

1222 • 2022
800
ANNI



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

DTG

Department of Management and Engineering

Design of Experiments

Luigi Salmaso

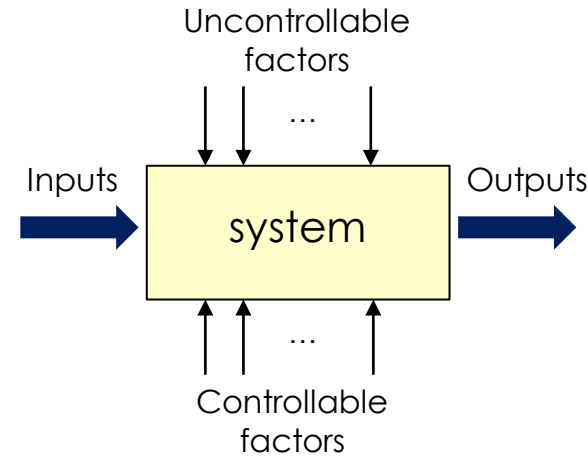
The practice of using statistics in industry is well established since decades.

It is a long-lasting story disseminated of successes:

- **1900–1930** Henry Ford: how to improve productivity and quality; in-process inspection
- **1907–1908** AT&T begins systematic inspection and testing of products and materials
- **1908** W. S. Gosset (writing as “Student”) introduces the t-distribution—results from his work on quality control at Guinness Brewery
- **1920s** AT&T Bell Laboratories forms a quality department, emphasizing quality, inspection, test and product reliability. General Electric in England uses statistical methods to control the quality of electric lamps.
- **1922-1923 R. A. Fisher publishes series of fundamental papers on designed experiments and their application to the agricultural sciences**

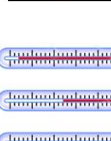
The basis of “**Design of Experiments**” (DOE) is the identification of a set of **factors**, which can potentially drive **process performance**, the selection of **reasonable levels** for each of these factors, the definition of a set of **combinations** of factor levels and the execution of **experiments** according to the defined experimental design.

The technique of **DOE** deals with the identification of the combinations of factor levels that **maximize information acquisition** on a specific phenomenon. Therefore, an implicit objective of DOE is to **minimize the number of experimental trials** to execute.



A process with 10 input factors on 5 levels, potentially leads to 10 million factorial combinations (experimental trials!) $\Rightarrow 5^{10} \sim 10M$

Example:
You are studying factors that could affect plastic strength during the manufacturing process.

Factors	Levels	Response variable
Type of Additive	A  B 	Plastic tensile strength 
Temperature	High (200°C) 	
	Medium (150°C) 	
	Low (100°C) 	

Why **statistically well-design** experiments?

- Reduce **time** to design/develop new products & processes
- Improve **performance** of existing processes and products
- Improve **reliability** of products (similar results under consistent conditions)
- Achieve product & process **robustness** (good performance under different field conditions)
- **Evaluation** of materials, design alternatives, **setting** component & system tolerances, etc.



Experiments are usually iterative

The learning process advances by formulating hypotheses, performing experiments to investigate them, formulating new ones on the basis of the results, and so on.

It is usually a mistake to design a single, large, comprehensive experiment at the start of the study.

A sequential approach employing a series of smaller experiments is usually a better strategy.

It is usually preferable to adopt a sequential approach of experimentation

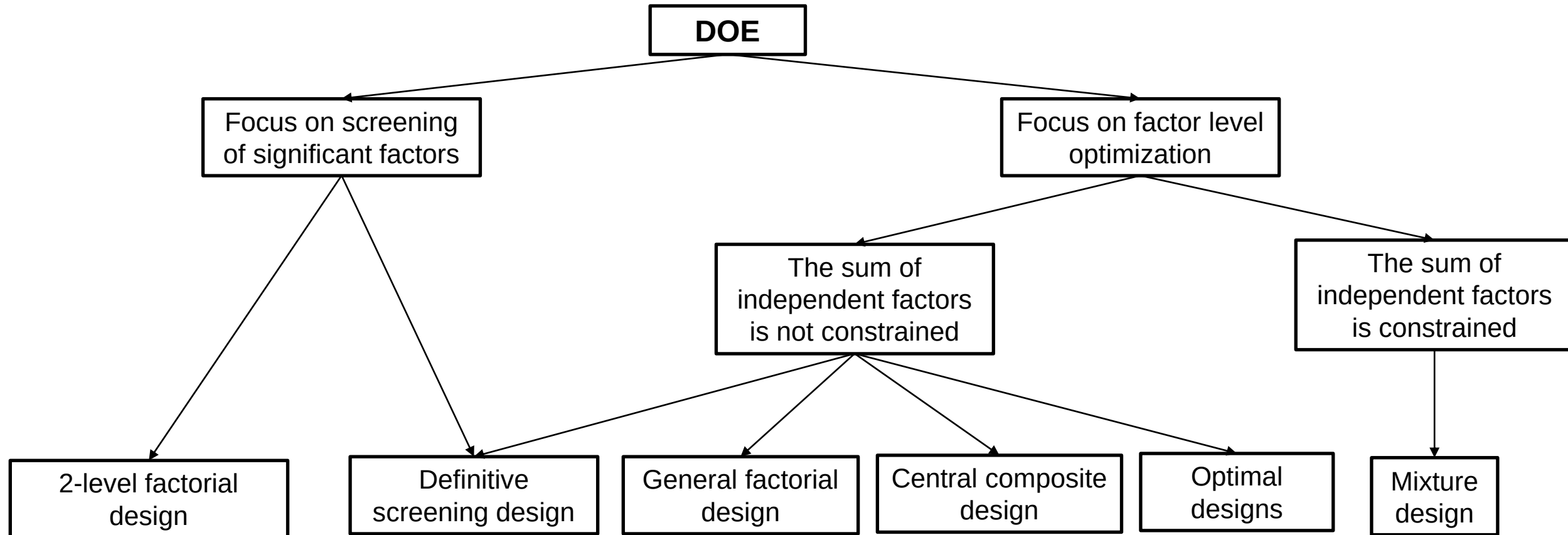
SCREENING EXPERIMENT

- *The objective is to detect the significant variables*
- *Among the significant variables, the most important ones can be identified*
- *2-level designs: investigate a lot of variables while keeping the number of treatments under control*



CONFIRMATORY EXPERIMENT

- *The objective is to optimize the factor settings*
- *Only significant factors are considered*
- *General factorial designs: factors on >2 levels to detect curvature*
- *Optimal designs to minimize the number of runs*



2-LEVEL FACTORIAL DESIGNS

OBJECTIVE:

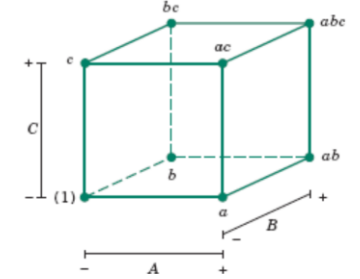
- Find out which factors are significant for the response(s).
- Assess the direction and entity of the impact of several factors on one or more responses.

FIELD OF APPLICATION:

- Applied in '**screening experiments**', when the behaviour of the phenomenon is not known and several potential factors should be investigated.

DETAILS:

- Only two levels are set for each factor. The levels chosen for the factors are problem-specific, but a guideline may be to choose the 'absence' of the factor as low level and a 'large' level of the factor as high level.
- Employed at the first step in the 'sequential approach' of experiments.
- It can be 'full' factorial design, when the resources available are many, or 'fractional' factorial design when the resources are limited.
- Replicated runs are necessary.



Example

Clostridium autoethanogenum (CA) is an anaerobic bacterium that produces *ethanol* from carbon monoxide (CO), in so-called syngas fermentation, being one of the few known microorganisms to do so.

It is of interest to find out which factors have an impact on *ethanol* production using CA.

To carry out the foreseen chemical reaction, the bacteria culture needs to grow in a medium, and several chemical compounds can be used that may potentially have an impact on the growth and ethanol production of the bacteria.

Specifically, the identified molecules are:

$NaCl$, NH_4Cl , KH_2PO_4 , $MgSO_4$, $CaCl_2$

Additional factors that can play a role are:

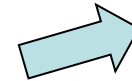
Yeast extract, pH



Example

The chemists identify a relevant **range** of variation for each factor. The range chosen should be large enough to ensure a detection of the effect of the factor, if this effect exists.

Factor	Low level (-1)	High level (+1)
<i>NaCl</i>	0,4 g/L	1 g/L
<i>NH₄Cl</i>	0,6 g/L	1 g/L
<i>KH₂PO₄</i>	0,04 g/L	0,1 g/L
<i>MgSO₄</i>	0,15 g/L	0,3 g/L
<i>CaCl₂</i>	0,02 g/L	0,04 g/L
<i>Yeast extract</i>	0,15 g/L	0,3 g/L
<i>pH</i>	4,5	5



A full factorial design considering all the combinations of factors levels amounts to $2^7 = 128$ experimental trials, that should be replicated.



In a **FULL FACTORIAL DESIGN** responses are measured at all combinations of factor levels



In a **FRACTIONAL FACTORIAL DESIGN** responses are measured at a subset (fraction) of all combinations of factor levels.

FULL FACTORIAL DESIGNS

All main effects and interactions can be estimated.

With many factors and/or many levels, they may result in a **prohibitive** number of runs.

OPTIMAL / FRACTIONAL FACTORIAL DESIGNS

Some of the effects will be **confounded**. Confounded effects cannot be estimated separately and are said to be **aliased**.

They reduce the number of runs to a **manageable** size minimizing time and cost.

FACTORIAL DESIGNS

Full



All effects estimated



Prohibitive No. runs



Optimal/Fractional



Some effects not estimated



Manageable No. runs

