2024-11-07 | Marianne Toft

novonesis

DOE for quantification of assay variation

Bio for Marianne Toft

- Education
  - Master in Food Science and Technology from UCPH (2000)
  - PhD in Chemometrics and Spectroscopy from UCPH (2004)
  - Master in Applied Statistics and Machine Learning from DTU (2021)
- Employment
  - Application specialist at MKS Umetrics (2004-2009)
  - Senior Statistician at Novozymes/Novonesis (2009-)

novonesis

Agenda

- Introduction to Novonesis and how we use DOE
- Introduction to Measurement System Analysis (MSA)
- MSA in QC assay development
- MSA for bread baking

With 100+ years of innovation as our foundation, we will keep delivering transformative solutions

novozymes®

CHR HANSEN

novonesis

Microbiology is at the heart of Novonesis

We leverage the power of our planet's microbes, enzymes and functional proteins with our innovative technology, enabling healthier lives and a healthier planet.

Across 30+ industries, our innovations both help businesses achieve their commercial targets and balance the needs of people and our planet

Biosolutions for planetary health

- Household care
- Bioag & plant health
- Bioenergy
- Animal health & nutrition
- Carbon capture
- Plastic recycling
- Leather & textiles
- Pulp & paper
- Professional cleaning
- Grain & starch

Biosolutions for food & health

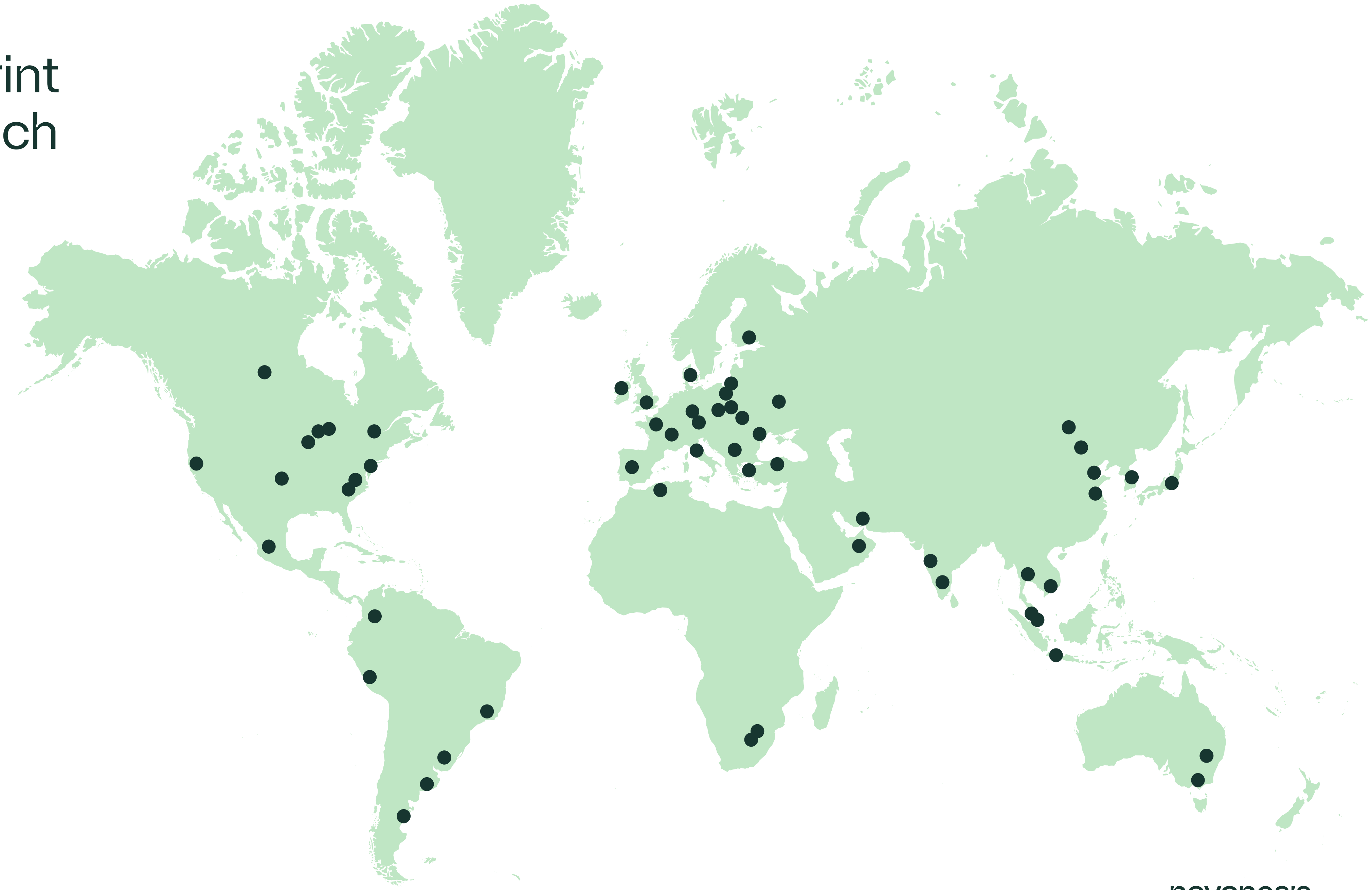
- Plant-based foods
- Dairy
- Infant nutrition
- Baking
- Dietary supplements
- Meat, seafood & culinary
- Functional foods
- Advanced protein solutions
- Beverages

Global footprint with local reach

~ 10K
Employees

~40
R&D and
application centers

>20
Manufacturing sites



novonesis

DOE is used in all areas of Novonesis


Basic R&D


Applied Research


Process dev and opt


Quality Control

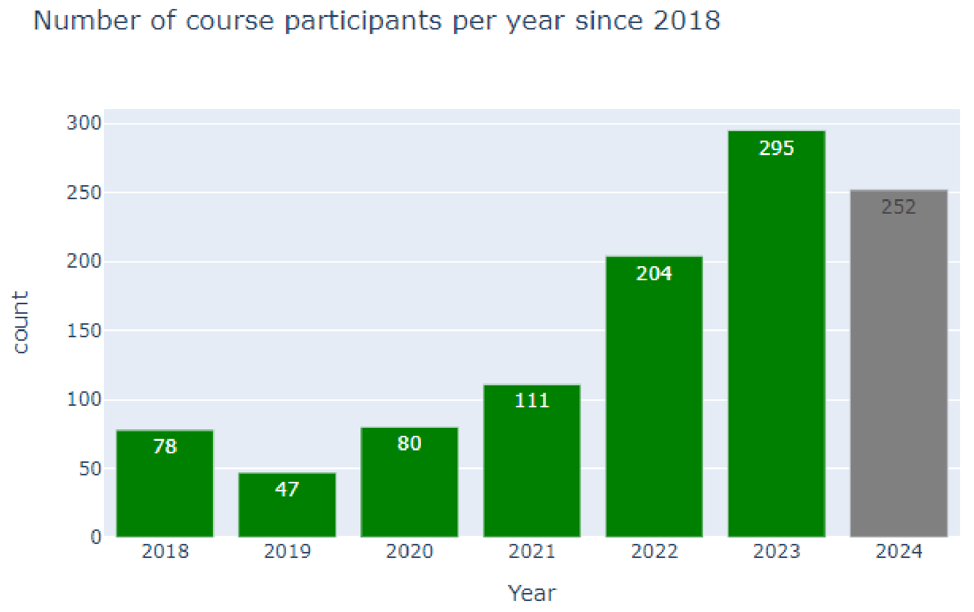

Quality Mangement

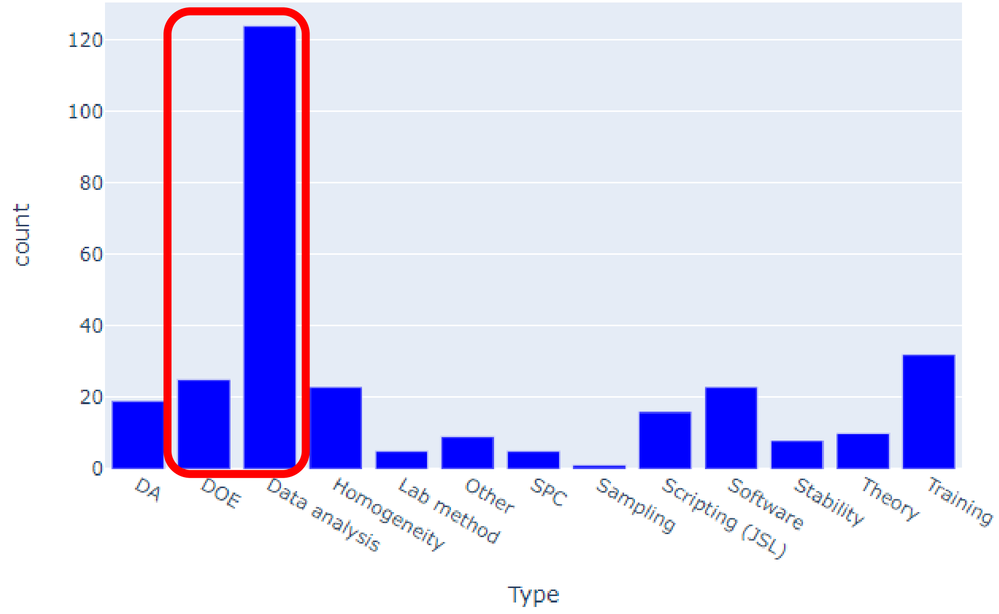

Tech Service

Learning DOE in Novonesis

- StatTeam teaches DOE using JMP
  - Virtual courses (Videos, exercises, Teams Q&A meetings)
  - 100 colleagues joined the DOE course in the last 2 years

- StatTeam support scientists after the DOE course
  - 150 tasks in the last 2 years related to help setting up DOE or analyzing data

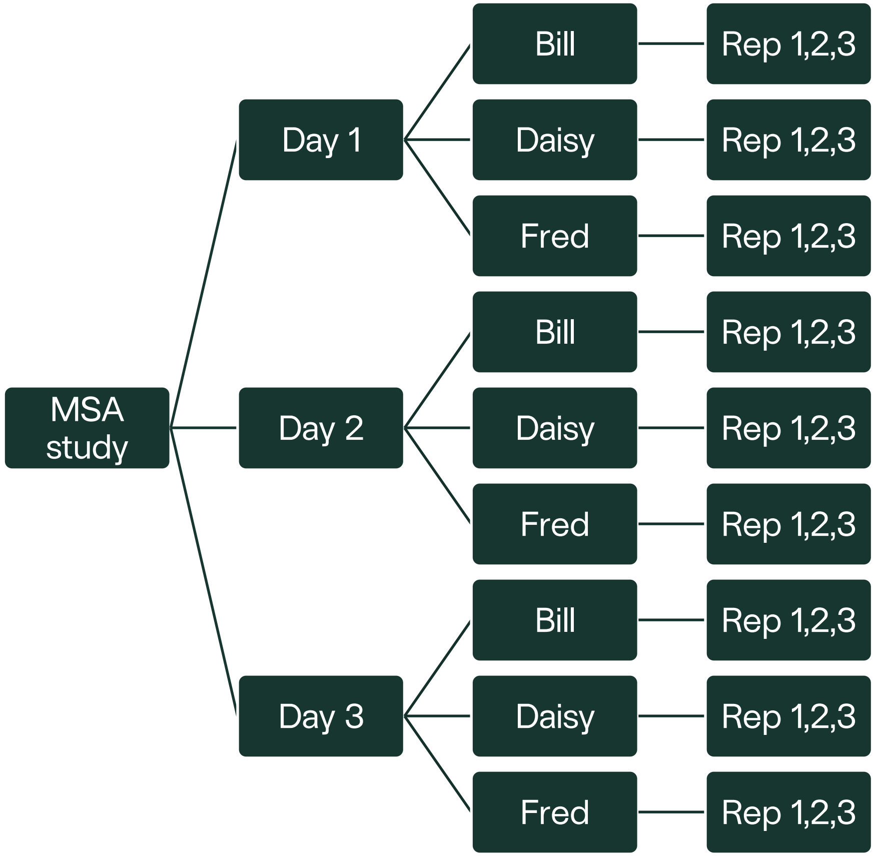




What is Measurement System Analysis (MSA)?

Wikipedia:
*A **measurement system analysis** (**MSA**) is a thorough assessment of a measurement process, and typically includes a specially designed* experiment *that seeks to identify the components of variation in that measurement process.*

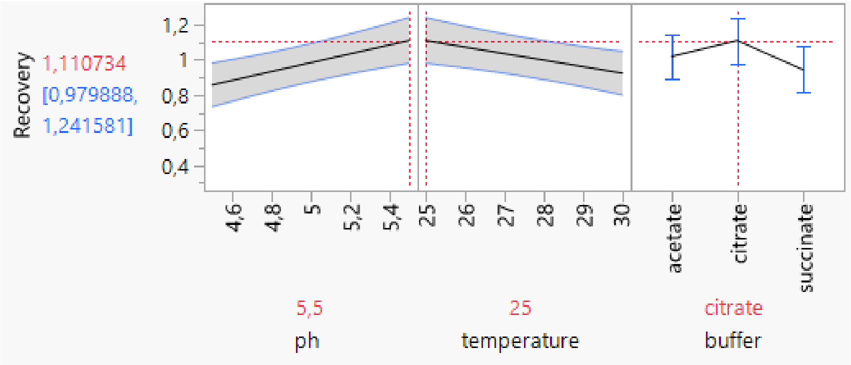
- Typically, simple Full Factorial DOEs are used, e.g.
  - 3 Days **x** 3 Technicians **x** 3 replicate measurement



Difference between classical DOE and DOE for MSA

Classical DOE

- Purpose is to solve a problem, find optimal settings, understand a process
- Factors are modeled as Fixed effects, meaning we get estimates of slope for continuous factors and estimates of levels for categorical factors and statistical tests for whether these have significant effect on the response or not
- Result = μ + pH + temperature + buffer + ε



DOE for MSA

- Purpose is to quantify the variation stemming from different sources in a measurement system process
- Factors are modeled as Random effects, meaning we are not interested in the levels themselves, but rather in the variation between levels of a factor.
- Result = μ + Day&Random + Operator&Random + ε

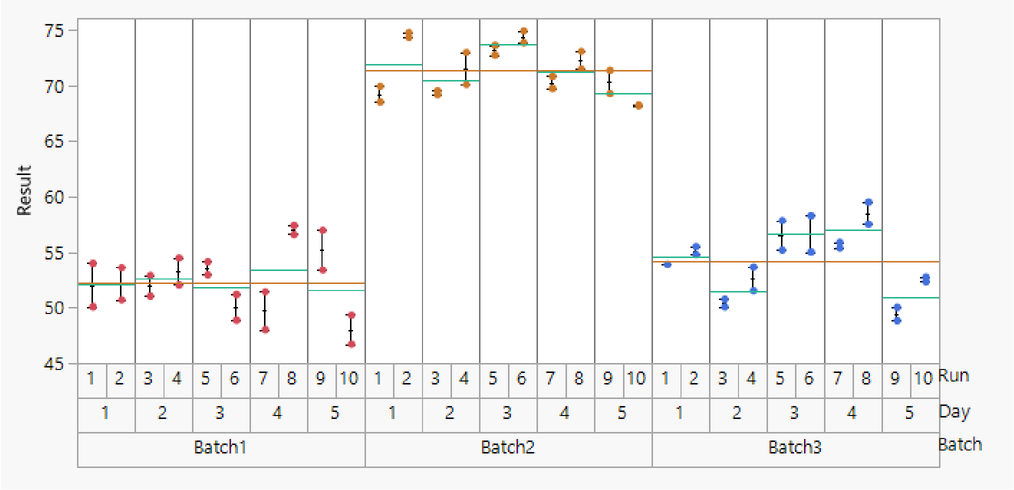
Random Effect	Variance	Pct of Total
Day	0,1153	7,43
Operator	0,1221	7,87
Residual	1,3137	84,70
Total	1,5511	100,00

novonesis

MSA examples from Novonesis

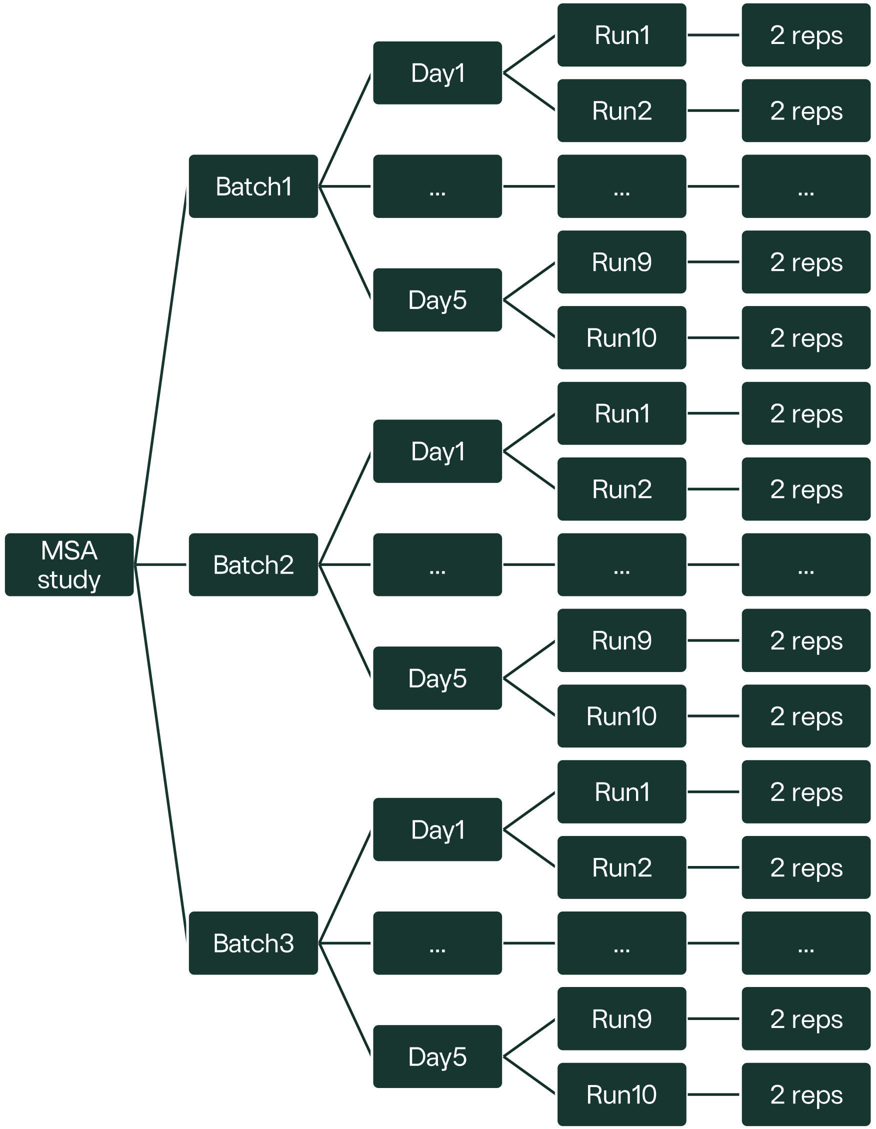
MSA in QC assay validation

- 3 batches **x** 5 days **x** 2 runs per day **x** 2 replicates per run



- Model: Result = μ + Batch + Day&Random + Run&Random + ε

Random Effect	Variance	Pct of Total	CV
Day	0,76	9,83	1,47
Run	2,16	28,05	2,48
Residual	4,78	62,11	3,69
Total	7,70	100,00	4,68



novonesis

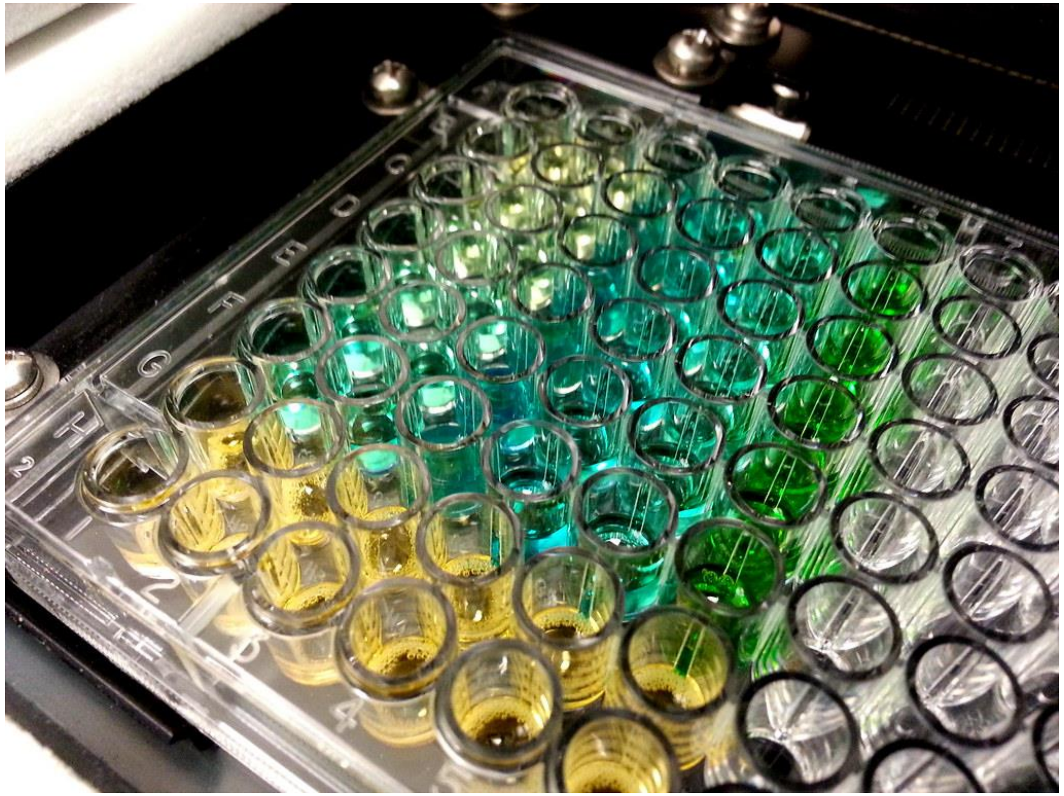
MSA in QC assay development

- New QC method, early development

- 3 technicians **x** 5 microtiter plates (96 wells)
  - 6 batches **x** weighed out 2 times **x** pipetted into 8 wells

- Model: Result = μ + Batch +
  Technician&Random +
  Plate[Technician]&Random +
  Weigh[Sample,Technician,Plate]&Random + ε

Random Effect	Variance	Pct of Total
Technician	0,0041	21,35
Plate[Technician]	0,0025	13,27
Weigh[Sample,Technician ,Plate]	0,0041	21,41
Residual	0,0083	43,97
Total	0,0190	100,00



MSA for a baking process

- Novonesis produces enzymes that are used in baking to improve various parameters of bread

- When developing new enzymes, we test them in baking trials

- To be able to optimally plan these trials, we need to quantify the variation in the different steps of the baking process

Baking process

Ingredients

Mixing

Bench time

Proofing

Baking

Mixing:

- There are 4 **mixers** (A/B/C/D).
- Each run of a mixer results in 1 dough.
- The baker can run the 4 mixers and thus produce 4 doughs in parallel, but then he will make all the breads/buns and bake them together.

Bench time:

- Each dough is divided into e.g. 6 loaves or 30 buns, which are all placed on one tray

Proofing:

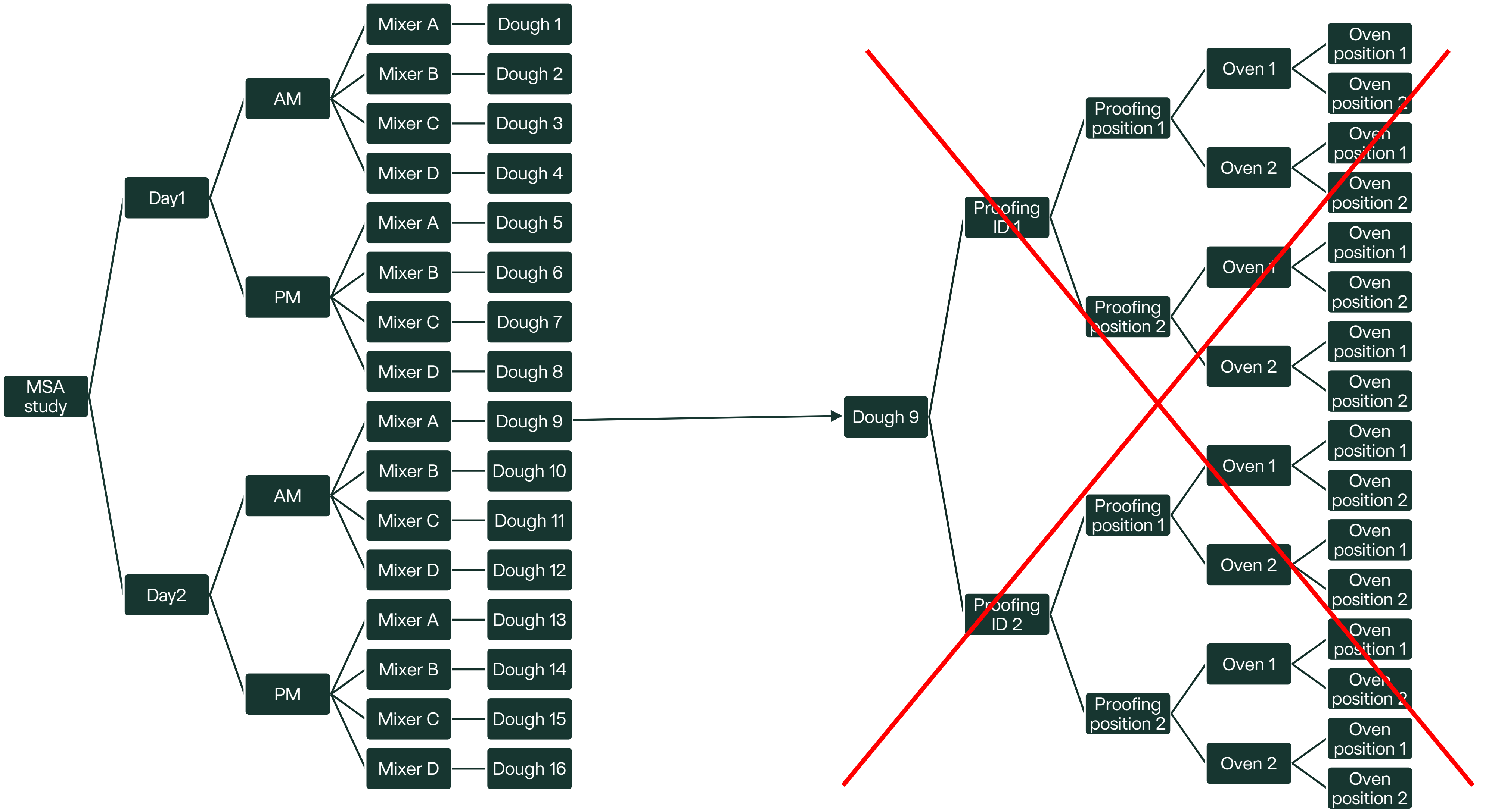
- There are 2 **Proofing chambers**
- They each have 2 **Proofing positions** (top shelf, bottom shelf)

Baking:

- There are 2 **Ovens**
- They each have 2 **Oven positions** (top shelf, bottom shelf)

This entire process takes ½ **day**, so the baker can make 2 ”**rounds**” per day, thus 8 mixer runs and 8 doughs in total per day

We cannot use a classical DOE for the MSA for baking



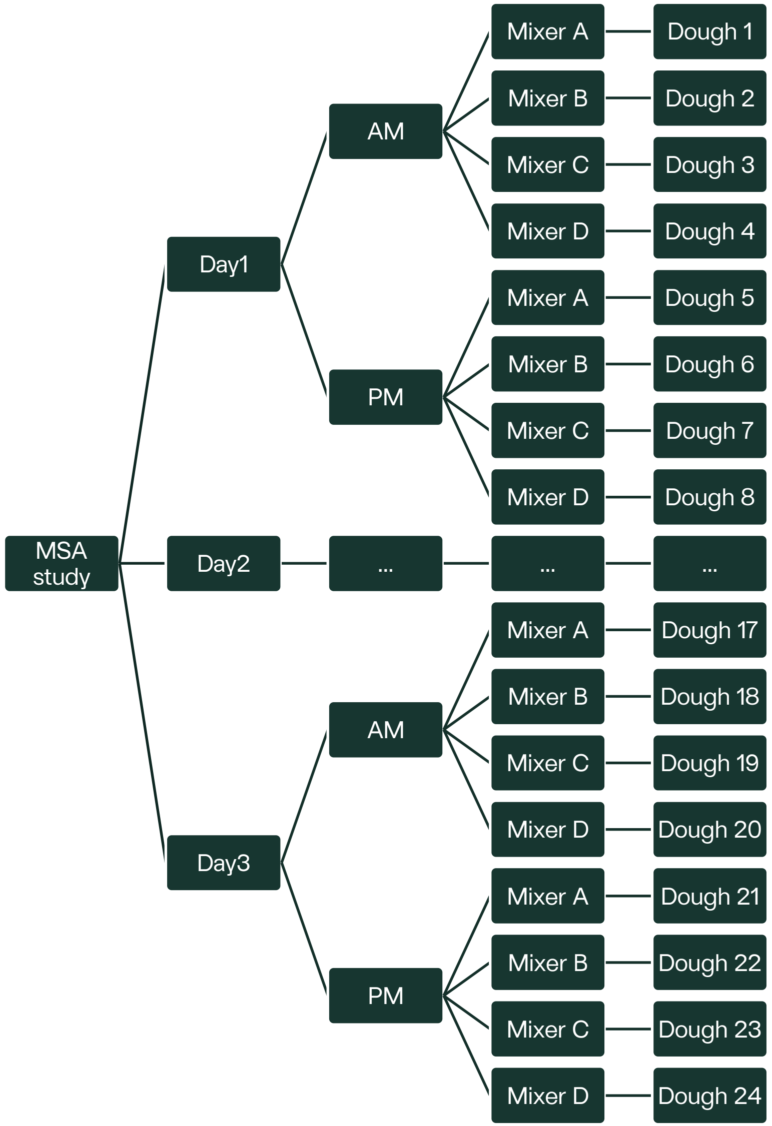
novonesis

We will use an Optimal design instead

- Optimal designs
  - Are very flexible
  - Can give exactly the number of runs we want
  - The model gives the design

- We need all combinations of Day, Round and Mixer
  - Exactly 24 runs, i.e. 24 doughs
  - AM and PM used once per Day
  - Mixer A, B, C, D used exactly once in each Round

- But for the other factors, we need the Optimal design to tell us which settings to use for every tray, so that we can estimate all the variance components



Tweaking the design

- We cannot specify the model we really want to fit
  - Random Effects model
  - Many terms should be nested
- We use crossing instead of nesting:
  - Add interactions to the model
  - Ensures balance and avoids confounding of factors
- But to fit this model we need too many runs
- Solution:
  - Day*Round, Day*Mixer and Round*Mixer ensure that we get a Full Factorial for these three factors
  - But all the other interactions we set to 'If Possible'
  - This we can do in 24 runs

Factors

Add Factor | Remove | Add N Factors 1

Name	Role	Changes	Values			
Day	Categorical	Easy	1	2	3	
Round	Categorical	Easy	AM	PM		
Mixer	Categorical	Easy	A	B	C	D
PID	Categorical	Easy	1	2		
PP	Categorical	Easy	1	2		
OID	Categorical	Easy	1	2		
OP	Categorical	Easy	1	2		

Model

Main Effects | Interactions | RSM | Cross

Name	Estimability
Intercept	Necessary
Day	Necessary
Round	Necessary
Mixer	Necessary
PID	Necessary
PP	Necessary
OID	Necessary
OP	Necessary
Day*Round	Necessary
Day*Mixer	Necessary
Day*PID	Necessary
Day*PP	Necessary
Day*OID	Necessary
Day*OP	Necessary
Round*Mixer	Necessary
Round*PID	Necessary
Round*PP	Necessary
Round*OID	Necessary
Round*OP	Necessary
Mixer*PID	Necessary
Mixer*PP	Necessary
Mixer*OID	Necessary
Mixer*OP	Necessary
PID*PP	Necessary
PID*OID	Necessary
PID*OP	Necessary
PP*OID	Necessary
PP*OP	Necessary
OID*OP	Necessary

Design Generation

Group runs into random blocks

Number of Replicate Runs: 0

Number of Runs:

Minimum 52

Default 60

User Specified 52

Make Design

Model

Main Effects | Interactions | RSM | Cross

Name	Estimability
Intercept	Necessary
Day	Necessary
Round	Necessary
Mixer	Necessary
PID	Necessary
PP	Necessary
OID	Necessary
OP	Necessary
Day*Round	Necessary
Day*Mixer	Necessary
Round*Mixer	Necessary
Day*PID	If Possible
Day*PP	If Possible
Day*OID	If Possible
Day*OP	If Possible
Round*PID	If Possible
Round*PP	If Possible
Round*OID	If Possible
Round*OP	If Possible
Mixer*PID	If Possible
Mixer*PP	If Possible
Mixer*OID	If Possible
Mixer*OP	If Possible
PID*PP	If Possible
PID*OID	If Possible
PID*OP	If Possible
PP*OID	If Possible
PP*OP	If Possible
OID*OP	If Possible

Design Generation

Group runs into random blocks

Number of Replicate Runs: 0

Number of Runs:

Minimum 22

Default 36

User Specified 24

Make Design

Bread MSA - Analysis and conclusions

- The resulting design with 24 doughs was run and the desired model was estimated
- Largest variance contributors are Bread, Day and OvenID
- We use this to plan future designs.
- To get as reliable conclusions for our test enzymes…
  - We should bake many breads from each dough and average the quality parameter measured
  - We should block by Day, so that the day-to-day variation can be modeled
  - We should have a closer look at the ovens – is there a consistent difference between the two or is it more an ”oven run” variation we are seeing

Random Effect	Pct of Total
Day	29
Round[Day]	0
Mixer[Day,Round]	2
Proof ID[Day,Round]	9
Proof pos[Day,Round]	2
Oven ID[Day,Round]	16
Oven pos[Day,Round]	0
Bread[Day,Round,Mixer]	41
Total	100

Other applications of these complex MSA's

- Testing enzymes for baking
  - In full scale baking of loaves or buns
  - Variance components: Days, Rounds, Mixers, ProofingID, ProofingPosition, OvenID, OvenPosition, Loaves

- Testing enzymes for washing powder
  - In full scale washing machines
  - Variance components: Days, Rounds, Machines, Swatches, Measurements

- Testing bioprotection microbes
  - On plants grown in a green house
  - Variance components: Table, Row of table, Column of table, Leaf, Measurement

- Testing probiotics for animals
  - On chicken in a barn
  - Variance components: North-South, East-West, Distance from center aisle, Pen, Animal

Thank you