Using multivariate experiments for faster learning

By: Martin Carignan & Vincent Béchard

Presented at the 2022 PaperWeek Canada

2022-02-10

Agenda

- Univariate trials
- Why “multivariate” matters?
- What about multivariate DOE?
- The power of model-driven DOE
- Final thoughts

Keywords you need to know

Univariate

Working with one variable at the time

Multivariate

Working with several variables altogether

Experiment

Manipulating some process variables (X) to study the impact on performance indicators (Y)

Martin Carignan

Somewhat strange statistician with useful tricks to help people and process feeling better

Vincent Béchard

Definitely bizarre mathematician surprisingly providing beneficial insights from data and models

Univariate trials

Concept of univariate trials

- Approach commonly used in the industry:
  - Using process expertise, select the most influential variables
  - Take one variable and select the desired trial points
  - Perform the trials:
    - Make sure everything is stable and constant in the process
    - Sequentially apply the trial points for this variable
    - Collect the process performance data
  - Repeat for the other variables until done

Goal: isolate the impact of a specific variable (since nothing else varies)

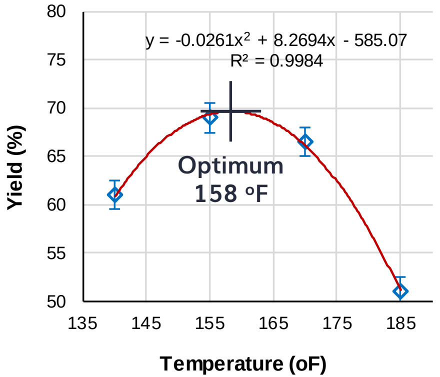
Illustration of the concept

≡ The yield of a batch process reaction depends on:

- Temperature, between 140ºF and 180ºF
- Batch time, between 0.5h and 2.5h

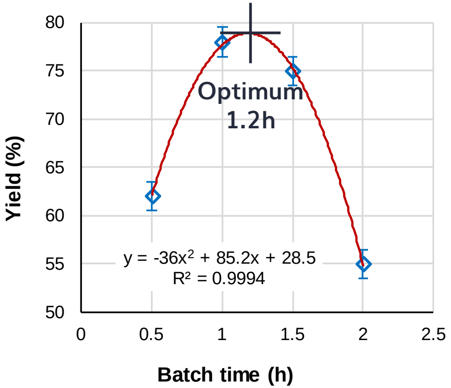
Step 1: temperature

Using actual batch time of 1.75



Step 2: batch time

Using the recommended 158 ºF

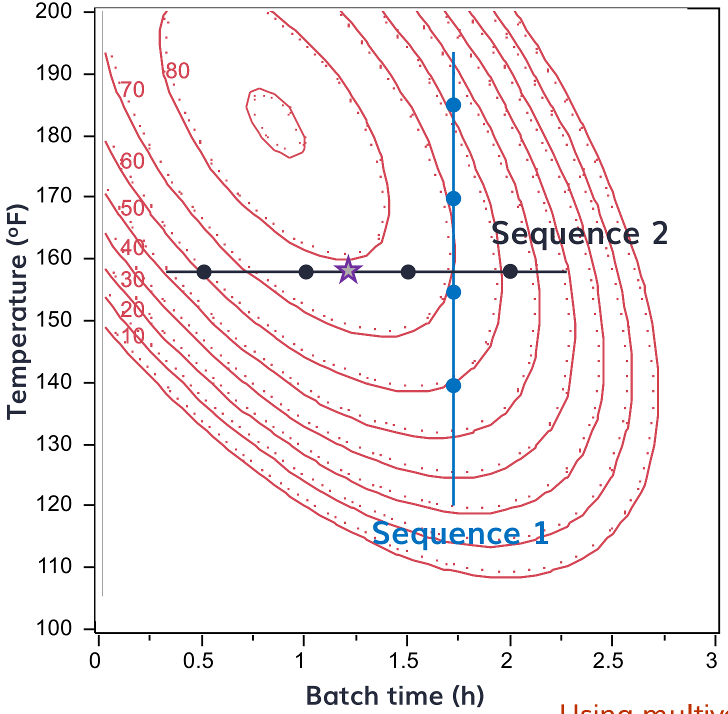


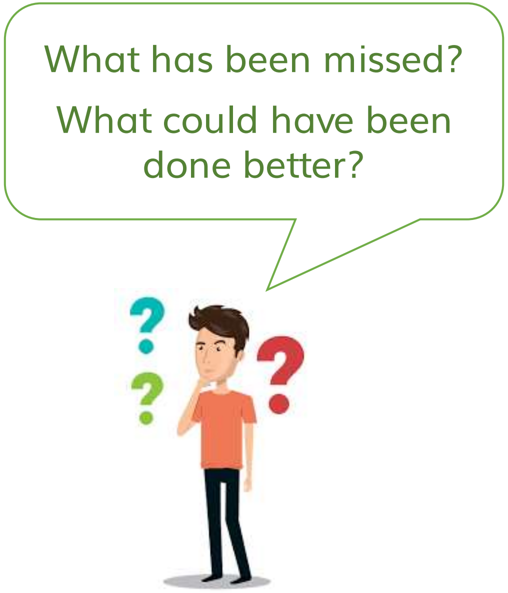
Solution

At 158 ºF for 1.2h, the maximum yield is 78%!

What did we really do?

≡ Sequence of 2 univariate trials:



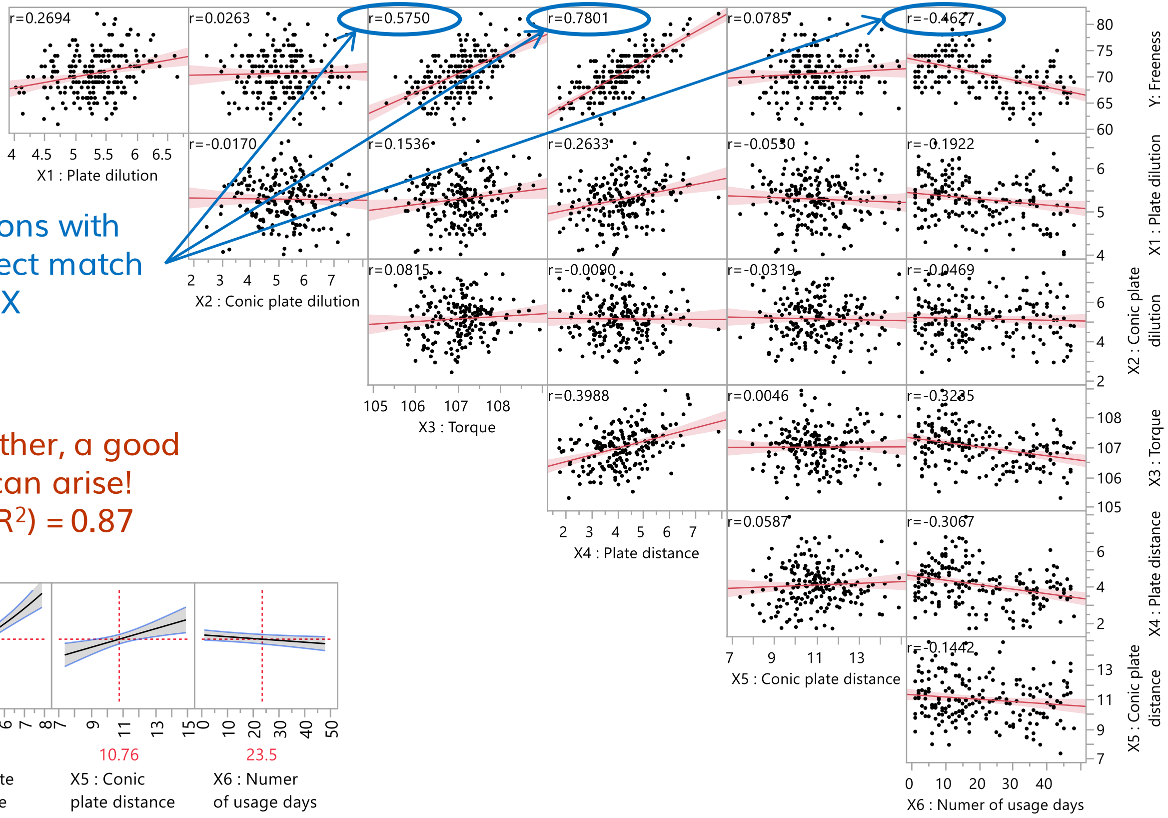


Why “multivariate” matters?

Why "multivariate" matters?

- In a typical process, many parameters have to be set simultaneously to improve the performance
  - Example with sheet breaks at the paper machine: TMP% in pulp, jet-to-wire ratio, roll pressure, headbox pH, …
  - How many data collection points on a machine????
- Question: what if several X variables are mildly correlated with the Y?
  - Dropping them: some signal could be lost
  - Keeping them: some noise could come with the signal
  - Could they work well together to explain the Y?

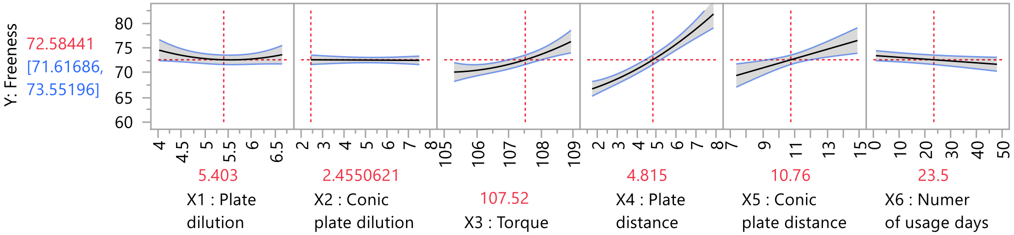
Illustration on “real” data



Interesting correlations with Freeness, but no perfect match with a single X

RSquare	0.750174
RSquare Adj	0.732329
Root Mean Square Error	2.056201
Mean of Response	70.65816
Observations (or Sum Wgts)	196

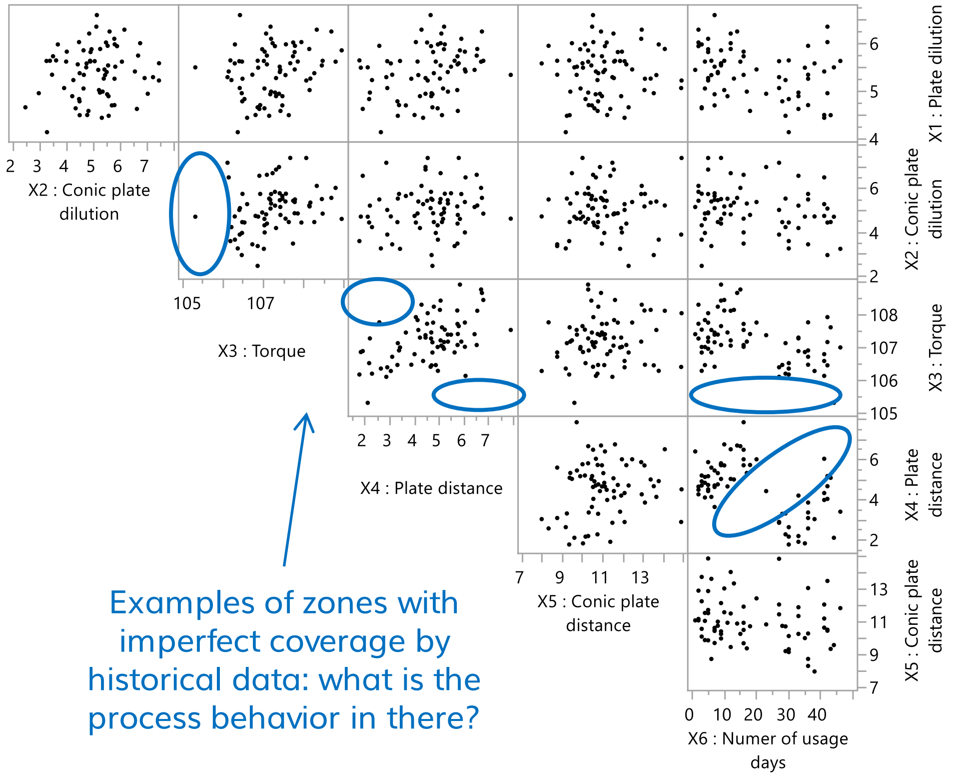
When together, a good model can arise!
$r = \text{sqrt}(R^2) = 0.87$



Using historical data to optimize

Potential pitfall #1

Have all the multivariate combinations been explored?

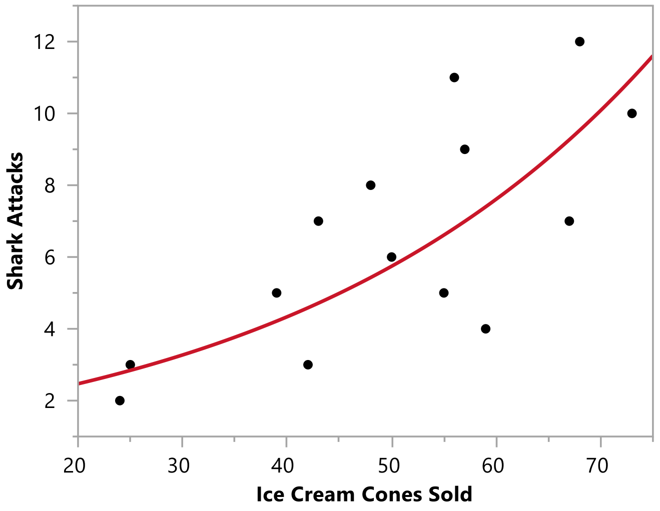


Using historical data to optimize

Potential pitfall #2

Are these “true” cause-and-effect relationships?

How was that proven?

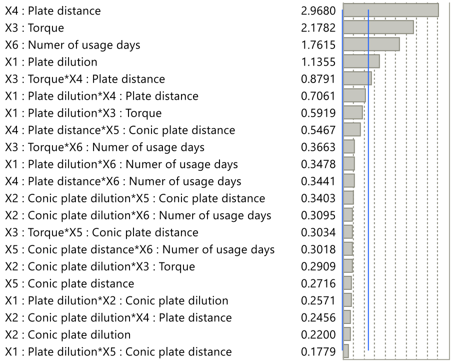


Really a cause-and-effect?

Who was crazy enough to try that!

The good news…

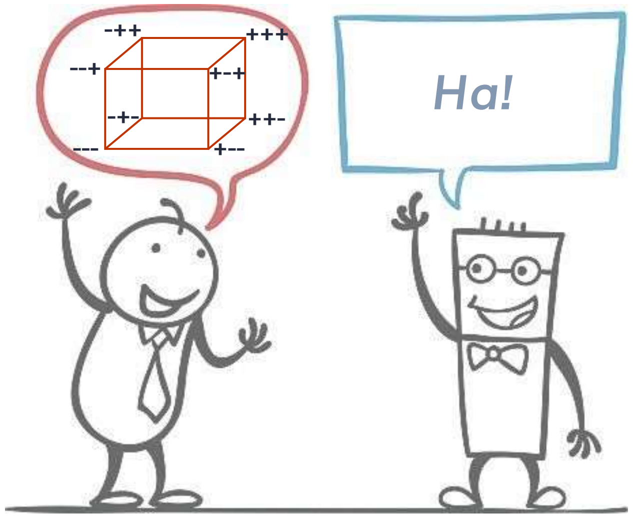
- Multivariate historical data can't replace a thorough experiment
- However, historical data can help screening the important factors to avoid testing them all
  - Use the Pareto principle to select the most influential ones
  - Anticipate the type of relationship (model) to expect



What about multivariate DOE?

What is a DOE ?

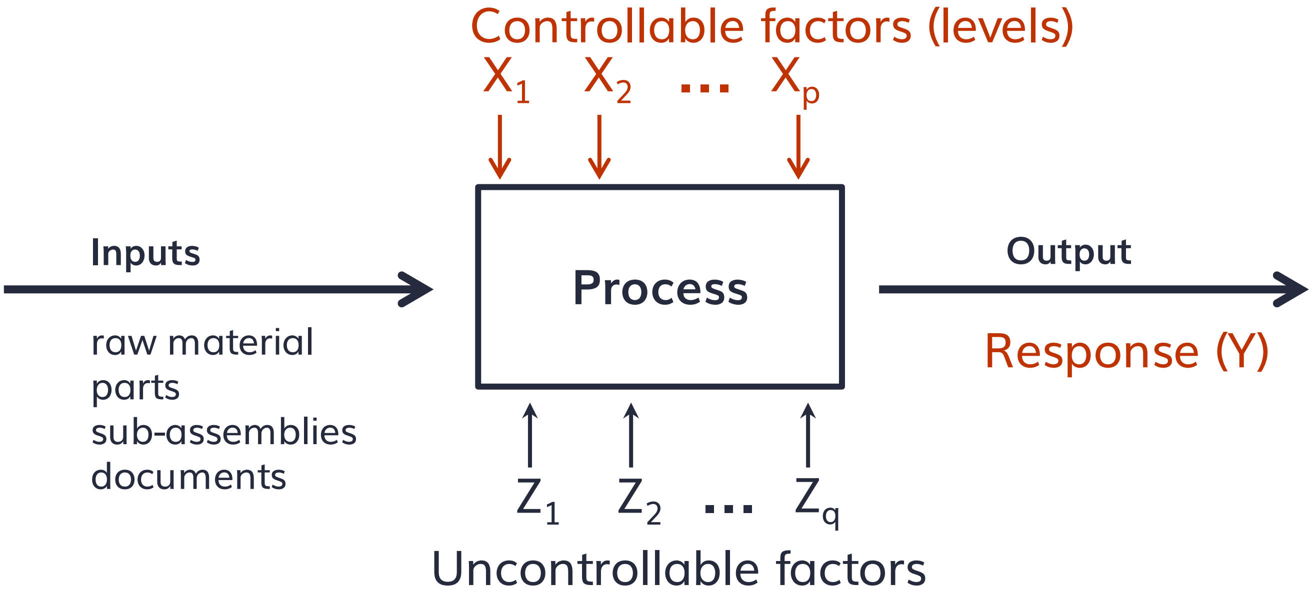
- Design of Experiments (DOE)
  = Dialog with the process
- Identifies cause-and-effect relationships
- Helps understanding complex processes



Important terminology

Factor	σ Parameter suspected to affect a response variable (or dependent variable or Y) σ Examples: temperature, pressure
Level	σ Values taken by a factor during an experiment σ Example: 100°C, 200°C, 300°C.
Treatment	σ Combination of factors' levels being investigated (a row in the data table!)
Experimental space	σ Material, equipment, environmental conditions and other conditions prevailing during an experiment

What is a process?



What is a process?

Controllable factors (levels)
- Torque of presses
- Distance between conical and flat plates
- Dilution at conical and flat plates
- Tonnage

Rejected pulp → **TMP**
Refining of rejects → **Refined pulp**

Quality parameters
- Freeness
- Tear
- Burst

Uncontrollable factor
Plate wear

Identify which factors and interactions have the biggest impact on strength and freeness

Basic experimental principles

- In order to perform efficient experiments on processes affected by uncontrollable factors, four basic principles must be considered
  - Randomization
  - Replication
  - Repetition
  - Blocking

Basic principles – Example

- Nadine wants to run a DOE in her paint booth.
- After some brainstorming with her team and data analysis they decide to experiment with
  - X1: fluid flow
  - X2: attack angle of the paint gun

  and measure their impact on paint thickness
- She wants to test a “high” (1) and “low” (-1) level for each factor for a total of 4 different treatments

Basic Principles – Replication

- Design replication is redoing all treatments (with reset; i.e. all runs have the same setup process)
  - Allows to estimate experimental error. This estimate will then be used to determine whether the effects observed are statistically significant.
  - When the sample mean is used to estimate the effect of a factor, each replicate then improves the precision of the estimate (more observations to calculate the sample mean)

Replication is associated to the improvement of the precision of treatments effect by redoing them

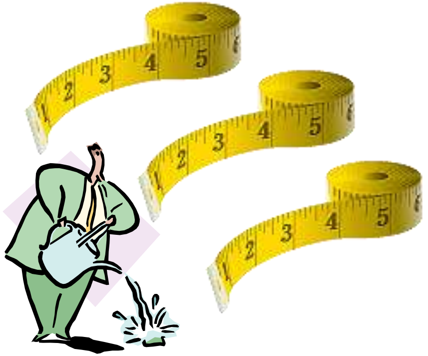
Basic Principles – Replication

Four replicates of each treatment

	Pattern	Fluid flow / débis fluide	Attack angle / Angle d'attaque	Thickness / Epaisseur
1	+−	1	-1	•
2	+−	1	-1	•
3	−+	-1	1	•
4	++	1	1	•
5	++	1	1	•
6	−+	-1	1	•
7	−−	-1	-1	•
8	−+	-1	1	•
9	+−	1	-1	•
10	−−	-1	-1	•
11	−−	-1	-1	•
12	−−	-1	-1	•
13	++	1	1	•
14	−+	-1	1	•
15	−−	-1	-1	•
16	++	1	1	•
17	++	1	1	•
18	−+	-1	1	•
19	+−	1	-1	•
20	+−	1	-1	•

Basic Principles – Repetition

- Repetition is making multiple measurements on the same treatment during the same run of a treatment (without resetting)
  - Allows to get a better estimate of the effect of a treatment during a single run (by averaging all the repeated measurements, the measurement error is decreased)



Repetition is associated to the improvement of the precision of treatments effect by reducing the measurement error

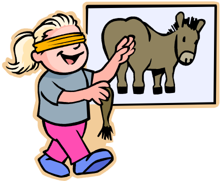
Basic Principles – Repetition

	Pattern	Fluid flow / débis fluide	Attack angle / Angle d'attaque	Thickness / Epaisseur	Thickness / Epaisseur 2	Thickness / Epaisseur 3	Thickness / Epaisseur 4	Thickness / Epaisseur 5
1	+−	1	-1	•	•	•	•	•
2	+−	1	-1	•	•	•	•	•
3	−+	-1	1	•	•	•	•	•
4	++	1	1	•	•	•	•	•
5	++	1	1	•	•	•	•	•
6	−+	-1	1	•	•	•	•	•
7	−−	-1	-1	•	•	•	•	•
8	−+	-1	1	•	•	•	•	•
9	+−	1	-1	•	•	•	•	•
10	−−	-1	-1	•	•	•	•	•
11	−−	-1	-1	•	•	•	•	•
12	−−	-1	-1	•	•	•	•	•
13	++	1	1	•	•	•	•	•
14	−+	-1	1	•	•	•	•	•
15	−−	-1	-1	•	•	•	•	•
16	++	1	1	•	•	•	•	•
17	++	1	1	•	•	•	•	•
18	−+	-1	1	•	•	•	•	•
19	+−	1	-1	•	•	•	•	•
20	+−	1	-1	•	•	•	•	•

Four repetitions at each treatment.

Basic Principles – Randomization

- Randomization means that both, the allocation of the experimental material and order of trials, are randomly determined and a reset is done between each run
  - Allows to "average out" the impact of the uncontrollable or extraneous factors that may be present. It minimizes the risk of bias (for example: working shift)
  - Allows to have observations (or errors) that are independently distributed (required by the statistical methods)



Basic Principles – Randomization

Each treatment is performed **in random order.**

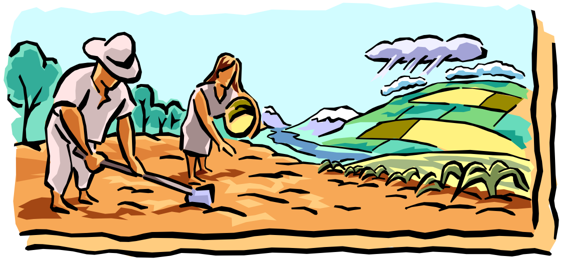
It is not all "Fluid flow" at low level done first then "Fluid flow" at high level.

Between each run, a reset is done.

	Pattern	Fluid flow / débis fluide	Attack angle / Angle d'attaque	Thickness / Epaisseur
1	+-	1	-1	•
2	+-	1	-1	•
3	-+	-1	1	•
4	++	1	1	•
5	++	1	1	•
6	-+	-1	1	•
7	--	-1	-1	•
8	-+	-1	1	•
9	+-	1	-1	•
10	--	-1	-1	•
11	--	-1	-1	•
12	--	-1	-1	•
13	++	1	1	•
14	-+	-1	1	•
15	--	-1	-1	•
16	++	1	1	•
17	++	1	1	•
18	-+	-1	1	•
19	+-	1	-1	•
20	+-	1	-1	•

Basic Principles – Blocking

- Blocking is a subdivision of the experimental space in subgroups, each subgroup being constituted of relatively homogeneous units. It is expected that the experimental error will be smaller within the blocks than within the whole experimental space.
  - The word comes from the domain of agriculture where a field was divided into common conditions : exposition to the wind, nearness of water, thickness of arable land.



Basic Principles – Blocking

If the experiment is performed over five days, five blocks could be created

All treatments are randomized within blocks

This is a randomized complete block design

	Random Block	Fluid flow / débis fluide	Attack angle / Angle d'attaque	Thickness / Epaisseur
1	1	-1	1	•
2	1	1	1	•
3	1	-1	-1	•
4	1	1	-1	•
5	2	-1	1	•
6	2	1	-1	•
7	2	-1	-1	•
8	2	1	1	•
9	3	1	-1	•
10	3	1	1	•
11	3	-1	1	•
12	3	-1	-1	•
13	4	1	-1	•
14	4	-1	-1	•
15	4	1	1	•
16	4	-1	1	•
17	5	-1	-1	•
18	5	-1	1	•
19	5	1	1	•
20	5	1	-1	•

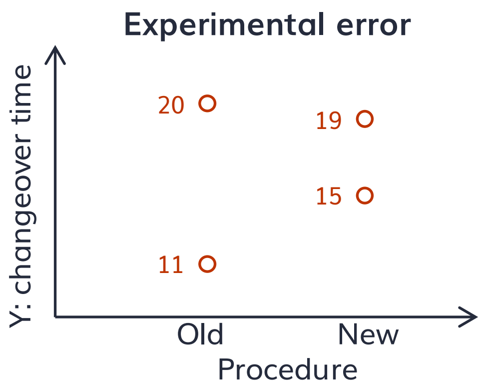
Basic Principles – Blocking

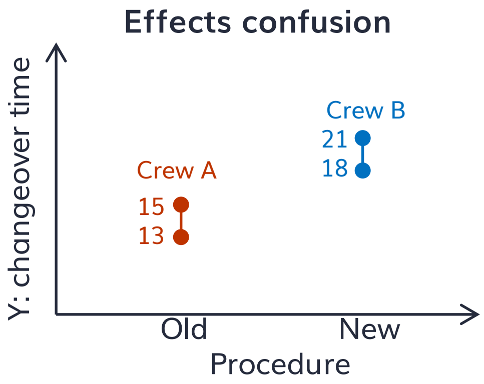
- Examples in industry :
  - paper grade
  - operators
  - day of production
  - Etc.
- Blocking allows to take into account factors which have an effect on the response variable but which are difficult or impossible to control. It is important to consider the blocks in the statistical analysis of the results.
  - There are experiments with complete or incomplete blocks where "incomplete" simply means all treatments do not occur within the same block.

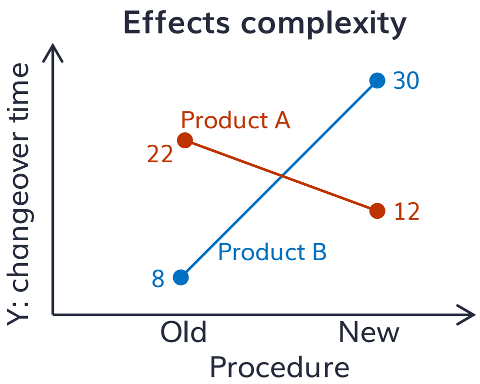
But why do we need replication, repetition, randomization and blocking?

To address the very practical challenges we face with experiments!

Challenges with experiments







Challenges with experiments

Challenge	Observation	Cause	Remedy
Experimental error	Same treatment (X), different results (Y)…	Because of uncontrollable and unknown factors (Z)	Replicate treatments! Repeat tests of the run! Use "blocking" factors
Confusion between effects	Two factors changed simultaneously	Which factor is really accountable for the impact?	Randomize trials
Complexity of effects studied	Effect(A+B) is different than Effect(A)+Effect(B)	Factors interact (synergy), impacts are not linear	Use the proper design approach and change more than one factor at a time

In other words...

Reduce the experimental error

Replications
Repetitions
Blocking

More trials!

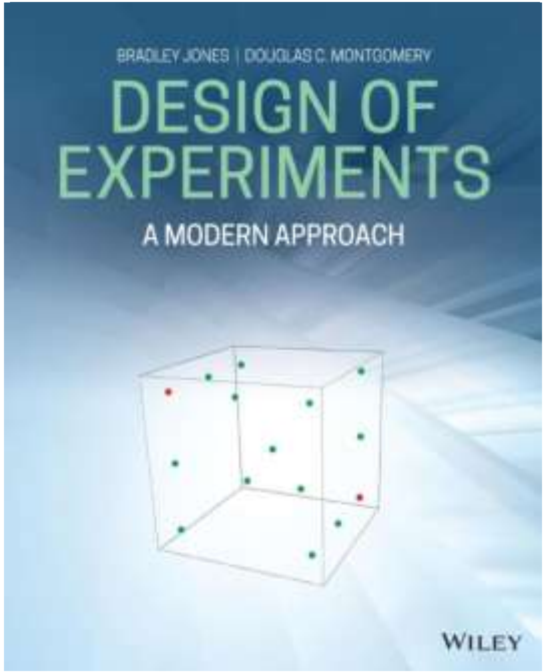
More planning complexity!

Study complex relationships

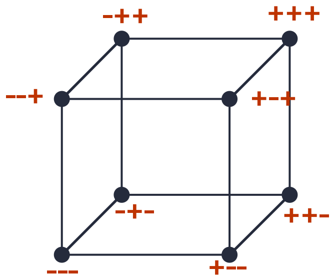
Avoid effect confusion

Vary more than one factor at the time

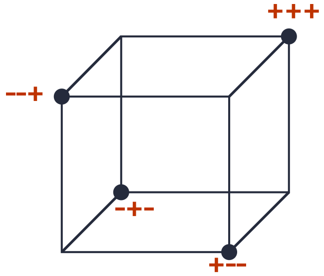
Statistical rescue!



Well-known designs

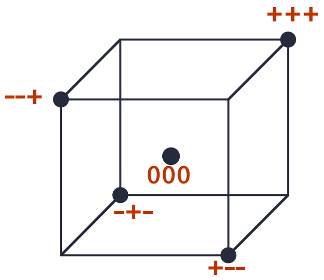


Full factorial designs



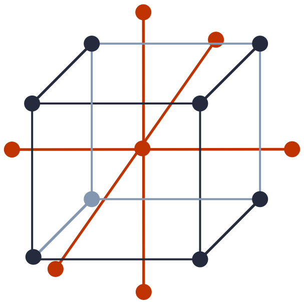
Fractional factorial designs

Screening designs
(main effects and interactions, selection of most important Xs)

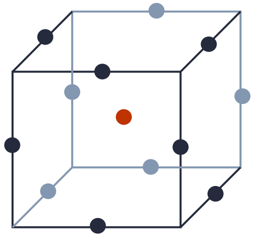


... with central point

Response surface
(interactions and quadratic effects, for optimization)



Central composite (CC)



Box-Behnken (BB)

Applying them in Pulp & Paper

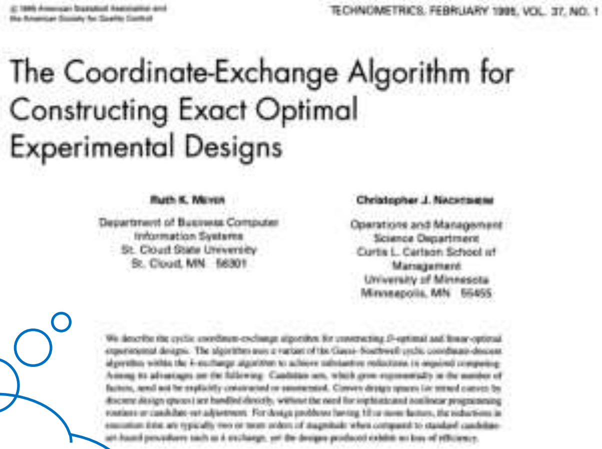
Why they are used

- Proven efficient multivariate experimental designs
- Widely taught since the 1920s, relatively simple to understand
- Could be designed by hand, very common in statistical software

Why they are not used

- Number of trials too high compared to budget$
- Data collection strategy forcing to perform trials not required for the desired effects
- Need a response surface design (CC and BB) to capture specific quadratic terms

Statistical rescue!


TECHNOMETRICS, FEBRUARY 1995, VOL. 37, NO. 1

The Coordinate-Exchange Algorithm for Constructing Exact Optimal Experimental Designs

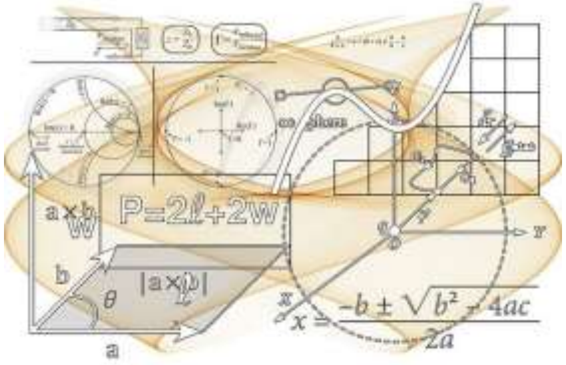
Ruth K. Meyer
Department of Business Computer Information Systems
St. Cloud State University
St. Cloud, MN 56301

Christopher J. Nachtsheim
Operations and Management Science Department
Curtis L. Carlson School of Management
University of Minnesota
Minneapolis, MN 55455

Novel strategy in the late 1990s to reduce the number of trials to the bare minimum!

But it's a bit more complicated science...

The power of model-driven DOE



The context

- Following a multivariate analysis made by a strange but apparently smart consultant, the company asks for a second advice to a loud-laughing statistician
  - The mandate consists in validating some interesting cause-and-effect relationships
  - These process parameters have been selected:
    - Lip opening
    - Headbox pressure
    - Calendar pressure
    - Calendar draw
    - Specific energy
    - High pressure steam

A “standard” strategy

- Choose the combination of 6 factors that seems the best, then try it
- If it doesn’t yield to the expected results, move a factor level in the best direction and try it
- If it doesn’t yield to the expected results, move another factor in the same fashion
- Etc..

A little less "standard" strategy

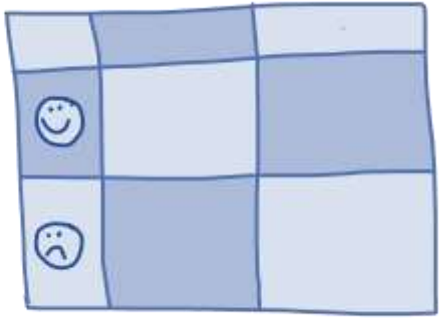
≡ Since there are 6 factors, we could consider the following experiments (without replications):

- Full factorial designs $2^6 = 64$ trials

- Fractional factorial design $2^{6-1} = 32$ trials (Resolution 5)
- Fractional factorial design $2^{6-2} = 16$ trials (Resolution 4)
- Fractional factorial design $2^{6-3} = 8$ trials (Resolution 3)

Pros and cons of the strategy

- PROS
  - Known designs
  - Known aliasing structure
  - Balanced plans



- CONS
  - Too many trials since we estimate terms with a negligible interest
  - Cannot test for specific terms – allows only testing for the presence of curvature if center points are added

A new philosophy

- Let's think the opposite way!
  - Prepare the model you wish to obtain
  - Define which trials are required to fit that model

(Serious) Note

A software is required to prepare custom designs. In this presentation we will use JMP.

jmp Statistical Discovery.™ From SAS.

Illustrated methodology

1. Identify the experiment's goal
   - Evaluate the impact of the 6 factors on the Y (paper quality %)
2. Select the factors, their levels and the difficulty (response time) to change from a level to the other during the experiment
   - X1: Lip opening [9.5 ; 11.5]
   - X2: Headbox pressure [28 ; 35] - **HARD**
   - X3: Calendar pressure [8 ; 43]
   - X4: Calendar draw [0.03 ; 0.40]
   - X5: Specific energy [3 ; 17]
   - X6: High pressure steam [700 ; 725]
3. Identify which parameters will have to remain constant during the experiment
4. Specify the desired model: main effects, interactions, quadratic terms
   - Main effects: X1, X2, X3, X4, X5, X6
   - Interactions: X1*X4, X1*X6, X2*X6
   - Quadratic: X4*X4
5. Build the plan
6. Verify the power

DOE Template

- A template to help in designing experiments

*To get the template, connect with Martin or Vincent on LinkedIn

Main Objective:

Secondary Objectives:

Response Variables :

Name	Units	Minimum	Maximum	Target	Response goal	Importance	Quality of measurement
Y1:							
Y2:							
Y3:							

Factor and Levels:

Name	Units	Low Level	Med Level	High Level	Other levels	Easiness to change	Risk of directly jumping from low to high level or vice-versa
X1:							
X2:							
X3:							
X4:							
X5:							

Factors held constant:

Name	Units	Desired setting	Acceptable variation
C1:			
C2:			
C3:			
C4:			

Noise factors

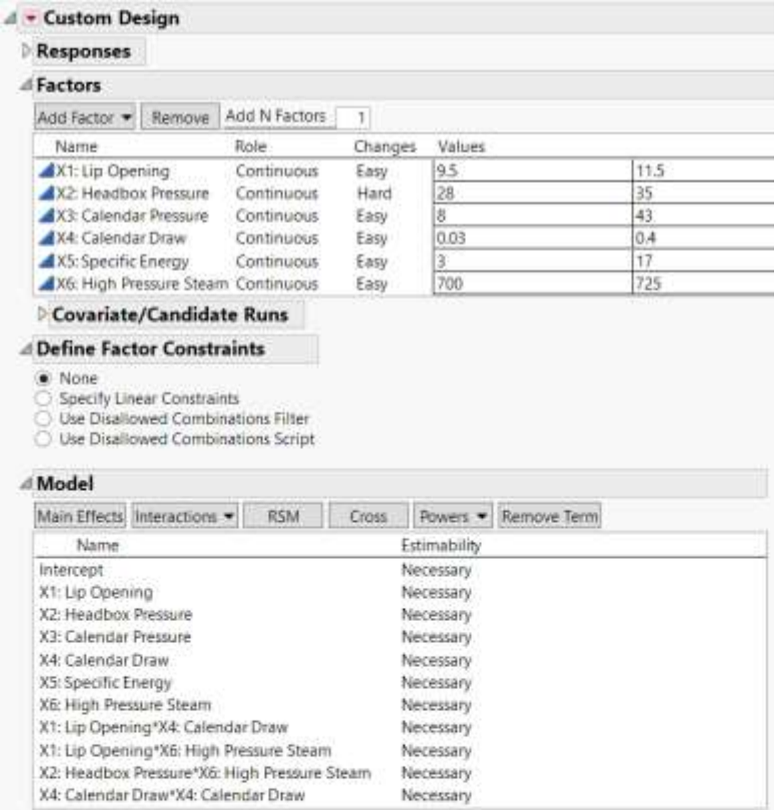
Name	Units	Will be measured during the experiment ? (y/n)
N1:		
N2:		
N3:		

Name	Units	Will be measured during the experiment ? (y/n)
N4:		
N5:		
N6:		

Model: Main effects +

Specify interactions and quadratic effects:

In JMP...


Custom Design
Responses
Factors
Add Factor
Remove
Add N Factors 1
Name Role Changes Values
X1: Lip Opening Continuous Easy 9.5 11.5
X2: Headbox Pressure Continuous Hard 28 35
X3: Calendar Pressure Continuous Easy 8 43
X4: Calendar Draw Continuous Easy 0.03 0.4
X5: Specific Energy Continuous Easy 3 17
X6: High Pressure Steam Continuous Easy 700 725
Covariate/Candidate Runs
Define Factor Constraints
None
Specify Linear Constraints
Use Disallowed Combinations Filter
Use Disallowed Combinations Script
Model
Main Effects
Interactions
RSM
Cross
Powers
Remove Term
Name Estimability
Intercept Necessary
X1: Lip Opening Necessary
X2: Headbox Pressure Necessary
X3: Calendar Pressure Necessary
X4: Calendar Draw Necessary
X5: Specific Energy Necessary
X6: High Pressure Steam Necessary
X1: Lip Opening*X4: Calendar Draw Necessary
X1: Lip Opening*X6: High Pressure Steam Necessary
X2: Headbox Pressure*X6: High Pressure Steam Necessary
X4: Calendar Draw*X4: Calendar Draw Necessary

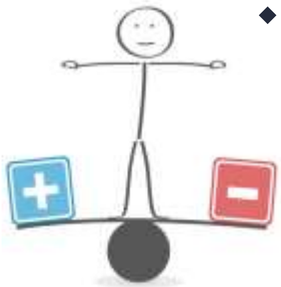
Pros & cons of model-driven DOE

PROS

- Minimizes the number of trials
- Triggers a discussion on the assumptions we want to test
- Makes it easy to add specific quadratic terms

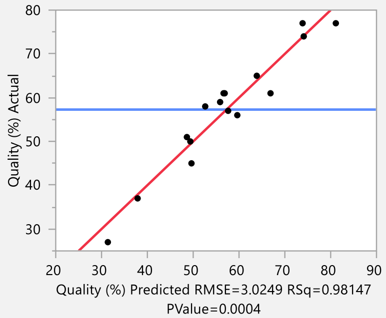
CONS

- Limits the chances to discover “by luck” an important effect
- The notion of “aliasing” has to be well understood
- The aliasing structure can be complex



Results analysis in JMP…

Actual by Predicted Plot



Effect Summary

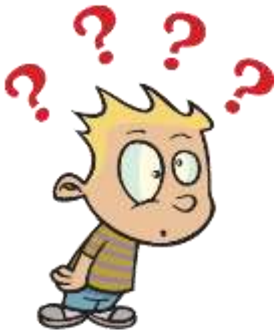
Source	LogWorth		PValue
X1: Lip Opening(9.5,11.5)	4.102		0.00008
X4: Calendar Draw*X4: Calendar Draw	2.607		0.00247
X1: Lip Opening*X4: Calendar Draw	2.439		0.00364
X6: High Pressure Steam(700,725)	2.371		0.00425
X3: Calendar Pressure(8,43)	2.309		0.00491
X1: Lip Opening*X6: High Pressure Steam	2.179		0.00662
X4: Calendar Draw(0.03,0.4)	1.872		0.01343 ^

Summary of Fit

RSquare	0.981466
RSquare Adj	0.965248
Root Mean Square Error	3.024897
Mean of Response	57.25
Observations (or Sum Wgts)	16

Think about it

- Do your trials allow statistically testing the assumptions you believe to be important?
  - If the answer is “NO”, using the model-driven optimal experiment design approach could be beneficial:
    - It forces using scientific thinking
    - It efficiently tests the most likely possible assumptions
    - It is not tied to a set of pre-defined models like “ordinary” DOEs (screening designs and response surfaces)



A theory is just a mathematical model to describe the observations.

Final thoughts

About testing variation

- Don't waste you DOE because of testing variation
  - The hard part is often to perform the experiment
  - It could be discouraging to realize after-the-fact that the results are not valid because of variation in the lab
  - To reduce that risk it is recommended to:
    - Try using a single tester and a single instrument
    - If not possible, use the DOE principles to distribute the tests among the testers/instruments to ensure no factor is confounded with testers/technician

Factor A	Factor B	Tester	Tester
-	+	X	Y
-	-	Y	X
+	+	Y	X
+	-	X	Y

About testing variation

≡ Factor A is confounded with the Tester

Factor A	Factor B	Tester
-	+	X
-	-	X
+	+	Y
+	-	Y

Factor A	Factor B	Tester
-	+	X
-	-	Y
+	+	Y
+	-	X

≡ The interaction AB is confounded with the Tester

≡ If you can do repetitions of the test, there is no confusion … and it improves measurement precision!

Factor A	Factor B	Tester	Tester
-	+	X	Y
-	-	Y	X
+	+	Y	X
+	-	X	Y

Sequential adaptative approach

≡ A series of smaller experiments is preferable to one bigger holistic “all eggs in the same basket” strategy:

time

Screening
- ≡ Many factors
- ≡ Limited information (list)

Optimization
- ≡ Few factors
- ≡ Lots of information (model, adjustments)

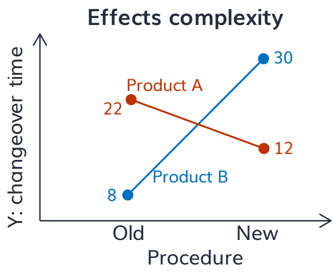
Confirmation
- ≡ Results validation (long-term sustainability)
- ≡ On-line or off-line

Sample size matters

- Prior to performing an experiment, it could be a good idea to estimate the "power", i.e. the probability the experiment will allow to detect an impact of δ from a factor
  - This probability is function of the expected noise in the results
  - The power can be improved by increasing the number of runs or by reducing the measurement error with repetitions of the test in the lab

I am still hesitant to perform a multivariate experiment

- I am not a statistician … how can I do this ???
  - With the help of a software and some coaching, it is possible to learn this rapidly
- It is too complex to change multiple factors at the same time
  - You can limit the number of factors if you want … but the only way to identify interactions is to change the factors simultaneously and … there are often interactions in a P&P process



- If you want to get into AI, multivariate experiments can be an efficient way to feed the "machine" with excellent data to get better results faster …

Planning saves effort!

A good planning and design
of the data collection process

+

A strong statistical analysis

=

**Much faster and
accurate convergence**

Différence is a society offering coaching, consulting and training services in statistic, data science, simulation and continuous improvement.

We promote the use of quantitative tools that can be applied at the different steps of an improvement and variability reduction project.

Powerful methods

Adapted approach

Combining hard work with fun

For more information, you can contact:

Martin Carignan, M.Sc., MBA
Principal Associate
514-795-8000
mcarignan@difference-gcs.com
linkedin.com/in/martincarignan

Vincent Béchard, MASc.
Associate, Analytical Decision Specialist
438-521-5829
vbechard@difference-gcs.com
linkedin.com/in/vincentbechard