Development of autonomous platforms for kinetic model identification

Federico Galvanin
Associate Professor of Chemical Engineering
Department of Chemical Engineering
University College London
http://www.homepages.ucl.ac.uk/~ucecfga/

Sargent Centre for Process Systems Engineering

CatalysisHub

Continuos Flow Reactor Technology for Industrial Applications (CFRT 14)
27-28 September 2023, Portmarnock, Dublin, Ireland

Outline

- **Introduction**
  - Problem definition
  - Automated Vs Autonomous reaction systems
  - Model identification procedures: key steps and bottlenecks
  - Autonomous reaction systems: synergies and collaborations
- **Autonomous Platforms for Kinetic Model Identification**
  - Key elements and features
  - Integration of standard and advanced Model-based Design of Experiments (MBDoE) techniques
  - Case Study 1: Benzoic Acid Esterification in a Capillary Microreactor System
  - Case Study 2: Catalytic Methane Oxidation in a Microreactor Platform
  - The problem of practical distinguishability under uncertainty
- **ML-Assisted Model Identification**
  - Kinetic Model Structure Identification using ANNs-based Optimal Experimental Design
  - ML-Assisted Reliability Mapping
- **Final remarks**

Introduction

Problem definition: optimisation of flow reaction systems[1,2]

Operating Conditions → **u** → **Reaction System**

Key Performance Indicators (KPIs):
- **Conversion, Selectivity, Yield**
- **Concentration of impurities/products …**

Reaction System → Black box model optimisation → **Update** → Operating Conditions

$$\mathbf{KPIs} = f(\mathbf{u})$$

Challenge: identifying **a reliable functional form *f* (model)** relating operating conditions (**u**) to **KPIs**

f can be represented by

1. a **data-driven** or **surrogate model** (ML)
2. a **physics-based**, **mechanistic model**
3. a **hybrid model**, combining **1)** and **2)**

Interested in solving the **inverse problem**

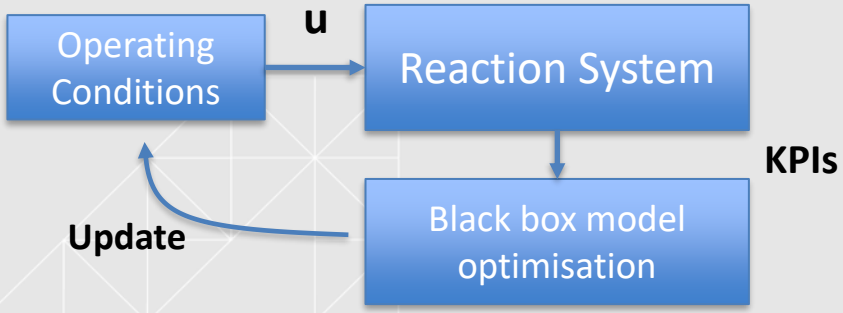
$$\mathbf{u} = f^{-1}(\mathbf{KPIs})$$

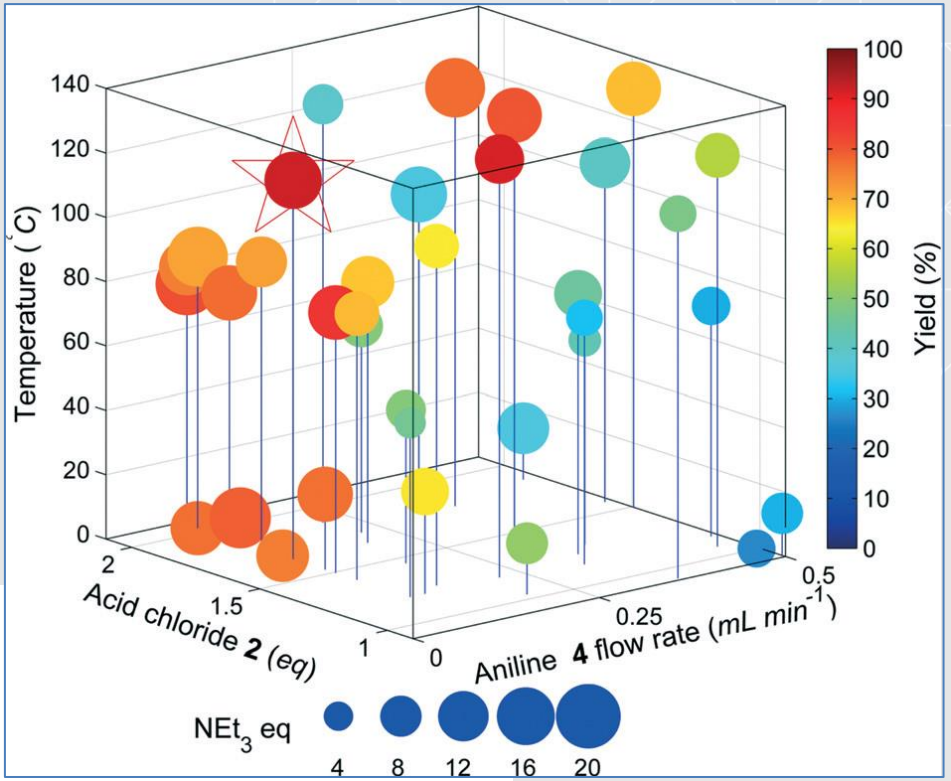
[1]Reizman, B. J., Jensen, K. F. (2012). An Automated Continuous-Flow Platform for the Estimation of Multistep Reaction Kinetics. Org. Process Res. Dev. 2012, 16, 11, 1770.

[2]Bourne, R. A. et al. (2016). Self-optimisation of the final stage in the synthesis of EGFR kinase inhibitor AZD9291 using an automated flow reactor. React. Chem. Eng., 1, 366.

Introduction

Problem definition: data-driven optimisation in automated flow reaction systems[1,2]





PROS

- Agile exploration of reaction conditions
- Fast evaluation/optimisation of KPIs

CONS

- Knowledge on reaction kinetics not guaranteed
- **Extrapolation to unexplored regions particularly challenging (critical for scale-up!)**
- May require many runs before reaching a convergence
- Extremely **sensitive to data quality**

[1]Reizman, B. J., Jensen, K. F. (2012). An Automated Continuous-Flow Platform for the Estimation of Multistep Reaction Kinetics. Org. Process Res. Dev. 2012, 16, 11, 1770.

[2]Bourne, R. A. et al. (2016). Self-optimisation of the final stage in the synthesis of EGFR kinase inhibitor AZD9291 using an automated flow reactor. React. Chem. Eng., 1, 366.

Introduction

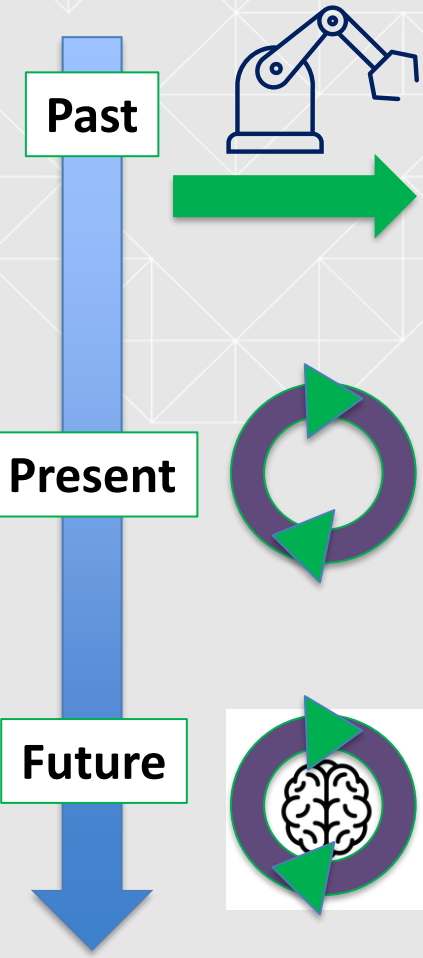
Reaction platforms: where are we heading to?

Reaction platforms[1] can be:

- **Automated**: systems allowing enhanced control and data acquisition, agile execution of scheduled **open-loop** experiments, **requiring minimum user intervention**

- **Autonomous**: automated reaction system enabling **closed-loop** operation, self-optimisation and experimental design with **no user intervention during operation**. High-level decisions still required *(Example: choice of design criteria/KPIs targets, stopping rules, etc.)*

- **Intelligent**: autonomous reaction system capable of **adapting through learning**, even in presence of disturbances from the environment. All decisions (including high-level ones) are directly taken by the platform (**no inputs from the user are required)**



[1]Barz, T., Kager, J., Herwig, C., Neubauer, P., Bournazou, M. N. C., Galvanin, F. (2022). Characterization of reactions and growth in automated continuous flow and bioreactor platforms—From linear DoE to model-based approaches, *Simulation and Optimization in Process Engineering*, 273-319.

Introduction

Autonomous Reaction Systems – Team Synergies (1)

Federico Galvanin

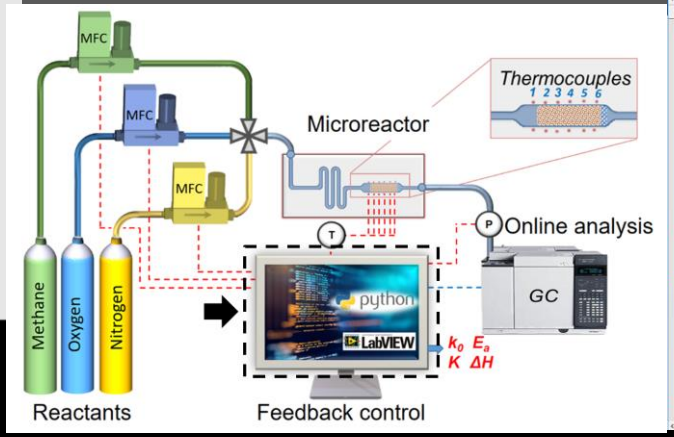
- Computational Algorithms for Model Identification
- Optimal Experimental Design
- Autonomous Reaction Systems
- Machine Learning Applications to Systems Modelling

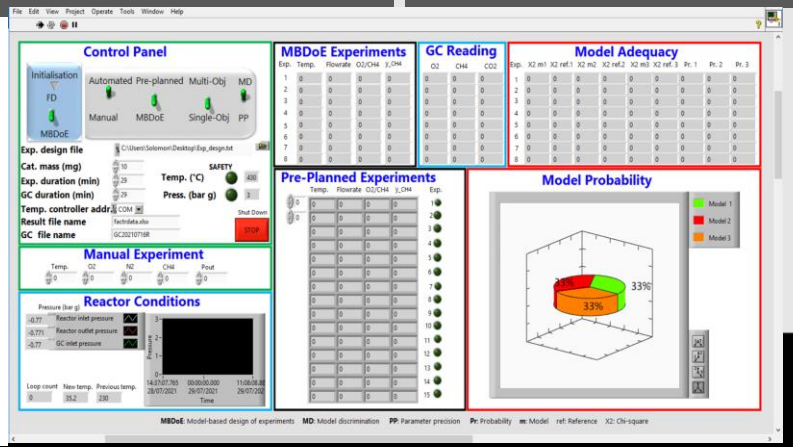
www.homepages.ucl.ac.uk/~ucecfga/

Asterios Gavriilidis

- Reaction Engineering
- Flow chemistry
- Design of automated reactor platforms
- Integration with online analysis tools

www.ucl.ac.uk/chemical-engineering/research/gavriilidis-lab





Development of autonomous platforms for kinetic model identification

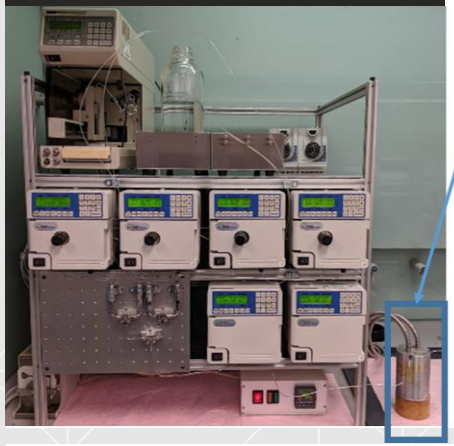
UCL ENGINEERING
Change the world

Introduction

Cloud-based Experimental Design and Analysis Service (EDAS)

UNIVERSITY OF LEEDS

AstraZeneca



Lab-scale flow reactor

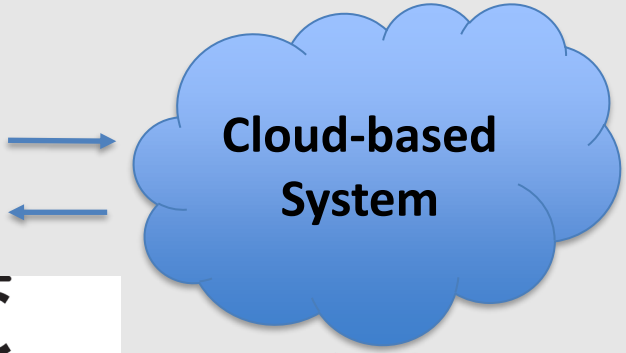
- Isothermal flow
- Plug flow
- Fast heating-cooling

EPSRC

Engineering and Physical Sciences Research Council

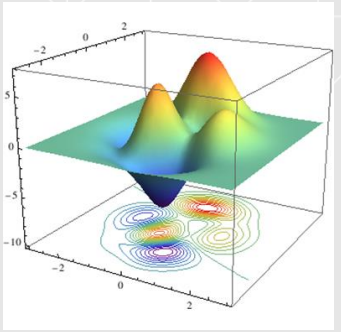
EP/R032807/1, Cognitive Chemical Manufacturing

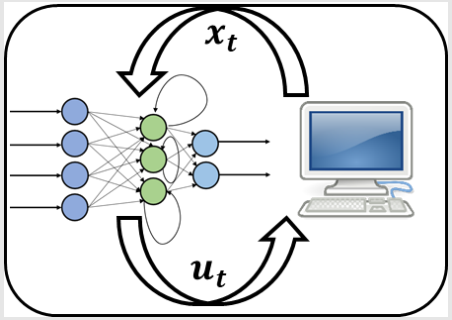
Target: Pharma Applications

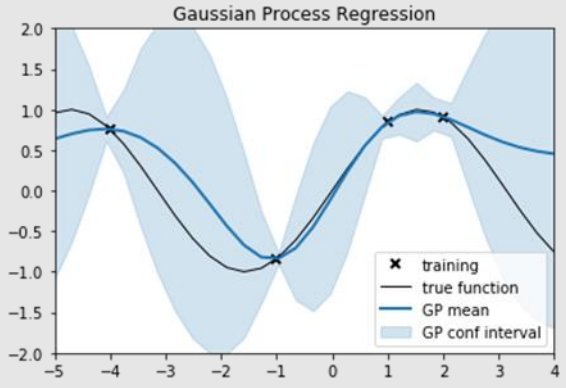


IBM

UCL







Research Fellows

Emmanuel Agunloye

Panagiotis Petsagkourakis

Richard Bourne (PI)

Identification of kinetic models

- Mechanistic modelling
- Hybrid modelling[1]

[1]Petsagkourakis, P., Galvanin, F. (2021). Safe model-based design of experiments using Gaussian processes, *Computers and Chemical Engineering,* 107339.

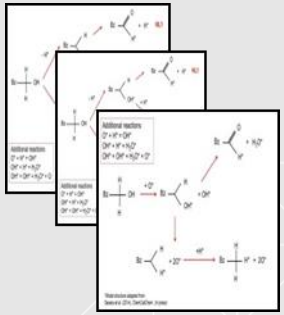
UCL ENGINEERING Change the world

Introduction

Modelling complex reaction systems: problem definition

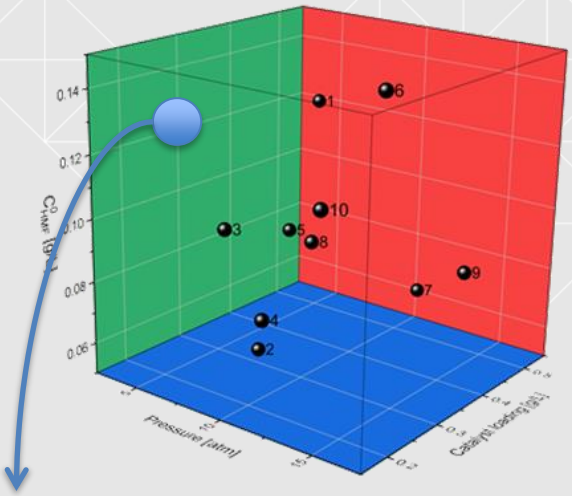
Chemistry-Surface Science

Potential kinetic mechanisms

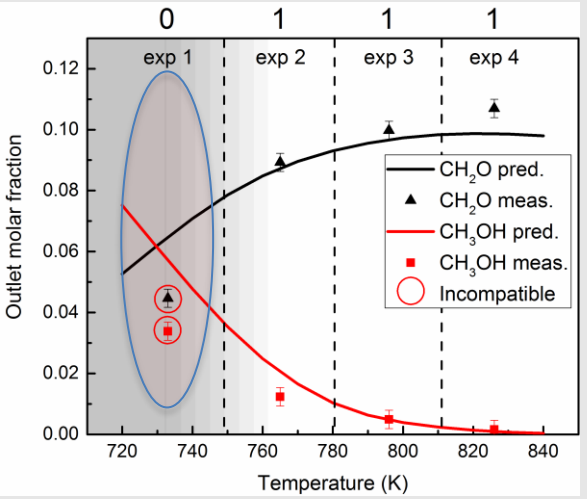


Formulation of alternative kinetic models

Kinetic model identification



Reactor Model → **<u>Reliable</u> phenomenological process model**

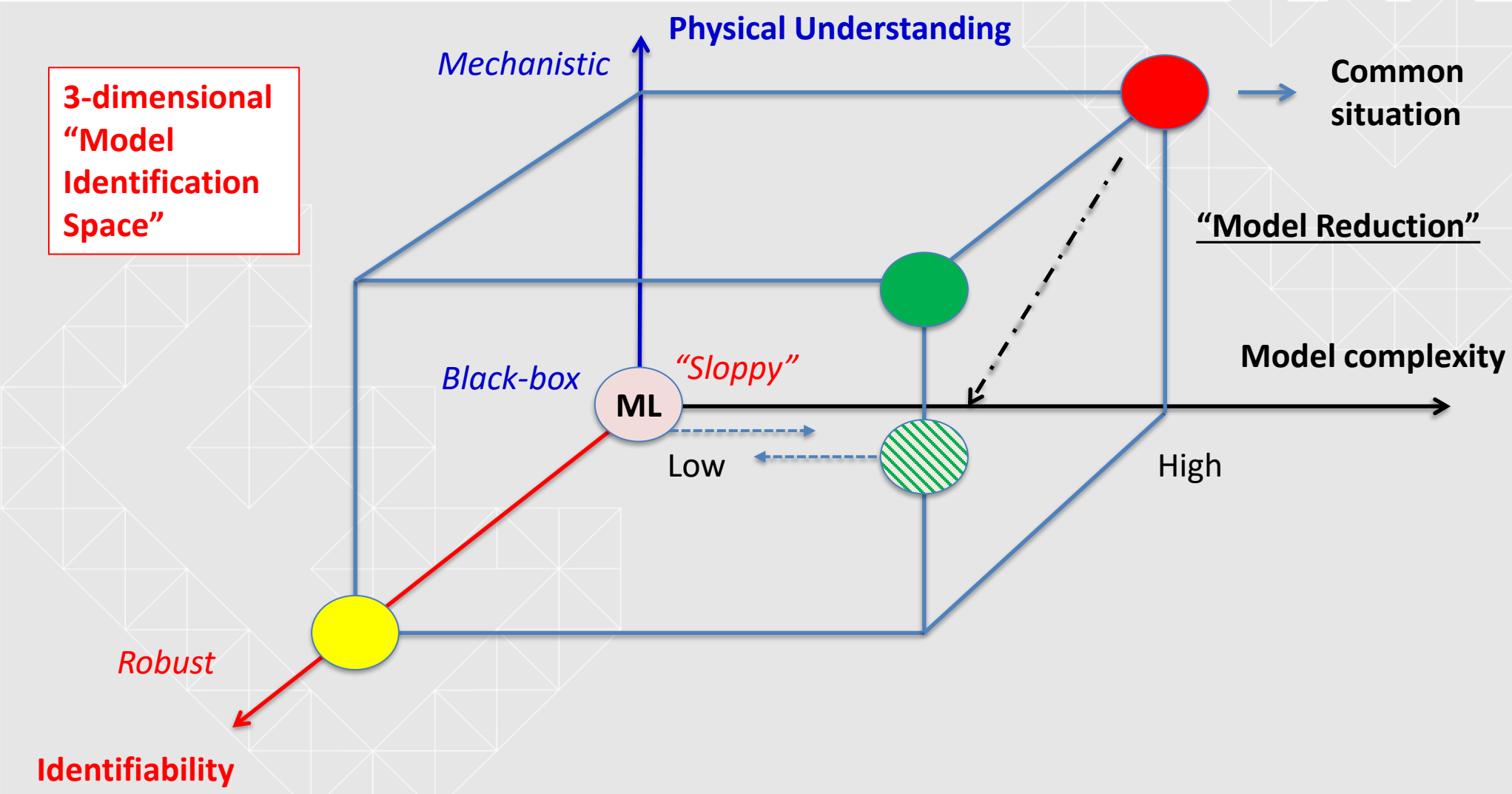


Model requirements

- **Adequate to represent the physical system**
- **Minimum variance of prediction(s) in the range of expected utilisation**
- **Robustness in <u>unexplored</u> regions of the design space**

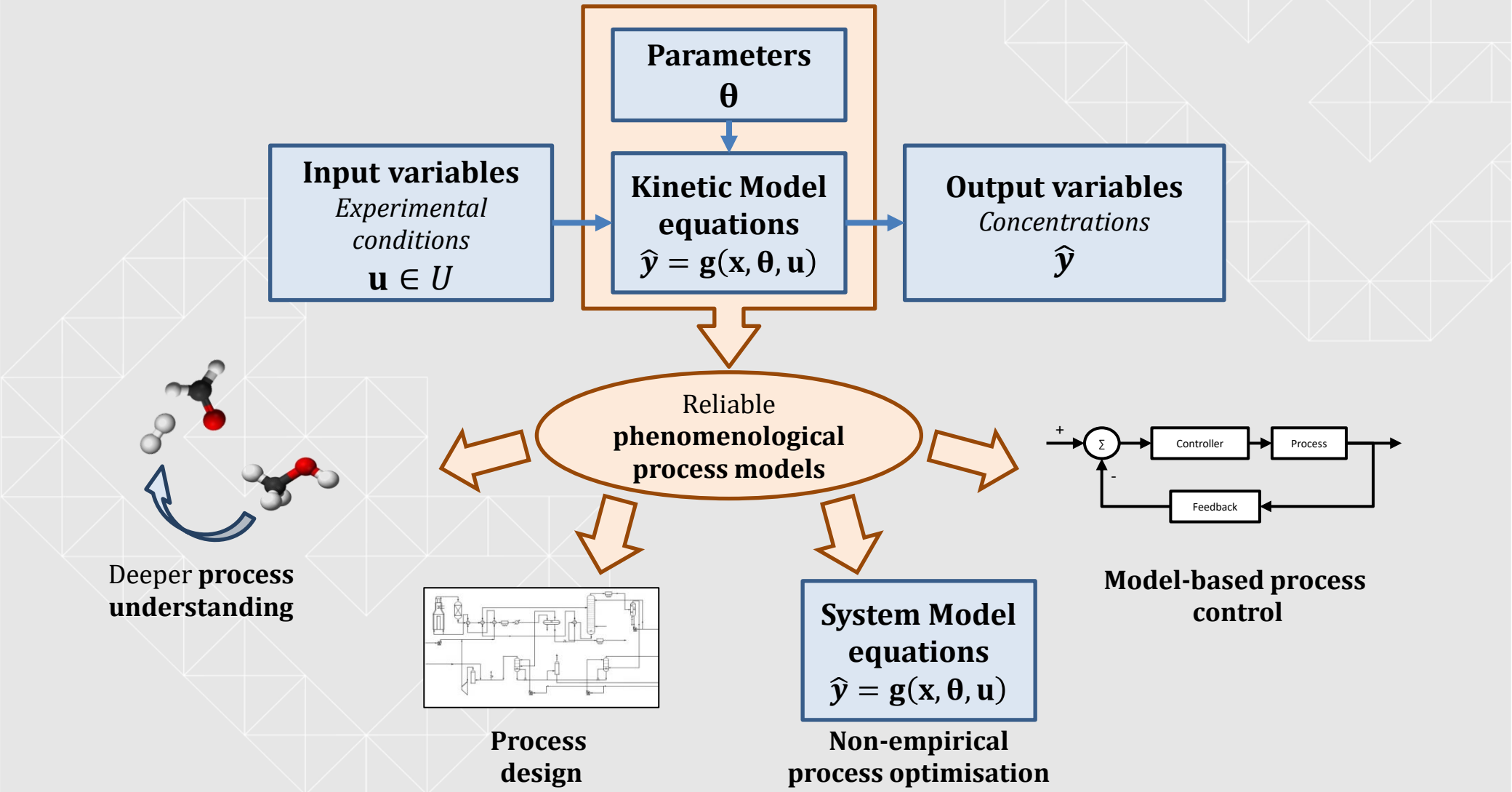
Modelling complex reaction systems

The three dimensions of model identification



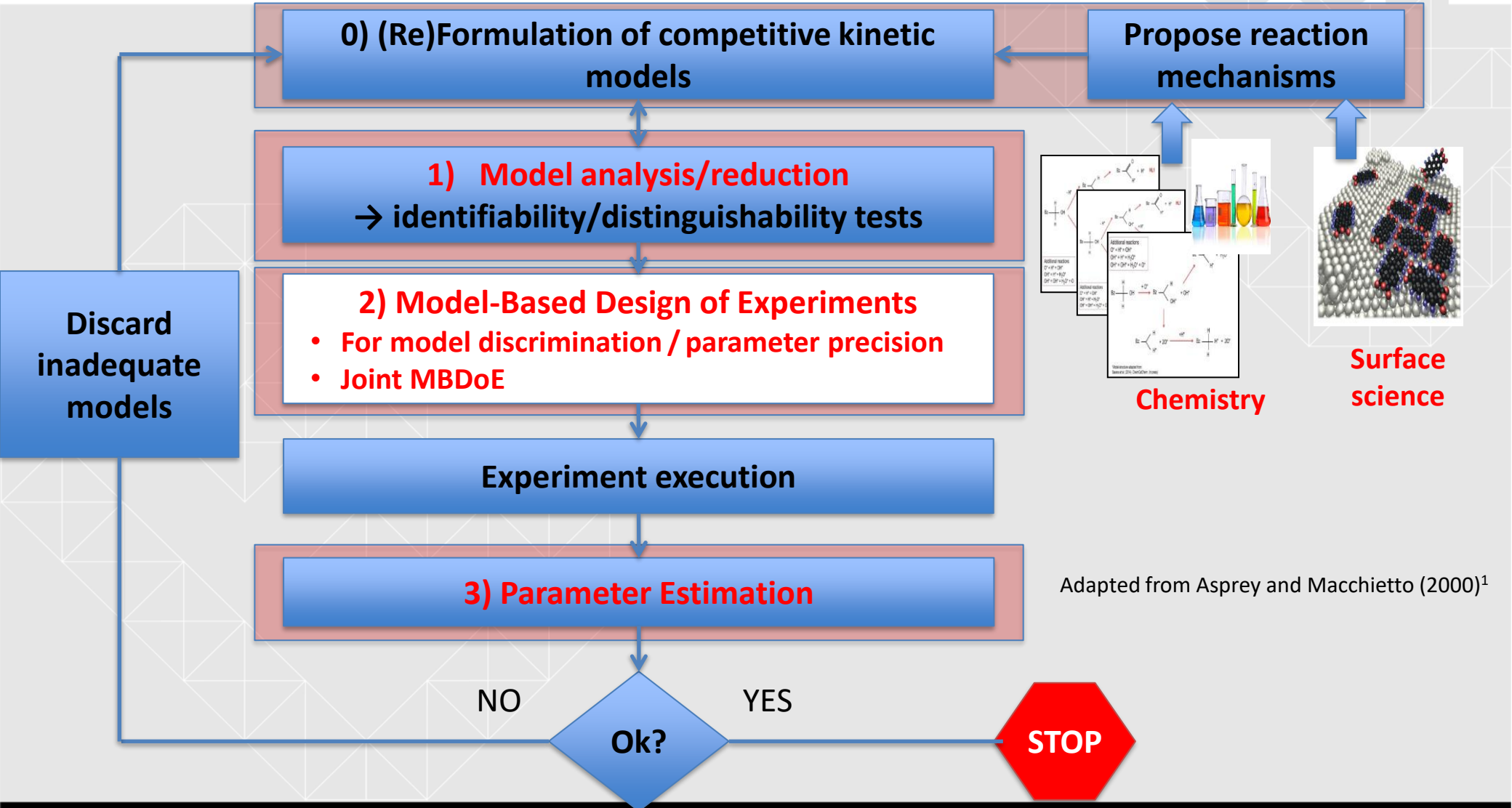
Introduction

The mathematical problem: identification of parametric models



Introduction

Kinetic model identification procedures: key steps and bottlenecks



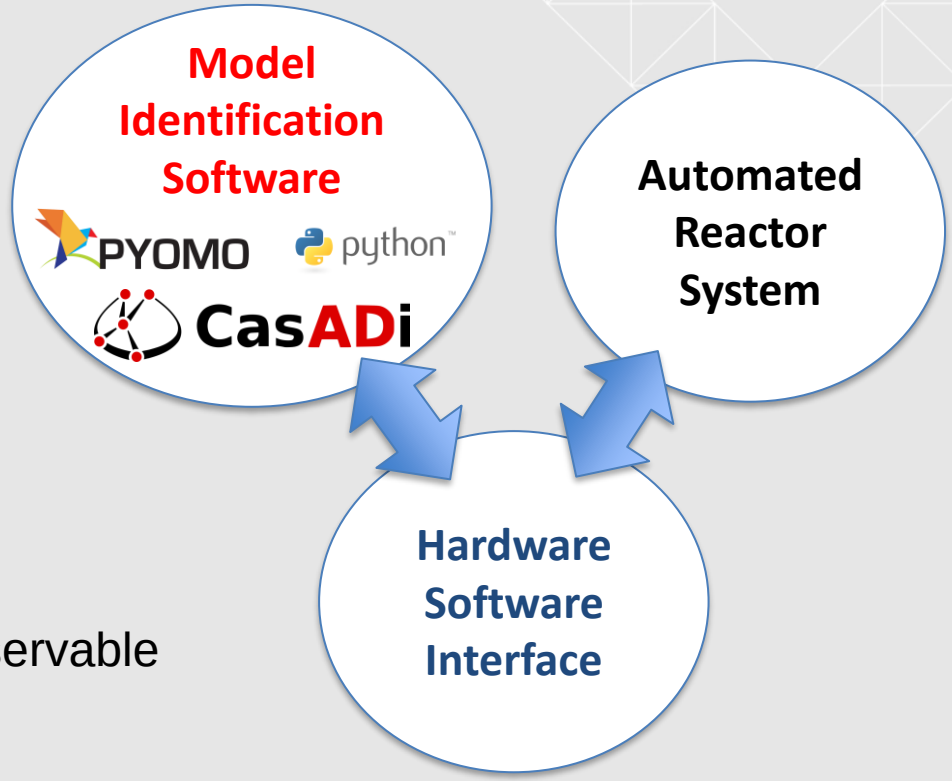
Adapted from Asprey and Macchietto (2000)[1]

[1] Asprey, S. P., Macchietto, S. (2000), Statistical tools for optimal dynamic model building, Computers & Chemical Engineering, 24, 1261.

Autonomous platforms for kinetic model identification

Key elements of Autonomous Reaction Platforms

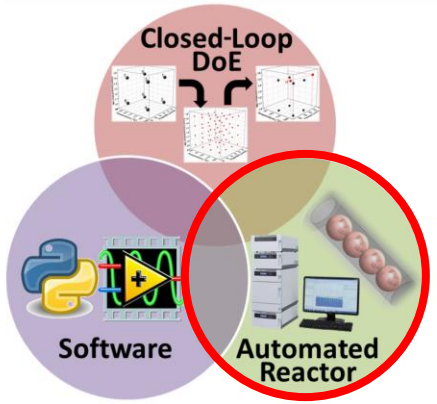
- **Model identification software**
  - DoE/Factorial design/User defined experiments -> Screening module
  - Model-based design of experiments for model discrimination (MBDoE-MD)[1]
  - Model-based design of experiments for parameter precision (MBDoE-PE)[2]
  - Joint design (j-MBDoE)[3]
  - Online model reparametrisation[4]
  - Robust parameter estimation
- **Hardware/Software interface**
  - Establishing communication protocols
  - Python-LabView-Devices integration
- **Automated reactor system**
  - Reactor design
  - Safe/reliable operation
  - Experiments must be reproducible and observable
  - Measurement/control system



[1]Waldron, C; et al. (2019), React. Chem. Eng., 4, 1623.
[2]Waldron, C et al. (2019). Ind. Eng. Chemi Res. , 58 (49) 22165.
[3]Galvanin, F. et al. (2016), Computers & Chemical Engineering , 61, 5791.
[4]Quaglio, M et al. (2019) ,Computers & Chemical Engineering , 124, 270.

Autonomous platforms for kinetic model identification

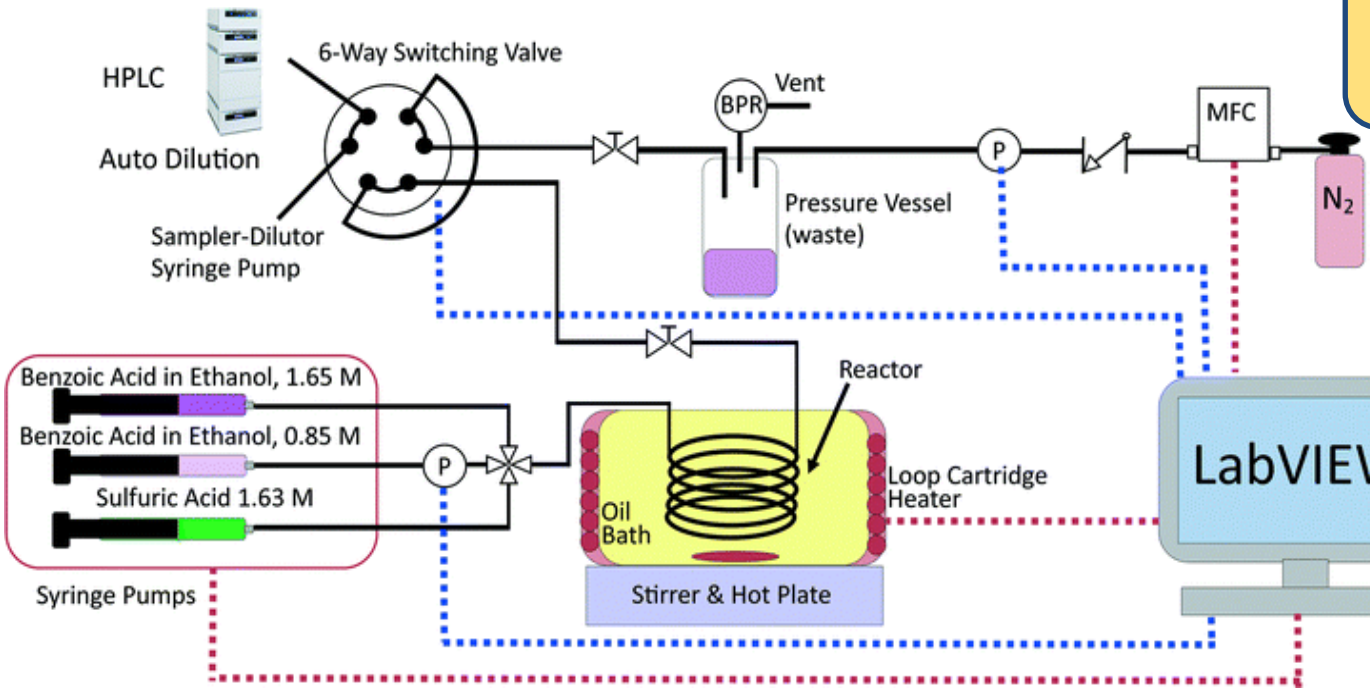
Framework example[1,2]



Automated reactor → **Autonomous (self-optimising) reactor**

Goal

Identification of reliable kinetic models with minimum time, materials and effort



Model Identification Algorithms

python

[1]Waldron C., Pankajakshan A., Quaglio M., Cao E., Galvanin F., Gavriilidis A. (2019), *An autonomous microreactor platform for the rapid identification of kinetic models*, Reaction Chemistry & Engineering, 4, 1623-1636.

[2]Quaglio M., Waldron C., Pankajakshan A., Cao E., Gavriilidis A., Fraga E. S., Galvanin F. (2019), *An online reparametrisation approach for robust parameter estimation in automated model identification platforms*, Computers & Chemical Engineering, 124, 270-284.

UCL ENGINEERING Change the world

Standard MBDoE techniques

MBDoE-MD: Optimal Design for Model Discrimination

Example of MBDoE-MD criterion[1]

"Discriminating power"

$$\underset{\boldsymbol{\varphi}}{\text{maximize}}\ T_{ij}(\boldsymbol{\varphi}) = \left[\hat{\mathbf{y}}\left(\boldsymbol{\varphi},\hat{\boldsymbol{\theta}}_i\right) - \hat{\mathbf{y}}\left(\boldsymbol{\varphi},\hat{\boldsymbol{\theta}}_j\right)\right]^{\mathrm{T}} \mathbf{V}_{ij}^{-1}(\boldsymbol{\varphi})\left[\hat{\mathbf{y}}\left(\boldsymbol{\varphi},\hat{\boldsymbol{\theta}}_i\right) - \hat{\mathbf{y}}\left(\boldsymbol{\varphi},\hat{\boldsymbol{\theta}}_j\right)\right]$$

Operating conditions optimised online **to maximise the divergence between model predictions**

Different criteria for model discrimination are integrated in the software

Design vector

$$\boldsymbol{\varphi} = [\mathbf{y}_0, \mathbf{u}, \mathbf{t}^{sp}, \mathbf{z}^{sp}, \tau]^{\mathrm{T}}$$

- $\mathbf{y}_0$ set of initial conditions on the measured variables (C_i)
- $\mathbf{u}$ set of manipulated inputs (T, P, F)
- $\mathbf{t}^{sp}$ set of time instants at which the measured variables are sampled
- $\mathbf{z}^{sp}$ set of time instants at which the measured variables are sampled
- τ the experiment duration (possibly)

[1]Buzzi-Ferraris G, Forzatti P. (1983), A new sequential experimental design procedure for discriminating among rival models. Chemical Engineering Science, 1, 38, 225.

Standard MBDoE techniques

MBDoE-PE: Optimal Design for Improving Parameter Precision

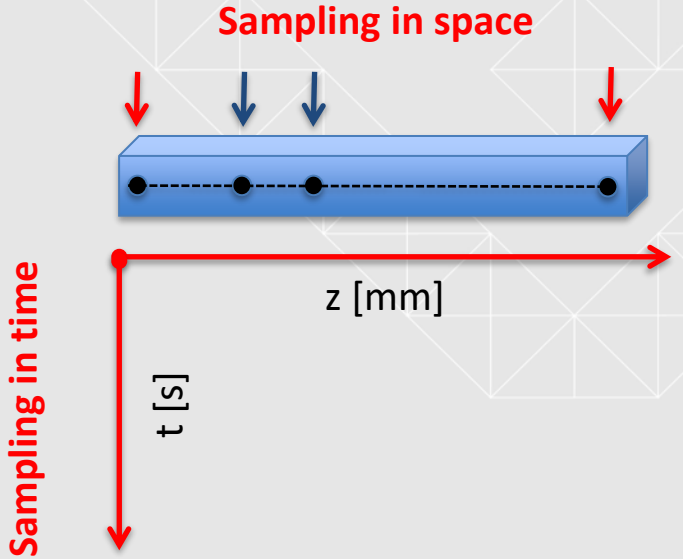
MBDoE formulation for flow reactor systems

Fisher information matrix (FIM)

INFO from **spatial domain**

$$\mathbf{H}_\theta(\boldsymbol{\theta},\boldsymbol{\varphi}) = \mathbf{H}_\theta^0 + \sum_{k=1}^{n_{sp}} \sum_{i=1}^{N_y} \sum_{j=1}^{N_y} s_{ij} \left[\frac{\partial \hat{y}_i(z_k,t_k)}{\partial \theta_l} \frac{\partial \hat{y}_j(z_k,t_k)^T}{\partial \theta_m} \right]_{l,m=1\ldots N_{\boldsymbol{\theta}}}$$

INFO from **time domain**



Optimal design for improving parameter estimation

$$\boldsymbol{\varphi}^{\mathrm{PE}} = \arg\min_{\boldsymbol{\varphi}\in D} \{\psi[\mathbf{V}_\theta(\boldsymbol{\theta},\boldsymbol{\varphi})]\} = \arg\min_{\boldsymbol{\varphi}\in D} \{\psi[\mathbf{H}_\theta^{-1}(\boldsymbol{\theta},\boldsymbol{\varphi})]\}$$

MBDoE-PE criterion (A, D, E, …)

Design vector

$$\boldsymbol{\varphi} = [\mathbf{y}_0, \mathbf{u}, \mathbf{t}^{sp}, \mathbf{z}^{sp}, \tau]^{\mathrm{T}}$$

- $\mathbf{y}_0$ set of initial conditions on the measured variables (C_i)
- $\mathbf{u}$ set of manipulated inputs (T, P, F)
- $\mathbf{t}^{sp}$ set of time instants at which the measured variables are sampled
- $\mathbf{z}^{sp}$ set of time instants at which the measured variables are sampled
- τ the experiment duration (possibly)

Advanced MBDoE techniques

Joint Model-based Design of Experiments (j-MBDoE)[1]

Multi-objective MBDoE formulation (MBDoE-MD/MBDoE-PE)

- Optimal design for discriminating between N_M competing kinetic models[2]
- Optimal design for improving the estimation of kinetic parameters[3]

j-MBDoE

$$\varphi^{MD} = \arg\max_{\varphi \in D}\left\{\psi^{MD}\right\} = \arg\max_{\varphi \in D}\left\{\sum_{M,N=1}^{N_M} P_M P_N \left[\sum_{i=1}^{N_y} \frac{\left(\hat{y}_{M,i} - \hat{y}_{N,i}\right)^2}{\sigma_{y,i}^2}\right]_{M,N}\right\}$$

Design of experimental conditions providing the **greatest difference between model predictions**

$$\psi^{PE} = \sum_{j=1}^{N_M} \left\|\mathbf{H}_j\right\| / N_M \leq \varepsilon$$

… ensuring at the same time the best possible **reduction of parametric uncertainty**

st $\varepsilon^{MIN} \leq \varepsilon \leq \varepsilon^{MAX}$ "ε-constraint method"

MBDoE for improving parameter estimation

MBDoE for model discrimination

P_i = probability of the i-th model to be the "true" model

$\hat{y}_{ji}$ = *i*-th predicted response of the *j*-th model

[1]Galvanin, F. et al. (2016), Comp. Chem. Eng, 61, 5791-5806
[2]Schwaab, M. et al. (2006), Chem. Eng. Sci., 61, 5791-5806
[3]Reizman, B. J., Jensen, K. F. (2012), Org. Process Des. Dev., 16, 1770-1782

Advanced MBDoE techniques

Explorative MBDoE based on G-optimality maps[1]

Challenges:

- MBDoE algorithms tend to be "**greedy**"
  → only highly informative regions of the design space are mapped in the online optimization
- Can we formulate an optimal experimental design problem allowing **agile experimental design space exploration**?

Potential solution

Explorative MBDoE based on G-optimality maps (***G-map eMBDoE***)

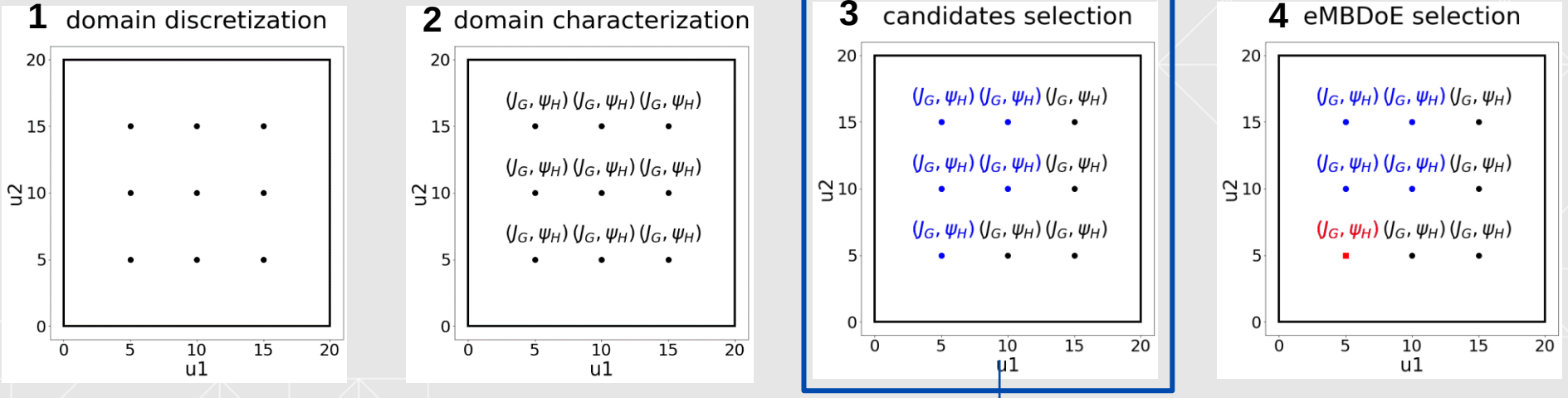
Francesca Cenci

Pierantonio Facco

[1]Cenci, F., Pankajakshan, A., Galvanin, F. (2023), An exploratory model-based design of experiments approach to aid parameters identification and reduce model prediction uncertainty, Computers & Chemical Engineering, 177, 108353, https://doi.org/10.1016/j.compchemeng.2023.108353

Advanced MBDoE techniques

Explorative MBDoE based on G-optimality maps[1]



- Explorative MBDoE based on G-optimality maps (*G-map eMBDoE*)
  - **Step 3**: eMBDoE first discriminates points based on a **G-optimality threshold** $J_G = \sum_{i=1}^{N_y} V_{y,i}$

$$J_G \geq J_{G,\text{thr}} J_{G,\max}$$

N_y number of response variables

ψ_H scalar measure of FIM

  - **Step 4**: Then, the experiment with the highest amount of information based on MBDoE-PE design criteria is selected

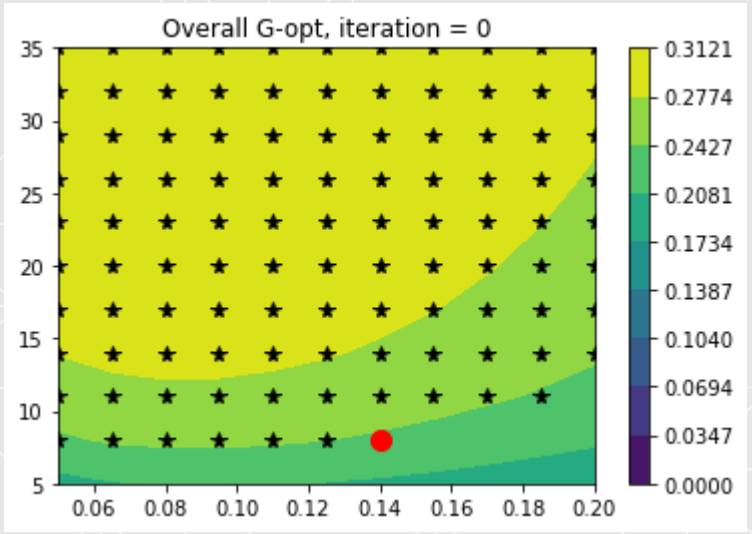
[1]Cenci, F., Bawa, S., Gavriilidis, A., Facco, P., Galvanin, F. (2023), An exploratory model-based design of experiments technique to aid parameters identification and reduce prediction uncertainty, Computer-Aided Chemical Engineering, 1-6.

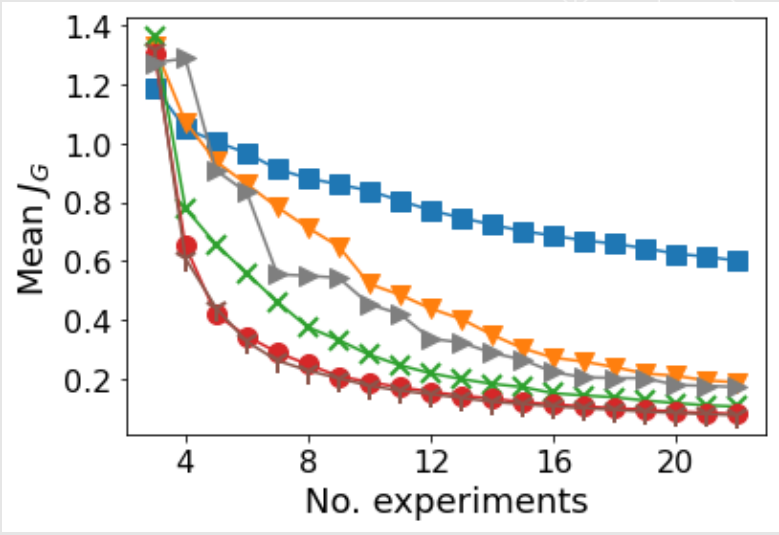
Advanced MBDoE techniques

Explorative MBDoE based on G-optimality maps[1]

G-map eMBDoE Settings

- $J_{G,thr} = 0.75$
- D-optimal experimental design



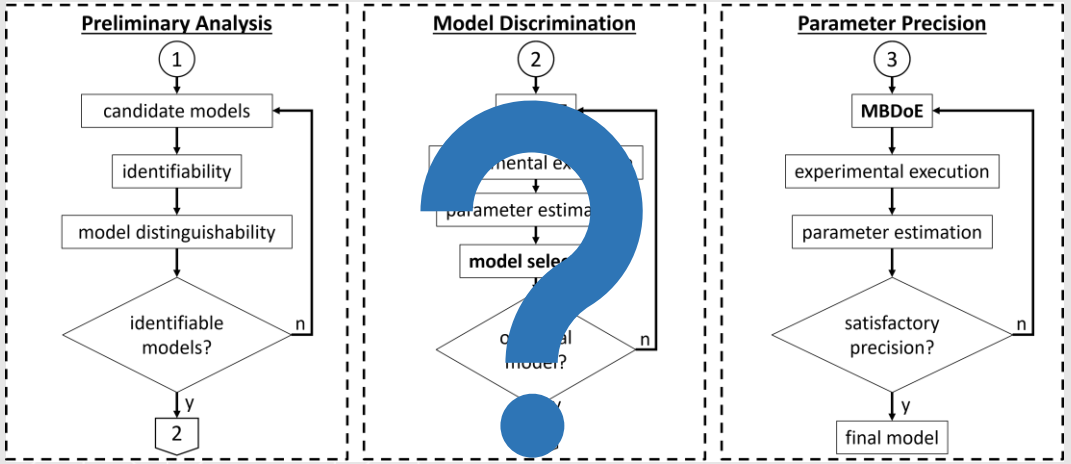


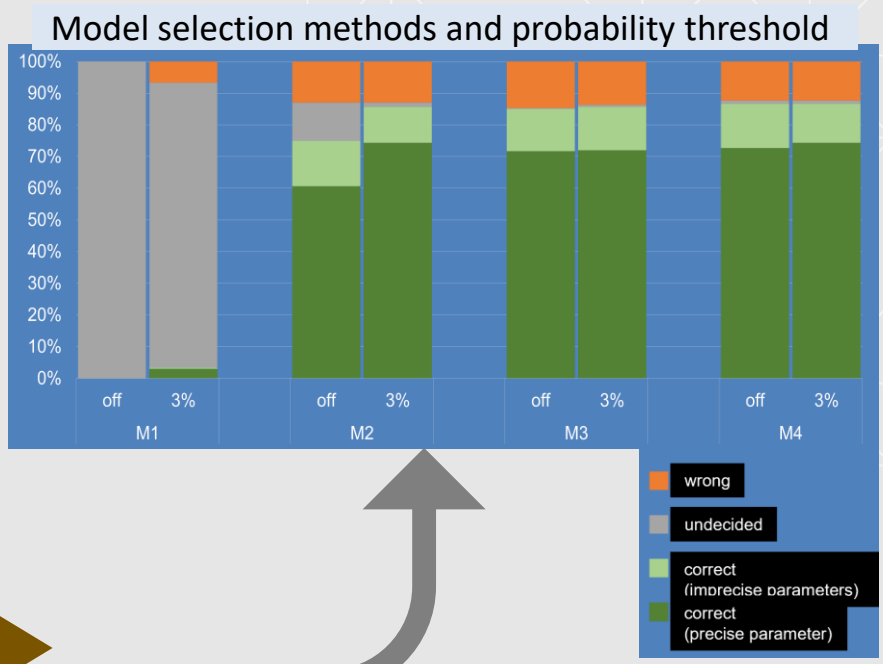
- Efficient reduction of prediction uncertainty in the design space compared to MBDOE and DoE
- Satisfactory estimation of the full set of model parameters, thanks to the MBDoE step 4
- Methods have been developed to automatically adjust the G-optimality threshold

[1]Cenci, F., Bawa, S., Gavriilidis, A., Facco, P., Galvanin, F. (2023), An exploratory model-based design of experiments technique to aid parameters identification and reduce prediction uncertainty, Computer-Aided Chemical Engineering, 1-6.

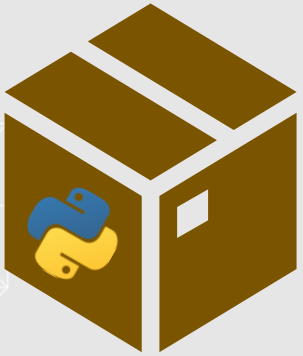
Advanced MBDoE techniques

Development of a holistic Python package for optimal selection of experimental design criteria[1]





→ **MBDoE-MD**: first step towards optimal selection of MBDoE settings



Maerthe Theresa Tillmann, MSc

- Computational tool for Holistic Model Identification (**HoliMI**[2])
- Different model selection/MBDoE-MD criteria implemented

[1]Tillmann, M. T., Galvanin, F. (2023), Development of a holistic Python package for optimal selection of experimental design criteria in kinetic model discrimination, Computer-Aided Chemical Engineering, 631-636.

[2] https://github.com/mtt9/HoliMI

Case study 1

Benzoic acid esterification in a capillary microreactor

Benzoic Acid Esterification with Ethanol using Sulfuric Acid (homogeneous)

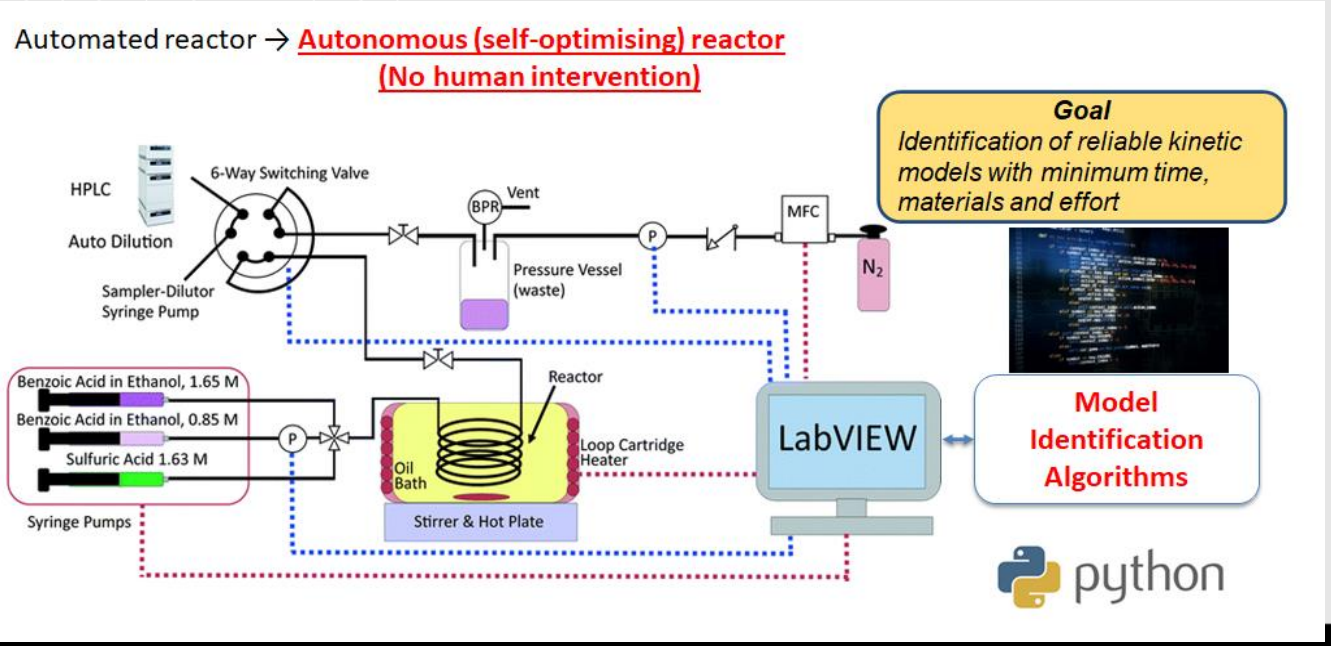
- Liquid phase reaction
- PTFE capillary reactor
- Online HPLC as measurement system

Objective

- **Identify a kinetic model from a sequence of steady state flow experiments**

Data fitting results for **two candidate kinetic models** using two identical preliminary steady-state factorial campaigns

Model	χ^2 ($\chi_{ref}^2 = 43.7$)	Result
$r = kC_{BA}$	16.2	Possible model
$r = kC_{BA}^2$	156	Reject model



Design vector in MBDoE

$\boldsymbol{\varphi} = [T, F, c_{IN,\,i}]$

- Temperature
- Flowrate
- Concentration of benzoic acid in the feed

[1]Waldron, C; Pankajakshan, A; Quaglio, M; Cao, E; Galvanin, F; Gavriilidis, A. (2019) An autonomous microreactor platform for the rapid identification of kinetic models. **Reaction Chemistry and Engineering** 4, 1623. 10.1039/C8RE00345A.

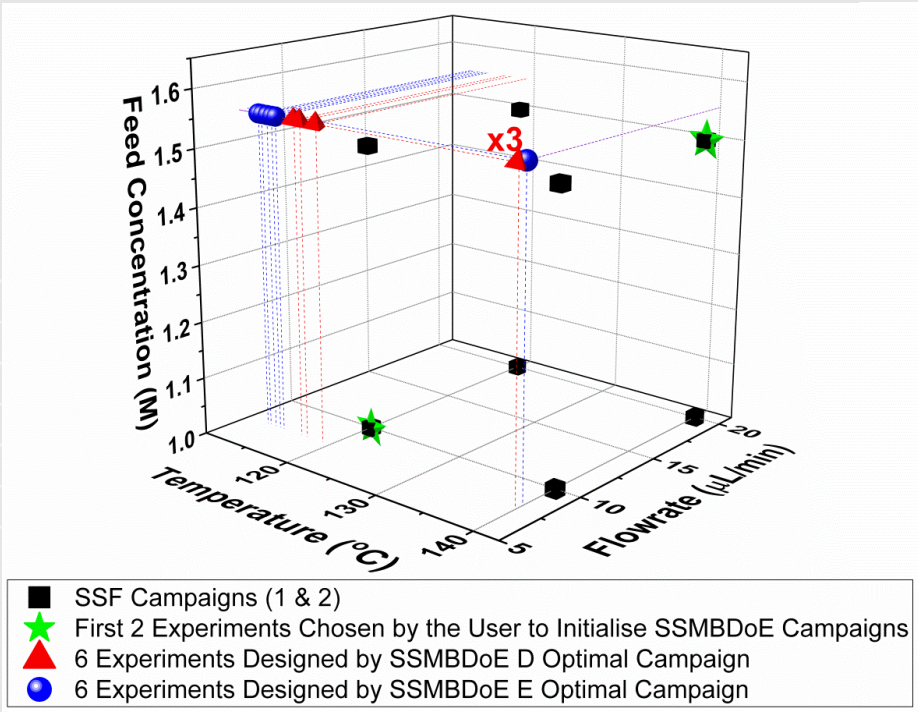
Case study 1: Benzoic acid esterification in a capillary microreactor

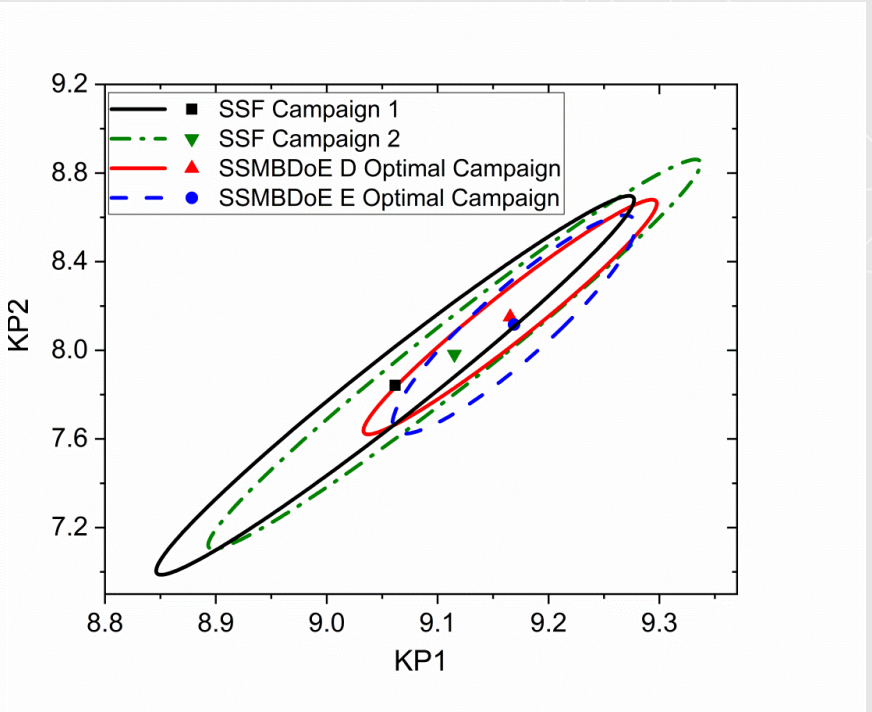
MBDoE VS DoE[1]

Steady-State Factorial (SSF) Vs Steady-State MBDoE (SSMBDoE) campaigns

Uncertainty of parameters KP1 and KP2 for the Steady-State Factorial (SSF) and Steady-State MBDoE (SSMBDoE) D- and E- optimal campaigns

$$r_{BzAcid} = -kC_{BzAcid} \qquad k = exp\left(-KP1 - \frac{KP2 * 10000}{R} * \left[\frac{1}{T} - \frac{1}{T_M}\right]\right)$$





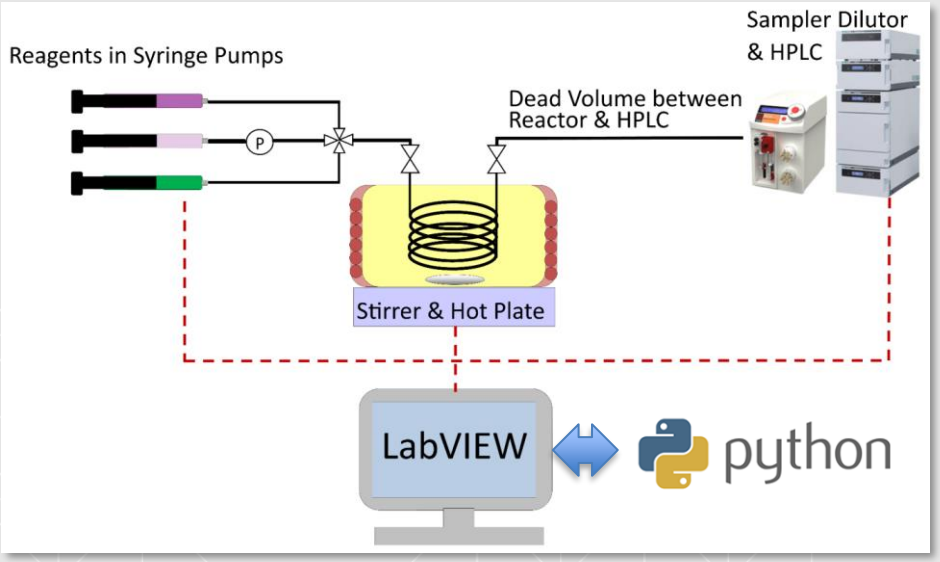
- Steady-state MBDoE experiments produced **significantly smaller uncertainty regions** than the ones designed by a full factorial DoE using the **same number of steady-state experiments**

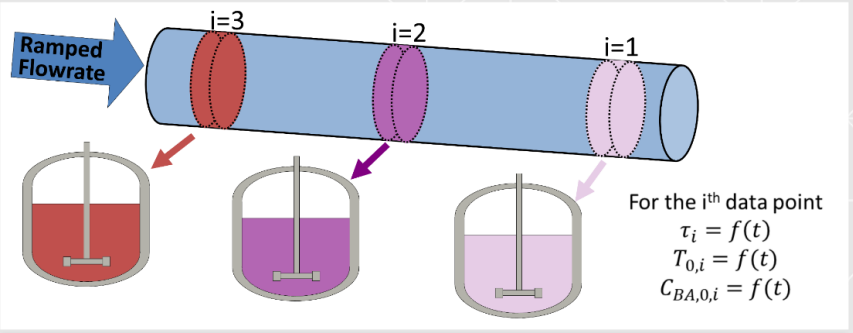
[1]Waldron, C; Pankajakshan, A; Quaglio, M; Cao, E; Galvanin, F; Gavriilidis, A. (2019) An autonomous microreactor platform for the rapid identification of kinetic models. **Reaction Chemistry and Engineering** 4, 1623. 10.1039/C8RE00345A.

Case study 1: Benzoic acid esterification in a capillary microreactor

Optimal design of transient flow experiments

Optimal design of transient flow experiments for the Identification of kinetic parameters[1]





$$\frac{dC_{BA,i}}{d\tau} = r_{BA,i} \quad \text{for } i = 1, \dots, N_{sp}$$

$$\frac{dT_i}{d\tau} = -\alpha_T \quad \text{for } i = 1, \dots, N_{sp}$$

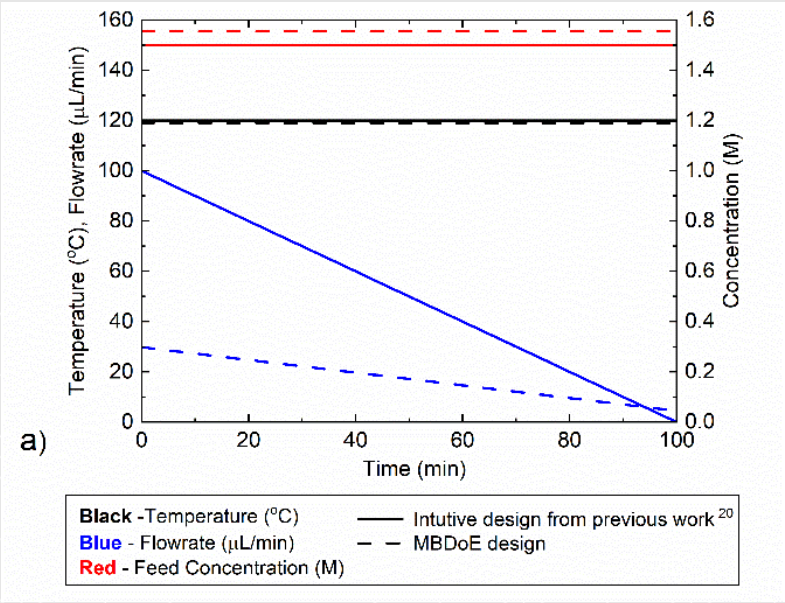
- The transient PFR can be described by a **system of ideal batch reactors**
- Dead volume between end of reactor and sampling needs to be known
- Transient experiments can provide **faster estimation of kinetic parameters**
- **Design vector for MBDoE:** $\varphi = [T, F, c_{\text{IN, BA}}, \alpha_T, \alpha_V, \alpha_C]$

[1]Waldron, C; Pankajakshan, A; Quaglio, M; Cao, E; Galvanin, F; Gavriilidis, A; (2020) Model-based design of transient flow experiments for the identification of kinetic parameters. **Reaction Chemistry & Engineering**, 5, 112. 10.1039/c9re00342h.

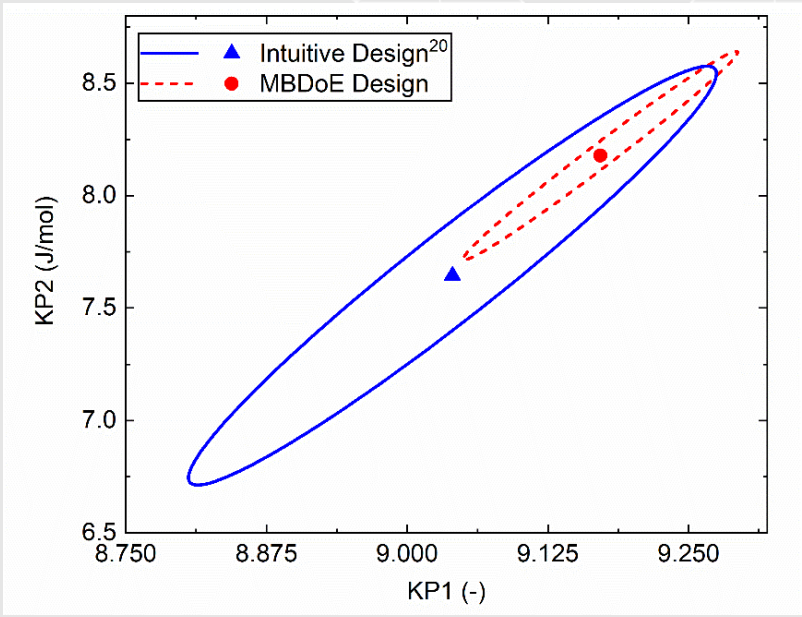
Case study 1: Benzoic acid esterification in a capillary microreactor

Optimal design of transient flow experiments[1]

Control variable profiles (temperature, flowrate and benzoic acid inlet concentration) designed **intuitively** and by **MBDoE**



Confidence ellipsoids of kinetic parameters for transient experiments designed intuitively and by MBDoE



- **MBDoE improves significantly the parameter precision obtained from transient experiments**
- Parameter precision obtained in transient from **a single run:**

  → similar to that obtained from **8 steady-state MBDoE runs**[1]
- … transient experiments required significantly less time and resources to run!

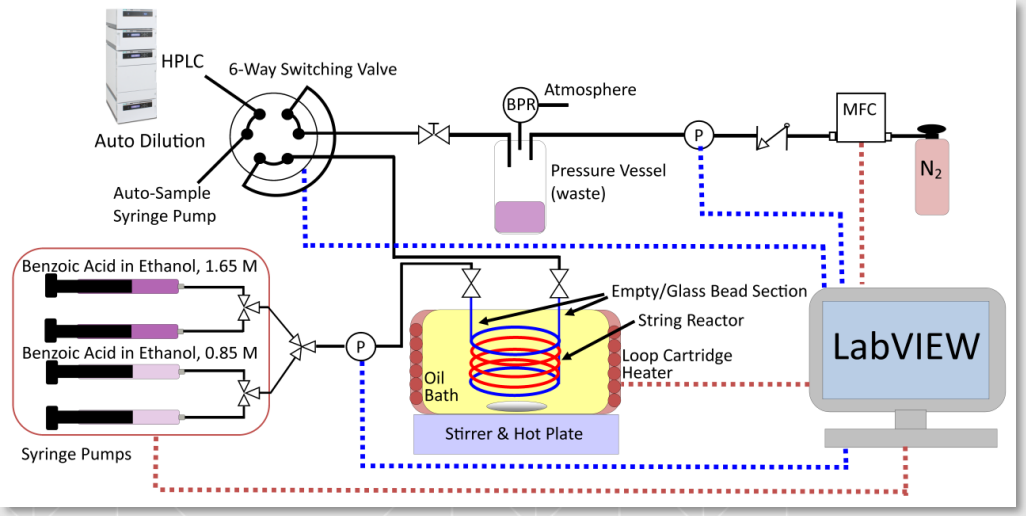
[1]Waldron, C; Pankajakshan, A; Quaglio, M; Cao, E; Galvanin, F; Gavriilidis, A; (2020) Model-based design of transient flow experiments for the identification of kinetic parameters. **Reaction Chemistry & Engineering**, 5, 112. 10.1039/c9re00342h.

Case study 1: Benzoic acid esterification in a capillary microreactor

Closed-loop identification including MBDoE-MD and MBDoE-PE

Closed Loop MBDoE for Kinetic Model Discrimination and Parameter Estimation[1]

Benzoic Acid Esterification on Amberlyst-15 (heterogeneous) with Ethanol



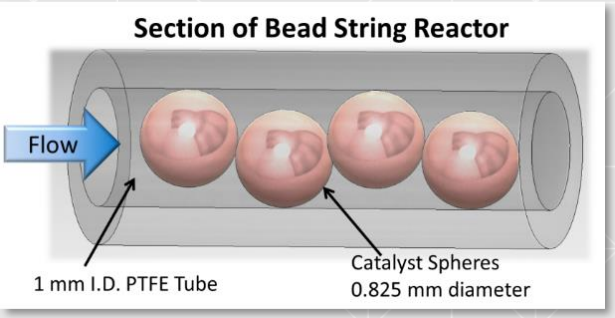
Four candidate kinetic models

$$r'_{BA} = -kC_{BA}C_{EtOH} \quad (M1)$$

$$r'_{BA} = \frac{-kC_{BA}C_{EtOH}}{(1 + K_W C_W)^2} \quad (M2)$$

$$r'_{BA} = \frac{-kC_{BA}C_{EtOH}}{(1 + K_W C_W + K_{EtOH}C_{EtOH})^2} \quad (M3)$$

$$r'_{BA} = \frac{-kC_{BA}C_{EtOH}}{(1 + K_{BA}C_{BA} + K_{EtOH}C_{EtOH} + K_W C_W + K_{EB}C_{EB})^2} \quad (M4)$$



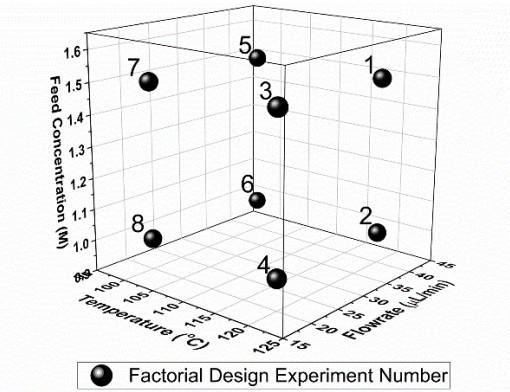
Bead-string reactor offers **small external mass transfer resistance**

[1]Waldron, C; Pankajakshan, A; Quaglio, M; Cao, E; Galvanin, F; Gavriilidis, A; (2019) Closed-Loop Model-Based Design of Experiments for Kinetic Model Discrimination and Parameter Estimation: Benzoic Acid Esterification on a Heterogeneous Catalyst. **Industrial & Engineering Chemistry Research** , 58, 2216. 10.1021/acs.iecr.9b04089.

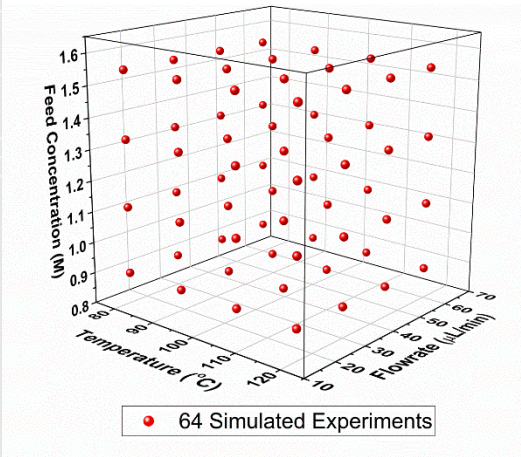
Case study 1: Benzoic acid esterification in a capillary microreactor

Closed-loop identification including MBDoE-MD and MBDoE-PE

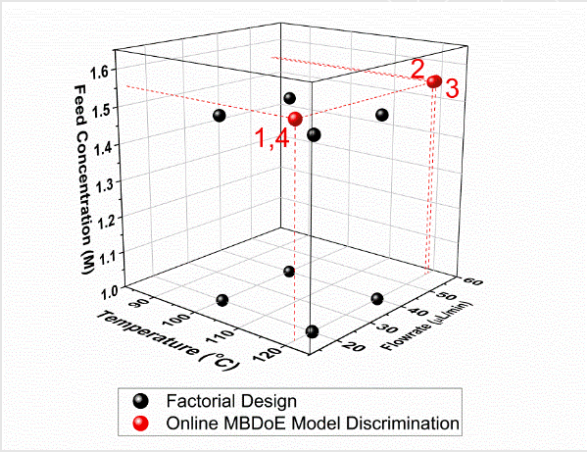
Closed Loop MBDoE for Kinetic Model Discrimination and Parameter Estimation



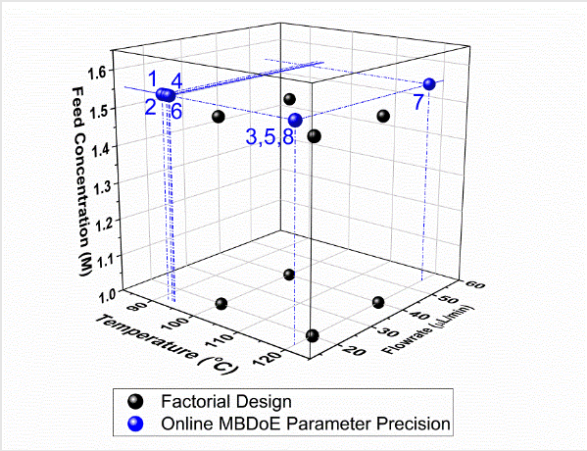
1. 8 steady-state experiments designed by a full factorial DoE method
   M1, M2, M3, M4



2. 64 simulated experiments to test **practical model identifiability based on rank of FIM**
   M1, M2



3. MBDoE-designed discriminating experiments (MBDoE-MD)
   M2



4. MBDoE designed experiments for improved parameter precision (MBDoE-PE)
   M2

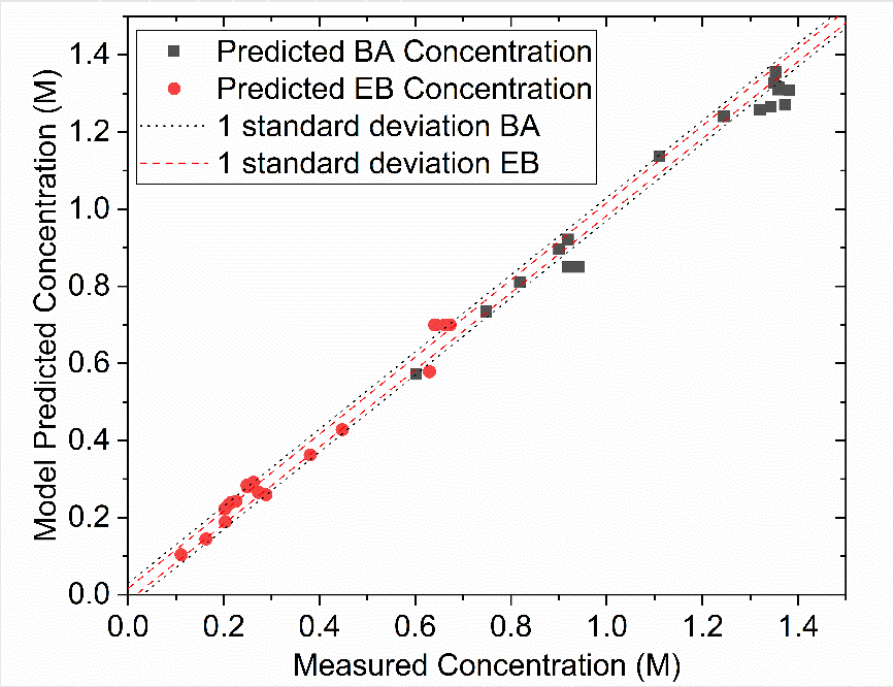
[1]Waldron, C; Pankajakshan, A; Quaglio, M; Cao, E; Galvanin, F; Gavriilidis, A; (2019) Closed-Loop Model-Based Design of Experiments for Kinetic Model Discrimination and Parameter Estimation: Benzoic Acid Esterification on a Heterogeneous Catalyst. **Industrial & Engineering Chemistry Research** , 58, 2216. 10.1021/acs.iecr.9b04089.

Case study 1: Benzoic acid esterification in a capillary microreactor

Closed-loop identification including MBDoE-MD and MBDoE-PE

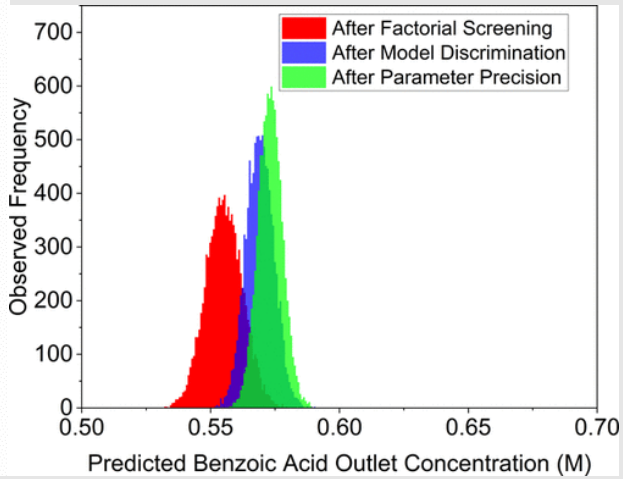
Identified (apparent) kinetic model (Model 2)

$$r'_{BA} = \frac{185.3 \exp\left(-\frac{68\,800}{RT}\right) \times C_{BA} C_{EtOH}}{(1 + 0.53 C_W)^2}$$



- Bead string reactor provided plug flow behaviour with efficient mass transfer
- The methodology proposed **minimized the number of experiments** required for kinetic model identification
- **It prevents the use of nonidentifiable models**
- MBDoE-MD distinguished between two candidate models
- MBDoE-PE improved parameter precision
- All kinetic experiments were **completed in 3 days**

… against 2 weeks for a full factorial DoE!



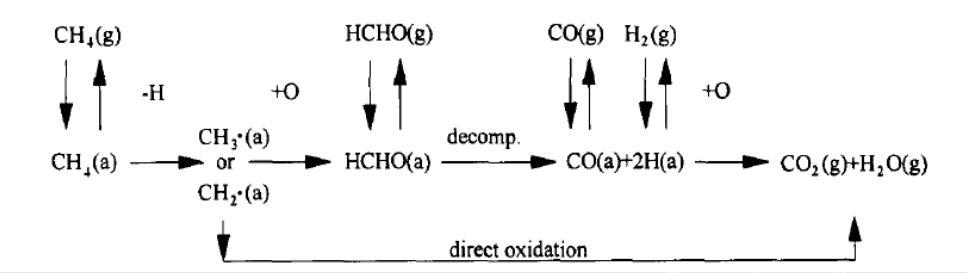
[1]Waldron, C; Pankajakshan, A; Quaglio, M; Cao, E; Galvanin, F; Gavriilidis, A; (2019) Closed-Loop Model-Based Design of Experiments for Kinetic Model Discrimination and Parameter Estimation: Benzoic Acid Esterification on a Heterogeneous Catalyst. **Industrial & Engineering Chemistry Research** , 58, 2216. 10.1021/acs.iecr.9b04089.

Case study 2: Catalytic methane oxidation

Catalytic methane oxidation in flow microreactor systems

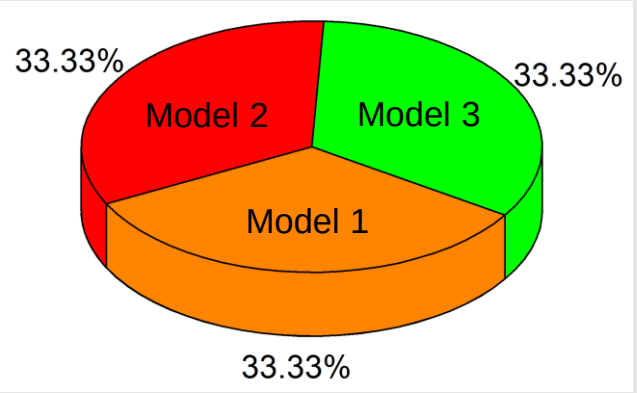
Overall chemical reaction of methane oxidation over Pd/Al2O3 catalyst[1]:

$$CH_4 + 2\,O_2 \xrightarrow{Pd/Al_2O_3} CO_2 + 2\,H_2O$$



Three candidate kinetic models are considered for the methane oxidation reaction after a **screening and identifiability study based on 12 candidate models** [2]

Model	Rate law
Model 1	$r_{CH_4} = kP_{avg}y_{CH_4}$
Model 2 (LH)	$r_{CH_4} = \dfrac{k_r K_{CH_4} P_{CH_4} \sqrt{K_{O_2} P_{O_2}}}{\left(1 + K_{CH_4} P_{CH_4} + \sqrt{K_{O_2} P_{O_2}}\right)^2}$
Model 3 (MVK)	$r_{CH_4} = \dfrac{k_1 k_2 P_{CH_4} P_{O_2}}{k_1 P_{O_2} + 2k_2 P_{CH_4} + (k_1 k_2 / k_3) P_{O_2} P_{CH_4}}$



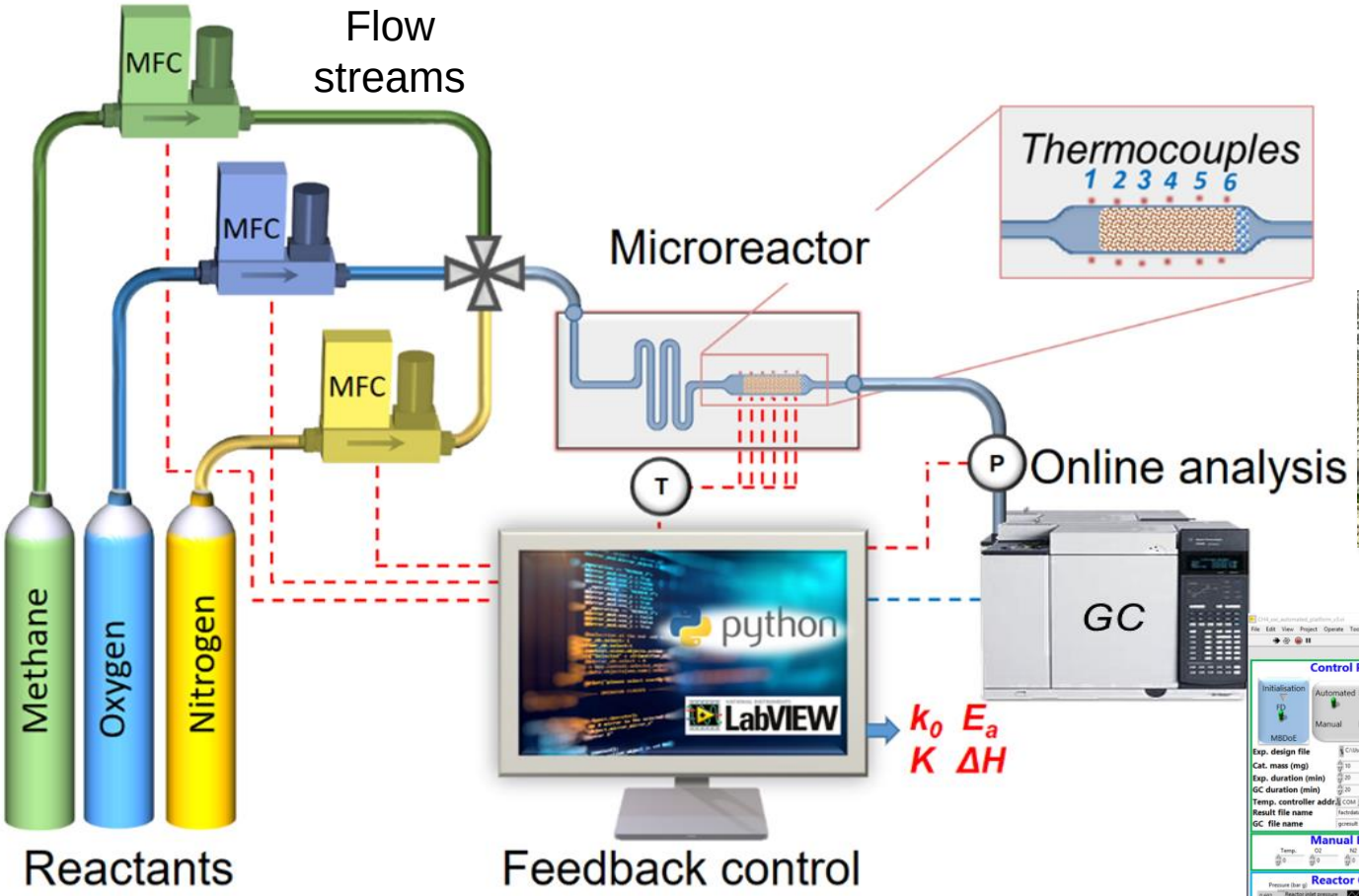
Initial probability of model correctness

[1] J. H. Lee, D. L. Trimm (1995), Catalytic combustion of methane, Fuel Processing Technology 42, 339.
[2] S. G. Bawa, A. Pankajakshan, C. Waldron, E. Cao, F. Galvanin, A, Gavriilidis (2023). *Chemistry–Methods*, *3*, 1.

28

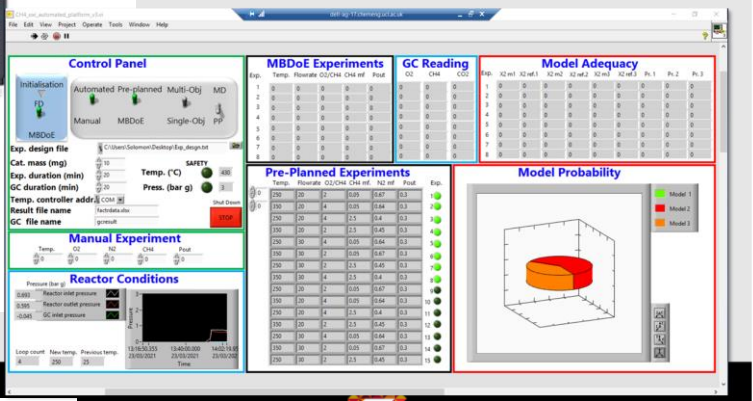
Case study 2: catalytic methane oxidation

The platform – microreactor technology



Reactor Setup

- Silicon-glass reactor
- 10 mg of 5 wt.% Pd/Al_2O_3
- ±3 °C axially at 400 °C
- Online analysis by GC

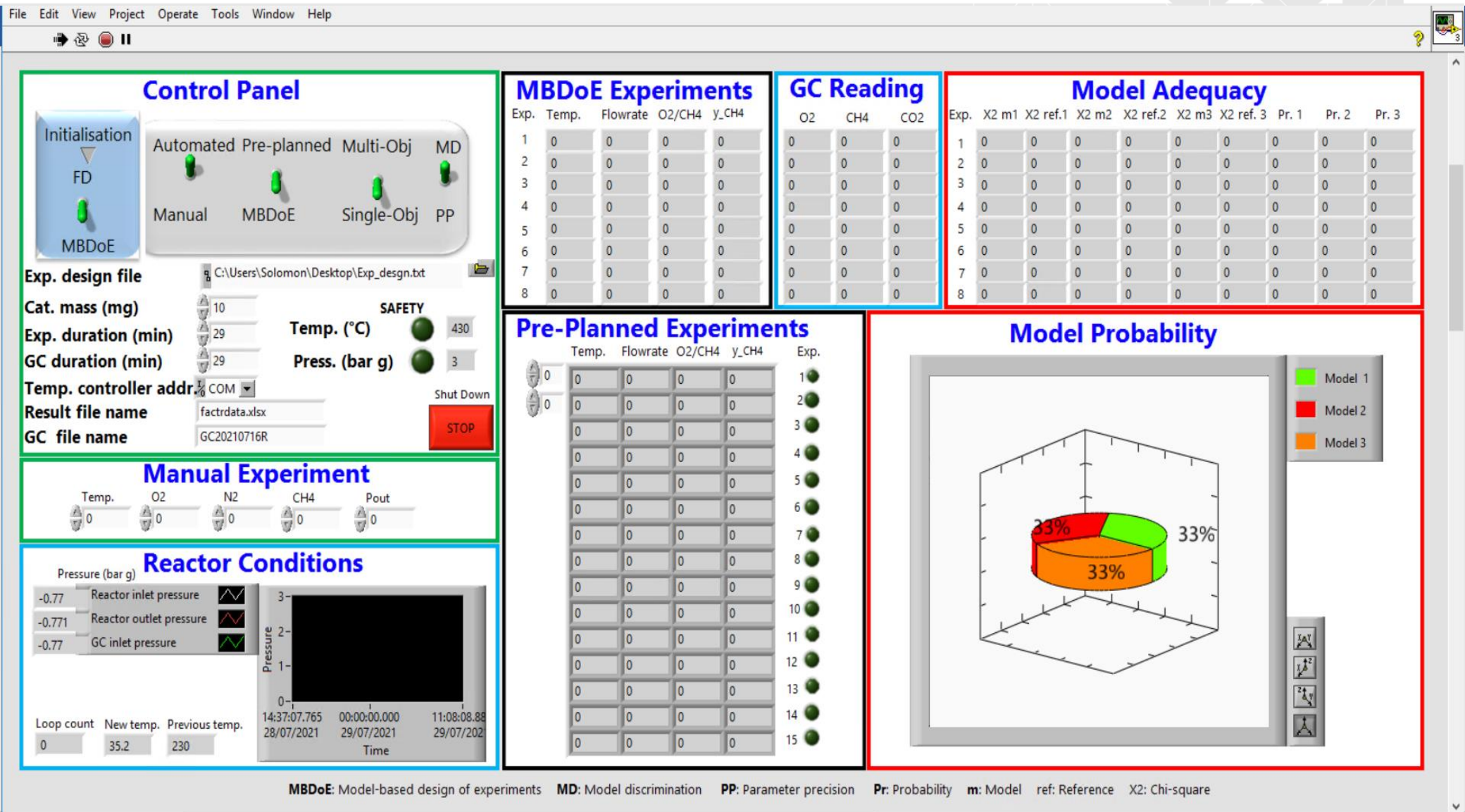


https://www.youtube.com/watch?v=kMCtQqbPixk

29

Case study: catalytic methane oxidation

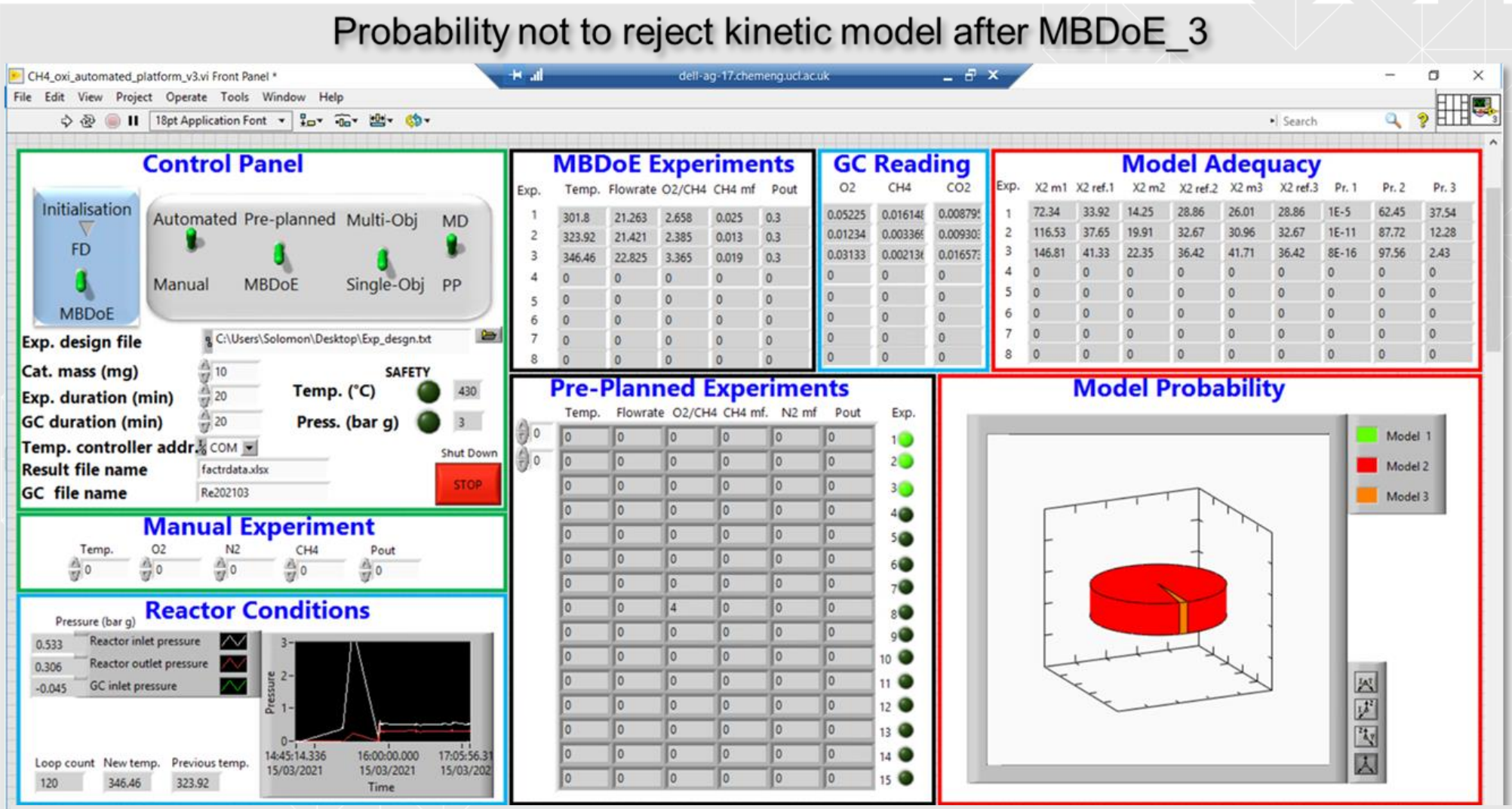
LabView Graphical User Interface and digital twin



https://www.youtube.com/watch?v=kMCtQqbPixk

Case study: catalytic methane oxidation

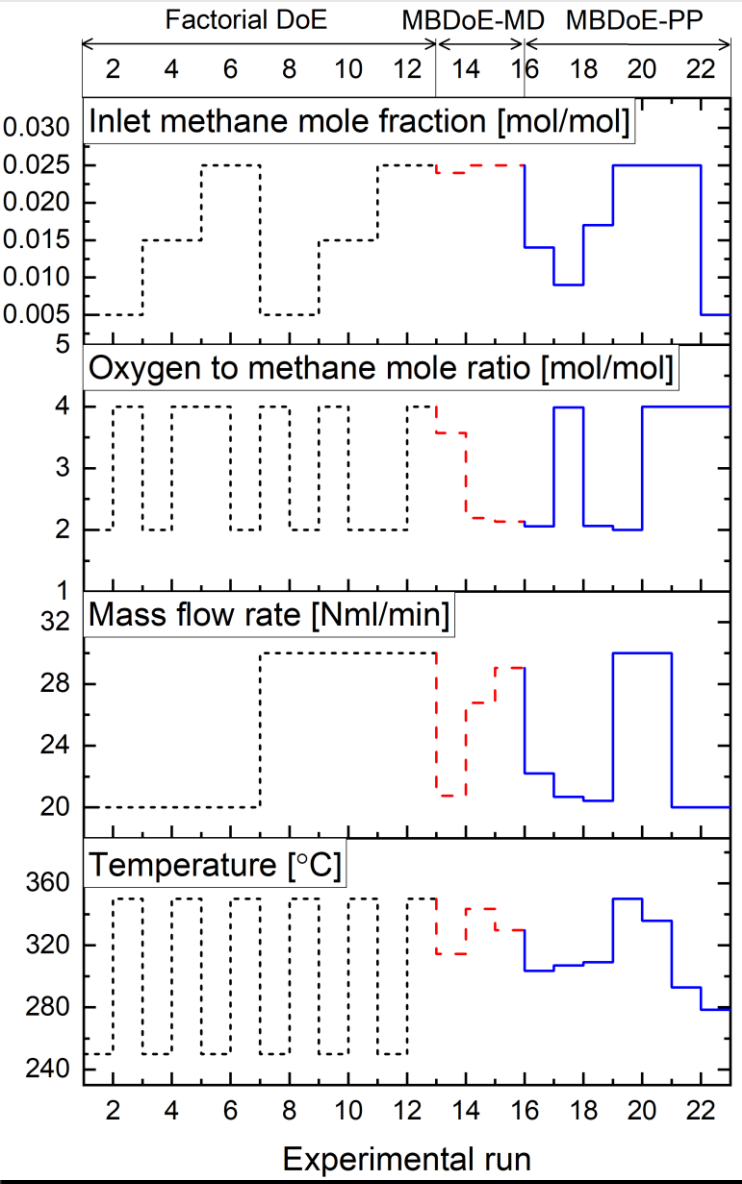
LabView Graphical User Interface and digital twin

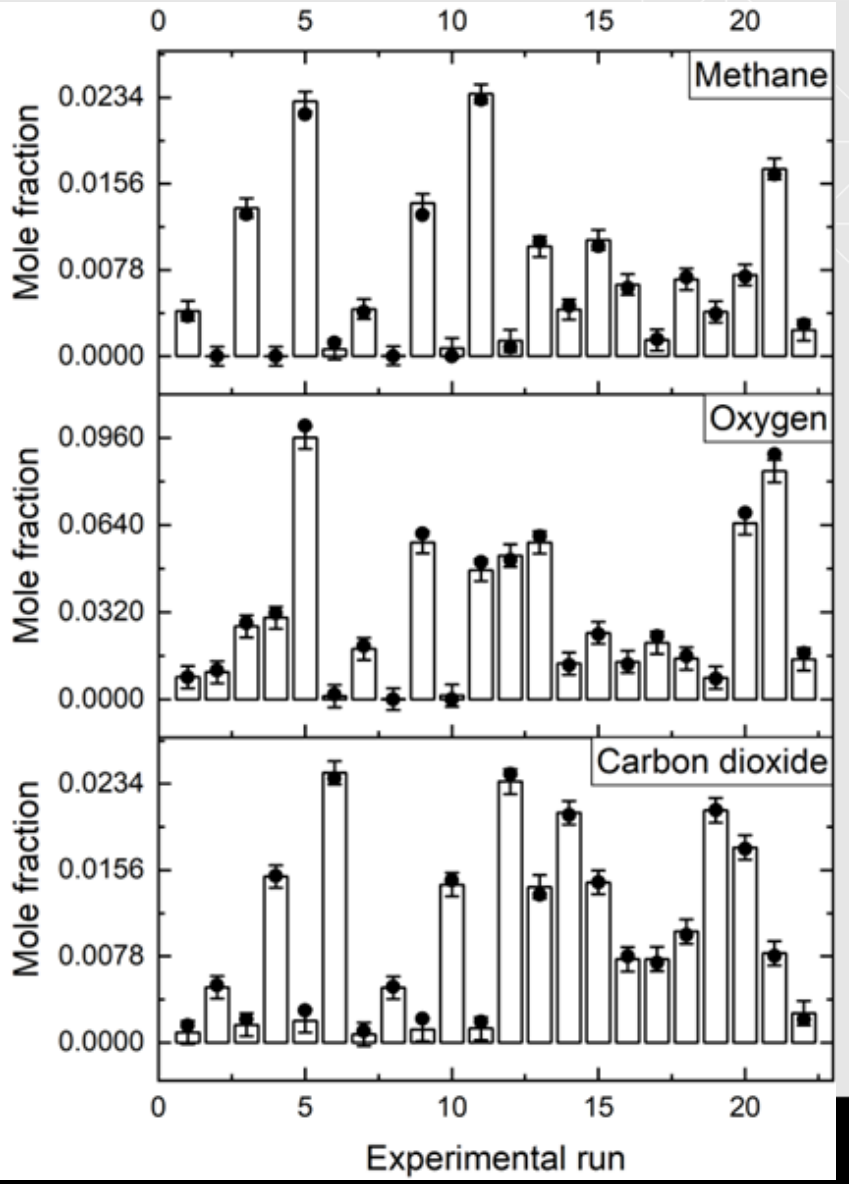


https://www.youtube.com/watch?v=kMCtQqbPixk

Case study: catalytic methane oxidation

Results from the autonomous platform: inputs/outputs





Campaign for kinetic model validation:

1. Preliminary set of DoE experiments for **pre-screening**
2. **MBDoE-MD runs**
3. **MBDoE-PE runs**

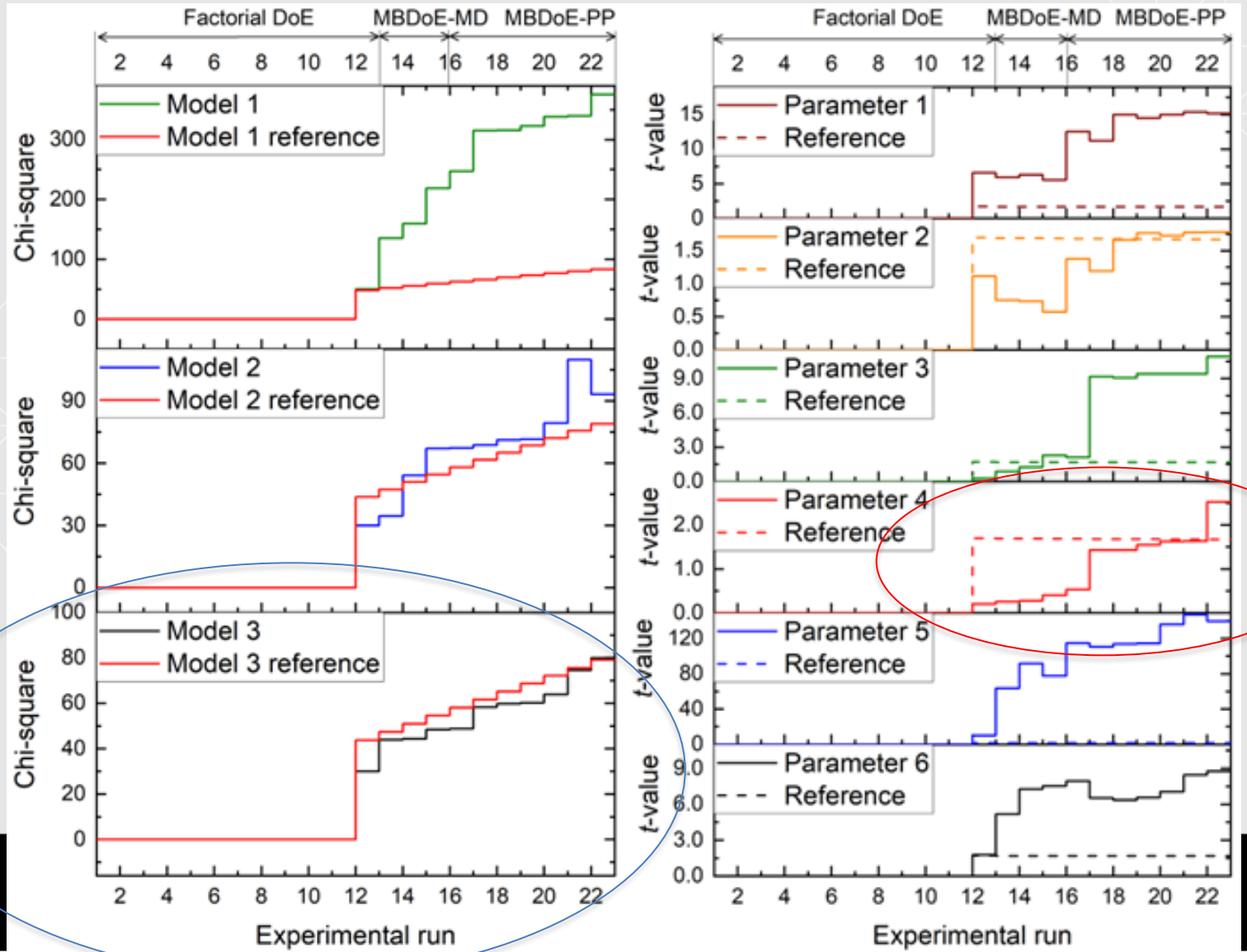
- Switch between steps 2) and 3) dictated by the **best model probability (>95%)**
- **STOP** when parameters are all precisely estimated (**t-test satisfied**)

Case study: catalytic methane oxidation

Results from the autonomous platform: model calibration

Model Adequacy

Parameter precision (Model 3)

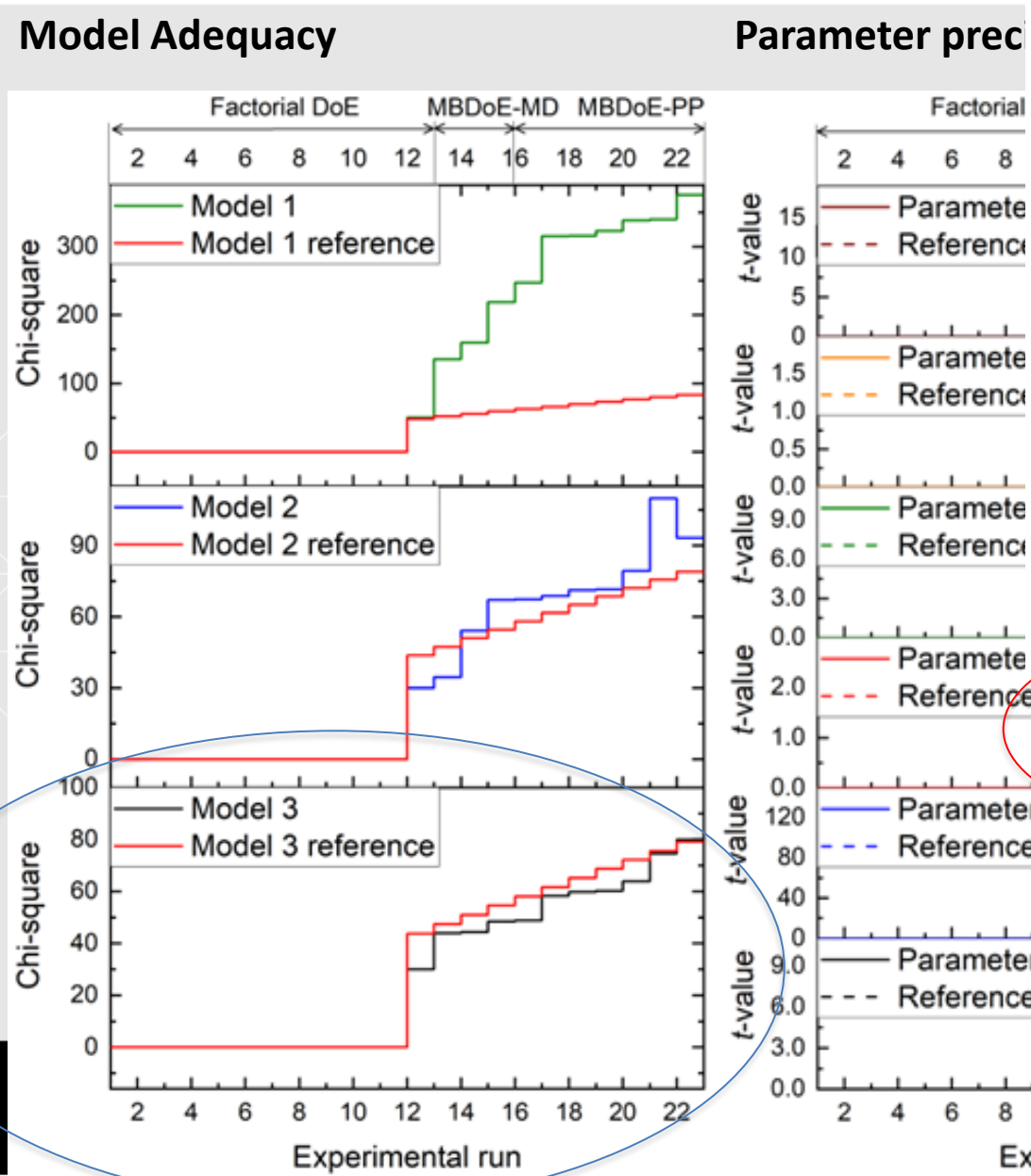


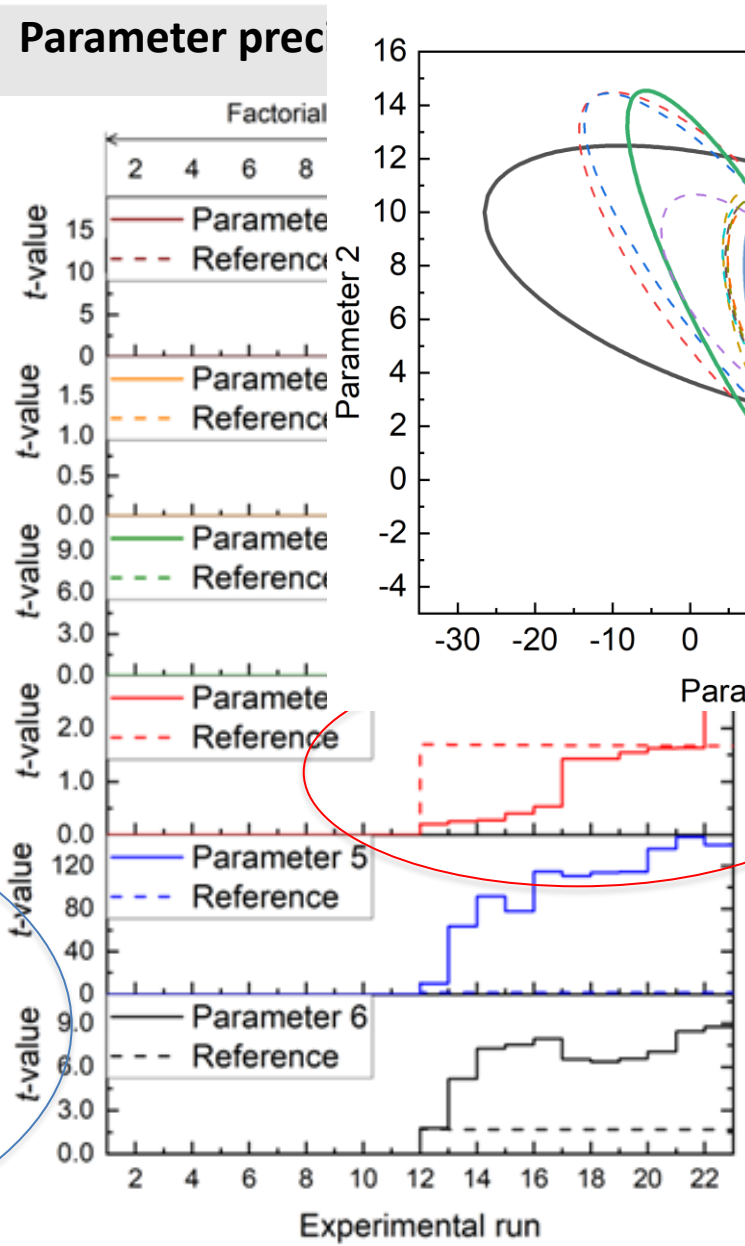
- Only **4 MBDoE-MD** experiments required for model discrimination!
- Additional **6 experiments** needed for a precise estimation of kinetic model parameters

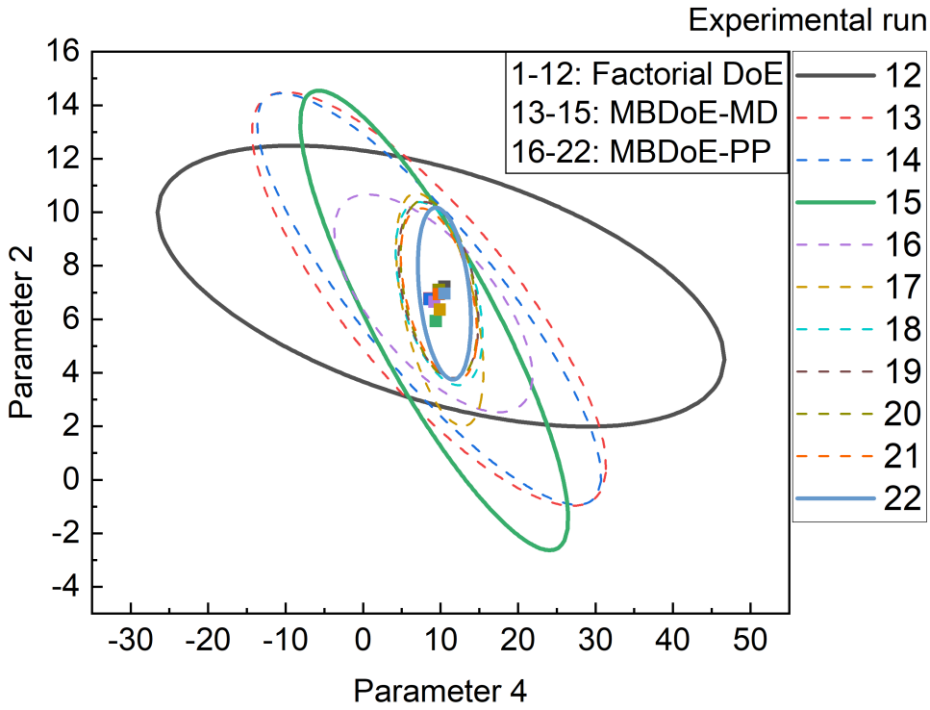
UCL ENGINEERING
Change the world

Case study: catalytic methane oxidation

Results from the autonomous platform: model calibration







UCL ENGINEERING
Change the world

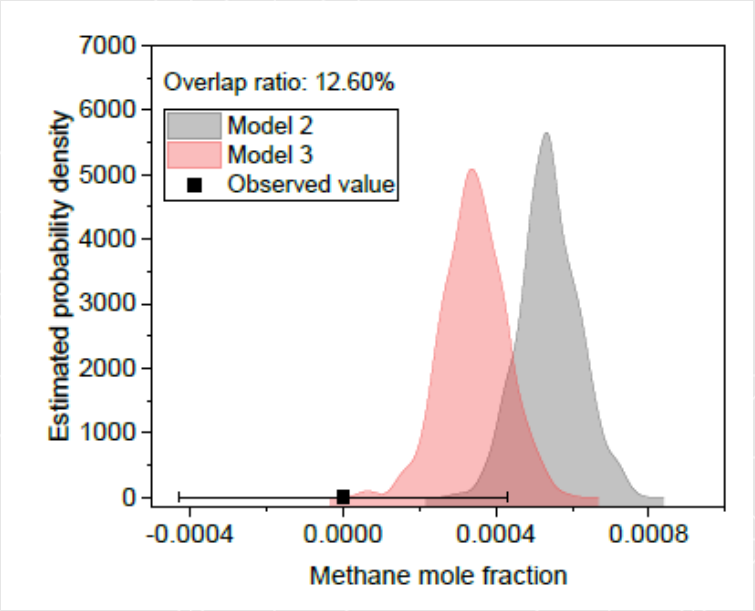
Case study: catalytic methane oxidation

The problem of practical model distinguishability[1]

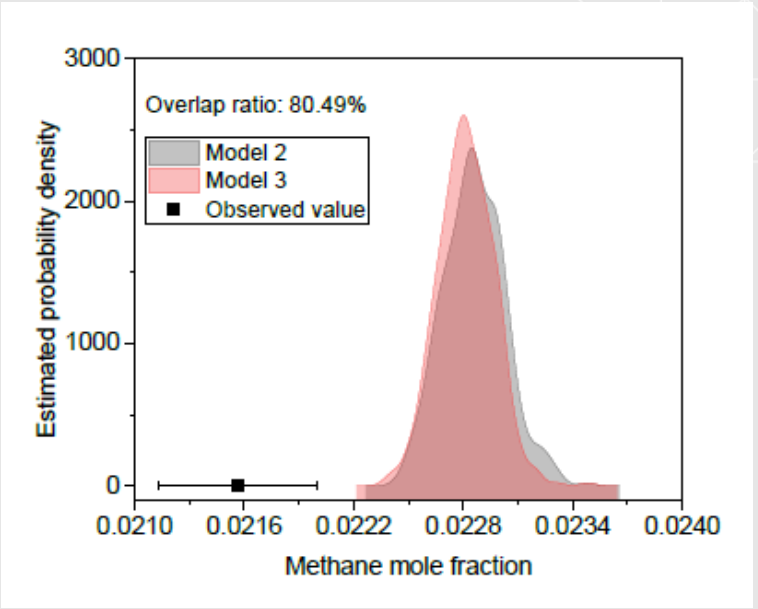
Results can be analysed retrospectively

- **Probability density plots of model responses**
- **Overlap ratio**

Arun Pankajakshan
Research Assistant



"Easily" distinguishable models



Challenging model discrimination

[1]Pankajakshan, A., Bawa, S., Gavriilidis, A., Galvanin, F. (2023). Autonomous kinetic model identification using optimal experimental design and retrospective data analysis: methane complete oxidation as a case study. Reaction Chemistry & Engineering (in press).

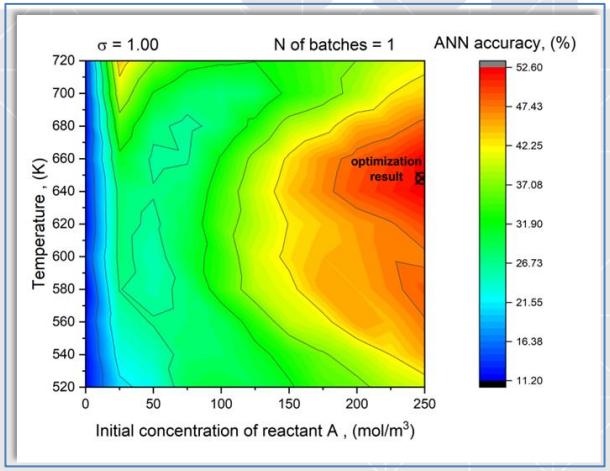
Machine Learning (ML)-Assisted Model Identification

Goals and features

Integration of machine learning (ML) techniques and MBDoE:

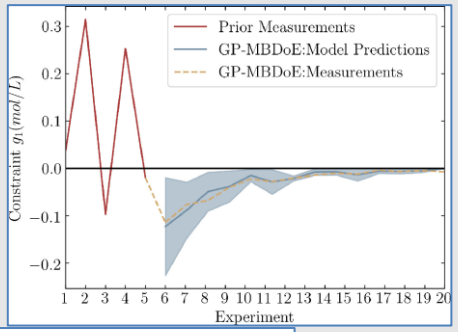
1. **To assist model discrimination and selection**

   Optimal Experimental Design for Kinetic Model Recognition Using Artificial Neural Networks (ANNs)[1,2]



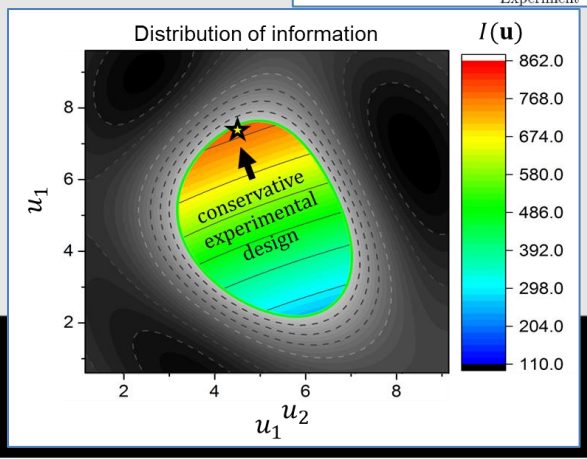
2. **To efficiently design experiments under uncertain scenarios**

   Safe model-based design of experiments using Gaussian processes[3]



3. **To identify design regions of model reliability**

   Machine Learning (ML)-Assisted Model Reliability Mapping[4]



[1] Quaglio et al., (2020), Computers & Chemical Engineering 135, 106759
[2] Sangoi et al. (2022), CACE, 49, 117
[3] Petsagkourakis, P., Galvanin, F. (2021), Computers & Chemical Engineering 151, 107339
[4] Quaglio et al. (2018), Chemometrics and intelligent laboratory systems 172, 58

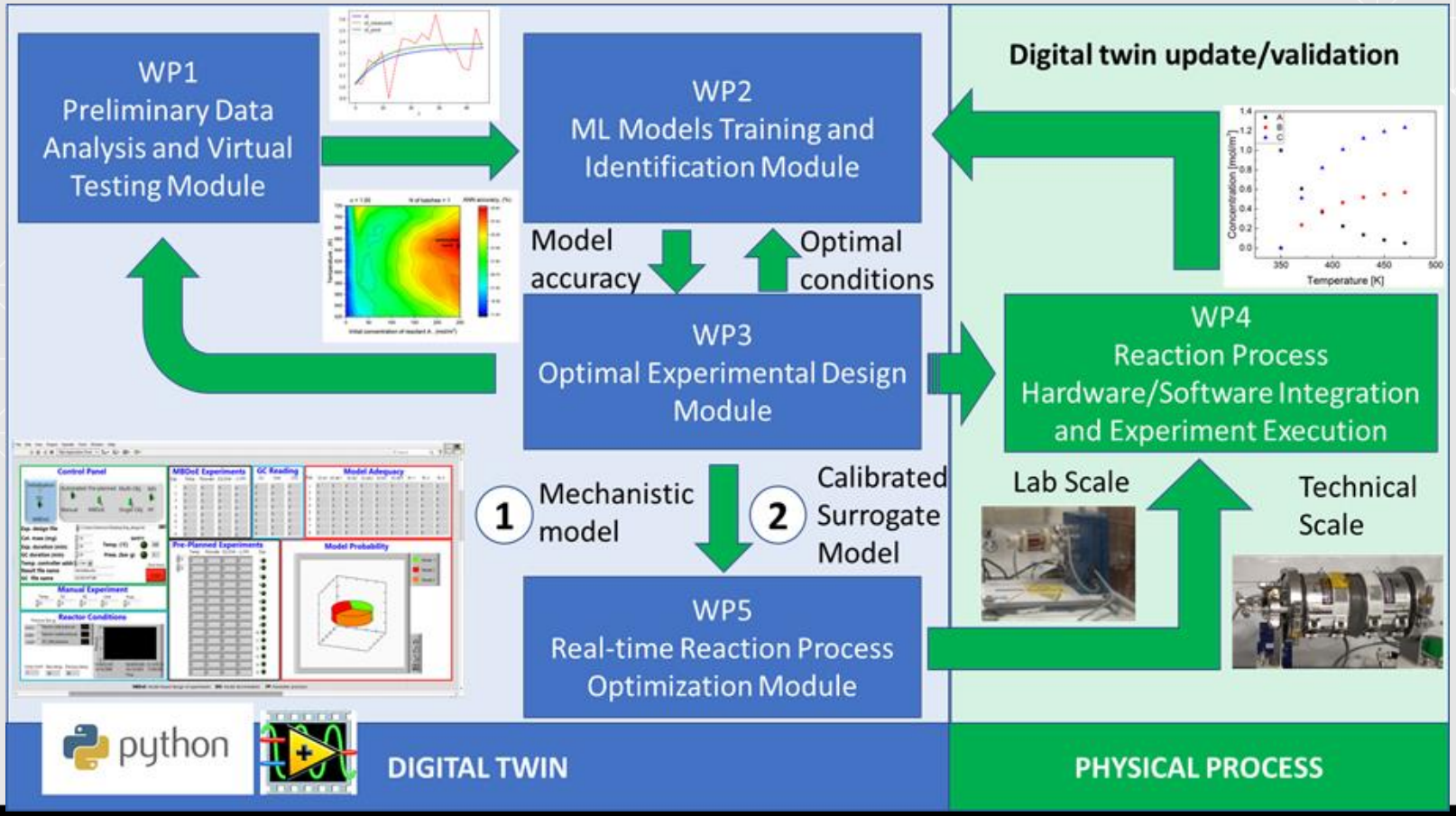
UCL ENGINEERING
Change the world

Machine Learning (ML)-Assisted Model Identification

ML-assisted digital twin platform for real-time optimisation of reaction systems

GOAL: Coupling physics-informed ML and optimal experimental design algorithms for reaction optimisation[1]

PI: Federico Galvanin
f.galvanin@ucl.ac.uk



EPSRC
Engineering and Physical Sciences Research Council

gsk

JM Johnson Matthey
Inspiring science, enhancing life

Quotient Sciences

Autichem

BASF

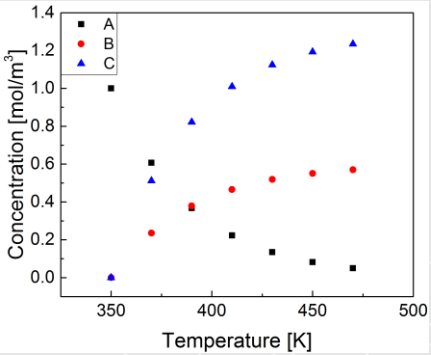
[1]https://gow.epsrc.ukri.org/NGBOViewGrant.aspx?GrantRef=EP/X024016/1

UCL ENGINEERING
Change the world

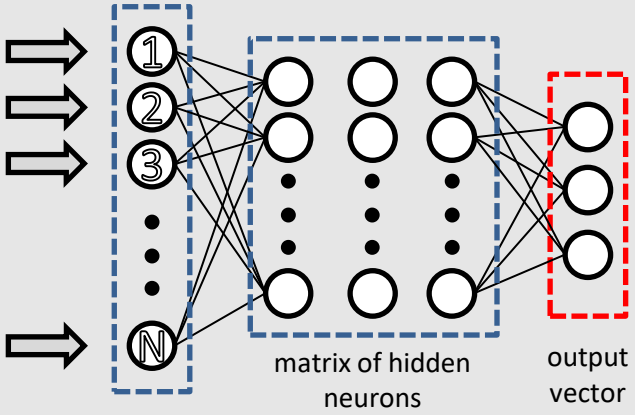
Kinetic Model Recognition using ANNs

Proposed approach

Experimental data



Artificial Neural Network



data fed to the ANN as input

Output

"KINETIC MODEL"

$$r_{A\to B} = kC_A$$

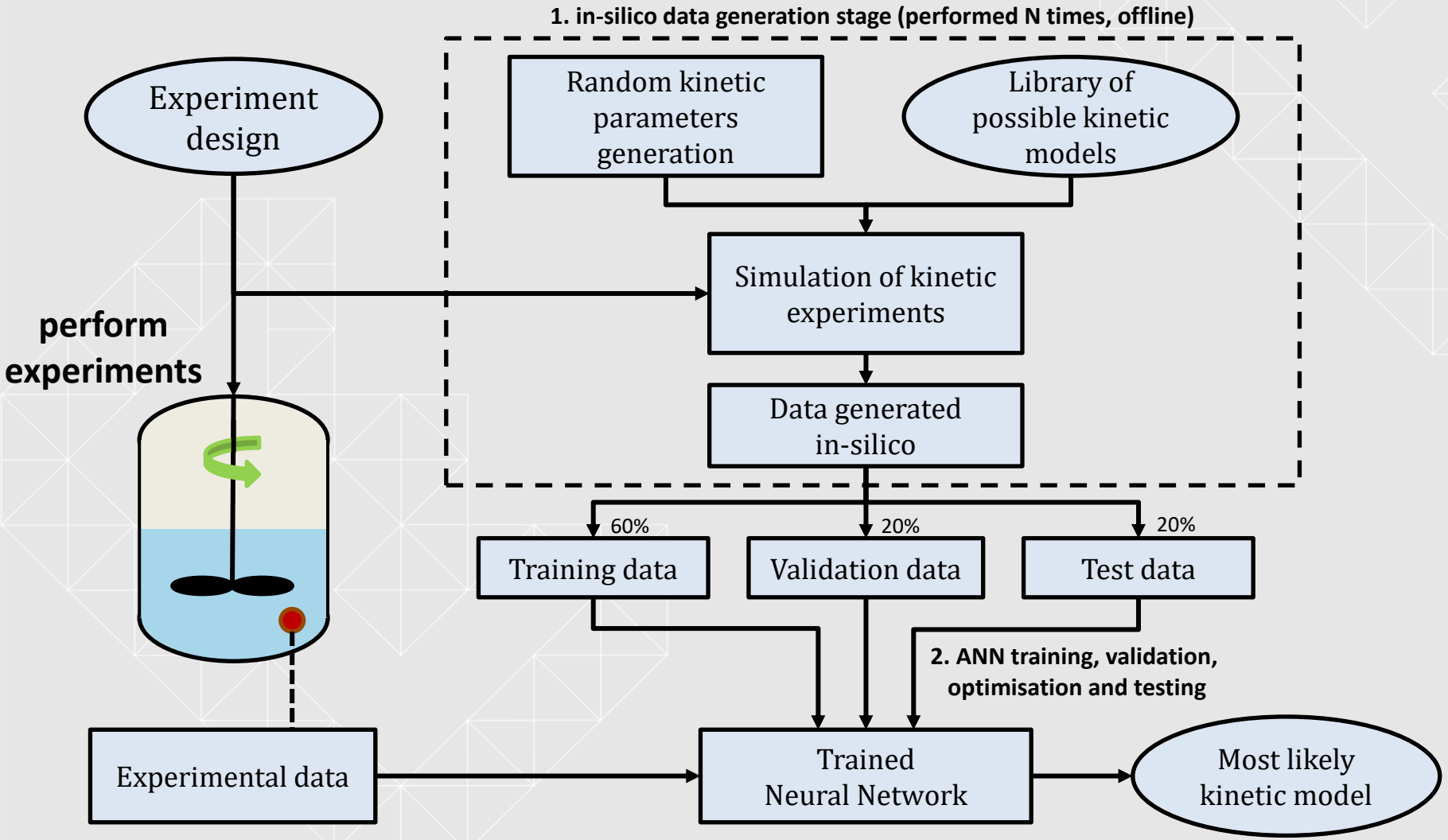
$$r_{A\to C} = kC_A^2$$

$$r_{B\to C} = 0$$

In the proposed approach, ANNs are used to model the **cognitive processes** of the **scientist's brain** rather than building a black-box representation of the physical system.

Kinetic Model Recognition using ANNs[1]

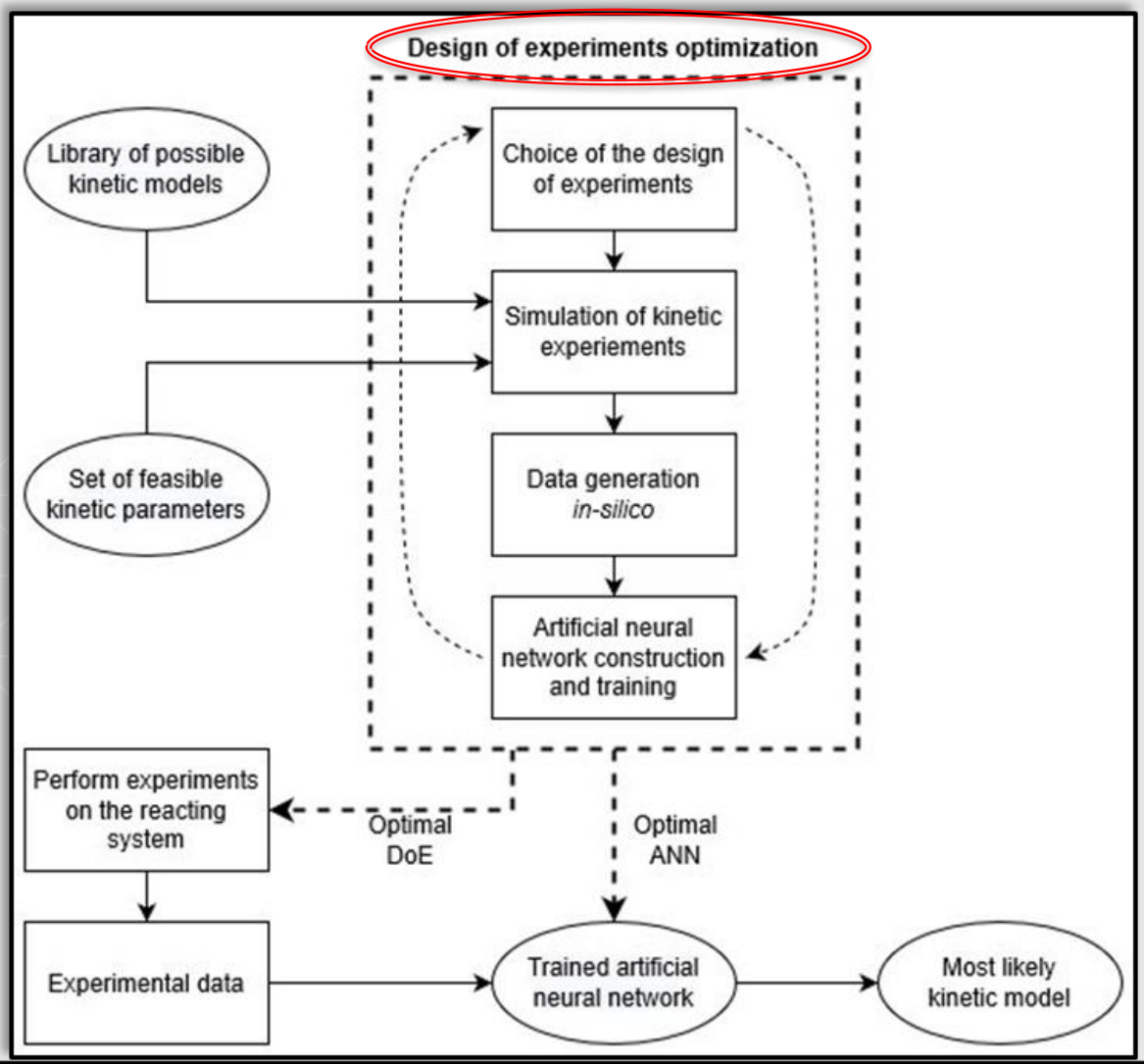
Proposed approach



[1]Quaglio, M., Roberts, L., Jaapar, M., Dua, V., Galvanin, F. (2020). An artificial neural network framework for classifying chemical reaction types from experimental data. *Computers and Chemical Engineering.*, 135, 106759.

ANNs-based Optimal Experimental Design

Proposed procedure



Input layer size:	3 x (total number of samples)
One hidden layer:	100 nodes activation function = *rectified linear unit*
Output layer:	8 nodes, corresponding to each kinetic model activation function = *softmax*
Accuracy:	$\frac{\text{number of kinetic models correctly classified}}{\text{total number of kinetic models in the test set}}$

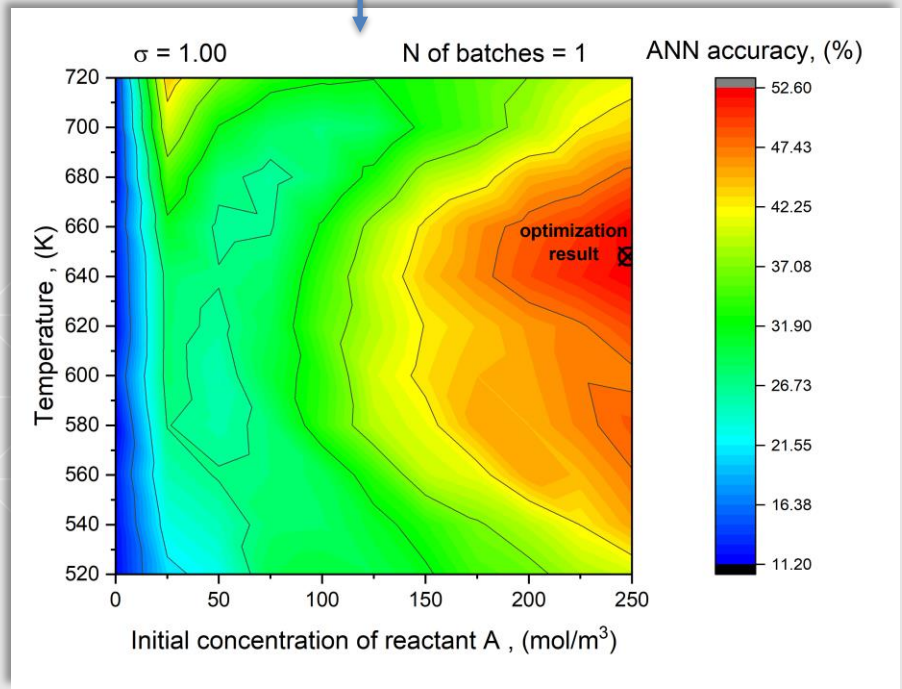
- **Objective function used in the optimal DoE**: **ANN test-accuracy**
- A **differential evolution algorithm** (DEA)[1] is chosen for the optimization → population-based algorithm inspired by the evolutionary theory

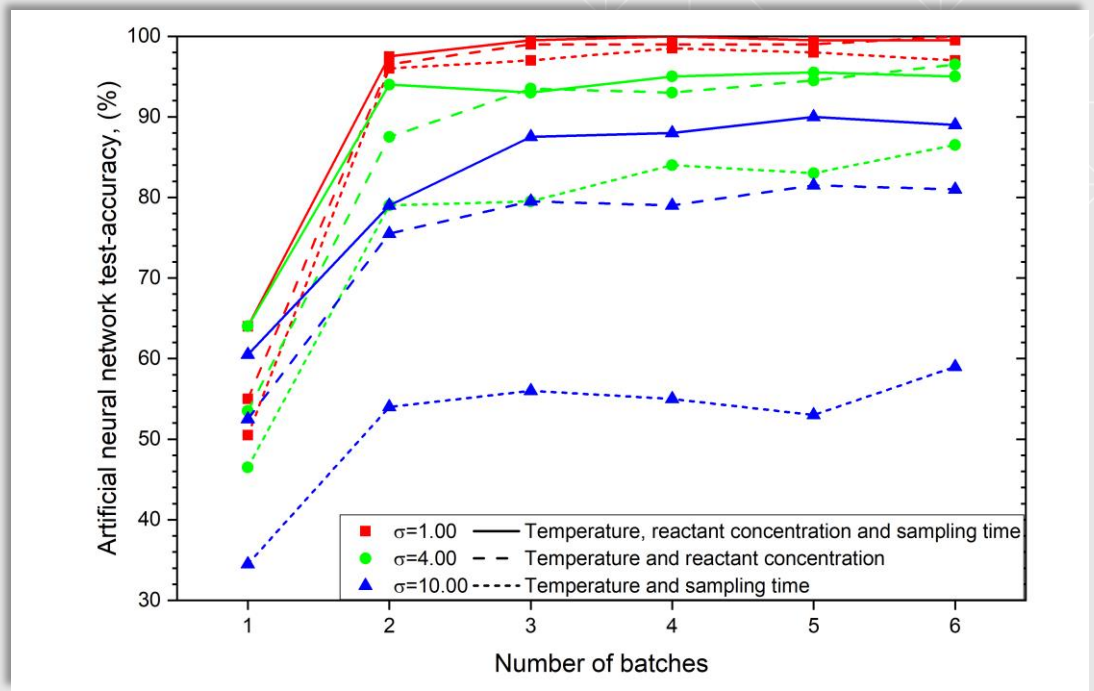
[1]Storn R., Price K. (1997). Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces", Journal of Global Optimization, 11, 4, 341-359. DOI: 10.1023/A:1008202821328

ANNs-based Optimal Experimental Design

Effect of design variables on ANN accuracy

Effect of **temperature** and **reactant A concentration** on ANN accuracy (1 sample)[1]





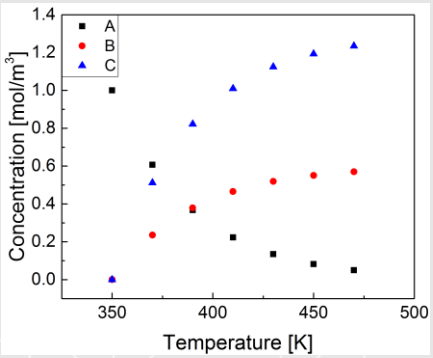
- **Very high accuracy** achieved with limited number of experiments in **low noise scenarios**
- **High accuracy achieved** (**close to 90%**) when temperature, reactant concentration and sampling time are optimised even in the most uncertain measurements scenario

[1]Sangoi, E., Quaglio, M., Bezzo, F., Galvanin, F. (2022). Optimal Design of Experiment Based on Artificial Neural Network Classifiers for Fast Kinetic Model Recognition. In: 14th International Symposium on Process Systems Engineering (PSE 2021, Kyoto).

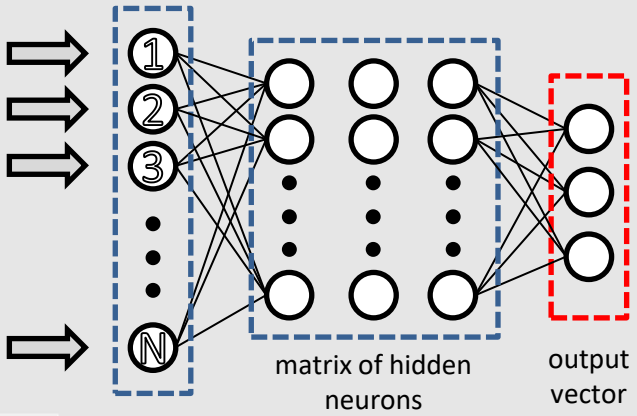
ANN-based Optimal Design Vs MBDoE-MD

Effect of dataset quality on discrimination performance[1]

Experimental data



Artificial Neural Network



data fed to the ANN as input

Output

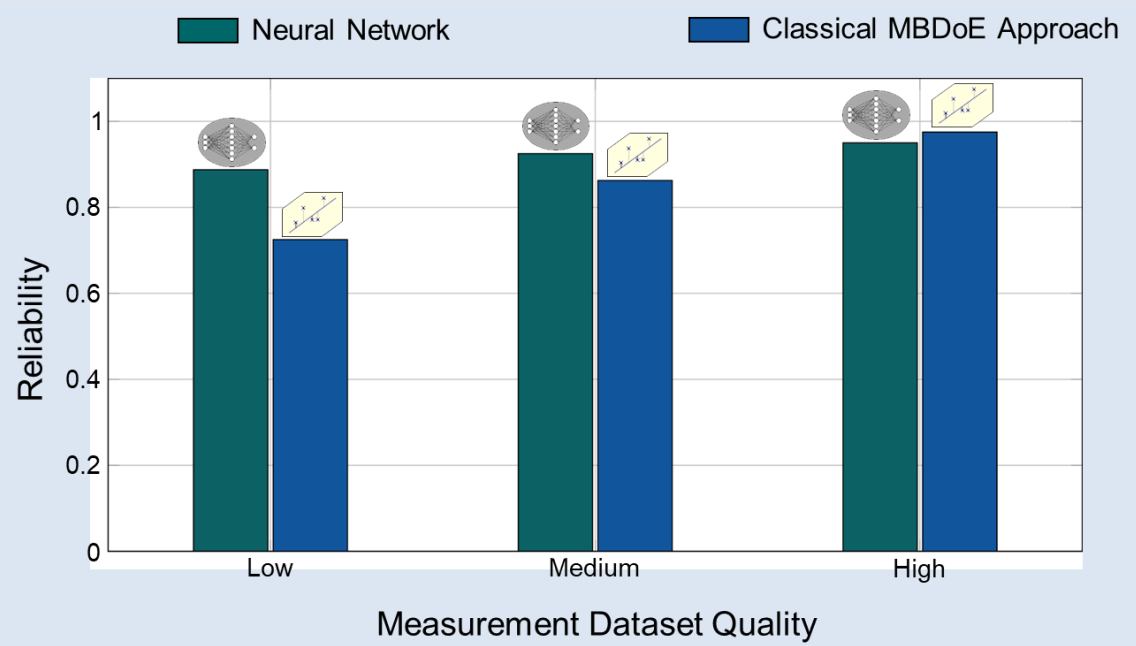
"KINETIC MODEL"

$$r_{A\to B} = kC_A$$

$$r_{A\to C} = kC_A^2$$

$$r_{B\to C} = 0$$

What if this approach is **compared to a standard MBDoE approach** for model discrimination?



[1]N. Riske (2021). Mechanistic Model Discrimination Using Model-Based Design of Experiments and Artificial Neural Networks (Master's Thesis).

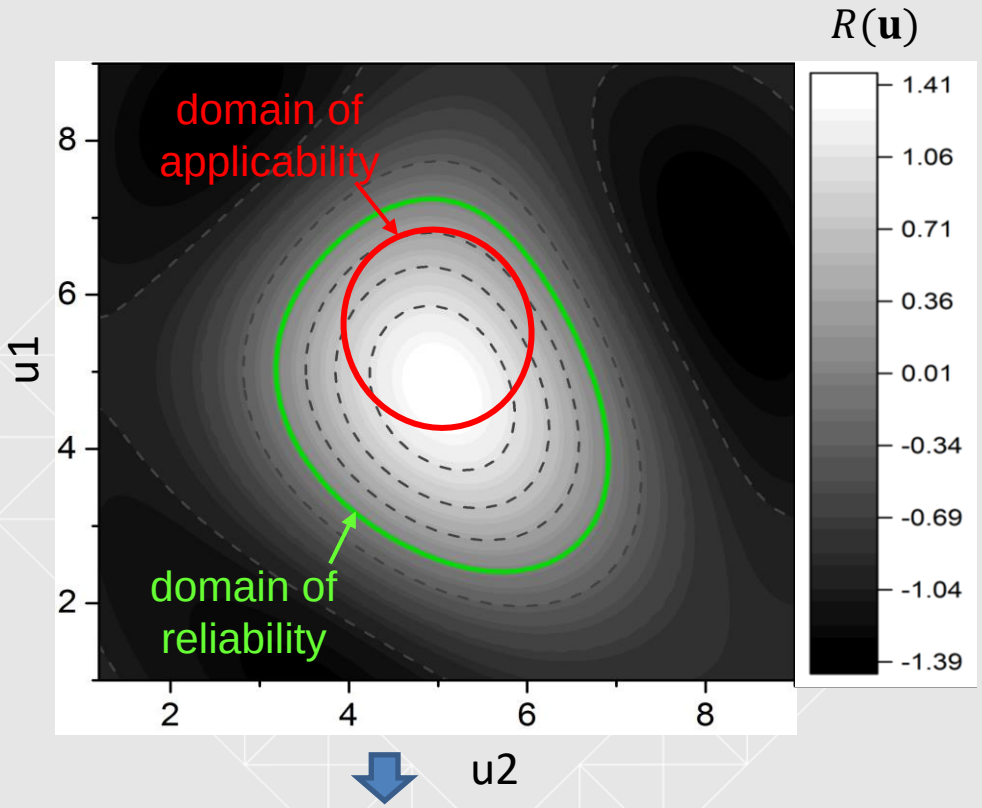
ML-Assisted Reliability Mapping

Domain of applicability VS Domain of model reliability

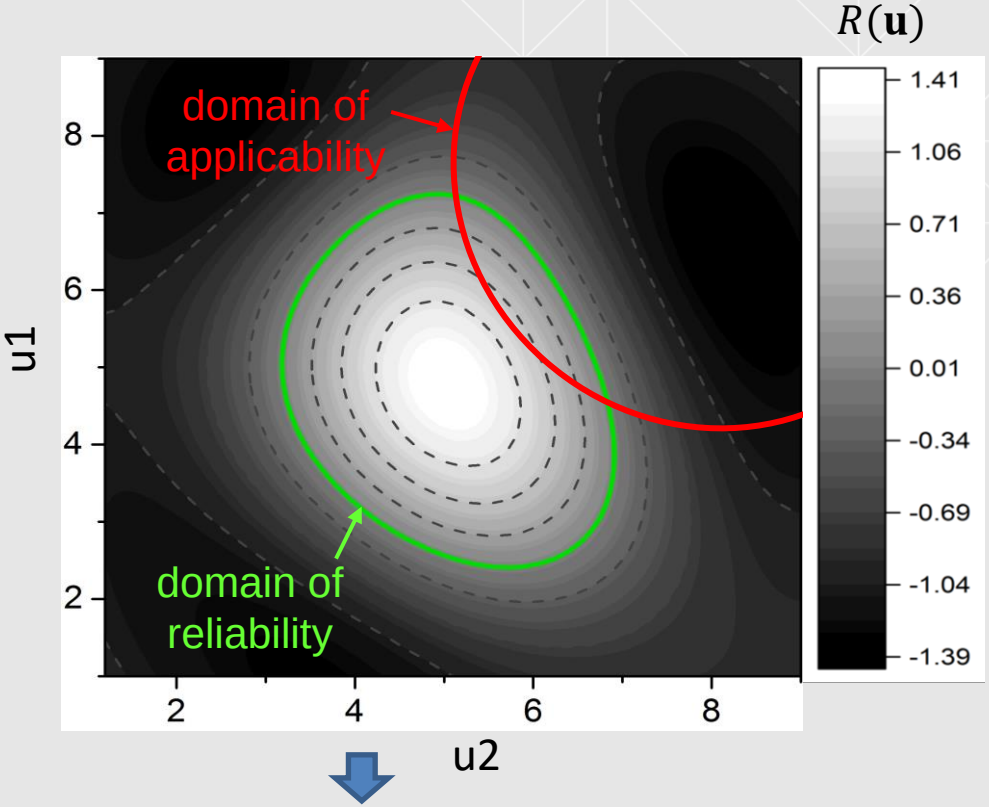
Domain of applicability Vs Domain of reliability

We need to identify **where** the model is reliable in the **design space**

This affects **all** the model-based activities



MODELLING ASSUMPTIONS ARE APPROPRIATE FOR THIS SPECIFIC APPLICATION



MODELLING HYPOTHESES MUST BE REFORMULATED

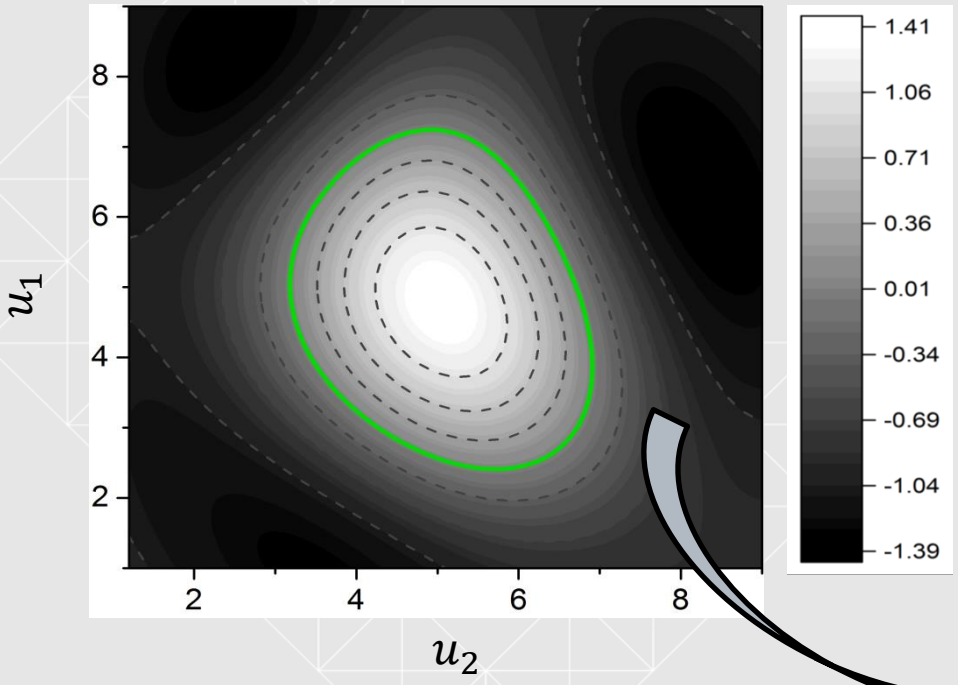
43

ML-Assisted Reliability Mapping

Constrained Model-Based Design of Experiments

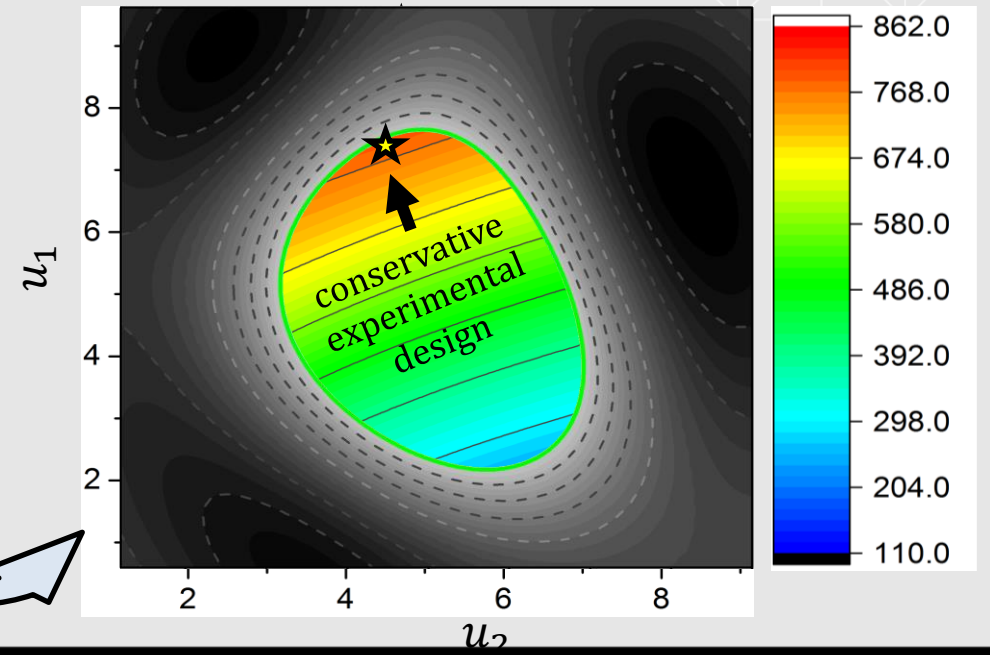
What will happen if the model is not reliable in the experimental design?
→ **Local mismatch on information prediction**
Solution: we can predict **where** the model is reliable (i.e. regions of reliability in the design space)
→ **reliability maps using Model-Based Data Mining (MBDM)**[1]

Reliability map[1]



Experimental design is constrained to conditions at $Reliability(\mathbf{u}) > 0$

Distribution of information



[1]Quaglio M., Fraga E., Cao E., Gavriilidis A. and Galvanin F. (2017). A model-based data mining approach for determining the domain of validity of approximated models. *Chemometrics and Intelligent Laboratory Systems,* 172, 58-67

ML-Assisted Reliability Mapping

Constrained MBDoE: iterating kinetic model identification procedure[1]

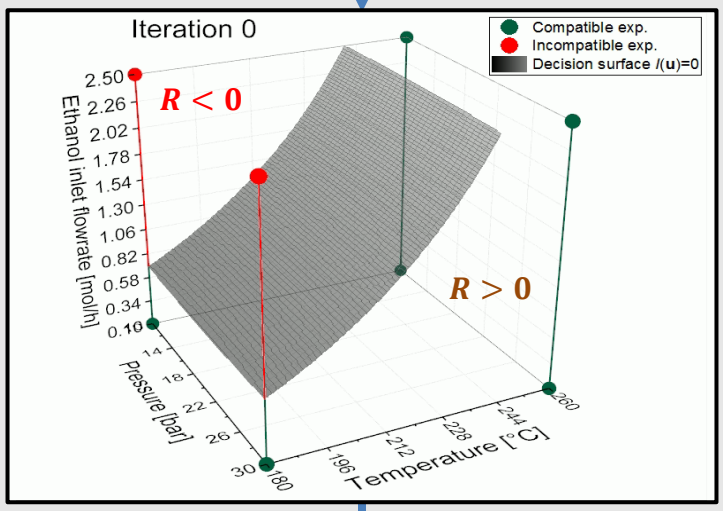
experimental data and related errors

additional experimental data

approximated model structure $\hat{y} = \mathbf{g}(\boldsymbol{\theta}, \mathbf{u})$

1. Model-Based Data Mining

2. Machine learning algorithm (SVM)



execute new experiment(s)

values for model parameters

parameter statistics?

N

Y

3. Constrained Model-Based Design of Experiments for Parameter Precision

STOP

[1]Quaglio M., Fraga E. S., Galvanin F. (2018). Constrained model-based design of experiments for the identification of approximated models, IFAC-PapersOnLine, 51(15), 515-520 .

Final remarks

- **Identification of kinetic models in Autonomous Reaction Platforms**
  - Fast identification of kinetics integrating i) **model discrimination** (MBDoE-MD); ii) **parameter precision** (MBDoE-PE); ii) **joint MBDoE-MD/MBDoE-PE**
  - Algorithms used in **online applications** and **cloud-based services (EDAS)**
  - **Online tracking of model adequacy** for mechanistic understanding
  - Techniques developed to handle practically non-identifiable («sloppy») models
  - Exploratory MBDoE based on G-optimality (**G-map eMBDoE**) now integrated
- **Kinetic model recognition using ANNs**
  - Possibility to leverage ML to identify the model structure

    → **ANN-based optimal experimental design**
  - Trade-off between noise and number of measurements to be carefully evaluated by design
- **ML-Assisted Model Reliability Mapping**
  - Systematic approach to model building using ML techniques
  - Integration in autonomous platforms

Future works

- **Identification of kinetic models in Autonomous Reaction Platforms**
  - Integration of tools for **online model diagnostics**[1]
  - Integration of **ANN-based model identification framework** in autonomous platforms, and potential combination with MBDoE
  - Integration of **generative algorithms for model identification**

    -> SINDY[2] integration
  - Integration of new MBDoE techniques including **exploratory MBDoE**[3]
  - Integration of **optimization-free methods** for online MBDoE[4]
  - Application to **large reaction networks systems**
  - **Application to flexible reaction systems (Taylor Vortex reactors)**

Wenyao Lyu
PhD student

Andrea Friso
PhD Student

[1]Quaglio, M., Fraga, E. S. & Galvanin, F. (2020), A diagnostic procedure for improving the structure of approximated kinetic models. Computers & Chemical Engineering, 133, 106659.
[2]Brunton, S. L., Proctor, J. L. & Kutz, J. N. (2016), Discovering governing equations from data by sparse identification of nonlinear dynamical systems. Proceedings of the national academy of sciences, 113, 3932.
[3]Cenci, F., Pankajakshan, A., Galvanin, F. (2023), An exploratory model-based design of experiments approach to aid parameters identification and reduce model prediction uncertainty, Computers & Chemical Engineering, 177, 108353, https://doi.org/10.1016/j.compchemeng.2023.108353
[4]Friso, A., Galvanin, F. (2023), An optimization-free Fisher information driven approach for online design of experiment, Computer-Aided Chemical Engineering, 52, 13.

Acknowledgements

Academics
Federico Galvanin (UCL)
Asterios Gavriilidis (UCL)
Eric Fraga (UCL)

Research Associates/Assistants
Panagiotis Petsagkourakis
Emmanuel Agunloye
Arun Pankajakshan

PhD Students
Marco Quaglio
Solomon Bawa
Enrico Sangoi
Conor Waldron
Francesca Cenci (UniPD)
Wenyao Lyu
Andrea Friso

MSc Students
Maerthe Theresa Tillmann

CatalysisHub

Sargent Centre for Process Systems Engineering

EPSRC
Engineering and Physical Sciences Research Council

- EP/R032807/1 “Cognitive chemical manufacturing”
- EP/V050796/1 “Fully Automated Platforms for Drug Nanocrystals Manufacturing via Continuous-Flow, Data-Driven Antisolvent Crystallization
- EP/X024016/1 “Development of a machine learning assisted digital twin platform for real-time optimization of reaction systems under uncertainty”

UNIVERSITY OF LEEDS

Thank you for listening!

Selected publications

1. Waldron, C; Pankajakshan, A, Quaglio, M, Cao, E; Galvanin, F, Gavriilidis, A; (2019) An autonomous microreactor platform for the rapid identification of kinetic models. **Reaction Chemistry and Engineering** 4, 1623. 10.1039/C8RE00345A.
2. Waldron, C; Pankajakshan, A, Quaglio, M, Cao, E; Galvanin, F, Gavriilidis, A; (2019) Model-based design of transient flow experiments for the identification of kinetic parameters. **Reaction Chemistry & Engineering**, 5, 112. 10.1039/c9re00342h.
3. Pankajakshan, A; Waldron, C; Quaglio, M; Gavriilidis, A; Galvanin, F; (2019) A Multi-Objective Optimal Experimental Design Framework for Enhancing the Efficiency of Online Model-Identification Platforms. **Engineering, 6**, 1049. 10.1016/j.eng.2019.10.003.
4. Quaglio, M; Waldron, C; Pankajakshan, A; Cao, E; Gavriilidis, A; Fraga, ES; Galvanin, F; (2019) An online reparametrisation approach for robust parameter estimation in automated model identification platforms. **Computers & Chemical Engineering** , 124, 270. 10.1016/j.compchemeng.2019.01.010.
5. Waldron, C; Pankajakshan, A; Quaglio, M; Cao, E; Galvanin, F; Gavriilidis, A; (2019) Closed-Loop Model-Based Design of Experiments for Kinetic Model Discrimination and Parameter Estimation: Benzoic Acid Esterification on a Heterogeneous Catalyst. **Industrial & Engineering Chemistry Research** , 58, 2216. 10.1021/acs.iecr.9b04089.
6. Quaglio, M.; Roberts, L.; Jaapar, M.; Dua, V.; Galvanin, F; (2020). An artificial neural network framework for classifying chemical reaction types from experimental data. **Computers and Chemical Engineering**, 135, 106759. 10.1016/j.compchemeng.2020.106759
7. Petsagkourakis, P.; Galvanin, F. (2021). Safe model-based design of experiments using Gaussian processes. **Computers & Chemical Engineering**, 151, 107339. 10.1016/j.compchemeng.2021.107339
8. Quaglio, M; Fraga, ES; Galvanin, F; (2020) A diagnostic procedure for improving the structure of approximated kinetic models. Computers & Chemical Engineering , 133 , 106659. https://doi.org/10.1016/j.compchemeng.2019.106659
9. Sangoi, E; Quaglio, M, F. Bezzo; Galvanin, F (2022). Optimal Design of Experiments Based on Artificial Neural Network Classifiers for Fast Kinetic Model Recognition. **Proceedings of the 14th International Symposium on Process Systems Engineering** – PSE 2021+ June 19-23, 2022, Kyoto, Japan. https://doi.org/10.1016/B978-0-323-85159-6.50136-6
10. Cenci, F., Bawa, S., Gavriilidis, A., Facco, P., Galvanin, F. (2023), An exploratory model-based design of experiments technique to aid parameters identification and reduce prediction uncertainty, Computer-Aided Chemical Engineering, 1-6. https://doi.org/10.1016/B978-0-443-15274-0.50001-9
11. Cenci, F., Pankajakshan, A., Galvanin, F. (2023), An exploratory model-based design of experiments approach to aid parameters identification and reduce model prediction uncertainty, Computers & Chemical Engineering, 177, 108353, https://doi.org/10.1016/j.compchemeng.2023.108353
12. Pankajakshan, A., Bawa, S., Gavriilidis, A., Galvanin, F. (2023). Autonomous kinetic model identification using optimal experimental design and retrospective data analysis: methane complete oxidation as a case study. **Reaction Chemistry & Engineering (in press)**. https://doi.org/10.1039/D3RE00156C