processes

MDPI

Type of the Paper (Article, Review, Communication, etc.)

Artificial Neural Network for Fast and Versatile Model Parameter Adjustment utilizing PAT signals of Chromatography Processes for Process Control under Production Conditions

Mourad Mouellef[1], Glaenn Szabo[1], Florian Lukas Vetter[1], Christian Siemers[2] and Jochen Strube[1]*

[1] Institute for Separation and Process Technology, Clausthal University of Technology, Leibnizstraße 15, D-38678 Clausthal-Zellerfeld, Germany; mouellef@itv.tu-clausthal.de (M.M.); vetter@itv.tu-clausthal.de (F.L.V.); zobel-roos@itv.tu-clausthal.de (S.Z.-R.)

[2] Institute for Electrical Information Technology, Clausthal University of Technology, Julius-Albert-Str. 4, D-38678 Clausthal-Zellerfeld, Germany; siemers@iei.tu-clausthal.de (C.S.)

* Correspondence: strube@itv.tu-clausthal.de

Abstract: Preparative chromatography is a well-established operation in the chemical and biotechnology manufacturing. Chromatography achieves high separation performances but often has to deal with the yield versus purity trade-off as the optimization criterium regarding through-put. The initial trade-off is often disturbed by the well-known phenomenon of chromatogram shifts over process lifetime and has to be corrected by operators via adjustment of peak fraction cutting. Nevertheless, with regards to autonomous operation and batch to continuous processing modes, any advanced process control strategy is needed to identify and correct shifts from the optimal operation point automatically. Previous studies already presented solutions for batch-to-batch variance and process control options with the aid of rigorous physico-chemical process modelling. These models can be implemented as distinct digital twins as well as statistical process operation data analysers. In order to utilize such models for advanced process control, the model parameters have to be up dated with aid of inline PAT data to describe the actual operational status. Also including any occurring operational change phenomenon and its relation to their physico-chemical root cause. Typical phenomena are fluid dynamic changes due to packing breakage, channelling or compression as well as mass transfer and phase equilibrium related separation performance decrease due to adsorbent ageing or feed and buffer composition changes. In order to track these changes an Artificial Neural Network (ANN) is trained in this work. The ANN training is in this first step based on the simulation results of a distinct and previously experimentally validated process model. The model is implemented in the open source tool CasADi for python. This allows the implementation of interfaces to e.g. process control systems with relatively low effort. Therefore, PAT signals can easily incorporated for the sufficient adjustment of the process model for appropriate process control. Further steps would be the implementation of optimization routines based on the PAT and ANN predictions to derive optimal operation points with the model.

Keywords: parameter estimation; machine learning; ion-exchange chromatography; chromatography modeling; artificial neural networks

Citation: Lastname, F.; Lastname, F.; Lastname, F. Title. *Processes* **2021**, *9*, x. https://doi.org/10.3390/xxxxx

Academic Editor: Firstname Lastname
Received: date
Accepted: date
Published: date

1. Introduction

The utilization of machine learning approaches in chromatography is a rising field of research, which ranges from extracting crucial process information from measurement data in real-time via partial least squares algorithms [1] to separation factor prediction for chromatography process optimization with the aid of artificial neural networks (ANNs) [2]. Special focus lies on the ANNs, which become more and more accessible to a wide field of researches via frameworks like Tensorflow [3] or the The MathWorks Inc.

MATLAB toolboxes for Artificial Intelligence. Some of these works already investigated the possibility of determining chromatography model parameters by utilizing ANNs. Their results show that a maximum of three different experiments is necessary to predict model parameters within milliseconds after the training process and thus shortening model development time. [4,5]. In addition, both groups suggested that the automated real-time model parameter estimation during chromatographic experiments should be possible, especially due to the short computations times. In this work, the possibility of utilizing the suggested approach in a production environment is investigated.

The previous approaches for performed in a Lab-environment there multiple experiments with high process parameter variation can be conducted for example the use of different salt gradients. Therefore, the available information in general and for ANN training is much higher as in a controlled production environment. In contrast production data is only known to be of some variance, but operator decisions cope mostly with this by adjusting the fraction cut points. Purity is kept, but non-optimal yield losses are accepted [6,7]. Additionally, accessible chromatograms in a production environment are limited to the latest chromatograms from previous batch runs or cycles in case of continuous/cyclic processes like the multicolumn countercurrent solvent gradient purification process (MCSGP) [8] or the integrated countercurrent chromatography [9]. Such operational data can be analysed and necessary actions predicted [10]. On the other hand, the acceptable variation of the process in a production environment is limited due to the regulatory defined design and control space[11]. For example, the input feed mixture, purity and yield requirements, the column itself or gradient length and steepness can be assumed constant over process lifetime. Therefore, the authors mainly consider typical column packing fouling/aging phenomena as governing effects on the chromatograms, which still cause an considerable economic impact on the process [12,13]. The phenomena observed are related to either fluid-dynamics or mass-transfer and phase equilibrium. Fluid-dynamic variations in fluid-distribution, packing breakage, channelling, swelling or compression can be described by axial dispersion and/or voidage adjustments. Whereas, any adsorbent ageing phenomena are related to mass transfer and/or phase equilibrium behaviour. To automatically mitigate these effects, models previously used for process development can be used to calculate process conditions such as new fraction cut points. However, the effects on fluid dynamics, column packing and absorption phase equilibria caused by aging should be taken into account in order to achieve optimal results. In order to perform this study a chromatography model was developed in Python[3] with the CasADi framework [14] based on previous works [11,15] for a preparative chromatography separation of a three component protein mixture. It is assumed that a valid PAT strategy for online concentration determination is implemented, whose feasibility is already shown in several publications [11,15–20]. The total concept for advanced process control is shown in Figure 1

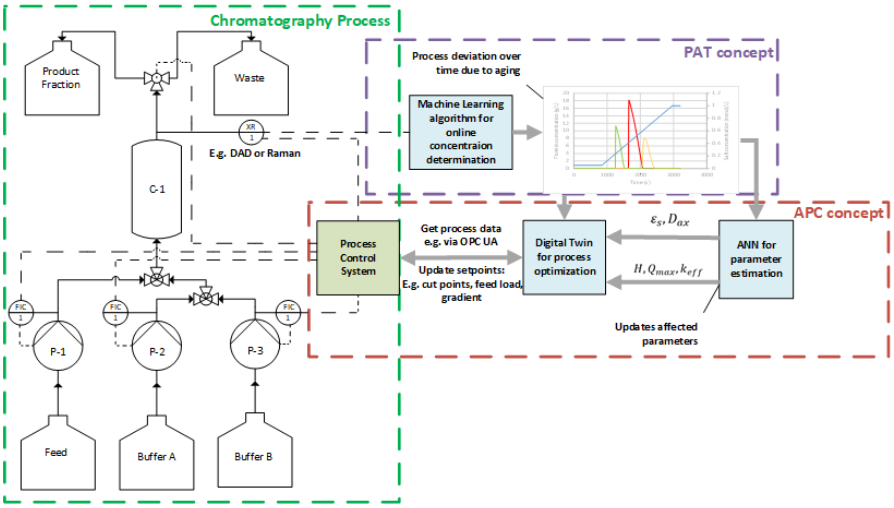


Figure 1 Advanced process control concept for process chromatography – batch and continuous

The PAT signal of the chromatogram is evaluated by a partial least squares algorithm (PLS). This data is transferred to the ANN, which estimates appropriate model parameter adjustments for the digital twin. The digital twin is implemented as rigorous physico-chemical process model. Such a process model with updated parameter adapts to the current operation state and is utilized to predict new operation set-points. For example to fit the necessary purity-yield requirements within the predefined control space.

As first step a sensitivity analysis is performed to identify the impact of relevant parameters on chromatograms, which may change over process lifetime. After that ANNs are trained from simulative generated data to predict relevant model parameters from a single chromatogram.

2. Materials and Methods

2.1. Chromatography Modeling

The necessary data for the ANN were generated through simulations. All simulations were performed in an Python 3.8 environment [21] with the CasADi Framework [22] on an Dell Optiplex 7010 System. All programming was done in the Spyder Integrated Development Environment (IDE) [23]. Commonly used chromatography models for mass transport by convection and dispersion are based on the general rate model Equation (1) or on the lumped pore diffusion model Equation (2) [24–26]. These equations describe the mass balance of the stationary phase [15]. This work utilized the lumped pore diffusion model.

$$\varepsilon_{p,i}\frac{\partial c_{p,i}}{\partial t} + (1-\varepsilon_{p,i})\frac{\partial q_i}{\partial t} = \frac{1}{r^2}\frac{\partial}{\partial r}\left[r^2\left(\varepsilon_{p,i}D_{p,i}\frac{\partial c_{p,i}}{\partial r} + (1-\varepsilon_{p,i})D_{S,i}\frac{\partial q_i^*}{\partial r}\right)\right] \quad (1)$$

$$\varepsilon_{p,i}\frac{\partial c_{p,i}}{\partial t} + (1-\varepsilon_{p,i})\frac{\partial q_i}{\partial t} = \frac{6}{d_P}\frac{(1-\varepsilon_S)}{\varepsilon_S}k_{eff,i}(c_i - c_{p,i}) \quad (2)$$

There the parameters mean diameter of the resin particle d_P, the porosity $\varepsilon_{P,i}$ and the voidage ε_S describe the column packing. The variables $c_{P,i}$ represents the concentration of the component in the pores of the resin, q_i the loading, c_i the concentration in the continuous phase and $k_{eff,i}$ the effective mass transport coefficient. The variable t represents the time. The boundary conditions are described by equation (3) for the column inlet and equation (4) for the column outlet. D_{ax} depicts the axial dispersion coefficient, L the length of the column and x the length domain. [24]

$$uc_{in,i}(t) = uc_i(t,0) - D_{ax}\frac{\partial c_i}{\partial x}(\mathrm{t},0) \quad (3)$$

$$\frac{\partial c_i}{\partial x}(t,L) = 0 \quad (4)$$

The mass transfer coefficient $k_{eff,i}$ is given by Equation (5). Here, $k_{f,i}$ is the film mass transfer coefficient, r_p the particle radius, and $D_{p,i}$ the pore diffusion coefficient [18]. The pore diffusion coefficient $D_{p,i}$ was calculated according to the correlation of Carta [27] and $k_{f,i}$ according to Wilson and Geankoplis [28].

$$\mathrm{k_{eff,i}} = \frac{1}{\frac{1}{k_{f,i}} + \frac{r_p}{D_{p,i}}} \quad (5)$$

In this work, the competitive Langmuir-isotherm Equation (6), which has already demonstrated its performance in (bio-) chromatography, was used [25,26,29–31]. The adsorption and desorption behaviour of the components can also be described by various other approaches [24,25,29,30,32–34].

$$\mathrm{q_i} = \frac{q_{max,i}K_{eq,i}c_i}{1 + \sum_j K_{eq,j}c_j} \quad (6)$$

With $q_{max,i}$ as the maximum loading capacity of component i and $K_{eq,i}$ as the Langmuir-coefficient of component i. To include the salt dependence into Equation (6) the Langmuir-coefficient can be written as shown in Equation (7). [24].

$$\mathrm{q_{max,i}K_{eq,i}} = \mathrm{H_i} \quad (7)$$

The salt dependence of the maximum loading $q_{max,i}$ and the Henry-coefficient H_i can then be expressed by Equations (8) and (9) and the empiric coefficients $a_{1,i}$, $a_{2,i}$, $b_{1,i}$ and $b_{2,i}$ for each component i [25,35].

$$\mathrm{q_{max,i}} = b_{1,i}c_{p,i} + b_{2,i} \quad (8)$$

$$\mathrm{H_i} = a_{1,i}{c_{p,i}}^{a_{2,i}} \quad (9)$$

The spatial discretization of the partial differential equations system followed a finite differences scheme.

2.2. Artificial Neural Networks

With the development of frameworks like Tensorflow [21] or the Mathworks Inc. MATLAB Artificial Intelligence toolboxes algorithms for machine learning in general and artificial neural networks (ANNs) became easily accessible for an even broader group of researches. This results in an even broader area of application including the chemical/pharmaceutical industry. Artificial neural networks (ANN) enabled computer-aided solutions of problems that were nearly impossible or difficult to solve with conventional algorithms within acceptable time limits. Typical applications are computer visions applications or natural language processing [36,37]. More relevant modeling contributions were the reduction in computational effort and/or the description of not yet (sufficiently) described physico-chemical relationships [38–40]. In general, artificial neural networks consist of interconnected neurons, which send information in the form of activations signals over weighted connections to other neurons which map inputs onto outputs [36] The process of finding the weights of these connection is called training. Usually this training is performed by supplying the neural network with input and output data from which it learns the underlying relationships via so-called backpropagation algorithms. Further information on ANNs can be found in Fausett [36] and Goodfellow [41]. After the training the ANN is capable of mapping previously unknown inputs onto outputs, be it classifica-

tion or regression tasks [36,41]. The activation functions can range from simple linear functions (equation (10)) over sigmoid functions (equation (11)) to more complex functions like the hyperbolic tangens (tanh) activation function (equation (12)) [36,42] with each of them having different benefits and drawbacks depending on the problem [36,42,43].

$$\mathrm{f(x)} = x \tag{10}$$

$$f(x) = \frac{1}{1 + e^{\sigma x}} \tag{11}$$

$$\mathrm{f(x)} = \tanh(x) = 1 - \frac{2}{e^{2x}} + 1 \tag{12}$$

In order to characterize ANN, a distinction between properties, architecture, activation functions, and training can be made [36,41,43]. ANNs are arranged as layers of neurons. Every ANNs consist at least of one input layer and one output layer. The layers in between are so called hidden layers and can resemble different structures or perform specific functions [36]. Common architectures are feed forward networks, in which all the information is fed forward from the input to the output nodes, and recurrent neural networks, in which the information can be cycled back to nodes of the same or previous layers [36,41,43]. An example of the fully connected feed forward network is given in Figure 2. The activation function maps the weighted inputs of a neuron to its output.

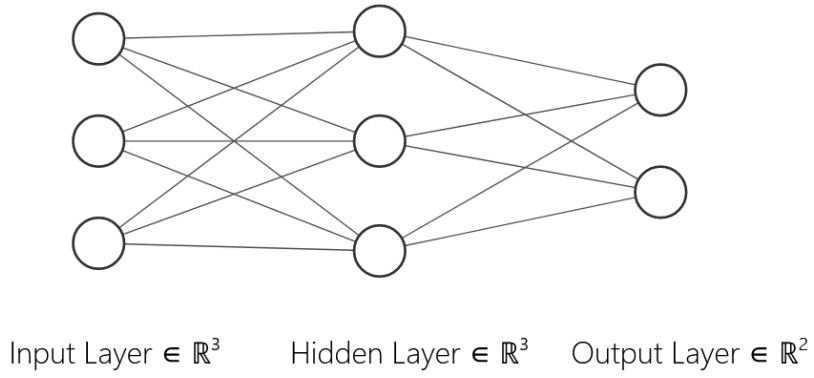


Figure 2 An example of a fully connected feed forward network with three input neurons, three neurons in the hidden layer, and two neurons in the output layer.

In this study all ANN calculation were performed on a Dell Optiplex 7010 Desktop PC. Programming was performed in the Spyder IDE [23]. For implementation, the Tensorflow 3 [3]backend of Keras v. 2.4.0 [44] with Python 3.8 [21]was used.

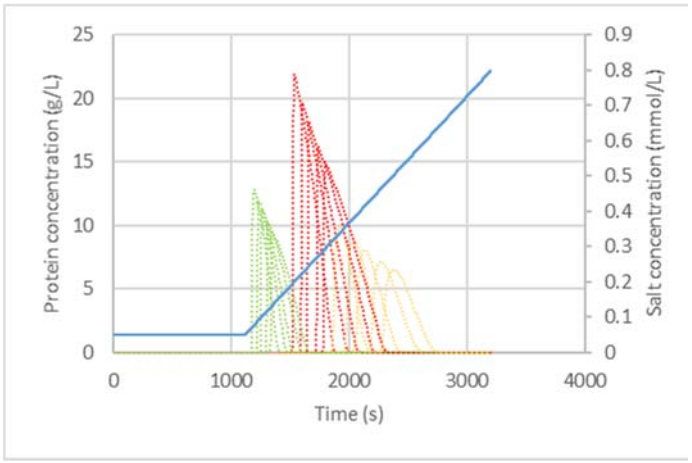
2.3. Model parameter choice and ANN dataset generation

As previously mentioned the dataset has to reflect the chromatography process conditions in a production environment. In this environment, the components are fixed and the option of multiple gradient experiments or arbitrary alternation of injection volumes is not available. This case stands in clear contrast to the previous work [5], there isotherm parameters for arbitrary 3 components mixtures were estimated. The data for this study was generated by an chromatography model based on the previous piloting case study of a monoclonal antibody manufacturing process from Kornecki et al. [15] and the work of Zobel-Roos et al. [26]. The model comprises three proteins, immunoglobulin G (IgG), a weak binding host cell protein (HCP1) and a strong binding host cell protein (HCP2). To reflect a production environment the column was up-scaled to a length of 15 cm and a diameter of 20 cm. The process parameters were increased to 1.6 L/min buffer flow, a 15 column volume (CV) gradient and an injection volume of 2.6 L with 5 g/L IgG, 2 g/L HCP1 and 2 g/L HCP2.

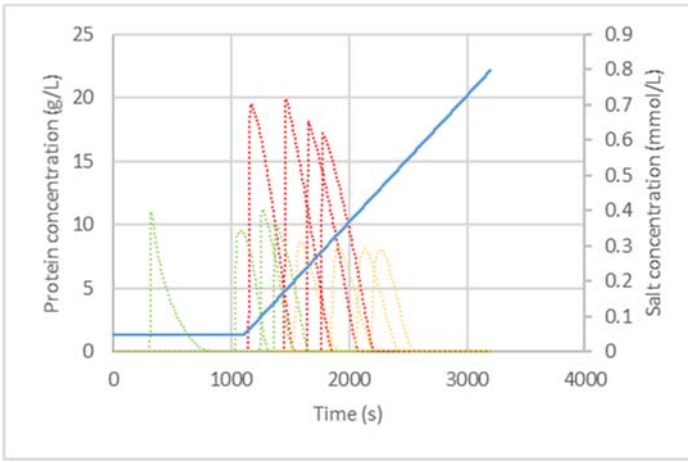
As stated before the ANN maps inputs on outputs. Therefore, the relevant outputs for the prior described use-case, deviation of model parameters due to column aging, must be identified. A subset of parameters can already be excluded or neglected based on expert knowledge. This comprises component properties like the molecular mass of the three proteins or the particle diameter of the packing and others, which should not alter during the process lifetime. Additionally some parameters can be excluded beforehand, because they can be substituted in others through correlations. This leads to an implicit consideration of these via other parameters. The molecular masses, tortuosity, steric factors, protein radius, particle diameter and molecular diffusion coefficient are considered constant [45].The film diffusion coefficient can be substituted into the mass transfer coefficient as shown in equation (5). The pore diffusion coefficient can be substituted in the particle porosity as shown in [27]. The influence of the remaining model parameters on the chromatograms is investigated through one parameter at a time sensitivity analysis. The variation range of each parameter is intentionally greater as it is expected by expert knowledge for the investigated use-case. The flow was excluded from this rule, because it is considered as a well controllable input parameter. An overview over the remaining parameters and their variation ranges for the sensitivity study is given in **Table 1**. The results of the sensitivity study is shown in **Figure *3***.

Table 1. Upper and lower limits of the varied parameters in the sensitivity analysis.

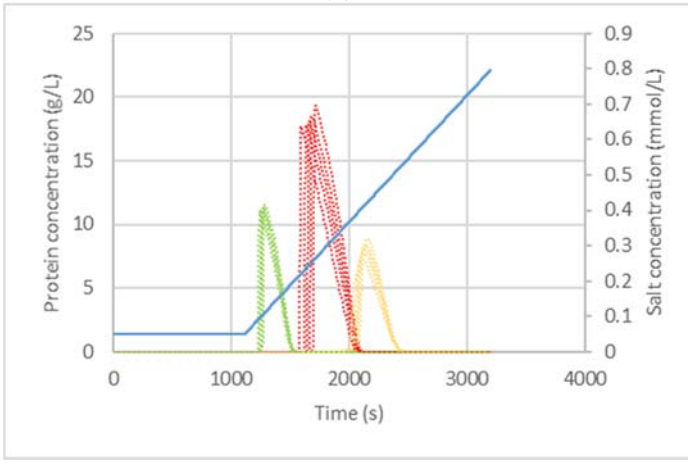
Parameter	Lower bound	initial	Upper bound
$a_{1,IgG}$	0.4	0.8	1.6
$a_{1,HCP1}$	0.8	1.6	3.2
$a_{1,HCP2}$	0.05	0.1	0.2
$a_{2,IgG}$	-3.6	-3.0	-1.5
$a_{2,HCP1}$	-3.6	-3.0	-1.5
$a_{2,HCP2}$	-3.6	-3.0	-1.5
$b_{1,IgG}$	-0.36	-0.24	-0.12
$b_{1,HCP1}$	-0.23	-0.15	-0.078
$b_{1,HCP2}$	-0.15	-0.099	-0.05
$b_{2,IgG}$	$0.13\ g/L$	$0.25\ g/L$	$0.5\ g/L$
$b_{2,HCP1}$	$0.13\ g/L$	$0.25\ g/L$	$0.5\ g/L$
$b_{2,HCP2}$	$0.05\ g/L$	$0.1\ g/L$	$0.2\ g/L$
$flow$	$1.568\ L/min$	$1600\ L/min$	$1.632\ L/min$
D_{ax}	$1.3x10^{-6}\ cm^2/s$	$1.3x10^{-3}\ cm^2/s$	$1.3\ cm^2/s$
r_p	$1x10^{-7}\ cm$	$1x10^{-5}\ cm$	$1x10^{-3}\ cm$
$k_{eff,IgG}$	$1x10^{-5}\ cm^2/s$	$1.5x10^{-2}\ cm^2/s$	$1x10^{2}\ cm^2/s$
$k_{eff,HCP1}$	$1x10^{-5}\ cm^2/s$	$1.2x10^{-2}\ cm^2/s$	$1x10^{2}\ cm^2/s$
$k_{eff,HCP2}$	$1x10^{-5}\ cm^2/s$	$2.7x10^{-2}\ cm^2/s$	$1x10^{2}\ cm^2/s$
ε_s	0.24	0.36	0.48
ε_p	0.3	0.78	0.78


Protein concentration (g/L)
Salt concentration (mmol/L)
Time (s)
0
5
10
15
20
25
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
0
1000
2000
3000
4000

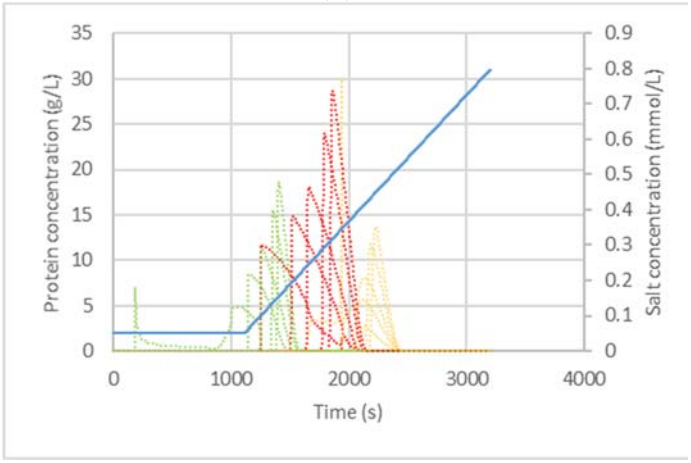
(**a**)


Protein concentration (g/L)
Salt concentration (mmol/L)
Time (s)
0
5
10
15
20
25
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
0
1000
2000
3000
4000

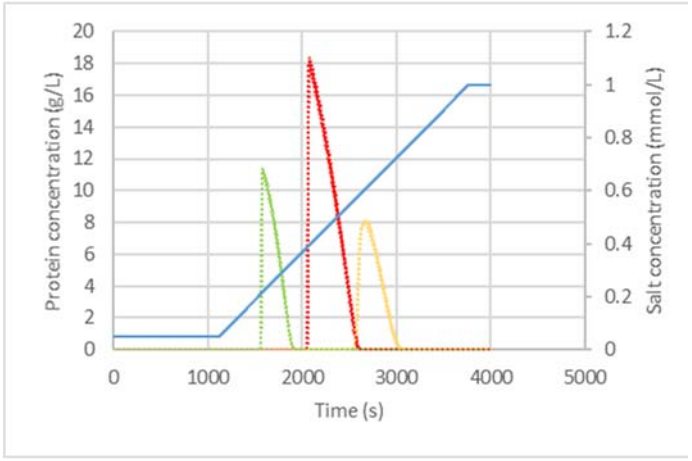
(**b**)


Protein concentration (g/L)
Salt concentration (mmol/L)
Time (s)
0
5
10
15
20
25
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
0
1000
2000
3000
4000

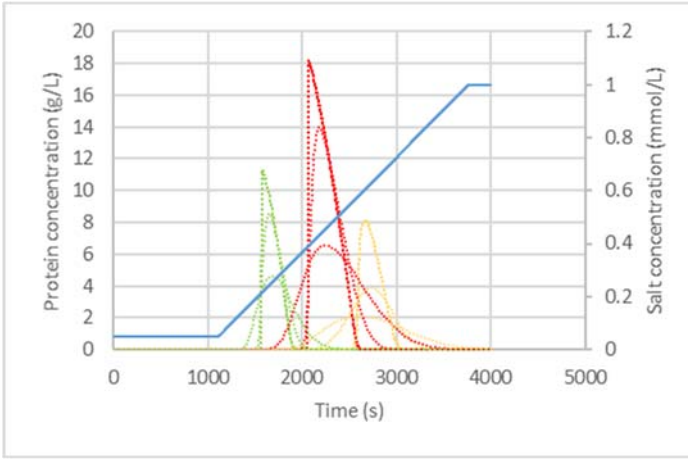
(**c**)


Protein concentration (g/L)
Salt concentration (mmol/L)
Time (s)
0
5
10
15
20
25
30
35
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
0
1000
2000
3000
4000

(**d**)


Protein concentration (g/L)
Salt concentration (mmol/L)
Time (s)
0
2
4
6
8
10
12
14
16
18
20
0
0.2
0.4
0.6
0.8
1
1.2
0
1000
2000
3000
4000
5000

(**e**)


Protein concentration (g/L)
Salt concentration (mmol/L)
Time (s)
0
2
4
6
8
10
12
14
16
18
20
0
0.2
0.4
0.6
0.8
1
1.2
0
1000
2000
3000
4000
5000

(**f**)

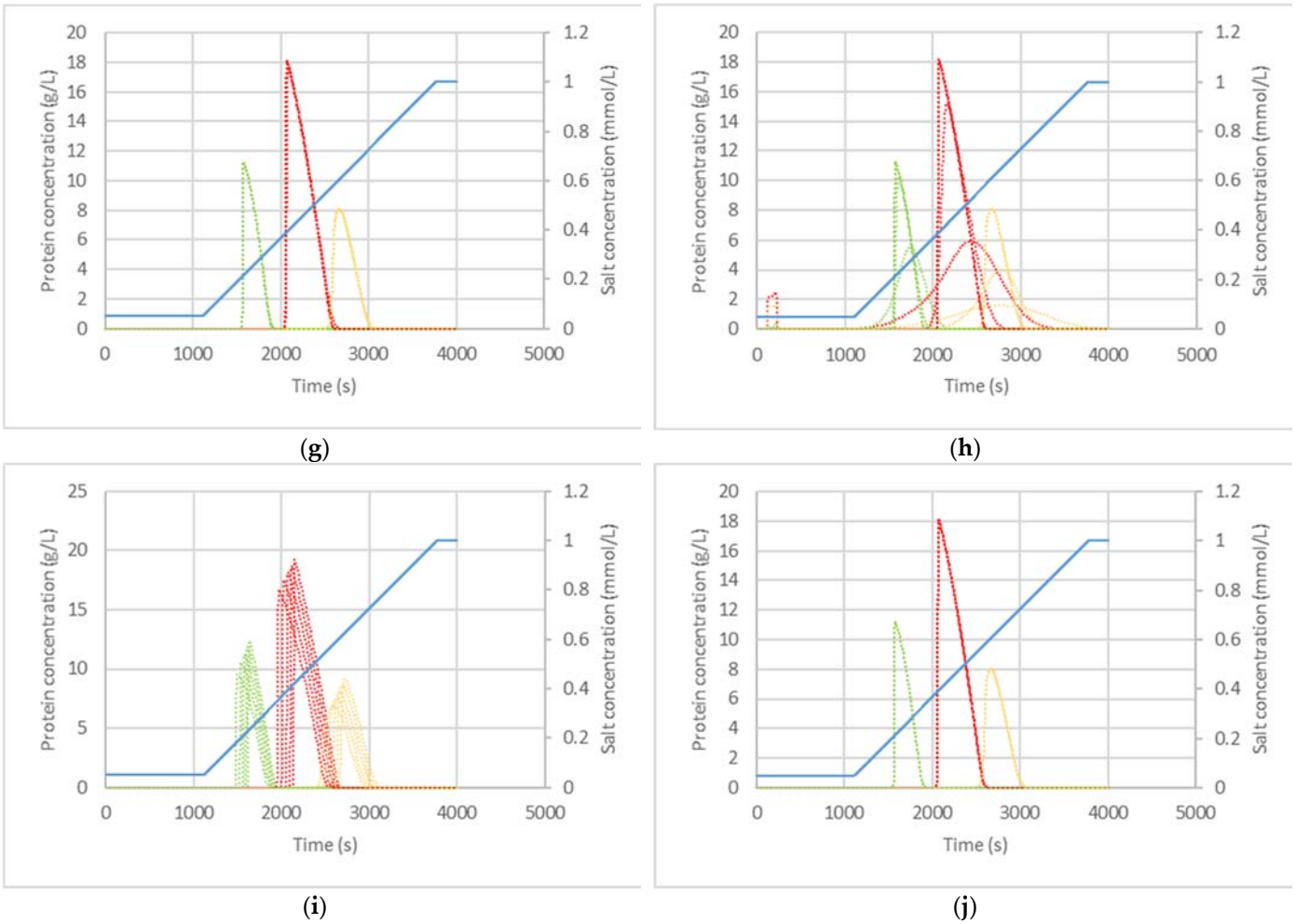


Figure 3. Shows the sensitivity analysis of the non a priori excluded parameters on each component. Green curves represent HCP1, red curves IgG, orange curves HCP2. Upper and lower variation ranges for the sensitivity analysis are given in Table 1. (**a**) Shows the impact on the chromatograms of parameters $a_{1,i}$, which describe the salt dependence of the henry coefficients. (**b**) Depicts the impact on the chromatograms of parameters $a_{2,i}$, which also describe the salt dependence of the henry coefficients. (**c**) Illustrates the impact of the parameters $b_{1,i}$, which describe the salt dependence of the maximum loading capacities. (**d**) Depicts the impact of the parameters $b_{2,i}$, which also describe the salt dependence of the maximum loading capacities. (**e**) Shows the impact of the volume flow. (**f**) Shows the impact of D_{ax}. (**g**) Depicts the impact of r_p. (**h**) Shows the impact of $k_{eff,i}$. (**i**) Illustrates the impact of ε_s. (**j**) Shows the impact of ε_p.

From **Figure 3** (a), (b) and (d) a clear impact of the salt dependency describing parameters a_1, a_2 and b_2 of the Langmuir-Isotherm (equation (6), (7), (8), (9)) on the chromatogram can be seen. The impact of b_1(c) is minor but still noticeable. Hence, all of these parameters are considered in the following steps. The flow in (e) has nearly no impact on the resulting chromatogram within the given variation range. In addition, D_{ax} shows a high impact of values greater $0.1\ cm^2/s$ in subplot (f). Because of this and the direct correlation of the flow and D_{ax}, which is described by Chung and Wen [46], both parameters are considered relevant in further steps. Following the results of (g) and (j) the pore radius r_p and porosity ε_p are not considered. The high impact of the voidage ε_s seen in (i) considered in the following steps. Also k_{eff} from subplot (h) is considered for now even though its values must decrease below $10^{-4}\ cm^2/s$ to show an impact on the chromatogram.

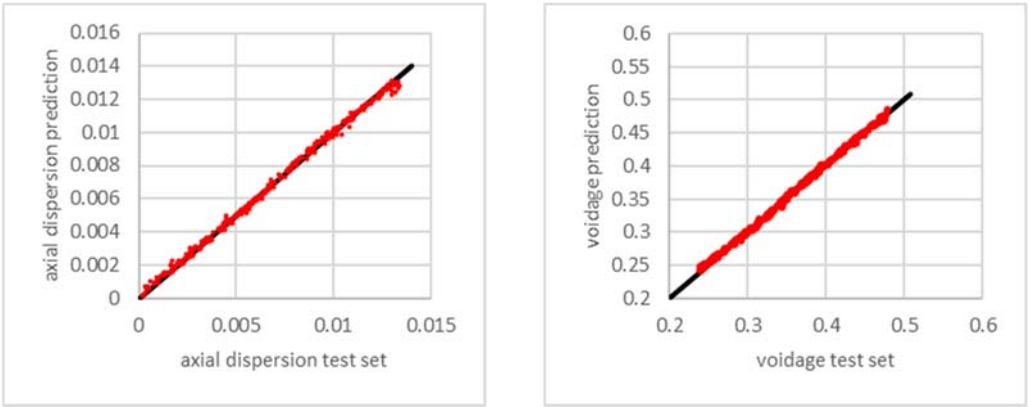
Because of the number of different parameters, the authors chose to split the set of these parameters into fluid dynamic and a column packing set and a phase equilibrium parameter set in the first step. The mass transfer coefficient is also moved to the fluid dynamic and column packing set, because the effects are similar. The reasoning is to facilitate the identification of parameters, which definitely need more than one batch experiment for an adequate estimation, and to reduce the total amount of needed simulations. The fluid dynamic and column packing set contains the parameter variation from flow, voidage, axial dispersion and the mass transfer coefficient. The other set contains the variations of a_1, a_2, b_1 and b_2 of all components. All parameter variations were uniformly distributed and were varied in the boundaries of **Table 1**. As input for the ANN the chromatogram of each component was reduced to 38 data points to reduce the complexity of the input data. Additionally, the maximum concentration of each peak was included in each entry. The data reduction scheme, its drawbacks and its benefits are explained in detail in the previous publication [5]. Accordingly, a single dataset entry consists of 39 data points. Another approach to reducing chromatogram information into fewer points can be found in Wang et al. [4].

3. Results

As mentioned earlier the two datasets were generated to examine the utilization of a single batch chromatography experiment for parameter estimation via ANNs. Therefore, two ANNs were trained in this process. With the first ANN the general possibility of estimating the packing and fluid dynamic parameters and the mass transfer coefficient of each component from a single chromatogram was investigated. That possibility for the phase equilibrium parameters was investigated with the second ANN. After the training process of all ANNs was completed, the corresponding parameters were predicted from training and validation set. The results were evaluated and used to generate a third dataset in which all-remaining parameters were varied. The prediction of a single data entry costs 200 milliseconds of computation time including the loading of the input data and of the ANN model itself.

3.1. Variation of packing and fluid dynamic parameters

This ANN was trained with 1000 chromatograms which for split to 70% training data and 30% validation data. The FF-ANN consists of 118 input neurons, 120 neurons with tanh activation in hidden layer 1 and 20% dropout probability, 80 neurons with tanh activation in hidden layer 2 with 20% dropout probability and 5 output neurons for D_{ax}, ε_s, $k_{eff,1}$, $k_{eff,2}$ and $k_{eff,3}$. Adam was chosen as optimizer. Training was performed over 10000 epochs with a batch size of 16. The results after training can be seen in **Figure 4**



(**a**) (**b**)

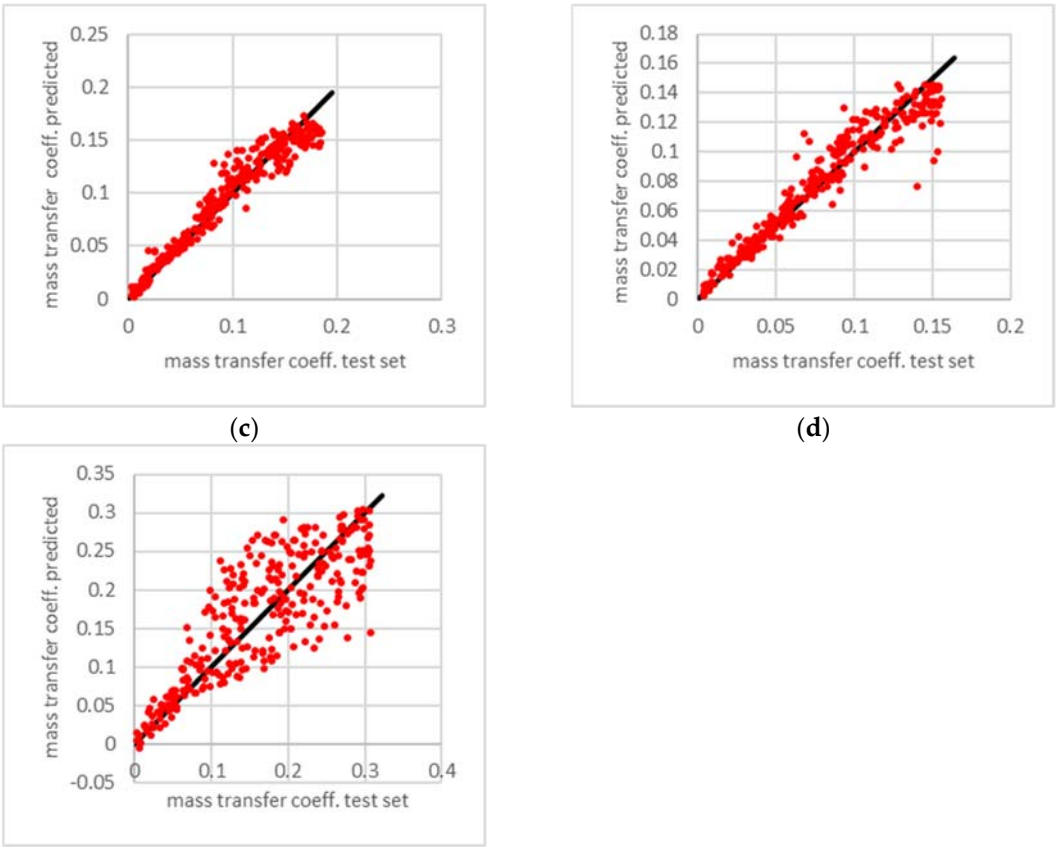


Figure 4: Predicted over original data of the test set. (a) depicts the results of the D_{ax} prediction with an R² over 99%, (b) depicts the the ε_s prediction with R² over 99%. The results of $k_{eff,IgG}$, $k_{eff,HCP1}$ and $k_{eff,HCP2}$ are illustraed in (c), (d) and (e) with R² of 91%, 88% and 59% respectively.

As shown in **Figure *4*** the parameters D_{ax} and ε_s can be predicted well from a single experiment, while the prediction performance of the k_{eff} values of all components seems insufficient with rising values of k_{eff}. As seen before in **Figure *3*** (h) the impact of values greater 10^{-4} cm^2/s is negligible. Therefore, the worse prediction accuracy can be explained by the low sensitivity of k_{eff} at higher values. To investigate this further the chromatograms were resimulated with the predicted parameters. An example chromatogram with the original parameters and the ANN predicted parameters is given in **Figure 5**.

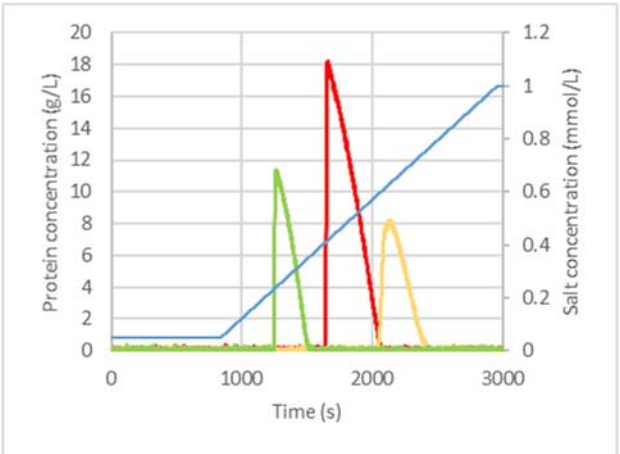


Figure 5: Chromatogram of the simulation with the original parameters (solid lines) and the resimulated chromatogram with the ANN predicted parameters (dashed lines). The red lines represent the target component IgG, the green line the side component group HCP1 and the yellow line the side component group HCP2. The blue line represents the salt gradient.

The predicted parameters have an error of lower than 5%, except for $k_{eff,HCP2}$. With values of $\sim 0.02\ cm^2/s$ (original) and $\sim 0.007\ cm^2/s$ the error is 172%. Despite the high error of $k_{eff,HCP2}$ the original chromatogram and the resimulated chromatogram perfectly match with coefficient of determination (R^2) of over 99%. The confidence interval at 95% level of the R^2 of the original and resimulated chromatograms of IgG is $[97.4,\ 99.9]$, of HCP1 is $97.5, 99.9]$ and of HCP2 is $[97.1, 99.4]$. Hence, this parameters can be predicted from a single chromatogram with k_{eff} only needing to be a rough estimate whether it is smaller or above $10^{-4}\ cm^2/s$.

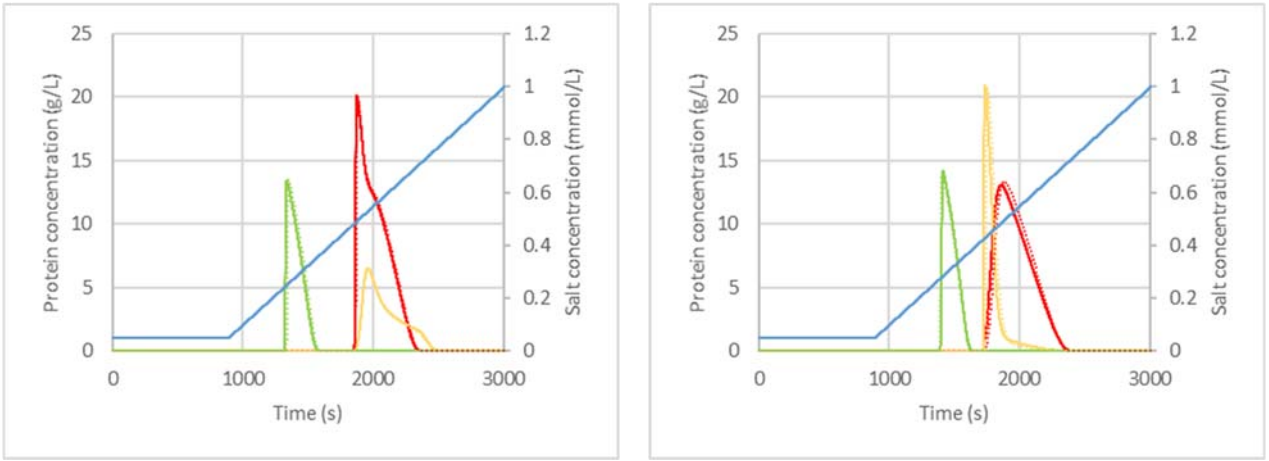
3.2. Variation of phase equilibrium parameters

To test the prediction ability of the phase equilibrium parameters a FF-ANN with 118 input neurons, 117 neurons with 20% dropout and tanh activation in hidden layer 1, 100 neurons with tanh activation function and 20% dropout and 12 output neurons with relu activation was implemented. As training algorithm the Adam optimizer was utilized. The dataset of 1000 simulative experiments was split into 70% training and 30% validation data. Training was performed over 10000 epochs with a batchsize of 16. The R^2 values of the original values over predicted values of the test set are shown in **Table 2**.

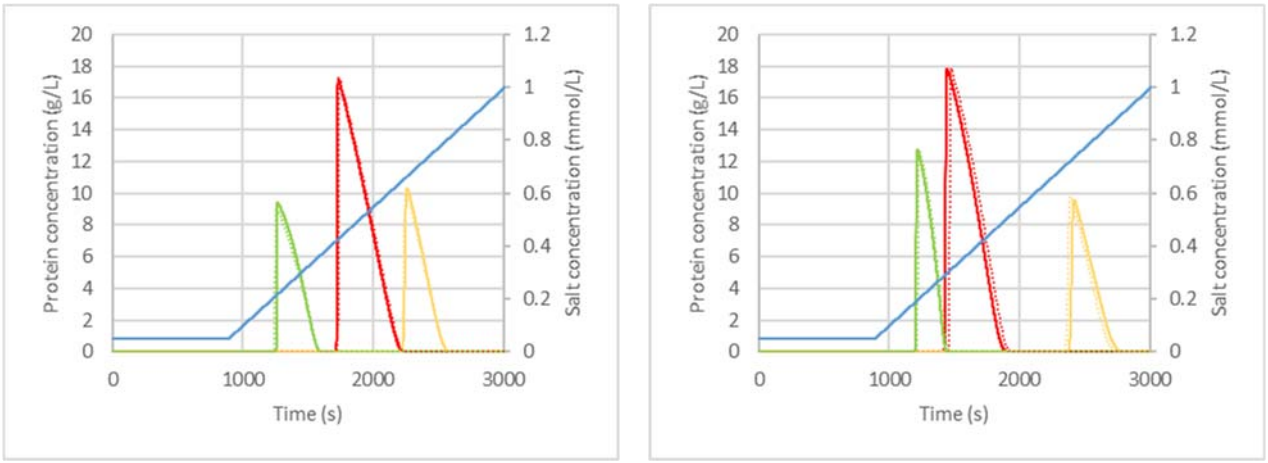
Table 2: Coefficient of determination of the phase equilibrium parameters from the test set.

	$a_{1,IgG}$	$a_{1,HCP1}$	$a_{1,HCP2}$	$a_{2,IgG}$	$a_{2,HCP1}$	$a_{2,HCP2}$	$b_{1,IgG}$	$b_{1,HCP1}$	$b_{1,HCP2}$	$b_{2,IgG}$	$b_{2,HCP1}$	$b_{2,HCP2}$
R^2	83%	85%	84%	78%	73%	91%	56%	10%	55%	80%	64%	92%

Again, it could be assumed that the prediction quality of all parameters is insufficient. Therefore, this hypothesis is tested by resimulating all chromatograms with the ANN predicted parameters. Some example chromatograms are given in **Figure 6**.



(**a**) (**b**)



(**c**) (**d**)

Figure 6 Comparison of four chromatograms with original and ANN predicted isotherm parameters from the validation set. In all subplots, the protein group HCP1 is green, the protein group HCP2 is yellow, and the IgG is red. Solid lines represent the chromatograms original isotherm parameters. Dashed lines represent the chromatograms with ANN predicted isotherm parameters: (**a**) Depicts strongly overlapping peaks with nearly maximum loading of IgG and HCP2 and strongly non-ideal Gauss-peaks with clear competitive behaviour. (**b**) Illustrates similar behaviour like (**a**) with peak switching. (**c**) and (**d**) show the typical "shark-fin" shape of highly loaded columns with baseline separation.

In contrast to the assumptions based on **Table 2** the resimulated chromatograms match the original chromatograms. This even applies to extreme conditions seen in subplot (a) and (b). Based on this, the results of the previous work [5] and the scenario in the next step the capability of predicting the phase equilibrium parameters is assumed to be sufficient.

3.3. Variation of phase equilibrium, fluid dynamic and packing parameters at once

After the possibility to predict the relevant parameters from one chromatogram was evaluated in the previous step a last dataset was created. For this dataset, the following scenario was set. The previously presented chromatography step starts as optimized process with the initial parameters of **Table 1**. It is assumed, that the parameters are affected by column aging/fouling and therefore move into the direction of decreasing performance of the column (except the flow) within the range of **Table 1**. The column is deemed insufficient as soon as the target protein's peak area (IgG) is overlapped by the side components' (HCP1, HCP2) peak area by more than 10%. Furthermore, it is assumed that a validated but noisy PAT concept for online concentration measurement is implemented. An overview is given in the previous publications [11,20]. Under these conditions, the covered parameter range of the dataset is reduced but extra complexity is added by the noise. The noisy initial chromatogram is shown in **Figure 7**.

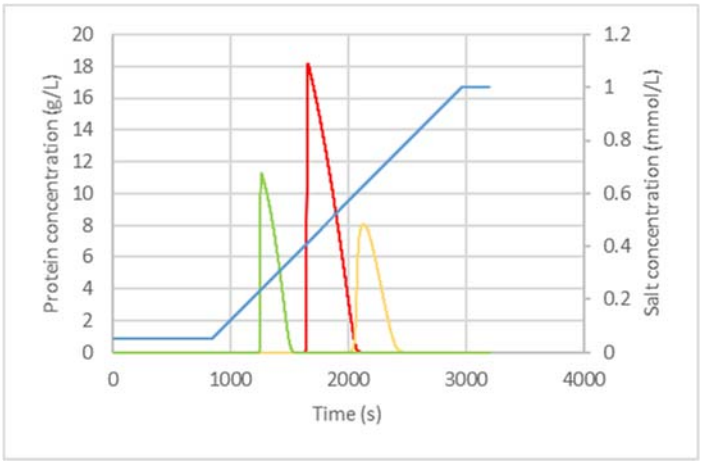


Figure 7: The initial noisy chromatogram before the separation performance decreases. The green curve shows HCP1, the red curve IgG, the yellow curve HCP2 and the blue curve shows the salt gradient.

To generate a dataset, which suffices overlapping area rule and does not add unnecessary complexity to the dataset, each chromatogram was evaluated after its simulation. If the 10% rule was violated, the chromatogram was discarded. This process was repeated until the dataset contained 1000 entries. Unexpectedly, no entries with k_{eff} values below $10^{-4} cm^2/s$ were generated. Because greater values had no impact on the chromatograms, the k_{eff} prediction was discarded after further investigation. Thus, the complexity of the prediction task can be reduced further. The resulting dataset was split in 70% training and 30% validation data. Afterwards a feed-forward ANN with 118 input neurons, 118 neurons with 20 % dropout and tanh activation function in hidden layer I, 100 neurons with 20% dropout and tanh activation layer and 14 output neurons with linear activation was trained. The neural network training utilized the Adam algorithm. It was trained over 30000 epochs with a batch size of 16.Again, the validation data was used to test the ANN performance. The predicted parameters were used to resimulate the validation set chromatograms. The results of the R^2 are summarized in the box plot in

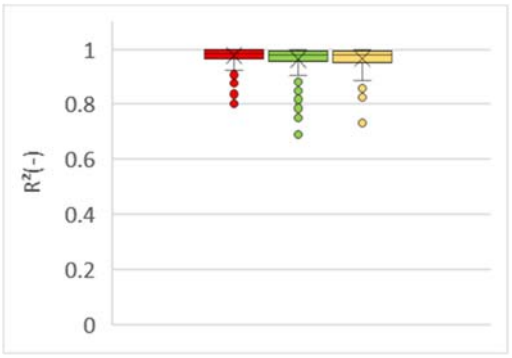


Figure 8: Box plots of the R^2 of the validation chromatograms with the predicted parameters over the chromatograms with the original parameters. The red box plot represents IgG, green represents HCP1 and yellow represents HCP2.

Except for a few outliers, a high prediction performance can be seen. The confidence intervals at 95% level for IgG, HCP1 and HCP2 are $[0.97, 0.98]$, $[0.96, 0.97]$ and $[0.96, 0.97]$ respectively. The performance may be increased by further optimizing the ANN structure, adding additional chromatograms to the dataset or applying smoothing algorithms to the input data instead of using the noisy raw data. Additionally more chromatograms from previous batches could be used to supply the ANN with more information and therefore increasing prediction performance.

4. Conclusions

In this study, an ANN for chromatography model parameter prediction under preparative production conditions was developed and evaluated. Instead of using a wide parameter range with multiple chromatograms like in previous works for a more screening like approach, only a single and noisy chromatogram was used for parameter prediction. After excluding irrelevant parameters for the set scenario via expert knowledge and one parameter at a time sensitivity studies the ANN was trained to predict the voidage, the axial dispersion coefficient, and phase equilibrium parameters. Although the prediction of the individual parameters indicated poor performance, simulating the chromatograms with predicted values showed that a high agreement between the original and the chromatograms with predicted parameters occurs. The authors explain this with the sensitivity of the parameters in certain areas. Therefore, the ANN can be used to track parameter deviations of the specified process. A schematic application was shown in Figure 1. The presented approach could be used as supportive tool in chromatography processes with model-based control and optimization in production environments. The influence of typical deviations caused by column aging or feed and buffer variations on the model parameters can be tracked by the ANN within 200 milliseconds. Therefore, a quick adjustment of the parameters to the current state of the process is possible by a chromatogram given by PAT. A drift of the model from reality can thus be counteracted, which leads to an increased performance of the control and optimization model over the whole process lifetime.

Author Contributions: Conceptualization: J.S.; methodology, design and evaluation: M.M., G.S.; chromatography modeling: F.L.V., writing, editing, and reviewing M.M, F.L.V., and J.S.; supervision: J.S. All authors have read and agreed to the published version of the manuscript.

Funding: The authors want to gratefully acknowledge the Bundesministerium für Wirtschaft und Energie (BMWi), especially Dr. Michael Gahr (Projektträger FZ Jülich), for funding the scientific work. We also kindly acknowledge the support by Open Access Publishing Fund of the Clausthal University of Technology.

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Data Availability Statement: Data cannot be made publicly available.

Acknowledgments: The authors would like to thank Dr. Reinhard Ditz, formerly of Merck KGaA, Darmstadt, for paper revision and fruitful discussions; Professor Andreas Potschka from the Institute of Mathematics at the Clausthal University of Technology for helpful input; as well as the ITVP lab-team, especially Frank Steinhäuser, Volker Strohmeyer and Thomas Knebel, for their efforts and support.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or the decision to publish the results.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or the decision to publish the results.

References

1. Zobel-Roos, S.; Mouellef, M.; Siemers, C.; Strube, J. Process Analytical Approach towards Quality Controlled Process Automation for the Downstream of Protein Mixtures by Inline Concentration Measurements Based on Ultraviolet/Visible Light (UV/VIS) Spectral Analysis. *Antibodies* **2017**, *6*, 24, doi:10.3390/antib6040024.
2. Golubović, J.; Protić, A.; Zečević, M.; Otašević, B.; Mikić, M. Artificial neural networks modeling in ultra performance liquid chromatography method optimization of mycophenolate mofetil and its degradation products. *J. Chemometrics* **2014**, *28*, 567–574, doi:10.1002/cem.2616.

3. 46. Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.S.; Davis, A.; Dean, J.; Devin, M.; et al. Tensorflow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. Available online: https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/45166.pdf (accessed on 19 December 2021).
4. Wang, G.; Briskot, T.; Hahn, T.; Baumann, P.; Hubbuch, J. Estimation of adsorption isotherm and mass transfer parameters in protein chromatography using artificial neural networks. *J. Chromatogr. A* **2017**, *1487*, 211–217, doi:10.1016/j.chroma.2017.01.068.
5. Mouellef, M.; Vetter, F.L.; Zobel-Roos, S.; Strube, J. Fast and Versatile Chromatography Process Design and Operation Optimization with the Aid of Artificial Intelligence. *Processes* **2021**, *9*, 2121, doi:10.3390/pr9122121.
6. *Ullmann's encyclopedia of industrial chemistry;* Wiley: Chichester, 2010, ISBN 3527306730.
7. Mestmäcker, F.; Schmidt, A.; Huter, M.; Sixt, M.; Strube, J. Systematic and Model-Assisted Process Design for the Extraction and Purification of Artemisinin from Artemisia annua L.—Part III: Chromatographic Purification. *Processes* **2018**, *6*, 180, doi:10.3390/pr6100180.
8. Ströhlein, G.; Aumann, L.; Müller-Späth, T.; Tarafder, A.; Morbidelli, M. CONTINUOUS PROCESSING: The Multiclomn Countercurrent Solvent Gradient Purification Process: A continuous chromatographic process for monoclonal antibodies without using Protein A. *BioPharm International* **2007**, *22*, 42–48.
9. Zobel, S.; Helling, C.; Ditz, R.; Strube, J. Design and Operation of Continuous Countercurrent Chromatography in Biotechnological Production. *Ind. Eng. Chem. Res.* **2014**, *53*, 9169–9185, doi:10.1021/ie403103c.
10. Helling, C.; Dams, T.; Gerwat, B.; Belousov, A.; Strube, J. Physical characterization of column chromatography: stringend control over equipment performance in biopharmaceutical production. *Trends in Chromatography* **2013**, *8*, 55–71.
11. Zobel-Roos, S.; Schmidt, A.; Mestmäcker, F.; Mouellef, M.; Huter, M.; Uhlenbrock, L.; Kornecki, M.; Lohmann, L.; Ditz, R.; Strube, J. Accelerating Biologics Manufacturing by Modeling or: Is Approval under the QbD and PAT Approaches Demanded by Authorities Acceptable Without a Digital-Twin? *Processes* **2019**, *7*, 94, doi:10.3390/pr7020094.
12. Zhang, J.; Siva, S.; Caple, R.; Ghose, S.; Gronke, R. Maximizing the functional lifetime of Protein A resins. *Biotechnol. Prog.* **2017**, *33*, 708–715, doi:10.1002/btpr.2448.
13. Rathore, A.S.; Bracewell, D.G.; Phatak, M.; Ma, G. Re-use of Protein A Resin: Fouling and Economics. *BioPharm International* **2015**, *28*, 28–33.
14. Andersson, J.A.E.; Gillis, J.; Horn, G.; Rawlings, J.B.; Diehl, M. CasADi: a software framework for nonlinear optimization and optimal control. *Math. Prog. Comp.* **2019**, *11*, 1–36, doi:10.1007/s12532-018-0139-4.
15. Kornecki; Schmidt; Lohmann; Huter; Mestmäcker; Klepzig; Mouellef; Zobel-Roos; Strube. Accelerating Biomanufacturing by Modeling of Continuous Bioprocessing—Piloting Case Study of Monoclonal Antibody Manufacturing. *Processes* **2019**, *7*, 495, doi:10.3390/pr7080495.
16. *Process analytical technology*: *Spectroscopic tools and implementation strategies for the chemical and pharmaceutical industries;* Bakeev, K.A., Ed., [Nachdr.]; Blackwell Publ: Oxford, 2006, ISBN 1-4051-2103-3.
17. Kessler, W. *Multivariate Datenanalyse für die Pharma-, Bio- und Prozessanalytik*: *Ein Lehrbuch*, 1. Aufl., 1. Nachdr; WILEY-VCH: Weinheim, 2008, ISBN 9783527312627.
18. Vetter, F.L.; Zobel-Roos, S.; Strube, J. PAT for Continuous Chromatography Integrated into Continuous Manufacturing of Biologics towards Autonomous Operation. *Processes* **2021**, *9*, 472, doi:10.3390/pr9030472.
19. Rathore, A.S.; Kapoor, G. Application of process analytical technology for downstream purification of biotherapeutics. *J. Chem. Technol. Biotechnol.* **2015**, *90*, 228–236, doi:10.1002/jctb.4447.
20. Helgers, H.; Schmidt, A.; Lohmann, L.J.; Vetter, F.L.; Juckers, A.; Jensch, C.; Mouellef, M.; Zobel-Roos, S.; Strube, J. Towards Autonomous Operation by Advanced Process Control—Process Analytical Technology for Continuous Biologics Antibody Manufacturing. *Processes* **2021**, *9*, 172, doi:10.3390/pr9010172.
21. van Rossum, G. *The Python language reference*, Release 3.0.1 [Repr.]; Python Software Foundation; SoHo Books: Hampton, NH, Redwood City, Calif., 2010, ISBN 1441412697.
22. Andersson, J.; Gillis, J. *Casadi 3.4.4;* Zenodo, 2018.
23. Spyder IDE. *Avaiable online: https://www.spyder-ide.org/ (accessed on 19 Decembre 2021).*
24. Guiochon, G. *Fundamentals of preparative and nonlinear chromatography*, 2. ed.; Elsevier Acad. Press: Amsterdam, 2006, ISBN 978-0123705372.

25. Zobel-Roos, S. Entwicklung, Modellierung und Validierung von integrierten kontinuierlichen Gegenstrom-Chromatographie-Prozessen. Dissertation; Shaker Verlag GmbH; Technische Universität Clausthal.
26. Zobel-Roos, S.; Mouellef, M.; Ditz, R.; Strube, J. Distinct and Quantitative Validation Method for Predictive Process Modelling in Preparative Chromatography of Synthetic and Bio-Based Feed Mixtures Following a Quality-by-Design (QbD) Approach. *Processes* **2019**, *7*, 580, doi:10.3390/pr7090580.
27. Carta, G.; Rodrigues, A.E. Diffusion and convection in chromatographic processes using permeable supports with a bidisperse pore structure. *Chemical Engineering Science* **1993**, *48*, 3927–3935, doi:10.1016/0009-2509(93)80371-V.
28. Wilson, E.J.; Geankoplis, C.J. Liquid Mass Transfer at Very Low Reynolds Numbers in Packed Beds. *Ind. Eng. Chem. Fund.* **1966**, *5*, 9–14, doi:10.1021/i160017a002.
29. Carta, G.; Jungbauer, A. *Protein Chromatography;* Wiley, 2010, ISBN 9783527318193.
30. Seidel-Morgenstern, A.; Guiochon, G. Modelling of the competitive isotherms and the chromatographic separation of two enantiomers. *Chemical Engineering Science* **1993**, *48*, 2787–2797, doi:10.1016/0009-2509(93)80189-W.
31. Leśko, M.; Åsberg, D.; Enmark, M.; Samuelsson, J.; Fornstedt, T.; Kaczmarski, K. Choice of Model for Estimation of Adsorption Isotherm Parameters in Gradient Elution Preparative Liquid Chromatography. *Chromatographia* **2015**, *78*, 1293–1297, doi:10.1007/s10337-015-2949-0.
32. Mollerup, J.M. A Review of the Thermodynamics of Protein Association to Ligands, Protein Adsorption, and Adsorption Isotherms. *Chem. Eng. Technol.* **2008**, *31*, 864–874, doi:10.1002/ceat.200800082.
33. Brooks, C.A.; Cramer, S.M. Steric mass-action ion exchange: Displacement profiles and induced salt gradients. *AIChE J.* **1992**, *38*, 1969–1978, doi:10.1002/aic.690381212.
34. Langmuir, I. THE ADSORPTION OF GASES ON PLANE SURFACES OF GLASS, MICA AND PLATINUM. *J. Am. Chem. Soc.* **1918**, *40*, 1361–1403, doi:10.1021/ja02242a004.
35. Seidel-Morgenstern, A. Experimental determination of single solute and competitive adsorption isotherms. *J. Chromatogr. A* **2004**, *1037*, 255–272, doi:10.1016/j.chroma.2003.11.108.
36. Fausett, L.V. *Fundamentals of neural networks*: *Architectures, algorithms, and applications;* Prentice Hall: Englewood Cliffs, N.J., 1994, ISBN 0-13-334186-0.
37. *Computational analysis and understanding of natural languages*: *Principles, methods and applications;* Gudivada, V.N.; Rao, C.R., Eds.; North-Holland an imprint of Elsevier: Amsterdam, Oxford, 2018, ISBN 978-0-444-64042-0.
38. Asprion, N.; Böttcher, R.; Pack, R.; Stavrou, M.-E.; Höller, J.; Schwientek, J.; Bortz, M. Gray-Box Modeling for the Optimization of Chemical Processes. *Chemie Ingenieur Technik* **2019**, *91*, 305–313, doi:10.1002/cite.201800086.
39. Hagge, T.; Stinis, P.; Yeung, E.; Tartakovsky, A.M. *Solving differential equations with unknown constitutive relations as recurrent neural networks*, 2017. Available online: https://arxiv.org/pdf/1710.02242.
40. Gao, W.; Engell, S. Neural Network-Based Identification of Nonlinear Adsorption Isotherms. *IFAC Proceedings Volumes* **2004**, *37*, 721–726, doi:10.1016/S1474-6670(17)31895-5.
41. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep learning;* MIT Press: Cambridge, Massachusetts, London, England, 2016, ISBN 978-0262035613.
42. Klambauer, G.; Unterthiner, T.; Mayr, A.; Hochreiter, S. Self-Normalizing Neural Networks. *Advances in Neural Information Processing Systems 30 (NIPS*.
43. Kruse, R.; Borgelt, C.; Braune, C.; Klawonn, F.; Moewes, C.; Steinbrecher, M. *Computational Intelligence;* Springer Fachmedien Wiesbaden: Wiesbaden, 2015, ISBN 978-3-658-10903-5.
44. Keras. Available online: https://keras.io (accessed on 19 December 2021).
45. Young, M.E.; Carroad, P.A.; Bell, R.L. Estimation of diffusion coefficients of proteins. *Biotechnol. Bioeng.* **1980**, *22*, 947–955, doi:10.1002/bit.260220504.
46. Chung, S.F.; Wen, C.Y. Longitudinal dispersion of liquid flowing through fixed and fluidized beds. *AIChE J.* **1968**, *14*, 857–866, doi:10.1002/aic.690140608.