sanofi

Comparison of Classical and Hybrid Methodologies for Model-based Design of Experiments

Amos Lu

Senior Data Scientist

Global CMC Development, Sanofi

amos.lu@sanofi.com

Outline

- Introduction and motivation for investigating model-based design of experiments (MBDoE)
- Benchmark development
  - Ground truth CHO model
  - Iterative test strategy
- Results
  - Classical DoE (resolution III and IV fractional factorial, quadratic response surface)
  - DataHow hybrid model + (Latin hypercube, model-based design of experiments)
- Conclusions

Introduction and Motivation

- Key mandate: add value to CMC development activities through data science

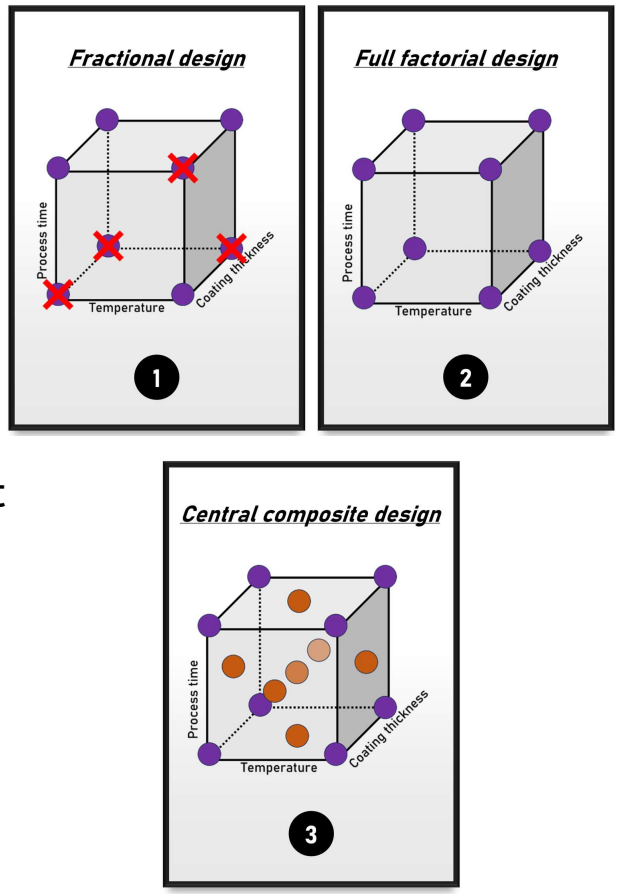
- How to get value from models (hybrid or otherwise)?
  - Reduction in number or cycles of experiments to get acceptable titer and product quality (PQ)
  - Improvement of titer, PQ, reduction of COGS through numerical optimization

- Goal: Speed up learning and optimization cycle through model-based design of experiments

State of the art: Traditional DoE

- Traditional DoE is the standard of practice
  - Fractional factorial and response surface methodologies (RSM)
  - Process scientists pick relevant factors and ranges
  - Design experiments based on classical “canned” designs
  - Experiments allow for factors to be refined
  - Higher resolution experiments may be designed for further refinement

- Benefits: Causal identification of input-output relationships
- Limitations of these models
  - Limited representation of nonlinearity and dynamics
  - Input-output relationship ignores mass balance and cell-specific rates
  - Model uncertainty not directly handled (indirect factor screening)

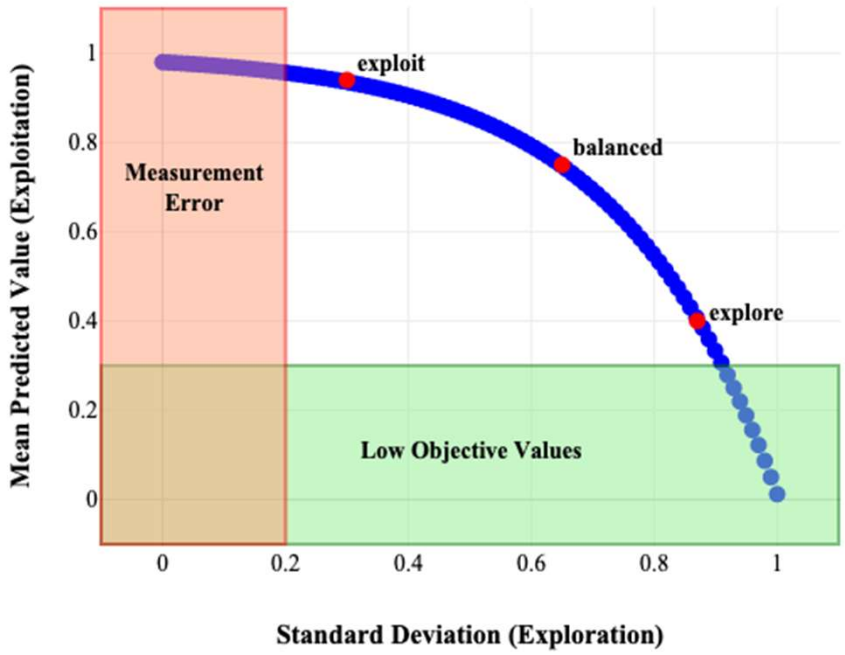


https://www.experimentaldesignhub.com/blog/advanced-doe-plans-part-1

Moving to a model-based DoE strategy

DATAHOWLAB

- Approach within DataHow
  - Latin hypercube sampling (LHS) for initial set
  - Multi-objective optimization to balance between uncertainty reduction (explore) and optimization (exploit)
- Benefits
  - Use of bioreactor mass balances converts problem into a smoother cell-specific derivative space
  - Naturally handles dynamics
  - Direct uncertainty description
- Unknowns
  - How do they compare to traditional DoE?

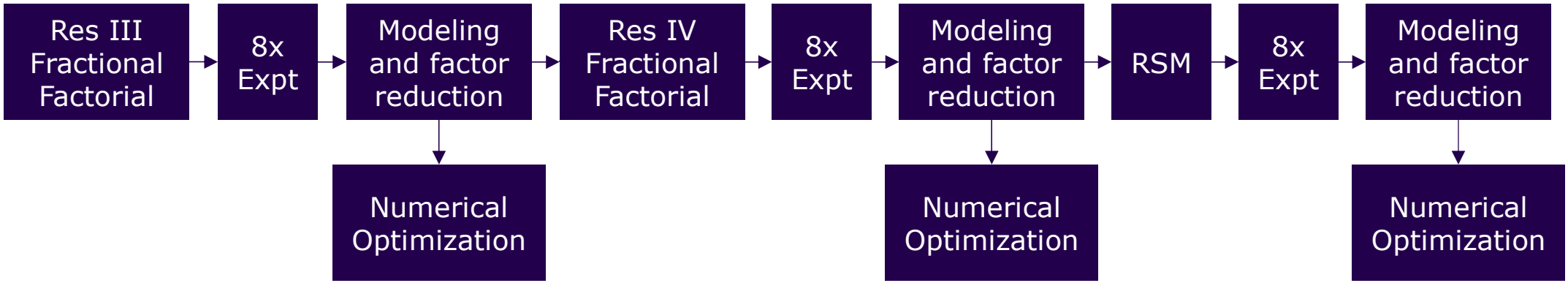


sanofi

Ideal Empirical Strategy

Does the DataHow hybrid MBDoE outperform classical DoE in a media optimization setting?

Control:

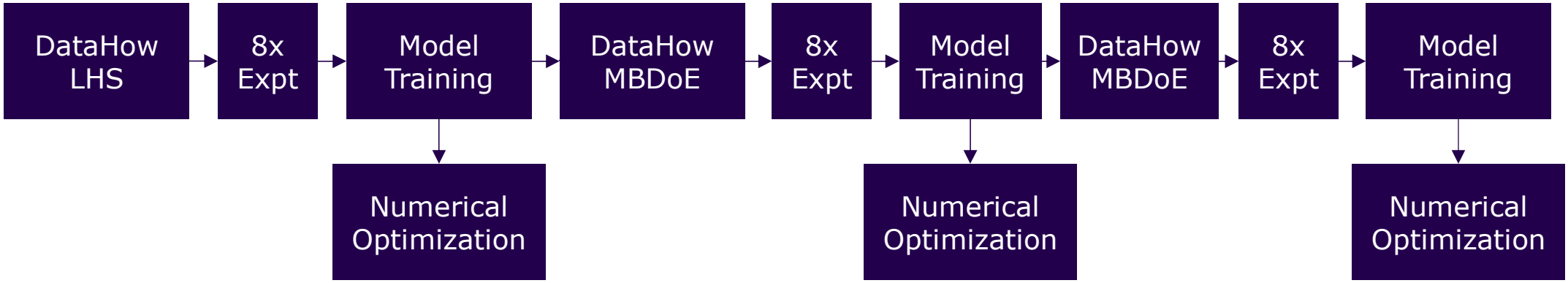


Record final optimized titer at each stage

Ideal Empirical Strategy

Does the DataHow hybrid MBDoE outperform classical DoE in a media optimization setting?

Treatment:



Record final optimized titer at each stage

Challenges with Empirical Comparison

- High experimental burden (3 cycles * 8 experiments * 2 arms) for a single comparison

- Difficult to study non-ideal behavior
  - Analytical and biological variability
  - Block effects
  - Contamination events

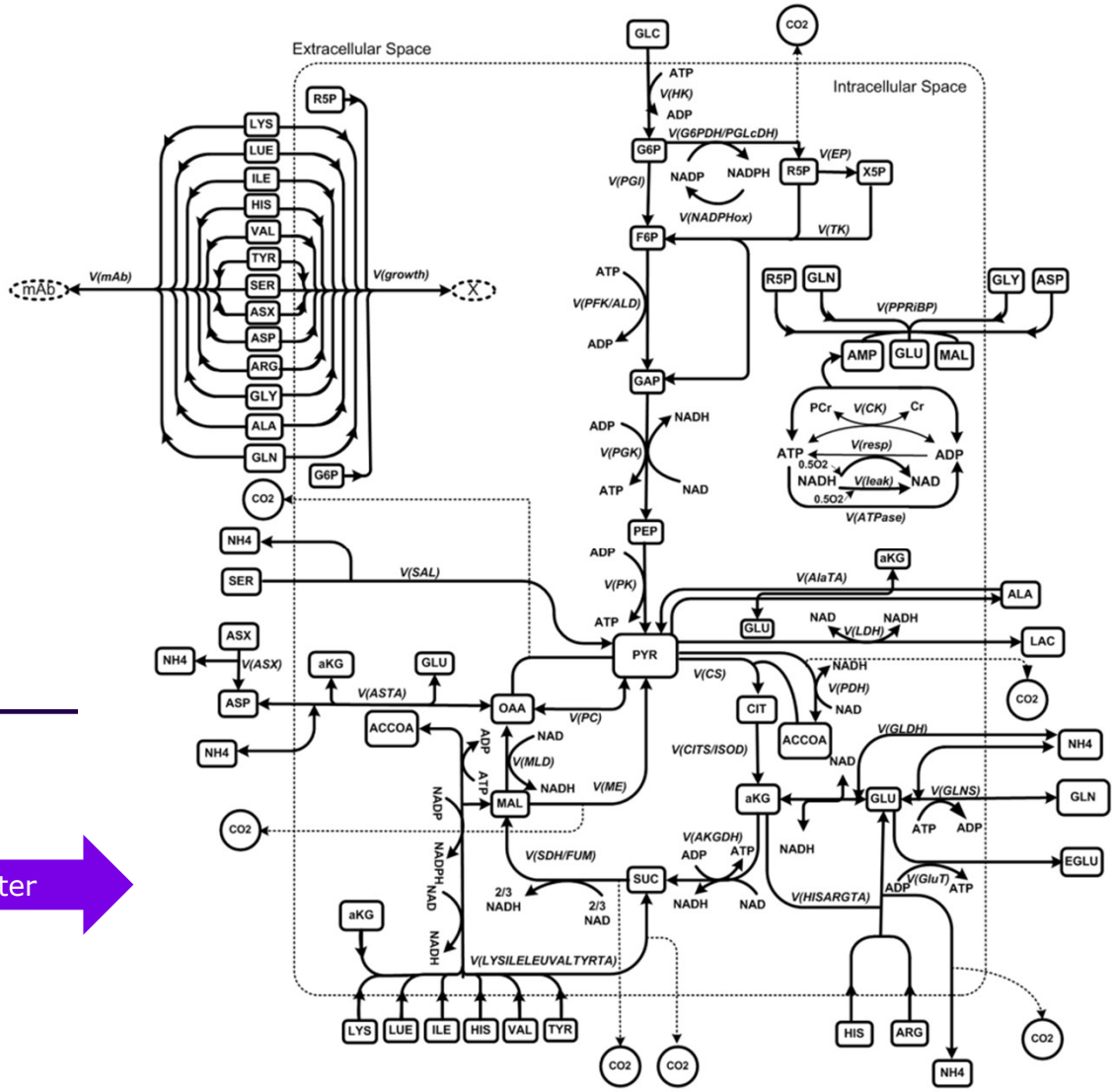
- No access to ground truth, difficult to explain *why* the models behave the way they do

- Solution: Develop a ground truth model that replicates experimental work
  - Model does not need to accurately replicate Sanofi cell culture results
  - Only needs to capture representative dynamics, nonlinearity, and smoothness
  - Model can be used for evaluation of future model-based DoE tools

Ground Truth Model

Considered a combination of two closely related models from literature (Robitaille/Ghorbaniaghdam) modeling mAb production from CHO cells.

Modified to be fed-batch system, include cell death, and extended to 14 days in order to be closer to real world cell cultures.



1) 2015, Robitaille, et al.; PLOS One
2) 2014, Ghorbaniaghdam, et al.; PLOS One

Ground Truth Model (Model Rates)

Reaction: GLC+ATP→G6P+ADP

$$v(HK) = v_{\max HK} * \frac{GLC}{K_{mGLC} + GLC} * \frac{\frac{ATP}{ADP}}{K_{m\frac{ATP}{ADP}} + \frac{ATP}{ADP}} * \frac{K_{iG6P}}{K_{iG6P} + G6P}$$

↳ Inhibition Terms

Reaction : PEP+ADP→PYR+ATP

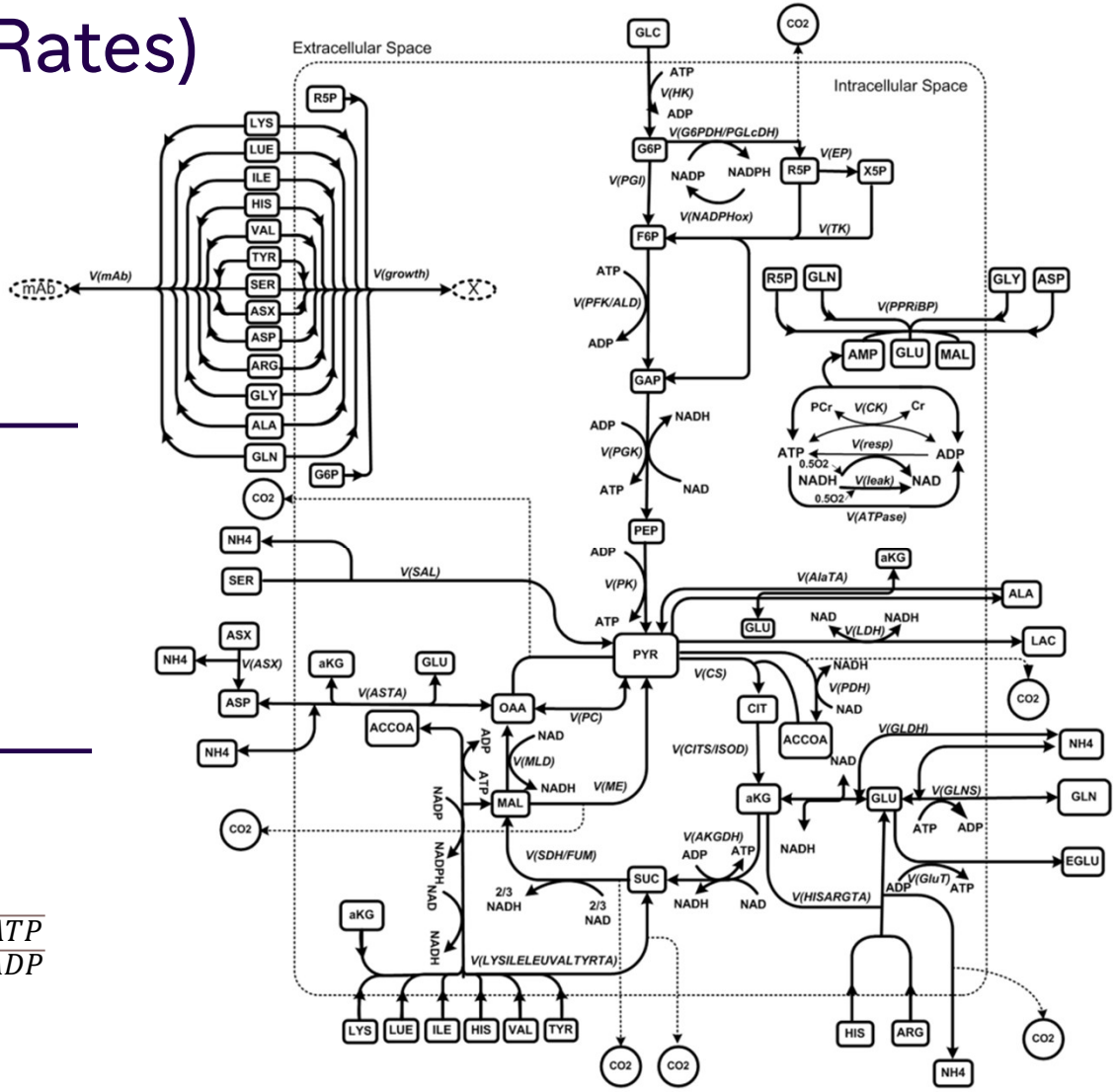
$$v(PK) = v_{\max PK} * \frac{PEP}{K_{mPEP} * \left(1 + \frac{Ka_{F6P}}{F6P}\right) + PEP} * \frac{\frac{ADP}{ATP}}{K_{m\frac{ADP}{ATP}} + \frac{ADP}{ATP}}$$

↳ Activation Terms

Reaction : 0.01GLU +0.01GLN+…+0.0145GLY+4ATP→mAb

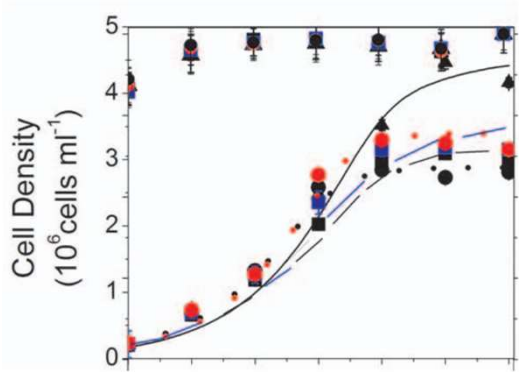
$$v(mAb) = v_{maxmAb} * \frac{GLN}{K_{mGLNmab} + GLN} * \cdots * \frac{GLY}{K_{mGLY} + GLY} \frac{\frac{ATP}{ADP}}{K_{m\frac{ATP}{ADP}} + \frac{ATP}{ADP}}$$

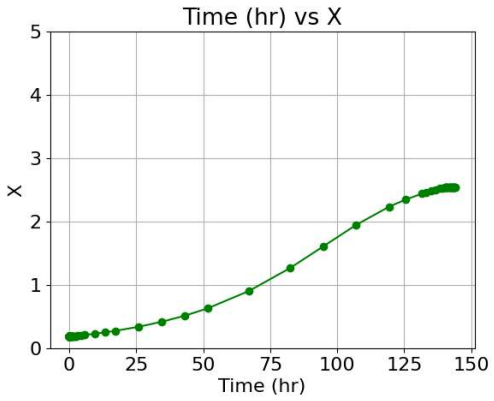
↳ 16 Michaelis-Menten Terms

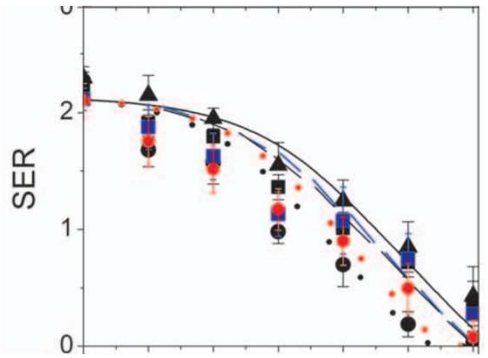


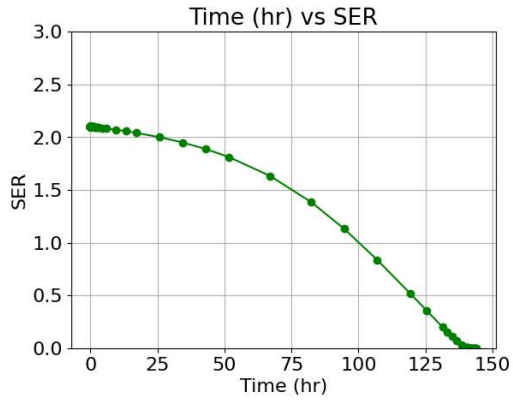
1) 2015, Robitaille, et al.; PLOS One
2) 2014, Ghorbaniaghdam, et al.; PLOS One

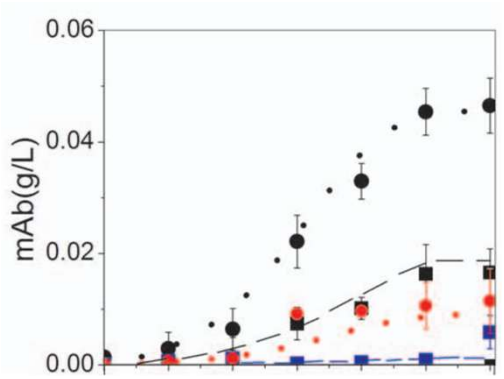
Ground Truth Model matches literature data

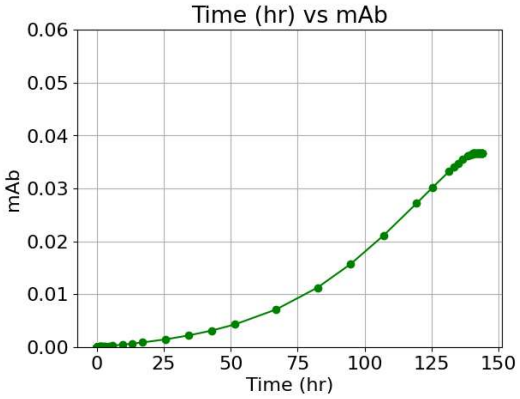


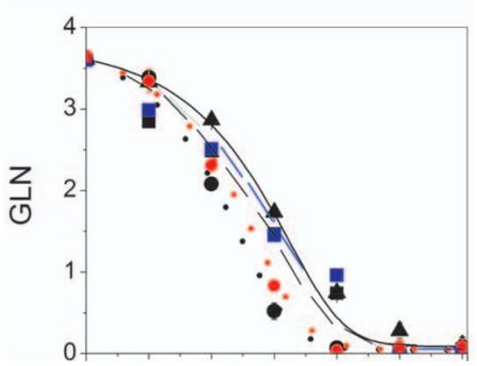


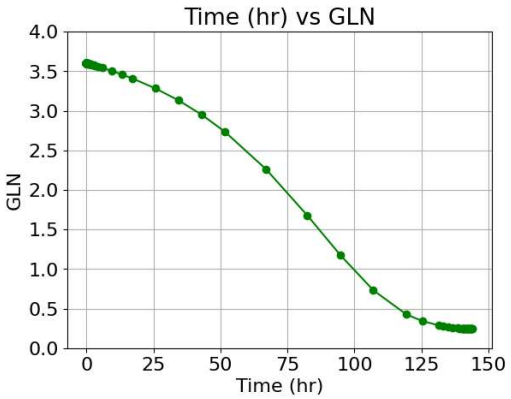












sanofi

Ground Truth Model (Modifications)

Modifications required to better reflect commercial cell culture

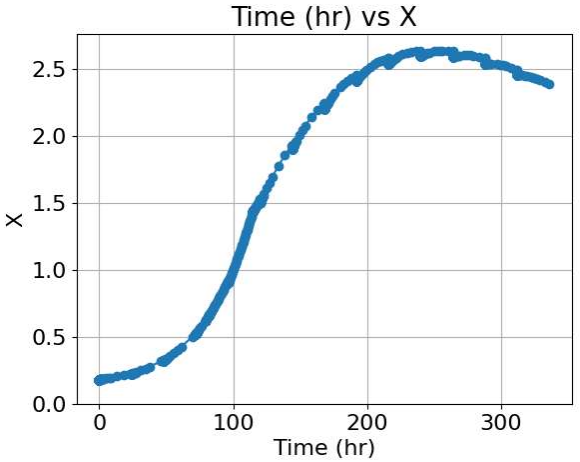
- Changing from a batch model to a bolus fed-batch model
- Cell death based on ammonia and lactate concentrations.

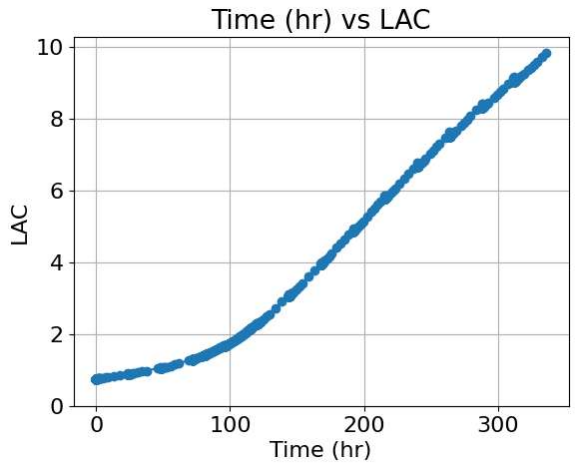
$$\mu_d = k_d \cdot \frac{[\mathrm{LAC}]}{\mathrm{KD}_{\mathrm{lac}} + [\mathrm{LAC}]} \cdot \frac{[\mathrm{AMM}]}{\mathrm{KD}_{\mathrm{amm}} + [\mathrm{AMM}]}$$

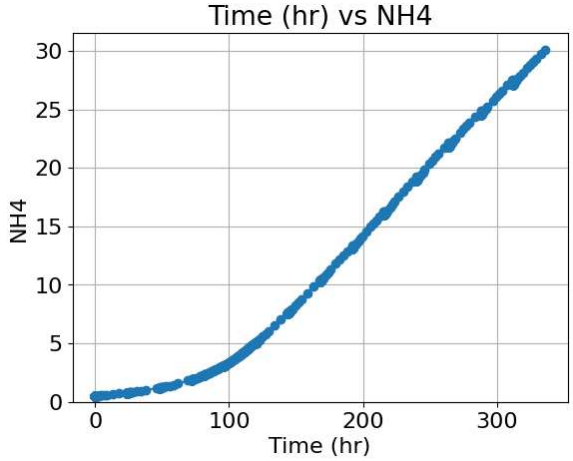
where

$$\begin{cases} \text{if } [\mathrm{AMM}] \leq 4.6\ \mathrm{mM} \text{ then } \mathrm{KI}_{\mathrm{amm}} \gg [\mathrm{AMM}], \\ \quad \mathrm{KD}_{\mathrm{amm}} \gg [\mathrm{AMM}] \\ \text{if } [\mathrm{LAC}] \leq 52\ \mathrm{mM} \text{ then } \mathrm{KI}_{\mathrm{lac}} \gg [\mathrm{LAC}],\ \mathrm{KI}_{\mathrm{lac}} \gg [\mathrm{LAC}] \end{cases}$$

3) 2009, Xing, et al.; Biotechnology Progress



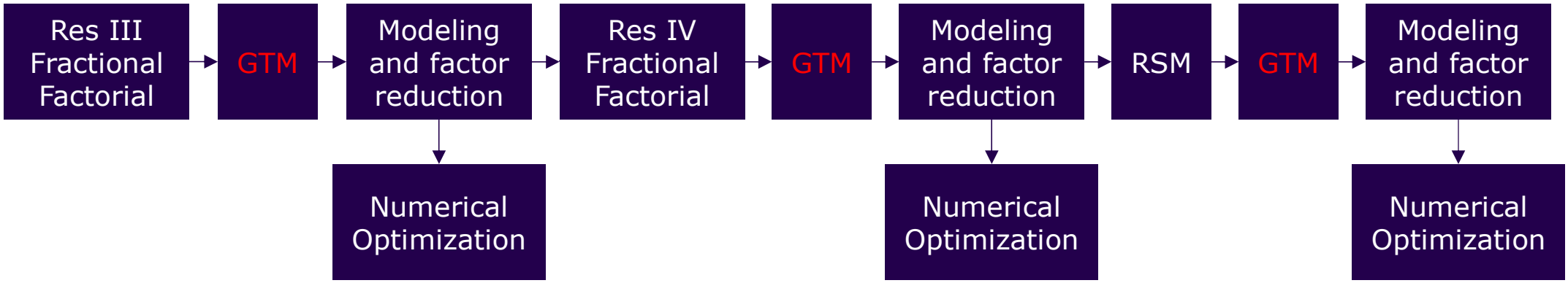




Actual Empirical Strategy

Does the DataHow hybrid MBDoE outperform classical DoE in a media optimization setting?

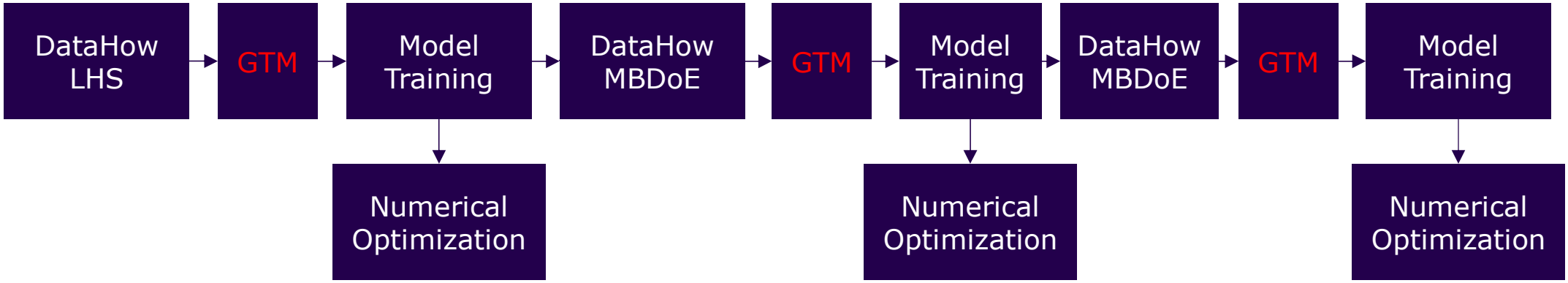
Control:



Actual Empirical Strategy

Does the DataHow hybrid MBDoE outperform classical DoE in a media optimization setting?

Treatment:



DoE Evaluations

DoE Parameter	Low	High
Feeding Start Time (hrs)	24	312
Basal Glutamine (mM)	0.01	7.2
Feed Glutamine (mM)	0.05	36
Glucose Setpoint (mM)	0.01	56
Basal Alanine (mM)	0.01	1.5
Feed Alanine (mM)	0.05	7.5

Choice of Parameter Ranges

- Glucose, glutamine, and alanine had interesting, nonlinear behavior
- Studied ranges of 0 - 2x of Gorbaniaghdam conditions
- Feed was 5x of the media conditions

Evaluation Criteria

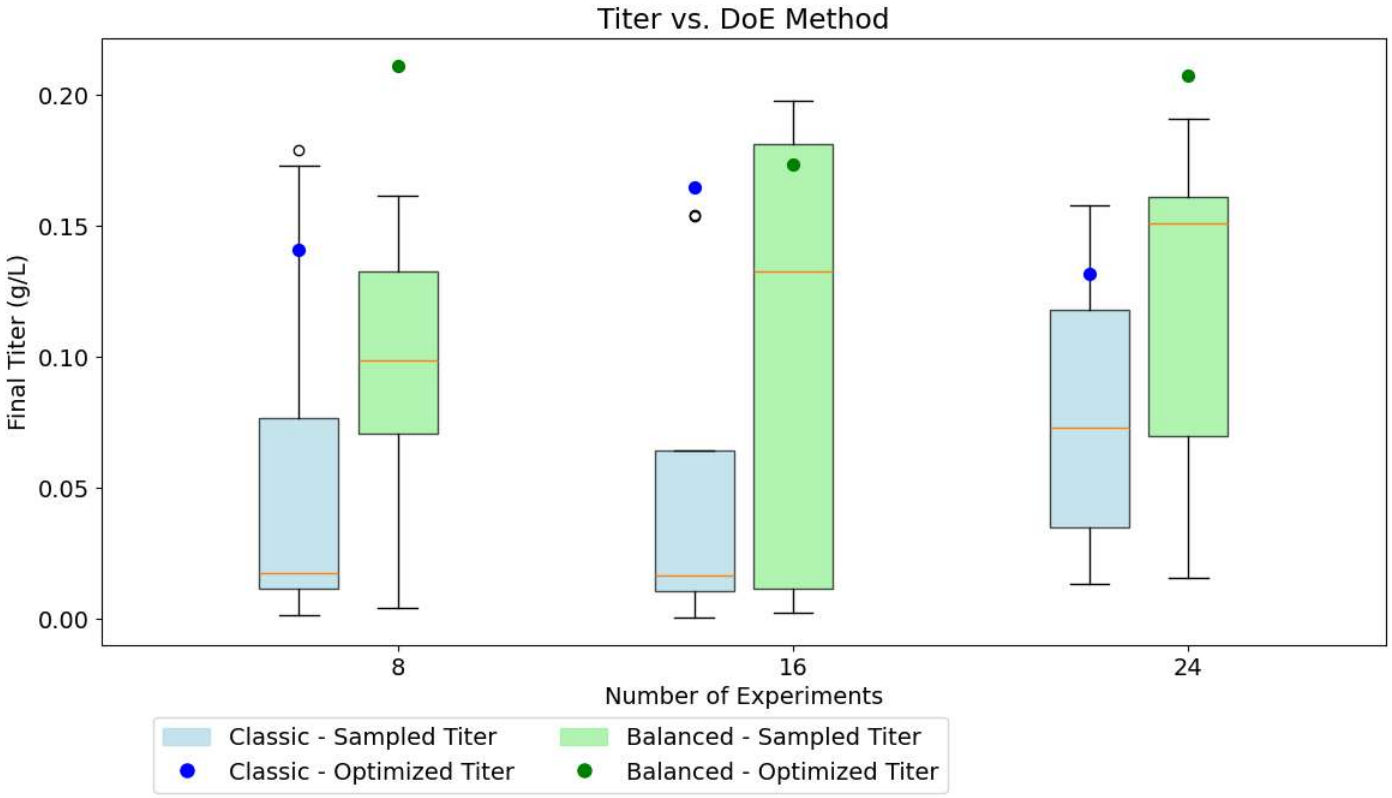
- (Highest) sampled titer:
  - GTM titer across the batch of 8 experiments
- Model-optimized titer
  - Classical/hybrid model trained and optimized to suggest single optimal run
  - GTM titer for that single optimized run

Results

Overall Results

- Hybrid-balanced approach has better final titers at each stage of design

Titer (g/L)	Highest Sampled	Model Optimum
Classic	0.179	0.165
Hybrid (Balanced)	0.198 (+11%)	0.211 (+28%)



Classical DoE

Hybrid MBDoE

Empty circles are box plot outliers

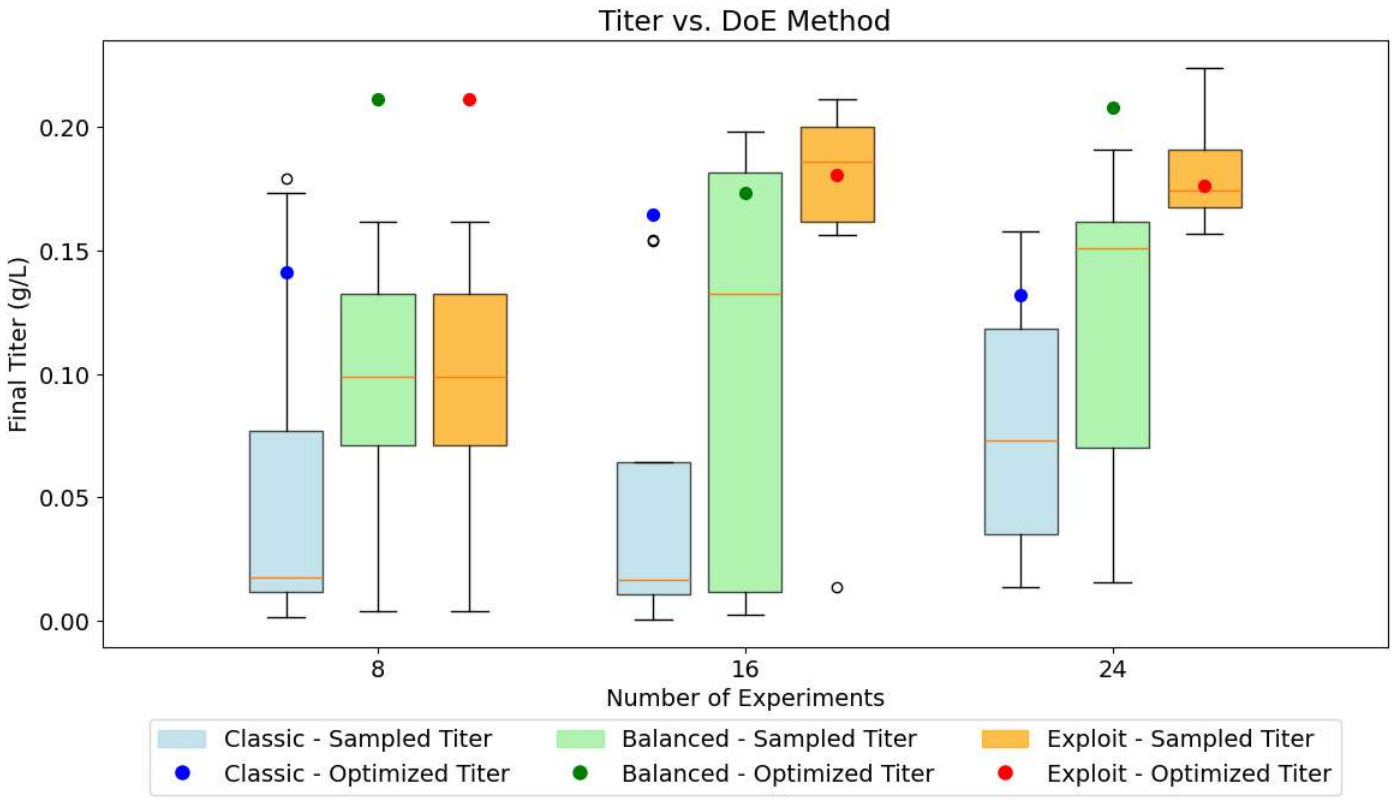
Results

Empty circles are box plot outliers

Overall Results

- Hybrid-exploit approach improves sampled titer at each stage of design

Titer (g/L)	Highest Sampled	Model Optimum
Classic	0.179	0.165
Hybrid (Balanced)	0.198 (+11%)	0.211 (+28%)
Hybrid (Exploit)	0.224 (+25%)	0.211 (+28%)



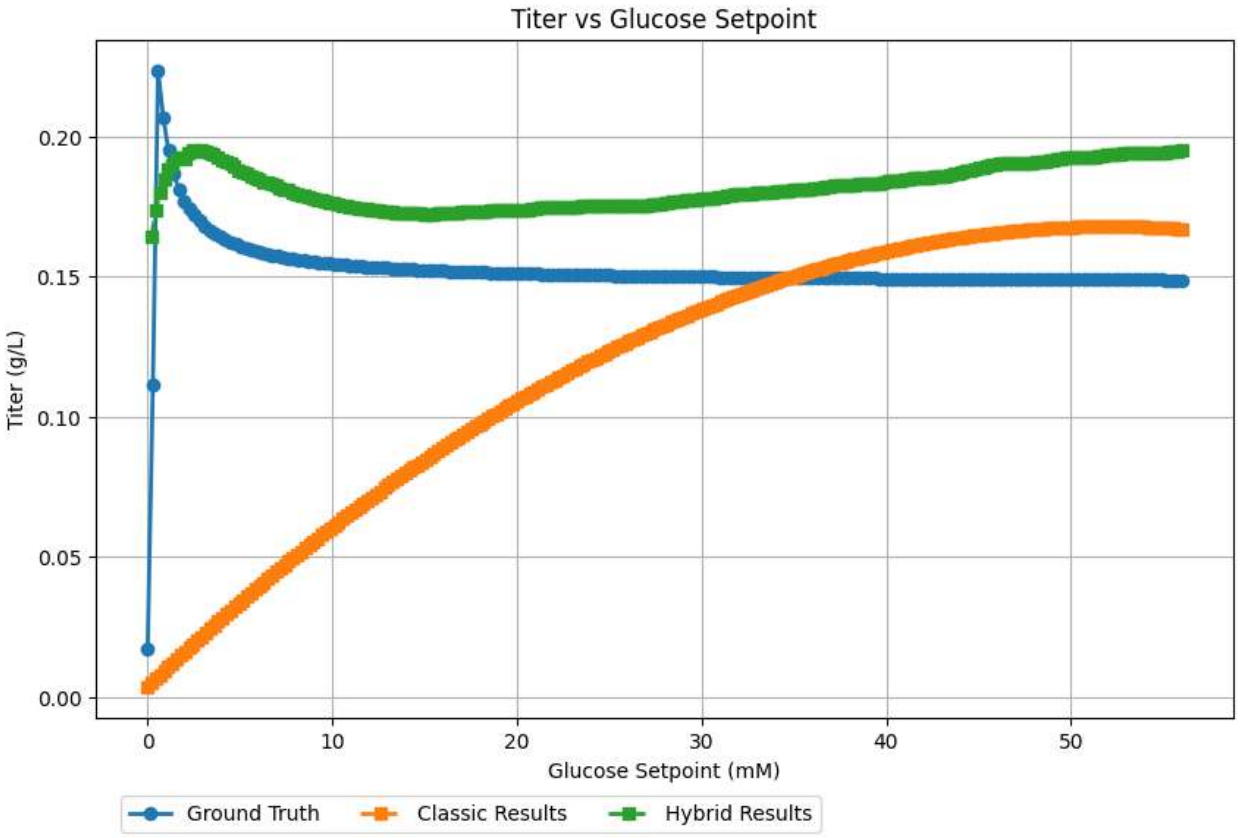
Classical DoE

Hybrid MBDoE

Results

Model System Mismatch Analysis

- Glucose setpoint has a very sharp peak near the optimum
- Classical models cannot capture that behavior using quadratic RSM models
- Hybrid methodology has the necessary complexity to pick up on the peak like behavior, resulting in better experiments.



Conclusions

- Role of modeling in CMC development for cell culture
  - Get to better processes faster
- Design of experiments using hybrid models is a rational approach
  - Cell-specific rates, mass balance, uncertainty description
- Developed a ground truth mechanistic model to evaluate DoE methodologies
  - Does not need to replicate in-house processes, just needs to capture smoothness and nonlinearity
- Hybrid model-based methodology outperforms classical DoE under the benchmark test
- Next steps
  - Experimental implementation on assets
  - Impact of non-ideal behavior (variability, contamination)

Acknowledgements

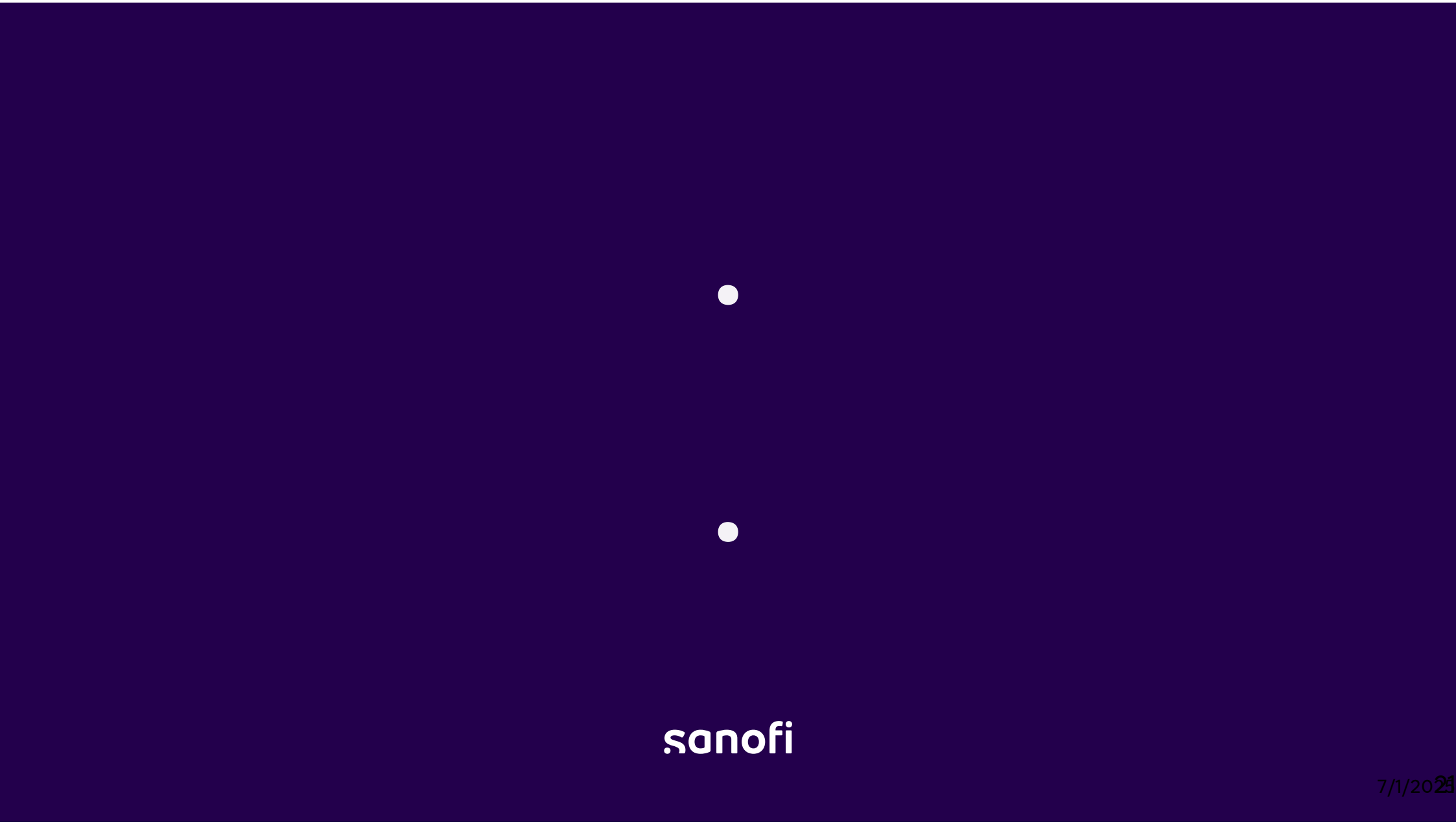
Michael Fouts - Data Science co-op student, West Virginia University

Sanofi

Antonio Barberio - Commercial Cell Culture Development
Ethan Penner - Commercial Cell Culture Development
Maurice Finger - Microbial Upstream Process Development
Denizhan Yilmaz - Microbial Upstream Process Development
Nian Liu – Data Science
Gabriele Bano – Data Science

DataHow

Tiago Dias
Tommaso Sardelli
Miguel Ressurreição
Michael Sokolov


sanofi