Petroleum 9 (2023) 647–657

Contents lists available at ScienceDirect

Petroleum

journal homepage: www.keaipublishing.com/en/journals/petlm

Forecasting oil production in unconventional reservoirs using long short term memory network coupled support vector regression method: A case study

Shuqin Wen [a], Bing Wei [a, *], Junyu You [b, **], Yujiao He [a], Jun Xin [c], Mikhail A. Varfolomeev [d]

[a] National Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Southwest Petroleum University, Chengdu, 610500, China
[b] School of Petroleum and Natural Gas Engineering, Chongqing University of Science and Technology, Chongqing, 401331, China
[c] Research Institute of Geological Exploration and Development, CNPC Chuanqing Drilling Engineering Limited Company, Chengdu, 610051, China
[d] Department of Petroleum Engineering, Kazan Federal University, Kremlevskaya Str. 18, Kazan, 420008, Russia

ARTICLE INFO

Article history:
Received 17 September 2022
Received in revised form
16 January 2023
Accepted 26 May 2023

Keywords:
Tight oil
Production forecast
LSTM-SVR
Residual correction

ABSTRACT

Production prediction is crucial for the recovery of hydrocarbon resources. However, accurate and rapid production forecasting remains challenging for unconventional reservoirs due to the complexity of the percolation process and the scarcity of available data. To address this problem, a novel model combining a long short-term memory network (LSTM) and support vector regression (SVR) was proposed to forecast tight oil production. Three variables, the tubing head pressure, nozzle size, and water rate were utilized as the inputs of the presented machine-learning workflow to account for the influence of operational parameters. The time-series response of tight oil production was the output and was predicted by the optimized LSTM model. An SVR-based residual correction model was constructed and embedded with LSTM to increase the prediction accuracy. Case studies were carried out to verify the feasibility of the proposed method using data from two wells in the Ma-18 block of the Xinjiang oilfield. Decline curve analysis (DCA) methods, LSTM and artificial neural network (ANN) models were also applied in this study and compared with the LSTM-SVR model to prove its superiority. It was demonstrated that introducing residual correction with the newly proposed LSTM-SVR model can effectively improve prediction performance. The LSTM-SVR model of Well A produced the lowest prediction root mean square error (RMSE) of 5.42, while the RMSE of Arps, PLE Duong, ANN, and LSTM were 5.84, 6.65, 5.85, 8.16, and 7.70, respectively. The RMSE of Well B of LSTM-SVR model is 0.94, while the RMSE of ANN, and LSTM were 1.48, and 2.32.

1. Introduction

Unconventional resources play a vital role in China's energy consumption, and tight oil/gas has gained increasing attention in recent years [1–3]. The amount of tight oil resources in China is approximately 200×10^8 t [4]. The accurate prediction of oil and gas production is important, as it serves as the basis for optimizing development schedules and improving project economics. The development of tight oil and gas reservoirs is characterized by complicated seepage mechanisms, and tight oil/gas recovery is influenced by various factors [5,6]. Traditional production prediction methods mainly include decline curve analysis (DCA) [7–10], flow material balance [11,12] and numerical simulations [13,14]. The DCA method is an empirical model, and conventional empirical models are not suitable for production forecasting of unconventional reservoirs because of their low prediction accuracy. This is because conventional empirical models consider simple parameters, which make it difficult to reflect the complex conditions of unconventional reservoirs. Moreover, excess parameters in improved models make them difficult to solve. Numerical simulations must consider numerous parameters, such as geological factors, fluid properties, and well operational parameters.

* Corresponding author.
** Corresponding author.
E-mail addresses: bwei@swpu.edu.cn (B. Wei), youjunyu2013@gmail.com (J. You).

Peer review under responsibility of Southwest Petroleum University.

Production and Hosting by Elsevier on behalf of KeAi

https://doi.org/10.1016/j.petlm.2023.05.004

Uncertainties in these parameters can lead to inaccurate simulation results. Therefore, it is necessary to develop new production prediction methods for unconventional reservoirs.

In recent years, with the development of machine learning (ML) and deep learning (DL) technology, artificial intelligence (AI) has been widely used in petroleum engineering domains, such as geological exploration, logging technology, and reservoir development [15–17]. ML techniques can comprehensively study the relationship between input data and output data and build high-quality models to mimic the input-output data structure.

Numerous studies have predicted production using machine learning algorithms. The prediction results were compared with different methods to verify the prediction accuracy of the algorithms. Martin et al. [18] proposed an automated data-driven approach for production forecasting using machine learning techniques, including random forest (RF), lasso regression, and support vector regression (SVR), to save time and cost. Noshi et al. [19] conducted research on production forecasts using ML consisting of a gradient boosting decision tree (GBDT), AdaBoost, and SVR. The forecast outcomes of the proposed models were compared with field data by cross-validation to express the generalization and precision performance.

Deep learning has also been applied in forecast production, which can imitate the human brain neural network to solve more complicated studies. Among these methods, artificial neural networks (ANN) and recurrent neural networks (RNN) are commonly used. In a previous study, an ANN model was used to predict the production of legacy and new wells in unconventional reservoirs. The production history, pressure data, and operational constraints were regarded as input variables in the model [20]. Han et al. [21] employed an ANN to forecast shale gas production in transient flow and compared the results with those of DCA methods. They concluded that the ANN predicted more accurate results than DCA methods. In the same year, Nnamdi and Adelaja [22] selected the key flow indicators (KFI) of production wells to establish ANN frameworks. This method was applied to three oil wells in the Niger Delta produced from separate reservoirs under different flow regimes. Linear regression and RNN can be used for production prediction without any geological model, fluid and rock properties, or flow characteristics of fluids in porous media [23] Extensive research has been carried out to obtain better prediction results, and researchers have used one algorithm to establish the input-output model and other algorithms to optimize the hyperparameters.

Production data is an obvious dynamic sequence, and the time-series data is gradually considered in production forecasting to improve prediction accuracy. Among numerous studies, LSTM shows a higher prediction potential owing to its structure of processing production data. A production prediction method based on data-mining technology and time-series analysis was proposed in 2017 [24]. Sun et al. [25] evaluated the production time series of a single well and multiple wells, and established corresponding prediction models based on the LSTM algorithm. The shale gas production and shut-in cycles were considered as input variables to establish the LSTM production prediction model for shale gas [26]. Gu et al. [27] employed LSTM to build a production forecast model and optimized the model hyperparameters to improve prediction accuracy. The application results indicated that the model could correctly predict the changing production trend. The LSTM production prediction was established to forecast the production of a sandstone water drive oilfield with medium and high permeability [28]. Liu et al. [29] proposed ensemble empirical mode decomposition (EEMD) based on the LSTM learning paradigm to achieve fast and accurate production forecasts.

Existing research has mainly focused on using different machine learning algorithms to analyze the characteristics of historical production data and establish the corresponding production prediction model. However, field production is influenced by various parameters such as geological factors and operational schedules. In particular, for unconventional reservoirs, which are characterized by low porosity and permeability, formation fluids cannot easily flow through the reservoir. Existing models do not perform well under such complex conditions. More importantly, the recorded production information can be incomplete, which leads to limited available data for production projection, thus lowering the prediction accuracy. Therefore, it is necessary to propose a new method to solve those problems for unconventional oil and gas reservoirs.

In this study, a prediction workflow for tight oil based on LSTM-SVR model is proposed, which considers the influence of the development operational schedules, as well as the structural features of the production data sequence. The LSTM model was first established to predict the initial production. The difference between the initial forecast production and the actual production was termed as the residual. The SVR model was then trained and applied to forecast the residual, and the predicted residual was taken back to the LSTM model. The result of the LSTM-SVR model was the final prediction. It was demonstrated that introducing the residual correction method and couple it with a hybrid LSTM-SVR model can improve the prediction performance when the available field measurements are limited.

This method can be applied to different types of reservoirs and can effectively improve prediction accuracy, thereby providing a practical example for forecasting the production of unconventional oil and gas reservoirs.

2. Reservoir description

In this study, we used data collected from an active tight oil reservoir located in the Ma-18 block of the Xinjiang oilfield. Fig. 1 shows the location of the study area and the stratigraphic information of the reservoir rock [30]. The Baikouquan Formation (T_1b) of the Ma-18 Block is located in the central depression of the Junggar Basin. It is subdivided into T_1b_1, T_1b_2 and T_1b_3 from bottom to top. The oil-bearing area of this formation is 82.04 km^2, and the crude oil is mainly distributed in T_1b_1 and T_1b_2. There is little difference in porosity among different layers, and the interlayer heterogeneity of the block is at the middle level. The variation coefficient of T_1b_1 is 3.38, and the permeability ratio is 333.1. The intra-formational heterogeneity is strong. Some key properties of the formation are listed in Table 1. For T_1b_1, the average porosity is 9.23%, average permeability is 0.39 mD, oil saturation is 52.7%, average oil density is 0.829 g/cm^3, and crude viscosity is 5.41 cP. For T_1b_2, the average porosity is 8.2%, average permeability is 0.16 mD, oil saturation is 44.7%, average oil density is 0.826 g/cm^3, and viscosity is 4.94 cP. Thus, the reservoir rock properties of T_1b_1 are better than those of T_1b_2.

3. Methodology

The main idea of the proposed method is as follows. First, the LSTM model is established to predict production. Then, the SVR model is employed to forecast the residual, and the prediction residual is compensated for in the forecast production as the final prediction result. The final predicted production of the LSTM model coupled with the SVR model is compared with that of the DCA, ANN, and LSTM models to prove the feasibility of the proposed method.

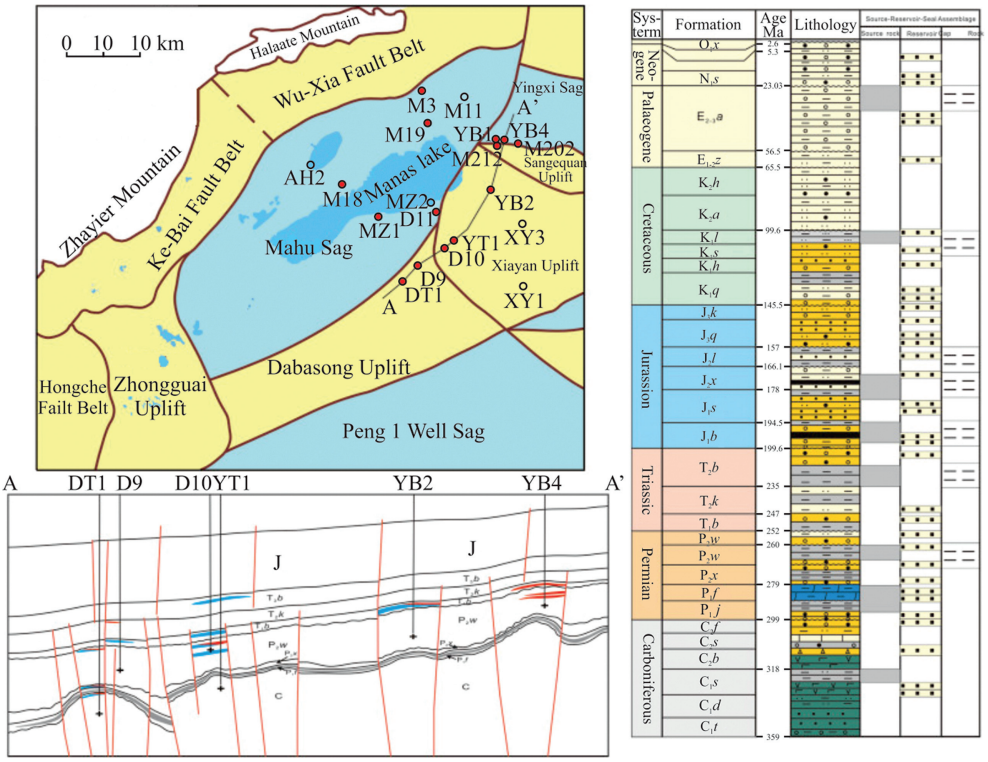

Fig. 1. Location of the study block and stratigraphic column of source rock and reservoir rock in the study area [30].

Table 1
Physical properties of the formation and fluid in the study block.

Section	K (mD)	φ (%)	S_o (%)	Oil density (g/cm^3)	μ (cp)
T_1b_1	0.39	9.23	52.7	0.829	5.41
T_1b_2	0.16	8.20	44.7	0.826	4.94

3.1. LSTM

A recurrent neural network (RNN) is a deep learning model structure, which is mainly used to solve time-series problems and problems of natural language processing. RNNs may encounter problems such as "vanishing gradient" and "exploding gradient", which means that the error gradient decreases or increases exponentially with augmentation of the number of network layers. The error gradient is the magnitude and direction of the calculated error when neural networks are trained. The model structure will be unstable under such conditions. LSTM is based on a simple RNN model, and was first proposed by Hochreiter [31]. The structure of the LSTM neural network is shown in Fig. 2. Compared to an RNN, an LSTM cell includes three gates called the memory units: an input gate, a forget gate, and an output gate. The unique memory units control information transmission between hidden layers, which can effectively solve the problem of "vanishing gradient" and "exploding gradient". The crucial aspects of LSTM are to ensure information update, network status, and information output.

At time step t, the input information passes through the input gate and the input value is I_t. The forget value F_t and output value O_t are as follows:

$$I_t = \sigma(W_{xi}x_t + W_{hi}H_{t-1} + W_{ci}C_{t-1} + d) \quad (1)$$

$$F_t = \sigma\left(W_{xf}x_t + W_{hf}H_{t-1} + W_{cf}C_{t-1} + d\right) \quad (2)$$

$$O_t = \sigma(W_{xo}x_t + W_{ho}H_{t-1} + W_{co}C_{t-1} + d) \quad (3)$$

where H_{t-1} is the value of the hidden layer at t-1 time step, C_{t-1} is the state value of the memory cell, σ is the activation function, d is the bias term, W is the weight coefficient matrix, and subscripts x,h,c,i,o and f represent the input layer, hidden layer, memory cell, input gate, output gate, and forget gate, respectively.

$$H_t = O_t \tan h(C_t) \quad (4)$$

The state value of the memory cell at time t is determined by

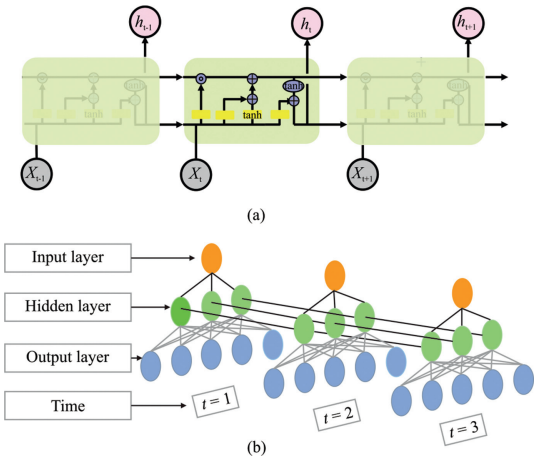

Fig. 2. Schematic diagram of a typical LSTM model: (a) the LSTM cell structure, (b) the LSTM neural network.

two parts: the value of the memory cell at the previous time step, and the candidate cell at the current time step:

$$C_t = f_t C_{t-1} + i_t \tan h(W_{xC}X_t + W_{hC}H_{t-1} + d) \quad (5)$$

where C_t is the state value of the memory cell and f_t is the value of the forget gate at time t.

3.2. SVR

SVR has the same theoretical basis as a support vector machine (SVM), which is typically used to solve linear and nonlinear regression problems. SVR can be used to study the structure of the objective function by learning the sample dataset. The regression model is formulated as follows:

$$G = \{(x_i, y_i) : x_i, y_i \in F\}_{i=1}^{n} \quad (6)$$

where x_i is the input variable and y_i is the output variable; $f(x) = w^T\Phi(x) + b$ is a regression function; $\Phi(x)$ is a nonlinear mapping method that maps x_i to a high-dimensional feature space; w is the weight vector, and b is the bias vector.

The quality of the regression function is measured by a loss function, which is usually an ε—insensitive loss function: $c(x,y,f(x)) = |y - f(x)|_\varepsilon$, where ε is a predetermined threshold. Fig. 3a plots the loss function of the univariate linear function, $f(x) = (w \cdot x) + b$. The dotted area is called the ε—band, which is shown in Fig. 3b. When the difference between the actual production y and the predicted production $f(x)$ at point x does not exceed ε, the predicted data sample is in the dotted area. The prediction at this point is considered to have no loss and indicates that the result is accurate. In contrast, the prediction has a loss at this point if the data sample point is outside the dotted area.

Building an SVR model mainly includes four steps.

(1) Selection of the training set: $T = \{(x_1,y_1),\ldots,(x_l,y_l)\} \in (x,y)^l$, where $x_i \in x \in R^n$, $y_i \in y \in R^n$, $i = 1,2,\ldots,l$.
(2) Determining the appropriate threshold ε and kernel function $K(x, x')$. The radial basis function (RBF) is selected as the kernel function of the SVR model. It has good generalization ability and fast learning convergence speed, which can approach any nonlinear function and form the regularity of the study.
(3) The dual problem is constructed and solved to obtain an optimal solution: $\overline{\alpha} = (\overline{\alpha_1}, \overline{\alpha_1}^*, \ldots, \overline{\alpha_l}, \overline{\alpha_l}^*)^T$.
(4) Construct the decision function $f(x) = \sum_{i=1}^{l}(\overline{\alpha_i}^* - \overline{\alpha_i})K(x_i,x_j) + \overline{b}$, where $\overline{\alpha_i}^* - \overline{\alpha_i}$ is not equal to zero, and the corresponding sample is the support vector.

Additionally, $\overline{a_i}$ and $\overline{a_i^*}$ are Lagrange multipliers, $\overline{b}$ is the threshold, K is the kernel function satisfying the Mercer condition $K(x_i,x_j) = \Phi(x_i)^T\Phi(x_j)$.

3.3. ANN

An ANN is an adaptive information processing system composed of a certain number of processing units that can model the interaction between neurons (Fig. 4). ANNs can abstract the human brain neural network from the perspective of information processing [32]. Weights represent the strength of connections between neurons, which are estimated by learned values from the input data based on an activation function. The estimations are calculated by reflecting the weights of different neurons through the learning of a cluster of neurons [21]. There are three different types of layers in an ANN model: an input layer, an output layer, and one or more hidden layers [33]. The number of hidden layers determines the complexity of the model. The ANN model can adjust the weights according to changes in the environment to meet the requirements of different network models.

In Fig. 4, X is the input vector, Y is the output vector, and w is the weights connecting different neurons.

4. The proposed method

The method proposed in this study mainly includes four procedures: preparation of the training database, data pre-processing, training the LSTM-SVR model, and result evaluation. The most important aspect was to first establish the LSTM model using time-series dynamic analysis. The input parameters of this LSTM model are head pressure, nozzle size, and water rate, and the output is oil rate. The SVR model was employed to predict the residual, which was the difference between the initial value predicted by the LSTM model and actual production. Subsequently, the prediction residual was compensated for in the initial prediction result of the LSTM model to obtain the final prediction results. The final predicted production was compared with the results of the DCA and ANN models to demonstrate the superiority of the proposed method. Fig. 5 shows a flowchart of the proposed method, and the four steps are as follows:

4.1. Preparation of the training database

The three characteristic variables including the tubing head pressure, nozzle size, and daily water production were trained as the inputs of the LSTM model from existing data. Subsequently, the three characteristic variables and residuals were used as the inputs of the SVR model to forecast the residual.

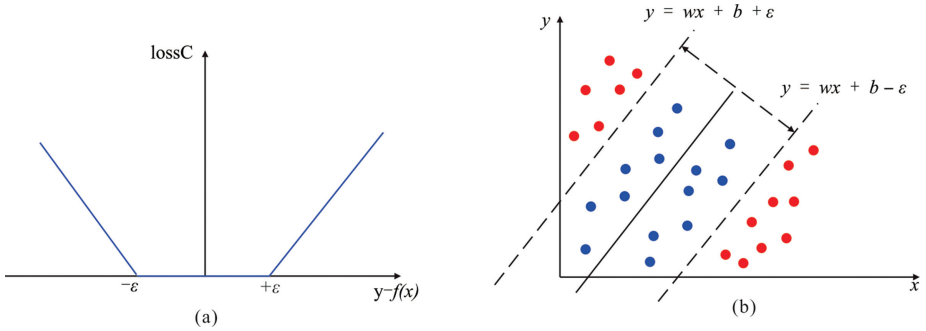

Fig. 3. The schematic diagram of the SVR: (a) loss function, (b) univariate linear loss function band.

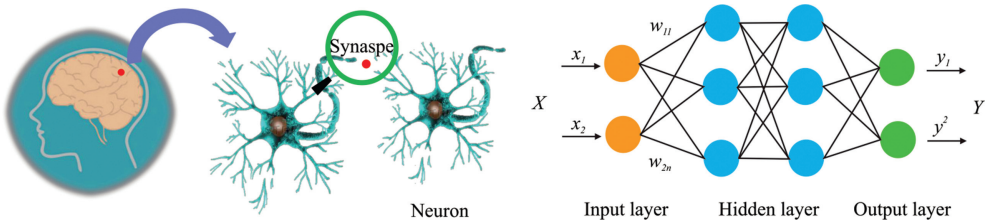

Fig. 4. Schematic diagram of ANN.

4.2. Data pre-processing

The input characteristic variables were standardized, and the output variable was the imputation of missing values. We adopted Z-score standardization. This standardized method can remove the discrete values of the characteristic variables, effectively eliminating the influence of data noise. The data were divided into two sets: a training set and a test set. The training set was used to train the LSTM model to predict the production and the test set was used to validate the accuracy of the prediction model.

4.3. Training the LSTM-SVR model

The training included three processes. First, the LSTM model was trained to predict the production. Then, the SVR model was applied to forecast the residual, and the prediction residual was compensated for in the forecast results of the LSTM model to obtain the final predicted production.

4.4. Evaluation of results

The LSTM, LSTM-SVR, DCA (Arps, PLE, and Duong), and ANN models were analyzed using the RMSE, which is an evaluation criterion to quantify the accuracy of each prediction model. The calculation of the RMSE is shown in Eq. (7).

$$RMSE = \sqrt{\frac{\Sigma_{i=1}^{n}(X_i - X_{pre})^2}{n}} \tag{7}$$

where X_i represents the actual daily oil production, X_{pre} represents the predicted production, and n represents the sample size.

5. Case study

Two wells in the Ma-18 block of the Xinjiang oilfield were selected to verify the proposed workflow, and the production data were trained and tested to prove the feasibility of the proposed method. The two types of factors that have the maximum impact on production are geological and operational parameters. The intra-formational heterogeneity of the selected blocks was strong. The operational parameters changed when the production wells were shut-in for well maintenance or nozzle replacement. Consequently, the declining trend of production was irregular (Fig. 6). Fig. 7 shows the changes in the nozzle size, tubing head pressure, and water production. The numerical simulation method was not considered because of its high cost and lack of key data. Production forecasting methods such as DCA and data-driven methods are more favorable. The historical production data of the reservoir were used as the training database. A total of 1190 data points were extracted from Well A, with production times ranging from June 2016 to September 2019. A total of 1194 data points were extracted from Well B, and the production time was from June 2017 to September 2020. Well A experienced two production stages: oil production first increased rapidly (June 2016 to August 2016) and then gradually decreased (August 2016 to May 2019). From day 1083, the nozzle size was changed from 3 mm to 4 mm, and the tubing head pressure was increased by 2 MPa. The oil production rate also had an obvious jump. From day 1088, the nozzle size was changed from 4 to 5 mm, the tubing head pressure remained constant, and the daily oil production decreased to 40 t. Subsequently, the daily oil production decreased gradually. Well B was in a stable production stage with a large fluctuation in output after the rapid increase.

5.1. Data preprocessing

The production data was pre-processed before training the prediction model. Preprocessing included two steps: interpolation for missing production values and normalization of input variables. When processing missing values because of a lack of records, the

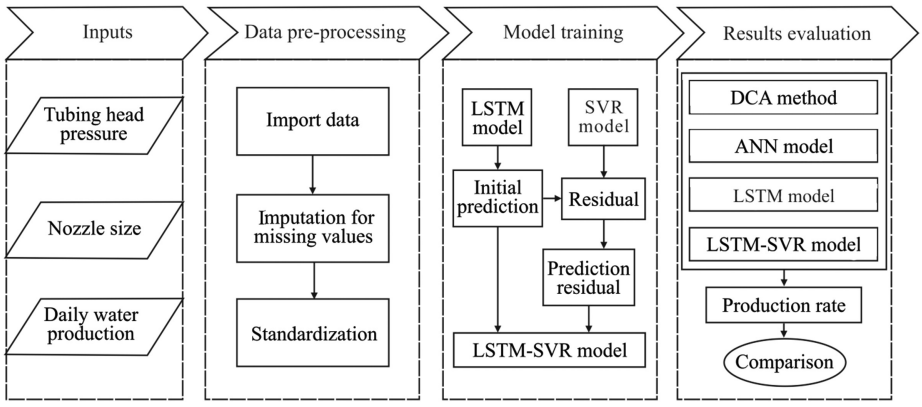

Fig. 5. The schematic diagram of the proposed workflow.

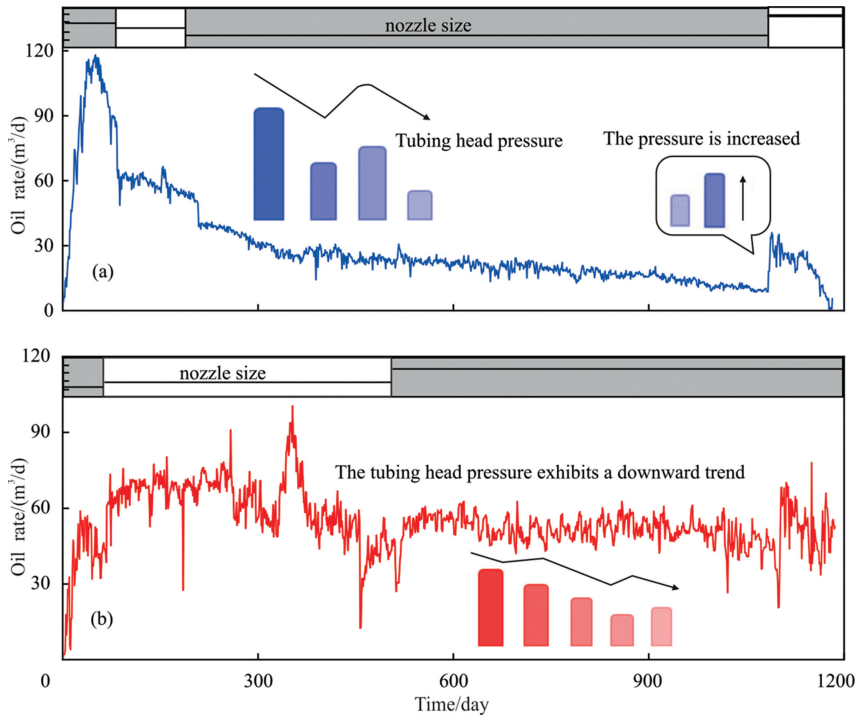

Fig. 6. History oil rate: (a) well A, (b) well B.

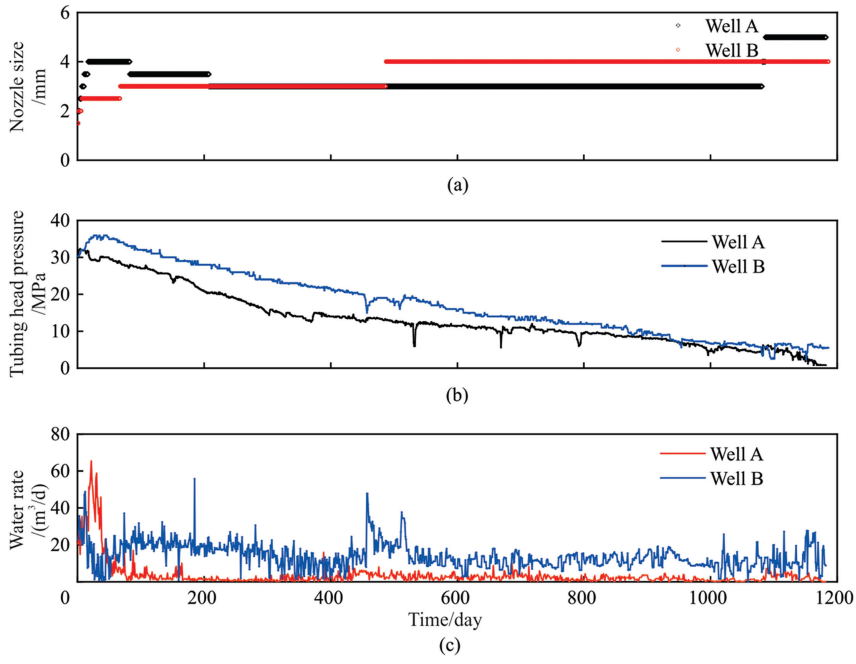

Fig. 7. Plots of input characteristic variables: (a) the oil nozzle size, (b) tubing head pressure, and (c) daily water production.

empty production values were filled with the average production over the weeks before and after the missing data. Considering the discrete values of the variables, Z-score normalization was adopted, and the calculation is shown in Eq. (8).

$$z_{ij} = \frac{x_{ij} - x_{mean}}{s_i} \tag{8}$$

where Z_{ij} is the standardized data, x_{ij} is the actual data, x_{mean} is the average of the actual data, and S_i is the variance.

After data preprocessing, there were 1182 and 1186 effective data points for Well A and Well B, respectively. Normalization effectively reduced the difficulty of training and the impact of the order of magnitudes on the prediction results.

5.2. DCA method

The DCA method [34,35] is one of the most commonly used methods to predict production. It is easy to implement and requires simple data. Three popular DCA models were chosen for forecasting

Table 2
The various models of the DCA method.

Arps model	$q = \frac{q_i}{(1 + bD_i t)^{1/b}}$	q_i is the initial production at the declining stage; D_i is the initial decline rate, d^{-1}, b is the decline index; exponential decline ($b = 0$); hyperbolic decline ($0 < b < 1$); harmonic decline ($b = 1$).
PLE model	$q = q_i \exp\left(-D_\infty t - \frac{D_i}{n} t^n\right)$	q_i is the initial production at the declining stage; D_∞ is the decline rate when time approaches infinity, d^{-1}, n is the time index;
Duong model	$q = q_i t^{-m} e^{\frac{a}{1-m}(t^{1-m} - 1)}$	q_i is the initial production at the declining stage; t is the time function; a is the decline coefficient; m is the index of the power function.

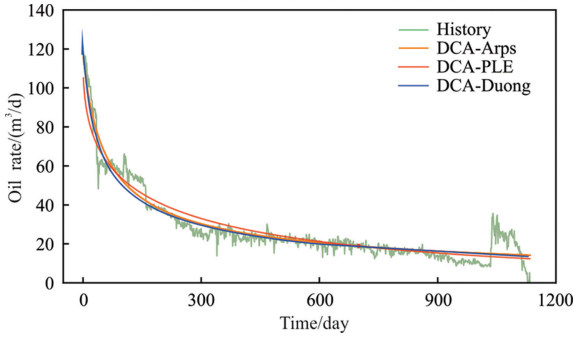

Fig. 8. The prediction results of DCA method.

oil production and were compared with the proposed methodology. These were the Arps model [7], PLE model [36] and Duong model [37,38]. According to the value of the decline index b, the Arps model was subdivided into exponential, hyperbolic, and harmonic declines (Table 2). The PLE model was proposed by Ilk [39] to improve the prediction accuracy by introducing more parameters, and was an extension of the Arps model. The Duong model was mainly applied to shale gas wells, which were studied based on linear fracture flow.

DCA method requires only production data for forecasting. They are based on empirical relations and cannot reflect the physical properties of underground flow. Fig. 8 shows the prediction results of Well A using DCA methods. These are able to fit the general declining trend of the production data. However, when there are changes in production measures (such as replacing the nozzle) due to which strong fluctuations occur, DCA methods cannot obtain good predictions. The production of Well A exhibits a long period of downward trend, where the fitting is effective. The fitting trends of the Arps and Duong models are basically the same. The RMSE of the Arps model is 5.84, indicating that the prediction result of this model is the best. The RMSE of the PLE model is 6.65, and the prediction accuracy is the lowest. In the first 1000 days, the operational parameters are stable without major changes. The production data are smoothed and exhibit a natural downward trend. Hence, the prediction effect is good. In the final 200 days, with the adjustment of the nozzle size from 3 mm to 5 mm, the oil production rate exhibits an obvious jump. It is difficult to predict the trend of change in production and the prediction is erroneous. The production of Well B is still in a stable stage with no downward trend. Thus, the DCA method has not been applied to forecast the production from Well B.

5.3. ANN model

The input layer of the ANN model included the same variables as those used in the LSTM model. The prediction result was compared with those of the DCA methods and the LSTM model. The hyperparameters of the ANN model were selected using a random search (RS) algorithm. The hyperparameters to be optimized included the number of hidden layers and the number of neurons in each hidden layer, which has been proven to effectively influence the prediction accuracy of trained ANN models [40]. The final optimized ANN model was selected for conducting the predictions. The number of hidden layers in the optimized ANN model was two and the number of neurons in each layer was five. The maximum number of training iterations was set to 5000. The database was divided into a training set and a test set at a ratio of 7:3, as shown in Fig. 9.

Fig. 10a and c shows the cross-plots of the trained ANN models of wells A and B, which demonstrate that an effective prediction model is obtained. In the first 260 days, the declining trend of Well A is predicted correctly, and the sudden change in production caused by changing the oil nozzle is accurately predicted. After 260 days, with the change in operational parameters, the production is promptly reduced to zero, which leads to a large error in the prediction. The final prediction error for Well A is 4.48. Well B is in a stable production stage, and the oil nozzle size remains unchanged during the test period. The main influencing factor is

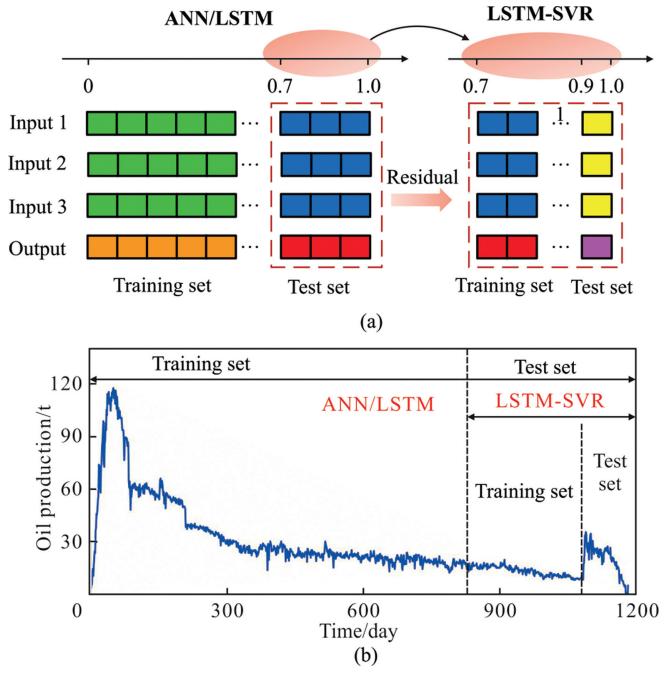

Fig. 9. (a) the schematic diagram of training and test dataset. For ANN and LSTM model, 70% of the data is used as training data, and the other 30% is used as test data; for LSTM-SVR, the first 70% of the raw data is used to train LSTM model; the left 30% of the data was then also divided into two parts based on a ratio of 7:3; (b) the schematic diagram of the data partition using Well A as an example.

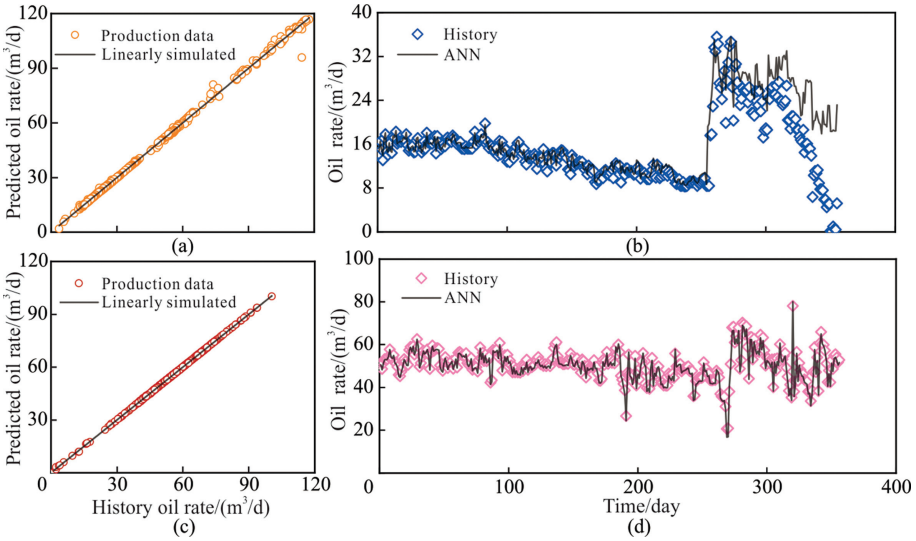

Fig. 10. The analysis and prediction results of the trained ANN model: (a) training cross-plot analysis of Well A, (b) the prediction results of Well A, (c) training cross-plot analysis of Well B, and (d) the prediction results of Well B.

the tubing head pressure in the input variables. Fig. 10d shows that the forecast production trend is consistent with the actual production. The RMSE of Well B is 1.1, and the prediction accuracy is effective.

5.4. LSTM model

Compared with the ANN model, the LSTM model considers the characteristics of the time series. The input variables of the LSTM model also consisted of the tubing head pressure, nozzle size, and daily water production to ensure the consistency of the prediction conditions. The parameters of the LSTM model were determined by RS, and the output with the best training performance was selected to predict production. Considering that our data sample was small, a complex LSTM structure may lead to over-fitted prediction results. Therefore, the searching range of the number of hidden layers was set to be [1,3], the range of the number of neurons was set to be [10, 20, …, 200]. Finally, it is observed that when the hidden layer is 1, the number of neurons is 50, and the number of training sessions was 5000, the prediction performance of the model was better. The ratio of the training to test sets was also 7:3.

Fig. 11 shows the conditions of the training and test results of the LSTM model. Compared to the ANN model, the LSTM model can correctly predict the declining trend of Well A according to the change in input variables after 260 days. However, the average forecast accuracy of the LSTM model is lower than that of the ANN model (Table 3). Hence, the prediction effect of the

Table 3
The RMSE of different methods.

Well number	DCA method			ANN model	LSTM model
	Arps	PLE	Duong		
A	5.84	6.65	5.85	4.48	4.77
B	_	_	_	1.10	1.16

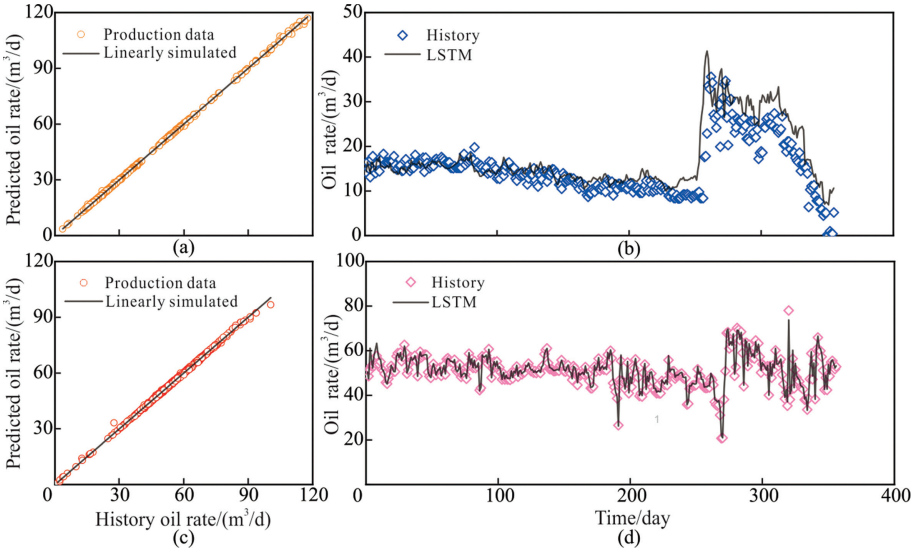

Fig. 11. The analysis and prediction results of the trained LSTM model: (a) training cross-plot analysis of Well A, (b) the prediction results of Well A, (c) training cross-plot analysis of Well B, and (d) the prediction results of Well B.

ANN model is better than that of the LSTM model before the production spurt. Well B is still in the stable production stage and exhibits a more regular production trend than Well A. As shown in Fig. 11, Well B has better prediction results than Well A. However, the prediction errors of both types of curves increase with the increase in the prediction time. Therefore, this problem can be solved by using the LSTM model to predict production.

Table 3 lists the RMSE values for the different methods. The RMSE of ANN and LSTM model of Well A are 4.48 and 4.77 respectively, while the RMSE of Arps, PLE and Doung model are 5.84, 6.65 and 5.85. It is apparent that the prediction accuracy of ML is significantly better than that of the DCA method. The data of Well B is not suitable for DCA methods, and the ML model of Well B has high prediction accuracy. Overall, Machine learning can fully mine the existing data features, which can effectively improve the prediction accuracy and provide a basis for the adjustment of the reservoir development plan.

5.5. LSTM-SVR model

The LSTM model can accurately obtain the characteristics of dynamic data because of the unique structure of the memory unit, which is suitable for handling events with time-series characteristics. Thus, it can be employed for complex production problems caused by well shut-in or equipment repairing. SVR has an efficient nonlinear fitting ability and a simple structure. Therefore, it was applied to predict the residuals of the LSTM model. The predicted residual was taken back to the LSTM model to improve the prediction accuracy. The training of LSTM model took up the first 70% of the raw data, the training of SVR-based residual correction model used 21% of the datasets, and test data was the last 9% (as shown in Fig. 9). The sample sizes of the test sets of wells A and B were 106 groups.

As shown in Fig. 12, the prediction accuracy can be improved after residual correction. The prediction trend of the LSTM-SVR model of Well A is closer to the actual production trend, in contrast to the LSTM model from Figs. 12a and c.LSTM-SVR model can accurately predict the production fluctuation stage (blue circle marked in Fig. 12). The RMSE of Well A is 7.70 before residual correction and 5.42 after residual correction in the same data sample. The operational parameters of Well B are continuous and stable, and do not change frequently. The production of Well B is in a stable stage and is more regular than that of Well A. Thus, the forecasting accuracy of Well B is better than that of Well A based on the prediction curves of both the ANN and LSTM models. Compared with the LSTM model, the extreme points of well B are more accurate after residual correction. The RMSE of the LSTM-SVR model is 0.94, while that of LSTM is 2.32 in the same conditions. The prediction accuracy of the LSTM model is improved significantly after residual correction.

Table 4 shows the RMSE of different models using the same test data as LSTM-SVR model. Compared with Well B, Well A shows higher prediction error no matter what model are used, which is due to the sharp changes in production data of Well A caused by the change in operational parameters. Additionally, in accordance with the observation from Tables 3 and 4, the prediction error of the LSTM model increases with increasing prediction time. The integral RMSE of the LSTM model of well A is 4.77. However, the RMSE of the final 30% prediction is 7.7. The overall RMSE of Well B is 1.16, whereas that of the last 30% prediction is 2.32. Both the RMSE values of the last 30% predictions are significantly greater than the integral RMSE.

To reflect the accuracy of the prediction results intuitively, Fig. 13 compares the prediction accuracy distribution of wells A and B using different models. The prediction accuracy is represented by the relative error. It is evident that the prediction accuracy of well B is more homogeneous than that of well A. The average prediction accuracy of the ANN model is higher than that of the LSTM model, and the accuracy of the LSTM model is improved effectively after residual correction. The prediction accuracy of Well A is increased by 7.13%, and that of Well B increases by 0.77%. This is because the prediction accuracy of the LSTM model of Well A is significantly lower than that of Well B, and the margin for improvement is higher than that of Well B.

Table 4
The RMSE of different model (based on the same test data as LSTM-SVR model).

Well number	ANN model	LSTM model	LSTM-SVR model
A	8.16	7.70	5.42
B	1.48	2.32	0.94

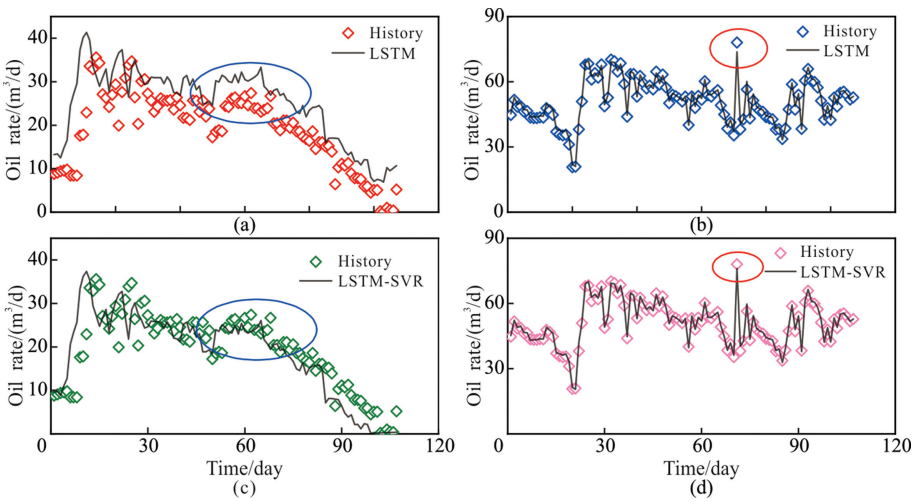

Fig. 12. The comparison of prediction results between LSTM model and LSTM-SVR model: (a) LSTM model prediction results of Well A, (b) LSTM model prediction results of Well B, (c) LSTM-SVR model prediction results of Well A, and (d) LSTM-SVR model prediction results of Well B. The circled parts show the improvements made after implementing residual correction.

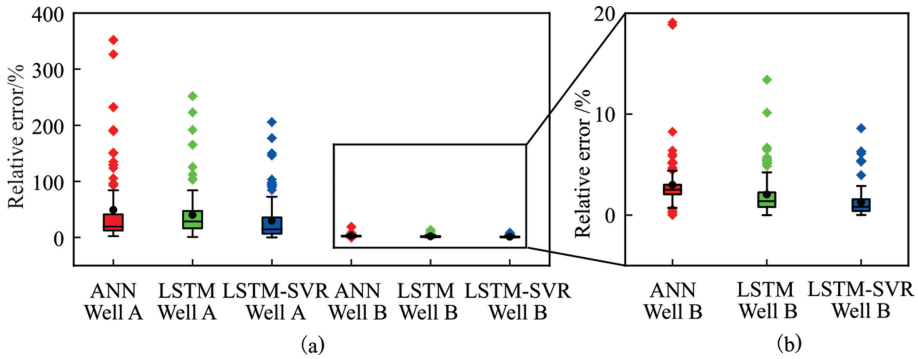

Fig. 13. (a) the comparison of ANN, LSTM and LSTM-SVR model (the prediction error is calculated based on test data), (b) the magnified relative error of Well B.

6. Conclusions

DCA methods cannot adjust the prediction results with changes in operational parameters, which possibly results in inaccurate forecasts. In addition, the error of the LSTM model increases with increase in the prediction time. A hybrid LSTM-SVR model is proposed to consider not only the characteristics of the time series of production data, but also the influence of operational parameters. Introducing residual correction mothed with LSTM-SVR model can effectively enhance the prediction accuracy. This forecasting approach builds an entire prediction process, which is practical to predict the production of unconventional oil and gas reservoirs. The main conclusions drawn from this study are as follows:

(1) The prediction result of LSTM-SVR model for Well A was compared with that of DCA methods, an ANN model and an LSTM model. The DCA method is appropriate for wells with declining production. Well B is still in stable production stage with no declining trend, and DCA methods cannot accurately predict production.
(2) Compared to the ANN model, the LSTM model can capture the characteristics of the time series. However, the prediction errors of the LSTM model are larger than that of the ANN model. The RMSE of the LSTM model of Well A and Well B are 4.77 and 1.16, respectively, while that of the ANN is 4.48 and 1.10.
(3) The developed method for tight oil production prediction based on LSTM-SVR model can effectively improve the forecasting accuracy. The RMSE of Well A of the LSTM model is 7.70, and that of LSTM-SVR is 5.42. The RMSE of the LSTM and LSTM-SVR models of Well B are 2.32 and 0.94, respectively. LSTM coupled SVR can accurately predict the change trend of production and extreme points and has broad application prospects.

Acknowledgments

The authors gratefully acknowledge the financial support of National Natural Science Foundation of China (52274041 and 51974265), Sichuan science fund for distinguished Young Scholars (2023NSFSC1954), the Ministry of Science and Higher Education of the Russian Federation under Agreement No. 075-15-2022-299 within the framework of the development program for a world-class Research Center “Efficient development of the global liquid hydrocarbon reserves”, Science and Technology Research Program of Chongqing Municipal Education Commission (KJQN202201510), Natural Science Foundation of Chongqing (CSTB2022NSCQ-MSX0403), Chongqing Municipal Support Program for Overseas Students Returning for Entrepreneurship and Innovation (2205012980950154), Scientific Research Funding Project of Chongqing University of Science and Technology (ckrc2021040). Thanks for the relevant data provided by Xinjiang oilfield.

References

[1] Y. Assef, A. Kantzas, P. Pereira Almao, Numerical modelling of cyclic CO_2 injection in unconventional tight oil resources; trivial effects of heterogeneity and hysteresis in Bakken formation, Fuel 236 (2019) 1512–1528, https://doi.org/10.1016/j.fuel.2018.09.046.
[2] S. Li, L. Sun, L. Wang, Z. Li, et al., Hybrid CO_2-N_2 huff-n-puff strategy in unlocking tight oil reservoirs, Fuel 309 (2022) 122198, https://doi.org/10.1016/j.fuel.2021.122198.
[3] B. Wei, L. Wang, T. Song, M. Zhong, M.A. Varfolomeev, Enhanced oil recovery by low-salinity water spontaneous imbibition (LSW-SI) in a typical tight sandstone formation of mahu sag from core scale to field scale, Petroleum 7 (3) (2021) 272–281, https://doi.org/10.1016/j.petlm.2020.09.005.
[4] C. Zou, et al., Geological concepts, characteristics, resource potential and key techniques of unconventional hydrocarbon: on unconventional petroleum geology, Petrol. Explor. Dev. 40 (4) (2013) 385–399, https://doi.org/10.11698/PED.2013.04.01.
[5] L. Sun, C. Zou, A. Jia, et al., Development characteristics and orientation of tight oil and gas in China, Petrol. Explor. Dev. 46 (6) (2019) 1073–1087, https://doi.org/10.1016/S1876-3804(19)60264-8.
[6] X. Zhou, Q. Yuan, Y. Zhang, et al., Performance evaluation of CO_2 flooding process in tight oil reservoir via experimental and numerical simulation studies, Fuel 236 (2019) 730–746, https://doi.org/10.1016/j.fuel.2018.09.035.
[7] J.J. Arps, Analysis of decline curves, Transactions of the AIME 160 (1) (1945) 228–247, https://doi.org/10.1016/j.petrol.2010.05.007.
[8] R.D. Carter, Type curves for finite radial and linear gas-flow systems: constant-terminal-pressure case, Soc. Petrol. Eng. J. 25 (5) (1985) 719–728, https://doi.org/10.2118/12917-PA.
[9] R.D. Hazlett, U. Farooq, D.K. Babu, A complement to decline curve analysis, SPE J. 26 (4) (2021) 2468–2478, https://doi.org/10.2118/205390-PA.
[10] K. Jongkittinarukorn, N. Last, F.H. Escobar, et al., A straight-line DCA for a gas reservoir, J. Petrol. Sci. Eng. 201 (2021) 108452, https://doi.org/10.1016/j.petrol.2021.108452.
[11] L. He, H. Mei, X. Hu, et al., Advanced flowing material balance to determine original gas in place of shale gas considering adsorption hysteresis, SPE Reservoir Eval. Eng. 22 (4) (2019) 1282–1292, https://doi.org/10.2118/195581-PA.
[12] I.S. Afanasyev, A.V. Timonov, I.V. Sudeev, et al., Analysis of multiple fractured horizontal wells application at Priobskoye field, SPE 162031, in: SPE Russian Oil and Gas Exploration and Production Technical Conference and Exhibition, 2012, https://doi.org/10.2118/162031-MS. Moscow, Russia.
[13] R. Velasco, P. Panja, M. Deo, Moving boundary approach to forecast tight oil production, AIChE J. 67 (2) (2021), e17012, https://doi.org/10.1002/aic.17012.
[14] Y. Wu, L. Cheng, L. Ma, S. Huang, et al., A transient two-phase flow model for production prediction of tight gas wells with fracturing fluid-induced formation damage, J. Petrol. Sci. Eng. 199 (2021) 108351, https://doi.org/10.1016/j.petrol.2021.108351.
[15] l Kuang, et al., Application and development trend of artificial intelligence in petroleum exploration and development, Petrol. Explor. Dev. 48 (1) (2021) 1–11, https://doi.org/10.11698/PED.2021.01.01.
[16] P. Gao, C. Jiang, Q. Huang, et al., Fluvial facies reservoir productivity prediction method based on principal component analysis and artificial neural network, Petroleum 2 (1) (2016) 49–53, https://doi.org/10.1016/j.petlm.2015.12.005.
[17] M.A. Ahmadi, Z. Chen, Comparison of machine learning methods for estimating permeability and porosity of oil reservoirs via petro-physical logs, Petroleum 5 (3) (2019) 271–284, https://doi.org/10.1016/j.petlm.2018.06.002.

[18] E. Martin, P. Wills, D. Hohl, et al., Using machine learning to predict production at a Peace River thermal EOR site, SPE-182696-MS, in: SPE Reservoir Simulation Conference, 2017, https://doi.org/10.2118/182696-MS. Montgomery, USA.
[19] C.I. Noshi, M.R. Eissa, R.M. Abdalla, An intelligent data driven approach for production prediction, OTC-29243-MS, in: Offshore Technology Conference, 2019, https://doi.org/10.4043/29243-MS. Houston, Texas, USA.
[20] Q. Cao, R. Banerjee, S. Gupta, et al., Data driven production forecasting using machine learning, SPE-180984-MS, in: SPE Argentina Exploration and Production of Unconventional Resources Symposium, 2016, https://doi.org/10.2118/180984-MS. Buenos Aires, Argentina.
[21] D. Han, S. Kwon, H. Son, J. Lee, Production forecasting for shale gas well in transient flow using machine learning and decline curve analysis, URTEC-198198-MS, in: SPE Asia Pacific Unconventional Resources Technology Conference, 2019, https://doi.org/10.15530/AP-URTEC-2019-198198. Brisbane, Australia.
[22] D.N. Nnamdi, V.O. Adelaja, Dynamic Production forecasting using artificial neural networks customized to historical well key flow indicators, SPE-198756-MS, in: SPE Nigeria Annual International Conference and Exhibition, Lagos, Nigeria, 2019, https://doi.org/10.2118/198756-MS.
[23] L. Kubota, D. Reinert, Machine learning forecasts oil rate in Mature Onshore Field jointly driven by water and steam injection, SPE-196152-MS, in: SPE Annual Technical Conference and Exhibition, 2019, https://doi.org/10.2118/196152-MS. Calgary, Alberta, Canada.
[24] A.F. Maqui, X. Zhai, A. Suarez Negreira, et al., A comprehensive workflow for near real time waterflood management and production optimization using reduced-physics and data-driven technologies, SPE-185614-MS, in: SPE Latin America and Caribbean Petroleum Engineering Conference, 2017, https://doi.org/10.2118/185614-MS. Buenos Aires, Argentina.
[25] J. Sun, X. Ma, M. Kazi, Comparison of decline curve analysis DCA with recursive neural networks RNN for production forecast of multiple wells, SPE-190104-MS, in: SPE Western Regional Meeting, 2018, https://doi.org/10.2118/190104-MS. Garden Grove, California, USA.
[26] K. Lee, J. Lim, D. Yoon, et al., Prediction of shale-gas production at Duvernay formation using deep-learning algorithm, SPE J. 24 (6) (2019) 2423–2437, https://doi.org/10.2118/195698-PA.
[27] j Gu, m Zhou, z Li, et al., Oil well production forecast with long-short term memory network model based on data mining, Special Oil Gas Reservoirs 26 (2) (2019) 77–81, https://doi.org/10.3969/j.issn.1006 - 6535.2019.02.013.
[28] h Wang, et al., Production prediction at ultra-high water cut stage via recurrent neural network, Petrol. Explor. Dev. 47 (5) (2020) 1009–1015, https://doi.org/10.11698/PED.2020.05.15.
[29] W. Liu, W.D. Liu, J. Gu, Forecasting oil production using ensemble empirical model decomposition based Long Short-Term Memory neural network, J. Petrol. Sci. Eng. 189 (2020) 107013, https://doi.org/10.1016/j.petrol.2020.107013.
[30] W. Jiang, A. Imin, X. Wang, et al., Geochemical characterization and quantitative identification of mixed-source oils from the baikouquan and lower wuerhe formations in the eastern slope of the mahu sag, junggar basin, NW China, J. Petrol. Sci. Eng. 191 (2020) 107175, https://doi.org/10.1016/j.petrol.2020.107175.
[31] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (8) (1997) 1735–1780, https://doi.org/10.1162/neco.1997.9.8.1735.
[32] M. Stundner, J.S. Al-Thuwaini, How data-driven modeling methods like neural networks can help to integrate different types of data into reservoir management, in: SPE Middle East Oil Show, Bahrain, 2001, https://doi.org/10.2118/68163-MS. SPE 68163.
[33] A. Ansari, M. Heras, J. Nones, M. Mohammadpoor, F. Torabi, Predicting the performance of steam assisted gravity drainage (SAGD) method utilizing artificial neural network (ANN), Petroleum 6 (4) (2020) 368–374, https://doi.org/10.1016/j.petlm.2019.04.001.
[34] I. Gupta, C. Rai, D. Devegowda, et al., Haynesville shale: predicting long-term production and residual analysis to identify well intereference and frac hits, SPE-195218-MS, in: SPE Oklahoma City Oil and Gas Symposium, 2019, https://doi.org/10.2118/195673-PA. Oklahoma City, Oklahoma, USA.
[35] D. Ilk, A.D. Perego, J.A. Rushing, et al., Integrating multiple production analysis techniques to assess tight gas sand reserves: defining a new paradigm for industry best practices, SPE 114947, in: CIPC/SPE Gas Technology Symposium 2008 Joint Conference, 2008, https://doi.org/10.2118/114947-MS. Calgary, Alberta, Canada.
[36] D. Ilk, J.A. Rushing, A.D. Perego, et al., Exponential vs. hyperbolic decline in tight gas sands: understanding the origin and implications for reserve estimates using Arps' decline curves, SPE 116731, in: SPE Annual Technical Conference and Exhibition, 2008, https://doi.org/10.2118/116731-MS. Denver, Colorado, USA.
[37] A.N. Duong, An unconventional rate decline approach for tight and fracture-dominated gas wells, CSUG/SPE 137748, in: Canadian Unconventional Resources and International Petroleum Conference, 2010, https://doi.org/10.2118/137748-MS. Calgary, Alberta, Canada.
[38] A.N. Duong, Rate-decline analysis for fracture-dominated shale reservoirs: Part 2, SPE-171610-MS, in: SPE/CSUR Unconventional Resources Conference, 2014, https://doi.org/10.2118/137748-PA. Calgary, Alberta, Canada.
[39] D. Ilk, A.D. Perego, J.A. Rushing, et al., Integrating multiple production analysis techniques to assess tight gas sand reserves: defining a new paradigm for industry best practices, SPE 114947, in: CIPC/SPE Gas Technology Symposium 2008 Joint Conference, 2008, https://doi.org/10.2118/114947-MS. Calgary, Alberta, Canada.
[40] J. You, W. Ampomah, Q. Sun, Development and application of a machine learning based multi-objective optimization workflow for CO_2-EOR projects, Fuel 264 (2020) 116758, https://doi.org/10.1016/j.fuel.2019.116758.