Systems & Control Transactions

Original Research Article
Peer Reviewed Conference Proceeding
ESCAPE 36 - European Symposium on Computer Aided Process Engineering
Sheffield, UK. 21-24 June 2026

PSE PRESS

Automatic k_La determination in stirred tank reactors by model-based design of experiments

Ana Helena V. Caetano[a]*, Krist V. Gernaey[a], and Julian Kager[a]

[a] DTU, Department of Chemical and Biochemical Engineering, Kgs. Lyngby, Denmark
* Corresponding Author: ahvca@kt.dtu.dk.

ABSTRACT

The volumetric gas-liquid mass transfer coefficient (k_La) is a key performance parameter in stirred tank reactors and is commonly determined through extensive experiments across the operational space. This work presents an automatic, closed-loop framework for k_La determination based on model-based design of experiments (MBDoE), in which agitation and aeration inputs are adapted in real time.During each experiment, dissolved oxygen data is collected and used to estimate the parameters of a Van't Riet k_La relation. The parameter uncertainty is quantified using the covariance matrix, and the experiments are iteratively selected based on D-optimality or E-optimality MBDoE, until a threshold of $RSE_i < 0.15$ is reached for all parameters. The MBDoE approach is evaluated through repeated runs and compared against random designs, full factorial (FF) design, and a full grid design.The results demonstrate that the closed-loop MBDoE framework can significantly reduce the number of experiments required to characterize k_La while maintaining accurate predictions of gas-liquid mass transfer over the design space. D-optimal MBDoE converges in fewer than four experiments on average, while E-optimal MBDoE converges in less than six experiments on average, translating into approximately a 60% reduction of the number of experiments as compared to a full factorial DoE design.

Keywords: Optimization, Modelling, Numerical Methods, Model-based design of experiments, Gas-liquid mass transfer, Stirred tank reactors

INTRODUCTION

k_La identification as an experiment design problem

The volumetric gas-liquid mass transfer coefficient (k_La) is a key performance parameter of stirred tank reactors and its determination can be vital to avoid oxygen limitations and to optimize bioprocesses. In practice, the k_La can be affected by several different factors such as tank geometry, operational conditions and properties of the mixture/fluid/medium [1]. Consequently, k_La is typically determined experimentally by running multiple experiments under different combinations of operational inputs, followed by fitting an empirical correlation to the experimental data [2].

Fundamentally, this is an experimental design problem. Each experiment results in different excitation profiles of the operational inputs and consequently in different information content. Therefore, each experiment provides different information about the underlying k_La model. Because experiments in stirred tanks can be time-consuming and resource-intensive, reducing the number of experiments required to obtain a reliable k_La estimate is important for rapid bioreactor characterization and process development.

Limitations of classical design strategies

Classical k_La characterization strategies rely on predefined experimental designs, such as full factorial designs and stepwise variation of the input variables [3]. Although these approaches can give a good coverage of the design space, they do not exploit the information gained during the experiments, *i.e.*, later experiments are not adapted based on what has already been learned, which can lead to redundancy.

Model-based design of experiments (MBDoE) defines a structured way to iteratively select the input trajectories of the experiments that maximize information about the system. Along the experiments there is a

constant and iterative retrofitting of the prototype model with new data, which ensures minimal experimental redundancy and inclusion of information gathered throughout the process [4]. Thus, a closed-loop approach is established, in which experimental conditions are adapted in real time based on continuously updated parameter estimates and uncertainty information.

Contributions of this work

In this work, we present an automatic, closed-loop framework for k_La determination in stirred tank reactors using MBDoE. The system iteratively performs experiments in which agitation and aeration are varied, the experiment is run, the data of the dissolved oxygen trajectory is collected in real time, the parameters of the Van't Riet k_La correlation [5] are estimated, and the next most informative experiment to perform is computed.

The methodology is evaluated on a stirred-tank k_La case study, using D-optimal and E-optimal MBDoE strategies and it is compared against random designs, full factorial DoE designs, and a grid of 100 experiments with uniform stepwise variation of the two inputs covering the full design space (referred to as full-grid). The performance is assessed in terms of the number of experiments required to reach a target level of parameter uncertainty, as well as the accuracy of the resulting k_La predictions across the design space.

MATERIALS AND METHODS

Experimental system

The experiments were conducted in a lab-scale Biostream glass stirred tank reactor, equipped with a top mounted Rushton impeller, a dissolved oxygen (DO) probe, mass flow controller for air and nitrogen supply, and a Biostream tower. The tower is connected to the Lucullus software (Securecell, Urdorf, Switzerland), thus facilitating data acquisition and remote control. A Rest API connection is used to communicate between the software and python scripts, thus allowing for a full closed-loop automated system.

The reactor was operated in a gas-liquid system using 2L of water without biological consumption, so that oxygen transfer is governed exclusively by physical mass transfer.

The agitation speed and the gas flow rate were used as the manipulated variables. For each experiment, the DO concentration was recorded continuously, from which the experimental k_La corresponding to the operational inputs was calculated. This was computed by performing a dynamic oxygen balance model (Eq. 1) and using nonlinear regression techniques.

$$\frac{dC}{dt} = k_L a\,(C^* - C) \tag{1}$$

Where C represents the DO concentration at time t and C^* is the DO concentration at saturation.

The experiments are performed using the gassing-out method [5], where nitrogen gas is sparged through the liquid to strip it of DO before a k_La experiment.

The design space is restricted to air flow [0.4 : 4.0] L/min and stirrer speed [100 : 1000] rpm.

k_La model, parameter estimation and uncertainty

The volumetric mass transfer coefficient is described using a Van't Riet correlation [5], which relates k_La to the power input P (directly related to agitation speed) and superficial gas velocity v_g (directly related to aeration rate), and reactor geometry, where V_R is the reactor volume (Eq. 2).

$$k_L a = K\left(\frac{P}{V_R}\right)^{\alpha} v_g^{\beta} \tag{2}$$

For each experiment, the model parameters (K, α, β) were estimated by minimizing the normalized root-mean-square error (NRMSE) between the measured and simulated k_La.

Parameter uncertainty was quantified by relative standard errors (RSE), computed based on the covariance matrix (Cov) of local sensitivities. The local sensitivity matrix (S) is computed based on central differences with a relative step of 1%. Standard errors (SE) are computed to quantify the spread of the estimate around its value and scaled by the parameter value (Eq. 3 – 4).

$$SE_i = \sqrt{Cov(\hat{\theta})_{ii}} \tag{3}$$

$$RSE_i = \left|\frac{SE_i}{\hat{\theta}_i}\right| \tag{4}$$

Closed-loop MBDoE workflow

The k_La identification was performed using an iterative, closed-loop MBDoE framework (Figure 1). Each run started from a predefined initial experiment, defined at the centre of the design space. After each experiment, the collected DO data were used to re-estimate the model parameters. Based on the calibrated model, sensitivities and uncertainties are computed with respect to each of the parameters. The updated model is then used to solve the MBDoE optimization problem and suggest the next most informative experiment. For this, a genetic algorithm is used, by sampling, *in silico* and across the design space, new experiments to perform. The objective is to maximize the expected information gain of the next experiment. The resulting inputs are then applied to the reactor, generating a new DO time series, and the procedure is repeated. The iteration can be terminated when all the model parameters are below a threshold of parameter uncertainties or a threshold of the maximum allowed number of experiments performed has been reached.

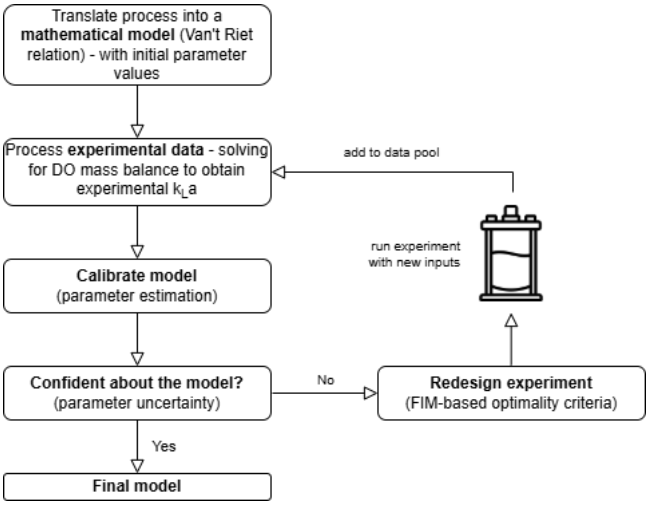

Figure 1. Closed-loop MBDoE worflow for the k_La model.

Design strategies

The information content given by the closed-loop MBDoE framework was quantified using two different optimality criteria: D-optimality and E-optimality. D-optimality or determinant optimality, aims to maximize the determinant of the Fisher Information Matrix (FIM), thereby minimizing the overall parameter uncertainty. E-optimality or eigenvalue optimality, aims to maximize the smallest eigenvalue of FIM, thus targeting a balanced identifiability across all parameters.

$$\Phi_{D-opt} = \det(\mathrm{FIM}) \tag{5}$$

$$\Phi_{E-opt} = \lambda_{min}(\mathrm{FIM}) \tag{6}$$

To benchmark the performance of the MBDoE framework, other alternative designs were evaluated, such as random uniform design, full factorial (FF) design and full grid design.

All design methods were evaluated using the same benchmark $RSE_i < 0.15$. For each method that does not rely on pre-setting the operational inputs, 10 independent runs were performed, so to account for stochastic effects. The FF DoE were performed in triplicates.

All final parameter estimates were evaluated against a full grid reference by computing the NRMSE of the predicted k_La surface over the design space.

RESULTS AND DISCUSSION

Number of experiments to converge

The primary performance metric evaluated was the number of experiments required to satisfy the stopping criterion $RSE_i < 0.15$ for all model parameters. Figure 2 summarizes the mean and standard deviation of the required number of experiments over 10 independent runs for each design strategy.

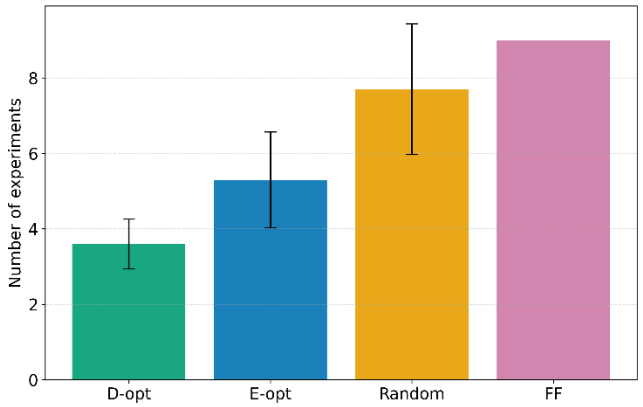

Figure 2. Number of experiments required to reach $RSE_i < 0.15$ for all parameters.

Both closed-loop MBDoE strategies substantially reduced the number of experiments necessary to reach the target precision as compared to the benchmark designs. In particular, the D-optimal MBDoE method converged in significantly fewer experiments with a mean of 3.60 ± 0.66 number of experiments. The E-optimal MBDoE strategy required 5.30 ± 1.27 experiments on average, while the random strategy required 7.70 ± 1.73 experiments to reach the same parameter uncertainty threshold. The full factorial design required 9 experiments by construction, while the full grid consisted of 100 experiments.

These results show a substantial reduction in experimental effort when using a closed-loop MBDoE design. D-optimal MBDoE reduced the number of experiments by 60% and 47% respectively, compared to a full factorial design and a random strategy. The relatively small standard deviation of the D-optimal MBDoE further indicates robustness to stochastic variation between runs.

Parameter uncertainty collapse

Figure 3 shows the evolution of the RSE of the parameters K, α, β of the Van't Riet relation as a function of the experiment number, averaged over 10 independent runs.

For the D-optimal and E-optimal MBDoE strategies, the parameter uncertainties decrease rapidly during the first few experiments. Random designs reduce uncertainty more slowly and with larger variability, reflecting the lack of adaptation to the collected data thus far.

These results show how real-time recalibration of parameter estimates pushes the MBDoE framework to select highly informative agitation and aeration inputs.

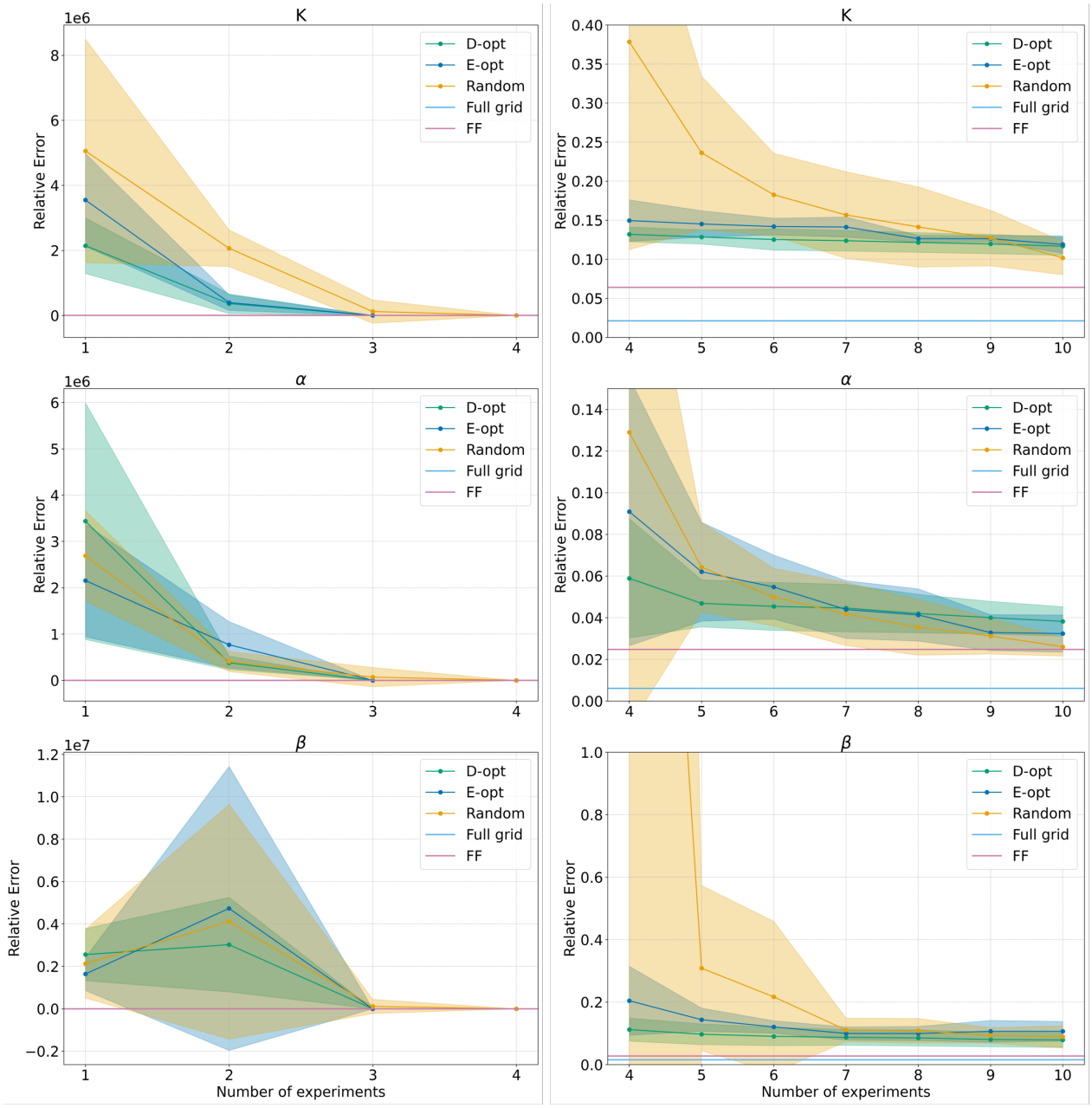

Figure 3. Evolution of the parameters RSE along the experiments for 10 independent runs.

k_La surface accuracy

Although the closed-loop algorithm terminates based on parameter uncertainty, the practical objective of the k_La characterization is an accurate prediction of this coefficient over the operational space. In order to assess this, we compare the k_La surfaces that were obtained by the different designs when they reached convergence.

For each design strategy, the final parameter estimates were averaged across the ten independent runs. This parameter set is the one used in the model (Eq. 2) to compute the k_La surface over the operational space of agitation and aeration. The resulting k_La surfaces are compared pointwise to the reference k_La surfaces by NRMSE (Figure 4).

The k_La surface references used are (a) the model fit to the full grid design of 100 experiments and (b) the model fit to the full factorial (FF) dataset. When compared against the full factorial reference, the D-optimal MBDoE strategy achieved the lowest deviation, with an NMRSE of 0.020. The E-optimal MBDoE and random strategies showed larger discrepancies, with NRMSE values of 0.092 and 0.122, respectively. This shows that the

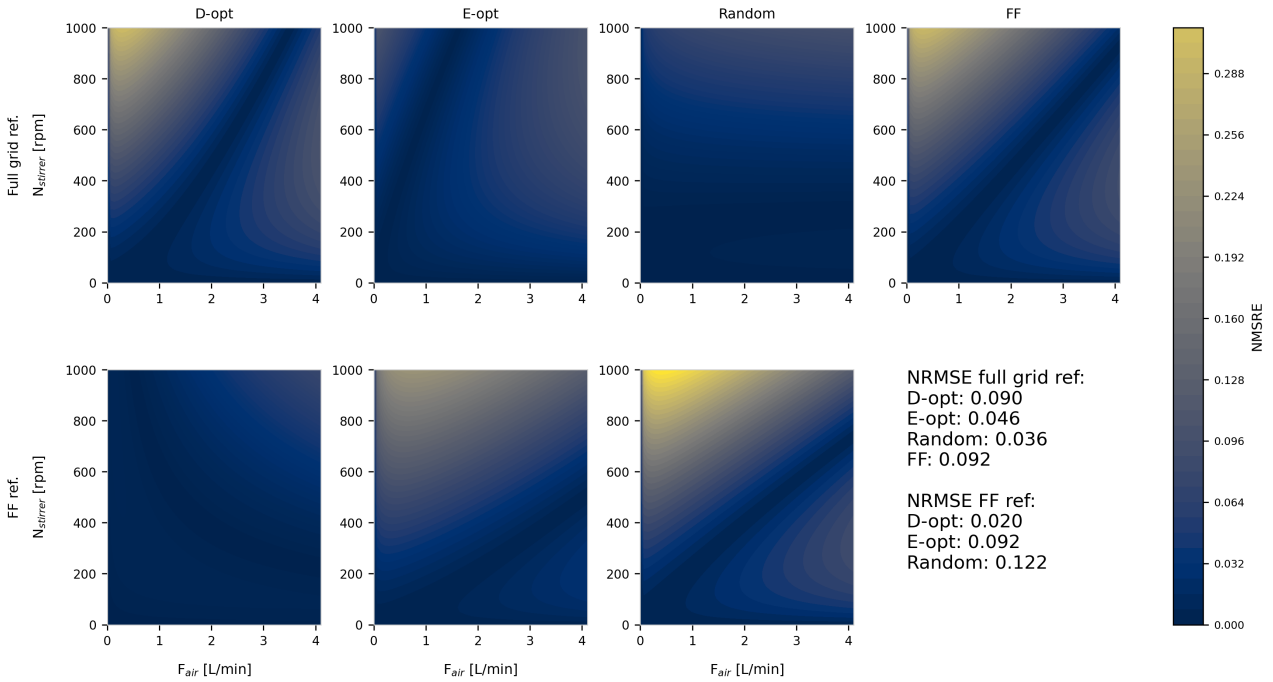

Figure 4. NRMSE of the predicted k_La surfaces in reference to the full grid design (upper row) and full factorial design (lower row). The aggregated NRMSE per design and per reference are listed on the bottom right.

closed-loop MBDoE approach achieves a similar k_La surface as the one learned by the FF dataset using substantially fewer experiments – 3.60 ± 0.66 as opposed to 9.

However, when comparing to the full grid reference, all design strategies show similar NRMSE ranges, although neither FF nor a MBDoE approach resulted in a better accuracy. Because different methods converge to different parameter estimates, these results further suggest that multiple parameter sets can produce identical k_La predictions across the design space.

Identifiability and model sloppiness

The differences observed between the two reference comparisons highlight a fundamental limitation of the Van't Riet model parametrization used in this study. Although this model provides an empirical description of k_La which is widely accepted, its parameters are structurally non-identifiable (over the explored design space)[2, 6].

This consequently results in the situation that distinct parameter vectors obtained by fitting the Van't Riet model to different experimental datasets, can generate very similar k_La surfaces. In such cases, low parameter uncertainty does not necessarily imply a better model. The accuracy of the predicted k_La surface can provide a more meaningful measure of the model quality.

Nevertheless, it is important to note that, in practical applications, the true parameter values are not known a priori and thus the objective of MBDoE is not to identify a unique "true" parameter set, but to obtain a model that can predict the k_La behaviour reliably with minimal experimental burden. Hence, the here proposed closed-loop MBDoE framework remains effective.

These results further indicate that the MBDoE stopping-criterion is better interpreted as an indication of prediction confidence rather than an indication of convergence to unique parameter vectors for a Van't Riet relation.

CONCLUSIONS

This work presented an automatic, closed-loop MBDoE framework for k_La determination in stirred tank reactors. By iteratively recalibrating the parameter estimates and uncertainty information from time-series data of dissolved oxygen, this approach autonomously selects agitation and aeration inputs that maximize information gain.

The results demonstrated that closed-loop MBDoE can substantially reduce the number of experiments necessary to reach a confidence threshold as compared to other design strategies, namely full factorial DoE, full grid over the operational space, and random inputs selection. In particular, a D-optimal MBDoE approach achieves the set uncertainty threshold of $RSE_i < 0.15$ with a 60% reduction in the required number of experiments as compared to a full factorial design.

Furthermore, comparison of the predicted k_La surfaces against full factorial and full grid designs showed

that, although the parameter values obtained differ between design strategies, the resulting k_La predictions over the design space are often comparable. This reflects the lack of identifiability of the Van't Riet relation and further suggests that the uncertainty-based stopping criterion indicates mostly a convergence of predictive behaviour rather than convergence to unique model parameters values.

Future and ongoing work shall extend the proposed framework to include additional operational variables, such as temperature, pressure, ionic strength, and viscosity, all important factors to gas-liquid mass transfer [1].

The methodology here suggested provides a foundation for autonomous reactor characterization, in particular for situations where experimental efficiency is of the most importance.

ACKNOWLEDGEMENTS

Financial support was provided by the Novo Nordisk Foundation Start Package grant for Bioprocess Control and Digitalization (NNF220C0081250).

AUTHOR IDENTIFIERS

Author ORCIDs:
Caetano AHV: 0009-0008-2951-4159
Gernaey KV: 0000-0002-0364-1773
Kager J: 0000-0001-7274-6678

REFERENCES

1. Garcia-Ochoa F, Gomez E. Bioreactor scale-up and oxygen transfer rate in microbial processes: an overview. Biotechnology Advances 27:153-176 (2009). https://doi.org/10.1016/j.biotechadv.2008.10.006
2. Schaepe S, Kuprijanov A, Sieblist C, Jenzsch M, Simutis R, Lübbert A. Kla of stirred tank bioreactors revisited. Journal of Biotechnology 168:576-583 (2013). https://doi.org/10.1016/j.jbiotec.2013.08.032
3. Aroniada M, Maina S, Koutinas A, Kookos IK. Estimation of volumetric mass transfer coefficient (kla)—review of classical approaches and contribution of a novel methodology. Biochemical Engineering Journal 155:107458 (2020). https://doi.org/10.1016/j.bej.2019.107458
4. Franceschini G, Macchietto S. Model-based design of experiments for parameter precision: state of the art. Chemical Engineering Science 63:4846-4872 (2008). https://doi.org/10.1016/j.ces.2007.11.034
5. Van't Riet, K. Review of Measuring Methods and Results in Nonviscous Gas-Liquid Mass Transfer in Stirred Vessels Introduction Stirred vessels are frequently employed to achieve a. *Ind. Eng. Chem. Process Des. Dev* 18, 357 (1979).
6. Heinrich M, Arutjunjan R, Timmer J. On the different flavours of practical identifiability. Current Opinion in Systems Biology 42:100556 (2025). https://doi.org/10.1016/j.coisb.2025.100556