Reaction Chemistry & Engineering

ROYAL SOCIETY OF CHEMISTRY

TUTORIAL ACCOUNT

Published on 03 September 2026
Licensed under CC-BY 4.0

Check for updates

Cite this: DOI: 10.1039/d6re00247a

Kinetic modeling and optimization of complex enzymatic reactions: a practical guide for chemists and biotechnologists

Maerthe Theresa Tillmann, [bc] Manuel Ibañez, [a] Kayla Reata Dittmer, [d] Jixia Lu, [ac] Gustavo Rodriguez Manotas, [c] Daniel Ohde, [d] Andreas Liese, [d] Jakob Burger, [a] Mirko Skiborowski [c] and Leandros Paschalidis *[ac]

Designing complex enzymatic reaction processes that involve multiple reaction steps, multiple phases, or different catalytic mechanisms is often challenging. This tutorial account provides a structured, accessible overview of modeling and optimization methods for such process designs. It covers fundamental concepts of modeling different types of enzymatic reaction processes, as well as methods for parameter estimation, simulation, and model validation. Building on this foundation, the article introduces parametric studies, approximations, and model-based optimization techniques for analyzing and optimizing these processes. Each methodological step is illustrated through a consistent application example and supported by openly available Python implementations. The article targets experimental researchers in chemistry and biotechnology, and provides practical guidance for developing, analyzing, and applying kinetic models. By integrating theoretical foundations with practical tools and workflows, this work aims to facilitate broader adoption of model-based insights and optimization in early-stage development of enzymatic reaction processes.

Received 17th June 2026,
Accepted 30th August 2026

DOI: 10.1039/d6re00247a

rsc.li/reaction-engineering

1 Introduction

Enzymatic reaction processes have been studied for more than a century.[1] A foundational milestone was the work of Leonor Michaelis and Maud Menten in 1913, who introduced a quantitative framework for describing enzyme kinetics.[2] Since then, significant progress has been made in nearly all aspects of enzyme-catalyzed processes, ranging from enzyme purification and immobilization[3–10] to advanced methods for enzyme engineering.[3,4,11–13] Particularly noteworthy recent advances include directed evolution methods for enzyme engineering and artificial intelligence approaches for studying enzyme structure, both of which have been recognized with Nobel Prizes in Chemistry.[14,15]

Enzymes are employed in a wide variety of applications, including the production of fine chemicals, pharmaceuticals, and bio-based materials.[3,4] However, a persistent limitation in their use as industrial catalysts is their relatively rapid deactivation and their sensitivity to conditions deviating from those required for optimal catalytic activity. Deactivation can be caused by elevated temperature, extreme pH values, or numerous other factors, and may be reversible or irreversible.[16–19] This rapid deactivation makes careful design of enzymatic processes essential. Even in cases where these limitations can be overcome or partially mitigated (*e.g.*, *via* immobilization or enzyme engineering), the design of economically viable enzymatic reaction processes remains challenging. In reaction systems involving more than one enzyme, the relative enzyme ratios and concentrations must be carefully chosen. In processes that require cofactors, both the cofactor concentration and the cofactor regeneration rate must be designed to enable the reaction process to reach high conversions without becoming rate-limiting. In multiphase systems, mass-transfer limitations or phase equilibria may significantly affect reaction performance, requiring careful selection of operating conditions.[20]

Complex reaction processes involving enzymes are often limited by kinetics or mass transfer rather than thermodynamic equilibrium.[1,21–26] Enzyme-catalyzed reactions frequently follow Michaelis–Menten-type kinetics with multiple substrates; the associated reaction networks, particularly those involving cofactor regeneration, can be highly complex. These systems often behave strongly

[a] *Laboratory of Chemical Process Engineering, Technical University of Munich, Campus Straubing for Biotechnology and Sustainability, Uferstrasse 53, 94315 Straubing, Germany*
[b] *Ehrfeld Mikrotechnik GmbH, Mikroforum Ring 1, 55234 Wendelsheim, Germany*
[c] *Institute of Process Systems Engineering, Hamburg University of Technology, Am Schwarzenberg-Campus 4, 21073 Hamburg, Germany.*
E-mail: leandros.paschalidis@tuhh.de
[d] *Institute of Technical Biocatalysis, Hamburg University of Technology, Denickestr. 15, 21073 Hamburg, Germany*

nonlinear and are difficult to design by intuition alone; process developers therefore benefit from modeling and model-based optimization. Careful selection of initial conditions often determines not only whether the reaction process will run optimally, but whether it will run at all.[20]

Alongside advances in biotechnology, the process systems engineering community has made significant progress in modeling and optimizing reaction processes.[27–31] Several studies have been published in recent years applying modeling and model-based optimization methods to enzymatic reaction systems.[32–37] However, these studies are mostly conducted by specialists in modeling and optimization and are often focused on specific applications. Application of the methods in experimental settings is often hindered by significant communication barriers between modeling experts and experimental experts working on biotechnological applications. Several open-source and academic tools for kinetic modeling, parameter estimation, sensitivity analysis, model-based experimental design and dynamic optimization are already available, such as COPASI,[38] CasADi,[39] MOPEDS[40] and Pyomo.DAE.[41] However, their uptake in the experimental research community remains limited, particularly among researchers primarily focused on enzymatic reaction systems and laboratory-based studies. Especially two main gaps persist in this direction that this tutorial aims to address. First, there is no comprehensive collection of examples of established modeling and optimization methods that illustrates how these can be effectively applied to the design of complex enzymatic reactions, accompanied by ready-to-run open-source implementations. Second, accessible explanations and practical tools to help experimentalists apply modeling and model-based optimization, or critically engage with modeling specialists, remain scarce.

This tutorial account collects modeling and computational methods that can be used to design enzymatic reaction processes. The discussion begins with some preliminary theoretical background and proceeds through modeling methods for different systems, model parameterization and validation, before turning to how models can be used to extract process insights and support model-based design and optimization. Each section is accompanied by an application example implemented in Python. The article concludes with recently published case studies that combine experimental and modeling investigations.

The article is aimed at chemists and biotechnologists seeking to apply modeling and optimization to their enzymatic reaction systems. It is written to be accessible to this audience, although some prior knowledge typically acquired at the Master's level is assumed (*i.e.* especially physical chemistry, calculus, differential equations and scientific computing). Considering that complete textbooks are devoted to parameter estimation[42,43] and model-based optimization[44] the current tutorial article focuses on a number of applications of practical relevance, while references and suggestions for additional in-depth reading are provided where appropriate.

2 Preliminaries

2.1 The inherent complexity of reaction processes

Reaction processes vary and are often very complex.[45] They can occur in the gas, liquid, or solid phase and may be catalyzed by homogeneous or heterogeneous catalysts. They may also be chemical, biochemical, electrochemical, or of another type. They can be extremely fast or extremely slow, with some reaching equilibrium in seconds while others require hours, days or even more time.[46] Enzymes are most often either dissolved in an aqueous solution or immobilized on a support (or within the pores of a support).[47,48] These reaction processes are often relatively slow in approaching equilibrium, and thus studying their kinetics is most relevant for their design.

As shown in Fig. 1, the phenomena involved in reaction processes are described across several physical domains, including reaction kinetics, thermodynamics, transport phenomena, transition state theory, and electronic structure theory (among others). Because these domains are interconnected, reaction processes are often difficult to decipher: the observed behavior is typically determined by their interplay rather than by any single factor. Naturally, a kinetic model cannot include all of this information, but is instead a simplification aimed at expressing reaction rates as functions of state variables.

In multiphase reaction systems, the overall behavior is often governed by the interplay of phase equilibrium, interphase mass transfer, and reaction kinetics. While gas–liquid equilibrium can provide important thermodynamic constraints, its direct use implicitly assumes that interphase mass transfer is sufficiently fast relative to the reaction rates. If this assumption is not valid, mass-transfer limitations must be modeled explicitly. In heterogeneously catalyzed systems, the coupling between fluid dynamics (and associated transport phenomena) and reaction kinetics often plays a key role. In electrochemical systems, the interplay between surface phenomena (such as electrode kinetics and interfacial charge transfer) and mass transfer is often dominant.[49] It is therefore important for researchers working on a given type of reaction process to be familiar with the intricacies of their system. Experimentalists are typically very familiar with their own reaction processes and their complexities, yet they often lack the language or tools to model and describe them in a formal way.

Note: aim

The focus of this tutorial account is to inform and instruct chemists and biotechnologists who are very familiar with the intricacies of their systems but lack the modeling and model-based optimization skills needed to study them computationally.

This journal is © The Royal Society of Chemistry 2026

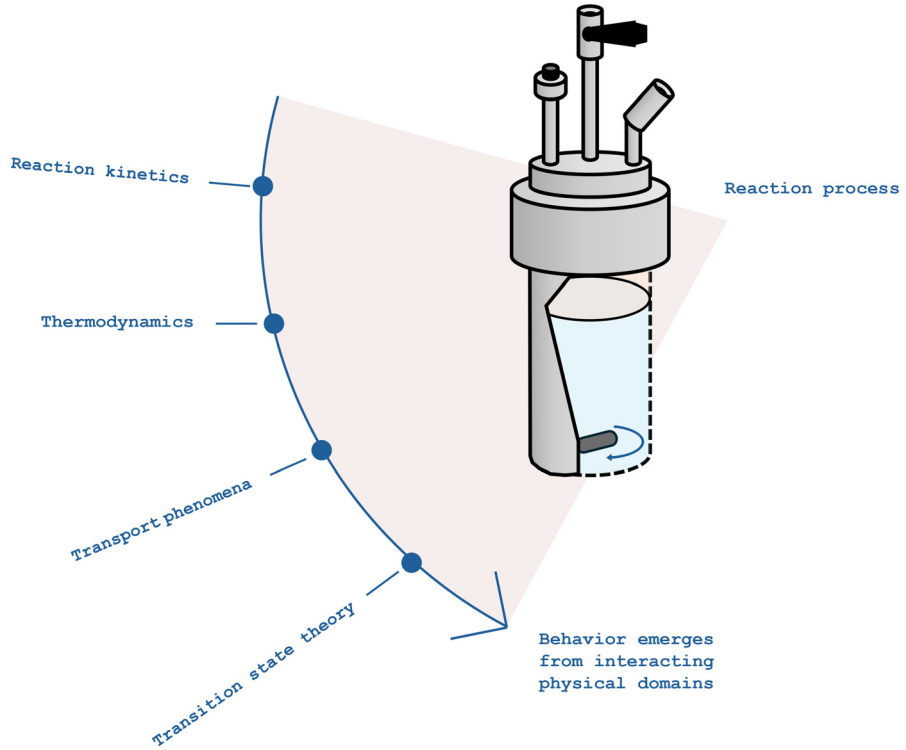


Fig. 1 The behavior of a reaction process emerges from multiple interacting physical domains. Adapted from Dittmer *et al.*[50] under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) license.

2.2 Enzyme kinetics

A substantial body of theory exists for describing enzymatic reaction kinetics.[1] The standard starting point is a mechanistic reaction scheme, from which a system of mass-action differential equations is derived for all species, including short-lived intermediates. Under suitable assumptions, most commonly the quasi-steady-state approximation, this high-dimensional dynamical system can be reduced to a single effective rate law. This reduction yields the well-known family of Michaelis–Menten-type models.[21] These models are widely used because their parameters can be estimated from experimental data (assuming sufficient experimental data exist to make them identifiable) and they remain interpretable in terms of binding and catalytic rates. Variants exist for one-, two-, and multi-substrate reactions. In many practical settings, however, the true molecular mechanism is either partially unknown or too complex to resolve, and empirical alternatives such as Hill kinetics are also commonly employed.[51]

Note: in practice

In practice, the detailed mechanism of an enzyme-catalyzed reaction is often unknown or cannot be uniquely identified from available data. In such cases, it is advisable to select the simplest model that adequately describes the observed dynamics and incorporates any known mechanistic features, such as inhibition or saturation effects.[52]

Within this framework, the Michaelis–Menten model provides the simplest description of enzyme-catalyzed conversion of substrate into product. It is based on the reaction scheme

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P. \tag{1}$$

The key simplification is the quasi-steady-state approximation, which assumes that the intermediate complex ES rapidly reaches a steady state. This leads to the classical rate law

$$r = \frac{V_{\max} S}{K_M + S}, \tag{2}$$

where $V_{\max} = k_2 E_0$ and $K_M = (k_{-1} + k_2)/k_1$. This expression naturally captures enzymatic saturation: at low substrate concentrations the reaction is approximately first-order in S, whereas at high concentrations the rate approaches the limiting value $V_{\max}$.

More complex enzymatic systems can be described by extending the same mechanistic principles. For example, in reactions involving two substrates, an ordered sequential mechanism is often assumed:

$$E + A \rightleftharpoons EA, \quad EA + B \rightleftharpoons EAB \rightarrow E + P. \tag{3}$$

Under simplifying assumptions, this yields an effective rate law of the form

$$r = \frac{V_{max} S_A S_B}{(K_A + S_A)(K_B + S_B)}. \quad (4)$$

This expression highlights that catalytic turnover requires both substrates to be sufficiently available: depletion of either A or B significantly reduces the reaction rate. In the limit of high substrate concentrations, the rate again approaches V_{max}, consistent with enzyme saturation.

A further important modification arises in the presence of inhibitors. For example, in competitive inhibition, the inhibitor I competes directly with the substrate for binding to the free enzyme:

$$E + I \rightleftharpoons EI. \quad (5)$$

This leads to the modified rate law

$$r = \frac{V_{max} S}{K_M\left(1 + \frac{I}{K_I}\right) + S}. \quad (6)$$

Here, the primary effect of inhibition is an increase in the apparent Michaelis constant, reflecting reduced effective substrate binding affinity. Consequently, higher substrate concentrations are required to achieve the same reaction velocity, while the maximal rate V_{max} remains unchanged in the limit of saturating substrate.[51]

2.3 Modeling philosophy for kinetically-limited reaction processes

A central part of modeling kinetically limited reaction processes is identifying the chemical reaction network. This is often done pragmatically, by determining which chemical species are detectable using the available analytical methods. The reaction network considered in the model must contain all relevant reactions, in the sense that it includes all reactions needed to describe the reaction process under consideration. This will mostly correspond to a simplified form of the actual

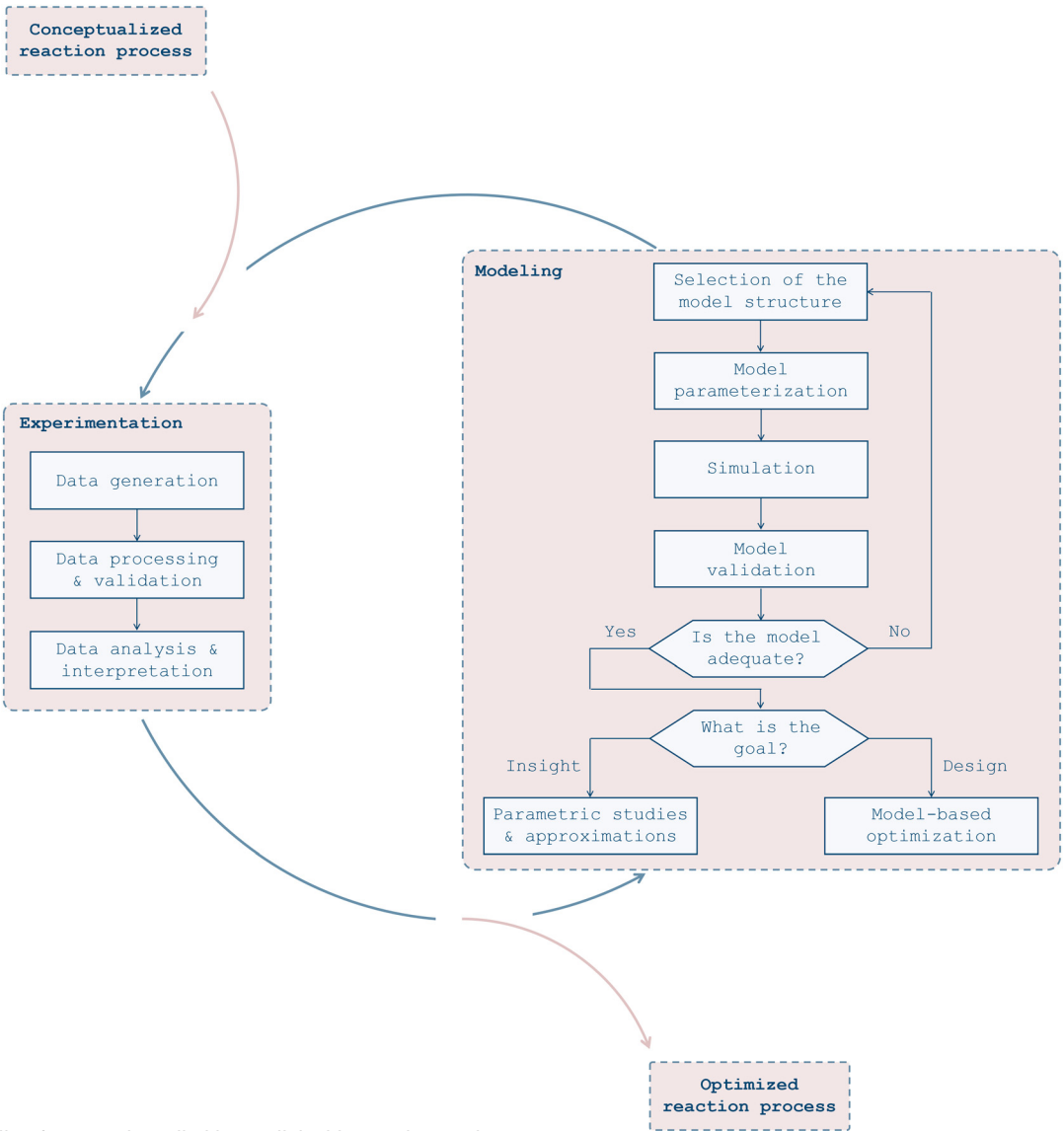


Fig. 2 A modeling framework applied in parallel with experimentation.

This journal is © The Royal Society of Chemistry 2026

chemical reaction network for two reasons: first, some reactions are omitted because they are irrelevant or unknown; and second, even the reactions that are considered are often treated in a simplified manner. For example, not all individual elementary steps are included, but are instead simplified using approaches such as the quasi-steady-state approximation, as in Michaelis–Menten-type models, or through other simplifications. Such simplifications are often necessary to create models with identifiable parameters, particularly when the available experimental data do not contain sufficient sensitivity to distinguish between all mechanistic contributions, reliably estimate strongly correlated parameters, or support the level of detail represented in the full model.

In addition, determining which further physical phenomena must be considered is of great importance, such as accounting for mass-transfer steps, treating certain parts of the process as being at equilibrium, occurring at a faster time scale than the rest of the process, or considering enzyme deactivation.

The general rule is that it is not possible to account for every single detail occurring within a reaction process when developing a kinetic model. What is required, however, is to account for every relevant effect, relevant in the sense that it influences the result we aim to achieve. For example, if the goal is to maximize space–time yield (also referred to as productivity) by varying reaction conditions A, B, and C, a good model should be able to represent the effect of conditions A, B, and C on the space–time yield with sufficient accuracy. It is less important that the model captures the effect of an unimportant condition (*e.g.*, condition D) on the space–time yield.

At the same time, the level of model detail must also be supported by the available experimental data. Even if a mechanistic effect is potentially relevant, it can only be estimated reliably if the measured data contain sufficient information and sensitivity to observe its contribution. For example, product inhibition can only be identified if product concentrations vary sufficiently and the resulting decrease in reaction rate can be distinguished from other effects, such as substrate depletion, enzyme deactivation, or mass-transfer limitations. Similarly, competitive substrate inhibition, enzyme inactivation, or reversible reaction steps can only be included meaningfully if the experimental design generates conditions under which these effects lead to observable changes in the measured response. If different mechanisms or parameters lead to indistinguishable model responses under the available experimental conditions, the corresponding parameters may become non-identifiable or strongly correlated.

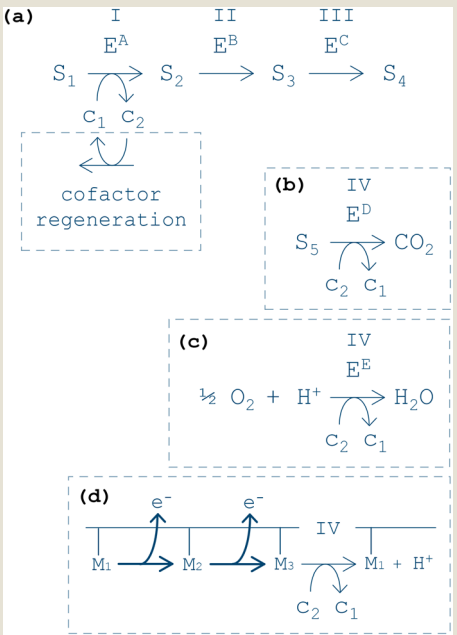
2.4 Application example

To illustrate the relevant material, a synthetic application example will be presented alongside the discussion. This differs from the case studies at the end of the article, which primarily focus on already published cases combined with experimental studies. The purpose of the synthetic example is to help the reader understand the methods more clearly and provide them with example code in Python.

The article is structured such that it follows a practical and applicable flow any experimentalist could apply in early-stage development of their enzymatic reaction process. Fig. 2 illustrates an iterative framework for the development and optimization of reaction processes through the integration of experimentation and process modeling. Experimental activities, including data generation, processing, validation, and interpretation, provide the basis for model development and refinement. In parallel, the modeling workflow involves model structure selection, parameterization, simulation, and validation, followed by an assessment of model adequacy. Once a validated model is obtained, it can be used either for parametric studies and process understanding or for model-based optimization. The continuous interaction between experimental and modeling activities ultimately enables the systematic optimization of the reaction process.

Application example: definition

The example relies on the following multi-enzymatic reaction network, which starts from the raw material S_1 and produces the main product S_4. It is to be carried out in a batch reactor. As the first reaction step requires a cofactor that cannot be added in stoichiometric amounts, an additional reaction is included for cofactor regeneration. Three different regeneration methods are shown: one producing CO_2 (b), one reducing oxygen (c), and one electrochemical method using mediator M (d).



3 Experimentation

Experimentation is the basis from which models are formed. Therefore, experimental data should comply with the principles of good scientific practice to ensure model validity. This includes reproducible experimental protocols, well-documented operating conditions, quantified measurement uncertainty, replicate

experiments, appropriate calibration procedures, mass-balance closure and the use of independent data for model validation.

As seen in Fig. 2, experimental investigation begins with the conceptualization of the reaction process. This stage includes several key steps:

- reviewing the literature on the reaction,
- formulating a hypothesis,
- defining clear objectives,
- identifying key variables and parameters (dependent, independent, controlled),
- developing an experimental plan,
- designing and implementing the experimental setup within the constraints of the available infrastructure and equipment,
- and developing the analytical method, including considerations of limits of detection, limits of quantitation, and other methodological constraints.[53]

Note: the importance of the experimental plan

Developing an effective experimental plan is essential in model-based studies because the quality of the data determines how well model parameters can be estimated and competing mechanistic hypotheses can be evaluated. Experimental conditions should be chosen to maximize information content, supporting parameter estimation, identifiability analysis, and discrimination between alternative kinetic models. Varying factors such as temperature, reactant concentrations, pH, or residence time can improve parameter sensitivity, reduce uncertainty, and enhance the validity and predictive power of the resulting kinetic model.[32,54–56]

Once the experimental setup has been established and preliminary trials have been completed, data suitable for modeling can be generated, processed, validated, analyzed, and interpreted.

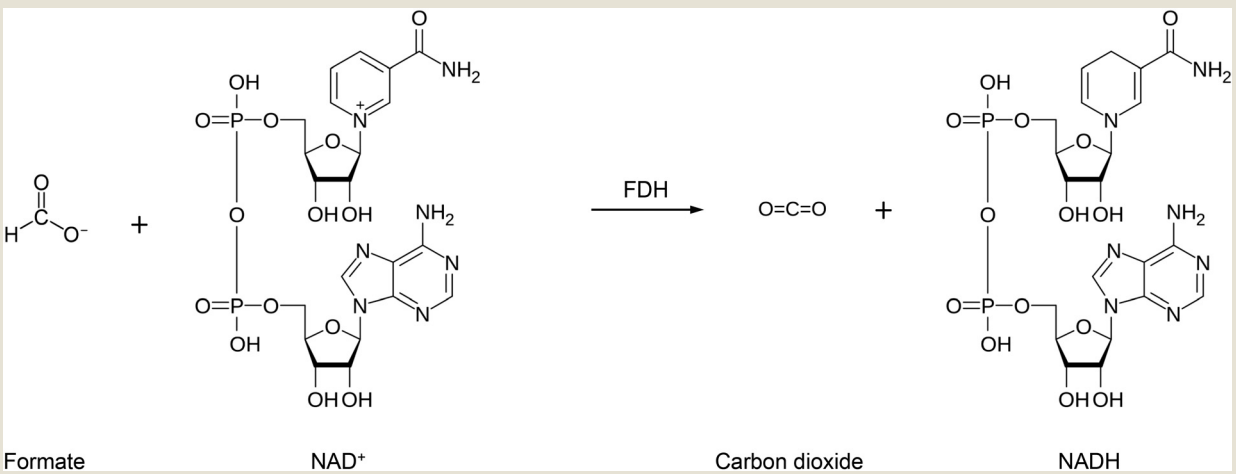
3.1 Data generation

Data generation comprises the systematic acquisition of reproducible experimental results under controlled conditions. The objective is to adequately capture the behavior of the enzymatic reaction in a reliable and quantitative manner. The parameter space (*e.g.*, pH, temperature, enzyme loading, substrate and cofactor concentration, buffer concentration, or the selected solvent) should be designed to capture all mechanistic behavior the model seeks to represent. Ranges relevant to the intended application are necessary but may not be sufficient for this purpose.

To ensure reproducibility and data reliability, experiments must be conducted under well-controlled and consistent conditions. Analytical instruments should be properly calibrated prior to measurement to guarantee sufficient accuracy, accounting for the reaction system, intended model use, and the signal-to-noise ratio. Each experiment should be performed in replicates (ideally a minimum of three) to enable statistical evaluation of standard deviations. In addition, appropriate control experiments must be included to distinguish enzymatic activity from non-enzymatic contributions. Data acquisition should be carried out with

Note: from chemical equations to kinetic model abstractions

The reactions shown in the application example are intentionally simplified to introduce the notation used throughout this tutorial. Using conventional chemical notation, reaction (b) of the application example would be represented as shown below.



NAD stands for nicotinamide adenine dinucleotide and FDH stands for formate dehydrogenase. In the application example, formate is represented with S_5 while the two cofactor forms are represented with c_1 and c_2. From the perspective of a chemist, this abstract notation can be interpreted as follows: each state, S_i or c_i, represents a collection of molecular species with associated concentrations. Molecules are transformed from one state to another through kinetic processes, such as chemical reactions, mass transfer, or other physicochemical phenomena. This abstraction allows a wide range of processes to be described within a common mathematical framework, independent of the specific underlying mechanism.

This journal is © The Royal Society of Chemistry 2026

sufficient temporal and analytical resolution to capture the relevant system dynamics. Finally, comprehensive metadata, including detailed experimental conditions, batch information, and equipment specifications, must be documented to ensure traceability and facilitate reproducibility.

3.2 Data processing and validation

Following data generation, raw experimental data must be processed and validated to ensure consistency, accuracy, and suitability for modeling. This step is critical, as errors introduced here directly propagate into the model and compromise its predictive capability. The processing of data depends heavily on the data format. Some typical processing in biocatalysis includes baseline correction, normalization of data and conversion of raw data into physically meaningful quantities (such as concentration from absorption or reaction rate from concentration changes over time). Time-resolved data may require smoothing, though this should be applied cautiously to avoid distorting underlying trends. It is also imperative to use machine-readable formats in order for scripts to analyze the data.

Validation involves assessing the data quality based on defined criteria. Outliers should be identified using statistical methods rather than being removed indiscriminately. Measurement uncertainty should be quantified from the replicates, and replicate agreement should be evaluated to assess reproducibility. When possible, mass balances may provide an additional validation of the data. If the measured concentrations do not satisfy the mass balance, data reconciliation may be required.[20,57] Any inconsistencies should be explicitly acknowledged and investigated further. At the end of this step, the dataset should be ready for analysis, and its uncertainty should be clearly characterized.

3.3 Data analysis and interpretation

The goal of data analysis is to extract mechanistic insight and identify relationships that inform model structure and parameterization. Initially, the analysis should be focused on visualizing trends. In the area of biocatalysis, analyzing substrate consumption or product formation as a function of time is typical. This helps identify kinetic regimes, potential rate limitations, and nonlinear behaviors. Subsequently, parameters can be calculated, such as kinetic parameters, reaction rates and inhibitory terms. Interpretation should always be grounded in biochemical or physical reasoning as well as parameter identifiability and statistical uncertainty, since nominal parameter values alone may suggest mechanistic effects or inhibitory relationships that are not sufficiently supported by the data. Specifically for enzymatic reactions, this includes considering enzyme stability, cofactor dynamics, mass transfer limitations, and potential side reactions. A look into the literature can offer insights into mechanistic effects, expected results, and possible reasons for deviations. Importantly, deviations from expected trends should not be dismissed but rather examined, as they can often provide insight into the system. The outcome of this step is a set of numerical results and an improved understanding of the system. This understanding directly informs the selection of an appropriate model structure and defines the scope and limitations of the subsequent modeling.

4 Selection of the model structure

4.1 Ideal reactor models

The type of reactor in which a reaction process takes place is the first aspect to consider when developing the model. In reaction engineering, three ideal reactor models are commonly used to approximate the behavior of real reactors, the batch reactor (BR), the plug-flow reactor (PFR), and the continuous stirred-tank reactor (CSTR).[46] These idealized models rely on simplifying assumptions. In this work, we will primarily consider isothermal reactors, *i.e.*, reactors in which the temperature remains constant. This assumption is justified either when the heat of reaction is negligible or when sufficient heat exchange ensures effective temperature control. Under isothermal conditions, the modeling of these ideal reactors can be based solely on material balances, assuming that non-ideal mixing effects can be neglected.

4.1.1 Reaction rate expressions. Consider a system with N_s species and N_r reactions. The rate of production of species i that participates with a known stoichiometric coefficient in a number of reactions is

$$R_i = \sum_{j=1}^{N_r} \nu_{ij} r_j, \tag{7}$$

where R_i denotes the net production rate of species i, r_j is the rate of reaction j, and ν_{ij} is the stoichiometric coefficient of species i in reaction j (positive for products and negative for reactants). The indices $i = 1,\ldots, N_s$ and $j = 1,\ldots, N_r$ label species and reactions, respectively.

For enzymatic systems, the reaction rate typically depends on the concentrations of substrates, products, inhibitors, and the enzyme itself. In practice, however, enzyme activity is influenced by a much broader set of factors. Environmental conditions such as pH, temperature, ionic strength, and pressure can significantly alter catalytic performance, as can the presence of cofactors, competing substrates, or stabilizing additives. Even seemingly minor changes, such as switching buffer systems, introducing gas sparging, or altering mixing intensity, may shift the observed kinetics.

Note: reaction engineering book suggestion

Octave Levenspiel has written a classic and timeless book in reaction engineering.[46]

Because of these sensitivities, kinetic parameters measured in controlled enzyme-assay experiments do not always translate directly to production. Nevertheless, assays

provide essential baseline information, and in many cases, they represent the best available characterization. As a result, we often rely on these measurements while acknowledging that real process conditions may lead to deviations from idealized kinetic behavior.

Application example: reaction rates

The rate of production of different species in the application example is shown below:

$$R_{S_1} = -r_{\mathrm{I}}, \tag{8}$$

$$R_{S_2} = r_{\mathrm{I}} - r_{\mathrm{II}} \tag{9}$$

$$R_{S_3} = r_{\mathrm{II}} - r_{\mathrm{III}}, \tag{10}$$

$$R_{S_4} = r_{\mathrm{III}}. \tag{11}$$

When the first cofactor regeneration method is used, the rate of production of the two cofactor form is shown below:

$$R_{c_1} = -r_{\mathrm{I}} + r_{\mathrm{IV}}, \tag{12}$$

$$R_{c_2} = r_{\mathrm{I}} - r_{\mathrm{IV}}. \tag{13}$$

4.1.2 Ideal batch reactor model. For a perfectly mixed reactor without inflow or outflow (*i.e.*, an ideal batch reactor), the mole balance for species i is

$$\frac{\mathrm{d}S_i}{\mathrm{d}t} = \sum_{j=1}^{N_\mathrm{r}} \nu_{ij} r_j, \tag{14}$$

where S_i denotes the concentration of species i, t is time, r_j is the rate of reaction j, and ν_{ij} is the stoichiometric coefficient of species i in reaction j (positive for products and negative for reactants). The indices $i = 1,\ldots, N_\mathrm{s}$ and $j = 1,\ldots, N_\mathrm{r}$ refer to species and reactions, respectively. A batch reactor generally operates under unsteady-state conditions, and it is therefore typically not modeled at steady state.

4.1.3 Ideal plug-flow reactor model. For a system with spatial variation along the axial coordinate z and superficial velocity u (*i.e.*, an ideal plug-flow reactor), the transient mole balance is

$$\frac{\partial S_i}{\partial t} + u\frac{\partial S_i}{\partial z} = \sum_{j=1}^{N_\mathrm{r}} \nu_{ij} r_j, \tag{15}$$

where z is the axial position along the reactor and u is the constant flow velocity.

At steady state, the transient term vanishes, giving

$$u\frac{\mathrm{d}S_i}{\mathrm{d}z} = \sum_{j=1}^{N_\mathrm{r}} \nu_{ij} r_j, \tag{16}$$

4.1.4 Ideal continuously-stirred tank reactor model. For a well-mixed system with inflow and outflow (*i.e.*, an ideal continuously-stirred tank reactor), the transient mole balance is

$$\frac{\mathrm{d}S_i}{\mathrm{d}t} = \frac{Q}{V}\left(S_{i,\mathrm{in}} - S_i\right) + \sum_{j=1}^{N_\mathrm{r}} \nu_{ij} r_j, \tag{17}$$

where Q is the volumetric flow rate, V is the reactor volume, and $S_{i,\mathrm{in}}$ is the inlet concentration of species i.

At steady state, all time derivatives vanish, and the balance reduces to

$$0 = \frac{Q}{V}\left(S_{i,\mathrm{in}} - S_i\right) + \sum_{j=1}^{N_\mathrm{r}} \nu_{ij} r_j, \tag{18}$$

These equations provide a general and compact framework for modeling reacting systems with multiple species and reactions. Most resulting models are systems of ordinary or partial differential equations (with the exception of the steady state CSTR equation), depending on whether spatial variation is present, which directly influences the choice of numerical solution methods.

Note: residence time analogy

It is worth noting that the equations describing a steady state plug-flow reactor closely resemble those of a batch reactor. In reaction engineering, a well-known analogy exists between time in batch reactors and the residence time in plug-flow reactors. The residence time is defined as

$$\tau = \frac{z}{u} = \frac{V}{Q}, \tag{19}$$

where z is the axial coordinate, u the superficial velocity, V the volume of the reactor and Q the volumetric flow rate. This analogy is useful to keep in mind: whether time progresses in a batch reactor or the residence time increases as fluid moves through a plug-flow reactor, the governing equations are mathematically equivalent. Consequently, both models yield identical solutions when expressed in their respective independent variables.

Even though these material balances represent the 'backbone' of the model, it is not yet complete. In order to model the system, initial conditions and reaction rate expressions are the minimum additions required.

Application example: material balances

For a batch reactor, the unsteady-state material balances for the chemical species are given by:

$$\frac{\mathrm{d}S_i}{\mathrm{d}t} = R_i \tag{20}$$

For a continuous stirred-tank reactor (CSTR), the unsteady-state material balances are:

$$\frac{\mathrm{d}S_i}{\mathrm{d}t} = \frac{Q}{V}\left(S_{i,\mathrm{in}} - S_i\right) + R_i \tag{21}$$

This journal is © The Royal Society of Chemistry 2026

At steady state, this simplifies to:

$$0 = \frac{Q}{V}(S_{i,\mathrm{in}} - S_i) + R_i \tag{22}$$

For a plug-flow reactor (PFR), the unsteady-state material balances are:

$$\frac{\partial S_i}{\partial t} + u\frac{\partial S_i}{\partial z} = R_i \tag{23}$$

At steady state, this reduces to:

$$u\frac{\mathrm{d}S_i}{\mathrm{d}z} = R_i \tag{24}$$

for all species $i \in \{1, 2, 3, 4, 5\}$.

Note: real reactor and reactor model

A clear distinction must be made between the actual reactor used in laboratory or industrial practice and its corresponding mathematical model. For instance, an experimental setup may include a membrane and be referred to as an enzymatic membrane reactor (EMR), whereas its behavior is often described using an idealized model such as a continuous stirred tank reactor (CSTR).[58] Similarly, this distinction extends to all aspects of process modeling. Physical phenomena, such as mass transfer, must be differentiated from the mathematical models used to describe them.

4.2 Reaction networks

The reaction network is the collection of all relevant reactions occurring inside a reactor. (It should be distinguished from the reaction mechanism, which describes the rearrangement of molecules at the electron level). In general, a reaction network comprises two types of reactions: fast reactions that reach equilibrium almost instantaneously, and kinetically limited reactions whose rates must be explicitly modeled.[59] The reaction network also defines the stoichiometry, clarifying the set of reactions in which each species participates. In reaction-network modeling, some reactions may be treated as effectively equilibrated if they occur much faster than the processes of interest. Other reactions should be described by kinetic rate laws when they noticeably influence the observed behavior and when the available data are sufficient to estimate the corresponding kinetic parameters reliably.[50]

Application example: reaction network

In the application example, all reactions are irreversible and kinetically limited. It is also assumed that the measured data provides sufficient information for all individual reaction steps.

4.3 Considerations for multi-phase systems

When multiple phases are present in a reaction process, mass transfer between these phases often has to be taken into account. A common example in biotechnological systems is the transfer of a gaseous substrate (*e.g.*, oxygen) into a liquid phase in which enzymatic reactions occur. Another example is the presence of two immiscible liquid phases, where species must transfer across the phase boundary to react.[60–62] In such situations, it is necessary to assess the relative time scales of mass transfer and reaction. If mass transfer is fast compared to the reaction, phase equilibrium can be assumed, and thermodynamic relations are sufficient to describe the distribution of species between phases. In contrast, if mass transfer and reaction occur on comparable time scales, mass transfer must be modeled explicitly, typically through the use of mass-transfer coefficients.[63]

Phase equilibrium describes how species distribute between phases after sufficient time has elapsed for thermodynamic equilibrium to be reached. At equilibrium, the chemical potential of each species is equal in all phases. In practice, this is often expressed using simplified relations. Building on distribution coefficients, such as Henry's constant in Henry's law, for gas–liquid systems:

$$S_i^* = H_i p_i, \tag{25}$$

where S_i^* is the equilibrium concentration in the liquid phase, p_i is the partial pressure in the gas phase, and H_i is Henry's solubility constant.[64] When applying such simplified relations, it has to be considered that they rely on assumptions of near-ideal behavior in the involved phases; in particular, non-ideality in the liquid and gas phase is neglected, which is typically justified at sufficiently low concentrations and low pressures. Correction *via* activity coefficient in the liquid phase is possible if needed.[65] Nevertheless, equations of state for both gas- and liquid-phase mixtures are available and are widely presented in modern thermodynamics textbooks.[66]

When bulk-phase equilibrium cannot be assumed, interphase mass transfer is often described using a linear driving force model based on film theory:

$$N_i = k_{\mathrm{L}} a (S_i^* - S_i), \tag{26}$$

where N_i is the molar flux of species i, k_{L} is the mass-transfer coefficient, a is the interfacial area per unit volume, S_i is the bulk concentration, and S_i^* is the hypothetical equilibrium concentration corresponding to the composition of the opposite phase, usually obtained from phase-equilibrium relations.[20,63]

Naturally, other situations in which mass transfer and phase equilibrium need to be considered might exist. In enzymatic reaction systems, often two immiscible liquid phases appear in which chemical species can distribute between. This is used, for example, to extract the product from the reaction solution. Similar to gas–liquid systems, these systems can be modeled either under the assumption of thermodynamic equilibrium, or under non-equilibrium conditions where mass transfer occurs across the interface at a finite rate. In the latter case, interfacial transport

limitations are explicitly considered, and the overall transfer is governed by resistance in both liquid phases.[63]

Dimensionless numbers can provide additional insight into the relative importance of competing physical phenomena. For example, the first Damköhler number, defined as the ratio of the characteristic reaction rate to the characteristic mass transfer rate, provides a convenient criterion for assessing whether the overall process is reaction-limited (Da $\ll$ 1) or mass-transfer-limited (Da $\gg$ 1).[63]

Application example: cofactor regeneration with NOX

When cofactor regeneration strategy (c) is applied, oxygen is supplied to the reactor by sparging gas through the liquid phase, resulting in a two-phase system consisting of gas bubbles and a liquid phase containing dissolved oxygen. For example, this is the case when using the NADH oxidase enzyme for cofactor regeneration. Because oxygen is continuously transferred from the gas to the liquid, the reactor is more accurately described as a semi-batch system rather than a purely batch reactor. The dynamic oxygen balance in the liquid phase can be written as

$$\frac{dS_{O_2}}{dt} = k_L a(S^*_{O_2} - S_{O_2}) - 0.5 r_{IV}, \tag{27}$$

where S_{O_2} is the dissolved oxygen concentration in the liquid phase, t is time, k_L is the liquid-side mass transfer coefficient, a is the interfacial area per unit reactor volume, $S^*_{O_2}$ denotes the hypothetical dissolved oxygen concentration that the liquid phase would attain if it were in equilibrium with the current oxygen partial pressure in the gas phase, and r_{IV} represents the volumetric rate of oxygen consumption (or production) due to the cofactor regeneration reaction. The quasi-equilibrium concentration is described by Henry's law as

$$S^*_{O_2} = H_{O_2} p_{O_2}, \tag{28}$$

where H_{O_2} is Henry's law constant relating the gas partial pressure to the dissolved oxygen concentration and p_{O_2} is the partial pressure of oxygen in the gas phase.

4.4 Considerations for electrochemical systems

Electrochemical systems require additional considerations compared to purely chemical reaction systems, as reactions take place at the interface between an electrode and an electrolyte. As a result, both interfacial (surface) phenomena and transport processes in the bulk phase must be accounted for.[67,68]

At the electrode surface, the reaction rate is governed by electrochemical kinetics, typically described by the Butler–Volmer equation.[69] These kinetics depend on the overpotential and on the local concentrations of reacting species at the interface. In addition, surface-specific effects such as adsorption and the structure of the electrical double layer may influence the reaction rate. A common expression for the interfacial reaction rate (current density) is given by

$$j = j_0 \left[\exp\left(\frac{\alpha_a F \eta}{RT}\right) - \exp\left(-\frac{\alpha_c F \eta}{RT}\right) \right], \tag{29}$$

where j is the current density, j_0 is the exchange current density, η is the overpotential, α_a and α_c are the anodic and cathodic transfer coefficients, F is Faraday's constant, R is the gas constant, and T is the temperature.

Faraday's law provides the fundamental link between electrochemical reaction rates and measurable electrical current. It states that the rate of conversion of a chemical species at an electrode is directly proportional to the electric current passing through the electrochemical interface. In differential form, the molar rate of reaction is related to the current density through

$$r_i = \frac{j}{nF}, \tag{30}$$

where r_i is the molar rate of production or consumption of species i per unit electrode area, j is the current density, n is the number of electrons transferred per reaction event, and F is Faraday's constant. This relationship arises because each mole of electrons carries a charge F, and therefore the total charge passed in the system directly determines how much chemical transformation occurs at the electrode surface. In a reactor context, this allows the electrochemical reaction rate to be coupled to the bulk mass balance, since the consumption or generation of species in the electrolyte can be written as a volumetric rate proportional to the applied current, scaled by the electrode area and reactor volume. Consequently, Faraday's law serves as the bridge between electrochemical kinetics, expressed through current–potential relationships such as the Butler–Volmer equation, and macroscopic species balances in electrochemical reactor models.

At the same time, reactants must be transported from the bulk electrolyte to the electrode surface, and products must be removed. This transport occurs *via* diffusion, convection, and migration in the electric field. A general expression for the flux of species i is given by the Nernst–Planck equation.

$$J_i = -D_i \nabla S_i - \frac{z_i D_i F}{RT} S_i \nabla \phi + S_i \mathbf{v} \tag{31}$$

where J_i is the flux of species i, D_i is the diffusion coefficient, S_i is the concentration of species i, z_i is the ionic charge number, F is the Faraday constant, R is the universal gas constant, T is the absolute temperature, ϕ is the electric potential, and $\mathbf{v}$ is the convective velocity vector. A simplified approach is often used in which the transport is described using a mass transfer coefficient between the bulk solution and the electrode surface:

$$N_i = k_m(S_{i,b} - S_{i,s}), \tag{32}$$

where N_i is the molar flux of species i, k_m is the overall mass transfer coefficient, $S_{i,b}$ is the bulk concentration, and $S_{i,s}$ is the concentration at the electrode surface.

This journal is © The Royal Society of Chemistry 2026

Application example: electrochemical cofactor regeneration

When the third cofactor regeneration method is used, additional considerations must be made to account for the electrochemical phenomena. An example of such a modeling methodology is presented below. Consider a batch reactor of volume V with an immobilized redox mediator on an electrode of area A_{el}. The mediator cycles between oxidized (M^{ox}) and reduced (M^{red}) forms, with θ the fraction in oxidized form. The reduced mediator regenerates the cofactor, converting c_2 into c_1. Applying the same assumptions as Housseini *et al.*[70] the mediator balance is

$$\frac{d\theta}{dt} = -k_e^{red}\theta + k_e^{ox}(1-\theta) + k_c c_{2,s}(1-\theta), \tag{33}$$

with overpotential $\eta = E - E_0$ and

$$k_e^{red} = k_0 \exp\left(-\frac{\alpha_c F\eta}{RT}\right), \tag{34}$$

$$k_e^{ox} = k_0 \exp\left(\frac{\alpha_a F\eta}{RT}\right), \tag{35}$$

where $k_0 = j_0/(nF)$ and $\alpha_a = 1 - \alpha_c$. Mass transfer between bulk and surface is

$$k_L(c_{2,b} - c_{2,s}) = k_c c_{2,s}(1-\theta), \quad \Rightarrow \tag{36}$$

$$c_{2,s} = \frac{k_L}{k_L + k_c(1-\theta)} c_{2,b}. \tag{37}$$

Bulk balances (batch operation):

$$\frac{dc_{2,b}}{dt} = -\frac{A_{el}}{V} k_c c_{2,s}(1-\theta), \tag{38}$$

$$\frac{dc_{1,b}}{dt} = +\frac{A_{el}}{V} k_c c_{2,s}(1-\theta). \tag{39}$$

The current is given by

$$I_{el} = I_{res} + nFA_{el}(k_e^{ox}(\eta)(1-\theta) - k_e^{red}(\eta)\theta). \tag{40}$$

Consequently, the overall rate of an electrochemical process is often determined by the interplay between surface reaction kinetics and mass transfer. It is the modeling specialist's task to determine which steps should be considered and explicitly modeled and which ones can be omitted, approximated, or assumed to occur at a much faster timescale. This coupling can be expressed, for example, through a species balance at the electrode:

$$\frac{dS_i}{dt} = N_i + R_i, \tag{41}$$

where R_i includes the contribution from the electrochemical reaction at the interface.[71]

If mass transfer is fast relative to the electrode kinetics, the system is controlled by reaction kinetics. Conversely, if the electrochemical reaction is fast, the process may become mass-transfer limited, in which case the rate is determined by transport to the interface. In many practical systems, both effects occur on comparable time scales and must be modeled simultaneously.[70] For enzymatic processes, electrochemistry is particularly important with regards to cofactor regeneration.[72–77]

4.5 Considerations for immobilized enzyme systems

Enzyme immobilization, *i.e.*, the attachment of enzymes onto a solid support, is a widely used strategy in biocatalysis because it enables recovery and reuse of the biocatalyst, facilitates its integration into continuous-flow reactors, and often improves enzyme stability by reducing unfolding, subunit dissociation, or autolysis.[78,79] A classic industrial example is glucose isomerase immobilized on porous supports for the continuous production of high-fructose corn syrup, one of the largest-scale biocatalytic processes ever implemented.

When enzymes are immobilized inside porous particles, substrate molecules must first cross a liquid film around the particle and then diffuse along the pore structure before reaching the active sites. The rate of product formation observed in the bulk of the reactor, usually referred to as the macro-kinetic or apparent reaction rate, therefore combines the intrinsic enzyme kinetics with the rates of these transport steps. When multiple enzymes of a cascade are involved, an additional degree of freedom appears for the design and modeling: the spatial immobilization distribution (SID), *i.e.*, how the different enzymes are arranged relative to each other across particles and inside the pores.

Two complementary modeling approaches are useful to describe these systems and are both presented in this section. The first is a spatially-resolved approach that tracks concentration profiles along the pore coordinate.[34] The second is a lumped approach that describes transport globally with a single mass-transfer coefficient.[80] Both approaches ultimately reduce the coupled transport-reaction problem inside the particle to a macroscopic rate expression that can be plugged directly into the reactor material balance from section 4.1.

Note: intrinsic kinetics of immobilized enzymes

Immobilization can alter the enzyme's intrinsic kinetics. Physical attachment to a support may distort the active site, restrict conformational mobility, or induce microenvironmental effects (*e.g.*, local pH shifts, steric hindrance, or partitioning). These phenomena directly alter k_{cat}, K_M, and inhibition behavior compared to the free enzyme.

4.5.1 Spatially-resolved approach. In this approach, the concentration profiles are tracked inside a single representative pore. The pore is usually approximated as one-dimensional, with length L, cross-sectional area A, open at one end ($x = 0$) and closed at the other ($x = L$).

The enzyme surface density inside the pore is prescribed according to the chosen SID. Recent experimental approaches can help give insights in these intricate immobilized systems.[81]

Under a quasi-steady-state assumption inside the pore, justified when the total amount of immobilized enzyme is small compared to the amount of substrate in the bulk, so that the profiles inside the pore relax much faster than the bulk concentrations change, the material balance for each species i inside a pore of type n reduces to a reaction–diffusion differential equation:

$$D_i \frac{\mathrm{d}^2 S_{i,n}}{\mathrm{d}x^2} = -\sum_j \nu_{ij} r_{j,n}, \tag{42}$$

where D_i is the effective diffusivity of species i, ν_{ij} is the stoichiometric coefficient of species i in reaction j, and $r_{j,n}$ is the intrinsic reaction rate. Mass transfer in the liquid film around the particle is assumed to be fast compared to the pore diffusion, so that the boundary conditions are:

$$S_{i,n}\big|_{x=0} = S_{i,\mathrm{bulk}}, \qquad \left.\frac{\mathrm{d}S_{i,n}}{\mathrm{d}x}\right|_{x=L} = 0. \tag{43}$$

Once the concentration profiles are known, the macroscopic (apparent) reaction rate is obtained by applying Fick's law at the pore entrance and summing over all pores in the reactor:

$$R_{i,\mathrm{macro}} = \sum_n N_n A D_i \left.\frac{\mathrm{d}S_{i,n}}{\mathrm{d}x}\right|_{x=0}, \tag{44}$$

where N_n is the number of pores of type n. It is this macroscopic rate $R_{i,\mathrm{macro}}$ that enters the reactor material balance in place of the intrinsic reaction rates discussed in section 4.1.

For first-order enzyme kinetics, the resulting ODE system can be solved analytically for a range of relevant SIDs, yielding closed-form expressions for the macroscopic rates. For more complex kinetics (*e.g.*, Michaelis–Menten with inhibition), a numerical solution of a boundary value problem is typically required. Full derivations of the analytical case are available in the SI of the original reference.[34] An extension of this framework to packed-bed reactors is also available.[35]

Note: spatial immobilization distributions

For a two-enzyme cascade, four simple SIDs are commonly considered: (α) individual immobilization, in which each porous particle contains only one enzyme type; (β) homogeneous co-immobilization, in which both enzymes are uniformly distributed in every pore; and two heterogeneous co-immobilization variants (γ, δ), in which the two enzymes are placed sequentially along the pore. The choice of SID changes the concentration profiles inside the pore, and consequently, the macroscopic rate.

4.5.2 Lumped approach. When the main interest is in comparing biocatalyst formulations rather than resolving intra-pore concentration profiles, a simpler lumped description is often used.[80] The particle is treated as a well-mixed region with an intraparticle concentration $S_{i,\mathrm{p}}$, distinct from the bulk concentration $S_{i,\mathrm{b}}$. Transport from the bulk into the particle is described by a linear driving-force expression with a global volumetric mass-transfer coefficient K_L:

$$N_i = K_\mathrm{L}(S_{i,\mathrm{b}} - S_{i,\mathrm{p}}). \tag{45}$$

Inside the particle, the intrinsic enzyme kinetics are evaluated at the intraparticle concentration $S_{i,\mathrm{p}}$, and the species balance in the particle reads:

$$\frac{\mathrm{d}S_{i,\mathrm{p}}}{\mathrm{d}t} = R_i + N_i, \tag{46}$$

with a coupled balance in the bulk that accounts for the particle volume fraction.

Note: initial rates *vs.* full reaction courses

A common practical question in biocatalyst design is whether to optimize enzyme ratios based on initial rates or on the time required to reach a target yield. These two criteria often give different answers, particularly when diffusional limitations are present. Optimizing on initial rates tends to overweight the enzyme that reduces the lag phase, while optimizing on cycle time produces a more balanced formulation. This distinction becomes especially relevant for co-immobilized biocatalysts.[80]

4.5.3 Choosing a method for immobilized systems. The spatially-resolved approach provides a more detailed physical picture and is especially useful when intra-particle concentration gradients are expected to be significant, when different SIDs are being compared, or when the relation between pore geometry, diffusivities, and observed rate must be made explicit. The lumped approach, in contrast, requires fewer parameters and is well-suited for rapid comparative studies and for dynamic optimization over the full reaction course.

In practice, the two approaches are complementary. The spatially-resolved approach can be used to understand and justify a particular SID, while the lumped approach can then be used for efficient process-level optimization. However, the choice between these approaches depends heavily on the information content of the available data. Spatially resolved models are most useful when the additional required information, such as effective diffusivities, particle geometry, or internal concentration gradients, can be measured, estimated independently, or identified from sufficiently informative data. If the available data do not support this level of detail, a lumped model with identifiable effective parameters may provide a more reliable basis for comparison and optimization.

This journal is © The Royal Society of Chemistry 2026

Application example: co-immobilization for reactions II and III

Consider reaction III of the application example cascade, where the intermediate S_3 is converted to S_4 by the enzyme catalyst E^C:

$$S_3 \xrightarrow{E^C} S_4 \tag{47}$$

To investigate the impact of mass-transfer limitations within an immobilized biocatalyst particle, we formulate a spatially-resolved model inside a single cylindrical support pore of length L.

The local intrinsic consumption rate r_{III} at any position x along the pore axis depends on the localized intermediate concentration S_3, governed by Michaelis–Menten kinetics with substrate affinity constant $K_{S_3,III}$ and maximum catalytic velocity $V_{max,III}$:

$$r_{III} = \frac{V_{max,III} \cdot S_3}{K_{S_3,III} + S_3} \tag{48}$$

Assuming a steady state balance between the axial molecular diffusion of S_3 through the pore fluid and its enzymatic consumption, the transport profile is defined by the following second-order ordinary differential equation (ODE):

$$D_{S_3} \frac{d^2 S_3}{dx^2} = \frac{V_{max,III} \cdot S_3}{K_{S_3,III} + S_3} \tag{49}$$

where D_{S_3} represents the effective pore diffusivity constant of the intermediate species.

The system is uniquely closed by specifying conditions at the pore boundary coordinates:

1. At the pore mouth ($x = 0$), the local concentration matches the well-mixed bulk liquid phase concentration:

$$S_3|_{x=0} = S_3, \text{ bulk} \tag{50}$$

2. At the dead end or symmetry center of the pore ($x = L$), a zero-flux condition is applied:

$$\left.\frac{dS_3}{dx}\right|_{x=L} = 0 \tag{51}$$

Evaluating this boundary value problem permits the determination of the net macroscopic reaction rate $R_{S_4,macro}$ observed at the reactor scale. By applying Fick's law across the entrance cross-sectional area A for all N active pores, the structural boundary term becomes:

$$R_{S_4,macro} = -R_{S_3,macro} = -N \cdot A \cdot D_{S_3} \left.\frac{dS_3}{dx}\right|_{x=0} \tag{52}$$

This apparent rate $R_{S_4,macro}$ replaces the intrinsic rate laws within the global, ideal macro-reactor material balances.

5 Kinetic model parameterization

5.1 From data to model

Enzymatic reaction processes are usually characterized with respect to their kinetics individually in so-called enzyme-assay experiments, where the reaction rate of the reaction as a function of enzyme, substrate, product, and inhibitor is quantified.[82] It should be noted that the reaction rate may differ if the enzyme is used under different reaction conditions. For example, the presence of a different buffer, other substrates or gas-sparging might affect the activity of the enzyme. However, it is not always possible to study the enzyme under the exact correct conditions, and we are therefore often constrained to work with the best available information.

Historically, enzyme kinetic parameters were frequently estimated using graphical methods such as the Lineweaver–Burk plot.[83] While these methods provide a simple means of visualizing Michaelis–Menten kinetics, modern computational approaches typically estimate kinetic parameters by directly fitting the nonlinear model to experimental data using nonlinear regression. Direct nonlinear regression generally provides more accurate and statistically robust parameter estimates and can be readily extended to complex kinetic models that cannot be analyzed using graphical linearization methods.

Fitting these data to a kinetic model helps us in two ways. First of all, we have a quantitative tool to estimate the reaction rate under different conditions, even those not tested experimentally, by interpolating or extrapolating to these conditions. The extrapolating aspect is especially relevant when the mechanistic model is correctly/carefully selected as it can assume that we can deviate from the conditions of the assay experiments and still get a reasonable accuracy in our model-based prediction. For example, when the reaction rate is linearly increasing with respect to enzyme amount (often the case with enzyme kinetics) our kinetic model can quite accurately extrapolate to reaction rates using different enzyme amounts.

It should be noted that multiple methods exist for parameter estimation, and the choice of method can influence the resulting parameter values. Different approaches also require different types of datasets, produced by differently designed experiments.[32,56] In the following, two approaches are presented. The reader is encouraged to assess the quality and reliability of the estimated parameters using statistical tools for goodness of fit, confidence intervals, and confidence ellipses.[43,52,54,84] These measures provide insight into parameter identifiability and potential correlations between parameters, further discussed in 7.3.

5.2 Nonlinear regression with initial rate data

Traditional approaches to parameter estimation in enzyme kinetics often relied on graphical methods or linearized forms of kinetic expressions. While such methods can

provide qualitative insight, they introduce distortions in the error structure and are generally less robust. Nowadays, nonlinear regression offers a more consistent and convenient way to estimate parameters directly from experimental data.

In nonlinear regression, model parameters are determined by minimizing the discrepancy between measured data and model predictions. Mathematically, this is formulated as an optimization problem:

$$\hat{\theta} = \arg\min_{\theta} \sum_{i=1}^{N} \left(y_i - f(x_i, \theta)\right)^2, \tag{53}$$

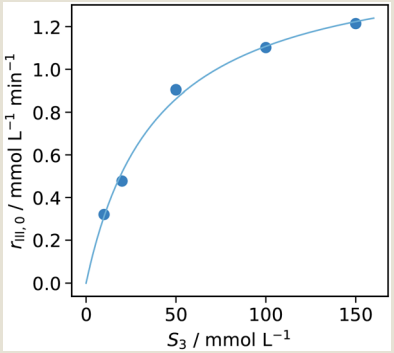
where y_i are the measured data points, x_i are the corresponding inputs (*e.g.*, substrate concentrations), $f(x_i, \theta)$ is the nonlinear model (*e.g.*, a kinetic rate expression), and θ is the vector of parameters to be estimated. This unweighted least-squares formulation treats all measurements equally; therefore, when different types of measurements are combined or when measurement accuracies differ considerably, a weighted least-squares formulation should be used, with weights typically chosen based on the measurement standard deviations, thereby approaching a maximum-likelihood formulation under appropriate error assumptions.

The minimization is typically performed using iterative numerical algorithms (such as Gauss–Newton or Levenberg–Marquardt), which update parameter estimates until the model predictions best fit the data in a least-squares sense. The solution of such parameter estimation problems is subject to defining upper and lower bounds and initial guesses for the parameters and may yield local optima in case of nonconvex problems. Despite its advantages, nonlinear regression must be applied with care. In enzyme kinetics, parameters in models such as the Michaelis–Menten equation are often strongly correlated. This means that changes in one parameter (*e.g.*, V_{max}) can be compensated by changes in another (*e.g.*, K_M), leading to multiple parameter combinations that describe the data equally well. As a result, a good fit to the data does not necessarily imply that the estimated parameters are physically meaningful or precise estimates.[32,85,86]

To assess the reliability of parameter estimates, confidence intervals must be considered. These are typically derived from the covariance matrix of the estimated parameters and provide a measure of uncertainty. Large confidence intervals or strong parameter correlations indicate that the parameters are not well identifiable from the available data. More on this in the statistical analysis section.[87,88]

Application example: initial rate data – script: initial rate.py

Initial rate data for the third reaction are shown for varying initial substrate concentrations below. The solid curve represents the best-fit model obtained from nonlinear regression applied to the measurements.



The best-fit parameters are summarized below.

Parameter	Value	Units
V_{max}	1.6	mmol L^{-1} min^{-1}
K_m	39.7	mmol L^{-1}

5.3 Nonlinear regression with progress curve data

Another method for estimating kinetic parameters is nonlinear regression based on progress curve data. In contrast to initial rate methods, this approach utilizes measurements of the time- or space-dependent evolution of species concentrations within a reactor. In this case, parameter estimation can be performed by fitting either an integrated form of the rate expression or, more generally, the full dynamic reactor model to the experimental data. Alternatively, concentration profiles may first be approximated from the progress-curve data to estimate local rates or concentration gradients, which can then be used as the basis for nonlinear regression of the kinetic parameters.[56] The reactor model typically consists of a system of differential (or differential-algebraic) equations that describe the material balances along with the relevant kinetic expressions.

Mathematically, this can be formulated as

$$\hat{\theta} = \arg\min_{\theta} \sum_{i=1}^{N} \left(S_i^{exp} - S_i^{model}(t_i, \theta)\right)^2, \tag{54}$$

where S_i^{exp} are the measured concentrations at time (or position) t_i, and $S_i^{model}(t_i, \theta)$ are the corresponding model predictions obtained by solving the governing equations for a given parameter set θ.[32,87] For example, in a batch reactor, the model prediction is obtained from the material balance

$$\frac{dS}{dt} = -r_S(S, \theta), \tag{55}$$

In the most general case, the model consists of a system of differential-algebraic equations (DAEs):

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, \mathbf{z}, \theta, t), \qquad \mathbf{0} = \mathbf{g}(\mathbf{x}, \mathbf{z}, \theta, t), \tag{56}$$

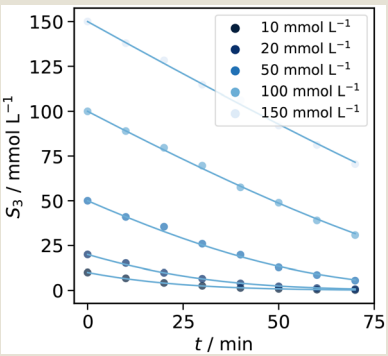
where $\mathbf{x}$ is the vector of all state variables, $\mathbf{z}$ is the vector of all control variables, θ is the vector of all parameters to be

This journal is © The Royal Society of Chemistry 2026

estimated, and t is time. This must be solved numerically during the parameter estimation procedure. This approach makes use of the larger information content of the time-resolved concentration profiles over the course of the reaction, which in particular provides information on effects that may only become observable during reaction progress, such as substrate depletion, product inhibition, enzyme deactivation, or equilibrium limitations, enabling an improved parameter identification.[82]

Application example: progress curve data – script: progress curve.py

Progress curve data for the third reaction are shown for varying initial substrate concentrations. The solid curve represents the best-fit model obtained from nonlinear regression applied to the progress curves.



The best-fit parameters are summarized below.

Parameter	Value	Units
V_{max}	1.4	mmol L^{-1} min^{-1}
K_m	24.7	mmol L^{-1}

Note: importance of parameter estimation method

Although the initial rate data were extracted from the progress curves used in the application example, the resulting parameter estimates differ substantially between the two estimation methods. This illustrates that reliable parameter estimation depends not only on the estimation method itself, but also on the underlying data, model formulation, experimental conditions, weighting, and numerical implementation.

6 Simulation

6.1 Bringing it all together

The development of the final model is achieved by combining the individual elements discussed above. The resulting model is typically a coupled set of equations that includes the reactor performance equations, reaction-rate expressions, and, where relevant, mass-transfer relations, phase equilibrium conditions, and electrochemical kinetics.

Application example: complete model

The complete dynamic model consists of material balances for all species, the corresponding reaction rate expressions, parameter values, and initial conditions.

$$\frac{dS_1}{dt} = -r_I, \tag{57}$$

$$\frac{dS_2}{dt} = r_I - r_{II}, \tag{58}$$

$$\frac{dS_3}{dt} = r_{II} - r_{III}, \tag{59}$$

$$\frac{dS_4}{dt} = r_{III}, \tag{60}$$

$$\frac{dc_1}{dt} = -\frac{dc_2}{dt} = -r_I + r_{IV}. \tag{61}$$

$$r_I = \frac{k_{cat,I} E_A S_1 c_1}{(K_{S_1,I} + S_1)(K_{c_1,I} + c_1)\left(1 + \frac{S_1}{K_{i,S_1,I}}\right)}, \tag{62}$$

$$r_{II} = \frac{k_{cat,II} E_B S_2}{(K_{S_2,II} + S_2)\left(1 + \frac{S_1}{K_{i,S_1,II}}\right)\left(1 + \frac{S_5}{K_{i,S_5,II}}\right)}, \tag{63}$$

$$r_{III} = \frac{k_{cat,III} E_C S_3}{K_{S_3,III} + S_3}, \tag{64}$$

$$r_{IV} = \frac{k_{cat,IV} E_D S_5 c_2}{(K_{S_5,IV} + S_5)(K_{c_2,IV} + c_2)\left(1 + \frac{S_1}{K_{i,S_1,IV}}\right)}. \tag{65}$$

The initial conditions are presented below.

Parameter	Value	Unit
$S_1(0)$	50	mmol L^{-1}
$S_2(0)$	0	mmol L^{-1}
$S_3(0)$	0	mmol L^{-1}
$S_4(0)$	0	mmol L^{-1}
$c_1(0)$	2	mmol L^{-1}
$c_2(0)$	0	mmol L^{-1}
$S_5(0)$	50	mmol L^{-1}

The parameter values for the kinetic models are shown below.

Parameter	Value	Unit
$k_{cat,I}$	1.2	mmol L^{-1} min^{-1} mg^{-1}
$k_{cat,II}$	0.3	mmol L^{-1} min^{-1} mg^{-1}
$k_{cat,III}$	1.4	mmol L^{-1} min^{-1} mg^{-1}
$k_{cat,IV}$	0.5	mmol L^{-1} min^{-1} mg^{-1}
$K_{S_1,I}$	30.2	mmol L^{-1}
$K_{c_1,I}$	1.1	mmol L^{-1}
$K_{i,S_1,I}$	4.3	mmol L^{-1}
$K_{S_2,II}$	8.2	mmol L^{-1}
$K_{i,S_1,II}$	8.0	mmol L^{-1}
$K_{i,S_5,II}$	14.3	mmol L^{-1}
$K_{S_3,III}$	24.7	mmol L^{-1}
$K_{S_5,IV}$	15.2	mmol L^{-1}
$K_{c_2,IV}$	3.1	mmol L^{-1}
$K_{i,S_1,IV}$	8.3	mmol L^{-1}

The complete system, together with appropriate initial and boundary conditions, constitutes a mathematical representation of the reaction process. It allows one to track the evolution of species concentrations (and, if applicable, other state variables such as

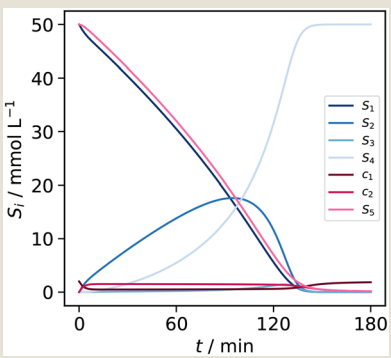
temperature or potential) within the reactor. Such a model can be used to analyze system behavior, predict performance, and support the design and optimization of reaction processes.

6.2 Solution and visualization

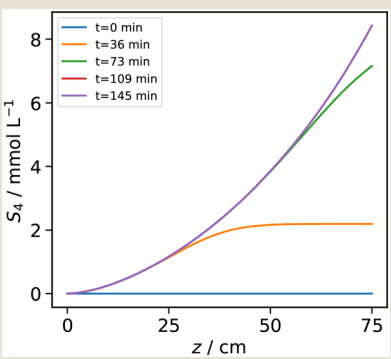
Ordinary differential equations describe the temporal or spatial evolution of state variables and are often solved using numerical integration methods such as the Runge–Kutta schemes, implicit solvers, or adaptive step-size methods. Partial differential equations (PDEs), which involve both time and spatial derivatives, can be tackled using the method of lines (discretizing space and integrating the resulting ODEs), finite difference, finite volume, or finite element approaches. Algebraic equations, including nonlinear systems, are typically solved using iterative root-finding algorithms such as the Newton–Raphson method or quasi-Newton methods.

Application example: batch and plug-flow reactor simulation – script: simulation.py

The unsteady-state batch and steady-state plug-flow reactor simulations are identical and therefore shown below in a single figure, with species concentrations on the *y*-axis and time (or residence time for the plug-flow reactor) on the *x*-axis. Full conversion is achieved before the 180 min mark. As dictated by the stoichiometry of the reaction network, the concentrations of S_1 and S_5 decrease in parallel throughout the reaction.



The unsteady-state PFR simulation is shown below, with S_4 concentrations on the *y*-axis (shown at different time points) and the length of the reactor on the *x*-axis. The length of the reactor is 75 cm. Notice that after 109 min, steady state is achieved.

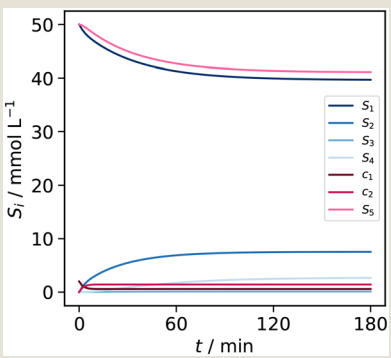


In Python, a variety of specialized functions exist for each class of problem, allowing rapid implementation and robust solutions for complex kinetic and transport models.[89] These tools are ready to use and abundant but should be used with care as they often include a variety of underlying choices, *e.g.*, different solvers, algorithms, tolerances, scaling options, *etc.* The reader is therefore advised to study the tools they apply, learn about the available options, and test whether the chosen numerical settings are appropriate for the specific problem. This is particularly important for nonlinear parameter estimation and optimization problems, where default settings may lead to poor convergence, misleading parameter estimates, or unnecessary computational effort. For advanced or critical applications, collaboration with modeling or optimization experts is advisable.

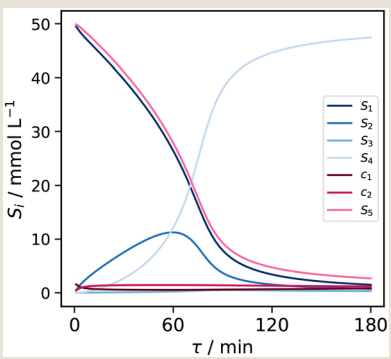
The steady state CSTR simulation is shown at the end of the following application example box, with species concentrations on the *y*-axis and the residence time on the *x*-axis. Although the time in a batch reactor and the residence time in a CSTR are not directly equivalent, the comparison provides useful insight into how reactor operation influences the reaction outcome. In the application example, the CSTR reaches high conversion at shorter effective times than the corresponding batch process. At first sight, this may not be immediately expected because the reaction network involves several interlinked reaction steps and inhibitory effects.

Application example: CSTR simulation – script: simulation.py

The unsteady-state CSTR simulation is shown below, with species concentrations at the outlet on the *y*-axis and the time on the *x*-axis. The selected residence time is 30 min. Notice that steady state is achieved after roughly 60 min.



The simulation of a CSTR in steady state is shown below as outlet concentrations of the different species over residence time. At a residence time of 30 min, the same concentrations are achieved as the ones achieved in the unsteady state simulation.



This journal is © The Royal Society of Chemistry 2026

The model-based evaluation helps to explain this behavior systematically: in the CSTR, substrates and products are continuously diluted and removed through the inlet and outlet streams, so that the reaction rates are governed by the steady-state outlet concentrations rather than by the concentration history observed in batch operation. This reduces substrate inhibition and thereby accelerates the overall conversion. The example illustrates how reactor models can be used not only to compare process alternatives, but also to identify the mechanistic reasons underlying differences in process performance.

7 Model validation

7.1 Goodness of fit

While parameter estimation ensures that the model fits the observed data well, this alone does not guarantee its predictive power under different reaction conditions. Validation answers two questions: does the model reproduce the experimental data, and are its parameter estimates reliable? We address the first with goodness-of-fit metrics and predictions on unseen experiments, and the second through identifiability and uncertainty analysis. Goodness-of-fit indicates whether the model can explain the observed data, while identifiability and uncertainty analysis indicate whether this explanation is sufficiently constrained by the data and provide a basis for assessing the reliability of predictions beyond calibration.

After fitting a model to experimental data, we must evaluate the fitting accuracy. Several statistical metrics serve this purpose, each with its own strengths and weaknesses. The list covers key examples but is not exhaustive; deeper variations exist.

1. Coefficient of determination (R^2): R^2 measures the proportion of variance in the dependent variable that can be attributed to the linear model of the independent variables. While commonly reported, the interpretation of the coefficient of determination is straightforward mainly for linear regression and can be highly misleading for nonlinear kinetic models, overparameterized models, or data sets with a poor signal-to-noise ratio. In particular, a high R^2 does not imply that the underlying model structure is mechanistically correct, nor does it guarantee that the estimated parameters are uniquely identifiable. Therefore, relying solely on this metric is prone to misinterpretation and should be supplemented by rigorous parameter diagnostics.[90,91]

2. Mean Absolute Error (MAE): MAE measures the average magnitude of errors in a set of predictions, without considering their direction. It is less sensitive to outliers compared to other metrics, but may not capture the full extent of model inaccuracies.

3. Mean Squared Error (MSE): MSE measures the average squared difference between the predicted and actual values. It is sensitive to the scale of the data and can be heavily influenced by outliers.

4. Visual inspection *via* parity plots: parity plots are a graphical tool where predicted model values are plotted against actual experimental data. Points lying on or near the diagonal indicate a good fit, while substantial deviations refer to model insufficiency. This method is particularly useful for identifying systematic errors or biases in the model.

7.2 Comparison of model to unseen data

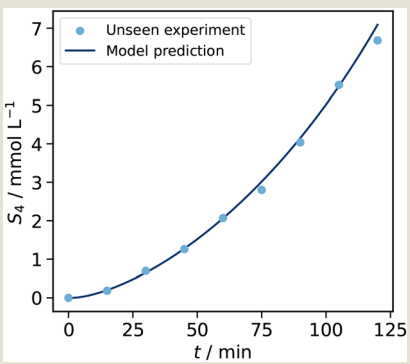
A robust validation step involves the comparison of the model's predictions with experimental data that were not used during the parameter fitting process. This step is crucial for assessing the model's predictive capability. When the model performs well on the fitting data but poorly on the unseen data, this overfitting often stems from over-parameterization. If the model's predictions match the unseen data, we consider the model validated. While no universal agreement criterion exists for validation on unseen data, the acceptable level of agreement should not be left solely to the process designer's subjective assessment. It should be defined in relation to the intended model use, measurement uncertainty, expected experimental variability, and decision-relevant tolerances, ideally using predefined validation metrics and acceptance criteria.

Note: deciding between competing kinetic mechanisms

Model comparison can help determine whether the available data discriminate between kinetic model structures associated with competing mechanistic hypotheses. For example, candidate models representing competitive, noncompetitive, or uncompetitive inhibition, or ping–pong and ordered or random Bi–Bi mechanisms, can be fitted independently to the same experimental dataset. Their relative statistical support can then be assessed using criteria such as the Akaike information criterion (AIC) or Bayesian information criterion (BIC). Model-based experimental design can further identify experimental conditions under which the candidate models yield distinct predictions, thereby improving model discrimination.[54]

Application example: comparison to unseen experiment – script: validation unseen data.py

Two progress curve experiments served as the fitting set: $S_1(0)$ = 50 mmol L^{-1} and $S_1(0)$ = 100 mmol L^{-1} (with $S_5(0) = S_1(0)$ and $c_1(0)$ = 2 mmol L^{-1}). The product S_4 was measured at nine time points over 120 min. The model describes both fitting experiments well with a mean absolute error of MAE_{train} = 0.19 mmol L^{-1}. A third experiment at $S_1(0)$ = 75 mmol L^{-1} was withheld entirely from parameter estimation. The model prediction for this unseen condition, obtained without any refitting, is compared to the experimental data below.



The prediction matches the data with MAE_{train} = 0.09 mmol L^{-1}. This agreement confirms the model's ability to generalize to conditions outside the fitting set, even though the validation points remain within the same range. To further assess robustness, the model could be challenged with unseen data requiring extrapolation beyond the fitting range.

7.3 Parameter identifiability and uncertainty analysis

The reliability of the model heavily depends on the quality of the parameter estimates. Adding more parameters to a model can improve the fit but also increases the risk of overfitting. Each parameter should contribute uniquely to the model's prediction. Despite a general plausibility check for the estimated parameter values, the following methods can help in analyzing the parameter estimates.

The Jacobian matrix J, which contains the first-order partial derivatives of the model outputs with respect to the parameters, is a key tool in the sensitivity analysis:

$$J_{i,j} = \frac{\partial y_i}{\partial \theta_j} \tag{66}$$

with y_i as i-model output and θ_j as j-th parameter. It relates the parameters with observables (*e.g.*, how V_{max} affects the substrate concentration). The parameter scale must be considered when comparing sensitivities in kinetic models, as differences in magnitude can distort the interpretation of sensitivity values. For instance, a parameter with a large absolute value (*e.g.*, K_m in millimolar range) may appear less sensitive than one with a small value (*e.g.*, k_{cat} in s^{-1}) purely due to their disparate scales, rather than their true influence on model dynamics. Reparameterization such as normalization, log-transformation, or non-dimensionalization ensures consistent scaling across parameters, enabling meaningful comparisons of their relative contributions to model behavior.

We can express the uncertainty in parameter estimates using confidence intervals. These intervals provide a range in which the true parameter value is expected to lie with a certain probability, typically 95%. If the parameter estimates are close to the true value, the confidence intervals can be approximated using the residuals of the model fit and the Jacobian matrix:[43,87]

$$\mathrm{CI} \approx f(\text{residuals}, J) \tag{67}$$

If the confidence interval spans a wide range, the respective parameter estimate is subject to large uncertainty. Such uncertainty typically indicates that the available data do not contain sufficient information to constrain the parameter independently, either because the model output is insufficiently sensitive to this parameter under the investigated conditions or because its effect is strongly correlated with that of other parameters. In case of highly correlated parameters, the model can achieve a good fit to the data through multiple combinations of parameter values, making it difficult to separate different effects. For a Michaelis–Menten model, the overestimation of V_{max} can be balanced by an overestimation of K_M, leading to a good fit but physically meaningless parameters.[52]

Confidence ellipses are a suitable tool to visualize the parameter estimate uncertainty and correlations between pairs of parameters. Under a local linear approximation and approximately normally distributed parameter uncertainty, the confidence region for a pair of parameter estimates forms an ellipse centered around the estimate in parameter space. Their computation relies on the covariance matrix of parameter estimates and the confidence level, usually 95%.[43,87] Narrow ellipses indicate precise parameter estimates and wide ellipses suggest high uncertainty, while diagonally elongated reveals strong correlation between parameters.

For the inspection of all possible parameter combinations at once, heatmaps help detect correlations. A heatmap is a graphical representation of the correlation matrix, where each entry shows the correlation coefficient between pairs of parameters with an easy-to-use color code. The correlation coefficient ρ_{ij} between parameters θ_i and θ_j is given by:

$$\rho_{ij} = \frac{\mathrm{Cov}(\theta_i, \theta_j)}{\sigma_{\theta_i}\sigma_{\theta_j}}, \tag{68}$$

where $\mathrm{Cov}(\theta_i, \theta_j)$ is the covariance between θ_i and θ_j, and σ_{θ_i} and σ_{θ_j} are their standard deviations. For ρ_{ij} close to 1, the parameters have a strong positive correlation. If close to −1, the parameters vary inversely. Values of 0 indicate no correlation. If significant correlations are detected, consider the following strategies:

1. Simplify the model: remove non-identifiable parameters or terms that contribute to correlations. For example, if two parameters are highly correlated, consider fixing one based on prior knowledge or literature values or using approximations 8.2.

2. Collect more data: design experiments to break correlations by varying inputs in a way that decouples the effects of correlated parameters. For example, varying substrate concentrations over a wider range can help distinguish V_{max} and K_M.

3. Reparameterization: transform the model to use combinations of parameters that are less correlated. For example, in the Michaelis–Menten model, using V_{max}/K_M (the catalytic efficiency) as a single parameter can sometimes reduce correlation issues.

Application example: confidence ellipses – script: confidence ellipses.py

To illustrate parameter correlation and the effect of experimental design, we isolate Reaction III ($S_3 \xrightarrow{E_C} S_4$, simple Michaelis–Menten kinetics) and fit $k_{cat,III}$ and $K_{S_3,III}$ to synthetic progress curve data. Three scenarios, each comprising exactly three progress curves, are compared so that data quantity is held constant and only data diversity varies:

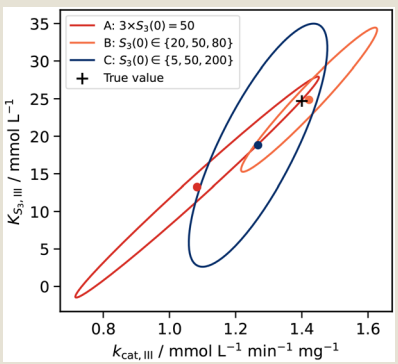
- Scenario A: three replicates of the same initial condition ($S_3(0)$ = 50 mmol L^{-1}). Replicates improve noise estimates but add no new kinetic information.

- Scenario B: three curves at $S_3(0) \in \{20, 50, 80\}$ mmol L^{-1}, all in a similar concentration regime.

This journal is © The Royal Society of Chemistry 2026

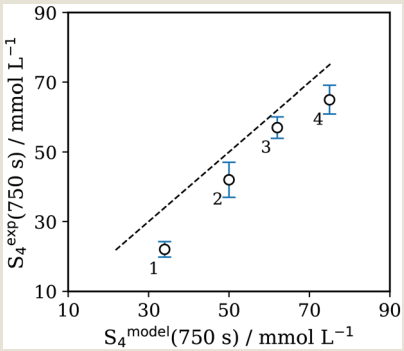
• Scenario C: three curves at $S_3(0) \in \{5, 50, 200\}$ mmol L^{-1}, spanning the linear ($S_3 \ll K_{S_3}$), intermediate, and saturating ($S_3 \gg K_{S_3}$) regimes.

The parameter correlation coefficients are ρ = 0.992 (A), 0.970 (B), and 0.801 (C). In Scenario A, the nearly degenerate ellipse shows that replicates alone cannot break the correlation between $k_{cat,III}$ and $K_{S_3,III}$: any overestimation of $k_{cat,III}$ can be compensated by an overestimation of $K_{S_3,III}$, leaving the model prediction nearly unchanged. Covering a wider substrate concentration range (Scenario C) forces the data to probe both the linear and saturated portions of the rate law independently, substantially reducing the correlation. However, because measurement noise is specified as a relative error, the highest initial concentration ($S_3(0)$ = 200 mmol L^{-1}) carries a proportionally large absolute uncertainty, which inflates the overall residual variance and leads to a larger confidence ellipse area despite the improved decorrelation. Confidence ellipses (at 95% confidence level) for $k_{cat,III}$ and $K_{S_3,III}$ under Scenarios A (red), B (orange), and C (blue) are shown below. The cross marks the true parameter values.



Application example: disagreement to experiment – script: parity.py

The measured product concentration can be compared with the predicted values using a parity plot, as shown in the figure below. The model consistently overpredicts the experimental results, indicating a systematic deviation between model and experiment. In other words, the model appears overly optimistic. It is the responsibility of the modeling specialist to determine whether this level of disagreement is acceptable or if further model refinement is required.



Importantly, the nature of the deviation can provide insight into potential model deficiencies. The model predicts a faster reaction progression than observed experimentally. This discrepancy may arise from stronger inhibition effects in the real system than those accounted for in the model, or from enzyme deactivation occurring during the reaction process, which is not adequately captured in the current formulation as well as possible non-idealities for the reactor.

7.4 Disagreement or agreement to experiment

If the model does not agree sufficiently well with experimental data, there are three key questions to consider:

1. Is the model parameterized sufficiently well?
2. Are the physics considered in the model adequate?
3. Are additional or better experimental data required to model the system adequately?

Here, "better" experimental data does not necessarily mean data with lower measurement error only, but data that are more informative for the specific modeling question. This may include measurements under additional operating conditions, a broader concentration or temperature range, replicate experiments to quantify uncertainty, higher temporal resolution, or the measurement of additional species or process variables that make relevant mechanisms observable and parameters identifiable. The limit of detection (LOD) and limit of quantification (LOQ) might also need to be checked or evaluated.

It is then the responsibility of the modeling specialist to answer these questions and determine the appropriate course of action. Much like properly designing and executing an experimental study, modeling a system accurately requires a certain level of skill that is developed by iteration and a better understanding of the underlying physics. When the model agrees to experiment sufficiently well, it can be used for studying or optimizing the reaction process under the considered assumptions. These always need to be stated and borne in mind when using the model. This also highlights the added value of the modeling process, by formalizing the assumptions and structuring the available understanding of the system.

8 Insights: parametric studies and approximations

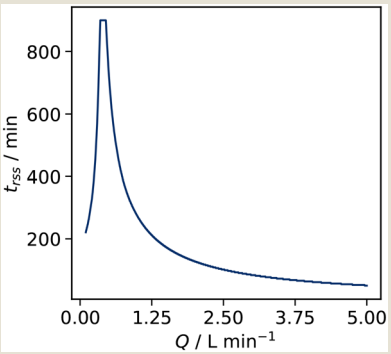
8.1 Sensitivity analysis and parametric studies

Once we have a sufficiently accurate model, we can use it to gain insights into the system. A straightforward approach for this is sensitivity analysis, which involves varying one or more parameters and observing their effect on a particular output. This can be performed by solving the model using a standard ODE or PDE solver and extracting the resulting values. The results can then be visualized in various ways, one common method is through contour plots, which allow

the simultaneous effect of two parameters on the selected output to be analyzed and presented clearly.

Application example: time to reach steady state in a CSTR – script: sensitivity analysis.py

Here, a sensitivity analysis is performed to calculate how long it takes a CSTR to reach steady state for different volumetric flow rates. The results of the sensitivity analysis are shown in the figure below. Notice that due to the complex dynamics of the system, the results do not change monotonically but instead display a maximum.



8.2 Useful approximations

Several approximations are commonly applied in reaction engineering to simplify complex kinetic models. In this section, the approximations most relevant for enzymatic reaction processes are introduced and briefly explained.[92–95]

8.2.1 Quasi-steady-state approximation. The quasi-steady-state approximation assumes that the concentration of reactive intermediates remains approximately constant over time. This, in turn means that the rate of the reactions producing it equal the rates of the reactions consuming it.

8.2.2 Quasi-equilibrium approximation. The quasi-equilibrium approximation assumes that certain steps in a reaction mechanism are sufficiently fast to reach thermodynamic equilibrium relative to the overall reaction. These steps can therefore be described using equilibrium relationships instead of rate expressions.

8.2.3 Rate-limiting step approximation. In multi-step reaction mechanisms, the overall reaction rate is often controlled by the slowest step, referred to as the rate-limiting step. Under this approximation, the overall rate can be approximated by the rate of this step.

8.2.4 Lumped parameter approximation. The lumped parameter approximation combines multiple physical or chemical effects into a single effective parameter or variable. This reduces model complexity by avoiding detailed spatial or mechanistic descriptions. For example, spatial variations in concentration within an ideal batch reactor may be neglected by assuming a uniform concentration. Similarly, multiple reaction steps may be represented by a single effective rate constant.

8.2.5 Large excess approximation. When one reactant is present in large excess compared to others, its concentration remains approximately constant during the reaction. This allows simplification of the reaction rate models and allows omitting this reactant from such models.

Application example: rate-limiting step

When cofactor regeneration is rate-limiting, the overall reaction rate of the reaction system is equal to the rate of the cofactor regeneration reaction. Under this approximation (*i.e.*, the rate-limiting step approximation for the cofactor regeneration reaction) all of the cofactor is in the c_2 form ($c_2^{\mathrm{IV,rl}} = c_1(0) + c_2(0)$, where the superscript IV,rl means that the fourth reaction is rate limiting) and the overall reaction rate of the reaction network is calculated from the following equation.

$$r_{\mathrm{overall}} = r_{\mathrm{IV}} = \frac{k_{\mathrm{cat,IV}} E_D S_5 c_2^{\mathrm{IV,rl}}}{\left(K_{S_5,\mathrm{IV}} + S_5\right)\left(K_{c_2,\mathrm{IV}} + c_2^{\mathrm{IV,rl}}\right)\left(1 + \frac{S_1}{K_{i,S_1,\mathrm{IV}}}\right)} \tag{69}$$

These approximations are essential tools in reaction engineering and can be a powerful part of an experimentalist's toolkit for obtaining rapid insights. They enable the analytical treatment of otherwise complex kinetic systems while retaining the dominant physical and chemical behavior. It should be noted that this is not an exhaustive list, and many other approximations exist.

9 Model-based optimization

9.1 Single and multi-objective problem formulations

In general terms, an objective is something we aim to achieve. In the context of optimization, an objective function is a quantitative measure that is either maximized or minimized. In industrial chemical processes, the ultimate goal is often to maximize economic profit. However, defining such an objective rigorously is challenging in practice, as it requires accurate knowledge of economic parameters (*e.g.*, raw material costs, product prices, utility costs), which may not be known *a priori* and can fluctuate over time. A possible formulation of a profit-related objective function is given by

$$\max_x O = w_1 \cdot \mathrm{STY} - w_2 \cdot \frac{E_{\mathrm{tot}}}{t} - w_3 \cdot \frac{C_{\mathrm{tot}}}{t} - \frac{C_{\mathrm{reactor}}}{t} - \frac{C_{\mathrm{utility}}}{t} \tag{70}$$

where STY denotes the space–time yield, E_{tot} represents the total enzyme consumption, and C_{tot} refers to the total cofactor consumption. The variable t corresponds to the process time, which is used to normalize consumption and cost terms on a time basis. The quantities C_{reactor} and C_{utility} denote the reactor-related and utility-related costs, respectively. The weighting factors w_1, w_2, and w_3 reflect the relative economic importance assigned to space–time yield, enzyme usage, and cofactor usage in the overall objective. At early stages of process development, it is often sufficient to consider simplified objectives. A common choice is the

This journal is © The Royal Society of Chemistry 2026

maximization of space–time yield (STY) subject to process constraints:

$$\max_x \ \mathrm{STY}(x) \quad \textbf{subject to} \quad x \in \mathscr{X} \tag{71}$$

where $\mathscr{X}$ denotes the feasible set defined by operational and physical constraints. However, in enzymatic processes, such formulations inherently involve trade-offs. For a fixed initial substrate concentration, the achievable STY depends strongly on the amount of enzyme and cofactor used. Increasing enzyme loading typically leads to higher STY, but also increases cost. This introduces a trade-off between productivity and resource consumption, which may potentially be represented by a proper weighting of the individual terms, as in eqn (70). If such an explicit weighting is infeasible, the optimal trade-offs between the objective functions can also be sought by solving a multi-objective optimization problem:

$$\max_x \ O_1 \ \min_x \{O_2, O_3\} \tag{72}$$

while such problems can be addressed by different methods including metaheuristics, such as genetic or evolutionary algorithms,[96] gradient-based optimization usually builds on a reformulation of the multi-objective optimization problem to a single objective problem by means of a scalarization strategy. The transformation of this multi objective optimization problem by means of the epsilon constraint method results in the following problem formulation:

$$\begin{aligned} \max_x \quad & O_1(x) \\ \textbf{subject to} \quad & O_2(x) \le \varepsilon_2, \\ & O_3(x) \le \varepsilon_3, \\ & x \in \mathscr{X}, \end{aligned} \tag{73}$$

where O_1 may represent productivity (*e.g.*, STY), and O_2, O_3 represent resource usage such as enzyme and cofactor consumption.[97–99] The solutions of the multi-objective optimization problem are refered to as Pareto optimal solutions, which are all possible optimal trade-offs between the different objectives. An individual solution is Pareto-optimal if no objective can be improved without worsening at least one other objective. The set of Pareto-optimal solutions is determined by solving the transformed problem for different values of ε_2 and ε_3. The final selection of an operating point from the Pareto front depends on external considerations, such as economic parameters or strategic priorities, which may not be explicitly included in the model.

Application example: objective selection

Simple objectives are often used in early-stage process development. In the following optimal control, superstructure and Bayesian optimization problems, the amount of product at the final time is considered as an objective to be maximized.

9.2 Optimal control

Once a dynamic process model in the form of differential-algebraic equations (DAEs) is available, and the objective function together with all relevant constraints has been defined, the problem can be formulated as an optimal control problem. This is particularly relevant for batch reactors and plug-flow reactors, where the system evolves over time or along a spatial coordinate.

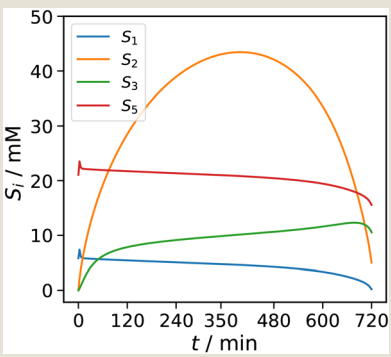
One way to solve the optimal control problem numerically is by transcribing the infinite-dimensional problem into a finite-dimensional nonlinear programming (NLP) problem. This is typically achieved *via* direct methods such as time discretization. Different discretization methods exist, backwards and forward Euler and collocation being two of the most important ones. These discretization approaches are implemented in modeling frameworks such as Pyomo, which automate the transformation of DAEs into algebraic constraints. After discretization, the optimal control problem is transformed into a large-scale NLP. Such NLPs are typically solved using gradient-based algorithms, for example, interior-point methods as implemented in solvers like IPOPT. The first-order optimality conditions for the NLP are given by the Karush–Kuhn–Tucker conditions.[44]

Note: optimal control problem

The goal of an optimal control problem is to determine optimal control trajectories (*e.g.*, feed rates, temperature profiles, or enzyme dosing strategies) that maximize or minimize a given objective while satisfying the system dynamics and constraints.[44]

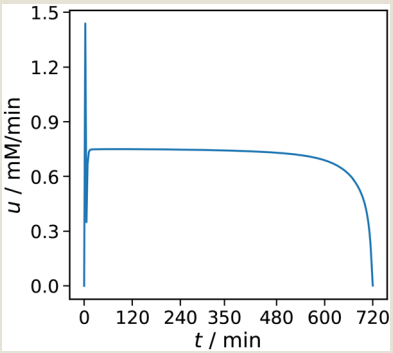
Application example: optimal dosing – script: optimal control.py

Due to the strong inhibitory effects of substrates S_1 and S_5, a process strategy is adopted in which these species are fed gradually during the course of the reaction. The resulting optimal control problem is formulated and solved using Pyomo, where the optimizer determines the optimal dosing profiles of S_1 and S_5 in order to maximize the product concentration at the final time point. The results are shown below.



The gradual addition of S_1 and S_5 is shown in the figure below. After a short oscillation at the start, it is kept constant throughout the process up until the end, where it decreases and eventually stops. It can be

observed that the concentrations of S_1 and S_5 are maintained approximately constant at around 5 mM and 20 mM, respectively. This behavior reflects the optimizer's attempt to operate in concentration regimes where substrate inhibition is minimized.



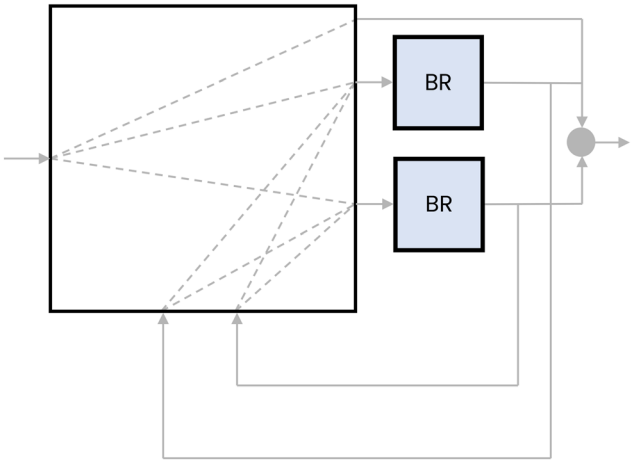


Fig. 3 Representation of a superstructure of two batch reactors.

It is important to note that the resulting NLP is generally non-convex. Consequently, numerical solvers such as IPOPT can only guarantee convergence to a local optimum, which depends on the initial guess provided. Global optimality is not guaranteed, and multiple initializations or global optimization techniques may be required to explore the solution space more thoroughly.

9.3 Superstructure optimization

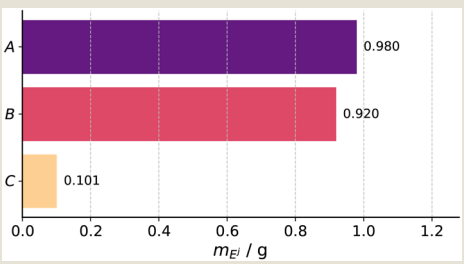
In many applications, the goal is not just to optimize a single reactor, but to determine the optimal configuration of an entire reactor network. This type of problem can be formulated as a superstructure optimization, where all plausible process configurations—different reactor types, interconnections, recycle streams, and bypasses—are included in a single overarching model. The optimizer then selects the optimal connections and operational settings to maximize the desired objective while satisfying process constraints.[100–102] Structural decisions can be modeled explicitly using binary variables, which leads to a mixed-integer nonlinear programming (MINLP) problem. In practice, continuous relaxations are often used, turning the problem into a nonlinear programming (NLP) formulation. While this simplifies the problem, NLPs are usually highly non-convex, so solvers may only find local optima. Techniques such as metaheuristics can help explore multiple starting points and guide the solution toward better or near-global optima (Fig. 3).

This approach allows simultaneous optimization of both reactor operation and network structure, making it a powerful tool for process design and intensification.

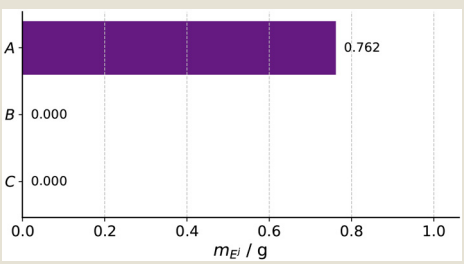
Separating enzymes from certain substrates can be important due to issues such as substrate inhibition, enzyme deactivation, and differing optimal pH or temperature conditions. In this context, superstructure optimization can be a valuable tool: it helps determine which enzymes to include in each reactor, which enzymes can operate together efficiently, and which should be separated to maintain activity and maximize overall process performance.

Application example: superstructure optimization – script: superstructure.zip

In the application example, multiple substrates inhibit different enzymes in the cascade, raising the question of how to distribute enzymes across reactors to minimize inhibition and maximize productivity. Assuming cofactor regeneration is sufficiently fast, a superstructure optimization can be used to determine optimal enzyme allocations across reactors. We consider two configurations: a single-reactor system and a two-reactor cascade. The objective is to maximize product formation from 100 mM of substrate S_1 using no more than 2 g L^{-1} total enzyme. For the single-reactor case, the optimized solution achieves a yield of approximately 25%, with the following enzyme titers.

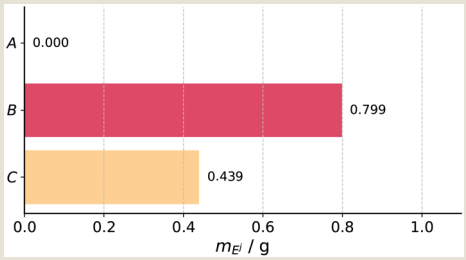


When splitting the enzymes in two reactors full conversions are achieved and the optimal enzyme titers are given below. The titers for the first reactor are shown below.



This journal is © The Royal Society of Chemistry 2026

The titers for the second reactor are shown below.



9.4 Hybrid optimization

Due to the non-convex nature of many nonlinear programming (NLP) problems arising in process and reactor optimization, gradient-based solvers such as IPOPT are generally only guaranteed to converge to local optima. To mitigate this limitation, metaheuristic optimization methods can be employed to explore the global solution space more effectively. These methods, including genetic algorithms, particle swarm optimization, and simulated annealing, do not rely on gradient information and instead use stochastic search strategies to escape local minima. While these methods do not guarantee convergence to the global optimum, they are effective at identifying high-quality solutions in complex landscapes and can be used in combination with deterministic NLP solvers in hybrid strategies, where a metaheuristic varies the initialization (exploration) while the optimum is sought by means of gradient-based optimization (exploitation).[103] This strategy is conceptually illustrated in Fig. 4, for which the workflow terminates due to some termination criteria, which is often defined on the basis of a maximum number of evaluations, but better reflects a lack of improvement over a continued number of iterations.

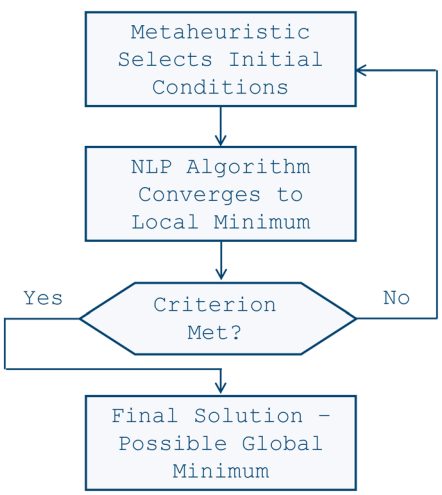
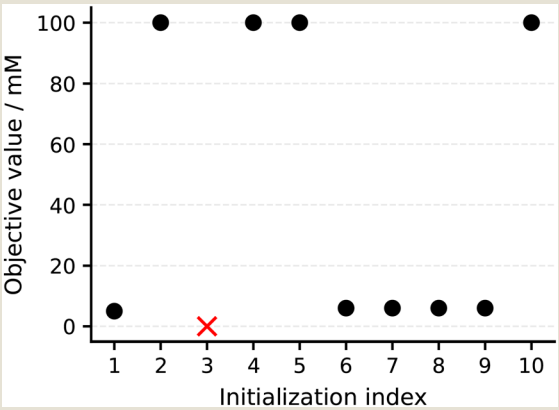


Fig. 4 A workflow for the solution of a hybrid optimization problem.

Application example: hybrid optimization

In the superstructure optimization shown in the previous section, we applied a multi-start strategy. This approach represents the simplest form of a metaheuristic: the optimizer is initialized from multiple, randomly generated starting points in an attempt to escape poor local optima and identify improved solutions. Below, we report the results obtained from each of the initializations, illustrating how even a basic multi-start strategy can enhance solution quality. Notice how different initializations result in very different results, while one initialization does not allow the optimizer to converge at all.



9.5 Bayesian optimization

Bayesian optimization is especially well-suited to problems where evaluating the objective function is costly, the response surface is non-convex, or gradients are unavailable or unreliable. As a black-box method, it requires no explicit mathematical form of the objective, only the ability to sample it at specific points.[104–106] Instead of relying on explicit derivatives, the method constructs a surrogate model of the objective function, typically based on Gaussian processes, and iteratively refines this model by selecting new evaluation points in a systematic way.[107] This allows for an efficient exploration of the decision space while minimizing the number of required model evaluations. In the present context, Bayesian optimization can support the identification of near-optimal operating conditions when the process model is expensive to evaluate, gradients are unavailable or unreliable, or conventional NLP solvers struggle with non-smooth or highly nonconvex behavior. It should therefore be regarded as a complementary optimization strategy rather than a general replacement for model-based NLP approaches. An overview of Bayesian optimization methods for enzymatic systems can be found in the work of Tachibana *et al.*[108]

Note: numerical methods book suggestions

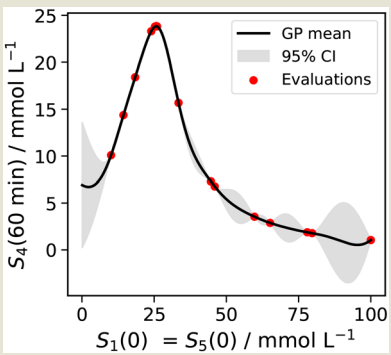
Two excellent resources for readers interested in exploring numerical methods and mathematical optimization in greater depth are the textbooks by Chapra and Martins.[109,110]

Application example: Bayesian optimization for initial substrate concentration selection – script: Bayesian optimization.py

The Bayesian optimization framework is used to solve a black-box, derivative-free optimization problem aimed at maximizing reactor performance with respect to the initial equimolar loading $S_1(0) = S_5(0) = S_0$. The objective is to maximize the product concentration after a fixed batch time of 60 minutes,

$$\max S_4(60\ \text{min}),$$

where $S_4(t)$ is obtained by numerical integration of a nonlinear system of coupled ODEs describing enzyme kinetics with inhibition and cascade reactions. Bayesian optimization constructs a Gaussian process surrogate model and sequentially selects new evaluations using an acquisition function that balances exploration and exploitation to efficiently identify the optimal S_0. The results of the Bayesian optimization are shown below.



The results show that adding roughly 25 mM of both S_1 and S_5 at the start of the reaction process leads to the maximum space–time yield at the end of the reaction process.

10 Case studies from literature

While the theoretical foundations of biochemical modeling and optimization frameworks provide the necessary mathematical tools, their true value is demonstrated through their application to complex, real-world biocatalytic processes. This section reviews several recent case studies from the literature that illustrate how these methodologies solve distinct engineering challenges. Specifically, we examine the application of black-box Bayesian optimization to design optimal pH control strategies in multi-phase reactive systems, the use of multi-objective optimization to navigate economic and operational trade-offs, and a sensitivity analysis study to select initial concentrations of intermediates. Together, these examples showcase how tailoring the optimization strategy to the specific kinetic and thermodynamic features of a system can significantly enhance process productivity and efficiency.

10.1 Bayesian optimization for optimal pH control

In a recent study, Dittmer *et al.*[50] developed a validated model that included both chemical equilibria for fast acid–base reactions, gas–liquid interfacial mass transfer, and

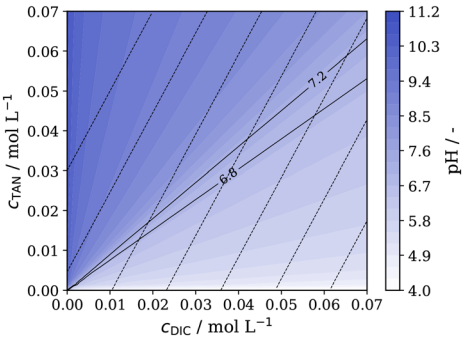


Fig. 5 Calculated pH as a function of total ammonia nitrogen (TAN) and dissolved inorganic carbon (DIC) concentrations at 35 °C. Generated using Python code from Dittmer *et al.*,[111] with extended axis ranges, based on the model reported in Dittmer *et al.*[50] The associated article is licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0) license,[50] and the code is available through TORE under the Public Domain Mark 1.0 (PDM 1.0).[111]

reaction kinetics for the enzymatic hydrolysis of urea. The reaction converts urea into ammonia and CO_2 according to:

$$(NH_2)_2CO + H_2O \xrightarrow{\text{Urease}} 2NH_3 + CO_2 \quad (74)$$

In aqueous systems, the released ammonia participates in the ammonia–ammonium acid–base equilibrium, increasing the pH. The model was used to design an optimal pH control strategy *via* the addition of CO_2 (to lower the pH) in the gas phase. Some representative results are given below. First, in Fig. 5, the pH in the reaction solution as a function of the concentration of total ammonia nitrogen and dissolved inorganic carbon is given. The dashed lines show the stoichiometry of the reaction, while the solid lines show the optimal pH region for the urease enzyme.

The model was used to determine optimal initial CO_2 pressures that maximize process productivity. Due to the complexity of the model, which makes it difficult to handle with standard nonlinear programming (NLP) solvers, Bayesian optimization was employed as an alternative approach. The resulting optimal conditions and corresponding performance are presented below (Fig. 6).

10.2 Trade-off exploration *via* multi-objective optimization

In a model-based study, Paschalidis *et al.*[20] demonstrated that trade-offs between the achieved space–time yield, enzyme consumption, and cofactor consumption exist. Intuitively, more enzyme and cofactor can help produce more product. These trade-offs were visualized in Pareto fronts, as shown in Fig. 7.

Each of these points corresponds to a Pareto-optimal process schedule; such an example of a process schedule is shown in Fig. 8.

This journal is © The Royal Society of Chemistry 2026

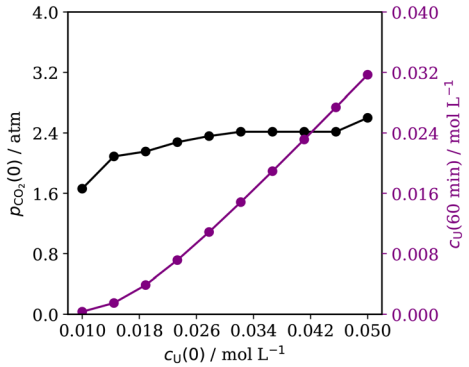


Fig. 6 Bayesian optimization results showing the optimal initial CO_2 pressure ($p_{CO_2}(0)$) and the resulting urea concentration after 60 min ($c_u(60\ min)$) for different initial urea concentrations ($c_u(0)$). Generated using Python code from Dittmer *et al.*,[111] with extended axis ranges, based on the model reported in Dittmer *et al.*[50] The associated article is licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0) license,[50] and the code is available through TORE under the Public Domain Mark 1.0 (PDM 1.0).[111]

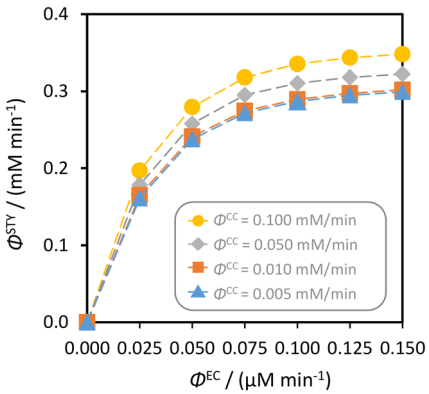


Fig. 7 Trade-offs between space–time yield, enzyme consumption, and cofactor consumption for a cascade that produces alpha-ketoglutarate. Adapted from *Biochemical Engineering Journal*, vol. 181, Paschalidis *et al.*, "Design of enzymatic cascade reactors through multi-objective dynamic optimization," article 108384, copyright (2022), with permission from Elsevier.[20]

10.3 Sensitivity analysis of intermediate addition

In another modeling study, Paschalidis *et al.*[112] investigated a strategy termed *boosting by intermediates*, which was first demonstrated experimentally by Sutiono *et al.*[113] In certain redox-neutral enzymatic cascades, process performance can be enhanced by adding reaction intermediates at the beginning of the reaction. This seemingly counterintuitive behavior highlights the strongly nonlinear interactions that can arise within complex enzymatic reaction networks. Fig. 9 shows the simplest version of such a redox neutral cascade.

Fig. 10 illustrates how increasing the initial intermediate concentration affects the residence time required to achieve complete conversion for the reaction network of Fig. 9.

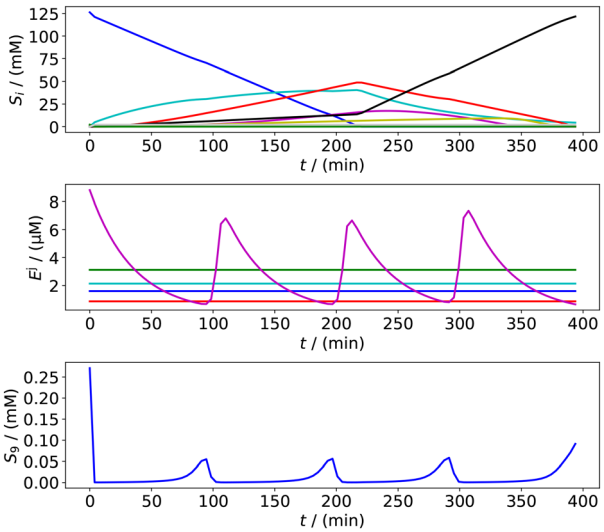


Fig. 8 A Pareto-optimal process schedule for the cascade. Adapted from *Biochemical Engineering Journal*, vol. 181, Paschalidis *et al.*, "Design of enzymatic cascade reactors through multi-objective dynamic optimization," article 108384, copyright (2022), with permission from Elsevier.[20] Figure generated using the code provided by the authors as SI for this publication, available under the MIT license at https://github.com/TUM-CS-CTV/CascadeMOO.

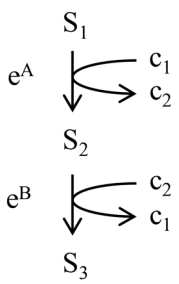


Fig. 9 One of the reaction networks studied by Paschalidis *et al.*[112] Adapted from Paschalidis *et al.*[112] under the Creative Commons Attribution 4.0 International (CC BY 4.0) license.

10.4 Other noteworthy case studies

A few additional case studies from the literature are worth noting. Recently, Chandra *et al.*[102] used a superstructure optimization algorithm to design an optimal network of batch reactors for a validated model of the Weimberg pathway. The authors found that distributing the enzymes across multiple reactors can improve productivity while mitigating inhibition effects. Another interesting study by Waluga *et al.*[32] demonstrated how different parameter estimation methods influence the resulting parameter values. Finally, Paschalidis *et al.*[34,35] published two studies on the modeling of immobilized enzyme systems and used parametric studies to derive further insights from the resulting models.

11 Conclusions

This tutorial presents a structured workflow for modeling and optimizing enzymatic reaction processes, covering all key steps

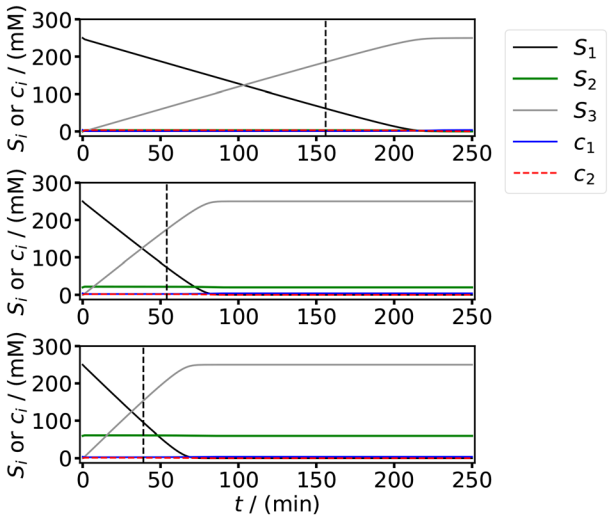


Fig. 10 Effect of the initial intermediate concentration on the space–time yield. Increasing the initial intermediate concentration (S_2) substantially enhances the space–time yield. Reproduced from Paschalidis *et al.*[112] under the Creative Commons Attribution 4.0 International (CC BY 4.0) license.

from experimentation to model-based optimization. The article does not introduce new modeling techniques, but demonstrates how established methods can be combined into a consistent, application-oriented workflow supported by implementation examples. Its main contribution is therefore to make these approaches more accessible to experimental researchers, support their application in early-stage process development, and facilitate informed collaboration with modeling and optimization experts.

The discussion shows that kinetic modeling of enzymatic systems requires careful selection of the reaction network, appropriate reactor representation (*e.g.*, batch, CSTR, plug-flow), and inclusion of relevant physical effects such as mass transfer, phase equilibria, or electrochemical phenomena where necessary. Rather than attempting to represent all underlying mechanisms, the model should capture the effects that are relevant for the specific objective of the study and the information that is available from the experimental data.

Parameter estimation is identified as a critical step, with the article illustrating how different approaches (*e.g.*, initial rate *vs.* progress curve fitting) can lead to different parameter values. This emphasizes the need to assess parameter identifiability and uncertainty, and to interpret fitted parameters with caution. Model validation against independent data is similarly important to ensure that the model is suitable for its intended use.

Once established, the model provides a basis for simulation, allowing the analysis of system behavior under different conditions. Building on this, parametric studies and approximations (such as quasi-steady-state or rate-limiting step assumptions) can be used to extract qualitative insights. Finally, model-based optimization methods—ranging from standard formulations to Bayesian and hybrid approaches—can support the exploration of operating conditions or process configurations.

Nomenclature

Latin symbols

A	Area
A_{el}	Electrode area
a	Interfacial area per unit reactor volume
c_i	Concentration of cofactor species i
$c_{i,\mathrm{b}}$	Bulk concentration of species i
$c_{i,\mathrm{s}}$	Surface concentration of species i
D_i	Diffusion coefficient of species i
E	Free enzyme concentration
E_0	Initial enzyme concentration
F	Faraday constant
H_i	Henry's law constant of species i
I_{el}	Electric current
I	Inhibitor concentration
j	Current density
j_0	Exchange current density
J_i	Flux of species i
k	Generic kinetic rate constant
k_0	Standard electrochemical rate constant
k_1, k_{-1}, k_2	Elementary reaction rate constants
k_{c}	Chemical reaction rate constant
$k_{\mathrm{e}}^{\mathrm{ox}}$	Electrochemical oxidation rate constant
$k_{\mathrm{e}}^{\mathrm{red}}$	Electrochemical reduction rate constant
k_{L}	Liquid-side mass-transfer coefficient
k_{m}	Overall mass-transfer coefficient
K_i	Inhibition constant
K_{M}	Michaelis constant
L	Characteristic pore length
M^{ox}	Oxidized mediator
M^{red}	Reduced mediator
N_i	Molar flux of species i
N_n	Number of pores of type n
N_{r}	Number of reactions
N_{s}	Number of chemical species
p_i	Partial pressure of species i
Q	Volumetric flow rate
r_j	Rate of reaction j
R	Universal gas constant
R_i	Net rate of production of species i
$R_{i,\mathrm{macro}}$	Apparent (macroscopic) production rate of species i
S_i	Concentration (state variable) of species i
$S_{i,\mathrm{in}}$	Inlet concentration of species i
$S_{i,\mathrm{bulk}}$	Bulk concentration of species i
S_i^*	Equilibrium concentration of species i
t	Time
T	Temperature
u	Superficial velocity
V	Reactor volume
$V_{\max}$	Maximum enzymatic reaction rate
x	Spatial coordinate within porous support
z	Axial reactor coordinate
z_i	Ionic charge number of species i

This journal is © The Royal Society of Chemistry 2026

Greek symbols

α	Charge-transfer coefficient
α_a	Anodic charge-transfer coefficient
α_c	Cathodic charge-transfer coefficient
η	Electrochemical overpotential
ν_{ij}	Stoichiometric coefficient of species i in reaction j
θ	Fraction of oxidized mediator
ϕ	Electric potential
τ	Residence time

Subscripts

b	Bulk phase
el	Electrode
in	Inlet
macro	Macroscopic (apparent)
ox	Oxidized form
red	Reduced form
s	Surface
0	Initial value or reference value

Abbreviations

AIC	Akaike information criterion
BIC	Bayesian information criterion
BR	Batch reactor
CSTR	Continuous stirred-tank reactor
DAEs	Differential-algebraic equations
EMR	Enzymatic membrane reactor
FDH	Formate dehydrogenase
MINLP	Mixed-integer nonlinear programming
NAD	Nicotinamide adenine dinucleotide
NADH	Reduced nicotinamide adenine dinucleotide
NLP	Nonlinear programming
NOX	NADH oxidase
ODE	Ordinary differential equation
PDE	Partial differential equation
PFR	Plug-flow reactor
SID	Spatial immobilization distribution
STY	Space–time yield

Conflicts of interest

There are no conflicts to declare.

Data availability

The Python code developed in this work is made publicly available in a TORE repository hosted by the Library of the Hamburg University of Technology (TUHH) (https://doi.org/10.15480/882.17328). The same code is also made available as a zip file as supplementary information (SI). See DOI: https://doi.org/10.1039/d6re00247a.

No other datasets were generated or analyzed in this tutorial account beyond those included in the manuscript.

Acknowledgements

Financial support from the Collaborative Research Center (CRC) "SMART Reactors for Future Process Engineering," funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number 503850735 – at the Hamburg University of Technology (TUHH), is gratefully acknowledged. Dr. Manuel Ibañez gratefully acknowledges financial support from the Walter Benjamin Programme of the Deutsche Forschungsgemeinschaft (DFG). Maerthe Theresa Tillmann gratefully acknowledges financial support from Ehrfeld Mikrotechnik GmbH. During manuscript preparation, the authors used ChatGPT (OpenAI, GPT-5.5) and Claude (Anthropic, Sonnet 4.6 and Opus 4.8) to assist with proofreading, grammar correction, and language refinement. Code for the application examples was written with assistance from Claude (Anthropic, Opus 4.7). All AI-assisted outputs were critically reviewed, tested where applicable, and edited by the authors, who retain full responsibility for the content of the manuscript, including all code, computational results, and scientific conclusions.

Notes and references

1 K. A. Johnson, *FEBS Lett.*, 2013, **587**, 2753–2766.
2 K. A. Johnson and R. S. Goody, *Biochemistry*, 2011, **50**, 8264–8269.
3 S. Patti, I. Magrini Alunno, S. Pedroni, S. Riva, E. E. Ferrandi and D. Monti, *ChemSusChem*, 2025, **18**, e202402007.
4 O. G. Ndochinwa, Q.-Y. Wang, O. C. Amadi, T. N. Nwagu, C. I. Nnamchi, E. S. Okeke and A. N. Moneke, *Heliyon*, 2024, **10**, e32673.
5 Y. R. Maghraby, R. M. El-Shabasy, A. H. Ibrahim and H. M. E.-S. Azzazy, *ACS Omega*, 2023, **8**, 5184–5196.
6 J. Bié, B. Sepodes, P. C. B. Fernandes and M. H. L. Ribeiro, *Processes*, 2022, **10**, 494.
7 J. M. Guisan, G. Fernandez-Lorente, J. Rocha-Martin and D. Moreno-Gamero, *Curr. Opin. Green Sustainable Chem.*, 2022, **35**, 100593.
8 S. Liu, M. Bilal, K. Rizwan, I. Gul, T. Rasheed and H. M. N. Iqbal, *Int. J. Biol. Macromol.*, 2021, **190**, 396–408.
9 Q. Zhu, Y. Zheng, Z. Zhang and Y. Chen, *Nat. Protoc.*, 2023, **18**, 3080–3125.
10 L. Huang, X. Li and Z. Li, *React. Chem. Eng.*, 2024, **9**, 1605–1628.
11 A. Sharma, G. Gupta, T. Ahmad, S. Mansoor and B. Kaur, *Food Rev. Int.*, 2021, **37**, 121–154.
12 Y. Wang, P. Xue, M. Cao, T. Yu, S. T. Lane and H. Zhao, *Chem. Rev.*, 2021, **121**, 12384–12444.
13 Z. Chen, Y. Yu, Y. Gao and Z. Zhu, *ACS Nano*, 2023, **17**, 13062–13080.
14 F. H. Arnold, *Angew. Chem., Int. Ed.*, 2019, **58**, 14420–14426.
15 J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S.

Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli and D. Hassabis, *Nature*, 2021, **596**, 583–589.
16 E. Erdem and J. M. Woodley, *ACS Catal.*, 2024, **14**, 18436–18441.
17 S. Weijers and K. Van't Riet, *Biotechnol. Adv.*, 1992, **10**, 237–249.
18 A. Sadana, *Biotechnol. Adv.*, 1988, **6**, 349–446.
19 J. Miłek, J. Liszkowska and M. Wróblewski, *Catalysts*, 2026, **16**(3), 281.
20 L. Paschalidis, B. Beer, S. Sutiono, V. Sieber and J. Burger, *Biochem. Eng. J.*, 2022, **181**, 108384.
21 Y. Ma and B. E. Eser, *Catalysts*, 2025, **15**(12), 1139.
22 R. H. Ringborg and J. M. Woodley, *React. Chem. Eng.*, 2016, **1**, 10–22.
23 A. Malvis Romero, L. Pesci, S. Kara and A. Liese, in *Enzyme Kinetics*, ed. K.-E. Jaeger, A. Liese and C. Syldatk, Springer International Publishing, Cham, 2024, pp. 61–90.
24 D. Ohde, S. Kühn and A. Liese, in *Enzyme Reactors and Process Control*, ed. K.-E. Jaeger, A. Liese and C. Syldatk, Springer International Publishing, Cham, 2024, pp. 91–121.
25 L. Pesci, S. M. Glueck, P. Gurikov, I. Smirnova, K. Faber and A. Liese, *FEBS J.*, 2015, **282**, 1334–1345.
26 D. Vasic-Racki, J. Bongs, U. Schörken, G. A. Sprenger and A. Liese, *Bioprocess Biosyst. Eng.*, 2003, **25**, 285–290.
27 I. E. Grossmann and I. Harjunkoski, *Comput. Chem. Eng.*, 2019, **126**, 474–484.
28 E. N. Pistikopoulos, A. Barbosa-Povoa, J. H. Lee, R. Misener, A. Mitsos, G. V. Reklaitis, V. Venkatasubramanian, F. You and R. Gani, *Comput. Chem. Eng.*, 2021, **147**, 107252.
29 K. Sivaramakrishnan, A. Puliyanda, D. T. Tefera, A. Ganesh, S. Thirumalaivasan and V. Prasad, *Ind. Eng. Chem. Res.*, 2019, **58**, 11149–11163.
30 M. S. Reis and P. M. Saraiva, *Comput. Chem. Eng.*, 2021, **155**, 107529.
31 M. Mowbray, M. Vallerio, C. Perez-Galvan, D. Zhang, A. Del Rio Chanona and F. J. Navarro-Brull, *React. Chem. Eng.*, 2022, **7**, 1471–1509.
32 T. Waluga, F. von Ziegner and M. Skiborowski, *Process Biochem.*, 2025, **151**, 1–13.
33 F. von Ziegner, G. Brauckmann, C. Witthoefft, P. Bubenheim and T. Waluga, *Chem Bio Eng.*, 2026, **3**(7), 641–649.
34 L. Paschalidis, S. Arana-Peña, V. Sieber and J. Burger, *React. Chem. Eng.*, 2023, **8**, 2234–2244.
35 L. Paschalidis, S. Arana-Peña, V. Sieber and J. Burger, *Chem. Ing. Tech.*, 2024, **96**, 741–748.
36 M. Česnik, M. Sudar, K. Hernández, S. Charnock, I. Vasić-Rački, P. Clapés and Z. Findrik Blažević, *React. Chem. Eng.*, 2020, **5**, 747–759.
37 H. Salami, C. E. Lagerman, P. R. Harris, M. A. McDonald, A. S. Bommarius, R. W. Rousseau and M. A. Grover, *React. Chem. Eng.*, 2020, **5**, 2064–2080.
38 S. Hoops, S. Sahle, R. Gauges, C. Lee, J. Pahle, N. Simus, M. Singhal, L. Xu, P. Mendes and U. Kummer, *Bioinformatics*, 2006, **22**, 3067–3074.
39 J. A. E. Andersson, J. Gillis, G. Horn, J. B. Rawlings and M. Diehl, *Math. Program. Comput.*, 2019, **11**, 1–36.
40 V. Kozachynskyi, M. Illner, E. Esche and J.-U. Repke, *Comput. Chem. Eng.*, 2024, **187**, 108746.
41 B. Nicholson, J. D. Siirola and J.-P. Watson, *et al.*, *Math. Program. Comput.*, 2018, **10**, 187–223.
42 R. C. Aster, B. Borchers and C. H. Thurber, *Parameter Estimation and Inverse Problems*, Academic Press, Boston, 2nd edn, 2012.
43 Y. Bard, *Nonlinear Parameter Estimation*, Academic Press, New York, 1974.
44 L. T. Biegler, *Nonlinear Programming: Concepts, Algorithms, and Applications to Chemical Processes*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2010.
45 S. Scott, B. Johnson, A. Taylor and M. Tinsley, *Chem. Eng. Sci.*, 2000, **55**, 209–215.
46 O. Levenspiel, *Chemical Reaction Engineering*, John Wiley & Sons, New York, 3rd edn, 1999.
47 M. Teshima, V. P. Willers and V. Sieber, *Curr. Opin. Biotechnol.*, 2023, **79**, 102868.
48 S. Arana-Peña, D. Carballares, R. Morellon-Sterlling, Á. Berenguer-Murcia, A. R. Alcántara, R. C. Rodrigues and R. Fernandez-Lafuente, *Biotechnol. Adv.*, 2021, **51**, 107584.
49 N. Plumeré and B. A. Johnson, *J. Am. Chem. Soc.*, 2024, **146**, 34771–34785.
50 K. R. Dittmer, L. Paschalidis, C. Ng-Brossard, D. Ohde, A. Liese and M. Skiborowski, *Ind. Eng. Chem. Res.*, 2026, **65**, 13617–13626.
51 A. Cornish-Bowden, *Fundamentals of Enzyme Kinetics*, Wiley-VCH, Weinheim, 4th edn, 2012.
52 T. Waluga and P. Nagshi, *AppliedChem*, 2026, **6**(2), 25.
53 V. Hilborne, *A Guide to Analytical Method Development*, Royal Society of Chemistry, 2026.
54 L. Hennecke, L. Schaare, M. Skiborowski and A. Liese, *React. Chem. Eng.*, 2024, **9**, 2984–2993.
55 C. Michalik, T. Schmidt, M. Zavrel, M. Ansorge-Schumacher, A. Spiess and W. Marquardt, *Chem. Eng. Sci.*, 2007, **62**, 5592–5597.
56 M. Zavrel, C. Michalik, T. Schwendt, T. Schmidt, M. Ansorge-Schumacher, C. Janzen, W. Marquardt, J. Büchs and A. C. Spiess, *Chem. Eng. Sci.*, 2010, **65**, 2491–2499.
57 P. Hassenstein, D. Niehaus, N. Taherkhani, J.-D. H. Krueger, D. Voß, D. Lumpp, B. Schroeter, L. Paschalidis, I. Smirnova, M. Skiborowski and J. Albert, *ACS Sustainable Chem. Eng.*, 2026, **14**, 7559–7574.
58 J. Ma, J. Chen, Q. Tian, X. Chen, Z. Huang, H. Wang, H. Shen, X. Huang, J. Yu, H. Wang, S. Zhang and Z. Yu, *React. Chem. Eng.*, 2025, **10**, 2748–2763.
59 O. Walz, C. Marks, J. Viell and A. Mitsos, *Comput. Chem. Eng.*, 2017, **98**, 143–153.
60 R. M. Lindeque and J. M. Woodley, *Biocatal. Biotransform.*, 2026, 1–6.
61 J. Wang, E. Erdem, L. Schmermund, X. Liang, V. Sieber and J. M. Woodley, *ACS Catal.*, 2025, **15**, 9277–9289.
62 I. P. Rosinha Grundtvig, S. Heintz, U. Krühne, K. V. Gernaey, P. Adlercreutz, J. D. Hayler, A. S. Wells and J. M. Woodley, *Biotechnol. Adv.*, 2018, **36**, 1801–1814.

 This journal is © The Royal Society of Chemistry 2026

63 R. B. Bird, W. E. Stewart, E. N. Lightfoot and D. J. Klingenberg, *Introductory Transport Phenomena*, John Wiley & Sons, 1st edn, 2015.
64 R. Sander, W. E. Acree, A. D. Visscher, S. E. Schwartz and T. J. Wallington, *Pure Appl. Chem.*, 2022, **94**, 71–85.
65 J. H. Vera, *Can. J. Chem. Eng.*, 2000, **78**, 429–430.
66 K. R. Cox and W. G. Chapman, *J. Am. Chem. Soc.*, 2001, **123**, 6745.
67 C. Batchelor-McAuley, E. Kätelhön, E. O. Barnes, R. G. Compton, E. Laborda and A. Molina, *ChemistryOpen*, 2015, **4**, 224–260.
68 S. Arshi, M. Nozari-Asbemarz and E. Magner, *Catalysts*, 2020, **10**, 1232.
69 E. J. Dickinson and A. J. Wain, *J. Electroanal. Chem.*, 2020, **872**, 114145.
70 W. E. Housseini, M. Etienne, E. Lojou, A. Walcarius and F. Lapicque, *Chem. Eng. Process.*, 2023, **187**, 109326.
71 J. S. Newman and K. E. Thomas-Alyea, *Electrochemical Systems*, John Wiley & Sons, Hoboken, NJ, 3rd edn, 2004.
72 Y. S. Lee, R. Gerulskis and S. D. Minteer, *Curr. Opin. Biotechnol.*, 2022, **73**, 14–21.
73 W. Liu and P. Wang, *Biotechnol. Adv.*, 2007, **25**, 369–384.
74 C. Zhang, H. Zhang, J. Pi, L. Zhang and A. Kuhn, *Angew. Chem., Int. Ed.*, 2022, **61**, e202111804.
75 M. Hou, J. Yuan, X. Dong, Y. Wang, S. Yang and J. Gao, *JACS Au*, 2024, **4**, 3581–3592.
76 S. García, C. Cocuzza, L. Wilson, M. Piumetti and C. Ottone, *Microporous Mesoporous Mater.*, 2025, **396**, 113706.
77 J. Wolf, R. Goy, J. A. Medlock, J. T. Kleinhaus, K. Pellumbi, L. Wickert, D. Siegmund and U.-P. Apfel, *ChemSusChem*, 2026, **19**, e202502221.
78 J. M. Bolivar, J. M. Woodley and R. Fernandez-Lafuente, *Chem. Soc. Rev.*, 2022, **51**, 6251–6290.
79 A. Liese and L. Hilterhaus, *Chem. Soc. Rev.*, 2013, **42**, 6236–6249.
80 D. Carballares, J. M. Woodley, R. Fernandez-Lafuente and J. M. Bolivar, *Chem. Eng. Sci.*, 2025, **318**, 122219.
81 M. R. Serial, L. Schmidt, M. Adrian, G. Brauckmann, S. Benders, V. Bueschler, A. Liese and A. Penn, *Biochem. Eng. J.*, 2026, **225**, 109909.
82 J. L. Rendón and J. P. Pardo, *Biomolecules*, 2025, **15**(5), 641.
83 H. Lineweaver and D. Burk, *J. Am. Chem. Soc.*, 2002, **56**, 658–666.
84 M. C. Sadino-Riquelme, J. Rivas, D. Jeison, R. E. Hayes and A. Donoso-Bravo, *Biotechnol. Bioeng.*, 2020, **117**, 1357–1366.
85 B. Choi, G. A. Rempala and J. K. Kim, *Sci. Rep.*, 2017, **7**, 17018.
86 H. J. Fromm, *Initial Rate Enzyme Kinetics*, Springer Science & Business Media, 2013.
87 P. Englezos and N. Kalogerakis, *Applied Parameter Estimation for Chemical Engineers*, CRC Press, Boca Raton, FL, 2001.
88 A. G. McDonald and K. F. Tipton, *Molecules*, 2022, **27**(1), 263.
89 J. C. Butcher, Numerical Differential Equation Methods, in *Numerical Methods for Ordinary Differential Equations*, ed. J. C. Butcher, John Wiley & Sons, Ltd, 2016, ch. 2, pp. 55–142, DOI: 10.1002/9781119121534.ch2.
90 A.-N. Spiess and N. Neumeyer, *BMC Pharmacol.*, 2010, **10**, 6.
91 X. Ying, *J. Phys.: Conf. Ser.*, 2019, **1168**, 022022.
92 A. Pellissier-Tanon, G. Morgado, L. Jullien and A. Lemarchand, Quasi-Steady-State and Partial-Equilibrium Approximations in Chemical Kinetics: One Stage Beyond the Elimination of a Fast Variable, 2021, https://www.cambridge.org/engage/coe/article-details/60c757d1bb8c1a808a3dc8fe.
93 M. R. Roussel, *Math. Biosci.*, 2019, **318**, 108274.
94 D. Y. Murzin, *Reactions*, 2020, **1**, 37–46.
95 J. Eilertsen, S. Schnell and S. Walcher, *Bull. Math. Biol.*, 2024, **86**, 68.
96 N. Gunantara, *Cogent Eng.*, 2018, **5**, 1502242.
97 M. Bortz, J. Burger, N. Asprion, S. Blagov, R. Böttcher, U. Nowak, A. Scheithauer, R. Welke, K.-H. Küfer and H. Hasse, *Comput. Chem. Eng.*, 2014, **60**, 354–363.
98 J. Burger, N. Asprion, S. Blagov, R. Böttcher, U. Nowak, M. Bortz, R. Welke, K.-H. Küfer and H. Hasse, *Chem. Ing. Tech.*, 2014, **86**, 1065–1072.
99 J. Burger and H. Hasse, *Chem. Eng. Sci.*, 2013, **99**, 118–126.
100 A. Lakshmanan and L. T. Biegler, *Ind. Eng. Chem. Res.*, 1996, **35**, 1344–1353.
101 V. P. Teixeira, M. A. da S. S. Ravagnani and C. B. B. Costa, *Chem. Eng. Commun.*, 2023, **210**(3), 361–380.
102 S. Chandra, L. Paschalidis, S. Kinau and M. Skiborowski, *Syst. Control Trans.*, 2026, **5**, 2176–2182.
103 F. I. Gómez-Castro and V. Rico-Ramírez, *Optimization in Chemical Engineering: Deterministic, Meta-Heuristic and Data-Driven Techniques*, Walter de Gruyter GmbH & Co KG, Berlin/Boston, 2025, p. 474.
104 F. Gisperg, R. Klausser, M. Elshazly, J. Kopp, E. P. Brichtová and O. Spadiut, *Biotechnol. Bioeng.*, 2025, **122**, 1313–1325.
105 M. Siska, E. Pajak, K. Rosenthal, A. del Rio Chanona, E. von Lieres and L. M. Helleckes, *Biotechnol. Bioeng.*, 2026, **123**, 805–830.
106 S. Desimpel, M. Dorbec, K. M. Van Geem and C. V. Stevens, *Chem. Soc. Rev.*, 2026, **55**, 2731–2775.
107 K. McBride and K. Sundmacher, *Chem. Ing. Tech.*, 2019, **91**, 228–239.
108 R. Tachibana, K. Zhang, Z. Zou, S. Burgener and T. R. Ward, *ACS Sustainable Chem. Eng.*, 2023, **11**, 12336–12344.
109 S. C. Chapra and R. P. Canale, *Numerical Methods for Engineers*, McGraw–Hill Education, 8th edn, 2021.
110 J. R. R. A. Martins and A. Ning, *Engineering Design Optimization*, Cambridge University Press, Cambridge, UK, 2022.
111 K. R. Dittmer, L. Paschalidis, C. Ng-Brossard, D. Ohde, A. Liese and M. Skiborowski, CO2-Driven pH Control in the Enzymatic Hydrolysis of Urea: A Coupled Modeling–Experiment Approach, 2026, DOI: 10.15480/882.17184.
112 L. Paschalidis, D. Fröschl, M. Ibañez, S. Sutiono, V. Sieber and J. Burger, *Biochem. Eng. J.*, 2024, **210**, 109440.
113 S. Sutiono, I. Zachos, L. Paschalidis, A. Pick, J. Burger and V. Sieber, *ACS Sustainable Chem. Eng.*, 2023, **11**, 6592–6599.