Available online at www.sciencedirect.com

Chemical Engineering Research and Design

journal homepage: www.elsevier.com/locate/cherd

Progressive learning for surrogate modeling of amine scrubbing CO_2 capture processes

*Yu-Da Hsiao, Chuei-Tin Chang**

Department of Chemical Engineering, National Cheng Kung University, Tainan, Taiwan

ARTICLE INFO

Article history:
Received 21 January 2023
Received in revised form 18 April 2023
Accepted 6 May 2023
Available online 10 May 2023

Keywords:
Surrogate modeling
Progressive learning
Multi-task learning
Artificial neural networks
Amine scrubbing
Post-combustion CO_2 capture

ABSTRACT

Amine scrubbing process is a promising approach for post-combustion CO_2 capture. To analyze and improve the process performance, various rigorous mathematical models have already been adopted for simulation purpose. If these high-fidelity models are utilized for more advanced applications, the required computational demand can be overwhelming. This drawback inevitably motivates the use of data-based surrogate models to relieve computation effort. However, due to the inherent complexities of any given amine scrubbing process, the task of taking enough simulation data for building an accurate model is still computationally expensive. Therefore, in this paper, an innovative modeling procedure using artificial neural networks is proposed to effectively alleviate the aforementioned data acquisition effort. To improve the data sampling efficiency while maintaining model accuracy, several concepts used in process synthesis and progressive learning have been adapted in this work. The proposed surrogate model was constructed for the specific purposes of predicting the CO_2 emission rate, the reboiler duty and the compression duty. In applying the proposed procedure, a large number of samples collected from repeated standalone simulation runs of absorber process were used to pretrain the corresponding surrogate model, and this model was then further fine-tuned and expanded according to samples collected from relatively few plantwide simulation runs. By using this modeling strategy, both plantwide sampling size and total data collection time can be effectively reduced. More specifically, a comparison between the conventional method and the current model-building strategy showed that over 23–64% of data acquisition time can be saved with the latter approach.

1. Introduction

Amine scrubbing processes have been extensively used in the hydrocarbon industry for sour gas treating (i.e., removal of H_2S and CO_2) to alleviate the environmental impacts (Hoff and Svendsen, 2013; Tikadar et al., 2020). Recently, due to concerns of increasing global warming, it is thus also regarded as an effective technology for post-combustion CO_2 capture (Hoff and Svendsen, 2013). However, the amine scrubbers are still suffered from practical issues such as intensive energy consumptions and high capital and operational costs. According to a previous analysis, the thermodynamic irreversibility of CO_2 capture via amine scrubbing is more than eight times of that of the CO_2 compression train (Ferrara et al., 2017). Therefore, numerous studies have been carried out in the past aiming to enhance the commercial feasibility of amine scrubbing process (Li et al., 2016). Reducing energy consumptions for solvent regeneration at the stripper section is widely considered to be one of the key issues for realizing financially feasible CO_2 capture (Khalifa et al., 2022). Proper heat recovery is critical for improving the energy efficiency and thermodynamic reversibility of the amine scrubbers (Liang et al., 2015; Lin and Rochelle, 2016). Many studies have attempted to improve the CO_2 capture process from chemical and/or engineering perspectives. The

* Corresponding author.
E-mail addresses: n38101036@gs.ncku.edu.tw (Y.-D. Hsiao), ctchang@mail.ncku.edu.tw (C.-T. Chang).

https://doi.org/10.1016/j.cherd.2023.05.016
0263-8762/

use of various alternative amine solvents and/or blends is the most popular approach (N.Borhani and Wang (2019)). On the other hand, modification of process configurations should be considered as another promising strategy to achieve better energy efficiency. Various process intensification and modification designs have already been developed, investigated and compared. These designs include modifications of absorber and stripper columns, and intensifications of the cross heat exchanger (Ahn et al., 2013; Damartzis et al., 2016; Khalifa et al., 2022; Le Moullec et al., 2014; Wang et al., 2015).

Since all amine scrubbing processes, which are fabricated with various configurations and solvents, need to be optimized at plant level to achieve the best CO_2 removal efficiency (Damartzis et al., 2016), it becomes necessary to construct reliable and accurate process models before further pilot-plant studies and subsequent commercial scale-up operations. In addition, these models may also be useful for real-time analysis, optimization and control (Chen et al., 2021). In amine scrubber design, it is well-known that a trade-off is always present between the energy consumption and the sour gas emission rates (Oh et al., 2016; Tikadar et al., 2020). Therefore, to achieve an appropriate balance between these two goals, multi-objective optimization methods may be applied to provide the process engineers with various feasible choices for proper decision making (Tikadar et al., 2020). The amine scrubbing process can of course be modelled with commercial simulators, e.g. Aspen HYSYS, and these proprietary software products are typically supported by rigorous models and correlations to facilitate steady-state and dynamic simulations. The simulation-based models of amine scrubbing processes have been included in complicated superstructure-based optimization problems for minimizing energy consumption (Oh et al., 2016). Although successful results have been reported, the optimization tasks are still tedious and time consuming since rigorous simulations of complex systems with highly non-linear constraints are computationally expensive. Lengthy convergence time is almost inevitable, while the numerical iteration algorithm could even fail (Alarie et al., 2021; Ochoa-Estopier et al., 2014). Therefore, there is undoubtedly another trade-off between model rigorousness and computation efficiency in the optimization calculations of the amine scrubbing process (Nuchitprasittichai and Cremaschi, 2013). Consequently, the data-based surrogate model that accurately replicates the non-linear input-output relationships of any given process should be considered as a valuable alternative (Zhu et al., 2022).

In the field of process systems engineering, the data-based surrogate models have been widely adopted for solving the optimization problems (Kajero et al., 2017; McBride and Sundmacher, 2019). The complex physiochemical equations can be converted into simplified functions embedded within a surrogate model. For example, the artificial neural networks (ANNs) can be thought of as mathematical models performing matrix-vector multiplications and nonlinear activations. To effectively train a data-driven model, large sets of diversely distributed simulation and/or experimental data are indispensable. Thus, it is necessary to collect sufficient number of samples with the help of first-principle models and/or pilot experiments (Nuchitprasittichai and Cremaschi, 2013; Ochoa-Estopier et al., 2014). The collected data can then be used for model training, validation and testing. If the resulting surrogate model accurately correlates the nonlinear input-output relationships of the target process system, its faster computation speeds can greatly relieve the aforementioned computational efforts (Ochoa-Estopier et al., 2014). It should be noted from the outset that various nonlinear surrogate models of the amine scrubbing processes have already been developed for both operational and economic assessments (Chung and Lee, 2020). Due to the recent progress of deep learning tools, fabrication of ANN models has become a popular approach for nonlinear data-driven modeling. A wide variety of ANNs have been utilized for modeling many different chemical processes, such as multicomponent fractionation, crude oil distillation units, hydrocrackers, and chemical synthesis process (Henao and Maravelias, 2011; Ochoa-Estopier et al., 2013; Osuolale and Zhang, 2016; Song et al., 2020), etc. For the amine scrubbers, multilevel ANN-based surrogate models were rigorously developed, investigated and compared by a large number of researchers with rigorous simulation studies (Goldstein et al., 2022; Henao and Maravelias, 2011). These ANN models have been used to correlate the key process features, such as the CO_2 capture levels, the rich loading and the specific reboiler duty (Sipöcz et al., 2011). The ANN-based surrogate models were also developed for steady-state and dynamic operations using bootstrap aggregated neural networks and deep belief networks (Li et al., 2015; Li et al., 2018). Such ANNs can also be used to replace objective functions for process optimization with various aqueous amine solvents or their blends (Nuchitprasittichai and Cremaschi, 2013).

Although the ANN-based surrogate models have been widely used to analyze and optimize the amine scrubbing processes, no published works were geared towards modeling of such scrubbers with ingeniously modified configurations. It can be expected that the simulation runs of these complex amine scrubbing plants can be even more time consuming and computationally expensive. This shortcoming makes data acquisition difficult and should be considered as one of the bottleneck steps for surrogate modeling (Kajero et al., 2017; Nuchitprasittichai and Cremaschi, 2013). As a result, the manageable sampling size should be determined a priori to reduce the total number of data acquisition intervals, and at the same time maintain the model accuracy. It should be noted that very few studies identified this practical problem in the past (Nuchitprasittichai and Cremaschi, 2013; Ochoa-Estopier et al., 2014), and to our best knowledge, no published studies attempted to address the related issues via model-building methodology. For efficient development of surrogate models, the ANNs can be built for independently mimicking the behaviors of absorber and stripper with data collected from the corresponding standalone simulation runs (Li et al., 2015; Sipöcz et al., 2011). However, not only the interactions among absorber, stripper and heat exchangers are clearly neglected in this approach, but also the key process mechanisms and implicit constraints may not be captured in the resulting models. Consequently, the predictions of the independent ANNs are bound to be inaccurate and thus unfit for practical applications. On the other hand, although less efficient during the data collection stage, the surrogate model trained with samples collected from plantwide simulations should be able to circumvent the above drawbacks.

To build a plantwide model, the most straightforward approach is to train an ANN from scratch using the expensive plantwide data. Another intuitive strategy may be to

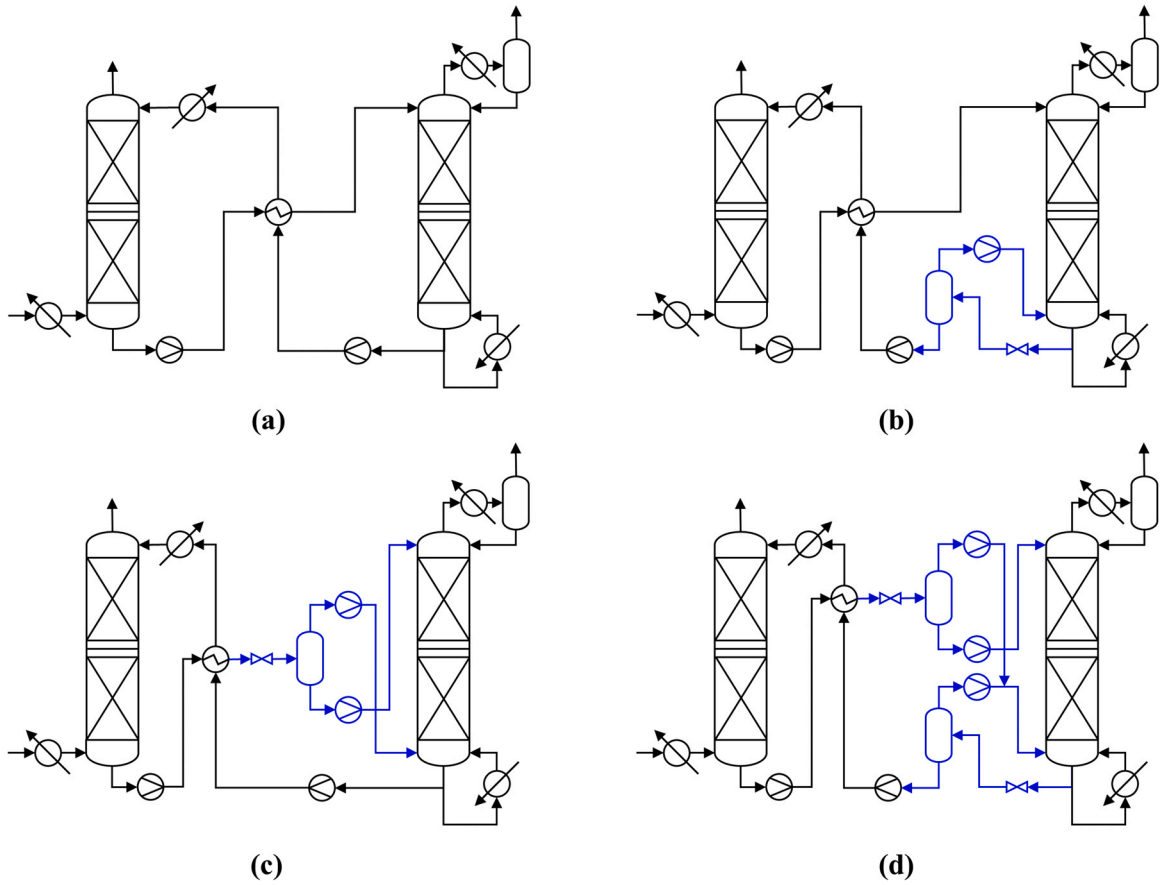

Fig. 1 – Process flow diagrams of various amine scrubbing CO_2 capture plants: (a) typical process, (b) lean vapor recompression (LVR), (c) rich vapor recompression (RVR), (d) LRVR.

simply mimic the commercial simulators, where the absorber model is first pre-trained and then cascaded in series by a stripper model. These two models are then bridged by variables characterizing the rich solvent conditions. Finally, the resulting cascaded model can be end-to-end trained with the plantwide data. However, to provide the stripper model with accurate and comprehensive information of the rich solvent conditions, additional outputs (i.e., rich temperature, rich flowrate, rich MEA concentration and rich CO_2 loading) should be explicitly learned and predicted by a relatively large absorber model. In addition, the prediction errors in these absorber model outputs may propagate to the stripper model and cause deterioration of the final learning outcomes. Furthermore, the overall system network structure can become unnecessarily complex since this model must be equipped with the global inputs (e.g., lean loading) and the local inputs of two columns (e.g., column pressure). Also, independently representing the effects of global inputs to absorber and stripper models may even give rise to the undesired parameter redundancy issues (Cheng et al., 2015), and the useful knowledge in the absorber model cannot be extracted and reused in the plantwide modeling step.

An innovative method is proposed in this paper to drastically lower the required simulation effort for ANN-based surrogate modeling of the amine scrubbing plant. To efficiently construct the model, the proposed method aims to cut down the data acquisition time by combining the concepts of process synthesis and progressive learning (Fayek et al., 2020; Henao and Maravelias, 2011). Since the computation loading of standalone absorber simulation is much lighter than its plantwide counterparts, it's preferable to make use of standalone simulations for the purpose of saving at least a portion of the total data acquisition time. Therefore, in the proposed procedure, a large amount of cheap absorber samples were collected from the standalone simulations, while a relatively smaller number of expensive samples were gathered from the large-scale plantwide simulations. The former is utilized to pre-train the absorber model, while the latter is used for transfer and progressive learning so as to migrate and expand the absorber model. Firstly, the pre-trained absorber model is fine-tuned so as to adapt to the plantwide domains. Secondly, an additional network block is then laterally and also layer-wisely connected to the original network, and extra input and output variables were also attached to the new model architecture. Finally, these newly added parameters can be trained with the expensive plantwide samples, while the original parameters were fixed and untrained. With the aforementioned steps, the required sampling size collected from plantwide simulation runs is expected to be easily manageable, while the accuracy of resulting model should still be kept at very high level. The aforementioned explicit learning tasks of the rich solvent conditions would seem to be unnecessary and can easily be replaced by imposing the interior lateral connections for sharing useful hierarchical features between tasks and to greatly simplify the network structures. This method was verified in this study by building static surrogate models of the amine scrubbing process with various configurations to show its general effectiveness.

2. Amine scrubbing process

2.1. Process descriptions

Amine scrubbing processes have long been adopted for sour gas sweetening in oil and gas industries to meet government regulations (Tikadar et al., 2020). To reduce carbon footprints, the same process is regarded as a promising means for post-combustion CO_2 capture in various industrial sectors. As illustrated in Fig. 1(a), the typical amine scrubber may be roughly divided into the absorbing and stripping sections. The former is usually operated at the atmospheric pressure (1 bar). The flue gas enters at the bottom of an absorption column and contacts counter-currently with the lean solvent introduced from the top. The CO_2 content in the flue gas is chemically absorbed by amine solvent in the column, in which the complex heat and mass transfers and exothermic chemical reactions take place. The treated gas stream is vented from the overhead, while the rich solvent with high CO_2 loading leaves from the bottom. The rich solvent is then regenerated in the pressurized stripper, where the captured CO_2 is desorbed and vented from the top as a vapor product with high CO_2 concentration. This vapor product is compressed again to a much higher pressure and then dehydrated to facilitate transportation. The lean solvent leaving from the bottom is recycled back to the top of absorber. For better energy efficiency, a large amount of heat is recovered at the lean/rich cross heat exchanger (LRHX), where the sensible heat in hot lean solvent is recovered by the cold rich solvent. In addition, it is a common practice to install an auxiliary cooler at the downstream of LRHX for the purpose of manipulating the temperature of lean solvent entering at the top of the absorber. To ensure operability, several pumps are also deployed to overcome the pressure drops in various parts of the system, such as vertical elevations, vessel nozzles and tubes of exchangers. The stripper operation is energy intensive since a large amount of heat should be supplied by the low pressure steam to its reboiler, and the overhead vapor should be partially condensed for reflux to maintain the column operation. In fact, it is quite challenging for to commercialize the aforementioned process due to high capital and operating costs.

Various improved processes have been developed to minimize the steam consumption or the total equivalent work with sophisticated designs such as absorber intercooling, exergetic integration and heat pumps (Le Moullec et al., 2014). The heat pumps in the improved design are typically realized by mechanical vapor recompression (MVR), where the vapor generated from flashing is recompressed to serve as an additional vapor source for the column. This practice results in less reboiling duty but more mechanical works (Li et al., 2022). To address this trade-off issue, numerous studies have tried to optimize the operating pressures to reduce the total equivalent work. For this reason, accurate surrogate models for these modified processes should also be made available so as to reduce the computation effort. However, to our best knowledge, no researches have tried to build these models systematically. In this paper, various modified configurations based on MVR techniques were studied, i.e., lean vapor recompression (LVR), rich vapor recompression (RVR) and lean/rich vapor recompression (LRVR). As shown in Fig. 1(b), (c) and (d), compared with the conventional process, additional drums, compressors and/or pumps are installed in these advanced configurations. The liquid streams with higher pressure from the bottom of absorber and/or stripper are flashed before feeding into their original downstream units. An amount of vapor is generated due the liquid flashing, and the vapor streams are then recompressed and injected to the bottom of the stripper to serve as an additional vapor source for column operation. Furthermore, the flashed rich solvent at lower temperature is pumped and fed to the top of the stripper, yielding in the reduction of condenser duty. The LVR using MEA solvent is the most studied MVR-based configurations, and it was reported to be able to save reboiler duty by 12.8–26.8% (Khalifa et al., 2022).

2.2. Process simulations

In this study, the amine scrubbing process was simulated with Aspen HYSYS. The selected amine solvent is monoethanolamine (MEA) with concentration around 30 wt% (Lin and Rochelle, 2016). Unlike the ideal process flow diagram depicted in Fig. 1, it should be noted that the water and MEA may escape from the systems with the vent gas and the CO_2 product. Hence, additional makeup streams are needed for maintaining water and MEA balances to ensure convergence. During simulation runs, the makeup quantities of their compositions were simultaneously calculated by a “Spreadsheet” module. The Acid Gas-Chemical Solvents property package in Aspen HYSYS was selected for rigorous calculations. This package supports property estimations of various amine solvents and their blends, and these estimates are produced on the basis of the Electrolyte Non-Random Two-Liquid (eNRTL) model and the Peng-Robinson equation of state. (Dubois and Thomas, 2018).

Both absorber and stripper were simulated with column modules using rate-based method and the default parameters. The mixed flow model were selected to characterize the rate-based columns. For both columns, due to the need for faster convergence, the “Efficiency” rate-based model of Aspen HYSYS was adopted, where the conventional equilibrium stages were assumed for the columns, and the rate-based efficiencies for CO_2 were estimated to account for the non-equilibrium behaviors in each stage (Dubois and Thomas, 2018). The Modified HYSYS Inside-Out algorithm with adaptive damping factor was selected as the column solvers (Oh et al., 2016). To satisfy the requirement of degree of freedom for simulating the stripper column, the condenser temperature and lean loading (mole ratio of CO_2 to MEA in the lean solvent) were chosen as the inputs to the simulator (Amrollahi et al., 2012), and the corresponding reboiler duty was then obtained through rigorous calculations. Also, the solvent circulation rate (mass flowrate of lean solvent) is another crucial parameter for plantwide simulations. The minimum approach temperature of lean/rich cross heat exchanger was set at 10 °C. For the process simulations of LVR and LRVR configurations, additional “Recycle” module should be attached to the outlet of recompressed lean vapor, which requires lengthier computation and longer data collection times.

2.3. Data acquisition

2.3.1. Automation of Aspen HYSYS

To generate enough samples for ANN-based surrogate modeling, HYSYS automation technique can be used to handle an extremely large number of repetitive steady-state

Table 1 – Operating ranges of the CO2 capture process.

Absorber	
Operating pressure	1.0 bar
Flue gas temperature	40–80 °C
Flue gas mass flowrate	30–100 kg/h
Flue gas CO_2 mole fraction	0.05–0.15
Lean solvent temperature	40–60 °C
Lean solvent mass flowrate	30–500 kg/h
Lean solvent MEA mass fraction	0.27–0.33
Lean solvent CO_2 loading	0.05–0.40
Stripper	
Operating pressure	1.5–3.0 bar
Condenser temperature	40–80 °C
Lean solvent CO_2 loading	0.05–0.40
Lean flashed pressure*	0.3–2.7 bar
Rich flashed pressure*	0.3–0.9 bar
Cross heat exchanger	
Minimum approach temperature	10 °C

*Parameters used only in MVR-based amine scrubbers.

simulations. This task is accomplished by interactive accesses using external codes via the component object model (COM) protocol. In this study, the COM protocol is supported by the Python win32com package. To effectively extract the general information of amine scrubbing process, the multivariate operating conditions were pseudo-randomly sampled within the given design ranges using the popular Latin hypercube sampling (LHS) technique (Kajero et al., 2017). It is expected that process mechanisms may be correctly extracted by feeding randomly varied parametric combinations, e.g., the flue gas temperature, flue gas flowrate, flue gas compositions, lean temperature, lean loading, solvent circulation rate, stripper pressure, condenser temperature and flashed pressure, etc., to the simulator. The detailed ranges of these specifications are listed in Table 1. Parameters related to solvent flashing were only used in the simulation runs of MVR-based processes. To identify possible improvements of the amine scrubbing processes on the basis of mathematical models, a wide operating range is required (Ahn et al., 2013). Therefore, as listed in Table 1, the design ranges of various parameters were deliberated enlarged in order to generate sufficiently general information for data-driven modeling. To satisfy material balances, the flue gas compositions sampled by the LHS technique were normalized before feeding into the simulator. Critical stream properties of flue gas, lean solvent, rich solvent, vent gas and CO_2 product were recorded during simulation runs, and the operating conditions and the design specifications of absorber, stripper and heat exchangers were also collected.

2.3.2. *Data acquisition time*

In this study, a computer with Intel Core i7–7700 CPU 3.60 GHz was used for data acquisition from process simulation results. During the data acquisition steps, the data collecting intervals were recorded to estimate the data acquisition speeds, and the results may be further used for the evaluation of the efficiencies of modeling methods detailed in the following sections. The data acquisition times for collection of 1000 samples are shown in Fig. 2. It should be noted that the specifications generated by LHS technique for simulations may result in divergence (Ochoa-Estopier et al., 2014), but the time intervals of such cases should still be counted. It can be clearly seen that the standalone

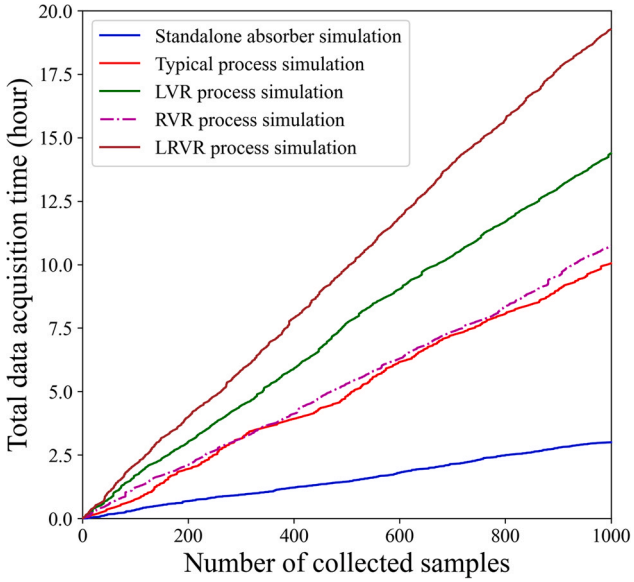

Fig. 2 – The data acquisition time needed by carrying out rigorous simulation runs for various modified processes.

simulation of absorber took the least amount of time, i.e., 3 h, to collect 1000 samples, while those of the other configurations were at least 3–7 times longer. The simulation results of the typical and RVR processes are similar, and they both yield the similar data acquisition speeds. It took approximately 10 h to collect 1000 samples for these two cases. On the other hand, notice that much greater time spans are needed to collect enough samples for modeling the LVR and LRVR processes. Specifically, it took around 15 and 20 h to collect 1000 samples for the corresponding model building tasks. This due to the fact that the simulated LVR and LRVR processes are more complex than the others. The results given in Fig. 2 also once again indicates that it is necessary to upgrade frameworks in practical applications, e.g., optimization (Zhu et al., 2022), for improving the calculation efficiency by incorporating with the surrogate models. In data acquisition stage for data-driven modeling, the sizes of plantwide datasets should therefore be properly determined. More specifically, the total data acquisition time should suppressed to an as-low-as-possible level (Nuchitprasittichai and Cremaschi, 2013).

3. Model building methodology

The proposed surrogate modeling procedure for amine scrubbing processes is introduced below. It can be observed from Fig. 3 that the complete process model is sequentially synthesized according to unit models. By following this approach, an innovative surrogate modeling procedure is proposed in this study to effectively reduce the required number of samples via progressive learning. Generally speaking, this model-building procedure consists of three successive steps, i.e., (1) pre-training, (2) fine-tuning and (3) progressive learning through model expansion (Rusu et al., 2016). In the following sub-sections, the descriptions of the above steps are given in further details.

3.1. *Pre-training of absorber model*

Traditional rigorous simulation of the integrated absorber-exchanger-stripper system calls for extremely lengthy and

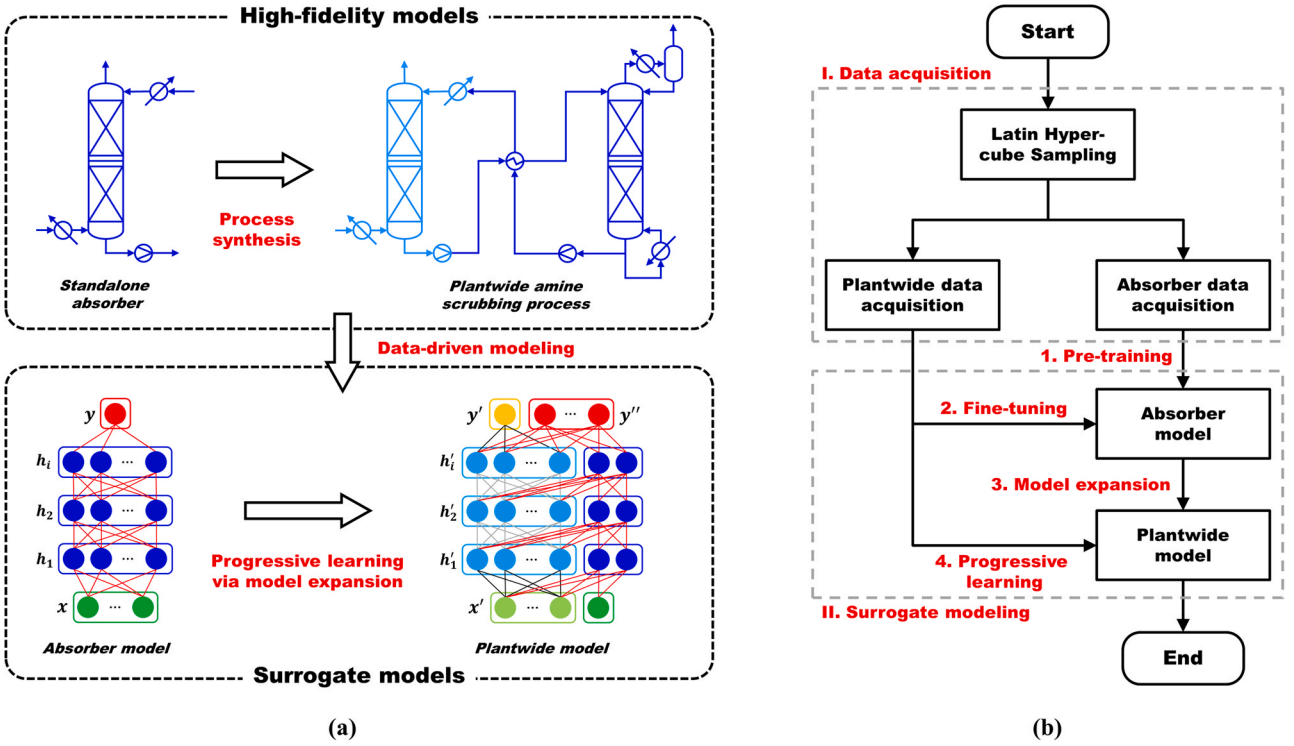

Fig. 3 – (a) conceptual analogy of process synthesis and progressive modeling via model expansion and (b) proposed flowchart for surrogate modeling of amine scrubbers.

iterative computation. Extra recycle loops may also be required if this process is modified to form one with an even more complicated configuration, e.g., LVR. Furthermore, if additional design specifications are imposed, e.g., 90% CO_2 removal rate, owing to the degree of freedom is saturated for simulations, repeated iterative adjustments would be required to reach the desired design targets. Consequently, the corresponding computation effort can be overwhelming. This and other computation issues obviously motivate the use of simpler surrogate models. However, the tedious task of data acquisition from first-principle models is still considered to be inevitable for the data-driven modeling approaches (Kajero et al., 2017; Ochoa-Estopier et al., 2014). Therefore, a novel model-building procedure has been developed in this work so as to shorten the data acquisition time while still resulted in the same or even more accurate predictions.

To facilitate faster convergence in simulating the amine scrubbing process, the plantwide flowsheet is first reduced to a standalone absorber (Alie et al., 2005). This practice certainly shorten the total computation time because of the fact that the interactions among different units in the given process via solvent recycling loops and LRHX are ignored completely. Since the standalone simulations are implicitly less constrained, the required total computation time should be much shorter and, naturally, the collection cost of the standalone samples is significantly cheaper. As mentioned before, the collected data from the standalone simulation are supposed to be utilized for supervised pre-training of absorber model. In this study, the multi-layer feedforward neural network (FNN) structures were adopted for ANN-based surrogate modeling. The mathematics of FNNs can be simply expressed in the following form:

$$\mathbf{K}_i = \phi(\mathbf{:}_i \mathbf{K}_{i-1}) \quad (1)$$

where $\mathbf{K}_i$ and $\mathbf{:}_i$ denote the hidden state vector and the weight matrix of layer i respectively, and ϕ is the linear or nonlinear activation function, e.g., sigmoid or hyperbolic tangent function. To mimic the input-output relationships of the rigorous models embedded in the process simulator, seven input variables, i.e., flue gas flowrate, flue gas temperature, flue gas CO_2 concentration, lean flowrate, lean temperature, lean MEA concentration and lean loading, have been selected for the surrogate model. On the other hand, the only output variable was chosen to be the CO_2 emission rate, which is determined by the CO_2 content in the gas stream vented from the absorber.

3.2. *Fine-tuning of absorber model*

From the practical perspective of process modeling, training models from scratch for each new tasks is tedious and inefficient. Therefore, transfer learning is an attractive technique to avoid such inefficiency. The core concept of transfer learning is to preserve the previously stored knowledge embedded in the source domain (i.e., the base models) and migrate the same information to the target domain (Pan and Yang, 2010). The useful knowledge stored in the base models is retained and transferred, while the invalid ones are updated and/or discarded. More specifically, transfer learning aims to reuse the base models in a systematic way by judiciously fine-tuning or preserving their weights and biases (Yosinski et al., 2014). With this strategy, it is expected that the required sampling size can be made smaller than that needed by the conventional methods. This data handling technique has been successfully applied to model various types of chemical processes, including multicomponent distillation column, non-isothermal continuous stirred tank reactor (CSTR), polymer injection molding, and steam cracking furnace (Bi et al., 2020; Chuang et al., 2018; Hsiao et al., 2021; Lu et al., 2009). These transfer learning-based

methods were developed to overcome the common problem of insufficient data in the field of chemical process engineering, and the main reason why transfer learning was successful in the above cases is that the high process similarities between the source domains and the target domains (Lu et al., 2009).

Due to the high process similarity, the absorber behaviors in standalone and plantwide simulations are believed to be closely related (Alie et al., 2005). Therefore, it is reasonable to expect that the knowledge gained from the standalone domain would be useful and transferable to the plantwide domains. However, there are still differences between the results obtained from standalone and plantwide simulations (Alie et al., 2005), which may be originated from the fact that the standalone simulation is not able to take unrecognizable implicit constraints related to solvent circulation into consideration. Such constraints, e.g., the recycling lean solvent conditions and the water/MEA balance, etc., are utilized in plantwide simulations, and exert great influences on model convergence and process behaviors. Although the above differences may be subtle, they are nonetheless still unknown in advance. As a result of these concerns based on the domain knowledge, it is considered in this study to be a conservative but reliable practice to fine-tune the model before the further training step in order to effectively capture the hidden information of implicit constraints. On the basis of this rationale, the second step of the proposed modeling procedure is designed to fine-tune the parameters of the pre-trained absorber model.

In this study, the transfer learning of pre-trained absorber model was accomplished by the parametric fine-tuning method (Yosinski et al., 2014). In particular, most of the network parameters of the pre-trained models are frozen, and only a small number of them were updated so as to adapt to the plantwide domain. More specifically, the parameters of the pre-trained absorber model were partially fine-tuned in a second supervised training procedure according to the plantwide data. Through this way, the network parameters stored in the pre-trained models may be preserved and transferred to construct the final models. From the technical aspect, with the above approach, the data-to-trainable parameter ratio can be effectively increased, and better model performances may be expected in comparison with the ones trained from scratch using same amount of samples. By following the aforementioned modeling strategy in the present study (Hsiao et al., 2021), the weights and biases of the first hidden layer and the output layers were updated in the standalone model, while the ones in the intermediate hidden layers are retained and transferred.

3.3. *Progressive learning through model expansion*

Progressive learning is an emerging field for deep learning, where the ANN architectures are progressively modified to integrate the new data with the help of previously gained knowledge (Fayek et al., 2020; Parisi et al., 2019). The idea of progressive learning is closely related to those of transfer learning, continual/lifelong learning and multi-task learning (Zhang and Yang, 2022). They are more advanced than their conventional counterparts, i.e., independent learning methods, where the latter train the models from scratch once new tasks are given (Fayek et al., 2020). The main targets of progressive learning are not only knowledge transfer but also sequential multi-task learning via expanding and/or pruning the architectures of the existing models. In the practical applications, there may exist several new process configurations (Yin et al., 2020), and the conventional transfer learning without modifying model architecture may not be applicable. Therefore, under such circumstances, the progressive learning approach should be adopted to expand the connections and parameters of the input layer, output layer and/or hidden layers. The model under construction may be able to progressively acquire the new system mechanisms on the basis of already available knowledge (Fayek et al., 2020; Parisi et al., 2019). Based on the ANN modularization techniques (Terekhov et al., 2015), expandable deep neural networks have been adopted to solve this problem by laterally expanding the number of parameters in a layer-wise manner with or without pruning existing parameters (Rusu et al., 2016; Yoon et al., 2018), and the number of input and/or output variables can be also extended according to the new input-output relationships (Yin et al., 2020).

In the last step of the proposed method, the progressive deep learning technique was adopted to expand the architecture in every layer of the fine-tuned absorber model. To prevent catastrophic interferences (Fayek et al., 2020; Parisi et al., 2019), the parameters of the fine-tuned absorber model are retained during the progressive learning step, and the expanded parameters are merely laterally connected to the existing model without pruning its original architecture (Rusu et al., 2016; Terekhov et al., 2015). The layer-wise matrix multiplications within the original network *block* still follows the Eq. (1), while the ones of the lateral connections between the original and expanded network blocks may be expressed as:

$$\mathbf{K}_i^{(k)} = \phi\left(\mathbf{:}_i^{(k)}\mathbf{K}_{i-1}^{(k)} + \sum_{j<k} \mathbf{8}_i^{(k:j)}\mathbf{K}_{i-1}^{(j)}\right) \tag{2}$$

where i, j and k are the indices of the hidden layers, the original block and the expanded block respectively. ϕ is the linear or nonlinear activation function. $\mathbf{K}_i^{(j)}$ and $\mathbf{K}_i^{(k)}$ denote the hidden state vectors of layer i of block j and k respectively, while $\mathbf{:}_i^{(k)}$ is the weight matrix of layer i of block k, and $\mathbf{8}_i^{(k:j)}$ is the newly initialized weight matrix of the lateral connection (:) from layer i-1 of block j to layer i of block k (Fayek et al., 2020; Rusu et al., 2016).

Through this practice, the knowledge gained for predicting CO_2 emission rate from the absorber model obtained in pre-training and fine-tuning steps may be effectively kept intact, and the useful inter-layer information may also be utilized to assist the learning of new process mechanisms. Practically, these modeling goals are realized by placing the lateral and layer-wise connections between a new block of FNN and the original ones to make proper use of the previously acquired knowledge. Furthermore, additional inputs (i.e., condenser temperature, stripper pressure and solvent flashed pressures) and two extra outputs (i.e., reboiler duty and mechanical vapor recompression duty) can also be attached along with the model expansion, and are then fitted in a supervised-learning manner. It can be observed from Table 2 that the input-output structures of the four models under study are quite different. Notice especially that the number of input variables increases with the complexity of process. The selected input and output variables of the above-mentioned processes are specified in Table 3.

Table 2 – The hyper-parameters for surrogate modeling of various amine scrubbers.

Hyper-parameters	Typical	LVR	RVR	LRVR
Number of input variables	9	10	10	11
Number of output variables	2	3	3	3
Maximum number of the modeling samples	1000	1000	1000	1000
The baseline method				
Number of hidden layers	2	2	2	2
Number of hidden neurons	10	15	15	20
Number of parameters	232	453	453	723
The proposed method				
Number of hidden layers	2	2	2	2
Number of hidden neurons	10 + 3	10 + 10	10 + 10	10 + 10
Number of parameters*	86	362	362	372

*The number of trainable parameters at the progressive learning stage.

4. Network architecture, training, and evaluation metrics

For comparison purpose, the baseline method was based on the independent learning approach, where the ANN models were directly trained from scratch with plantwide samples. The model predictions of the proposed method have been compared with those obtained with the baseline method. The hyper-parameters for various modeling tasks are listed in Table 2. After the preliminary cross validations by the collected datasets, two-layer FNNs were utilized for all cases in this study, and the maximum sample number was 1000. Each hidden layer in the absorber model accommodates 10 neurons, and they are laterally connected by an additional network block with 3 or 10 neurons at each layer. For the baseline cases, two-layer FNNs with 10–20 hidden neurons were used. Since a large amount of parameters were frozen, the number of trainable parameters of the proposed model was much smaller than that of the baseline method. The activation function used in this study is the scaled exponential linear unit (SELU), while the linear function is used for the output layers. Since SELU was adopted for nonlinear activations, Lecun normal initializer was selected, and the data were preprocessed via standard-deviation normalization. The learning rates for all training tasks were set to be 0.001, and the Adam optimizer was selected to minimize the overall losses.

The main goal of the numerical experiments reported below is to cross validate the model prediction results obtained by the baseline and the proposed methods. For all modeling tasks, a total of 1000 samples were selected and fixed to serve as the testing sets, while the remaining samples were treated as the modeling sets. In each set, 70% samples were used for training and 30% for validation. In order to quantitatively evaluate the performances of various ANN models, two common statistical metrics were used, i.e., the R-squared (R^2) score and the root-mean-square error (RMSE). The former provides the degree of linearity between predicted values and the ground truth, while the latter shows the deviations of predicted values from the real values. The formulas for calculating these metrics are given below in Eqs. (3) and (4).

$$R^2 = 1 - \frac{\sum_n (\hat{y}_n - y_n)^2}{\sum_n (\bar{y} - y_n)^2} \tag{3}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N} (\hat{y}_n - y_n)^2} \tag{4}$$

Modeling of the amine scrubbing process should be considered as a multi-output regression problem. Due to the heterogeneous physical meanings, it would be necessary to appropriately compare the averaged performances on different outputs, i.e., the CO_2 emission rate, the reboiler duty and the MVR duty. Therefore, in this study, the averaged mean squared error (MSE) of normalized data was adopted to compare the overall performances of different models (Xu et al., 2020). The formula for averaged MSE is given in Eq. (5), where M denotes the number of output variables.

$$avg.\ MSE = \frac{1}{M}\sum_{m=1}^{M}\left(\frac{1}{N}\sum_{n=1}^{N}(\hat{y}_n - y_n)^2\right) \tag{5}$$

5. Results and discussions

5.1. *Pre-training and fine-tuning of absorber model*

Fig. 4 shows the testing results of the pre-trained absorber models obtained with various different datasets. It can be clearly seen that the absorber model pre-trained by 750 samples may be able to accurately predict the CO_2 emission rates in various scenarios, where the R^2 scores were around 0.99. This is due to the fact that the absorber configurations under study were highly similar (Alie et al., 2005). In other words, Fig. 4 indicates that the proposed absorber model was able to accurately predict the CO_2 emission rate with a fixed number of input features regardless of the modifications introduced into the stripper section. Therefore, the standalone absorber model may be easily portable during the later fine-tuning and progressive learning steps. This practice guarantees the model quality for predicting the CO_2 emission rates. In this study, the absorber model pre-trained with 750 samples was selected as the base model for the subsequent modeling steps. In the data acquisition stage, only 2.310 h were needed to collect 750 absorber samples with standalone simulation, which is only around 12–23% of the time spent for the plantwide simulations. As given in Figs. 5–8, the prediction performance of absorber model may be further enhanced in the fine-tuning step. With only 400 new plantwide data used for fine-tuning, 1.6–6.5% of the prediction errors were effectively eliminated. Such improvements are quite significant, especially under the condition that the initial accuracies were already at extremely high levels.

Table 3 – Input-output structures of various surrogate models.

Models	Absorber	Typical	LVR	RVR	LRVR
Input variables					
Flue gas temperature	○	○	○	○	○
Flue gas mass flowrate	○	○	○	○	○
Flue gas CO_2 mole fraction	○	○	○	○	○
Lean solvent temperature	○	○	○	○	○
Lean solvent mass flowrate	○	○	○	○	○
Lean solvent MEA mass fraction	○	○	○	○	○
Lean solvent CO_2 loading	○	○	○	○	○
Condenser temperature		○	○	○	○
Stripper pressure		○	○	○	○
Lean solvent flashed pressure			○		○
Rich solvent flashed pressure				○	○
Output variables					
CO_2 emission rate	○	○	○	○	○
Reboiler duty		○	○	○	○
Vapor recompression duty			○	○	○

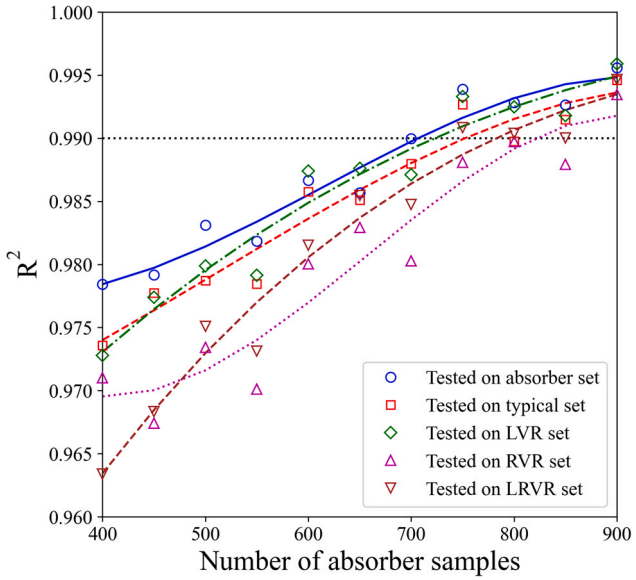

Fig. 4 – Testing results of pre-trained absorber model with different datasets.

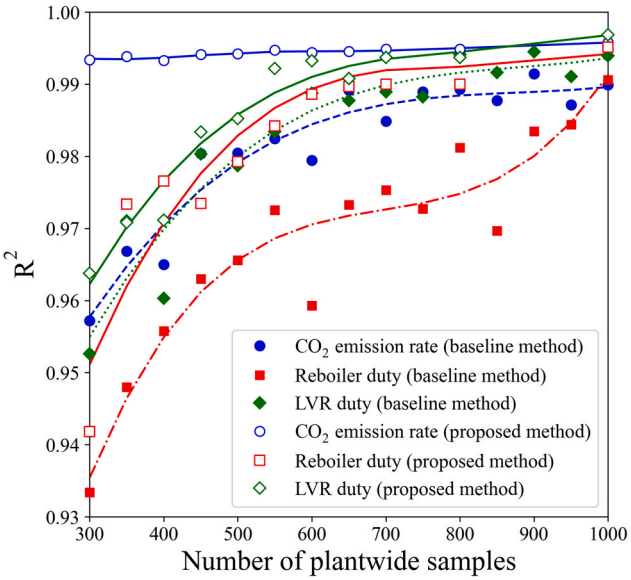

Fig. 6 – Testing results of surrogate model of LVR process.

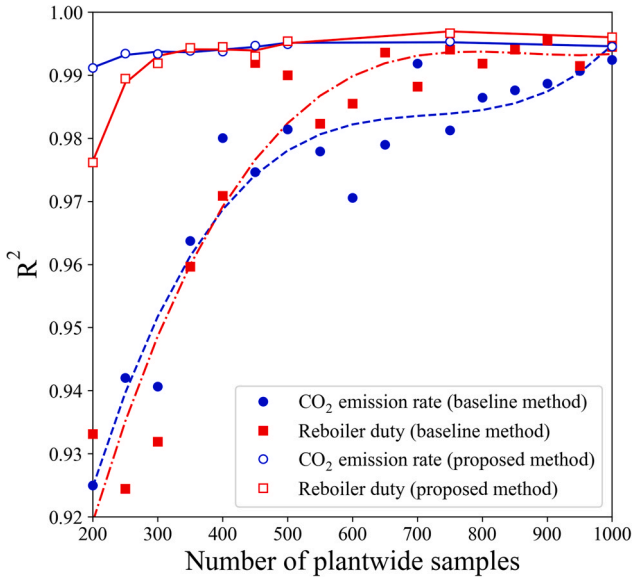

Fig. 5 – Testing results of surrogate model of typical process.

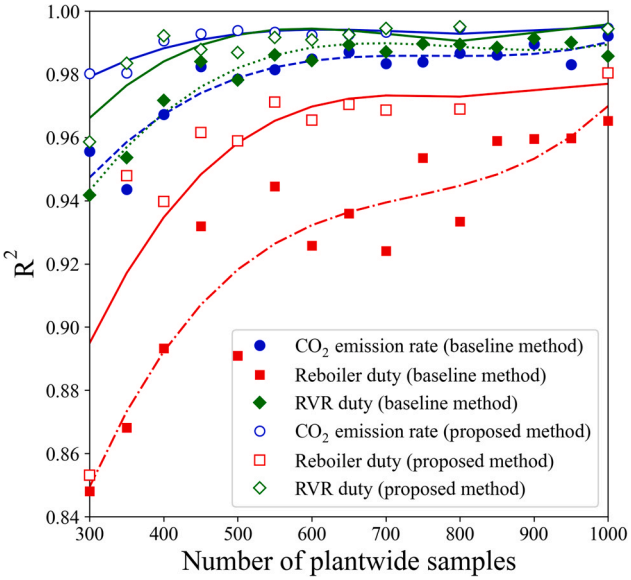

Fig. 7 – Testing results of surrogate model of RVR process.

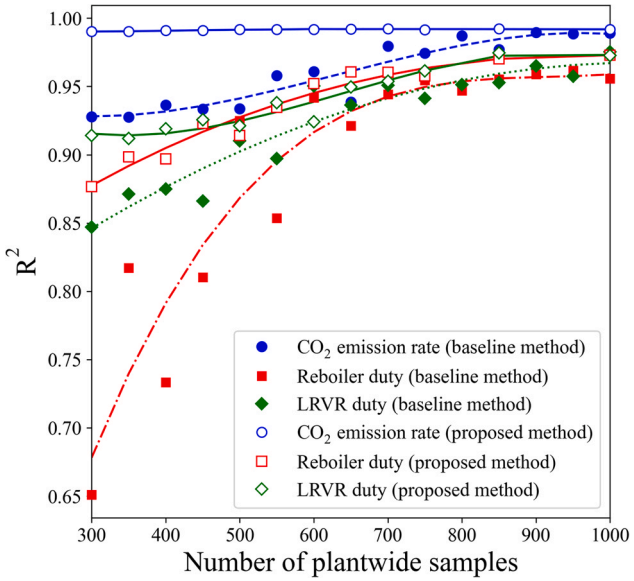

Fig. 8 – Testing results of surrogate model of LRVR process.

5.2. *Surrogate models of various processes*

5.2.1. *Surrogate models of typical process*

The testing results of surrogate models of the typical process constructed using different methods are given in Fig. 5. These models were tested by 1000 validation samples which have not been used in the training steps. It can be clearly seen that the proposed method outperforms the baseline method. The prediction accuracies of both CO_2 emission rate and reboiler duty achieved by the proposed models were relatively higher than those of their baseline counterparts. When only 250 samples were used, the testing R^2 scores of the proposed model were around 0.99, while the ones of the baseline model were around 0.92–0.94. Although the testing R^2 scores of the baseline models can be enhanced significantly to around 0.98 when 500–600 samples were available, the testing R^2 scores of the proposed models were meanwhile greater than 0.99. If it is desired to achieve testing R^2 score approaching 0.99 with the baseline method, 1000 modeling samples (or more) may have to be secured, while only 350 samples were enough via the suggested procedure. The above observations indicate over 64% of data acquisition time was reduced. Moreover, it can be noticed that the testing performances of the proposed models were much more stable than their conventional counterparts.

5.2.2. *Surrogate models of LVR process*

The testing results of LVR models can be found in Fig. 6. The learning task of each MVR process was more difficult than that of the typical process due to an additional output, i.e., the compression duty. For the proposed models, the testing R^2 scores on the CO_2 emission rate were stably greater than 0.99 regardless of the number of modeling samples, while the ones of the baseline models reached 0.99 only when 900–1000 samples were available. On the other hand, for the reboiler duty and the LVR duty, the testing performances of proposed models were enhanced much more noticeably if compared with their baseline counterparts. When 550 samples were available, the prediction accuracy of the proposed model was similar to that of the baseline model trained with 1000 samples, which implies that over 25% of data acquisition time can be reduced by using the suggested method. If the number modeling samples was increased greater than 600, the proposed models were clearly superior, especially the predictions of the reboiler duty were much more accurate than those of the baseline models.

5.2.3. *Surrogate models of RVR process*

In the case of RVR process, the proposed models also performed better if same amount of plantwide samples were used for modeling. Similar trends in prediction results of typical and LVR processes, the testing R^2 scores of the proposed models were again higher even when the number of modeling samples was reduced to around 400. It can be observed that the proposed model trained with 550 samples was able to perform better than the baseline model trained with 1000 samples, where the values of averaged MSE were approximately 0.015 and 0.019 respectively. For the proposed model trained with 550 samples, the testing RMSEs of the CO_2 emission rate, the reboiler duty and the RVR duty were 0.237 kg/h, 1.867 kW and 0.041 kW respectively, while those for the baseline model trained with same number of samples were 0.395 kg/h, 2.524 kW and 0.053 kW respectively. These indicators represent 40%, 26% and 23% reductions in the corresponding prediction errors. Furthermore, they imply that over 24% of data acquisition time was reduced by using the proposed methodology.

5.2.4. *Surrogate models of LRVR process*

The model predictions of the LRVR process are given in Fig. 8. It can be observed that the steps of absorber pre-training and fine-tuning had positive effects on the enhancement of model performances, especially when the number of modeling samples were relatively few. When only 300 samples were available, the testing R^2 scores of the CO_2 emission rate, the reboiler duty and the LRVR duty were estimated to be 0.990, 0.877 and 0.914 respectively, and the corresponding RMSEs were 0.194 kg/h, 3.612 kW and 0.161 kW respectively. Notice that these results were much better than those of the baseline model, where the RMSE values were respectively 0.536 kg/h, 5.891 kW and 0.216 kW. By comparing the above two sets of indicators, it can be deduced that approximately 63.8%, 38.7% and 25.5% of prediction errors were eliminated. The proposed model trained with 750 samples was able to produce predictions with accuracy similar to that accomplished by a baseline model trained with 1000 samples. The corresponding averaged MSEs were 25.1×10^{-3} and 28.5×10^{-3} respectively, and these results imply that around 23.1% of data acquisition time was saved.

5.3. *Further discussions*

5.3.1. *Enhancement of model accuracy*

In this subsection, the effectiveness of the proposed method on model accuracy enhancement is demonstrated. From the results given previously in Section 5.2, it can be observed that the proposed models possess the properties of higher start, higher slope and higher asymptote if compared with the corresponding baseline models trained with same number of samples. Consequently, one can intuitively deduce that application of the proposed method results in higher accuracy than the baseline method. However, notice that it is necessary for the proposed models to be pre-trained, which incurs penalty on data acquisition time (i.e., 2.310 h). Thus, it would be unfair to compare the models trained with same number of plantwide samples but created with different methods.

Table 4 – Model performances with similar data acquisition times.

Process	Models (No. of samples)	Time (hour)	RMSEs			Averaged MSE (normalized)	
			CO_2 emission rate (kg/h)	Reboiler duty (kW)	MVR duty (kW)		
Typical	Baseline (1000)	10.048	0.192	0.699	–	6.47×10^{-3}	-39.6%
	Proposed (750)	9.906	0.149	0.545	–	3.91×10^{-3}	
LVR	Baseline (1000)	14.384	0.181	1.008	0.038	9.62×10^{-3}	-14.4%
	Proposed (800)	14.005	0.161	0.935	0.036	8.23×10^{-3}	
RVR	Baseline (1000)	10.737	0.254	2.000	0.054	19.20×10^{-3}	-27.0%
	Proposed (800)	10.672	0.207	1.890	0.032	14.01×10^{-3}	
LRVR	Baseline (1000)	19.264	0.204	2.056	0.085	25.09×10^{-3}	-20.9%
	Proposed (850)	19.025	0.173	1.690	0.086	19.85×10^{-3}	

Table 5 – Total data acquisition times with similar model performances.

		Methods	Typical	LVR	RVR	LRVR
Number of plantwide samples		Baseline	1000	1000	1000	1000
		Proposed	350	550	550	750
RMSEs	CO_2 emission rate (kg/h)	Baseline	0.192	0.181	0.254	0.204
		Proposed	0.170	0.162	0.237	0.181
	Reboiler duty (kW)	Baseline	0.699	1.008	2.000	2.056
		Proposed	0.712	1.141	1.869	2.013
	MVR duty (kW)	Baseline	–	0.038	0.054	0.085
		Proposed	–	0.040	0.041	0.107
Averaged mean squared error (normalized)		Baseline	6.47×10^{-3}	9.62×10^{-3}	19.20×10^{-3}	25.09×10^{-3}
		Proposed	5.72×10^{-3}	10.81×10^{-3}	15.45×10^{-3}	28.54×10^{-3}
Total data acquisition time (hr)		Baseline	10.048	14.384	10.737	19.264
		Proposed	3.613	10.777	8.115	14.832
			-64.0%	-25.1%	-24.4%	-23.1%

Therefore, rather than models trained with the same number of plantwide samples, models trained with similar total data acquisition time were adopted for comparison from a different angle.

As shown in the Table 4, with similar data acquisition times, the proposed models had superior performances over their baseline counterparts. 14–40% reductions in the averaged MSEs was achieved for different amine scrubbing processes. It can also be seen that the complexities of process configurations may affect the model performances. With the simplest flowsheet, the averaged MSE of typical process can be reduced 39.6%. On the other hand, the proposed models of more complex processes, e.g., LVR and LRVR, both realized relatively small reductions on the averaged MSE values. On the other hand, the RMSE values achieved with the proposed models for various output variables, i.e., CO_2 emission rate, reboiler duty and MVR duty, were also equal to or even better than those of the corresponding baseline models. From the above results, it can be concluded that the proposed model-building strategy can be regarded as a novel and effective tool for enhancing the model performances if similar computation efforts are desired for the modeling tasks.

5.3.2. *Reduction of data acquisition time*

As mentioned before, the main purpose of this study is to reduce the data acquisition time for surrogate modeling of the amine scrubbing process by following the proposed modeling procedure. Therefore, in this subsection, in contrast to the previous one, models with similar prediction accuracies were compared. As shown in Table 5, the baseline models were trained with a sample size of 1000, while the proposed models were trained with 350–750 plantwide samples according to different target systems. Pre-trained by 750 absorber samples, the proposed models were able to obtain target accuracies based on smaller numbers of modeling samples if compared with their baseline counterparts. Take the LVR case as an example, it only took 550 samples to construct a model with an averaged MSE around 0.01 by the proposed method, while 1000 samples are needed for the conventional approach to reach similar performance. This finding indicates around 25% of data acquisition time were reduced. For the case of the typical process, 64% of data acquisition time was eliminated, while 23–25% were reduced for the MVR cases.

6. Conclusions

The amine scrubbing processes are regarded as attractive options for post-combustion CO_2 capture. However, since rigorous simulation of this complex processes could result in computation inefficiency for more advanced applications (such as optimization), the use of surrogate models become attractive in such scenarios. However, for the data-based modeling methods, it is inevitable to take an extremely large amount of time for data acquisition (Kajero et al., 2017). In this study, an innovative surrogate modeling strategy, which makes use of the concepts of process synthesis and progressive learning, is proposed to reduce the data acquisition time and also enhance the model performances. In this proposed method, the surrogate models were first pre-trained and fine-tuned with absorber samples collected from standalone simulations, and the models were then successively expanded to accommodate the knowledge for inferencing the output variables related to the stripper

sections. These surrogate models were constructed to predict the CO_2 emission rate, the reboiler duty and the mechanical compression duty. To achieve the designated degree of prediction accuracies, the results obtained in extensive case studies show that 25–65% of plantwide sampling sizes and 23–64% of data acquisition time may be reduced by using the proposed modeling procedure. Furthermore, the predictive errors were reduced by 14–40% when similar computation efforts were made on data acquisition.

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.

References

Ahn, H., Luberti, M., Liu, Z., Brandani, S., 2013. Process configuration studies of the amine capture process for coal-fired power plants. J. Greenh. Gas. Control 16, 29–40.

Alarie, S., Audet, C., Gheribi, A.E., Kokkolaras, M., Le Digabel, S., 2021. Two decades of blackbox optimization applications. Eur. J. Comput. Optim. 9.

Alie, C., Backham, L., Croiset, E., Douglas, P.L., 2005. Simulation of CO_2 capture using MEA scrubbing: a flowsheet decomposition method. Energy Convers. Manag 46, 475–487.

Amrollahi, Z., Ystad, P.A.M., Ertesvåg, I.S., Bolland, O., 2012. Optimized process configurations of post-combustion CO_2 capture for natural-gas-fired power plant – Power plant efficiency analysis. J. Greenh. Gas. Control 8, 1–11.

Bi, K., Beykal, B., Avraamidou, S., Pappas, I., Pistikopoulos, E.N., Qiu, T., 2020. Integrated modeling of transfer learning and intelligent heuristic optimization for steam cracking process. Ind. Eng. Chem. Res. 59, 16357–16367.

Chen, X., Wu, K., Bai, A., Masuku, C.M., Niederberger, J., Liporace, F.S., Biegler, L.T., 2021. Real-time refinery optimization with reduced-order fluidized catalytic cracker model and surrogate-based trust region filter method. Comput. Chem. Eng. 153.

Cheng, Y., Yu, F.X., Feris, R.S., Kumar, S., Choudhary, A., Chang, S.-F., 2015. An exploration of parameter redundancy in deep networks with circulant projections. Proc. IEEE Int. Conf. Comput. Vis. 2857–2865.

Chuang, Y.-C., Chen, T., Yao, Y., Wong, D.S.H., 2018. Transfer learning for efficient meta-modeling of process simulations. Chem. Eng. Res. Des. 138, 546–553.

Chung, W., Lee, J.H., 2020. Input–output surrogate models for efficient economic evaluation of amine scrubbing CO_2 capture processes. Ind. Eng. Chem. Res. 59, 18951–18964.

Damartzis, T., Papadopoulos, A.I., Seferlis, P., 2016. Process flowsheet design optimization for various amine-based solvents in post-combustion CO_2 capture plants. J. Clean. Prod. 111, 204–216.

Dubois, L., Thomas, D., 2018. Comparison of various configurations of the absorption-regeneration process using different solvents for the post-combustion CO_2 capture applied to cement plant flue gases. J. Greenh. Gas. Control 69, 20–35.

Fayek, H.M., Cavedon, L., Wu, H.R., 2020. Progressive learning: A deep learning framework for continual learning. Neural Netw. 128, 345–357.

Ferrara, G., Lanzini, A., Leone, P., Ho, M.T., Wiley, D.E., 2017. Exergetic and exergoeconomic analysis of post-combustion CO_2 capture using MEA-solvent chemical absorption. Energy 130, 113–128.

Goldstein, D., Heyer, M., Jakobs, D., Schultz, E.S., Biegler, L.T., 2022. Multilevel surrogate modeling of an amine scrubbing process for CO_2 capture. AIChE J. 68.

Henao, C.A., Maravelias, C.T., 2011. Surrogate-based superstructure optimization framework. AIChE J. 57, 1216–1232.

Hoff, K.A., Svendsen, H.F., 2013. CO_2 absorption with membrane contactors vs. packed absorbers-Challenges and opportunities in post combustion capture and natural gas sweetening. Energy Procedia 37, 952–960.

Hsiao, Y.-D., Kang, J.-L., Wong, D.S.H., 2021. Development of robust and physically interpretable soft sensor for industrial distillation column using transfer learning with small datasets. Processes 9.

Kajero, O.T., Chen, T., Yao, Y., Chuang, Y.-C., Wong, D.S.H., 2017. Meta-modelling in chemical process system engineering. J. Taiwan Inst. Chem. Eng. 73, 135–145.

Khalifa, O., Alkhatib, I.I.I., Bahamon, D., Alhajaj, A., Abu-Zahra, M.R.M., Vega, L.F., 2022. Modifying absorption process configurations to improve their performance for post-combustion CO_2 capture – What have we learned and what is still missing? Chem. Eng. J. 430.

Le Moullec, Y., Neveux, T., Al Azki, A., Chikukwa, A., Hoff, K.A., 2014. Process modifications for solvent-based post-combustion CO_2 capture. J. Greenh. Gas. Control 31, 96–112.

Li, F., Zhang, J., Oko, E., Wang, M., 2015. Modelling of a post-combustion CO_2 capture process using neural networks. Fuel 151, 156–163.

Li, F., Zhang, J., Shang, C., Huang, D., Oko, E., Wang, M., 2018. Modelling of a post-combustion CO_2 capture process using deep belief network. Appl. Therm. Eng. 130, 997–1003.

Li, K., Leigh, W., Feron, P., Yu, H., Tade, M., 2016. Systematic study of aqueous monoethanolamine (MEA)-based CO_2 capture process: Techno-economic assessment of the MEA process and its improvements. Appl. Energy 165, 648–659.

Li, T., Yang, C., Tantikhajorngosol, P., Sema, T., Liang, Z., Tontiwachwuthikul, P., Liu, H., 2022. Comparative desorption energy consumption of post-combustion CO_2 capture integrated with mechanical vapor recompression technology. Sep. Purif. Technol. 294.

Liang, Z., Gao, H., Rongwong, W., Na, Y., 2015. Comparative studies of stripper overhead vapor integration-based configurations for post-combustion CO_2 capture. J. Greenh. Gas. Control 34, 75–84.

Lin, Y.-J., Rochelle, G.T., 2016. Approaching a reversible stripping process for CO_2 capture. Chem. Eng. J. 283, 1033–1043.

Lu, J., Yao, K., Gao, F., 2009. Process similarity and developing new process models through migration. AIChE J. 55, 2318–2328.

McBride, K., Sundmacher, K., 2019. Overview of surrogate modeling in chemical process engineering. Chem. Ing. Tech. 91, 228–239.

N.Borhani, T., Wang, M., 2019. Role of solvents in CO_2 capture processes: The review of selection and design methods. Renew. Sustain. Energy Rev. 114.

Nuchitprasittichai, A., Cremaschi, S., 2013. An algorithm to determine sample sizes for optimization with artificial neural networks. AIChE J. 59, 805–812.

Ochoa-Estopier, L.M., Jobson, M., Smith, R., 2013. Operational optimization of crude oil distillation systems using artificial neural networks. Comput. Chem. Eng. 59, 178–185.

Ochoa-Estopier, L.M., Jobson, M., Smith, R., 2014. The use of reduced models for design and optimisation of heat-integrated crude oil distillation systems. Energy 75, 5–13.

Oh, S.-Y., Binns, M., Cho, H., Kim, J.-K., 2016. Energy minimization of MEA-based CO_2 capture process. Appl. Energy 169, 353–362.

Osuolale, F.N., Zhang, J., 2016. Energy efficiency optimisation for distillation column using artificial neural network models. Energy 106, 562–578.

Pan, S.J., Yang, Q., 2010. A survey on transfer learning. IEEE Trans. Knowl. Data Eng. 22, 1345–1359.

Parisi, G.I., Kemker, R., Part, J.L., Kanan, C., Wermter, S., 2019. Continual lifelong learning with neural networks: A review. Neural Netw. 113, 54–71.

Rusu, A.A., Rabinowitz, N.C., Desjardins, G., Soyer, H., Kirkpatrick, J., Kavukcuoglu, K., Pascanu, R., Hadsell, R., 2016, Progressive neural networks. arXiv preprint.

Sipöcz, N., Tobiesen, F.A., Assadi, M., 2011. The use of artificial neural network models for CO_2 capture plants. Appl. Energy 88, 2368–2376.

Song, W., Mahalec, V., Long, J., Yang, M., Qian, F., 2020. Modeling the hydrocracking process with deep neural networks. Ind. Eng. Chem. Res. 59, 3077–3090.
Terekhov, A.V., Montone, G., O'Regan, J.K., 2015. Knowledge transfer in deep block-modular neural networks. Conference on Biomimetic and Biohybrid Systems. Springer,, pp. 268–279.
Tikadar, D., Gujarathi, A.M., Guria, C., 2020. Multi-objective optimization of industrial gas-sweetening operations using economic and environmental criteria. Process Saf. Environ. Prot. 140, 283–298.
Wang, M., Joel, A.S., Ramshaw, C., Eimer, D., Musa, N.M., 2015. Process intensification for post-combustion CO_2 capture with chemical absorption: A critical review. Appl. Energy 158, 275–291.
Xu, D., Shi, Y., Tsang, I.W., Ong, Y.S., Gong, C., Shen, X., 2020. Survey on multi-output learning. IEEE Trans. Neural Netw. Learn. Syst. 31, 2409–2429.
Yin, L., Gao, Q., Zhao, L., Wang, T., 2020. Expandable deep learning for real-time economic generation dispatch and control of three-state energies based future smart grids. Energy 191.
Yoon, J., Yang, E., Lee, J., Hwang, S.J., 2018, Lifelong learning with dynamically expandable networks. arXiv preprint.
Yosinski, J., Clune, J., Bengio, Y., Lipson, H., 2014. How transferable are features in deep neural networks. Adv. Neural Inf. Process. Syst. 27.
Zhang, Y., Yang, Q., 2022. A survey on multi-task learning. IEEE Trans. Knowl. Data Eng. 34, 5586–5609.
Zhu, W., Ye, H., Yang, Y., Zou, X., Dong, H., 2022. Simulation-based optimization of a multiple gas feed sweetening process. ACS Omega 7, 2690–2705.