Received: 13 November 2018 | Revised: 23 January 2019 | Accepted: 2 February 2019

DOI: 10.1002/aic.16568

AIChE JOURNAL

PROCESS SYSTEMS ENGINEERING

Reduced dynamic modeling approach for rectification columns based on compartmentalization and artificial neural networks

Pascal Schäfer | Adrian Caspari | Kerstin Kleinhans | Adel Mhamdi | Alexander Mitsos

AVT - Aachener Verfahrenstechnik, Process Systems Engineering, RWTH Aachen University, Aachen, Germany

Correspondence
Alexander Mitsos, AVT Process Systems Engineering, RWTH Aachen University, 52056 Aachen, Germany.
Email: amitsos@alum.mit.edu

Funding information
Bundesministerium für Bildung und Forschung, Grant/Award Number: Kopernikus SynErgie

Abstract

The availability of reduced-dimensional, accurate dynamic models is crucial for the optimal operation of chemical processes in fast-changing environments. Herein, we present a reduced modeling approach for rectification columns. The model combines compartmentalization to reduce the number of differential equations with artificial neural networks to express the nonlinear input–output relations within compartments. We apply the model to the optimal control of an air separation unit. We reduce the size of the differential equation system by 90% while limiting the additional error in product purities to below 1 ppm compared to a full-order stage-by-stage model. We demonstrate that the proposed model enables savings in computational times for optimal control problems by ~95% compared to a full order and ~99% to a standard compartment model. The presented model enables a trade-off between accuracy and computational efficiency, which is superior to what has recently been reported for similar applications using collocation-based reduction approaches.

KEYWORDS
air separation, artificial neural networks, optimal control, rectification, surrogate modeling

1 | INTRODUCTION

Today's market environments are characterized by an increasing volatility due to fluctuations in the availability and prices of feedstocks and energy sources. Adapting the operating strategy of the chemical and process industry to such new circumstances is a crucial challenge.[1] One of the most important topics in this context is the optimization of the electricity consumption of processes subject to fluctuating prices, which is referred to as demand-side management (DSM).

Over the past decade, the economic benefits of DSM measures have been evaluated for a variety of processes using quasi-stationary process models (cf. the recent review of Zhang and Grossmann[2]). Although the assumption of quasi-stationarity is valid for many processes, for example, desalination[3] or electrolysis,[4,5] many chemical processes are characterized by non-negligible response times or time delays. In particular, this holds for processes with large liquid hold-ups in rectification sequences, such as the cryogenic air separation.[6] Nevertheless, its high energy demand combined with an inherent possibility to store energy in form of cryogenic liquid products makes the air separation a highly important process in the context of DSM. We will therefore use the air separation unit (ASU) as a motivating and highly relevant example. However, the developed method is not limited to this particular process, but it is rather applicable to all rectification columns.

The vast majority of literature concerning the operation of ASUs subject to fluctuating electricity prices accounts for dynamic responses only indirectly by restricting the transitions between quasi-stationary operating points.[7–10] Recent work aims at directly accounting for dynamics when making scheduling decisions through the use of low-order dynamic models,[11,12] which is also extended to a closed-loop scheduling formulation.[13] This framework furthermore allows for

explicit consideration of the tracking behavior of a subordinated model predictive controller (MPC) in scheduling of ASUs.[14] In related works, Jamaludin and Swartz[15] and Li and Swartz[16] further propose to introduce the optimality conditions of subordinated MPC into the formulation of the optimal operation. However, this latter approach has not yet been applied to ASUs or other large-scale distillation systems.

Moreover, an integration of the economic perspective directly into the control problem, which is referred to as economic model predictive control (eMPC), is possible and regarded as highly promising in the context of optimal operation in fast-changing environments.[17] For a comprehensive review of successful applications of eMPC, the reader is referred to the work of Ellis et al.[18] However in eMPC, the real-time applicability is crucial, as, particularly in the case of ASUs, large-scale nonlinear dynamic optimization problems need to be solved repeatedly. One opportunity to address the real-time challenge lies in the use of fast update schemes, which avoid the solution of a nonlinear problem. Corresponding control schemes have already been applied to ASUs.[19–21] A second possibility is the use of reduced-dimensional dynamic models to lower the computational demand for solving the dynamic optimization problems. However, previous work also documents the importance of highly accurate process models for control purposes in increasingly transient operating scenarios.[22] Consequently, highly accurate but computationally tractable process models are desired, in particular for rectification columns that easily comprise thousands of equations when modeled rigorously. Numerous approaches have been proposed in the literature, for example, the brief review in the remainder of this article. Recent work from Cao et al,[23] for example, applies a collocation-based approach for model reduction that was originally proposed several decades earlier.[24–26] Therein, required column states are reconstructed by means of interpolation between few collocation points to establish a low-dimensional ASU model. The authors later successfully apply this type of model in several case studies, for example, the optimal operation of a nitrogen plant subject to fast-changing electricity prices[27] or the optimization-based planning of preemptive actions to manage upcoming changes in customer demands.[28]

Nevertheless, there is a continued demand for highly accurate process models that substantially improve the computational efficiency in dynamic optimization. We thus present a novel modeling approach for rectification columns. Results and discussions focus on ASUs but are not limited to them. We combine the idea of compartmentalization[29–31] with surrogate modeling via artificial neural networks (ANNs). The model thus combines physical insights with machine learning techniques that replace computationally inefficient model parts. The use of the physically inspired compartmentalization approach mitigates the shortcomings of purely data-driven models, such as extrapolability. The proposed model shows advantageous properties compared to collocation-based models (CBMs). More precisely, it inherently satisfies integral balance equations around column sections as well as the entire column at any time. Besides, a substantial reduction of the number of differential equations can be achieved, which is beneficial for dynamic optimizations.

The remainder of this article is organized as follows: in the next section, a brief introduction into the ASU considered is given. Furthermore, we present a review on reduced-dimensional dynamic modeling of rectification columns. Thereafter, we show the derivation of our proposed model based on compartmentalization and ANNs. A specific section is dedicated to the fitting procedure for the ANNs. The performance of the model is evaluated in a case study concerning the optimal operation of the plant described subject to an historic time series of fluctuating German electricity spot prices.

2 | AIR SEPARATION UNITS

As stated in the introduction, we select the ASU as a case study due to its high relevance in the context of increasingly transient process operation. The key principle of this process is a cryogenic distillation to retrieve pure air components. This work focuses on a process for producing highly pure gaseous nitrogen for an on-site customer, such as a chemical plant (Figure 1). Similar process designs are well investigated in the literature.[11–14,21,27,32,33]

Incoming feed air is compressed to 6–10 bar and subsequently chilled to ambient temperature. Afterward, the feed air is cooled down to cryogenic temperatures in a counter current heat exchanger. Part of the feed air stream is withdrawn from the heat exchanger and expanded in a turbine. The rest of the feed air is completely liquefied.

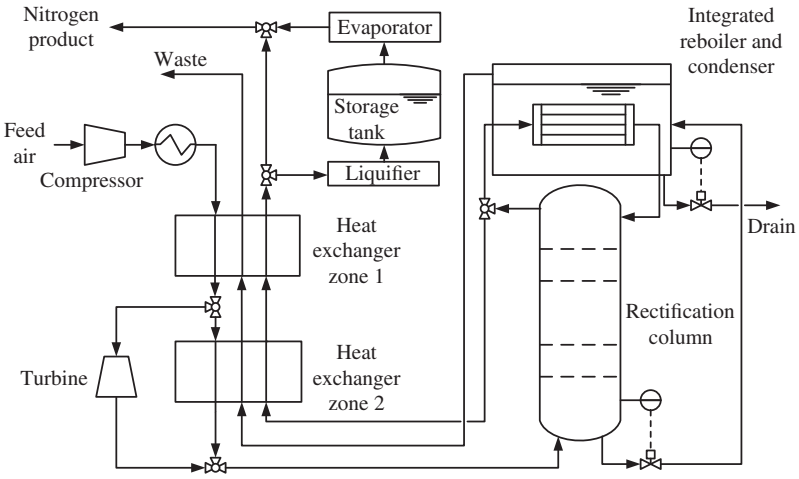

FIGURE 1 Flowsheet of the nitrogen purification plant considered

Both vapor and liquid air are fed to the bottom stage of the cryogenic rectification column. This column comprises approximately 40 theoretical stages to achieve the desired product purities. On top of the column, gaseous nitrogen is withdrawn. The heat released in the condenser of the column is used in the integrated reboiler, which is fed with the bottom product of the column and operates at a lower pressure. Note that the column thus comprises an internal heat integration. A liquid waste stream is used to control the level of the reboiler. Both the gaseous nitrogen product stream and the gaseous waste stream from the reboiler pass the heat exchanger and thereby cool down the feed stream to the column. This thermal coupling of the column's input and outputs makes the ASU an interesting case study for control purposes.

Furthermore, the process has a liquid storage tank enabling that a part of the gaseous product stream can be liquefied and stored. If required, liquid product can be taken from the storage tank and revaporized to satisfy the gaseous nitrogen demand of the customer, thus increasing the plant's ability to perform load changes. Moreover, the buffer tank can be utilized for performing DSM, that is, when electricity is cheap, production can be increased to fill the tank. Vice versa, if electricity is expensive, production is decreased and the demand is satisfied by revaporizing liquid from the tank.

3 | MODELING THE RECTIFICATION COLUMNS

3.1 | Full-order stage-by-stage model

A rigorous full-order stage-by-stage model (FSM) forms the basis of the reduced models considered and is also used as a benchmark. The FSM consists of dynamic balance equations for each of the interconnected stages. In the following, we explicitly present the differential algebraic equation (DAE) system for the FSM. This model is standard for rectification columns but we show it to facilitate describing the derivation of the reduced models. Note that the derivation of the reduced models is largely independent of the actual assumptions made for the FSM.

In this work, we apply the following assumptions for the FSM that have been widely used and validated in literature for modeling of ASUs. Thus, we also consider the FSM a suitable replacement of a real rectification column in an ASU.

1. The vapor holdup on a column stage is negligible compared to the liquid holdup.[19,23,34,35]
2. Exiting liquid and vapor streams on each stage are in thermodynamic equilibrium.[19,23,34,35]
3. The liquid on each stage is ideally mixed.[19,23,34,35]
4. The dynamics of the temperature are much faster than the mass exchange, that is, a pseudo steady-state energy balance can be used. Note that this also reduces the differential index of the equation system from two to one.[23,35]
5. There is a constant pressure difference between neighboring stages.[19,35]
6. The mass holdup in the condenser is negligible.[19,23,34,35]
7. There is a linear relation for stage hydraulics with constant proportional factor k^d on each stage.[19,23,35]

Following these assumptions, the DAE system for a single column stage i with three components (N_2, O_2, and Ar) reads:

$$\dot{M}_i = L_{i-1} + V_{i+1} - L_i - V_i \tag{1}$$

$$M_i \dot{x}_i^j = L_{i-1}\left(x_{i-1}^j - x_i^j\right) + V_{i+1}\left(y_{i+1}^j - x_i^j\right) - V_i\left(y_i^j - x_i^j\right), \quad j = N_2, O_2 \tag{2}$$

$$x_i^{N_2} + x_i^{O_2} + x_i^{Ar} = 1 \tag{3}$$

$$0 = L_{i-1}\left(h_{i-1}^L - h_i^L\right) + V_{i+1}\left(h_{i+1}^V - h_i^L\right) - V_i\left(h_i^V - h_i^L\right) \tag{4}$$

$$\begin{pmatrix} \boldsymbol{y_i} \\ T_i \end{pmatrix} = \text{px-flash}(\boldsymbol{x_i}, p_i) \tag{5}$$

$$h_i^L = \text{h}^\text{L}(\boldsymbol{x_i}, T_i, p_i) \tag{6}$$

$$h_i^V = \text{h}^\text{V}(\boldsymbol{y_i}, T_i, p_i) \tag{7}$$

$$L_i = k_d M_i \tag{8}$$

with M_i denoting the holdup, L_i and V_i the liquid and vapor streams leaving stage i, x_i^j and y_i^j as mole fractions of component j in liquid and vapor phases, respectively (which can be combined as vectors $\boldsymbol{x_i}$ and $\boldsymbol{y_i}$), T_i denoting the temperature, p_i the pressure, and h_i^L and h_i^V as liquid and vapor enthalpies. The specific functions px - flash(·), $\text{h}^\text{L}(\cdot)$, and $\text{h}^\text{V}(\cdot)$ depend on the thermodynamic models used. In case of additional feeds to the stage or product withdrawals, the corresponding streams are added to the balance equations.

The entire column model is formed by connecting neighboring stages and adding the equation systems for condenser and reboiler.

3.2 | Reduced modeling of rectification columns

Existing literature techniques for reduced modeling of rectification columns can be divided into three major types: (a) nonlinear wave theory,[36,37] (b) compartmentalization,[29–31] and (c) collocation-based reduction.[24–26] All approaches have already been applied to dynamic simulation as well as operational optimization and/or control of ASUs.[23,27,28,35,38–41] We set the focus on compartmentalization and collocation in the following. These two reduction techniques are derived from the rigorous FSM presented above.

3.2.1 | Collocation-based model

CBMs represent an intuitive approach to reduce the number of equations. Therein, the equation systems for the N column stages are replaced by equation systems for K the so-called collocation points

(CPs). With $K \ll N$, a significant reduction in the number of model equations is achieved. All required remaining column states are reconstructed by means of interpolation between the CPs. Thereby, time-consuming stage-to-stage calculations are avoided in the CBM, which is believed to increase the computational efficiency of the CBM. For instance, Cao et al divide the column into multiple finite elements (FEs) with two CPs per FE.[23] Consequently, within each FE, a 2nd order Lagrange polynomial is used. Note that the CPs can but not necessarily need to coincide with actual column stages.

The size of the entire DAE system representing the CBM scales directly with the number of CPs used. Technically, a user-defined reduction of computational demand thus can be achieved. At the same time, the number of CPs used however poses limits on the realizable accuracy of transient responses.[42] Consequently, a balancing between model reduction (low number of CPs desired) and model accuracy (high number of CPs desired) is necessary. When applying collocation-based reduction approaches to rectification columns of ASUs, Cao et al find a reduction in the equation system by 60–70% to be possible while not introducing errors in product purities of more than 1 ppm. Reported reductions of simulation times are in a similar order (50–60%).[23] One important advantage of the CBM is that its implementation derived from an existing FSM is straightforward. The time for offline preparation is thus comparably small.

However, approximation of the column profile by interpolation polynomials can introduce nonphysical behavior.[42] In particular, integral balance relations around multiple column stages or the entire column are not guaranteed to be satisfied. Importantly, this accounts for both steady state as well as dynamic simulations and thus poses limits on the degree of model reduction that can be achieved by using CBMs.

3.2.2 | Compartment model

In the compartment model (CM), one of the compartments is defined to consist of multiple single stages. Without loss of generality, balance equations for one single stage, the so-called sensitivity stage, can be replaced by the overall compartment balances. Assuming that stages $N_{c,1}$ to $N_{c,2}$ form compartment c (Figure 2), we obtain

$$\dot{M}_c = L_{N_{c,1}-1} + V_{N_{c,2}+1} - L_{N_{c,2}} - V_{N_{c,1}} \tag{9}$$

$$\begin{aligned} M_c \dot{x}_c^j \quad &= L_{N_{c,1}-1}\left(x^j_{N_{c,1}-1} - x_c^j\right) + V_{N_{c,2}+1}\left(y^j_{N_{c,2}+1} - x_c^j\right) \\ &- L_{N_{c,2}}\left(x^j_{N_{c,2}} - x_c^j\right) - V_{N_{c,1}}\left(y^j_{N_{c,1}} - x_c^j\right), \quad j = N_2, O_2 \end{aligned} \tag{10}$$

$$x_c^{N_2} + x_c^{O_2} + x_c^{Ar} = 1 \tag{11}$$

$$\begin{aligned} 0 \quad &= L_{N_{c,1}-1}\left(h^L_{N_{c,1}-1} - h_c^L\right) + V_{N_{c,2}+1}\left(h^V_{N_{c,2}+1} - h_c^L\right) \\ &- L_{N_{c,2}}\left(h^L_{N_{c,2}} - h_c^L\right) - V_{N_{c,1}}\left(h^V_{N_{c,1}} - h_c^L\right). \end{aligned} \tag{12}$$

In Equations 9–12, the compartment states M_c, x_c^j, and h_c^L are introduced and defined as follows:

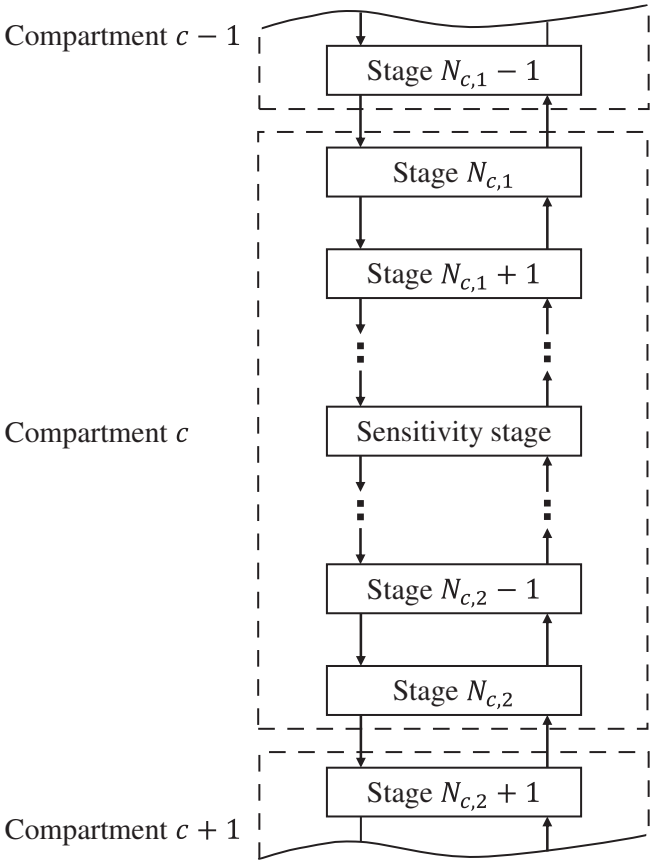

FIGURE 2 Illustration of the compartmentalization approach applied. Dashed lines depict the compartment boundaries

$$M_c = \sum_{i=N_{c,1}}^{N_{c,2}} M_i \tag{13}$$

$$x_c^j = \frac{1}{M_c}\sum_{i=N_{c,1}}^{N_{c,2}} M_i x_i^j, \quad j = N_2, O_2, \mathrm{Ar} \tag{14}$$

$$h_c^L = \frac{1}{M_c}\sum_{i=N_{c,1}}^{N_{c,2}} M_i h_c^L. \tag{15}$$

Assuming the compartments to be sufficiently large, single-stage dynamics can be neglected compared to overall compartment dynamics. Consequently, single-stage balance equations are assumed stationary. Thus, the entire equation system for compartment c consists of Equations 9–15 plus Equations 5–8 for stages $j = N_{c,1} \ldots N_{c,2}$ and the steady-state versions of Equations 1–4 for stages $j = N_{c,1} \ldots N_{c,2}$ except for the sensitivity stage. The overall column model arises from connecting neighboring compartments. If the DAE system for single stages is of differential index one, so is the equation system for an entire compartment.

Major findings from Bian et al[35] indicate that:

1. For a column with 60 stages, 10–15 single stages can be combined to one compartment without substantially influencing the quality of the dynamic response.

2. For highest accuracy, the sensitivity stage is to be placed at the center of the compartment.
3. Stages with additional feed streams or product withdrawals are recommended to be modeled separately and not included into one compartment.

In contrast to the CBM, the CM preserves physical insights in form of integral balance relations around single compartments and therefore also the entire column. Furthermore, as an obvious result of its derivation from the FSM, the CM has the same steady state as the FSM.

Regarding the computational efficiency of the CM, a substantial reduction of the number of differential equations can be achieved without considerable influence on the accuracy as has been demonstrated by Khowinij et al[40] and Bian et al.[35] However, the total number of equations (differential plus algebraic ones) remains (nearly) constant. In other words, the compartmentalization only changes the structure of the DAE but not its size. Therefore, the actual reduction of computational times will strongly be affected by the selection of the integrator, that is, if the integrator deals well with the new DAE system, a substantial reduction might be achieved (cf., the work of Linhart and Skogestad[43]). In the work of Bian et al,[35] however, no significant reduction of computational times was observed. Similar results are obtained in this work, that is, no reductions in computational times are found when using state-of-the-art integrators. In fact, our results even indicate an increase in computational times (cf., Table 3).

4 | PROPOSED MODEL BASED ON COMPARTMENTALIZATION AND ANNS

4.1 | Model formulation

We derive the proposed model from the representation of the original compartment model as an index-1 DAE system:

$$\dot{\hat{\boldsymbol{x}}}(t) = \hat{\boldsymbol{f}}(\hat{\boldsymbol{x}}(t), \hat{\boldsymbol{u}}(t), \hat{\boldsymbol{y}}(t)) \tag{16}$$

$$\boldsymbol{0} = \hat{\boldsymbol{g}}(\hat{\boldsymbol{x}}(t), \hat{\boldsymbol{u}}(t), \hat{\boldsymbol{y}}(t), \hat{\boldsymbol{z}}(t)). \tag{17}$$

In Equations 16 and 17, we introduce the following notation: differential compartment states are denoted by $\hat{\boldsymbol{x}}(t)$. Compartment inputs, which are handed over by the neighboring compartments, are denoted by $\hat{\boldsymbol{u}}(t)$, which corresponds to states of column stages $N_{c,1} - 1$ and $N_{c,2} + 1$ (cf., Figure 2). Compartment outputs, which are required in the equation system of neighboring compartments, are denoted $\hat{\boldsymbol{y}}(t)$, which corresponds to the states of column stages $N_{c,1}$ and $N_{c,2}$ (cf., Figure 2). The remaining (algebraic) compartment states are denoted $\hat{\boldsymbol{z}}(t)$.

When solving Equations 16 and 17, the main computational effort is spent in the solution of the highly nonlinear algebraic equation system (17)[43] that mainly originates from the thermodynamic relations (Equations 5–8). To reduce the computational effort, Linhart and Skogestad[44] propose interpolation between tabulated presolved solutions (18).

$$\begin{pmatrix} \hat{\boldsymbol{y}}(t) \\ \hat{\boldsymbol{z}}(t) \end{pmatrix} = \hat{\boldsymbol{g}}^{-1}(\hat{\boldsymbol{x}}(t), \hat{\boldsymbol{u}}(t)) \tag{18}$$

We propose to use a continuous surrogate model for the solution of the algebraic Equation 17 to further improve the computational efficiency of the model and make it readily applicable in state-of-the-art frameworks for dynamic optimization. We apply ANNs, as they have been shown to be powerful surrogate models for replacing thermodynamic relationships.[45–48] Furthermore, ANNs can be evaluated efficiently in a forward manner.[49] Likewise, the use of backpropagation algorithms (as a special case of automatic differentiation for ANNs) allows for efficient gradient calculations even for very complex network structures.[50] However, the latter opportunity is currently omitted in our computational framework. Instead, analytical Jacobians are used for the ANN function, which at some point will presumably cause limitations in problem complexities.

Sophisticated computer codes are readily available for efficient training of the ANNs. In particular, ANNs can also be fitted very efficiently to large data sets, which arise from a sampling of the high-dimensional input space.[51]

The DAE system representing the ANN-based compartment model (ANNCM) reads:

$$\dot{\hat{\boldsymbol{x}}}(t) = \hat{\boldsymbol{f}}(\hat{\boldsymbol{x}}(t), \hat{\boldsymbol{u}}(t), \hat{\boldsymbol{y}}(t)) \tag{19}$$

$$\hat{\boldsymbol{y}}(t) = \hat{\boldsymbol{g}}_{\mathrm{ANN}}(\hat{\boldsymbol{x}}(t), \hat{\boldsymbol{u}}(t)). \tag{20}$$

We highlight that model formulation (19) and (20) is only one possibility of an ANN-based compartmentalization approach. The choice of this system, however, seems appealing as it shows an analogy to the dynamic modeling of simple flash units, that is, the model outputs can be calculated as an explicit function of the model inputs (cf., a Tx-flash). Others possibilities for using ANNs exist as well. For instance, using a surrogate model for the ordinary differential equation (ODE) form of the index-1 DAE system (16) and (17) represents an alternative. Likewise, the ANN could also be used to replace (specific parts of) the mapping (17).

The ANNCM inherits the ability of the original CM to substantially reduce the number of differential equations, which is a desired attribute with regard to the computational efficiency. Furthermore, in the ANNCM, we facilitate the solution of the algebraic equation system describing the input–output relations within the compartments. More precisely, we confine to solving the system only for those variables which are relevant for the compartment balances and introduce explicit functions for that purpose. Like the CM, the ANNCM still considers physical insights in form of integral balance relations.

In contrast to the CM, the steady state of the ANNCM does not necessarily coincide with the steady state of the FSM. However, assuming a sufficiently good estimation accuracy of the function $\hat{\boldsymbol{g}}_{\mathrm{ANN}}(\cdot)$, deviations

are expected to be small. One disadvantage of the ANNCM is that it requires a significant preparation effort for obtaining $\hat{\boldsymbol{g}}_{\mathrm{ANN}}(\cdot)$. For instance, the wall time for the training of the most complex single ANNs used in the case study was 10–20 h by applying the presented procedure. This preparation effort represents a shortcoming compared to CBMs that can be set up much more easily. However, the offline preparation is only done once offline and in an automated way. The CPU time savings are in contrast online and repeatedly seen.

Extrapolability of the ANN functions cannot be expected in general. It is thus crucial to cover the required operating range well. Further note that using the hyperbolic tangent activation function of the ANN intrinsically introduces nonconvexities that need to be handled within the optimization problem.[52] Consequently, in the case of overfitting the relation (20), these nonconvexities might also introduce multiple unphysical local minima to the optimization problem. However, this risk can be mitigated by selecting a suitable ANN representation during training, for example, through applying commonly used regularization approaches such as the early stopping criterion[49] as done in this work or more sophisticated approaches such as dropout.[53] The results of the case study conducted in this work further indicate that the discussed issues might be of minor relevance in practice.

4.2 | Training procedure

Obtaining the required function $\hat{\boldsymbol{g}}_{\mathrm{ANN}}(\cdot)$ in Equation 20 consists of the three following major steps.

4.2.1 | Sampling the input space

As discussed earlier, it is crucial that the relation $\hat{\boldsymbol{g}}_{\mathrm{ANN}}(\cdot)$ produces accurate predictions across the entire operating range of the compartment section. We therefore aim to provide input samples (i.e., samples of $\hat{\boldsymbol{x}}(t)$ and $\hat{\boldsymbol{u}}(t)$) that cover this range well. If the range was known exactly, problems with extrapolation would not arise. However, identification of the actual operating range is not straightforward. For this purpose, the following approach is proposed: first, characteristic load changes across the entire operating range of the rectification column are performed by simulating the FSM. Thereby, lower and upper bounds for differential compartment states $\hat{\boldsymbol{x}}(t)$ and input variables $\hat{\boldsymbol{u}}(t)$ for each compartment can be obtained. Afterward, the space spanned by the bounds can be sampled by means of appropriate techniques, for example, the well-known latin hypercube method.[54] Here, we recommend to add adequate safety factors to the bounds of $\hat{\boldsymbol{x}}(t)$ and $\hat{\boldsymbol{u}}(t)$.

4.2.2 | Solving the nonlinear equation system

In a second step, Equation 17 is solved for all samples of $\hat{\boldsymbol{x}}(t)$ and $\hat{\boldsymbol{u}}(t)$. As this corresponds to the solution of nonlinear equation systems, various computer codes can be used for solution. As all models in the case study are implemented in the modeling language Modelica and made accessible to our in-house software package for dynamic simulation and optimization DyOS[55] as a functional mock-up unit (FMU), we herein use the internal capabilities of DyOS for solving nonlinear algebraic equation systems.

4.2.3 | Fitting the ANNs

As by solving the nonlinear equation system, the required input–output samples for Equation 18 are obtained, the function $\hat{\boldsymbol{g}}_{\mathrm{ANN}}(\cdot)$ can be fitted using state-of-the-art computer codes for training of ANNs. For instance, in this work, we use MATLAB R2018a's Neural Network Toolbox (Mathworks, Inc.). We investigated different network architectures and obtained best results by representing each compartment output separately with one multiple-input-single-output ANN with one hidden layer. For good performance of the ANNs, we further recommend to use the same scaling procedure for composition variables as applied by Cao et al.[23] More specifically, we use the logarithm of molar fractions (for O_2 and Ar) or the logarithm of one minus the molar fraction (for N_2) within the ANN instead of the molar fractions directly. For determining the number of neurons per ANN, an iterative procedure is applied, which increases the number of neurons within the hidden layer of the ANNs until a minimum performance criterion is met.

We highlight that potential improvements of the fitting procedure exist in each of the three steps. For instance, when sampling the input space, extra weight could be put to specific operating points (e.g., nominal one, limiting product specifications) to reproduce them better. Furthermore, tailored solution methods for solving the nonlinear equation system might be applied to improve performance. In particular, the resulting algebraic equation system shows substantial similarities with steady-state simulations of distillation columns. Thus, corresponding software packages might be a promising alternative. Finally, the procedure for determining the number of neurons might be replaced by more advanced ones and more complex network structures could be explored.

5 | CASE STUDY: OPTIMAL CONTROL OF A NITROGEN PLANT

5.1 | Formulation of the optimal control problem

The case study focuses on optimal operation of the plant depicted in Figure 1 over 24 h. For this purpose, we consider historic German electricity spot prices on the day-ahead market from January 4, 2018 (Figure 3) which can be retrieved from the European Power Exchange (www.epexspot.com). Note that this day was very promising for DSM due to strong fluctuations in the electricity price with even negative prices at the beginning. Furthermore, we consider a constant withdrawal of 155 $\frac{\mathrm{mol}}{\mathrm{s}}$ gaseous nitrogen by an on-site customer.

The case study corresponds to the solution of the optimal control problem (OCP) (21)–(24):

$$\min_{\text{Controls}} \quad \phi(t_f) \tag{21}$$

s.t. Process mode (22)

Process constraints (23)

Constraints on controls (24)

The optimal control inputs minimize the objective ϕ at final time t_f subject to a process model represented by a DAE system and constraints on process variables as well as control inputs. The single components of the OCP are introduced in the following.

5.1.1 | Objective function and operational constraints

The overall objective is to operate the plant in a way that all production targets are satisfied with as low electricity costs as possible. For calculating the electricity cost ϕ (21), we consider the current power consumption of the main air compressor $P_{MAC}(t)$ and the liquefier $P_{LIQ}(t)$ minus the power generation of the expansion turbine $P_{TURB}(t)$ multiplied with the instantaneous electricity price $c_E(t)$:

$$\dot{\phi}(t) = c_E(t)(P_{MAC}(t) + P_{LIQ}(t) - P_{TURB}(t)). \quad (25)$$

Due to delivery commitments, the gaseous product demand is to be fulfilled at every point in time either via the column withdrawal or re-vaporization of the tank content. In order to prevent emptying the tank, the stored volume must reach its nominal value at the end of the considered horizon of 24 h. Besides, there is a maximum impurity in the product stream that is not allowed to be exceeded at any time. Further path constrained variables are driving temperature differences at various points of the plant. Moreover, it is required to keep the temperature of the air flow at the turbine inlet above the boiling point and the temperature behind the second heat exchanger zone below the boiling point. All constraints on variables (23) are listed in Table 1. Note that a wide range of additional operational constraints have been used in the literature, such as flooding limits.[11,27] For the sake of simplicity, these are disregarded here. However, their integration would be straightforward. Assuming that the constrained variable can be calculated as a function of column states in the FSM (as is the case for the flooding velocity), we can also use a surrogate model for its calculation inside a compartment, which essentially corresponds to adding one ANN per additional constrained variable.

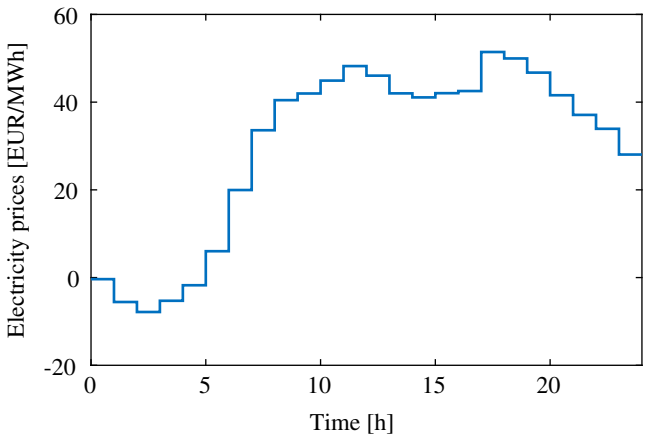

FIGURE 3 Electricity prices for the entire horizon of the case study

Four manipulated variables are available to control the process, which include the feed air stream taken from the environment (F_{MAC}), the fraction of the incoming stream that is liquefied in the second zone of the heat exchanger (F_{LIQ}), the product stream withdrawn from the column top (F_{TOP}), and the stored product stream (F_{TANK}). Respective bounds on the manipulated variables (24) are listed in Table 2 and correspond to an operating range of ±10%.

5.1.2 | Process models

We differentiate between the model used in the OCP (22) and the model used as replacement for the actual plant. The only difference between the two lies in the way that the rectification column is modeled. For all other unit operations, we use the same models as in Caspari et al.[21] for both solving the OCP and simulating the dynamic response of the plant replacement. Details can be found in the Supporting Information. In the plant replacement model, we use an FSM for the rectification column. The FSM comprises 40 theoretical stages, a steady-state condenser model and a single flash unit for the reboiler. For solving the OCP, we apply the presented ANNCM and compare it to the standard CM. Furthermore, we also consider an ideal case without prediction errors in the OPC as benchmark, that is, we also apply the FSM to the OCP.

We investigated different compartmentalization schemes for the CM and gained valuable insights into the influence of the number of compartments on model accuracy, which are used for setting up the ANNCM: Confining to compartments of almost equal size, we found the use of four compartments (i.e., approximately 10 stages per compartment which is in good agreement with the results of Bian et al)[35] to be best suited for the application. This corresponds to a reduction in the differential equation system describing the rectification column by 90%. Increasing the number of compartments to five, that is, a lower degree of reduction, does not show significant increases in accuracy. In contrast, decreasing the number to three, that is, a higher degree of reduction, leads to an undesired inverse-response behavior when simulating step changes. As a further adjustment to the original CM, we include the condenser into the upper compartment of the ANNCM. In contrast, feed stage and reboiler are still modeled separately. Details on the number of neurons used for fitting all required input–output functions within the compartments are also given in the Supporting Information.

5.1.3 | Implementation

Solution of the OPC 21–24 as well as simulations of the plant model are realized using the in-house software package DyOS,[55] which has access to the process models implemented in Modelica and exported as FMUs. DyOS solves the OPC in a sequential approach using a

TABLE 1 Constrained variables in the OCP with their respective bounds

Symbol	Description	Bounds
$1 - y_1^{N_2}$ (ppm)	Product impurity	[0, 20]
ΔT_{IRC} (K)	Driving temperature difference (integrated reboiler and condenser)	[0, ∞]
$\Delta T_{PHX1}^{N_2,in}$ (K)	Driving temperature difference (inlet heat exchanger zone 1)	[0, ∞]
$\Delta T_{PHX1}^{WASTE,in}$ (K)	Driving temperature difference (inlet heat exchanger zone 1)	[0, ∞]
$\Delta T_{PHX1}^{N_2,out}$ (K)	Driving temperature difference (outlet heat exchanger zone 1)	[0, ∞]
$\Delta T_{PHX1}^{WASTE,out}$ (K)	Driving temperature difference (outlet heat exchanger zone 1)	[0, ∞]
$\Delta T_{PHX2}^{N_2,out}$ (K)	Driving temperature difference (outlet heat exchanger zone 2)	[0, ∞]
$\Delta T_{PHX2}^{WASTE,out}$ (K)	Driving temperature difference (outlet heat exchanger zone 2)	[0, ∞]
T_{TURB}^{in} (K)	Temperature at turbine inlet (required to be appropriately superheated)	[115, ∞]
T_{PHX2}^{out} (K)	Temperature behind heat exchanger (required to be appropriately subcooled)	[0, 95]
V_{IRC} (%)	Relative volume (integrated reboiler and condenser)	[95, 105]
V_{TANK} (%)	Relative volume (product tank)	[0, 200]

Abbreviations: OCP: optimal control problem.

single-shooting algorithm.[56,57] Therein, the sequential quadratic programming (SQP) code SNOPT[58] is applied as solver for nonlinear programming (NLP) and both LIMEX[59] and NIXE[60] are applied as integrators. Settings used for the NLP solver and the integrators can be found in Supplementary Material. We parameterize each control vector with 12 equally sized, piecewise-constant intervals.

5.2 | Numerical results

All calculations are conducted on an Intel® Xeon® CPU E5-2630 v2 (Intel Corporation, Santa Clara, USA) with 2.60 GHz and 128 GB RAM. For solution of the NLPs, SNOPT version 7.2–4 is used.

TABLE 2 Manipulated (control) variables for the OCP with their respective bounds. Bounds for F_{MAC}, F_{LIQ}, and F_{TOP} correspond to an operating range of ±10%

Symbol	Description	Bounds
F_{MAC} ($\frac{mol}{s}$)	Inlet flow main air compressor	[301.5, 368.5]
F_{LIQ} ($\frac{mol}{s}$)	Inlet flow zone 2 of heat exchanger	[36.75, 44.90]
F_{TOP} ($\frac{mol}{s}$)	Rectification column product withdrawal	[139.5, 170.5]
F_{TANK} ($\frac{mol}{s}$)	Inlet flow product tank	[0, 20]

Abbreviations: OCP: optimal control problem.

TABLE 3 Average CPU times over 10 runs for solving the OCP using the different rectification column models

	CPU time (LIMEX) (s)	CPU time (NIXE) (s)
FSM	4,909	2,490
CM	12,099	14,928
ANNCM	123	162

Abbreviations: FSM: full-order stage-by-stage model; CM: compartment model; ANNNCM: ANN-based compartment model; ANN: artificial neural network.

5.2.1 | Accuracy of the reduced models

In Figure 4, we present the optimal profiles for selected controls obtained by solving the OCP using different models for the rectification column. As can be seen, there is an almost perfect agreement between the optimal control profiles obtained using the ANNCM, the CM, and the FSM within the OCP. In fact, the highest relative error compared to the optimal control profile from using the FSM is 0.07% for the ANNCM and 0.04% for the CM.

We further compare the actual responses of selected constrained variables from simulating the plant replacement model (with the FSM) applying the optimal control profiles to the corresponding predictions made by using both the CM and the ANNCM in the OCP. We thereby assess the error introduced through the reduction approaches. However, one should consider them in relation to the agreement between the original model (from which the reduced models are derived, that is, the FSM in this case) and actual responses of a physical plant. With regard to control applications, this so-called model-plant mismatch is commonly handled efficiently in a closed-loop manner using feedback of the process. Simply speaking, if errors that are introduced during model reduction are substantially below the inherent model-plant mismatch, the accuracy of the reduced model is not critical. For the ASU application, we assume a mismatch between plants and FSMs in the order of few ppm for the most important (and also the limiting) process variable—the nitrogen product purity—to be realistic.

The results given in Figure 5 validate the good prediction capabilities of the standard CM. Our results are in good agreement with findings from Bian et al,[35] indicating that, for columns comprising around 40 stages, 10 single stages can easily be combined into one compartment without significant influences on the dynamic response. We find that for the impurity in the product stream the root mean square error (RMSE) between the prediction using the CM and the actual plant response is ~0.5 ppm, which corresponds to an error of ~2.5% relative to the impurity limit of 20 ppm. Following the discussion above, this accuracy appears adequate for control purposes, as we assume that a sufficiently large safety factor $\gg$2.5% would be added in practice for bounds on crucial process variables. Due to the derivation of the ANNCM, the accuracy of the CM represents the actual benchmark for the ANNCM, that is, a better agreement between the ANNCM and the FSM than between the CM and the FSM could only occur by chance through vanishing of two error sources.

We further see from Figure 5 that—as expected—the ANNCM inherits the capabilities of the CM to model the transient behavior of

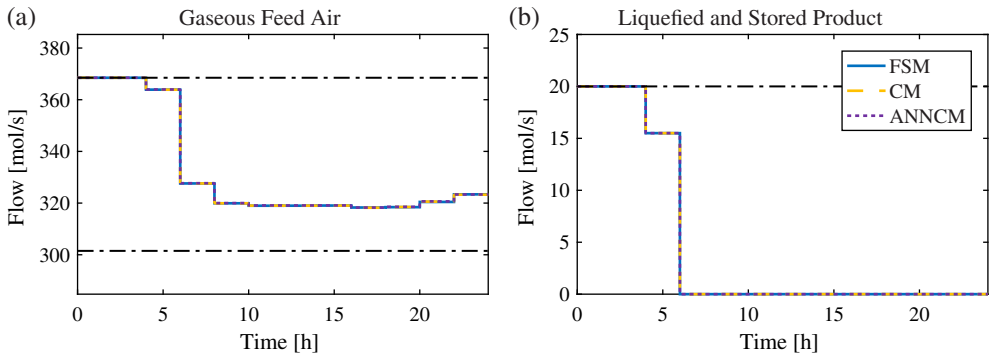

FIGURE 4 Selected optimal control profiles obtained by applying different process models in the OCP. Solid line trajectories are optimized using the FSM, dashed lines using the CM, and dotted lines using the proposed ANNCM. Thin dashed-dotted lines are the respective bounds. (a) Feed air stream of the main air compressor. (b) Liquefied and stored product stream. Abbreviations: OCP: optimal control problem; CM: compartment model; FSM: full-order stage-by-stage model; ANNNCM: ANN-based compartment model

the column well while reducing the number of differential equations by 90% compared to the FSM. More precisely, for the predictions of the ANNCM, an RMSE of 0.67 ppm in the product impurity is found, that is only slightly worse than the RMSE for the CM. The small deviations between the performance of the ANNCM and the CM can be explained by the remaining prediction errors of the embedded ANNs. These errors are present both in transient and in (almost) stationary operation, where the CM achieves a perfect agreement between predictions and actual responses. Because ANNs are universal approximators,[49] the prediction errors of the ANNs (i.e., the deviations between the ANNCM and the CM) can be decreased to arbitrary size, if sufficiently many hidden neurons are applied and data points

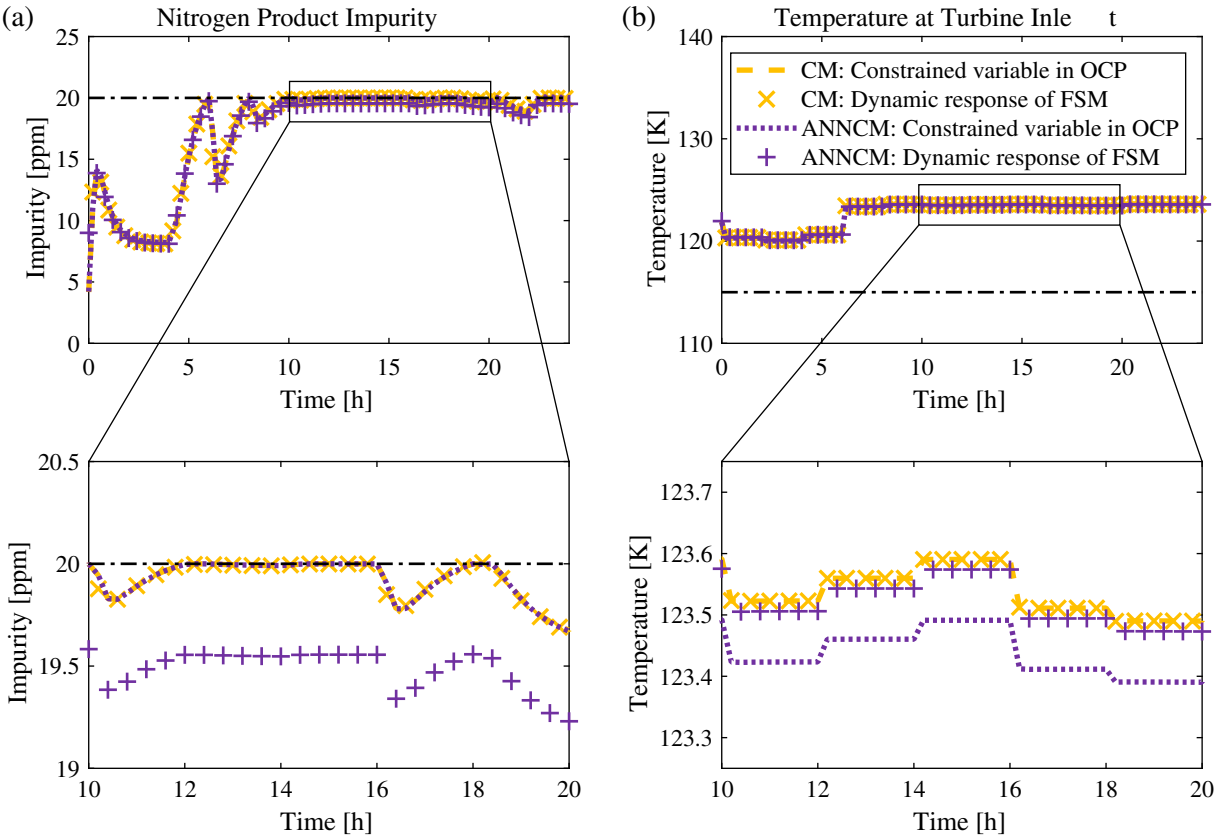

FIGURE 5 Comparison of predicted trajectories of selected constrained variables obtained by using the reduced models in the OCP to the actual simulated dynamic response of the plant replacement (FSM). Dashed lines shows predictions using the CM in the OCP, dotted lines the predictions using the ANNCM. Symbols depict the dynamic responses of the FSM when applying the obtained optimal control profiles. Thin dashed-dotted lines are the respective bounds for the variables. Lower figures show detailed views marked in the upper ones. (a) Impurity in the product stream. (b) Absolute temperature at turbine inlet. Abbreviations: OCP: optimal control problem; CM: compartment model; FSM: full-order stage-by-stage model; ANNNCM: ANN-based compartment model

for training are used. However, in practice, prediction accuracy of the ANNs is limited by the actual number of neurons applied and the size of the data set used for training. Using the obtained ANNs (see, Supporting Information), this additionally introduced error is always below 0.5 ppm for the product purity and thus already in the same order as the inherent limitations of the prediction accuracy of the CM (see above). We highlight that the prediction errors of the ANNCM are thus also significantly smaller than assumed safety factors and that the actual model-plant mismatch between real plants and the FSM might presumably be at least in comparable orders of magnitude. Qualitatively similar results are obtained for different process variables as well. For instance, the RMSE for predicting the temperature at the turbine inlet, which depends on the molar streams and enthalpies of the product and waste streams, is even below 0.1 K for the ANNCM.

5.2.2 | Solution times

Table 3 shows CPU times spent for solving the OCP using different reduced models and the FSM, averaged over 10 runs. Results are presented separately for using the integrators LIMEX and NIXE. We find that the ANNCM enables high-quality prediction capabilities (Figures 4 and 5) while substantially reducing the computational effort for solving the OCP by more than one order of magnitudes compared to using an FSM. This holds for both integrators tested. However, if LIMEX is used, the benefits appear even stronger (factor of ~40 when using LIMEX vs. ~15 when using NIXE). A similar behavior can be observed for the comparison of the ANNCM to the standard CM. Again, the relative improvement is higher when using LIMEX. Noticeably, for both integrators tested the use of the standard CM in our numerical set-up leads to a significant increase in computational times, that is, solution times using the CM are higher by factors of 2.5 (LIMEX) and 6 (NIXE), respectively, compared to the FSM. As discussed in beforehand, this behavior has also been observed in the literature.[43]

An extensive study might become necessary in the future for verification and generalization of the exceptionally good computational efficiency of our reduction approach to other solution methods for dynamic optimization (in particular, simultaneous solution methods).[61]

5.2.3 | Discussion and comparison with literature

The work of Khowinij et al[40] and Bian et al[35] for the application of standard CMs as well as the work of Cao et al[23,27] for the application of CBMs appear suitable for comparing the results to previous literature, as similar processes are considered. Furthermore, these works apply FSMs using the same assumptions as we did for benchmarking the respective reduction approaches. Khowinij et al[40] report a fivefold reduction of computational times using MATLAB's ode15s (Mathworks, Inc.) as integrator for a single nitrogen column. For another column producing both nitrogen and oxygen tested by Bian et al[35] using the same methods, only insignificant reductions were observed. In comparison, our results even show an increase in computational times by applying a CM instead of an FSM when using state-of-the-art integrators. However, we do not exclude the possibility that highly tailored integration methods might be able to efficiently exploit the structure of the CM. We however validate that CMs (as long as not too many stages are combined within one compartment) are exceptionally good in modeling dynamic responses and thus efficiently reduce the number of differential equations. We highlight that this fact is believed to be the key for the computational performance of the presented ANNCM.

Concerning the use of CBMs, Cao et al[27] report a reduction in simulation times in gPROMS (PSE, Ltd.) of ~70% when reducing the model size by ~60% for a similar case study (optimal operation of a nitrogen plant with a liquid buffer subject to fluctuating electricity prices). The authors further report absolute errors in relevant product purities of below 1 ppm for that model,[23] which appears comparable to the accuracy of the presented ANNCM. Unfortunately, the authors do only report time reductions for simulating the process and not for solving the OCP. However, as a first approximation, we estimate savings in solving the OCP to be of a similar order of magnitude. We thus expect the proposed ANNCM to enable a substantially better balancing between model accuracy and computational demand than what could be achieved by using CBMs instead.

6 | CONCLUSION

We propose a reduced model for rectification columns that is based on the compartmentalization approach and combines it with ANNs. The proposed ANNCM takes advantage from the principle of compartmentalization to efficiently reduce the required number of differential equations. Time-consuming stage-to-stage calculations within the single compartments are avoided by using ANNs as surrogate models for the input–output behavior.

We successfully apply the ANNCM in a case study for optimal flexible operation of an ASU subject to fluctuating electricity prices. The ANNCM comprises only 10% of the differential equation for describing the rectification column compared to the original FSM. The induced prediction errors compared to the FSM are below 1 ppm in relevant product purities. A large part of this error is introduced by the compartmentalization, the proportion of the additional error introduced by using the ANNs is smaller. The computational efficiency of the ANNCM is compared to both using an FSM for the rectification columns and the standard CM (without surrogate modeling). We demonstrate that by using the ANNCM for OPCs, solution times can be substantially reduced compared to using the FSM and the CM by more than one order of magnitude while obtaining high-quality results with good agreement to the FSM. The ANNCM thus enables an exceptionally good balancing between accuracy and computational efficiency. We admit that this advantage comes at the price off an increased offline preparation effort, which is required once for setting up the model. However, in particular for online applications like eMPC, we expect this effort to be justified. The small deviations between the predictions of the ANNCM and the actual response of

the plant can likely be treated efficiently in a closed-loop framework utilizing feedback of the plant, particularly considering the solution times presented above.

Important research potentials can be identified that should be addressed in the future. From the modeling perspective, a systematic building of the ANNCM should be considered. In particular, this holds for the decision which and how many stages to unify in one compartment. Furthermore, the existing trade-off between size of the ANN (and thus accuracy of ANN predictions) and its influence on computational demand should be analyzed. Besides, the model performance should be validated in additional case studies considering other rectification (or maybe even absorption) processes. Finally, the concept of ANN-based compartmentalization can in principle also be applied to different unit operations. In particular, we see strong similarities between modeling of (rectification) columns and distributed modeling of heat exchangers. Consequently, adaption of the approach to heat exchangers might be promising.

On the application side, the good balancing between accuracy and computational efficiency makes the model highly interesting for online applications, that is, for eMPC where cyclic feedback of the process is utilized for rejecting disturbances. In order to retain low computational times that ensure real-time applicability, a combination of the proposed model reduction technique with sophisticated structures for eMPC might be required particularly, when larger ASUs are focused that also comprise the production of oxygen and argon via a second and a third rectification column. More specifically, the use of fast eMPC schemes with sensitivity-based updates[62–64] appears promising.

ACKNOWLEDGMENTS

The authors gratefully 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. Furthermore, the authors thank Anna-Maria Ecker, Andreas Peschel and Bernd Wunderlich from Linde Engineering as well as Robert Kender from TU München for valuable discussions concerning the modeling of cryogenic rectification columns.

AUTHOR CONTRIBUTIONS

P. Schäfer conceived the presented approach, implemented the process model for the case study and wrote the manuscript. A. Caspari provided the algorithmic framework for solving the dynamic optimization problem. K. Kleinhans performed the ANN-based model reduction (data generation and training of the ANNs). A. Mhamdi and A. Mitsos provided advice during the development of the model, the formulation of the case studies, and the preparation of the manuscript, guided the effort, and edited the manuscript.

NOTATION

Acronyms

ANN	artificial neural network
ANNCM	(Proposed) ANN-based compartment model
ASU	air separation unit
CBM	collocation-based model
CM	(Standard) compartment model
CP	collocation point
DAE	differential algebraic equation
DSM	demand-side management
FMU	functional mock-up unit
FSM	full-order stage-by-stage model
(e)MPC	(economic) model predictive control
NLP	Nonlinear programming
ODE	ordinary differential equation
OPC	optimal control problem
RMSE	root mean square error
SQP	sequential quadratic programming

Index sets

C	compartments
i	column stages
j	components

Parameters

C_E	electricity cost
k_d	hydraulic proportional factor
t_f	final time

Variables

F_{LIQ}	liquefied feed air
F_{MAC}	total feed air flow through main air compressor
F_{TANK}	stored and liquefied product
F_{TOP}	top product of rectification column
h^L	molar enthalpy of liquid phase
h^V	molar enthalpy of vapor phase
M	molar mass
L	liquid stream leaving a column stage
p	pressure of a column stage
P_{LIQ}	net power consumption of liquifier
P_{MAC}	net power consumption of main air compressor
P_{TURB}	net power production of expansion turbine
T	temperature of a column stage
$T^{\mathrm{in}}_{\mathrm{TURB}}$	temperature of feed air at turbine inlet
$T^{\mathrm{out}}_{\mathrm{PHX2}}$	temperature of feed air behind heat exchanger
V_{IRC}	relative volume of integrated reboiler and condenser
V_{TANK}	relative volume of liquid product tank
$\hat{u}$	generalized compartment input
V	vapor stream leaving a column stage
x	liquid molar fraction
$\hat{x}$	generalized differential compartment state
y	vapor molar fraction
$\hat{y}$	generalized compartment output

$\hat{z}$	generalized algebraic compartment state
ϕ	objective function (electricity costs)
ΔT_{IRC}	temperature difference in integrated reboiler and condenser
$\Delta T_{PHX1}^{N_2,in}$	temperature difference between product and feed air (zone 1 inlet)
$\Delta T_{PHX1}^{WASTE,in}$	temperature difference between waste and feed air (zone 1 inlet)
$\Delta T_{PHX1}^{N_2,out}$	temperature difference between product and feed air (zone 1 outlet)
$\Delta T_{PHX1}^{WASTE,out}$	temperature difference between waste and feed air (zone 1 outlet)
$\Delta T_{PHX2}^{N_2,out}$	temperature difference between product and feed air (zone 2 outlet)
$\Delta T_{PHX2}^{WASTE,out}$	temperature difference between waste and feed air (zone 2 outlet)

ORCID

Pascal Schäfer https://orcid.org/0000-0002-3268-8976
Adel Mhamdi https://orcid.org/0000-0003-0192-0864
Alexander Mitsos https://orcid.org/0000-0003-0335-6566

REFERENCES

1. Mitsos A, Asprion N, Floudas CA, et al. Challenges in process optimization for new feedstocks and energy sources. *Comput Chem Eng*. 2018;113:209-221.
2. Zhang Q, Grossmann IE. Enterprise-wide optimization for industrial demand side management: fundamentals, advances, and perspectives. *Chem Eng Res Des*. 2016;116:114-131.
3. Ghobeity A, Mitsos A. Optimal time-dependent operation of seawater reverse osmosis. *Desalination*. 2010;263:76-88.
4. Brée L, Perrey K, Bulan A, Mitsos A. Demand side management and operational mode switching in chlorine production. *AIChE J*. 2018. https://doi.org/10.1002/aic.16352
5. Schäfer P, Westerholt H, Schweidtmann AM, Ilieva S, Mitsos A. Model-based bidding strategies on the primary balancing market for energy-intense processes. *Comput Chem Eng*. 2019;120:4-14.
6. Miller J, Luyben WL, Belanger P, Blouin S, Megan L. Improving agility of cryogenic air separation plants. *Indus Eng Chem Res*. 2008;47:394-404.
7. Mitra S, Grossmann IE, Pinto JM, Arora N. Optimal production planning under time-sensitive electricity prices for continuous power-intensive processes. *Comput Chem Eng*. 2012;38:171-184.
8. Mitra S, Pinto JM, Grossmann IE. Optimal multi-scale capacity planning for power-intensive continuous processes under time-sensitive electricity prices and demand uncertainty. Part I: modeling. *Comput Chem Eng*. 2014;65:89-101.
9. Zhang Q, Grossmann IE, Heuberger CF, Sundaramoorthy A, Pinto JM. Air separation with cryogenic energy storage: optimal scheduling considering electric energy and reserve markets. *AIChE J*. 2015;61:1547-1558.
10. Zhang Q, Sundaramoorthy A, Grossmann IE, Pinto JM. A discrete-time scheduling model for continuous power-intensive process networks with various power contracts. *Comput Chem Eng*. 2016;84:382-393.
11. Pattison RC, Touretzky CR, Johansson T, Harjunkoski I, Baldea M. Optimal process operations in fast-changing electricity markets: Framework for scheduling with low-order dynamic models and an air separation application. *Indus Eng Chem Res*. 2016;55:4562-4,584.
12. Kelley MT, Pattison RC, Baldick R, Baldea M. An MILP framework for optimizing demand response operation of air separation units. *Appl Energy*. 2018;222:951-966.
13. Pattison RC, Touretzky CR, Harjunkoski I, Baldea M. Moving horizon closed-loop production scheduling using dynamic process models. *AIChE J*. 2017;63:639-651.
14. Dias LS, Pattison RC, Tsay C, Baldea M, Ierapetritou MG. A simulation-based optimization framework for integrating scheduling and model predictive control, and its application to air separation units. *Comput Chem Eng*. 2018;113:139-151.
15. Jamaludin MZ, Swartz CLE. Dynamic real-time optimization with closed-loop prediction. *AIChE J*. 2017;63:3896-3,911.
16. Li H, Swartz CLE. Dynamic real-time optimization of distributed MPC systems using rigorous closed-loop prediction. *Comput Chem Eng*. 2018. https://doi.org/10.1016/j.compchemeng.2018.08.028
17. Baldea M, Harjunkoski I. Integrated production scheduling and process control: A systematic review. *Comput Chem Eng*. 2014;71:377-390.
18. Ellis M, Durand H, Christofides PD. A tutorial review of economic model predictive control methods. *J Process Control*. 2014;24:1156-1,178.
19. Huang R, Zavala VM, Biegler LT. Advanced step nonlinear model predictive control for air separation units. *J Process Control*. 2009;19:678-685.
20. Huang R, Biegler LT. Economic NMPC for energy intensive applications with electricity price prediction. *Comput Aided Chem Eng*. 2012;31:1612-1616.
21. Caspari A, Faust JMM, Schäfer P, Mhamdi A, Mitsos A. Economic nonlinear model predictive control for flexible operation of air separation units. *IFAC-PapersOnLine*. 2018;51:295-300.
22. Schäfer P, Bering LF, Caspari A, Mhamdi A, Mitsos A. Nonlinear dynamic optimization for improved load-shifting agility of cryogenic air separation plants. *Comput Aided Chem Eng*. 2018;44:547–552.
23. Cao Y, Swartz CLE, Flores-Cerrillo J, Ma J. Dynamic modeling and collocation-based model reduction of cryogenic air separation units. *AIChE J*. 2016;62:1602-1,615.
24. Cho YS, Joseph B. Reduced-order steady-state and dynamic models for separation processes. Part II. Application to nonlinear multicomponent systems. *AIChE J*. 1983;29:270-276.
25. Stewart WE, Levien KL, Morari M. Simulation of fractionation by orthogonal collocation. *Chem Eng Sci*. 1985;40:409-421.
26. Seferlis P, Hrymak AN. Optimization of distillation units using collocation models. *AIChE J*. 1994;40:813-825.
27. Cao Y, Swartz CLE, Flores-Cerrillo J. Optimal dynamic operation of a high-purity air separation plant under varying market conditions. *Indus Eng Chem Res*. 2016;55:9956-9,970.
28. Cao Y, Swartz CLE, Flores-Cerrillo J. Preemptive dynamic operation of cryogenic air separation units. *AIChE J*. 2017;63:3845-3,859.
29. Benallou A, Seborg DE, Mellichamp DA. Dynamic compartmental models for separation processes. *AIChE J*. 1986;32:1067-1,078.
30. Horton RR, Bequette BW, Edgar TF. Improvements in dynamic compartmental modeling for distillation. *Comput Chem Eng*. 1991;15:197-201.
31. Lévine J, Rouchon P. Quality control of binary distillation columns via nonlinear aggregated models. *Automatica*. 1991;27:463-480.
32. Zhu GY, Henson MA, Megan L. Low-order dynamic modeling of cryogenic distillation columns based on nonlinear wave phenomenon. *Sep Purif Technol*. 2001;24:467-487.
33. Cao Y, Swartz CLE, Baldea M, Blouin S. Optimization-based assessment of design limitations to air separation plant agility in demand response scenarios. *J Process Control*. 2015;33:37-48.

34. Roffel B, Betlem BHL, Ruijter JAF. First principles dynamic modeling and multivariable control of a cryogenic distillation process. *Comput Chem Eng*. 2000;24:111-123.
35. Bian S, Khowinij S, Henson MA, Belanger P, Megan L. Compartmental modeling of high purity air separation columns. *Comput Chem Eng*. 2005;29:2096-2,109.
36. Marquardt W. Nonlinear model reduction for binary distillation. In: McGreavy C, ed. *Dynamics and Control of Chemical Reactors and Distillation Columns*. Oxford: Pergamon:123-128 1988 IFAC Symposia Series. https://doi.org/10.1016/B978-0-08-034917-6.50023-4
37. Kienle A. Low-order dynamic models for ideal multicomponent distillation processes using nonlinear wave propagation theory. *Chem Eng Sci*. 2000;55:1817-1828.
38. Liu X, Zhou Y, Cong L, Zhang J. Nonlinear wave modeling and dynamic analysis of internal thermally coupled distillation columns. *AIChE J*. 2012;58:1146-1156.
39. Fu Y, Liu X. An advanced control of heat integrated air separation column based on simplified wave model. *J Process Control*. 2017;49:45-55.
40. Khowinij S, Henson MA, Belanger P, Megan L. Dynamic compartmental modeling of nitrogen purification columns. *Sep Purif Technol*. 2005;46:95-109.
41. Chen Z, Henson MA, Belanger P, Megan L. Nonlinear model predictive control of high purity distillation columns for cryogenic air separation. *IEEE Trans Control Syst Technol*. 2010;18:811-821.
42. Carta R, Tola G, Servida A, Morbidelli M. Performance of collocation models for simulating transient multistage separation units. *Comput Chem Eng*. 1995;19:1141-1,151.
43. Linhart A, Skogestad S. Computational performance of aggregated distillation models. *Comput Chem Eng*. 2009;33:296-308.
44. Linhart A, Skogestad S. Reduced distillation models via stage aggregation. *Comput Chem Eng*. 2010;65:3439-3,456.
45. Sharma R, Singhal D, Ghosh R, Dwivedi A. Potential applications of artificial neural networks to thermodynamics: vapor–liquid equilibrium predictions. *Comput Chem Eng*. 1999;23:385-390.
46. Chouai A, Laugier S, Richon D. Modeling of thermodynamic properties using neural networks: application to refrigerants. *Fluid Phase Equilibria*. 2002;199:53-62.
47. Nentwich C, Engell S. Application of surrogate models for the optimization and design of chemical processes. In: Proceedings from the 2016 International Joint Conference on Neural Networks (IJCNN): 1291–1296IEEE; 2016.
48. Schweidtmann AM, Huster WR, Lüthje JT, Mitsos A. Deterministic global process optimization: Accurate (single-species) properties via artificial neural networks. *Comput Chem Eng*. 2019;121:67-74.
49. Bishop CM. *Pattern recognition and machine learning (information science and statistics)*. Berlin, Heidelberg: Springer-Verlag; 2006.
50. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521:436-444.
51. Haykin SS. *Neural networks and learning machines*. 3rd ed. Upper Saddle River, NJ: Pearson Education; 2009.
52. Schweidtmann AM, Mitsos A. Deterministic global optimization with artificial neural networks embedded. *J Optim Theory Appl*. 2019;180:925–948.
53. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. *J Mac Learn Res*. 2014;15:1929-1958.
54. McKay MD, Beckman RJ, Conover WJ. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*. 1979;21:239-245.
55. Caspari A, Bremen AM, Faust JMM, et al. DyOS - a framework for optimization of large-scale differential algebraic equation systems. Comput Aided Chem Eng (corresponding to the proceedings of the 29th european symposium on computer aided process engineering). 2018.
56. Sargent RWH, Sullivan GR. The development of an efficient optimal control package. In: Stoer J, ed. *Optimization Techniques*. Berlin, Heidelberg: Springer 1978:158-168.
57. Kraft D. On converting optimal control problems into nonlinear programming problems. In: Klaus S, ed. *Computational mathematical programming*. Berlin, Heidelberg: Springer; 1985:261-280.
58. Gill P, Murray W, Saunders M. SNOPT: an SQP algorithm for large-scale constrained optimization. *SIAM Rev*. 2005;47:99-131.
59. Schlegel M, Marquardt W, Ehrig R, Nowak U. Sensitivity analysis of linearly-implicit differential-algebraic systems by one-step extrapolation. *Appl Numer Math*. 2004;48:83-102.
60. Hannemann R, Marquardt W, Naumann U, Gendler B. Discrete first- and second-order adjoints and automatic differentiation for the sensitivity analysis of dynamic models. *Proc Comput Sci*. 2010;1:297-305. ICCS 2010.
61. Cuthrell JE, Biegler LT. On the optimization of differential-algebraic process systems. *AIChE J*. 1987;33:1257-1,270.
62. Würth L, Hannemann R, Marquardt W. Neighboring-extremal updates for nonlinear model-predictive control and dynamic real-time optimization. *J Process Control*. 2009;19:1277-1,288.
63. Zavala VM, Biegler LT. The advanced-step NMPC controller: optimality, stability and robustness. *Automatica*. 2009;45:86-93.
64. Yang X, Biegler LT. Advanced-multi-step nonlinear model predictive control. *J Process Control*. 2013;23:1116-1,128.

SUPPORTING INFORMATION

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

How to cite this article: Schäfer P, Caspari A, Kleinhans K, Mhamdi A, Mitsos A. Reduced dynamic modeling approach for rectification columns based on compartmentalization and artificial neural networks. *AIChE J*. 2019;65:e16568. https://doi.org/10.1002/aic.16568