A Combined Neural Network Forecasting Approach for CO_2-Enhanced Shale Gas Recovery

Zhenqian Xue[1†], Yuming Zhang[1†], Haoming Ma[1,2*], Yang Lu[3], Kai Zhang[4], Yizheng Wei[5], Sheng Yang[1], Muming Wang[1], Maojie Chai[1], Zhe Sun[1], Peng Deng[1], and Zhangxin Chen[1,6*]

[1]Department of Chemical and Petroleum Engineering, University of Calgary
[2]Hildebrand Department of Petroleum and Geosystems Engineering, University of Texas at Austin
[3]School of Computing and Communications, Lancaster University
[4]School of Earth Resources, China University of Geosciences (Wuhan)
[5]Computer Modelling Group
[6]Eastern Institute of Technology

Summary

Intensive growth of geological carbon sequestration has motivated the energy sector to diversify its storage portfolios, given the background of climate change mitigation. As an abundant unconventional reserve, shale gas reservoirs play a critical role in providing sufficient energy supply and geological carbon storage potentials. However, the low recovery factors of the primary recovery stage are a major concern during reservoir operations. Although injecting CO_2 can resolve the dual challenges of improving the recovery factors and storing CO_2 permanently, forecasting the reservoir performance heavily relies on reservoir simulation, which is a time-consuming process. In recent years, pioneered studies demonstrated that using machine learning (ML) algorithms can make predictions in an accurate and timely manner but fails to capture the time-series and spatial features of operational realities. In this work, we carried out a novel combinational framework including the artificial neural network (ANN, i.e., multilayer perceptron or MLP) and long short-term memory (LSTM) or bi-directional LSTM (Bi-LSTM) algorithms, tackling the challenges mentioned before. In addition, the deployment of ML algorithms in the petroleum industry is insufficient because of the field data shortage. Here, we also demonstrated an approach for synthesizing field-specific data sets using a numerical method. The findings of this work can be articulated from three perspectives. First, the cumulative gas recovery factor can be improved by 6% according to the base reservoir model with input features of the Barnett shale, whereas the CO_2 retention factor sharply declined to 40% after the CO_2 breakthrough. Second, using combined ANN and LSTM (ANN-LSTM)/Bi-LSTM is a feasible alternative to reservoir simulation that can be around 120 times faster than the numerical approach. By comparing an evaluation matrix of algorithms, we observed that trade-offs exist between computational time and accuracy in selecting different algorithms. This work provides fundamental support to the shale gas industry in developing comparable ML-based tools to replace traditional numerical simulation in a timely manner.

Introduction

In light of the low-carbon energy transition, natural gas, being the most environmentally friendly fossil fuel, has garnered wide attention in future energy supply. In recent years, with the depletion of conventional gas reserves and the maturity of hydraulic fracturing technology, the energy sector has undergone a notable shift to unconventional resources (e.g., shale gas) (Song et al. 2019; Belyadi et al. 2019; Sambo et al. 2023). Previous studies also suggested that shale reservoirs might serve as a promising candidate for long-term CO_2 storage, aiding in a reduction in greenhouse gas emissions (Friedlingstein et al. 2022; Yang et al. 2024). CO_2-enhanced shale gas recovery (CO_2-ESGR) is a viable solution to improve a shale gas recovery factor as well as mitigate climate change impacts as demonstrated in the previous studies (Liu et al. 2019; Li and Elsworth 2019; Ma et al. 2022; Deng et al. 2023). Because the industrialization of CO_2-ESGR still remains at an infant stage, limited demonstration or pilot projects have been recorded to date (Zhang et al. 2023).

There are two major challenges during the primary recovery of shale gas reservoirs. As such, less than 30% of the original gas in place (OGIP) can be recovered. First, fracture closures in response to decreasing pore pressure reduce an available pore space for gas to flow through, limiting the overall productivity (Cai and Dahi Taleghani 2019; Dahi Taleghani et al. 2020). Second, extracting absorbed phase methane from the shale matrix requires a careful pressure management strategy, especially in reservoir conditions (Iddphonce and Wang 2021; Yang et al. 2022; Liao et al. 2023). CO_2-ESGR has the capability to desorb methane by means of the competitive adsorption mechanism of CO_2 over CH_4 (Yang and Liu 2020). As such, its recovery factor can be improved significantly by up to approximately 20% (Li and Elsworth 2019; Ma et al. 2022). Furthermore, the CO_2 retained in shale gas reservoirs according to the structural and adsorption trapping mechanisms unlocks a great potential to sequester greenhouse gases at a large scale (Yang et al. 2019). Although CO_2-ESGR has been studied experimentally and numerically at multiple scales to understand its driven mechanisms and operational strategies, an effective prediction of reservoir performance from both CH_4 recovery and CO_2 sequestration perspectives can provide valuable insights for operators at an early stage. Conventional analytical methods (e.g., decline curve analysis) and numerical methods have capabilities in reservoir performance forecasting (Yang et al. 2018; Ma et al. 2019); however, owing to the complexity of geological and operational conditions and computational limitations, these approaches always require considerable additional assumptions to forecast the CO_2-ESGR performance.

*Corresponding author; email: haoming.ma@ucalgary.ca; zhachen@ucalgary.ca

†These authors contributed equally to this article.

Original SPE manuscript received for review 1 December 2023. Revised manuscript received for review 10 April 2024. Paper (SPE 219774) peer approved 29 April 2024.

In recent years, the deployment of artificial intelligence (AI) in the energy industry offered great potential in terms of reservoir characterization and production forecasting (Abdelfattah et al. 2021; Liu et al. 2022; Olukoga and Feng 2022; Rehman and Lal 2023; Wang et al. 2023a). In detail, Du et al. (2023) developed a Bi-LSTM data-driven model to forecast coalbed methane production curves under different well patterns. Huang and Chen (2021) demonstrated that the MLP, a subcategory of ANNs, can be deployed to estimate the field performance of steam-assisted gravity drainage by comparing it with other supervised learning algorithms. Xue et al. proposed an MLP and differential evolution–based optimization framework for enhanced geothermal systems, then presented a time-series prediction of geothermal energy production by using LSTM, and furtuer investigated the economic feasibilities of large-scale geothermal reservoirs (Xue et al. 2023b, 2023c, 2024a, 2024b; Xue and Chen 2023a). Wang et al. (2023b) proposed an application of a sparrow search algorithm to forecast shale gas productivity. Iskandar and Kurihara (2022) proposed an LSTM-based data-driven forecast approach by comparing it with autoregressive and MLP algorithms. In summary, research progress confirmed that developing AI-based predictive tools can be an alternative to conventional numerical approaches. However, selecting the most appropriate AI algorithms and training data-driven models for a given problem are necessities in that regard.

Previous studies have demonstrated well that predicting the reservoir performance using LSTM is highly dependent on historical production data, while using ANN can only address the variabilities of geological and operational properties. As such, Shi et al. (2021) proposed a notable investigation that combined the LSTM with the MLP combinational neural network to predict the productivity of multilateral well geothermal systems, which illustrated a possibility to simultaneously incorporate time-series-based operational inputs, but the input determination strategies and the model applicability are not well documented. Therefore, the knowledge gap in using a combinational framework for CO_2-ESGR remains unclear according to the existing literature. To the best of our knowledge, there is no comprehensive feasibility understanding to benchmark the CO_2-ESGR performance by combining the MLP and LSTM algorithms such that the impacts of multiple properties can be addressed within each timestep.

In this paper, this knowledge gap has been filled in two ways. First, we carried out a CO_2-ESGR numerical model and validated it with our previous study to generate a large data set for training purposes according to different sensitive input features. Second, we proposed a combinational framework by coupling MLP with LSTM/Bi-LSTM to predict the CO_2-ESGR performance from both hydrocarbon recovery and CO_2 storage perspectives. The proposed model has been compared with the optimized MLP- and LSTM-based forecasting models individually to demonstrate its accuracy improvement. As such, the limitations of using either MLP or LSTM have been resolved by combining data complexity from both spatial and time horizons. This study provides novel insights to predict the performance of CO_2-ESGR for operators by proposing a combinational AI-based framework, which is applicable to multidimensional data sets efficiently and chronologically, fully comparable with the conventional numerical methods.

Methodology

Data Preparation. A dual-porosity, dual-permeability heterogeneous reservoir model using the Computer Modelling Group (CMG)-GEM software is developed and validated according to our previous study for the basic CO_2-ESGR case across 30 years with the reservoir properties from Barnett shale and constant pressure boundaries for both injector and producer (Ma et al. 2022). A porosity distribution for each gridblock is simulated based on the Monte Carlo approach, and the corresponding permeability is calculated based on the Kozeny-Carman equation (Costa 2006; Ma et al. 2022). The impacts of CO_2/CH_4 competitive adsorption and fracture closure have been fully addressed in the numerical model (Ma et al. 2022). To assess the performance of CO_2-ESGR from both gas recovery and CO_2 storage perspectives, two objectives are adopted in this work, the cumulative incremental recovery factor (Eq. 1) and the CO_2 retention factor (Eq. 2), where $\mathrm{RF}_{\mathrm{CO_2-ESGR}}$ represents the cumulative recovery factor and P stands for the daily production rate after injecting CO_2, as we observed that there is negligible gas produced immediately before the CO_2 injection. The $R_{\mathrm{ret,CO_2}}$ defines the retention factor of CO_2, which is the ratio of total CO_2 retained to the reservoir over the total injection in either mass or volumetric units:

$$\mathrm{RF}_{\mathrm{CO_2-ESGR}} = \frac{\sum P}{\mathrm{OGIP}}, \tag{1}$$

$$R_{\mathrm{ret,CO_2}} = \frac{M_{\mathrm{CO_2,Injected}} - M_{\mathrm{CO_2,produced}}}{M_{\mathrm{CO_2,Injected}}}. \tag{2}$$

Based on the previous studies related to the CO_2-ESGR, 13 parameters are determined as the sensitive inputs for the data-driven models because these variables illustrate different influences on gas recovery and carbon sequestration performance (Ma et al. 2023; Wang et al. 2018, 2023b; Golipour et al. 2019; Hui et al. 2023). The details of these properties in the base model are shown in **Table 1.** For a better understanding of gas production and carbon sequestration performance, three objectives, including the CO_2 mass production rate, CH_4 mass production rate, and CO_2 storage rate, are determined as the outputs for the data-driven models. Based on the determination of inputs and targets, the CMG-CMOST software is used to generate the data samples by randomly selecting values of inputs from predetermined ranges and calculating corresponding targets, and a total of 5,000 simulations were conducted using Monte Carlo simulation. Due to the simulation errors, the results of 3,963 simulations are collected as the data samples for data-driven models. The overall workflow of this study is shown in **Fig. 1.** To predict the targets at different timesteps, except for these determined 13 sensitive parameters, time is also treated as the input feature for the ANN model. In total, 14 inputs and three outputs are determined for the ANN model, which is shown in **Table 2.** Based on the subsequent data splitting, model training, and model testing, the predictions of the optimal ANN model are then treated as the input for the LSTM and Bi-LSTM models. Similarly, with the training and testing of the LSTM and Bi-LSTM models, the determined ANN-LSTM model and combined ANN and Bi-LSTM (ANN–Bi-LSTM) model are constructed, respectively, for the time-series predictions of gas production and carbon sequestration, and their model performance is then evaluated by an additional created scenario.

Combined Neural Network Model. In conventional applications of ML in the oil and gas industry, the ANN (also known as MLP) has shown the capability of regressing and clustering complex nonlinear patterns between inputs and outputs by involving multiple hidden layers, where each layer may contain a number of neurons that assign weights of inputs from the previous layer and pass them to the next layer. LSTM and Bi-LSTM are two deep-learning algorithms that have been applied in predicting sequential production data in previous studies. LSTM is a recurrent neural network architecture developed to solve supervised learning tasks. With the composition of a cell, an input gate, an output gate, and a forget gate in each unit, LSTM is designed to map the sequential inputs and outputs by considering both short and long memories. Unlike LSTM, Bi-LSTM is an extension of the traditional LSTM, which combines the traditional LSTM with bidirectional processing, specifically, the forward and backward propagations. Compared with LSTM, Bi-LSTM unlocks the predictive

Parameters	Value	Unit
Geological Properties		
Pressure gradient	12.2	kPa/m
Temperature gradient	0.023	C/m
Matrix average porosity	0.041	%
Natural fracture average porosity	0.0004	%
Engineering Parameters		
Injector bottomhole pressure	35	MPa
Producer bottomhole pressure	15	MPa
Well spacing	300	m
Hydraulic fracture half-length	150	m
Hydraulic fracture conductivity	3	md·m
Competitive Adsorption		
CO_2 Langmuir pressure constant	8.65	MPa
CO_2 Langmuir volume constant	5.2	m^3/ton
CH_4 Langmuir pressure constant	11.2	MPa
CH_4 Langmuir volume constant	1.2	m^3/ton

Table 1—Sensitive input parameters of the reservoir model and features of adsorption.

capability of deploying the past and future context from both directions, which allows a better predictive capability in general. The structures of common LSTM and Bi-LSTM are shown in **Figs. 2 and 3,** respectively. However, the operational constraints (e.g., a Langmuir adsorption constant multiplier, pressure gradient, temperature gradient, etc.) that demonstrate significant effects on CO_2-ESGR performance are not combined in the current sequential CO_2-ESGR performance prediction. On the other hand, by using conventional deep-learning algorithms (i.e., ANN), the historical production data are the only input feature and only the effect of a time constraint can be considered, while the nonlinear relationships between different working parameters and the production of CO_2-ESGR are ignored. Competing with the traditional reservoir modeling approach requires the reservoir performance to be forecasted by capturing both spatial and temporal features. Furthermore, the existing studies are limited to site-specific cases so that consistent comparisons cannot be conducted across all forecasting approaches. Developing generalized reservoir performance predictive approaches using neural network models necessitated to interpret panel data sets reasonably. Consequently, the predicted results by using either the ANN or the recurrent neural network may not lead to a promising accuracy. To address this issue, an ANN-LSTM and an ANN–Bi-LSTM are proposed in this study to estimate CH_4 and CO_2 production and CO_2 storage capacity.

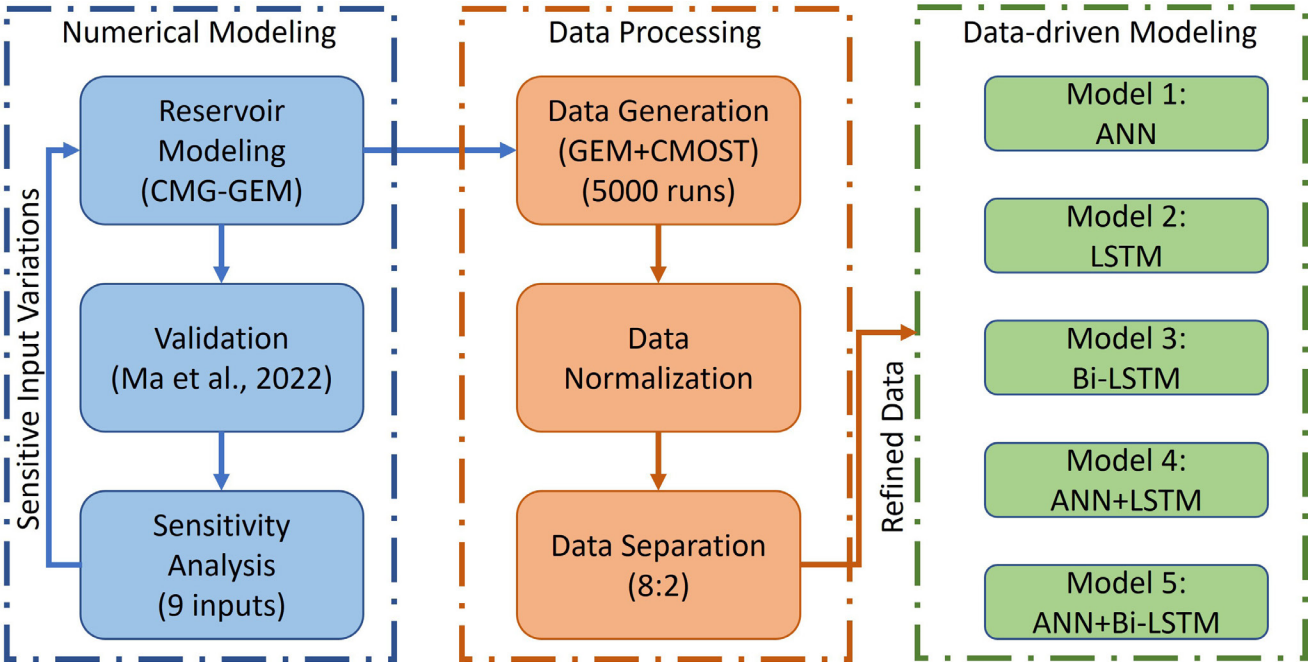

Fig. 1—Descriptive overview of this study.

Input Features				Targets	
No.1	Time	No.8	Producer bottomhole pressure	No.1	CO_2 production
No.2	Pressure gradient	No.9	Hydraulic fracture half-length	No.2	CH_4 production
No.3	Temperature gradient	No.10	Hydraulic fracture conductivity	No.3	CO_2 storage
No.4	Matrix porosity	No.11	CO_2 Langmuir pressure constant		
No.5	Natural fracture porosity	No.12	CO_2 Langmuir volume constant		
No.6	Well spacing	No.13	CH_4 Langmuir pressure constant		
No.7	Injector bottomhole pressure	No.14	CH_4 Langmuir volume constant		

Table 2—Input features and targets for ANN models.

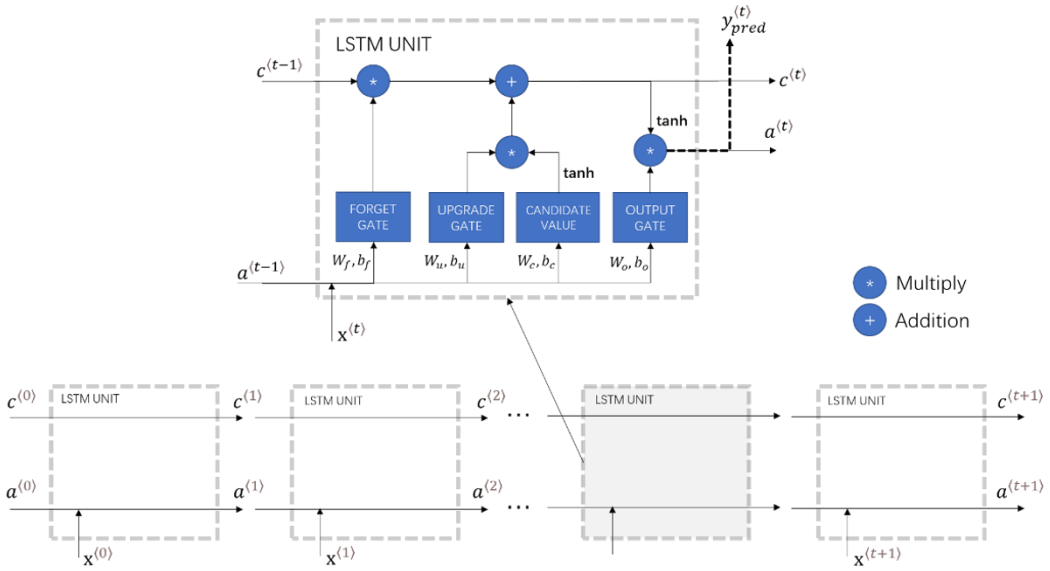

Fig. 2—Schematic illustration of LSTM.

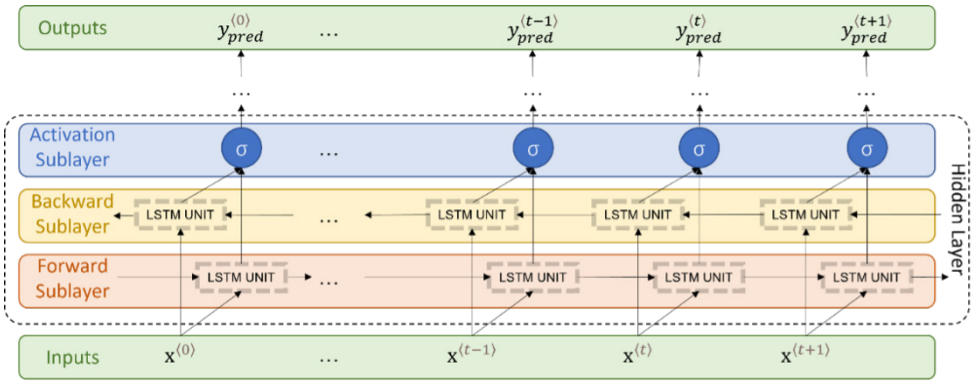

Fig. 3—Schematic illustration of Bi-LSTM.

Fig. 4 shows the schematic of combined neural networks. As illustrated, initially, ANN/MLP is utilized to generalize the complex correlations between various working properties and production in each timestep and produce the time-series performance results $y_1^{'}, y_2^{'}, \ldots, y_i^{'}$ consequently. Following that, LSTM and LSTM/Bi-LSTM are developed to memorize the time-series data with long-term dependencies. Specifically, combined with the historical production data $x^0, x^1, \ldots, x^t$, the MLP results are transferred into LSTM and Bi-LSTM separately as an input matrix. Finally, after the optimization of combined neural networks, the production data are predicted sequentially starting from the timestep of $t + 1$. The combined framework resolved the limitations of conventional data-driven models of using either MLP or LSTM by performing on a panel data set. However, there are also two limitations of the proposed framework. First, a combined model requires a longer computational time compared with an MLP- or LSTM-based model individually, but still faster than the numerical simulation. Second, a reasonable assumption of reservoir performance outputs for the initial timestep will be required to activate the LSTM/Bi-LSTM algorithms.

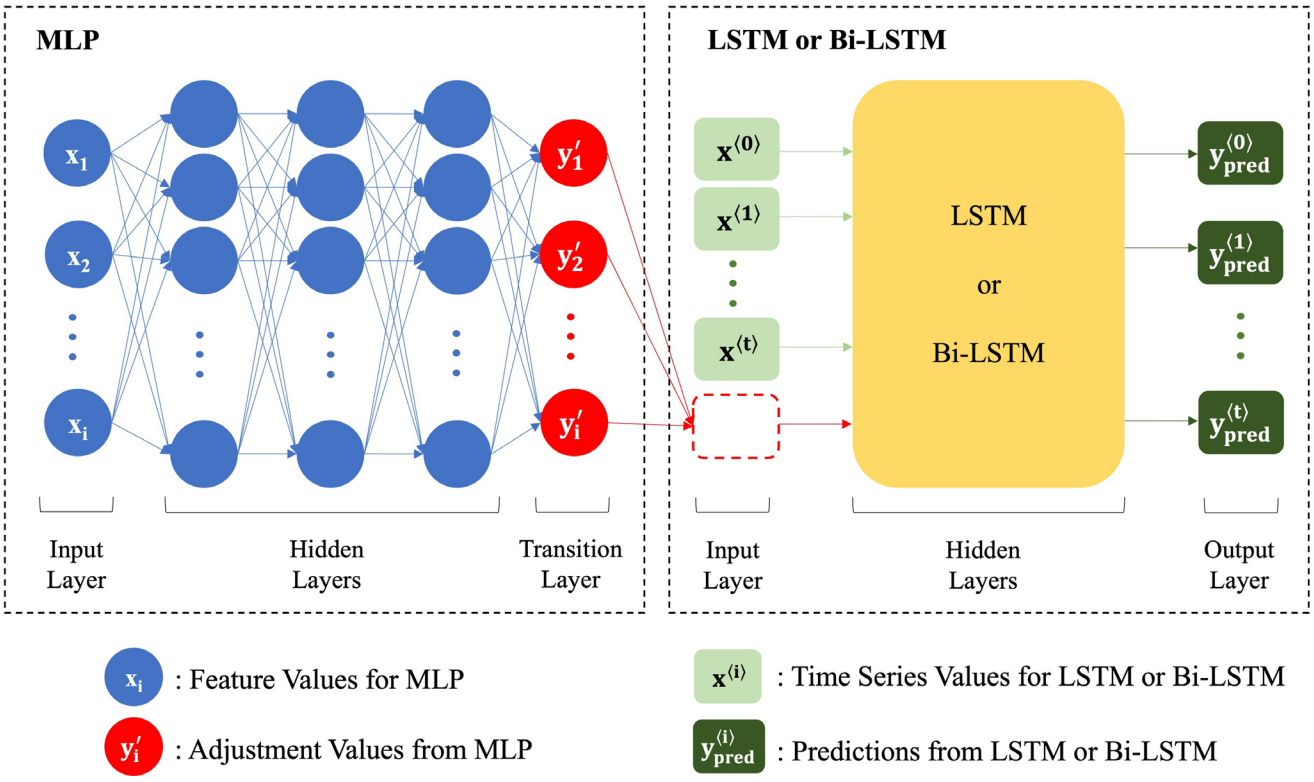

Fig. 4—Schematic illustration of combined neural networks.

Based on the structure of combined neural network models, the overall construction process of these models can be concluded in the following steps: First, the collected synthetic data set from the previous section is randomly split into a training set and a testing set with a ratio of 8:2 after the normalization and then used for the training and testing of an ANN model. The optimal ANN model is determined after the hyperparameter tuning during this process. Second, the predictions obtained from the resulting ANN model are treated as the inputs for the LSTM and Bi-LSTM models, where 80% of the data from the initial timestep are utilized for model training and the remaining 20% of the timestep data are used for model testing. Similarly, the optimal LSTM and Bi-LSTM models result from the optimization of the hyperparameters, and thus the optimal ANN-LSTM and ANN–Bi-LSTM models are developed successfully. Finally, to evaluate the performance of the determined combined neural network models, a new scenario is constructed with different values of sensitive parameters from the data samples by using the CMG-GEM software, and then the results of the ANN-LSTM and ANN–Bi-LSTM models are compared with those of this new scenario.

Hyperparameters Optimization. The optimization of the data-driven models is the process of hyperparameter tuning, which can substantially impact the accuracy of predictions. Common methods, such as grid search and random search, are usually time-consuming and result in low computational efficiency. In this work, we adopted the tree-structured Parzen estimator (TPE) algorithm, which is a type of the Bayesian optimization approach designed to optimize the sequential data-driven models in a faster and more time-efficient way. A TPE algorithm applies a Gaussian mixture model to find the hyperparameter values that can minimize the objective function. The conditional probability based on the Bayes theory is first defined in Eq. 3 (Watanabe 2023):

$$p(x|y) = \begin{cases} \ell(x) & \text{if } y < y^* \\ g(x) & \text{if } y \geq y^*, \end{cases} \tag{3}$$

where x is a hyperparameter that needs to be optimized; y and y^* represent an objective function and a loss threshold separately; $\ell(x)$ is the density formed by using the observations such that the corresponding loss was less than the threshold; and $g(x)$ is the density formed by using the remaining observations. The TPE algorithm is operated based on the optimization of the expected improvement (EI), as shown in Eq. 4 (Ozaki et al. 2022; Watanabe 2023). By construction, $\gamma = p(y < y^*)$, $p(x) = \int p(x|y)\, p(y)\, dy = \gamma \ell(x) + (1-\gamma)\, g(x)$, and $\int_{-\infty}^{y^*} (y^* - y)\, p(x|y)\, p(y)\, dy = \gamma y^* \ell(x) - \ell(x) \int_{-\infty}^{y^*} p(y)\, dy$, the final $\mathrm{EI}_{y^*}(x)$ can be expressed by Eq. 5. In an optimization process, the hyperparameter configuration is determined iteratively with the purpose of reaching the greatest EI within the desired number of iterations. Specifically, a new hyperparameter configuration (x^*) is obtained by minimizing $g(x)/\ell(x)$, and the resulting x^* is returned into the function for the iterative fitting and minimization of $g(x)/\ell(x)$ until the desired iteration number is reached:

$$\mathrm{EI}_{y^*}(x) = \int_{-\infty}^{y^*} (y^* - y)\, p(y|x)\, dy = \int_{-\infty}^{y^*} (y^* - y) \frac{p(x|y)\, p(y)}{p(x)} dy, \tag{4}$$

$$\mathrm{EI}_{y^*}(x) = \frac{\gamma y^* \ell(x) - \ell(x) \int_{-\infty}^{y^*} p(y)\, dy}{\gamma \ell(x) + (1-\gamma)\, g(x)} \propto \left(\gamma + \frac{g(x)}{\ell(x)} (1-\gamma) \right)^{-1}. \tag{5}$$

Evaluation Matrix. Four evaluation metrics were selected in this study to measure the accuracy of the data-driven models compared with the original outputs from the simulations. Eqs. 6 through 9 define the mean square error (MSE), mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R^2) (Iskandar and Kurihara 2022):

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - y_i^{\mathrm{pred}}\right)^2, \tag{6}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - y_i^{\mathrm{pred}}\right|, \tag{7}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - y_i^{\mathrm{pred}}\right)^2}, \tag{8}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - y_i^{\mathrm{pred}}\right)^2}{\sum_{i=1}^{n}\left(y_i - y_i^{\mathrm{avg}}\right)^2}, \tag{9}$$

where y_i represents the original values, y_i^{pred} is the predictive values, y_i^{avg} denotes the arithmetic average of true values, and n denotes the sample size. Lower values of MSE, MAE, and RMSE indicate more precious and accurate predictions. R^2 is a coefficient that ranges from 0 to 1, with values approaching 1 indicating a better fitting performance.

Results and Discussion

Data Collection and Preprocessing. In this section, the fundamental numerical modeling of CO_2-ESGR carried out by CMG-GEM has been outlined. As mentioned, the injector and the producer are designed with constant pressure boundaries. **Fig. 5a** shows that gas is produced at its daily maximum rate for 22 years and then starts declining afterward until Year 35. Thus, the slope of the cumulative production factor decreases correspondingly in **Fig. 5b.** A total recovery of 19.5% of OGIP is observed during the primary recovery stage. From Year 35, CO_2 is being injected into the reservoir for ESGR purposes and the daily gas production rate has been improved for another 7 years to its maximum at 20 tons/D until the CO_2 breakthrough, which result in the production rate declining again. The CO_2 is being produced after the breakthrough, which leads to an increasing trend of CO_2 production rate in **Fig. 5a**, and a sharp decline of CO_2 retention factor in **Fig. 5b**, whereas the retention factor remains at 100% before the CO_2 breakthrough. As mentioned, the retention factor represents the effectiveness of CO_2 storage in shale reservoirs. As such, before the breakthrough, all injected CO_2 is stored successfully in the reservoir and the retention factor drops sharply to about 25% after the CO_2 breakthrough. As a result, the CO_2-ESGR process recovered 9% additional OGIP and stored a total of 60% injected CO_2 during the 50 years of operation including the primary recovery.

From the results of the base reservoir model, it can be concluded that the gas production and carbon sequestration performance of the CO_2-ESGR during the operation present nonlinear relationships within a time change due to the complex correlations between geological and engineering parameters and the production. Therefore, conventional LSTM or Bi-LSTM algorithms that can only study the changes of the target in a time horizon are not abundant enough to predict time-series CO_2-ESGR performance; on the contrary, it is necessary to incorporate the comprehension of the relationships between the sensitive inputs and outputs. Therefore, the ANN-LSTM and ANN–Bi-LSTM models are proposed in this study.

Performance of Combinational Neural Networks. Table 3 demonstrates the optimal hyperparameter configurations of four developed models determined after the TPE-based optimizations separately, including the LSTM, Bi-LSTM, ANN-LSTM, and ANN–Bi-LSTM models. The LSTM and Bi-LSTM models are constructed to compare the performance between conventional deep-learning algorithms and the proposed combined neural networks. In this section, the prediction performance of ANN-LSTM and ANN–Bi-LSTM on three different output variables are first studied. The detailed comparison among these four models is illustrated in the next section.

An additional case with a different sensitive parameter configuration from generated data samples is constructed through CMG-GEM to evaluate the resulting ANN-LSTM and ANN–Bi-LSTM models. In addition, the use of historical data is considered another constraint to investigate its influence on the performance of the combined neural networks. Specifically, the scenarios with history data indicate that data have been preprocessed by mixing the ANN predictions and the results of the primary recovery stage from the CMG-GEM. For those without history data, the ANN predictions are used as a sole input feature.

Table 4 shows the results of evaluation indicators of ANN-LSTM and ANN–Bi-LSTM with and without consideration of history data. **Figs. 6 and 7** demonstrate the comparison between the prediction results of these two models and the true results of the created additional case separately. From the results, all the ANN-LSTM and ANN–Bi-LSTM models illustrate a promising time-series prediction accuracy based on their lower than 0.04 MAE values in all three different production targets. More specifically, ANN–Bi-LSTM presents better estimation results than the ANN-LSTM model according to its lower MSE, MAE, and RMSE values **(Table 4).** Besides, their predicted production profiles are more consistent with the true production curves **(Figs. 6 and 7).** Compared with ANN-LSTM, the bidirectional structure in ANN–Bi-LSTM allows a model to not only analyze the former production information but also incorporate future data in studying. Therefore, the ANN–Bi-LSTM model can study the data in the whole period and comprehensively understand the changes in each timestep and, consequently, bring more accurate estimations in three production targets. Regarding the influence of history data, two combined neural networks all illustrate better prediction performance when the history data are provided, where lower values of all the evaluation metrics are demonstrated. It can be explained by the true history production data that can effectively eliminate the variance generated results from ANN in the prediction processes of LSTM and Bi-LSTM. Specifically, the results of ANN are treated as the inputs in the following LSTM or Bi-LSTM models when there is no history data, and the prediction variances are retained and even enlarged eventually in the estimation process of sequential production data. Conversely, the provided history data can be considered another input feature, and, therefore, not only the nonlinear relationships between working parameters and production can be studied by the ANN structure but also the changes of targets in time can be better analyzed by the LSTM and Bi-LSTM neural networks. Importantly, the ANN-LSTM and ANN–Bi-LSTM models only take less than 60 seconds to predict around 18 years of three production profiles, which significantly improves the prediction efficiency by 120 times compared with the numerical simulation method (more than 7,200 seconds).

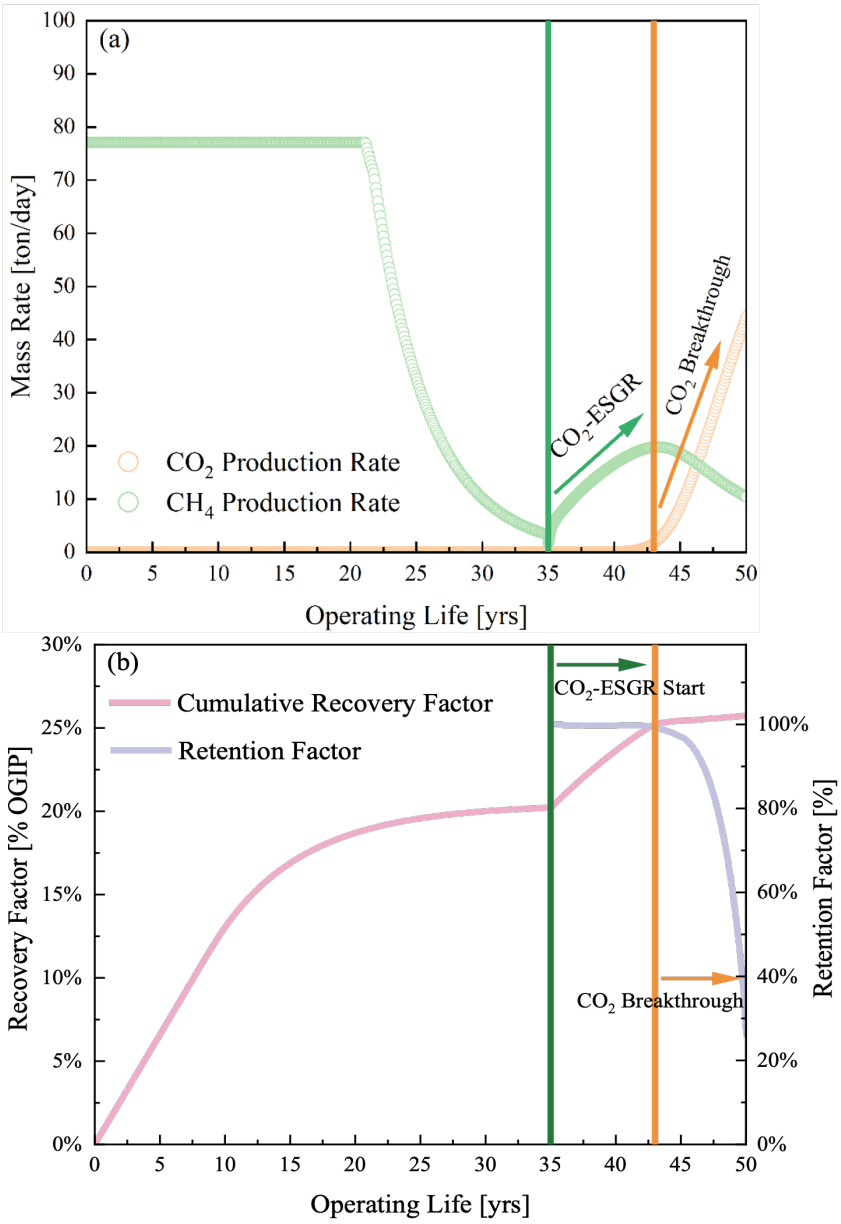

Fig. 5—Results for the base reservoir case from gas production and CO_2 sequestration perspectives.

In conclusion, the proposed combined neural networks can provide promising time-series prediction results with high accuracies and, meanwhile, considerably save the prediction time compared with the conventional numerical simulation method. Regarding different neural network structures, the ANN–Bi-LSTM is highly recommended in estimating the gas production of CO_2-ESGR and CO_2 storage. In addition, although the proposed models present better predictions when historical data are provided, the models without the historical data can also produce the results with a variance at a low level. Importantly, only 1 month of history data is needed in the proposed neural networks to provide for utilization and prediction. Therefore, the proposed combined models demonstrate a wide application range, where the field that just started to develop or even has not been extracted can effectively apply these models to predict future production profiles.

Comparison with Conventional Neural Networks. In this section, the sequential prediction performance of the combined neural networks is compared with that of the conventional neural networks (LSTM and Bi-LSTM). Differently, without the construction of ANN, the history data are necessary to be provided as the only input in the LSTM and Bi-LSTM models. Therefore, the results obtained from an additional created case are transferred into these two models for training and testing, and the underlying relationships between operations and production cannot be studied. The optimal hyperparameter configurations of the LSTM and Bi-LSTM models are presented in **Table 3.** The detailed comparison of indicators between the combined neural networks and conventional deep learning technologies is shown in **Fig. 8.**

From the results, a larger prediction variance can be observed in both LSTM and Bi-LSTM in contrast to ANN-LSTM and ANN–Bi-LSTM, and the LSTM presents the worst performance in predicting all three targets concluded by its highest MSE, MAE, and RMSE values. In addition, it should be noted that the proposed combined neural networks illustrate a higher estimation accuracy than the conventional LSTM and Bi-LSTM even without the history data provided, which means the complex nonlinear correlations between working

Hyperparameter	Value Range	Optimal Values			
		LSTM	Bi-LSTM	ANN-LSTM	ANN-Bi-LSTM
ANN					
No. hidden layers	[1, 5]			2	2
No. units	[1, 1,000]			(45,33)	(45,33)
Activation function	[tanh, ReLU, sigmoid]			ReLU	ReLU
Learning rate	[1×10^{-8}, 0.1]			0.001	0.001
LSTM					
No. hidden layers	[1, 10]	1		1	
No. units	[1, 1,000]	114		116	
Dropout	[0.01, 0.3]	0.18		0.19	
Learning rate	[1×10^{-8}, 0.1]	0.01		0.01	
Activation function	[tanh, ReLU, sigmoid]	tanh		tanh	
Bi-LSTM					
No. hidden layers	[1, 10]		1		1
No. units	[1, 1,000]		68		68
Dropout	[0.01, 0.3]		0.15		0.15
Learning rate	[1×10^{-8}, 0.1]		0.001		0.001
Activation function	[tanh, ReLU, sigmoid]		tanh		tanh

Table 3—Optimized hyper-parameters based on the TPE algorithm.

parameters and production contribute more than the time constraint in predicting these three targets. On the other hand, although LSTM and Bi-LSTM show slightly better estimation efficiency without the construction of the ANN model, the improvement of predictive ability in ANN-LSTM and ANN–Bi-LSTM demonstrates a higher application possibility in the industry. In terms of prediction efficiency, the ANN–Bi-LSTM necessitates a longer computational time than other approaches because of the forward and backward propagation features of the Bi-LSTM. For example, in this research, the ANN–Bi-LSTM takes 55 seconds on average, whereas the ANN-LSTM takes 35 seconds for each run. As references, using LSTM and Bi-LSTM took 27 seconds and 31 seconds, respectively. Even though it might cause overfitting issues when applying ANN–Bi-LSTM as mentioned in previous sections, its high potential to demonstrate a better predictive ability allows it to be the best option in predicting sequential CO_2-ESGR gas production and carbon storage amount. However, the data sets we deployed are processed based on reservoir simulations, which do not include the field uncertainties. Therefore, the ANN–Bi-LSTM is outperforming all scenarios. The feasibility of ANN–Bi-LSTM using the field data is suggested to be further investigated to address the overfitting issues in future work. In conclusion, the ANN–Bi-LSTM illustrated its predictive capabilities in predicting all desired outputs in a time-efficient manner.

		CO_2 Production	CH_4 Production	CO_2 Storage
ANN-LSTM				
With history data	MSE	0.00012	0.000040	0.0025
	MAE	0.0036	0.0029	0.020
	RMSE	0.011	0.0063	0.05
Without history data	MSE	0.00014	0.000043	0.0027
	MAE	0.0039	0.0031	0.023
	RMSE	0.012	0.0066	0.055
ANN–Bi-LSTM				
With history data	MSE	0.00008	0.000033	0.0020
	MAE	0.0030	0.0021	0.014
	RMSE	0.007	0.0058	0.043
Without history data	MSE	0.00011	0.000038	0.0023
	MAE	0.0035	0.0026	0.018
	RMSE	0.009	0.0061	0.046

Table 4—Errors for the optimized model.

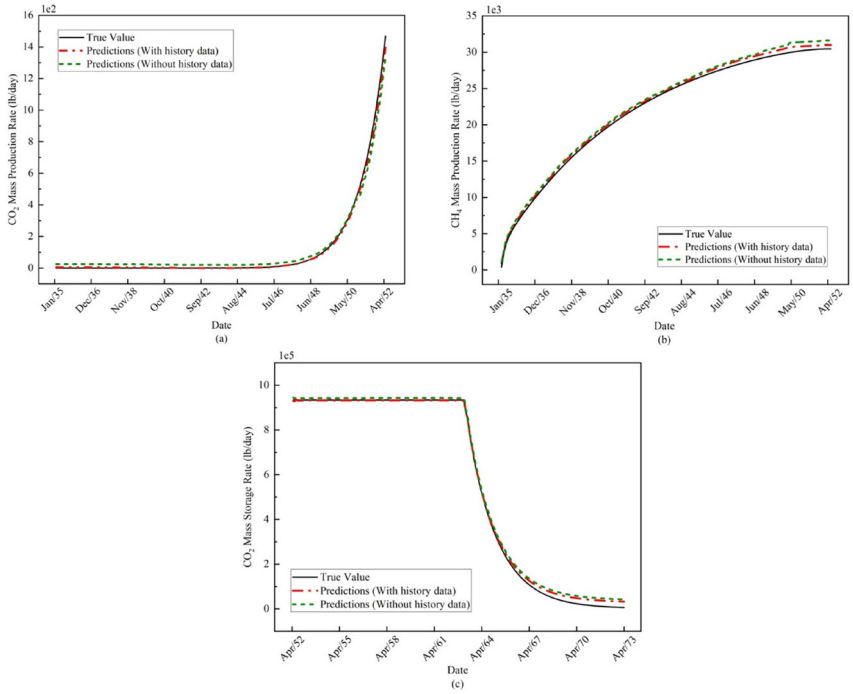

Fig. 6—Predictions of ANN-LSTM with and without history data: (a) CO_2 production rate; (b) CH_4 production rate; and (c) CO_2 storage rate.

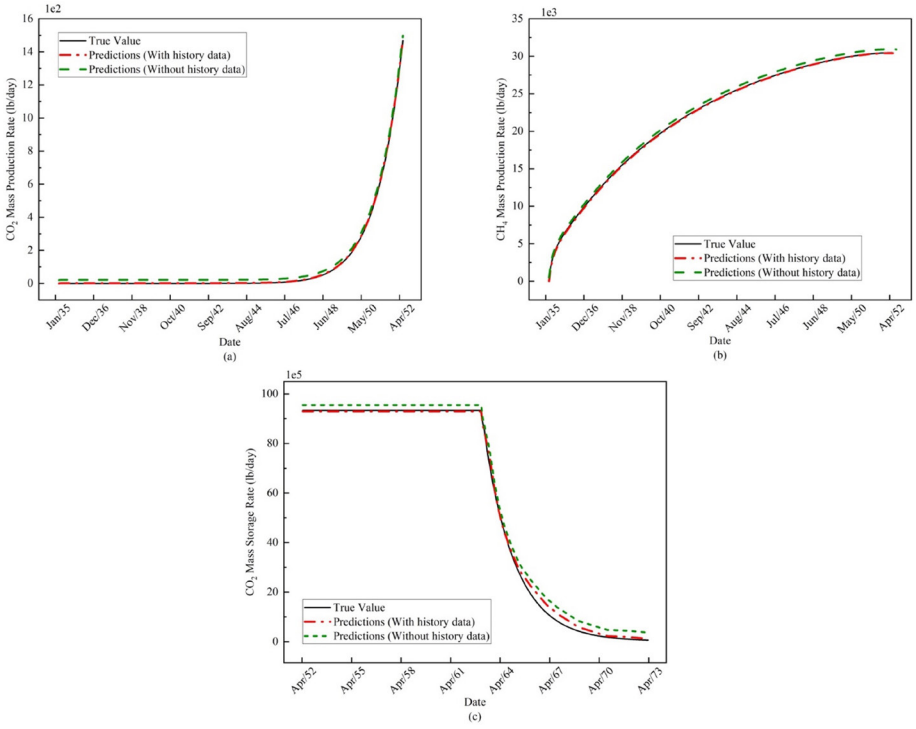

Fig. 7—Predictions of ANN–Bi-LSTM with and without history data: (a) CO_2 production rate; (b) CH_4 production rate; and (c) CO_2 storage rate.

Discussion

ML has been widely applied in the oil and gas industry. In terms of shale gas production prediction, previous studies mainly focused on utilizing ML algorithms to predict the ultimate productivity. For example, Yi et al. (2024) used an extreme gradient boost model to predict the estimated ultimate recovery of unexploited shale gas wells and obtained a mean relative error of 9.8% for their predictions. Although the 9.8% error is acceptable for the productivity prediction, the estimated ultimate recovery is unable to show a potential future production profile. Although there is no thorough study of future sequential production prediction on shale gas reservoirs, deep learning has been utilized successfully in calculating time-series oil production, natural gas production, and geothermal energy production (Du et al. 2022;

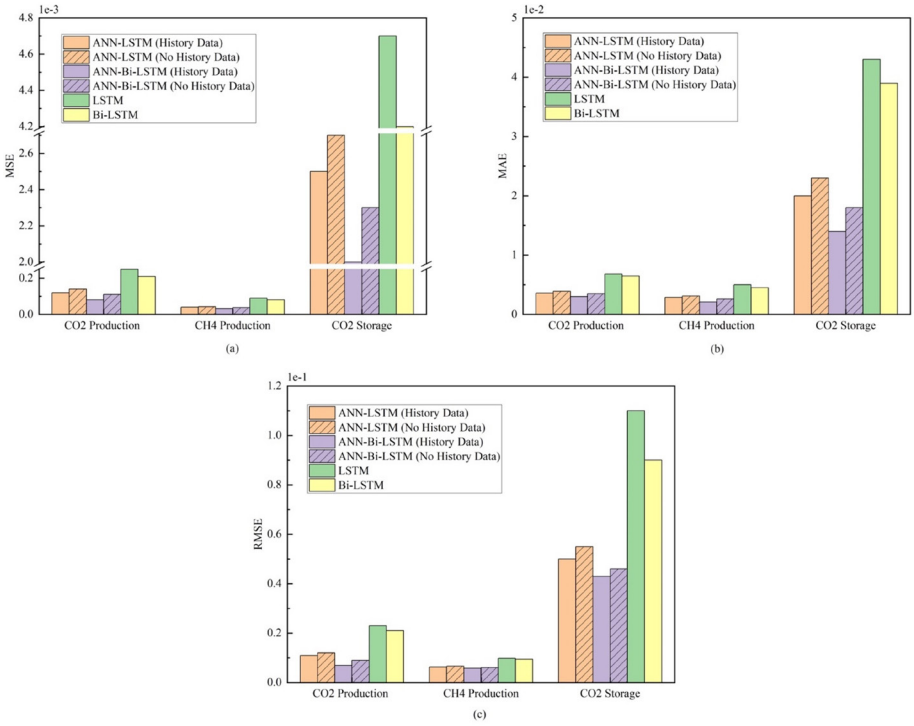

Fig. 8—Indicator values of six neural networks: (a) MSE; (b) MAE; and (c) RMSE.

Xue and Chen 2023a). For instance, Ning et al. (2022) compared the performance of LSTM and autoregressive integrated moving averages in forecasting the oil production of a well in the DJ Basin, and the RMSE values of these two models were all higher than 2.9. However, current related studies basically treated the available production data as the input to estimate future production, which can only study the underlying relationships of the production in a time horizon. This is possibly the main reason why an extremely high accuracy is hard to obtain in existing works. It is known that the productivity of a shale gas reservoir is significantly relevant to geological and engineering properties, and, therefore, it is necessary to incorporate the understanding of the correlations between sensitive parameters and productivity into a sequential shale gas production estimation. In this study, by simultaneously characterizing the complex nonlinear relationships between different sensitive factors and the targets and comprehending the correlations of the production at different time-steps, the proposed combined neural networks provide promising accuracy in predicting gas productions and CO_2 storage amounts with the RMSE values of all lower than 0.05. In addition, CO_2-ESGR is receiving more attention due to its combined benefits of enhancing gas recovery and sequestering CO_2 in a reservoir. Therefore, the performance investigation from both the gas production and carbon storage perspectives is essential but the previous application of ML or deep learning in shale gas reservoirs has not considered these two perspectives at the same time. The ANN–Bi-LSTM model in this study can simultaneously predict a gas production rate and a CO_2 storage rate, which can provide the operators with a comprehensive understanding of extraction performance. Importantly, with the application of the ANN, a combined neural network model is able to accomplish the estimation of future production without providing historical production data. From the results, only a slightly worse performance is demonstrated in the proposed model when the historical data are not provided. This achievement solves the drawback of the traditional deep-learning applications in time-series predictions, which is that the historical data are mandatory to provide. On the other hand, it can effectively widen the application range of AI technologies in shale gas reservoirs, especially in a field that has not been extracted and the production data is not available. This high-accuracy prediction tool can help the operators understand the potential production performance, thereby facilitating the determination of the operational strategy or the optimization of the field. In short, the proposed combined neural network models are a high-accuracy prediction tool for a shale gas reservoir that can comprehend complex underground relationships and estimate future production profiles, which can also be an effective framework to be applied in other industries.

Conclusions

In this research, a novel framework of applying combined ANN and LSTM algorithms is developed to predict the reservoir performance of CO_2-ESGR in a time- and computational-efficient way. The designed framework can be applied to field operations as an alternative forecasting tool for numerical modeling. Although previous studies have investigated the feasibility of applying deep-learning algorithms to this research topic, a major challenge is that the predictions can only be inferred in a time-series or spatial manner, which results in the fact that ML-based approaches are still not comparable with traditional numerical simulations. The designed framework in this research resolved this challenge by incorporating both spatial and time-series features of the synthetic data set established by CMG-GEM. However, the limitation of this framework is that the training data sets heavily rely on the numerical simulation data set, which does not account for the field variabilities during the operations. Another limitation is that the designed framework is only valid for a site-specific case of the Barnett shale reservoir. A generalized framework that can interpret the field variabilities is recommended in future investigations. The main findings of this research can be concluded as follows:

1. A compositional reservoir model of CO_2-ESGR incorporating the competitive adsorption mechanisms of CO_2/CH_4 is carried out. By the end of the project life, the cumulative gas recovery factor has been improved by 6%. The CO_2 cumulative retention factor has declined to 40% after the CO_2 breakthrough, demonstrating that there is an economic trade-off between shale gas recovery and carbon storage if the carbon policy is considered.

2. The feasibility of using combined ANN-LSTM/Bi-LSTM as an alternative for reservoir simulation is demonstrated. The time-series and spatial features can therefore be incorporated into the forecasting model development. Results show that the proposed approach performed comparably with the CMG-GEM in terms of predicting CH_4 recovery rate and CO_2 retention factor, but with less computational time (~120 times faster than CMG-GEM).
3. A scenario-based comparative investigation, including six scenarios of different algorithms by investigating MSE, MAE, and RMSE, is carried out. Results depicted that ANN–Bi-LSTM outperformed other algorithms with a computational time of 55 seconds on average. The trade-off between computational time and model accuracy exists according to our observation. The overfitting issue of using Bi-LSTM may exist but cannot be addressed without the field data interpreting the uncertainties.

Acknowledgments

This research has been made possible by contributions from the Natural Sciences and Engineering Research Council (NSERC)/Energi Simulation Industrial Research Chair in Reservoir Simulation, the Alberta Innovates (iCore) Chair in Reservoir Modeling, and the Energi Simulation/Frank and Sarah Meyer Collaboration Centre.

References

Abdelfattah, T., Nasir, E., Yang, J. et al. 2021. Data Driven Workflow to Optimize Eagle Ford Unconventional Asset Development Plan Based on Multidisciplinary Data. Paper presented at the SPE Annual Technical Conference and Exhibition, Dubai, UAE, 21–23 September. https://doi.org/10.2118/206276-MS.

Belyadi, H., Ebrahim, F., and Fatemeh, B. 2019. *Hydraulic Fracturing in Unconventional Reservoirs: Theories, Operations, and Economic Analysis*. Cambridge, Massachusetts, USA: Gulf Professional Publishing.

Cai, Y. and Dahi Taleghani, A. 2019. Pursuing Improved Flowback Recovery after Hydraulic Fracturing. Paper presented at the SPE Eastern Regional Meeting, Charleston, West Virginia, USA, 15–17 October. https://doi.org/10.2118/196585-MS.

Costa, A. 2006. Permeability-Porosity Relationship: A Reexamination of the Kozeny-Carman Equation Based on A Fractal Pore-space Geometry Assumption. *Geophys Res Lett* **33** (2). https://doi.org/10.1029/2005GL025134.

Dahi Taleghani, A., Cai, Y., and Pouya, A. 2020. Fracture Closure Modes during Flowback from Hydraulic Fractures. *Num Anal Meth Geomechanics* **44** (12): 1695–1704. https://doi.org/10.1002/nag.3086.

Deng, P., Chen, Z., Peng, X. et al. 2023. Optimized Lower Pressure Limit for Condensate Underground Gas Storage Using a Dynamic Pseudo-Component Model. *Energy* **285**: 129505. https://doi.org/10.1016/j.energy.2023.129505.

Du, S., Wang, J., Wang, M. et al. 2023. A Systematic Data-Driven Approach for Production Forecasting of Coalbed Methane Incorporating Deep Learning and Ensemble Learning Adapted to Complex Production Patterns. *Energy* **263**: 126121. https://doi.org/10.1016/j.energy.2022.126121.

Du, J., Zheng, J., Liang, Y. et al. 2022. A Hybrid Deep Learning Framework for Predicting Daily Natural Gas Consumption. *Energy* **257**: 124689. https://doi.org/10.1016/j.energy.2022.124689.

Friedlingstein, P., O'Sullivan, M., Jones, M. W. et al. 2022. Global Carbon Budget 2022. *Earth Syst Sci Data* **14** (11): 4811–4900. https://doi.org/10.5194/essd-14-4811-2022.

Golipour, H., Mokhtarani, B., Mafi, M. et al. 2019. Systematic Measurements of CH4 and CO2 Adsorption Isotherms on Cation-Exchanged Zeolites 13X. *J Chem Eng Data* **64** (10): 4412–4423. https://doi.org/10.1021/acs.jced.9b00473.

Huang, Z. and Chen, Z. 2021. Comparison of Different Machine Learning Algorithms for Predicting the SAGD Production Performance. *J Pet Eng* **202**: 108559. https://doi.org/10.1016/j.petrol.2021.108559.

Hui, G., Chen, Z., Wang, Y. et al. 2023. An Integrated Machine Learning-Based Approach to Identifying Controlling Factors of Unconventional Shale Productivity. *Energy* **266**: 126512. https://doi.org/10.1016/j.energy.2022.126512.

Liao, Q., Zhou, J., Xian, X. et al. 2023. Competition Adsorption of CO2/CH4 in Shale: Implications for CO2 Sequestration with Enhanced Gas Recovery. *Fuel* **339**: 127400. https://doi.org/10.1016/j.fuel.2023.127400.

Li, Z. and Elsworth, D. 2019. Controls of CO2–N2 Gas Flood Ratios on Enhanced Shale Gas Recovery and Ultimate CO2 Sequestration. *J Pet Eng* **179**: 1037–1045. https://doi.org/10.1016/j.petrol.2019.04.098.

Liu, H., Cui, L., Liu, Z. et al. 2022. Using Machine Learning Method to Optimize Well Stimulation Design in Heterogeneous Naturally Fractured Tight Reservoirs. Paper presented at the SPE Canadian Energy Technology Conference, Calgary, Alberta, Canada, 16–17 March. https://doi.org/10.2118/208971-MS.

Liu, J., Xie, L., Elsworth, D. et al. 2019. CO2/CH4 Competitive Adsorption in Shale: Implications for Enhancement in Gas Production and Reduction in Carbon Emissions. *Environ Sci Technol* **53** (15): 9328–9336. https://doi.org/10.1021/acs.est.9b02432.

Ma, H., Yang, Y., Xue, Z. et al. 2023. Comparative Machine Learning Frameworks for Forecasting CO2/CH4 Competitive Adsorption Ratios in Shale. Paper presented at the SPE Western Regional Meeting, Anchorage, Alaska, USA, 22–25 May. https://doi.org/10.2118/212994-MS.

Ma, H., Yang, Y., Zhang, Y. et al. 2022. Optimized Schemes of Enhanced Shale Gas Recovery by CO2-N2 Mixtures Associated with CO2 Sequestration. *Energy Convers Manag* **268**: 116062. https://doi.org/10.1016/j.enconman.2022.116062.

Ma, Y., Yue, C., Li, S. et al. 2019. Study of CH4 and CO2 Competitive Adsorption on Shale in Yibin, Sichuan Province of China. *Carbon Resour Convers* **2** (1): 35–42. https://doi.org/10.1016/j.crcon.2018.11.005.

Iddphonce, R. and Wang, J. 2021. Investigation of CO2 and CH4 Competitive Adsorption during Enhanced Shale Gas Production. *J Pet Eng* **205**: 108802. https://doi.org/10.1016/j.petrol.2021.108802.

Iskandar, U. P. and Kurihara, M. 2022. Time-Series Forecasting of a CO2-EOR and CO2 Storage Project Using a Data-Driven Approach. *Energies* **15** (13): 4768. https://doi.org/10.3390/en15134768.

Ning, Y., Kazemi, H., and Tahmasebi, P. 2022. A Comparative Machine Learning Study for Time Series Oil Production Forecasting: ARIMA, LSTM, and Prophet. *Comput Geosci* **164**: 105126. https://doi.org/10.1016/j.cageo.2022.105126.

Olukoga, T. A. and Feng, Y. 2022. Determination of Miscible CO2 Flooding Analogue Projects with Machine Learning. *J Pet Eng* **209**: 109826. https://doi.org/10.1016/j.petrol.2021.109826.

Ozaki, Y., Tanigaki, Y., Watanabe, S. et al. 2022. Multiobjective Tree-Structured Parzen Estimator. *Jair* **73**: 1209–1250. https://doi.org/10.1613/jair.1.13188.

Rehman, A. N. and Lal, B. 2023. Machine Learning in CO2 Sequestration. In *Machine Learning and Flow Assurance in Oil and Gas Production*, eds. B. Lal, C. B. Bavoh, and J. K. S. Sayani, Vol. 7, 119–140. Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-24231-1.

Sambo, C., Liu, N., Shaibu, R. et al. 2023. A Technical Review of CO2 for Enhanced Oil Recovery in Unconventional Oil Reservoirs. *Geo Sci Eng* **221**: 111185. https://doi.org/10.1016/j.petrol.2022.111185.

Shi, Y., Song, X., and Song, G. 2021. Productivity Prediction of a Multilateral-Well Geothermal System Based on a Long Short-Term Memory and Multi-Layer Perceptron Combinational Neural Network. *Appl Energy* **282**: 116046. https://doi.org/10.1016/j.apenergy.2020.116046.

Song, X., Guo, Y., Zhang, J. et al. 2019. Fracturing with Carbon Dioxide: From Microscopic Mechanism to Reservoir Application. *Joule* **3** (8): 1913–1926. https://doi.org/10.1016/j.joule.2019.05.004.

Wang, M., Hui, G., Pang, Y. et al. 2023a. Optimization of Machine Learning Approaches for Shale Gas Production Forecast. *Geo Sci Eng* **226**: 211719. https://doi.org/10.1016/j.geoen.2023.211719.

Wang, T., Tian, S., Li, G. et al. 2018. Molecular Simulation of CO2/CH4 Competitive Adsorption on Shale Kerogen for CO2 Sequestration and Enhanced Gas Recovery. *J Phys Chem C* **122** (30): 17009–17018. https://doi.org/10.1021/acs.jpcc.8b02061.

Wang, R., Wang, L., Chen, W. et al. 2023b. Surrogate-Assisted Evolutionary Optimization of CO_2 -ESGR and Storage . *Energy Fuels* **37** (19): 14800–14810. https://doi.org/10.1021/acs.energyfuels.3c01682.

Watanabe, S. 2023. Tree-Structured Parzen Estimator: Understanding Its Algorithm Components and Their Roles for Better Empirical Performance. arXiv:2304.11127 (preprint; last revised 26 May 2023). https://doi.org/10.48550/arXiv.2304.11127.

Xue, Z. and Chen, Z. 2023a. Deep Learning Based Production Prediction for an Enhanced Geothermal System (EGS). Paper presented at the SPE Canadian Energy Technology Conference and Exhibition, Calgary, Alberta, Canada, 15–16 March. https://doi.org/10.2118/212754-MS.

Xue, Z., Yao, S., Ma, H. et al. 2023b. Thermo-Economic Optimization of an Enhanced Geothermal System (EGS) Based on Machine Learning and Differential Evolution Algorithms. *Fuel* **340**: 127569. https://doi.org/10.1016/j.fuel.2023.127569.

Xue, Z, Zhang, K., Zhang, C. et al. 2023c. Comparative Data-Driven Enhanced Geothermal Systems Forecasting Models: A Case Study of Qiabuqia Field in China. *Energy* **280**: 128255. https://doi.org/10.1016/j.energy.2023.128255.

Xue, Z., Ma, H., Sun, Z. et al. 2024a. Technical Analysis of a Novel Economically Mixed CO2-Water Enhanced Geothermal System. *J Clean Prod* **448**: 141749. https://doi.org/10.1016/j.jclepro.2024.141749.

Xue, Z., Ma, H., Wei, Y. et al. 2024b. Integrated Technological and Economic Feasibility Comparisons of Enhanced Geothermal Systems Associated with Carbon Storage. *Appl Energy* **359**: 122757. https://doi.org/10.1016/j.apenergy.2024.122757.

Yang, Y. and Liu, S. 2020. Review of Shale Gas Sorption and Its Models. *Energy Fuels* **34** (12): 15502–15524. https://doi.org/10.1021/acs.energyfuels.0c02906.

Yang, Y., Liu, S., and Clarkson, C. 2022. Quantification of Temperature-Dependent Sorption Isotherms in Shale Gas Reservoirs: Experiment and Theory. *SPE J.* **27** (05): 3001–3019. https://doi.org/10.2118/205897-PA.

Yang, Y, Liu, S., and Ma, H. 2024. Impact of Unrecovered Shale Gas Reserve on Methane Emissions from Abandoned Shale Gas Wells. *Sci Total Environ* **913**: 169750. https://doi.org/10.1016/j.scitotenv.2023.169750.

Yang, S., Wu, W., Xu, J. et al. 2018. Modeling of Methane/Shale Excess Adsorption Under Reservoir Conditions. *SPE Res Eval & Eng* **21** (04): 1027–1034. https://doi.org/10.2118/180076-PA.

Yang, S., Wu, K., Xu, J. et al. 2019. Roles of Multicomponent Adsorption and Geomechanics in the Development of an Eagle Ford Shale Condensate Reservoir. *Fuel* **242**: 710–718. https://doi.org/10.1016/j.fuel.2019.01.016.

Yi, J., Qi, Z., Li, X. et al. 2024. Spatial Correlation-Based Machine Learning Framework for Evaluating Shale Gas Production Potential: A Case Study in Southern Sichuan Basin, China. *Appl Energy* **357**: 122483. https://doi.org/10.1016/j.apenergy.2023.122483.

Zhang, K., Lau, H. C., and Chen, Z. 2023. The Contribution of Carbon Capture and Storage to Canada's Net-Zero Plan. *J Clean Prod* **404**: 136901. https://doi.org/10.1016/j.jclepro.2023.136901.