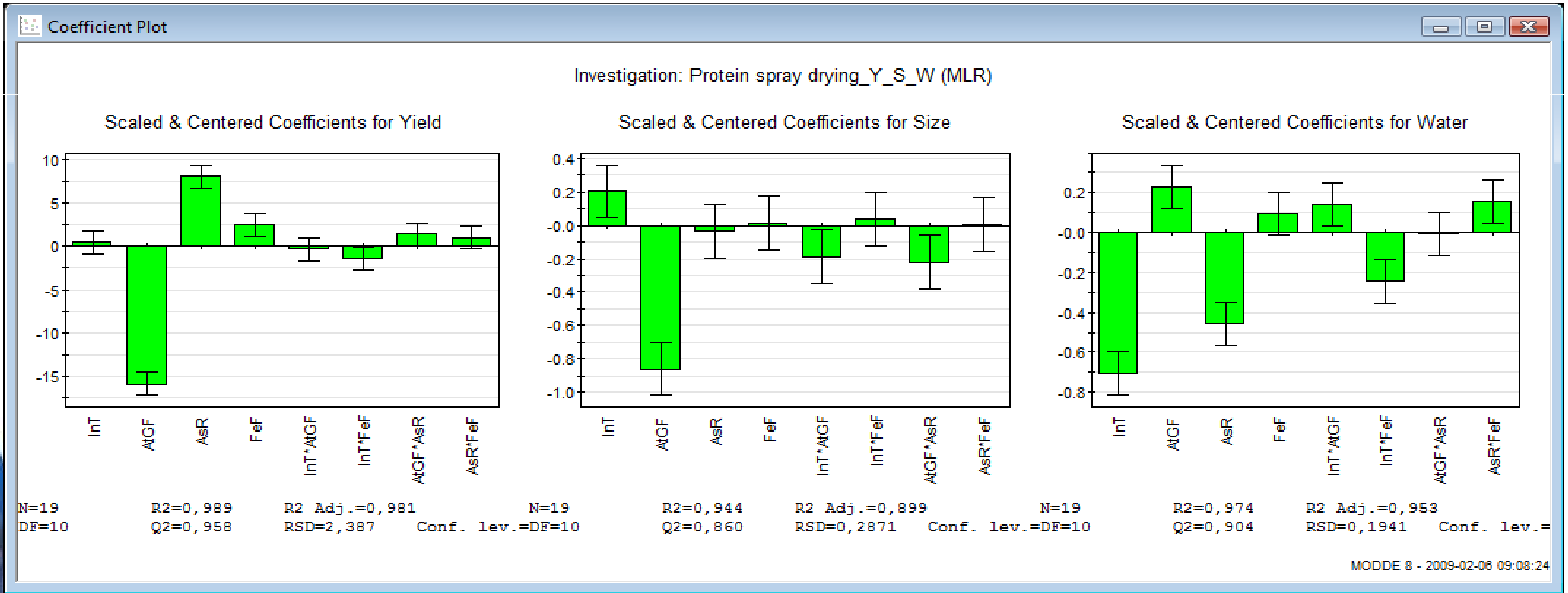
Regression analysis and model diagnostics

- No deviators (”outliers”) in Residuals N-plot



Model interpretation

- Atomization gas flow, Aspiration rate and Feedflow influence Yield
- Inlet temperature and atomization gas flow influence Size
- Inlet temperature, atomization gas flow and Aspiration rate influence Water

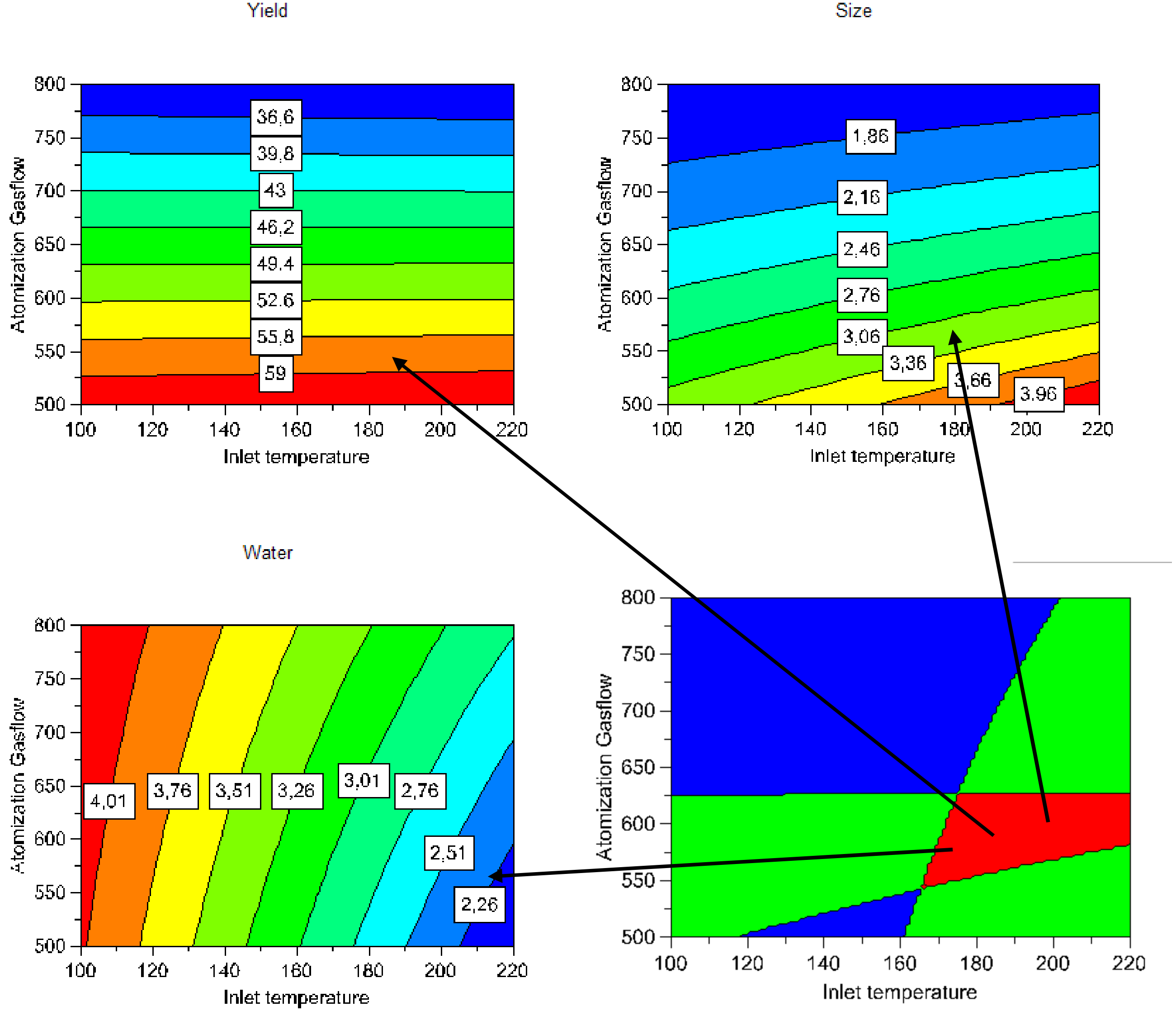


Prediction/ Contour plot wizard

Prediction / Sweet spot plot

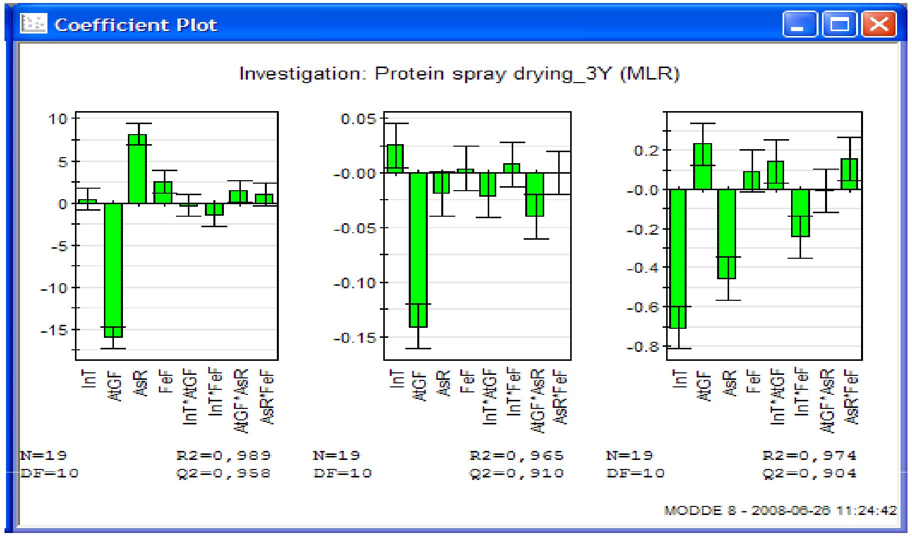
Protein spray drying - Use of model

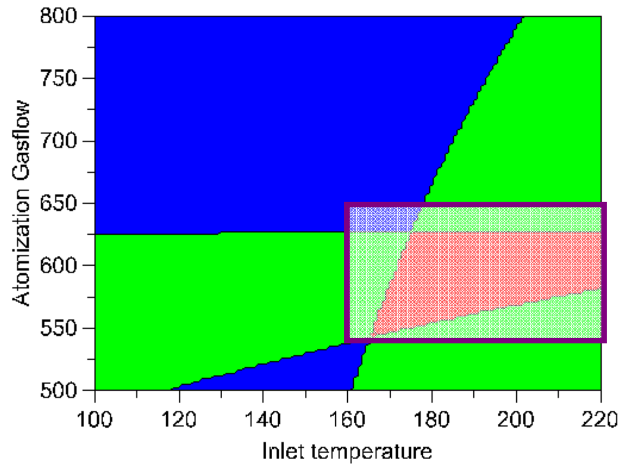
- Maximize yield
- Size < 3.3
- Minimize water



Conclusions

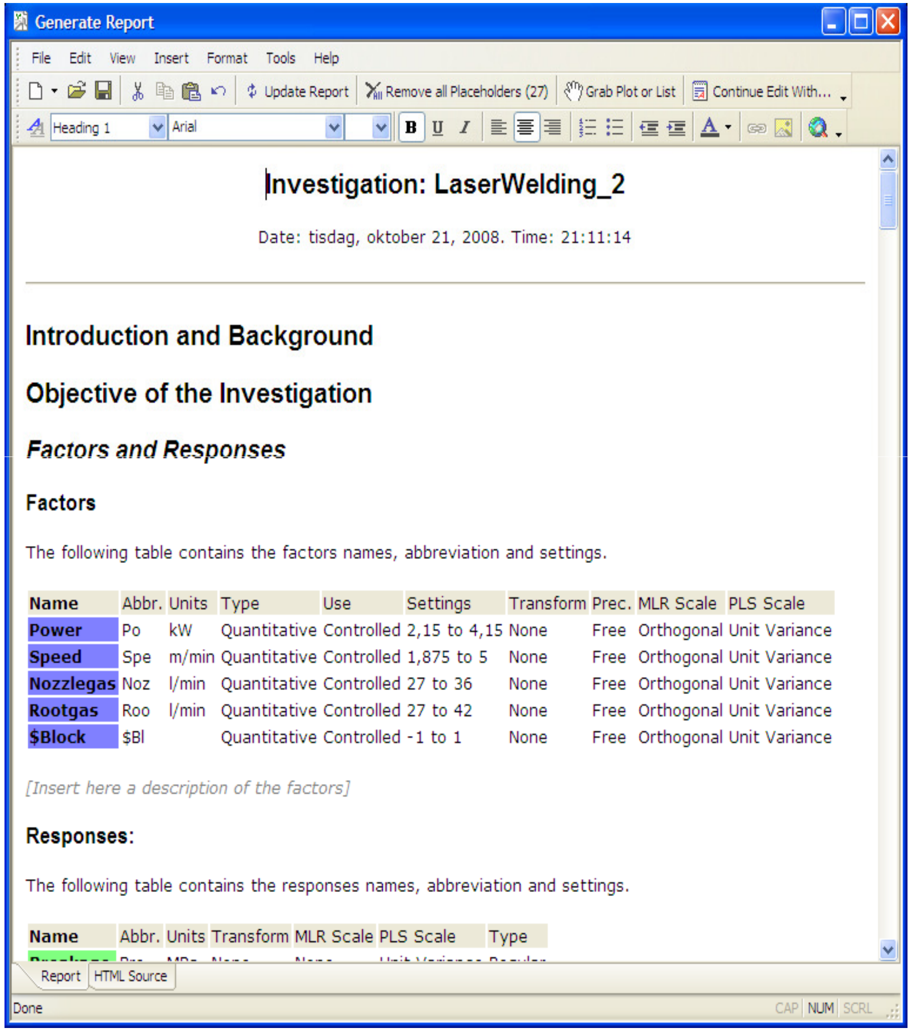
- All four investigated factors were influential
  - One response or the other…
  - Yield: AtGasflow, aspiration rate
  - Size: AtGasflow, InletTemp
  - Water: Inlet Temp, AtGasflow, aspiration rate
- None of the four factors could be eliminated
- An optimization design next step





Report generator in MODDE

- Report generator creates a html-report according to template
- Templates can simply be custom made with plots, tables and graphs of choice
- Room for comments and text written by experimenter
- Same report format each time!
- Entire DOE process, from factor definition to report, can be fully automized for standard experiments
  - Adapt production to a new batch of raw materials



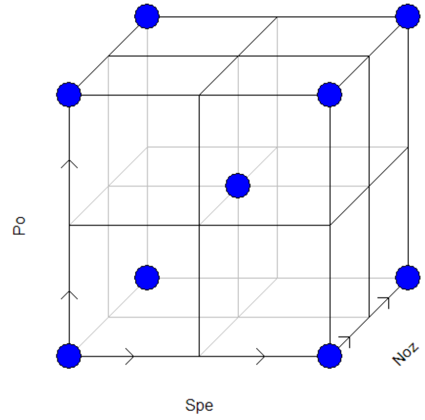
Summary

- Data analysis of DOE-data comprises three stages
  - evaluation of raw data
    - done to understand and clean data, and speed up regression modelling
  - regression analysis and model interpretation
    - done to derive the predictively most relevant model with meaningful mechanistic interpretation
  - use of model
    - done to find out the impact of the model: What does it mean? Where should new experiments be positioned?

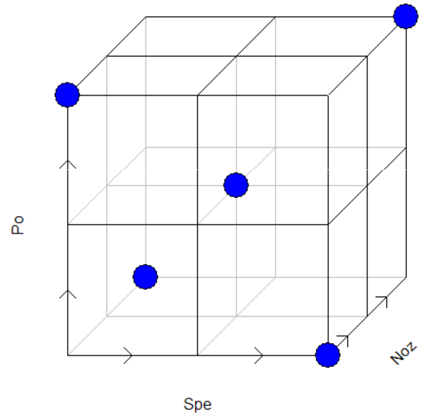
UMETRICS
AN MKS COMPANY

Design types Screening

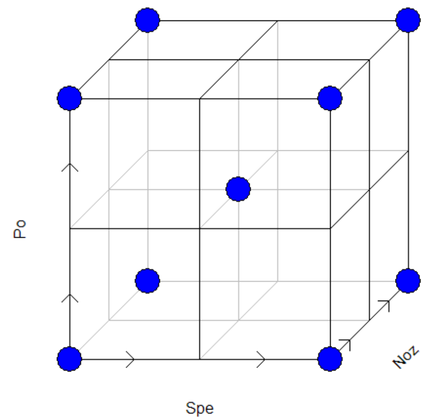
- Factorial design



- Fractional factorial design



- Plackett-Burman
  - For 3 factors identical to factorial design



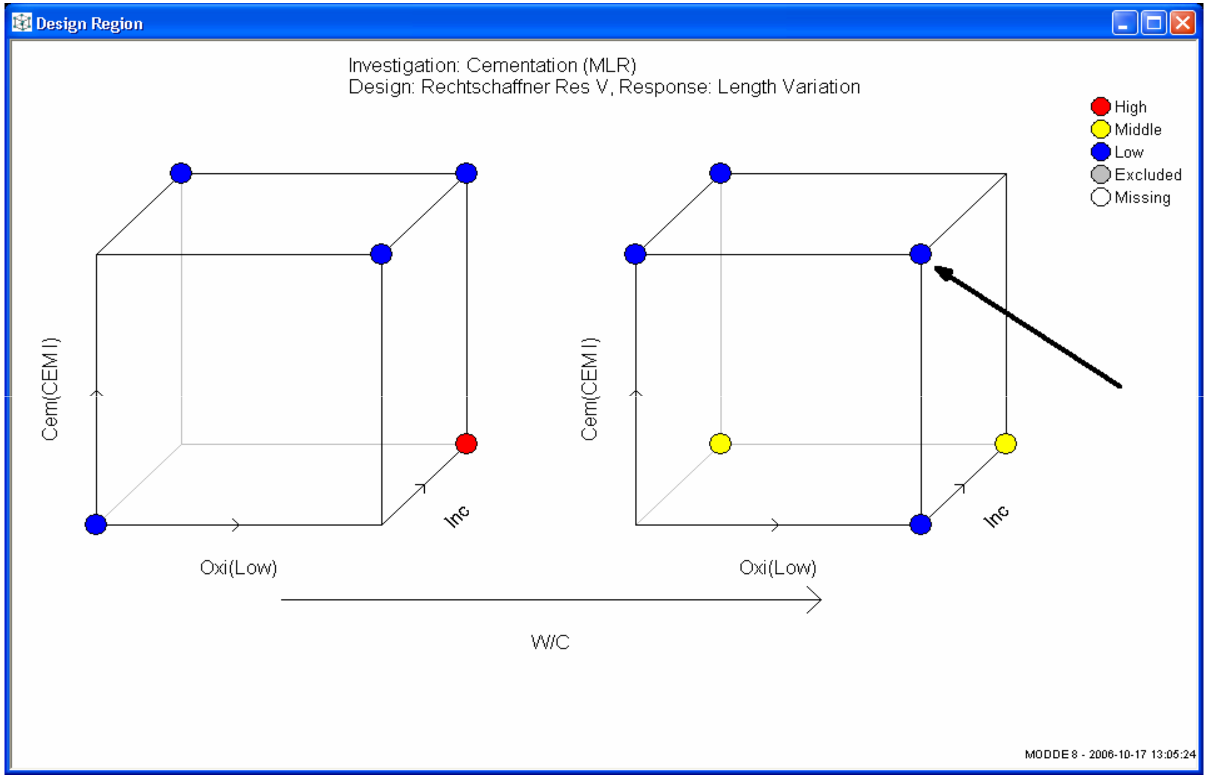
Rechtschaffner designs
(saturated fractional factorials with 2 or 3 levels)

- Interactions non-confounded with other interactions and main effects, i.e., resol. V; K <= 10,
- Efficient in situations when interactions estimates are critical (e.g. Pharma R&D, Chem R&D, Engn R&D)
- Example with 5 factors in 16 runs (identical to half 2^5)
- MODDE will add 3 ctr pts as default

run	generator	x1	x2	x3	x4	x5
1	– – – – –	–	–	–	–	–
2	– + + + +	–	+	+	+	+
3		+	–	+	+	+
4		+	+	–	+	+
5		+	+	+	–	+
6		+	+	+	+	–
7	+ + – – –	+	+	–	–	–
8		+	–	+	–	–
9		+	–	–	+	–
10		+	–	–	–	+
11		–	+	+	–	–
12		–	+	–	+	–
13		–	+	–	–	+
14		–	–	+	+	–
15		–	–	+	–	+
16		–	–	–	+	+

Geometry of Rechtschaffner design with K = 4

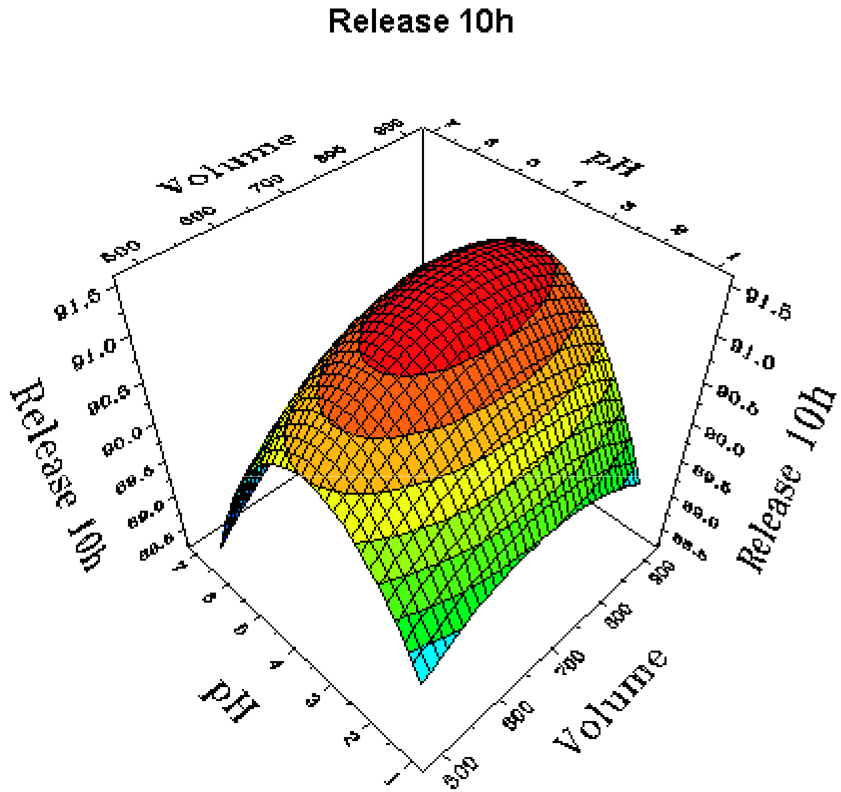
- Saturation means that the number of terms in the model equals the number of factorial points in the design, e.g.
  3 factors ↔ 7 runs,
  4 factors ↔ 11 runs,
  5 factors ↔ 16 runs,
  6 factors ↔ 22 runs;
  7 factors ↔ 29 runs,
  etc.
- Resolution V means that two-factor interactions are not confounded with other two-factor interactions.



OBJECTIVE OPTIMIZATION

Optimization- Introduction

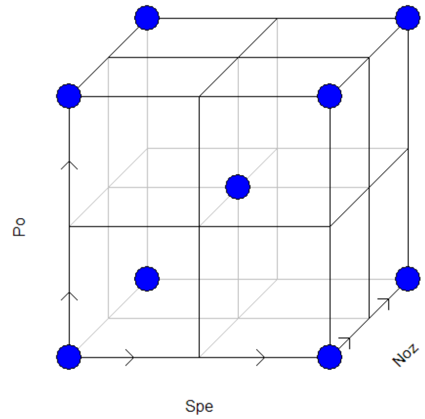
- Carried out when influential parameters are identified
  - We do not ask *if* a factor is relevant (screening), but *how* (optimization)
- Aim: Gain detailed knowledge about the factor influences
- Goal: To identify the factor combination at which the desired response profile is fulfilled (or almost so)

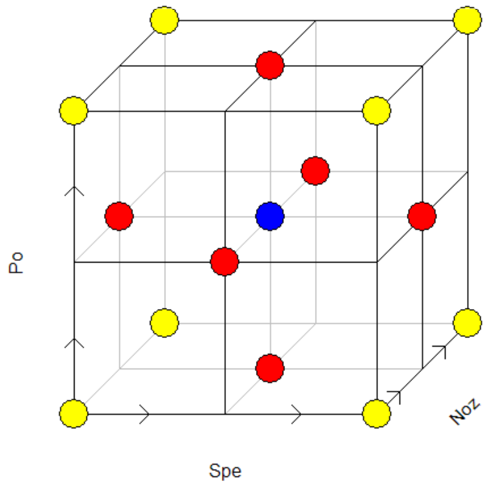


RSM: Response surface modelling (methodology)

Nature of optimization designs

- Detailed knowledge!
  - More experiments than screening
- Good RSM designs must give data with the ability
  - to allow estimation of model parameters with low uncertainty
  - to give rise to a model with small prediction error
  - have prediction error independent of direction
  - to permit a judgement of the adequacy of the model
  - to encode as few experiments as possible





Optimization Example – Loading conditions

- Optimization of loading conditions on Capto S
- Capto S is a chromatography medium for IEC
  - GE healthcare
- Objective:
  - Identify best loading conditions to get the highest dynamic binding capacity at 10% breakthrough ($QB_{10\%}$).

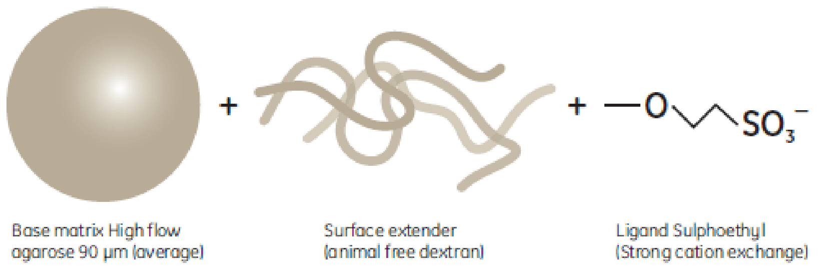


Fig 1. Capto S is based on a high flow agarose base matrix modified with dextran and strong cation exchange groups for optimal rigidity, capacity and mass transfer.

Ref: GE Healtcare Application note 28-4078-16 AA

Capto S- Factors and responses

- Objective: Investigate relationship between factors Conductivity, Residence time and pH with respect to QB10%

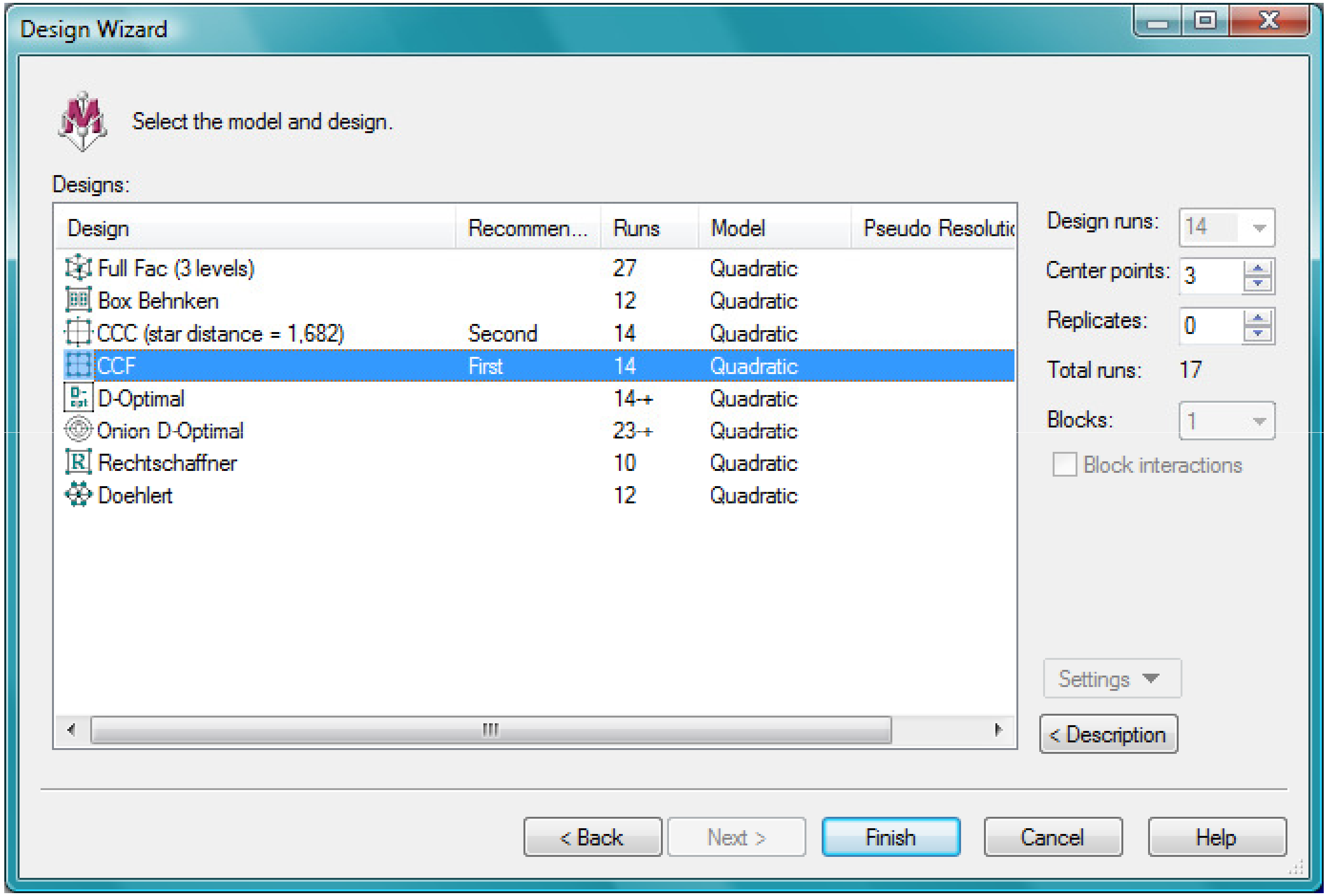
Factors

	Name	Abbr.	Units	Type	Use	Settings	Transform	Prec.	MLR Scale	PLS Scale
1	pH	pH		Quantitative	Controlled	4,5 to 5,5	None	Free	Orthogonal	Unit Variance
2	Residence time	Res	Min	Quantitative	Controlled	2 to 6	None	Free	Orthogonal	Unit Variance
3	Conductivity	Cond	ms/cm	Quantitative	Controlled	5 to 15	None	Free	Orthogonal	Unit Variance

Responses

	Name	Abbr.	Units	Transform	MLR Scale	PLS Scale	Type
1	QB10%	QB1	mg/ml medium	None	None	Unit Variance	Regular
				Double-click	*here to*	*add a new*	*response*

Capto S- Select design

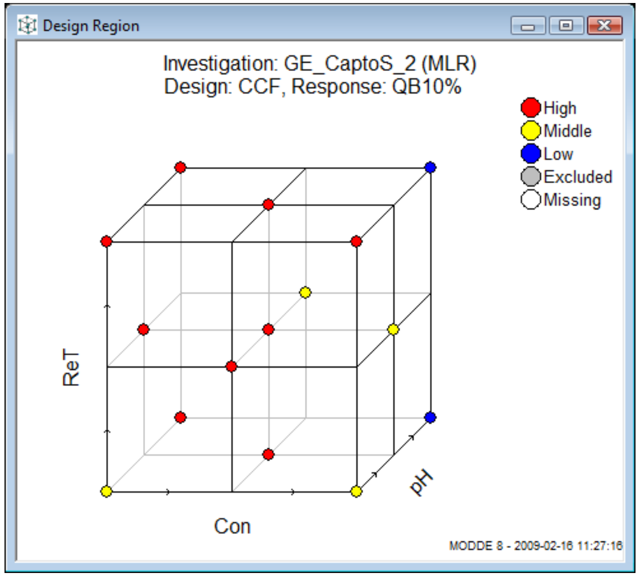


Selected Design

- Optimization design
  - Detailed knowledge about few factors and their relation to a response
- CCF design
  - Central Composite face

Worksheet

	1	2	3	4	5	6	7	8
1	Exp No	Exp Name	Run Order	Incl/Excl	Residence time	Cond	pH	QB10%
2	1	N1	1	Incl	2	5	4,5	76
3	2	N2	7	Incl	6	5	4,5	116
4	3	N3	11	Incl	2	15	4,5	99
5	4	N4	15	Incl	6	15	4,5	105
6	5	N5	16	Incl	2	5	5,5	126
7	6	N6	13	Incl	6	5	5,5	149
8	7	N7	17	Incl	2	15	5,5	0
9	8	N8	10	Incl	6	15	5,5	8
10	9	N9	9	Incl	2	10	5	125
11	10	N10	2	Incl	6	10	5	121
12	11	N11	14	Incl	4	5	5	139
13	12	N12	8	Incl	4	15	5	54
14	13	N13	4	Incl	4	10	4,5	119
15	14	N14	5	Incl	4	10	5,5	84
16	15	N15	12	Incl	4	10	5	121
17	16	N16	3	Incl	4	10	5	137
18	17	N17	6	Incl	4	10	5	143



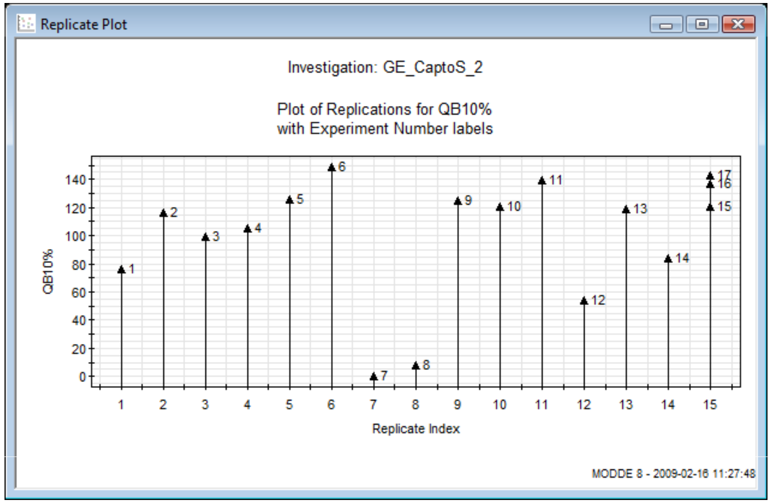
Evaluate

	1	2
1		Evaluation of MLR model
2		All factors are orthogonally scaled
3	Condition number	3,97813
4	Worksheet runs	17
5	Model terms	7
6	DF residual	10
7	DF lack of fit	8
8	DF pure error (repl. runs)	2

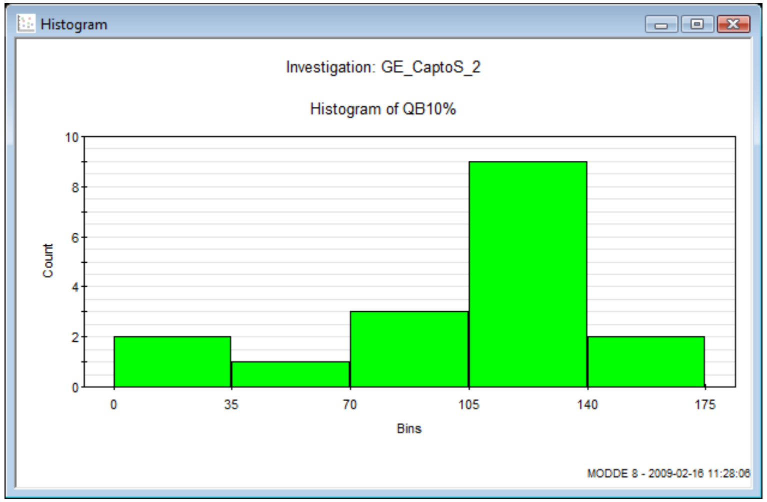
UMETRICS
AN MKS COMPANY

Raw data evaluation

- Replicate plot shows:
  - Small replicate errors
  - Replicates (centre points) positioned high in response interval
    - Indicates non-linearities

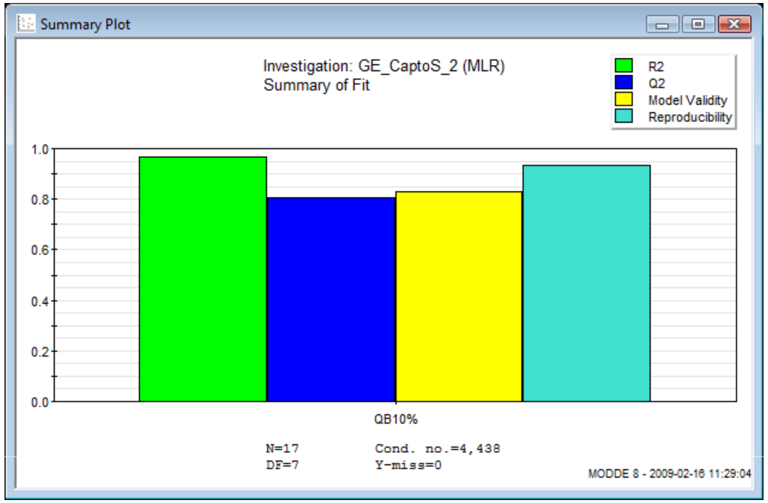


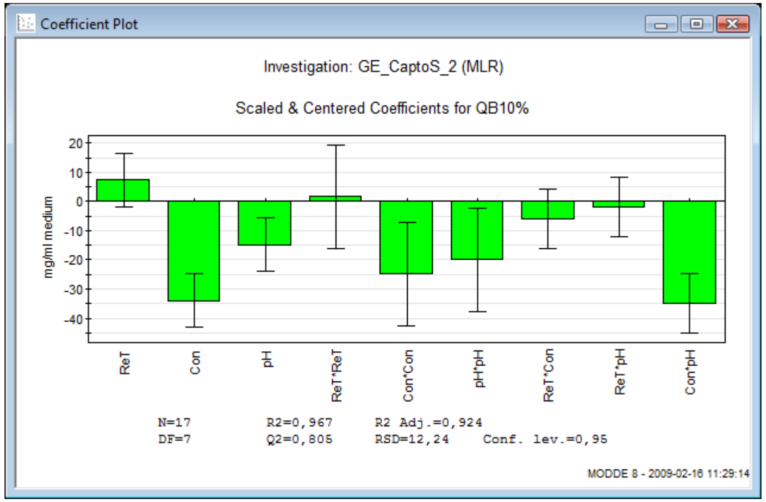
- No pre-treatment necessary



Regression Modeling and interpretation

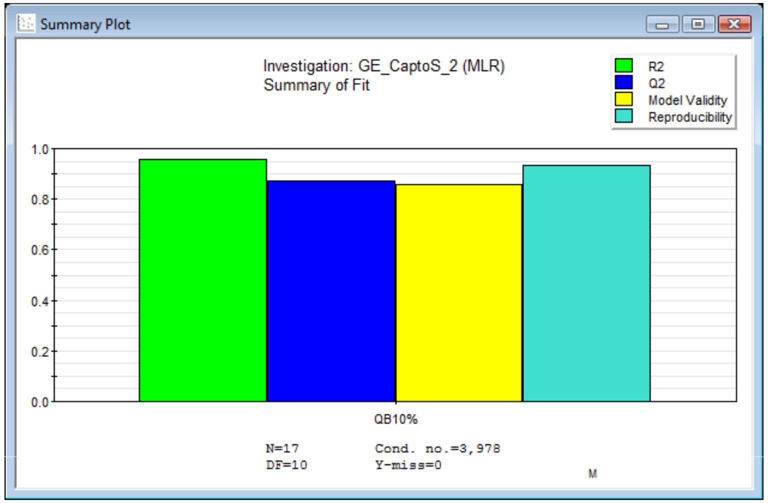
- Model fitted using MLR
- High R^2 and Q^2
- No Lack of fit
- High reproducibility
- Some model pruning possible

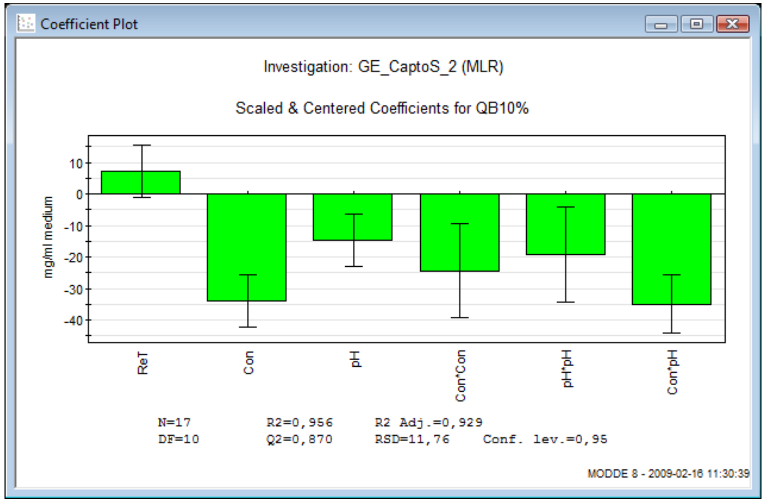




Model pruning and interpretation

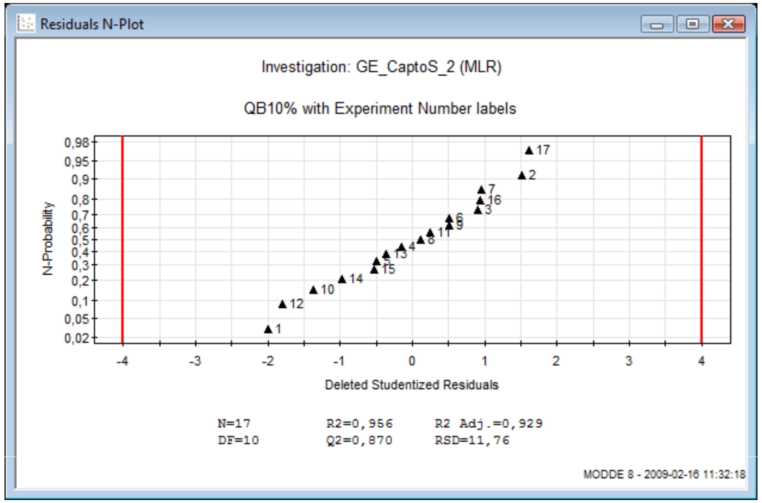
- Two Interaction terms removed
  - ReT*Con
  - ReT*pH
- One quadratic term
  - ReT*ReT
- Model slightly improved
  - Increase in Q^2 from 0,8 to 0,87



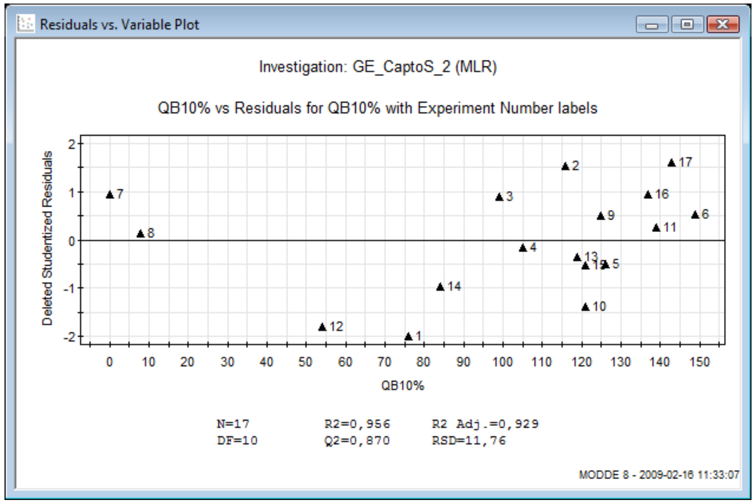


Regression Modeling and interpretation

- No deviating experiments in residuals N-plot



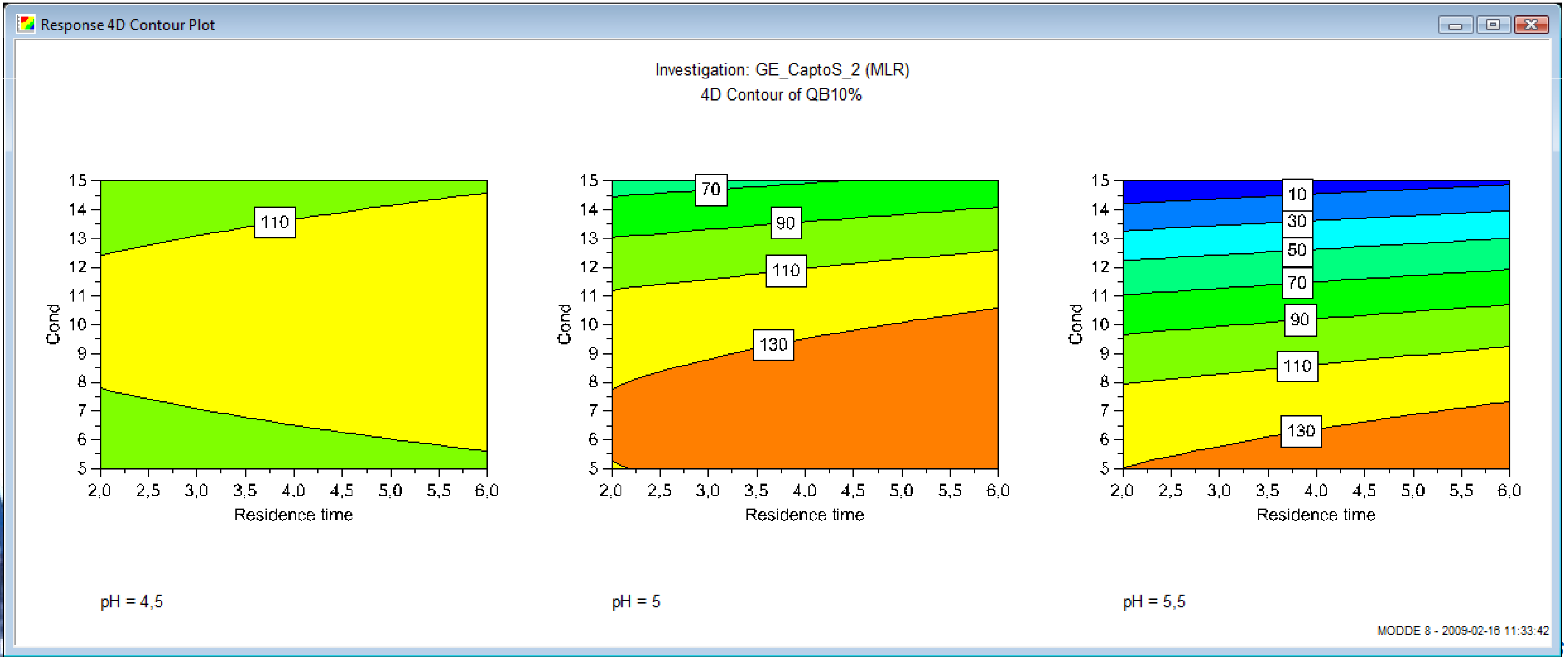
- Trends in residuals can be investigated using Residuals vs variable plot or Residual vs run order
  - Check for time trends



UMETRICS
AN MKS COMPANY

Model visualization and use of model

- Interpretation!
- Contour plots
- Response surfaces
- Next step?

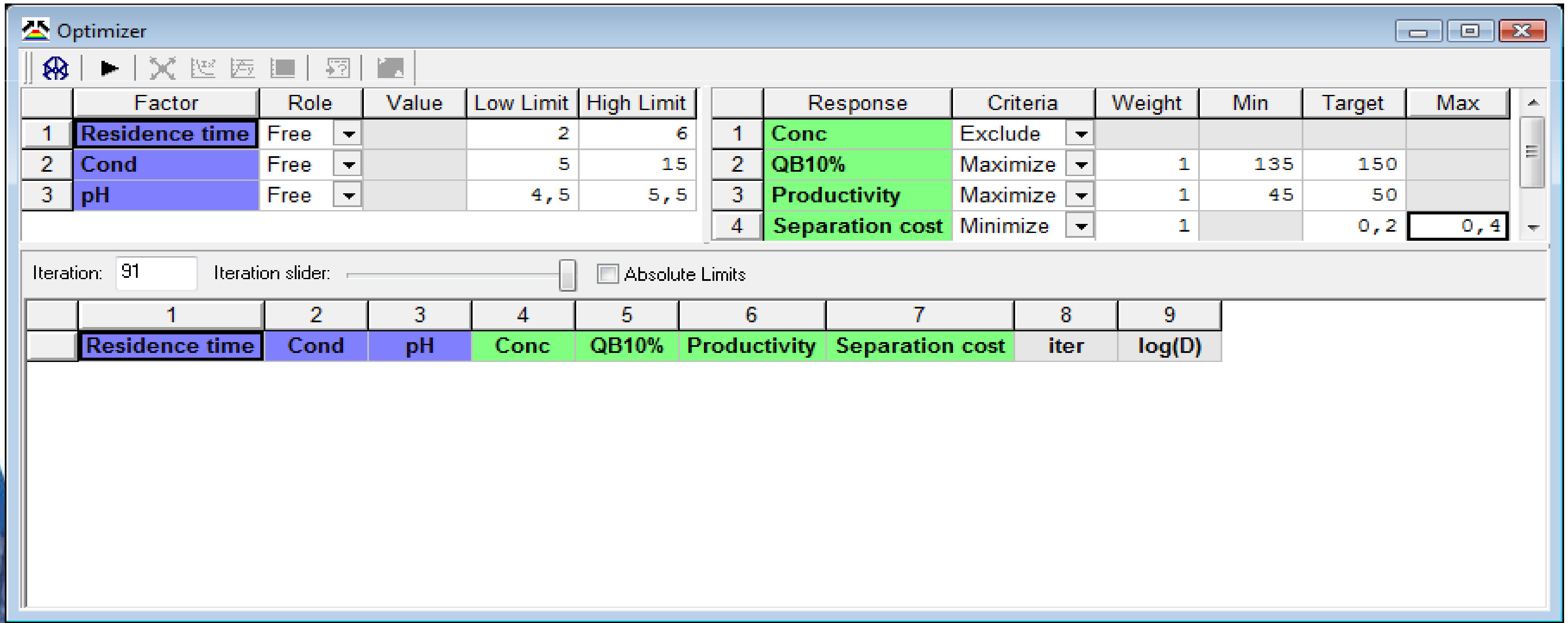


Optimizer possibility in MODDE 8

- When more than 1-2 responses an algorithmic approach is used to find optimum
- Three additional responses for Capt S example:
  - Concentration
  - Productivity
  - Separation cost

MODDE Optimizer

- Where is the optimal region?
- Most beneficial for multi-y models
- Define allowed factor ranges and wanted response values



MODDE Optimizer

Click 1, Click 2

Optimizer

Run Optimizer

		Role	Value	Low Limit	High Limit
1	Residence time	Free		2	6
2	Cond	Free		5	15
3	pH	Free		4,5	5,5

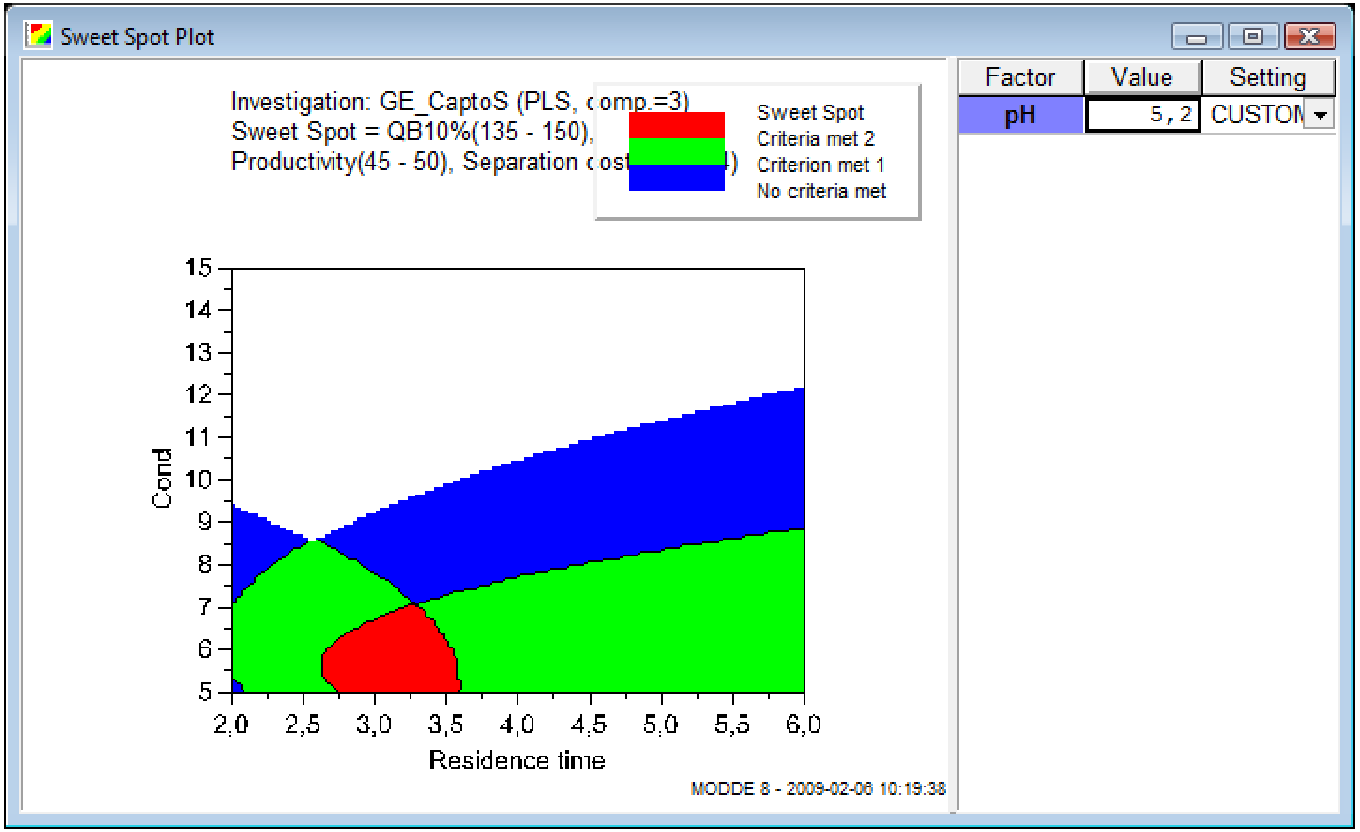
	Response	Criteria	Weight	Min	Target	Max
1	Conc	Exclude				
2	QB10%	Maximize	1	135	150	
3	Productivity	Maximize	1	45	50	
4	Separation cost	Minimize	1		0,2	0,4

Iteration: 91 Iteration slider: Absolute Limits

	1	2	3	4	5	6	7	8	9
	Residence time	Cond	pH	Conc	QB10%	Productivity	Separation cost	iter	log(D)
1	3,3964	5,6014	5,1937	5,0415	137,528	45,5319	0,3683	91	-0,1351
2	3,3462	5,8102	5,1719	5,0496	137,279	45,7352	0,3695	86	-0,1417
3	3,353	5,6011	5,1884	5,042	137,358	45,6874	0,3694	54	-0,1403
4	5,1973	11,1655	4,6173	4,9714	129,305	38,1902	0,366	52	0,4352
5	3,0792	5,6491	5,1823	5,0462	136,417	46,545	0,3756	63	-0,1614
6	3,353	5,6011	5,1884	5,042	137,358	45,6874	0,3694	54	-0,1403
7	3,1413	5,6847	5,1856	5,0468	136,645	46,3341	0,374	80	-0,1575
8	3,115	5,6782	5,1822	5,0469	136,542	46,4302	0,3748	55	-0,1594

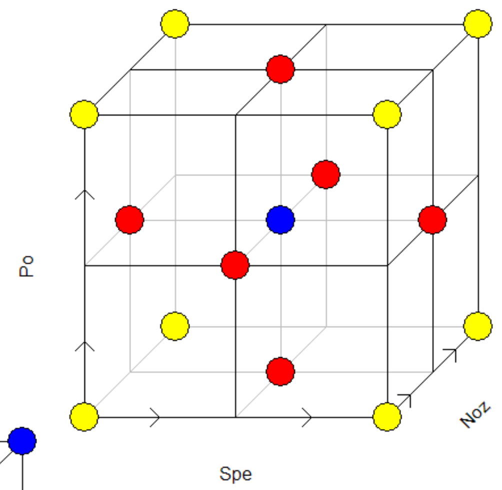
Sweet spot plot

- Red area- Sweet spot
- All requirements are fulfilled

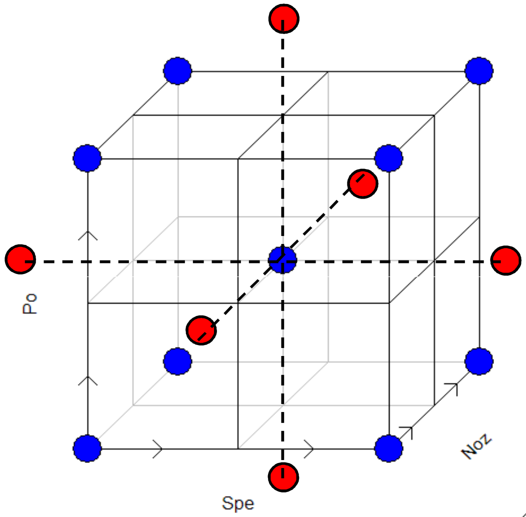


Design types Optimization/ RSM

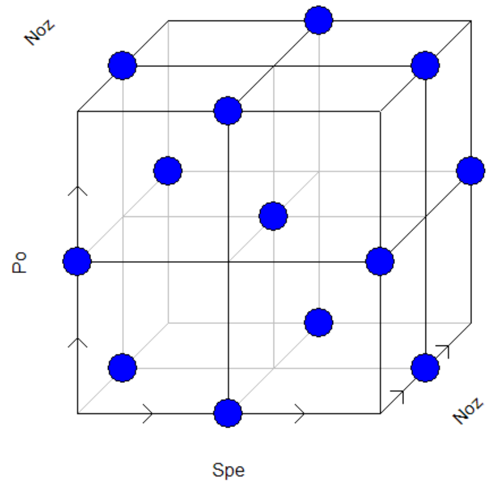
- Central composite face (CCF)



- Central composite circumscribed (CCC)

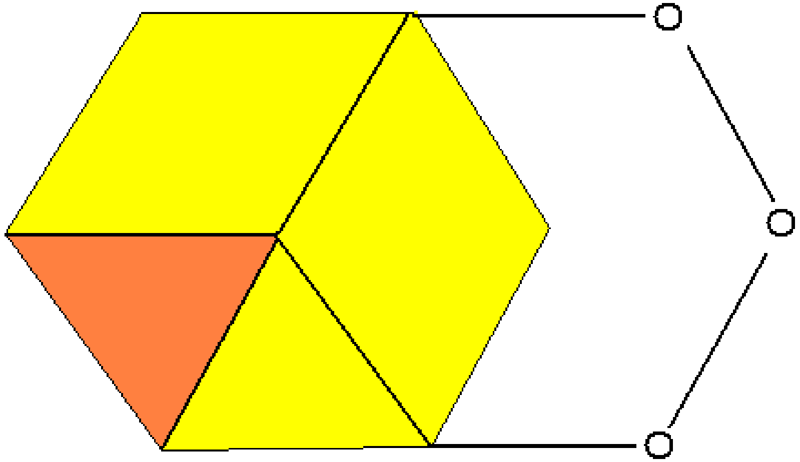


- Box- Behnken



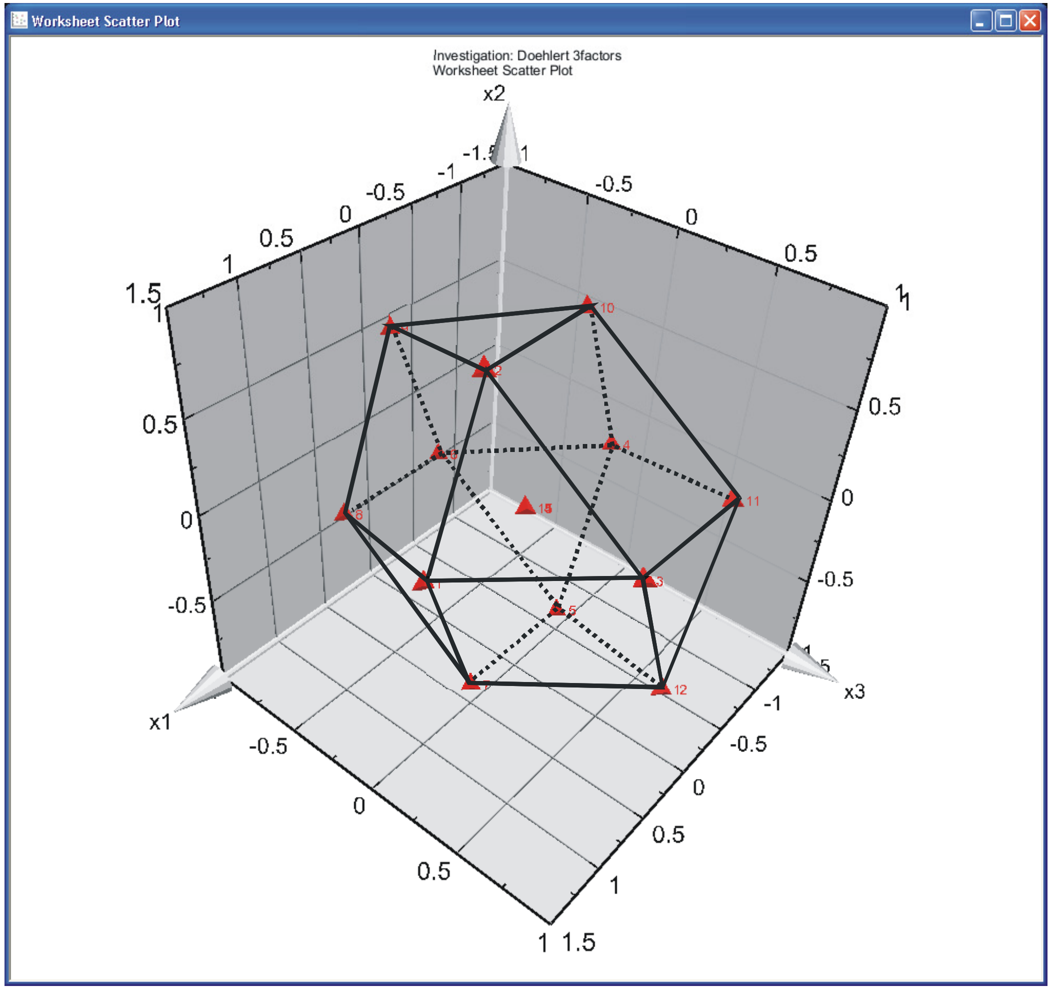
Doehlert designs (buildable RSM designs), K<= 10

- Common and favored in France (Phan Tan Luu, Goupy, …), efficient = small N
- Often applied in synthetic organic chemistry (e.g., Pharma R&D)
- Useful for sequential designs (buildable & addable)
- Based on hyper-hexagons and simplexes
- Support full quadratic models (RSM)



Doehlert design in 2 factors with 6 runs + ctr pts, extended to a new design by 3 runs + additional new ctr pts (blue)

Geometry of Doehlert designs for K = 2 and 3

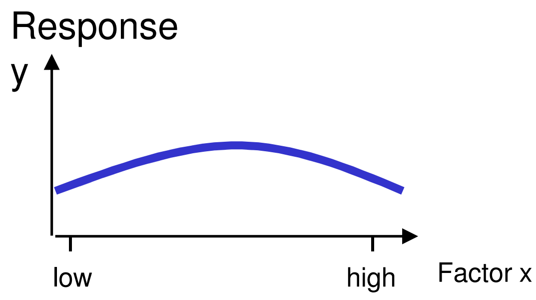


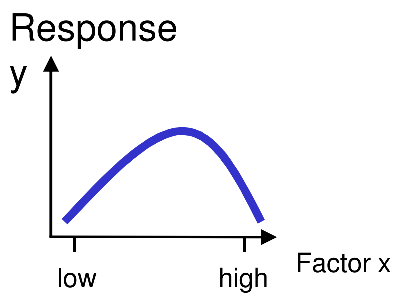
OBJECTIVE ROBUSTNESS TESTING

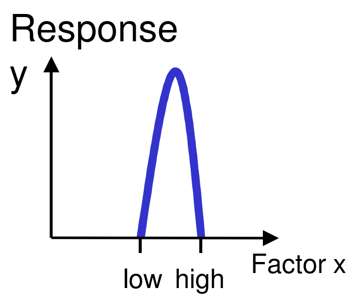
Introduction to robustness testing

- Main objective: PROVE robustness
  - Design space
- Investigate the system's sensitivity to "small" changes in critical factors
- Robustness is relative
- "Small factor changes ???"
  - variation that may normally occur in the laboratory
  - variation in raw materials, equipment, ...

- *Set point*: factor combination which is currently used for running the system







Introduction to robustness testing

- The objectives in robustness testing are:
  - to identify responses which are robust to small factor changes
  - to identify responses that are sensitive to small factors changes
  - to understand which factors that need to be better controlled to achieve robustness

- The ideal result in a robustness testing study is identical response values for each trial ⇔ low-resolution screening design supporting a linear model useful
  - Robustness assumptions based on knowledge about system proven

Four limiting cases of robustness testing

- Nature of robustness
  - Is regression model significant, or not?
  - Are responses inside or outside specifications?

- Four limiting cases
  - Inside specification/Significant model
  - Inside specification/Non-significant model
  - Outside specification/Significant model
  - Outside specification/Non-significant model

Robustness Testing Example - HPLC

- Five factors were varied in 12 runs
- Responses: capacity factors of two analytes and resolution between two adjacent peaks
  - Specification: Res1 should be >1.5 (complete baseline separation)

Factors

	Name	Abbr.	Units	Type	Use	Settings
1	AcN	Ac	%	Quantitative	Controlled	25 to 27
2	pH	pH		Quantitative	Controlled	3.8 to 4.2
3	Temp	Te	°C	Quantitative	Controlled	18 to 25
4	OSA	OS	mM	Quantitative	Controlled	0.09 to 0.11
5	Column	Co		Qualitative	Controlled	ColA, ColB

Responses

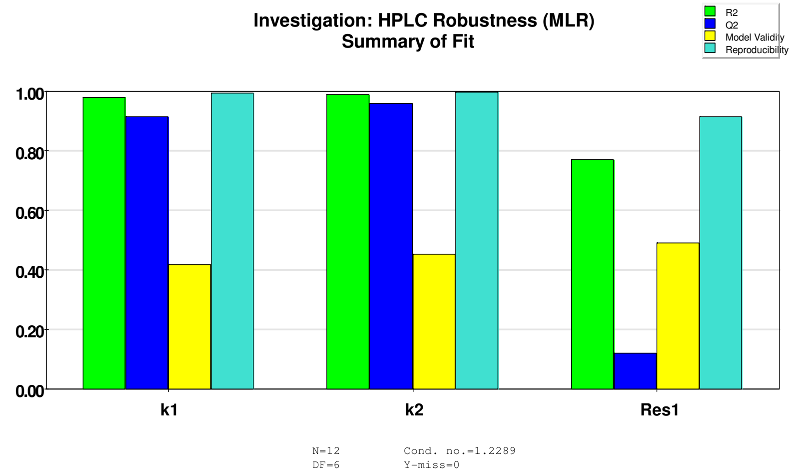
	Name	Abbr.	Units
1	k1	k1	
2	k2	k2	
3	Res1	Re1	

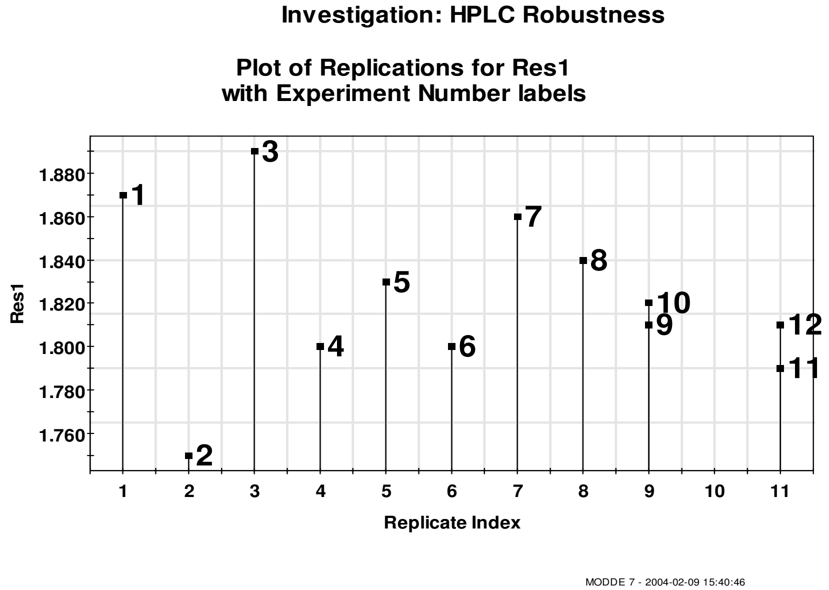
Worksheet

	1	2	3	4	5	6	7	8	9	10	11	12
1	Exp No	Exp Name	Run Order	Incl/Excl	AcN	pH	Temp	OSA	Column	k1	k2	Res1
2	1	N1	3	Incl	25	3.8	18	0.11	ColB	2.2906	3.3421	1.87
3	2	N2	8	Incl	27	3.8	18	0.09	ColA	1.7547	2.6802	1.75
4	3	N3	2	Incl	25	4.2	18	0.09	ColB	2.3933	3.4705	1.89
5	4	N4	1	Incl	27	4.2	18	0.11	ColA	1.823	2.8013	1.8
6	5	N5	6	Incl	25	3.8	25	0.11	ColA	2.1456	3.1599	1.83
7	6	N6	7	Incl	27	3.8	25	0.09	ColB	1.5031	2.4845	1.8
8	7	N7	5	Incl	25	4.2	25	0.09	ColA	2.2289	3.2715	1.86
9	8	N8	4	Incl	27	4.2	25	0.11	ColB	1.5994	2.6193	1.84
10	9	N12	9	Incl	26	4	22	0.1	ColA	2.0661	3.0592	1.81
11	10	N12	10	Incl	26	4	22	0.1	ColA	2.0253	3.0285	1.82
12	11	N12	11	Incl	26	4	22	0.1	ColB	2.0243	2.9903	1.79
13	12	N12	12	Incl	26	4	22	0.1	ColB	2.0131	3.0068	1.81

HPLC application - Conclusion

- Inside specification
  - Res 1 always exceeds 1.5
- Non-significant model
  - ANOVA (not shown)
  - $Q^2 = 0.12$
- **Res 1 is robust within investigated factor levels**



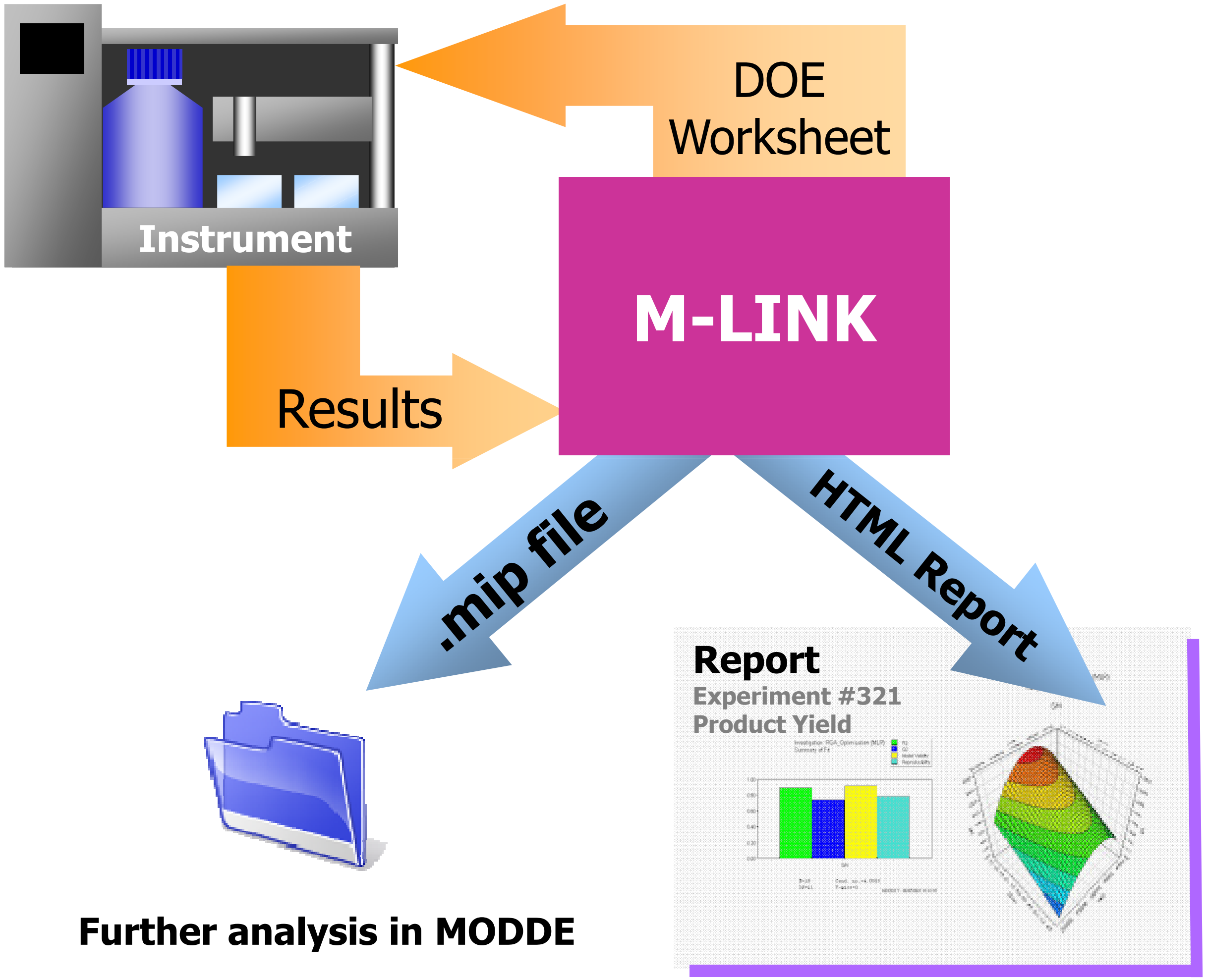


Automatic optimization and robustness testing

- Robustness testing can be a every-day routine
  - Can be automatic
- Using com-interface to MODDE
- M-Link
- Automatically plan, execute and evaluate standard operations
  - Lab and/ or Production environment
  - Semi or fully automatic
  - Built in to analytical instruments, manufacturing tools etc

Typical M-link system

- M-Link interface between instrument for defining experiment plan and retrieving results
- Automated Analysis and reporting
- Result export via HTML report or as MIP (MODDE) file
- No need for expert knowledge



DESIGN TYPE: MIXTURE DESIGN

Design of Mixture Experiments

- Experiments where the response Y is a function of the **proportions** of the ingredients in the mixture and **not** of the amounts of the ingredients

$$Y = F(X_1, X_2, X_3, ...X_p) + \varepsilon$$

- Response Y: octane rating of gasoline, crushing strength of a tablet, smoothness of a cream, ...

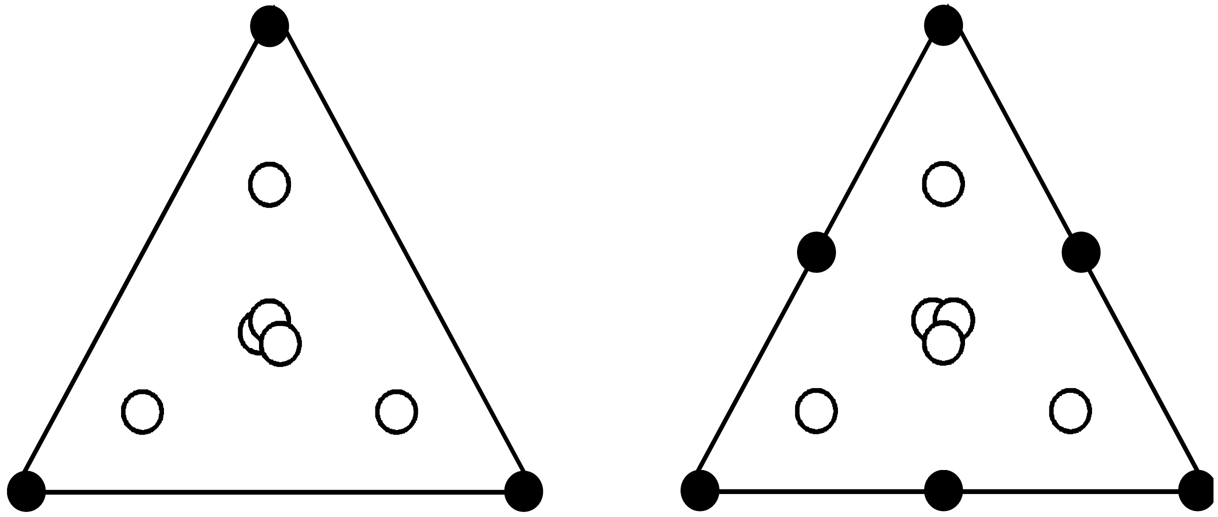
- The response depends only on the relative proportions of the ingredients of the mixture

$$\Sigma X_k = 1$$

- We can express the relative proportions as fractions or percentages

UMETRICS
AN MKS COMPANY

Design of Mixture Experiments



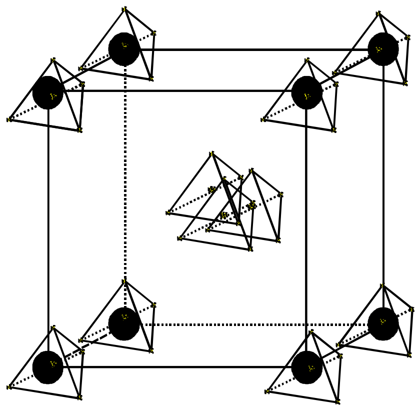
Linear Quadratic

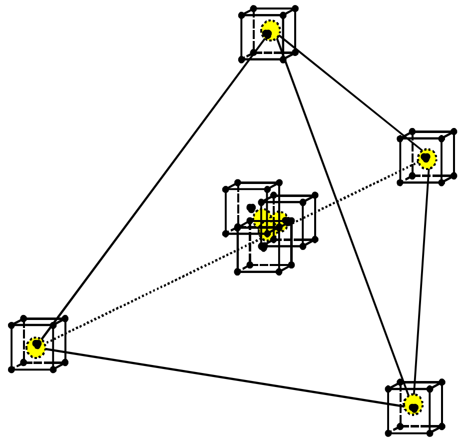
- Experimental domain is a simplex (or polyhedron)

- Experimental region has dimensionality k-1, where k is the number of mixture factors

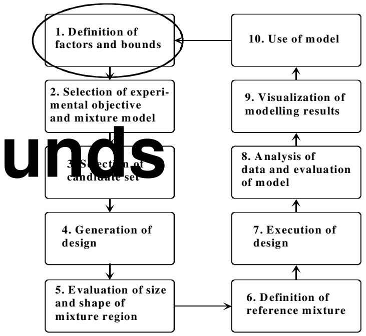
Process and mixture factors together

Process and Mixture Factors





Example: Tablet- Definition of factors and bounds



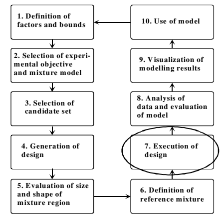
- **Aim**: To investigate tablet preparation and find out which factors that regulate the release rate of an active substance

- **Mixture Factors**:
  - Cellulose (0 - 1)
  - Lactose (0 - 1)
  - Phosphate (0 - 1)
  - All factors sum to 100% (mixture constraint)
  - Bounds display consistency

- **Constraint**:
  - No other extra constraint

- **Response**:
  - Release rate of the active substance (to be maximized)

Tablet: Execution of design

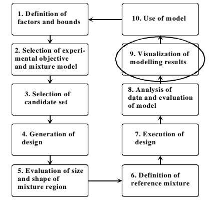


- Important to carry out experiments in random order
- This is done in order to break down any systematic time trend to become a non-important and random unsystematic variation

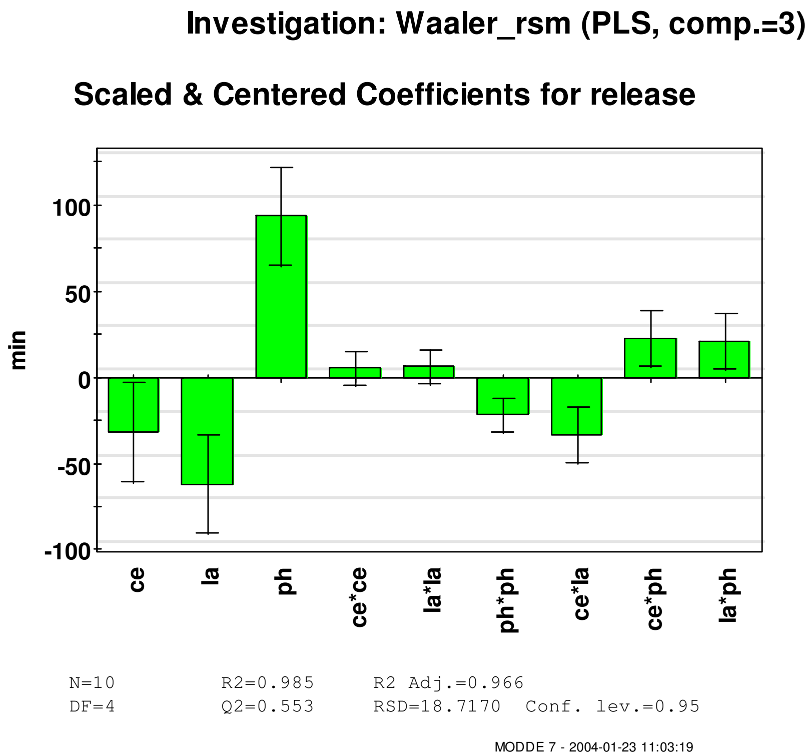
Worksheet

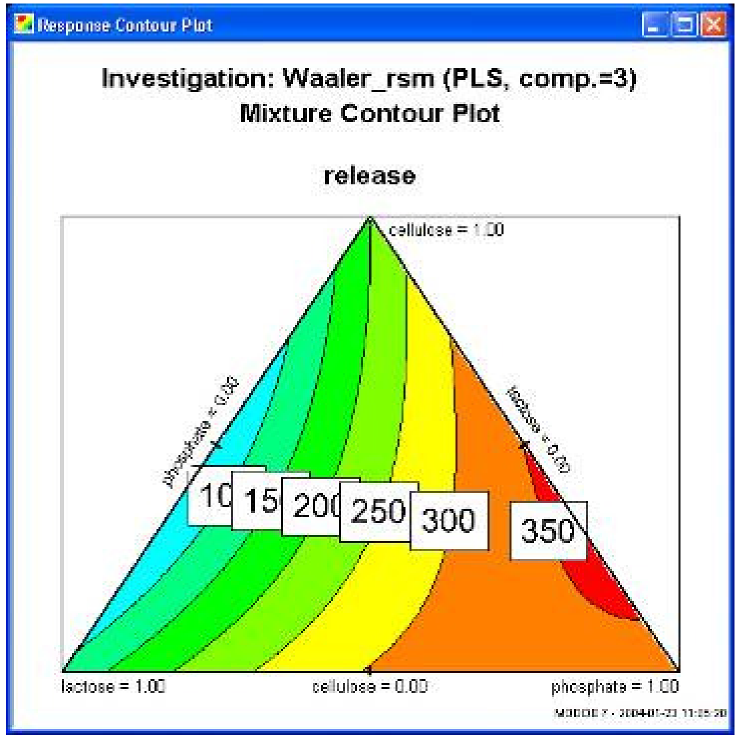
	1	2	3	4	5	6	7	8
1	Exp No	Exp Name	Run Order	Incl/Excl	cellulose	lactose	phosphate	release
2	1	N1	10	Incl	1	0	0	197
3	2	N2	7	Incl	0	1	0	110
4	3	N3	4	Incl	0	0	1	324
5	4	N4	9	Incl	0.5	0.5	0	67
6	5	N5	2	Incl	0.5	0	0.5	362
7	6	N6	6	Incl	0	0.5	0.5	312
8	7	N7	1	Incl	0.666667	0.166667	0.166667	206
9	8	N8	3	Incl	0.166667	0.666667	0.166667	171
10	9	N9	8	Incl	0.166667	0.166667	0.666667	344
11	10	N10	5	Incl	0.333333	0.333333	0.333333	214

UMETRICS AN MKS COMPANY



Tablet: Visualization of modelling results





Regression coefficients

Tri-linear contour plot

D-OPTIMAL DESIGNS

Introduction to D-optimal design

- A D-optimal design is a computer generated selection design
  - Requires candidate set

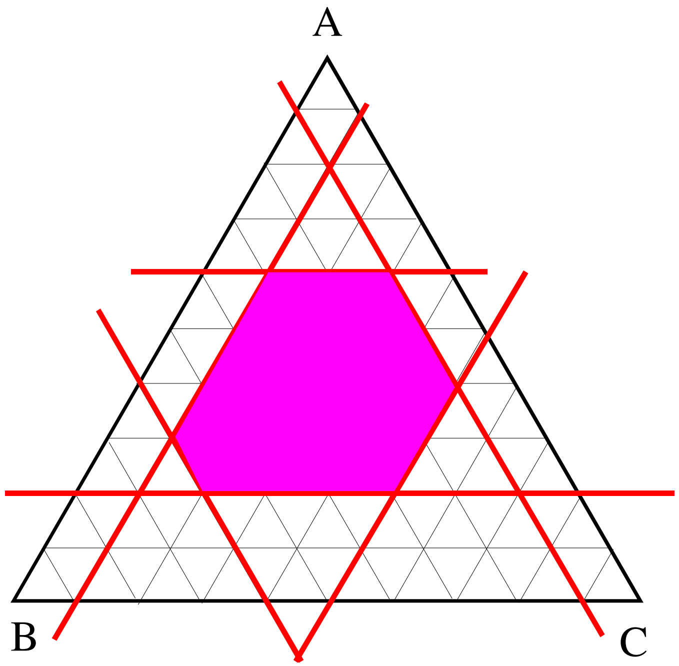
- A D-optimal design can be tailored to support an irregular experimental region, or a very complex problem set-up (process + mixture)

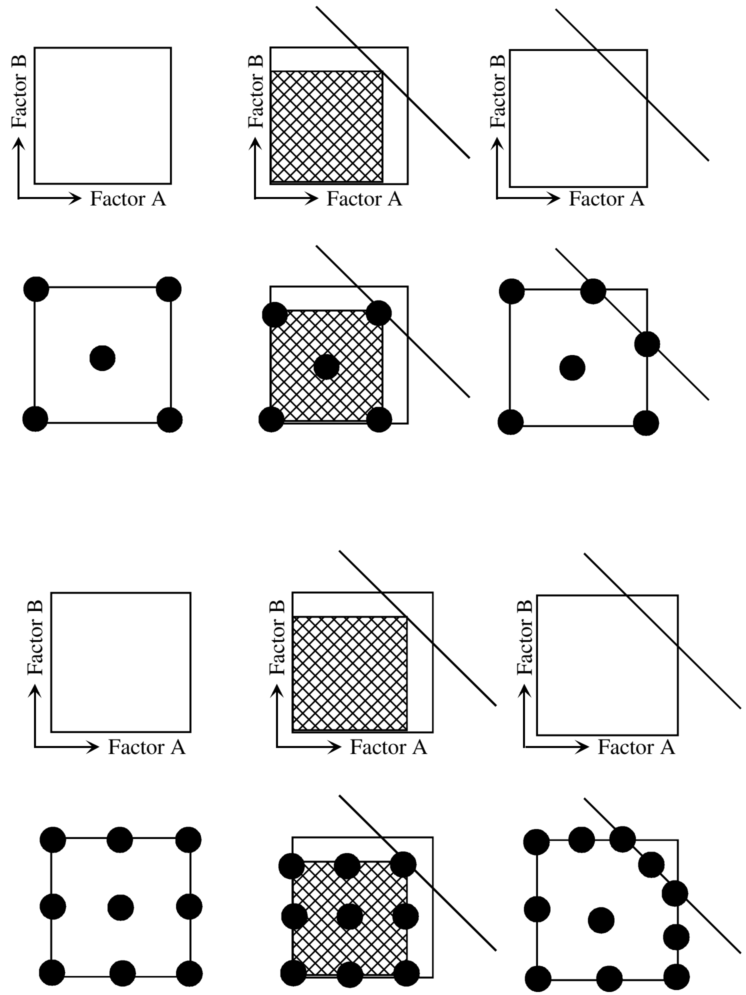
- For a given model, $Y = X\beta + \varepsilon$, the following can be said regarding the D-optimal approach:
  - the selected runs maximize the determinant of the matrix X'X
  - these experiments span the largest volume possible in the experimental region

UMETRICS
AN MKS COMPANY

When to use D-optimal design - Irregular regions

- Irregular experimental region in
  - screening
  - optimization
  - mixture design

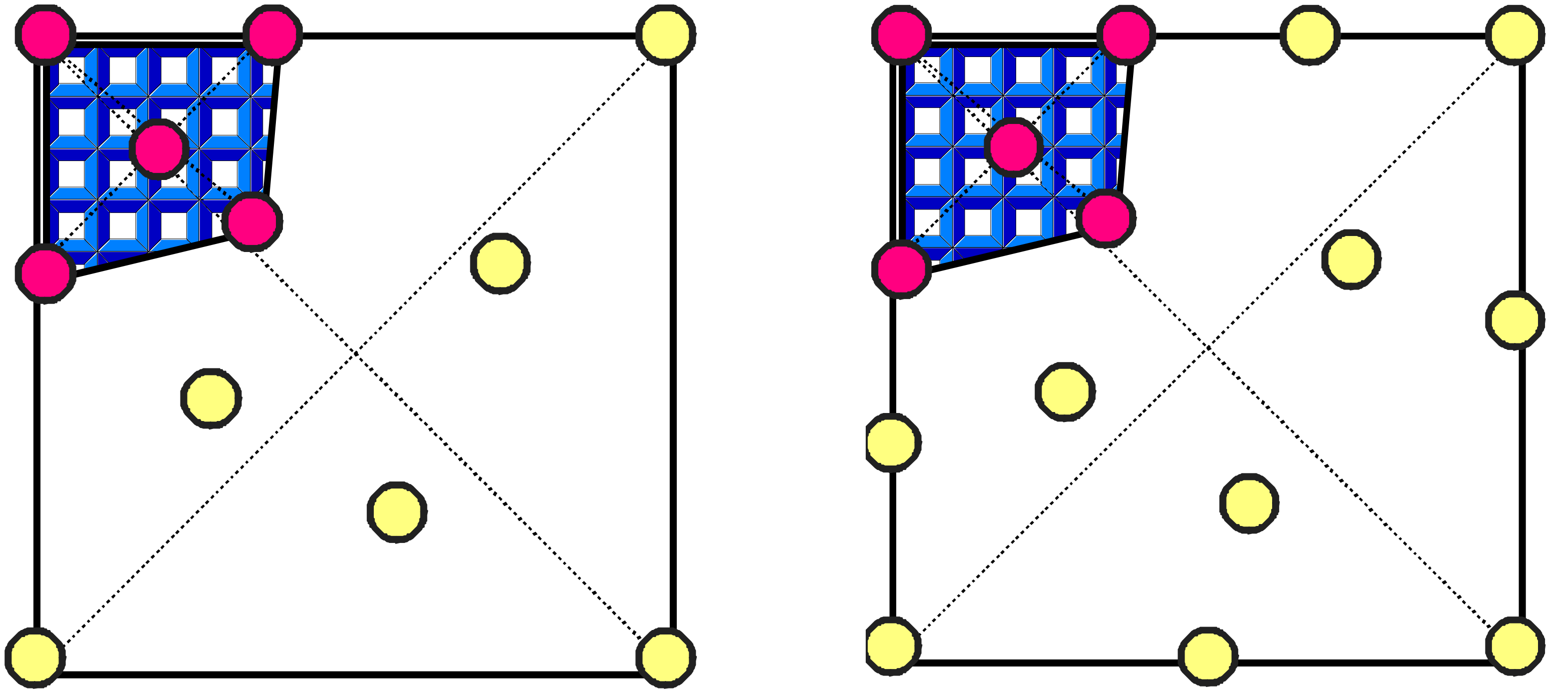




When to use D-optimal design - Inclusions

• Inclusions of existing experimental information

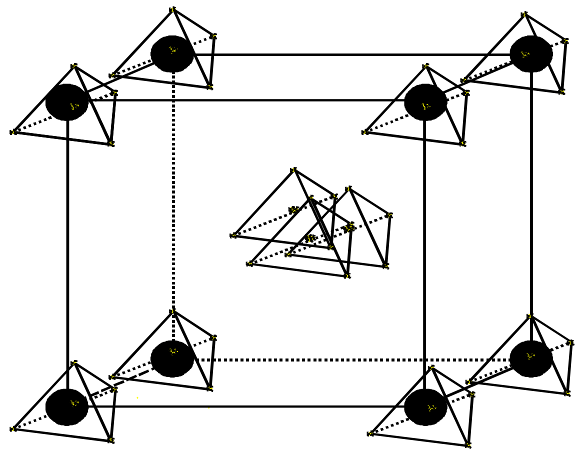
- screening
- optimization

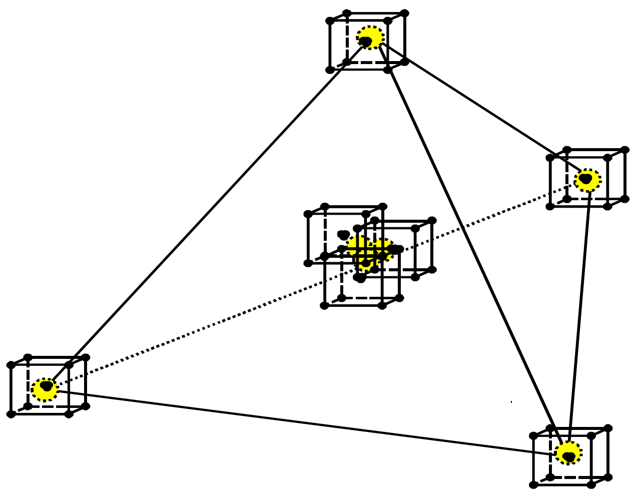


When to use D-opt. design - Process and Mixture Factors

- When making a combined design for process and mixture factors

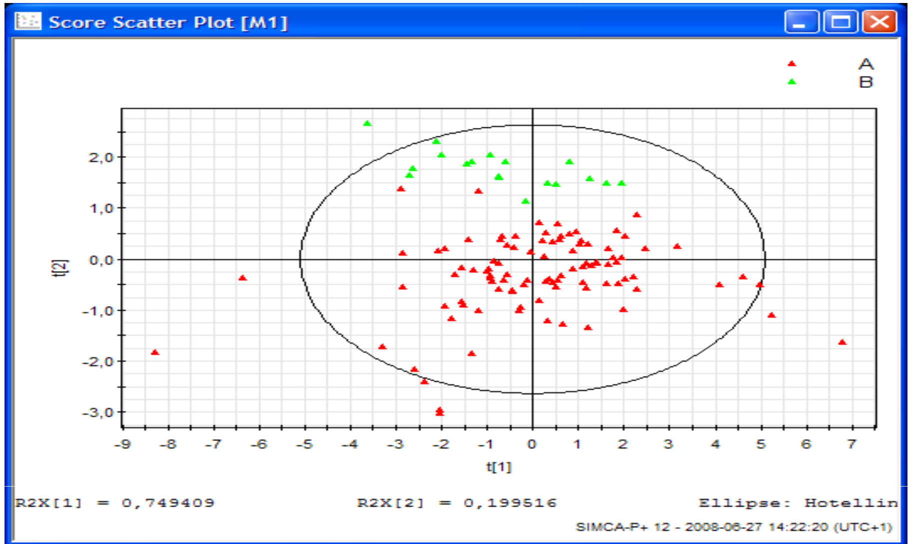
- LoafVolume is a typical example where D-optimal design could have been utilized

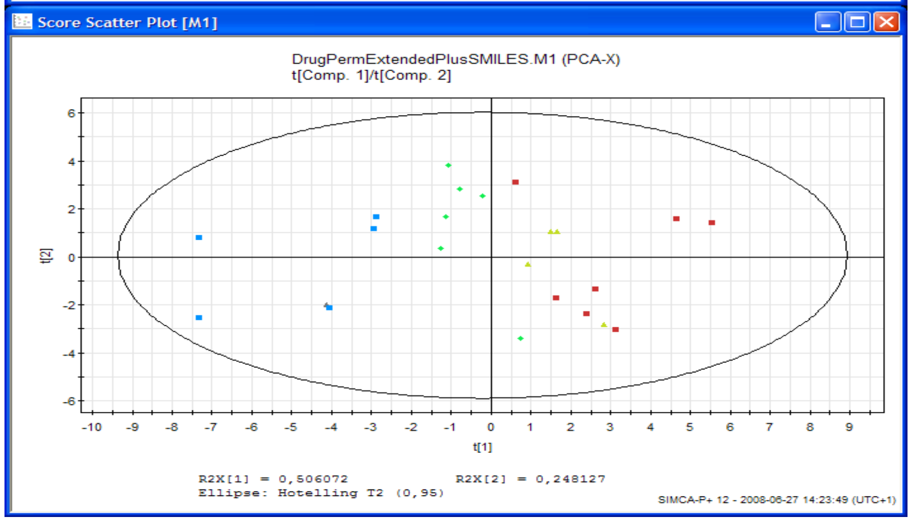




When to use D-opt. design – multivariate design

- Selection from non-continuous factors
  - Score space
    - Molecular descriptors
    - Spectral data
    - Human data
- Select representative sub set from Historical data bases
  - Spectral data
  - Process data
- Some things cannot be designed
  - Humans
  - Animals
  - Molecules
  - Other forces of nature





Summary - Key features of DOE

- How to make experiments efficiently
  – Span the experimental domain with the aid of a suitable experimental design
- How to analyze the data
  – Use good statistical tools to evaluate experimental results
- How to interpret the results
  – With the use of user-friendly PC-based graphical facilities
- How to convert modelling results into concrete actions/decisions
  – MODDE optimizer & verifying experiments

UMETRICS
AN MKS COMPANY

Applications and application specific design types

RED-MUP:

DOE FOR MULTI WELL PLATES (96, 384, 1536)

Introduction

- Traditional DOE (G E P Box) implies ***minimization of experiments***
- DOE in the 21st century implies ***minimization of work*** by using robots and automatization of experiments
- Rectangular Experimental Design for Multi-Unit Platforms (RED-MUP)
- A RED-MUP design consists of two partial designs (the vertical and horizontal designs) multiplied together.

Example: 96 well plates

- Basic idea: Factors varied in complete rows and complete columns

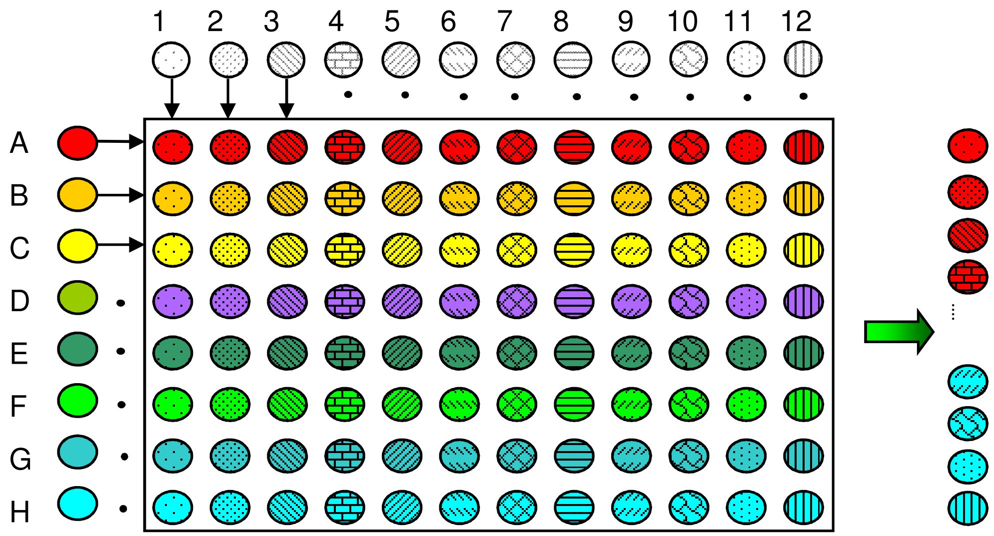
	1	2	3	4	5	6	7	8	9	10	11	12
A	1 2 3	1 2 3	1 2 3	1 2 3	1 2 3	1 2 3	1 2 3	1 2 3	1 2 3	1 2 3	1 2 3	1 2 3
	4 5 6	44 5 6	4 55 6	44 5 66	4 55 66	44 55 66	4 5 6	4 55 6	44 55 6	4 5 66	44 5 66	44 55 66
B	11 2 3	11 2 3	11 2 3	11 2 3	11 2 3	11 2 3	11 2 3	11 2 3	11 2 3	11 2 3	11 2 3	11 2 3
	4 5 6	44 5 6	4 55 6	44 5 66	4 55 66	44 55 66	4 5 6	4 55 6	44 55 6	4 5 66	44 5 66	44 55 66
C	1 22 3	1 22 3	1 22 3	1 22 3	1 22 3	1 22 3	1 22 3	1 22 3	1 22 3	1 22 3	1 22 3	1 22 3
	4 5 6	44 5 6	4 55 6	44 5 66	4 55 66	44 55 66	4 5 6	4 55 6	44 55 6	4 5 66	44 5 66	44 55 66
D	11 22 3	11 22 3	11 22 3	11 22 3	11 22 3	11 22 3	11 22 3	11 22 3	11 22 3	11 22 3	11 22 3	11 22 3
	4 5 6	44 5 6	4 55 6	44 5 66	4 55 66	44 55 66	4 5 6	4 55 6	44 55 6	4 5 66	44 5 66	44 55 66
E	1 2 33	1 2 33	1 2 33	1 2 33	1 2 33	1 2 33	1 2 33	1 2 33	1 2 33	1 2 33	1 2 33	1 2 33
	4 5 6	44 5 6	4 55 6	44 5 66	4 55 66	44 55 66	4 5 6	4 55 6	44 55 6	4 5 66	44 5 66	44 55 66
F	11 2 33	11 2 33	11 2 33	11 2 33	11 2 33	11 2 33	11 2 33	11 2 33	11 2 33	11 2 33	11 2 33	11 2 33
	4 5 6	44 5 6	4 55 6	44 5 66	4 55 66	44 55 66	4 5 6	4 55 6	44 55 6	4 5 66	44 5 66	44 55 66
G	1 22 33	1 22 33	1 22 33	1 22 33	1 22 33	1 22 33	1 22 33	1 22 33	1 22 33	1 22 33	1 22 33	1 22 33
	4 5 6	44 5 6	4 55 6	44 5 66	4 55 66	44 55 66	4 5 6	4 55 6	44 55 6	4 5 66	44 5 66	44 55 66
H	11 22 33	11 22 33	11 22 33	11 22 33	11 22 33	11 22 33	11 22 33	11 22 33	11 22 33	11 22 33	11 22 33	11 22 33
	4 5 6	44 5 6	4 55 6	44 5 66	4 55 66	44 55 66	4 5 6	4 55 6	44 55 6	4 5 66	44 5 66	44 55 66

Factors	Low	High	
X1	1	11	Row-wise
X2	2	22	Row-wise
X3	3	33	Row-wise
X4	4	44	Column-wise
X5	5	55	Column-wise
X6	6	66	Column-wise
X7	7	77	Column-wise

Model terms : Linear, interaction
Cond. No. : 2

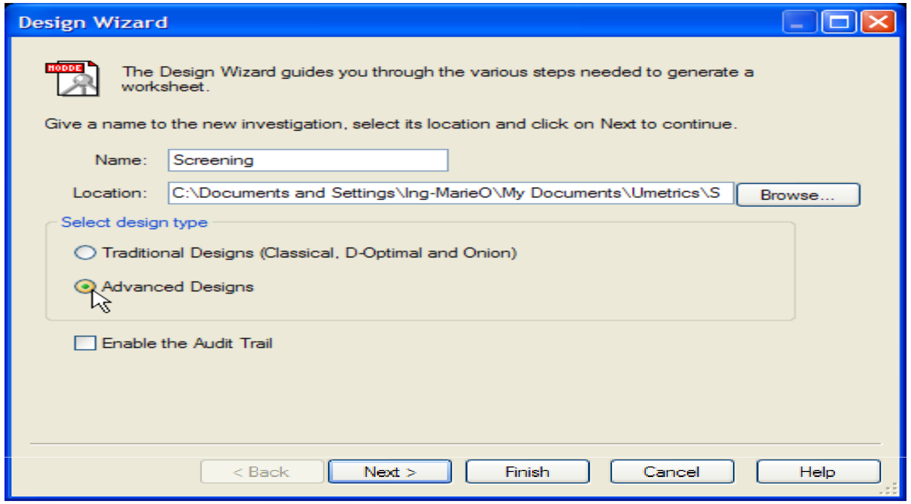
The RED-MUP approach

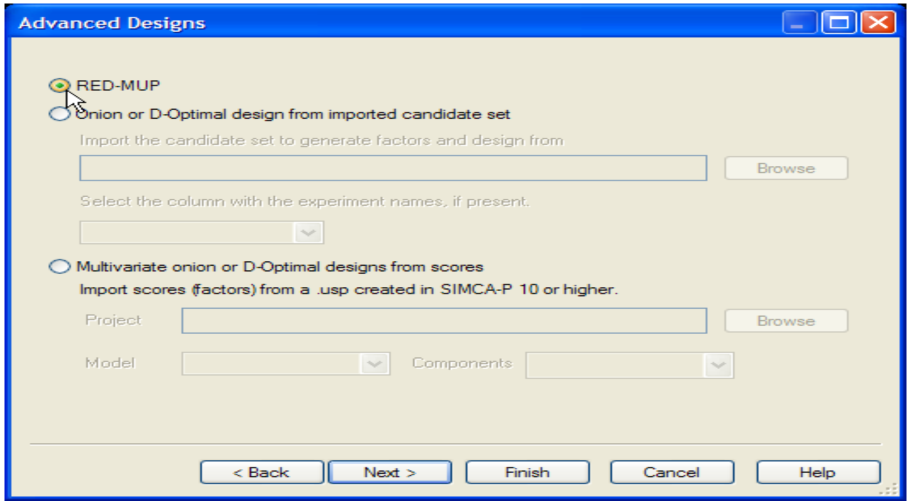
- Combination of two DOEs
- Limits workload and manual pipetting
    - Stem-solutions
- DOEs represent pipetting-schemes
- Result: Full matrix!



Creating a RED-MUP protocol

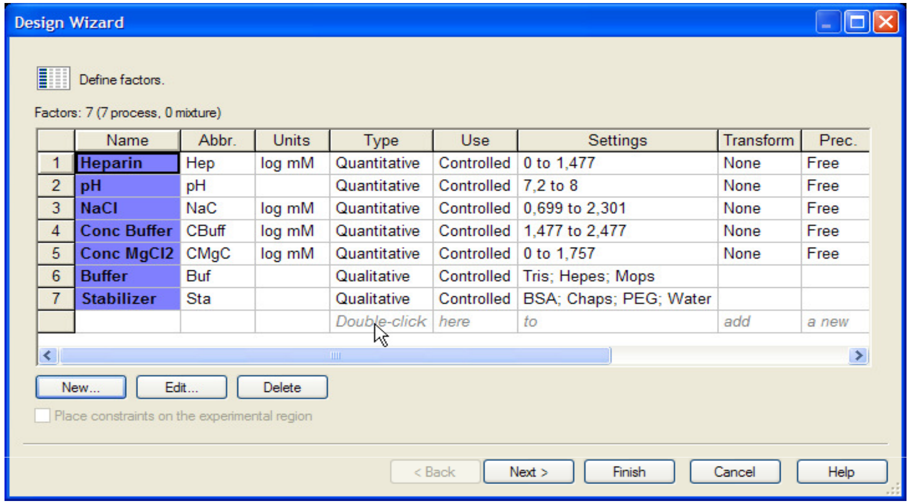
- Design wizard
- Step-by-step set-up
- Define project name
- Select ”Advanced Designs”
- Select RED-MUP





Factor and response definition

- Factor definition
  - Define experimental factors, i.e. pH, Concentrations of additives etc.
  - Define levels for the factors

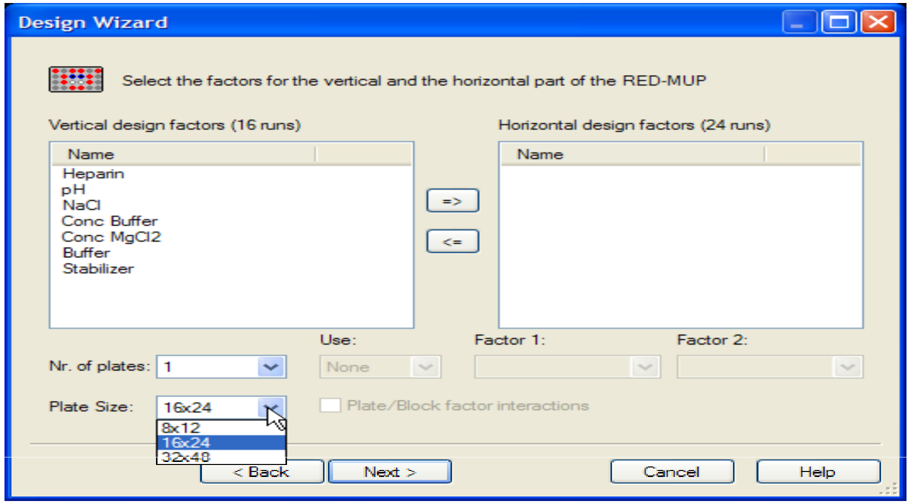


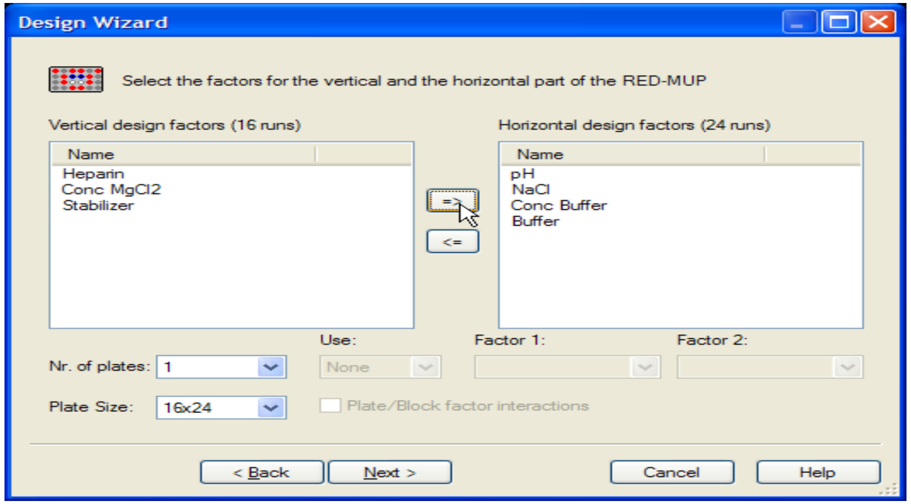
	Name	Abbr.	Units	Type	Use	Settings	Transform	Prec.
1	Heparin	Hep	log mM	Quantitative	Controlled	0 to 1,477	None	Free
2	pH	pH		Quantitative	Controlled	7,2 to 8	None	Free
3	NaCl	NaC	log mM	Quantitative	Controlled	0,699 to 2,301	None	Free
4	Conc Buffer	CBuff	log mM	Quantitative	Controlled	1,477 to 2,477	None	Free
5	Conc MgCl2	CMgC	log mM	Quantitative	Controlled	0 to 1,757	None	Free
6	Buffer	Buf		Qualitative	Controlled	Tris; Hepes; Mops		
7	Stabilizer	Sta		Qualitative	Controlled	BSA; Chaps; PEG; Water		
				Double-click	here	to	add	a new

- Response definition

Divide factors into two groups

- Define plate size
- Number of plates
- Blocked plates



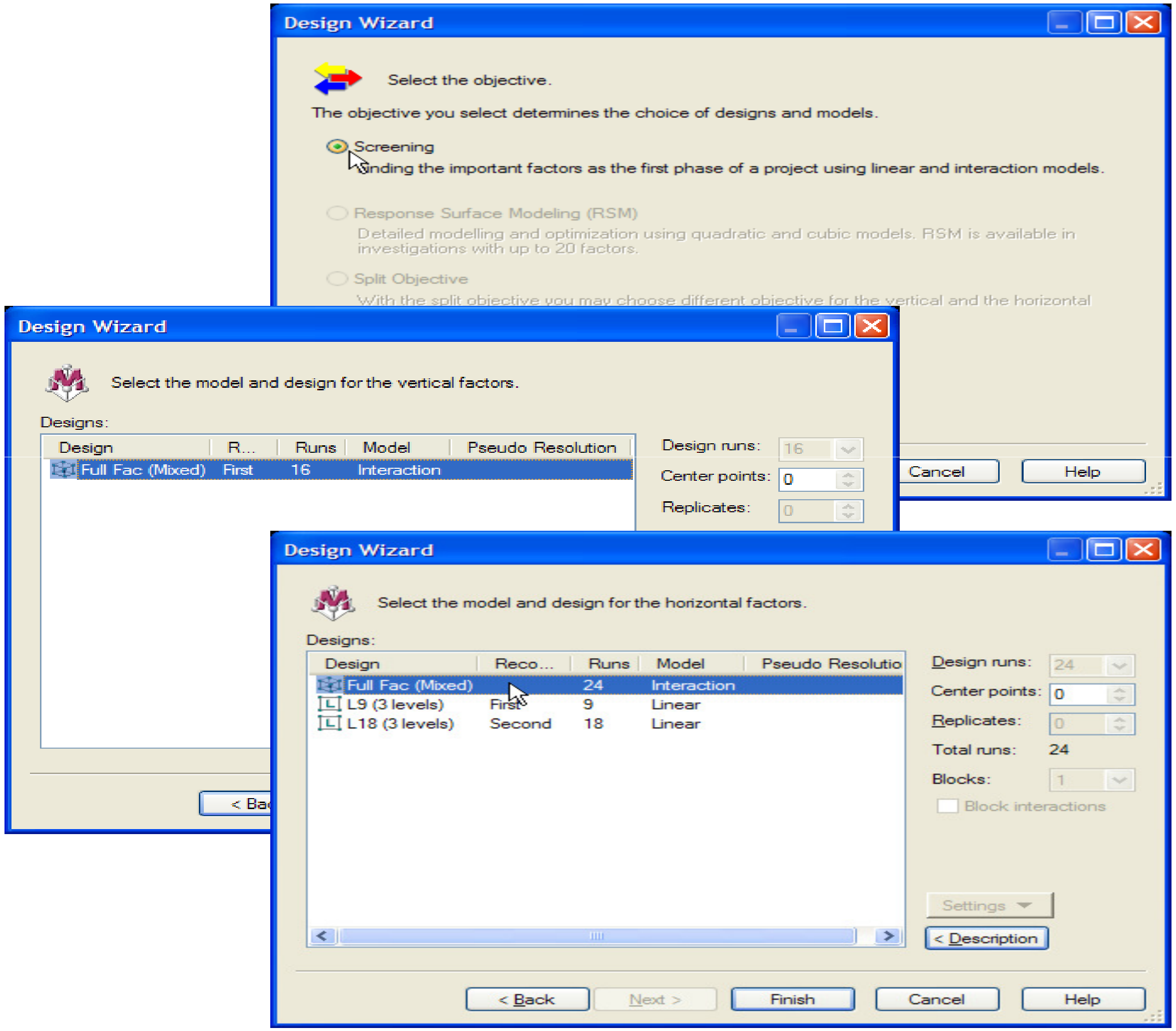


- Divide factors into vertical and horizontal design
  - Some factors have to be in the same design, such as Buffer and concentration of buffer

UMETRICS AN MKS COMPANY

Select designs

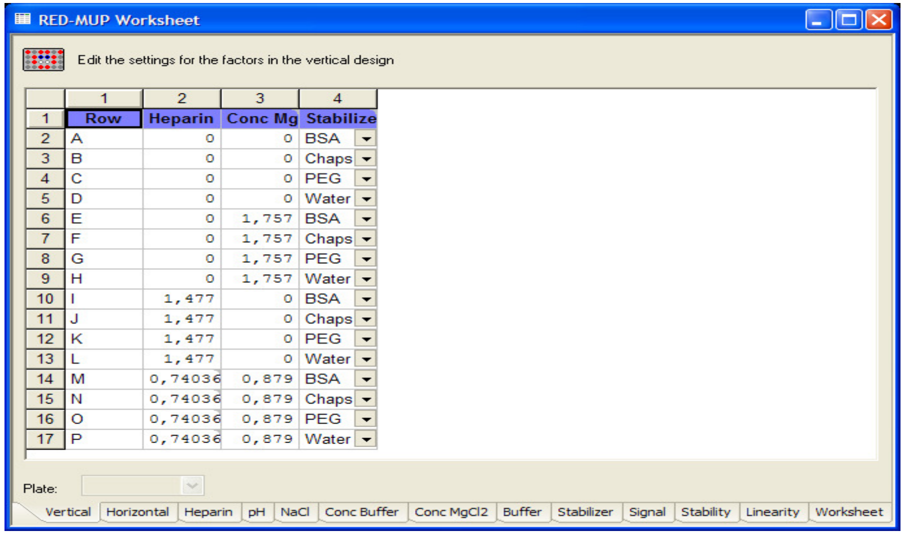
- Select objective
  - Screening
  - Optimization (RSM)
  - Split
- Select vertical design
- Select horizontal design



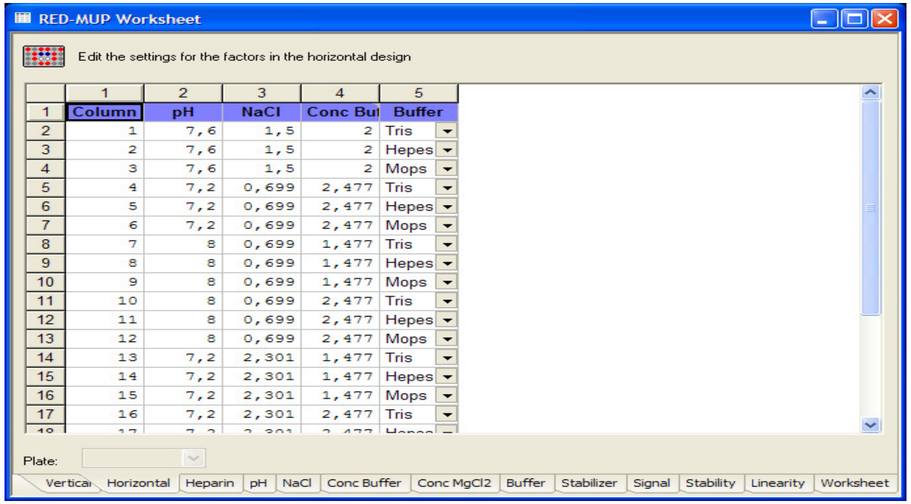
RED-MUP worksheet

- Consists of tabs for:
  - Vertical design
  - Horizontal design
  - Separate factors
  - Responses
  - Complete worksheet

- Complete worksheet is linked to tabs for separate designs and responses



	1	2	3	4
1	Row	Heparin	Conc Mg	Stabilize
2	A	0	0	BSA
3	B	0	0	Chaps
4	C	0	0	PEG
5	D	0	0	Water
6	E	0	1,757	BSA
7	F	0	1,757	Chaps
8	G	0	1,757	PEG
9	H	0	1,757	Water
10	I	1,477	0	BSA
11	J	1,477	0	Chaps
12	K	1,477	0	PEG
13	L	1,477	0	Water
14	M	0,74036	0,879	BSA
15	N	0,74036	0,879	Chaps
16	O	0,74036	0,879	PEG
17	P	0,74036	0,879	Water



	1	2	3	4	5
1	Column	pH	NaCl	Conc Bu	Buffer
2	1	7,6	1,5	2	Tris
3	2	7,6	1,5	2	Hepes
4	3	7,6	1,5	2	Mops
5	4	7,2	0,699	2,477	Tris
6	5	7,2	0,699	2,477	Hepes
7	6	7,2	0,699	2,477	Mops
8	7	8	0,699	1,477	Tris
9	8	8	0,699	1,477	Hepes
10	9	8	0,699	1,477	Mops
11	10	8	0,699	2,477	Tris
12	11	8	0,699	2,477	Hepes
13	12	8	0,699	2,477	Mops
14	13	7,2	2,301	1,477	Tris
15	14	7,2	2,301	1,477	Hepes
16	15	7,2	2,301	1,477	Mops
17	16	7,2	2,301	2,477	Tris

Response values

- Response values can be inserted in plate format (8x12, 16x24 or 32x48 matrix) in the separate tabs or:

RED-MUP Worksheet

Edit the settings for the response Signal

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A	38,1	28,8	23,3	24,9	29,5	19,5	59,2	36,1	34,9	32,6	20,9	10,4	25,7	19,1	17,3	17,5	14,2	13,4	36,1	20,6	17,7	20,3	12,1	17,7
B	51,5	35,7	29,6	38,3	33,5	23,9	68,6	52,7	46,3	51,2	22	16,4	29,3	27,8	24,6	28,4	25,2	21,7	38,6	36	28,4	35,2	17,6	28,1
C	39,1	42,4	35,2	42,8	46	29	88,4	59,7	59,2	61,4	28,7	20,9	38,2	30,7	27,4	26,2	26,8	15,2	47,4	35,9	41,1	53,9	20,7	21,6
D	60,5	25,9	24	34,1	21,5	23	66	53,5	59,9	40	21,2	10,6	36,3	19,4	17,4	17,7	15,1	12,9	28,9	22,1	20,2	25,3	17,4	32,2
E	35,5	31,2	23,4	31,4	26,7	20,6	45,8	37,1	27,5	29,6	24	17,4	20,5	26	20,7	17,2	19,9	7,2	26,5	22,7	19,1	15,2	10,9	18,8
F	44	30,3	31,8	36,1	28,2	21,5	53,3	39,1	42,2	45,3	22,9	19,1	29,9	26,9	23,6	26,9	23,2	18,7	35,2	30,1	27,3	29,3	17,7	24,4
G	56,5	38,6	32,9	40,5	33,6	26,2	67,7	45,3	43,5	54,4	26	16,6	35,6	28,4	26,4	29,8	27,4	22,4	46,6	34,2	28,3	34,9	21,9	31
H	45,6	26,2	30,5	41,1	29,6	24,7	60,1	46,1	34,8	53,4	27,2	15,9	30,3	15,6	23,2	19,3	17,8	16,5	23,1	31,2	22,4	23,8	15,9	18,1
I	46,3	34,8	32	31,5	31,2	25,9	66	56,6	31	24,9	33,6	14,4	25,1	22,4	31,2	22,4	18,7	12,3	27,8	21,8	18,3	23,5	11,1	19,5
J	54,4	39	37,9	39,3	38,1	24,6	78,6	56,5	44,7	54	26,3	18,3	64,5	25,2	27,6	29,1	25,4	17,7	43,2	31,5	26,5	37,8	21,9	30,6
K	66,8	46,8	40,3	48,5	46	30,5	85,6	60,7	52,3	66,1	37,4	24,6	40,4	29	34,7	33,9	31,2	24,7	58,2	43,2	36,4	42,8	26,4	38,1
L	74	43,4	37,9	42,1	49,8	29	100	63,1	48,4	57,2	28,6	17,4	44,6	29,1	26,1	32,6	34,2	22,8	49,8	33,1	30	29,3	17,3	31,3
M	39,9	42,7	21,4	36,6	34,2	25,9	49,9	38,2	23,3	31,5	26,9	17,6	41,9	26,6	24	30,2	22,1	6,9	34,2	13,8	17,8	26,8	15,4	19,8
N	68,8	42,9	34	41	35	26,9	60,8	47,9	38,7	46,2	29,6	17,4	34,7	34,5	26,5	34,3	29,1	24,1	43,7	34,9	27,9	40,4	21,4	30,6
O	40,9	56	37,6	44,7	46,3	35,3	65,9	58,6	44,3	65	38,2	22	39,2	31,7	31,7	47,9	41,2	27	50,9	44,7	38	51,7	27,3	33,8
P	57,5	43,5	35,8	38,4	41,8	27	84,9	40,3	52,9	75,2	36,8	24,1	48,3	31,4	24,2	34,8	25,8	22,9	41,1	34,7	29,4	36,5	21	25,3

Plate:

Vertical | Horizontal | Heparin | pH | NaCl | Conc Buffer | Conc MgCl2 | Buffer | Stabilizer | Signal | Stability | Linearity | Worksheet

- As a vector in the complete worksheet

RED-MUP Worksheet

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Exp No	Exp Name	Run Order	Incl/Excl	Heparin	pH	NaCl	Conc Buffer	Conc MgCl2	Buffer	Stabilizer	Signal	Stability	Linearity
2	1	N1	1	Incl	0	7,6	1,5	2	0	Tris	BSA	38,1	94,4949	82,9924
3	2	N2	2	Incl	0	7,6	1,5	2	0	Tris	Chaps	51,5	90,0937	87,6649
4	3	N3	3	Incl	0	7,6	1,5	2	0	Tris	PEG	39,1	178,9	108,592
5	4	N4	4	Incl	0	7,6	1,5	2	0	Tris	Water	60,5	78,5617	85,5993
6	5	N5	5	Incl	0	7,6	1,5	2	1,757	Tris	BSA	35,5	116,948	120,938
7	6	N6	6	Incl	0	7,6	1,5	2	1,757	Tris	Chaps	44	88,0979	83,1213
8	7	N7	7	Incl	0	7,6	1,5	2	1,757	Tris	PEG	56,5	132,553	84,8036
9	8	N8	8	Incl	0	7,6	1,5	2	1,757	Tris	Water	45,6	32,9288	88,4555
10	9	N9	9	Incl	1,477	7,6	1,5	2	0	Tris	BSA	46,3	109,135	86,6911
11	10	N10	10	Incl	1,477	7,6	1,5	2	0	Tris	Chaps	54,4	96,2662	85,6684
12	11	N11	11	Incl	1,477	7,6	1,5	2	0	Tris	PEG	66,8	115,29	80,4411
13	12	N12	12	Incl	1,477	7,6	1,5	2	0	Tris	Water	74	43,793	76,8159
14	13	N13	13	Incl	0,740363	7,6	1,5	2	0,879	Tris	BSA	39,9	98,9043	87,1128
15	14	N14	14	Incl	0,740363	7,6	1,5	2	0,879	Tris	Chaps	68,8	61,5928	75,3273
16	15	N15	15	Incl	0,740363	7,6	1,5	2	0,879	Tris	PEG	40,9	110,214	132,314
17	16	N16	16	Incl	0,740363	7,6	1,5	2	0,879	Tris	Water	57,5	53,7963	81,1424
18	17	N17	17	Incl	0	7,6	1,5	2	0	Hepes	BSA	28,8	91,4942	83,118
19	18	N18	18	Incl	0	7,6	1,5	2	0	Hepes	Chaps	35,7	95,0723	91,7244
20	19	N19	19	Incl	0	7,6	1,5	2	0	Hepes	PEG	42,4	116,208	94,4676
21	20	N20	20	Incl	0	7,6	1,5	2	0	Hepes	Water	25,9	130,057	67,629
22	21	N21	21	Incl	0	7,6	1,5	2	1,757	Hepes	BSA	31,2	66,5368	83,6654

Vertical | Horizontal | Heparin | pH | NaCl | Conc Buffer | Conc MgCl2 | Buffer | Stabilizer | Signal | Stability | Linearity | Worksheet

UMETRICS
AN MKS COMPANY

Example: Enzyme activity assay

- Master thesis: Per Rosén, Gothenburg Univeristy

- Increase signal
- Increase stability
- Increase linearity

- 3-step process
  - Screening:
    - Identify important factors
  - Optimization:
    - Identify optimal settings for factors
    - Identify optimal experimental region
  - Robustness testing:
    - Find the most critical, i.e. sensitive factors
    - Find less sensitive factors

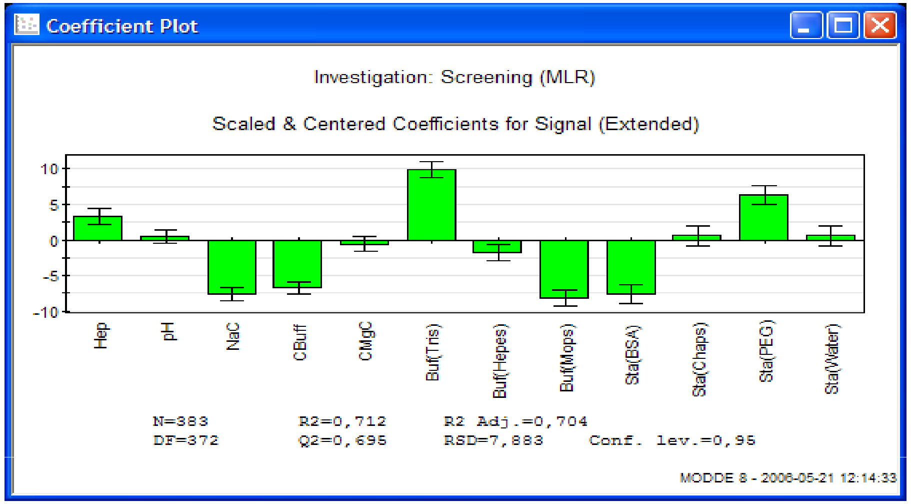
- 384-well microtiter plates

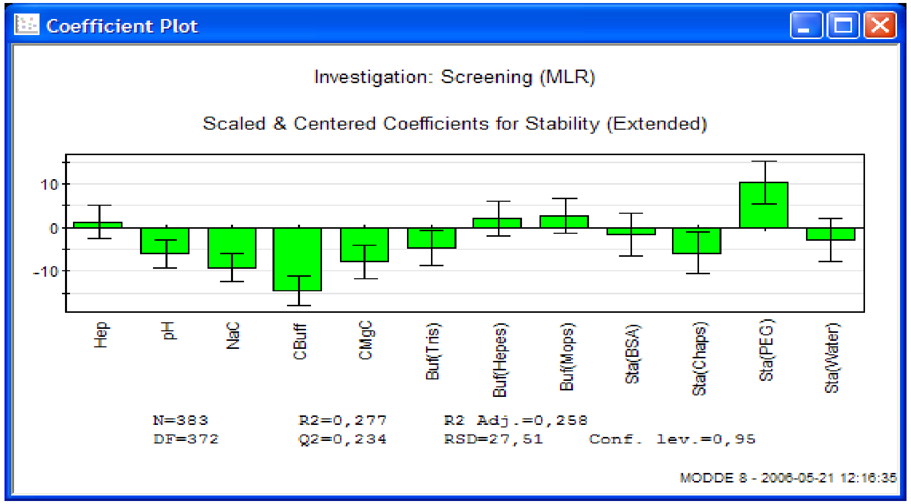
Screening

- Qualitative factors
  - Stabilizer (v)
  - Buffer (h)
- Quantitative factors
  - Conc Heparin (v)
  - Conc NaCl (h)
  - pH (h)
  - Conc Buffer (h)
  - Conc $MgCl_2$ (v)

- Vertical design (16 exp): reduced 2^2 x 4
- Horizontal design (24 exp): reduced 2^3 x 3

- Buffer- Tris
- Stabilizer- PEG

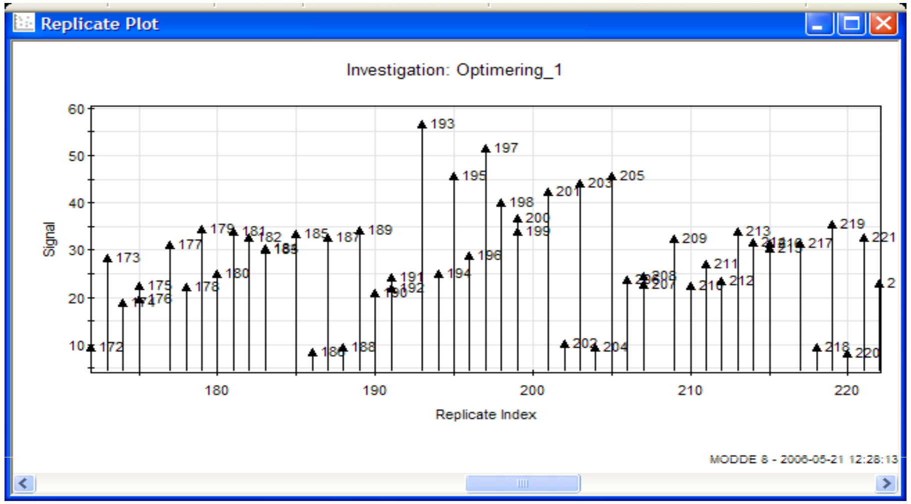


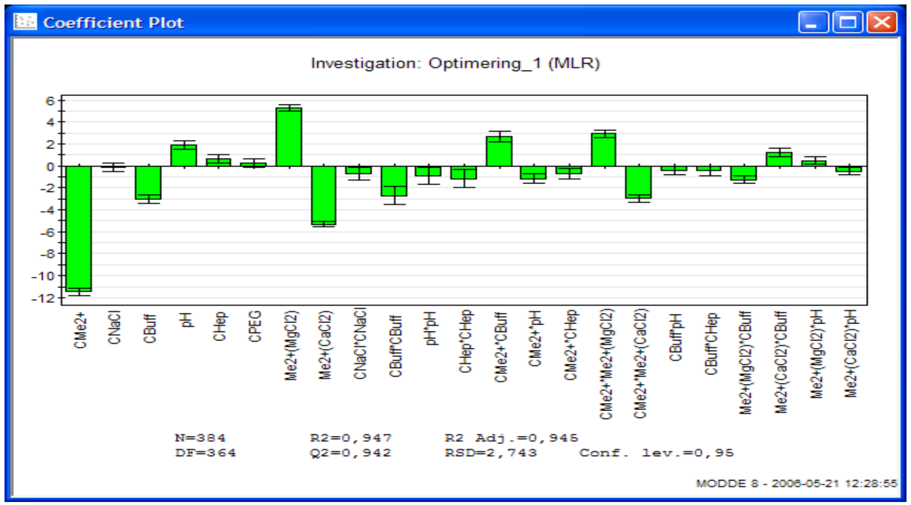


UMETRICS
AN MKS COMPANY

Optimization

- Qualitative factor
  - Type Me^{2+} (v)
- Quantitative factors
  - Conc Me^{2+} (v)
  - Conc NaCl (v)
  - Conc Buffer (h)
  - pH (h)
  - Conc Heparin (h)
  - Conc PEG (h)

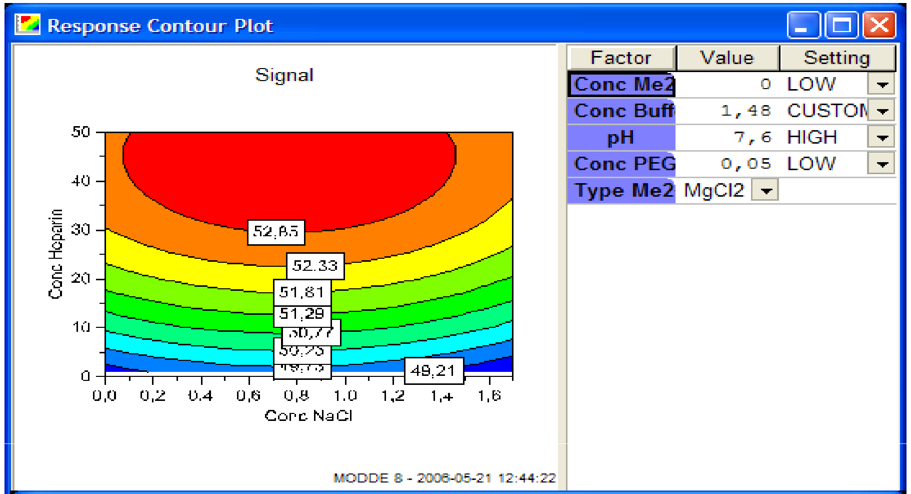


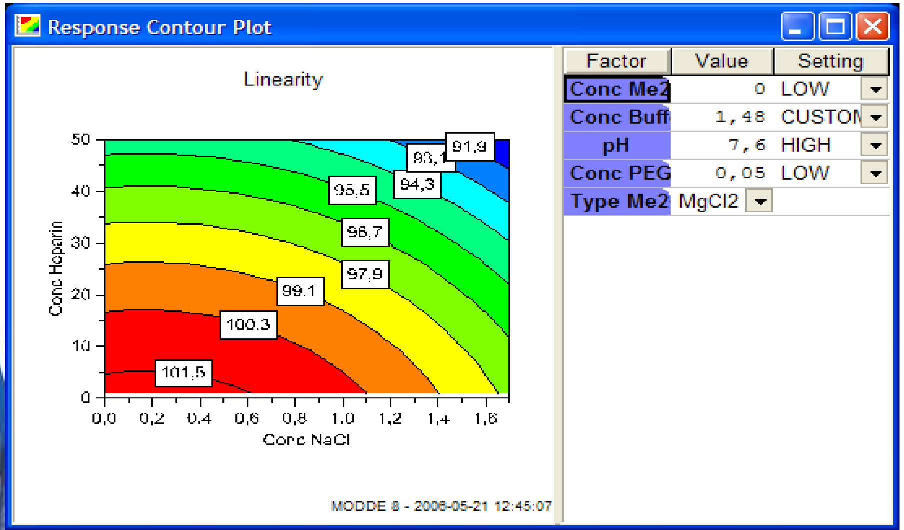


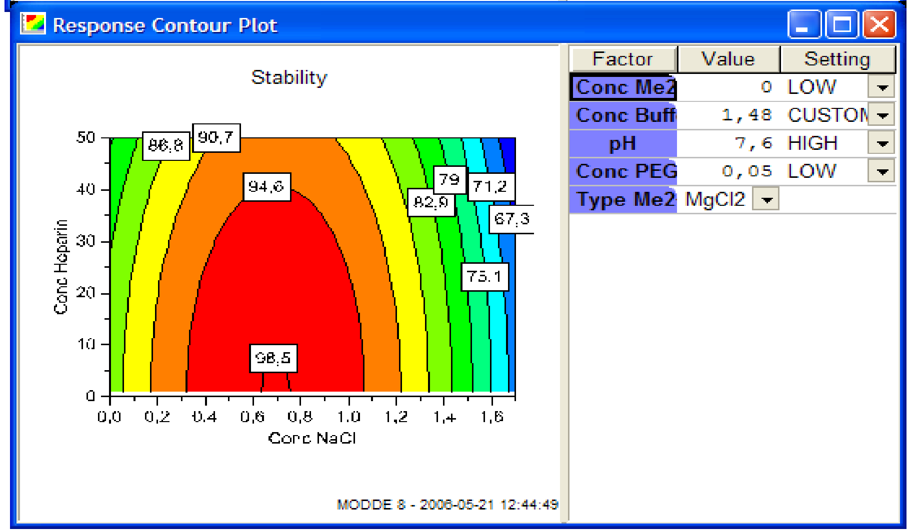
- Vertical design: reduced CCF (CCF-2 axial points) x 2
- Horizontal design: reduced CCF (Factorial reduced)

Results- Optimization

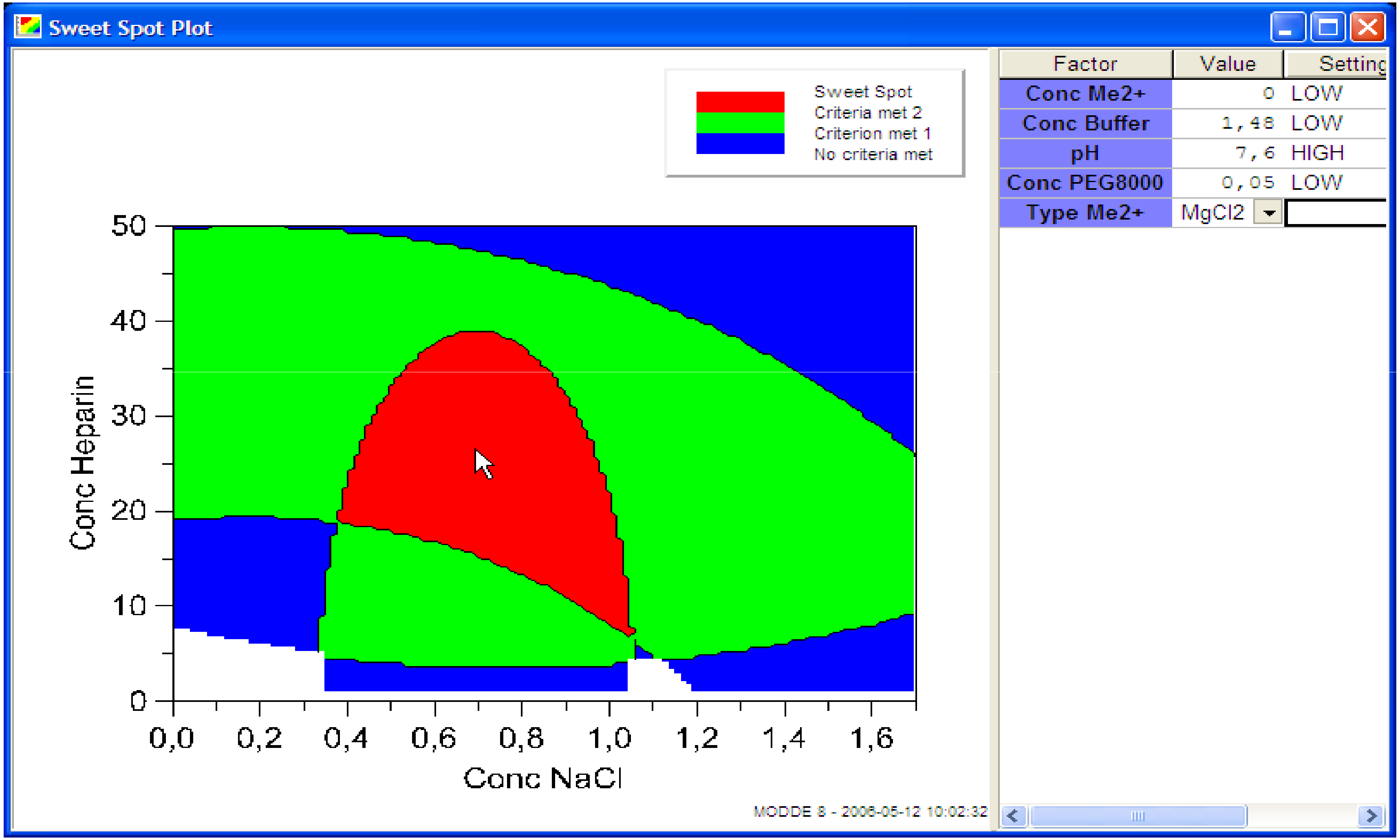
- Generally good models
- Signal: R^2= 0.94, Q^2= 0.94
- Stability: R^2= 0.60, Q^2= 0.55
- Me^{2+}: $MgCl_2$







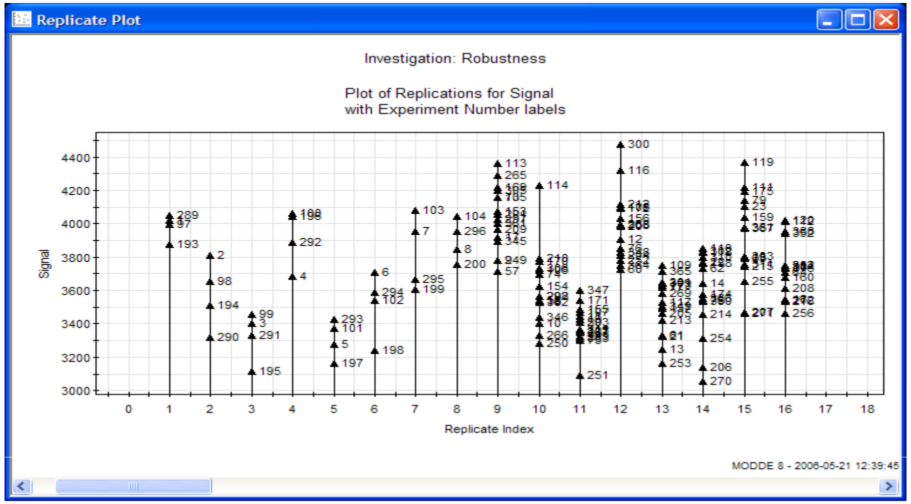
Sweet spot plot

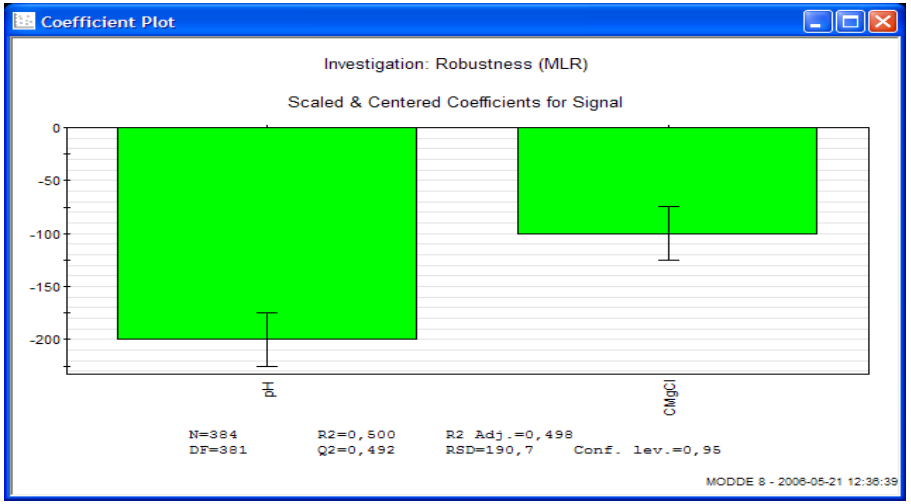


Factor	Value	Setting
Conc Me2+	0	LOW
Conc Buffer	1,48	LOW
pH	7,6	HIGH
Conc PEG8000	0,05	LOW
Type Me2+	MgCl2	

Robustness testing

- Factors
  - pH (v)
  - Conc $MgCl_2$ (v)
  - Conc PEG (h)
  - Conc Heparin (h)
- 4 x 96 well design
- Vertical design: CCF -1 axial point
- Horizontal design: CCF
- Two factors are sensitive to changes
  - pH
  - Conc $MgCl_2$





Results- Enzyme activity assay

- Comparison between optimized assay conditions and start conditions showed:
  - Almost 4 times increase of signal strength
  - The Stability, or decrease of enzyme activity, was improved from -52 % to -17 %, measured over a 23 h time period.
  - Deviation from linearity was decreased from 43 % to 11 %

Conclusions

- DoE is a valuable tool for assay optimization
  - More reliable data- better decisions, ”right” decisions
- RED-MUP minimizes workload and time required to perform DoE in rectangular formats
- Biological systems are difficult!
  - Need to understand how they are working- DoE necessary
  - Edge-effects, cross-talk
- Planning necessary
  - Practicality is important in the laboratory
    - Minimize experimental mistakes
- Approach applicable to many other applications where multi-well plates are used

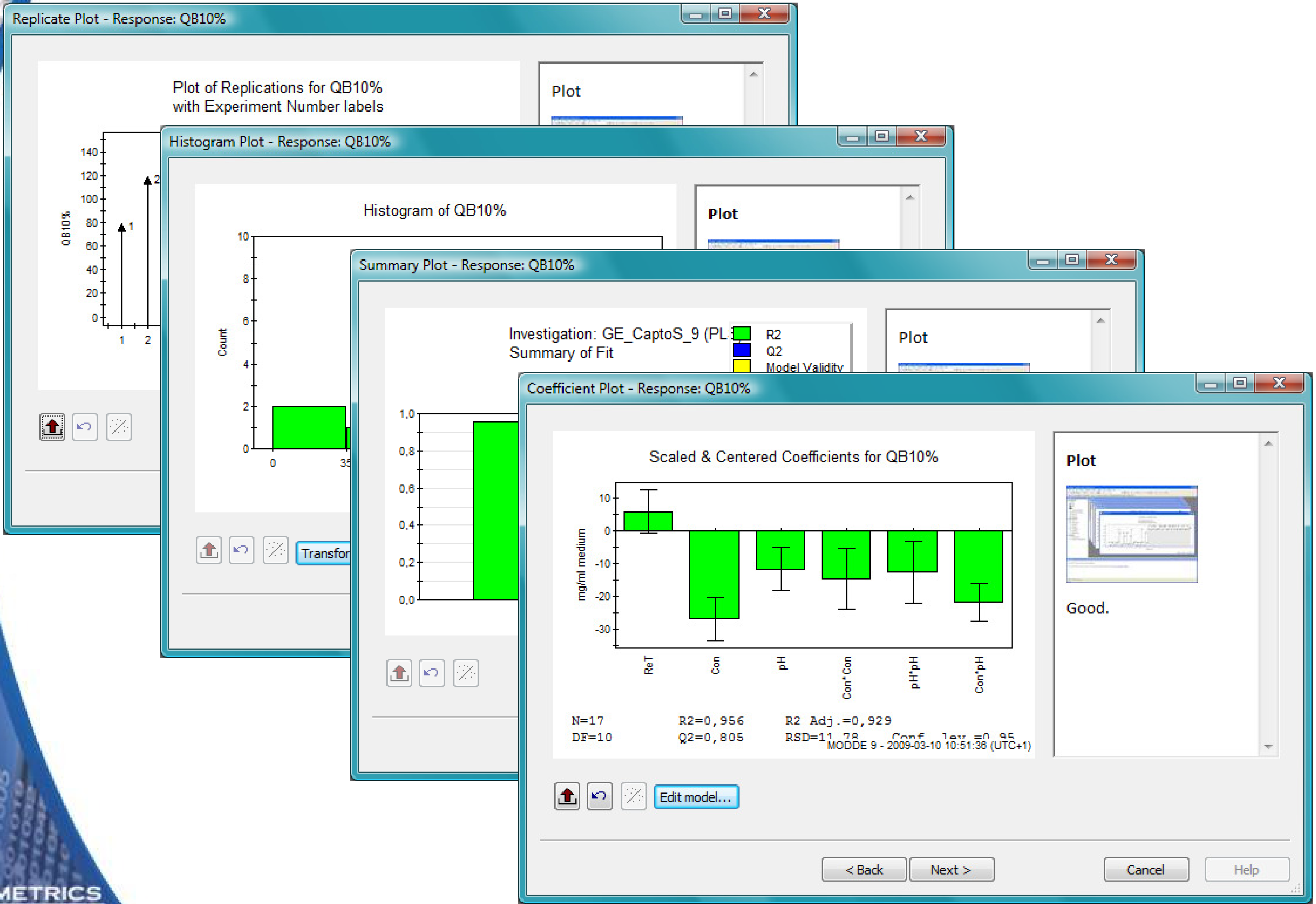
NEWS IN MODDE 9

News in MODDE 9

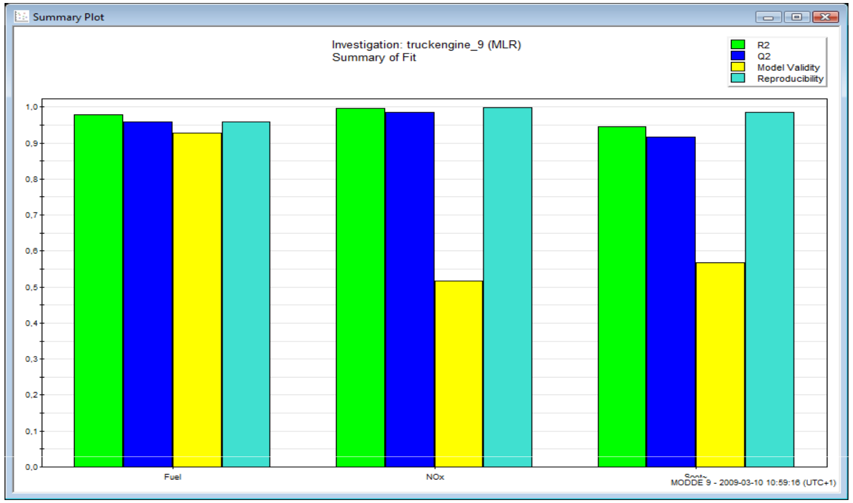
- Analysis wizard
  - Increase user friendliness
- Multiple models
  - Possibility to fit separate models for each response
- New graphics!
- Visualization of Design space
- Release Q3

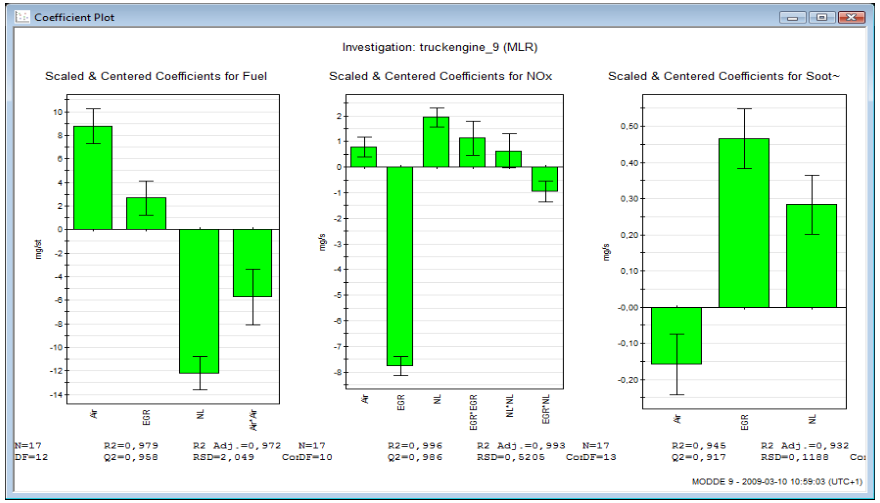
UMETRICS
AN MKS COMPANY

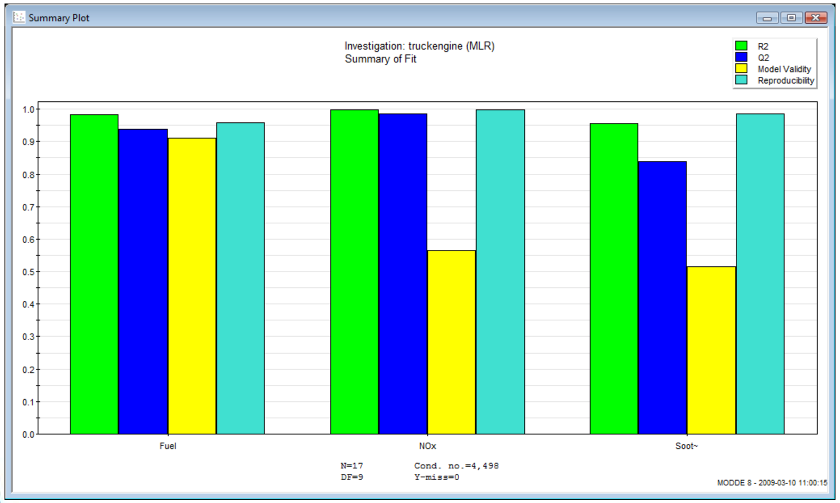
Analysis wizard

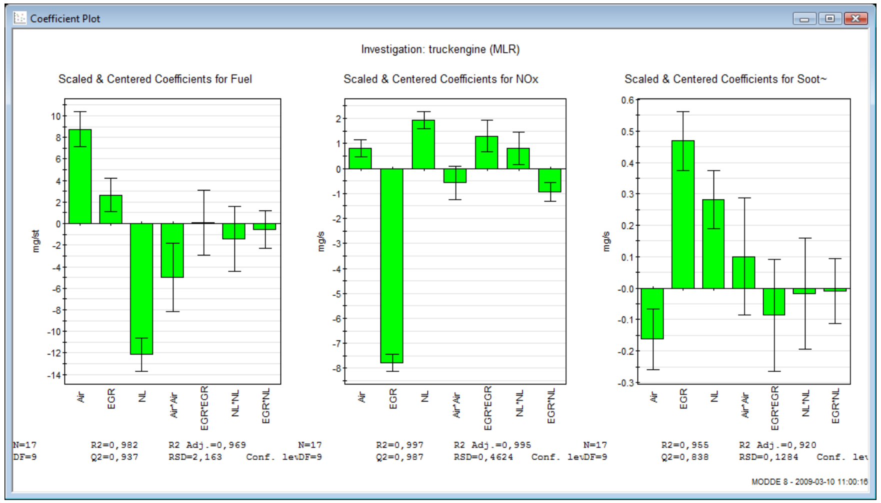


Separate models for each response





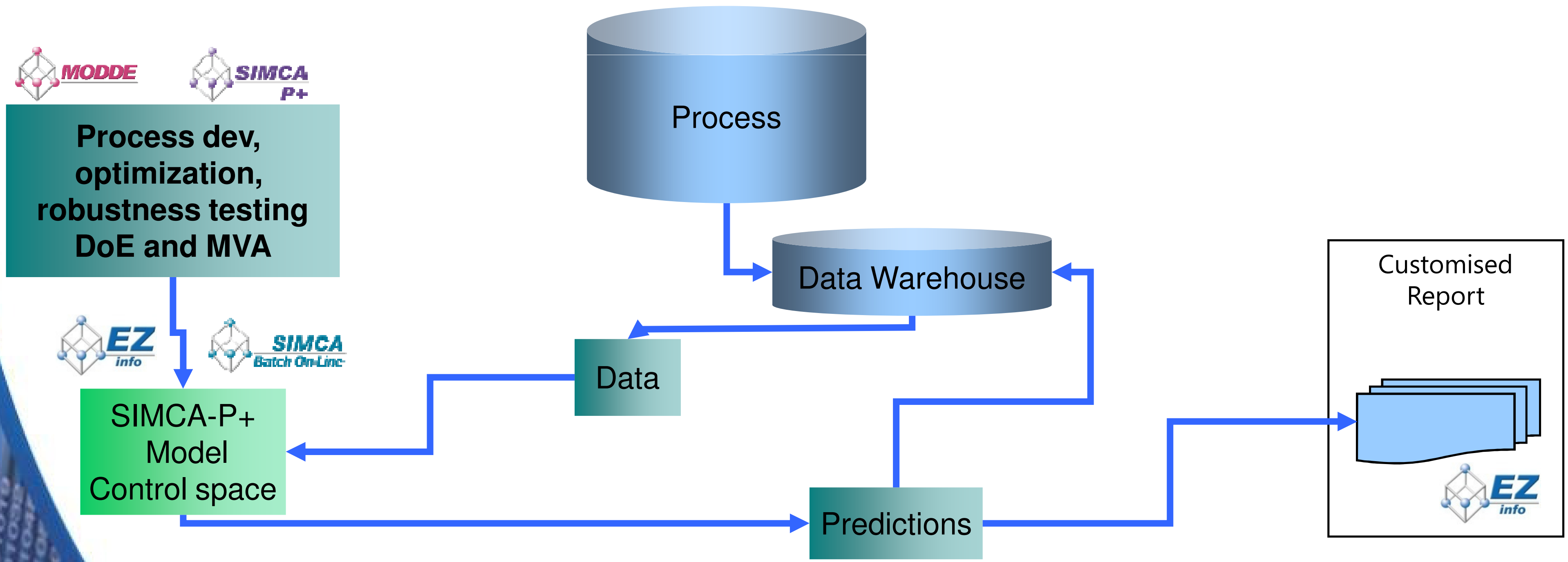




UMETRICS
AN MKS COMPANY

M-link interface

- Design of experiments is the guaranteed way to get an answer to a question
- Also in Multivariate data analysis DoE if of high importance
- Ultimate (and more and more used) workflow in production environment:



Summary

- Benefits of DoE
  - Systematic approach leading to maximum knowledge
- Importance of problem formulation
  - DoE does not have to be "perfect" but executable
- Three primary theoretical objectives of DoE
  - Screening, optimization and robustness testing
- New design types in MODDE
- Custom made design types
- "Multivariate" design

UMETRICS
AN MKS COMPANY

Contact information

Ing-Marie Olsson, Ph.D.

Application Specialist

ing-marie.olsson@umetrics.com

Umetrics AB

Stortorget 21, 4tr

SE- 211 34 Malmö

Sweden

Phone: +46 40664 25 87
Mobile: +46 73684 91 87
Fax: +46 40664 25 85

Umetrics website: http://www.umetrics.com

UMETRICS
AN MKS COMPANY