Good modeling practice for the calibration of ion exchange breakthrough prediction

Daniel Illana González[a,b] *, Mariane Yvonne Schneider[a,b] , Juan Pablo Gallo[a,b], Ingmar Nopens[a,b], Elena Torfs[c]

[a] BIOMATH, Department of Data Analysis and Mathematical Modelling. Ghent University, Coupure Links 653, 9000 Gent, Belgium.
[b] CAPTURE, Centre for Advanced Process Technology for Urban REsource recovery, Frieda Saeysstraat 1, 9052 Gent
[c] model*EAU*, Département de génie civil et de génie des eaux, Université Laval, pavillon Adrien-Pouliot, 1065, av. de la Médecine, Québec, G1V 0A6, Canada.
* Corresponding author. E-mail: *Daniel.IllanaGonzalez@UGent.be*

Abstract

Ion exchange (IX) is a key technology in resource recovery for demineralization and fit-for-purpose water production thanks to its ion-selective recovery features. A major bottleneck in the optimization of the IX process is the accurate prediction of ion breakthrough times, which has the potential to save on regeneration chemicals by maximizing resin utilization. The models used to predict ion breakthrough times are often unreliable due to poor calibration methods and high uncertainty in parameter estimates. A well-calibrated model for ion breakthrough prediction can provide important insights into the process and enable optimization and model-based control with the goal of improving the overall efficiency and sustainability of the process. Therefore, we performed local and global sensitivity analyses to identify the design and operational parameters that contribute most to the prediction of breakthrough curves. The global sensitivity analysis allowed us to select a small subset of parameters for calibration, which showed that only two parameters, namely the maximum adsorption capacity isotherm parameter and the resin particle size, need to be thoroughly calibrated to obtain an accurate prediction of the breakthrough curve. We also showed that uncertainty quantification for model calibration is important to establish the reliability of the predictions. Validation of the model was carried out using experimental data. Hence, we propose a sound calibration procedure, based on good modeling practice, that encompasses both sensitivity and uncertainty analyses and provides a basis for the optimization of the IX process with the aim of improving the accuracy of breakthrough prediction.

Keywords

Calibration protocol, fixed-bed column, global sensitivity analysis, ion exchange, uncertainty analysis

Highlights

1. Model parameters most affecting breakthrough are revealed by sensitivity analysis.
2. Breakthrough prediction improved by calibrating resin capacity and particle size.
3. Proper calibration of parameter subsets is reliant on uncertainty of the estimates.
4. Uncertainty quantification provides reliability bands for breakthrough prediction.
5. Calibration protocol is applicable to other physico-chemical adsorption processes.

Abbreviations

AIC	Akaike Information Criterion
CI	Confidence Interval
FMI	Fisher Information Matrix
GSA	Global Sensitivity Analysis
IX	Ion Exchange
LSA	Local Sensitivity Analysis
ODE	Ordinary Differential Equation
PDE	Partial Differential Equation
RMSE	Root Mean Square Error
WSSE	Weighted Sum of Square Error

1. Introduction

The global issue of water scarcity is placing increasing pressure on the supply and management of water resources. As a result, resource recovery has emerged as a strategy to increase the circularity of resources and water supplies. Ion exchange (IX) is a key technology in this endeavor, as it enables fit-for-purpose water demineralization through its ion-selective recovery properties (Kabdaşlı and Tünay, 2018; Jegatheesan et al., 2021; Taghvaie Nakhjiri et al., 2022). IX, like other adsorption processes, is typically operated as a continuous process in fixed-bed packed columns and used as an end-of-pipe removal treatment (Inglezakis and Zorpas, 2012). Notable applications include the recovery of nutrients and by-products in the food production industry (Kammerer et al., 2011); the removal of heavy metals such as nickel from urban wastewater (Ma et al., 2019), petrochemical wastewater (Cechinel et al., 2018) and other industrial waters such as chromium-rich textile waste (Wang et al., 2015); the recovery of precious metals from industrial wastewater (Taghvaie Nakhjiri et al., 2022); the removal of silica for the production of ultrapure water in various industries, including paper mills and electronics (Chen et al., 2022); or the removal of pharmaceuticals (Chu and Hashim, 2023a) and micropollutants such as PFAS in drinking water production (Smith et al., 2023).

The challenges of the IX technology include the optimization of the cost of chemicals associated with resin regeneration and the disposal of generated waste streams (Crittenden et al., 2012), as well as the accurate prediction of target ion breakthrough times for downstream compliance during operation and control (Inglezakis and Zorpas, 2012). Consequently, the optimization of IX processes would allow for a reduction in the energetic and material impacts of water treatments, thereby enhancing their environmental and economic sustainability and further increasing their potential for resource recovery.

A number of mechanistic and empirical models have been developed with the objective of improving the understanding of the IX process and addressing the aforementioned challenges by identifying the breakthrough time. A myriad of IX models can be found in the literature, including transport and dispersion mechanisms derived from conservation laws. These include surface diffusion (Ma et al., 2019), pore diffusion (Zhang et al., 2015), and other intraparticle diffusion kinetic models (Wang and Guo, 2022). Similarly, numerous equilibrium isotherm models have been developed beyond the classical Langmuir and Freundlich models. A comprehensive review of the most commonly used models can be found in LeVan and Carta (2008), and more recently in Wang and Guo (2023), (Wang et al., 2024).

Therefore, in view of the considerable number of existing models, no unified modeling procedure exists, and as a result, a standardized calibration protocol for adsorption models is still missing. This is evidenced by the numerous existing reports on inconsistencies, mistakes, and misconceptions in the modeling of the IX process (Chu, 2023; Haupert et al., 2021; Hu et al., 2021; Lima et al., 2021; Mudhoo and Pittman, 2023; Tran et al., 2017; Xiao et al., 2018), which result in models with limited reliability and prediction power. We identified the following five misconceptions as common in modeling tasks: (i) overfitting, (ii) absence of uncertainty quantification, (iii) use of linearized models, (iv) misconceptions regarding sensitivity analyses, which are performed locally and for single parameters, and (v) confusion with the uncertainty analysis.

In order to achieve reliable and powerful prediction capabilities and establish good modeling practice for the calibration and reproducibility of IX models, we followed general recommendations by Saltelli et al. (2019) regarding sensitivity and uncertainty analysis, which are valid across disciplines. Similar guidelines have been successfully established for the calibration of diverse processes in the wastewater field (Rieger et al., 2012; Vanrolleghem et al., 2003). In a recent study, Chu and Hashim (2023b) employed rigorous model selection techniques to evaluate the performance of competing models with varying numbers of fitting parameters. Nevertheless, essential considerations in the calibration process, such as verification of the model structure, a detailed description and uncertainty quantification of the parameter estimation procedure, and an analysis of the data quality and resulting performance, are still absent in numerous studies. Our work is based on the model proposed by Zhang et al. (2015), which considered an advection-diffusion-reaction model for fixed-bed ion exchange columns. This allowed us to address a few of the aforementioned limitations and to propose a framework for model calibration based on good modeling practice.

A review of the above articles from literature essentially shows that in IX, models can be effectively used to predict ion breakthrough in a multitude of applications. However, the calibration methodologies employed by various authors exhibit inconsistencies in rigor resulting in overfitting, limited reproducibility and high uncertainty regarding the predictions of these models. Consequently, these models frequently fail when subjected to validation or extrapolation to future time series, due to the considerable uncertainty associated with the calibrated parameters. The limited extrapolation capability, coupled with the complexity and non-linear nature of the operation, restricts the utility of these models as a monitoring and control tool for water treatment applications. A standardized model calibration procedure is currently missing.

This work is accompanied by open-access code (https://github.com/UGentBiomath/IX-GMP) and presents a framework for model calibration based on good modeling practice. The work includes a description of a simple one-component IX model for fixed-bed operation and employs sensitivity and uncertainty analyses to verify the model structure and facilitate rigorous calibration of an identifiable parameter set. The resulting protocol is applicable to other IX modeling studies and allows for scenario analysis and optimization of the system under study.

2. Materials and Methods

2.1 Model definition

We implemented a dynamic model describing the transport of ions through a fixed-bed IX column reactor. We selected a two one-dimensional partial differential equation (PDE) model by Zhang et al. (2015) in order to develop our model. The first PDE describes the liquid phase, while the second PDE describes the solid diffusion. As the goal of this modeling study is to optimize the process, computational efficiency is a significant factor. Consequently, we simplified the model with the following modifications: the simplified model considers the same one-dimensional set of equations, comprising an advection-dispersion-reaction partial differential equation (PDE) for the liquid phase, but proposes a linear driving force mass transfer ordinary differential equation (ODE) for the solid diffusion, as proposed by LeVan and Carta (2008). Therefore, we replaced the second PDE with an ODE.

Both models compute the concentration profile of ions in the liquid and solid phases along the column, thereby predicting the breakthrough of these ions following saturation of the resin. A visual inspection of Figure 8 shows that our simplified model using parameters from Zhang et al. (2015) (designated as “uncalibrated” in the figure) does indeed produce the anticipated breakthrough curve of the data presented by Zhang et al. (2015). Therefore, the simplified model can be employed as a fast, real-time optimization and control tool to improve the efficiency of IX operations. This is made possible by the fast concurrent computation of both scales of the process, namely the ion concentration in the liquid and solid phases, which are respectively referred to as the macro and micro-scale by Zhang et al. (2015). Subsequently, both scales are translated into a single output, namely the ion transport in the fluid along the column. Similarly, the solid phase can be analyzed for the purposes of regeneration, although this is beyond the scope of the present work.

2.1.1 Model structure

Figure 1 depicts the IX process as a fixed-bed resin-packed column, illustrating the main variables involved in the dynamic transport of ions along the column and through the liquid and solid phases.

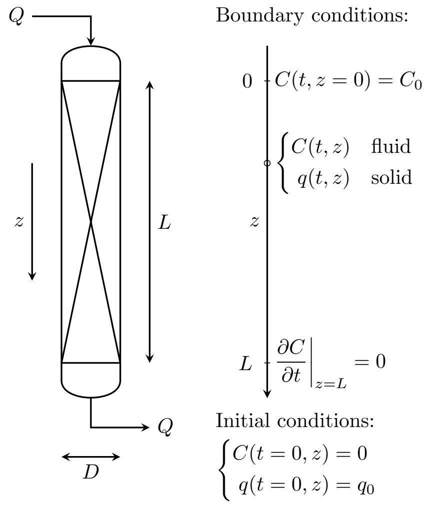

Figure 1. Schematic representation of a typical fixed-bed IX column and description of the main variables of the 1-D model with appropriate initial and boundary conditions. The parameters are detailed in Table 1.

The evolution of the concentration profiles of a given ion in the column can be obtained by performing a mass balance. This results in the PDE (1), which must be solved after the appropriate initial and boundary conditions have been defined. Furthermore, the solution involves the discretization of the space and time domains:

$$\frac{\partial C}{\partial t} = D_z \frac{\partial^2 C}{\partial z^2} - u \frac{\partial C}{\partial z} + \frac{\rho_b}{\varepsilon} \frac{\partial q}{\partial t} \tag{1}$$

In the above expression, C represents the concentration of ions present in the liquid phase at any given point. The variable q denotes the concentration of ions in the solid phase. The term u is the fluid velocity, which is assumed to be constant. D_z is the axial liquid dispersion coefficient. ρ_b is the bulk bed density, while ε denotes the bed porosity. In this context, the independent variables t and z are used to refer to time and column height, respectively. Furthermore, the following assumptions have been made: a uniform velocity profile and intraparticle diffusion in the column; and a controlling equilibrium between the liquid and solid phases (negligible resistance to mass transfer in the boundary layer, as indicated by Zhang et al., 2015). The velocity can be calculated by assuming a constant inlet flow and a constant void bed cross-section in the column, according to the following equation: $u = \frac{Q}{A\varepsilon}$.

The last term in Equation (1) denotes the change in ion concentration due to adsorption or desorption in the solid phase. Assuming a linear driving force between the liquid and solid phases (LeVan and Carta, 2008), the second equation of the model can be expressed as an ordinary differential equation at each discretized point within the spatial domain:

$$\frac{\partial q}{\partial t} = k_{ldf}(q^* - q)$$

(2)

where k_{ldf} is the mass transfer coefficient and q^* is the concentration in the resin phase surface in equilibrium with the liquid phase concentration, C, which can be calculated using an equilibrium isotherm model. Equations (1) and (2) must be integrated and solved concurrently.

The two most frequently utilized equilibrium isotherm models in the literature are the Freundlich and Langmuir models (O'Neal and Boyer, 2013). The Freundlich isotherm model (Freundlich, 1907) assumes multilayer adsorption described by a power law:

$$q^* = K_f \cdot C^{n_f} = K_f \cdot C^{1/n}$$

(3)

where K_f and n_f are parameters to be calibrated. The Langmuir model (Langmuir, 1918) assumes monolayer adsorption, with q_{max} representing the maximum adsorption capacity, and K_L a second parameter to be calibrated:

$$q^* = q_{max}\frac{K_L C}{1 + K_L C}$$

(4)

In this study, we used the Langmuir isotherm model due to the greater interpretability of its parameters. Nevertheless, the Freundlich model has also been successfully applied to describe the equilibrium between two phases (Sengupta and Pandit, 2011; O'Neal and Boyer, 2013). It should be noted that other, more complex models exist; however, the inclusion of additional parameters complicates the calibration process and is therefore not considered here.

The mass transfer coefficient k_{ldf} in Equation (2) is a function of the intraparticle diffusion coefficient, D_p, and the particle radius, r_p, according to a pore or solid diffusion mechanism (LeVan and Carta, 2008). This can be expressed as follows:

$$k_{ldf} = 15\frac{D_p}{r_p^2}$$

(5)

The solid and liquid phases are related through the bed density given by $\rho_b = \rho_p(1 - \varepsilon)$, where ρ_p is the resin density and ε is the void fraction in the column left by the settled resin, also known as bed porosity.

2.1.2 Model parameters and variables

Table 1 lists the variables considered for the present model structure, classified according to their function: output variable, input variable, or parameter. The latter can be further subdivided into the following categories: design parameters, which include column dimensions and the physicochemical properties of the resin; operational parameters; and equilibrium parameters. The last column indicates the source of the values: whether they were measured experimentally, calculated numerically, fixed by the process or equipment specifications related to design decisions and reported in manufacturers' data sheets, or obtained from the literature. Parameters for which there is no direct measurement or other means of obtaining a value are estimated from output measurements in the calibration process. For each parameter, the values reported in the literature are listed in Table 2, together with the ranges and sources. Even those parameters which have been assumed to be constant for design reasons (manufacturer's data sheets, operating rules), are considered uncertain for the purposes of the sensitivity analysis, in order to provide a comprehensive overview of potential model uncertainties.

The liquid phase concentration is frequently determined at the inlet (C_0) and outlet (C) of the column. Direct measurement of the concentration profile within the column is not feasible without introducing multiple sample extractions, which would affect the total column volume. However, these profiles are predicted by the model. The measurement of solid-phase concentrations is challenging; however, they can be calculated through a mass balance by means of Equation (2). The initial solid phase concentration, q_0, is assumed to be zero for both fresh resin and regenerated resins where fouling is considered to be negligible. As the original capacity is never fully restored, this initial condition could also be regarded as a variable in cases where regeneration is being investigated. For the purposes of this study, however, it is assumed that q_0 is equal to zero.

Table 1. Variables and parameters considered in the simplified IX model. The symbols are in accordance with the commonly used nomenclature in the literature. For detailed values, see Table 2.

Name	Symbol	Unit	Type	Source
Outputs				
Liquid-phase concentration	C	mmol/L		Computed by Eq. (1)
Solid-phase concentration	q	mmol/g		Computed by Eq. (2)
Inputs				
Inlet liquid-phase concentration	C_0	mmol/L		Directly measured*
Initial solid-phase concentration	q_0	mmol/g		Initial condition in Eq. (2)
Volumetric flow	Q	L/min		Manipulated/fixed*
Parameters				
Bed length	L	m	Design	Directly measured*
Bed diameter	D	m	Design	Column manufacturer*
Resin bead particle size	r_p	m	Design	Resin manufacturer*
Resin bead particle density	ρ_p	g/L	Design	Resin manufacturer*
Bed porosity (void fraction)	ε	–	Operation	Literature*
Axial dispersion coefficient	D_z	m^2/s	Operation	Literature*
Intraparticle diffusion coefficient	D_p	m^2/s	Operation	Literature*
Sorption capacity	q_{max}	mmol/g	Equilibrium	Literature*
Langmuir constant	K_L	L/mmol	Equilibrium	Literature*

*Determination of these parameters is uncertain and therefore could be considered for calibration.

The design parameters include the bed dimensions, which determine the volume available for both resin and effluent in the column, as well as the particle size and the density of the solid phase (inversely proportional to the specific volume occupied by the resin). The effective volume available for the liquid fraction is a function of the bed porosity, which is in turn determined by the compaction of the resin in the column and the swelling induced by the liquid. This parameter has been extensively studied, and a range of variability can be found in the literature depending on the packing shape and size (see Table 2). The actual liquid capacity of the bed is calculated as follows: $V = \varepsilon L A = \varepsilon L \frac{\pi}{4} D^2$. In the literature, the bed volume is typically referred to as a measure of the time elapsed during the process, calculated as the treated effluent volume per unit time divided by the effective bed volume. In addition, dispersion coefficients are frequently encountered in the literature in the form of empirical correlation functions of the Reynolds number and bed configuration. The determination of intraparticle coefficients is an experimental process, and they are therefore considered to be uncertain.

Table 2. Value and range of variability reported in the literature for the parameters listed in Table 1.

	Reported value	Variability range	Source
Design parameters			
Bed diameter, D	0.1 m	(±20%)	Zhang et al., 2015
Bed length, L	10^{-2} m	(±20%)	Zhang et al., 2015
Operational parameters			
Inlet P concentration, C_0	20 mmol/L	(±20%)	Zhang et al., 2015
Flow, Q	3.1 cm/min	(±20%)	Zhang et al., 2015
Porosity, ε	0.37	0.3–0.44 (±20%)	Yoshida et al., 1985
Resin size (diameter), $d_p = 2r_p$	$7.5 \cdot 10^{-4}$ m	$3\text{-}12 \cdot 10^{-4}$ m (±60%)	Sengupta and Pandit, 2011
Resin density, ρ_p	389 g/L	(±20%)	O'Neal and Boyer, 2013
Intraparticle diffusivity, D_p	$5.3 \cdot 10^{-10}$ m^2/s	(±20%)	Sengupta and Pandit, 2011
Maximum capacity, q_{max}	0.291 mmol/g*	(±20%)	O'Neal and Boyer, 2013
Langmuir constant, K_L	1.18 L/mmol*	(±20%)	O'Neal and Boyer, 2013

* Values for synthetic fresh urine.

2.1.3 Model implementation and numerical solution

The model equations have been formulated as a set of one-dimensional differential equations, thereby enabling the model to be employed as a fast optimization tool for the IX process. It is assumed that the velocity profiles of the fluid across the column are uniform and that there is no significant existence of wide preferential channels due to the low diameter-to-length ratio (D/L) of the column.

In order to predict the one-dimensional spatial variation of pollutant concentration along the IX column, a discretization of the column length was considered. The solution of the nonlinear system of equations can be computationally intensive. Therefore, a discretization error analysis was conducted (see Appendix for a detailed discussion) and an appropriate discretization step of 100 grid points and a time step of 0.1 seconds was selected to achieve sufficient accuracy without substantial computational effort.

The model was implemented in Python 3 (Van Rossum and Drake, 2009) and makes use of available scientific packages *Numpy* (Harris et al., 2020), *Scipy* (Virtanen et al., 2020), *Pandas* (McKinney, 2010; The pandas development team, 2020), and *Matplotlib* (Hunter, 2007). The code will be made available (upon publication) at https://github.com/UGentBiomath/IX-GMP.

The equations were solved using the finite difference method and the method of lines for the spatial and temporal discretization of the PDE with the *scikit-finite-diff* package (Cellier and Ruyer-Quil, 2019). Figure 2a illustrates the concentration profiles along the length of the column as a function of time in relation to the inlet concentration, C_0. The arrow indicates the typical temporal evolution of the profiles resulting from the transport of ions along the column and into the solid phase. The breakthrough curve (Figure 2b) reflects the evolving ion concentration in the effluent at the column outlet, illustrating the gradual depletion of the resin and the breakthrough of the ion, which is adsorbed onto the solid phase until the resin is completely saturated.

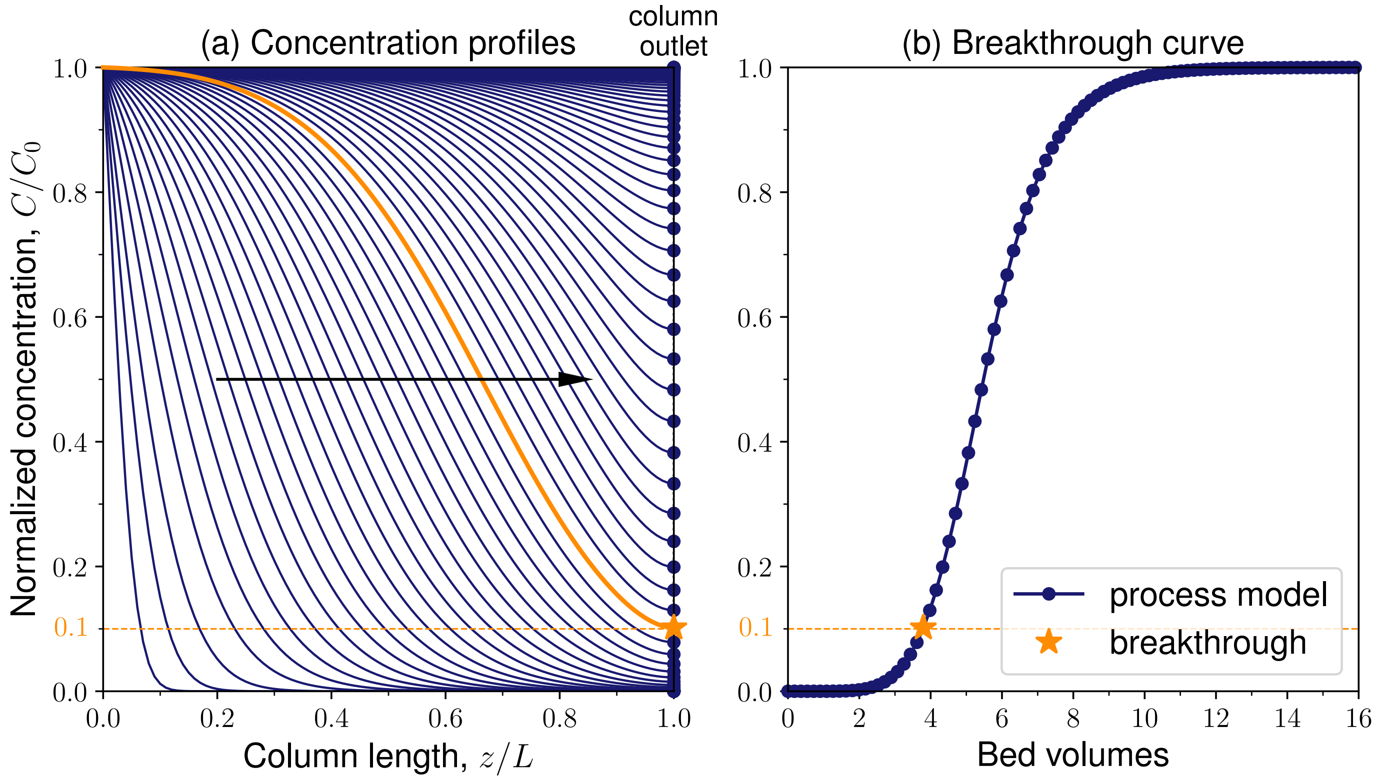

Figure 2. Concentration profiles (a) and breakthrough curve (b) obtained by solving the model equations (1) to (4) with the appropriate initial and boundary conditions as specified by the scheme in Figure 1. The breakthrough time (highlighted) is defined as the time for 10% of initial concentration in the outlet.

2.2 Sensitivity analysis

The objective of the sensitivity analysis is to identify those parameters that exert the greatest influence on model output, with the aim of reducing the variability or uncertainty in the model output by accurately determining their values. By identifying which parameters are sensitive, it is possible to reduce the cost of experiments by focusing efforts on measuring those parameters (Saltelli et al., 2007).

2.2.1 Local sensitivity analysis

A local sensitivity analysis (LSA) quantifies which model parameters have a greater influence on the model output(s) for a given value of each parameter (Saltelli et al., 2007). The sensitivity function is defined as the change in the model output resulting from a small change or perturbation in the value of a single parameter. A numerical approximation of this, commonly referred to as the derivative method, is the first-order forward finite approximation:

$$S_{\theta_i}^{j} = \frac{\partial y_j}{\partial \theta_i} \cdot \frac{\theta_i}{y_j} \approx \frac{y_j(\theta_i + \Delta\theta_i) - y_j(\theta_i)}{\Delta\theta_i} \cdot \frac{\theta_i}{y_j} \tag{6}$$

where y_j represents the model output, θ_i denotes the parameter subject to variation, and Δ_i is the perturbation value. In the present study, a perturbation value of $10^{-5} \cdot \theta$ was used for all parameters. The sensitivity function is normalized to accommodate for disparate scales or magnitudes for both the output and the parameters. It is important to note that the results of a local sensitivity analysis may vary significantly depending on the specific location within the parameter space under investigation. Hence, the analysis may be repeated at different locations to obtain further information, or a global sensitivity analysis may be performed (see section Global sensitivity analysis).

The local sensitivity analysis is employed as an initial model check on the parameters' behavior, requiring minimal computational efforts and offering insight into the extent to which varying parameters influence the model output. Furthermore, it identifies potential correlations between parameters and operational regions where parameter sensitivity is higher and correlation minimal. These regions are of particular interest for experimental data collection.

2.2.2 Global sensitivity analysis

A global sensitivity analysis (GSA) aims to quantify the relative importance of parameters in determining the variability of a model output over a wide range of parameter values (Saltelli et al., 2007). In contrast to a local analysis, a systematic exploration of the parameter space can provide a more comprehensive understanding of the overall influence of the different parameters on the model output. To this end, sensitivity indices can be calculated in order to identify these effects.

First-order indices describe the main effect of parameters on the variance of model outputs. They can be calculated as the variability in model output resulting from the variation of a single parameter in isolation, relative to the total variability attributable to changes in all parameters, expressed as follows:

$$S_i = \frac{V[E(Y|X_i)]}{V(Y)} \tag{7}$$

where V represents the variability of model output Y, E denotes the expected value, and X_i is a specific parameter value. In contrast, total effects comprise the cumulative impact of a parameter on the model output, accounting for both first-order and higher-order effects stemming from interactions or non-linearities between parameters. A disparity between total and first-order effects indicates the existence of higher-order interactions:

$$S_{T_i} = 1 - \frac{V[E(Y|X_{\sim i})]}{V(Y)} \tag{8}$$

The GSA is calculated at a specific time point. In the present study, the time at which 10% breakthrough is observed was selected as the most informative for GSA. The sensitivity indices for the 10% breakthrough concentration with respect to the model parameters listed in Table 2 were calculated by Sobol uniform sampling using the *SALib* sensitivity analysis library (Herman and Usher, 2017). Furthermore, second-order sensitivity indices were calculated and employed to ascertain the existing correlation between parameters. The Sobol sampling of parameters consisted of 5,632 uniform values, selected within the ranges of variability outlined in Table 2. Subsequently, a Monte Carlo-type of simulation was employed to ascertain the variability in model output derived from the simulation results corresponding to each parameter subset sampling. The primary disadvantage of GSA is its high computational cost, which is a consequence of the substantial number of model evaluations required.

One advantage of an LSA over a GSA is that it can facilitate the preliminary screening and reduction of the parameter set prior to a more comprehensive global analysis.

Once the most sensitive model parameters have been identified, they can be estimated through the process known as model calibration. In order to achieve this, it is necessary to obtain experimental data.

2.3 Experimental data for calibration and validation

The experimental data used for the calibration of our model are derived from measurements of breakthrough curves for ion exchange of phosphorus recovery from fresh urine, as documented by O'Neal and Boyer (2015). The data set describes the breakthrough of phosphate ions from synthetic fresh urine as it traverses a fixed-bed column containing a specific phosphate-selective HAIX-Fe resin. Further details regarding the column tests can be found in the referred work.

The authors acknowledge a limitation in the data set, namely the absence of measurements of sulphate and other ions present in the urine effluent. These ions compete with phosphates in their adsorption onto the resin; however, the analysis of competition and affinity effects is beyond the scope of the present work and not captured by the model. Therefore, the influence of these ions on the predictive power of the model will not be addressed. The required parameters for the modeling of this system are enumerated in Table 2, accompanied by the sources from which they were derived. As illustrated by Figure 8, the predicted breakthrough curve exhibits a high degree of agreement with the measured data. Further details can be found in the Results and Discussion section.

2.4 Model calibration

The goal of model calibration is to identify the optimal set of parameter values that fit the experimental data, which can be considered as an optimization problem. The quality of the fit is quantified by an objective function, which is then minimized. The weighted sum of squared errors (WSSE) was selected as the objective function for model calibration:

$$J(\theta) = \sum_{\mathrm{i}=1}^{N} (\hat{y}_i(\theta) - y_i)^T W (\hat{y}_i(\theta) - y_i) \tag{9}$$

where $\hat{y}_i(\theta)$ represents the model prediction of the output y_i corresponding to the parameter set of values θ, y_i denotes the measurement of the output y, W is the square matrix of weights associated with different outputs or time moments, and N is the number of measurements. In this analysis, we assumed W to be the identity matrix, thereby assigning equal weight to each time point of the single measured model output.

There are numerous techniques for minimizing the objective function, including the well-known Nelder-Mead simplex method (Nelder and Mead, 1965). The Levenberg-Marquardt method (Levenberg, 1944; Marquardt, 1963) was selected as a robust and efficient method well suited for unconstrained nonlinear least squares fitting problems. Further details regarding the implementation can be found in Gavin (2019). A recent overview of calibration methods for computer simulation is provided by Sung and Tuo (2024).

Unconstrained optimization methods may suggest values for the parameters that are outside of their range of physical validity. Consequently, we also employed a robust constrained trust region method (Branch et al., 1999) to evaluate the performance of constrained parameter subsets in comparison to that of the unconstrained algorithm. Both constrained and unconstrained methods are readily available in the optimization library of the *Scipy* Python package (Virtanen et al., 2020). For the constrained optimization method, the parameter values were bounded by their ranges of validity. The initial parameter estimates, along with their respective ranges of variability, are presented in Table 2. The relative tolerance, which serves as a stopping criterion for the iterative optimization algorithm, was set to 10^{-4} in order to reduce the number of model evaluations. This value was selected following a verification process, during which it was determined that it produced results that were not significantly different from those obtained with lower tolerance. Furthermore, the scale or magnitude of each parameter was provided to the algorithm, thereby reducing the number of iterations. The accuracy of the solution was evaluated by comparing the model output to the experimental data provided by Zhang et al. (2015).

2.5 Uncertainty of the parameter estimation

The estimated parameters have an associated uncertainty due to the presence of noise in the experimental data set used for calibration. Accordingly, parameter estimates are only meaningful when accompanied by their confidence region, within which the actual true value is situated at a specified confidence level (Donckels, 2009). The confidence region can be based on an approximation of the contour of the WSSE objective function, as this provides a measure of the fit to the experimental data. In the case of linear models, the contour is exact. However, for nonlinear models, it is common practice to employ a linear approximation of the parameter estimation covariance matrix, providing a lower bound for this region (Marsili-Libelli et al., 2003):

$$\{\theta : WSSE(\theta) \leq c \cdot WSSE(\hat{\theta})\} \tag{10}$$

A linear approximation results in the Fisher Information Matrix (FMI), whose inverse is employed for the approximation of the error covariance matrix, can be computed as follows (Dochain and Vanrolleghem, 2001):

$$FMI = \sum_{i=1}^{N} S_\theta' \cdot W^{-1} \cdot S_\theta = \sum_{i=1}^{N} \left(\frac{\partial \hat{y}}{\partial \theta}(t_i)\right)' \cdot W^{-1} \cdot \left(\frac{\partial \hat{y}}{\partial \theta}(t_i)\right) \tag{11}$$

where W is the covariance matrix of the measurement errors associated with the measured variables, and S_θ denotes the parameter sensitivity matrix of all outputs with respect to each parameter, as calculated by Equation (5) for N_p parameters and N experimental time points. The error covariance matrix can be used to construct a confidence region for the parameter estimates, δ_i, with a specified level of confidence 1-α (Marsili-Libelli et al., 2003):

$$\delta_i = t_{N-N_p}^{\alpha/2} \cdot \sqrt{\sigma_{i,i}^2}$$

(12)

where t represents the $\alpha/2$ quartile of the Student's t distribution for a given confidence level α and $N - N_p$ degrees of freedom, with N denoting the number of data points, N_p the number of estimated parameters, and $\sigma_{i,i}^2$ the variance of parameter i taken from the error covariance matrix. The diagonals of the covariance matrix thus provide the variances of the errors in the parameter estimates, whereas the off-diagonal elements are the covariances of the errors and offer a measure of the correlation between the different parameters. The linear correlation between two parameters can be estimated based on the following equation:

$$r_{i,j} = \sqrt{\frac{\sigma_{i,j}}{\sigma_{i,i}^2 \cdot \sigma_{j,j}^2}}$$

(13)

The linear correlation is approximately -1 or 1 for pairs of parameters with a high degree of correlation, whereas a value of approximately zero indicates a low correlation.

Once the confidence region for the parameter values has been established, the Monte Carlo technique can be employed to sample the parameters within the region and thereby obtain the expected variability in the model output. A normal distribution was assumed for all parameters, and a total of 640 values was sampled for use in the Monte Carlo simulations. The plotting of the most frequent values can assist in establishing an uncertainty band around the mean value, thereby providing a confidence interval band for prediction of the breakthrough curve.

2.6 Fitness comparison

The root mean square error (RMSE) was employed as a metric for assessing the goodness of fit, or the discrepancy between the experimental data and the calculated breakthrough concentrations. This was due to the fact that the RMSE represents a more accurate average of the distance between the data and the model prediction than the WSSR of the objective function. The RMSE can be calculated as follows:

$$RMSE = \sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{N}} = \sqrt{\frac{WSSR}{N}}$$

(14)

The RMSE offers a straightforward and readily understandable representation of the overall model's error, using the same units as the measured variable, even for unitless comparisons. On the other hand, the RMSE is sensitive to both outliers and overfitting; consequently, it diminishes when additional parameters are incorporated into the model. In order to facilitate a comparative analysis of model fitness for different numbers of calibrated parameters, it is possible to consider criteria that balance the goodness of fit to experimental data with the number of model parameters. This approach allows to penalize overfitting when increasing parameters are incorporated into the model structure. One such established method is Akaike's Information Criterion or AIC (Akaike, 1974):

$$AIC = N \cdot \log\left(\frac{WSSR}{N}\right) + 2 \cdot N_p \tag{15}$$

In Equation (15), the first term will decrease for overparametrized candidate models due to overfitting, while the second term will penalize the added complexity when more parameters are considered. Lower values of AIC are indicative of a superior model, while lower values of RMSE indicate overfitting.

2.7 Model validation

To confirm the predictive power of the calibrated model, a new data set was selected from the same column breakthrough experiments as the calibration tests (see Section 2.3 for a description of the data). For calibration, the parameter q_0 in our model was assumed to be zero, corresponding to fresh resin. The second cycle was chosen to validate the model and therefore the measured data correspond to a regenerated column. Hence, we assumed a regeneration efficiency of 95% for the resin, which was calculated as the amount of phosphorus recovered from the total adsorbed in the column, and the remaining amount was assumed to be retained or fouled in the resin, as also reported by the authors (O'Neal and Boyer, 2015; Zhang et al., 2015). From the adsorption data, fouling is estimated at 1.7 mg out of the total 31.9 mg adsorbed, resulting in a 5.3% fouled amount or 94.7% regeneration efficiency. We then assumed an initial concentration profile in the resin, parameterized by q_0 in our model.

3. Results and Discussion

3.1 Sensitivity analysis

3.1.1 Local sensitivity analysis

The specific set of model parameters to be analyzed is presented in Table 2. The parameter q_0 was omitted from the list, as it is assumed to be equal to zero in this study and therefore its sensitivity to changes would be insignificant. Figure 3(a) shows the variation over time, expressed in bed volumes, of the relative sensitivity values for the breakthrough concentration with respect to all parameters in the selected subset, with the most sensitive parameters highlighted. The less sensitive parameter D_p is also highlighted as a reference. As can be seen, the sensitivity to changes in parameter values is practically zero at the start of the column operation and increases exponentially when the first breakthrough is detected, around two bed volumes. A peak in sensitivity is reached for all parameters around four to six bed volumes and decreases again exponentially until the end of the operation when the resin is saturated. The occurrence of peaks for all parameters around the same time is a clear indication of the correlation between all parameters, which shows that unique determination of values is complicated when all parameters are estimated simultaneously. Since some sensitivities appear to change sign over time, an averaged sensitivity allows to compare the overall trend. Figure 3(b) shows the averaged local sensitivity values over 16 bed volumes for all parameters with their corresponding signs.

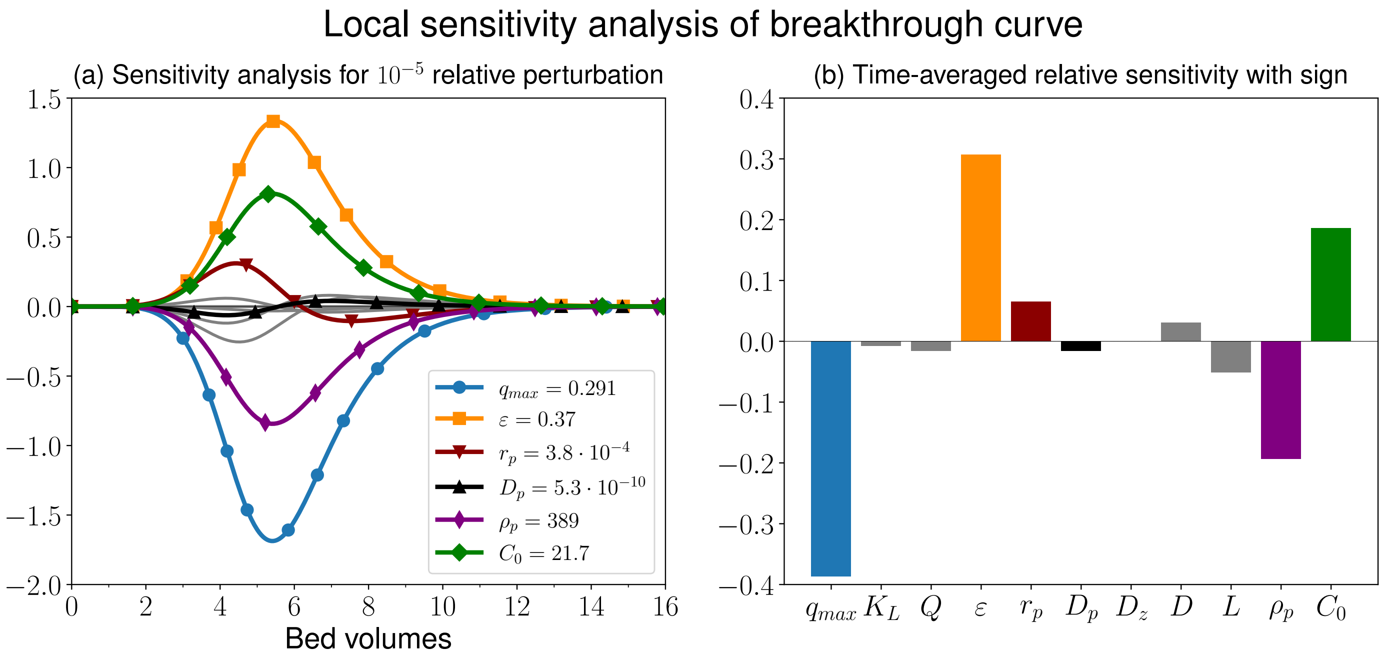

Figure 3. Local sensitivity analysis (LSA) of the breakthrough curve for the parameters listed in Table 1: time variation in bed volumes for a 0.001% relative perturbation (a), and time average with sign (b). The most sensitive parameters (q_{max}, ε, ρ_p, C_0, r_p, D_p) are highlighted, with r_p, D_p inversely correlated.

In descending order of sensitivity, the most sensitive parameters are q_{max}, ε, ρ_p, and C_0. Most of these parameters exhibit positive or negative sensitivity to varying degrees, indicating that they affect the model output in a consistent manner but with differing magnitudes. However, parameters r_p and D_p exert an inverse influence on breakthrough, as evidenced by their inverse correlation, illustrated in Figure 3(a) and in accordance with Equation (5). The particle size is more than twice as sensitive as the intraparticle diffusivity due to the quadratic influence, with the greatest effect observed around four bed volumes, where both parameters are most sensitive.

Changes in sensitivities are also indicative of nonlinearity. In order to illustrate the impact of varying sensitivities, it is possible to plot the effect of a change in different parameters on the model output. Figure 4 shows the impact of a 10% increase in four of the most sensitive parameters and its effect on the breakthrough concentration. An increase in a parameter with negative sensitivity (e.g., q_{max}) will result in a delay in breakthrough, as it has a negative influence on the concentration. Conversely, a positive sensitivity (e.g., with respect to ε or C_0) will result in an earlier occurrence of breakthrough for an increase in the parameter value. Figure 4 also demonstrates that C_0 is a less sensitive parameter than ε, indicating that a 10% increase in the latter will have a more pronounced effect on breakthrough. An analogous analysis may be performed for the remaining, less sensitive parameters. As can be observed, a change in the less sensitive parameter r_p is most pronounced around 4 bed volumes, where the parameter exhibits the greatest sensitivity to changes, as illustrated in Figure 3. However, the overall effect of this one parameter is minor in comparison to an increase in the adsorption capacity (q_{max}), as demonstrated by Figure 3: an increase in the bed height (parameter L) would result in greater resin availability too, which in turn delays breakthrough. However, this increase has overall less effect than a comparable change in the adsorption capacity. This serves to illustrate once more the nonlinear nature of the process and the importance of taking into account these effects for the estimation of parameters.

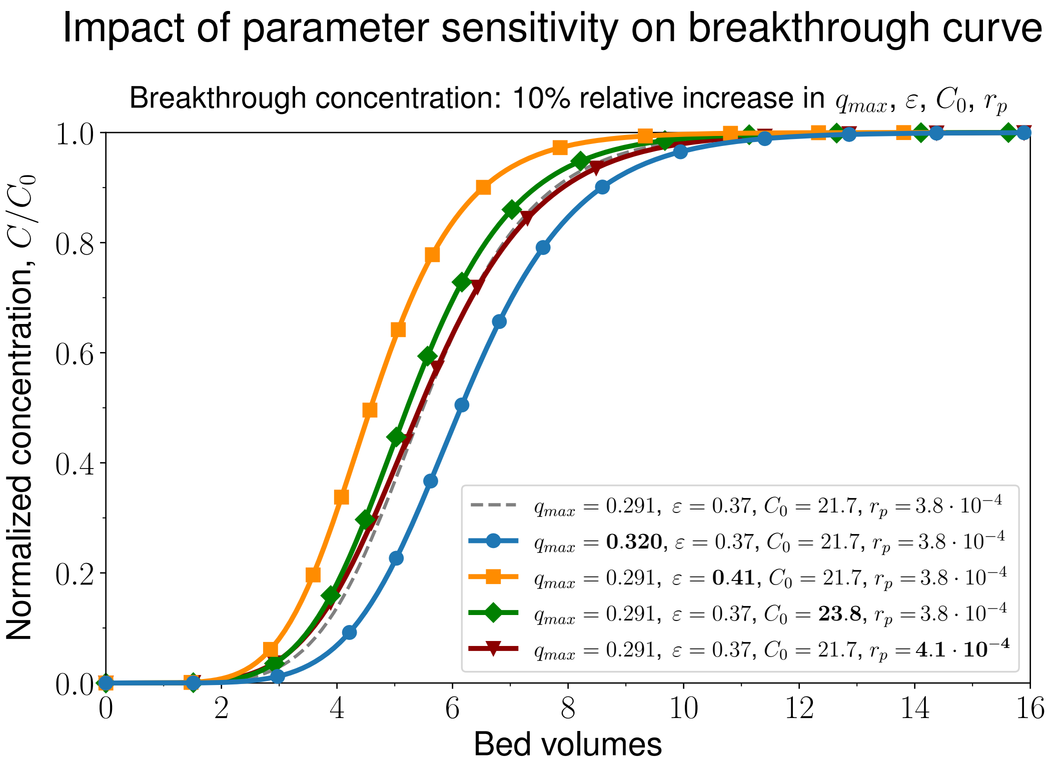

Figure 4. Effect of a 10% increase in the most sensitive parameters (q_{max}, ε, C_0 and r_p) on the breakthrough curve: the adsorption capacity (q_{max}) has a marked negative impact on the breakthrough concentration, while the other three parameters have a positive effect, albeit with varying magnitudes.

Following Figure 3 and Figure 4, we can establish a local ranking of the parameter sensitivities for the model structure considered. It should be noted, however, that this is only a local ranking and should therefore be interpreted and used with care. In the case of nonlinear models, parameter sensitivities may vary considerably when evaluated in different regions of the parameter space. Accordingly, a global sensitivity analysis is conducted in the following section to ascertain an overall parameter ranking for the model calibration. The results of the LSA serve two distinct purposes. First and foremost, the LSA results offer insights into the parameter sensitivities as a function of time, thereby enabling the identification of an optimal time point for conducting a GSA analysis. Based on the LSA analysis, the time at which a 10% breakthrough occurs is identified as an informative time instant to perform the GSA.

Secondly, the LSA results are employed to conduct an initial screening of the parameters, whereby parameters exhibiting low sensitivity can be excluded from a subsequent analysis. As the GSA analysis is considerably more computationally intensive, a preliminary screening through LSA can significantly accelerate the global analysis. The parameters exhibiting the lowest local sensitivity are K_L, Q, D_p and D_z. As the inlet flow rate Q is a variable that can be manipulated, it was decided that it should be retained for a comprehensive analysis in order to ascertain its influence. Furthermore, the isotherm parameter K_L is frequently calibrated in conjunction with q_{max}, and thus it was also determined that it should be included in the GSA. However, in the event of computational constraints, this category of parameters may be excluded from the subsequent analysis, given their negligible impact. It is therefore possible to assign a fixed value within their range of variability to those parameters that have little influence on the output and produce little variance, without compromising the estimation process for the remaining parameters.

A sensitivity analysis may also serve to identify the experimental conditions that will yield the most accurate parameter estimation by maximizing the effect on the model output, which is the goal of *optimal experimental design* or OED (Donckels, 2009). The objective may be to generate experimental data with high information content, thereby further reducing the uncertainty of the parameter estimates. The LSA analysis conducted in the present study indicates that experiments where the breakthrough concentration surpasses at least 50% of the initial concentration represent the most informative experiments, exhibiting the highest parameter sensitivity. Nevertheless, the optimization of the experimental design for parameter estimation falls outside of the scope of the present work.

3.1.2 Global sensitivity analysis

Selection of parameter subset for GSA

Based on the LSA results, two parameters, as detailed in Table 2, were excluded from the GSA: q_0, which was assumed to be zero in this study, and D_z, whose quasi-zero sensitivity corresponds to the common plug-flow assumption and the reason why the diffusion term in Equation (1) is often neglected in practice. Figure 5 illustrates the Sobol sensitivity indices for 10% breakthrough concentrations across the ten selected parameters. Both the total and first-order sensitivity indices are plotted for each parameter, accompanied by uncertainty bands. First-order sensitivity indices are analogous to local sensitivities, but they are calculated over the entire parameter space. Two parameters with the largest indices, D and L, are regarded as design parameters, both having a considerable impact on breakthrough, as they determine the total volume of effluent treated and the volume of resin available.

Subsequently, r_p and q_{max} are the parameters that exert the greatest influence on the breakthrough curve. The bed porosity (ε) has a smaller effect than anticipated by LSA when the full range of parameter values is considered. Conversely, the inlet flow rate, Q, has a more pronounced effect on the breakthrough concentration than anticipated by LSA. This is due to an increase in the value of Q reduces the residence time in the column, which in turn leads to a lower adsorption efficiency. The Langmuir constant K_L has a negligible effect on the breakthrough concentration across the entire range of values considered. Consequently, the equilibrium parameter q_{max}, which is considerably more sensitive, should be given precedence in model calibration. This implies that both isotherm parameters should not be estimated simultaneously. Lastly, we selected an arbitrary cut-off value of 0.02 for parameters exhibiting minimal global sensitivity. The density and the intraparticle diffusivity have low indices, indicating that these parameters can be excluded from the calibration process initially.

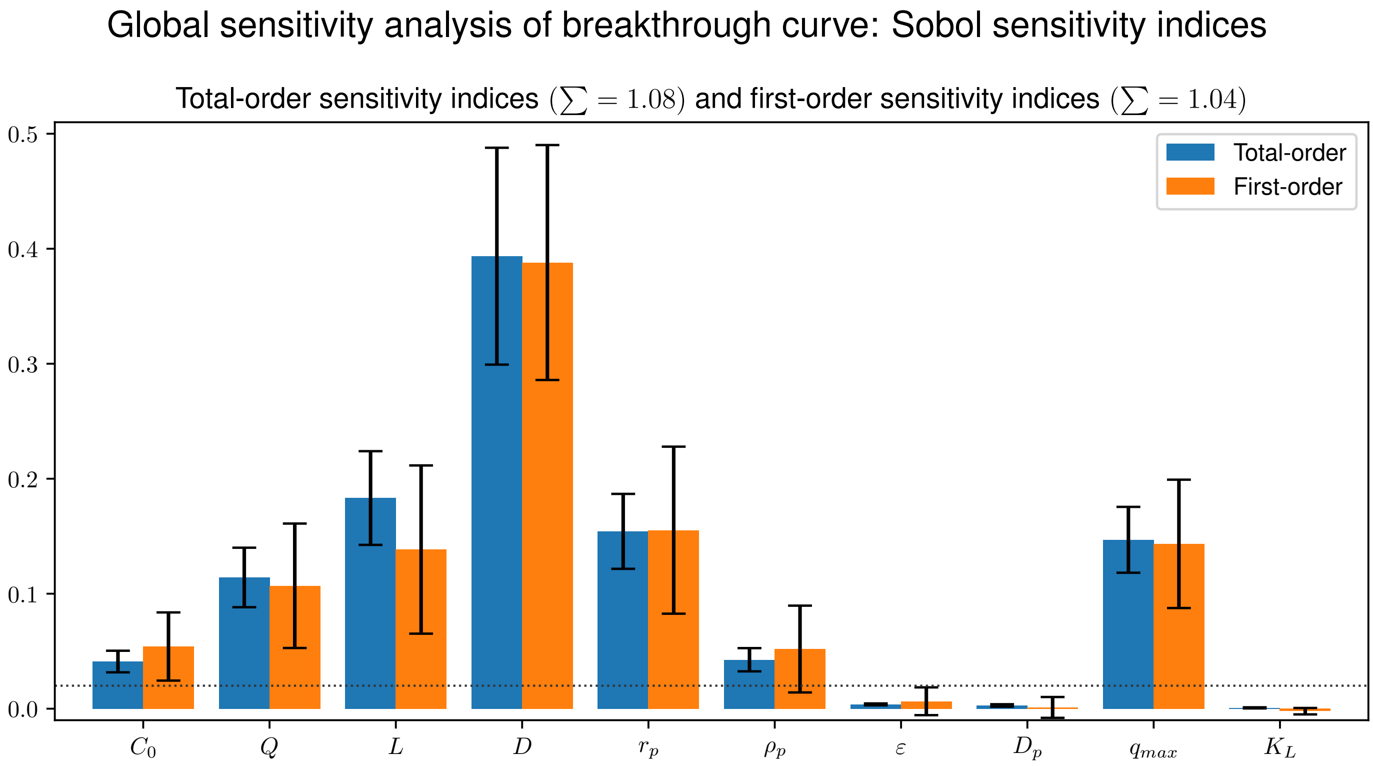

Figure 5. Sobol global sensitivity indices for the parameters listed in Table 2: total-order indices (blue) and first-order indices (orange). The analysis was performed for a breakthrough concentration of 10%.

Figure 5 additionally illustrates the total-order sensitivity indices for each parameter. The observation that the sum of the first-order indices is 1.04, which is slightly different from 1, suggests that this model exhibits some degree of nonlinearity. This is also evidenced by the discrepancy between total-order and first-order effects, which suggests the presence of higher-order interactions or correlations between each pair of parameters. The parameters that exhibit the most significant discrepancy between total and first-order indices indicate the highest degree of interaction with other parameters. The *SAlib* library further allows the calculation of second-order sensitivity indices with supplementary parameter values sampled, thus necessitating additional computational resources and effort. Figure 6 depicts the second-order sensitivity indices, which illustrate the interactions between pairs of parameters. The largest interactions are observed for the isotherm parameters, indicating that the adsorption may be particularly influenced by these parameters.

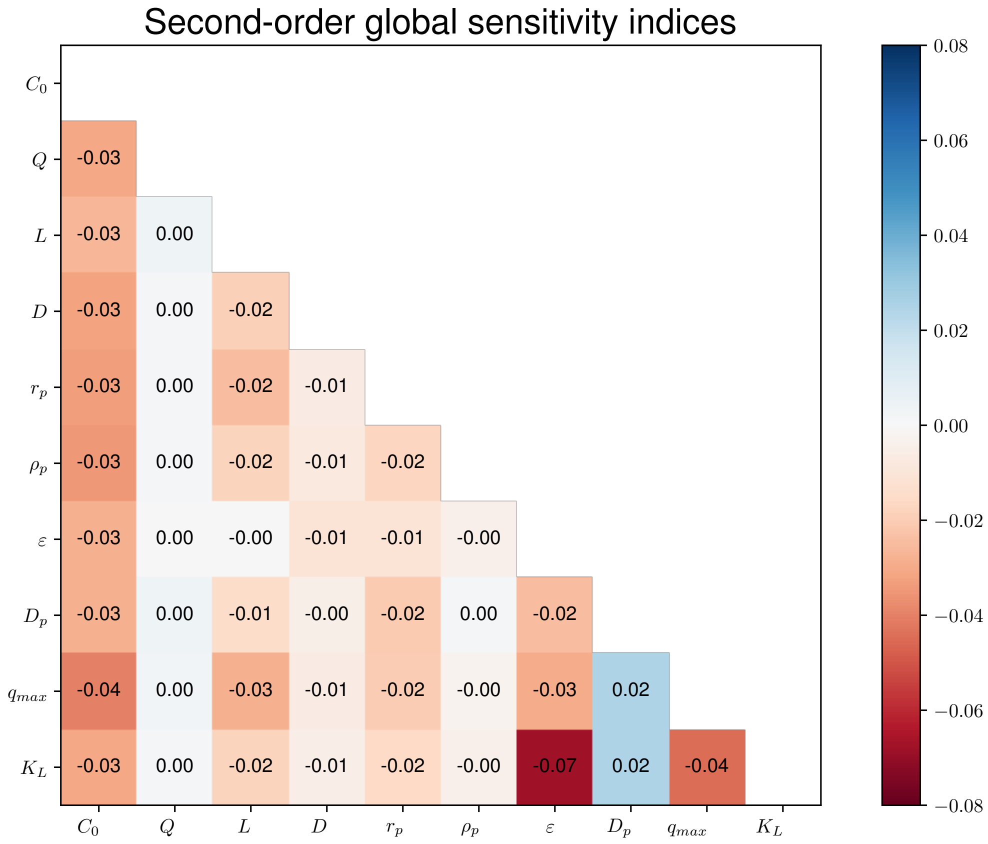

Figure 6. Second-order sensitivity indices. Parameters q_{max}, K_L, ε and C_0 show the largest interactions.

Refined parameter subset for GSA

Based on these results, a more refined analysis can be conducted, focusing exclusively on the most sensitive parameters. The design parameters D and L can now be excluded, given that their values are frequently fixed for existing columns. With regard to the manipulable variable Q, its value is typically fixed for operational reasons. The parameters ρ_p and D_p have overall a negligible influence on the breakthrough concentration and will thus be excluded from further consideration. In this study, the inlet concentration C_0 is maintained at a constant value, despite the fact that it is often variable in the effluent. Accordingly, it will be retained for subsequent analysis to ascertain the impact of its variability on the breakthrough prediction. Furthermore, parameter r_p exhibits the most significant influence on the breakthrough concentration, likely due to the extensive range of validity considered, 60%. In the subsequent analysis, the range was limited to 20% in order to compensate for the parameter's high sensitivity and to remove the distortion caused by this.

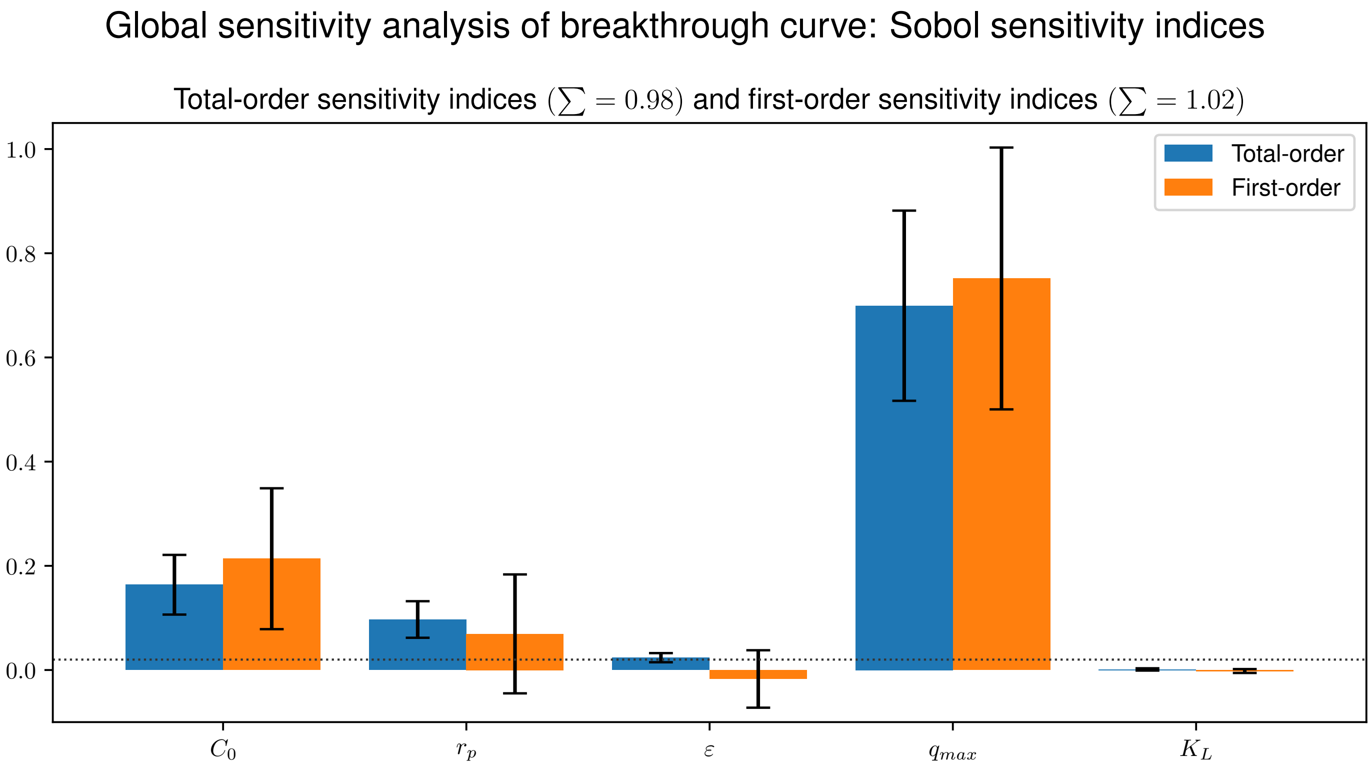

Figure 7. Refined global sensitivity analysis of breakthrough concentration with respect to parameters C_0, ε, r_p, q_{max}, and K_L: total-order and first-order indices. All parameters are subject to 20% variation. The calibration of K_L with q_{max} is common, but the low sensitivity suggests this might not be optimal.

Figure 7 depicts the results of the refined analysis. 448 additional simulations were performed. As can be observed, the particle size r_p no longer demonstrates the greatest sensitivity to changes, having been superseded by the isotherm parameter q_{max}. This illustrates that the range of validity under consideration has a significant influence on the sensitivity ranking and, as a result, has the potential to modify it. Accordingly, if the values of the parameters are constrained to the selected range of 20% variability, the first parameter to be estimated would be q_{max}, as this exerts the greatest influence on the breakthrough concentration. Additionally, it can be observed that both ε and K_L exert a negligible impact on the breakthrough concentration. Accordingly, these two parameters should not be calibrated initially. Moreover, the inlet concentration, C_0, which is assumed to be constant in this study, may have a relatively minor impact on the prediction of breakthrough. Consequently, the particle size, r_p, should also be estimated. These results suggest that the calibration of these two parameters, specifically q_{max} and r_p, should be conducted initially.

3.2 Model calibration

In light of the results of the refined GSA presented in Figure 7, it becomes evident that only a limited subset of parameters will undergo calibration. The final subset will exclude parameter C_0 due to the relatively minor influence on breakthrough and assumed constant. Furthermore, the inclusion of K_L serves to illustrate the fact that this parameter exhibits a markedly low sensitivity, which ultimately precludes the possibility of estimating it with a high degree of accuracy. This is because such an estimation would necessitate considerable changes in its value in order to produce a significant change in the objective function. We identified thus the final subset of parameters to be estimated with the calibration data set defined in section 2.3 and RMSE and AIC as fitness criteria. The results are summarized in Table 3 below.

Table 3. Results of the parameter estimation for up to four parameters in different calibration subsets. The values indicate the estimated parameter values, the corresponding 95% confidence intervals calculated with FIM (relative percentage), RMSE and AIC as fitness criteria. Initial estimates given by θ_0.

N_p [a]	N_f [b]	Min. [c]	q_{max}	r_p	ε	K_L	RMSE	AIC
1	10	LM, TR	0.323 ± 1.4%				0.0262	-143.6
2	15	LM, TR	**0.326 ± 1.4%**	**$4.3 \cdot 10^{-4}$ ± 8.4%**			**0.0207**	**-151.3**
3	**34**	LM	0.276 ± **72.4%**	$4.0 \cdot 10^{-4}$ ± **31.5%**	0.33 ± **43.1%**		**0.0206**	**-149.3**
	35	TR	0.312 ± **87.7%**	$4.2 \cdot 10^{-4}$ ± **36.9%**	0.36 ± **53.2%**		**0.0205**	**-149.4**
4	**52**	LM	**0.430** ± **10^4 %**	$4.9 \cdot 10^{-4}$ ± **34.8%**	**0.44** ± **56.1%**	**10.4** ± **10^7 %**	**0.0206**	**-147.2**
	40	TR	0.312 ± **10^5 %**	$4.2 \cdot 10^{-4}$ ± **46.3%**	0.36 ± **60.7%**	1.25 ± **10^7 %**	**0.0205**	**-147.4**
θ_0 (initial values)			0.291	$3.8 \cdot 10^{-4}$	0.37	1.18	0.0864	
Bounds		LM	-	-	-	-		
		TR	0.2-0.35 (±20%)	$1.5\text{-}6 \cdot 10^{-4}$ (±20%)	0.3-0.44 (±20%)	0.9-1.4 (±20%)		

[a] Number of model parameters in the calibration subset.

[b] Number of function evaluations of the minimization algorithm.

[c] Minimization algorithm. LM: Levenberg-Marquardt (unconstrained); TR: Trust Region method (constrained).

The calibration of only the most sensitive parameter, namely q_{max}, resulted in a 70% improvement in fit (RMSE of 0.0262 vs 0.0864) compared to the initial uncalibrated value of 0.291 (given by θ_0 in Table 3). The 95% confidence interval was estimated at ±1.4%, expressed as a relative percentage of the optimal value. After approximately 10 model evaluations, both minimization methods yielded identical results. Furthermore, the determined value of 0.323 is more closely aligned with the total phosphate loading on the resin of 10.2 mg P/g or 0.329 mmol P/g, as reported by O'Neal and Boyer (2015). Subsequently, the introduction of a second parameter, r_p, to the estimation resulted in a further reduction in the root mean square error (RMSE) to 0.0207, representing a 21% decrease. Moreover, the estimated value and uncertainty for q_{max} remained largely unchanged. The calculated 95% confidence interval for this second parameter is 8.4%, indicating that both parameters can be estimated with high confidence based on the available experimental data of breakthrough concentrations and with good agreement with the observed values.

Figure 8(a) illustrates the fit of the model predictions to the experimental data of Zhang et al. (2015), showcasing both calibrated and uncalibrated parameter values. The calibrated values for parameter q_{max} from Table 3 are slightly higher than the initial reported values in the literature, resulting in a rightward shift of the breakthrough curve with respect to the uncalibrated model. Figure 8(b) depicts the residuals, defined as the absolute difference between the experimental data and both the calibrated and uncalibrated model predictions. As can be observed, calibration essentially reduces the residuals in the zone where the calibrated parameters are most sensitive, namely after breakthrough, around 4-10 bed volumes. This is the case for all parameter subsets, although for two and three calibrated parameters, the residuals are almost indistinguishable. The calibration of additional parameters does not significantly improve the fit to the experimental data, due to the phenomenon of overfitting or overparameterization. The uncalibrated model displays a tendency to overpredict the breakthrough concentration. However, all models exhibit a comparable deficiency in accurately reproducing the initial time instants of the experimental data. This is a consequence of the low sensitivity of the parameters in this zone, which renders the model unable to accurately capture this aspect of the curve.

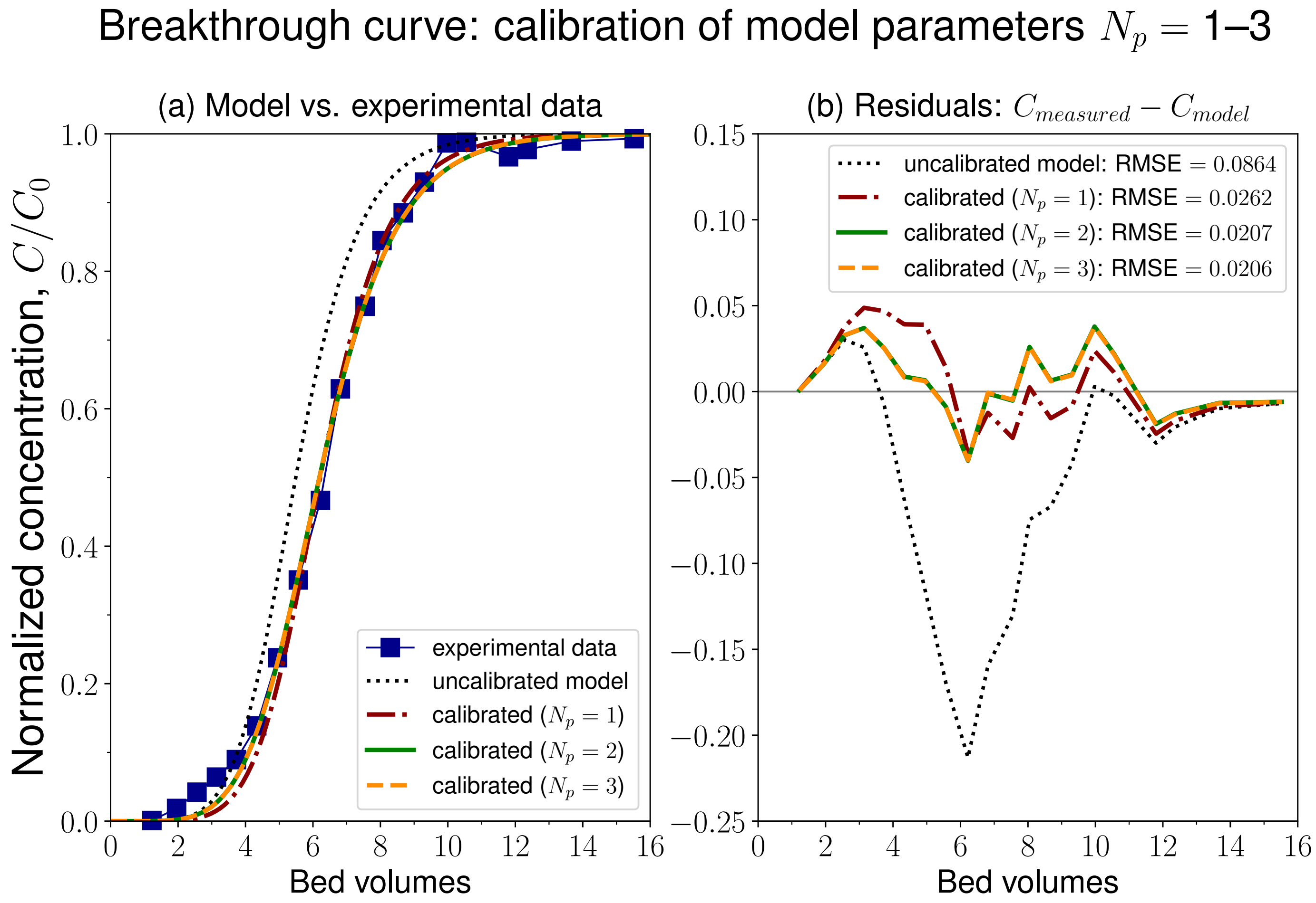

Figure 8. Calibration of up to three model parameters (q_{max}, r_p, ε): model fit to experimental data (a) and residual calculation (b). Data for fresh urine by Zhang et al. (2015). Increasing RSME is a better fit.

Fitting an additional parameter, ε, as illustrated in Table 3, shows that the uncertainties associated with both q_{max} and r_p become significantly more pronounced, resulting in much larger confidence intervals for all parameters. Such uncertain estimation does not provide added value to the model prediction. This is corroborated by the larger value of AIC, which is -149.3, with the lowest value of -151.3 corresponding to the calibration of only two parameters. The calibration of a fourth parameter results in a further deterioration in the precision of the estimation, as evidenced by the considerably broader confidence intervals (CIs) for all parameters, including those that were previously estimated with a high degree of accuracy. The incorporation of further parameters into the calibration subset results in overfitting without a significant reduction in RMSE. The substantial uncertainty suggests that the estimation problem has become ill-conditioned. This occurs when the choice of minimization method yields disparate solutions within the parameter space, with less sensitive parameters estimated at the expense of the certainty in the value of the most sensitive ones. The use of an unconstrained method yielded a value for ε of 0.44, which is at the limit of validity for this parameter. Similarly, parameter K_L is estimated as 10.4, which is nine times higher than the value reported in the literature. Nevertheless, an alternative constrained method does yield valid values of 0.36 and 1.25, respectively, albeit with an increased uncertainty (the confidence intervals for all parameters are higher for this method) and an increased amount of model evaluations (40 vs 52). This is due to the fact that the constrained method employs information regarding the validity region of parameters to determine the step size and direction at each iteration of the optimization process. While unconstrained methods may converge more rapidly when in proximity to a local minimum, they may not yield physically realistic parameter estimates.

The AIC indicates that reducing the RMSE is preferable for model selection, whereas the increase in the number of parameters is less favorable and therefore penalized. Consequently, the optimal model, as determined by the lowest value of the AIC, is obtained by calibrating only the two most sensitive parameters, q_{max} and r_p. This example demonstrates the importance of exercising caution when estimating parameters with low sensitivity, as for subsets of more than two parameters, there exist infinite combinations of parameters that produce the same fit. Hence, these results confirm that only the most sensitive model parameters should be included in the calibration process. Furthermore, a robust calibration protocol is essential, where sensitivity and correlation of parameters are evaluated prior to calibration, to ensure the development of a reliable model with minimal uncertainty and optimal predictive power.

3.3 Uncertainty analysis

In addition to the quality of the parameter estimation provided by the confidence regions from Table 3, an uncertainty analysis of the model with two calibrated parameters is presented in Figure 9. The calibrated model output of the breakthrough curve is enclosed by a 95% confidence band. As illustrated, the uncertainty in the prediction is minimal at the initial and final stages of the operation but increases after breakthrough at approximately 2-4 bed volumes and subsequently decreases near the saturation point at around 10-12 bed volumes. Consequently, the width of the uncertainty band is dependent upon the quality of the model calibration step and thus determines the reliability of the model prediction.

For two calibrated parameters, the uncertainty band is narrow and closely surrounds the calibration data set, indicating a highly accurate prediction. However, the initial instants of the process are not captured by the model as indicated in Section 3.2. Nevertheless, the 10% breakthrough time is in good agreement with the value of 3.8 bed volumes, with an error of 7.8%. Furthermore, the model accurately predicts the saturation time, with an error of 2.2%. More details are provided in Table 4, see Section 3.4.

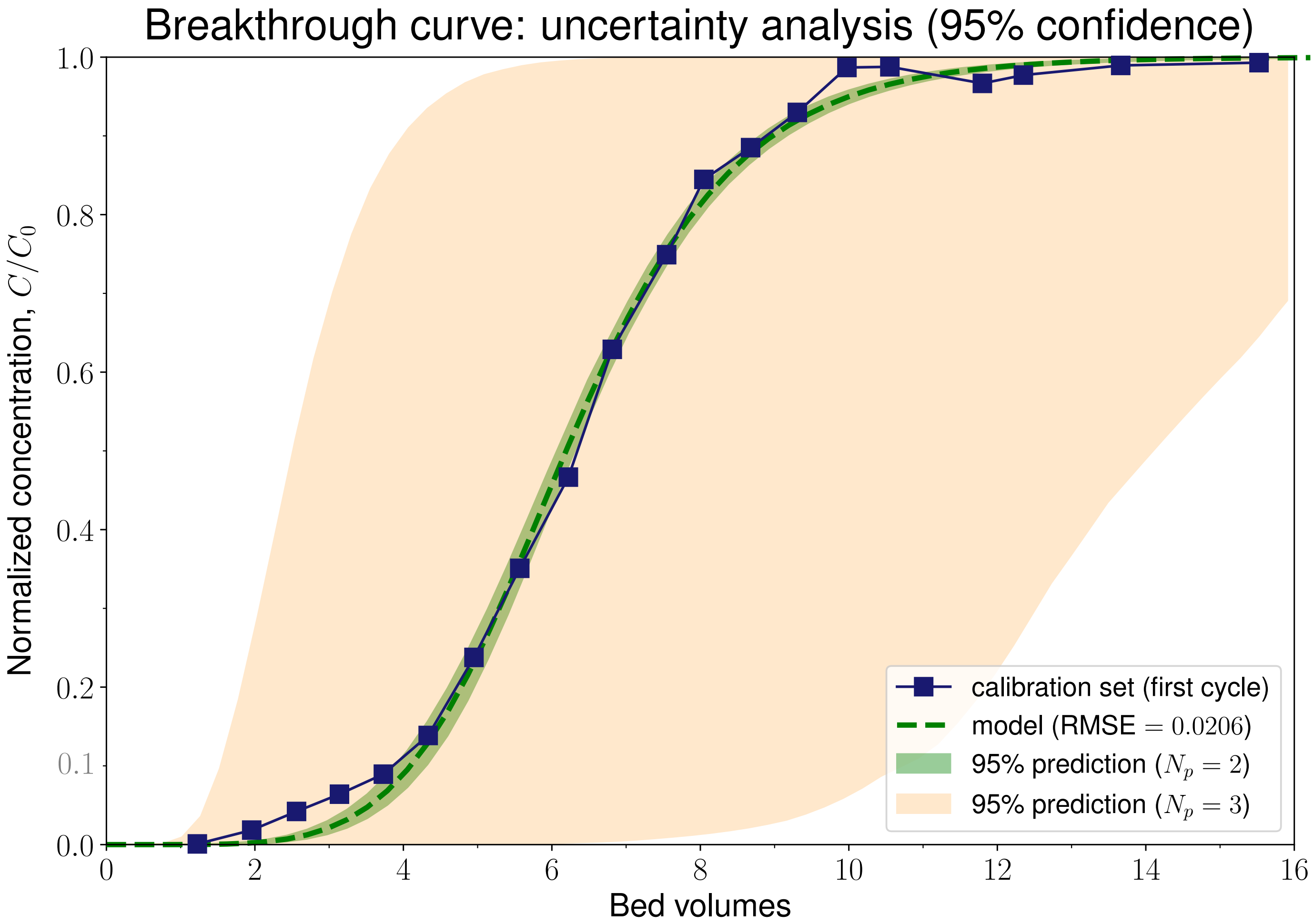

Figure 9. Uncertainty analysis of the breakthrough curve for a 95% confidence in the prediction interval. The uncertainty bands for two and three calibrated parameters are generated by the corresponding confidence intervals of Table 3. The lowest uncertainty is obtained with two calibrated parameters.

3.4 Model validation

Following calibration of the model with the accurate determination of the most sensitive parameters, the predictive power of the model can be tested against the validation data set. Figure 10 depicts the breakthrough curve of the second cycle, which demonstrates considerable agreement with the experimental data. As illustrated by the inset, the initial time instants of the experiment are accurately represented by the model, with a precise prediction of the 10% breakthrough time occurring at approximately 3.2 bed volumes (see Table 4). As demonstrated, the breakthrough time decreases as a consequence of resin fouling. Moreover, the saturation time, defined as the time required for a breakthrough concentration of 90% of the initial concentration, also decreases with fouling. However, the error in determining this point is less in agreement with the experimental data. Potential causes include interactions of the adsorbed ions onto the resin, which could result in increased adsorption capacity when the resin is close to saturation. This, in turn, could give rise to a delay in the breakthrough of the ions. It is possible that this trend may persist after additional regeneration cycles. However, further analysis is required to confirm this hypothesis, which is beyond the scope of the present work.

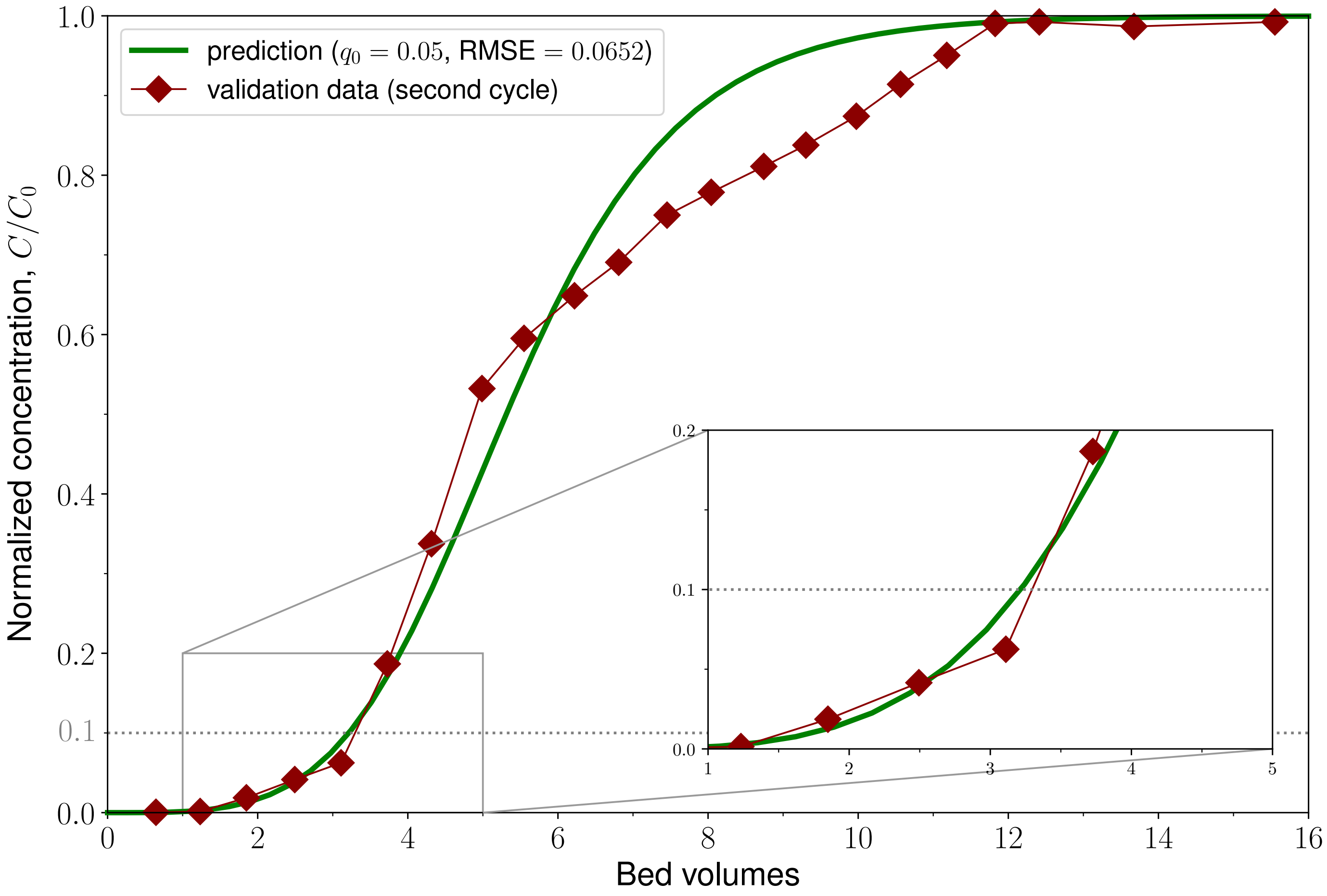

Figure 10. Model validation for 95% regenerated resin using column data from (O'Neal and Boyer, 2015). The inset shows that the initial time instants around 10% breakthrough are well captured by the model.

Table 4. Breakthrough and saturation times (in bed volumes) from the model and the experimental data. The first cycle corresponds to the calibration data set and the second cycle to the validation data set.

	Breakthrough time, t_{10}			Saturation time, t_{90}		
	Data	Model	Rel. error	Data	Model	Rel. error
First cycle (calibration)	3.8	4.1	7.8%	8.8	9.0	2.2%
Second cycle (5% fouled, validation)	3.2	3.2	1.0%	10.3	8.1	21.0%
Third cycle (50% fouled, not illustrated)	2.4	2.2	6.7%	9.3	7.2	22.6%

4. Conclusions

We present a framework for calibration of a dynamic model of fixed-bed column IX operation based on good modelling practice that can be followed for future modelling studies and practical model implementations. The local and global sensitivity analyses allowed us to identify the design and operational parameters that contribute most to the prediction of breakthrough curves. Specifically, the local sensitivity analysis revealed which time intervals during IX operation provide the most information for model calibration, thus enabling an initial screening to select highly sensitive parameters. The global sensitivity analysis then allowed us to select a small subset of parameters for calibration. This approach showed that only two parameters, namely the maximum adsorption capacity isotherm parameter and the resin particle size, need to be thoroughly calibrated to obtain an accurate prediction of the breakthrough curve. Additionally, we demonstrated that uncertainty quantification for model calibration is crucial for establishing the reliability of the predictions. To this end, we conducted a confidence interval analysis of the parameter estimates and applied model selection criteria, such as AIC, to identify the optimal fit. Validation of the model was carried out using experimental data. Hence, we propose a sound calibration procedure, based on good modeling practice, that encompasses both sensitivity and uncertainty analysis and provides a basis for the optimization of the IX process with the aim of improving the accuracy of breakthrough prediction.

Acknowledgements

We would like to express our gratitude to Benjamin Claessens and Ivaylo Hitsov Plamenov for their invaluable contribution of expertise on ion exchange processes and the development of the model.

Funding

This work was supported by Ghent University. MYS was supported by the Swiss National Science Foundation (P500PT_211132).

Appendix

Discretization analysis with respect to numerical parameters

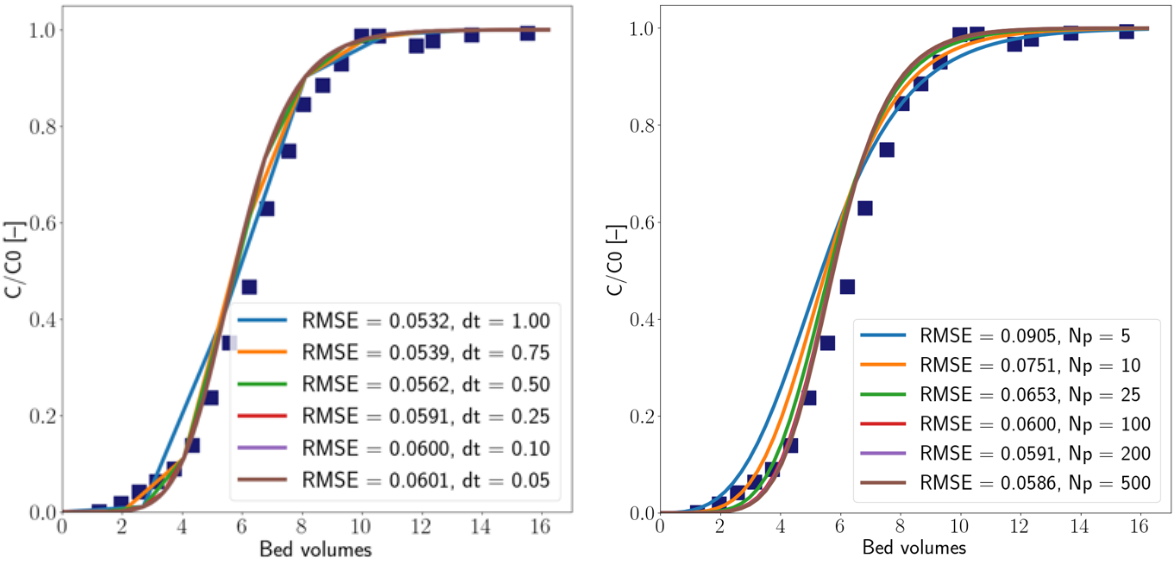

Figure S1. Discretization analysis of model output (breakthrough concentration) with respect to numerical parameters: time step (left) and number of discretized column elements (right). A time step dt = 0.1 and Np = 100 elements provide sufficient accuracy in the numerical solution to predict breakthrough without significantly adding computational effort. This was assessed by subsequently performing simulations with time steps of 1, 0.75, 0.5, 0.25, 0.1 and 0.05 seconds with 100 grid points, followed by 5, 10, 25, 100, 200 and 500 grid points with a 0.1 second time step, and observing indistinguishable simulation results. All simulations require a computation time of under 1 minute.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Autom. Control 19, 716–723. https://doi.org/10.1109/TAC.1974.1100705

Branch, M.A., Coleman, T.F., Li, Y., 1999. A Subspace, Interior, and Conjugate Gradient Method for Large-Scale Bound-Constrained Minimization Problems. SIAM J. Sci. Comput. 21, 1–23. https://doi.org/10.1137/S1064827595289108

Cechinel, M.A.P., Mayer, D.A., Mazur, L.P., Silva, L.G.M., Girardi, A., Vilar, V.J.P., de Souza, A.A.U., Guelli U. de Souza, S.M.A., 2018. Application of ecofriendly cation exchangers (*Gracilaria caudata* and *Gracilaria cervicornis*) for metal ions separation and recovery from a synthetic petrochemical wastewater: Batch and fixed bed studies. J. Clean. Prod. 172, 1928–1945. https://doi.org/10.1016/j.jclepro.2017.11.235

Cellier, N., Ruyer-Quil, C., 2019. scikit-finite-diff, a new tool for PDE solving. J. Open Source Softw. 4, 1356. https://doi.org/10.21105/joss.01356

Chen, S., Bai, S., Ya, R., Du, C., Ding, W., 2022. Continuous silicic acid removal in a fixed-bed column using a modified resin: Experiment investigation and artificial neural network modeling. J. Water Process Eng. 49, 102937. https://doi.org/10.1016/j.jwpe.2022.102937

Chu, K.H., 2023. Fixed Bed Adsorption of Water Contaminants: A Cautionary Guide to Simple Analytical Models and Modeling Misconceptions. Sep. Purif. Rev. 52, 75–97. https://doi.org/10.1080/15422119.2022.2039196

Chu, K.H., Hashim, M.A., 2023a. Adsorptive removal of pharmaceutical contaminants: Accurate description of tailing breakthrough curves. J. Environ. Chem. Eng. 11, 111025. https://doi.org/10.1016/j.jece.2023.111025

Chu, K.H., Hashim, M.A., 2023b. Removal of antibiotics through fixed bed adsorption: Comparison of different breakthrough curve models. J. Water Process Eng. 56, 104512. https://doi.org/10.1016/j.jwpe.2023.104512

Crittenden, J.C., Trussell, R.R., Hand, D.W., Howe, K.J., George Tchobanoglous, 2012. Ion Exchange, in: MWH's Water Treatment: Principles and Design, Third Edition. John Wiley & Sons, Ltd, pp. 1263–1334. https://doi.org/10.1002/9781118131473.ch16

Dochain, D., Vanrolleghem, P.A., 2001. Dynamical modelling and estimation in wastewater treatment processes. IWA Publ, London.

Donckels, B., 2009. Optimal experimental design to discriminate among rival dynamic mathematical models (dissertation). Ghent University.

Freundlich, H., 1907. Über die Adsorption in Lösungen. Z. Für Phys. Chem. 57U, 385–470. https://doi.org/10.1515/zpch-1907-5723

Gavin, H.P., 2019. The Levenberg-Marquardt algorithm for nonlinear least squares curve-fitting problems. Dep. Civ. Environ. Eng. Duke Univ. August 3.

Harris, C.R., Millman, K.J., van der Walt, S.J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M.H., Brett, M., Haldane, A., del Río, J.F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., Oliphant, T.E., 2020. Array programming with NumPy. Nature 585, 357–362. https://doi.org/10.1038/s41586-020-2649-2

Haupert, L.M., Pressman, J.G., Speth, T.F., Wahman, D.G., 2021. Avoiding pitfalls when modeling removal of per- and polyfluoroalkyl substances by anion exchange. AWWA Water Sci. 3, e1222. https://doi.org/10.1002/aws2.1222

Herman, J., Usher, W., 2017. SALib: An open-source Python library for Sensitivity Analysis. J. Open Source Softw. 2, 97. https://doi.org/10.21105/joss.00097

Hu, Q., Pang, S., Wang, D., Yang, Y., Liu, H., 2021. Deeper Insights into the Bohart–Adams Model in a Fixed-Bed Column. J. Phys. Chem. B 125, 8494–8501. https://doi.org/10.1021/acs.jpcb.1c03378

Hunter, J.D., 2007. Matplotlib: A 2D Graphics Environment. Comput. Sci. Eng. 9, 90–95. https://doi.org/10.1109/MCSE.2007.55

Inglezakis, V.J., Zorpas, A., 2012. Fundamentals of Ion Exchange Fixed-Bed Operations, in: Dr., I., Luqman, M. (Eds.), Ion Exchange Technology I: Theory and Materials. Springer Netherlands, Dordrecht, pp. 121–161. https://doi.org/10.1007/978-94-007-1700-8_4

Jegatheesan, V., Shu, L., Jegatheesan, L., 2021. Producing fit-for-purpose water and recovering resources from various sources: An overview. Environ. Qual. Manag. 31, 9–28. https://doi.org/10.1002/tqem.21780

Kabdaşlı, I., Tünay, O., 2018. Nutrient recovery by struvite precipitation, ion exchange and adsorption from source-separated human urine – a review. Environ. Technol. Rev.

Kammerer, J., Carle, R., Kammerer, D.R., 2011. Adsorption and Ion Exchange: Basic Principles and Their Application in Food Processing. J. Agric. Food Chem. 59, 22–42. https://doi.org/10.1021/jf1032203

Langmuir, I., 1918. THE ADSORPTION OF GASES ON PLANE SURFACES OF GLASS, MICA AND PLATINUM. J. Am. Chem. Soc. 40, 1361–1403. https://doi.org/10.1021/ja02242a004

LeVan, M.D., Carta, G., 2008. Adsorption and ion exchange. McGraw-Hill, New York.

Levenberg, K., 1944. A method for the solution of certain non-linear problems in least squares. Q. Appl. Math. 2, 164–168.

Lima, E.C., Sher, F., Guleria, A., Saeb, M.R., Anastopoulos, I., Tran, H.N., Hosseini-Bandegharaei, A., 2021. Is one performing the treatment data of adsorption kinetics correctly? J. Environ. Chem. Eng. 9, 104813. https://doi.org/10.1016/j.jece.2020.104813

Ma, A., Abushaikha, A., Allen, S.J., McKay, G., 2019. Ion exchange homogeneous surface diffusion modelling by binary site resin for the removal of nickel ions from wastewater in fixed beds. Chem. Eng. J. 358, 1–10. https://doi.org/10.1016/j.cej.2018.09.135

Marquardt, D.W., 1963. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. J. Soc. Ind. Appl. Math. 11, 431–441. https://doi.org/10.1137/0111030

Marsili-Libelli, S., Guerrizio, S., Checchi, N., 2003. Confidence regions of estimated parameters for ecological systems. Ecol. Model. 165, 127–146. https://doi.org/10.1016/S0304-3800(03)00068-1

McKinney, W., 2010. Data Structures for Statistical Computing in Python - SciPy Proceedings [WWW Document]. URL https://proceedings.scipy.org/articles/Majora-92bf1922-00a (accessed 9.26.24).

Mudhoo, A., Pittman, C.U., 2023. Adsorption data modeling and analysis under scrutiny: A clarion call to redress recently found troubling flaws. Chem. Eng. Res. Des. 192, 371–388. https://doi.org/10.1016/j.cherd.2023.02.033

Nelder, J.A., Mead, R., 1965. A Simplex Method for Function Minimization. Comput. J. 7, 308–313. https://doi.org/10.1093/comjnl/7.4.308

O'Neal, J.A., Boyer, T.H., 2015. Phosphorus recovery from urine and anaerobic digester filtrate: comparison of adsorption–precipitation with direct precipitation. Environ. Sci. Water Res. Technol. 1, 481–492. https://doi.org/10.1039/C5EW00009B

O'Neal, J.A., Boyer, T.H., 2013. Phosphate recovery using hybrid anion exchange: Applications to source-separated urine and combined wastewater streams. Water Res. 47, 5003–5017. https://doi.org/10.1016/j.watres.2013.05.037

Rieger, L., Gillot, S., Langergraber, G., Ohtsuki, T., Shaw, A., Takacs, I., Winkler, S., 2012. Guidelines for using activated sludge models, First published. ed, Scientific and technical report series. IWA publishing, London New York.

Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., Li, S., Wu, Q., 2019. Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. Environ. Model. Softw. 114, 29–39. https://doi.org/10.1016/j.envsoft.2019.01.012

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., 2007. Global Sensitivity Analysis. The Primer, 1st ed. Wiley. https://doi.org/10.1002/9780470725184

Sengupta, S., Pandit, A., 2011. Selective removal of phosphorus from wastewater combined with its recovery as a solid-phase fertilizer. Water Res. 45, 3318–3330. https://doi.org/10.1016/j.watres.2011.03.044

Smith, S.J., Wahman, D.G., Kleiner, E.J., Abulikemu, G., Stebel, E.K., Gray, B.N., Datsov, B., Crone, B.C., Taylor, R.D., Womack, E., Gastaldo, C.X., Sorial, G., Lytle, D., Pressman, J.G., Haupert, L.M., 2023. Anion Exchange Resin and Inorganic Anion Parameter Determination for Model Validation and Evaluation of Unintended Consequences during PFAS Treatment. ACS EST Water 3, 576–587. https://doi.org/10.1021/acsestwater.2c00572

Sung, C.-L., Tuo, R., 2024. A review on computer model calibration. WIREs Comput. Stat. 16, e1645. https://doi.org/10.1002/wics.1645

Taghvaie Nakhjiri, A., Sanaeepur, H., Ebadi Amooghin, A., Shirazi, M.M.A., 2022. Recovery of precious metals from industrial wastewater towards resource recovery and environmental sustainability: A critical review. Desalination 527, 115510. https://doi.org/10.1016/j.desal.2021.115510

The pandas development team, 2020. pandas-dev/pandas: Pandas. https://doi.org/10.5281/zenodo.3509134

Tran, H.N., You, S.-J., Hosseini-Bandegharaei, A., Chao, H.-P., 2017. Mistakes and inconsistencies regarding adsorption of contaminants from aqueous solutions: A critical review. Water Res. 120, 88–116. https://doi.org/10.1016/j.watres.2017.04.014

Van Rossum, G., Drake, F.L., 2009. Python 3 Reference Manual. CreateSpace, Scotts Valley, CA.

Vanrolleghem, P.A., Insel, G., Petersen, B., Sin, G., De Pauw, D., Nopens, I., Dovermann, H., Weijers, S., Gernaey, K., 2003. A comprehensive model calibration procedure for activated sludge models, in: Weftec 2003. Water Environment Federation, pp. 210–237.

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., Van Der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., Van Mulbregt, P., SciPy 1.0 Contributors, Vijaykumar, A., Bardelli, A.P., Rothberg, A., Hilboll, A., Kloeckner, A., Scopatz, A., Lee, A., Rokem, A., Woods, C.N., Fulton, C., Masson, C., Häggström, C., Fitzgerald, C., Nicholson, D.A., Hagen, D.R., Pasechnik, D.V., Olivetti, E., Martin, E., Wieser, E., Silva, F., Lenders, F., Wilhelm, F., Young, G., Price, G.A., Ingold, G.-L., Allen, G.E., Lee, G.R., Audren, H., Probst, I., Dietrich, J.P., Silterra, J., Webber, J.T., Slavič, J., Nothman, J., Buchner, J., Kulick, J., Schönberger, J.L., De Miranda Cardoso, J.V., Reimer, J., Harrington, J., Rodríguez, J.L.C., Nunez-Iglesias, J., Kuczynski, J., Tritz, K., Thoma, M., Newville, M., Kümmerer, M., Bolingbroke, M., Tartre, M., Pak, M., Smith, N.J., Nowaczyk, N., Shebanov, N., Pavlyk, O., Brodtkorb, P.A., Lee, P., McGibbon, R.T., Feldbauer, R., Lewis, S., Tygier, S., Sievert, S., Vigna, S., Peterson, S., More, S., Pudlik, T., Oshima, T., Pingel, T.J., Robitaille, T.P., Spura, T., Jones, T.R., Cera, T., Leslie, T., Zito, T., Krauss, T., Upadhyay, U., Halchenko, Y.O., Vázquez-Baeza, Y., 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods 17, 261–272. https://doi.org/10.1038/s41592-019-0686-2

Wang, J., Guo, X., 2023. Adsorption kinetics and isotherm models of heavy metals by various adsorbents: An overview. Crit. Rev. Environ. Sci. Technol. 53, 1837–1865. https://doi.org/10.1080/10643389.2023.2221157

Wang, J., Guo, X., 2022. Rethinking of the intraparticle diffusion adsorption kinetics model: Interpretation, solving methods and applications. Chemosphere 309, 136732. https://doi.org/10.1016/j.chemosphere.2022.136732

Wang, W., Li, M., Zeng, Q., 2015. Adsorption of chromium (VI) by strong alkaline anion exchange fiber in a fixed-bed column: Experiments and models fitting and evaluating. Sep. Purif. Technol. 149, 16–23. https://doi.org/10.1016/j.seppur.2015.05.022

Wang, Y., Wang, C., Huang, X., Zhang, Q., Wang, T., Guo, X., 2024. Guideline for modeling solid-liquid adsorption: Kinetics, isotherm, fixed bed, and thermodynamics. Chemosphere 349, 140736. https://doi.org/10.1016/j.chemosphere.2023.140736

Xiao, Y., Azaiez, J., Hill, J.M., 2018. Erroneous Application of Pseudo-Second-Order Adsorption Kinetics Model: Ignored Assumptions and Spurious Correlations. Ind. Eng. Chem. Res. 57, 2705–2709. https://doi.org/10.1021/acs.iecr.7b04724

Yoshida, H., Kataoka, T., Ikeda, S., 1985. Intraparticle mass transfer in bidispersed porous ion exchanger part I: Isotopic ion exchange. Can. J. Chem. Eng. 63, 422–429. https://doi.org/10.1002/cjce.5450630310

Zhang, J., Amini, A., O'Neal, J.A., Boyer, T.H., Zhang, Q., 2015. Development and validation of a novel modeling framework integrating ion exchange and resin regeneration for water treatment. Water Res. 84, 255–265. https://doi.org/10.1016/j.watres.2015.07.027