Hybrid Semi-parametric Modeling in Separation Processes: A Review

Kevin McBride[1], Edgar Ivan Sanchez Medina[2], and Kai Sundmacher[1,2,*]

DOI: 10.1002/cite.202000025

Separations of mixtures play a critical role in chemical industries. Over the last century, the knowledge in the area of chemical thermodynamics and modeling of separation processes has been substantially expanded. Since the models are still not completely accurate, hybrid models can be used as an alternative that allows to retain existing knowledge and augment it using data. This paper explores some of the weaknesses in the current knowledge in separations design, simulation, optimization, and operation, and presents many examples where data-driven and hybrid models have been used to facilitate these tasks.

Keywords: Chemical separation, Hybrid modeling, Machine learning, Thermodynamics

Received: February 26, 2020; *accepted:* April 15, 2020

1 Introduction

If all the different separations performed throughout human history are considered, several important examples like the separation and purification of metals from earth, the production of salt from sea water, capturing alcohol by distillation, and the cultivation of dyes from plant matter are found [1]. Each one of these has had a significant effect on not only technological progress, but also on the course of history. These acts of taking a mixture of substances and separating them result as a consequence of nature's tendency to increase in entropy. Much of the technological development as a species has been motivated by a drive to confront this chaos in an attempt to increase order. Perhaps the most relevant examples today are found in oil industry where refineries use large distillation sequences to separate crude oil into many different products used throughout all levels of the economy. There is no doubt that the separation of chemical substances or raw materials has long played an important role in the advancement of human species and is largely responsible for our modern standard of living.

The aim of chemical separation is to achieve the purification or enrichment of one or more substances from a mixture of several other substances. This is applied to all forms of conventional matter: gas, liquid, and solid. Whereas our ancestral forefathers labored under conditions on a much smaller scale, the modern chemical engineer of the industrial era is interested in performing these separations at a larger scale and in a more economically feasible manner. Many common and ancient separation methods have been adapted for use at this scale, such as distillation, adsorption, liquid-liquid extraction, and crystallization, especially in the previous century. During this time, chemical engineers have amassed a vast body of knowledge concerning chemical separations including thermodynamic information and theory, pure-component properties, and other relevant physical properties required to design and operate these processes.

Despite the great progress made, there is a lack of fundamental understanding in many areas that often precludes pursuing the optimal separation strategy or realizing that another separation route may be possible. This is important in modern chemical plants, where 60 to 80 % of costs originate in separations [2]. When looking for avenues to decrease operational costs, it would seem natural to want to reduce the expenditures for separations. But the question is, why other possibilities are not simply considered. A simple reason might be the lack of physical property data for a significant component in the mixture. Perhaps the theory used to design a planned unit operation is insufficient to describe the real behavior of the system leading to suboptimal performance, not to mention the difficulties that may arise in operating such a unit. Sometimes this lack of understanding leads the designers to fall back to easier decisions, such as using distillation for a separation. And why not, it is a well-known technology with quick and accurate models being readily available. This does not come as a surprise since the more mature a technology is, the more often it is used. Fig. 1 shows the general trend between how well-known a

[1]Dr.-Ing. Kevin McBride, Prof. Dr.-Ing. Kai Sundmacher
sundmacher@mpi-magdeburg.mpg.de
Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, 39106 Magdeburg, Germany.

[2]Edgar Ivan Sanchez Medina, Prof. Dr.-Ing. Kai Sundmacher
Otto-von-Guericke University Magdeburg, Chair for Process Systems Engineering, Universitätsplatz 2, 39106 Magdeburg, Germany.

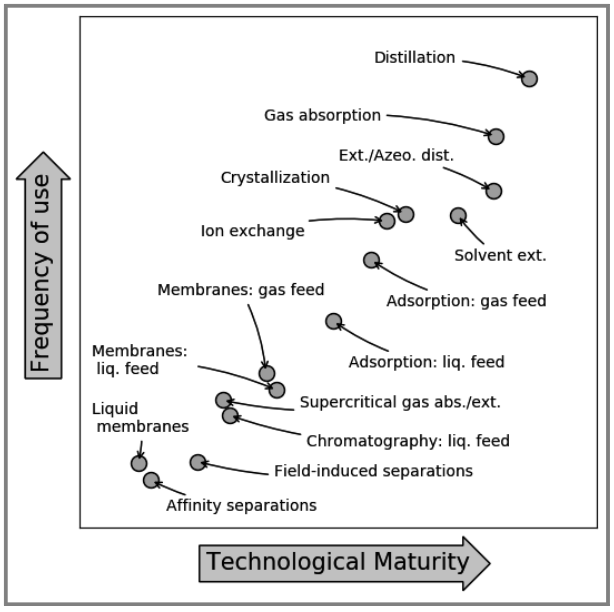

Figure 1. Correlation between separation technology maturity and frequency of use. Adapted from [1].

technology is and how often it is employed. If it is truly the intention to design better separation units and to possibly move away from the current reliance on distillation, the focus has to be on gaining more knowledge in order to more confidently design alternative separation technologies.

This begs the question as to which form of knowledge is best suited for this task. One compelling solution is to use process data to compensate for the lack of system information. For this reason, there is now a push to include separation technologies within the data-driven modeling paradigm. Closely related to this is the use of hybrid models, which are a combination of first principles and data-driven models, another area quickly gaining importance and popularity [3].

This contribution examines the role data and modeling play in the future of separation technologies, continuing a discussion started at the 58th Tutzing Symposium on this topic. In the first part, several open issues in the design of separators and the role of chemical thermodynamic models are introduced and briefly discussed. In the second half of the paper, several methods being proposed to handle these issues using data are reviewed. Finally, the paper is concluded with a recommendation and a hope that further discussions about where future efforts can be applied in the big data era in separations technology, or fittingly Separations 4.0, will continue.

2 Current Challenges in Separation Processes

The most important aspect in designing a separation unit operation is knowing the phase behavior. Without the generation of more than one phase, there can be no separation process [2]. The phase behavior is based on the chemical thermodynamics, which can be predicted using an equation of state (EoS) or an excess Gibbs energy (g^E) model. Many such models have been developed in the last century [4], starting with the van der Waals equation, but there is still no model that is generally applicable to all chemical components and mixtures. This means that the decision about which model to use is critical; the model must be accurate enough to adequately describe the phase behavior based on the measurable characteristics of the system. If not, it may be predicted that no separation should occur when in reality it does, or the reverse, where a feasible separation is predicted for a system that exhibits no such behavior.

However, it is often not the choice of the thermodynamic model that is the most important aspect, but instead the quality and nature of the experimental data available to fit the model parameters. The problem here is that much of the data required to accurately model all of the possible chemical systems has not been collected (Fig. 2). In fact, according to a rough estimate by [2], there is only about 1.2 % of the necessary vapor-liquid equilibria (VLE) data available for all binary mixtures of interest. It would be unreasonable to undertake the task of collecting all of the missing data. For this reason, much of the effort in thermodynamic modeling has focused on creating predictive methods. These are often group contribution methods because of the lower amount of experimental data needed to fit group-based parameters compared to the fitting of molecule-specific parameters more commonly found in EoS. The advantage of using group contribution methods is that the molecules can be derived from constituent groups, allowing a small subset of groups to represent the properties of a much larger set of molecules. However, even state-of-the-art group contribution methods like modified UNIFAC (UNIQUAC functional-group activity coefficients) Dortmund [5] still lack many of the binary interaction parameters because the experimental data is simply not there.

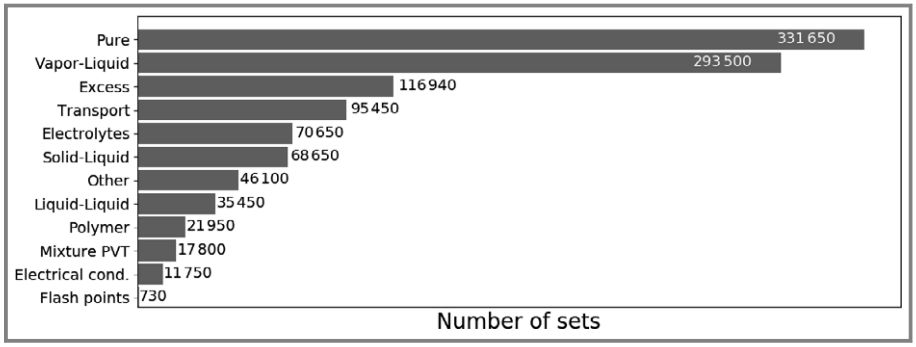

Figure 2. Thermodynamic data available in the Dortmund Data Bank 2019 [6].

Having experimental data is only one part of the equation; the model parameters need to be appropriately fit. This involves solving an optimization problem, which requires a carefully selected objective function to ensure reasonable results are obtained. Usually it is not enough to rely on a single class of data, such as concentration or pressure measurements, but also excess enthalpy and activity coefficients at infinite dilution [2]. Care must also be taken when evaluating the solution to the parameter fitting: it may be that only a local optimum has been found instead of the global optimum. This may lead to a suboptimally fit model that produces erroneous predictions and shows unsuspected behaviors [7].

Additionally, many of the experiments performed to collect the thermodynamic data needed to fit the parameters do not follow any experimental plan. There exist often very large confidence intervals or strong correlations between parameters. This makes the identification of parameters more difficult and leads to calibrations that are not adequate for use in accurate process designs and operation. With as much knowledge about the design of experiments that has been collected in the last century, its implementation should be standard [8]. Some authors are moving exactly in this direction, e.g., Dechambre et al. [9] recently presented a new method for designing experiments to improve the parameter estimation for the non-random two-liquid model (NRTL) and universal quasichemical (UNIQUAC) models used in liquid-liquid equilibria (LLE) predictions. An increase in such efforts could potentially reduce the number of experiments while increasing the quality of the information obtained. This in turn will lead to better parameters. An alternative is the introduction of high-throughput experiments for collecting thermodynamic data. Some examples are the recent use of Raman spectroscopy [10] or the use of bench-top nuclear magnetic resonance (NMR) devices [11] for measuring LLE. Successful industrial examples can also be found in the contributions by the research and development department at Dow Chemical Company [12–14].

These modern methods to increase the rate at which new data is collected is paramount. The strong focus on distillation has led to a proportional concentration of VLE data (it was already mentioned that even here only a small fraction of data exists). Thus, it is also of no surprise that there are large deficits in the needed experimental data for other, potentially beneficial, separation technologies such as micellar systems, supercritical fluids, ionic liquids, or systems containing complex biomolecules. Not only is there not enough data available for these substance classes, there are also very few models, if at all, that can be used to predict their thermodynamic behavior. For many of these systems, the use of predictive methods like conductor-like screening model for real solvents (COSMO-RS) [15] or UNIFAC-ionic liquids (UNIFAC-IL) [16] becomes necessary to find a starting point when considering their use in a new separation design.

Although critically important, the thermodynamic equilibrium is not the only criteria that influences a separation process. Separations are achieved by enhancing the rate of mass transfer of all species by bulk movement within a particular phase. This separation is constrained by the mass transfer and the extent is controlled by thermodynamic equilibrium. Other physical influences, such as fluid mechanics and heat transfer, also play an important role. At the end, the extent of the separation between chemical species depends on the ability to exploit the differences between the distinct product phases. This includes the molecular, thermodynamic, and transport properties of the individual components.

Therefore, the design of a unit operation for separation is necessarily part of a multi-scale model, such as the one seen in Fig. 3. In the molecular level, molecular dynamics, statistical mechanics, and density functional theory are commonly applied. While these models can usually describe the molecular level well enough, the use of these models is not sufficient for process design. Therefore, phase equilibrium models, rate-based models, and other constitutive equations need to be applied due to the influence of the bulk substance. The next level contains models for energy and mass balances as well as other phenomena like solute partitioning. The culmination of these models spanning across many levels is the final unit operation model.

It is quite common to make several assumptions to decrease the complexity of this multi-scale problem by assuming that the mixtures are homogeneous and that the process is at steady state. As a result, models for stationary

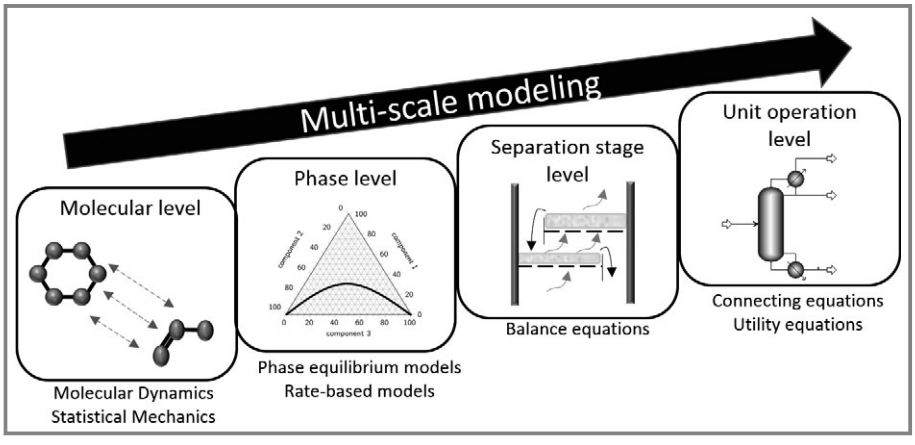

Figure 3. Different levels implicated in the modeling of separation units.

systems are much better developed than those describing dynamic systems (just consider equilibrium models compared to rate-based models). This is commonly seen in the modeling of separation processes where the design often hinges on the simple determination of the phase equilibrium. Two prominent examples of where this assumption fails are found in the modeling of the absorption of hydrogen chloride from exhaust gas using water as an extraction solvent [2] and in the distillation of the ternary system composed of water, ethanol, and acetone [17]. In the prior, it is predicted that after a single stage, a satisfactory separation at equilibrium is achieved. It turns out that large columns are needed because the mass transfer between the desired phases is very slow and equilibrium is not reached in practice. In the latter, if MESH (material balances, equilibrium relationships, summation equations, and heat balances) equations are used, the designed distillation column needs 25 stages to achieve the desired purity; by contrast, if the mass transfer is considered, 39 stages are needed to perform the separation. Usually an efficiency factor is used to account for the inaccuracies or ignored effects not covered in the equilibrium model, e.g., tray efficiency in distillation. Perhaps a better solution to the multi-scale problem is to use surrogate models instead of simplifications. The surrogate models serve to bridge the gap between levels and, thereby, simplify the problem without losing too much information [18].

Surrogates can also be used to replace the complex chemical thermodynamics for use in process simulation or optimization to obtain a reasonable trade-off between accuracy and computational expediency. The constitutive equations used to describe chemical thermodynamics and equilibrium are often complex, highly nonlinear equations that need to be solved implicitly. This does not often conform well to the standard optimization algorithms used in process optimization. The user may experience failure to converge to a feasible solution, experience long calculation times, and if, however, successful in creating a useful model, it may not be suitable for real-time optimization (RTO) or use in nonlinear model predictive control (NMPC) due to the long calculation time required. Not only are the equations difficult to converge, they may have several local optima as well, which may also lead to poor performance.

The quality of a simulation or design process also depends heavily on the prediction accuracy of the chosen thermodynamic models. One area that is particularly sensitive to errors in thermodynamic predictions is in distillation when the relative volatility of the key components is close to 1. In one glaring example from [2], a small deviation in the relative volatility leads to a much larger error in the height of the column, i.e., a 5 % error in the relative volatility around 1.10 leads to over 100 % error in the number of minimum trays. This is not a small discrepancy and reinforces the need for accurate predictions, in this case vapor pressure. Compounding this is the difficulty in estimating vapor pressure, compelling Gmehling et al. [2] to state that "estimation of vapor pressure is one of the most difficult problems in thermodynamics".

Perhaps even more problematic are predictions of LLE. Here, g^E models are used to predict activity coefficients due to their easier implementation than EoS and that usually pressure is not considered in LLE calculations. Currently, there is no method that can be used to reliably predict LLE [2]. The problem with these models lies in the fact that the activity coefficient must handle both the composition and temperature dependencies of the system. Similar to the problem with mixing rules used in EoS, LLE predictions start to deviate substantially when more than two components are present in a mixture in significant quantities to effect the equilibrium. Temperature dependencies can be very difficult to model, especially if measurements have not been conducted at various temperatures, and the representation of composition effects may not be as straightforward as many models assume [19]. One interesting example to consider is the system of tetrahydrofuran, water, and phenol, for which no g^E model can explain its phase behavior [20].

Additionally, there may be other phenomena that are not adequately described, such as foaming, fouling, or the influence of impurities. A correct description of these influences is necessary for an accurate design and operation of a real process. At the end of the spectrum, models that describe process behaviors in the presence of failures are almost completely lacking. The use of these models would allow a process simulator to converge to a solution (albeit a nonfunctioning process) instead of leading to convergence error. This additional information could lead to safer designs as well as provide more insight into the appropriate parameters for process control and operation.

In the big-data era, there are now two major options for improving the prediction accuracy of thermodynamic equations. The first avenue is continuing to increase the level of knowledge about chemical systems, so that better theories and thereby models can be produced. It would be desirable to find a unique EoS with accurate and broadly applicable mixing rules that describes all types of mixtures. Such a model would provide perfect values for all thermodynamic properties of pure components and mixtures. This is still far away, but many advancements have been made in this area in the last couple of decades, e.g., volume-translated Peng-Robinson (VTPR) [21], perturbed chain statistical associating fluid theory (PC-SAFT) [22], and variants. The second avenue is actively improving the existing models by incorporating more data in the form of data-driven and hybrid models. Thus, the remainder of this paper examines recent advances in the application of these models to separation processes.

3 Data Sources in Separations

The world is facing an enormous increase in the amount of data available in several fields. The data sources available in

chemical engineering can be classified into five major groups: 1) plant data, 2) laboratory data, 3) simulation data, 4) literature data, and 5) business data (Fig. 4).

The amount of data available during the operation of a chemical plant is extremely large. The chemical process industry has been progressively adopting machine learning tools to explore the information available in the data [23]. A recent review discussed that the relevance that data has in the chemical industry is likely to increase even more in the future [24]. However, most of this data (usually obtained by sensors) is highly correlated. Or, most of the data is collected from continuous processes at steady state and is not very useful for determining process dynamics. Therefore, some filtering or pre-processing of the raw data is paramount before it can be used for modeling. In fact, data-driven models have been applied successfully in the industry for process monitoring [13, 25–28], inferential sensors [29], and advanced control schemes [30].

Unlike data from chemical plants, the data obtained from laboratory experiments are characterized by better control and systematic management. Highly correlated data can be avoided using optimal experimental design techniques. However, the velocity of data generation in traditional laboratories is significantly slower compared to the plant case. As previously mentioned, high-throughput digitalized laboratories present a viable way to tackle this problem.

Simulation data can be generated in large amounts due to the continuously increasing computational power. More reliable and complex models can be solved more quickly, which opens a new data source for the training of surrogate models. The uncertainty found in this type of data is typically caused by numerical noise, which can be controlled more easily compared to noisy data from real-world systems. Many examples of the use of simulation data for surrogate training are available in the literature [31, 32]. For simulation data, it can be distinguished between two types of sampling techniques: one-shot (or stationary) and adaptive (or sequential) designs. The former refers to a predetermined space-filling strategy, e.g., Latin hypercube sampling [33], Sobol [34] and Halton [35] sequences. The latter approach starts from a simple one-shot design, which is iteratively augmented by employing exploration-exploitation methods, e.g., expected improvement function [36], bumpiness function [37], LOLO-Voronoi [38], smart sampling algorithm [39], and mixed adaptive sequential sampling [40]. It has been shown that, most of the times, adaptive sampling outperforms one-shot strategies in getting similar level of accuracy with less samples [38].

Another rich source of data comes from experimental results published in the scientific literature. In many applications, access to this type of data can save considerable amounts of time and resources for on-going research. However, some issues here may be that data is lacking for some components or mixtures of interest, or that the experimental conditions differ substantially such that the validity range is restricted to a narrower domain than desirable. Conveniently, a great amount of the thermodynamic data found in the literature has been collected and made available in several databases, such as the Dortmund Data Bank [6] and the NIST database [41].

Regarding business data, the reader is referred to the work of Chiang et al. [24], which reviews several successful examples of the integration of planning and scheduling tasks with process control by using this information through surrogate models.

4 Data-Driven Models in Separation Processes

Data-driven models have been used in the field of separation processes as a tool to overcome the lack of knowledge of the system or to alleviate the computational burden of expensive calculations. Among the data-driven surrogates used in separation process applications, the most common are artificial neural network (ANN), partial least squares

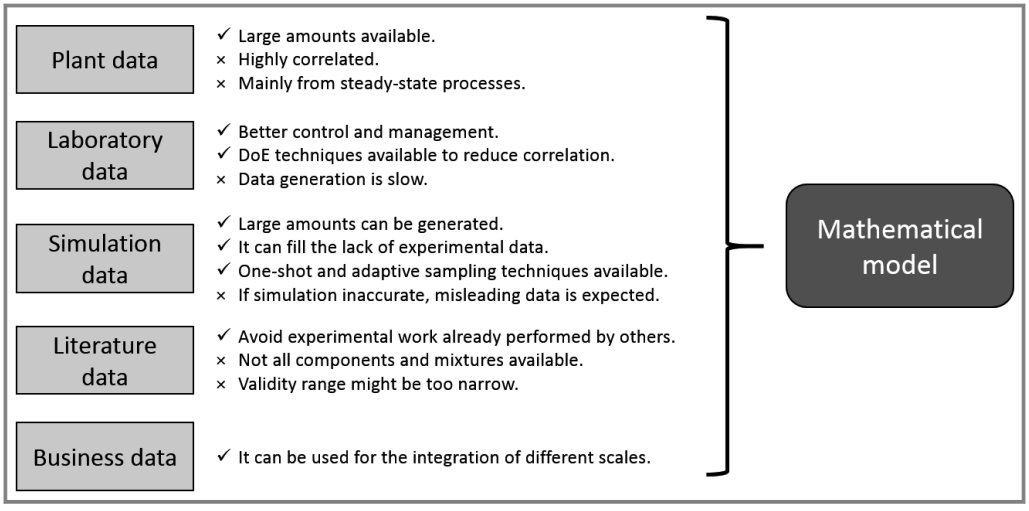

Figure 4. Data sources for separation process modeling and their advantages (✓) and disadvantages (×).

(PLS), Gaussian process (GP), and polynomials for regression purposes and support vector machine (SVM) for classification. Other techniques such as principal component analysis (PCA) for dimensionality reduction and fuzzy logic (FL) for the inclusion of heuristic knowledge have been also applied. Recent reviews on surrogate modeling within chemical engineering can be found in the literature [42, 43].

Data-driven models act as an input/output relationship, which has several advantages and disadvantages compared to pure mechanistic models (Tab. 1). On the one hand, data-driven models do not require expert knowledge to be developed, which make them very flexible. Moreover, this type of models usually has a fast evaluation time, which is desirable for optimization. Because they are significantly cheaper than mechanistic models, data-driven models can replace complex parametric models alleviating the computational burden of the entire model. On the other hand, in order to build this type of model, enough data has to be available, which is not the case in every application. Besides, data-driven models are known to be poor extrapolators, which limit their validity to the domain in which the model was trained. Given that the data-driven model acts as a black box, very little understanding of the system under consideration can be attained.

Table 1. Advantages and disadvantages of data-driven models compared to purely mechanistic models.

Advantages	Disadvantages
Expert knowledge is not needed for development.	Enough high-quality data has to be available to build an accurate model.
Less computationally expensive.	Poor extrapolation capabilities.
	Very little abstract understanding is achieved.

The selection of the surrogate form depends on the goal of the model and the available information of the system. This task commonly relies only on the modeler criterion, given that clear-cut guidelines do not exist for this, but some general advice can be found in the work of Forrester et al. [44].

Automatic surrogate generation approaches can also be found in the literature. The surrogate modeling (SUMO) Matlab toolbox [45] builds a set of candidate structures, which are then discriminated using techniques such as cross-validation and Akaike information criteria. Another example is the automatic learning of algebraic models for optimization (ALAMO) framework [46], which relies on adaptive sampling techniques and basis functions to construct a surrogate which aims to balance the trade-off between accuracy and complexity. ALAMO was recently extended to restrict the output space of the surrogate by allowing the implementation of physical knowledge in the form of constraints [47]. The performance of this framework is shown in several chemical engineering applications in [47, 48]. Algorithms for global optimization of constrained gray-box computational problems (ARGONAUT) follow similar ideas in the context of global optimization of combined data-driven and first-principle models [49–52].

Despite the selected form for the surrogate model, its performance relies mostly on the data quality (which depends on the source) and quantity. While the quantity of the data is restricted by the available measurement devices, the quality of the data depends heavily on the process conditions under which the plant is operated. Ideally, the lowest number of high-quality samples (that provides enough information about the structure of the true function) possible has to be gotten to ensure the high accuracy and low cost of the surrogate.

Purely data-driven surrogate models have been applied to separation process applications for either computational burden reduction or modeling complex phenomena. For instance, the optimization of pressure swing adsorption systems has been performed employing ANNs as surrogates [53–55]. The mechanistic models that describe this type of units consist on full partial differential algebraic equations (PDAE) systems that are computationally expensive. However, by employing ANNs, the optimization time has been reduced from several hours to just a few seconds, showing the benefits that data-driven models have in the optimization of complex systems if appropriately formulated. Moreover, convolutional neural networks and deep autoencoders have been used as a fault diagnosis tool for distillation processes [31]. Fault diagnosis in distillation is very complex to model using mechanistic models and in practice most of the time relies entirely on the operator experience. Nevertheless, by using data-driven models, it is possible to construct models that help decision makers on such complex tasks.

Despite the benefits of using data-driven models in separation processes, one should not solely rely upon them and discard all process knowledge accrued over the years. The reason for this is that both approaches (mechanistic and data-driven) are complementary as the weaknesses of the one can be strengthened by the features of the other. This type of hybrid model has been shown to outperform purely mechanistic and purely data-driven models in several applications [56–59].

5 Overview of the Use of Hybrid Models in Separation Processes

5.1 Background

In Sect. 4, data-driven surrogates have been compared to pure mechanistic models showing several advantages and disadvantages (Tab. 1). Hybrid modeling attempts to combine the strengths of pure data-driven and pure mechanistic models in an effort of integrating all available information of the system. In fact, some highly accurate EoS can be considered hybrid models. The reason for this is that they use a combination of a first principles model to describe the ideal

behavior of the system combined with a data-driven model determined by selecting the best subset of terms from a larger set of possibilities using a stochastic optimization algorithm [2]. However, in this paper, the focus is on models that combine nonparametric and parametric sub-models. Models of this kind are called hybrid semi-parametric models (called simply hybrid models for briefness in this paper). The difference depends on the form of the data-driven sub-model, which can have a predefined number of parameters whose values need to be determined (e.g., linear regression) or a number of parameters that is unknown beforehand (non-parametric models, e.g., ANN).

As stated by Glassey and von Stosch [60], the high benefit-to-cost ratio of hybrid modeling while dealing with complex systems makes it a very promising approach. S. Zendehboudi et al. [61] provided a recent review on hybrid modeling applications in various areas including reaction engineering and separation processes. The advantages of hybrid models compared to their individual components can be summarized with the following three major groups.

- Structure: Within the hybrid model, the white box provides a base for understanding the physical system, while the black box part only replaces the phenomena in the system that it is expensive or not well understood. For this reason, hybrid models have better interpretability than black box models and the number of model parameters needed is less [62]. Moreover, given that the complex or unknown phenomena are covered by a data-driven surrogate, the model synthesis is less complex compared to pure mechanistic models [58, 63].
- Performance: Compared to both black box and white box models, it has been shown in the literature that, in most cases, hybrid models produce more accurate results both in steady and dynamic systems [56–59]. The reason for this is their better extrapolation capabilities compared to pure data-driven models and their better handling of unknown/complex features compared to white box models.
- Data dependency: Compared to a pure data-driven model, a hybrid model requires less data [64]. This is because the extent of the predictions that the data-driven model has to perform is reduced by the inclusion of mechanistic knowledge [62].

Several challenges of hybrid modeling can also be identified. For instance, the inclusion of fundamental knowledge into data-driven models constrains the feasible region of prediction [60]. Therefore, if this knowledge is incorrect, poorer predictions from the hybrid model are expected when compared to a pure data-driven model. In addition, the best combination of parameters for each sub-model still needs to be identified. This is a challenging task given that training techniques for hybrid models have additional challenges compared with the parameter identification of its individual sub-models [65].

The use of hybrid models has gained popularity in the past 20 years. In fact, as shown by Mohd Ali et al. [3], the use of pure data-driven models in chemical process systems has been decreasing. On the contrary, the number of publications on hybrid modeling has being rising steadily. They forecast that this trend will continue in the future.

The general approach for creating hybrid models for use in separation unit design follows those usually applied to hybrid models in that they can be arranged into serial or parallel arrangements [62]. Agarwal [66] provided a complete theoretical description of these possible arrangements for hybrid models. In the serial arrangement, the data-driven sub-model (black box) can be connected to the mechanistic sub-model (white box) either before or after it (Fig. 5). The former serial arrangement has encountered more applications in process engineering than the latter one [65]. On the other hand, the parallel arrangement has been applied whenever the mechanistic model is available, but, for some reason, it is not completely accurate, i.e., model mismatch [59, 67, 68].

5.2 Literature Review

Separations are important tasks in every chemical process. The design and operation of separators have been traditionally carried out using first-principle and empirical models. The accuracy of these models is of immense importance for the proper design, operation, and optimization of the separation tasks. Nevertheless, the physical phenomena involved in separation processes are highly complex. This makes the development of reliable mechanistic models very difficult. Hybrid models have been applied in different separation systems showing an overall improvement in the modeling capabilities compared to traditional methods. In the following, several applications of hybrid models used in separation processes are reviewed.

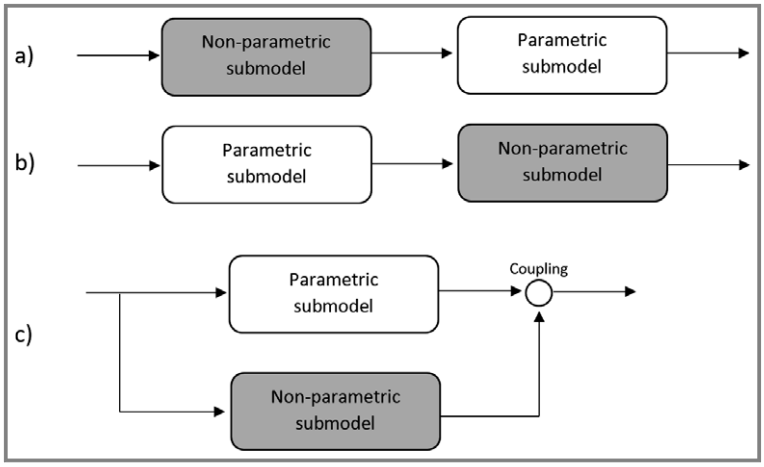

Figure 5. Hybrid semi-parametric model arrangements. a) Serial arrangement with non-parametric model first. b) Serial arrangement with parametric model first. c) Parallel arrangement of sub-models.

5.2.1 Distillation

Distillation is by far the most mature and commonly used separation technique in industry [1]. In a distillation column, the vapor and liquid phases come into contact on each stage, which facilitates the mass transfer between phases. The species are separated through their difference in volatility. This difference causes the separation of the components to be performed at dissimilar extents along the column, making the separation of components possible. The amount of energy consumed by distillation units is very significant, e.g., more than 20 % of the energy consumed in refineries are consumed by crude distillation units [69]. Therefore, due to their energetic relevance and popularity, the modeling of distillation columns has been of great interest in the past years.

Historically, distillation columns were initially designed using graphical methods, e.g., McCabe-Thiele or Ponchon-Savarit. However, with the increase of computational power, computer-aided methods started to gain popularity for column design [1]. Nevertheless, despite the computational power available nowadays, several complications for the design, control, and optimization of distillation units are still present. Subsequently, several of these limitations are discussed, along with some hybrid model applications found in the literature that try to address them.

Rigorous models required expensive phase equilibrium and mass transfer calculations, which results in very expensive models that cannot be easily integrated into an optimization framework (especially online optimization). Hybrid models have been applied with the aim of solving this problem and have achieved promising results. For instance, Safavi et al. [70] used wavelet neural networks to model the separation factor in a continuous binary distillation column. This surrogate separation factor model was integrated with the mass balances of the column, and it was shown that the computational cost of the model was reduced significantly. The hybrid model was then successfully integrated into an online optimization study. A similar approach was followed by Engell and Dadhe [71], who reduced the computational burden of a reactive distillation model by approximating the VLE with a radial basis function neural network, making their integration in an NMPC possible.

An interesting example that elucidates the capability of hybrid modeling to reduce computational costs was presented by Mahalec and Sanchez [72]. They applied PLS regression to approximate the true boiling point curve of an in-silico atmospheric pipe still column. This reduced the size of the entire rigorous model from around 10 000 nonlinear equations to around 100 almost linear equations maintaining similar levels of accuracy. Later, they optimized the resulting hybrid model showing that it converged to a better optimum than the rigorous mechanistic model (possibly due to the reduction of the nonlinearities in the model). This work was later extended by applying additional simplifications, which made the hybrid model independent from the tray temperatures (conserving the high accuracy) which made it suitable for RTO, planning and scheduling applications [73]. This last example reveals that hybrid models can be used to unify different scales (which traditionally are treated with independent models) within a single model. In this way, the mismatch between the individual models is eliminated. A recent review shows the importance of data for the integration of models among different scales [74]. Recent works have shown the capability of GP for global optimization of distillation columns [75, 76]. These tools can be easily implemented into hybrid schemes. Hybrid ANNs have been recently used for RTO in distillation columns [77].

Purely data-driven models ignore the extensive expert knowledge that has been gained in separation technology design and operation. This is not something seriously worth considering and underlines the importance of using other approaches, such as the hybrid modeling cases discussed here. In the operation of distillation columns, expert knowledge is especially useful during the start-up and shutdown stages. The nature of these operation modes is highly complex, which make their integration into a single mechanistic model very difficult. However, by using FL, the start-up behavior of a bath distillation column has been included into a single model that describes the complete operation cycle [78].

Process integration, e.g., reactive distillation, imposes additional complexity for the modeling of distillation units. Different types of ANNs have been used to replace expensive calculations in reactive distillation models [71, 79, 80]. Furthermore, in practice, composition measurement devices can be very expensive to purchase and maintain, and they usually have a time-delay response that prevents proper control of the distillation column. Khazraee and Jahanmiri [80] successfully applied an adaptive neuro-fuzzy inference system to establish a relationship between unmeasured variables (chemical composition) and measured variables (temperature) for control. ANNs have been also coupled with mechanistic models to solve the problem of missing data achieving robust and accurate models through data augmentation [81]. Further opportunities for hybrid model applications include fault diagnosis [31] and identification of more realistic columns [82].

5.2.2 Crystallization

Crystallization is a common separation technique primarily found in the pharmaceutical and food industries. Its use recounts to very old times, e.g., obtainment of sodium chloride, and is still broadly used for the production of many goods used in everyday life. This separation process occurs when crystalline solids are formed from a homogeneous fluid phase (usually a liquid phase), through the mechanisms of nucleation, growth, and agglomeration [1]. These mechanisms are highly complex, and many physical phenomena related to them are still poorly understood [83]. Further-

more, the estimation of product quality variables, e.g., average particle size, are often determined offline. This makes the modeling of crystallizers very complicated. Hybrid models have been used in both serial and parallel arrangements to compensate for this lack of knowledge.

For instance, sugar crystallization processes are known to be strongly nonlinear and nonstationary [63]. The complexity of these processes is captured by the growth, nucleation, and agglomeration kinetic rate expressions. ANN and SVM have been connected in serial arrangement to mass, energy, and population balances to describe industrial crystallizers [58, 63, 84]. The roll of the surrogates in these applications has been precisely to capture the behavior of the growth, nucleation, and agglomeration kinetics. It has been shown that hybrid models are more flexible (i.e., high-quality predictions at different operating conditions) than traditional modeling techniques (fitting empirical equations using nonlinear programming (NLP) optimization methods) and are significantly more consistent with experimental data [63, 84].

A common approach to overcome the complexity of modeling crystallizers is to neglect the agglomeration phenomena. However, it has been shown that this assumption leads to significant errors [58]. Georgieva et al. [58] removed this assumption by integrating an ANN to approximate the agglomeration kinetics along with the kinetics for growth and nucleation. The resulting hybrid model was compared to a fully mechanistic model, the former showing better agreement with the experiments. This work was later extended by showing its application to various control strategies to achieve an optimal supersaturation trajectory [85, 86]. In these last references, given that experimental values of the kinetic constants are not available, sensitivity equations were used to propagate the error signal back to the data-driven sub-model to achieve the training of the hybrid model.

PLS regression has been also used to improve the model performance of crystallization processes. This technique has been used in a parallel arrangement to integrate the residuals between the real system and the model by Hermanto et al. [68] and Zhang et al. [87] in control applications. It has been shown that by filtering the propagated error in the control of batch crystallizers (to remove possible noise caused by sporadic disturbances, i.e., only-one batch disturbances), the model generalization is improved significantly [87, 88]. Residuals in crystallizer models have been also integrated using fuzzy predictive adaptive resonance theory (ARTMAP) networks [89], and by employing self-learning systems, the model has been updated continuously to account for incrustation [90] resulting in better control strategies. Recently, soft sensors have been designed using Gaussian mixture regression for the online estimation of variables that are normally determined offline, e.g., average particle size [91].

5.2.3 Extraction

Extraction processes normally involve the use of an external liquid phase to remove, at least partially, one or more components from a mixture that can also be liquid (liquid-liquid extraction) or solid (solid-liquid extraction or leaching) [1]. Furthermore, some applications use a solvent in supercritical conditions, which is known as supercritical extraction. Common problems for the modeling of extraction systems are lack of knowledge of thermodynamic parameters for some substances (e.g., critical constants of solid components are usually unknown [92]), kinetic models for leaching systems are difficult to obtain and generalize, and LLE calculations are computationally expensive.

Kamali and Mousavi [59] used ANN to account for the discrepancies between the predicted solubility (using Peng-Robison) of α-pinene in supercritical CO_2 and experimental data. They reported that the hybrid model outperforms the dense gas model with Peng-Robinson (mechanistic model) and the pure ANN in terms of predictability. Similarly, recursive PLS and SVM for regression have been used in leaching processes to account for unmodeled phenomena, resulting in more accurate and cheaper models compared to pure mechanistic ones [67, 93]. Leaching kinetic rates have been also modeled by ANN and coupled with first principles increasing the model's accuracy and generalization [94].

As mentioned before, LLE calculations are often very expensive which prevents accurate models to be applied within optimization frameworks. Traditionally, accurate equilibrium conditions are calculated using expensive EoS such as PC-SAFT [95] or g^E models such as modified UNIFAC Dortmund [96]. GP and quadratic polynomials have been used to reduce the complexity caused by using such expensive thermodynamic models, achieving fast and reliable surrogates [97]. In the last reference, a case study is presented, where a thermomorphic solvent system is optimized using surrogates for the partitioning behavior of the catalyst and the two-phase region of the system with the mentioned techniques. The results show a more robust and economically viable reaction-extraction system compared to the referenced case. This work was later extended to reduce the model dimensionality further, by modeling the binodal curve (instead of the complete two-phase region) and using numerical continuation [75]. The resulting hybrid model was later globally optimized.

Recent studies have shown that, by employing GP and ANN trained on data obtained from the PC-SAFT in an LLE system, accurate predictions of phase equilibrium can be obtained around 36 times faster than by using the original PC-SAFT [98]. This surrogate model was later augmented by introducing an SVM that classifies the validity of the operating conditions (ensuring only operating conditions within the two-phase region are considered) [32]. The resulting classification/regression surrogate was later used in the optimization of a simplified hydroformylation of

1-dodecene process, demonstrating how expensive thermodynamic packages can be embedded into an optimization problem of a process flow sheet. By using the original PC-SAFT model, the problem converges to an optimum in more than 48 hours; in contrast, by employing the surrogate, the optimum is found in less than 2 seconds. The integration of classification (one-phase or two-phase region) and regression (within the two-phase region) avoided discontinuity in the regression model as described in [99].

5.2.4 Flotation

Flotation processes work under the principle of difference in hydrophobicity of substances [100]. This difference is often increased by using surfactants or wetting agents. The flotation technique is commonly used in the metallurgical industry for valuable mineral separation and in the wastewater treatment sector for the removal of fats and oils. Air bubbles are injected into the mixture causing the hydrophobic substances to float to the surface, where they can be removed mechanically. Due to the lack of understanding of the phenomena involved and the vast number of interactions between many variables, e.g., agitation intensity, residence time of bubbles, bubbles size, particles size, and surface hydration, constructing an accurate dynamic model for flotation systems remains challenging [100].

Several applications of hybrid models used in flotation systems can be found in the literature. ANN and PCA have been integrated with first principle models to approximate the flotation kinetics of dynamic metallurgical systems showing great generalization capabilities across different processes, resulting in flexible and robust models [94, 101]. A parallel arrangement of several ANNs to a mechanistic model for the activated sludge process is also reported in the literature, reporting good accuracy [102]. Dong et al. [103] used a hybrid model with PCA and SVM for the switching and control of an industrial coal flotation process achieving an increase in the clean coal recovery compared to other control strategies. An extensive review on the use of hybrid models and data-driven surrogates in flotation processes can be found in [104].

5.2.5 Drying

Drying refers to the removal of moisture from a system, which is mostly accomplished by evaporation or sublimation. However, condensation and sorption can also be used as drying mechanisms if the system is a fluid [1]. The most important complication in modeling drying systems is the accurate determination of the rates of heat and mass transfer. Usually, an accurate calculation of these phenomena may result in solving systems of nonlinear PDAE that might be computationally expensive to solve. Hence, hybrid models have also proved to be suitable for the modeling of drying units.

Heat and mass transfer coefficients have been estimated via ANNs within mechanistic dynamic models of a direct flow rotary dryer and a batch fluidized bed dryer [57]. The results showed that better predictability and flexibility is achieved when using the hybrid model compared to its individual sub-models. In terms of control of drying units, a simple first principles model has been coupled with an ANN to obtain an accurate and easier-to-solve model [105], which can be used in online control.

5.2.6 Filtration

Filtration processes are pressure-driven methods that rely on membranes through which the different components of the mixture to be separated move at different speeds [1]. Filtration is frequently used in the pharmaceutical, food, and water treatment industries. The most important problems in filtration processes are the proper design and maintenance of the filtration membrane. The former is related with the physical properties of the membrane itself, e.g., type of membrane material and physical structure, the latter is mainly related with the incrustation (or scaling, fouling) process. The understanding of the kinetics related with fouling and the concentration gradient during the filtration are two of the most challenging tasks for modeling this type of processes.

Mechanistic models include several parameters that are determined by empirical correlations and experimental data. Piron et al. [106] use ANNs for the determination of such parameters in a microfiltration system of suspensions of baker's yeast achieving accurate results compared to experimental data. Membrane fouling has been also modeled using hybrid ANN, demonstrating accurate results for a water treatment plant [107].

5.2.7 Discovery of Materials for Separation Processes

The discovery of new materials is currently one of the most important areas of research, since new materials can lead to processes that are more efficient, better healthcare and water/food availability. This is also true for the improvement of separation tasks, where it may lead to new separation technologies even further away from what is well known, like distillation. The reason is that many separation techniques rely on the material properties (other than the mixture to be separated) to perform the split (e.g., solvents in extraction, adsorbents, membrane materials). Therefore, the discovery of new materials for separations remains a big challenge, but one with much potential.

Historically, materials have been discovered mainly by trial and error by individuals working in traditional laboratories. This method is prone to errors, extremely slow, and inefficient given the immensity of the material design space. For this reason, enormous effort has been allocated to the development of computational techniques and robotics to explore the design space more proficiently. A recent review on this topic is found in [108].

It has been shown that SVMs are able to classify a large space of metal-organic frameworks (MOFs) according to their CO_2 adsorption capacity. This serves as a pre-screening step to reduce the number of MOFs to consider for a subsequent more detailed screening [109]. The results have shown that the search space can be reduced by up to one order of magnitude, which imposes great time savings in the discovery path. Furthermore, linear regression and PCA have been used for the discovery of new solvents employing quantum chemical calculations [110]. This solvent screening method has been later integrated to the process optimization of a Diels-Alder reaction into a unique mixed-integer nonlinear programming (MINLP) optimization problem [111]. This integration shows that the optimum process includes a less efficient (in terms of reaction extent) solvent that benefits more the solvent-product separation (compared to the traditional way of selecting a solvent independently of the process-wide optimization framework). Therefore, the potential of unifying the material design task into a higher-scale problem becomes tractable by employing hybrid models.

6 Conclusion

This article presents a discussion of various issues still found in the design of separations and what role data can play in helping to resolve them. This comes about due to the two primary means available for increasing model accuracy: either with better theories or with more data. Better theories are needed to create better first-principle models, which would provide higher accuracy combined with better extrapolation. This also translates into a much better understanding of the physical phenomena that take place inside of a separation unit. The other proposed path to improving separation process design and control involves the integration of more data sources into the models via surrogate and hybrid modeling. In this article, it is argued that the latter is the most promising because it combines process knowledge with that what is not known or understood. In this way, the advantages of both approaches can be used. To support this idea, many examples were presented, where data-driven and hybrid models were used to increase the accuracy of separation modeling, simulation, and design. We hope that this discussion will continue and that the examples provided in this article will inspire the further development of these methods.

This work is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – TRR 63 "Integrated Chemical Processes in Liquid Multiphase Systems" (subproject B9) – 56091768. E. I. Sanchez Medina is also affiliated with the International Max Planck Research School for Advanced Methods in Process and Systems Engineering (IMPRS ProEng) at the Max Planck Institute for Dynamics of Complex Technical Systems in Magdeburg.

Kevin McBride is a research associate in the group Process Systems Engineering at the Max Planck Institute for Dynamics of Complex Technical Systems, headed by Prof. Dr.-Ing. Kai Sundmacher. He studied Chemical Engineering at Texas Tech University before obtaining his diploma at Clausthal University of Technology. He received his Ph.D. from the Otto-von-Guericke University Magdeburg. His research focuses on computer-aided molecular design of solvents and on modeling chemical production processes.

Edgar Ivan Sanchez Medina is a Ph.D. candidate in the group Process Systems Engineering, headed by Prof. Dr.-Ing. Kai Sundmacher, at the Otto-von-Guericke University Magdeburg. He is also part of the International Max Planck Research School for Advanced Methods in Process and Systems Engineering at the Max Planck Institute for Dynamics of Complex Technical Systems. He studied Chemical Engineering at Universidad Nacional Autonoma de Mexico and received his M.Sc. from Imperial College London. His research focuses on hybrid semi-parametric modeling for process optimization.

Kai Sundmacher studied Mechanical and Process Engineering at the Leibniz University Hannover and at the TU Clausthal (Ph.D. 1995, habilitation 1998). After a postdoctoral phase at the University of Newcastle, UK, he was appointed to the Chair of Process Systems Engineering at Otto-von-Guericke University Magdeburg in 1999. Since 2001, he has also been director of the Max Planck Institute for Dynamics of Complex Technical Systems in Magdeburg. His research focuses on modeling, analysis, and synthesis of complex processes in chemical production, biotechnology, and chemical energy conversion.

Abbreviations

ALAMO	automatic learning of algebraic models for optimization
ANN	artificial neural network
ARGONAUT	algorithms for global optimization of constrained gray-box computational problems
ARTMAP	predictive adaptive resonance theory
COSMO-RS	conductor-like screening model for real solvents
EoS	equation of state
FL	fuzzy logic
g^E	excess Gibbs energy
GP	Gaussian process
LLE	liquid-liquid equilibria
MESH	material balances, equilibrium relationships, summation equations, and heat balances
MINLP	mixed-integer nonlinear programming
MOF	metal-organic framework
NLP	nonlinear programming
NMPC	nonlinear model predictive control
NMR	nuclear magnetic resonance
NRTL	non-random two-liquid model
PCA	principal component analysis
PC-SAFT	perturbed chain statistical associating fluid theory
PDAE	partial differential algebraic equations
PLS	partial least squares
RTO	real-time optimization
SUMO	surrogate modeling
SVM	support vector machine
UNIFAC	UNIQUAC functional-group activity coefficients
UNIFAC-IL	UNIFAC-ionic liquids
UNIQUAC	universal quasichemical
VLE	vapor-liquid equilibria
VTPR	volume-translated Peng-Robinson

References

[1] J. D. Seader, E. J. Henley, *Separation Process Principles*, 2nd ed., John Wiley & Sons, Hoboken, NJ **2006**.

[2] J. Gmehling, B. Kolbe, M. Kleiber, J. Rarey, *Chemical Thermodynamics for Process Simulation*, 1st ed., Wiley-VCH, Weinheim **2012**.

[3] J. Mohd Ali, M. A. Hussain, M. O. Tade, J. Zhang, *Expert Syst. Appl.* **2015**, *42 (14)*, 5915–5931. DOI: https://doi.org/10.1016/j.eswa.2015.03.023

[4] S. M. Walas, *Phase Equilibria in Chemical Engineering*, Butterworth, Stoneham, MA **1985**.

[5] U. Weidlich, J. Gmehling, *Ind. Eng. Chem. Res.* **1987**, *26 (7)*, 1372–1381. DOI: https://doi.org/10.1021/ie00067a018

[6] *Dortmund Data Bank*, DDBST Dortmund Data Bank Software & Separation Technology GmbH, Oldenburg **2019**. www.ddbst.de/

[7] O. Kahrs, W. Marquardt, *Comput. Chem. Eng.* **2008**, *32 (4–5)*, 694–705. DOI: https://doi.org/10.1016/j.compchemeng.2007.02.014

[8] G. E. P. Box, J. S. Hunter, W. G. Hunter, *Statistics for Experimenters*, John Wiley & Sons, Hoboken, NJ **2005**.

[9] D. Dechambre, L. Wolff, C. Pauls, A. Bardow, *Ind. Eng. Chem. Res.* **2014**, *53 (50)*, 19620–19627. DOI: https://doi.org/10.1021/ie5035573

[10] J. Thien, C. Peters, T. Brands, H. J. Koß, A. Bardow, *Ind. Eng. Chem. Res.* **2017**, *56 (46)*, 13905–13910. DOI: https://doi.org/10.1021/acs.iecr.7b03230

[11] A. Friebel, A. Fröscher, K. Münnemann, E. von Harbou, H. Hasse, *Fluid Phase Equilib.* **2017**, *438*, 44–52. DOI: https://doi.org/10.1016/j.fluid.2017.01.027

[12] T. R. Boussie et al., *J. Am. Chem. Soc.* **2003**, *125 (14)*, 4306–4317. DOI: https://doi.org/10.1021/ja020868k

[13] J. Liu, C. Han, M. McAdon, J. Goss, K. Andrews, *Microporous Mesoporous Mater.* **2015**, *206*, 207–216. DOI: https://doi.org/10.1016/j.micromeso.2014.11.030

[14] C. E. Mohler, R. L. Kuhlman, C. A. Witham, M. K. Poindexter, in *AADE Natl. Tech. Conf. Exhib.*, Houston, TX, April **2011**.

[15] A. Klamt, *J. Phys. Chem.* **1995**, *99 (7)*, 2224–2235. DOI: https://doi.org/10.1021/j100007a062

[16] E. I. Alevizou, G. D. Pappa, E. C. Voutsas, *Fluid Phase Equilib.* **2009**, *284 (2)*, 99–105. DOI: https://doi.org/10.1016/j.fluid.2009.06.012

[17] R. Taylor, R. Krishna, H. Kooijman, *Chem. Eng. Prog.* **2003**, *99 (7)*, 28–39.

[18] Q. Chen, I. E. Grossmann, *Annu. Rev. Chem. Biomol. Eng.* **2017**, *8 (1)*, 249–283. DOI: https://doi.org/10.1146/annurev-chembioeng-080615-033546

[19] J. Rarey, *Ind. Eng. Chem. Res.* **2005**, *44 (19)*, 7600–7608. DOI: https://doi.org/10.1021/ie050431w

[20] K. Řehák, J. Matouš, J. P. Novák, *Fluid Phase Equilib.* **1995**, *109 (1)*, 113–129. DOI: https://doi.org/10.1016/0378-3812(95)96894-3

[21] J. Ahlers, J. Gmehling, *Fluid Phase Equilib.* **2001**, *191 (1–2)*, 177–188. DOI: https://doi.org/10.1016/S0378-3812(01)00626-4

[22] J. Gross, G. Sadowski, *Fluid Phase Equilib.* **2000**, *168 (2)*, 183–199. DOI: https://doi.org/10.1016/S0378-3812(00)00302-2

[23] S. J. Qin, *AIChE J.* **2014**, *60 (9)*, 3092–3100. DOI: https://doi.org/10.1002/aic.14523

[24] L. Chiang, B. Lu, I. Castillo, *Annu. Rev. Chem. Biomol. Eng.* **2017**, *8 (1)*, 63–85. DOI: https://doi.org/10.1146/annurev-chembioeng-060816-101555

[25] J. Liu, *Control Eng. Pract.* **2007**, *15 (7)*, 769–778. DOI: https://doi.org/10.1016/j.conengprac.2005.12.005

[26] Y. Liu, Z. Zhang, J. Chen, *Chem. Eng. Sci.* **2015**, *137*, 140–151. DOI: https://doi.org/10.1016/j.ces.2015.06.005

[27] S. J. Qin, *Annu. Rev. Control* **2012**, *36 (2)*, 220–234. DOI: https://doi.org/10.1016/j.arcontrol.2012.09.004

[28] J. Yu, S. J. Qin, *AIChE J.* **2008**, *54 (7)*, 1811–1829. DOI: https://doi.org/10.1002/aic.11515

[29] B. Lu, I. Castillo, L. Chiang, T. F. Edgar, *Chemom. Intell. Lab. Syst.* **2014**, *135*, 90–109. DOI: https://doi.org/10.1016/j.chemolab.2014.03.020

[30] K. D. Starr, H. Petersen, M. Bauer, *IFAC-PapersOnLine* **2016**, *49 (7)*, 526–532. DOI: https://doi.org/10.1016/j.ifacol.2016.07.396

[31] C. Li, D. Zhao, S. Mu, W. Zhang, N. Shi, L. Li, *Chin. J. Chem. Eng.* **2019**, *27 (3)*, 598–604. DOI: https://doi.org/10.1016/j.cjche.2018.12.021

[32] C. Nentwich, C. Varela, S. Engell, in *Proc. Int. Jt. Conf. Neural Networks*, IEEE, Piscataway, NJ **2019**, 1–8. DOI: https://doi.org/10.1109/IJCNN.2019.8851816

[33] M. D. McKay, R. J. Beckman, W. J. Conover, *Technometrics* **1979**, *21 (2)*, 239–245. DOI: https://doi.org/10.1080/00401706.1979.10489755

[34] I. M. Sobol, *USSR Comput. Math. Math. Phys.* **1967**, *7 (4)*, 86–112. DOI: https://doi.org/10.1016/0041-5553(67)90144-9

[35] J. H. Halton, *Numer. Math.* **1960**, *2 (1)*, 84–90. DOI: https://doi.org/10.1007/BF01386213

[36] D. R. Jones, M. Schonlau, W. J. Welch, *J. Global Optim.* **1998**, *13*, 455–492. DOI: https://doi.org/10.1023/A:1008306431147

[37] H. M. Gutmann, *J. Global Optim.* **2001**, *19 (3)*, 201–227. DOI: https://doi.org/10.1023/A:1011255519438

[38] K. Crombecq, D. Gorissen, D. Deschrijver, T. O. M. Dhaene, *SIAM J. Sci. Comput.* **2011**, *33 (4)*, 1948–1974. DOI: https://doi.org/10.1137/090761811

[39] S. S. Garud, I. A. Karimi, M. Kraft, *Comput. Chem. Eng.* **2017**, *96*, 103–114. DOI: https://doi.org/10.1016/j.compchemeng.2016.10.006

[40] J. Eason, S. Cremaschi, *Comput. Chem. Eng.* **2014**, *68*, 220–232. DOI: https://doi.org/10.1016/j.compchemeng.2014.05.021

[41] *NIST Chemistry WebBook*, National Institute of Standards and Technology, Gaithersburg, MD **2019**. https://webbook.nist.gov/

[42] A. Bhosekar, M. Ierapetritou, *Comput. Chem. Eng.* **2018**, *108*, 250–267. DOI: https://doi.org/10.1016/j.compchemeng.2017.09.017

[43] K. McBride, K. Sundmacher, *Chem. Ing. Tech.* **2019**, *91 (3)*, 228–239. DOI: https://doi.org/10.1002/cite.201800091

[44] A. Forrester, A. Sóbester, A. J. Keane, *Engineering Design via Surrogate Modelling: A Practical Guide*, 1st ed., John Wiley & Sons, Chichester **2008**.

[45] D. Gorissen, I. Couckuyt, P. Demeester, T. Dhaene, K. Crombecq, *J. Mach. Learn. Res.* **2010**, *11*, 2051–2055.

[46] A. Cozad, N. V. Sahinidis, D. C. Miller, *AIChE J.* **2014**, *60 (6)*, 2211–2227. DOI: https://doi.org/10.1002/aic.14418

[47] A. Cozad, N. V. Sahinidis, D. C. Miller, *Comput. Chem. Eng.* **2015**, *73*, 116–127. DOI: https://doi.org/10.1016/j.compchemeng.2014.11.010

[48] Z. T. Wilson, N. V. Sahinidis, *Comput. Chem. Eng.* **2017**, *106*, 785–795. DOI: https://doi.org/10.1016/j.compchemeng.2017.02.010

[49] B. Beykal, F. Boukouvala, C. A. Floudas, E. N. Pistikopoulos, *Comput. Chem. Eng.* **2018**, *116*, 488–502. DOI: https://doi.org/10.1016/j.compchemeng.2018.02.017

[50] B. Beykal, F. Boukouvala, C. A. Floudas, N. Sorek, H. Zalavadia, E. Gildin, *Comput. Chem. Eng.* **2018**, *114*, 99–110. DOI: https://doi.org/10.1016/j.compchemeng.2018.01.005

[51] F. Boukouvala, C. A. Floudas, *Optim. Lett.* **2017**, *11 (5)*, 895–913. DOI: https://doi.org/10.1007/s11590-016-1028-2

[52] F. Boukouvala, M. M. F. Hasan, C. A. Floudas, *J. Global Optim.* **2017**, *67 (1–2)*, 3–42. DOI: https://doi.org/10.1007/s10898-015-0376-2

[53] H. R. Sant Anna, A. G. Barreto, F. W. Tavares, M. B. de Souza, *Comput. Chem. Eng.* **2017**, *104*, 377–391. DOI: https://doi.org/10.1016/j.compchemeng.2017.05.006

[54] K. T. Leperi, D. Yancy-Caballero, R. Q. Snurr, F. You, *Ind. Eng. Chem. Res.* **2019**, *58 (39)*, 18241–18252. DOI: https://doi.org/10.1021/acs.iecr.9b02383

[55] S. G. Subraveti, Z. Li, V. Prasad, A. Rajendran, *Ind. Eng. Chem. Res.* **2019**, *58 (44)*, 20412–20422. DOI: https://doi.org/10.1021/acs.iecr.9b04173

[56] N. Bhutani, G. P. Rangaiah, A. K. Ray, *Ind. Eng. Chem. Res.* **2006**, *45 (23)*, 7807–7816. DOI: https://doi.org/10.1021/ie060247q

[57] F. A. Cubillos, P. I. Alvarez, J. C. Pinto, E. L. Lima, *Powder Technol.* **1996**, *87 (2)*, 153–160. DOI: https://doi.org/10.1016/0032-5910(95)03083-2

[58] P. Georgieva, M. J. Meireles, S. Feyo de Azevedo, *Chem. Eng. Sci.* **2003**, *58 (16)*, 3699–3713. DOI: https://doi.org/10.1016/S0009-2509(03)00260-4

[59] M. J. Kamali, M. Mousavi, *J. Supercrit. Fluids* **2008**, *47 (2)*, 168–173. DOI: https://doi.org/10.1016/j.supflu.2008.08.005

[60] J. Glassey, M. von Stosch, *Hybrid Modeling in Process Industries*, CRC Press, Boca Raton, FL **2018**.

[61] S. Zendehboudi, N. Rezaei, A. Lohi, *Appl. Energy* **2018**, *228*, 2539–2566. DOI: https://doi.org/10.1016/j.apenergy.2018.06.051

[62] M. L. Thompson, M. A. Kramer, *AIChE J.* **1994**, *40 (8)*, 1328–1340. DOI: https://doi.org/10.1002/aic.690400806

[63] P. Lauret, H. Boyer, J. C. Gatina, *Control Eng. Pract.* **2000**, *8 (3)*, 299–310. DOI: https://doi.org/10.1016/S0967-0661(99)00151-3

[64] M. von Stosch et al., *Biotechnol. J.* **2014**, *9 (6)*, 719–726. DOI: https://doi.org/10.1002/biot.201300385

[65] M. von Stosch, R. Oliveira, J. Peres, S. Feyo de Azevedo, *Comput. Chem. Eng.* **2014**, *60*, 86–101. DOI: https://doi.org/10.1016/j.compchemeng.2013.08.008

[66] M. Agarwal, *Int. J. Syst. Sci.* **1997**, *28 (1)*, 65–81. DOI: https://doi.org/10.1080/00207729708929364

[67] R.-D. Jia, Z.-Z. Mao, Y.-Q. Chang, L.-P. Zhao, *Chem. Eng. Res. Des.* **2011**, *89 (6)*, 722–728. DOI: https://doi.org/10.1016/j.cherd.2010.09.015

[68] M. W. Hermanto, R. D. Braatz, M.-S. Chiu, *AIChE J.* **2011**, *57 (4)*, 1008–1019. DOI: https://doi.org/10.1002/aic.12331

[69] J. Pellegrino, S. Brueske, T. Carole, A. Howard, *Energy and Environmental Profile of the U.S. Petroleum Refining Industry*, Energetics Inc., Columbia, MD **2007**.

[70] A. A. Safavi, A. Nooraii, J. A. Romagnoli, *J. Process Control* **1999**, *9 (2)*, 125–134. DOI: https://doi.org/10.1016/S0959-1524(98)00041-9

[71] S. Engell, K. Dadhe, *IFAC Proc. Vol.* **2001**, *34 (22)*, 354–359. DOI: https://doi.org/10.1016/s1474-6670(17)32964-6

[72] V. Mahalec, Y. Sanchez, *Comput. Chem. Eng.* **2012**, *45*, 15–26. DOI: https://doi.org/10.1016/j.compchemeng.2012.05.012

[73] G. Fu, Y. Sanchez, V. Mahalec, *AIChE J.* **2016**, *62 (4)*, 1065–1078. DOI: https://doi.org/10.1002/aic.15086
[74] C. Tsay, M. Baldea, *Ind. Eng. Chem. Res.* **2019**, *58 (36)*, 16696–16708. DOI: https://doi.org/10.1021/acs.iecr.9b02282
[75] C. Kunde, T. Keßler, S. Linke, K. McBride, K. Sundmacher, A. Kienle, *Processes* **2019**, *7 (10)*, 753. DOI: https://doi.org/10.3390/pr7100753
[76] T. Keßler, C. Kunde, N. Mertens, D. Michaels, A. Kienle, *SN Appl. Sci.* **2019**, *1 (1)*, 11. DOI: https://doi.org/10.1007/s42452-018-0008-9
[77] Z. Zhang, Z. Wu, D. Rincon, P. D. Christofides, *Mathematics* **2019**, *7 (10)*, 890. DOI: https://doi.org/10.3390/math7100890
[78] P. F. Van Lith, B. H. L. Betlem, B. Roffel, *Comput. Chem. Eng.* **2003**, *27 (7)*, 1021–1030. DOI: https://doi.org/10.1016/S0098-1354(03)00067-X
[79] L. Chen, Y. Hontoir, D. Huang, J. Zhang, A. J. Morris, *Control Eng. Pract.* **2004**, *12 (7)*, 819–826. DOI: https://doi.org/10.1016/j.conengprac.2003.09.006
[80] S. M. Khazraee, A. H. Jahanmiri, *Chin. J. Chem. Eng.* **2010**, *18 (4)*, 703–710. DOI: https://doi.org/10.1016/S1004-9541(10)60278-9
[81] A. Sabharwal, W. Y. Svrcek, D. E. Seborg, *IFAC Proc. Vol.* **1999**, *32 (2)*, 6799–6805. DOI: https://doi.org/10.1016/s1474-6670(17)57161-x
[82] E. Abdul Jaleel, K. Aparna, *Evol. Syst.* **2019**, *10 (2)*, 149–166. DOI: https://doi.org/10.1007/s12530-018-9220-5
[83] M. Fujiwara, Z. K. Nagy, J. W. Chew, R. D. Braatz, *J. Process Control* **2005**, *15 (5)*, 493–504. DOI: https://doi.org/10.1016/j.jprocont.2004.08.003
[84] Y. Meng, S. Yu, J. Zhang, J. Qin, Z. Dong, G. Lu, H. Pang, *J. Food Eng.* **2019**, *257*, 44–55. DOI: https://doi.org/10.1016/j.jfoodeng.2019.03.026
[85] P. Georgieva, S. Feyo de Azevedo, *Int. J. Chem. Mol. Eng.* **2007**, *1 (12)*, 145–154. DOI: https://doi.org/10.5281/zenodo.1078553
[86] P. Georgieva, S. Feyo de Azevedo, in *Computational Intelligence Techniques for Bioprocess Modelling, Supervision and Control* (Eds: M. do Carmo Nicoletti, L. C. Jain), Springer, Berlin **2009**.
[87] S. Zhang, F. Wang, D. He, R. Jia, *Powder Technol.* **2012**, *224*, 253–259. DOI: https://doi.org/10.1016/j.powtec.2012.03.001
[88] Z. Xiong, J. Zhang, *J. Process Control* **2005**, *15 (1)*, 11–21. DOI: https://doi.org/10.1016/j.jprocont.2004.04.005
[89] G. A. Carpenter, S. Grossberg, N. Markuzon, J. H. Reynolds, D. B. Rosen, *IEEE Trans. Neural Networks* **1992**, *3 (5)*, 698–713. DOI: https://doi.org/10.1109/72.159059
[90] C. Valencia Peroni, M. Parisi, A. Chianese, *Chem. Eng. Res. Des.* **2010**, *88 (12)*, 1653–1658. DOI: https://doi.org/10.1016/j.cherd.2010.01.038
[91] S. Zhang, F. Chu, G. Deng, F. Wang, *IEEE Access* **2019**, *7*, 118749–118763. DOI: https://doi.org/10.1109/access.2019.2936542
[92] G. Soave, *J. Supercrit. Fluids* **2000**, *19 (1)*, 19–24. DOI: https://doi.org/10.1016/S0896-8446(00)00074-7
[93] G. Hu, Z. Mao, D. He, F. Yang, *Comput. Chem. Eng.* **2011**, *35 (12)*, 2611–2617. DOI: https://doi.org/10.1016/j.compchemeng.2011.02.012
[94] M. A. Reuter, J. S. J. Van Deventer, T. J. Van Der Walt, *Chem. Eng. Sci.* **1993**, *48 (7)*, 1281–1297. DOI: https://doi.org/10.1016/0009-2509(93)81009-K
[95] J. Gross, G. Sadowski, *Ind. Eng. Chem. Res.* **2001**, *40 (4)*, 1244–1260. DOI: https://doi.org/10.1021/ie0003887
[96] J. Gmehling, R. Wittig, J. Lohmann, R. Joh, *Ind. Eng. Chem. Res.* **2002**, *41 (6)*, 1678–1688. DOI: https://doi.org/10.1021/ie0108043
[97] K. McBride, N. M. Kaiser, K. Sundmacher, *Comput. Chem. Eng.* **2017**, *105*, 212–223. DOI: https://doi.org/10.1016/j.compchemeng.2016.11.019
[98] C. Nentwich, J. Winz, S. Engell, *Ind. Eng. Chem. Res.* **2019**, *58 (40)*, 18703–18716. DOI: https://doi.org/10.1021/acs.iecr.9b02758
[99] C. Nentwich, S. Engell, *Comput. Chem. Eng.* **2019**, *126*, 204–217. DOI: https://doi.org/10.1016/j.compchemeng.2019.04.006
[100] K. A. Matis, *Flotation Science and Engineering*, Marcel Dekker, New York **1995**.
[101] F. A. Cubillos, E. L. Lima, *Miner. Eng.* **1997**, *10 (7)*, 707–721. DOI: https://doi.org/10.1016/s0892-6875(97)00050-2
[102] M. Côté, B. P. A. Grandjean, P. Lessard, J. Thibault, *Water Res.* **1995**, *29 (4)*, 995–1004. DOI: https://doi.org/10.1016/0043-1354(95)93250-W
[103] Z. Dong, R. Wang, M. Fan, X. Fu, *PLoS One* **2017**, *12 (10)*, e0186553. DOI: https://doi.org/10.1371/journal.pone.0186553
[104] I. Jovanović, I. Miljanović, T. Jovanović, *Miner. Eng.* **2015**, *84*, 34–63. DOI: https://doi.org/10.1016/j.mineng.2015.09.020
[105] A. Saraceno, M. Aversa, S. Curcio, *Food Bioprocess Technol.* **2012**, *5 (5)*, 1694–1705. DOI: https://doi.org/10.1007/s11947-010-0477-3
[106] E. Piron, E. Latrille, F. René, *Comput. Chem. Eng.* **1997**, *21 (9)*, 1021–1030. DOI: https://doi.org/10.1016/S0098-1354(96)00332-8
[107] T. M. Hwang, H. Oh, Y. J. Choi, S. H. Nam, S. Lee, Y. K. Choung, *Desalination* **2009**, *247 (1–3)*, 210–221. DOI: https://doi.org/10.1016/j.desal.2008.12.025
[108] T. Zhou, Z. Song, K. Sundmacher, *Engineering* **2019**, *5 (6)*, 1017–1026. DOI: https://doi.org/10.1016/j.eng.2019.02.011
[109] M. Fernandez, P. G. Boyd, T. D. Daff, M. Z. Aghaji, T. K. Woo, *J. Phys. Chem. Lett.* **2014**, *5 (17)*, 3056–3060. DOI: https://doi.org/10.1021/jz501331m
[110] T. Zhou, Z. Qi, K. Sundmacher, *Chem. Eng. Sci.* **2014**, *115*, 177–185. DOI: https://doi.org/10.1016/j.ces.2013.11.020
[111] T. Zhou, K. McBride, X. Zhang, Z. Qi, K. Sundmacher, *AIChE J.* **2015**, *61 (1)*, 147–158. DOI: https://doi.org/10.1002/aic.14630

