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

Big Data Cybernetics Gang

NTNU

General course information

- Lecturers:
  - Frank Westad
  - Øivind Riis
- Lecture time: Thursdays 12:15-14:00
- 10 lectures, see Blackboard
- 2-4 Exercises
- hands-on analysis

Reference group - VERY IMPORTANT

- at least 3 students
- will do 4 meetings (1 after the exam)
- shall represent the whole class $\implies$ you will have meetings among yourselves too
- shall lead to a referansegrupperapport containing suggestions for improvements

Design of Experiments

Objectives with this course (specialization topic):

- Understand the principles of Design of Experiments (DoE)
- Understand the use of ANOVA in analysing results from DoE (and in general)
- Be able to decide on the right design given the problem at hand
- How to apply DoE for metamodelling
- How DoE falls into a framework of Quality by Design (QbD) and Process Analytical Technology (PAT)

Lecture overview

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

Exercises/assignments and hands-on analysis

The Design-Expert®software from the company Stat-Ease is available for your use during this course
The classroom serial number is 8300-5935-5835-CLAS

`http://www.stat-ease.com/`

Follow these steps to register, download, and install the program to your personal computer:

- Create an account on the Stat-Ease website register this license to your account
- Download and install the software on your computer

Introduction

Design of Experiments

- We claim: Everybody working in quantitative sciences should know about the principles of DoE
- DoE is useful for:
  - discovering what are the important parameters in a system/process
  - identify if there are interaction effects and higher order relationships between input and output
  - metamodelling based on physical models
  - speeding up simulation systems
  - finding the best process settings given changes in raw material, equipment, sensors etc.

Why is not DoE widely used?

A survey showed:

The main barriers that hinder the widespread use of DoE are

- low managerial commitment
- engineers' general weakness in statistics.

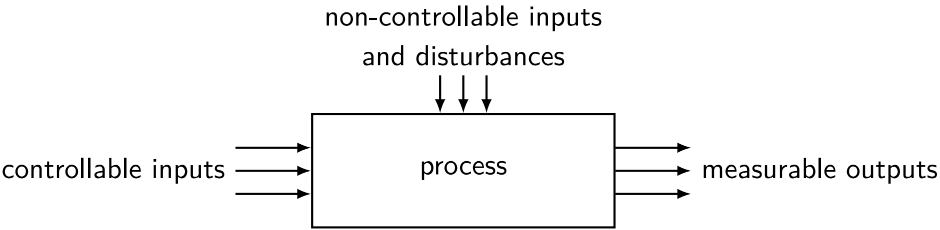
The overall 16 barriers were classified in three groups:

- business barriers
- **educational barriers**
- technical barriers

Although DoE is commonly found in statistics and quality literature, it is clearly underused in industry.

Standing assumption:

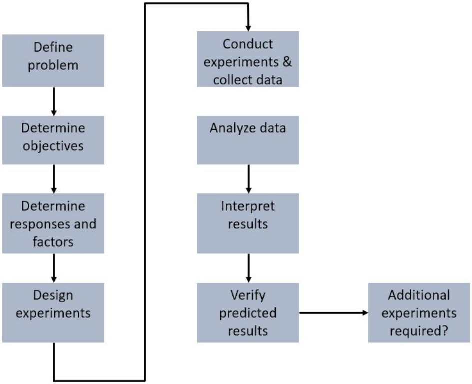
We want to model a system, and we can control some of its inputs / decide how to measure



What is Design of Experiments?

Design of Experiments (DoE) is the pre-planned, **systematic variation of controllable** experimental factors that **induce a response** in a system. The factors are measured in such a way that the **minimum effort** is required to gain a **maximum amount of information**

The experimental design cycle



If needed: Iterate by starting from top or at a certain step in the cycle

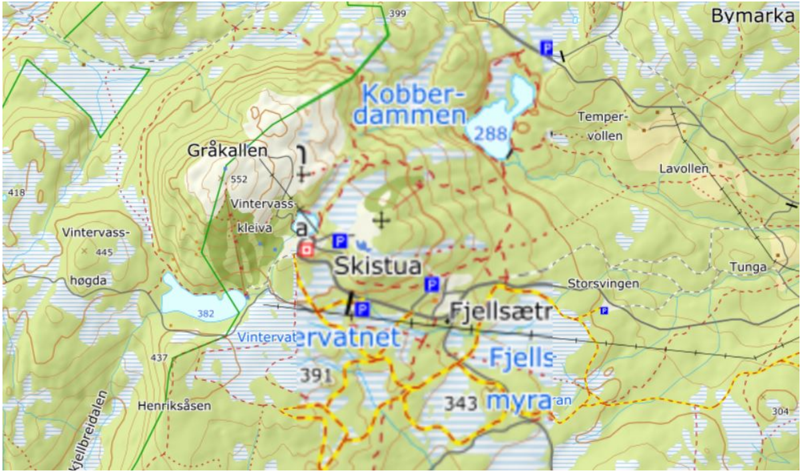
One variable at the time (OVAT)

In order to establish a relationship between
cause and effect,
each cause must be investigated
separately,
all other conditions being fixed.

An excursion to Gråkallen

Assume you want to hike to the top of Gråkallen

How would you reach the top?



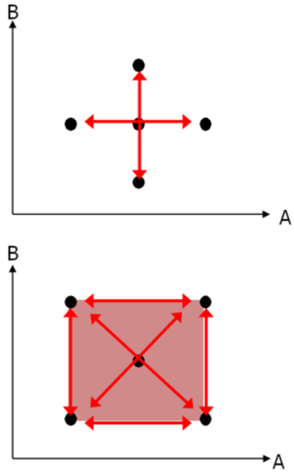
Experimental design versus OVAT

Traditional approach: One Variable At a Time

- One variable at a time (OVAT)
- Cannot detect interactions
- Inefficient (serial processing)

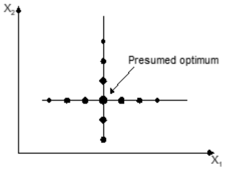
Experimental design

- All experiments are used to estimate effects of A and B
- Interactions can be estimated
- Precision can be estimated
- Maximizes information with minimum runs

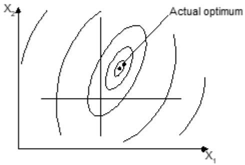


How to span the experimental space

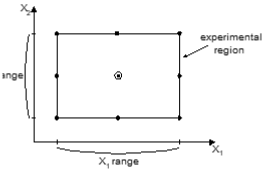
The Classical Approach:



(OVAT) What can go wrong?



How can we do it better?



The DoE process

- Identify opportunity and define objective
- State objective in terms of measurable responses
  1. Define the minimal change (Δy) that is important to detect for each response (signal)
  2. Estimate experimental error (σ) for each response (noise)
  3. Use the signal to noise ratio (Δy / σ) to estimate power
- Select the input factors to study.
- Select an appropriate design

DoE is in most cases a sequential process

In most cases the experiments must be performed as a sequence of trials

- Screening: A design with 6-12 factors with the purpose of identifying the important ones
- Advanced screening: Investigate possible interaction effects of a smaller number of factors
- Optimization: Find the optimum settings with a more precise model

NB! With proper DoE one can re-use experiments from the previous steps!

Advantages of Experimental Design

	OVAT	**Experimental Design**
Main effects	Not estimated	Estimated
Interactions	Not detected	Detected and estimated
Experimental variability	100% impact	Reduced impact
Number of experiments	Unknown	Known per step
Best solution	???	Spotted
If no solution	???	Detected
Several responses	Difficult	As easy as one response
New objectives	Start all over again	Re-using existing results

Main types of designs

Type of design	*Objective*
Fractional factorial	Find main effects
Full factorial	Find main effects and interactions
Optimization designs	Find optimal settings for a response surface
Mixture designs	Find the optimal recipe of a mixture
Optimal designs	Designs with constraints

A small example

Assume you want to bake the best cake ever

Which are the factors you can change?

What characterize the quality?