Green Chemical Engineering 6 (2025) 233–248

Contents lists available at ScienceDirect

Green Chemical Engineering

journal homepage: www.keaipublishing.com/en/journals/green-chemical-engineering

Article

Pressure swing adsorption process modeling using physics-informed machine learning with transfer learning and labeled data

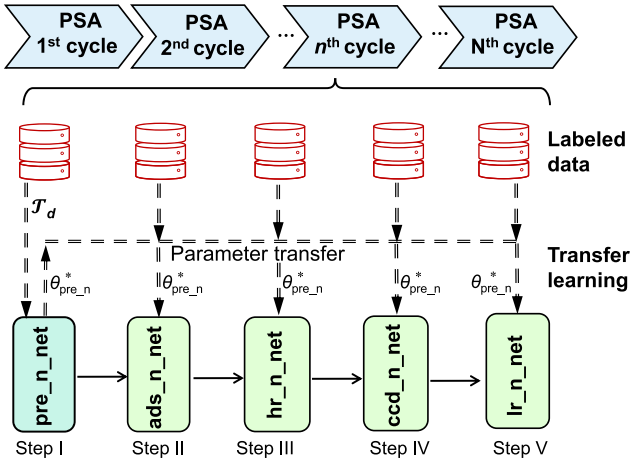
Zhiqiang Wu [a,1], Yunquan Chen [a,1], Bingjian Zhang [a,b], Jingzheng Ren [c], Qinglin Chen [b,d], Huan Wang [e,f,*], Chang He [b,d,*]

[a] School of Materials Science and Engineering, Sun Yat-Sen University, Guangzhou, 510275, China
[b] Guangdong Engineering Center for Petrochemical Energy Conservation, Guangzhou, 510275, China
[c] Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong Special Administrative Region, China
[d] School of Chemical Engineering and Technology, Sun Yat-sen University, Zhuhai, 519082, China
[e] State Key Laboratory of Heavy Oil Processing, China University of Petroleum-Beijing, Beijing, 102249, China
[f] CNNC No. 7 Research and Design Institute Co. Ltd., Taiyuan, 030012, China

HIGHLIGHTS

- Physics-informed machine learning was used to capture the PSA's spatiotemporal dynamics.
- Integrating labeled data into transfer learning can effectively capture the time-varying characteristics of the data.
- The system's network representation was decomposed into five lightweight sub-networks.
- Domain decomposition accelerated the transfer learning.
- The addition of labeled data, even in small quantities, enables PINN to reconstruct field distributions.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Physics-informed machine learning
Pressure swing adsorption
Transfer learning
Labeled data
Partial differential equations

ABSTRACT

Pressure swing adsorption (PSA) modeling remains a challenging task since it exhibits strong dynamic and cyclic behavior. This study presents a systematic physics-informed machine learning method that integrates transfer learning and labeled data to construct a spatiotemporal model of the PSA process. To approximate the latent solutions of partial differential equations (PDEs) in the specific steps of pressurization, adsorption, heavy reflux, counter-current depressurization, and light reflux, the system's network representation is decomposed into five lightweight sub-networks. On this basis, we propose a parameter-based transfer learning (TL) combined with domain decomposition to address the long-term integration of periodic PDEs and expedite the network training process. Moreover, to tackle challenges related to sharp adsorption fronts, our method allows for the inclusion of a

* Corresponding author.
E-mail addresses: wanghuan@cnsi.ltd (H. Wang), hechang6@mail.sysu.edu.cn (C. He).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.gce.2024.08.004
Received 25 June 2024; Received in revised form 1 August 2024; Accepted 13 August 2024
Available online 14 August 2024

specified amount of labeled data at the boundaries and/or within the system in the loss function. The results show that the proposed method closely matches the outcomes achieved through the conventional numerical method, effectively simulating all steps and cyclic behavior within the PSA processes.

1. Introduction

Pressure swing adsorption (PSA) has garnered significant interest as a promising option for the separation of hydrogen and the capture of carbon dioxide from flue gas streams [1,2]. Compared to thermally driven separations, PSA not only enhances energy efficiency and environmental sustainability, but facilitates easier design in a compact, modular, and portable device [3]. Despite these advantages, designing and operating PSA processes remain a significant challenge. One major obstacle to the extensive use of PSA in the process industry is the deficiency of tools that can be easily utilized, offer the required flexibility, and maintain a high level of dependability. This stems from the complexity of understanding and modeling mixed-gas adsorption, which involves both the operational dynamics and the physics of adsorption [4]. For example, PSA exhibits a dynamic nature and cyclic behavior as a unit operation, which typically follows the Skarstrom cycle that consists of individual steps, such as pressurization, adsorption, and depressurization [5,6]. The computational effort required to accurately address the strong non-linearity and non-idealities of the adsorption dynamics involved in the PSA cycle is quite expensive and prohibitively time-consuming. Hence, it is crucial to develop accurate mathematical models grounded in fundamental principles that can efficiently capturing the complex spatiotemporal dynamics inherent in PSA operation while notably alleviating the computational burden.

Partial differential equations (PDEs) offer a precise structure for mathematically representing the spatial and temporal dynamics of variables within a continuous domain. The mathematical model of PSA process comprises coupled nonlinear PDEs distributed in both space and time, with periodic initial and boundary conditions defining each step in the PSA cycle. These PDEs dictate the transfer of mass and energy within the system, and must be concurrently resolved with the differential-algebraic equations that delineate the transport phenomena and adsorption process both between and within particles [4]. The process of modeling the PSA involves iteratively solving a series of such PDEs until reaching a cyclic steady state. In the past, a range of numerical methods have been suggested to address PSA modeling and optimization issues, such as finite differences [7], orthogonal collocation [1], finite elements [8], and finite volumes [9]. Among these methods, two finite volume methods, namely the total variation diminishing (TVD) and the weighted essentially nonoscillatory (WENO), are notable for their capability to manage sharp adsorption fronts in the bed and minimize non-physical oscillations in the adsorption dynamics [10,11]. In particular, the TVD method simultaneously achieves high-order accuracy and low-order stability by introducing flux limiters, while the WENO method ensures the non-oscillatory nature of numerical solutions through weighted functions.

To support the utilization of these methods, several commercial process simulation tools are accessible to engineers and researchers, such as Aspen Adsorption [12], AVEVA Process Simulation [13], and Dynamic Adsorption Column Simulation [14]. While using these commercial tools helps to build and analyze PSA models more efficiently and conveniently, they are coded in a black-box environment and lack sufficient modeling flexibility and freedom. Practitioners can overcome these constraints by developing open-source process simulators based on published numerical algorithms and existing solution techniques. For example, several in-house PSA dynamic models of varying complexity have been successfully developed, such as GRAMS [15] and toPSAil [3]. However, when it comes to numerical models, it is important to note that a substantial amount of computational time is required to fine-tune model parameters/configurations with domain-specific expertise. Furthermore, they fall far short of meeting the demands for online prediction and real-time optimization in the current process industry [16].

A strategy for diminishing the intricacy of the modeling challenge involves the application of model reduction methodologies, which can be further used as surrogate models in dealing with modeling and optimization issues. Agarwal et al. [1] used a reduced-order model based on proper orthogonal decomposition to create cost-effective low-order models that can replace those obtained from spatial discretization. Uebbing et al. [17] employed an equilibrium model that enables the reformulation of PDEs into algebraic equations of substantially lower order, resulting in much simpler representations for PSA models and enhancing the computational efficiency of optimization problems. Recently, with the increasing attention to machine learning, data-driven surrogate models have emerged as significant alternatives to replace computationally expensive numerical models. Hao et al. [18] developed surrogate models grounded in Gaussian processes to replicate the complex, original process models of PSA. These models were then synergized with a composite optimization strategy that integrates elements of both stochastic and deterministic approaches, thereby streamlining the quest for the Pareto optimal boundary. Leperi [19] and Sundaram [20] et al. demonstrated that artificial neural networks can serve as effective surrogate models for swiftly simulating each phase of diverse PSA cycles. Rebello and Nogueira [21] designed a multiple-input, single-output framework consisting of two deep neural network models to forecast key performance indicators for CO_2 capture in PSA units. Oliveira et al. [22] found that LSTM-based deep neural networks can serve as dependable predictors of the system's dynamic behavior, making them a promising option for the advancement of control, optimization, and measurement techniques. One of the major challenges associated with the data-driven method is the need for extensive and high-quality training data, which are typically obtained through experimental measurements or numerical simulations [23,24]. In theory, they are regarded as black-box models, lacking physical interpretability, generalizability, and a direct connection to the underlying physical laws governing the system [25,26].

Recently, an innovative scientific machine learning method, known as physics-informed machine learning [27,28], has been designed to cultivate surrogate models that can be endowed with an understanding of the fundamental physical principles and specialized knowledge within a given domain. An illustration of this principle is the group of physics-informed neural networks (PINNs) [29,30], which mathematically integrate the governing PDEs into the neural network's loss function. In contrast to conventional numerical approaches, physics-informed machine learning harnesses the capabilities of automatic differentiation [31] for mesh-free numerical computation, thus circumventing the issue of the curse of dimensionality [32,33]. During the network's training phase, a loss function is formulated by evaluating the discrepancies in the governing equations, utilizing these discrepancies as penalty factors to limit the range of potential solutions. Transforming the task of resolving PDEs into an optimization challenge that involves minimizing the loss function is the essence of this approach [34,35]. The benefits of PINN go beyond just facilitating the resolution of intricate scientific and engineering challenges; they have additionally emerged as a central area of interest within the burgeoning domain of scientific machine learning [35]. Recent statistical data from academic research highlight the successful application of PINN across various domains, such as fluid mechanics [36,37], heat transfer [38,39], chemical reactions [40,41], biomedicine [42,43], materials science [44,45], solid mechanics [46,47], and fracture mechanics [48,49].

While significant progress has been achieved in modeling differential equations and clearly defined physical processes through physics-

informed learning, their limitations are exposed when dealing with PDE systems characterized by high dimensions or pronounced nonlinearity [27,50]. As previously mentioned, PSA processes involve sharp adsorption fronts, cyclic behavior, and internal recycling, which can result in stiff and ill-conditioned PDEs. This requires the sustained integration of these PDEs while managing the challenges posed by time-varying boundaries and initial conditions within the problem domain [51]. The resulting computational challenge can be partially mitigated by designing specialized structures, such as multi-PINN [52], segregated network [53], and Deep M&Mnet [54]. Additionally, domain decomposition can be utilized in conjunction with transfer learning [48,55] or pre-training [56,57] techniques, this enables the transfer of knowledge and application of models from one domain or task to another, facilitating their use in new contexts. While these combined methods are theoretically feasible, to the best of our knowledge, there is currently a lack of systematic research specifically focused on modeling PSA processes using physics-informed machine learning.

Considering the issues previously mentioned, this research presents an alternative to traditional numerical solutions by integrating physics-informed machine learning with transfer learning and labeled data, aiming to develop a one-dimensional, spatiotemporal model of the PSA processes. To mitigate the computational complexity arising from the long-term integration of PDEs, we initially implement domain decomposition to partition the physical domain into a series of subdomains, where each subdomain corresponds to a distinct step in the Skarstrom cycle. Five interconnected and lightweight sub-networks are meticulously designed to resolve the underlying solutions of PDEs in the steps of pressurization, adsorption, heavy reflux, counter-current depressurization, and light reflux. This not only enables a more accurate representation of the underlying physics, but also facilitates the utilization of parameter-based transfer learning. Furthermore, to enhance the model's understanding of the underlying physics such as sharp adsorption fronts, our method permits the inclusion of observational and measurement data in the loss function during training.

2. Model description

2.1. Modified Skarstrom PSA cycle

In this study, the PSA process is aimed at separating a mixture gas containing 15% CO_2 and 85% N_2. PSA utilizes a packed-bed chromatographic column, in which CO_2 is selectively adsorbed onto a solid sorbent under high pressure. Subsequently, the sorbent undergoes a regeneration process by reducing the pressure, which results in the release of the concentrated CO_2 that was previously adsorbed, now available as the extracted product. The commonly used cycle configuration of PSA [58] includes the modified Skarstrom cycle, the fractionated vacuum swing adsorption cycle, and the five-step cycle. Here, the modified Skarstrom cycle is employed due to its ability to reduce energy demands for CO_2/N_2 separation and its compatibility with most metal-organic framework (MOF) materials [59]. As depicted in Fig. 1, it consists of the following five steps.

(1) **Pressurization:** it begins at the first adsorber bed, where the pressure gradually increases from low pressure (P_L) to high pressure (P_H) as the compressed feed gas is introduced into the column;
(2) **Adsorption**: once pressurization is finished, the valve at the column top is opened, allowing the feed gas to be adsorbed through the adsorbent in the column. The light product N_2 is recovered at the column top, and the heavy product CO_2 is mainly adsorbed at the other end;
(3) **Heavy reflux**: after the adsorption reaches a predetermined time, the feed gas is switched to the high CO_2 concentration stream obtained during the light reflux step, which can enhance the CO_2 recovery rate;

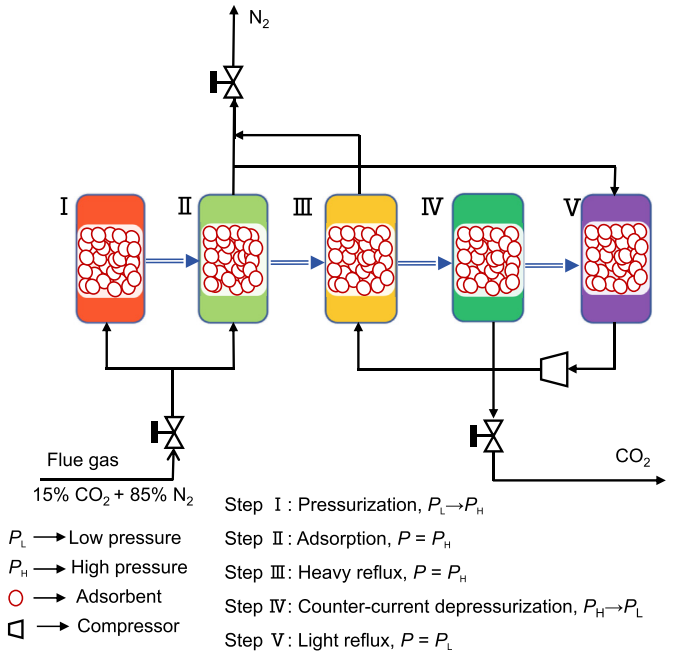


Fig. 1. Diagram of a pressure swing adsorption process using a modified Skarstrom cycle.

(4) **Counter-current depressurization**: following the re-reflux step, the pressure in the column gradually drops to P_L by turning off the valve at the column top, allowing the discharge of CO_2 at the column bottom;
(5) **Light reflux**: the concentrated N_2 is transported to the top of the column. A portion of unadsorbed CO_2 is used as the feed in the heavy reflux step.

2.2. Governing equation

In practical operation, the periodic switching between the above-mentioned five steps I ~ V of the PSA cycle is achieved by switching the valve at both ends of the column. From a mathematical modeling perspective, the transitions between steps are accomplished by changing the boundary and initial conditions. To achieve a cyclic steady state, we follow a united approach [60], where all beds undergo identical steps so that a single bed is required to describe the adsorption dynamics of the multibed cycle. As the cyclic steady state is achieved, the boundary conditions of each step can reflect the results of combined heat and mass

Table 1
The boundary conditions for the steps of the modified Skarstrom cycle.

Step	Bottom ($x^* = 0$)	Top ($x^* = 1$)
Pressurization	$P = P_L \rightarrow P_H$ $y_i = y_{i,feed}$ $T = T_{feed}$	$\partial P/\partial x = 0$ $\partial y_i/\partial x = 0$ $\partial T/\partial x = 0$
Adsorption	$P = 1.02P_H$ $y_i = y_{i,flue}$ $T = T_{flue}$	$P = P_H$ $\partial y_i/\partial x = 0$ $\partial T/\partial x = 0$
Heavy reflux	$P = 1.02P_H$ $y_i = y_{i,lr}\|_{x^*=0^+}$ $T = T_{lr}\|_{x^*=0^+}$	$P = P_H$ $\partial y_i/\partial x = 0$ $\partial T/\partial x = 0$
Counter-current depressurization	$P = P_H \rightarrow P_L$ $\partial y_i/\partial x = 0$ $\partial T/\partial x = 0$	$\partial P/\partial x^* = 0$ $\partial y_i/\partial x = 0$ $\partial T/\partial x = 0$
Light reflux	$P = P_L$ $\partial y_i/\partial x = 0$ $\partial T/\partial x = 0$	$P > P_L$ $y_i = y_{i,ads}\|_{x^*=1^-}$ $T = T_{ads}\|_{x^*=1^-}$

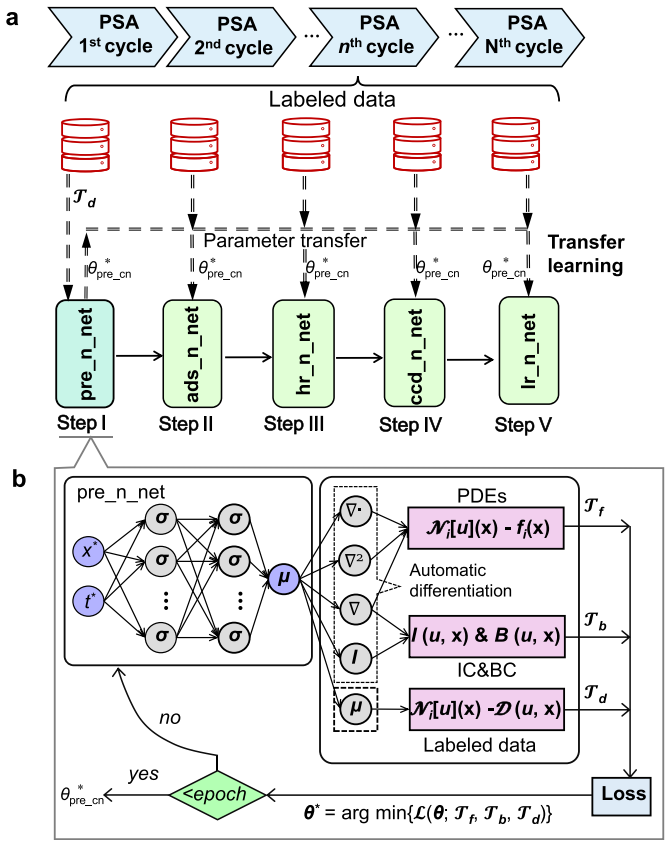


Fig. 2. Schematic diagrams of (a) the PSA process modeling by TL-LD PINN and (b) pressurization sub-network.

transfer occurring at specific times for each step, *e.g.*, the spatiotemporal distributions of pressure, composition, and temperature at both ends of the bed [61]. The detailed boundary conditions are outlined in Table 1.

To streamline the spatiotemporal model and circumvent the need for computationally demanding methods, the adsorption process of the PSA cycle can be described using a one-dimensional mathematical model with the following assumptions [62,63]:

1) gas phase behavior can be described by the ideal gas law;
2) no radial variations in pressure, concentration, and temperature;
4) gas diffusion within the adsorbent can be linearly modeled;
5) the adsorption properties of adsorbent, void fraction, and gas viscosity remain constant;
6) thermal equilibrium exists between the gas phase and the adsorbent;
7) heat transfer occurs at the column wall, and the temperature outside the column wall remains unchanged;
8) the column exhibits a plug flow that is dispersed along the axis.

Based on the above assumptions, a coupled nonlinear system of PDEs based on the mass and energy balances [64] can be obtained. In this system, the total mass balance can be represented by:

$$\frac{1}{P}\frac{\partial P}{\partial t}-\frac{1}{T}\frac{\partial T}{t}=-\frac{T}{P}\frac{\partial}{\partial x}\left(u_x\frac{P}{T}\right)-\frac{(1-\varepsilon)}{\varepsilon}\frac{RT}{P}\sum_{i=1}^{n}\frac{\partial q_i}{\partial t} \tag{1}$$

where P and T represent the total pressure of the system and column temperature; u_x, ε, and q_i represent the surface gas velocity, bed voidage, and molar loading of component i in the solid phase, respectively.

It is worth noting that there may be significant differences of orders of magnitude between the state variables and input parameters of the governing PDEs in terms of spatial and temporal scales during the PSA cycle process. A term with a large order of magnitude during the training process can contribute more to the loss function of the network. Meanwhile, in physics-informed machine learning, which frequently utilizes gradient descent-based optimizers, there is a tendency to prioritize terms with larger magnitudes, potentially neglecting those with smaller magnitudes. These variables (or parameters) may vary significantly during network training, which in turn prolongs convergence time and affects the training results. To ensure that the PDEs and boundary conditions are well constrained, it is necessary to nondimensionalize the key variables and input parameters by the following equation:

$$P^*=\frac{P}{P_0};T^*=\frac{T}{T_0};T_w^*=\frac{T_w}{T_0};h_i=\frac{q_i}{q_{s0}};u_x^*=\frac{u_x}{u_0};t^*=\frac{tu_0}{L};h_i^*=\frac{q_i^*}{q_{s0}};T_a{}^*=\frac{T_a}{T_0};P_L{}^*=\frac{P_L}{P_0};x^*=\frac{x}{L} \tag{2}$$

where P^*, T^*, T_w^*, h_i, u_x^*, t^*, h_i^*, T_a^*, P_L^*, and x^* denote the dimensionless results including pressure, temperature, wall temperature, molar loading of component i in the solid phase, superficial gas velocity, time, equilibrium molar loading of component i in the solid phase, ambient temperature, low pressure, and bed length, respectively.

Substituting the nondimensional variables from Eq. (2) into Eq. (1) transforms the total mass balance equation into:

$$\frac{\partial P^*}{\partial t^*}=\left(-P^*\frac{\partial u_x^*}{\partial x^*}-u_x^*\frac{\partial P^*}{\partial x^*}+u_x^*\frac{P^*}{T^*}\frac{\partial T^*}{\partial x^*}\right)-\psi T^*\sum_i\frac{\partial h_i}{\partial t^*}+\frac{P^*}{T^*}\frac{\partial T^*}{\partial t^*} \tag{3}$$

where $\psi=\frac{1-\varepsilon}{\varepsilon}\frac{RT_0q_{s0}}{P_0}$, the system total pressure can be obtained according to Eq. (3).

Additionally, we calculate the component balance of the system based on the molar fraction of CO_2 (y_1) in the gas phase, with the nondimensional equation as shown by:

$$\frac{\partial y_i}{\partial t^*}=\frac{1}{Pe}\left(\frac{\partial^2 y_i}{\partial x^{*2}}+\frac{1}{P^*}\frac{\partial P^*}{\partial x^*}\frac{\partial y_i}{\partial x^*}-\frac{1}{T^*}\frac{\partial y_i}{\partial x^*}\frac{\partial T^*}{\partial x^*}\right)-u_x^*\frac{\partial y_i}{\partial x^*}+\frac{\psi T^*}{P^*}\left((y_i-1)\frac{\partial h_i}{\partial t^*}+y_i\frac{\partial h_{N_2}}{\partial t^*}\right) \tag{4}$$

where $Pe=\frac{u_0L}{D_L}$, $D_L=0.7D_m+r_pu_0$. The molar fraction of N_2 in the gas phase $y_2=1-y_1$. Within the column, the temperature variation between the gas and solid phases can be computed from:

$$\frac{\partial T^*}{\partial t^*}=\pi_1\frac{\partial^2 T^*}{\partial x^{*2}}-\pi_2u_x^*\frac{\partial T}{\partial x^*}+\sum_i(\pi_{3i}+\pi_4T^*)\frac{\partial h_i}{\partial t^*}-\pi_5(T^*-T_w^*) \tag{5}$$

where

$$\pi_1=\frac{K_Z}{(\rho_gC_{p,g}\varepsilon+(1-\varepsilon)(C_{p,s}\rho_s+C_{p,a}q_{s0}))u_0L} \tag{5–1}$$

$$\pi_2=\frac{\rho_gC_{p,g}\varepsilon}{\rho_gC_{p,g}\varepsilon+(1-\varepsilon)(C_{p,s}\rho_s+C_{p,a}q_{s0})} \tag{5–2}$$

$$\pi_{3i}=\frac{(1-\varepsilon)(-\Delta H_i)q_{s0}}{T_0(\rho_gC_{p,g}\varepsilon+(1-\varepsilon)(C_{p,s}\rho_s+C_{p,a}q_{s0}))} \tag{5–3}$$

$$\Delta H_i=\Delta U_i-RT^*T_0 \tag{5–4}$$

$$\pi_4=\frac{(1-\varepsilon)(C_{p,g}-C_{p,a})q_{s0}}{\rho_gC_{p,g}\varepsilon+(1-\varepsilon)(C_{p,s}\rho_s+C_{p,a}q_{s0})} \tag{5-5}$$

$$\pi_5=\frac{2h_{in}L}{r_iu_0(\rho_gC_{p,g}\varepsilon+(1-\varepsilon)(C_{p,s}\rho_s+C_{p,a}q_{s0}))} \tag{5–6}$$

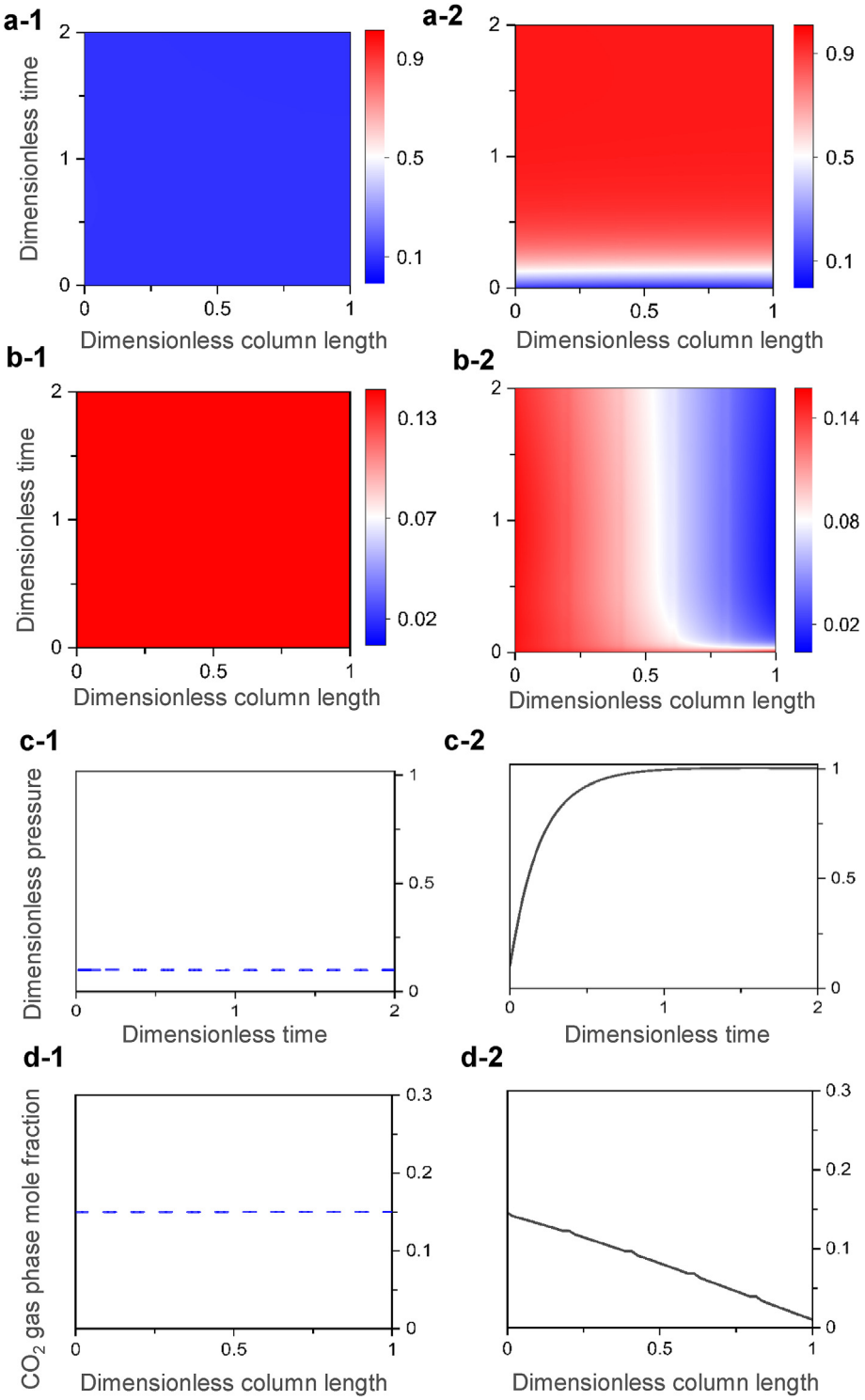


Fig. 3. Spatiotemporal distributions of (a) pressure and (b) CO_2 mole fraction in the pressurization step of the 1st PSA predicted by PINN with (right) and without (left) labeled data. The corresponding spatial profiles at the end of this step are shown in (c) and (d). The solid line represents the predicted values with labeled data, while the dashed line indicates the results without labeled data.

$$\rho_g = \frac{P^* P_H}{RT^* T_0} \tag{5–7}$$

The gas-solid mass transfer behavior can be described by the linear driving force model [62] shown in Eq. (6):

$$\frac{\partial h_i}{\partial t^*} = \frac{k_i L}{u_0}\left(h_i^* - h_i\right) \tag{6}$$

Utilizing the Ergun equation, one can determine the pressure decrease that occurs within a column [64]:

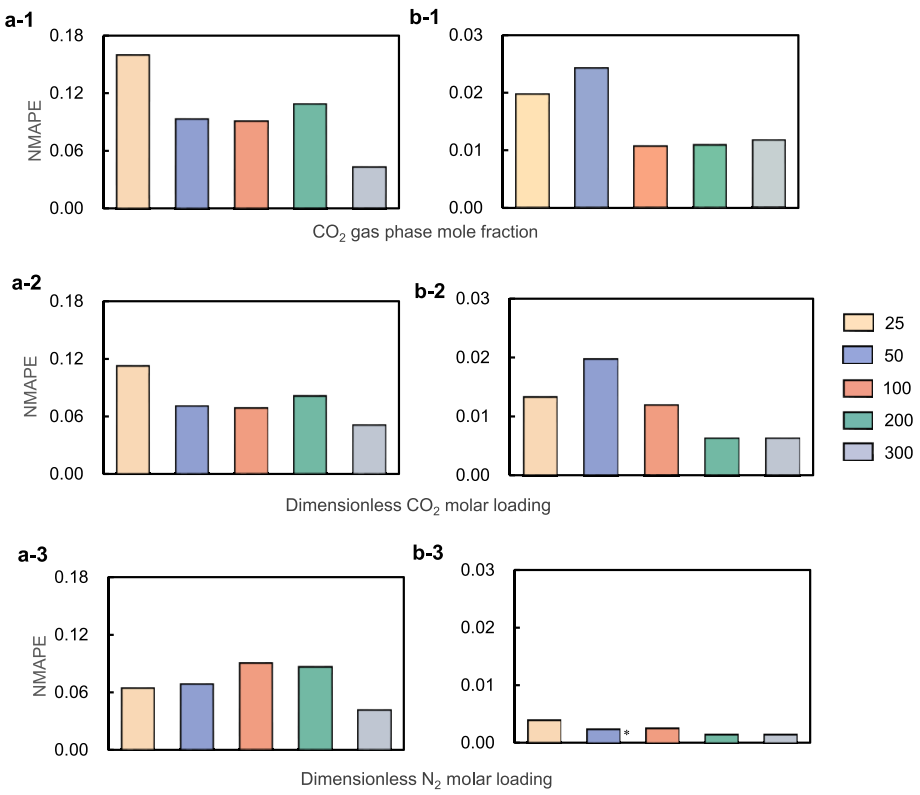


Fig. 4. NMAPEs of the key state variables for the pressurization step of the 1st cycle with the inclusion of $|\Gamma_L| = 25, 50, 100, 200, 300$ sets. (a) and (b) represent the labeled data originating solely from the boundary (Γ_{LB}) as well as from both the boundary and the interior ($\Gamma_{LB} + \Gamma_{LI}$). 1, 2, and 3 represent the CO_2 mole fraction, CO_2 molar loading, and N_2 molar loading, respectively.

$$-\frac{\partial P^*}{\partial x^*} = \frac{150\mu(1-\varepsilon)^2 L u_0}{4r_p^2 P_0 \varepsilon^3} u_x^* + \frac{1.75(1-\varepsilon) L u_0^2}{2r_p P_0 \varepsilon^3}\left(\sum_i MW_i y_i \rho_g\right) u_x^* |u_x^*| \tag{7}$$

2.3. The selection of adsorbent

Here, MOF-177 was selected as the adsorbent. The dual-site Langmuir model was used to describe the adsorption process of CO_2 and N_2 using MOF-177. To simulate the competitive isotherms of two gases in the adsorption process, it is assumed that site 1 with a strong adsorption capacity only adsorbs CO_2, while the weaker site 2 is computationally adsorbed by CO_2 and N_2. The adsorption isotherms of CO_2 and N_2 are fitted by using data from Ref. [59], which are as forms:

$$q_{CO_2}^* = \frac{q_{sat,CO_2}^1 \cdot B_{CO_2}^1 \cdot P_{CO_2}}{1 + B_{CO_2}^1 \cdot P_{CO_2}} + \frac{q_{sat,CO_2}^2 \cdot B_{CO_2}^2 \cdot P_{CO_2}}{1 + B_{CO_2}^2 \cdot P_{CO_2} + B_{N_2}^2 \cdot P_{N_2}} \tag{8–1}$$

$$q_{N_2}^* = \frac{q_{sat,N_2}^2 \cdot B_{N_2}^2 \cdot P_{N_2}}{1 + B_{CO_2}^2 \cdot P_{CO_2} + B_{N_2}^2 \cdot P_{N_2}} \tag{8–2}$$

where $q_{sat,i}^s$ and B_i^s represent the saturation loading and isotherm parameters of the component i at site s, respectively. Considering the effect of temperature on the loading during the adsorption of CO_2 (component 1) and N_2 (component 2), the isotherm parameters were fitted by using the following Arrhenius form:

$$B_i^* = b_i^* \exp\left[\frac{-\Delta U_i^*}{RT}\right] \tag{9}$$

where ΔU_i^* represents the adsorption internal energy of component i. The isotherm parameters of MOF-177 are as follows [59]: $q_{sat,1}^1 = 48$ mol/kg, $q_{sat,1}^2 = 0$, $b_1^1 = 8.06 \times 10^{-10}$ kPa^{-1}, $b_1^2 = 0$, $\Delta U_1^1 = \Delta U_1^2 = 14$ kJ/mol, $q_{sat,2}^2 = 48$ mol/kg, $b_2^2 = 1.11 \times 10^{-9}$ kPa^{-1}, $\Delta U_2^2 = 10$ kJ/mol.

3. Methodology

3.1. TL-LD PINN method

As shown in Section 2.1, the mathematical model of the PSA process consists of a set of coupled PDEs. These PDEs were used to describe the total mass balance (Eq. (3)), the component mass balance (Eq. (4)), the gas-solid phase temperature gradient (Eq. (5)) in the column, and the gas-solid mass transfer equation (Eq. (6)) based on the linear driving force model. In addition, we used the competitive dual-site Langmuir model to describe the adsorption behavior between CO_2 and N_2, and used Eq. (8–1) and Eq. (8–2) to represent the adsorption loads of CO_2 and N_2, respectively. The temperature effect on adsorption load is further considered by fitting the Arrhenius equation Eq. (9). The combined solutions of these differential equations can be expressed as $\boldsymbol{u}(\boldsymbol{x})$=$\{p^*(\boldsymbol{x}), y_i^*(\boldsymbol{x}), q_1^*(\boldsymbol{x}), q_2^*(\boldsymbol{x}), T^*(\boldsymbol{x})\}$; herein $\boldsymbol{x} = (x^*, t^*)$ is the input space-time coordinate, and their generic relationship can be expressed by:

$$f\left(\boldsymbol{x}; \frac{\partial u}{\partial x^*}, \frac{\partial u}{\partial t^*}, \frac{\partial u}{\partial x^{*2}}\right) = 0, \boldsymbol{x} \in \Omega \tag{10}$$

The corresponding boundary conditions and initial conditions can be expressed by:

$$\boldsymbol{B}(u, \boldsymbol{x}) = 0, \boldsymbol{x} \in \partial\Omega_x \tag{10–1}$$

$$\boldsymbol{I}(u, \boldsymbol{x}) = 0, \boldsymbol{x} \in \partial\Omega_\tau \tag{10–2}$$

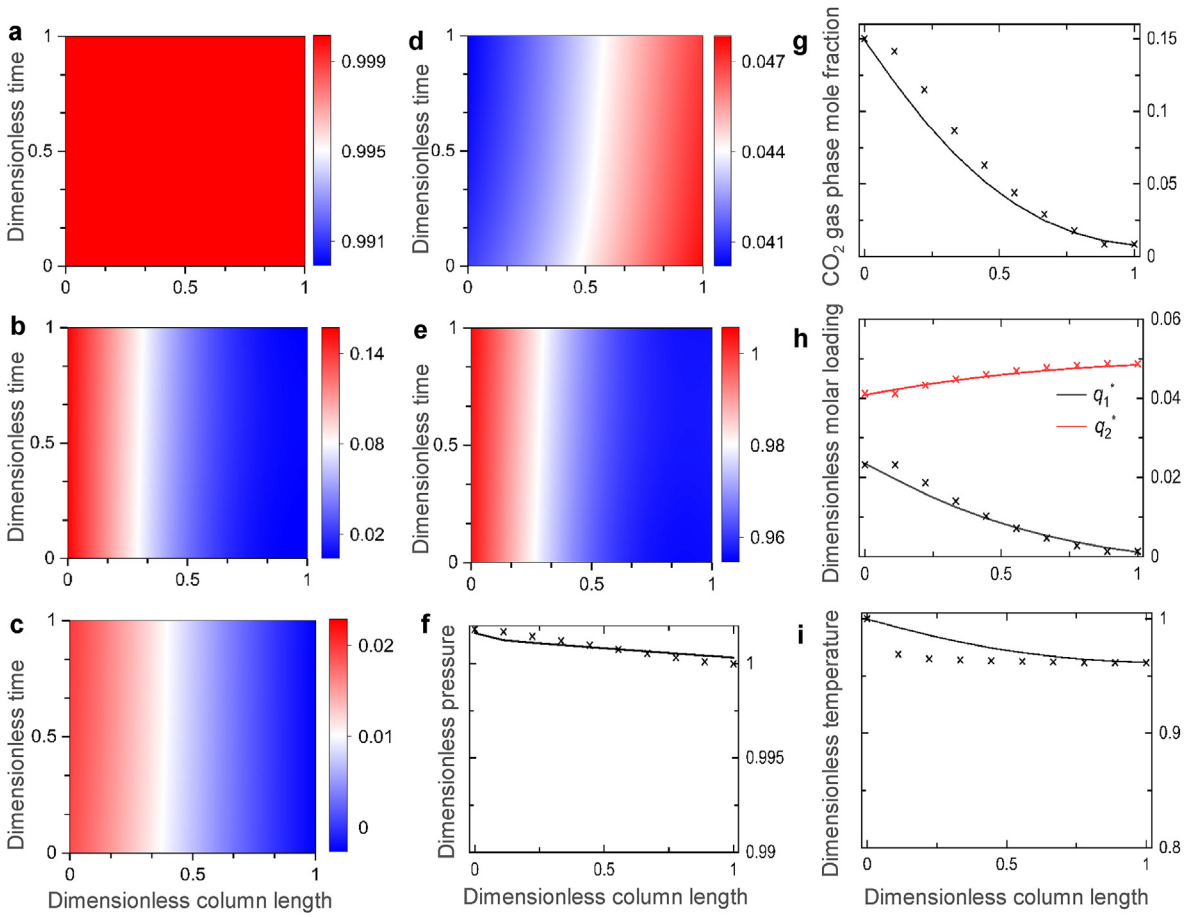


Fig. 5. The distributed state variables in the adsorption step of the 1st cycle. (a) Pressure, (b) CO_2 mole fraction, (c) CO_2 molar loading, (d) N_2 molar loading, and (e) temperature; (f)~(i) presents the corresponding spatial profiles at the end of this step obtained through TL-LD PINN and WENO-FV methods. Here, q_1* and q_2* are dimensionless CO_2 molar loading and N_2 molar loading. The solid line represents the predicted values of the TL-LD method, and the " × " indicates the results of the WENO-FV method.

where $\partial\Omega_x$ and $\partial\Omega_\tau$ represent the space boundary and time boundary of the domain required for defining constraints, respectively. $\boldsymbol{B}(u, \boldsymbol{x})$ and $\boldsymbol{I}(u, \boldsymbol{x})$ are the general forms of the boundary and initial condition, respectively.

As mentioned in the Introduction, PSA processes are characterized as cyclic, dynamic, and multiphysics systems governed by a set of coupled nonlinear PDEs, with frequent step-switching steps, and each step contains complex boundary and initial conditions. Accurately capturing the time-varying characteristics in each step using physics-informed learning presents the following two main challenges:

(I) The cyclic behavior and internal recycling within PSA processes result in an increased time domain and time-varying boundary conditions, which a large number of spatiotemporal degrees of freedom are needed to train the network. Consequently, this demands an extensive and prolonged integration of PDEs, a task that surpasses the ability of a single network to encompass the full temporal range.

(II) The PSA process involves sharp adsorption fronts where the presentence of steep concentration and pressure gradients within the adsorption bed can pose significant computational challenges. When simulating PSA using a PINN-based PDE solver, special treatment of these sharp fronts is necessary to ensure the model's accuracy and solution stability.

To address these issues, we proposed an integrated method that can combine standard PINN with transfer learning and labeled data techniques (TL-LD PINN) to enhance the learning and generalization capabilities for the PSA process modeling. The specific implementation of this method shown in Fig. 2 will be detailed in the following sections.

3.1.1. *Transfer learning*

To address the aforementioned issue I, a domain decomposition technique is proposed to segment the physical domain of the system into several subdomains, with each subdomain corresponding to a specific step in the Skarstrom cycle. The goal of this decomposition is to transform the complex, long-term periodic PDE-solving task into a series of easily solvable short-term tasks. This allows for a more targeted and accurate representation of the underlying physics for the subsequent network training process. The utilization of domain decomposition is visualized in Fig. 2a. Taking the *n*th cycle as an example, the network representation of the system can be decomposed into five interconnected and lightweight sub-networks: *pre_n_net*, *ads_n_net*, *hr_n_net*, *ccd_n_net*, and *lr_n_net*. They are designed to address the underlying solutions of PDEs in the steps of pressurization, adsorption, heavy reflux, counter-current depressurization, and light reflux, respectively. To satisfy the continuity constraints, a specific quantity of randomly selected points on the boundary is designated after each step. These points are then utilized for training the subsequent network. The optimal solution for the state variables after each training step is allocated as the initial conditions for the state variables in the following step, facilitating a seamless transition between steps.

In this model, the PSA process is considered to occur periodically within a single bed. In each step of the PSA cycle, there are underlying

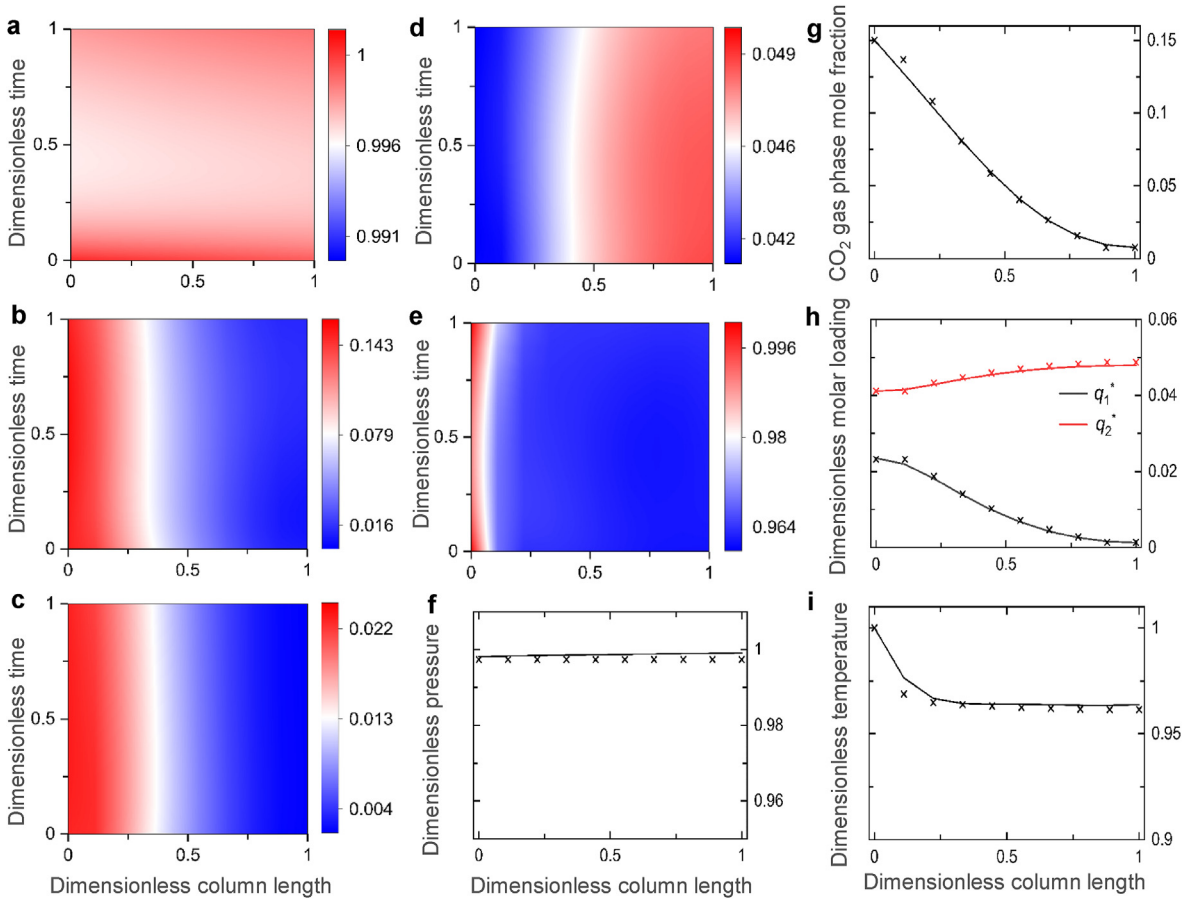


Fig. 6. The distributed state variables in the heavy reflux step of the 1st cycle. (a) Pressure, (b) CO_2 mole fraction, (c) CO_2 molar loading, (d) N_2 molar loading, and (e) temperature; (f)~(i) presents the corresponding spatial profiles at the end of this step obtained through TL-LD PINN and WENO-FV methods. Here, q_1* and q_2* are dimensionless CO_2 molar loading and N_2 molar loading. The solid line represents the predicted values of the TL-LD method, and the " × " indicates the results of the WENO-FV method.

physical laws and governing equations that exhibit similarities. Mathematically, the operation transition between these steps is essentially a change in boundary/initial conditions. Transfer learning is an important machine learning technique that leverages the similarities in tasks, enabling knowledge from a pre-trained model (source model) to be transferred to a new one (target model), thus expediting the network training and enhancing the efficiency of convergence. Herein, a parameter-based transfer learning approach [55] was utilized, which leverages the overlap in the feature space between the source domain and the target domain, and through shared parameters, it helps the target model quickly capture the common characteristics. As shown in Fig. 2b, a standard PINN is initially utilized to approximate the solution in the pressurization step during the *n*th PSA cycle. The obtained pressurization subnetwork (referred to as *pre_n_net*) functions as the source model. The knowledge gathered from the source model is encoded into the shared parameters $\boldsymbol{\theta}^*_{\text{pre\_n}}$ which are then used to initialize the target models with slightly varying boundaries and initial conditions. After the solutions for all target models have been thoroughly understood, we integrate the source and target models in accordance with their chronological order.

3.1.2. *Labeled data*

As mentioned in Section 3.1.1, the presence of sharp adsorption fronts can lead to strong non-linearity and non-idealities for PSA modeling. To the best of our knowledge, accurately describing the resulting steep concentration and pressure gradients has far exceeded the model capabilities of standard physics-informed learning methods. This also poses a significant challenge for traditional numerical methods. Researchers have employed the weighted essentially nonoscillatory (WENO) scheme to assist the finite volume method in discretizing the PDEs into easily solvable ordinary differential equations (ODEs) [11,65]. However, it is note that it may be difficult to directly integrate the WENO scheme into PINN architecture.

Physics-informed machine learning provides the ability to incorporate measured or simulation data with prior knowledge, enabling the solution of incomplete or parameter-deficient sets of PDEs. When modeling the PSA process using PINN, to better enable the model to learn the underlying physical laws of the state variables, our approach allows for adding observational and measurement data in the loss function during the training process. As shown in Fig. 2b, we restrict each subnetwork (*e.g.*, *pre_n_net*) to comply with the physics imposed by the PDEs and boundary/initial conditions. Additionally, we incorporate a specific quantity of labeled data obtained from finite volume simulation combined with the WENO scheme (referred to as WENO-FV in this study). We denote the training dataset as $\boldsymbol{\Gamma} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{|\Gamma|}\}$, where $|\Gamma|$ is the sample size. It consists of four parts: the interior sampling point set ($\Gamma_E \subset \Omega$), the boundary sampling point set ($\Gamma_B \subset \partial\Omega$), the initial sampling point set ($\Gamma_I \subset \partial\Omega$), and the labeled data set ($\Gamma_L \subset \partial\Omega$). For each step of the PSA cycle, Γ_E and Γ_B represent the sampling points within the domain and on the boundary of that step, with the sampling sizes denoted as $|\Gamma_E|$ and $|\Gamma_B|$, respectively. For all steps except for pressurization, where the initial conditions of the state variables depend on the final state of the previous step, we add additional initial sampling points $\Gamma_I \subset \partial\Omega$ to assist the training process, with a size of $|\Gamma_I|$. Besides, the labeled data Γ_L refers to a vector set generated by the WENO-FV method, containing all state variables in $\boldsymbol{u}(\boldsymbol{x})$. It is noted that labeled data may fall on the boundary of

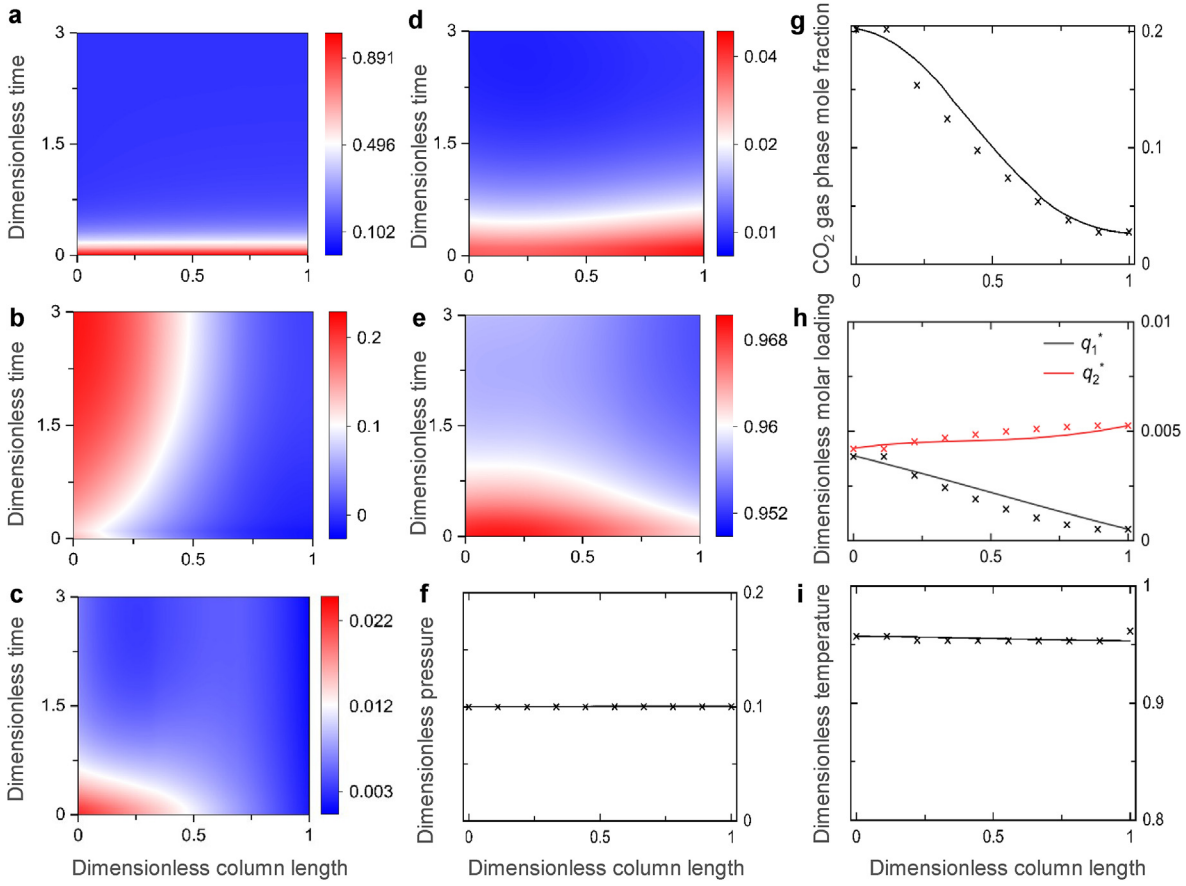


Fig. 7. The distributed state variables in the counter-current depressurization step of the 1st cycle. (a) Pressure, (b) CO_2 mole fraction, (c) CO_2 molar loading, (d) N_2 molar loading, and (e) temperature; (f)~(i) presents the corresponding spatial profiles at the end of this step obtained through TL-LD PINN and WENO-FV methods. Here, q_1* and q_2* are dimensionless CO_2 molar loading and N_2 molar loading. The solid line represents the predicted values of the TL-LD method, and the " × " indicates the results of the WENO-FV method.

the domain or with the interior of the domain, denoted as Γ_{LB} and Γ_{LI}, respectively, $\Gamma_L = \{\Gamma_{LB}, \Gamma_{LI}\}$.

During the iterative training, PINN aims at constructing a surrogate (subnetwork, $\boldsymbol{u}_{net}$ ($\boldsymbol{x}$; $\boldsymbol{\theta}$)) for a system of PDEs using a deep neural network architecture, where $\boldsymbol{\theta} = \{\boldsymbol{W}^{\sigma}, \boldsymbol{b}^{\sigma}\}$ is a collection of network weight matrices $\boldsymbol{W}$ and bias vectors $\boldsymbol{b}$. By optimization, the surrogate continuously approximates the combined solution $\boldsymbol{u}(\boldsymbol{x})$ of the PDE system. Finding the optimal solution involves searching for a set of optimal parameters θ* that can maximally satisfy all constraints from the governing equations, the boundary and initial conditions, as well as the added labeled data of the present step. Mathematically, this is achieved by constructing a composite loss function, $\mathbf{L}$, describing the residuals between $\boldsymbol{u}_{net}$ ($\boldsymbol{x}$; $\boldsymbol{\theta}$) and $\boldsymbol{u}(\boldsymbol{x})$. Here, $\mathbf{L}$ consists of four terms as follows: the PDEs loss ($\mathbf{L}_E$), the boundary condition loss ($\mathbf{L}_B$), the initial condition loss ($\mathbf{L}_I$), as well as the loss linked to the labeled data ($\mathbf{L}_D$), as expressed by:

$$\mathbf{L}(\theta;\Gamma) = \lambda_E \mathbf{L}_E + \lambda_B \mathbf{L}_B + \lambda_I \mathbf{L}_I + \lambda_D \mathbf{L}_D \tag{11}$$

$$\mathbf{L}_E(\theta;\Gamma_E) = \frac{1}{|\Gamma_E|}\left|\left|_{x\in\Gamma_E} f\left(x;\frac{\partial\hat{u}}{\partial x^*},\frac{\partial\hat{u}}{\partial t^*},\frac{\partial\hat{u}}{\partial x^{*2}}\right)\right|\right|_2^2 \tag{11–1}$$

$$\mathbf{L}_B(\theta;\Gamma_B) = \frac{1}{|\Gamma_B|}\sum_{x\in\Gamma_B}\left|\left|B(\hat{u};\boldsymbol{x})\right|\right|_2^2 \tag{11–2}$$

$$\mathbf{L}_I(\theta;\Gamma_I) = \frac{1}{|\Gamma_I|}\sum_{x\in\Gamma_I}\left|\left|I(\hat{u};\boldsymbol{x})\right|\right|_2^2 \tag{11–3}$$

$$\mathbf{L}_D(\theta;\Gamma_D) = \frac{1}{|\Gamma_D|}\sum_{\boldsymbol{x}\in\Gamma_D}\left|\left|\hat{u}(\boldsymbol{x};\theta) - u(\boldsymbol{x};\theta)\right|\right|_2^2 \tag{11–4}$$

where λ_E, λ_B, λ_I, and λ_D are the weights assigned to the $\mathbf{L}_E$, $\mathbf{L}_B$, $\mathbf{L}_I$, $\mathbf{L}_D$, respectively.

During the training process, our objective is to find the optimal parameters θ that minimize the loss function (θ; $\mathbf{\Gamma}$). To achieve this, we employ a widely-used gradient-based optimizer, Adam [66], to minimize the loss function, which takes the following form:

$$\theta = arg_{\theta=\{w,b\}}\ \min L(\theta;\mathbf{\Gamma}) \tag{12}$$

3.2. Hyperparameter setting

The developed TL-LD PINN model is designed to construct a deep neural network that can map the input space-time coordinate $\boldsymbol{x} = (x^*, t^*)$ to the output variable vector $\boldsymbol{u}(\boldsymbol{x})$. The neural network utilizes a feed-forward architecture with 4 hidden layers, each consisting of 70 nodes for function approximation. To address the vanishing gradient problem and support the computation of multiple derivatives, the hyperbolic tangent (Tanh) function is employed as the activation function, given that the system of PDEs involves second-order differential operators. The initial learning rate is established at 0.001, and the learning rate decays exponentially during training, which is as follows:

$$lr = lr_0 \cdot \beta^{\text{epoch}} \tag{13}$$

where lr and β represent the learning rate per epoch and the decay factor, respectively. In this study, $lr_0 = 0.001$, $\beta = 0.999$, and the maximization

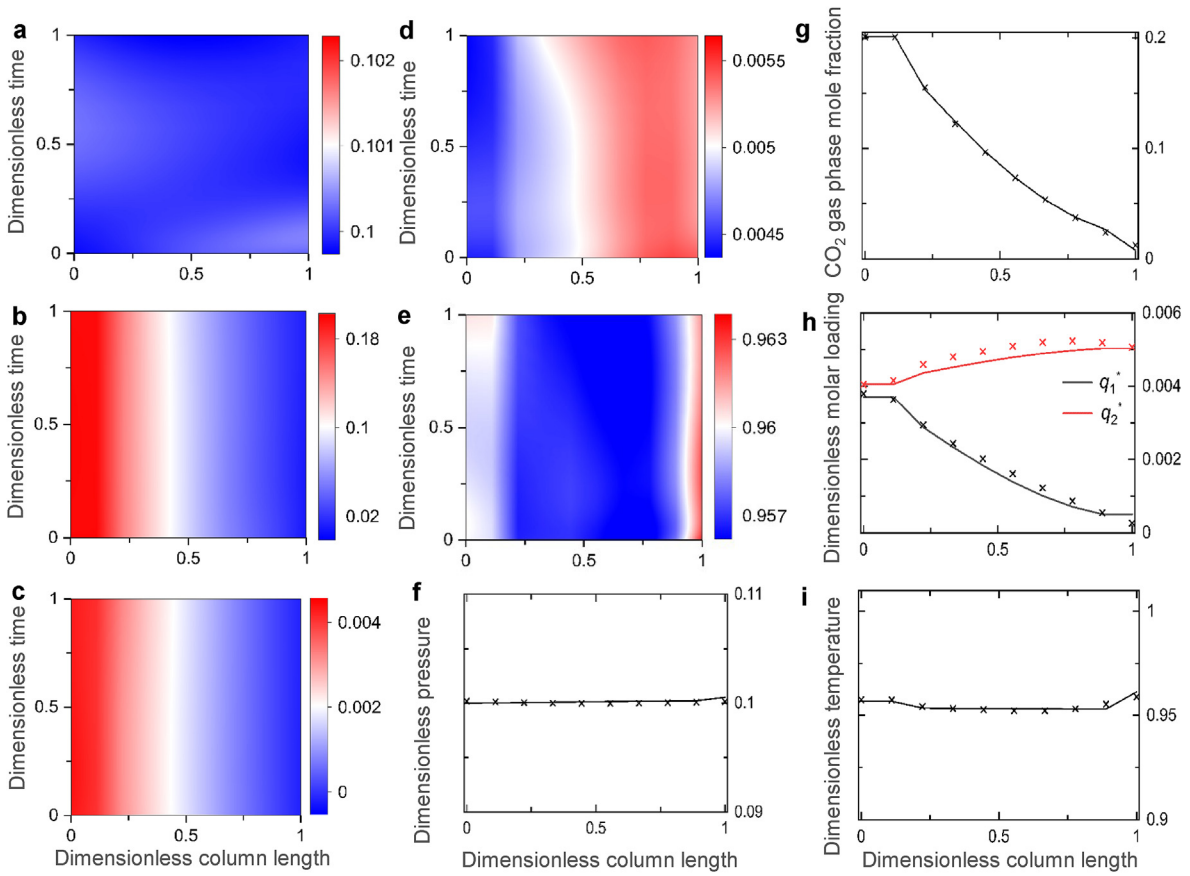


Fig. 8. The distributed state variables in the light reflux step of the 1st cycle. (a) Pressure, (b) CO_2 mole fraction, (c) CO_2 molar loading, (d) N_2 molar loading, and (e) temperature; (f)~(i) presents the corresponding spatial profiles at the end of this step obtained through TL-LD PINN and WENO-FV methods. Here, q_1* and q_2* are dimensionless CO_2 molar loading and N_2 molar loading. The solid line represents the predicted values of the TL-LD method, and the " x " indicates the results of the WENO-FV method.

number of epochs is 50,000.

4. Results and discussion

In this work, the PINN models were developed using Python and built on the PyTorch framework. They were run on workstations with dual Intel Xeon E5-2695v4 CPUs at 2.1 GHz, 96 GB of RAM, and a pair of NVIDIA GeForce RTX 3070 GPUs, all operating under the Linux Ubuntu system. The validation and labeled data were generated using numerical solutions obtained from MATLAB running on a Windows 10Pro operating system.

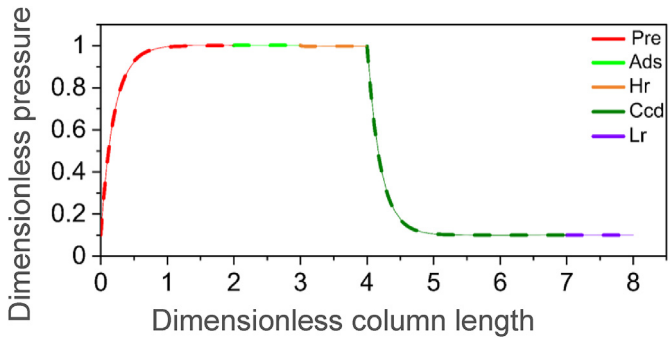


Fig. 9. The variation of pressure over time during the 1st cycle of PSA. The solid line represents the predicted values from the TL-LD method, while the dashed line indicates the results from the WENO-FV method.

4.1. The effect of labeled data

In this section, we first conduct a systematic comparison of the predictive accuracy of the PINN method with and without labeled data, using the pressurization step of the initial cycle ($n = 1$) as an illustrative example. Fig. 3 (a-1, b-1) shows the spatiotemporal distributions of pressure and CO_2 mole fraction. It is evident that, without the inclusion of labeled data, the pressure and CO_2 mole fraction remained constant throughout the spatiotemporal domain. For example, the spatial profiles after this step, shown in Fig. 3 (c-1, d-1), reveal a stable value of $p^* = 0.1$ bar and $y^* = 15\%$, indicating the operation of pressurization has a negligible impact on the PSA process. This significant deviation contrasts the expected impact of the pressurization in the practical application of PSA. By contrast, as shown in Fig. 3 (a-2, b-2), with the integration of labeled data, the resulting pressure and CO_2 concentration exhibit noticeable and reasonable variations. The dimensionless magnitudes of these state variables span from 0.1 to 1 and 0.01 to 0.15, respectively. In Fig. 3 (c-2, d-2), the pressure exponentially increases from low pressure ($p^* = 0.1$) to high pressure ($p^* = 1$) as the feed gas is fully injected into the column. Meanwhile, the CO_2 mole fraction linearly drops from 15% at the column bottom to approximately 0.6% at the column top, which

Table 2
The MSEs for the predicted state variables in the light reflux step with and without transfer learning.

	CO_2 mole fraction	CO_2 loading	N_2 loading
LD PINN	2.74×10^{-4}	1.93×10^{-7}	5.75×10^{-8}
TL-LD PINN	7.80×10^{-5}	6.99×10^{-8}	6.66×10^{-8}

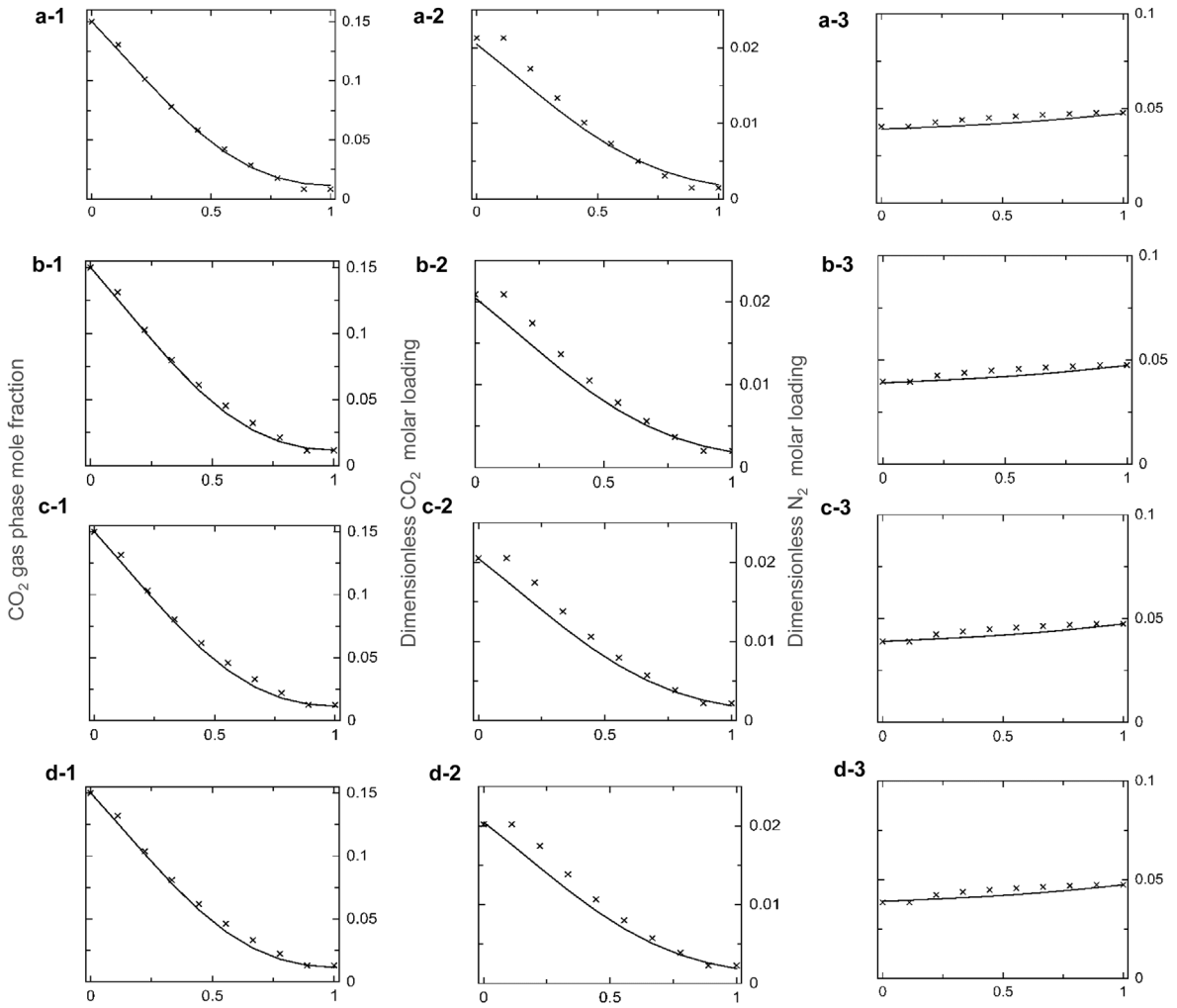


Fig. 10. Spatial profiles of the key state variables at the end of the pressurization step in the multicycle PSA process. (a), (b), (c), and (d) represent the 2nd, 4th, 6th, and 8th cycles of the PSA, respectively; 1, 2, and 3 represent CO_2 mole fraction, CO_2 molar loading, and N_2 molar loading, respectively. The solid line represents the predicted values of the TL-LD method, and the " × " indicates the results of the WENO-FV method.

aligns closely with the actual change trend observed in PSA processes. This improvement indicates that the use of labeled data allows the PINN to accurately learn the trends of the key state variables throughout the pressurization step. By leveraging the labeled data, the PINN method gains deeper insights into the sharp front in PSA and establishes connections between these state variables, enhancing its ability to make accurate predictions regarding their spatiotemporal distributions.

To further investigate the effect of labeled data on the model accuracy, we selected $|\Gamma_L| = 25$, 50, 100, 200, and 300 sets of input-output mappings obtained from the WENO-FV method. These labeled data are separately incorporated into the loss term of *pre_1_net* to improve network training. Herein, a widely-used performance metric, namely the normalized mean absolute percentage error (NMAPE) [67], is used to quantitatively measure the deviation of the field distributions of state variables in contrast to the WENO-FV solution, as indicated by

Table 3
MSEs of key state variables for different pressurization steps.

Cycle	CO_2 mole fraction	CO_2 loading	N_2 loading
1st cycle	5.03×10^{-5}	1.83×10^{-6}	1.81×10^{-7}
2nd cycle	6.50×10^{-6}	2.57×10^{-6}	6.03×10^{-6}
4th cycle	1.20×10^{-5}	2.48×10^{-6}	5.25×10^{-6}
6th cycle	1.54×10^{-5}	2.34×10^{-6}	5.07×10^{-6}
8th cycle	1.73×10^{-5}	2.23×10^{-6}	5.05×10^{-6}

$$\mathrm{NMAPE} = \frac{1}{N_i}\sum_{i=1}^{N_i}\frac{\left|\phi_{i,\mathrm{LD}} - \phi_{i,\mathrm{LD-PINN}}\right|}{\max(\phi_{i,\mathrm{LD}}) - \min(\phi_{i,\mathrm{LD}})} \tag{14}$$

where $\phi_{i,\mathrm{LD}}$ and $\phi_{i,\mathrm{LD\text{-}PINN}}$ represent the distributed state variables predicted by using WENO-FV method and PINN method corresponding to the ith point in the validation dataset, $N_i = 50$ is the number of validation datasets.

From the left hand of Fig. 4, it is seen that the inclusion of labeled data on the boundary, even in a limited amount (*e.g.*, $|\Gamma_L| = 25$ sets), empowers the PINN method to reconstruct the field distributions of key state variables. As the amount of labeled data increases from 25 sets to 200 sets, the overall NMAPEs for CO_2 mole fraction and CO_2 loading decrease from from 15.9% to 10.8% and from 11.3% to 8.2%, respectively. However, the NMAPE for N_2 loading has an approximately 40% increase from 6.4% to 9.0%. As $|\Gamma_L| = 300$, the NMAPE for N_2 loading remarkably drops to 4.2%, whereas for the CO_2 mole fraction and CO_2 loading it decreases to 4.3% and 5.1%, respectively. In another sampling scheme, the labeled data includes both the boundary data and interior data, as shown in the right hand of Fig. 4. The NMAPEs exhibit a continuous and significant decrease, with all statistics falling below 3%. Particularly for N_2 loading, the NMAPEs reduce to a range of 0.1%–0.4%. Overall, in this sampling scheme, the prediction accuracy achieves a satisfactory level even with a sparse sample size. During the training process of other subnetworks, we thus simultaneously consider the boundary and interior

data in the labeled data with a size of $|\Gamma_L| = 50$.

4.2. The effect of transfer learning

In section 4.1, the PINN model with labeled data has effectively simulated the spatiotemporal distributions of state variables along the column in the pressurization step. The obtained optimal network parameters of the pressurization subnetwork (*pre_1_net*) are used to initialize the subsequent subnetworks *via* transfer learning. Fig. 5 (a-e) shows the detailed spatiotemporal distribution of the main state variables in the adsorption step of the 1st cycle. The profiles of various state variables along the column at the end of this step, obtained from both the TL-LD PINN and the WENO-FV methods, are shown in Fig. 5 (f-i). From this figure, it can be seen that the CO_2 mole fraction and loading gradually decrease from the bottom ($x^* = 0$) to the top ($x^* = 1$) of the adsorption column, with values decreasing from around 0.15 and 0.025 at the bottom to around 0.006 and 0.001 at the top, respectively. In contrast, the N_2 loading increases from around 0.04 at the bottom to around 0.05, while the pressure is maintained near a high value ($p^* = 1$) during this step. The dimensionless temperature is $T^* = 1$ at the boundary of the bottom of the adsorption column, while it remains around $T^* = 0.96$ internally. In all, as compared to the reference WENO-FV method, the TL-LD PINN method accurately describes the spatiotemporal profiles of all state variables during the adsorption step.

After reaching the predetermined adsorption time, the heavy product obtained in the light reflux step is recycled to the column as feed gas, initiating the heavy reflux step. Fig. 6 presents the spatiotemporal profiles of the state variables along the column in this step. Due to the switch in feedstock being the only difference between these two steps, with the operating conditions and objectives remaining essentially consistent, it can be observed in Figs. 5 and 6 that the spatiotemporal distributions of the key state variables are also largely similar. The only difference is that in the adsorption step, the pressure along the bottom to the top of the column is unevenly distributed, gradually increasing from 0.96 to 1, while in the reflux step, the overall pressure distribution remains relatively constant at around 1. Besides, during this step, the concentration of CO_2 at the bottom of the column increases as the CO_2 mole fraction in the refluxed heavy product is higher than that in the feed gas. Although the reflux of heavy products generated from the adsorption of feed gas results in a reduction in the amount of CO_2 ultimately collected in the cycle, it can significantly enhance the maximum CO_2 purity achievable by the adsorbent.

After the heavy reflux step concludes, the valve at the column's end is shut, leading to the pressure at the bottom of the column to quickly decrease from $p^* = 1$ to $p^* = 0.1$, initiating the counter-current depressurization step. From Fig. 7, it can be observed that in this step, the CO_2 mole fraction and loading gradually decrease, with values dropping from around 0.2 and 0.004 at the bottom to around 0.03 and 0.0005 at the top of the column, respectively. In contrast, the N_2 loading increases from around 0.004 at the bottom to 0.005. The reason for the increase in CO_2 mole fraction at the column bottom is the reflux of heavy products in the reflux step. Throughout this step, the pressure is maintained near low pressure ($p^* = 0.1$), while the temperature remains between 0.95 and 0.97 at both the boundary and interior. After the counter-current depressurization step, the light products generated in the feed step are sent to the top of the column, while the heavier products re-enter the column as the feed gas, constituting the light reflux step. From Fig. 8, it is evident that the pressure is maintained near low pressure ($p^* = 0.1$), while the temperature remains around 0.96 at both the boundary and interior. The CO_2 mole fraction and loading exhibit an obvious spatial variation, with values dropping from around 0.2 and 0.004 at the bottom to approximately 0.008 and 0.0005 at the top of the column, respectively. In contrast, the N_2 loading increases from around 0.004 at the bottom to near 0.005.

To more intuitively demonstrate the benefits of transfer learning in step switching, we capture the profiles of state variable values at the end of each step and compare them with the WENO-FV solution. As shown in Fig. 9, it is evident that the dimensionless pressure profiles obtained from the TL-LD PINN model closely align with those from the WENO-FV solution across all steps involved in the 1st cycle. In addition, we compared the differences in model prediction accuracy between these two methods, using the light reflux step as an example in Table 2. To simplify, the mean squared error (MSE) is used as an indicator to evaluate the prediction accuracy of monitored state variables at the end of this step, as given by

$$MSE = \frac{1}{N_i}\sum_{i=1}^{N_i} |\boldsymbol{u}_{\text{net}}(\boldsymbol{x};\theta) - \boldsymbol{u}(\boldsymbol{x})|^2 \tag{15}$$

where $\boldsymbol{u}_{\text{net}}(\boldsymbol{x};\theta)$ represents the predicted state variables by using the PINN method, $\boldsymbol{u}(\boldsymbol{x})$ represents the reference results of state variables by using the WENO-FV method, $N_i = 50$ is the number of validation datasets.

It is evident that the TL-LD PINN method leads to a notable drop in MSE when predicting the key state variables, except for the N_2 loading from Table 2. The slight increase in the MSE of N_2 loading could be due to the small scale of the dimensionless value ($< 10^{-3}$), where even minor fluctuations during training could significantly affect its error, resulting in varied outcomes. As a result, the integration of transfer learning has enhanced both training efficiency and prediction accuracy.

Additionally, we conducted a comparison of the simulation time between the WENO-FV method and the LD-TL PINN model. The WENO-FV method requires 0.6996 s for a single cycle, whereas the LD-TL PINN model completes one cycle simulation in 0.029 s. Nevertheless, training the LD-TL PINN model takes 7 h.

To summarize, in this section, we effectively modeled the first cycle of the PSA process using the TL-LD PINN method. Upon comparison, it is clear that the prediction results (both in terms of time and space profiles) from the TL-LD PINN model closely align with those from the WENO-FV solution throughout the entire process. This alignment indicates that the proposed method can precisely capture the trends in key state variable variations during the first cycle.

4.3. Multicycle simulation

The results in section 4.2 demonstrate that the TL-LD PINN solution aligns well with the results obtained from the solution of the WENO-FV model, taking the 1st cycle of the PSA process as an illustrative case. To assess the stability of this method in the cyclic process, Fig. 10 further depicts the spatial profiles of the key state variables along the column at the end of the pressurization step for the 2nd, 4th, 6th, and 8th cycles of the PSA. It is evident from the figure that the predicted profiles closely match the reference curves obtained from the WENO-FV method.

To visually assess the accuracy of the TL-LD PINN model for predicting the multi-cycle PSA process, we computed the errors between the main state variables predicted by the TL-LD PINN and WENO-FV methods for various pressurization steps across cycles. Table 3 presents the MSEs of dimensionless state variables for different cycles. From the table, it can be observed that the MSE of dimensionless state variables for each cycle is generally low at the end of the pressurization step. Specifically, the MSEs for CO_2 and N_2 loadings are on the order of 10^{-6}, while the MSE for CO_2 mole fraction is slightly higher, around the order of 10^{-5}. This discrepancy could stem from the significantly higher magnitude of CO_2 mole fraction compared to CO_2 and N_2 loadings. During the training process, gradient descent may be biased towards terms associated with CO_2 mole fraction, leading to deviations in the predicted results.

5. Conclusion

PSA processes are characterized by cyclic, dynamic, and multiphysics systems that are governed by a set of coupled nonlinear PDEs. Besides, they involve frequent step-switching steps, with each step containing complex boundary and initial conditions. In this work, our goal was to evaluate the feasibility of using a physics-informed machine learning

approach for modeling PSA processes, with a focus on its ability to effectively incorporate transfer learning and labeled data. To approximate the latent solutions of the steps within each running cycle, the network representation of the system was decomposed into five interconnected and lightweight sub-networks, namely *pre_n_net*, *ads_n_net*, *hr_n_net*, *ccd_n_net*, and *lr_n_net*. Each sub-network was dedicated to a particular step within the modified Skarstrom cycle. On this basis, this work specifically focused on addressing two main computational obstacles: (I) the long-term integration of periodic PDEs and (2) the handling of sharp adsorption fronts. To tackle these obstacles, we proposed a systematic framework that combined the standard PINN with transfer learning and labeled data to boost the learning and generalization capacities of physics-informed machine learning. The results of the pressurization step in the 1st cycle showed that the inclusion of labeled data, even with a very small amount, empowered the PINN to well-reconstruct the field distributions of key state variables. Meanwhile, with the help of transfer learning, the proposed method effectively simulated the remaining steps of the PSA cycle, and the model's prediction results align with those obtained from the WENO-FV method. Finally, the results of the multicycle simulation also demonstrated the potential of the TL-LD PINN model in handling cyclic problems.

In conclusion, we offered an alternative method for utilizing deep learning in the modeling and design of gas separation processes. By considering the intricate dynamic variations of the PSA cycle, designing specific neural network representations for each step, and accounting for the continuity constraints between steps, this method can gain a deep understanding of the performance, efficiency, and limitations of the adsorption dynamics in the PSA processes. On this basis, further study can explore system-level optimization and fine-tuning operational parameters to ensure their feasibility in online prediction, analysis, and optimization applications for advanced PSA designs.

CRediT authorship contribution statement

Zhiqiang Wu: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Investigation. **Yunquan Chen:** Writing – review & editing, Validation, Software, Methodology, Investigation. **Bingjian Zhang:** Writing – review & editing, Validation. **Jingzheng Ren:** Validation, Supervision. **Qinglin Chen:** Writing – review & editing, Validation. **Huan Wang:** Validation, Formal analysis. **Chang He:** Writing – review & editing, Validation, Supervision, Methodology, Formal analysis, Conceptualization.

Declaration of competing interests

The author declares the following financial interests/personal relationships which may be considered as potential competing interests: Huan Wang is currently employed by CNNC No. 7 Research and Design Institute Co. Ltd. The other 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.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (Nos. 22078373 and 22078372).

Notation

ε	Bed porosity
μ	Gas viscosity (Pa·s)
π	Dimensionless group of energy balance equation
ρ_g	Gas density (mol/m^3)
ρ_s	Adsorbent density (kg/m^3)
D_m	Molecular diffusion coefficient of CO_2-N_2 mixture (m^2/s)
D_L	Axial dispersion coefficient (m^2/s)
k_i	Mass transfer coefficient of component i (s^{-1})
MW_i	Molecular weight of component i (kg/mol)
K_Z	Effective gas thermal conductivity (W/m/K)
R	Gas constant (J/mol/K)
$C_{p,g}$	Specific heat capacity of gas (J/mol/K)
$C_{p,s}$	Specific heat capacity of adsorbent (J/mol/K)
$C_{p,a}$	Adsorption relative heat capacity (J/mol/K)
L	Column length (m)
P	pressure (Pa)
P_0	Adsorbent pressure (Pa)
P^*	Dimensionless pressure
q_i	Molar load of component i in the solid phase (mol/kg)
q_i^*	Equilibrium molar load of component i in a solid phase (mol/kg)
q_{s0}	Molar load ratio factor (mol/kg)
r_i	Column radius (m)
r_0	Outside column radius (m)
r_p	Adsorbent particle radius (m)
t	Time (s)
t^*	Dimensionless time
T	Temperature (K)
T^*	Dimensionless temperature
T_a	Environment temperature (K)
T_0	Feed temperature (K)
T_w^*	Dimensionless wall temperature
u_0	Velocity scaling factor (m/s)
u_x	Surface gas velocity (m/s)

u_x^*	Dimensionless surface gas velocity
U	Internal energy (J/mol)
H	Enthalpy (J/mol)
h_i	The solid component balances the molar load
h_i^*	The solid component has no dimensional equilibrium molar load
y_i	Molar fraction of component i in the gas phase
x	Column length coordinate (m)
x^*	No dimensional length coordinates

Subscript and index symbol

PSA	Pressure swing adsorption
PINN	Physical information neural network
PDE	Partial differential equation
Ω	Training space
$\partial\Omega_x$	Spatial boundary
$\partial\Omega_\tau$	Time boundary
LD	Labeled data
TL	Transfer learning
pre	Pressurization
ads	Adsorption
lr	Light reflux
ccd	Counter-current depressurization
hr	Heavy reflux
i	Component index
P_L	Low pressure
P_H	High pressure
θ	Neural network training parameter set
lr	Training iteration learning rate
lr_0	Initial learning rate of training
β	Learning rate attenuation factor
NMAPE	Normalized mean absolute percentage error
MSE	Mean square error
n	Represents the nth cycle
Γ	Training point set
Γ_E	Internal sampling point
Γ_B	Boundary sampling point
Γ_I	Initial sampling point
Γ_L	Labeled data sampling points
Γ_{LB}	Boundary labeled data sampling points
Γ_{LI}	Internal labeled data sampling points
m	Total number of sampling points

Loss function symbol

L	Total loss function
L_E	Control equation loss function
L_B	Boundary condition loss function
L_I	Initial condition loss function
L_D	Labeled data residual loss function
λ_E	Control equation loss function weight
λ_B	Boundary condition loss function weight
λ_I	Initial condition loss function weight
λ_D	Labeled data residual loss function weight

Appendix

Table A-1
Operating parameters values used in the simulations.

Parameters	Values	Parameters	Values
Column length L, m	1	Viscosity of gas, Pa·s	1.72×10^{-5}
Initial pressure P_0, kPa	10	Void fraction	0.37
Adsorption pressure P_{ads}, kPa	100	Molecular diffusivity D_m, m^2/s	1.2995×10^{-5}
Purge pressure P_l, kPa	10	Thermal conduction in gas phase Kz, W/m/k	0.09
Time of pressurization t_{pre}, s	20	Specific heat of gas C_{pg}, J/mol/k	30.7

(*continued on next page*)

Table A-1 (*continued*)

Parameters	Values	Parameters	Values
Time of adsorption t_{ads}, s	10	Specific heat of adsorbed phase Cpa, J/mol/k	30.7
Time of heavy reflux t_{hr}, s	10	Molecular weight of CO_2 MW_{CO2}, kg/mol	0.04402
Time of depressurization t_{ccd}, s	30	Molecular weight of N_2 MW_{N2}, kg/mol	0.02802
Time of light reflux t_{lr}, s	10	Radius of the pellets r_p, m	0.001
Inlet velocity v_0, m/s	0.1	Specific heat capacity of the adsorbent C_{ps}, J/kg/K	1070
Flue gas inlet temp. T_{feed}, K	313.15	Molar loading scaling factor q_{s0}, mol/kg	5.84
Discrete volume	10	Mass transfer coefficient for CO_2 K_{CO2}, 1/s	0.1631
Adsorbent density ρ_s, kg/m^3	427	Mass transfer coefficient for N_2 K_{N2}, 1/s	0.2044

References

[1] A. Agarwal, L.T. Biegler, S.E. Zitney, Simulation and optimization of pressure swing adsorption systems using reduced-order modeling, Ind. Eng. Chem. Res. 48 (2009) 2327–2343.

[2] V. Subramanian Balashankar, A.K. Rajagopalan, R. de Pauw, A.M. Avila, A. Rajendran, Analysis of a batch adsorber analogue for rapid screening of adsorbents for postcombustion CO_2 capture, Ind. Eng. Chem. Res. 58 (2019) 3314–3328.

[3] T. Kim, J.K. Scott, Dynamic modeling and simulation of pressure swing adsorption processes using toPSAil, Comput. Chem. Eng. 176 (2023) 108309.

[4] M.D. Sees, T. Kirkes, C.-C. Chen, A simple and practical process modeling methodology for pressure swing adsorption, Comput. Chem. Eng. 147 (2021) 107235.

[5] P. Cruz, J.C. Santos, F.D. Magalhães, A. Mendes, Cyclic adsorption separation processes: analysis strategy and optimization procedure, Chem. Eng. Sci. 58 (2003) 3143–3158.

[6] Y. Chen, H. Ahn, Optimization strategy for enhancing the product recovery of a pressure swing adsorption through pressure equalization or co-current depressurization: a case study of recovering hydrogen from methane, Ind. Eng. Chem. Res. 62 (2023) 5286–5296.

[7] C.-W. Shu, Essentially non-oscillatory and weighted essentially non-oscillatory schemes, Acta Numer. 29 (2020) 701–762.

[8] Y. Liu, Z. Fang, H. Li, S. He, A mixed finite element method for a time-fractional fourth-order partial differential equation, Appl. Math. Comput. 243 (2014) 703–717.

[9] R. Haghpanah, A. Majumder, R. Nilam, A. Rajendran, S. Farooq, I.A. Karimi, M. Amanullah, Multiobjective optimization of a four-step adsorption process for postcombustion CO_2 capture *via* finite volume simulation, Ind. Eng. Chem. Res. 52 (2013) 4249–4265.

[10] J.C. Knox, A.D. Ebner, M.D. LeVan, R.F. Coker, J.A. Ritter, Limitations of breakthrough curve analysis in fixed-bed adsorption, Ind. Eng. Chem. Res. 55 (2016) 4734–4748.

[11] G.-S. Jiang, C.-W. Shu, Efficient implementation of weighted ENO schemes, J. Comput. Phys. 126 (1996) 202–228.

[12] A.T. Inc, Aspen Adsorption and Aspen Chromatography, Bedford, MA, 2024.

[13] C. Depew, Making Pressure Swing Adsorption (PSA) Technology Evaluation Easier and Faster, 2020.

[14] ProSim, Dynamic Adsorption Column Simulation, ProSim DAC, 2024.

[15] A. Arora, S.S. Iyer, M.M.F. Hasan, GRAMS: a general framework describing adsorption, reaction and sorption-enhanced reaction processes, Chem. Eng. Sci. 192 (2018) 335–358.

[16] E.A. Costa, C.M. Rebello, L. Schnitman, J.M. Loureiro, A.M. Ribeiro, I.B.R. Nogueira, Adaptive digital twin for pressure swing adsorption systems: integrating a novel feedback tracking system, online learning and uncertainty assessment for enhanced performance, Eng. Appl. Artif. Intell. 127 (2024) 107364.

[17] J. Uebbing, L.T. Biegler, L. Rihko-Struckmann, S. Sager, K. Sundmacher, Optimization of pressure swing adsorption *via* a trust-region filter algorithm and equilibrium theory, Comput. Chem. Eng. 151 (2021) 107340.

[18] Z. Hao, A. Caspari, A.M. Schweidtmann, Y. Vaupel, A.A. Lapkin, A. Mhamdi, Efficient hybrid multiobjective optimization of pressure swing adsorption, Chem. Eng. J. 423 (2021) 130248.

[19] K.T. Leperi, D. Yancy-Caballero, R.Q. Snurr, F. You, 110th anniversary: surrogate models based on artificial neural networks to simulate and optimize pressure swing adsorption cycles for CO_2 capture, Ind. Eng. Chem. Res. 58 (2019) 18241–18252.

[20] N. Sundaram, Training neural networks for pressure swing adsorption processes, Ind. Eng. Chem. Res. 38 (1999) 4449–4457.

[21] C.M. Rebello, I.B.R. Nogueira, Optimizing CO_2 capture in pressure swing adsorption units: a deep neural network approach with optimality evaluation and operating maps for decision-making, Sep. Purif. Technol. 340 (2024) 126811.

[22] L.M.C. Oliveira, H. Koivisto, I.G.I. Iwakiri, J.M. Loureiro, A.M. Ribeiro, I.B.R. Nogueira, Modelling of a pressure swing adsorption unit by deep learning and artificial Intelligence tools, Chem. Eng. Sci. 224 (2020) 115801.

[23] M. Alauddin, F. Khan, S. Imtiaz, S. Ahmed, P. Amyotte, Integrating process dynamics in data-driven models of chemical processing systems, Process, Saf. Environ. Prot. 174 (2023) 158–168.

[24] X. Meng, Z. Li, D. Zhang, G.E. Karniadakis, PPINN: parareal physics-informed neural network for time-dependent PDEs, Comput. Methods Appl. Mech. Eng. 370 (2020) 113250.

[25] M.R. Dobbelaere, P.P. Plehiers, R. Van de Vijver, C.V. Stevens, K.M. Van Geem, Machine learning in chemical engineering: strengths, weaknesses, opportunities, and threats, Engineering 7 (2021) 1201–1211.

[26] N. Sharma, Y.A. Liu, A hybrid science-guided machine learning approach for modeling chemical processes: a review, AIChE J. 68 (2022) e17609.

[27] G.E. Karniadakis, I.G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, L. Yang, Physics-informed machine learning, Nat. Rev. Phys. 3 (2021) 422–440.

[28] M. Raissi, A. Yazdani, G.E. Karniadakis, Hidden fluid mechanics: learning velocity and pressure fields from flow visualizations, Science 367 (2020) 1026–1030.

[29] L. Lu, X. Meng, Z. Mao, G.E. Karniadakis, DeepXDE: a deep learning library for solving differential equations, SIAM Rev. 63 (2021) 208–228.

[30] M. Raissi, P. Perdikaris, G.E. Karniadakis, Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, J. Comput. Phys. 378 (2019) 686–707.

[31] A. Lawrence, P. Thollander, M. Andrei, M. Karlsson, Specific Energy Consumption/Use (SEC) in energy management for improving energy efficiency in industry: meaning, usage and differences, Energies 12 (2019) 1–22.

[32] M. Li, Cyclic simulation and energy assessment of closed-circuit RO (CCRO) of brackish water, Desalination 545 (2023) 116149.

[33] S. Markidis, The old and the new: can physics-informed deep-learning replace traditional linear solvers? Front. Big Data 4 (2021) 669097.

[34] J. Blechschmidt, O.G. Ernst, Three ways to solve partial differential equations with neural networks—a review, GAMM-Mitteilungen 44 (2021) e202100006.

[35] S. Cuomo, V.S. Di Cola, F. Giampaolo, G. Rozza, M. Raissi, F. Piccialli, Scientific machine learning through physics–informed neural networks: where we are and what's next, J. Sci. Comput. 92 (2022) 88.

[36] S. Cai, Z. Mao, Z. Wang, M. Yin, G.E. Karniadakis, Physics-informed neural networks (PINNs) for fluid mechanics: a review, Acta Mech. Sin. 37 (2021) 1727–1738.

[37] Y. Yang, Y. Mesri, Learning by neural networks under physical constraints for simulation in fluid mechanics, Comput. Fluids 248 (2022) 105632.

[38] N. Zobeiry, K.D. Humfeld, A physics-informed machine learning approach for solving heat transfer equation in advanced manufacturing and engineering applications, Eng. Appl. Artif. Intell. 101 (2021) 104232.

[39] M.M. Billah, A.I. Khan, J. Liu, P. Dutta, Physics-informed deep neural network for inverse heat transfer problems in materials, Mater. Today Commun. 35 (2023) 106336.

[40] S.I. Ngo, Y.-I. Lim, Solution and parameter identification of a fixed-bed reactor model for catalytic CO_2 methanation using physics-informed neural networks, Catalysts 11 (2021) 1304.

[41] Z. Wu, M. Li, C. He, B. Zhang, J. Ren, H. Yu, Q. Chen, Physics-informed learning of chemical reactor systems using decoupling–coupling training framework, AIChE J 70 (2024) e18436.

[42] G. Kissas, Y. Yang, E. Hwuang, W.R. Witschey, J.A. Detre, P. Perdikaris, Machine learning in cardiovascular flows modeling: predicting arterial blood pressure from non-invasive 4D flow MRI data using physics-informed neural networks, Comput. Methods Appl. Mech. Eng. 358 (2020) 112623.

[43] F. Sahli Costabal, Y. Yang, P. Perdikaris, D.E. Hurtado, E. Kuhl, Physics-informed neural networks for cardiac activation mapping, Front. Physiol. 8 (2020) 42.

[44] Y. Chen, L. Lu, G.E. Karniadakis, L. Dal Negro, Physics-informed neural networks for inverse problems in nano-optics and metamaterials, Opt. Express 28 (2020) 11618–11633.

[45] Z. Fang, J. Zhan, Deep physical informed neural networks for metamaterial design, IEEE Access 8 (2020) 24506–24513.

[46] Y. Diao, J. Yang, Y. Zhang, D. Zhang, Y. Du, Solving multi-material problems in solid mechanics using physics-informed neural networks based on domain decomposition technology, Comput. Methods Appl. Mech. Eng. 413 (2023) 116120.

[47] A.M. Roy, R. Bose, V. Sundararaghavan, R. Arróyave, Deep learning-accelerated computational framework based on physics informed neural network for the solution of linear elasticity, Neural Network 162 (2023) 472–489.

[48] S. Goswami, C. Anitescu, S. Chakraborty, T. Rabczuk, Transfer learning enhanced physics informed neural network for phase-field modeling of fracture, Theor. Appl. Fract. Mech. 106 (2020) 102447.

[49] C.J.G. Rojas, J.L. Boldrini, M.L. Bittencourt, Parameter identification for a damage phase field model using a physics-informed neural network, Theor. Appl. Mech. Lett. 13 (2023) 100450.

[50] A.S. Krishnapriyan, A. Gholami, S. Zhe, R.M. Kirby, M.W. Mahoney, Characterizing possible failure modes in physics-informed neural networks, in: M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, J. Wortman Vaughan (Eds.), Advances in Neural Information Processing Systems 34, Curran Associates, Inc, New York, 2021, pp. 26548–26560.

[51] Y. Chen, Z. Wu, B. Zhang, J. Ren, C. He, Q. Chen, Transfer learning and pretraining enhanced physics-informed machine learning for closed-circuit reverse osmosis modeling, Desalination 580 (2024) 117557.
[52] A.D. Jagtap, G.E. Karniadakis, Extended physics-informed neural networks (XPINNs): a generalized space-time domain decomposition based deep learning framework for nonlinear partial differential equations, Commun. Comput. Phys. 28 (2020) 2002–2041.
[53] R. Laubscher, Simulation of multi-species flow and heat transfer using physics-informed neural networks, Phys. Fluids 33 (2021) 087101.
[54] S. Cai, Z. Wang, L. Lu, T.A. Zaki, G.E. Karniadakis, DeepM&Mnet: inferring the electroconvection multiphysics fields based on operator approximation by neural networks, J. Comput. Phys. 436 (2021) 110296.
[55] Z. Wu, B. Zhang, H. Yu, J. Ren, M. Pan, C. He, Q. Chen, Accelerating heat exchanger design by combining physics-informed deep learning and transfer learning, Chem. Eng. Sci. 282 (2023) 119285.
[56] R. Mattey, S. Ghosh, A novel sequential method to train physics informed neural networks for Allen Cahn and Cahn Hilliard equations, Comput. Methods Appl. Mech. Eng. 390 (2022) 114474.
[57] J. Guo, Y. Yao, H. Wang, T. Gu, Pre-training strategy for solving evolution equations based on physics-informed neural networks, J. Comput. Phys. 489 (2023) 112258.
[58] S.P. Reynolds, A. Mehrotra, A.D. Ebner, J.A. Ritter, Heavy reflux PSA cycles for CO_2 recovery from flue gas: Part I. Performance evaluation, Adsorption 14 (2008) 399–413.
[59] D. Yancy-Caballero, K.T. Leperi, B.J. Bucior, R.K. Richardson, T. Islamoglu, O.K. Farha, F. You, R.Q. Snurr, Process-level modelling and optimization to evaluate metal–organic frameworks for post-combustion capture of CO_2, Mol. Syst. Des. Eng. 5 (2020) 1205–1218.
[60] S.G. Subraveti, S. Roussanaly, R. Anantharaman, L. Riboldi, A. Rajendran, Techno-economic assessment of optimised vacuum swing adsorption for post-combustion CO_2 capture from steam-methane reformer flue gas, Sep. Purif. Technol. 256 (2021) 117832.
[61] J.A. Delgado, A.E. Rodrigues, Analysis of the boundary conditions for the simulation of the pressure equalization step in PSA cycles, Chem. Eng. Sci. 63 (2008) 4452–4463.
[62] D. Ko, R. Siriwardane, L.T. Biegler, Optimization of pressure swing adsorption and fractionated vacuum pressure swing adsorption processes for CO_2 capture, Ind. Eng. Chem. Res. 44 (2005) 8084–8094.
[63] E. Gleuckauf, J. Coates, The influence of incomplete equilibrium on the front boundary of chromatograms and the effectiveness of separation, J. Chem. Soc. 78 (1947) 1315–1321.
[64] K.T. Leperi, R.Q. Snurr, F. You, Optimization of two-stage pressure/vacuum swing adsorption with variable dehydration level for postcombustion carbon capture, Ind. Eng. Chem. Res. 55 (2016) 3338–3350.
[65] L.F. Shampine, M.W. Reichelt, The MATLAB ODE suite, SIAM J. Sci. Comput. 18 (1997) 1–22.
[66] D.P. Kingma, J. Ba, Adam: a method for stochastic optimization, arXiv Preprint, arXiv: 1412.6980 (2014).
[67] R. Laubscher, Simulation of multi-species flow and heat transfer using physics-informed neural networks, Phys. Fluids 33 (2021) 087101.