A Python/Numpy-based package to support model discrimination and identification

Seyed Zuhair Bolourchian Tabrizi[a,b], Elena Barbera[a], Wilson Ricardo Leal da Silva[b], Fabrizio Bezzo[a]

Università degli Studi di Padova

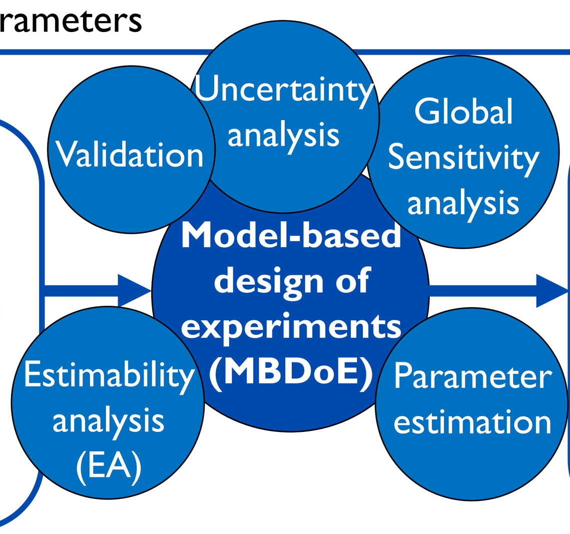
[a] Department of Industrial Engineering, University of Padova, Italy
[b] FLSmidth Cement, Green Innovation, Denmark

CAPE >_ lab UNIVERSITY OF PADOVA

1. INTRODUCTION AND MOTIVATION

- Mechanistic models are cornerstones of optimisation, control, and scale-up in PSE.
- Achieving high fidelity in these structures is challenging during the model identification step. This challenge depends on:
  - Informed selection of the most representative model
  - Precise estimation of model parameters

- Experimental data is often expensive and limited.
- Mechanistic models are often affected by:
  - structural ambiguities
  - parameter correlations



- Model-Based Design of Experiment to maximise information for:
  - Model discrimination [1]
  - Parameter precision (**MBDoE-PP**) [1]

ISSUE

- MBDoE applications have been limited and mostly in the academic sector.
- The lack of comprehensive and user-friendly packages is one reason hindering its widespread usage. No existing package (e.g. *PYOMO.DOE* [2]) offers a **comprehensive** framework that:

1. Integrates the essential techniques for model discrimination and parameter calibration;
2. Incorporates physical constraints into the design space; and
3. Remains modular and easily interfaces with external simulators.

OBJECTIVE

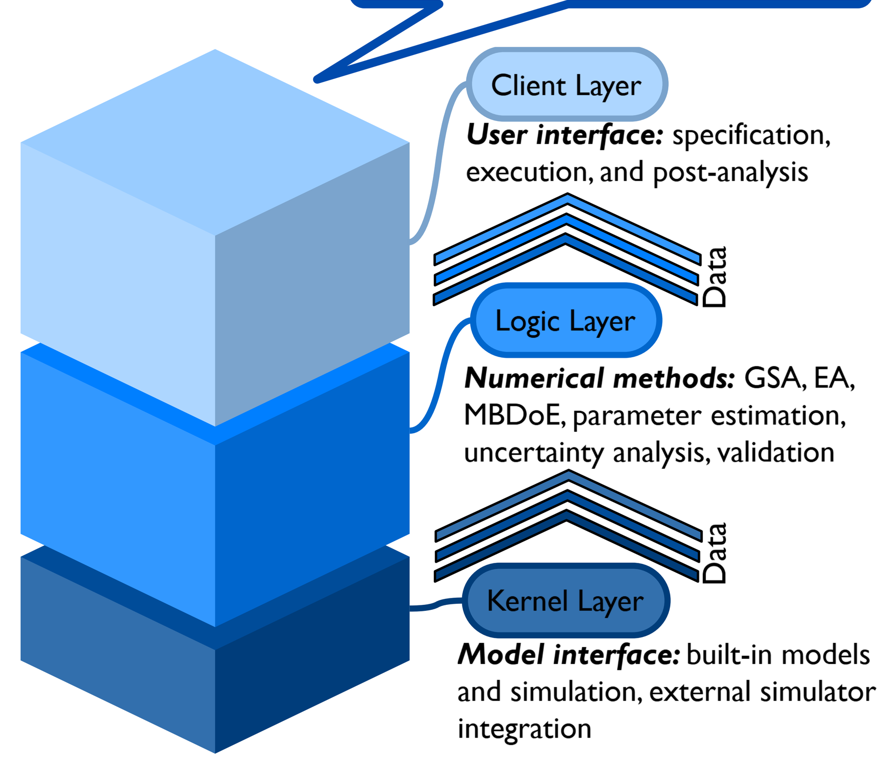
Developing ***MIDDoE*** – Model-(based) Identification, Discrimination, and Design of Experiments – as a comprehensive, user-friendly, and modular open-source tool in *Python* to bridge the gaps in MBDoE digitalisation

2. MIDDoE ARCHITECTURE

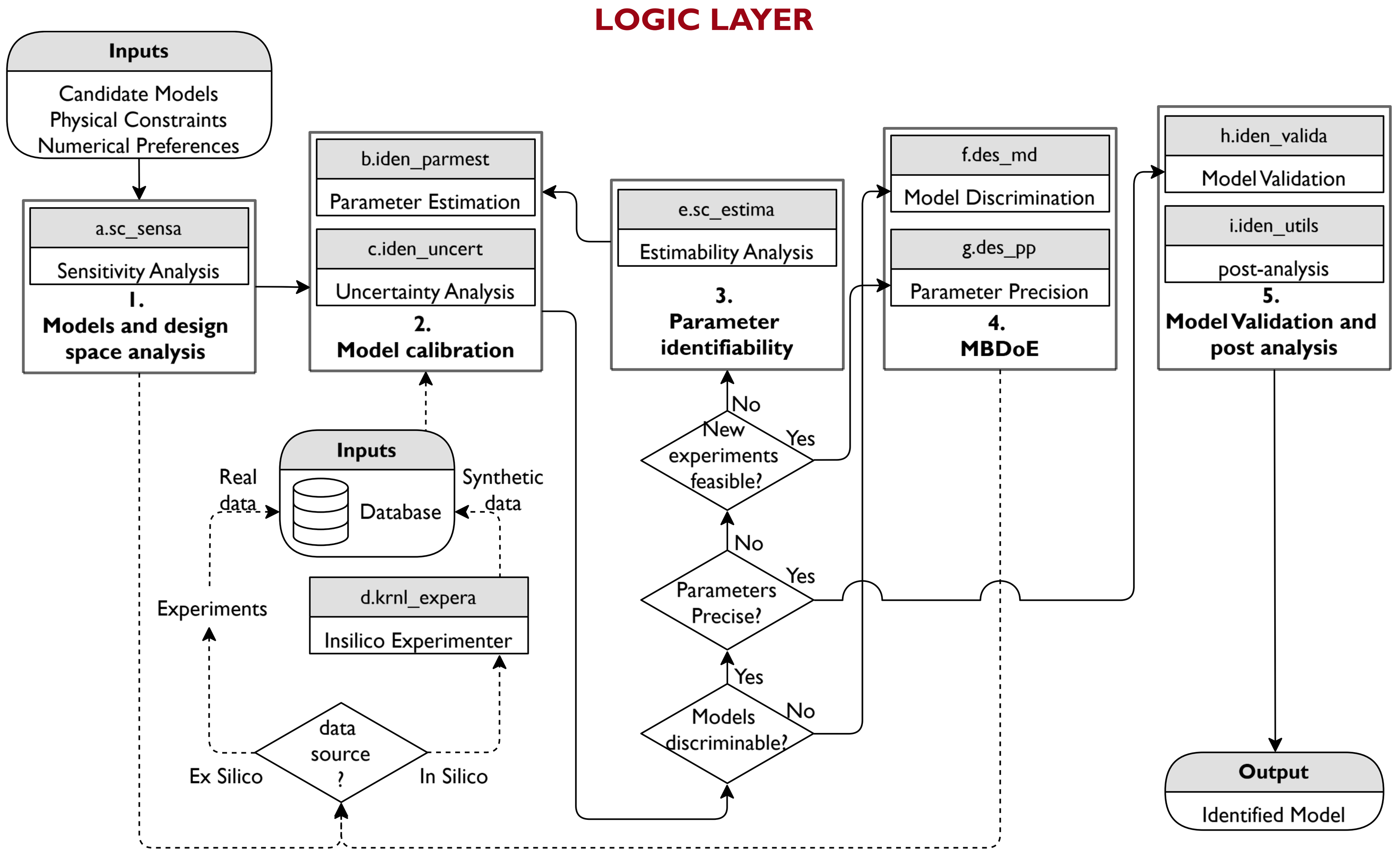
Get it: `pip install middoe`

- *MIDDoE is developed* in 3 layers
  - **Kernel layer** includes a built-in DAE simulator for stiff/non-stiff systems and custom black-box interfacing.
  - **Client layer** is developed for experimenters to run various tasks easily, excluding intensive programming.
  - **Logic layer** provides MBDoE and various numerical capabilities to support it.

The **Logic layer** includes supportive techniques for MBDoE, including **GSA-Sobol** and **Estimability Analysis (EA)** [3], used to investigate the parameter space and select unidentifiable parameters



LOGIC LAYER



- MBDoE in ***MIDDoE*** includes **model discrimination** and **parameter precision**, supporting different optimal design objectives.
  - It is capable of handling **non-convexities** in the optimisation core
  - It can embed **physical constraints** into the design space, such as:
    - Control Vector Parametrisations (**CVPs**) for linear and piecewise-constant profiles;
    - Enforced **increasing**, **decreasing**, or **relaxed** CVPs;
    - Forced **control/uncontrollable** time spans; and
    - Synchronisation of **sampling** across multiple measurements.

3. CASE STUDY

PROCESS AND SCENARIOS

- An insilico model calibration of Monod model for a semi-batch fermentation reactor in 2 different scenarios:

Sc1 **Experimental data available (2 batches); new experiments cannot be performed**

Sc2 **No data available; 4 new experiments can be designed**

FORMULATION

$$\frac{\partial y_1(t)}{\partial t} = \left(r - u_1(t) - \theta_4\right) \cdot y_1(t)$$

$$\frac{\partial y_2(t)}{\partial t} = -\frac{r \cdot y_1(t)}{\theta_3} \cdot u_1(t)\left(u_2 - y_2\right)$$

$$r = \left(\frac{\theta_1 \cdot x_2}{\theta_2 + x_2}\right) \cdot exp\left(-\frac{x_2}{\theta_5}\right)$$

DESIGN SPACE

Variable	Range	CVP	Definition
$y_{1,0}$	$1-10$	–	Biomass initial concentrations
$y_{2,0}$	$1-10$	–	Substrate initial concentrations
$\mathbf{u_1}$	$0.05-0.2$	Constant	Time-variant dilution rate
$\mathbf{u_2}$	$5-35$	Constant	Time-variant feed substrate concentration

TRUE PARAMETERS

Parameter	value	Definition
θ_1	0.31	Maximum specific growth rate
θ_2	0.11	Michaelis constant
θ_3	0.65	Yield coefficient
θ_4	0.25	Biomass loss rate
θ_5	5.00	Substrate inhibition constant

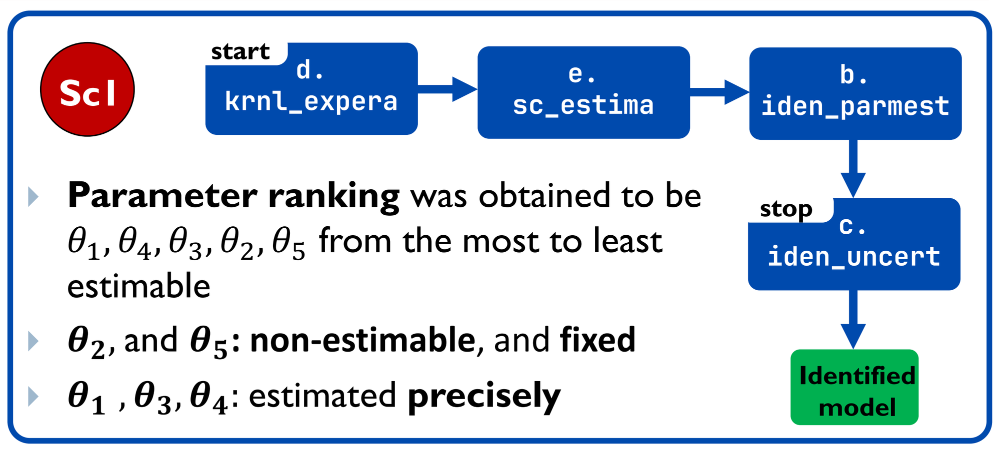
DESIGN CONSTRAINTS

- CVPs have 5 steps, and $\mathbf{u_2}$ is forced to be decreasing.
- Minimum switching/sampling point intervals are 0.5 h.
- Minimum signal perturbation for $\mathbf{u_1}$, and $\mathbf{u_2}$ are 0.01, and 5.

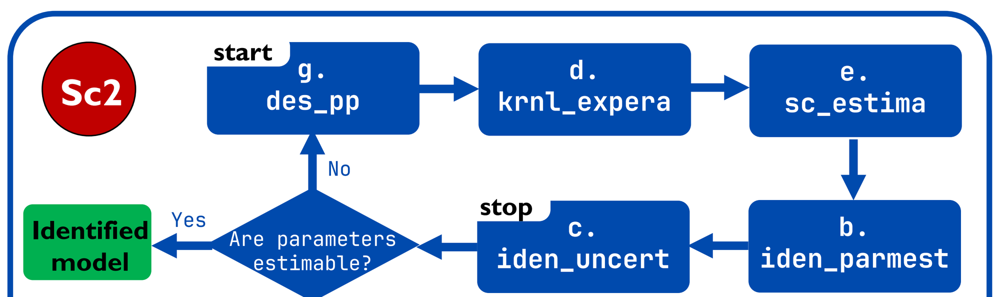
INSILICO EXPERIMENTS

- 6 sampling points per batch
- 10 h process time
- 5% normally distributed relative noise

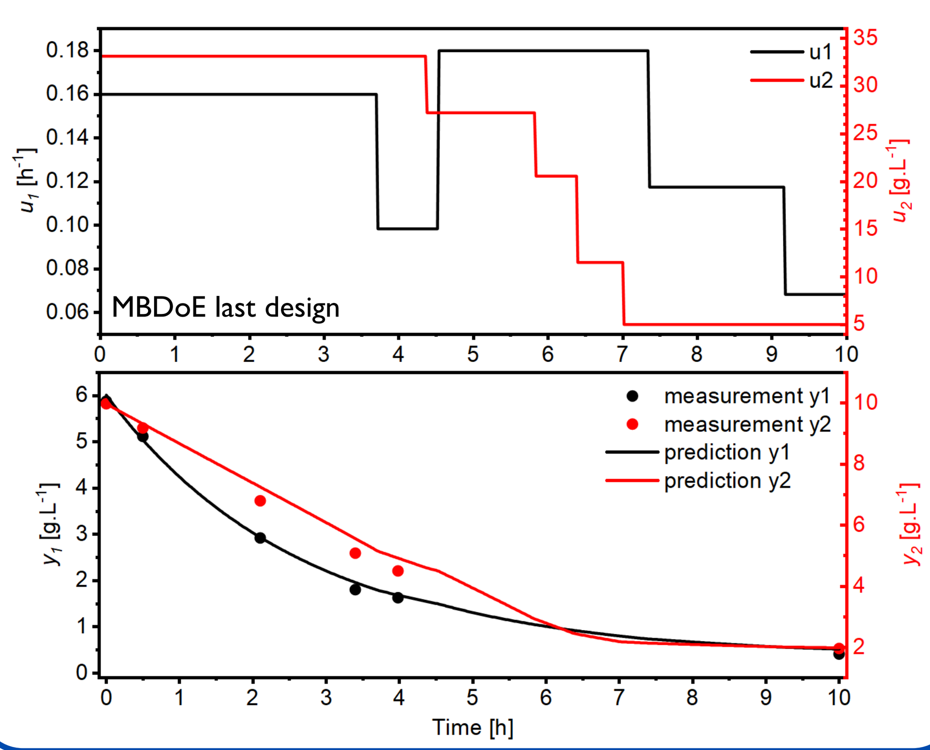
4. RESULTS

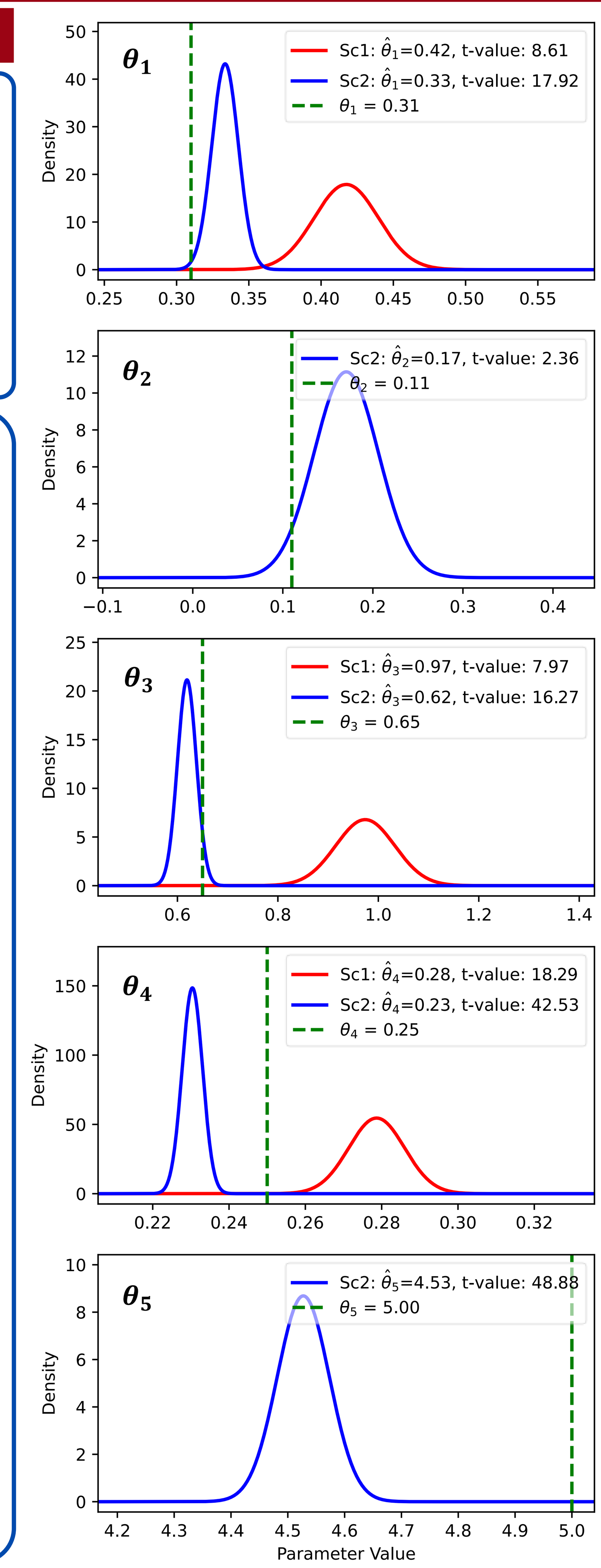


- **Parameter ranking** was obtained to be $\theta_1, \theta_4, \theta_3, \theta_2, \theta_5$ from the most to least estimable
- $\boldsymbol{\theta_2}$, and $\boldsymbol{\theta_5}$: **non-estimable**, and fixed
- $\boldsymbol{\theta_1}$, $\boldsymbol{\theta_3}, \boldsymbol{\theta_4}$: estimated **precisely**



- **4 MBDoE-PP** designs were sequentially executed.
- After each design, **EA** identified the estimable parameters.
- Following the 1st experiment, **only** $\boldsymbol{\theta_4}$ was deemed estimable.
- As new experiments were added, more parameters became estimable, and by the 4th experiment, **all** were **precisely** estimated.
- This scenario could provide more **accurate** estimations.





5. CONCLUSIONS

- ***MIDDoE*** is a comprehensive *Python* library to accelerate model identification.
- It provides essential steps to support essential numerical techniques by investigating model structure and design space.
- Modular structure enables easy interfacing with external simulators, or using built-in solvers.

6. ACKNOWLEDGEMENT

This work is part of the CO2Valorize project that has received funding from the European Union's Horizon Europe research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 101073547.

7. REFERENCES

[1] Asprey, S. P. & Macchietto, S. Statistical tools for optimal dynamic model building. Comput Chem Eng 24, (2000).
[2] Wang, J. & Dowling, A. W. Pyomo.DOE: An open-source package for model-based design of experiments in Python. AIChE Journal 68, (2022).
[3] Moshiritabrizi, I., Abdi, K., McMullen, J. P., Wyvratt, B. M. & McAuley, K. B. Parameter estimation and estimability analysis in pharmaceutical models with uncertain inputs. AIChE Journal (2023).

Contact:
Prof. Fabrizio Bezzo - E-mail: fabrizio.bezzo@unipd.it