Comparing in silico flowsheet optimization strategies in biopharmaceutical downstream processes

Keulen, Daphne; Apostolidi, Myrto; Geldhof, Geoffroy; Le Bussy, Olivier; Pabst, Martin; Ottens, Marcel

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

RESEARCH ARTICLE

Bioseparations and Downstream Processing

BIOTECHNOLOGY PROGRESS

Comparing in silico flowsheet optimization strategies in biopharmaceutical downstream processes

Daphne Keulen[1] | Myrto Apostolidi[1] | Geoffroy Geldhof[2] | Olivier Le Bussy[2] | Martin Pabst[1] | Marcel Ottens[1]

[1]Department of Biotechnology, Delft University of Technology, Delft, The Netherlands

[2]GSK, Technical Research & Development, Rixensart, Belgium

Correspondence
Marcel Ottens, Department of Biotechnology, Delft University of Technology, Van der Maasweg 9, 2629 HZ Delft, the Netherlands.
Email: m.ottens@tudelft.nl

Present address
Myrto Apostolidi, Heineken, Leiden, The Netherlands.

Funding information
GlaxoSmithKline Biologicals SA

Abstract

The challenging task of designing biopharmaceutical downstream processes is initially to select the type of unit operations, followed by optimizing their operating conditions. For complex flowsheet optimizations, the strategy becomes crucial in terms of duration and outcome. In this study, we compared three optimization strategies, namely, simultaneous, top-to-bottom, and superstructure decomposition. Moreover, all strategies were evaluated by either using chromatographic Mechanistic Models (MMs) or Artificial Neural Networks (ANNs). An overall evaluation of 39 flowsheets was performed, including a buffer-exchange step between the chromatography operations. All strategies identified orthogonal structures to be optimal, and the weighted overall performance values were generally consistent between the MMs and ANNs. In terms of time-efficiency, the decomposition method with MMs stands out when utilizing multiple cores on a multiprocessing system for simulations. This study analyses the influence of different optimization strategies on flowsheet optimization and advices on suitable strategies and modeling techniques for specific scenarios.

KEYWORDS
artificial neural networks, chromatography, filtration, mechanistic modeling, superstructure-based optimization

1 | INTRODUCTION

Downstream processing is of major importance for delivering the required quality and quantity of a biopharmaceutical product, which has to meet the strict standards by regulatory authorities.[1] The downstream process is a substantial expense of the overall manufacturing costs, therefore, an efficient and cost-effective process is crucial. One of the major, most costly, and essential purification techniques is chromatography, which is capable to achieve very high product purities.[2] Eventually, the combination of purification steps will determine the overall process performance. Therefore, developing a purification process is a challenging task, involving many variables, such as type and sequential order of purification techniques, operating conditions, and costs.[3,4] A comprehensive overview of the different strategies in downstream process development together with the latest breakthroughs was given recently by Keulen, et al.[5] Finding an optimal purification process at an early stage of the process design are desirable in terms of costs, quality, and development time. Flowsheet optimization evaluates all process possibilities in silico, which can support the decision-making for an early process design. For many

years, flowsheet optimization has been applied to design chemical processes, therefore, it is well-known in the field of process systems engineering.[6,7]

Around the 1970s, the first articles were published about process design synthesis.[8,9] Siirola, et al.[8] developed a general computer-aided process synthesizer that was able to select process equipment and the system configurations. Umeda, et al.[9] presented an integrated optimization approach to optimize two alternative routes for a distillation system. Over the past five decades, the field of superstructure-based optimizations has evolved greatly, along with the intensified computing possibilities.[10] Mencarelli, et al.[6] provides an adequate overview of superstructure-based optimization history, superstructure representation types, and modeling strategies. Most superstructure-based optimizations applied in chemical engineering are related to reactor networks,[11] distillation processes,[12] and heat exchangers.[13] Several programs are available to perform a chemical superstructure-based optimization, for example, P-graph,[14] Pyosyn,[15] and Super-O.[16] As most of these chemical process simulations are based on first-principle models this can be computationally time-consuming, therefore the interest in employing surrogate models for optimization purposes increased. In 1998, Altissimi, et al.[17] already showed the value of replacing a first-principle model with a surrogate model for optimization purposes. Afterwards, more research followed on using surrogate or meta-models for superstructure or complex optimization purposes.[18–22]

Despite the biopharmaceutical industry only emerged about 40 years ago, this industry is advancing rapidly and shifting toward Industry 4.0.[23–25] Industry 4.0 desires to entirely digitalize the manufacturing process, aiming to implement and combine model-based process development techniques with efficiently stored monitored data. Hence, realizing the utilization of Digital Twins, which are digital models of the real process and enable to directly control the real process.[26,27] In this way, more knowledge can be acquired about the processes, which is in compliance with the Quality by Design guidelines.[28,29] A general biopharmaceutical process consists of an upstream and downstream part, in which the downstream part focuses on the purification of the biopharmaceutical. The purification steps can be subdivided into capture, intermediate, and polishing steps as shown in Figure 1. The main purpose of the capture and intermediate steps is to concentrate, isolate, and stabilize the product, and remove the majority of the impurities. While the subsequent polishing steps target high purity values.[2]

Chromatographic MMs have been around for several years, and industry is gradually adopting these methods.[30,31] Lately, advances have been made to faster and more efficiently determine the adsorption isotherms, which are needed as input parameters for the mechanistic model.[32–34] Likewise, several research has been published to determine adsorption isotherms for complex mixtures.[35–37] And more recently, Disela, et al.[38] characterized the host cell proteome of two universal *E. coli* strains based on mass spectrometry data, which approach can be used for initial decision-making on process development. Not only the techniques and methods to determine the adsorption isotherm are making progress, also the MMs are advancing in terms of speed and accuracy. Meyer, et al.[39] applied a computational more efficient method for the spatial discretization and obtained a
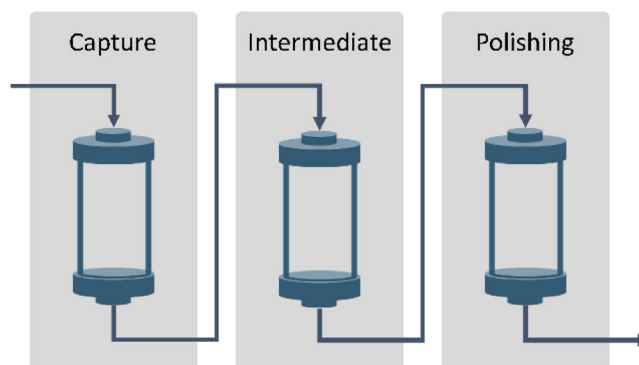

FIGURE 1 Simplified schematic overview of the chromatography steps in a biopharmaceutical downstream process, the sequence can also have less or more chromatography operations depending on the process. The capture step aims to concentrate, isolate, and stabilize the product, together with the intermediate step, their target is to remove the bulk impurities. The main purpose of the polishing step is to attain high product purities.

speed-improvement of at least 20 times, for higher precision it even improves over 100 times compared to the open-software CADET.[40] Their chromatography model was recently extended by Breuer, et al.,[41] which applied a similar method to the particle mass balance. Rao, et al.[42] developed a 3-D model to simulate the chromatography process with very high precision to acquire knowledge about the complex transport mechanism. Moreover, hybrid modeling, using artificial intelligence (AI) in combination with mechanistic modeling, can overcome certain limitations of both modeling techniques.[43,44] Narayanan, et al.[45] employed artificial neural networks (ANNs) for fitting the solid-phase mass balance, which reduced the model complexity, and an improved accuracy compared to the conventional mechanistic model was observed. Accordingly, this progress in experimentally determining model-parameters, improving the MMs, and making use of hybrid modeling, is advantageous for digitalization of the downstream process and likewise for optimization purposes.

As described previously, flowsheet optimization enables to screen the overall design space and finding the optimal purification process at an early development stage. Process systems engineering recognized the added value of superstructure-based optimization for chemical processes. Also for biochemical processes, it is essential to optimize the integrated processing steps to discover the most optimal process globally.[46] Liu and Papageorgiou[47] developed a data-driven optimization framework to find the best process according to economical and certain performance objectives. However, the data for each optional processing steps is already provided and not generated internally. This type of optimization is known as biopharmaceutical manufacturing process optimization, usually based on mixed integer programming techniques.[48–51] Though, these optimizations do not use detailed mechanistic modeling techniques, they are either data-driven or using surrogate models to represent the unit operations. In the work of Nfor, et al.,[3] a top-to-bottom optimization approach is performed that evaluates the performance of each unit operation at each level and disregards the least promising options. As the influence of sequential steps is not incorporated in this approach, it might overlook the most

promising sequence(s). Therefore, Huuk, et al.[4] performed an integrated two-step ion-exchange chromatography optimization. Subsequently, Pirrung, et al.[52] performed a flowsheet optimization having a maximum of three chromatography steps (e.g., cation exchange, hydrophobic interaction, and mixed-mode) including a buffer exchange if needed, and simultaneously optimizing each flowsheet. In their work, ANNs functioned as surrogate model for the MMs during the global optimization to find starting conditions for the local optimization, and so reducing the overall optimization time. However, the ANNs were infrequently able to find realistic results and for the subsequent local optimization the MMs were used, which was the most time-consuming part of the overall optimization.[52,53] In our previous work, we extended this method by including the mass of each component as a variable and using more data to increase the ANN accuracy.[54] Subsequently, we compared ANNs, functioning as surrogate models, versus MMs for flowsheet optimization to select the 'most promising sequences' during the global optimization. Only the 'most promising sequences' were further optimized through local optimization using MMs. The ANNs selected three out of four best flowsheets and reduced the overall computational time by 50%. However, for more complex flowsheet optimizations (e.g., including more unit operations and/or larger sequences) or when considering more components, not only the modeling technique (e.g., MMs or surrogate models) matters, but also the optimization strategy might play a significant role in the overall flowsheet optimization. Hence, what optimization strategy is most useful in terms of outcome, complexity, and time-efficiency?

In this article, we compared three different optimization strategies: simultaneous optimization, top-to-bottom approach, and superstructure decomposition, to evaluate which strategy would be most beneficial in terms of outcome, complexity, and time-efficiency when performing a
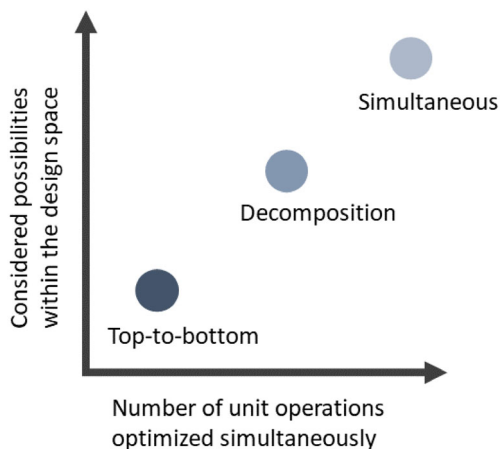

FIGURE 2 Visualization of the difference between the chosen optimization strategies; top-to-bottom, superstructure decomposition, and the simultaneous strategy. The x-axis shows the number of unit operations being optimized simultaneously during flowsheet optimization. While, the y-axis correspondingly shows that more options in the design space are explored when the connection between chromatography steps is also considered, which is not taken into account for the top-to-bottom approach as it individually optimizes each chromatography step.

complex flowsheet optimization. Simultaneous optimization involves optimizing all parameters simultaneously, top-to-bottom approach optimizes parameters sequentially from the initial to the final unit operation, and decomposition of the superstructure involves breaking down the process into smaller parts and optimizing each part separately. These strategies were chosen based on the difference in number of unit operations being optimized simultaneously and so the overall considered possibilities within the design space as indicated in Figure 2. For example, the top-to-bottom approach might overlook promising sequences, as it lacks a focus on optimizing the connections between chromatography steps. Additionally, for each optimization strategy the MMs and the ANNs are employed to evaluate their performance on a more complex optimization. In this complex flowsheet optimization, we included an optional buffer exchange between the chromatography steps, described by a filtration MM. This gives a total combination of 39 flowsheets to be evaluated.

2 | MATERIALS AND METHODS

2.1 | Flowsheet optimization workflow

First, the superstructure was generated considering a maximum of three chromatography steps and a dilution or buffer exchange by Tangential Flow Filtration (TFF) between the chromatography operations. This gives a maximum sequence of five unit operations, and at least one unit operation is needed for the purification. To generate this superstructure, confirming the defined conditions, the mathematical problem is formulated as.

$$y = [y_1, y_2, ..., y_n] \tag{1}$$

$$s.t. \sum y \geq 1 \tag{2}$$

$$\begin{aligned} &For\ i\ is\ odd: \\ &y_i = 1, 2, 3 \\ &y_i \neq y_{i+2}\ for\ all\ y_i > 0 \end{aligned} \tag{3}$$

$$\begin{aligned} &For\ i\ is\ even: \\ &y_i = 4, 5 \end{aligned} \tag{4}$$

$$\begin{aligned} &For\ i = 2, 3, ..., n: \\ &if\ y_i > 0,\ then\ y_{i-1} > 0, \end{aligned} \tag{5}$$

where y is the process configuration, in which n, in this case $n = 5$, is the length of the vector. The variable $y_i \in \{0, 1, 2, 3, 4, 5\}$ represents the value of the ith element of vector y. The first statement, Equation 1, defines the set of all possible vectors y, where each element is an integer number between 0 and 5, which in this study represents the considered unit operations: none, CEX, AEX, HIC, dilution, and filtration, respectively. The second statement, Equation 2, guarantees that the sequence includes at least one unit operation. The third and fourth statements, Equation 3 and 4, specify that only at odd positions in the sequence, a chromatography step is present, while for even
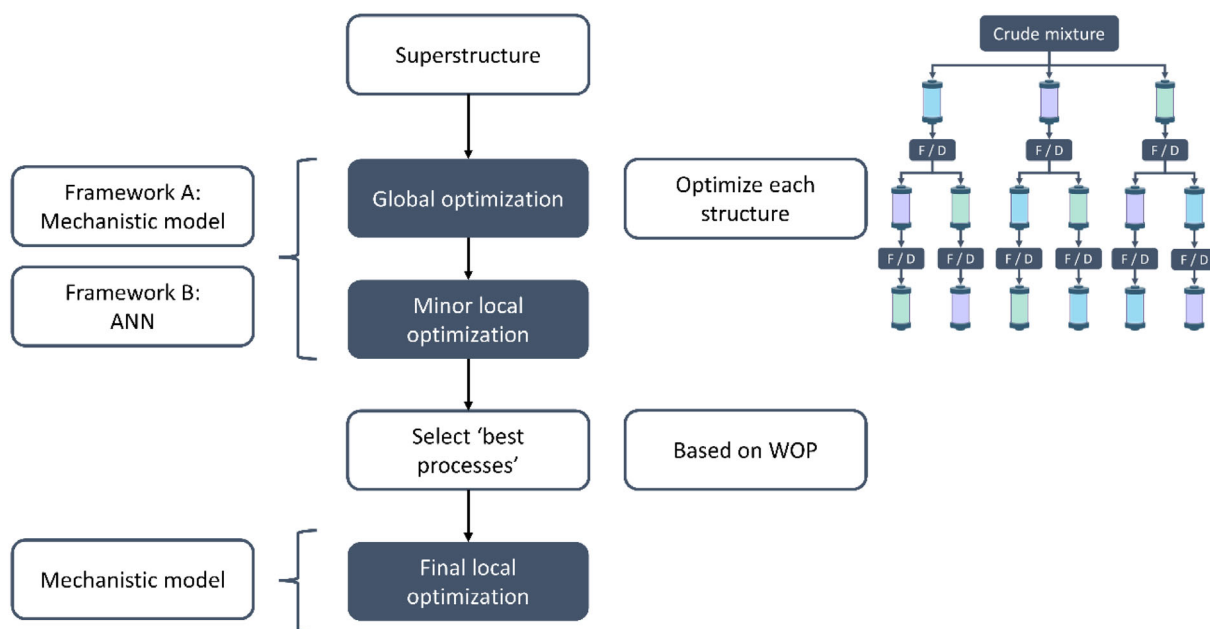
FIGURE 3 Within the superstructure, as depicted in the upper right figure, each flowsheet is initially optimized globally to identify the most optimal processes. F/D indicates the option to have either a filtration (F) or dilution step (D). Subsequently, these selected processes are finetuned using a final local optimization step. For the global optimization, Framework A uses MMs and Framework B uses ANNs.

number positions, either a dilution or a filtration step is employed. Furthermore, statement three ensures that each chromatography mode appears only once in the sequence. The conditional constraint in Equation 4 is applicable to all positions in the sequence, except the first position. It enforces that any occupied position in the sequence must be preceded by another occupied position. This guarantees that there are no isolated modes in the sequence and requires all modes to be connected.

The flowsheets consisting of a filtration operation to perform the buffer exchange step are modeled as a nested optimization,[55] which means that the outer optimization involves matching the chromatography steps with their respective variables, while the inner optimization focuses on optimizing the filtration step. So, for each evaluation of the outer optimization, the filtration step is always optimized internally. As the filtration model is less complex and described by ordinary differential equations (ODEs) with respect to time, it has a significantly shorter solving time compared to the chromatography model. The same flowsheet optimization workflow, as presented in our previous article,[54] was applied as shown in Figure 3. First, a global and minor local optimization was performed according to certain objective(s) and constraint(s), these are described in 2.5. Case study. For this part, either MMs or ANNs were used for the chromatography steps. After this global and minor local optimization, the most promising sequences were selected based on the weighted overall performance (WOP), which is described as follows:

$$WOP = 0.5 * purity + 0.3 * yield + 0.2 * (100 - buffer\ consumption),$$
$$(6)$$

where the calculation of purity (%) involves dividing the product amount by the total amount of proteins present in the product-pool. The yield (%) is determined by the total amount of product recovered divided by the loaded amount of product. The buffer consumption typically ranges from 1 to 50 (L/$g_{product}$). Subtracting this buffer consumption from 100 aligns it with the purity and yield ranges, and ensures that higher WOP values indicate less buffer consumption.

The selected processes were further locally optimized using the simultaneous strategy with MMs, the outcome of preceding minor local optimization was used as initial guess for the final local optimization. This flowsheet optimization workflow was applied to all three optimization strategies, the difference is the manner of solving the superstructure. Each strategy was evaluated for using either the MMs or ANNs for the global and minor local optimization. The strategies (e.g., simultaneous optimization, top-to-bottom approach, and decomposition of the superstructure) are separately described in the following sections.

2.1.1 | Strategy I: Simultaneous flowsheet optimization

The simultaneous flowsheet optimization is the same as applied in Keulen, et al.[54] In this strategy, all parameters are optimized simultaneously, which means that the total number of variables linearly increases with the number of chromatography steps present in the sequence, as shown in Figure 4. For example, if five optimization variables are considered and the sequence consists of two
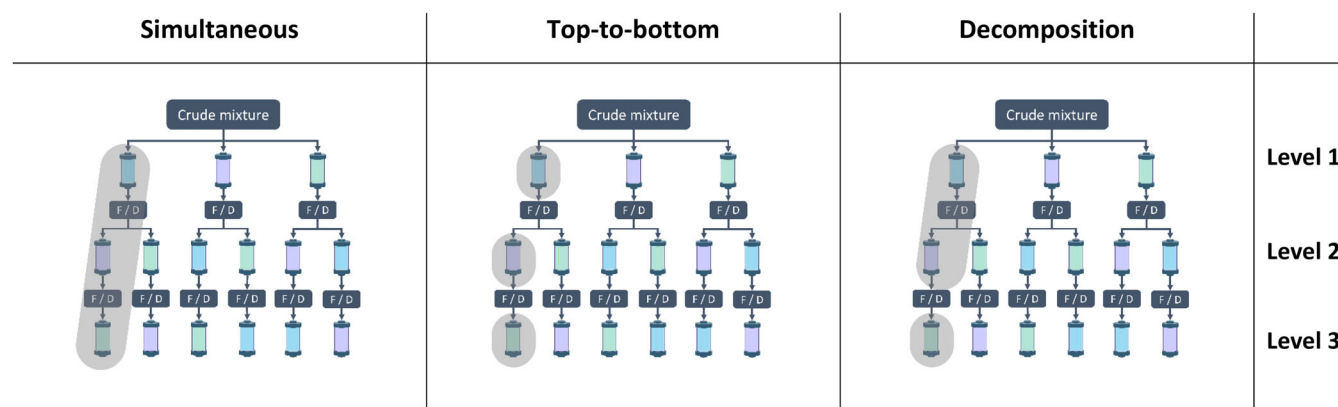
FIGURE 4 Schematic representation of each optimization strategy indicating with gray planes which unit operations are optimized simultaneously during the optimization. F/D indicates the option to have either a filtration (F) or dilution step (D).

chromatography steps, 10 variables have to be optimized in total. For three unit operations, this will lead to 15 variables to be optimized.

2.1.2 | Strategy II: Top-to-bottom approach

The top-to-bottom approach, based on the work of Nfor, et al.,[3] evaluates the superstructure by each level, see Figure 4. The first level optimizes the first unit operation individually. After optimizing the first level, the initial constraint assesses whether the optimal process has been achieved (e.g., purity >99% and yield >95%). The second constraint assesses if the optimized unit operation satisfies the minimal requirements (e.g., purity >20% and yield >40%) to continue to the next level, otherwise the flowsheets, starting with this type of unit operation, will be disregarded. All options present in the second level are also optimized individually. Subsequently, the overall sequence of two chromatography steps, including the dilution or the filtration operation, are simulated. The outcome of these flowsheets is evaluated by the previously described constraints, however, the second constraint is only satisfied if the purity and yield are higher than 40%. If the optimal process has not been identified yet, the optimization will continue to the third level, which operates in the same manner as the second level. If the optimal performance is not achieved after three levels, the best out of all these evaluated flowsheets can still be chosen, as all outcomes are stored. The constraints between the levels can be easily adapted to a different number and/or different performance measurements to be assessed.

2.1.3 | Strategy III: Superstructure decomposition

In the previous study, we observed that approximately 60% of the total optimization time, whether employing MMs or ANNs, was dedicated to optimizing sequences of three unit operations.[54] This aligns with the fact that the maximum number of function evaluations increases with the number of variables to be optimized.[56] Accordingly, the question raised; does the third unit operation have a significant impact on the previous unit operations? Followed by, is it really necessary to optimize the whole process simultaneously or can we decompose the superstructure when optimizing larger sequences? In chemical engineering, different formats of decomposing the superstructure have been applied.[6,57–59] In this study, this strategy is a combination of the simultaneous and top-to-bottom approach as shown in Figure 4. The superstructure is ordered in such a way that the sequences consisting of the same first three unit operations are sequential in order of length. The sequence consisting of three unit operations is optimized first, subsequently, the outcome of the third unit operation (e.g., second chromatography step) is used as input for the last chromatography step, which is optimized individually. After individually optimizing the third chromatography step, the overall sequence of five unit operations is simulated, similar to the workflow of top-to-bottom approach. In this way, only a maximum of two chromatography steps is optimized simultaneously, making the overall optimization more time-efficient compared to simultaneous optimization.

2.2 | Chromatography

2.2.1 | Mechanistic model

The same chromatographic MM from previous work was used in Ref. [54]. The equilibrium transport dispersive model in combination with the linear driving force described the dynamic adsorption behavior during the chromatographic separation process as.

$$\frac{\partial C_i}{\partial t} + F\frac{\partial q_i}{\partial t} = -u\frac{\partial C_i}{\partial x} + D_{L,i}\frac{\partial^2 C_i}{\partial x^2}, \tag{7}$$

$$\frac{\partial q_i}{\partial t} = k_{ov,i}\left(C_i - C^*_{eq,i}\right), \tag{8}$$

$$k_{ov,i} = \left[\frac{d_p}{6k_{f,i}} + \frac{d_p^2}{60\varepsilon_p D_{p,i}}\right]^{-1}, \tag{9}$$

TABLE 1 Overview of hyperparameters for each chromatography mode and the applied parameter space.

	CEX	AEX	HIC
Hyperparameters			
Batch size	512	512	512
Epochs	200	500	500
Number of hidden layers	2	2	2
Number of neurons	50	50	50
Learning rate	0.01	0.01	0.01
Parameter space			
Gradient length (CV)	1–10	1–10	1–10
Loading factor (CV)	0.05–5	0.05–5	0.05–5
Mass (g)	2e-5–0.39	2e-5–0.39	2e-5–0.39
Loading (g/L CV)	4e-6–7.8	4e-6–7.8	4e-6–7.8
Initial salt concentration (mM)	1–200	1–200	350–500
Final salt concentration (mM)	100–1200	100–1200	5–200
Lower cut point (%)	1–80	1–80	1–80
Upper cut point (%)	20–99	20–99	20–99

where C_i is the concentration in the liquid phase, q_i the concentration in the solid phase, and $C_{eq,i}^*$ is the liquid phase concentration in equilibrium with the solid phase. The phase ratio, F, is defined as $F = (1 - \varepsilon_b)/\varepsilon_b$, where ε_b is the bed porosity. u represents the interstitial velocity of the mobile phase and D_L is the axial dispersion coefficient. Time and space are indicated by t and x respectively. $k_{ov,i}$ is the overall mass transfer coefficient defined as a summation of the separate film mass transfer resistance and the mass transfer resistance within the pores.[60] Here, d_p is the particle diameter, ε_p is the intraparticle porosity, and D_p is the effective pore diffusivity coefficient. The first term represents the film mass transfer resistance, $k_f = D_f Sh/d_p$, in which D_f is the free diffusivity and Sh is the Sherwood number. More information on the MM can be found in a previous study.[61] Moreover, we used the linear multicomponent mixed-mode isotherm, as formulated by Nfor, et al.[62] and described in Appendix A.

2.2.2 | Artificial neural networks

The ANNs were created as described previously.[54] In this work, we applied the same input variables (e.g., mass of each component, amount of loading in column volume (CV), gradient length, initial and final salt concentrations, and the lower and upper cut points in percentage of the peak maximum) and output variables (e.g., mass of each component, volume, salt concentration and each cut point in CV, salt concentration). The parameter space was based on prior-knowledge of biopharmaceutical downstream processes.[63] The data consisted of 10.000 sample points divided into 70% for training, 15% for validation, and 15% for testing. Based on previous work, the same hyperparameters were used as starting point for developing the ANNs. Out of 10 trained, validated, and tested ANNs, the best one was chosen based on R^2 and root mean squared error (RMSE) values. An overview

of the final used hyperparameters and applied parameter space is given in Table 1.

2.3 | Filtration mathematical model

An ultrafiltration/diafiltration (UF/DF) mathematical model was developed to describe the buffer exchange if a filtration step was used between the chromatography steps. This model consists of first-order differential equations involving the feed solution volume (V) and added diluent volume (V_w) over time, and the solute concentrations (C_i) and the salt concentration (C_s) over time.[64] The system of mass balances for which the proteins are completely retained by the membrane is written as follows:

$$\frac{dV}{dt} = (\alpha - 1)JA, \tag{10}$$

$$\frac{dV_w}{dt} = \alpha JA, \tag{11}$$

$$\frac{dC_i}{dt} = \frac{C_i}{V}(\sigma_i - \alpha)JA, \tag{12}$$

$$\frac{dC_s}{dt} = \frac{C_s}{V}(\sigma_s - \alpha)JA, \tag{13}$$

where J is the permeate flux and A is the membrane area. σ_i and σ_s are the rejection coefficients, in this work all proteins were significantly larger than the membrane pores, hence σ_i was equal to one. While the salts could flow through and therefore σ_s was equal to zero. α is the ratio between the diluent flowrate (u) and the permeate flowrate and given as

$$\alpha = \frac{u}{JA}. \qquad (14)$$

The operation was performed in an ultrafiltration with variable volume diafiltration (UFVVD), therefore α can range between 0 and 1. A value close to zero indicates to operate in an UF mode, while close to one DF occurs. The flux was defined by the osmotic pressure model as.

$$J = \frac{\Delta P_{TM} - \Delta \pi}{\mu * R_m}, \qquad (15)$$

where $\Delta \pi$ denotes the osmotic pressure difference and μ is the solution viscosity. ΔP_{TM} is the transmembrane pressure, which denotes the pressure difference between both sides of the membranes and acts as the driving force for the flux through the membrane. In the osmotic pressure model, the solute wall concentration is considered as a variable and increases usually over time, therefore the osmotic pressure changes, which directly impacts the flux negatively over time. The initial solute wall concentration, $C_{i,w,0}$, is predicted by solving the following equation:

$$k_0 \ln \frac{C_{i,w,0}}{C_{i,0}} = \frac{\Delta P_{TM} - \Delta \pi}{\mu \, Rm}, \qquad (16)$$

where $C_{i,0}$ represent the initial concentrations in solution and k_0 is the initial mass transfer coefficient. The change of the wall concentration over time was included in the mass balance systems as.[64]

$$\frac{dC_{i,w}}{dt} = \frac{\frac{k_0}{C_i} - \ln \frac{C_{i,w}}{C_i} \frac{dk}{dC_i}}{\frac{k_0}{C_{i,w}} + \frac{1}{\mu \, Rm} \frac{\Delta \pi}{dC_{i,w}}} \frac{dC_i}{dt}, \qquad (17)$$

where the change of osmotic pressure is found by differentiating Equation 17 with respect to $C_{i,w}$. Similarly, differentiating the mass transfer to C_i gives dk/dc_i. The mass transfer coefficient is viscosity dependent and given as follows:[64]

$$k = k_0 \left(\frac{\mu}{\mu_0} \right)^{-\frac{1}{\delta}}, \qquad (18)$$

where μ is the solution viscosity and μ_0 is the viscosity of the pure solvent. In Appendix B, additional information is provided on the transmembrane pressure, osmotic pressure, second virial coefficient (B_{22}), the mass transfer correlations, and determination of the initial membrane resistance through a water flux wet experiment. Moreover, the filtration model was validated for an UF/DF wet experiment using a Bovine Serum Albumin (BSA) solution, more information can also be found in Appendix B.

2.4 | Numerical methods

The same numerical methods as applied in previous work were used, only minor adjustments were made.[54] All codes are written in Python (version 3.8.5). An overview of the Python libraries used is provided in Appendix C. The computations were performed on a Dell Precision 5820 Tower XCTO having a 3.7G Intel Xeon processor of 3.7 GHz, 10C, and a 8GB Nvidia Quadro. Multiple cores were used to execute the simulations most efficiently; however, the number of cores varied depending on the simulation.

Dynamic chromatography column model.

The Method of Lines is applied for the spatial discretization, using a fourth-order central difference scheme for both first and second-order derivatives with respect to space, to transfer partial differential equations into ODEs with respect to time. The LSODA (Livermore Solver for Ordinary Differential Equations) algorithm from the *scipy.integrate* package is used to solve the ODEs, this method automatically switches between the nonstiff Adams method and the stiff BDF method.[65]

Optimization.

The *scipy.optimize* package was employed for the optimization, whereas the *differential_evolution* algorithm was used for the global optimization and Nelder–Mead algorithm for the local optimization. For global optimization, the maximum number of iterations was 6 and a population size of 5 for MMs, while for ANNs, the maximum number of iterations was 8 with a population size of 8. Latin hypercube sampling was used to initialize the population. The initial local optimization had a maximum of 5 iterations. The relative and function tolerances for both global and local optimizations were set to 1e-2. The final local optimization allowed a maximum of 50 iterations. Limited ANN accuracy can lead to varied mass predictions and affect the performance measurements. Over predicted masses were set to the injected mass. The lower cut point ranged from 1% to 80% of the peak maximum, while the upper cut point ranged from 20% to 99% of the peak maximum. Initial salt concentrations were between 1 and 150 mM for CEX and AEX, 100–500 mM for HIC using MM, and 350–500 mM for HIC using ANN. Final salt concentrations were between 160 and 1200 mM for CEX and AEX, 5–300 mM for HIC using MM, and 5–200 mM for ANN. The gradient length varied from 1 to 10 CV. For optimizing the filtration operation, the Nelder–Mead algorithm with standard settings was employed.

Artificial neural networks.

The Keras Module (version 2.10.0) of TensorFlow (version 2.10.1) were used to create the ANNs, these are open-source libraries compatible with the Python programming language. The ANN structure, optimized with a learning rate of 0.01 using *keras.optimizers.Adam* and defined using *keras.models.Model*, employed data scaling via the *sklearn.preprocessing.MinMaxScaler* module. The optimizer loss function used the 'mean_squared_error' metric. Randomized data was generated by applying the Latin hypercube sampling method from the *pyDOE* package.

2.5 | Case study

The case study focused on a monoclonal antibody product of interest and referred to as protein 1, and eight impurities (referred to as

TABLE 2 Input parameters of each protein used for the flowsheet optimization, in which K_{eq} is the equilibrium constant, v is the stoichiometric coefficient of salt counter ions (characteristic charge), and n is the hydrophobic interaction stoichiometric coefficient. Protein 1 = monoclonal antibody, protein 2 = Moesin, protein 3 = Chitotriosidase, protein 4 = Legumain, protein 5 = Thioredoxin reductase, protein 6 = Bovine Serum Albumin, protein 7–9 = artificial proteins.

Protein			1	2	3	4	5	6	7	8	9
Initial concentration		g/L	1.5	0.9	0.8	1.2	1.2	0.8	0.3	0.9	1.4
Molecular weight		kDa	145.6	68	51.5	56.2	54.5	56.2	70	60	90
CEX	K_{eq}	(−)	8.5	500.8	604.2	0	8.5	8.5	8.5	15	8.5
	v	(−)	2.6	2.5	2.6	0	2.6	2.6	3.0	2.5	3.0
AEX	K_{eq}	(−)	0.5	0.5	0.9	3.9	3.9	3.9	0.5	0.5	2.5
	v	(−)	4.0	4.0	1.7	2.9	2.9	2.9	4.0	4.0	3.0
HIC	K_{eq}	(−)	9.3	1.6	10.4	9.3	1.6	10.4	3.0	9.3	9.3
	n	(−)	9.3	1.6	10.4	9.3	1.6	10.4	3.0	9.3	9.3

proteins 2 to 9), using data from a prior study[35] and additional artificial data as shown in Table 2. No data was available for BSA on HIC-resin, based on performed column gradient experiments, we estimated the isotherm parameters to be equal to protein 3 (Chitotriosidase), as both proteins elute at the end of the gradient. More details can be found in Appendix A, as well as details about the resin parameters. The artificial data ensured at least three chromatography modes were required to purify the product of interest. Accordingly, a comprehensive comparison between the different optimization strategies could be accomplished. The chromatography column size (20.1 mL) was set in compliance to the size of the filtration unit operation. The linear flowrate of the chromatography process was 150 cm/h and the loading factor was 2.0 CV.

The validated filtration model for BSA was used to make valid assumptions for the simulation of other proteins during the overall flowsheet optimization. All proteins had similar or higher molecular weights compared to BSA, therefore full retention by the membrane was assumed for all proteins. The yield of the filtration operation was set to 95% for compensation of the lost material by adding an additional unit operation. The same constants for determining the B_{22} value, as given in Appendix B, were assumed for the other proteins. However, due to the low protein concentrations evaluated in this case-study, the B_{22} has no significant influence on the DF operation. Here, a DF mode ($\alpha = 0.99$) was employed to exchange buffers, for example, adapt salt conditions, between the chromatography steps. Therefore, only the time is a variable and the optimization problem was formulated as.

$$minf(t) = |C_{s,model}(t) - C_{s,desired}| \quad (19)$$

$$s.t. V(t_0) = V_0; C_i(t_0) = C_{i,0}; C_s(t_0) = C_{s,0}, \quad (20)$$

where t is the time variable to be optimized. $C_{s,model}$ is the model-predicted final salt concentration to be equalized to the desired final salt concentration, $C_{s,desired}$. The desired final salt concentration is in this case the initial salt concentration of the next chromatography operation.

For the flowsheet optimization, the global and local objective were formulated as.

$$minf(x) = (100 - yield(x)) + 2*(100 - purity(x)) + eluent\ consumption(x) \quad (21)$$

$$s.t.\ h(x) = 0 \quad (only\ applies\ to\ MM) \quad (22)$$

$$0 \leq x \leq 1, \quad (23)$$

where $f(x)$ is the objective function to be minimized, all variables (x) were normalized between 0 and 1 for enhanced optimization purposes (Equation 23). Additionally applicable when using MMs is to satisfy the equality equations $h(x)$, such as the mass balances and equilibrium relations (Equation 22). The optimizing variables (x) for the chromatography steps were: the gradient elution length, initial and final salt concentrations, and the lower and upper cut points. The performance measurements (e.g., yield, purity, buffer consumption) were evaluated across the entire purification process, with purity being assigned twice the weight due to its critical importance in biopharmaceutical purifications. Minimizing buffer consumption indirectly addresses the costs, batch throughput, and productivity concerns. The cost of lost feed is related to yield. Finally, the selected optimal flowsheets and their conditions from the global and minor local optimization were used as input for the final local optimization.

For both the global and local optimizers the following requirements were applied:

- Evaluation of the subsequent unit operation is only performed if the prior unit operation exceeds a yield of 5%, preventing solver failure due to excessively low concentration values.
- If the product pool's salt concentration is larger than the initial salt concentration of the next unit operation, either a dilution or filtration step is performed, depending on the flowsheet being evaluated.
- If the product pool's salt concentration is smaller than the initial salt concentration of the next unit operation, a spiking dilution step using a salt stock concentration of 5 M is performed.

- When using ANNs, the loading factor should be within the range of 0.05 and 5 CV to ensure compatibility with the data range for which the ANNs were developed. Otherwise, this option is indicated with not-a-number (Nan).

3 | RESULTS & DISCUSSION

3.1 | Filtration model validation

The filtration model was validated for the UF/DF experiment of BSA as shown in Figure 5. A good agreement between the experimental protein concentration and the model was found, $R^2 = 0.99$ and a low standard deviation of 0.03. Also the salt reduction over time is accurately predicted, $R^2 = 0.97$ and a standard deviation of 6.25. The alpha parameter was fitted to be 0.405, instead of the initial determined 0.7, as the permeate flowrate appeared to be not entirely constant throughout the process.

3.2 | Artificial neural networks

The quantitative evaluations showed that the desired values of $R^2 > 0.90$ and RMSE < 0.04, based on previous research,[54] were reached for almost all ANNs (Table 3). Converting the normalized RMSE values into absolute RMSE values gives an error value between 9.3% and 14.1% for protein 1, and for the volume between 3.6% and 11%. As justified in previous research, we considered an error rate of 15% to be acceptable, and to confidently identify the most optimal flowsheets while disregarding the less promising ones during flowsheet optimization. The generated data is focused around the product peak, resulting in some proteins that never elute or appear in the product pool. Therefore, training accurate ANNs is challenging due to their consistently low output values, inducing low R^2 values. Nevertheless, the absolute RMSE values also remain low ($< 8 \cdot 10^{-5}$). Given our understanding that these proteins will never be present in the product pool, we can assume they would always be removed. The most challenging proteins to remove are the ones eluting around the product peak, and therefore these are considered as the critical proteins for that chromatography mode. For AEX these are the proteins: 2, 3, 7, and 8, while for CEX these are the proteins: 5, 6, 7, 8, and 9, and for HIC the proteins: 4, 8, and 9.

TABLE 3 Quantitative evaluation for all proteins and volume on each chromatography mode. The RMSE is given as a normalized number. The product pool volume and salt concentration are included as these are needed for connecting the unit operations and calculating certain performance measurements.

|	AEX		CEX		HIC	
	R^2	RMSE	R^2	RMSE	R^2	RMSE
Protein 1	0.99	0.016	0.99	0.020	0.98	0.022
Protein 2	0.99	0.020	−0.10	0.028	0.00	0.005
Protein 3	0.94	0.028	0.00	0.052	−0.14	0.328
Protein 4	−0.41	0.018	0.61	0.021	0.98	0.024
Protein 5	−1.08	0.014	0.99	0.021	0.00	0.010
Protein 6	−1.11	0.006	0.99	0.023	0.03	0.327
Protein 7	0.99	0.020	0.98	0.026	0.03	0.019
Protein 8	0.99	0.017	0.93	0.021	0.98	0.025
Protein 9	0.55	0.006	0.98	0.024	0.97	0.029
Volume	0.93	0.052	0.94	0.042	0.89	0.035
Salt	0.98	0.018	0.98	0.02	0.97	0.022
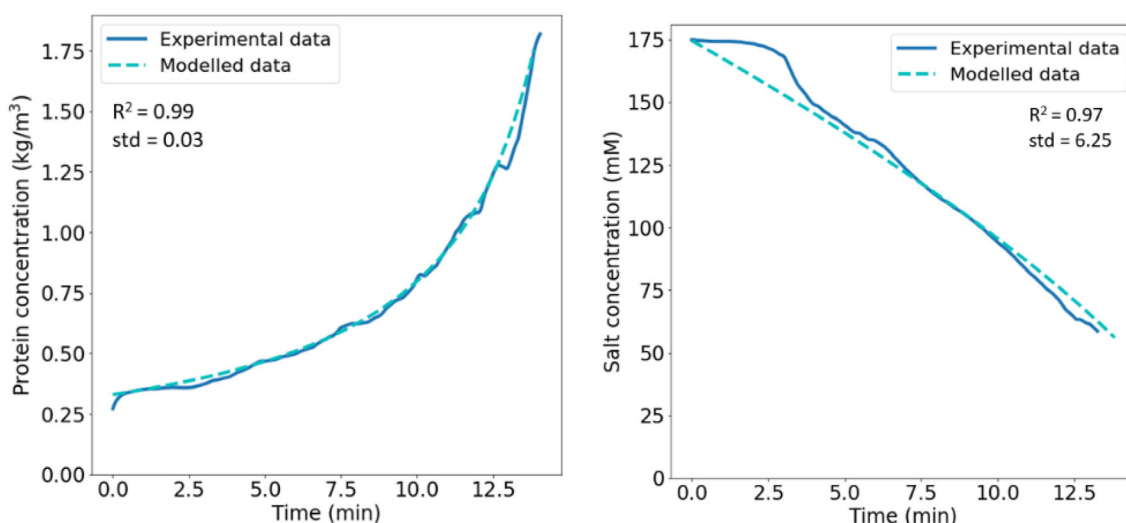
FIGURE 5 Left: Model prediction of the protein concentration, containing BSA, over time compared to the experimental values. Right: Model prediction of the salt concentration over time compared to the experimental values. The initial protein concentration was 0.3 kg/m³, the initial salt concentration contained 175 mM NaCl. The initial volume was 100 mL, the flowrate was 20 mL/min. The transmembrane pressure was 0.142 MPa.

3.3 | Flowsheet optimization

The flowsheet optimization workflow is designed to initially identify the global optima for each flowsheet. Subsequently, the most promising candidates can be further optimized locally, while the less promising ones may be disregarded. In this way, the number of flowsheets to be evaluated locally can be drastically reduced and correspondingly decreasing the overall optimization time. Optimizing a complex flowsheet involves finding global optima, therefore, a stochastic and heuristic algorithm was employed to increase the chance of finding most of the global optima.[66]

We compared three optimization strategies, namely, simultaneous, top-to-bottom, and decomposition, in terms of time-efficiency, complexity, and final results. Each optimization strategy was executed following the optimization workflow, as described in 2.1. Flowsheet optimization workflow, by either using MMs or ANNs. The flowsheet optimization was performed for a superstructure of three chromatography modes with a dilution or a filtration operation between the chromatography steps to function as a buffer exchange. In total, 39 flowsheets were evaluated. The maximum number of iterations using MMs was reduced compared to previous work to perform the flowsheet optimization within a reasonable amount of time, details can be found in 2.4. Numerical methods.[54] Similarly for ANNs, the number of iterations was adapted to guarantee a fair comparison between both workflows. The overall performance of each flowsheet is evaluated using the WOP value as described in 2.1. Flowsheet optimization workflow. In this work, the WOP is determined by the purity, yield, and buffer consumption. Based on the highest WOP value for all strategies using MMs, two best flowsheets were selected for which both MM and ANN results are shown in Table 4. All results of the global optimized flowsheet for all strategies, using MMs or ANNs, can be found in Appendix D. Note, when the salt concentration in the pool is lower than the initial salt concentration of the subsequent chromatography step, a dilution with a stock salt solution is performed, as described in 2.5. Case study. This also applies to flowsheets positioned with a filtration step, and can be confirmed by evaluating the optimized variables for the salt conditions. Moreover, in the top-to-bottom strategy using ANNs, Nan occurred when the

loading factor of a second or third chromatography step was out-of-range for the ANNs, as stated in the requirements in 2.5. Case study.

The strategies top-to-bottom and decomposition found the same best flowsheet (AEX – D – HIC – D – CEX), while the simultaneous strategy found a different one (CEX – D – HIC – D – AEX), as highlighted in Table 4. The flowsheet (AEX – D – HIC – D – CEX) was selected as an optimal candidate in all strategies when using ANNs. In overall, the ANNs found more optimal flowsheets (WOP > 96) compared to MM results. This is mainly attributed to an overestimation of the yield, which depends on the ANN accuracies for each protein (Appendix D). The Swarmplot, in Figure 6, shows the WOP values for the structures of one, two, or three chromatography steps in a sequence by either using MMs or ANNs. The different strategy outcomes are merged into the number of chromatography steps. Moreover, we clearly observe the same increasing trend when considering more chromatography steps for both ANNs and MMs. For one and two chromatography steps, the WOP value is a bit overestimated by the ANNs, mainly due to the overestimation of the yield as pointed out previously. The range for WOP values of three chromatography steps is about equal, only more flowsheets were estimated with a higher WOP value when using ANNs.

The selected best flowsheets, for each optimization strategy with MMs, were further locally optimized using the simultaneous strategy with MMs, as shown in Figure 7. Noticeably, the solver objective is to discover the ideal salt conditions within sequential chromatography steps, thereby eliminating the need for filtration and so obtaining enhanced yields and reducing buffer consumptions. Often, an orthogonal structure is applied in industrial processes, meaning that ion exchange and hydrophobic interaction chromatography are alternated.[2] Here, the two selected best flowsheets also have an orthogonal structure. However, from the global optimization results, other promising sequences, with a WOP > 96, are not necessarily orthogonal. For the final local optimization, a maximum number of 50 iterations was set to minimize the computational time, which took about 8 h. From the final results in Figure 7, it can be observed that there is a clear trade-off between purity and yield, for example the purity result of the simultaneous strategy is reduced, while the yield

TABLE 4 Performance measurement results of the global and minor local optimization results for the selected two best flowsheets from the MM modeling workflow, the ANN results are also provided. The selected best flowsheets for each strategy are highlighted.

Structure	Strategy	Purity (%)		Yield (%)		Buffer consumption (L/g)		WOP	
		MM	ANN	MM	ANN	MM	ANN	MM	ANN
CEX – D – HIC – D – AEX	Simultaneous	99.7	98.9	96.0	98.1	8.59	7.29	97*	97
	Top-to-bottom	92.7	89.6	92.3	100.0	7.63	4.20	93	94
	Decomposition	99.2	90.2	81.4	95.4	5.02	6.70	93*	92*
AEX – D – HIC – D – CEX	Simultaneous	99.9	99.6	89.3	100.0	6.48	7.23	95**	98
	Top-to-bottom	99.3	98.3	95.7	100.0	5.77	8.70	97	97
	Decomposition	99.1	97.3	96.2	100.0	6.04	10.76	97	97

Note: The filtration is a spiking dilution step as explained in 2.5. Case study for *Flowsheet [1-5-3-4-2] in table D.1, D.3, and D.6 and for **Flowsheet [2-5-3-4-1] in table D.1.

FIGURE 6 The WOP value of each flowsheet determined by each optimization strategy is compared for one, two, and three chromatography steps, and between using either MMs or ANNs as modeling workflow.
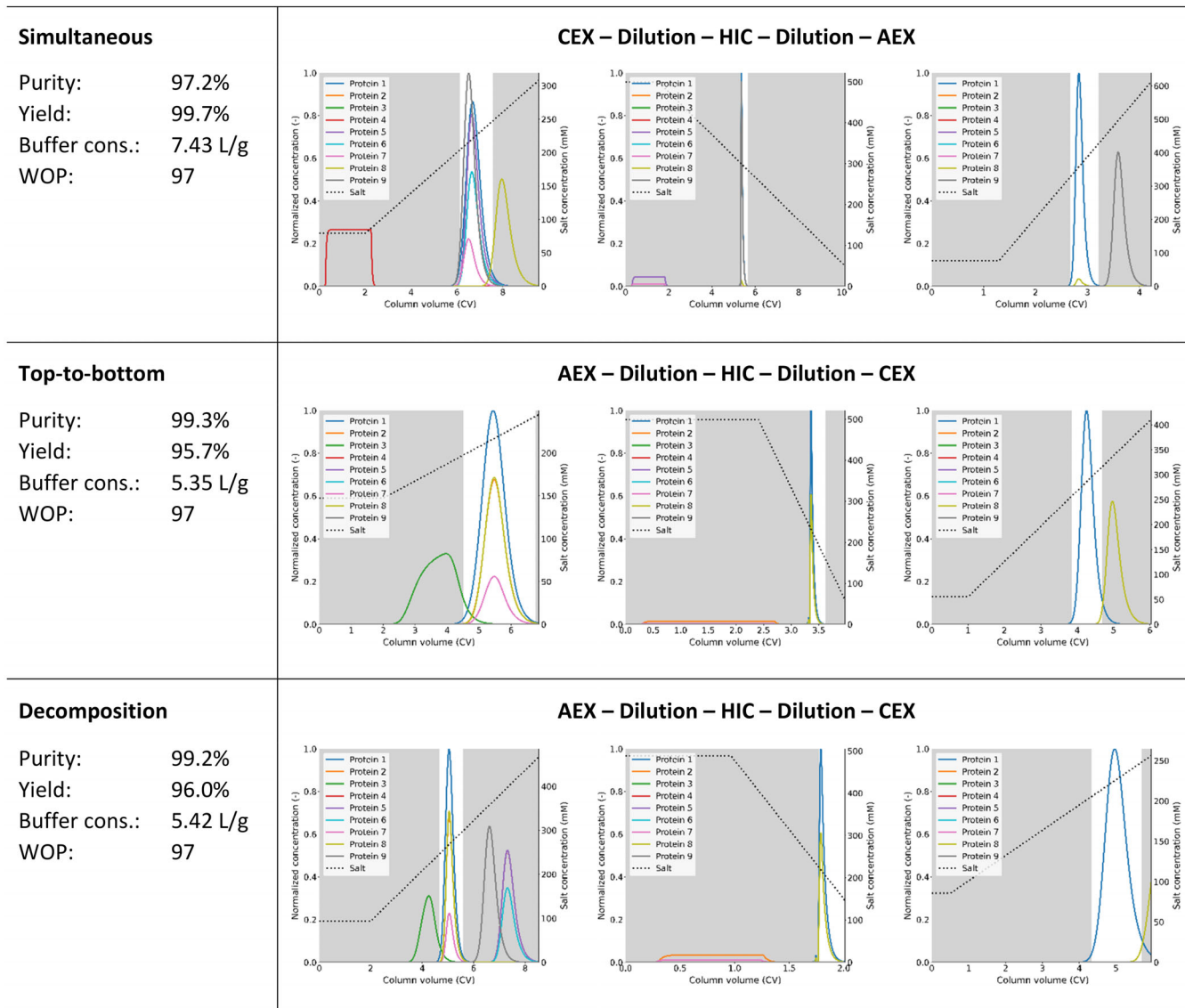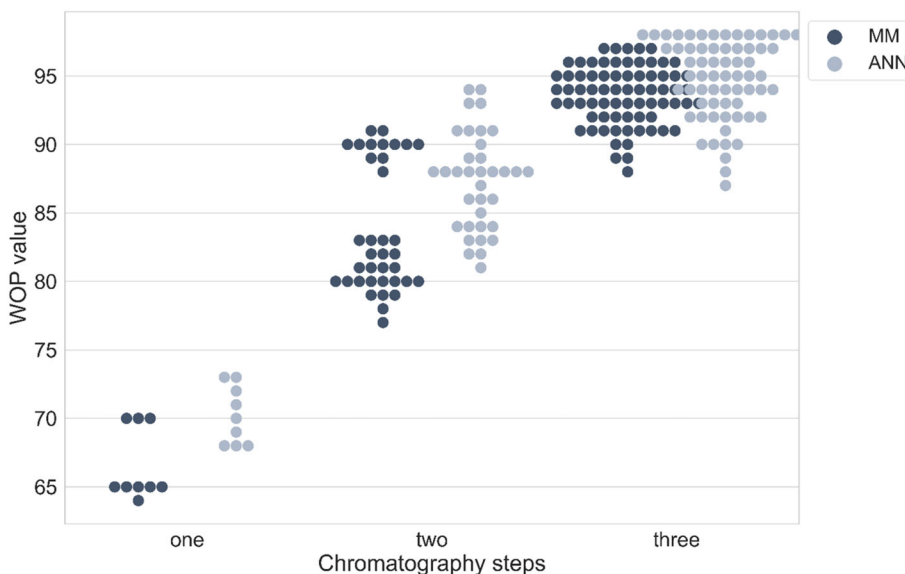

Simultaneous

Purity: 97.2%
Yield: 99.7%
Buffer cons.: 7.43 L/g
WOP: 97

CEX – Dilution – HIC – Dilution – AEX

Top-to-bottom

Purity: 99.3%
Yield: 95.7%
Buffer cons.: 5.35 L/g
WOP: 97

AEX – Dilution – HIC – Dilution – CEX

Decomposition

Purity: 99.2%
Yield: 96.0%
Buffer cons.: 5.42 L/g
WOP: 97

AEX – Dilution – HIC – Dilution – CEX

FIGURE 7 Final local optimization results using the simultaneous strategy with MMs. The global results of the best flowsheets for each strategy using MMs are used as input for the final local optimization. The maximum number of iterations was 50.
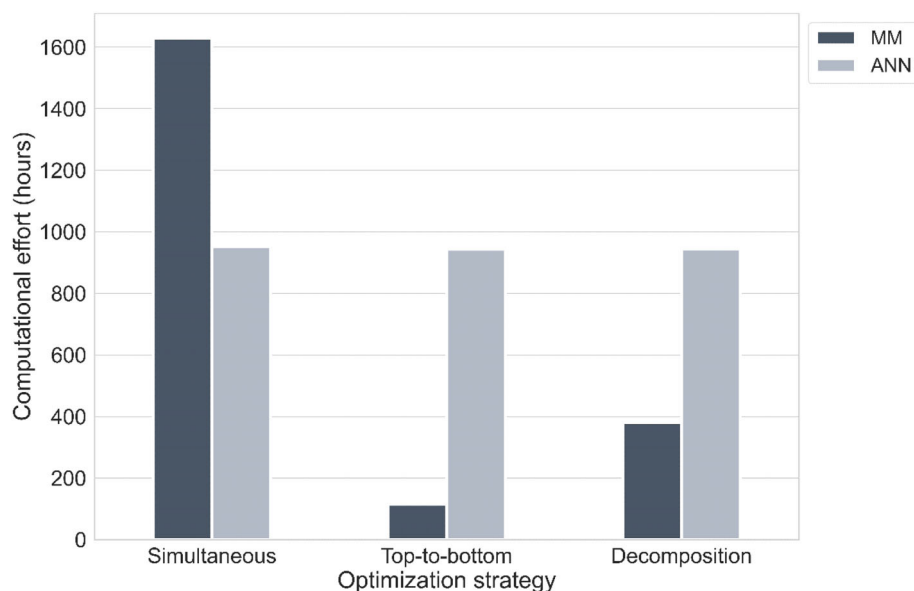
FIGURE 8 Comparison of the overall computational effort between the optimization strategies and modeling workflows. The computational hours represent the total (sequential) amount of hours needed for each strategy.
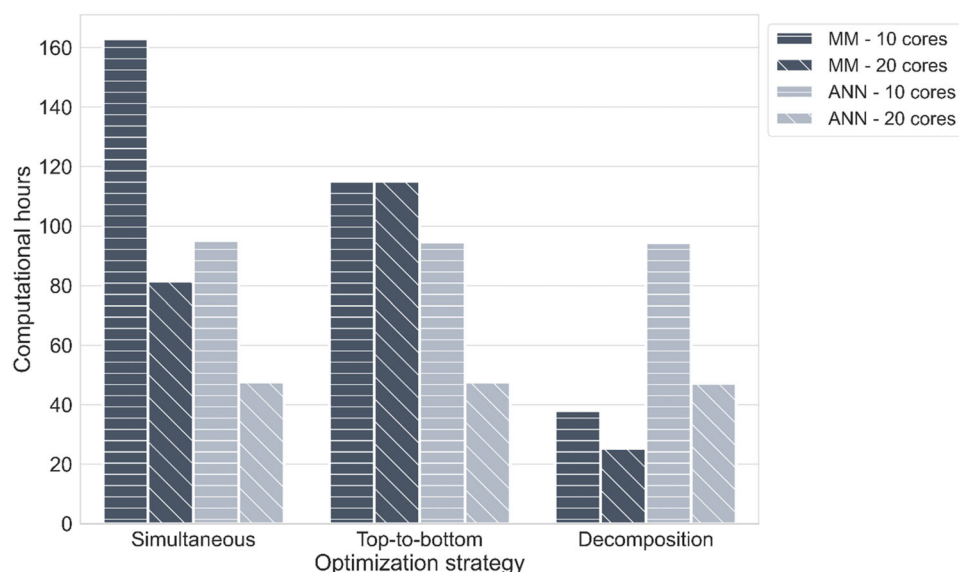

FIGURE 9 Comparing the computational hours for each optimization strategy and modeling workflow when using 10 or 20 cores.

increases, when comparing to the global optimized results. The buffer consumption was reduced in all strategies, but the overall WOP value was not improved for all strategies. So, to really improve the outcome, more iterations are needed. Or if a certain performance measurement, such as the purity, is a severe constraint (>99%), this can be applied to only local or both global and local optimization.

For comparing the overall computational effort, the total amount of hours for each strategy and workflow (MMs or ANNs) are evaluated and shown in Figure 8. However, the overall flowsheet optimization workflow applied parallelization whenever possible. The ANN-time involves the data-generation (using MMs), ANN development, and running the optimization, though, 99% of the time is devoted to the data-generation. The MM only includes the optimization time. The simultaneous strategy with MMs is obviously the most computationally intensive, whereas the top-to-bottom with MMs requires the least amount of computational effort.

Nowadays, more advanced computers consist of at least 10 or even 20 cores, and as a consequence the simultaneous and decomposition strategy can be executed way more time-efficiently. The decomposition can be parallelized maximally 15 times, as sequences of three chromatography steps depend on the two-chromatography step sequences. Whereas, the simultaneous strategy can be split into the number of flowsheets to be evaluated, in this case 39. Similarly for the ANN workflow, where, in principle, infinite codes can run simultaneously to generate data. Only the top-to-bottom strategy with MMs cannot be parallelized, as decisions are made sequentially between the various levels of chromatography steps. Figure 9 shows the effect of using 10 or 20 cores on each strategy and workflow. The decomposition strategy with MMs is the most time-efficient when making optimal use of the cores. In this case study, ANNs are significantly more time-efficient for the simultaneous strategy and for the top-to-bottom strategy when using 20 cores.

TABLE 5 Suggestions for deciding the type of optimization strategy and/or modeling workflow (ANNs or MMs) for certain scenarios/case studies.

Optimization problem	Time
Optimization objectives and constraints	*Depending on available number of cores.*
• Objective(s) and constraint(s) are clear: MMs, however, depending on superstructure size	If multiple cores can be used:
• Different objective(s) and constraint(s) to be evaluated: ANNs	• Limited time: Decomposition strategy
	• Extended time: Simultaneous strategy
Superstructure size Number of chromatography modes (type of resins) to be considered:	
• 3 chromatography modes: MMs	
• 4 chromatography modes: ANNs $+$ MMs	
• 5 chromatography modes: ANNs $+$ MMs	

Flexibility of method	Complexity
Optimizing variables	*In terms of coding and knowledge*
• Decided variables: MMs and/or ANNs	• All optimization strategies are about equal in development complexity, as the general optimization workflow is similar to all of them for both ANNs and MMs
• Undecided variables: MMs easier to use, or make more general ANNs with various input variables, or generate multiple ANNs	• Developing the ANNs adds more complexity to the overall approach
	• Advanced knowledge is required on the various MMs employed, the overall optimization workflow, developing the ANNs, all the algorithms/solvers used for the optimization and ANNs
Apply different objectives for different steps Decomposition strategy, this strategy can apply different objectives for the first step (capture step) and second and/or third steps (polishing steps).	*In terms of solving*
	• Least complex: Top-to-bottom, as it individually solves each unit operation
	• Most complex: Simultaneous, challenging to find the optimal solution for a sequence of more than 3 unit operations having at least 5 variables per unit operation. Increasing the number of unit operations in the sequence or the number of variables will significantly increase the complexity to solve the problem

Evidentially, the optimization strategy plays a significant role in the overall computational effort. But, if the optimization strategy and workflow are parallelized most efficiently, the difference in computational time between the strategies decreases, ranging from about 1 to 7 days. In this case-study, all strategies found multiple and similar optimal flowsheets. However, to obtain the most optimal conditions when connecting several unit operations, the simultaneous strategy is still recommended. In this flowsheet optimization evaluation, ANNs did not appear to be more time-efficient. Presumably, if more resins and/or larger sequences are considered and at least 20 cores can be used, it is expected that the ANNs exceeds the time-efficiency compared to MMs. This would be an interesting evaluation for a follow-up. Moreover, ANNs are very fast in executing the flowsheet optimization, which can be advantageous when evaluating different scenarios for the optimization problem. In general, multiple factors determine which optimization strategy and workflow (MMs or ANNs) might be optimal for a specific case-study, such as, the objective(s) and constraint(s), the size of the superstructure, and/or the computer power. The overview in Table 5 can help to make decisions for a flowsheet optimization approach. In future work, exploring different optimization algorithms, such as Bayesian optimization and genetic algorithms, could improve both efficiency and robustness. These methods can complement Differential Evolution by offering alternative approaches for navigating complex search landscapes.

4 | CONCLUSIONS

In this study, we compared three optimization strategies to determine the most effective approach for complex flowsheet optimization based on their outcomes, time-efficiency, and complexity. Each strategy, for example, simultaneous, top-to-bottom, and decomposition of the superstructure, was evaluated by either using MMs or ANNs for the global optimization. This complex flowsheet optimization consisted of 39 flowsheets, including an optional buffer exchange between the chromatography steps. The filtration mathematical model was validated for an UF/DF step using BSA. The protein concentration achieved an R^2 of 0.99 and a standard deviation of 0.03, and the salt concentration achieved an R^2 of 0.97 and a standard deviation of 6.25. Therefore, this model was assumed to be valid and applicable to the other proteins during flowsheet optimization, which had a similar or higher molecular weight than BSA. For the ANNs, all critical proteins, which are present around the product peak, reached an $R^2 > 0.93$, and the product of interest achieved an $R^2 > 0.98$ and RMSE < 0.022.

Subsequently, flowsheet optimization using MMs identified the same optimal flowsheet (AEX – D – HIC – D – CEX) for both top-to-bottom and decomposition strategies, the ANNs predicted the same WOP for this sequence. The simultaneous strategy with MMs identified a different sequence (CEX – D – HIC – D – AEX), which was not selected as one of the best by the other two strategies, giving a WOP threshold of at least 96. In general, the WOP values were predicted within a similar range when using either ANNs or MMs. In the case of orthogonal sequences, the solver often determined the optimal salt conditions to exclude the filtration step and instead employed a dilution/spiking step, and so reducing buffer

consumptions and enhancing yields. Leveraging the multi-core processing capabilities, commonly available in contemporary computers, minimizes the duration of the flowsheet optimization between the strategies. When using multiple cores, the superstructure decomposition method employed with MMs is the most time-efficient approach. Utilizing ANNs is only significantly more time-efficient when employing the simultaneous strategy, and top-to-bottom approach when utilizing 20 cores. Furthermore, if various optimization problems should be evaluated, ANNs are valuable for their fast flowsheet optimization, taking under an hour with multiple cores. All strategies are about equal in terms of complexity to develop the software. However, the combination with ANNs adds a layer of complexity because more knowledge is required on different aspects.

This study points out the importance of different optimization strategies and modeling techniques for complex flowsheet optimizations. Since numerous factors play a role, the decision-making table can support to find the most suitable type of strategy and modeling technique for a certain case study. Flowsheet optimization is crucial during the early conceptual process design to decrease costs and development time. Moreover, at the initial stage of a development process, limited sample material is available and knowledge about the sample purification has yet to be acquired. All strategies, whether employing MMs and ANNs, successfully identified multiple optimal flowsheets. Moreover, due to efficient parallelization, the difference in computational time between the strategies was minimized. Though, the decomposition of the superstructure strategy with MMs proved to be most time-efficient. Furthermore, it has the advantage to apply different objectives for specific steps during the purification process, enhancing its versatility and utility in biopharmaceutical process development. In this article, we have demonstrated the broad application and capability of this approach. This method is flexible and can be easily adapted to specific case study requirements and process limitations. It serves as an initial screening tool to guide further optimization and refinement.

AUTHOR CONTRIBUTIONS

Daphne Keulen: Conceptualization; methodology; software; data curation; investigation; formal analysis; visualization; writing – original draft; writing – review and editing; validation; project administration. **Myrto Apostolidi:** Software; data curation; validation; writing – review and editing. **Geoffroy Geldhof:** Supervision; writing – review and editing. **Olivier Le Bussy:** Supervision; writing – review and editing. **Martin Pabst:** Conceptualization; supervision; writing – review and editing. **Marcel Ottens:** Conceptualization; supervision; funding acquisition; writing – review and editing; project administration.

ACKNOWLEDGMENTS

CONFLICT OF INTEREST STATEMENT

All authors have declared the following interests: Geoffroy Geldhof and Olivier Le Bussy are employees of the GSK group of companies. The other authors declare no conflict of interests.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Daphne Keulen https://orcid.org/0000-0001-8086-333X

REFERENCES

1. Jagschies G, Łącki KM. Chapter 4 – process capability requirements. In: Jagschies G, Lindskog E, Łącki K, Galliher P, eds. *Biopharmaceutical Processing*. Elsevier; 2018:73-94.
2. Łącki KM. Chapter 16 – introduction to preparative protein chromatography. In: Jagschies G, Lindskog E, Łącki K, Galliher P, eds. *Biopharmaceutical Processing*. Elsevier; 2018:319-366.
3. Nfor BK, Ahamed T, van Dedem GWK, et al. Model-based rational methodology for protein purification process synthesis. *Chem Eng Sci*. 2013;89:185-195. doi:10.1016/j.ces.2012.11.034
4. Huuk TC, Hahn T, Osberghaus A, Hubbuch J. Model-based integrated optimization and evaluation of a multi-step ion exchange chromatography. *Sep Purif Technol*. 2014;136:207-222. doi:10.1016/j.seppur.2014.09.012
5. Keulen D, Geldhof G, Bussy OL, Pabst M, Ottens M. Recent advances to accelerate purification process development: a review with a focus on vaccines. *J Chromatogr A*. 2022;1676:463195. doi:10.1016/j.chroma.2022.463195
6. Mencarelli L, Chen Q, Pagot A, Grossmann IE. A review on superstructure optimization approaches in process system engineering. *Comput Chem Eng*. 2020;136:106808. doi:10.1016/j.compchemeng.2020.106808
7. Chen Q, Grossmann IE. Recent developments and challenges in optimization-based process synthesis. *Ann Rev Chem Biomol Eng*. 2017; 8(1):249-283. doi:10.1146/annurev-chembioeng-080615-033546
8. Siirola JJ, Powers GJ, Rudd DF. Synthesis of system designs: III. Toward a process concept generator. *AIChE J*. 1971;17(3):677-682. doi:10.1002/aic.690170334
9. Umeda T, Hirai A, Ichikawa A. Synthesis of optimal processing system by an integrated approach. *Chem Eng Sci*. 1972;27(4):795-804. doi:10.1016/0009-2509(72)85013-9
10. Westerberg AW. A retrospective on design and process synthesis. *Comput Chem Eng*. 2004;28(4):447-458. doi:10.1016/j.compchemeng.2003.09.029
11. Achenie LKE, Biegler LT. A superstructure based approach to chemical reactor network synthesis. *Comput Chem Eng*. 1990;14(1):23-40. doi:10.1016/0098-1354(90)87003-8
12. Bauer MH, Stichlmair J. Superstructures for the mixed integer optimization of nonideal and azeotropic distillation processes. *Comput Chem Eng*. 1996;20:S25-S30. doi:10.1016/0098-1354(96)00015-4
13. Short M, Isafiade AJ, Fraser DM, Kravanja Z. Synthesis of heat exchanger networks using mathematical programming and heuristics in a two-step optimisation procedure with detailed exchanger design. *Chem Eng Sci*. 2016;144:372-385. doi:10.1016/j.ces.2016.01.045

14. Friedler F, Aviso KB, Bertok B, Foo DCY, Tan RR. Prospects and challenges for chemical process synthesis with P-graph. *Curr Opin Chem Eng*. 2019;26:58-64. doi:10.1016/j.coche.2019.08.007

15. Chen Q, Liu Y, Seastream G, Siirola JD, Grossmann IE. Pyosyn: a new framework for conceptual design modeling and optimization. *Comput Chem Eng*. 2021;153:107414. doi:10.1016/j.compchemeng.2021.107414

16. Bertran M-O, Frauzem R, Zhang L, Gani R. A generic methodology for superstructure optimization of different processing networks. In: Kravanja Z, Bogataj M, eds. *Computer Aided Chemical Engineering*. Elsevier; 2016:685-690.

17. Altissimi R, Brambilla A, Deidda A, Semino D. Optimal operation of a separation plant using artificial neural networks. *Comput Chem Eng*. 1998;22:S939-S942. doi:10.1016/S0098-1354(98)00185-9

18. Chambers M, Mount-Campbell CA. Process optimization via neural network metamodeling. *Int J Prod Econ*. 2002;79(2):93-100. doi:10.1016/S0925-5273(00)00188-2

19. Henao CA, Maravelias CT. Surrogate-based superstructure optimization framework. *AIChE J*. 2011;57(5):1216-1232. doi:10.1002/aic.12341

20. Fernandes FAN. Optimization of Fischer-Tropsch synthesis using neural networks. *Chem Eng Technol*. 2006;29(4):449-453. doi:10.1002/ceat.200500310

21. Schweidtmann AM, Mitsos A. Deterministic global optimization with artificial neural networks embedded. *J Optim Theory Appl*. 2019;180(3):925-948. doi:10.1007/s10957-018-1396-0

22. Nascimento CAO, Giudici R, Guardani R. Neural network based approach for optimization of industrial chemical processes. *Comput Chem Eng*. 2000;24(9):2303-2314. doi:10.1016/S0098-1354(00)00587-1

23. Reinhardt IC, Oliveira DJC, Ring DDT. Current perspectives on the development of industry 4.0 in the pharmaceutical sector. *J Ind Inf Integr*. 2020;18:100131. doi:10.1016/j.jii.2020.100131

24. Silva F, Resende D, Amorim M, Borges M. A field study on the impacts of implementing concepts and elements of industry 4.0 in the biopharmaceutical sector. *J Open Innov: Technol, Mark, Complex*. 2020;6(4):175. doi:10.3390/joitmc6040175

25. Bisschops M, Cameron L. Process intensification and industry 4.0: mutually enabling trends. *Process Control, Intensification, and Digitalisation in Continuous Biomanufacturing*. Wiley; 2022:209-229.

26. Chen Y, Yang O, Sampat C, Bhalode P, Ramachandran R, Ierapetritou M. Digital twins in pharmaceutical and biopharmaceutical manufacturing: a literature review. *Processes*. 2020;8(9):1088. doi:10.3390/pr8091088

27. Portela RMC, Varsakelis C, Richelle A, et al. When is an in silico representation a digital twin? A biopharmaceutical industry approach to the digital twin concept. *Digital Twins*. Springer; 2020:35-55.

28. FDA. PAT Guidance for Industry – A Framework for innovative Pharmaceutical Development, Manufacturing and Quality Assurance. www.fda.gov/regulatory-information/search-fda-guidance-documents/pat-framework-innovative-pharmaceutical-development-manufacturing-and-quality-assurance

29. Yu LX. Pharmaceutical quality by design: product and process development, understanding, and control. *Pharm Res*. 2008;25(4):781-791. doi:10.1007/s11095-007-9511-1

30. Felinger A, Guiochon G. Comparison of the kinetic models of linear chromatography. *Chromatographia*. 2004;60(1):S175-S180. doi:10.1365/s10337-004-0288-7

31. Kumar V, Lenhoff AM. Mechanistic modeling of preparative column chromatography for biotherapeutics. *Ann Rev Chem Biomol Eng*. 2020;11(1):235-255. doi:10.1146/annurev-chembioeng-102419-125430

32. Shekhawat LK, Tiwari A, Yamamoto S, Rathore AS. An accelerated approach for mechanistic model based prediction of linear gradient elution ion-exchange chromatography of proteins. *J Chromatogr A*. 2022;1680:463423. doi:10.1016/j.chroma.2022.463423

33. Saleh D, Wang G, Müller B, et al. Straightforward method for calibration of mechanistic cation exchange chromatography models for industrial applications. *Biotechnol Prog*. 2020;36(4):e2984. doi:10.1002/btpr.2984

34. Hess R, Yun D, Saleh D, et al. Standardized method for mechanistic modeling of multimodal anion exchange chromatography in flow through operation. *J Chromatogr A*. 2023;1690:463789. doi:10.1016/j.chroma.2023.463789

35. Nfor BK, Ahamed T, Pinkse MWH, et al. Multi-dimensional fractionation and characterization of crude protein mixtures: toward establishment of a database of protein purification process development parameters. *Biotechnol Bioeng*. 2012;109(12):3070-3083. doi:10.1002/bit.24576

36. Close EJ, Salm JR, Bracewell DG, Sorensen E. A model based approach for identifying robust operating conditions for industrial chromatography with process variability. *Chem Eng Sci*. 2014;116:284-295. doi:10.1016/j.ces.2014.03.010

37. Gétaz D, Stroehlein G, Butté A, Morbidelli M. Model-based design of peptide chromatographic purification processes. *J Chromatogr A*. 2013;1284:69-79. doi:10.1016/j.chroma.2013.01.118

38. Disela R, Bussy OL, Geldhof G, Pabst M, Ottens M. Characterisation of the *E. Coli* HMS174 and BLR host cell proteome to guide purification process development. *Biotechnol J*. 2023;18(9):2300068. doi:10.1002/biot.202300068

39. Meyer K, Leweke S, von Lieres E, Huusom JK, Abildskov J. ChromaTech: a discontinuous Galerkin spectral element simulator for preparative liquid chromatography. *Comput Chem Eng*. 2020;141:107012. doi:10.1016/j.compchemeng.2020.107012

40. Leweke S, von Lieres E. Chromatography analysis and design toolkit (CADET). *Comput Chem Eng*. 2018;113:274-294. doi:10.1016/j.compchemeng.2018.02.025

41. Breuer JM, Leweke S, Schmölder J, Gassner G, von Lieres E. Spatial discontinuous Galerkin spectral element method for a family of chromatography models in CADET. *Comput Chem Eng*. 2023;177:108340. doi:10.1016/j.compchemeng.2023.108340

42. Rao JS, Püttmann A, Khirevich S, et al. High-definition simulation of packed-bed liquid chromatography. *Comput Chem Eng*. 2023;178:108355. doi:10.1016/j.compchemeng.2023.108355

43. Narayanan H, von Stosch M, Feidl F, Sokolov M, Morbidelli M, Butté A. Hybrid modeling for biopharmaceutical processes: advantages, opportunities, and implementation. Review. *Front Chem Eng*. 2023;5:1-10. doi:10.3389/fceng.2023.1157889

44. von Stosch M, Oliveira R, Peres J, de Azevedo SF. Hybrid semiparametric modeling in process systems engineering: past, present and future. *Comput Chem Eng*. 2014;60:86-101. doi:10.1016/j.compchemeng.2013.08.008

45. Narayanan H, Seidler T, Luna MF, Sokolov M, Morbidelli M, Butté A. Hybrid models for the simulation and prediction of chromatographic processes for protein capture. *J Chromatogr A*. 2021;1650:462248. doi:10.1016/j.chroma.2021.462248

46. Kiss AA, Grievink J. Process systems engineering developments in Europe from an industrial and academic perspective. *Comput Chem Eng*. 2020;138:106823. doi:10.1016/j.compchemeng.2020.106823

47. Liu S, Papageorgiou LG. Optimal antibody purification strategies using data-driven models. *Engineering*. 2019;5(6):1077-1092. doi:10.1016/j.eng.2019.10.011

48. Natali JM, Pinto JM, Papageorgiou LG. Efficient MILP formulations for the simultaneous optimal peptide tag design and downstream processing synthesis. *AIChE J*. 2009;55(9):2303-2317. doi:10.1002/aic.11913

49. Polykarpou EM, Dalby PA, Papageorgiou LG. Optimal synthesis of chromatographic trains for downstream protein processing. *Biotechnol Prog*. 2011;27(6):1653-1660. doi:10.1002/btpr.670

50. Liu S, Papageorgiou LG. Multi-objective optimisation for biopharmaceutical manufacturing under uncertainty. *Comput Chem Eng*. 2018;119:383-393. doi:10.1016/j.compchemeng.2018.09.015

51. Simeonidis E, Pinto JM, Lienqueo ME, Tsoka S, Papageorgiou LG. MINLP models for the synthesis of optimal peptide tags and

downstream protein processing. *Biotechnol Prog*. 2005;21(3):875-884. doi:10.1021/bp049650n

52. Pirrung SM, Berends C, Backx AH, van Beckhoven RFWC, Eppink MHM, Ottens M. Model-based optimization of integrated purification sequences for biopharmaceuticals. *Chem Eng Sci X*. 2019; 3:100025. doi:10.1016/j.cesx.2019.100025

53. Pirrung SM, van der Wielen LAM, van Beckhoven RFWC, van de Sandt EJAX, Eppink MHM, Ottens M. Optimization of biopharmaceutical downstream processes supported by mechanistic models and artificial neural networks. *Biotechnol Prog*. 2017;33(3):696-707. doi:10.1002/btpr.2435

54. Keulen D, van der Hagen E, Geldhof G, Le Bussy O, Pabst M, Ottens M. Using artificial neural networks to accelerate flowsheet optimization for downstream process development. *Biotechnol Bioeng*. 2023;121:2318-2331. doi:10.1002/bit.28454

55. Tanartkit P, Biegler LT. A nested, simultaneous approach for dynamic optimization problems—I. *Comput Chem Eng*. 1996;20(6):735-741. doi:10.1016/0098-1354(95)00206-5

56. SciPy. scipy.optimize.differential_evolution – SciPy v1.11.2 Reference Guide. https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.differential_evolution.html

57. Kocis GR, Grossmann IE. A modelling and decomposition strategy for the minlp optimization of process flowsheets. *Comput Chem Eng*. 1989;13(7):797-819. doi:10.1016/0098-1354(89)85053-7

58. Daichendt MM, Grossmann IE. Integration of hierarchical decomposition and mathematical programming for the synthesis of process flowsheets. *Comput Chem Eng*. 1998;22(1):147-175. doi:10.1016/S0098-1354(97)88451-7

59. Liñán DA, Bernal DE, Ricardez-Sandoval LA, Gómez JM. Optimal design of superstructures for placing units and streams with multiple and ordered available locations. Part I: a new mathematical framework. *Comput Chem Eng*. 2020;137:106794. doi:10.1016/j.compchemeng.2020.106794

60. Ruthven DM. *Principles of Adsorption and Adsorption Processes*. John Wiley & Sons; 1984.

61. Nfor BK, Zuluaga DS, Verheijen PJT, Verhaert PDEM, van der Wielen LAM, Ottens M. Model-based rational strategy for chromatographic resin selection. *Biotechnol Prog*. 2011;27(6):1629-1643. doi:10.1002/btpr.691

62. Nfor BK, Noverraz M, Chilamkurthi S, Verhaert PDEM, van der Wielen LAM, Ottens M. High-throughput isotherm determination and thermodynamic modeling of protein adsorption on mixed mode adsorbents. *J Chromatogr A*. 2010;1217(44):6829-6850. https://10.1016/j.chroma.2010.07.069

63. Fellner M, Delgado A, Becker T. Functional nodes in dynamic neural networks for bioprocess modelling. *Bioproc Biosyst Eng*. 2003;25(5):263-270. doi:10.1007/s00449-002-0297-6

64. Foley GA. *Membrane Filtration: A Problem Solving Approach with MATLAB*. Cambridge University Press; 2013.

65. Petzold L. Automatic selection of methods for solving stiff and non-stiff Systems of Ordinary Differential Equations. *SIAM J Sci Statist Comput*. 1983;4(1):136-148. doi:10.1137/0904010

66. Dominico G, Parpinelli RS. Multiple global optima location using differential evolution, clustering, and local search. *Appl Soft Comput*. 2021;108:107448. doi:10.1016/j.asoc.2021.107448

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.