

Multi-scale membrane process optimization with high-fidelity ion transport models through machine learning

Deniz Rall^{a,b,*}, Artur M. Schweidtmann^{c,*}, Maximilian Kruse^a, Elizaveta Evdochenko^a,
Alexander Mitsos^c, Matthias Wessling^{a,b,**}

^a*RWTH Aachen University, AVT.CVT - Aachener Verfahrenstechnik, Chemical Process Engineering,
Forckenbeckstrasse 51, 52074 Aachen, Germany*

^b*DWI - Leibniz Institute for Interactive Materials,
Forckenbeckstrasse 50, 52074 Aachen, Germany*

^c*RWTH Aachen University, AVT.SVT - Aachener Verfahrenstechnik, Process Systems Engineering,
Forckenbeckstrasse 51, 52074 Aachen, Germany*

Abstract

Innovative membrane technologies optimally integrated into large separation process plants are essential for economical water treatment and disposal. However, the mass transport through membranes is commonly described by nonlinear differential-algebraic mechanistic models at the nano-scale, while the process and its economics range up to large-scale. Thus, the optimal design of membranes in process plants requires decision making across multiple scales, which is not tractable using standard tools. In this work, we embed artificial neural networks (ANNs) as surrogate models in the deterministic global optimization to bridge the gap of scales. This methodology allows for deterministic global optimization of membrane processes with accurate transport models – avoiding the utilization of inaccurate approximations through heuristics or short-cut models. The ANNs are trained based on data generated by a one-dimensional extended Nernst-Planck ion transport model and extended to a more accurate two-dimensional distribution of the membrane module, that captures the filtration-related decreasing retention of salt. We simultaneously design the membrane and plant layout yielding optimal membrane module synthesis properties along with the optimal plant design for multiple objectives, feed concentrations, filtration stages, and salt mixtures. The developed process models and the optimization solver are available open-source, enabling computational resource-efficient multi-scale optimization in membrane science.

Keywords: Nanofiltration, Data-driven surrogate model, E_nPE_n model, Artificial neural network, Deterministic global superstructure optimization

*These authors contributed equally to this work

**Corresponding author: manuscripts.cvt@avt.rwth-aachen.de

1. Introduction

Global population growth and increasing urbanization require novel economically viable concepts for drinking water treatment and water disposal [1]. This challenge includes smart operations, such as data-driven urban water management [2], but also new treatment concepts and technologies [3, 4]. Synthetic membranes are an essential technological basis for this separation processes [5, 6]. Innovative membrane technologies help to develop new approaches for the reuse of sustainable raw material sources from urban and industrial waste water [7, 8] as well as producing smart process water tailored to its application [9]. These ambitious aspirations require a versatile modeling environment and rigorous optimization methodologies for membrane systems made for case-specific customized processes.

Nowadays, ion selectivity becomes increasingly essential for sustainable drinking water treatment and electrochemical processes [10, 11]. A transition followed from the complete removal of all minerals from drinking water through reverse osmosis membranes towards more selective nanofiltration membranes. The selectivity of ionic components enables the recovery of sustainable raw material sources from urban and industrial waste water [12], such as phosphorus [13]. The development of ion separation membranes is driven by the advent of the layer-by-layer (LbL) nanofiltration technology [14–16]. The layer-by-layer (LbL) nanofiltration technology is characterized by oppositely charged polymer layers applied to a conventional highly porous ultrafiltration membrane. The layer-wise manufacturing technique enables precise tuning of selectivity and permeability [17–19] and a scale-up from the laboratory to industrial scale [20]. However, tailoring of the selectivity of charged and uncharged species [21–23] is the predominant challenge. Moreover, exploiting all synergies in membrane development through the simultaneous design of LbL membranes, the process structure, and its operation is an open research question [24].

Detailed mechanistic models describe the mass transport and separation by membranes at the nano-scale. There exists a variety of different membrane models for ion separation [25–29]. These mechanistic models range from highly accurate 3D models to low fidelity heuristics, as shown in Figure 1. Most notably, a trade-off exists between the fidelity of the approach and its 3D computational resource-efficiency [30]. Accurate mechanistic models describe ion transport mechanisms in multiple dimensions, leading to partial differential equations (PDE) that require special solvers for the solution, i.e., computational fluid dynamics (CFD) for 3D. Lumping dimensions reduces the dimensionality to 2D and 1D models that can be evaluated at lower computational effort at the cost of fidelity. Finally, heuristics describe lumped systems (0D) and can be assessed at little computational effort, but their predictive capability is insufficient to describe the complex ion transport. Moreover, only experiments obtain high fidelity data directly. This results in the dilemma that increased accuracy leads to enormous computational effort and cannot be used in the context of process optimization by using standard tools.

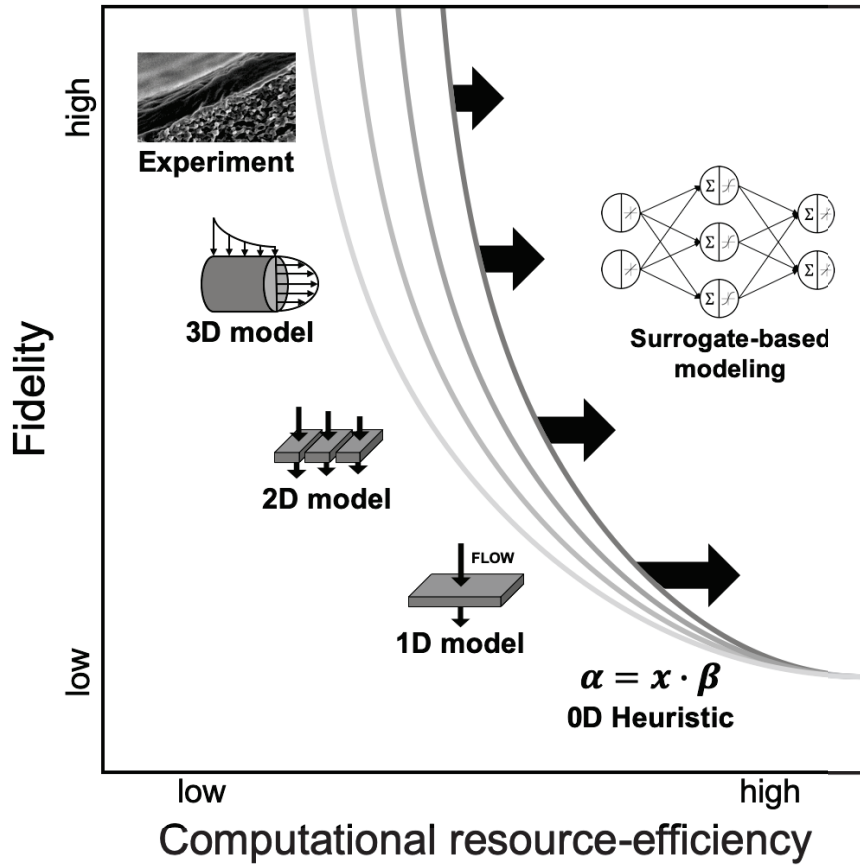


Figure 1. ANN-based surrogate models push the boundary of computational resource-efficiency: Fidelity and computational resource-efficiency trade-off of different descriptive approaches. The descriptive approaches include experiments, simulation studies, and heuristics. Adapted from [30]

Economically viable and sustainable process design requires decision making at large process scales (potentially also considering environmental effects [31]). At the same time, physical models of the complex ion transport are described at small-scales. This multi-scale problem has strong implications for optimal decision making: On the one hand, superstructure process optimization with 3D, 2D, and 1D models embedded lead to (mixed-integer) optimization problems with PDE constraints. These mixed-integer dynamic optimization (MIDO) problems [32–34] are very difficult-to-solve and their deterministic global optimization is currently limited to small-scale problems [35]. Thus, process optimization using this approach is currently not possible. On the other hand, heuristics or short-cut models can be embedded computational resource-efficiently in complex optimization superstructures [36–42]. However, these simplified heuristics or short-cut models do not sufficiently describe the influence of membrane parameters, such as the synthesis protocols or process environment, on the membrane performance. Simultaneous development of the membrane synthesis, along with the process using heuristics or short-cut models, is not possible.

Overcoming the trade-off between fidelity and computational resource-efficiency is necessary for deterministic global membrane process optimizations with the simultaneous development of the membrane synthesis. One way to

overcome this issue is to replace the expensive simulation by a surrogate model [43]. Here, the accurate model is evaluated offline, creating a training data set that is subsequently learned by a supervised machine learning algorithm. This results in a data-driven surrogate model that approximates the accurate model. Then, the data-driven surrogate model can be combined with further data-driven and mechanistic models yielding a hybrid mechanistic / data-driven model that can be optimized [44]. The use of machine learning models enables bridging scales, e.g., in material design [45–50] or process systems engineering [51, 52]. In addition, data-driven models have been used in various disciplines for process optimization (e.g., [53–58]). Further, we have found that hybrid mechanistic / data-driven models with artificial neural networks (ANNs) embedded can be optimized efficiently by a reduced-space deterministic global optimization method [59]. In the past, this new optimization approach has already been used successfully for the optimization of energy processes [60–64].

Many publications in the field of membrane science address the implementation of neural networks in membrane development [65–67] or prediction of membrane operation [68–70]. Recently, we extended the use of ANN surrogate models in membrane science to describe and optimize LbL-based membrane systems by identifying superior membrane synthesis protocols based on the delicate trade-off between ion retention characteristics and permeability [21]. Next, a coupling of the ANNs into a hybrid mechanistic / data-driven model enables an optimization strategy to simultaneously design the performance of LbL nanofiltration membrane modules and the separation process. This yields membrane synthesis protocols and membrane processes that are optimally tailored to the desired separation task [24]. The results suggest that simultaneous membrane synthesis and process optimization design achieve immediately favorable results with lower impurities at comparable costs. But in the previous work, this optimization has been limited to lab data. In particular, only a single-stage membrane process has been considered because the available experimental data is only valid for a fixed feed concentration and single salt solutions. Thus, an extension to accommodate for higher feed concentrations, multiple-staged processes, and salt mixtures is highly desired.

In this work, we propose a surrogate-based approach to bridge the gap between mechanistic ion transport models at the nano-scale and optimal process design through deterministic global superstructure optimization at a large-scale. Thereby, we facilitate ANN surrogate models trained on data generated by a one-dimensional ion transport model, which are subsequently embedded in a state-of-the-art membrane process optimization model.

We use the extended Nernst-Planck model, called pEnPEn, for describing the ion transport through the membrane [28, 71, 72]. The model pEnPEn describes ion transport through multi-layered geometries (i.e., LbL nanofiltration membranes) consisting of n electrolyte layers (En) with n polyelectrolyte layers, i.e., membranes, (PEn). An applied pressure (p) acts as the driving force of the separation process. The aforementioned one-dimensional ion transport model is evaluated offline, creating an extensive data set that describes the membrane's ionic retention based on membrane-specific parameters (such as, the layer charge X and the layer thickness Δx) and process-specific parameters (such as, the transmembrane velocity v , and ionic feed concentration c_j). A subsequent training of the data creates an ANN surrogate model.

Next, the surrogate models are exploited towards a more accurate two-dimensional distribution of the membrane module even to capture the filtration-related decreasing retention of salt. Therefore, individual ANNs are arranged in series to resolve the membrane in the direction of flow. The data that is generated by this series of ANNs are used to create a two-dimensional surrogate model, which is then embedded in the membrane process optimization model. This enables the complex mechanistic model to be applied computational resource-efficiently in an optimization context, despite the partial differential-algebraic form of the original model.

The paper is organized as follows: First, we present the data generation and the learning of ANN surrogate models. Second, the more accurate two-dimensional distribution of the membrane module is presented through ANN surrogate modeling. Third, the hybrid mechanistic / data-driven multi-scale process model is applied to a case study optimizing nanofiltration membrane plants for multiple objectives, feed concentrations, filtration stages, and salt mixtures. The membrane module synthesis properties are optimized along with the superstructure of the membrane plant. In particular, process configurations of the plant (i.e., number of filtration stages, recirculating of streams, pump power) as well as the membrane synthesis properties (i.e., the layer charge X and the layer thickness Δx) are degrees of freedom and are optimized simultaneously.

This work sets the foundation for computational resource-efficient multi-scale modeling integrating simulation studies of the complex ion transport in the decision making process at large process scale optimizations. Finally, all developed models and the optimization solvers are available open-source, making it a viable tool for future research and industrial applications.

2. Methodology

In this section, the workflow of the data-driven approach is described to bridge the scales between (i) high fidelity ion transport models on a nano-scale and (ii) deterministic global membrane process optimization on a large-scale. First, the high fidelity simulation of the ion transport through the membrane on the nano-scale is replaced by a surrogate model. This surrogate model based replacement is performed by evaluating the high fidelity ion transport model offline and creating a training data set that is subsequently learned by a supervised machine learning algorithm (here ANNs). Then, the resulting data-driven ANN-based surrogate model that approximates the accurate model is embedded in a hybrid mechanistic / data-driven model and optimized to find optimal membrane plant process configurations. In summary, this procedure avoids PDE constraints in the formulation of the deterministic global optimization problem. Thereby, this method enables deterministic optimization of membrane processes with surrogate models at a low computational cost similar to lumped systems (i.e., heuristics or short cut models) but significantly more accurate.

The workflow commences as depicted in Figure 2. (1) The high fidelity ion transport model (Section 2.1) is evaluated offline to create an extensive data set, that (2) is used to train the data-driven ANN-based surrogate models on the data (Section 2.2). (3) Moreover, a surrogate model, including a two-dimensional distribution of the membrane module is created for higher fidelity. The most precise two-dimensional distribution and additional data generation are described in Section 2.3, and the implementation of the two-dimensional surrogate model is described in Section 2.4. (4) Next, the individual surrogate models are integrated with a mechanistic process model to a hybrid mechanistic / data-driven process model. In Section 2.5, we explain the overall mechanistic process model, including cost correlations. (5) Finally, the hybrid model is optimized using a reduced-space formulation [59] and our open-source deterministic global solver MAiNGO [73]. In Section 2.6, we briefly describe the optimization method.

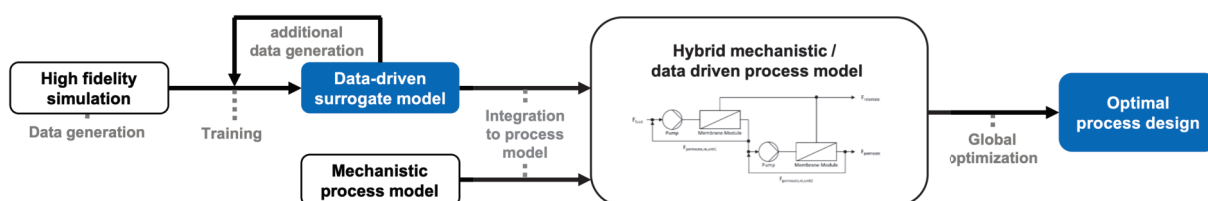


Figure 2. The workflow of the data-driven approach. The input data is created through the high fidelity ion transport model simulation study. Next, an ANN takes this data to learn the ionic separation performance of the membrane based on the model, creating a data-driven ANN-based surrogate model. Two-dimensional distribution of the membrane module is constructed for higher fidelity of the surrogate model. For this purpose, the surrogate model is used to generate additional data that is trained in a new ANN surrogate model. Next, this ANN surrogate model is embedded in a mechanistic process model. Finally, the optimal process design is solved by deterministic global process optimization.

All tools necessary to rebuild and extend this method are open-source available to everyone. The process models and optimization problems are available open-source (<http://permalink.avt.rwth-aachen.de/?id=506639>). The training data, the trained artificial neural networks, and the optimization results are available in the elec-

tronic supplementary material of this article (published by the journal). The artificial neural network models and training scripts are available open-source within the “MeLOn - Machine Learning Models for Optimization” toolbox [74] (<https://git.rwth-aachen.de/avt.svt/public/MeLOn/>). Furthermore, the deterministic global solver MAiNGO is also available open-source (<https://git.rwth-aachen.de/avt.svt/public/maingo>).

2.1. Data generation by high fidelity ion transport model

In this work, the initial data set to train the ANN surrogate model is created by a simulation study using a high fidelity mechanistic ion transport model. We use a one-dimensional extended Nernst-Planck model, called pEnPEn [28, 71], to describe the ion transport through multi-layered geometries. LbL nanofiltration membranes can be modeled with this model [72] by introducing the layer thickness Δx of the separation layer with a homogeneous layer charge X as shown in Figure 3.

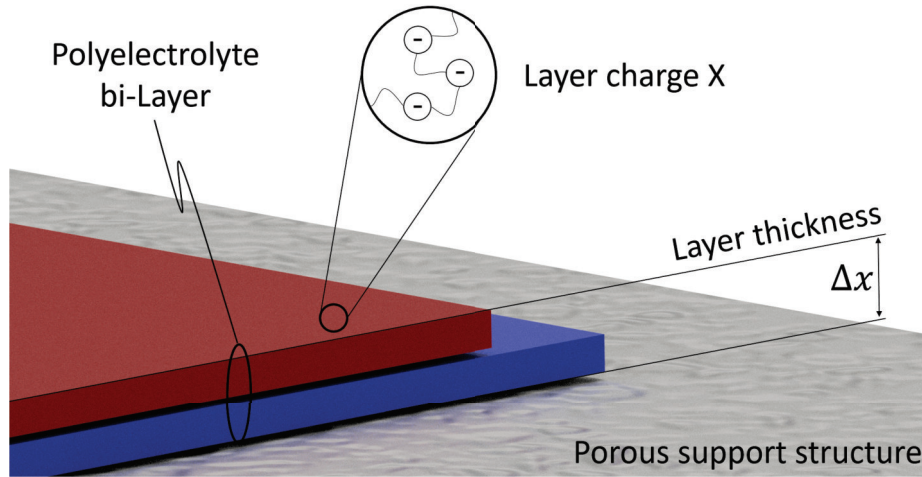


Figure 3. Overview of model parameters of the extended ion transport model pEnPEn [28, 71, 72]. The layer charge X , and the layer thickness Δx of the separation layer determine the separation performance. These values are inputs to the ANN and serve as a degree of freedom during simultaneous optimization of membrane module synthesis properties along with the superstructure of the membrane plant.

The model is evaluated offline, creating a large data set that describes the membrane’s ionic retention based on membrane-specific parameters (such as, the layer charge X and the layer thickness Δx) and process-specific parameters (such as, the transmembrane velocity v , ionic feed concentration c_j). The generation of data is automated, utilizing a wrapper script. This generates a set of training data inputs/outputs by a Latin hypercube sampling design for the ANN surrogate model. The variables are listed as follows:

- The **layer charge** X [mol m^{-3}] is a membrane-specific variable and **input** to the ANN. For single salts the layer charge varies in the interval $[-500, -100] \subset \mathbb{R}$. For salt mixtures this interval needs to be extended to $[-1000, -10] \subset \mathbb{R}$ to account for the more challenging separation task. The membrane charge needs to be higher as compared to single salts due to the consideration of higher ionic concentrations for salts mixtures.

- The separation **layer thickness** Δx [nm] is a membrane-specific variable and **input** to the ANN. The separation layer thickness varies in the interval $\in [75, 150] \subset \mathbb{R}$ according to experimental observations of [20].
- The **transmembrane velocity** v [$\mu\text{m s}^{-1}$] is a process-specific variable and **input** to the ANN. The retention of ions strongly depends on the transmembrane velocity, i.e., recovery rate, and varies in this study in the interval $[0, 50] \subset \mathbb{R}$.
- The **salt feed concentration** c_j [mol m^{-3}] is a process-specific variable and **input** to the ANN. The retention of ions strongly depends on the feed concentration. The range of salt feed concentration for a single membrane unit is $[1, 50] \subset \mathbb{R}$. When using two membrane filtration units in series, an extension of the data for feed concentrations below 1 mol m^{-3} is needed. Therefore, the data is extended using an additional Latin hypercube sampling to extend the data towards concentrations in the interval $[0.01, 50] \subset \mathbb{R}$.
- The **ionic retention of ion j** R_j [%] is the **output** to the ANN.

The one-dimensional model is established for the membrane itself, along with two adjacent diffusive layers. The model is implemented in COMSOL Multiphysics® [75] and executed via the MATLAB [76] API. The partial differential extended Nernst-Planck equation is simplified to a second-order ordinary differential equation by considering only the steady-state behavior of the charged species. Moreover, the pressure-driven convective flux is described via the first-order ordinary differential equation Darcys law. The boundary conditions are applied for the pressure at the inlet and outlet of the domain. Additionally, charge neutrality across the membrane is assumed. The resulting problem formulation involves the coupling of two different ordinary differential equations and is highly non-linear. It is solved with an adaptive Newton-Rapson algorithm. The systems matrix of the linearized problem is inverted using the MUMPS algorithm. This modeling approach is generic and thus applies to all combinations of salts used in this study.

Due to convergence issues in the range of low initial salt feed concentrations (i.e., at high local gradients in the second-order ordinary differential equation), a load ramping technique is applied to the excess charge parameter on the membrane surface. Input combinations to the one-dimensional ion transport model for which the solver did not converge are excluded from the training set.

2.2. Training of the data-driven surrogate model

In this study, we utilize supervised machine learning techniques for the creation of surrogate models. In literature, there exists a large variety of machine learning techniques, including Gaussian processes, artificial neural networks, and decision trees for correlations of varying complexity. In this work, we implement ANNs as surrogate models, as they can capture the ion transport in membranes based on synthesis parameters, as demonstrated in our previous work [21, 24]. ANNs are black box models that are frequently used for unsupervised and supervised learning tasks, including pattern recognition, classification, and regression [77]. ANNs learn training data and can approximate an

underlying input-output function. The evaluation of ANNs is given by an explicit algebraic function that depends on the architecture of the ANN, i.e., the connection of neurons [78]. This function is embedded in the process optimization problem and is efficiently solved by a reduced-space deterministic global optimization method [59].

In this work, a shallow feed-forward multi-layered perceptron ANN including one hidden layer is utilized to recreate the one-dimensional extended Nernst-Planck mechanistic ion transport model (pEnPEn) from the previous section. A hyperbolic tangent transfer function is employed in the hidden- and output layer. The structure of the ANN is depicted in Figure 4. Inputs to the ANN are the layer charge X , the layer thickness Δx , the transmembrane velocity v , and the salt feed concentration c_j . Output to the ANN is the ionic retention of ion j R_j .

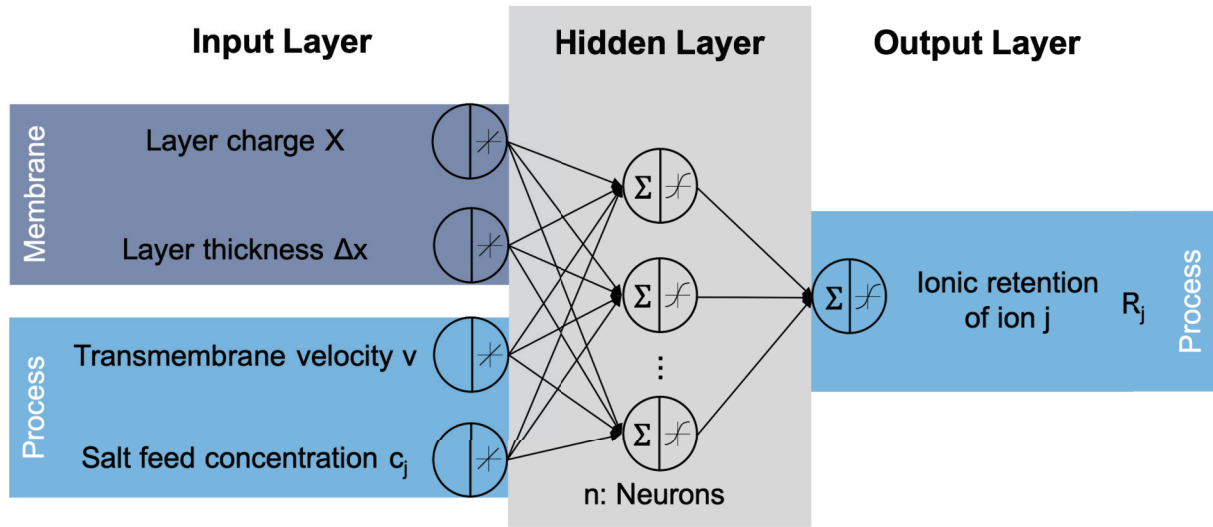


Figure 4. Multi-layered perceptron ANN with one hidden layer used this study as a data-driven surrogate model to recreate the ion transport through nanofiltration membranes. The training is based on data generated by a one-dimensional extended Nernst-Planck mechanistic ion transport model (pEnPEn). The membrane-specific parameters (such as the layer charge X and the layer thickness Δx) and process-specific parameters (such as the transmembrane velocity v and salt feed concentration c_j) are inputs to the ANN. The ionic retention of ion j R_j is the output to the ANN.

The main drawback of the data-driven modeling approach using ANNs is its limited extrapolation capability. Meaning that any data-driven model is not valid anymore outside of its training data domain (i.e., its convex envelope). Therefore, this results in a large data requirement for the training of ANNs. In our study, the training data of the ANNs is generated by a mechanistic ion transport model that allows us to have any number of data points (within computational reason) for training and validation. A consideration of probabilities, i.e., through Gaussian processes, is omitted because we use a data set extracted from a mechanistic model. The training data set we use for this study is small compared to many other machine learning applications, such as deep artificial neural networks with millions of parameters that are frequently used for image recognition. Thus, we select a shallow artificial neural network architecture with one hidden layer ranging from 6 to 14 hidden neurons because the input dimensionality and training data set are much smaller. Furthermore, the training of the ANNs is optimized to minimize adverse effects using small

data sets including: (a) multiple runs of ANN trainings to mitigate sporadic fluctuations in artificial neural network performance, (b) a randomized data set for training, validation and test to account for random effects due to small test data, and (c) the use of a k-fold cross-validation to prevent overshooting between data sets and avoid overfitting of the data set.

The data preparation and training procedure are based on our previous work [21, 24], and adapted accordingly. The architecture of ANNs is flexible and can be adjusted by setting its hyperparameters, e.g., number of layers, for application. In general, the architecture should be selected appropriately to avoid undesired under- and overfitting effects. The available data set is split into 70 % training, 15 % validation, and 15 % test set at random. Early stopping is used to prevent overshooting during all ANN training. All ANNs used in this work are created in the MATLAB environment [76, 79]. The training procedure commences in the following steps: First, the hidden layer size is determined by k-fold crossvalidation with a range of $n \in [1, 50] \subset \mathbb{N}$ hidden neurons. The minimum of the characteristic bias/variance trade-off bathtub curve of the mean-squared-error (MSE) determines the best number of hidden neurons n and thereby the hidden layer size. Second, ANNs are trained manifold with the previously determined hidden layer size $n \pm 3$ using a split of 70 % training, 15 % validation, and 15 % test set at random. Third, the best performing candidate of the ANN is selected for the following case studies.

Two ANNs are utilized as one dimensional data-driven surrogate model considering a single salt Na_2SO_4 or either a salt mixture of NaCl and Na_2SO_4 . Here, only the predominant ion species (i.e., the ions Cl^- and SO_4^{2-}) of the filtrated salt mixture are trained as inputs to the ANN. This consideration is valid due to electroneutrality in the solution and membrane. The monovalent cation Na^+ is calculated based on a mass balance. For single salts the ANN has one output (i.e., Na_2SO_4 retention) and for salt mixtures, the ANN has two outputs (i.e., Cl^- retention and SO_4^{2-} retention). Negative salt retentions are possible [80] and are mapped by the ANN. A list of properties of both ANNs is summarized in Table 1.

Table 1. One-dimensional data-driven surrogate models used in this work to recreate the one-dimensional extended Nernst-Planck high fidelity mechanistic ion transport model (pEnPEN). The number of neurons n of each ANN, the amount of training data set entries, and the section in which it is used is given.

Network name	Hidden neurons n	Description	Training points	Used in Section
$\text{ANN}_{\text{Na}_2\text{SO}_4, k=1, 1 \text{ Unit}}$	6	Single salt	998	2.3, 3.3
$\text{ANN}_{\text{NaCl}, \text{Na}_2\text{SO}_4, k=1, 1 \text{ Unit}}$	8	Salt mixture	1,000	2.3

2.3. Two-dimensional distribution of the membrane module

Up until now, the membrane unit is considered where only the changes in the orthogonal direction of the membrane are considered by the one-dimensional surrogate model of the previous section. However, due to the retention of salt at the membrane, the feed concentration in the module increases with the length of the membrane. The feed concentration of the salt strongly influences the retention of the nanofiltration membrane. Changes in feed concentration

in the direction of flow are not considered so far.

In Figure 5, a two-dimensional distribution of the membrane module is proposed in flow direction through ANN-based surrogate models. Here, the membrane length is divided into several elements k and solved for the conservation equations on every element. An individual ANN describes each element. In Figure 5 A-B the strong influence of the feed concentration of the salt on the retention is evident for a selected membrane case.

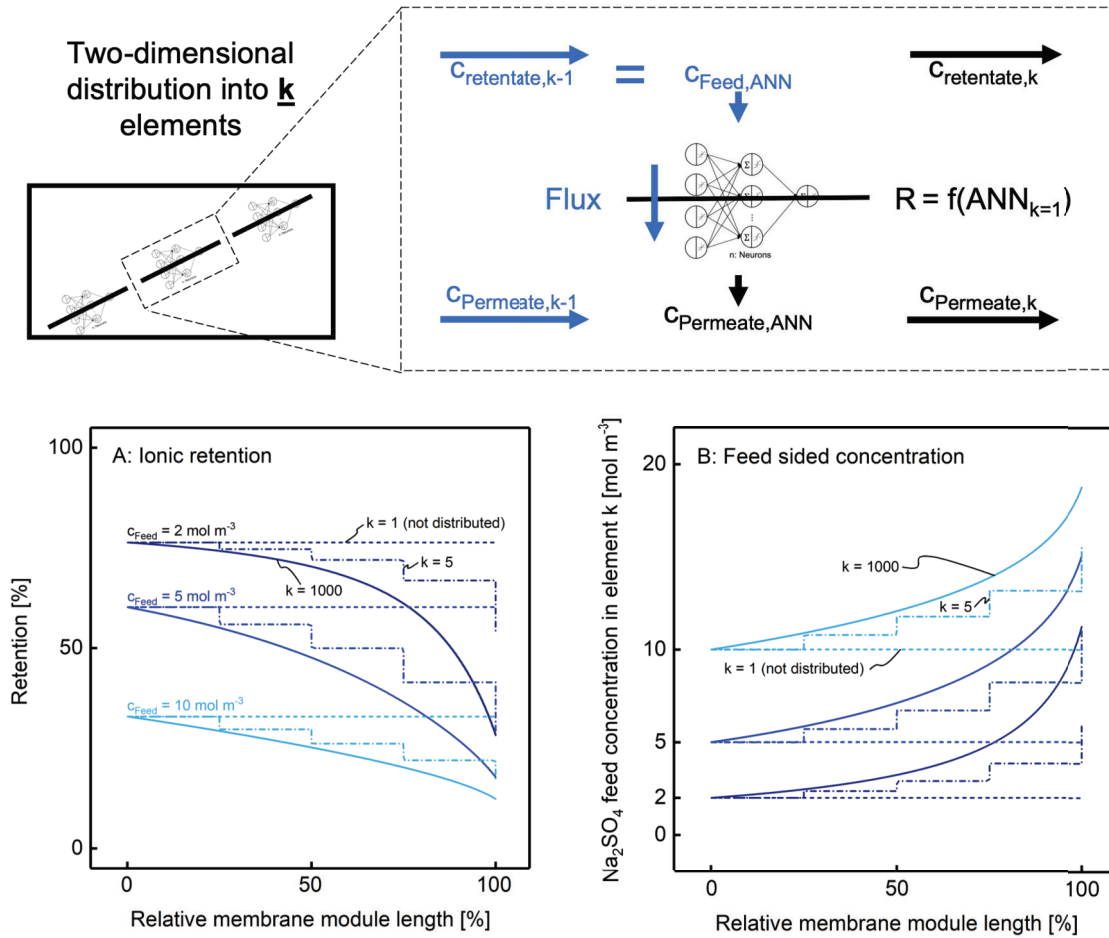


Figure 5. Two-dimensional distribution of the membrane module in flow direction through ANN-based surrogate models. Shown is the negative effect of the feed concentration on the salt retention of the membrane. Only the two-dimensional distributed cases do not underestimate this local reduction of ionic retention in the direction of flow. In A), the retention of one selected case and membrane are shown for three different initial feed concentrations and varying element numbers k . In b), the corresponding feed sided concentration is shown resulting from the filtration-related increase of feed concentration.

Two effects contribute to the reduction of salt retention. First, the retention decreases with increasing initial salt feed concentration. This is demonstrated in Figure 5 A-B for the concentrations of 2 mol m^{-3} , 5 mol m^{-3} , and 10 mol m^{-3} . Second, a filtration-related reduction of the retention of salt is observed in the direction of flow. During

filtration, the salt concentration increases at the retentate side. This negative effect of the increasing feed concentration on the salt retention of the membrane is differently pronounced for the shown three cases. A one-dimensional distributed membrane underestimates this effect and leads to inaccurate results. With an increasing number of elements k , the approximation approaches the most accurate concentration profile. The relative gap between a one-dimensional distributed membrane and a two-dimensional distributed membrane is calculated for the specific case in Table 2. Here, the relative gap ranges from +43% to +240% dependent on the initial feed concentration. These results are case-specific but apply to any other case.

Table 2. The Feed concentration increases along the membrane length during filtration. The relative gap of the permeate concentration is calculated based on a two-dimensional distributed membrane with $k = 1000$ elements and a one-dimensional distributed membrane with $k = 1$ element for different feed concentrations. The gap occurs due to the decreasing retention with increasing feed concentration which is not accounted for in a one-dimensional distributed membrane.

Variable name	Unit	$c_{feed} = 10 \text{ mol m}^{-3}$	$c_{feed} = 5 \text{ mol m}^{-3}$	$c_{feed} = 2 \text{ mol m}^{-3}$
Two-dimensional distributed ($k = 1000$)	$[\text{mol m}^{-3}]$	6.71	1.99	0.47
One-dimensional distributed ($k = 1$)	$[\text{mol m}^{-3}]$	9.62	4.58	1.61
Absolute difference	$[\text{mol m}^{-3}]$	2.92	2.53	1.14
Relative difference	$[\%]$	+43.50	+129.75	+239.86

Next, the most accurate number of elements for a single salt or salt mixture process optimization is identified. This is necessary to determine the number of two-dimensional distribution elements k that are necessary to describe the change of states along the membrane surface accurately. Therefore, the number of elements is increased until the relative gap of the concentration is below 1%, where it can be considered sufficiently accurate. The relative difference (or residuum) is calculated by the relative gap between the two-dimensional distributed ($k > 1$), and one-dimensional distributed ($k = 1$) permeate concentration leaving the membrane module. This procedure is repeated for different membrane configurations and observed covering the whole application range. Two exemplary cases and the progression of the residuum over the number of elements used are visualized in Figure 6. Here, a maximum number of $k = 56$ elements are needed for single salt applications, and $k = 51$ elements are required for salt mixtures.

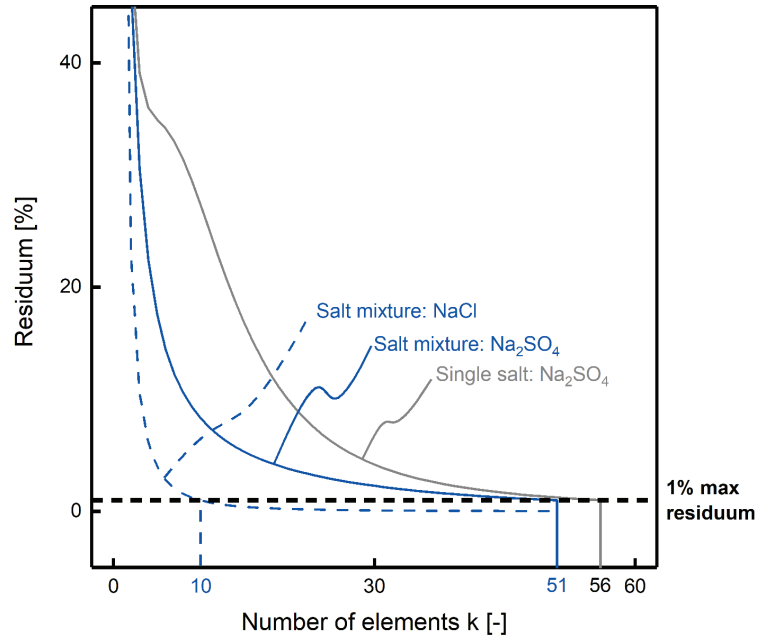


Figure 6. Two-dimensional distribution of the module in flow direction: The membrane unit is split into multiple elements k . The number of elements is increased until the relative gap of filtration performance through the permeate concentration is below 1%.

2.4. Training of the surrogate model with two-dimensional distribution of the membrane

Resolving the membranes through two-dimensional distribution is achieved by subdividing it into a fixed number of elements and arranging the individual ANNs in series. Then, the mass conservation laws and the ANN computations are employed on every element. This creates a local output state from the input for each element. The technique is recursively applied to all elements along the direction of flow.

Now, this repeated computation of the local element quantities can be directly embedded into the process of plant optimization. However, this nesting of ANNs can lead to weak McCormick relaxations and can adversely affect the computational performance of global optimization [81]. An alternative approach is to train a new ANN surrogate model that includes these recursions as a black box creating a single surrogate model for the two-dimensional modeling of the nanofiltration module. The training of an ANN-based surrogate model, including a two-dimensional distribution of the membrane module, is described in the following.

A special case of data generation and ANN training arises when creating the surrogate model, including a two-dimensional distribution. The one-dimensional surrogate models trained on the one-dimensional extended Nernst-Planck ion transport model are connected serially one after the other and evaluated offline. This series of surrogate models create an additional data set for the new two-dimensional surrogate model. This extended two-dimensional surrogate model contains additional influential variables as the volume flows of the entire module need to be considered. A partial resolution of the recursive equation (exemplary for the single salt case) results in the following description

$$c_{in}(k) = c_{in}(k-1) \cdot \left(1 + \frac{\gamma}{K - (k-1) \cdot \gamma} R(k-1) \right) \quad (1)$$

where $c_{in,j}$ denotes the feed concentration in element k , K the total number of elements, R_j the ionic retention, and γ the flow ratio which denotes the ratio of feed and permeate flow along with the nanofiltration unit. Where γ is defined as

$$\gamma = \frac{F_{\text{permeate}}}{F_{\text{feed}}}. \quad (2)$$

This flow ratio of the membrane module determines the performance of the unit. Therefore, the two-dimensional surrogate model is extended by the additional input of γ to account for the extra parameter.

For obtaining the training data of the two-dimensional surrogate model, an additional Latin hypercube sampling for a random two-dimensional distribution is applied, and the data created offline. Due to the low computational effort to calculate such recursive ANNs in series, $k = 1000$ elements is chosen for the following process optimizations. Therefore, the relative difference between the two-dimensional distributed ($k > 1$) and one-dimensional distributed ($k = 1$) permeate concentration leaving the membrane module is always below 1%. For example, in Figure 6 only a maximum number of $k = 56$ elements is needed.

The training procedure is performed as previously stated in Section 2.2. Now, inputs to the ANN are the layer charge X , the layer thickness Δx , the transmembrane velocity v , the salt feed concentration c_j , and the additional parameter γ for the flow ratio. The flow ratio γ is valid in the range $\in [0.18; 1] \subset \mathbb{R}$ under the consideration of the intercorrelation of γ (i.e., the permeate flow can never be higher than the feed flow). Output to the ANN is the ionic retention of ion j R_j . A list of properties of the new two-dimensional surrogate models can be found in Table 3.

Table 3. Two-dimensional surrogate models used in this work to recreate the two-dimensional distribution of the membrane in the flow direction. The number of neurons n of each ANN, the amount of training data set entries, and the section in which it is used is given.

Network name	Hidden neurons n	Description	Training points	Used in Section
ANN _{Na₂SO₄, $k=1000$, 1 Unit}	6	Single salt	998	3.3, 3.4
ANN _{Na₂SO₄, $k=1000$, 2 Units}	6	Single salt	1,198	3.4
ANN _{NaCl, Na₂SO₄, $k=1000$, 1 Unit}	14	Salt mixture	1,000	3.5

2.5. Hybrid mechanistic / data-driven process superstructure model

The individual surrogate models are integrated with a mechanistic process model to a hybrid mechanistic / data-driven process model. The hybrid process superstructure model comprises of a mechanistic process model to describe the process plant and of surrogate models to describe the ionic transport based on membrane and process-specific

variables. The mechanistic process model consists of component mass balances, a pump model, and cost correlations. The superstructure of the process plant is shown in Figure 7. The model is adapted from our previous work [24].

Either a single-stage nanofiltration unit or a two-stage connection of nanofiltration units is chosen for the process plant. A pump is installed in the inlet stream of each nanofiltration unit to provide the main driving force. Each nanofiltration unit consists of multiple membrane modules, N_{module} in parallel. Each nanofiltration unit (unit i) has a feed stream $F_{feed,i}$. A concentrated retentate stream $F_{retentate,i}$ and a diluted permeate stream $F_{permeate,i}$ each leave the nanofiltration unit. The permeate stream $F_{permeate,i}$ is calculated based on the membrane's ionic retention performance, permeability, and pressure difference Δp applied:

$$F_{permeate,i} = Q \cdot A_{unit,i} \cdot \Delta p \quad (3)$$

where Q is the permeability of the membrane unit, $A_{unit,i} = N_{module,i} \cdot A_{module}$ is the total membrane area of the membrane unit. The size of a single membrane module in a nanofiltration unit is set to $A_{module} = 60 \text{ m}^2$ in accordance with multiple membrane modules available on the market [82, 83]. The permeability of the membrane is estimated based on a linear correlation of the membrane thickness Δx with the range $Q \in [5; 35] \subset \mathbb{R}$ in accordance with laboratory results obtained in our previous work [20].

$$Q = 0.4 \cdot \Delta x - 25 \quad (4)$$

Additionally, the layer charge X is assumed to increase linearly with membrane thickness Δx . A mass balance coupled with the ionic retention of ion j $R_{i,j}$ for ion j closes the governing equations.

$$F_{feed,i} c_{feed,i,j} = F_{retentate,i} c_{retentate,i,j} + F_{permeate,i} c_{permeate,i,j} \quad (5)$$

$$R_{i,j} = 1 - \frac{c_{permeate,i,j}}{c_{feed,i,j}} \quad (6)$$

The ionic retention of ion j $R_{i,j}$ is computed by the ANNs as a function of membrane-specific variables and process-specific variables as described in Section 2.2 and Section 2.4.

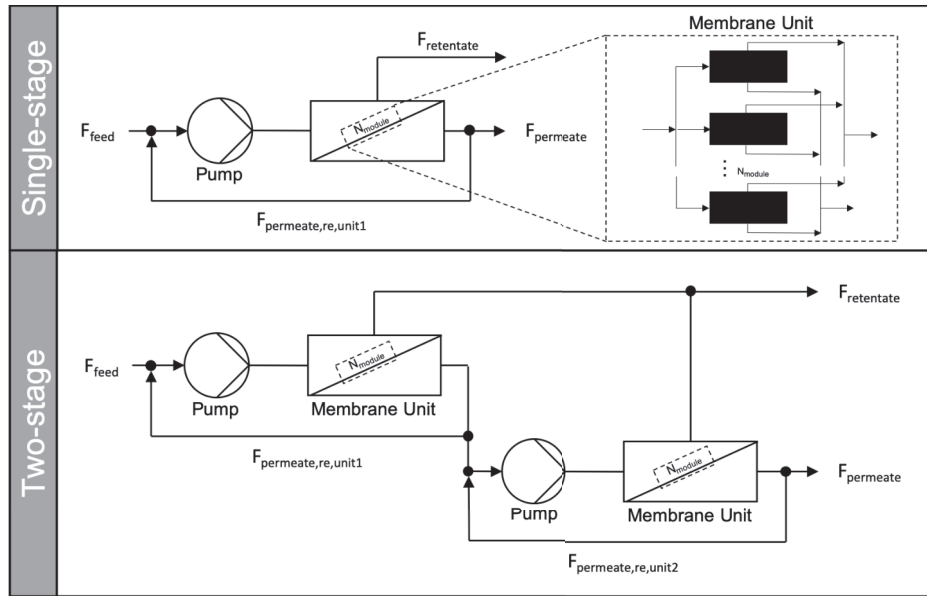


Figure 7. Layouts of the membrane process plant consisting of either a single nanofiltration unit or a two-stage connection of nanofiltration units. Permeate flow recirculation is allowed but does not have to be used. Each of the membrane units is equipped with a feed sided pump. A global feed, and retentate and permeate are connected by mixing units and represent the process input and outputs

The cost model for this study is adapted from our previous work [24]. Here, the Verberne cost model [84–86] is used. The total investment cost $C_{investment}$ is calculated based on the sum of the civil C_{civil} , mechanical $C_{mechanical}$, electromechanical $C_{electro}$, and membrane $C_{membrane}$ investment cost

$$C_{investment} = C_{civil} + C_{mechanical} + C_{electro} + C_{membrane}. \quad (7)$$

The annual operation cost is calculated based on the sum of the depreciation cost $C_{depreciation}$, the energy cost C_{energy} , the maintenance cost $C_{maintenance}$, and specific cost $C_{specific}$.

$$C_{operation} = C_{depreciation} + C_{energy} + C_{maintenance} + C_{specific} \quad (8)$$

Appropriate cost parameters and depreciation periods are accounted for by the individual investment cost [24].

This study does not consider the effects of osmotic pressure differences on the reduction of the driving force (here transmembrane pressure). This reduction of the driving force means that the presented results underestimate the pumping power because of the high salt concentrations used in this study. Considering the previously described cost model, it might be more economically viable to implement more membrane modules per membrane unit than using a higher pressurized feed stream. There always exists a trade-off between the costs by increased pumping power and the capital costs of more membrane area. Thus, the integration of the osmotic pressure into the proposed framework would be relevant and important for future research. In particular, future work could consider osmotic pressure differences in the ion transport models by accounting for the osmotic pressure in the two-dimensional model

used for data generation (c.f. Section 2.3-2.4). However, this is beyond the scope of the current study. Considering the effects of osmotic pressure differences on the reduction of the driving force demand for a more detailed investigation as well as validation of the model against existing methods.

2.6. Numerical optimization approach

Finally, the hybrid model is optimized using a reduced-space formulation and our open-source deterministic global solver MAiNGO [73]. The optimization problems in this work are formulated as a multi-objective mixed-integer nonlinear program (MINLP):

$$\begin{aligned}
 \min_{\mathbf{x}, \mathbf{y}} \quad & \begin{pmatrix} f_1(\mathbf{x}, \mathbf{y}) \\ f_2(\mathbf{x}, \mathbf{y}) \end{pmatrix} \\
 \text{s.t.} \quad & g_i(\mathbf{x}, \mathbf{y}) = 0, \quad i = 1, \dots, I \\
 & h_j(\mathbf{x}, \mathbf{y}) \leq 0, \quad j = 1, \dots, J \\
 & \mathbf{x} \in X \subset \mathbb{R}^n, \quad \mathbf{y} \in Y \subset \mathbb{Z}^m
 \end{aligned} \tag{9}$$

where $g_i(\mathbf{x}, \mathbf{y})$ are equality constraints, $h_j(\mathbf{x}, \mathbf{y})$ are inequality constraints, $f_1(\mathbf{x}, \mathbf{y})$, $f_2(\mathbf{x}, \mathbf{y})$ are the objective functions, $\mathbf{x}$ are continuous optimization variables, and $\mathbf{y}$ is the integer optimization variable (i.e., number of membrane modules $N_{module,i}$). All model equations for the mechanistic process model are described in Subsection 2.5. The ionic separation performance is mapped by ANNs as described in Subsection 2.1, which are formulated in a reduced-space formulation [59]. The objective functions are (i) minimal annual operation costs ($f_1(\mathbf{x}, \mathbf{y}) = C_{operation}$) and (ii) minimal permeate concentration ($f_2(\mathbf{x}, \mathbf{y}) = c_{permeate}$).

The MINLP is solved for the two objectives by the ϵ -constraint method. Here, the multi-objective problem is reformulated to multiple single-objective problems. Thereby, one objective is minimized, and the other objective is enforced to be less or equal to a parameter ϵ . This procedure is repeated for different ϵ yielding Pareto-optimal points. All optimization problems in this work are implemented and solved by our open-source deterministic global solver MAiNGO [73].

3. Case study & results

In this section, the results of the process optimizations, including the ANN-based surrogate models, are discussed for a specific case scenario (cf. Section 3.1). In Section 3.2 the methodology to integrate multi-scale membrane plant optimization using LbL nanofiltration membranes is presented. Then, the performance of the overall process is evaluated when including a more accurate two-dimensional surrogate model by an optimal two-dimensional distribution of the membrane in Section 3.3. Here, the single filtration unit is considered by (A) using a one-dimensional distribution, (B) (1) by using a three-element two-dimensional distribution which is directly implemented in the optimization framework, and (B) (2) by using a new surrogate model that includes these two-dimensional distributions over $k = 1000$ cells in a black box creating a computational resource-efficient two-dimensional modeling of the nanofiltration module. In Section 3.4, the filtration performance of the process is evaluated when including a second stage. Finally, in Section 3.5, salt mixtures are considered and the process optimized.

3.1. Case scenario for membrane plant optimization

The case scenario and cost correlations are adapted from our previous work [24]. The membrane plant is optimized based on a drinking water purification meeting given quality specifications for a small town with 10,000 inhabitants [87]. In this study, water softening is considered by retaining Na_2SO_4 salt in the membrane plant. The drinking water demand amounts to a peak demand of approximately $224 \text{ m}^3 \text{ h}^{-1}$. All processes dimensions are set to meet the permeate volume flow of this peak demand. Demand-side management is not considered within the scope of this work.

3.2. Bridging the gap of multiple scales by surrogate model-based membrane plant optimization

In this section, the results of the surrogate model-based membrane plant optimization are presented. The optimization is performed for a single-stage nanofiltration unit (cf., Figure 7) using the $\text{ANN}_{\text{Na}_2\text{SO}_4, k=1, 1 \text{ Unit}}$ surrogate model (cf., Table 1) to describe the ionic retention of the membrane unit. The whole process is optimized for ranging feed concentrations of Na_2SO_4 and solved for the two objectives – (i) minimal annual operation costs and (ii) minimal Na_2SO_4 permeate concentration. The Na_2SO_4 feed concentration range from 5 mol m^{-3} to 20 mol m^{-3} . The results of this optimization are shown in Figure 8 depicting a smooth, convex Pareto front.

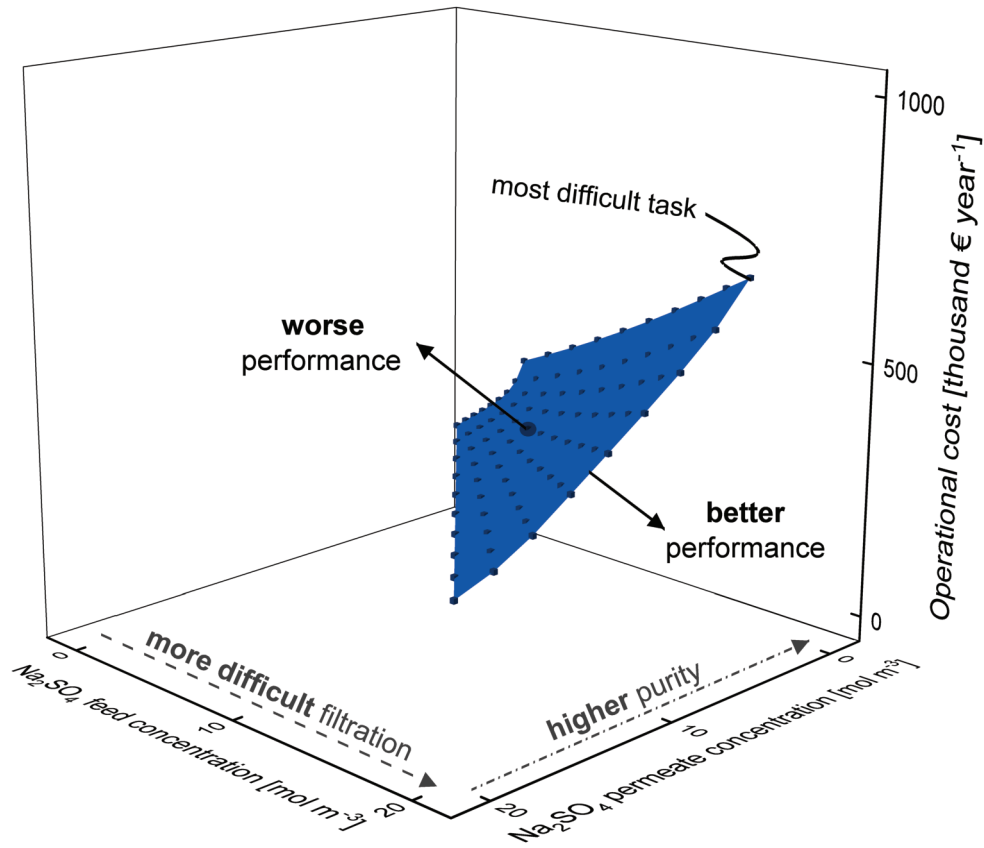


Figure 8. Surrogate model-based membrane plant optimization for a single-stage nanofiltration unit using the $ANN_{Na_2SO_4, k=1, 1 \text{ Unit}}$ surrogate model to describe the ionic retention of the membrane unit for Na_2SO_4 feed concentration ranging from 5 mol m^{-3} to 20 mol m^{-3} . The optimization results in a smooth, convex Pareto front. The annual operating costs increase along with higher purity and with the more difficult filtration task.

Ideally, an optimal membrane plant would operate at zero costs and zero permeate concentrations for all feed concentrations. This is called the utopia point. However, in reality, higher feed concentrations (i.e., more difficult filtration) and lower permeate concentrations (i.e., higher purities) always lead to higher operational costs. Hence, we face an inherent trade-off between conflicting objectives. This trade-off can be addressed by multi-objective optimization. The solution of a multi-objective optimization problem is a Pareto-front between the objective functions. The points on the Pareto front correspond to processes where neither one of the objectives can be improved anymore without worsening another objective. Thus, for each of the Pareto efficient process points, one cannot archive higher purity at a given feed concentration without spending the additional operational cost.

Considering the calculated Pareto front in Figure 8, indeed the annual operating costs increase along with higher purity (i.e., lower Na_2SO_4 permeate concentration) and more difficult filtration task (i.e., higher Na_2SO_4 feed concentration). The achievable permeate purity strongly depends on the process' salt feed concentration. Here, the maximum process purity ranges from 0.63 mol m^{-3} for a feed concentration of 5 mol m^{-3} to 3.6 mol m^{-3}

for a feed concentration of 20 mol m^{-3} . The latter is the most difficult separation task of this optimization resulting in an annual operation cost of approximately 700 thousand € year⁻¹. A table of the optimal solution points can be found in the supplementary data to this publication.

In the presented work, we demonstrate that nano-scale ion transport models can be used in process optimization. Besides, some essential effects for the design of membrane processes result from the observation of the simulated data. Therefore, the qualitative differences between the process configurations are assessed in the following. The tables to this discussion can be found in the supplementary data to this publication. The optimal process configuration towards higher purity (and higher annual operation cost) follows a stringent pattern regarding the involved process parameters. The process recirculation (cf., Figure 7 single stage) is fully exploited to minimize the filtration units feed concentration. Permeate flow recirculation is used to keep the feed concentration low to maintain the membrane's retention high. This process indeed seems to be non-intuitive because the purified permeate stream is recirculated back to the feed stream. However, the feed concentration strongly influences the retention of the nanofiltration membrane and must, therefore, remain low to sustain high ionic retention of the membrane. It becomes apparent that it is economically more viable to recirculate the permeate stream than investing in more membrane area or higher membrane charges associated with low permeability. Moreover, the membrane thickness and surface charge are also increased throughout the optimization to yield a similar, but quantitatively less significant, effect. The increase of the membrane thickness leads to a decreased permeability, which has to be compensated by a larger membrane surface or higher pressure gradient. This can be observed from a large number of membrane modules utilized for high-purity process configurations. Additionally, a large membrane surface facilitates a suitable mass transfer along with a medium transmembrane velocity, which also has a significant impact on retention.

For processes with lower cost and purity, the membrane's surface is drastically reduced to eliminate this cost factor. In particular, no recirculating is performed to minimize the operation costs for the involved feed-sided pressure pump. The preceding observations apply to all following discussed process configurations. For the filtration of salt mixtures, small restrictions have to be made that shall be considered later.

Furthermore, the shape and position of the calculated Pareto front can be used to determine the performance of the membrane system. In Figure 8, a curved Pareto front towards the utopia point indicates overall a better performance. Whereas a curved Pareto away from the utopia point means a worse performance of the process up to the angular end of the feed- and permeate concentration plane so that no filtration takes place. This visual effect may be used when comparing different membrane systems embedded in the same process of plant optimization. Taking a closer look at the next Section 3.3, the shape and position of the calculated Pareto front indicates the severe underestimation of the costs when only including a one-dimensional distribution of the membrane.

The optimization problem includes 6 optimization variables, 1 equality constraint, and 9 inequality constraints. The optimization was executed on 4 cores in parallel on a high performance computing cluster. The average CPU time summed over all cores for the optimization of a Pareto point was 2.1 CPU seconds (see Table 4). Thus, the computation of the Pareto front, which is approximated by 100 points, took roughly 35 CPU minutes in total.

3.3. Optimization including the two-dimensional distribution of the membrane

In the next step, the performance of the overall process is evaluated when including a more accurate two-dimensional surrogate model by an optimal two-dimensional distribution of the membrane. A membrane module is considered by (A) using a one-dimensional distribution as presented in the previous Section 3.2), and compared to (B) (1) using a three-element two-dimensional distribution by three $\text{ANN}_{\text{Na}_2\text{SO}_4, k=1, 1 \text{ Unit}}$ surrogate models (cf., Table 1) which are directly implemented in the optimization framework, and (B) (2) by using a new surrogate model $\text{ANN}_{\text{Na}_2\text{SO}_4, k=1000, 1 \text{ Unit}}$ (cf., Table 3) that includes these two-dimensional distributions as a black box creating a computational resource-efficient two-dimensional modeling of the nanofiltration module.

Again, one single-stage membrane unit (c.f., Figure 7) is considered for the optimization of a process plant for ranging feed concentrations of Na_2SO_4 and solved for the two objectives - (i) minimal annual operation costs and (ii) minimal Na_2SO_4 permeate concentration. The resulting Pareto fronts of the optimizations are shown in Figure 9. Here, for comparability, the results of Figure 8 are displayed in Figure 9 (Membrane not distributed, i.e. $k = 1$) as well. A table of the optimal solution points can be found in the supplementary data to this publication.

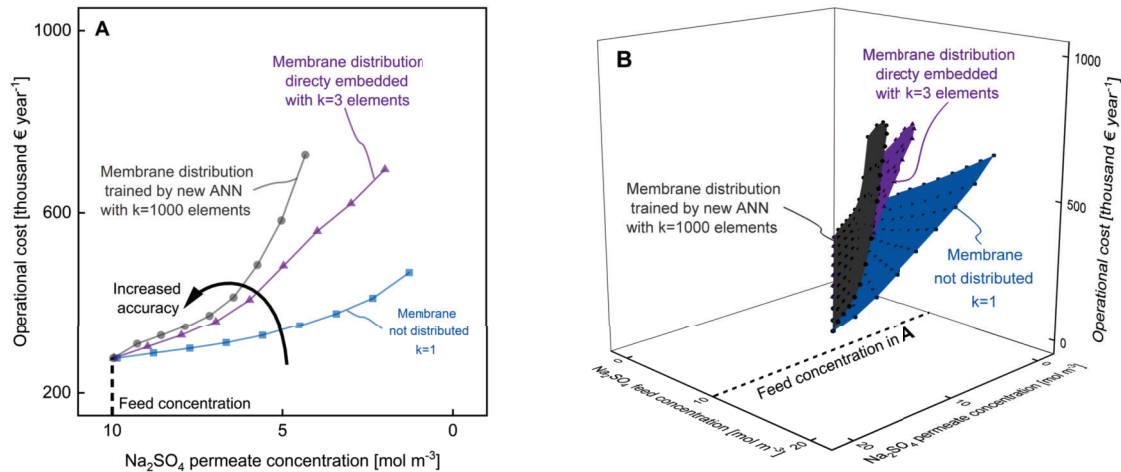


Figure 9. Surrogate model-based membrane plant optimization for a single-stage nanofiltration for Na_2SO_4 filtration. Two implementations of the two-dimensional distribution are presented: (B) (1) by direct implementation of the two-dimensional distribution by three $\text{ANN}_{\text{Na}_2\text{SO}_4, k=1, 1 \text{ Unit}}$ surrogate models (B) (2) by the recursive back box surrogate model approach using $\text{ANN}_{\text{Na}_2\text{SO}_4, k=1000, 1 \text{ Unit}}$. For comparability, the results of Figure 8 are displayed using one $\text{ANN}_{\text{Na}_2\text{SO}_4, k=1, 1 \text{ Unit}}$ surrogate model with one-dimensional distribution. Showing A) the results for a feed concentration of 10 mol m^{-3} and B) for ranging feed concentrations between 5 mol m^{-3} and 20 mol m^{-3} .

(B) (1) Directly embedding the two-dimensional distribution in the optimization framework results in a dramatic increase in run time for only a few two-dimensional distribution elements. As shown in Table 4, the average CPU time to solve one Pareto point using $k = 3$ elements is $9.1 \cdot 10^4$ CPU seconds corresponding to over 25 CPU hours. To reduce the wall-clock time, we solved this optimization parallel on 48 cores. (Note that the CPU times are summed over all cores.) Thus, the maximum number of elements for directly embedding the two-dimensional distribution in

the optimization framework is $k = 3$ elements using a single nanofiltration unit. This low number of elements is not sufficient to accurately describe the changes along the membrane length, as presented in Section 2.3. The results suggest clearly that the solution of a large number of elements, i.e., $k > 3$, would not be feasible using this method.

(B) (2) In contrast, our proposed recursive black-box approach using the $\text{ANN}_{\text{Na}_2\text{SO}_4, k=1000, 1 \text{ Unit}}$ surrogate model requires on average $2.4 \cdot 10^2$ CPU seconds to solve a Pareto point (summed over all cores). Thus, the complete Pareto front, including 100 points, is solved within 7 CPU hours. As we use 48 cores in parallel, this corresponds to a wall-clock time of under 10 minutes. Thus, the method significantly reduced the computational effort compared to (B) (1) and makes the integration of $k = 1000$ elements feasible. Compared to the consideration of $k = 1$, we can observe an acceptable increase in CPU time, which is mainly due to the additional input of the ANN describing the flow ratio γ .

Table 4. Computational performance comparison of the proposed optimization methods. For each method, 100 optimization problems are solved, i.e., one for each Pareto point. Not all solved points are displayed in the figure. We provide average, variance, minimal, and maximal CPU times overall optimization problems. Note that CPU times for optimization include preprocessing and branch-and-bound time summed over all parallel cores in seconds.

#	Name	Section	k [-]	CPU Cores	CPU time summed over all cores [s]			
					Average	Variance	min	max
A	$\text{Na}_2\text{SO}_4, k = 1, 1 \text{ Unit}$	3.2	1	4	$2.1 \cdot 10^1$	$2.4 \cdot 10^2$	$7.7 \cdot 10^{-1}$	$5.6 \cdot 10^1$
B (1)	$\text{Na}_2\text{SO}_4, k = 3, 1 \text{ Unit}$	3.3	3	48	$9.1 \cdot 10^4$	$2.5 \cdot 10^{10}$	$2.3 \cdot 10^0$	$7.7 \cdot 10^5$
B (2)	$\text{Na}_2\text{SO}_4, k = 1000, 1 \text{ Unit}$	3.3	1000	48	$2.4 \cdot 10^2$	$1.1 \cdot 10^5$	$1.7 \cdot 10^0$	$1.5 \cdot 10^3$

A severe underestimation of the costs is observed in Figure 9 A when not including a two-dimensional distribution of the membrane. This severe underestimation of the costs is prevailing for all feed concentration of Na_2SO_4 from 5 mol m^{-3} to 20 mol m^{-3} as indicated by the location of the three Pareto surfaces shown in Figure 9 B. The use of a more precise two-dimensional distribution of the membrane reveals a worse performance than the idealized process only with one-dimensional distribution in Figure 9. The worsening in process performance can be explained by the filtration-related decreasing retention of salt at the retentate side. For positive retention (which is always the case for the single salt filtration examined here), the concentration on the retentate side of the filtration unit steadily increases along the membrane surface. However, this leads to decreasing retention and, therefore, a worse filtration performance. Additionally, a back-coupling through recirculation of the permeate feed stream back to the membrane's unit feed stream is observed. For the two-dimensional distribution of the membrane with $k = 1000$ elements, for instance, the maximum achievable performance for a feed concentration of 5 mol m^{-3} is 1.19 mol m^{-3} and therefore presents an increase in impurity as compared to the idealized case using the one-dimensional distribution. This effect is even more pronounced when considering higher feed concentrations.

3.4. Extension of the optimization process by a second filtration unit

Next, the two-dimensional surrogate model approach is utilized to extend the optimization process by a second filtration unit (cf., Figure 7 two stage). The $\text{ANN}_{\text{Na}_2\text{SO}_4, k=1000, 1 \text{ Unit}}$ surrogate model is taken from the previous

Section 3.3. For the two-stage process the $ANN_{Na_2SO_4, k=1000, 2 \text{ Units}}$ surrogate model is chosen (cf., Table 3). The whole process is optimized for ranging feed concentrations of Na_2SO_4 from 5 mol m^{-3} to 20 mol m^{-3} and solved again for the two objectives - (i) minimal annual operation costs and (ii) minimal Na_2SO_4 permeate concentration. The results of the optimization are shown in Figure 10. A table of the optimal solution points can be found in the supplementary data to this publication.

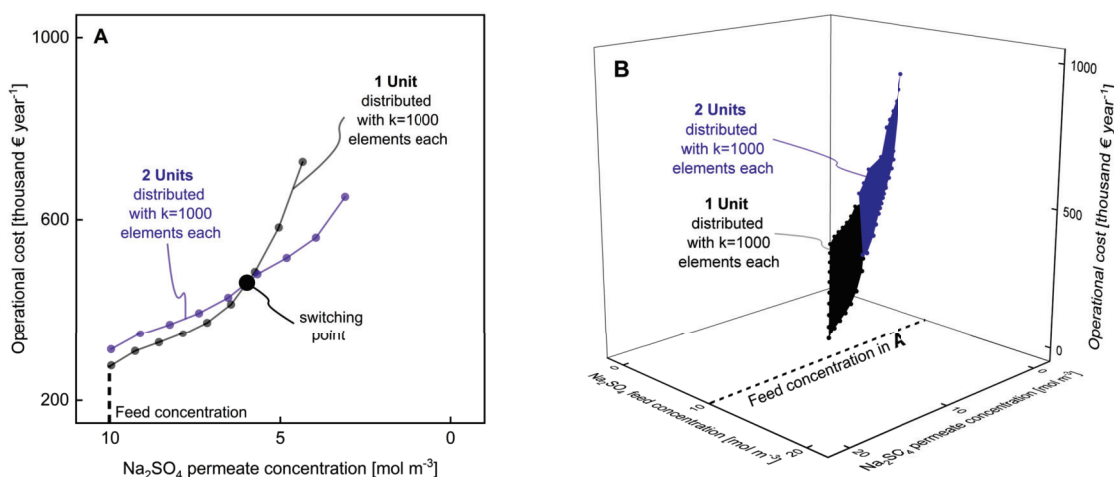


Figure 10. Surrogate model-based membrane plant optimization for a single-stage and two-stage serial connection of nanofiltration units for Na_2SO_4 filtration. The two-dimensional surrogate model approach is chosen using $ANN_{Na_2SO_4, k=1000, 1 \text{ Unit}}$ for one nanofiltration unit, and the $ANN_{Na_2SO_4, k=1000, 2 \text{ Units}}$ surrogate model for the two-stage serial connection. Showing A) the results for a feed concentration of 10 mol m^{-3} and B) for ranging feed concentrations between 5 mol m^{-3} and 20 mol m^{-3} .

This combination of a two-stage series of filtration units results in a strongly enhanced filtration performance. This tendency is particularly intensified as the feed concentration of the second unit is already the purified concentration of the first unit. Due to the strong correlation of the feed concentration and retention, the second filtration unit is particularly useful. Notably, the improvement is much more significant for lower feed concentrations. For higher initial feed concentrations, however, the effect of utilizing two filtration units is not as predominant.

There exists a distinct switching point, as highlighted in Figure 10 A, where the economic viability of either one or two nanofiltration unit changes. In the range of less difficult filtration operations, a single-stage option is preferable. More difficult filtration operations with higher purity, however, are cheaper under the inclusion of a second unit (or even impossible for a single unit). The costs of a second filtration unit fall short in the overall economic estimation. This distinct switching point is case-specific and dependent on the feed concentration of Na_2SO_4 . As shown in Figure 10 B for all feed concentration of Na_2SO_4 from 5 mol m^{-3} to 20 mol m^{-3} a distinct switching point exists highlighted by the change of color (here from black to purple).

The superstructure optimization includes 12 continuous and 2 integer optimization variables, 6 equality constraints, and 14 inequality constraints. We run the optimization for the two-unit superstructure with $k = 1$ elements

per unit on 48 cores (result not displayed in Figure 10). Here, the average CPU time summed over all cores for one Pareto point was $7.7 \cdot 10^3$ seconds. This is an increase by a factor of 360 compared to the one unit process with a $k = 1$ element. Due to the increased computational effort, the results with two units and $k = 1000$ have been solved using a local optimization approach with 10 multistarts. This method cannot guarantee global optimality. Therefore, only the results with two units and $k = 1$ can guarantee global optimality, but are not meaningful as they severely underestimate the overall process costs. The local optimization took on average 25 CPU seconds for a Pareto point.

3.5. Extension of the optimization process by salt mixtures

Next, the two-dimensional surrogate model approach is extended by salt mixtures. Due to computational effort, only one single-stage is considered in the optimization (cf., Figure 7). The $\text{ANN}_{\text{NaCl}, \text{Na}_2\text{SO}_4, k=1000, 1 \text{ Unit}}$ surrogate model is utilized in the process optimization framework. The whole process is optimized for ranging feed concentrations of Na_2SO_4 from 5 mol m^{-3} to 20 mol m^{-3} . The feed concentration is extended by the addition of NaCl by fixed composition ratios. More precise the ratios for $\text{NaCl}:\text{Na}_2\text{SO}_4$ are 1:2, 1:1, and 2:1. For example in Figure 11 A, for the ratio $\text{NaCl}:\text{Na}_2\text{SO}_4 = 1:2$, 5 mol m^{-3} of NaCl is added to the feed stream when considering a feed of 10 mol m^{-3} Na_2SO_4 . The optimization is again solved for the two objectives - (i) minimal annual operation costs and (ii) minimal Na_2SO_4 permeate concentration. The results of the optimization for the given ratios for $\text{NaCl}:\text{Na}_2\text{SO}_4$ are shown in Figure 11. A table of the optimal solution points can be found in the supplementary data to this publication.

To reduce computational efforts, the optimization problem is solved for fewer operating conditions (Pareto points), leading to a more coarse approximation of the Pareto front. The shape and different process configurations show qualitatively similar behavior concerning the optimization of the single salt processes in Figure 9. In particular, the operational costs are in the same order of magnitude for both cases.

Taking a closer look at the different Pareto fronts for the individual salt mixtures, they reveal differences in filtration operation. On the one hand, the Pareto fronts for ratios 1:1 and 2:1 ($\text{NaCl}:\text{Na}_2\text{SO}_4$) in Figure 11 C-D present similar behavior to the previous results. On the other hand, the optimization points for a ratio of 1:2 in Figure 11 B are clustered in a much narrower region of feasible points in the Pareto front. This process infeasibility is mainly observed for higher feed concentrations. This behavior can be ascribed to the more complex relations between the membrane retention, and its influencing process-specific parameters. These relations are neither monotone (as for the single salt case), nor is the retention itself limited to positive values (the Donnan effect induces negative retention). Most notably filtration operation with high retention is still possible in Figure 11 D even at very high salt concentrations (i.e., 40 mol m^{-3} of Na_2SO_4 and 20 mol m^{-3} of Na_2SO_4) by considering the Donnan effect. Hence an accurate description of the underlying physical processes becomes quite cumbersome, making the surrogate-based model for the determination of optimal operation points even more valuable. The increased complexity of the process for salt mixtures also manifests in clustered points on the Pareto front. Such points are Pareto-dominated by preceding configurations, which means a process in that particular setting gets more expensive under looser purity constraints.

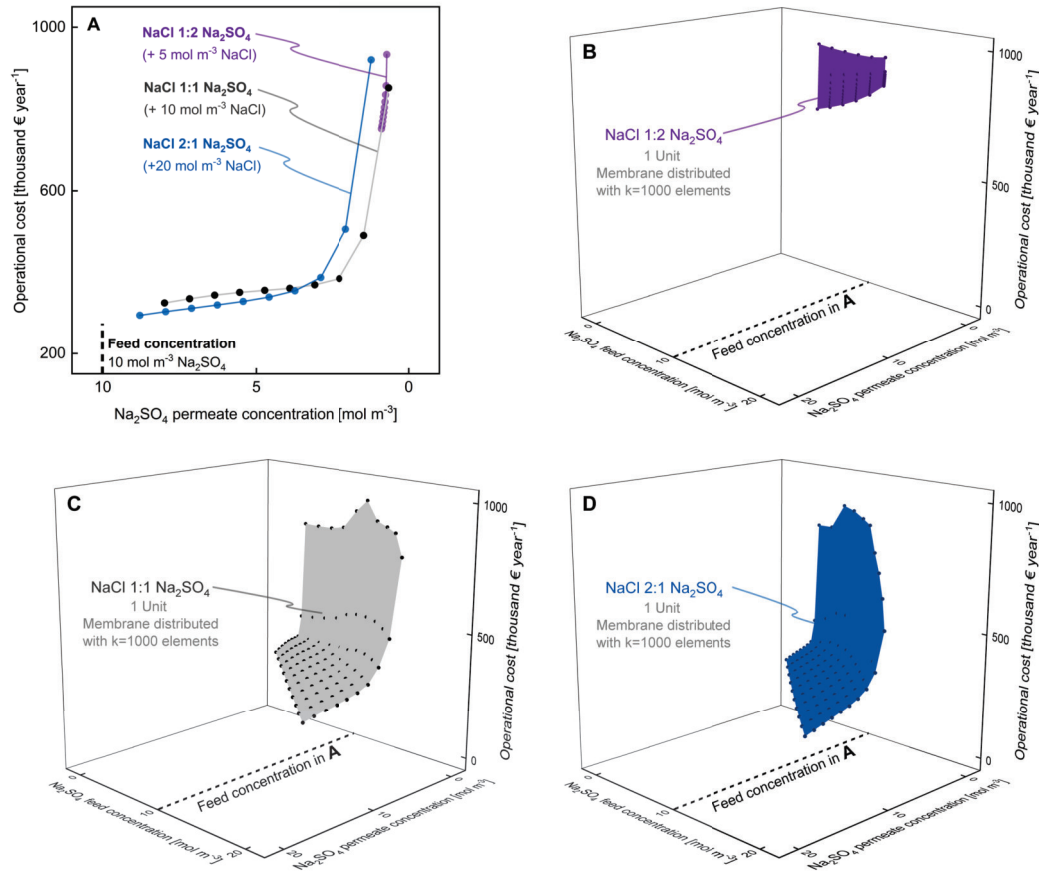


Figure 11. Surrogate model-based membrane plant optimization for a single-stage nanofiltration unit for filtration of NaCl and Na₂SO₄ salt mixture. The two-dimensional surrogate model approach is chosen using the ANN_{NaCl, Na₂SO₄, k=1000, 1 Unit} surrogate model. Showing A) the results for a Na₂SO₄ feed concentration of 10 mol m⁻³ and for ranging feed concentrations between 5 mol m⁻³ and 20 mol m⁻³ for salt mixture ratios (NaCl:Na₂SO₄) of B) 1:2, C) 1:1, and D) 2:1.

The optimization for salt mixtures includes 6 continuous and 1 integer optimization variables, 2 equality constraints, and 13 inequality constraints. We run the optimization for one single-stage unit and $k = 1000$ elements on 48 cores. The average CPU time summed over all cores for one Pareto point was $3.3 \cdot 10^3$, $1.1 \cdot 10^5$, and $1.0 \cdot 10^4$ seconds for 1:1, 1:2, and 2:1 salt mixtures (NaCl:Na₂SO₄), respectively. These results demonstrate that the CPU time for optimization is not only dependent on the structure of the optimization problem but also the case study.

4. Conclusion

We present a computational resource-efficient methodology to integrate multi-scale modeling in membrane science. We propose a machine learning approach to integrate accurate physical models of ion transport, valid on the nano-scale, into large-scale superstructure optimization of membrane processes. Thereby, the multi-scale decision-making process allows us to simultaneously design the membrane synthesis properties along with the process design and operation. All models and the solver itself are open-source and freely available, enabling multi-scale modeling in membrane science.

In particular, we train ANN surrogate models on simulated data from a differential-algebraic one-dimensional extended Nernst-Planck ion transport model, called pEnPEN. These surrogate models are then exploited to extend the framework towards a more accurate two-dimensional distribution of the membrane module to capture the filtration-related decreasing retention of salt. Here, individual ANNs are arranged in series to resolve the membrane in the direction of flow. Next, the two-dimensional surrogate models are embedded in a mechanistic membrane process model leading to a hybrid mechanistic / data-driven model. The results demonstrate that:

- A consideration of a one-dimensional distribution of the membrane module leads to a severe underestimation of process costs. Only the extension towards a two-dimensional distribution captures the adverse effects of filtration-related decreasing retention of salt.
- Non-intuitive interconnections of the membrane plant layout are optimal for high permeate purities, in which the permeate flow recirculates back to keep the feed concentration low.
- A distinct point between a single-stage and a two-stage process exists, where the economic viability of either one or two nanofiltration unit changes. Costs of a second filtration unit fall short in the overall economic estimation for more difficult filtration operations.
- For salt mixtures, it is much more complex to find optimal operating points, as further effects (i.e., Donnan exclusion) have an additional impact on the separation process. Therefore, it is all the more important that surrogate-based modeling bridges the gap between the scales for application of high fidelity ion transport models in superstructure optimization of membrane processes.

This work stimulates to exploit this open-source optimization framework by further research and in industrial applications, such as in drinking water treatment and water disposal. Additional surrogate models may include other membrane-based platforms. Furthermore, the concept of surrogate-based modeling should be exploited to integrate driving force diminishing effects acting on nano-scale, such as the osmotic pressure difference. Future work in deterministic global optimization and model reduction is desirable to push the computational limits of this approach further.

Acknowledgement

D.R., E.E., and M.W. acknowledge the support through the German Federal Ministry of Education and Research (BMBF) under the project UO-Rohrfabrikation (03XP0100E) and the project EffluenNF (02WIL1486) and the European Union's Horizon 2020 research and innovation program (grant agreement no. 694946). A.M.S. and A.M. gratefully acknowledge funding by the excellence initiative of the German federal and state governments and they acknowledge the financial support of the Kopernikus project SynErgie by the Federal Ministry of Education and Research (BMBF) and the project supervision by the project management organization Projektträger Jülich (PtJ). Funded by the Excellence Initiative of the German federal and state governments.

References

- [1] T. A. Larsen, S. Hoffmann, C. Lüthi, B. Truffer, M. Maurer, Emerging solutions to the water challenges of an urbanizing world, *Science* 352 (6288) (2016) 928–933.
- [2] S. Eggimann, L. Mutzner, O. Wani, M. Y. Schneider, D. Spuhler, M. Moy de Vitry, P. Beutler, M. Maurer, The potential of knowing more: A review of data-driven urban water management, *Environmental Science & Technology* 51 (5) (2017) 2538–2553.
- [3] S. Al Aani, T. Bonny, S. W. Hasan, N. Hilal, Can machine language and artificial intelligence revolutionize process automation for water treatment and desalination?, *Desalination* 458 (2019) 84–96.
- [4] M. Bagheri, A. Akbari, S. A. Mirbagheri, Advanced control of membrane fouling in filtration systems using artificial intelligence and machine learning techniques: A critical review, *Process Safety and Environmental Protection* 123 (2019) 229–252.
- [5] S. P. Nunes, P. Z. Culfaz-Emecen, G. Z. Ramon, T. Visser, G. H. Koops, W. Jin, M. Ulbricht, Thinking the future of membranes: Perspectives for advanced and new membrane materials and manufacturing processes, *Journal of Membrane Science* 598 (2020) 117761.
- [6] N. Ghaffour, T. M. Missimer, G. L. Amy, Technical review and evaluation of the economics of water desalination: current and future challenges for better water supply sustainability, *Desalination* 309 (2013) 197–207.
- [7] C. Abels, F. Carstensen, M. Wessling, Membrane processes in biorefinery applications, *Journal of Membrane Science* 444 (2013) 285–317.
- [8] C. Niewersch, A. B. Bloch, S. Yüce, T. Melin, M. Wessling, Nanofiltration for the recovery of phosphorus - development of a mass transport model, *Desalination* 346 (2014) 70–78.
- [9] R. R. Nair, E. Protasova, S. Strand, T. Bilstad, Membrane performance analysis for smart water production for enhanced oil recovery in carbonate and sandstone reservoirs, *Energy & Fuels* 32 (4) (2018) 4988–4995.
- [10] J. R. Werber, C. O. Osuji, M. Elimelech, Materials for next-generation desalination and water purification membranes, *Nature Reviews Materials* 1 (5) (2016) 16018.
- [11] T. Luo, S. Abdu, M. Wessling, Selectivity of ion exchange membranes: A review, *Journal of Membrane Science* 555 (2018) 429–454.
- [12] M. A. Shannon, P. W. Bohn, M. Elimelech, J. G. Georgiadis, B. J. Marinas, A. M. Mayes, Science and technology for water purification in the coming decades, *Nature* 452 (7185) (2008) 301.
- [13] K. Remmen, B. Müller, J. Köser, M. Wessling, T. Wintgens, Phosphorus recovery in an acidic environment using layer-by-layer modified membranes, *Journal of Membrane Science* 582 (2019) 254–263.
- [14] W. Liu, S. Wijeratne, L. Yang, M. Bruening, Porous star-star polyelectrolyte multilayers for protein binding, *Polymer* 138 (2018) 267–274.
- [15] J. J. Harris, J. L. Stair, M. L. Bruening, Layered polyelectrolyte films as selective, ultrathin barriers for anion transport, *Chemistry of Materials* 12 (7) (2000) 1941–1946.
- [16] R. Malaisamy, M. L. Bruening, High-flux nanofiltration membranes prepared by adsorption of multilayer polyelectrolyte membranes on polymeric supports, *Langmuir* 21 (23) (2005) 10587–10592.

- [17] W. Cheng, C. Liu, T. Tong, R. Epsztein, M. Sun, R. Verduzco, J. Ma, M. Elimelech, Selective removal of divalent cations by polyelectrolyte multilayer nanofiltration membrane: role of polyelectrolyte charge, ion size, and ionic strength, *Journal of Membrane Science* 559 (2018) 98–106.
- [18] S. Ilyas, R. English, P. Aimar, J.-F. Lahitte, W. M. De Vos, Preparation of multifunctional hollow fiber nanofiltration membranes by dynamic assembly of weak polyelectrolyte multilayers, *Colloids and Surfaces A: Physicochemical and Engineering Aspects* 533 (2017) 286–295.
- [19] D. Menne, C. Üzümlü, A. Koppelman, J. E. Wong, C. van Foeken, F. Borre, L. Dähne, T. Laakso, A. Pihlajamäki, M. Wessling, Regenerable polymer/ceramic hybrid nanofiltration membrane based on polyelectrolyte assembly by layer-by-layer technique, *Journal of Membrane Science* 520 (2016) 924–932.
- [20] D. Menne, J. Kamp, J. E. Wong, M. Wessling, Precise tuning of salt retention of backwashable polyelectrolyte multilayer hollow fiber nanofiltration membranes, *Journal of Membrane Science* 499 (2016) 396–405.
- [21] D. Rall, D. Menne, A. M. Schweidtmann, J. Kamp, L. von Kolzenberg, A. Mitsos, M. Wessling, Rational design of ion separation membranes, *Journal of Membrane Science* 569 (2019) 209–219.
- [22] O. Labban, C. Liu, T. H. Chong, J. H. Lienhard, Relating transport modeling to nanofiltration membrane fabrication: navigating the permeability-selectivity trade-off in desalination pretreatment, *Journal of Membrane Science* 554 (2018) 26–38.
- [23] Y. I. Dirir, Y. Hanafi, A. Ghoufi, A. Szymczyk, Theoretical investigation of the ionic selectivity of polyelectrolyte multilayer membranes in nanofiltration, *Langmuir* 31 (1) (2014) 451–457.
- [24] D. Rall, A. M. Schweidtmann, B. M. Aumeier, J. Kamp, J. Karwe, K. Ostendorf, A. Mitsos, M. Wessling, Simultaneous rational design of ion separation membranes and processes, *Journal of Membrane Science* 600 (2020) 117860.
- [25] H. Lonsdale, U. Merten, R. Riley, Transport properties of cellulose acetate osmotic membranes, *Journal of Applied Polymer Science* 9 (4) (1965) 1341–1362.
- [26] R. Schlögl, Membrane permeation in systems far from equilibrium, *Berichte der Bunsengesellschaft für Physikalische Chemie* 70 (4) (1966) 400–414.
- [27] A. Yaroshchuk, M. L. Bruening, E. E. L. Bernal, Solution-diffusion–electro-migration model and its uses for analysis of nanofiltration, pressure-retarded osmosis and forward osmosis in multi-ionic solutions, *Journal of Membrane Science* 447 (2013) 463–476.
- [28] R. Femmer, M. Martí-Calatayud, M. Wessling, Mechanistic modeling of the dielectric impedance of layered membrane architectures, *Journal of Membrane Science* 520 (2016) 29–36.
- [29] W. R. Bowen, J. S. Welfoot, Modelling the performance of membrane nanofiltration - critical assessment and model development, *Chemical Engineering Science* 57 (7) (2002) 1121–1137.
- [30] Y. Jin, Surrogate-assisted evolutionary computation: Recent advances and future challenges, *Swarm and Evolutionary Computation* 1 (2) (2011) 61–70.
- [31] A. A. Lapkin, Chemical engineering science and green chemistry—the challenge of sustainability, *Handbook of Green Chemistry: Online* (2010) 1–16.
- [32] B. Chachuat, A. B. Singer, P. I. Barton, Global methods for dynamic optimization and mixed-integer dynamic optimization, *Industrial & Engineering Chemistry Research* 45 (25) (2006) 8373–8392.
- [33] A. B. Singer, P. I. Barton, Global optimization with nonlinear ordinary differential equations, *Journal of Global Optimization* 34 (2) (2006) 159–190.
- [34] C. Wesselhoeft, D. A. Ham, R. Misener, Algorithms for mixed-integer optimization constrained by partial differential equations, in: *Computer Aided Chemical Engineering*, Vol. 44, Elsevier, 2018, pp. 799–804.
- [35] S. Sager, M. Claeys, F. Messine, Efficient upper and lower bounds for global mixed-integer optimal control, *Journal of Global Optimization* 61 (4) (2015) 721–743.
- [36] B. Ohs, J. Lohaus, M. Wessling, Optimization of membrane based nitrogen removal from natural gas, *Journal of Membrane Science* 498 (2016) 291–301.
- [37] R. Zarca, A. Ortiz, D. Gorri, L. T. Biegler, I. Ortiz, Optimization of multistage olefin/paraffin membrane separation processes through rigorous

- modeling, *AIChE Journal* 65 (6).
- [38] S. Lee, M. Binns, J.-K. Kim, Automated process design and optimization of membrane-based CO₂ capture for a coal-based power plant, *Journal of Membrane Science* 563 (2018) 820–834.
 - [39] A. Böcking, V. Koleva, J. Wind, Y. Thiermeyer, S. Blumenschein, R. Goebel, M. Skiborowski, M. Wessling, Can the variance in membrane performance influence the design of organic solvent nanofiltration processes?, *Journal of Membrane Science* 575 (2019) 217–228.
 - [40] P. Mores, A. Arias, N. Scenna, J. Caballero, S. Mussati, M. Mussati, Membrane-based processes: optimization of hydrogen separation by minimization of power, membrane area, and cost, *Processes* 6 (11) (2018) 221.
 - [41] S. Alsayegh, J. Johnson, B. Ohs, J. Lohaus, M. Wessling, Systematic optimization of H₂ recovery from water splitting process using membranes and N₂ diluent, *International Journal of Hydrogen Energy* 42 (9) (2017) 6000–6011.
 - [42] A. Ghobeity, A. Mitsos, Optimal design and operation of desalination systems: new challenges and recent advances, *Current Opinion in Chemical Engineering* 6 (2014) 61–68.
 - [43] D. A. White, W. J. Arrighi, J. Kudo, S. E. Watts, Multiscale topology optimization using neural network surrogate models, *Computer Methods in Applied Mechanics and Engineering* 346 (2019) 1118–1135.
 - [44] M. von Stosch, R. Oliveira, J. Peres, S. F. de Azevedo, Hybrid semi-parametric modeling in process systems engineering: past, present and future, *Computers & Chemical Engineering* 60 (2014) 86–101.
 - [45] T. Zhou, Z. Song, K. Sundmacher, Big data creates new opportunities for materials research: a review on methods and applications of machine learning for materials design, *Engineering* 5 (6) (2019) 1017 – 1026.
 - [46] C. Tsay, M. Baldea, 110th anniversary: using data to bridge the time and length scales of process systems, *Industrial & Engineering Chemistry Research* 58 (36) (2019) 16696–16708.
 - [47] A. Prakash, S. Sandfeld, Chances and challenges in fusing data science with materials science: the working group 3D data science is headed by Prof. Dr. Stefan Sandfeld., *Practical Metallography* 55 (8) (2018) 493–514.
 - [48] B. Sanchez-Lengeling, A. Aspuru-Guzik, Inverse molecular design using machine learning: generative models for matter engineering, *Science* 361 (6400) (2018) 360–365.
 - [49] C. A. Henao, C. T. Maravelias, Surrogate-based superstructure optimization framework, *AIChE Journal* 57 (5) (2011) 1216–1232.
 - [50] J. F. Unger, C. Könke, Neural networks as material models within a multiscale approach, *Computers & Structures* 87 (19–20) (2009) 1177–1186.
 - [51] J. H. Lee, J. Shin, M. J. Realff, Machine learning: overview of the recent progresses and implications for the process systems engineering field, *Computers & Chemical Engineering* 114 (2018) 111–121.
 - [52] V. Venkatasubramanian, The promise of artificial intelligence in chemical engineering: is it here, finally, *AIChE J* 65 (2) (2019) 466–78.
 - [53] M. Mistry, D. Letsios, G. Krennrich, R. M. Lee, R. Misener, Mixed-integer convex nonlinear optimization with gradient-boosted trees embedded, *arXiv preprint arXiv:1803.00952*.
 - [54] F. Boukouvala, J. Li, X. Xiao, C. A. Floudas, Data-driven modeling and global optimization of industrial-scale petrochemical planning operations, in: 2016 American Control Conference (ACC), IEEE, 2016, pp. 3340–3345.
 - [55] A. Cozad, N. V. Sahinidis, D. C. Miller, Learning surrogate models for simulation-based optimization, *AIChE Journal* 60 (6) (2014) 2211–2227.
 - [56] I. Fahmi, S. Cremaschi, Process synthesis of biodiesel production plant using artificial neural networks as the surrogate models, *Computers & Chemical Engineering* 46 (2012) 105–123.
 - [57] E. A. Del Rio-Chanona, X. Cong, E. Bradford, D. Zhang, K. Jing, Review of advanced physical and data-driven models for dynamic bioprocess simulation: case study of algae–bacteria consortium wastewater treatment, *Biotechnology and Bioengineering* 116 (2) (2019) 342–353.
 - [58] P. Schäfer, A. Caspari, K. Kleinhans, A. Mhamdi, A. Mitsos, Reduced dynamic modeling approach for rectification columns based on compartmentalization and artificial neural networks, *AIChE Journal* 65 (5).
 - [59] A. M. Schweidtmann, A. Mitsos, Deterministic global optimization with artificial neural networks embedded, *Journal of Optimization Theory and Applications* 180 (3) (2019) 925–948.

- [60] A. M. Schweidtmann, W. R. Huster, J. T. Lüthje, A. Mitsos, Deterministic global process optimization: accurate (single-species) properties via artificial neural networks, *Computers & Chemical Engineering* 121 (2019) 67–74.
- [61] W. R. Huster, A. M. Schweidtmann, A. Mitsos, Working fluid selection for organic rankine cycles via deterministic global optimization of design and operation, *Optimization & Engineering* (2019) 1–20.
- [62] A. M. Schweidtmann, D. Bongartz, W. R. Huster, A. Mitsos, Deterministic global process optimization: flash calculations via artificial neural networks, in: *Computer Aided Chemical Engineering*, Vol. 46, Elsevier, 2019, pp. 937–942.
- [63] W. R. Huster, A. M. Schweidtmann, A. Mitsos, Impact of accurate working fluid properties on the globally optimal design of an organic rankine cycle, in: *Computer Aided Chemical Engineering*, Vol. 47, Elsevier, 2019, pp. 427–432.
- [64] P. Schäfer, A. M. Schweidtmann, P. H. Lenz, H. M. Markgraf, A. Mitsos, Wavelet-based grid-adaptation for nonlinear scheduling subject to time-variable electricity prices, *Computers & Chemical Engineering* (2019) 106598.
- [65] S. Madaeni, N. T. Hasankiadeh, A. Kurdian, A. Rahimpour, Modeling and optimization of membrane fabrication using artificial neural network and genetic algorithm, *Separation and Purification Technology* 76 (1) (2010) 33–43.
- [66] H. Al-Zoubi, N. Hilal, N. Darwish, A. W. Mohammad, Rejection and modelling of sulphate and potassium salts by nanofiltration membranes: neural network and Spiegler–Kedem model, *Desalination* 206 (1–3) (2007) 42–60.
- [67] M. Wessling, M. Mulder, A. Bos, M. Van Der Linden, M. Bos, W. Van Der Linden, Modelling the permeability of polymers: a neural network approach, *Journal of Membrane Science* 86 (1–2) (1994) 193–198.
- [68] E. A. Roehl Jr, D. A. Ladner, R. C. Daamen, J. B. Cook, J. Safarik, D. W. Phipps Jr, P. Xie, Modeling fouling in a large RO system with artificial neural networks, *Journal of Membrane Science* 552 (2018) 95–106.
- [69] F. Salehi, S. M. Razavi, Modeling of waste brine nanofiltration process using artificial neural network and adaptive neuro-fuzzy inference system, *Desalination and Water Treatment* 57 (31) (2016) 14369–14378.
- [70] R. Soleimani, N. A. Shoushtari, B. Mirza, A. Salahi, Experimental investigation, modeling and optimization of membrane separation using artificial neural network and multi-objective optimization using genetic algorithm, *Chemical Engineering Research and Design* 91 (5) (2013) 883–903.
- [71] R. Femmer, A. Mani, M. Wessling, Ion transport through electrolyte/polyelectrolyte multi-layers, *Scientific Reports* 5 (2015) 11583.
- [72] L. Evdochenko, R. Femmer, J. Kamp, Y. Xu, V. Nikonenko, M. Wessling, Direct numerical simulations of ion transport through polyelectrolyte multilayer nanofiltration membranes, Submitted.
- [73] D. Bongartz, J. Najman, S. Sass, A. Mitsos, MAiNGO – McCormick-based algorithm for mixed-integer nonlinear global optimization, Tech. rep., *Process Systems Engineering (AVT.SVT)*, RWTH Aachen University (2019).
- [74] A. M. Schweidtmann, A. Mitsos, Melon: Machine learning models for optimization, <https://git.rwth-aachen.de/avt.svt/public/MeL0n/> (2020).
- [75] COMSOL, Inc, COMSOL Multiphysics® (Version 5.3), <https://www.comsol.com>, accessed 19 June 2019.
- [76] MATLAB, version 9.3 (R2017b), The MathWorks Inc., Natick, Massachusetts, 2017.
- [77] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, *Nature* 521 (7553) (2015) 436–444.
- [78] J. E. Dayhoff, J. M. DeLeo, Artificial neural networks: opening the black box, *Cancer: Interdisciplinary International Journal of the American Cancer Society* 91 (S8) (2001) 1615–1635.
- [79] D. W. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, *Journal of the society for industrial and applied mathematics* 11 (2) (1963) 431–441.
- [80] A. E. Yaroshchuk, Negative rejection of ions in pressure-driven membrane processes, *Advances in Colloid and Interface Science* 139 (1–2) (2008) 150–173.
- [81] D. T. Doncevic, A. M. Schweidtmann, Y. Vaupel, P. Schäfer, A. Caspari, A. Mitsos, Deterministic global nonlinear model predictive control with neural networks embedded, Submitted.
- [82] Pentair Xflow Products, <https://xflow.pentair.com/en/products/>, accessed: 1st of April 2019.
- [83] DU PONT Ultrafiltration Products, <https://www.dupont.com/water/ultrafiltration.html>, accessed: 1st of April 2019.

- [84] A. Verberne, J. Wouters, Membraanfiltratie voor de drinkwaterbereiding: economische optimalisatie van ontwerpparameters, *H2O* 26 (14) (1993) 383–387.
- [85] S. Sethi, M. R. Wiesner, Cost modeling and estimation of crossflow membrane filtration processes, *Environmental Engineering Science* 17 (2) (2000) 61–79.
- [86] W. Ang, D. Nordin, A. Mohammad, A. Benamor, N. Hilal, Effect of membrane performance including fouling on cost optimization in brackish water desalination process, *Chemical Engineering Research and Design* 117 (2017) 401–413.
- [87] A. Baur, P. Fritsch, W. Hoch, G. Merkl, J. Rautenberg, M. Weiß, B. Wricke, *Mutschmann/Stimmelmayer Taschenbuch der Wasserversorgung* (2014).