Hybrid Models for the simulation and prediction of chromatographic processes for protein capture

Harini Narayanan , Tobias Seidler , Martin Luna ,
Michael Sokolov , Massimo Morbidelli , Alessandro Butté

PII: S0021-9673(21)00372-1
DOI: https://doi.org/10.1016/j.chroma.2021.462248
Reference: CHROMA 462248

To appear in: *Journal of Chromatography A*

Received date: 7 January 2021
Revised date: 29 April 2021
Accepted date: 7 May 2021

Please cite this article as: Harini Narayanan , Tobias Seidler , Martin Luna , Michael Sokolov , Massimo Morbidelli , Alessandro Butté , Hybrid Models for the simulation and prediction of chromatographic processes for protein capture, *Journal of Chromatography A* (2021), doi: https://doi.org/10.1016/j.chroma.2021.462248

Title: Hybrid Models for the simulation and prediction of chromatographic processes for protein capture

Authors: Harini Narayanan[1], Tobias Seidler[1], Martin Luna[1], Michael Sokolov[2], Massimo Morbidelli[3], Alessandro Butté[2, *]

[1]Institute of Chemical and Bioengineering, Department of Chemistry and Applied Biosciences, ETH Zurich, Zurich, Switzerland

[2]DataHow AG, Zurich, Switzerland

[3]Dipartimento di Chimica, Materiali e Ingegneria Chimica, Giulio Natta, Politecnico di Milano

Corresponding Author: Alessandro Butté, DataHow AG, Zurich, Switzerland

E-mail: a.butte@datahow.ch

Highlights

- Mechanistic models are limited by the need to have appropriate assumption about the phenomena
- Hybrid model learns unknown phenomena in flexibly and in an un-biased manner from data, using neural network.
- Subsequently, they show 2-3 times higher accuracy compared to industrial benchmark in interpolation and extrapolation
- Errors made by hybrid model are close to the analytical errors made in target quantification
- Tailored hybrid models generate representations for the underlying principles

Abstract

The biopharmaceutical industries are continuously faced with the pressure to reduce the development costs and accelerate development time scales. The traditional approach of heuristic-based or platform process-based optimization is soon getting obsolete, and more generalized tools for process development and optimization are required to keep pace with the emerging trends. Thus, advanced model-based methods that can reduce the can ensure accelerated development of robust processes with minimal experiments are necessary. Though mechanistic models for chromatography are quite popular, their success is limited by the need to have accurate knowledge of adsorption isotherms and mass transfer kinetics. As an alternative, in this work, a hybrid modeling approach is proposed. Thereby, the chromatographic unit behavior is learned by a combination of neural network and mechanistic model while fitting suitable experimental breakthrough curves. Since this approach does not require identifying suitable mechanistic assumptions for all the phenomena, it can be developed with lower effort. Thus, allowing the scientists to concentrate their focus on process development. The performance of the hybrid model is compared with the mechanistic Lumped kinetic Model for in-silico data and experiments conducted on a system of industrial relevance. The flexibility of the hybrid modeling approach results in about three times higher accuracies compared to Lumped Kinetic Model. This is validated for five different isotherm models used to simulate data, with the hybrid model showing about two to three times lower prediction errors in all the cases. Not only in prediction, but we could also show that the hybrid model is more robust in extrapolating

across process conditions with about three times lower error than the LKM. Additionally, it could be demonstrated that an appropriately tailored formulation of the hybrid model can be used to generate representations for the underlying principles such as adsorption equilibria and mass transfer kinetics.

Keywords: Hybrid Models, Chromatography, Capture, Artificial Neural Network

1. Introduction

The progress in molecular engineering and upstream processing (USP) [1] has respectively led to increased diversity of therapeutic formats and the titers. Subsequently, downstream processing (DSP) is currently the bottleneck of biomanufacturing [2]. DSP is dominated by chromatographic processes [3,4] whose development is still not sufficiently efficient or streamlined [5]. Most process development relies on heuristics and platform processes, which are suitable only for a limited range of biologics, such as the widely investigated monoclonal antibodies (mAbs) [6]. A more generalized procedure for chromatographic process development is required to keep pace with the advancements in molecular and USP technologies. Due to the lacking of *a-priori* knowledge about the process and product, experimental approaches are required to characterize the system [6]. Trial-and-error and investigating all variables one by one are inefficient methods and, thus, attention towards high throughput screening systems and experimental design is increasing [7–12].

In this context, mathematical models can immensely reduce the number of experiments to be performed and can strategically guide process development and optimization [13][6], resulting in accelerated development time scales and robust processes. Furthermore, mathematical models are promising tools to support scale-up [14] as well as operation in the manufacturing phase. In combination with UV or Raman sensors, they can be used to monitor breakthrough [15], resin aging [16], and in advanced control schemes [17–19]. In

this regard, the ability to predict the chromatographic behavior of proteins is highly desirable [20]. Thus, currently, the field is faced with challenges to build robust modeling methods that can later be used for several process applications such as optimization, monitoring, control and maintenance.

In general, modeling approaches can be grouped into two types of approaches: (i) Data-Driven (or Empirical) modeling approaches that are purely based on data, and (ii) Mechanistic modeling approaches that are based on understanding of the underlying phenomena. For chromatographic processes, empirical modeling approaches such as Design of Experiments (DoE), Response Surface Methodology (RSM) [21,22], supervised machine learning [23–26], and reinforcement learning [27] have been used for optimization. Additionally, empirical or data-driven modeling approaches have been used to model product recovery [28], retention time [29], and also for Quantitative Structure-Activity Relationship Modeling (QSAR) [30,31] for chromatographic processes. However, given the well-established process understanding, mechanistic modeling of bio-chromatography is predominant and has currently reached a remarkable maturity level [12,32]. These approaches are highly attractive in terms of its ability to model dynamic and spatial profiles and extrapolation capabilities which makes it an ideal tool for process development or optimization [13,33–39], monitoring [15], and advanced control [17–19].

Mechanistic modeling of chromatographic processes includes models of varying complexity [12,40,41]. These range from rigorous Generalized rate model (GRM) [12,32,35,40–43] to lesser complex ones such as the pore model and the Lumped Kinetic Model (LKM), with the equilibrium dispersive model being the simplest variant [5,20,43]. The type of model chosen usually depends on the target application. The challenge with mechanistic models is to identify the exact underlying physicochemical phenomena and postulate suitable mathematical formulation to represent them in the models. To gain this understanding requires significant professional expertise combined with experimental observations (both of which require effort and time to obtain). Further, once general model structures have been identified, the involved parameters must be evaluated, whose number typically increases with model complexity. The most reliable approach is to evaluate them independently through dedicated experimental investigations, which requires additional experimental effort and time. A simpler alternative is to postulate assumptions about suitable equilibrium and transport models and estimate the corresponding parameter values by fitting experimentally observable quantities such as the elution or breakthrough curves for chromatography. However, the so obtained models and the correspondingly estimated parameters may not

reflect the actual thermodynamic and transport properties [20]. For instance, if the models postulated for describing the mass transfer kinetics and the isotherm differ significantly from the real ones, the overall chromatographic models may have trouble fitting the experimental data and extrapolating to different process conditions. Thus, approaches are needed that can provide a robust and accurate model with minimal experimental effort, and simultaneously assist in finding the correct assumptions about the system during the process development phase. Additionally, for industrial application, it is required that the modeling framework is generalizable that can be quickly adapted to various systems and products based on data.

In this work, a hybrid modeling approach to chromatographic processes is proposed. Herein no assumptions are made about the adsorption isotherm, the inter- or intraphase transport kinetics. Instead, a neural network (NN) is used to learn the relationship directly while fitting data. However, the convection and axial diffusion of liquid phase concentration is still modeled like in mechanistic models (such as lumped kinetic model). Previously, a combination of mechanistic model and NN has been used for other applications. For instance, Wang et al (2017) [44] estimated mechanistic model parameters from experimental chromatograms using NNs. On the other hand, in Pirrung et al (2017) [24], NNs were used as a surrogate model to the mechanistic model to facilitate the optimization of the order and the number of different chromatographic units to be used. However, in Creasy et al 2015 [20], a similar approach to this work was suggested, but using interpolation techniques for batch adsorption data to learn about the adsorption isotherm model. On the other hand, in this work, a novel formulation of hybrid modeling is presented wherein the liquid phase concentration is modeled like in a typical lumped kinetic model while the dynamics of the solid phase is learned using a NN. In particular, two different ways of formulating such a hybrid model (c.f. Section 2.2.2) are discussed.

The proposed approach can be applied to any type of chromatography. However, we selected the capture process to demonstrate the development and performance of hybrid models trained on breakthrough (BT) profiles. Protein A resins used in capture processes make up for a considerable share of the total capital cost of the production facilities [45,46]. Efforts in various directions are being made to reduce this cost. Thereby efficient modeling strategies would be of immense help to accelerate the selection, development, and optimization. In this work, firstly, the performance of the hybrid model is compared to that of the mechanistic LKM using an *in-silico* dataset generated based on a Langmuir adsorption isotherm. The advantage of hybrid models in extrapolation and thus for process optimization and design is

illustrated. Further, the ability of hybrid models to accurately predict BT runs generated with different underlying isotherms is demonstrated and compared to the LKM. Finally, the use of hybrid models for a capture system of industrial relevance is discussed based on the corresponding experimental observations.

2. Materials and Methods

2.1 Dataset

The application of hybrid modeling is demonstrated for two case studies: one *in-silico* based on a simulated dataset and the other corresponding to an industrial protein-A chromatography process.

2.1.1 Data Organization

Data from breakthrough (BT) runs, which is the concentration of the protein measured at the column outlet in time, is used for modeling. The process or operating conditions (referred to as "factors" in DoE) are essentially the variables set by design and that stay constant throughout the process operation. The operating conditions varied in the different BT runs include the flowrate (Q) and inlet or feed concentration (C_{in}). The column parameters like the diameter and length of the column are kept fixed. The first row of Figure 1A and 1B tabulates the column and particle dimensions used to obtain data for the *in-silico* and experimental case, respectively. These include the cross-section area of the column (A_{col}), length of the column (L_{col}), and particle diameter (d_p).

2.1.2 In-silico Dataset

In order to test the different properties of the hybrid model such as the extrapolation capability, the ability to predict internal column profiles, and the flexibility to cope with different isotherms, artificial data is generated *in-silico* using a variation of the shrinking core model developed by Steinebach et al [18]. This is a pore model that considers a moving front inside the beads, separating the solid phase into a fully saturated shell and an empty core. The change in shell size has an impact on the mass transfer rate. Besides the liquid concentration and the pore concentration, the model includes two different solid phase concentrations. This accounts for two binding sites in the protein A chain, with the second one being accessible only after the first one is occupied. While the original formulation of the model in [18] uses the Langmuir isotherm, modifications have been made to include various empirical isotherms. The *in-silico* model was implemented using MATLAB 2019b, and a detailed description of the model is presented in the supporting information file. White noise with a 2% standard deviation is imposed on the simulated BT profiles to account for the measurement errors. To this aim, a draw is made from the normal distribution with a standard

deviation of 2% of the true variable value (simulated liquid phase outlet concentration at the given time). The thus obtained value is added to the true simulated value to obtain noisy measurements.

Nine BT runs (BTC #1 – BTC #9) are simulated following a three-level full factorial design with the flowrate (Q) and the inlet concentration (C_{in}) as factors (c.f. Figure1A, "Calibration"). Four additional BT runs are each simulated inside and outside of the experimental domain to test the interpolation (c.f. Figure 1A, "Test Interpolation") and extrapolation (c.f. Figure 1A, "Test Extrapolation") capabilities of the models, respectively. The extrapolation is performed 25% outside the training domain. In the same table, the duration of the BT run (t_{exp}) and the sample collection (t_{samp}) is also reported.

2.1.3 Experimental Dataset

The implementation and performance of the hybrid modeling framework are also demonstrated for a set of experimental data reported by Feidl et al. [15]. In these experiments, a monoclonal antibody (mAb) solution is loaded to a column containing protein A resin until breakthrough. The outlet stream is collected, and the mAb concentration is measured using HPLC, with the standard deviation of the analytical method being 0.007 mg/ml. The details of the experimental procedure and analytics can be found in the original literature [15]. Data of 15 BT curves are available with flowrates, Q being varied between 0.5 to 1.5 mL/min and mAb inlet concentration, C_{in} varied between 0.3 to 0.6 mg/mL as shown in Figure 1B. Nine (BTC #1 – BTC #9) out of the 15 experiments were performed following a three-level full factorial design for flowrates and inlet concentrations. The first six of these BT runs (BTC #1 – BTC #6) correspond to lower and upper boundaries of C_{in}, and all the three levels for Q are used for model training (Figure 1B, "Calibration"). While the other three BT runs (BTC #7 – BTC#9) corresponding to the center point of C_{in}, spanning all the three levels of Q, are used for model testing (Figure 1B, "Test Interpolation"). The remaining six experiments (BTC #10 – BTC #15), used to evaluate the model, were performed in a different column, but with flow rate and inlet concentration inside the training space (Figure 1B, "Test Extrapolation"). In the same table, the duration of the BT run (t_{exp}) and the sample collection (t_{samp}) is also reported.

2.2 Methodology

In this work, all computations have been performed using MATLAB R2019b. Both datasets are divided into two parts: (1) the calibration set used to develop the model and (2) an independent test set used to assess the performance of the calibrated models. It is here noteworthy that in BT runs, only the concentration of protein in the liquid phase at the outlet

is available at different times. Thus, the minimization of errors observed in the outlet liquid phase concentration is used as the objective function. Further, the model performance is evaluated based on root mean squared error in prediction (RMSEP) of the outlet protein concentration in the liquid phase. For the *in-silico* dataset, nine BT runs are used to calibrate (Figure 1A, "Calibration") the model, and four BT runs are used as the test set for interpolation (Figure 1A, "Test Interpolation") and extrapolation (Figure 1A, "Test Extrapolation") each. For the industrial experimental case, six BT runs are used to calibrate the model (Figure 1B, "Calibration"), and three runs performed in the same column are used to test the model (Figure 1B, "Test Interpolation"). Additionally, six experiments performed in a different column are used to test the model performance in extrapolation (Figure 1B, "Test Extrapolation"). Two variants of the hybrid model are considered in this work: the Lumped-Hybrid (referred to as the "Lumped-Hybrid") and the Mass Transfer Isotherm-Hybrid (referred to as "MTI- Hybrid") as described in Section 2.2.2. The performance of the hybrid model is compared to that of a mechanistic Lumped Kinetic Model, from here on referred to as LKM. LKM is probably the most commonly used mechanistic model for liquid chromatography [47] as it requires a relatively limited effort to be developed while maintaining a reasonable predictive power. Hence LKM was chosen as the comparison benchmark for this work. The modeling procedure of LKM and hybrid models is described in the following section and summarized schematically in Figure 2A.

2.2.1 Lumped Kinetic Model (LKM)

The LKM is a modeling approach for chromatography well-established in the literature. It considers two phases where the protein can be located: a mobile liquid phase and a stationary solid phase. For the 1-dimensional model (considering no radial dispersion), the mass balances of the two phases are:

$$\frac{dC}{dt} = -v\,\frac{dC}{dx} + D_l\frac{d^2C}{dx^2} - \frac{1-\varepsilon}{\varepsilon}\frac{dq}{dt} \quad (1)$$

$$\frac{dq}{dt} = k_m(q^* - q) \quad (2)$$

where t is the time and x is the space coordinate along the column axis; C and q are the protein concentrations in the liquid and solid phase, respectively, ε is the bed porosity, v is the interstitial velocity and D_l is the axial diffusivity. A detailed description of the LKM is reported in literature [15] and is summarized in the SI. There are in total seven model parameters in the LKM that is estimated from the fitting of BT runs: A, ε, three parameters from the empirical relationship of mass transfer coefficient (k_{max}, S_1, S_2) and two parameters from the isotherm model (H, q_{sat}). Bayesian Optimization solver, based on the in-built

function *bayesopt ()*, is used to estimate the parameters (A, ε, S_1, S_2, H, q_{sat}, k_{max}) of the LKM model (c.f. Figure 2A). The routine *ode15s ()* is used to solve the system of ordinary differential equations, obtained by discretizing the original PDEs in space. To avoid local minima in the objective function, ten models with ten random initializations is performed. The model with the least root mean squared error (RMSE) in calibration (RMSEC) is chosen and investigated with respect to its prediction error on the test set (RMSEP).

2.2.2 Hybrid Model

The LKM assumes a Langmuir isotherm model for adsorption, a linear driving force, and an empirical relationship to estimate the mass-transfer coefficient. Like any other mechanistic model, the LKM performs as good as the underlying assumption: an error in the assumptions (e.g., the mass transfer rate and the adsorption isotherm model) as well as in the simplified representation of the complex physicochemical system would hamper its performance. Besides reducing its extrapolation capabilities, this would also limit the model flexibility needed to capture the real behavior of the process. Thus, in the Lumped-Hybrid approach, the overall adsorption kinetics has been freed from all these assumptions. Subsequently, the right-hand-side of the equation modeling the solid phase concentration (Eq (2)) is substituted by a NN depending on both q and C as follows:

$$\frac{dq}{dt} = NN\left(q(x,t)\,, C(x,t)\right) \tag{11}$$

where the different variables indicate the same quantities as in LKM. On the other hand, equation modeling the liquid phase (Eq (1)) remains the same as in the LKM. Note that for protein chromatography the interphase mass transfer resistance is typically located inside the porous particle, i.e., related to the intraparticle transport, which is independent of the linear velocity or flowrate. Accordingly, the NN trained in this work does not account for a dependency on the liquid flowrate, Q. However, for other relevant chromatographic processes involving small molecules, appropriate NNs including functionality with respect to Q in addition to q and C should be considered.

As indicated in Eq (11), the NN is a function of q and C at a given point in space and time. In order to choose the optimal NN structure, hyperparameter tuning using grid search is performed [48]. SI Table 3 tabulates the optimal NN structure identified for the different cases demonstrated in this work. Similar to the numerical solution of the LKM, the solution of the Lumped-Hybrid approach involves solving an optimization problem while iteratively integrating a system of PDEs (c.f. Figure 2A). Essentially, a non-linear parameter optimization problem is solved to obtain the optimal weights that minimize the difference

between the measured BT profile and the model predicted ones. The difference is that, in the LKM, the parameters of the mechanistic model are optimized, while in the Lumped-Hybrid model, the NN weights are optimized. The routine *ode15s()* is used to solve the system of PDEs discretized in space while *fmincon()* is used for the optimization of A, ε and the NN weights. Similar to the LKM approach, multiple shooting with 10 random initializations are performed to avoid local convergence. The model with the least RMSEC is selected, and its performance on the test set is evaluated using RMSEP.

Besides the Lumped-Hybrid, an alternative formulation of the hybrid model is developed to enable the understanding of mass transfer kinetics and adsorption equilibria. This approach still assumes a linear driving force like the LKM, but the adsorption isotherm and the mass transfer kinetics are now described by two independent NN as follows:

$$\frac{dq}{dt} = NN_{mt}\left(q(x,t)\right) \cdot \left[NN_{iso}\left(C(x,t)\right) - q(x,t)\right] \tag{12}$$

In this model, referred to as MTI-Hybrid, while the NN for the mass transfer term (NN_{mt}) depends on the solid phase concentration (q), the NN for the adsorption isotherm (NN_{iso}) depends on the liquid phase concentration (C). Figure 2B shows a visual representation of both the hybrid model formulations developed in this work.

Unlike the Lumped-Hybrid model where the NN weights are randomly initialized, a different approach is used for initializing the weights of the two networks in the MTI-Hybrid. This is for two reasons: (1) avoid instabilities in optimization since the first network depends on a variable (q) that is not measured and hence not considered in the calculation of the objective function during the fitting procedure; (2) there can be multiple combinations of the NN_{mt} and NN_{iso} terms that can fit equally well the BT curves but might not be physically meaningful. Thus, we first estimate the adjustable parameters of the LKM by fitting the BT runs. Next, we use the empirical relationship for mass transfer (c.f. Eq (SI 26)), with the so obtained values of the adjustable parameters, to generate data for tuning the NN_{mt}. Similarly, the fitted Langmuir isotherm model (c.f. Eq (SI 27)) is used to generate the data for tuning the NN_{iso}. The two NNs, NN_{mt} and NN_{iso}, are trained preliminarily on the data generated using the aforementioned procedure. The optimal weights thus identified are used to initialize the NNs in the MTI-Hybrid model framework. Subsequently, the MTI-Hybrid model is calibrated like the Lumped-Hybrid model by solving an optimization problem while iteratively integrating a system of PDEs. Same routines as the Lumped-Hybrid approach, *ode15s()* and *fmincon()* are

used to solve the system of discretized PDEs and optimization, respectively. Additionally, the optimal NN structure is obtained through hyperparameter optimization using grid search [48].

3. Results and Discussion

3.1 Hybrid model validation: *In-silico* data generated using Langmuir adsorption isotherm

As mentioned above, the performances of the Lumped-Hybrid model and LKM are first discussed, with reference to an *in-silico* dataset generated using the shrinking core model (as described in Section 2.1.2) with a Langmuir adsorption isotherm. The models are trained on nine BT runs simulated using the process conditions indicated as "Calibration" in Figure 1A. Subsequently, the performance of the models is tested on BT runs simulated using process conditions that fall inside the training region indicated as "Test Interpolation" in Figure 1A. In Figure 3A, the predictions of the Lumped-Hybrid model and LKM are compared with the *in-silico* observed values for the "Test Interpolation" conditions. The Lumped-Hybrid model predictions lie very close to the diagonal, indicating good agreement with the *in-silico* observed data. There is no bias in the predictions made by the Lumped-Hybrid model, while this is observed for the LKM. It overpredicts at low concentration and underpredicts at higher concentration. This can be confirmed from Figure 3B, which compares the prediction of the Lumped-Hybrid model and LKM for an exemplary BT run from the "Test Interpolation" set. Here the LKM predictions rise earlier and attain saturation later, while the Lumped-Hybrid model predictions follow better the actual BT.

Additionally, the predictions made by the Lumped-Hybrid model do not vary much across the diagonal (Figure 3A), indicating they are very close to the measurement error. Quantitatively, the overall RMSEP averaged across all BT runs in the "Test Interpolation" set, and all sampling time points are 0.065 mg/mL and 0.0196 mg/mL for the LKM and Lumped-Hybrid model, respectively. The RMSEP of the Lumped-Hybrid model is very close to the average added noise of about 0.015 mg/mL, which marks the realistic prediction limit of a model. The three times higher RMSEP of the LKM reflects the qualitative observations made earlier. The LKM is not capable of capturing the effects imposed by the more detailed model, i.e., the shrinking core model, used to generate the data.

The LKM considers two phases for the protein inside the column: the moving liquid phase and the stationary solid phase inside the beads (where the protein binds to the binding sites). In contrast, the more complex shrinking core model additionally includes the liquid phase inside the pores of the beads. It also differentiates between two types of binding sites in the solid phase. In addition to that, the shrinking core model considers the effect of the coverage

of the binding sites from the external radius of the beads to its center, on the protein diffusion inside the pores (thus giving the name of "shrinking core"). As a result, the LKM may perform well in describing a single BT run, but the mathematical structure imposed by the mechanistic assumptions is not flexible enough to capture the different diffusive and binding phenomena of the shrinking core model across the entire training set. Thus, the optimizer can only find a parameterization that is a compromise among the different BT curves. The hybrid model, on the other hand, learns an aggregate behavior of the effects imposed by the shrinking core model using the flexibility of the NN, even with the same number of differential equations as the LKM.

Further, Figure 3 C compares the internal column profiles predicted by the LKM and the Lumped-Hybrid model with the actual simulated column profiles for different times. It is worth highlighting that both the LKM and the Lumped-Hybrid model are trained only on the outlet concentration and not on the internal column profiles. It can be observed that the Lumped-Hybrid model outperforms the LKM in predicting the internal adsorption fronts at all time points, except for t = 15 min, where both LKM and Lumped-Hybrid models deviate from the actual internal profile. Similarly, as for the BT curves, the LKM overpredicts low concentrations and underpredicts higher concentrations, thus underestimating the steepness of the fronts. On the other hand, the Lumped-Hybrid model predictions closely follow the actual adsorption front except for the initial time point. Accurate prediction of the internal column profiles is crucial, for instance, for control applications. This is because, during the capture step, the column must be loaded to its maximum capacity (avoiding any breakthrough) to maximize the resin utilization without a drop in the yield. Thus, having a correct description of the internal loading is mandatory. This is particularly important in the case of continuous multicolumn units, like Capture SMB or Partial Countercurrent Columns systems (PCC) [32], where the final state of a column in a given phase of the process cycle represents the initial state in the following one.

3.2 Potential of Hybrid model in Extrapolation

Since the hybrid model exploits the synergy between mechanistic models and data-driven models, they are considered to be particularly reliable in extrapolations, for example, illustrated for upstream cell culture processes by Narayanan et al [49]. In the following, we investigate the performance of the same LKM and Lumped-Hybrid model trained in section 3.1 in predicting the observed BT runs corresponding to a combination of flowrate (Q) and

inlet concentration (C_{in}) that fall outside of the training *in-silico* dataset (c.f. Figure 1A, "Test Extrapolation").

It is seen that for low Q and high C_{in} (Figure 4A), both LKM and Lumped-Hybrid BT profile raises a bit earlier than the observed curve. However, the Lumped-Hybrid model correctly predicts the steepness and attainment of breakthrough coherent with the observed curve. In contrast, the LKM predicts a less steep BT profile, thus, underestimating the higher concentrations and attaining the breakthrough later than observed. It is further noted that also for the other three conditions (Figure 4B through 4D), LKM fails to capture the appropriate steepness of the BT curve. Additionally, for low C_{in} (Figure 4C and 4D), the LKM predicts a rise in the outlet concentration much earlier than observed and, though it achieves the saturation concentration at a similar time for low flowrate (Figure 4C), there is a significant mismatch in the rest of the BT profile. On the contrary, the Lumped-Hybrid model is capable of predicting the entire BT profile consistent with the observed profile with minimal deviation.

Overall, it can be concluded that the Lumped-Hybrid model is superior in extrapolation compared to the LKM for all four cases, as also indicated by the corresponding average RMSEPs of 0.044 mg/mL (Lumped-Hybrid) and 0.102 mg/mL (LKM). Additionally, when extrapolating 50% outside the training domain, the RMSEP by the Lumped-Hybrid model is 0.066 mg/mL, which is comparable to the RMSEP of LKM in interpolation. However, for this case, the RMSEP made by LKM is 0.158 mg/mL (c.f. SI Figure 1). This highlights the robustness and accuracy of the hybrid modeling approach in extrapolating across process conditions which is extremely important to guide decision making in process development and control in process operation.

3.3 Performance of Hybrid model for unknown adsorption isotherms

A distinct advantage of the proposed Lumped-Hybrid modeling approach over mechanistic models is that it does not require to formulate any hypothesis about the physicochemical mechanisms involved in the process. In this section, it is discussed with reference to the adsorption equilibria which can be described through different model adsorption isotherms. For this, we compare the Lumped-Hybrid model against the LKM for BT runs generated *in-silico*, as discussed in section 2.1.2, but using different underlying adsorption isotherms: Langmuir-Freundlich (LF), Temkin (TM), Toth (TH), and Bi Langmuir (BL), to generate different training and test datasets using the same process condition reported in Figure 1A ("Calibration" and "Test Interpolation"). It is to be noted that since the real adsorption isotherm is not known *a-priori*, the LKM is formulated based on the common practice of

using the Langmuir adsorption isotherm. On the other hand, no such assumptions are to be made for the Lumped-Hybrid model, which is flexible to learn from the data observed.

From Figure 5, it can be observed that the Lumped-Hybrid model provides a better representation of the data, similar to the case of the Langmuir isotherm discussed above, independently of the adsorption isotherm used to generate the reference data. In particular, the LKM presents a significant systematic deviation from the diagonal in the observed versus predicted plot for all isotherms (Figure 5, the plot in the upper left corner). For LF, TM, and TH (Figure 5A through C), the LKM overpredicts the low and underpredicts the high concentration values. However, for the BL case, LKM overpredicts very low and high concentrations while it underpredicts the intermediate concentrations (Figure 5D). These behaviors are reflected in the BT profiles shown in Figure 5 for an exemplary experiment from the "Test Interpolation" set (Figure 1A). Similar to the Langmuir isotherm case, the LKM is unable to capture the actual steepness of the BT profiles and underestimates the steepness strongly for the LF case and considerably for the BT runs generated using the TM and TH isotherms. For the BL-based BT run, the LKM fails to get the shape of the curve.

In contrast, the Lumped-Hybrid model is capable of reliably modeling and predicting the BT curves generated using different underlying adsorption isotherms, owing to the flexibility offered by the NN that allows it to learn from the observed BT experiments. In Figure 5, it is seen that the observed versus predicted plots do not exhibit any systematic bias nor large variances. This is also reflected in the exemplary BT profiles of the test experiment for the different isotherms. This results for the Lumped-Hybrid model in a RMSEP of 0.0189 mg/mL, 0.0173 mg/mL, 0.0216 mg/mL and 0.0188 mg/mL for the LF, TM, TH and BL isotherm, respectively. Whereas the corresponding RMSEP values for the LKM are about two to three times higher: 0.0765 mg/mL, 0.0303 mg/mL, 0.0458 mg/mL and 0.0354 mg/mL for the LF, TM, TH and BL, respectively. Similar to the Langmuir isotherm discussed in Section 3.1, the RMSEP of the Lumped-Hybrid model is very close to the measurement noise, which marks the realistic prediction limit of a model. This remarkable performance of the Lumped-Hybrid model indicates the potential of such an approach in the case of complex systems, with various components that may follow different adsorption equilibria, possibly competing with each other, as in the case of bio-chromatography.

3.4 Hybrid model to suggest plausible underlying isotherm

An advanced potential of hybrid modeling, besides the presented predictive power, is to use this approach to generate some insights about the underlying physicochemical phenomena,

such as the adsorption equilibria. For this, alternative approaches to the Lumped-Hybrid model discussed above, such as the MTI-Hybrid approach mentioned in Section 2.2.3, should be considered. Here, independent NNs are used to learn about adsorption equilibria as well as mass transfer kinetics. Following the training procedure described in section 2.2.3, MTI-Hybrid models are trained on the BT profiles generated by the *in-silico* model using different isotherms for "Calibration" conditions (Figure 1A). In other words, the same datasets used in sections 3.1 and 3.3 to train the Lumped-Hybrid model are now used to train the MTI-Hybrid model. The NN_{iso} (c.f. Eq 12, and Figure 2B) network learnt as part of the MTI-Hybrid model training can then be used to generate data from which the adsorption equilibrium can be recovered. This is illustrated in Figure 6 where the isotherms learnt by the MTI-Hybrid model are compared to the original isotherms used in the *in-silico* model, in terms of the normalized solid phase concentration, $\theta = \frac{q^*}{q^*_{max}}$ as a function of the equilibrium liquid phase concentration, c.

It can be observed that the MTI-Hybrid model is capable of capturing most of the trends of the isotherms in terms of slope and curvature. A considerable deviation is found only for the Bi-Langmuir case. However, for all cases, the RMSEP of the MTI-Hybrid model is comparable to that of the Lumped-Hybrid model (c.f., SI Table 3). It is worth noting that the *in-silico* data simulator, the shrinking core model, considers two adsorption sites in the solid phase, while the MTI-Hybrid only has one. Thus, the model has to describe the adsorption phenomena of two sites which are different (since the second site is hindered by the binding on first site). In addition to that, the *in-silico* model considers the diffusion of the protein inside the resin pores. In contrast, the MTI-Hybrid model only uses the bulk liquid and the mean solid concentration values, which can be challenging if the intra-particular gradients are important. These mismatches between the model and the simulator give rise to differences in the learned isotherm. It is here crucial to point out that the MTI-Hybrid model is not trained on the adsorption equilibrium data but on the BT experiments, which makes the learning of the isotherm even more difficult since it has to be indirectly done while fitting the BT profiles. Despite this, the MTI-Hybrid was able to generate a good approximation of the actual isotherm used in the in-silico model. One could think of using this approach to generate new mechanistic hypotheses about the system and support the development of mechanistic models.

3.5 An industrially relevant set of experimental data

In the previous sections, *in-silico* data are used since these provided the possibility to test different properties of the hybrid model such as the extrapolation capability, the ability to predict internal column profiles, and the flexibility to cope with different isotherms, which would be impossible to study with a true experimental system. In this section, we investigate the performance of hybrid modeling using a set of experimental data of industrial relevance. The LKM and Lumped-Hybrid models are trained on six BT runs indicated as the "Calibration" set in Figure1B. The performance of the LKM and the Lumped-Hybrid model are compared for the interpolation data set (c.f., Figure 1B, "Test Interpolation"), where the experimental BTs were performed with process conditions that lie inside the training design space and in particular using the same column. The obtained results are shown in Figures 7A and 7B for the LKM and the Lumped-Hybrid model. It is seen that the LKM underpredicts the lower concentration and overpredicts the higher concentration values for two out of the three BT curves, while for the other, it slightly overpredicts the saturation concertation. The Lumped-Hybrid model, on the other hand, predicts well all the three BT curves demonstrating no systematic bias or large variances. Quantitatively, the average RMSEP of the Lumped-Hybrid model is 0.0072 mg/mL, which is considerably lower than the RMSEP of LKM that is around 0.009 mg/mL.

Finally, the LKM and the Lumped-Hybrid model predictions are compared, in Figures 7C and 7D, respectively, with the experimental BT curves measured in a different column. Therefore, this corresponds to an extrapolation test, although the operating conditions are still chosen within the design space used for training, as indicated in Figure 1B "Text Extrapolation". It can be observed that for BTC-14, 15 in Figure 7D, the Lumped-Hybrid model can predict the BT profile accurately, while the LKM (Figure 7C) predicts an earlier breakthrough. For BTC-10,11, and 12, the Lumped-Hybrid model predicts an earlier raise, although it manages to get the final breakthrough point and the overall shape appropriately, and the BTC predictions appear to be shifted along the time-axis. LKM, on the other hand, also predicts an early rise but is not able to predict neither the shape of the experimental BT profile nor the BT time and concentration. For BTC-13, both the Lumped-Hybrid model and LKM show an offset from the actual experimental profile. Thus summarizing, while the Lumped-Hybrid model presents only a translational offset along the time axis and predicts the shape of the BT curve correctly, the LKM predicts a different shape overall for the experimental BT curve. It is worth recalling that the experiments considered here were conducted in a different column than the one used in the training of the models and, in particular, in the evaluation of the bed porosity parameter, ε. This could be a plausible

explanation for the observed translation along the time axis, although to a different extent, in both models.

4. Conclusion and Summary

Developing proper chromatographic models is a complex process. This involves the definition the selection and the validation of proper mechanistic equations to describe for instance the adsorption process and the diffusion to and into the particles. In addition to this, obtaining reliable estimates of the mechanistic model parameters is often quite tedious and requires multiple and specific experimental techniques. Such effort clearly increases with increasing model and process complexity. For this reason, in industrial applications, the Lumped Kinetic Models (LKM) is often used as a reasonable compromise between model complexity and reliability. Despite their widespread use, assumptions must be laid to obtain a suitable set of equations that describes adsorption equilibria and mass transfer kinetics. This restricts the performance of the models to be only as good as the imposed assumptions.

To overcome this issue, in this work, a novel hybrid modeling approach for chromatographic processes is presented. This model is not intended as a more accurate version of the proposed mechanistic models available in the literature, but as a technique which can account for model complexity without the need of model assumptions and specific experimental procedures to estimate physical parameters. In our hybrid model, the features of the process are learnt in an unbiased way by the machine learning part of the model. In this context, the comparison with a LKM make sense as hybrid models are supposed to replace simple models, like LKM, especially in the industrial environment, where models are simply used to support process development and optimization.

The performance and potential of the hybrid models are investigated on: (i) *in-silico* data generated using a detailed shrinking core model and (ii) an experimental dataset generated in a system of industrial relevance. The error in prediction for the hybrid model was around 0.02 mg/mL for the *in-silico* data set, which is close to the 2% Gaussian noise imposed, and 0.0072 mg/mL for the experimental data set, which is close to the reported 0.007 mg/mL experimental error of the analytical HPLC. This indicates that the performance of hybrid model marks the realistic predictive limit of the model that can be developed using this data.

The hybrid model outperformed the LKM in the predictive accuracy by about two to three times for the five different isotherms tested in this work, namely, Langmuir, Langmuir-Freundlich, Temkin, Toth and Bi-Langmuir. While the LKM showed systematic biases and higher variance, the hybrid models had lower bias and variance in all cases. Additionally, the hybrid models also presented a higher accuracy and robustness in extrapolation along the

process conditions resulting in three times lower error in prediction than the LKM. The conclusion reached for the *in-silico* dataset has been confirmed using an experimental data set referring to a system of industrial relevance.

An important additional result of this study is that, through an alternative formulation of the hybrid models (i.e., the MTI-Hybrid) it is possible to derive quantitative information about the mathematical structure of the models describing the underlying adsorption equilibrium and mass transfer kinetics from breakthrough experiments. Subsequently, this information could be used to support the improvement of mechanistic models.

Overall, we have shown that hybrid models exhibit the predictive capabilities and robustness required to model the dynamics of chromatographic processes. These models can be used for process optimization, monitoring, control and even to improve process understanding. For the purpose of real-time monitoring and control, such models can also be coupled with spectroscopic measurements via Extended Kalman filters as discussed in our recent work for upstream processes [50]. It is also evident that results similar to those obtained in this work in the context of protein A capture processes are expected for other chromatographic processes based on different stationary phases.

A limitation of the hybrid model in the current form is the time required for training the model, that involves solving iteratively a system of discretized ODEs for each optimization call. To circumvent this drawback, alternative numerical methods could be considered for model training.

Declaration of Interest Statement

The authors have no conflicts of interest to declare.

Credit Author Statement

Harini Narayanan – Conceptualization, Methodology, Formal Analysis, Validation, Writing – Original Draft, Visualization

Tobias Seidler – Methodology, Software, Formal Analysis

Martin Luna – Conceptualization, Formal Analysis, Writing – Review& Editing

Michael Sokolov – Writing – Review& Editing

Massimo Morbidelli - Writing – Review& Editing

Alessandro Butté - Conceptualization, Writing – Review& Editing, Supervision

Reference

[1] P. Gronemeyer, R. Ditz, J. Strube, Trends in Upstream and Downstream Process Development for Antibody Manufacturing, Bioengineering. 1 (2014) 188–212. https://www.mdpi.com/2306-5354/1/4/188.

[2] A.T. Hanke, M. Ottens, Purifying biopharmaceuticals: Knowledge-based chromatographic process development, Trends Biotechnol. 32 (2014) 210–220. doi:10.1016/j.tibtech.2014.02.001.

[3] H.F. Liu, J. Ma, C. Winter, R. Bayer, Recovery and purification process development for monoclonal antibody production, MAbs. 2 (2010) 480–499. doi:10.4161/mabs.2.5.12645.

[4] J. Smiatek, A. Jung, E. Bluhmki, Towards a Digital Bioprocess Replica: Computational Approaches in Biopharmaceutical Development and Manufacturing, Trends Biotechnol. (2020) 1–13. doi:10.1016/j.tibtech.2020.05.008.

[5] V. Kumar, A.M. Lenhoff, Mechanistic Modeling of Preparative Column Chromatography for Biotherapeutics, Annu. Rev. Chem. Biomol. Eng. 11 (2020) 235–255. doi:10.1146/annurev-chembioeng-102419-125430.

[6] P. Baumann, J. Hubbuch, Downstream process development strategies for effective bioprocesses: Trends, progress, and combinatorial approaches, Eng. Life Sci. 17 (2017) 1142–1158. doi:10.1002/elsc.201600033.

[7] M. Bensch, P. Schulze Wierling, E. von Lieres, J. Hubbuch, High throughput screening of chromatographic phases for rapid process development, Chem. Eng. Technol. 28 (2005) 1274–1284. doi:10.1002/ceat.200500153.

[8] A. Susanto, K. Treier, E. Knieps-Grünhagen, E. Von Lieres, J. Hubbuch, High throughput screening for the design and optimization of chromatographic processes: Automated optimization of chromatograpic phase systems, Chem. Eng. Technol. 32 (2009) 140–154. doi:10.1002/ceat.200800350.

[9] D. Pfister, L. David, M. Holzer, R.M. Nicoud, Designing affinity chromatographic processes for the capture of antibodies. Part I: A simplified approach, J. Chromatogr. A. 1494 (2017) 27–39. doi:10.1016/j.chroma.2017.02.070.

[10] T. Afroz, C.L. Beisel, Understanding and exploiting feedback in synthetic biology, Chem. Eng. Sci. (2013) 1–12. doi:10.1016/j.ces.2013.02.017.

[11] A. Ljunglöf, K. Eriksson, T. Frigård, Rapid process development for purification of a MAb, Bioprocess Int. 9 (2011) 62–68.

[12] G. Carta, A. Jungbauer, Protein Chromatography, 2010.

[13] C.K.S. Ng, H. Osuna-Sanchez, E. Valéry, E. Sørensen, D.G. Bracewell, Design of high productivity antibody capture by protein A chromatography using an integrated experimental and modeling approach, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 899 (2012) 116–126. doi:10.1016/j.jchromb.2012.05.010.

[14] S. Gerontas, M. Asplund, R. Hjorth, D.G. Bracewell, Integration of scale-down experimentation and general rate modelling to predict manufacturing scale chromatographic separations, J. Chromatogr. A. 1217 (2010) 6917–6926.

doi:10.1016/j.chroma.2010.08.063.

[15] F. Feidl, S. Garbellini, M.F. Luna, S. Vogg, J. Souquet, H. Broly, M. Morbidelli, A. Butté, Combining mechanistic modeling and raman spectroscopy for monitoring antibody chromatographic purification, Processes. 7 (2019). doi:10.3390/pr7100683.

[16] F. Feidl, M.F. Luna, M. Podobnik, S. Vogg, J. Angelo, K. Potter, E. Wiggin, X. Xu, S. Ghose, Z.J. Li, M. Morbidelli, A. Butté, Model based strategies towards protein A resin lifetime optimization and supervision, J. Chromatogr. A. 1625 (2020) 461261. doi:10.1016/j.chroma.2020.461261.

[17] C. Grossmann, G. Ströhlein, M. Morari, M. Morbidelli, Optimizing model predictive control of the chromatographic multi-column solvent gradient purification (MCSGP) process, J. Process Control. 20 (2010) 618–629. doi:10.1016/j.jprocont.2010.02.013.

[18] F. Steinebach, M. Angarita, D.J. Karst, T. Müller-Späth, M. Morbidelli, Model based adaptive control of a continuous capture process for monoclonal antibodies production, J. Chromatogr. A. 1444 (2016) 50–56. doi:10.1016/j.chroma.2016.03.014.

[19] M. Morari, C. E. Garcia, D. M. Prett, J. J. Lee, Model Predictive Control, Pearson Professional Education, New Jersey, United States, 2004.

[20] A. Creasy, G. Barker, Y. Yao, G. Carta, Systematic interpolation method predicts protein chromatographic elution from batch isotherm data without a detailed mechanistic isotherm model, Biotechnol. J. 10 (2015) 1400–1411. doi:10.1002/biot.201500089.

[21] M. Toueille, A. Uzel, J.F. Depoisier, R. Gantier, Designing new monoclonal antibody purification processes using mixed-mode chromatography sorbents, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 879 (2011) 836–843. doi:10.1016/j.jchromb.2011.02.047.

[22] A. Osberghaus, S. Hepbildikler, S. Nath, M. Haindl, E. von Lieres, J. Hubbuch, Optimizing a chromatographic three component separation: A comparison of mechanistic and empiric modeling approaches, J. Chromatogr. A. 1237 (2012) 86–95. doi:10.1016/j.chroma.2012.03.029.

[23] S. Liu, L.G. Papageorgiou, Optimal Antibody Purification Strategies Using Data-Driven Models, Engineering. 5 (2019) 1077–1092. doi:10.1016/j.eng.2019.10.011.

[24] S.M. Pirrung, L.A.M. van der Wielen, R.F.W.C. van Beckhoven, E.J.A.X. van de Sandt, M.H.M. Eppink, M. Ottens, Optimization of biopharmaceutical downstream processes supported by mechanistic models and artificial neural networks, Biotechnol. Prog. 33 (2017) 696–707. doi:10.1002/btpr.2435.

[25] D. Nagrath, A. Messac, B.W. Bequette, S.M. Cramer, A Hybrid Model Framework for the Optimization of Preparative Chromatographic Processes, Biotechnol. Prog. 20 (2004) 162–178. doi:10.1021/bp034026g.

[26] G. Wang, T. Briskot, T. Hahn, P. Baumann, J. Hubbuch, Root cause investigation of deviations in protein chromatography based on mechanistic models and artificial neural networks, J. Chromatogr. A. 1515 (2017) 146–153. doi:10.1016/j.chroma.2017.07.089.

[27] S. Nikita, A. Tiwari, D. Sonawat, H. Kodamana, A.S. Rathore, Reinforcement learning based optimization of process chromatography for continuous processing of biopharmaceuticals, Chem. Eng. Sci. 230 (2021) 116171. doi:10.1016/j.ces.2020.116171.

[28] P.R. Patnaik, Applications of neural networks to recovery of biological products, Biotechnol. Adv. 17 (1999) 477–488. doi:10.1016/S0734-9750(99)00013-0.

[29] T. Shojaeimehr, F. Rahimpour, Retention time modeling of short-chain aliphatic acids in aqueous ion-exclusion chromatography systems under several conditions using computational intelligence methods (artificial neural network and adaptive neuro-fuzzy inference system), J. Liq. Chromatogr. Relat. Technol. 41 (2018) 810–817. doi:10.1080/10826076.2018.1518846.

[30] T. Yang, C.M. Breneman, S.M. Cramer, Investigation of multi-modal high-salt binding ion-exchange chromatography using quantitative structure-property relationship modeling, J. Chromatogr. A. 1175 (2007) 96–105. doi:10.1016/j.chroma.2007.10.037.

[31] Š. Ukić, M. Novak, A. Krilić, N. Avdalović, Y. Liu, B. Buszewski, T. Bolanča, Development of gradient retention model in ion chromatography. Part III: Fuzzy logic QSRR approach, Chromatographia. 78 (2015) 889–898. doi:10.1007/s10337-015-2845-7.

[32] D. Pfister, L. Nicoud, M. Morbidelli, Continuous Biopharmaceutical Processes, Cambridge University Press, 2018. doi:10.1017/9781108332897.

[33] D. Baur, M. Angarita, T. Müller-Späth, M. Morbidelli, Optimal model-based design of the twin-column CaptureSMB process improves capacity utilization and productivity in protein A affinity capture, Biotechnol. J. 11 (2016) 135–145. doi:10.1002/biot.201500223.

[34] D. Baur, M. Angarita, T. Müller-Späth, F. Steinebach, M. Morbidelli, Comparison of batch and continuous multi-column protein A capture processes by optimal design,

Biotechnol. J. 11 (2016) 920–931. doi:10.1002/biot.201500481.

[35] D. Baur, J. Angelo, S. Chollangi, T. Müller-Späth, X. Xu, S. Ghose, Z.J. Li, M. Morbidelli, Model-assisted process characterization and validation for a continuous two-column protein A capture process, Biotechnol. Bioeng. 116 (2019) 87–98. doi:10.1002/bit.26849.

[36] A. Sellberg, M. Nolin, A. Löfgren, N. Andersson, B. Nilsson, Multi-flowrate Optimization of the Loading Phase of a Preparative Chromatographic Separation, Comput. Aided Chem. Eng. 43 (2018) 1619–1624. doi:10.1016/B978-0-444-64235-6.50282-5.

[37] J. Gomis-Fons, M. Yamanee-Nolin, N. Andersson, B. Nilsson, Optimal loading flow rate trajectory in monoclonal antibody capture chromatography, J. Chromatogr. A. 1635 (2021). doi:10.1016/j.chroma.2020.461760.

[38] A. Löfgren, M. Yamanee-Nolin, S. Tallvod, J.G. Fons, N. Andersson, B. Nilsson, Optimization of integrated chromatography sequences for purification of biopharmaceuticals, Biotechnol. Prog. 35 (2019) 1–7. doi:10.1002/btpr.2871.

[39] S. Ghose, D. Nagarath, B. Hubbard, C. Brooks, S.M. Cramer, Erratum: Use and optimization of a dual-flowrate loading strategy to maximize throughput in protein-A affinity chromatography (Biotechnology Progress (2004) 20 (830-840)), Biotechnol. Prog. 20 (2004) 1614. doi:10.1021/bp040029x.

[40] G. Guiochon, Preparative liquid chromatography, J. Chromatogr. A. (2002). doi:10.1016/S0021-9673(01)01471-6.

[41] L.K. Shekhawat, A.S. Rathore, An overview of mechanistic modeling of liquid chromatography, Prep. Biochem. Biotechnol. 49 (2019) 623–638. doi:10.1080/10826068.2019.1615504.

[42] S. Leweke, E. von Lieres, Chromatography Analysis and Design Toolkit (CADET), Comput. Chem. Eng. 113 (2018) 274–294. doi:10.1016/j.compchemeng.2018.02.025.

[43] S. Qamar, S. Bashir, S. Perveen, A. Seidel-Morgenstern, Relations between kinetic parameters of different column models for liquid chromatography applying core-shell particles, J. Liq. Chromatogr. Relat. Technol. 42 (2019) 16–30. doi:10.1080/10826076.2019.1570522.

[44] G. Wang, T. Briskot, T. Hahn, P. Baumann, J. Hubbuch, Estimation of adsorption isotherm and mass transfer parameters in protein chromatography using artificial neural networks, J. Chromatogr. A. 1487 (2017) 211–217. doi:10.1016/j.chroma.2017.01.068.

[45] G.R. Bolton, K.K. Mehta, The role of more than 40 years of improvement in protein A chromatography in the growth of the therapeutic antibody industry, Biotechnol. Prog. 32 (2016) 1193–1202. doi:10.1002/btpr.2324.

[46] D.G.B. Anurag S. Rathore, Mili Pathak, Guijun Ma, Re-use of protein A resin: fouling and economics, BioPharm Int. 4 (2015).

[47] A. Felinger, G. Guiochon, Comparison of the kinetic models of linear chromatography, Chromatographia. 60 (2004) 175–180. doi:10.1365/s10337-004-0288-7.

[48] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning The Elements of Statistical Learning, (2017). doi:10.1198/jasa.2004.s339.

[49] H. Narayanan, M. Sokolov, M. Morbidelli, A. Butté, A new generation of predictive models–the added value of hybrid models for manufacturing processes of therapeutic proteins, Biotechnol. Bioeng. (2019).

[50] H. Narayanan, L. Behle, M.F. Luna, M. Sokolov, G. Guillén-Gosálbez, M. Morbidelli, A. Butté, Hybrid-EKF: Hybrid Model coupled with Extended Kalman Filter for real-time monitoring and control of mammalian cell culture, Biotechnol. Bioeng. (2020) 1–12. doi:10.1002/bit.27437.

Figure Legends

Figure 1: Process conditions in the BT runs together with indication of the column and particle dimension (Column Parameter), experimental time (t_{exp}) and sampling time (t_{samp}) for *in-silico* (A) and experimental dataset (B).

(A)

Column Parameter		A_{col} = 1.76 cm^2, L_{col} = 10 cm, d_p = 85 µm			
Partition	**S.No.**	**Q (mL/min)**	**C_{in} (mg/mL)**	**t_{exp} (min)**	**t_{samp} (min)**
Calibration	BTC #1	6	0.8	200	5
	BTC #2	8	0.8	150	5
	BTC #3	10	0.8	150	5
	BTC #4	6	1	150	5
	BTC #5	8	1	150	5
	BTC #6	10	1	150	5
	BTC #7	6	1.2	150	5
	BTC #8	8	1.2	125	5
	BTC #9	10	1.2	125	5
Test Interpolation	BTC #10	7	0.9	150	5
	BTC #11	9	0.9	150	5
	BTC #12	7	1.1	150	5
	BTC #13	9	1.1	125	5
Test Extrapolation	BTC #14	4.5	0.6	600	5
	BTC #15	12.5	0.6	150	5
	BTC#16	4.5	1.5	150	5
	BTC #17	12.5	1.5	125	5

(B)

Column Parameter		A_{col} = 0.196 cm^2, L_{col} = 10 cm, d_p = 85 µm			
Partition	**S.No.**	**Q (mL/min)**	**C_{in} (mg/mL)**	**t_{exp} (min)**	**t_{samp} (min)**
Calibration	BTC #1	1.5	0.30	230	10
	BTC #2	1.0	0.30	340	10
	BTC #3	0.5	0.30	690	10
	BTC #4	0.5	0.60	200	3
	BTC #5	1.5	0.60	80	3.5
	BTC #6	1.0	0.60	120	3.5
Test Interpolation (Same Col.)	BTC #7	1.0	0.43	240	4
	BTC #8	1.5	0.43	160	4
	BTC #9	0.5	0.43	480	4
Test Extrapolation (Diff Col.)	BTC #10	1.0	0.34	200	8
	BTC #11	1.5	0.34	230	4
	BTC #12	1.0	0.34	250	6
	BTC #13	1.5	0.42	130	2
	BTC #14	1.0	0.42	120	3.5
	BTC #15	1.0	0.42	170	3.5

Figure 2: (A) Generalized schematic representation of the modeling framework workflow for training LKM and hybrid model (B) Schematic visualization of the two hybrid model formulations: Lumped-Hybrid and MTI-Hybrid. It is noted that the NN architecture represented here is an illustration and does not indicate the optimal NN structure used in the hybrid models.

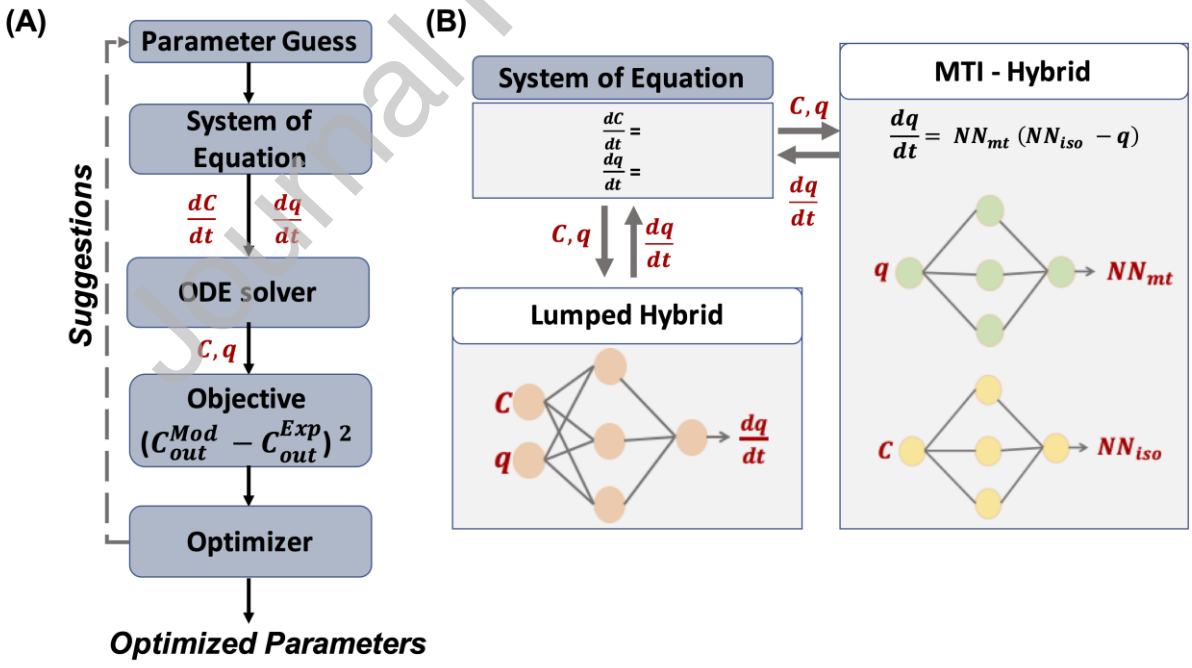

Figure 3: (A) In silico observed concentration values against the predictions by LKM and Lumped-Hybrid model for the test set, "Test Interpolation" (c.f. Figure 1A). (B) Comparison of LKM and Lumped-Hybrid model predictions demonstrated using an

exemplary BT run from the "Test Interpolation" set (c.f. Figure 1A). (C) Concentration profiles inside the column at different instances of time as predicted by the LKM and Lumped-Hybrid model compared against the *in-silico* profile inside the column.

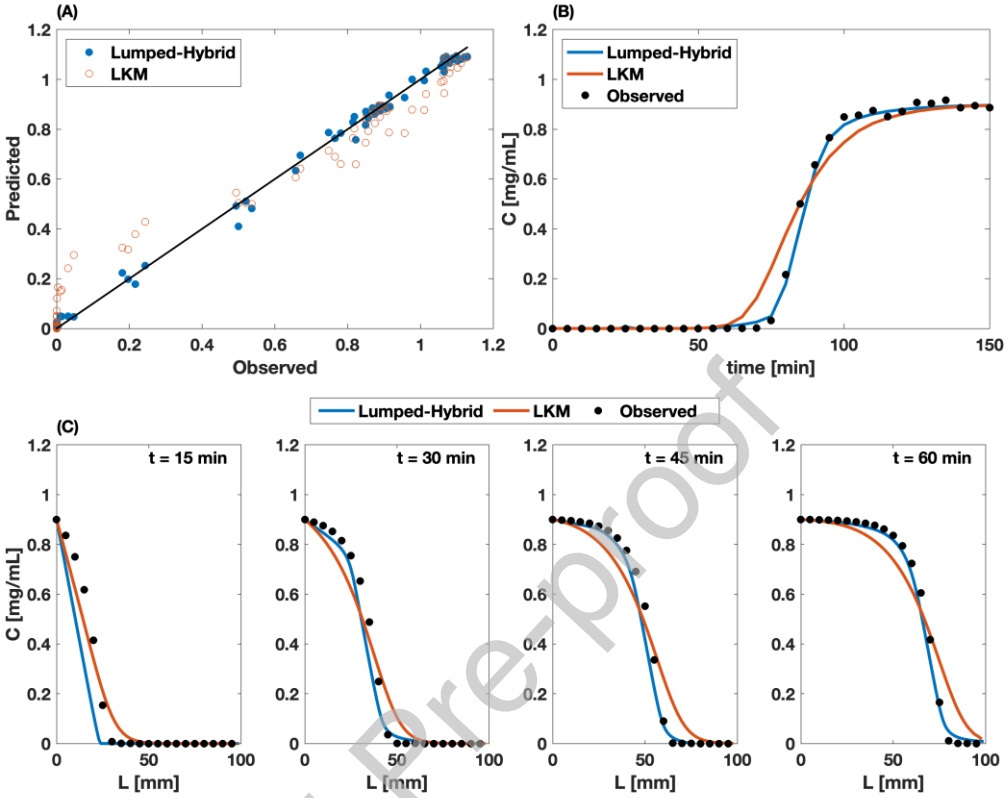

Figure 4: Comparison of the BT runs predicted by the LKM and Lumped-Hybrid model with the *in-silico* extrapolation data set, "Test Extrapolation", in Figure 1A.

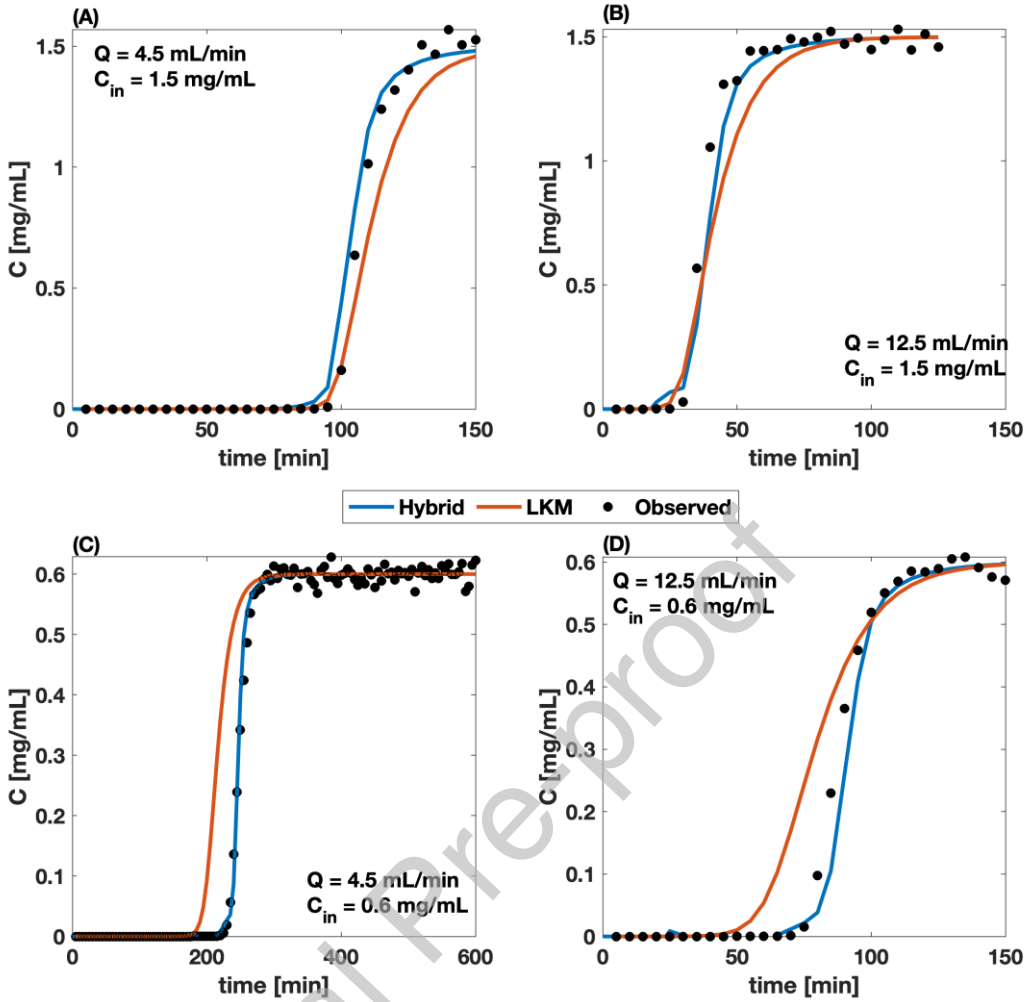

Figure 5: Comparison of the performance of Lumped-Hybrid model and LKM in predicting *in-silico* BT experiments generated using different underlying adsorption isotherms: (A) Langmuir- Freundlich, (B) Temkin, (C) Toth and (D) Bi Langmuir for process condition indicated in Figure 1A "Test Interpolation". Parity plots are shown in the upper left corner of each figure where the x-axis indicates the observed concentrations and the y-axis represent the model predicted concentrations.

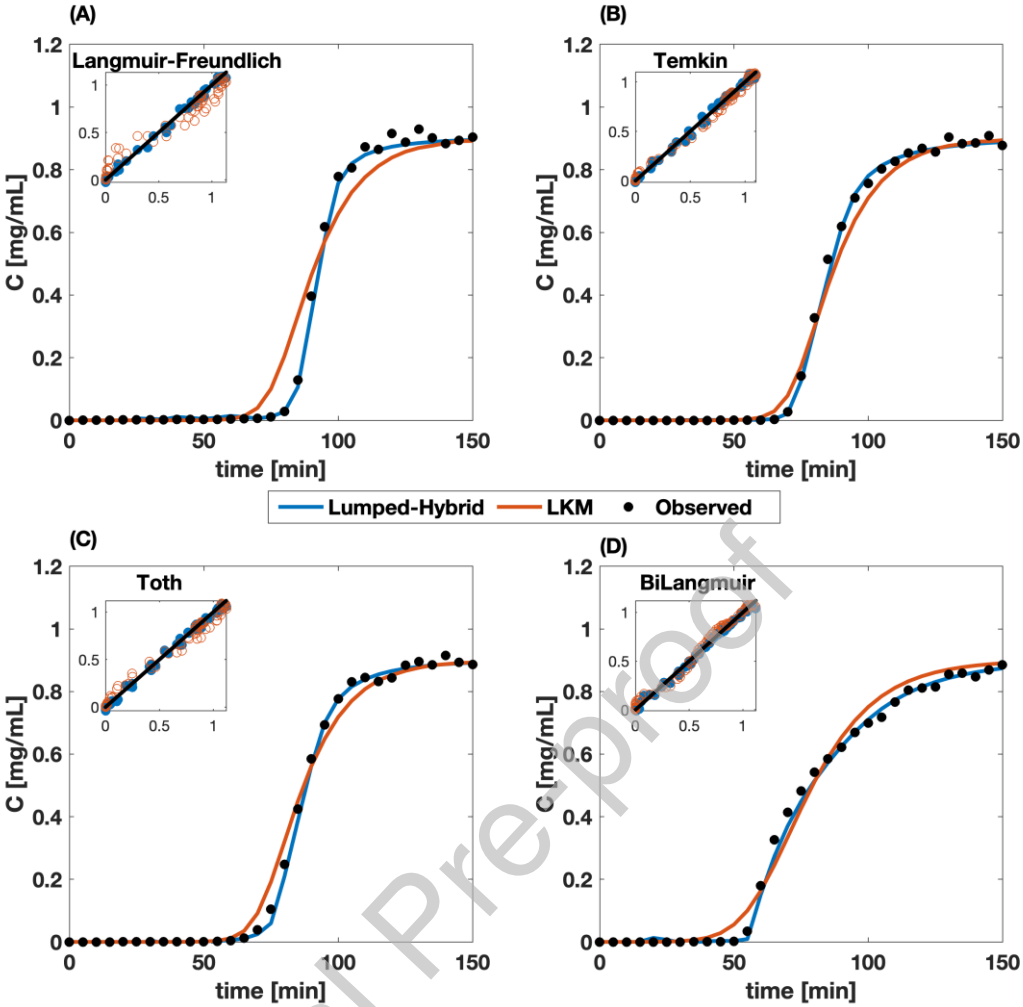

Figure 6: Comparison of the adsorption isotherms simulated by the MTI-Hybrid model and the original ones used in generating the *in-silico* BT runs.

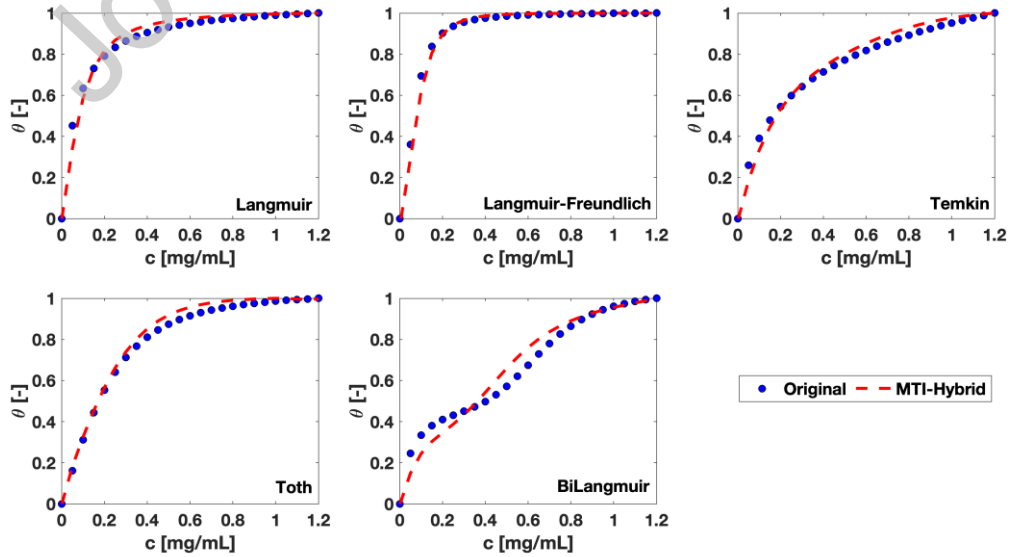

Figure 7: (A, B) Comparison of LKM and Lumped-Hybrid model predictions with experimental BT curves for the "Test Interpolation (Same Col.)" in Figure 1B. (C, D) Same comparison for the "Test Extrapolation (Different Col.)" in Figure 1B. In the upper left corner of each figure is shown the corresponding parity plots where the x-axis indicates the observed concentrations and the y-axis represent the model predicted concentrations