

Available online at www.sciencedirect.com

Chemical Engineering Research and Design

journal homepage: www.elsevier.com/locate/cherdIChemE
ADVANCING
CHEMICAL
ENGINEERING
WORLDWIDE

Data-driven simultaneous process optimization and adsorbent selection for vacuum pressure swing adsorption

Sun Hye Kim, Héctor Octavio Rubiera Landa, Suryateja Ravutla, Matthew J. Realff, Fani Boukouvala*

School of Chemical & Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0100, USA

ARTICLE INFO

Article history:

Received 27 May 2022

Received in revised form 3 October 2022

Accepted 4 October 2022

Available online 10 October 2022

Keywords:

Surrogate-based optimization

Machine learning

Vacuum pressure-swing adsorption

Post-combustion CO₂ capture

Process intensification

Neural networks

ABSTRACT

Technologies for post-combustion carbon capture are essential for the reduction of greenhouse gas emissions to the atmosphere. However, they are still associated with high costs and energy consumption. Intensified processes for carbon capture have the potential to overcome these challenges due to their higher efficiency, lower capital cost, and increased operational flexibility. This work investigates simultaneous optimization of process conditions and adsorbent selection for a modular Vacuum Pressure-Swing Adsorption system designed for CO₂ capture. Both surrogate-based Nonlinear Programming and Mixed-Integer Nonlinear Programming approaches are applied and compared in terms of computational efficiency and solution accuracy. Moreover, process performance results are examined by applying several data analytics techniques to gain insights into the material-process correlations. Data-driven classifiers and neural networks can accurately predict whether a material is likely to satisfy purity, recovery, and energy constraints when operated at optimal process conditions.

© 2022 Institution of Chemical Engineers. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Emission of CO₂ has been recognized as an important environmental issue and one of the major contributing causes of climate change (Choi et al. 2009, Bhowan and Freeman, 2011, Hasan et al. 2014, Hasan et al. 2015, Ben-Mansour et al. 2016, Bui et al. 2018). In 2017, the amount of CO₂ emissions in the U.S. totaled 6457 million metric tons ("U.S. EPA's Inventory of U.S.", 2019). Carbon Capture and Storage (CCS) has been proposed to reduce CO₂ emissions but several existing techniques are currently associated with high cost and large energy consumption (Yang, Xu et al. 2008, Hasan et al. 2013, Ben-Mansour et al. 2016, Leperi et al. 2019). One promising technology for post-combustion carbon capture is adsorption using solid sorbents due to its relatively high

separation efficiency, low energy cost in comparison to absorption-based technologies, and its potential for modularization (DeWitt et al. 2019).

During adsorption, separation units containing solid adsorbent sequester CO₂ from flue gas through a dynamic cyclic operation (Ruthven, 1984, Ebner and Ritter, 2009). During each cycle, CO₂ is captured and separated from the rest of the mixture; the adsorbent is then regenerated, and the cycle is repeated (Ruthven, 1984). Depending on how the adsorbent is regenerated, adsorption-based technology, which is typically carried out in packed-bed adsorbers, can be grouped into three main operational modes: (1) pressure-swing adsorption (PSA), (2) vacuum-swing adsorption (VSA), and (3) temperature-swing adsorption (TSA) (Ruthven, 1984, Ruthven et al. 1994). Significant research efforts have focused on improving these packed-bed adsorbers by overcoming pressure-drop limitations, mitigating adsorption enthalpy, and improving mass transfer to make the operation more cost- and energy-efficient (Samanta et al. 2012, Rezaei et al. 2014, DeWitt et al. 2019, Jang et al. 2019, Sinha and Realff, 2019). Parallel to these

* Corresponding author.

E-mail address:

fani.boukouvala@chbe.gatech.edu (F. Boukouvala).

<https://doi.org/10.1016/j.cherd.2022.10.002>

0263-8762/© 2022 Institution of Chemical Engineers. Published by Elsevier Ltd. All rights reserved.

efforts, the development of new adsorbents with high operating capacity and adsorptive selectivity has also been investigated extensively (Darunte et al. 2019). A wide spectrum of sorbents are commercially available and under development, including activated carbons, zeolites, and metal-organic frameworks (MOFs) (Yazaydin et al. 2009, Huck et al. 2014, Ben-Mansour et al. 2016, Findley et al. 2018, Darunte et al. 2019). It has been repeatedly demonstrated that the choice of an adsorbent is critical for the successful design of an adsorption process that targets low cost and minimal energy penalty for CCS (Choi et al. 2009, Sinha and Realff, 2019).

Recent studies have revealed that process performance is intricately linked to the choice of an adsorbent, implying that the adsorbent selection and process optimization should be considered simultaneously (Hasan et al. 2013, Khurana and Farooq, 2016, Khurana and Farooq, 2017, Subramanian Balashankar, 2019). This requires the formulation of a complex simulation-based optimization problem because cyclic adsorption processes are typically described by detailed mathematical models consisting of Partial Differential-Algebraic Equations (PDAEs). Conventional equation-based optimization has been used in (Kikkinides et al. 1993, Ko et al. 2005, Agarwal et al. 2010) but may pose limitations when model complexity is high. Hence, the use of stochastic sampling-based and/or surrogate-based techniques have been proposed. In (Fiandaca et al. 2009, Haghpanah et al. 2013, Haghpanah et al. 2013, Khurana and Farooq, 2016; 2017, Leperi et al. 2019) a stochastic optimization algorithm (e.g., Genetic Algorithm (GA)) is used to determine the optimal process operating conditions. However, stochastic algorithms tend to require many simulation evaluations; thus, it may not be suitable when the computer simulation is computationally expensive. To overcome this limitation, surrogate-based optimization techniques have been proposed (Forrester, 2009, Boukouvala et al. 2017, Beykal et al. 2020, Dias and Ierapetritou, 2020). Different terminology has been used in the literature such as ‘meta-model’, ‘Machine-Learning model’, and ‘Surrogate model’ to refer to approximations of data or higher-order models. To avoid any confusion, we will use the term ‘Surrogate model’ throughout the paper. A surrogate model approximates the input-output relationship of a high-fidelity simulation with reduced complexity (Boukouvala and Floudas, 2017, Bhosekar and Ierapetritou, 2018, Kim and Boukouvala, 2019, McBride and Sundmacher, 2019). The surrogate model is subsequently used within algorithms that iteratively sample-fit-optimize surrogates to find optimal solutions and several such surrogate-based optimization techniques exist in the literature (Boukouvala et al., 2017; Cozad et al., 2014; Hüllen et al., 2019; Regis, 2020; Williams and Cremaschi, 2021; Kim and Boukouvala, 2020; Regis, 2020).

Recent advances in simulation based optimization have attracted the attention of many researchers and led to the development of algorithms that can handle mixed-integer inputs in high-dimensional spaces (Kim and Boukouvala, 2020, Regis, 2020; Sun et al. 2020, Williams and Cremaschi, 2021). However, there are some challenges that remain, especially when the sampling costs are limited and the inputs contains both discrete and continuous variables (Sun et al. 2020). For example, the performance of one of the popular direct-search solvers NOMAD was compared w.r.t. the quality of the solution and the number of simulation runs to other surrogate-based optimization algorithms in (Kim and Boukouvala, 2020, Regis,

2020, Sun et al. 2020) and was found to have high sampling costs. This further encouraged to look for alternatives, such as surrogate-based optimization algorithms.

Realizing the complexity of dynamic PSA simulations, many researchers have employed surrogate modeling to represent outputs of interest produced by the simulation for optimization (Sundaram, 1999, Li et al. 2016, Sant Anna, et al. 2017, Leperi et al. 2019, Ye, Ma et al. 2019, Xiao, Li et al. 2020, Xiao et al. 2021). Several works studying adsorbent selection and process optimization also exist in the literature. Hasan et al. (Hasan et al. 2013) proposes an adsorbent screening framework using a combined material characterization and process optimization procedure for both PSA and VSA processes. A kriging-based grey-box optimization approach is used, and the minimum cost of capture and compression is obtained for the final optimal design satisfying purity and recovery constraints. Khurana and Farooq present a two-stage adsorbent screening framework for a VSA process for carbon capture (Khurana, 2016). They apply a neural network-based classification model to preliminarily screen adsorbents based on purity-recovery targets; an extensive optimization study is then performed to rank adsorbents and determine the best operating conditions. The same authors perform a study of integrated material-design optimization, where material properties are represented as cleverly extracted features that span the material property space (Khurana and Farooq, 2017). Leperi et al. (Leperi et al. 2019) perform a full PSA modeling and optimization using GA and introduced a general evaluation metric (GEM) for the screening of MOFs. More recently, Balashankar et al. (Subramanian Balashankar and Rajendran, 2019) coupled a genetic algorithm-based process optimization with a detailed VSA model to develop a zeolite screening framework. Khurana and Farooq (Khurana and Farooq, 2019) extended their previous work to include a costing framework and Yancy, Liperi et al., performed process-level optimization and economic analysis of 15 MOFs from the literature (Yancy-Caballero, Leperi et al. 2020). A screening analysis using machine learning techniques was performed in (Burns et al. 2020) to screen 1632 experimentally characterized MOFs utilizing GA for process optimization. Some of the recent works such as (Nogueira et al. 2022) utilize surrogate based optimization framework using Neural Networks for simultaneous material screening and process optimization and (Pai et al. 2022) use Neural Networks predicting the performance of adsorbents. Many such works involving performance-based screening have been listed in (Farmahini, Krishnamurthy et al. 2021).

Despite these efforts, challenges still remain for the industrial-scale deployment of adsorptive separations for carbon capture, namely the cost of adsorbents, pressure-drop limitations observed in conventional industrial-size beds, and high adsorption enthalpy (DeWitt et al. 2019). Modular process intensification (PI) offers the opportunity to effectively overcome these technical challenges (Stankiewicz and Moulijn, 2000, Baldea et al. 2017, Kim, Park et al. 2017). Energy consumption is highly correlated with the size of a chemical process, implying large-scale processes are potentially more inefficient (Stankiewicz and Moulijn, 2000, Lutze et al. 2010). By reducing the size, one can enhance mass and heat transfer and reduce pressure drop, which may lead to 30% energy savings and 20% lower operating costs (Baldea et al. 2017). Another major advantage of modularization concerns module fabrication and deployment. In contrast to

centralized facilities that require most of their processes to be built on-site, modules can be mass-produced, leading up to 40% less capital expenditure and shorter module construction and deployment time (Baldea et al. 2017). These modules can be added or removed depending on plant capacity, allowing flexible deployment. However, in order to fully exploit these advantages, modules need to be optimally designed so that they can be operated at variable process conditions with the chosen adsorbent. Moreover, the potential of operating integrated modules each with different material for adsorption is a promising approach to process feeds with varying flowrates, and this is a problem that would require simultaneous optimization of operation and material selection.

Responsible for about 30% of electricity production in the USA, coal-fired power plants could benefit greatly from modular PI by exploiting several advantages intrinsic to modular deployment and production. First, modularization may allow CCS to be viable by allowing incremental deployment, reducing the initial capital requirements, and permitting technological learning, thereby resulting in an overall reduction in risk (Yang et al. 2008, Yu et al. 2012, Weber et al. 2018). Second, there is significant variability in the CO₂-content emitted by coal-fired plants across the United States. Since the modules are mass-produced and are smaller in size, they could be transported and operated to best-match power-plant capacity and seasonal emission characteristics, offering advantageous operational flexibility.

The purpose of this work is two-fold. First, we propose a technique to design modular CO₂ capture systems for coal-fired power plants using VPSA coupled with thermally-modulated fiber composite adsorbents. Surrogate-based optimization is used to determine the optimal adsorbent and process conditions simultaneously to meet at least 95% purity and 90% recovery targets. From a computational aspect, in this work we compare the feasibility and compute-cost savings of simultaneous optimization of material selection and operation using surrogate-based techniques versus a brute-force parallel optimization of operation for each material separately. To the best of our knowledge, this is the first attempt to use a surrogate based MINLP algorithm considering 75 adsorbents and 5 operating conditions for the simultaneous optimization of adsorbent and process conditions. In the second part, we present a purely machine learning-based classification and regression approach to gain further insight into adsorbent-process performance through the use of the large amount of data obtained from the simulation-based optimization study. In this part, we present useful reduced-order correlations between material and process operating conditions, that are revealed when analyzing the large amount of locally optimal solutions obtained by our optimization algorithm.

The paper is organized as follows. Section 2 provides a background on VPSA cycle modeling as well as an overview of the surrogate-based optimization algorithm with a specific focus on a neural network surrogate model. In Section 3, adsorbent selection and process optimization are performed using surrogate-based NLP and MINLP algorithms, and a comparison between the two proposed methodologies is presented. Using the optimal process data for adsorbents, Section 4 presents the use of dimensionality reduction, classification, and regression techniques to gain further insight into adsorbent-process performance and trade-offs. We end with conclusions and future outlook.

2. Background: process modeling and optimization

2.1. VPSA cycle modeling

2.1.1. Process description

In this work, we consider a VPSA process for a single module with a length of 1 m and a diameter of 1/6 m. This fixed-bed adsorber is packed with thermally-modulated fiber composites that consist of a porous polymeric matrix spun as a fiber embedded with adsorbent particles (ADS) and phase-change material (PCM). The fraction of the polymer matrix that can be loaded with solids is fixed at a minimum of 25% to guarantee correct fiber manufacturing, while allowing suitable amounts of adsorbent and PCM. In this way, an optimal trade-off between ADS and PCM is achieved in the process design. We present a detailed thermal management study in (Rubiera et al. 2020) that addresses this type of contactors based on our recent experimental results (DeWitt et al. 2019). The application of this kind of fiber composite reduces pressure drop in the packed-bed, thus enabling higher gas-velocity operation and better modulation of the heat generated by adsorption, rendering productive cycle operation with optimal adsorbent utilization (DeWitt et al., 2018). Adsorption equilibria are modeled with the extended dual-site Langmuir equation (eDSL) as required by the selection of adsorbents applied in the optimization (see Section 2.3). The heats of adsorption for N₂ and CO₂ required for process modeling are estimated for each adsorbent from the corresponding eDSL equation parameter values. The isosteric heats of adsorption were calculated using Clausius-Clapeyron equation. We have assumed that the heats of adsorption were constant throughout the loading range. We utilize a thermally modulated fiber bed, and therefore neglect the dependence of heats of adsorption on temperature (see Supporting Information). The VPSA cycle consists of the following four steps: counter-current pressurization (ccPr), high-pressure adsorption (Ad), co-current blow-down (coBd), and counter-current evacuation (ccEv). Fig. 1 illustrates this cycle, which allows recovering CO₂ during the evacuation step at high purity. This is enhanced by pressurization with the light-product (LPP) stream, which is rich in N₂.

Extensive studies have been performed on the adequacy of this kind of cycle configuration for CO₂ capture applications – see e.g. (Haghpanah et al. 2013) and references listed therein. Other cycle designs that consider multi-column operation and column-interaction through pressure equalization steps are also available for this kind of separation but are outside of the scope of this work (Xiao et al. 2008).

2.1.2. Mathematical Modeling and Numerical Solution

The detailed full-order dynamic adsorber model for non-isothermal operation consists of a set of partial differential-algebraic equations (PDAEs) in one spatial dimension, which is transformed to an ordinary differential equation (ODE) system in time by application of the first-order upwind discretization scheme (UDS), a finite-volume discretization that is suitable due to its numerical robustness—method of lines (MOL) approach (Schiesser, 1991). The time integration of this ODE system proceeds by applying the Backward Differentiation Formulae of Gear (Gear and Gear, 1971). This model is coded and solved entirely in MATLAB (MATLAB, 2018), applying the 'ode15s' numerical integration function at default tolerances. We apply suitable boundary conditions (BCs)

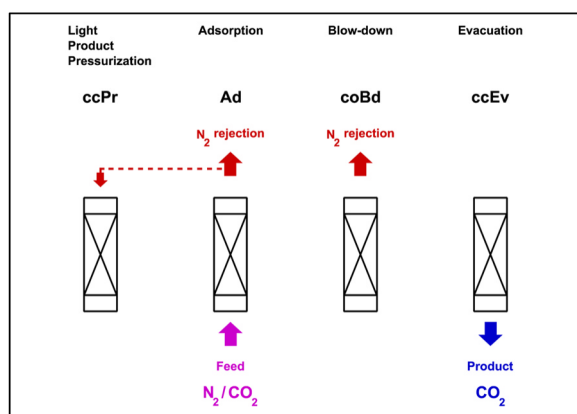


Fig. 1 – Schematic representation of the 4-step VPSA cycle with light-product pressurization (LPP).

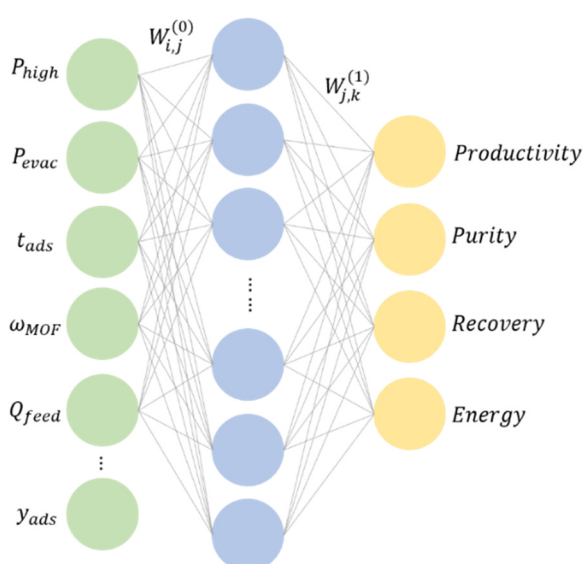


Fig. 2 – Illustration of a neural network with an input, one hidden, and an output layer.

and initial conditions (ICs) to represent the cyclic operation displayed in Fig. 1. The complete mathematical model is documented in the [Supporting Information](#).

2.1.3. Process Optimization

Several degrees of freedom are available to optimize the 4-step VPSA cycle. We select a subset of these decision variables, which have the strongest impact on VPSA process performance. Specifically, these are (1) Adsorption step pressure ($2.5 \leq P_{high} \leq 20[atm]$), (2) Evacuation step pressure ($0.01 \leq P_{evac} \leq 0.5[atm]$), (3) Feed gas flow rate ($0.001 \leq Q_{feed} \leq 0.0075 \left[\frac{m^3}{s} \right]$), (4) Weight-fraction of adsorbent in fiber ($0.15 \leq \omega_{ads} \leq 0.5 \left[\frac{kg \text{ adsorbent}}{kg \text{ solid fiber}} \right]$) and (5) Adsorption time ($15 \leq t_{ads} \leq 120[sec]$).

The adsorption and evacuation pressures determine the magnitude of the pressure-swing applied to the process. Feed gas flow rate and adsorption step time control the feed throughput and are therefore critical in the positioning of concentration and temperature fronts along the axial flow direction of the adsorber, once cyclic steady-state (CSS) operation has been attained. Lastly, the weight fraction of PCM

in the composites establishes the extent to which the heat generated by adsorption is modulated. The adsorption enthalpy varies between the adsorbents considered, leading to values of PCM content that are specific to each adsorbent and the optimal VPSA operation conditions. The blow-down step pressure is fixed at 1 atm, therefore avoiding the application of two different evacuation pressure levels, which in practice would increase the technical complexity of the applied evacuation system. We select values of the valve-coefficients for the pressure-changing steps (ccPr, coBd & ccEv) that minimize total cycle time. The five decision variables described above are critical in determining the cycle operation and are essential for its optimization, as we discuss below.

In contrast to the conventional fixed bed packed with particles and without any thermal modulation, we explore the process operation at higher adsorption pressures for a few reasons: (1) one of the intrinsic advantages of thermally modulated fiber composite packed beds includes minimal pressure drop, which allows operating at higher gas flow rates; and (2) the heat generated by the adsorption process is managed by the phase-change material, therefore avoiding large excursions in temperature along the bed. We have verified recently in the work of DeWitt et al. that this thermal management strategy is attainable experimentally (DeWitt et al. 2019). We set the lower bound for the operating pressure as (0.01 atm) for this application of flue gas carbon capture, because it has been found to be challenging to achieve vacuum pressures with the state-of-the-art evacuation systems below this value (Pai et al. 2021). Undoubtedly, one disadvantage of operating in elevated pressures is higher energy consumption, for which we only impose an upper bound constraint. To explore the potential of this technology, in this work we provide the optimizer with a larger range of pressures and allow the solver to identify optimal pressures and PCM content in the fiber blend that maximize productivity.

In order to evaluate process performance, we use four standard metrics for separation processes: productivity $\left[\frac{mol \text{ CO}_2 \text{ in product stream}}{kg \text{ adsorbent} \cdot s} \right]$, product purity $\left[\frac{mol \text{ CO}_2 \text{ in product stream}}{total \text{ mol in product stream}} \right]$, product recovery, $\left[\frac{mol \text{ CO}_2 \text{ in product stream}}{mol \text{ CO}_2 \text{ in Ad step, in}} \right]$ and specific energy consumption $\left[\frac{kWh}{tonne \text{ CO}_2} \right]$. Additional complexity in the simulation and optimization of VPSA processes results from having to evaluate these metrics at CSS. The transient period before this condition is attained can take several hundreds or even thousands of cycles in some cases, with the associated computational burden. Moreover, a challenging feature encountered by process optimization of cyclic adsorption processes lies in the fact that there exists an inherent trade-off between performance objectives. This is traditionally addressed by applying multi-objective optimization techniques (Rangaiah and Bonilla-Petriciolet, 2013). The alternative approach we present herein formulates a constrained single-objective optimization task that maximizes the CO₂ productivity of the VPSA cycle, while introducing purity, recovery, and specific energy consumption as constraints to the problem. This is a suitable approach, since there exist well-established specifications adopted worldwide by government agencies that a separation technology for CO₂ capture should fulfill: CO₂ purity of at least 95% and CO₂ recovery of 90%. The energy constraint is needed since the specific energy consumption is directly correlated to the operating cost of the process.

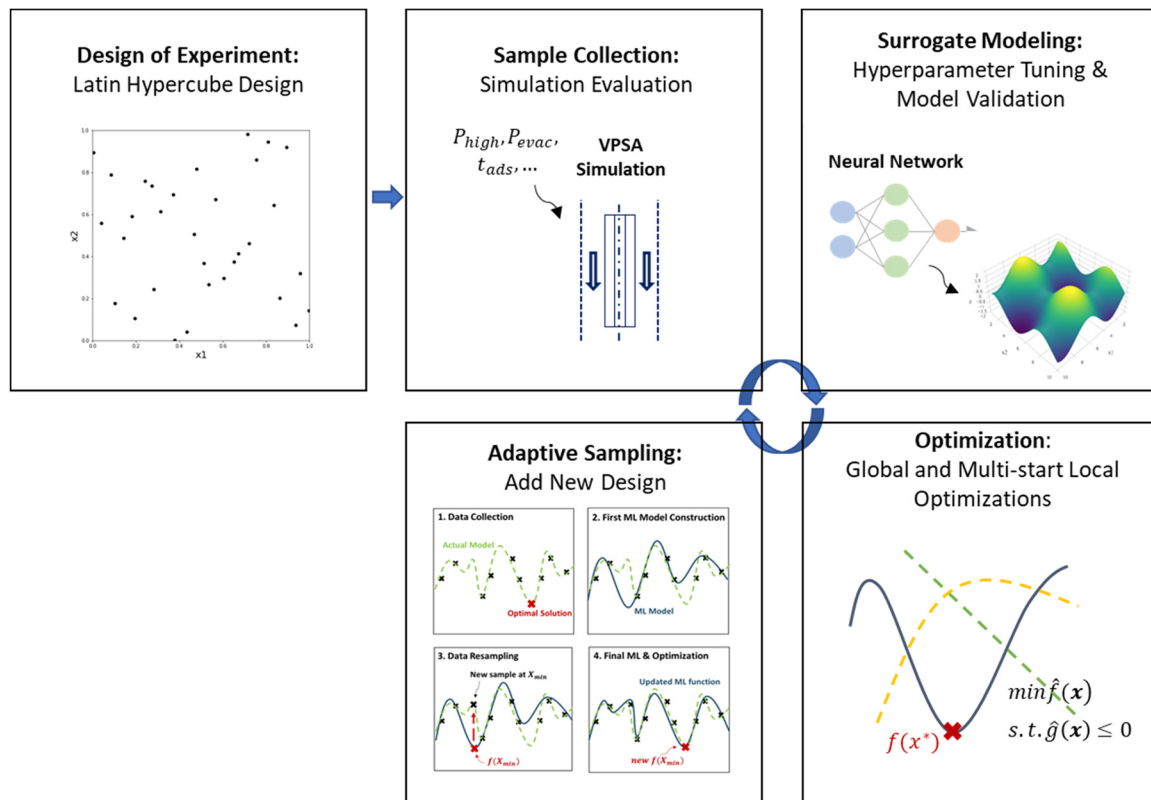


Fig. 3 – Overview of surrogate-based optimization for VPSA process optimization. These steps are repeated until one of the convergence criteria is met.

2.2. Overview of surrogate-based optimization

Due to the complexity of the VPSA-cycle simulation, surrogate-based optimization is an attractive alternative to equation-based optimization. We briefly introduce the general surrogate-based optimization framework and later illustrate in detail how it is adapted for the investigated VPSA system. We start by formulating the general optimization problem as follows (P1):

$$(P1) \min f(x, y)$$

$$s. t. \quad g_c(x, y) \leq 0, \quad c = 1, \dots, C$$

$$x_i^l \leq x_i \leq x_i^u, \quad x_i \in \mathbb{R}, y_j \in \{0, 1\}, \quad i = 1, \dots, I, \quad j = 1, \dots, J$$

where x_i represents a continuous variable, x_i^l and x_i^u are lower and upper bounds of x_i , respectively; y_j represents a binary variable; $f(x, y)$ represents the objective function, and $g_c(x, y)$ represents each unknown constraint. Using surrogate modeling, we seek to obtain $\hat{f}(x, y)$ and $\hat{g}_c(x, y)$ that approximate $f(x, y)$ and $g_c(x, y)$. If only continuous variables exist (i.e., $J = \emptyset$), P1 is a constrained nonlinear programming (NLP) problem; otherwise, P1 is a constrained mixed-integer (binary) nonlinear programming (MINLP) problem.

Surrogate-based optimization has been studied extensively for NLPs and several algorithms currently exist (Forrester and Keane, 2009, Hasan et al. 2013, Cozad et al. 2014, Boukouvala et al. 2017, Bhosekar and Ierapetritou, 2018, Hüllen et al. 2019, Kim and Boukouvala, 2019). Unlike surrogate-based optimization for NLP, only a few works currently exist on surrogate-based optimization for MINLP (Holmström et al. 2008, Mülle et al. 2013, Rashid et al. 2013, Müller 2016, Kim and Boukouvala, 2020). Existing surrogate-based optimization algorithms generally consist of four main steps: (1)

initial sampling, (2) surrogate model construction, (3) optimization, and (4) adaptive sampling to locate promising solutions and update the surrogate model. When binary variables are present in addition to continuous variables, a black-box MINLP (bb-MINLP) algorithm described in (Kim and Boukouvala, 2020) can be used. The algorithm follows a decomposition approach: first, surrogate-based optimization is performed to locate the optimal binary solution y^* ; the binary variables are then fixed at y^* to proceed with a NLP surrogate-based optimization search with respect to continuous variables x . We have previously shown in (Kim and Boukouvala, 2020) that this approach is capable of locating the global solutions of MINLP benchmark problems with up to 12 continuous and 8 binary variables, and some of its key features are briefly described below.

First, an efficient sampling strategy must be used to construct an initial design of experiment (DOE). A good space-filling design should satisfy two conditions: 1) samples should be uniformly, but not regularly, distributed in the search space, and 2) when sample points are projected onto each variable axis, the projections should not overlap (Forrester and Keane, 2009). Several DOE techniques currently exist, such as Latin Hypercube Design (McKay, Beckman et al. 1979), and Sobol Sequences. In this work, we fix our initial design to LHD because it has been validated in previous works as a good starting point (Forrester and Keane, 2009). While different DOE techniques could have minor impacts on the overall result, this is outside the scope of this work. For bb-MINLP, we have previously shown that creating a balanced LHD for each discrete level is the best sampling strategy, and we have proposed heuristics on the size of the initial LHD for each discrete level such that we minimize the overall sampling requirements (Kim and Boukouvala, 2020).

Second, a surrogate model is constructed for the objective and all constraints. There currently exist several types of surrogate models, ranging from simpler models, such as linear and quadratic regression, to more complicated models, such as Gaussian Process (GP), Neural Network (NN), and Support Vector Regression (SVR) models (Jones et al. 1998, Smola and Schölkopf 2004, Heaton, 2008, Forrester and Keane, 2009). The surrogate models, $\hat{f}(x, y)$ and $\hat{g}_c(x, y)$, are low-fidelity representations of an actual high-fidelity simulation $f(x, y)$ and $g_c(x, y)$, thereby reducing the complexity of the optimization model. When binary variables are present, the surrogate model can be constructed either using a relaxed approach or a mixed-integer approach (Kim and Boukouvala, 2020). A relaxed surrogate model assumes all model inputs are continuous; a mixed-integer surrogate model uses a data preprocessing technique called “one-hot encoding” to explicitly handle the discrete search space. The details of relaxed and mixed-integer surrogate models have been published in (Kim and Boukouvala, 2020). In this work, we use a relaxed NN surrogate model to achieve a balance between model accuracy and complexity. While a mixed-integer surrogate model is more accurate than a relaxed model, the use of one-hot encoding increases the dimension of the problem by two as each binary variable is converted to two dummy variables. For this work, we expect that the cost of doubling the dimension would be more significant than the slight loss of model accuracy when relaxing the discrete inputs. Due to the large dimensionality of the binary space, the mixed-integer sampling and surrogate fitting approach would be intractable, therefore, the relaxed approach was chosen. This is a limitation of the current implementation of our method, and the general recommendation is to apply a “relaxed-surrogate” option for any problem with more than 15–20 binary variables.

Within a neural network, the input layer represents input variables, the output layer represents the objective and all constraints, and these two layers are connected by a hidden layer. The number of hidden layers plays an important role in the overall NN architecture and the number of neurons in the hidden layer affect the fitting. If too few neurons are used in the hidden layers, it might result in underfitting and similarly using too many neurons in the hidden layers might result in overfitting. Most of the conventional problems require no more than two hidden layers (Hush and Horne, 1993, Xiao et al. 2020). It should be noted that determining the number of hidden layers and the neurons in the hidden layers is subject to the availability of the data. Although very extensive grid search or optimization can be performed to identify the optimized hyperparameters given the training data, here we utilize the heuristic rule (Heaton, 2008), where the number of nodes is 2/3 of the number of input nodes plus the number of output nodes. A hyperbolic tangent function ($\tanh(x)$) is used as activation function for the input-hidden layer. For the final layer, the identity activation function ($\sigma(x) = x$) is used (Schweidtmann and Mitsos, 2018). These decisions reduce the computational cost of the overall algorithm, since surrogates are used as approximate models to guide further sampling. The final functional form of a neural network with a single hidden layer is:

$$\hat{f}_{NN}(x) = \sigma_i \left(\sum_l W_l^{(1)} \sigma_h \left(\sum_h W_h^{(0)} x + b^{(0)} \right) + b^{(1)} \right)$$

where h and l represent the number of nodes in hidden and output layers, respectively, σ represents an activation

function, $W^{(n)}$ and $b^{(n)}$ are the weights and bias values for input-hidden ($W^{(0)}$ and $b^{(0)}$) and hidden-output ($W^{(1)}$ and $b^{(1)}$) layers.

After neural network surrogate models are constructed, these models are then optimized using deterministic local and global optimization solvers, and the accuracy of the obtained solutions (x^*, y^*) are validated. Adaptive sampling is used to update the surrogate models and further refine the optimal solution. Steps 2–4 are repeated until we find a solution that meets one of the termination criteria: (1) constraint violation is less than 10^{-5} , (2) no improvement in the objective value over ten consecutive iterations, and (3) the maximum number of sample evaluation is reached. In the case where inputs are mixed-integer (bb-MINLP), a MINLP search is first performed and a binary optimal solution is identified; with the binary solution fixed, the NLP stage is then performed to refine the solution with respect to the continuous variables. On the other hand, when all inputs are continuous (bb-NLP), a single search procedure is required. Details of the algorithm have been published in (Boukouvala and Floudas, 2017, Kim and Boukouvala, 2019, Kim and Boukouvala, 2020).

2.3. Problem Formulation and Surrogate-Based Optimization for MINLP and NLP

In this section, we provide an in-depth explanation on surrogate modeling and optimization specific to the VPSA process. We consider 75 adsorbents investigated by Khurana & Farooq (Khurana and Farooq, 2016) and originally addressed by Huck et al. (Huck et al. 2014) for post-combustion CO₂ capture. The applied extended dual-site Langmuir adsorption isotherm model in their work (Khurana and Farooq, 2016) consists of 15 parameters specific to each adsorbent.

One main contribution of this work is the use of black-box MINLP algorithm for efficient process optimization and adsorbent selection. Unlike the brute-force approach, where each adsorbent is optimized separately using black-box NLP algorithm, binary variables are used to represent an adsorbent such that simultaneous process optimization and adsorbent selection can occur. Even though the bb-MINLP solver initially requires a set of LHD for each adsorbent, further exploration of the material (i.e., binary space) is limited to adsorbents (i.e., nodes) that have promising behavior. Thus, the hypothesis is that this simultaneous optimization approach will overall exploit process-material relationships and find optimal solutions with fewer samples by avoiding to sample for non-promising adsorbents. At the end of the MINLP stage of the algorithm, all materials are ranked based on a heuristic that incorporates both average feasibility and objective function values. The algorithm then automatically selects which of the materials (i.e., binary combinations) to further refine in the NLP stage. Although we do not make any claims about global guarantees of optimality using our approach, in this case the algorithm was able to identify the optimal material option. More details about the algorithmic heuristics can be found in (Kim and Boukouvala, 2020). In order to test the performance of the bb-MINLP algorithm with respect to solution accuracy and computational efficiency, we solve this problem using both the bb-MINLP and brute-force NLP approaches. The following optimization formulation (P2) is used to identify the optimal VPSA design that maximizes productivity subject to 95%

purity and 90% recovery constraints with an upper bound on the specific energy consumption of 2000 kWh/tonne CO₂:

$$(P2) \max \text{Productivity} = f(P_{\text{high}}, P_{\text{evac}}, t_{\text{ads}}, \omega_{\text{ads}}, Q_{\text{feed}}, y_{\text{ads}}^k)$$

$$s. t. \text{Purity} = g_1(P_{\text{high}}, P_{\text{evac}}, t_{\text{ads}}, \omega_{\text{ads}}, Q_{\text{feed}}, y_{\text{ads}}^k) \geq 0.95$$

$$\text{Recovery} = g_2(P_{\text{high}}, P_{\text{evac}}, t_{\text{ads}}, \omega_{\text{ads}}, Q_{\text{feed}}, y_{\text{ads}}^k) \geq 0.9$$

$$\text{Specific Energy Consumption} = g_3(P_{\text{high}}, P_{\text{evac}}, t_{\text{ads}}, \omega_{\text{ads}}, Q_{\text{feed}}, y_{\text{ads}}^k) \leq 2000$$

$$2.5 \leq P_{\text{high}} \leq 20, \quad 0.01 \leq P_{\text{evac}} \leq 0.5, \quad 15 \leq t_{\text{ads}} \leq 120,$$

$$0.15 \leq \omega_{\text{ads}} \leq 0.5, \quad 0.001 \leq Q_{\text{feed}} \leq 0.0075$$

$$y_{\text{ads}}^k \in \{0, 1\}, \quad \sum_{k \in \text{adsorbents}} y_{\text{ads}}^k = 1, \quad k = 1, \dots, K$$

Binary variables y_{ads}^k represent an adsorbent k , where $k \in$ set of 75 adsorbents (i.e., $K = 75$). If $y_{\text{ads}}^{k'} = 1$, then adsorbent k' is selected. An additional constraint $\sum_k y_{\text{ads}}^k = 1$ is needed to allow the selection of a single adsorbent. For the brute-force approach, $k = \emptyset$; hence, the problem reduces to NLP.

3. Results and discussion: adsorbent selection and process optimization

In this section, we present the optimization results for adsorbent selection and process optimization using both bb-NLP and bb-MINLP approaches. The bb-NLP approach performs a separate optimization for *each* adsorbent. The bb-MINLP approach, on the other hand, conducts a single optimization for simultaneous adsorbent selection and process optimization. These results are compared with respect to computational efficiency and solution accuracy.

The optimization framework is written in Python. The VPSA simulation is coded in Matlab version 2018b (MATLAB, 2018), and a Python-Matlab interface is applied. Since the simulation is computationally expensive, 5 processors are used. A neural network surrogate model is constructed with the Python module 'scikit-learn', and the optimization is performed using GAMS. At each iteration, all local and global optimal solutions are collected using DICOPT (Drud, 1994) (or CONOPT (Drud, 1994) for NLP) and BARON (Tawarmalani and Sahinidis, 2005) solvers, respectively. The maximum allowed computation time is 50 h. The optimization is repeated 3 times, and the best result is reported. Because the surrogate model we use is non-convex, this repetition was done to choose the best possible solution out of the three runs. Although we cannot guarantee the global solution will be identified, multistart techniques increase the chance of locating a better local solution. We have observed that in certain cases, different local solutions are very similar, while in others there is significant variation especially in the input space (up to 90% relative error), often without significantly affecting the optimal objective function. This variation is an indication that the landscape is nonconvex, and/or that there may be multiple combinations of inputs that provide similar objective function values. This dependence on initialization is a challenge for many local sampling-based methods, and as a result, multi-start techniques are advised when computational resources allow it. All results and analysis of their variation are provided in the Supplementary File–Jupyter notebook. One way to reduce the variation is to generate an

accurate surrogate, which can later be used for optimization with deterministic solvers. But it should be noted that generating an accurate surrogate is itself associated with computationally expensive hyperparameter tuning and requires large amount of data. The requirement on large amount of data can be mitigated to a certain extent if the known physics about the system is used in training the surrogate model. This approach is usually referred to as Hybrid Modeling and Physics-Informed Machine Learning and has attracted a lot of attention over the past few years (Bradley et al. 2022). However, in the present study we focus more on surrogate-based MINLP approach and leave Hybrid Modeling as a possibility for improvement and future work.

Table 1 summarizes the main difference between MINLP and NLP strategies. It should be noted that each NLP problem has a dimensionality of 5 and is provided with 51 initial samples, while the MINLP formulation has a dimensionality of 80 and is provided with 3875 samples (51 for each material). This implies that the two approaches are given an “equivalent” number of initial samples, if normalized by the dimensionality. This allows us to compare the performance of the methods.

3.1. bb-NLP Optimization: Brute-Force Approach

Process optimization is performed separately for all 75 adsorbents using a brute-force NLP approach. For each adsorbent, an initial LHD of size 51 ($10^l + 1$, where $l = 5$) is created. A single layer neural network is constructed with 5 input, 8 hidden, and 4 output nodes using the same heuristic (Heaton, 2008) to predict 4 process performance metrics (i.e., productivity, purity, recovery, and specific energy consumption). Latin Hypercube sampling was selected to ensure the space was maximally covered with our initial sampling budget. We also observe that a small fraction of our initial samples are feasible with respect to all constraints of P2, which implies that the feasible region of the search space is very small and finding a feasible optimal solution is challenging.

Table 2 shows the optimal process conditions and the performance of all feasible adsorbents ranked in the order of decreasing productivity. Fig. 4 shows scatterplots of optimal productivity vs. 5 process inputs of all 75 adsorbents. From the optimization result of 75 adsorbents (Table 2. Supporting Information and Fig. 4), we can notice a few trends. In terms of the operating pressure, feasible adsorbents tend to cluster at high P_{high} and low P_{evac} . Operating the VPSA cycle at a higher P_{high} allows increased adsorption of CO₂ during the pressurization step; operating at a lower P_{evac} improves the recovery of CO₂ and adsorbent regeneration, while sustaining the CO₂ product purity target. For the feed gas flow rate, feasible adsorbents tend to cluster at lower Q_{feed} , since Q_{feed} should allow sufficient contact time for the gas uptake to occur. In the case of the adsorption time, t_{ads} , the trend is less obvious, since both feasible and infeasible adsorbents tend to converge to the lower bound of t_{ads} , enabling shorter cycle times and therefore better productivity results. While the overall trend is also less obvious for the content of adsorbent in the fiber composite ω_{ads} when all 75 adsorbents are considered, we found an interesting trend among only the feasible adsorbents. As shown in Table 2, Supporting Information the productivity of an adsorbent is inversely proportional to ω_{ads} . In fact, two adsorbents that significantly

Table 1 – Comparison of MINLP and NLP strategies.

Approach	# Problems	Dimension	LHD size	ANN structure	Optimization solvers
bb-MINLP	1	80	3825	80–58–4	Baron, Dicopt
bb-NLP	75	5	51	5–8–4	Baron, Conopt

Table 2 – Optimal process conditions for 26 feasible adsorbents that meet 95% Purity - 90% Recovery constraints. The adsorbents are listed in the order of decreasing productivity. Detailed cyclic-steady state profiles for the Top 3 adsorbents in this list can be consulted in [Supporting Information](#) along with the profiles for zeolite 13X and Cu-BTC, both commercially available adsorbents and also listed below. In the last row, the optimal result of the simultaneous MINLP approach is listed for the optimal material identified by the solver.

Optimal material and solution identified by bb-NLP solver:									
Optimal Process Conditions						Performance Measurements			
	P_{high} (atm)	P_{vac} (atm)	Q_{feed} (1e-3 m ³ /s)	t_{ads} (s)	ω_{ads}	Productivity (1e-2 mol/kg's)	Purity (%)	Recovery (%)	Energy ($\frac{MJ}{kg CO_2}$)
Zn-MOF-74	20	0.069	1.78	17.85	0.150	4.91	95.10	90.06	860
UTSA-16	19.92	0.088	1.05	25.29	0.150	4.25	95.57	90.22	828
MgX	20	0.036	1.03	24.88	0.150	2.66	95.11	91.18	857
Co-MOF-74	20	0.029	1.05	38.16	0.150	2.43	95.11	90.06	876
ZIF-39-DIA	19.83	0.072	1.38	15.42	0.331	1.96	95.61	90.06	881
NAB	20	0.052	1.02	15.00	0.280	1.85	95.17	90.04	904
h8155527	20	0.074	1.22	15.00	0.348	1.74	95.05	92.51	833
ZIF-82	20	0.033	1.00	30.01	0.291	1.59	95.06	90.09	909
13x	13.46	0.054	1.20	29.17	0.293	1.56	95.13	90.39	708
HMOF-992	20	0.031	1.00	26.74	0.298	1.56	95.16	90.12	950
ZIF-69	20	0.035	1.06	27.25	0.341	1.44	95.05	90.24	921
ZIF-116-MER	20	0.035	1.00	42.49	0.360	1.38	95.40	90.01	929
ZIF-78	20	0.042	1.00	34.69	0.362	1.31	95.00	90.20	875
CaX	20	0.025	1.00	36.01	0.226	1.25	95.05	90.00	877
NaA	20	0.078	1.00	32.90	0.478	1.25	95.52	91.08	829
Al-X	20	0.035	1.00	31.58	0.313	1.22	95.03	91.01	854
CuBTtri	20	0.040	1.02	68.39	0.470	1.21	95.76	90.35	909
ZIF-36-FRL	18.46	0.047	1.00	26.47	0.365	1.16	95.01	90.09	820
CuBTC	20	0.010	1.00	42.32	0.231	1.15	95.06	95.08	995
Na-X	19.88	0.035	1.06	36.51	0.408	0.96	95.40	90.11	865
Mg-X	15.82	0.051	1.00	51.58	0.500	0.88	95.39	90.04	763
h8124767	8.89	0.032	2.20	39.71	0.400	0.82	95.18	91.45	572
ZIF-68	18.62	0.017	1.00	36.24	0.445	0.79	95.03	92.54	932
ZIF-81	19.24	0.018	1.00	40.50	0.500	0.66	95.23	90.02	936
h8291835	19.83	0.015	1.13	18.12	0.500	0.65	95.10	90.21	996
NaX	13.68	0.029	1.00	71.78	0.500	0.57	95.02	92.72	710
Optimal material and solution identified by bb-MINLP solver:									
Zn-MOF-74	19.83	0.064	1.23	24.73	0.15	4.81	95.97	90.33	853

outperform the rest of the adsorbents – Zn-MOF-74 and UTSA-16 – converge to the lower range of ω_{ads} , and as the optimal productivity decreases, ω_{ads} increases. Since the fiber loading factor is fixed, decreasing adsorbent content in the fiber composite is equivalent to increasing its PCM content. This achieves better temperature modulation (Rubiera Landa et al. 2020). Hence, for adsorbents capable of high CO₂ uptake, the presence of PCM in the fiber composites can further enhance adsorption by effectively managing the heat generated. The optimum P_{high} values reported by the algorithm for all the materials vary from 2.5 – 20 atm. But a closer look shows that for most of the feasible adsorbent materials the optimum P_{high} found is near to 20 atm. This is because high adsorption pressure helps improve the recovery and productivity. But it should be noted that there is also a tradeoff because, high adsorption pressures might result in high energy consumption and hence higher cost.

3.2. Bb-MINLP Optimization: Simultaneous Approach

Previously, we used a surrogate-based NLP algorithm to conduct process optimization of 75 adsorbents separately. This optimal set of results provides a valuable insight into adsorbent performance when coupled with data analytics techniques, which we explore in Section 4. However, requiring $K = 75$ optimization problems to be solved to convergence, this brute-force approach is computationally expensive. However, if we treat this as a bb-MINLP problem, binary variables, y_{ads}^k , are used to represent 75 adsorbents, and (P2) is solved as a single problem. To create a balanced initial sample set, an initial LHD design with a size of $10/\sqrt{K} + 1$ is constructed for each adsorbent k . These LHD sets are then combined to construct a final sample set with $K(10/\sqrt{K} + 1)$ points. With y_{ads}^k , the neural network now has $K + 5$ input

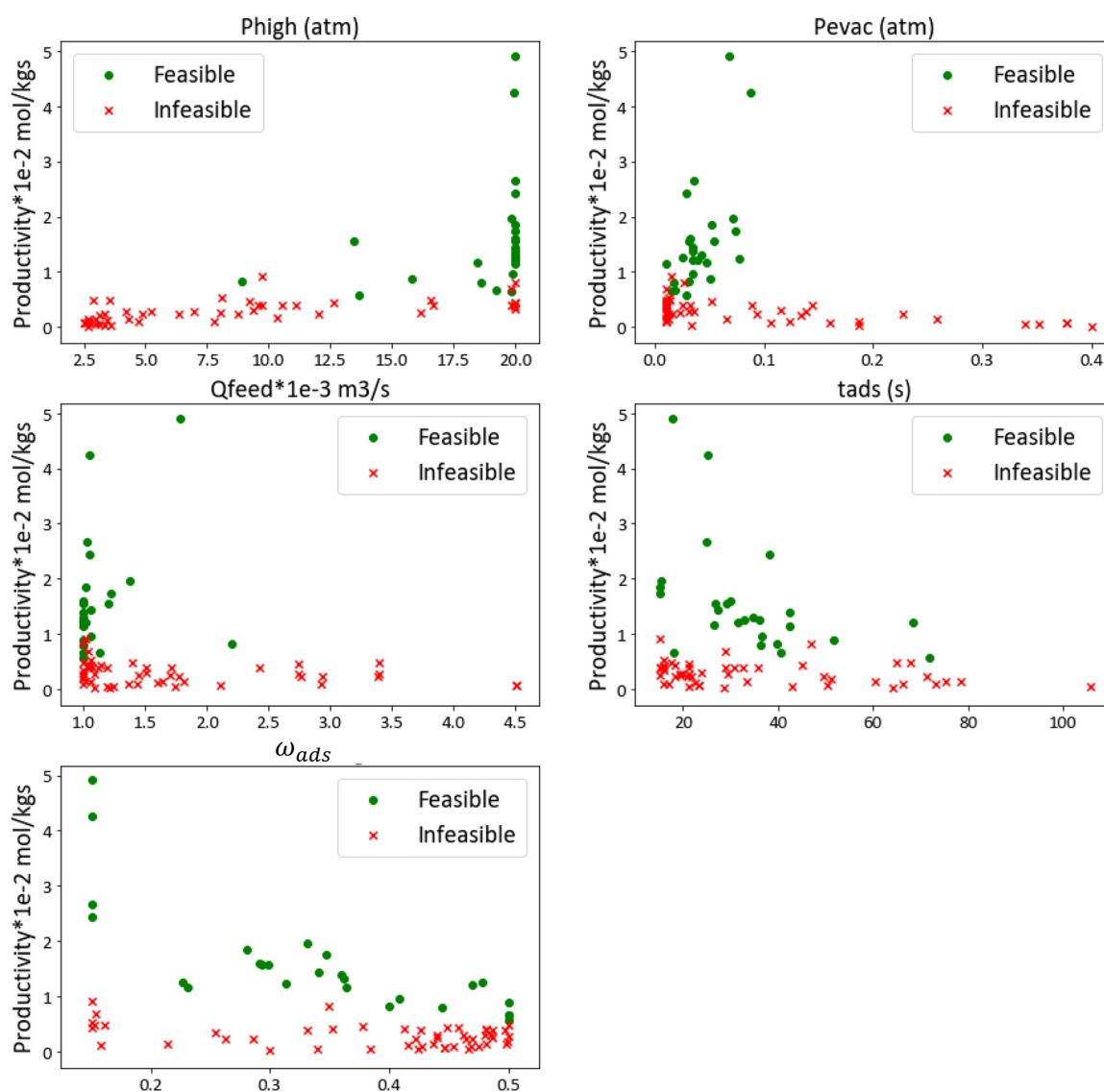


Fig. 4 – Optimal productivity is plotted with 5 process inputs (P_{high} , P_{vac} , Q_{feed} , t_{ads} , and ω_{ads}). Feasible adsorbents (green) satisfy the required purity-recovery-energy constraints.

nodes, 4 output nodes, and one hidden layer. The MINLP search step preliminarily determines the most promising adsorbent y_{ads}^k , while the NLP search is then performed for the most promising adsorbent to further refine the solution with respect to process conditions: P_{high} , P_{vac} , t_{ads} , Q_{feed} , ω_{ads} . In the last row of Table 2. Supporting Information, the optimization result of the simultaneous MINLP approach is shown. The optimal adsorbent is Zn-MOF-74 with the optimal productivity of 0.0481 mol/kg.s. Note that for all the bb-NLP and bb-MINLP solutions, the final purity and recovery values converged slightly above 95% and 90%, respectively. This is potentially due to small approximation errors of regression models that lead to solutions that are not exactly on the feasible boundary.

3.3. Comparison between bb-NLP and bb-MINLP Approaches

A comprehensive comparison of bb-NLP and bb-MINLP approaches with respect to sampling and computational requirements is presented in this section. The bb-NLP

approach is computationally expensive, solving 75 individual surrogate-based optimization tasks. While this computational cost can be alleviated to some extent by using parallel computing, this computational resource may not always be available. The bb-MINLP approach is computationally more efficient since it only requires a single surrogate-based optimization to be performed and exploits the material-process search space efficiently by avoiding re-sampling of non-promising materials.

Table 3 shows the comparison of the sampling requirement and computation time between bb-NLP and bb-MINLP approaches. For the bb-NLP approach, we first report the total computation time for the all 75 adsorbents, assuming that the optimization is performed sequentially. The computation times of each stage of this optimization strategy — sampling, model fitting, and optimization — are shown. As expected, the bb-MINLP approach requires about 30 times less computation time than the bb-NLP approach. When the computation time of the three stages is compared, the most computationally expensive stage consists of collecting samples from the VPSA simulation. In fact, most of the

Table 3 – Comparison of computational cost between bb-NLP and bb-MINLP approaches. For bb-NLP, both the total and average computation time are reported to sequentially optimize all 75 adsorbents and a single adsorbent, respectively.

	# samples	Sampling (hr)	Modeling (hr)	Optimization (hr)	Total
bb-NLP (total)	29335	893.76	9.53	5.76	909.04
bb-NLP (average)	391	11.92	0.13	0.08	12.12
bb-MINLP	5447	13.41	1.72	11.46	26.59

computation time is spent during the sample collection stage, and the model fitting and optimization stages are significantly faster.

In Table 3, the average computation time per adsorbent is also reported for bb-NLP. Looking at the average allows us to compare the computational efficiency of the algorithm for a single optimization. On average, the bb-NLP approach requires less number of samples and computation time for a single adsorbent. This is expected since the bb-NLP approach constitutes a simpler problem with 5 input variables only. The bb-MINLP approach requires more samples and computation time because it is a more challenging problem with 80 input variables. In particular, the optimization stage contributes to about 43% of the total computation time, while that of the brute-force approach is less than 1%. This is also expected since the deterministic optimization of a MINLP problem is more difficult than that of the NLP problem. Nevertheless, when the total computation requirement is considered, we can conclude that the bb-MINLP approach is computationally more efficient overall.

It must be noted that if high-performance computing capabilities are available, the brute-force approach can be parallelized, which would lead to significant compute cost savings. However, what we think is interesting in the simultaneous approach is the fact that the optimizer is able to screen through infeasible adsorbents in an automated way (i.e., via the MINLP search), and as a result it requires a significantly less amount of data to find optimal solutions.

In terms of solution accuracy, both the bb-MINLP and bb-NLP approaches identified Zn-MOF-74 as the best adsorbent with productivity 0.0481 mol/kg.s and 0.0491 mol/kg.s , respectively. While the MINLP search stage correctly identifies the optimal adsorbent, the NLP stage can lead to a slightly different optimal result, which is typical when applying surrogate-based approaches due to stochasticity caused by different sampling locations and model training.

4. Analysis of Adsorbent-Process Interaction using Data Analytics and Machine Learning

In Section 3, we collect all results obtained from the previous analysis (optimal process conditions and module design) for all 75 adsorbents. The different adsorbent materials are characterized by different extended Dual-Site Langmuir equilibrium parameters. Our aim here is to perform some analysis on the merged process-material data and gain further insight into the correlations between process design and adsorbent. We first compute a correlation matrix to observe how all process features and isotherm features are correlated. We then perform Principal Component Analysis (PCA) to handle highly correlated variables and observe the importance of process and adsorbent features. Finally, we construct machine learning-based classification and regression models that allow us to predict adsorbent feasibility and performance.

4.1. Correlation Matrix of Isotherm and Process Features

A correlation matrix is constructed to observe how the input variables are correlated. Both the isotherm features (i.e., 12 extended Dual-Site Langmuir equation parameters; adsorbent density; & isosteric heats of adsorption) and process features (i.e., optimal operating conditions determined in Section 3.1) of all 75 adsorbents are included in this analysis. Investigating the correlation between independent variables is an important stage in machine learning since multi-collinearity can potentially create difficulty in estimating model parameters (Alin, 2010).

Fig. 5 shows a correlation matrix generated from the 75 adsorbents, the correlation coefficients are calculated using Spearman's correlation. As expected, adsorption equilibrium parameters tend to be highly correlated (e.g., $qsat_{11}$ and $qsat_{12}$) since these parameters are typically estimated by nonlinear fitting of experimental data and solving a system of equations (Khurana and Farooq, 2016). Thus, we can observe higher correlation values in the upper left quadrant of the correlation matrix. On the other hand, operation features tend to be less correlated, because these are design variables that are optimized. Some moderate correlation is still observed among P_{high} , P_{vac} and Q_{feed} as observed in the lower right quadrant of the correlation matrix.

4.2. Principal Component Analysis Results

Principal Component Analysis is a dimensionality reduction technique used for data visualization and handling data multi-collinearity (Hastie, Tibshirani et al. 2013). To observe any pattern or clustering among feasible and infeasible adsorbents, we combine 15 adsorbent features (i.e., isotherm parameters) and 5 optimal process features for 75 adsorbents and perform linear PCA. Note that we have tested several nonlinear kernels (e.g., polynomial, radial basis function, sigmoid, and cosine), but they did not improve the visualization of feasible/infeasible adsorbents. These results are not included in this paper but can be found in the provided Supplementary Material – Jupyter notebook. Fig. 6 shows the cumulative explained variance vs. the number of principal components (PCs). With just 6 PCs, we can explain about 80% of the data variance, which means that we can potentially decrease the dimension of the problem from 20 to 6 and still accurately capture most of the data variance.

For linear PCA, each PC is represented as a linear combination of original features: $PC_p = \sum_{f \in F} w_f (x - \mu_f)$, where p represents the selected PC and F represents all 20 isotherm and process features. We can analyze the importance of each feature by analyzing its weight w_f , where a feature with larger $|w_f|$ is considered more important. To compare the importance of adsorbent and process features for the first six PCs, we computed their percentage contribution, expressed by: $\%Contribution = 100 \cdot \frac{\sum_{f \in F'} |w_f|}{\sum_{f \in F} |w_f|}$, where F' represents a set of

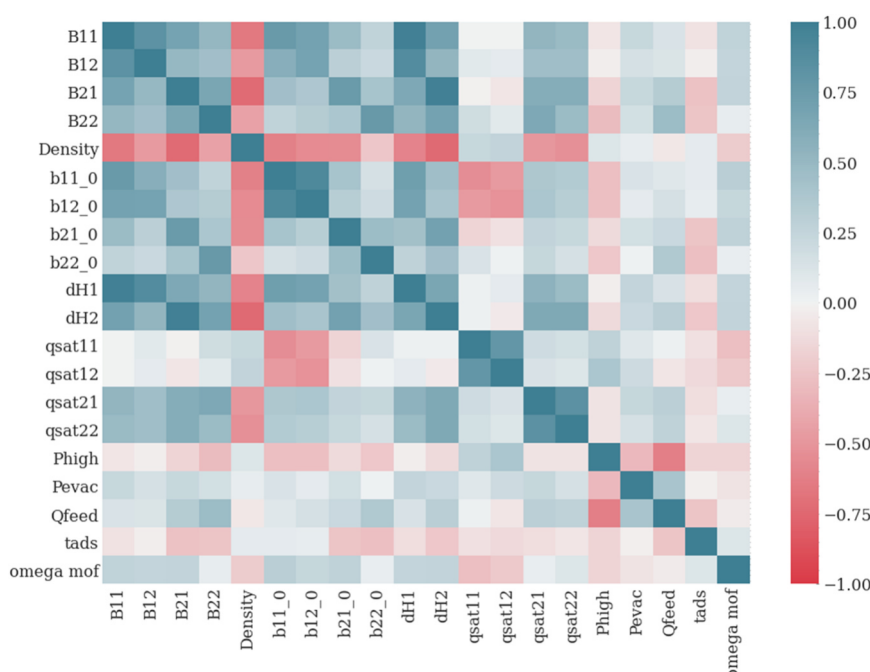


Fig. 5 – Correlation matrix of 15 adsorbent features and 5 process features. The adsorbent features include: 12 eDSL equilibrium parameters ('B11', 'B12', 'B21', 'B22', 'b11_0', 'b12_0', 'b21_0', 'b22_0', 'qsat11', 'qsat12', 'qsat21', 'qsat22'); adsorbent density ('Density'); & isosteric heats of adsorption ('dH1' & 'dH2'). The 5 process features are: adsorption pressure level ('Phigh'); evacuation pressure ('Pevac'); feed volumetric flow rate ('Qfeed'); adsorption time ('tads'); & weight fraction of adsorbent in the fiber composite ('omega_ads'). Darker colors (both blue and red) represent higher correlation.

either adsorbent or process features. Fig. 6 also shows the computed feature contribution. For the first PC, adsorbent features explain about 70% of data variance; for the second PC, adsorbent features explain about 50% of data variance. While adsorbent features seem to contribute slightly more to the PCs than the process features, it is difficult to conclude the existence of a dominant feature. These results support our claim that both adsorbent and process features contribute to the overall data variance and are both important for the VPSA process design. Finally, the first two PCs are plotted to visualize the data in a 2-dimensional space and observe any pattern among feasible and infeasible adsorbents. From Fig. 7, we can observe some clustering of feasible adsorbents in the lower left quadrant of the PC space as well as a linear separation between feasible and infeasible adsorbents. In fact, we have observed that the performance

of an adsorbent improves as we approach the lower left quadrant in the PC space.

4.3. SVM classification model for adsorbent feasibility

Using the PCs obtained in the previous section, we construct a classification model that predicts adsorbent feasibility, assuming that we are operating an adsorbent at optimal operating conditions. As shown in Fig. 7, the feasible and infeasible adsorbents are linearly separable in the 2-dimensional PC space. To exploit this trend, we construct a linear Support Vector Machine (SVM) classification model to classify adsorbent feasibility using 2 PCs. A linear SVM model seeks to find a linear hyperplane that can separate two classes of points and it has a regularization hyperparameter C , which can be tuned using a grid search (Smola and

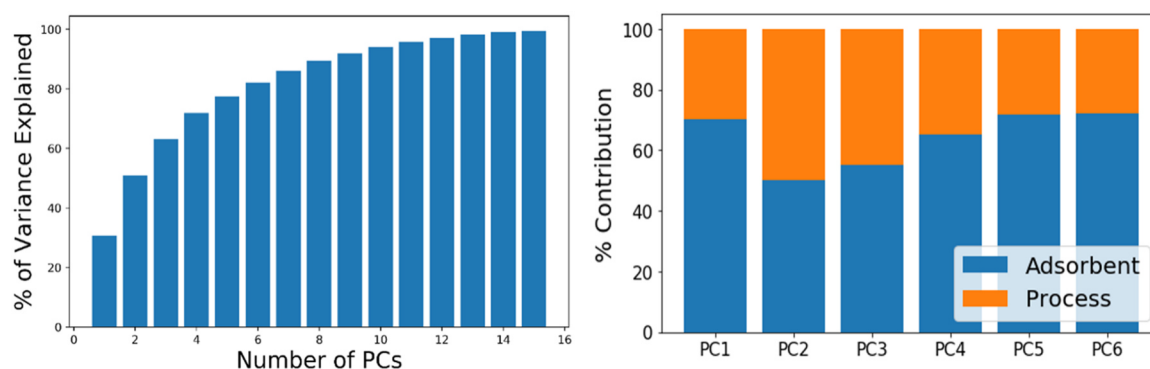


Fig. 6 – Left: Percentage of variance explained vs. the number of principal components. Right: Feature % contribution of the first 6 PCs.

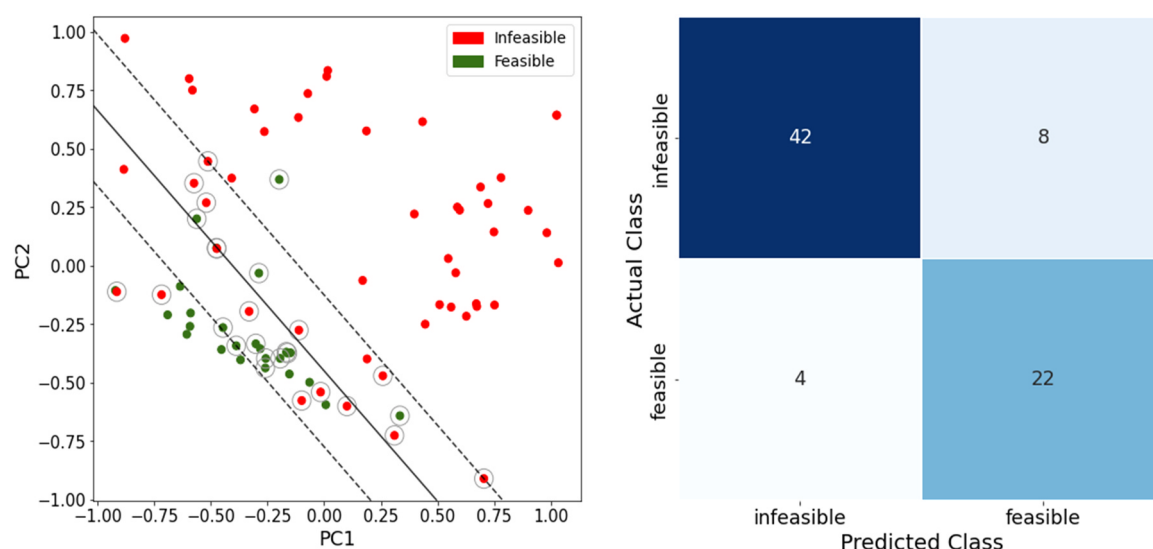


Fig. 7 – PCA is performed and the first two PCs are plotted, where feasible adsorbents that satisfy all constraints are indicated by green dots. The decision boundary found via support vector classification is also displayed by a black solid line. And to the right we show a confusion matrix of the classification.

Schölkopf 2004). For model training, 80% of the data is used with 5-fold cross validation, and 20% of the data set is set aside to test how well the model generalizes to a new set of data. The resulting linear SVM model is also shown in Fig. 7. For the training set, the SVM model accuracy is 82%, and this error results from the fact that the points are not perfectly linearly separable. For the test set, the model accuracy is 94%, which implies good generalizability. While we can generally improve the model accuracy by increasing the number of PCs or considering nonlinear classification techniques, neither of these two approaches significantly improved our results. Specifically, increasing the number of PCs led to less than 1% improvements, while consideration of a nonlinear rbf kernel led to less than 5% improvement, and not perfect separation of feasible to infeasible materials (see Supplementary Material – Jupyter notebook). As a result, we favor simplicity and interpretability of PCs in this work, and report results for the linear case.

4.4. Adsorbent performance prediction model

In the previous section, we constructed a SVM-based adsorbent feasibility classification model using 2 PCs. While this model is sufficient when one is interested in determining whether an adsorbent is feasible or not, it is not sufficient to provide detailed information on the performance of the VPSA system in terms of product productivity, purity, recovery, and specific energy consumption. Thus, we construct a neural network-based adsorbent performance prediction model to predict the VPSA system performance given isotherm parameters and the optimal operating conditions.

Previously in Section 4.1, we have determined that isotherm features (i.e., extended dual-site Langmuir isotherm parameters) tend to be highly correlated, while process features exhibit less correlation. Hence, we perform linear PCA on just the isotherm features to reduce the dimension from 15 to the selected number of PCs and handle existing correlations between isotherm parameters. All of the obtained data from the optimization runs (i.e., initial samples, local and best solutions) were used to train this final model. Out of

a total of 224 samples, 50 points were filtered out prior to training because they were highly infeasible solutions with respect to energy and these would be outliers that would induce bias in the regression model. These isotherm-based PCs are then combined with optimal process features to construct a prediction model. Fig. 8 summarizes the result of linear PCA. Even with just 2 PCs, we can capture 60% of data variance; with 5 PCs, we can capture 80% of data variance. When the first two PCs are plotted (Fig. 8), we did not observe any pattern among feasible and infeasible adsorbents. This further enhances our previous claim that both operation and adsorbent features are important in predicting the feasibility of an adsorbent; thus, both features must be considered.

After reducing the dimensionality of isotherm features using PCA, the first two principal components are combined with optimal process features, which results in 7 inputs (i.e., $PC_{ads,1}$, $PC_{ads,2}$, P_{high} , P_{vac} , Q_{feed} , t_{ads} , and ω_{ads}). These inputs are then used to train a single multi input-output neural network model that can predict product productivity, purity, recovery, and energy consumption. For the selection of hyperparameters of the NN, a grid search is used with 5-fold cross-validation to select the most appropriate activation function—i.e., hyperbolic tangent (tanh) or rectified linear unit (relu); the number of hidden layers (from 1 to 4 layers), and the number of nodes per hidden layer (varied from 5 to 12 per layer). We found that the most challenging output to predict was energy, and we hypothesize that this is due to the fact that the energy constraint is relatively easy to satisfy and is often non-active in the optimal solution. The optimal architecture of the NN from grid search contained 4 hidden layers with hyperbolic tangent activation and 11, 8, 11, 10 nodes per hidden layer respectively. Fig. 9 shows the parity plot of neural network model, and the goodness-of-fit given is by R^2 , and for the train-set were found to be (0.920, 0.918, 0.959, 0.896) and for the test set were found to be (0.951, 0.929, 0.979, 0.958) respectively for productivity, purity, recovery and energy. We have observed that increasing the number of PCs does not significantly improve the model. Therefore, applying only 2 PCs is a good compromise that balances model accuracy and complexity.

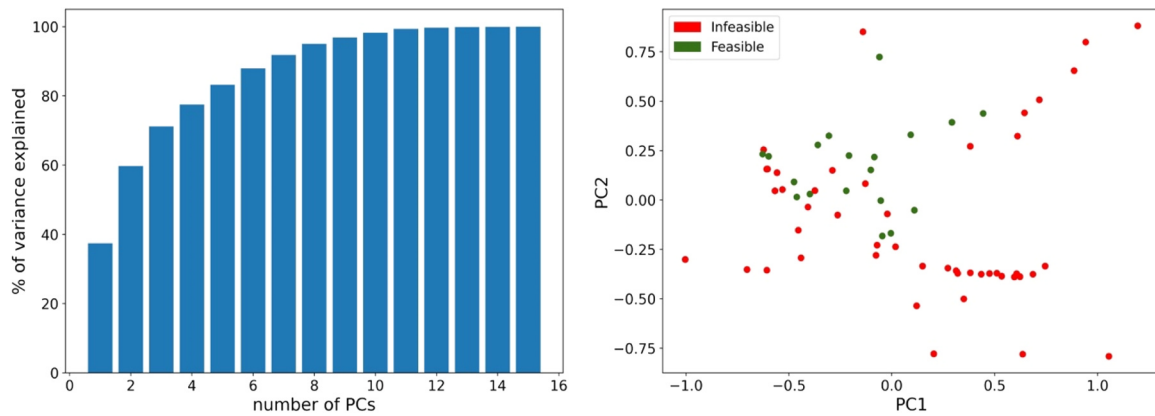


Fig. 8 – PCA on only isotherm parameters (PC_{ads}): (a) Percentage of variance explained vs. the number of PCs, and (b) 2-d representation of PC space.

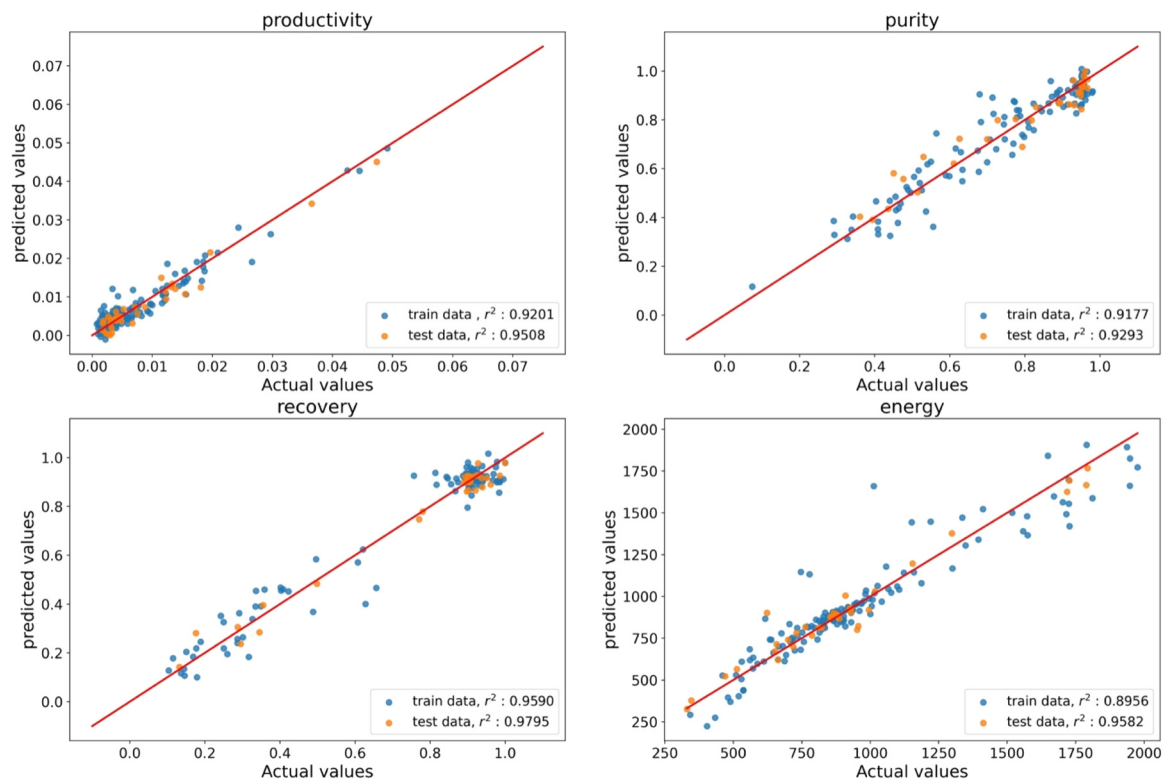


Fig. 9 – Parity plot of productivity, purity, recovery, and energy from the neural network model. The y-axis is the predicted value from a neural network model, and the x-axis is the actual simulation output.

To check the robustness of the model, we randomized the train-test split and fitted the training data using the best architecture we found in the grid search for 6 test cases. We found that the average of R^2 the train-set were found to be (0.928, 0.916, 0.961, 0.905) and for the test set were found to be (0.928, 0.932, 0.968, 0.916) respectively for productivity, purity, recovery and energy. This indicates a good model performance and that the model is not overfitting. This model can be used to preliminarily evaluate the feasibility of an adsorbent using isotherm and process features.

5. Conclusions and future perspectives

In this work, we propose a surrogate-based optimization approach for the design of modular VPSA systems for post-combustion CO_2 capture. The VPSA cycle considers the

application of fixed-beds packed with thermally modulated fiber composites, which enhances mass transfer, minimizes pressure drop, and allows intrinsic thermal management, thereby intensifying carbon capture efficiency. To achieve optimal performance, we consider both the adsorbent selection and process operation conditions to design a modular VPSA system. We investigate two different approaches to formulate and solve the optimization problem: 1) the bb-NLP approach, where the process optimization is performed for each adsorbent, and 2) the bb-MINLP approach, where the adsorbent selection and process optimization are performed simultaneously. For the bb-NLP approach, 75 adsorbents are considered and 26 feasible adsorbents are identified. The bb-MINLP approach performed simultaneous optimization of adsorbent selection and process condition by using binary variables to represent adsorbents. When all 75 adsorbents

are compared, the latter approach is more efficient with respect to sampling and computational requirements.

In addition to the design of a module, we also demonstrate how machine learning classification and regression techniques can be applied to identify feasible adsorbents and predict the performance from a purely data-driven perspective. Principal Component Analysis is used to reduce the dimension of the problem and analyze the importance of adsorbent and process features. Our result indicates that no dominant feature exists and both the adsorbent and process features are important, implying adsorbent selection is highly linked to process performance. A clustering of feasible adsorbents has been observed, and an accurate linear SVM classification is developed. Finally, we construct a neural network-based performance prediction model to predict four outputs of the VPSA simulation (i.e., purity, recovery, energy consumption, and productivity). These classification and neural network models provide valuable preliminary insight into different adsorbents and process performance.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

H. O. Rubiera Landa and M. J. Realff were partially supported by the US DOE through grant: DE-FE0026433 for project: 'Enabling 10 mol/kg swing capacity via heat-integrated sub-ambient pressure swing adsorption'. S.H. Kim and F. Boukouvala would like to acknowledge funding from DOE/AIChE through the RAPID Manufacturing Institute (Synopsis Project; 102623) and NSF-CBET (1805724, 1944678). Any opinions, findings, conclusions or recommendations expressed herein are those of the authors and do not necessarily reflect the view of the DOE.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.cherd.2022.10.002](https://doi.org/10.1016/j.cherd.2022.10.002).

References

- "U.S. EPA's Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990–2017." 2019, from <https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks>.
- Agarwal, A., Biegler, L.T., Zitney, S.E., 2010. A superstructure-based optimal synthesis of PSA cycles for post-combustion CO₂ capture. *AIChE J.* 56 (7), 1813–1828.
- Alin, A., 2010. Multicollinearity WIREs. *Comput. Stat.* 2 (3), 370–374.
- Baldea, M., Edgar Thomas, F., Stanley Bill, L., Kiss Anton, A., 2017. Modular manufacturing processes: status, challenges, and opportunities. *AIChE J.* 63 (10), 4262–4272.
- Ben-Mansour, R., Habib, M.A., Bamidele, O.E., Basha, M., Qasem, N.A.A., Peedikakkal, A., Laoui, T., Ali, M., 2016. Carbon capture by physical adsorption: Materials, experimental investigations and numerical modeling and simulations – a review. *Appl. Energy* 161, 225–255.
- Beykal, B., Onel, M., Onel, O., Pistikopoulos, E.N., 2020. A data-driven optimization algorithm for differential algebraic equations with numerical infeasibilities. *AIChE J.* 66 (10), e16657.
- Bhosekar, A., Ierapetritou, M., 2018. Advances in surrogate based modeling, feasibility analysis, and optimization: a review. *Comput. Chem. Eng.* 108, 250–267.
- Bhown, A.S., Freeman, B.C., 2011. Analysis and status of post-combustion carbon dioxide capture technologies. *Environ. Sci. Technol.* 45 (20), 8624–8632.
- Boukouvala, F., Floudas, C.A., 2017. ARGONAUT: algorithms for Global Optimization of coNstrained grey-box compUTational problems. *Optim. Lett.* 11 (5), 895–913.
- Boukouvala, F., Hasan, M.M.F., Floudas, C.A., 2017. Global optimization of general constrained grey-box models: new method and its application to constrained PDEs for pressure swing adsorption. *J. Glob. Optim.* 67 (1), 3–42.
- Bradley, W., Kim, J., Kilwein, Z., Blakely, L., Eydenberg, M., Jalvin, J., Laird, C., Boukouvala, F., 2022. Perspectives on the Integration between First-principles and Data-driven Modeling. *Comput. Chem. Eng.*, 107898.
- Bui, M., Adjiman, C.S., Bardow, A., Anthony, E.J., Boston, A., Brown, S., Fennell, P.S., Fuss, S., Galindo, A., Hackett, L.A., Hallett, J.P., Herzog, H.J., Jackson, G., Kemper, J., Krevor, S., Maitland, G.C., Matuszewski, M., Metcalfe, I.S., Petit, C., Puxty, G., Reimer, J., Reiner, D.M., Rubin, E.S., Scott, S.A., Shah, N., Smit, B., Trusler, J.P.M., Webley, P., Wilcox, J., Mac Dowell, N., 2018. Carbon capture and storage (CCS): the way forward. *Energy Environ. Sci.* 11 (5), 1062–1176.
- Burns, T.D., Pai, K.N., Subraveti, S.G., Collins, S.P., Krykunov, M., Rajendran, A., Woo, T.K., 2020. Prediction of MOF Performance in Vacuum Swing Adsorption Systems for Postcombustion CO₂ Capture Based on Integrated Molecular Simulations, Process Optimizations, and Machine Learning Models. *Environ. Sci. Technol.* 54 (7), 4536–4544.
- Choi, S., Drese, J.H., Jones, C.W., 2009. Adsorbent materials for carbon dioxide capture from large anthropogenic point sources. *ChemSusChem* 2 (9), 796–854.
- Cozad, A., Sahinidis, N.V., Miller, D.C., 2014. Learning surrogate models for simulation-based optimization. *AIChE J.* 60 (6), 2211–2227.
- Darunte, L.A., Sen, T., Bhawanani, C., Walton, K.S., Sholl, D.S., Realff, M.J., Jones, C.W., 2019. Moving beyond adsorption capacity in design of adsorbents for CO₂ capture from ultradilute feeds: kinetics of CO₂ adsorption in materials with stepped isotherms. *Ind. Eng. Chem. Res.* 58 (1), 366–377.
- DeWitt, S.J.A., Rubiera Landa, H.O., Kawajiri, Y., Realff, M., Lively, R.P., 2019. Development of phase-change-based thermally modulated fiber sorbents. *Ind. Eng. Chem. Res.* 58 (15), 5768–5776.
- Dias, L.S., Ierapetritou, M.G., 2020. Integration of planning, scheduling and control problems using data-driven feasibility analysis and surrogate models. *Comput. Chem. Eng.* 134, 106714.
- Drud, A.S., 1994. CONOPT—a large-scale GRG code. *ORSA J. Comput.* 6 (2), 207–216.
- Ebner, A.D. and J.A. Ritter (2009). "State-of-the-art adsorption and membrane separation processes for carbon dioxide production from carbon dioxide emitting industries."
- Farmahini, A.H., Krishnamurthy, S., Friedrich, D., Brandani, S., Sarkisov, L., 2021. Performance-based screening of porous materials for carbon capture. *Chem. Rev.* 121 (17), 10666–10741.
- Fiandaca, G., Fraga, E., Brandani, S., 2009. A multi-objective genetic algorithm for the design of pressure swing adsorption." *Eng. Optim.* 41 (9), 833–854.
- Findley, J.M., Ravikovitch, P.I., Sholl, D.S., 2018. The effect of aluminum short-range ordering on carbon dioxide adsorption in zeolites. *J. Phys. Chem. C.* 122 (23), 12332–12340.
- Forrester, A.I.J., Keane, A.J., 2009. Recent advances in surrogate-based optimization. *Prog. Aerosp. Sci.* 45 (1), 50–79.
- Gear, C.W., Gear, W.C., 1971. Numerical Initial Value Problems in Ordinary Differential Equations. Prentice-Hall.

- Haghighpanah, R., Majumder, A., Nilam, R., Rajendran, A., Farooq, S., Karimi, I.A., Amanullah, M., 2013. Multiobjective Optimization of A Four-step Adsorption Process for Postcombustion CO₂ capture via finite volume simulation. *Ind. Eng. Chem. Res.* 52 (11), 4249–4265.
- Haghighpanah, R., Nilam, R., Rajendran, A., Farooq, S., Karimi, I.A., 2013. Cycle synthesis and optimization of a VSA process for postcombustion CO₂ capture. *AIChE J.* 59 (12), 4735–4748.
- Hasan, M.M.F., Boukouvala, F., First, E.L., Floudas, C.A., 2014. Nationwide, regional, and statewide CO₂ capture, utilization, and sequestration supply chain network optimization. *Ind. Eng. Chem. Res.* 53 (18), 7489–7506.
- Hasan, M.M.F., First, E.L., Boukouvala, F., Floudas, C.A., 2015. A multi-scale framework for CO₂ capture, utilization, and sequestration: CCUS and CCU. *Comput. Chem. Eng.* 81, 2–21.
- Hasan, M.M.F., First, E.L., Floudas, C.A., 2013. Cost-effective CO₂ capture based on in silico screening of zeolites and process optimization. *Phys. Chem. Chem. Phys.* 15 (40), 17601–17618.
- Hastie, T., Tibshirani, R., Friedman, J., 2013. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York.
- Heaton, J., 2008. *Introduction to Neural Networks for JAVA*. Heaton Research, Inc.
- Holmström, K., Quttineh, N.-H., Edvall, M.M., 2008. An adaptive radial basis algorithm (ARBF) for expensive black-box mixed-integer constrained global optimization. *Optim. Eng.* 9 (4), 311–339.
- Huck, J.M., Lin, L.-C., Berger, A.H., Shahrak, M.N., Martin, R.L., Bhowan, A.S., Haranczyk, M., Reuter, K., Smit, B., 2014. Evaluating different classes of porous materials for carbon capture. *Energy Environ. Sci.* 7 (12), 4132–4146.
- Hüllen, G., Zhai, J., Kim, S.H., Sinha, A., Realff, M.J., Boukouvala, F., 2019. Managing uncertainty in data-driven simulation-based optimization. *Comput. Chem. Eng.*, 106519.
- Hush, D.R., Horne, B.G., 1993. Progress in supervised neural networks. *IEEE Signal Process. Mag.* 10 (1), 8–39.
- Jang, H.-Y., Johnson, J.R., Ma, Y., Mathias, R., Bhandari, D.A., Lively, R.P., 2019. Torlon® hollow fiber membranes for organic solvent reverse osmosis separation of complex aromatic hydrocarbon mixtures. *AIChE J.* 65 (12), e16757.
- Jones, D.R., Schonlau, M., Welch, W.J., 1998. Efficient global optimization of expensive black-box functions. *J. Glob. Optim.* 13 (4), 455–492.
- Khurana, M., Farooq, S., 2016. Adsorbent screening for post-combustion CO₂ capture: a method relating equilibrium isotherm characteristics to an optimum vacuum swing adsorption process performance. *Ind. Eng. Chem. Res.* 55 (8), 2447–2460.
- Khurana, M., Farooq, S., 2017. Integrated adsorbent-process optimization for carbon capture and concentration using vacuum swing adsorption cycles. *AIChE J.* 63 (7), 2987–2995.
- Khurana, M., Farooq, S., 2019. Integrated adsorbent process optimization for minimum cost of electricity including carbon capture by a VSA process. *AIChE J.* 65 (1), 184–195.
- Kikkinides, E.S., Yang, R.T., Cho, S.H., 1993. Concentration and recovery of carbon dioxide from flue gas by pressure swing adsorption. *Ind. Eng. Chem. Res.* 32 (11), 2714–2720.
- Kim, S.H. and F. Boukouvala (2019). "Machine learning-based surrogate modeling for data-driven optimization: a comparison of subset selection for regression techniques." *Optimization Letters*.
- Kim, S.H. and F. Boukouvala (2020). "Surrogate-Based Optimization for Mixed-Integer Nonlinear Problems." *Computers & Chemical Engineering*.
- Kim, Y.-h, Park, L.K., Yiacoymi, S., Tsouris, C., 2017. Modular chemical process intensification: a review. *Annu. Rev. Chem. Biomol. Eng.* 8 (1), 359–380.
- Ko, D., Siriwardane, R., Biegler, L.T., 2005. Optimization of pressure swing adsorption and fractionated vacuum pressure swing adsorption processes for CO₂ capture. *Ind. Eng. Chem. Res.* 44 (21), 8084–8094.
- Leperi, K.T., Chung, Y.G., You, F., Snurr, R.Q., 2019. Development of a general evaluation metric for rapid screening of adsorbent materials for postcombustion CO₂ capture. *ACS Sustain. Chem. Eng.* 7 (13), 11529–11539.
- Leperi, K.T., Yancy-Caballero, D., Snurr, R.Q., You, F., 2019. 110th anniversary: surrogate models based on artificial neural networks to simulate and optimize pressure swing adsorption cycles for CO₂ capture. *Ind. Eng. Chem. Res.* 58 (39), 18241–18252.
- Li, W., Wei, S., Jiao, W., Qi, G., Liu, Y., 2016. Modelling of adsorption in rotating packed bed using artificial neural networks (ANN). *Chem. Eng. Res. Des.* 114, 89–95.
- Lutze, P., Gani, R., Woodley, J.M., 2010. Process intensification: a perspective on process synthesis. *Chem. Eng. Process.: Process Intensif.* 49 (6), 547–558.
- MATLAB, 2018. Natick, Massachusetts. The MathWorks Inc.
- McBride, K., Sundmacher, K., 2019. "Overview of surrogate modeling in chemical process engineering." *Chem. Ing. Tech.* 91 (3), 228–239.
- McKay, M.D., Beckman, R.J., Conover, W.J., 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21 (2), 239–245.
- Müller, J., 2016. MISO: mixed-integer surrogate optimization framework. *Optim. Eng.* 17 (1), 177–203.
- Müller, J., Shoemaker, C., Piche, R., 2013. SO-MI: a surrogate model algorithm for computationally expensive nonlinear mixed-integer black-box global optimization problems. *Comput. Oper. Res.* 40 (5), 1383–1400.
- Nogueira, I.B.R., Dias, R.O.M., Rebello, C.M., Costa, E.A., Santana, V.V., Rodrigues, A.E., Ferreira, A., Ribeiro, A.M., 2022. A novel nested loop optimization problem based on deep neural networks and feasible operation regions definition for simultaneous material screening and process optimization. *Chem. Eng. Res. Des.* 180, 243–253.
- Pai, K.N., Nguyen, T.T.T., Prasad, V., Rajendran, A., 2022. Experimental validation of an adsorbent-agnostic artificial neural network (ANN) framework for the design and optimization of cyclic adsorption processes. *Sep. Purif. Technol.* 290, 120783.
- Pai, K.N., Prasad, V., Rajendran, A., 2021. Practically achievable process performance limits for pressure-vacuum swing adsorption-based postcombustion CO₂ capture. *ACS Sustain. Chem. Eng.* 9 (10), 3838–3849.
- Rangaiah, G.P., Bonilla-Petriciolet, A., 2013. *Multi-Objective Optimization in Chemical Engineering: Developments and Applications*. Wiley.
- Rashid, K., Ambani, S., Cetinkaya, E., 2013. An adaptive multi-quadratic radial basis function method for expensive black-box mixed-integer nonlinear constrained optimization. *Eng. Optim.* 45 (2), 185–206.
- Regis, R.G., 2020. *High-dimensional constrained discrete expensive black-box optimization using a two-phase surrogate approach*. Springer.
- Regis, R.G. (2020). *Large-Scale Discrete Constrained Black-Box Optimization Using Radial Basis Functions*. 2020 IEEE Symposium Series on Computational Intelligence (SSCI).
- Rezaei, F., Subramanian, S., Kalyanaraman, J., Lively, R.P., Kawajiri, Y., Realff, M.J., 2014. Modeling of rapid temperature swing adsorption using hollow fiber sorbents. *Chem. Eng. Sci.* 113, 62–76.
- Rubiera Landa, H.O., R.P. Lively, Y. Kawajiri and M. Realff (2020). Theoretical investigation of vacuum pressure swing adsorption process applying thermally-modulated fiber composite adsorbents.
- Ruthven, D.M., 1984. *Principles of Adsorption and Adsorption Processes*. Wiley.
- Ruthven, D.M., Farooq, S., Knaebel, K.S., 1994. *Pressure Swing Adsorption*. VCH Publishers.
- Samanta, A., Zhao, A., Shimizu, G.K.H., Sarkar, P., Gupta, R., 2012. Post-combustion CO₂ capture using solid sorbents: a review. *Ind. Eng. Chem. Res.* 51 (4), 1438–1463.

- Sant Anna, H.R., Barreto, A.G., Tavares, F.W., de Souza, M.B., 2017. Machine learning model and optimization of a PSA unit for methane-nitrogen separation. *Comput. Chem. Eng.* 104, 377–391.
- Schiesser, W.E., 1991. *The Numerical Method of Lines: Integration of Partial Differential Equations*. Academic Press.
- Schweidtmann, A.M. and A. Mitsos (2018). "Deterministic Global Optimization with Artificial Neural Networks Embedded." *Journal of Optimization Theory and Applications*.
- Sinha, A., Realff, M.J., 2019. A parametric study of the techno-economics of direct CO₂ air capture systems using solid adsorbents. *AIChE J.* 65 (7), e16607.
- Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. *Stat. Comput.* 14 (3), 199–222.
- Stankiewicz, A. and J.A. Moulijn (2000). *Process Intensification: Transforming Chemical Engineering*.
- Subramanian Balashankar, V., Rajendran, A., 2019. Process optimization-based screening of zeolites for post-combustion CO₂ capture by vacuum swing adsorption. *ACS Sustain. Chem. Eng.* 7 (21), 17747–17755.
- Sun, Y., Sahinidis, N.V., Sundaram, A., Cheon, M.-S., 2020. Derivative-free optimization for chemical product design. *Curr. Opin. Chem. Eng.* 27, 98–106.
- Sundaram, N., 1999. Training neural networks for pressure swing adsorption processes. *Ind. Eng. Chem. Res.* 38 (11), 4449–4457.
- Tawarmalani, M., Sahinidis, N.V., 2005. A polyhedral branch-and-cut approach to global optimization. *Math. Program.* 103 (2), 225–249.
- Weber, R.S., Holladay, J.E., Jenks, C., Panisko, E.A., Snowden-Swan, L.J., Ramirez-Corredores, M., Baynes, B., Angenent, L.T., Boysen, D., 2018. Modularized production of fuels and other value-added products from distributed, wasted, or stranded feedstocks. *WIREs Energy Environ.* 7 (6), e308.
- Williams, B., Cremaschi, S., 2021. Selection of surrogate modeling techniques for surface approximation and surrogate-based optimization. *Chem. Eng. Res. Des.* 170, 76–89.
- Xiao, J., Li, C., Fang, L., Böwer, P., Wark, M., Bénard, P., Chahine, R., 2020. Machine learning-based optimization for hydrogen purification performance of layered bed pressure swing adsorption. *Int. J. Energy Res.* 44 (6), 4475–4492.
- Xiao, J., Mei, A., Tao, W., Ma, S., Bénard, P., Chahine, R., 2021. Hydrogen purification performance optimization of vacuum pressure swing adsorption on different activated carbons. *Energies* 14 (9), 2450.
- Xiao, P., Zhang, J., Webley, P., Li, G., Singh, R., Todd, R., 2008. Capture of CO₂ from flue gas streams with zeolite 13X by vacuum-pressure swing adsorption. *Adsorption* 14 (4), 575–582.
- Yancy-Caballero, D., Leperi, K.T., Bucior, B.J., Richardson, R.K., Islamoglu, T., Farha, O.K., You, F., Snurr, R.Q., 2020. Process-level modelling and optimization to evaluate metal-organic frameworks for post-combustion capture of CO₂. *Mol. Syst. Des. Eng.* 5 (7), 1205–1218.
- Yang, H., Xu, Z., Fan, M., Gupta, R., Slimane, R.B., Bland, A.E., Wright, I., 2008. Progress in carbon dioxide separation and capture: a review. *J. Environ. Sci.* 20 (1), 14–27.
- Yazaydin, A.Ö., Snurr, R.Q., Park, T.-H., Koh, K., Liu, J., LeVan, M.D., Benin, A.I., Jakubczak, P., Lanuza, M., Galloway, D.B., Low, J.J., Willis, R.R., 2009. Screening of metal-organic frameworks for carbon dioxide capture from flue gas using a combined experimental and modeling approach. *J. Am. Chem. Soc.* 131 (51), 18198–18199.
- Ye, F., Ma, S., Tong, L., Xiao, J., Bénard, P., Chahine, R., 2019. Artificial neural network based optimization for hydrogen purification performance of pressure swing adsorption. *Int. J. Hydrog. Energy* 44 (11), 5334–5344.
- Yu, C.-H., Huang, C.-H., Tan, C.-S., 2012. A review of CO₂ capture by absorption and adsorption aerosol and air quality. *Research* 12 (5), 745–769.