TTK31 - Design of Experiments (DoE), metamodelling and Quality by Design (QbD)
Autumn 2021

Big Data Cybernetics Gang

NTNU

Lecture overview

1. DoE: Introduction and motivation
2. ANalysis Of VAriance (ANOVA)
3. Factorial designs and Fractional factorial designs
4. **Response surface designs and model selection**
5. Optimal designs
6. Metamodelling
7. Combining DoE with multivariate analysis/machine learning
8. QbD – PAT
9. Practical examples of DoE related to cybernetics

Topics for today

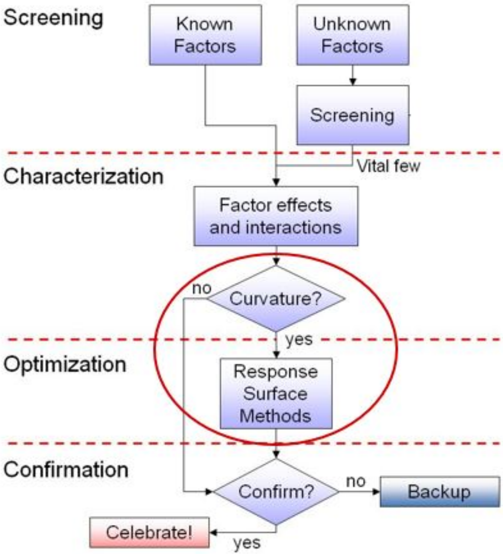
- Response surface designs
- Model selection
- Demo: Response surface design
- If time allows: Rerun of helicopter design from last week

What if you cannot perform all experiments in one day? Some words about blocking

- Blocking is a technique used to mathematically remove the variation caused by some identifiable change during the experimental campaign
- E.g. if you extend a factorial design to an optimization design, and performing the new experiments the week after
- It is assumed that the block variable does not interact with the factors.
- You can look at the alias structure to see which effects have been “lost to blocks".

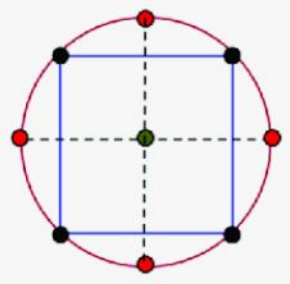
Response surface designs

Strategy for experimentation



Optimization with two design factors

Optimization with two design factors



Optimization designs

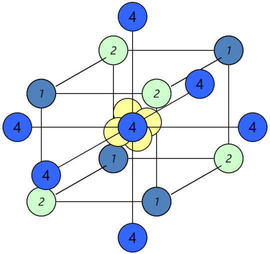
- Objective
  - Model the response surface with *accuracy*, so as to know the *precise shape* of the response surface and find the *optimal* values
- The model terms
  - Include main effects
  - Include interactions
  - Include squared and/or cubic terms
- Design types
  - Central Composite designs
  - Box-Behnken designs
  - Optimal designs (in situations with constraints or to minimize the number of runs)

While a two-level design with center points cannot estimate individual pure quadratic effects, it can detect them effectively

Central Composite Designs (CCD)

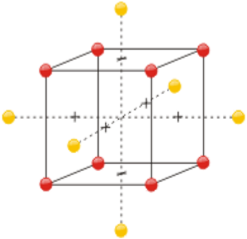
A full factorial 2-level design can be extended to a *Central Composite Design* by adding star points

- Good for modelling a response surface
- 5 levels for each variable
- Can be built as an extension of a full factorial
- Additional points are called star or axial points

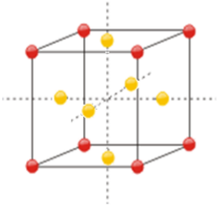


1. Fractional factorial, 2. Full factorial, 3. Centre points, 4. Axial points

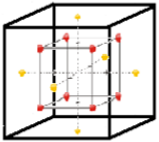
Types of CCD



Circumscribed central composite (CCC)



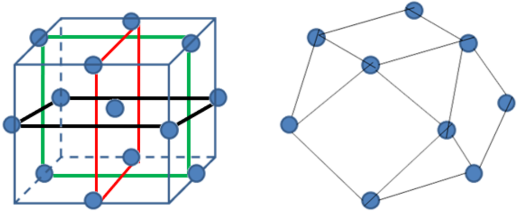
Faced central) composite (CCF)



Inscribed central composite (CCI)

Bob-Behnken designs

- 3 levels for each variable
- Slightly fewer runs than CCD
- Extreme combinations are avoided



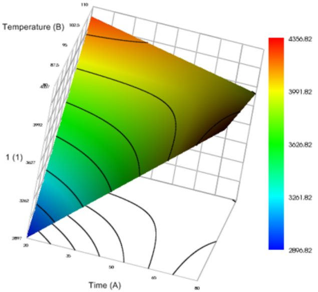
How to select ranges of variation

- Wide enough to generate response variation
- Narrow enough to avoid huge non-linearities
- Useful tip: Start with two extreme combinations
  - If too extreme results: narrow down
  - If different enough results: OK
  - If too close results: Check center sample
    - If close to the others: widen up
    - If different: curvature (narrow down)

Response surface methodology

$$y = b_0 + \sum b_i x_i + \sum b_i x_i^2 + \sum \sum b_{ij} x_i x_j + f$$

- Purpose: Closely approximate the true shape of the response surface
- Quadratic model
- Method: MLR/ANOVA
- Predict the response value(s) for any combination of the design variable settings in the experimental region
- Find the variable settings that give desired response value(s) in the experimental region (optimization)

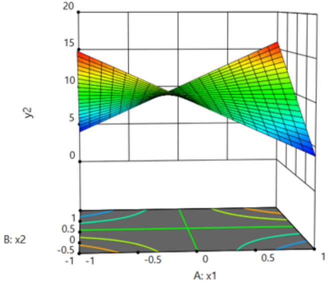


Plots for main results and diagnostics

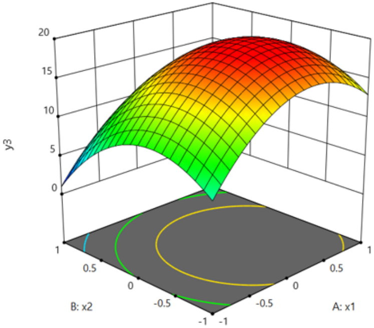
- Plots of main effects interactions (e.g. half-normal plots)
- Response surface (contour, 3D surface)
- Predicted vs. Reference
- Residuals
- Various statistical diagnostics, e.g. R^2

Examples of response surfaces - I

Two-factor interaction

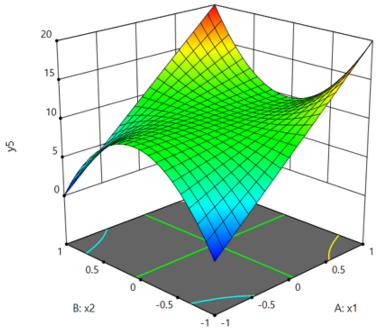


Squared terms

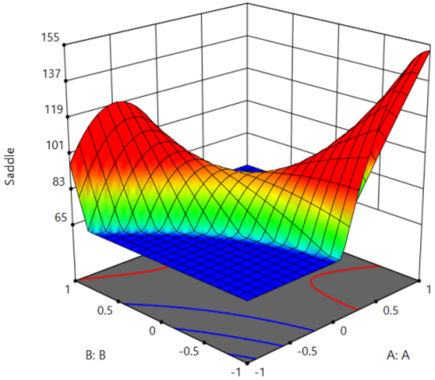


Examples of response surfaces - II

Cubic model: Inflection



Cubic model: Saddle



Example: CCD with three design factors

Std	Run	Factor 1 A:Time min.	Factor 2 B:Temperature deg C	Factor 3 C:Catalyst %	Response 1 Conversion %	Response 2 Activity
1	6	40	80	2	74	53.2
2	20	50	80	2	51	62.9
3	7	40	90	2	88	53.4
4	2	50	90	2	70	62.6
5	4	40	80	3	71	57.3
6	3	50	80	3	90	67.9
7	15	40	90	3	66	59.8
8	14	50	90	3	97	67.8
9	8	36.591	85	2.5	81	59.2
10	11	53.409	85	2.5	75	60.4
11	12	45	76.591	2.5	76	59.1
12	10	45	93.409	2.5	83	60.6
13	9	45	85	1.6591	76	53.6
14	1	45	85	3.3409	79	65.9
15	16	45	85	2.5	85	60
16	17	45	85	2.5	97	60.7
17	18	45	85	2.5	55	57.4
18	13	45	85	2.5	81	63.2
19	5	45	85	2.5	80	60.8
20	19	45	85	2.5	91	58.9

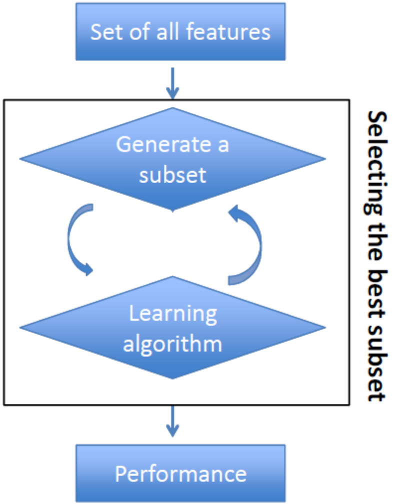
ANOVA for a quadratic model with replicated center samples

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	1603.32	9	178.15	1.52	0.2625	not significant
A-Time	0.0871	1	0.0871	0.0007	0.9788	
B-Temperature	160.19	1	160.19	1.36	0.2700	
C-Catalyst	155.25	1	155.25	1.32	0.2771	
AB	36.12	1	36.12	0.3075	0.5914	
AC	1035.12	1	1035.12	8.81	0.0141	
BC	120.12	1	120.12	1.02	0.3358	
A^2	42.42	1	42.42	0.3611	0.5613	
B^2	20.25	1	20.25	0.1723	0.6868	
C^2	51.61	1	51.61	0.4393	0.5224	
Residual	1174.88	10	117.49			
Lack of Fit	127.38	5	25.48	0.1216	0.9814	not significant
Pure Error	1047.50	5	209.50			
Cor Total	2778.20	19				

Model selection

Model selection

- Several subsets of features are searched through, and the data modelling accuracy is evaluated.
- Advantages:
  - Allow detection of possible interactions between variables
  - Often finds a good subset for a specific learning algorithm
- Disadvantages:
  - Risk of overfitting when the number of observations is low
  - Computationally demanding when the number of features is large



Examples of various approaches

- **Forward selection**
  - Starts with having no feature in the mode
  - Iteratively adds the feature which best improves the model
  - Stops when no improvement is observed upon addition of features
- **Backward elimination**
  - Starts with all the features
  - Iteratively removes the least significant feature
  - Stops when no improvement is observed upon removal of features
- Stepwise Selection:
  - Re-considers all dropped/added variables for reintroduction/drop-out in each step
- Randomized wrapper methods:
  - Include some degree of randomness in the selection

NB! Good approach for orthogonal designs, with non-orthogonal designs and with covariates present must be careful with interpretation

Criteria for including or excluding model terms

- Adjusted R-squared
- p-values
- Akaike's Information Criterion (AIC) or AICc
- Bayesian Information Criterion (BIC)

When selecting your model, it is suggested to use multiple methods and criteria

More details

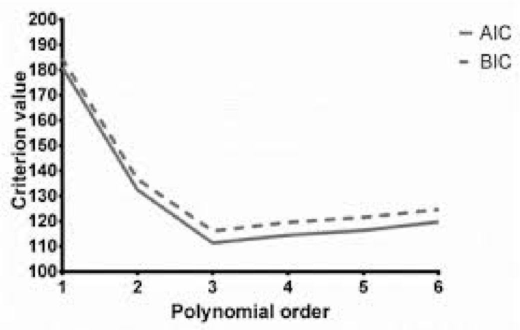
- For all methods:
  - Use Type II or Type III Sum of Squares
  - Select model complexity (main effects, interactions, squares,...)
- For p-values, AIC and BIC: Forward or backward
- for p-values and Adjusted R-squared: Cut-off for excluding model terms

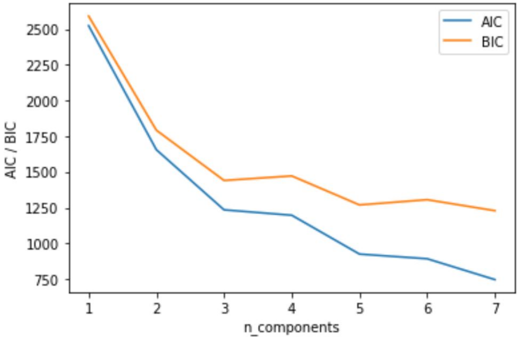
In general: What is the least significant difference in R-squared that should be applied to say that one model is better than another?

AIC and BIC

- AIC rewards goodness of fit, but it also includes a penalty that is an increasing function of the number of estimated parameters
- When the sample size is small, there is a substantial probability that AIC will select models that have too many parameters
- To address such potential overfitting, AICc was introduced: AICc is AIC with a correction for small sample sizes
- BIC is similar to AIC but applies a different penalty for the number of parameters

Comparison of AIC and BIC





Demo: Optimization design

- Objective: Optimize the conversion of a chemical reaction
- A Central Composite Design
- Design factors: Time, temperature, catalyst (%)
- Responses: Conversion (%), activity
- Procedure:
  - Set up the design
  - Inspect raw data
  - Analyze data
  - Numerical optimization