Advances in
Geo-Energy Research

Vol. 8, No. 3, p. 159-169, 2023

Original article

Gas well performance prediction using deep learning jointly driven by decline curve analysis model and production data

Liang Xue[1,2]*, Jiabao Wang[1,2], Jiangxia Han[1,2], Minjing Yang[1,2], Mpoki Sam Mwasmwasa[1,2], Felix Nanguka[3]

[1]*National Key Laboratory of Petroleum Resources and Engineering, China University of Petroleum, Beijing 102249, P. R. China*

[2]*Department of Oil-Gas Field Development Engineering, College of Petroleum Engineering, China University of Petroleum, Beijing 102249, P. R. China*

[3]*Tanzania Petroleum Development Corporation, Dar es Salaam 2774, Tanzania*

Keywords:
Gas well performance
long short-term memory neural network
decline curve analysis model
deep learning

Cited as:
Xue, L., Wang, J., Han, J., Yang, M., Mwasmwasa, M. S., Nanguka, F. Gas well performance prediction using deep learning jointly driven by decline curve analysis model and production data. Advances in Geo-Energy Research, 2023, 8(3): 159-169.
https://doi.org/10.46690/ager.2023.06.03

Abstract:
The prediction of gas well performance is crucial for estimating the ultimate recovery rate of natural gas reservoirs. However, physics-based numerical simulation methods require a significant effort to build a robust model, while the decline curve analysis method used in this field is based on certain assumptions, hence its applications are limited due to the strict working conditions. In this work, a deep learning model driven jointly by the decline curve analysis model and production data is proposed for the production performance prediction of gas wells. Due to the time-series characteristics of gas well production data, the long short-term memory neural network is selected to establish the architecture of artificial intelligence. The existing decline curve analysis model is first implicitly incorporated into the training process of the neural network and then used to drive the neural network construction along with the actual gas well production historical data. By applying the proposed innovative model to analyze the conventional and tight gas well performance predictions based on field data, it is demonstrated that the proposed long short-term memory neural network deep learning model driven jointly by the decline curve analysis model and production data can effectively improve the interpretability and predictive ability of the traditional long short-term memory neural network model driven by production data alone. Compared with the data-driven model, the jointly driven model can reduce the mean absolute error by 42.90% and 13.65% for a tight gas well and a carbonate gas well, respectively.

1. Introduction

The prediction of gas well performance is one of the key tasks of natural gas development, as it can provide the scientific basis for setting up the development plan of natural gas reservoirs. The accurate forecasting of gas reservoir production dynamics is crucial to the rational management of gas production systems, as well as the dynamic evaluation of reservoir reserves and the adjustment of the future development plan, ultimately affecting the overall efficiency of gas reservoir development. Gas well is the basic unit of a gas reservoir development system, and monitoring devices at the wellhead and bottom can provide information about changes in the reservoir production and formation properties. A deeper understanding of gas well production performance is beneficial to the effective identification of the dynamic characteristics of reservoir development as a whole. Therefore, it is necessary to establish an accurate method to analyze the production dynamics of gas wells and establish the tendency of gas reservoir development to guide scientific-based and rational exploitation.

Several methods have been developed to dynamically forecast gas well production to satisfy the needs of the practical development process of gas reservoirs. Decline curve analysis

Yandy Scientific Press

*Corresponding author.
E-mail address: xueliang@cup.edu.cn (L. Xue); wangjiabao724@163.com (J. Wang); hanj459@163.com (J. Han); 2019210383@student.cup.edu.cn (M. Yang); mwasamwasampoki143@gmail.com (M. S. Mwasmwasa).
2207-9963
Received April 11, 2023; revised May 2, 2023; accepted May 18, 2023; available online May 22, 2023.

(DCA) and numerical simulation are examples of widely used methods. Arps (1945) proposed three DCA methods by studying the relationship between well production rate and production time based on the bottom-hole pressure of the well. These methods were used to analyze the production rate changes of oil and gas wells in the boundary-controlled flow stage and to establish the typical decline curve chart. However, Arps's methods are only suitable for production prediction before the gas well production stage reaches the boundary-controlled flow stage. To address this issue, Fetkovich (1973) developed semi-analytical type curves that combine analytically derived transient flow regime stems with Arps's empirically derived boundary dominated flow hyperbolic decline stems. Blasingame and Poe (1993) utilized a superposition time function to handle varying rate/pressure conditions, and implemented pseudo-pressure functions to account for the pressure-dependence of fluid properties. This method can objectively and realistically reflect the actual situation of gas reservoirs under the changing bottom-hole flow pressure condition. Valkó and Lee (2010) presented the extended exponential DCA model, and Duong (2011) proposed a new DCA model. Both of the above models have been applied successfully in predicting the production of shale gas and tight gas. Shabro et al. (2011) studied the effects of no-slip and slip flow, Knudsen diffusion and Langmuir desorption in numerical simulation to predict shale gas production. Frooqnia et al. (2011) simulated the fluid flow in the wellbore based on fluid mechanics theory and logging data to estimate the formation permeability. Despite the above research advances, the conventional methods used in gas well performance prediction often have limitations and shortcomings, which are mainly due to the complexity of the numerical simulation process constrained by specific reservoir conditions and production system requirements.

In recent years, the rapid development of neural network algorithms has provided new directions for the dynamic prediction of gas well production. These algorithms use deep learning methods and big data analysis techniques to overcome the shortcomings of traditional methods, such as the needs for multiple sets of typical curve graphs for single well interpretation and the multi-scale nature of the modeling process. Deep learning methods have been gradually introduced into the field of production DCA and have achieved good results. Mollaiy and Shahbazian (2011) proposed a new method based on a feed-forward artificial neural network and imperialist competitive algorithm to predict the oil flow rate of wells. Kuzma et al. (2014) constructed a prediction model that can capture oil and gas seepage rules using a small amount of actual information combined with a generative model and statistical methods to make accurate predictions based on production experience. Jia and Zhang (2016) applied time series analysis and neural network models to forecast the future production of a gas well in the Barnett shale field, and compared it with the Arps decline curve. Sagheer and Kotb (2019) proposed a deep learning approach capable of addressing the limitations of traditional forecasting approaches, and it achieved accurate predictions. Song et al. (2020) developed a long short-term memory (LSTM) neural network-based model to infer the production of fractured horizontal wells in a volcanic reservoir, which could address the limitations of traditional method and obtain accurate predictions. Temizel et al. (2020) analyzed the data of tight shale reservoirs using the LSTM neural network, in which the operational interferences to the well were taken into account to ensure that the machine learning model was not impacted by interferences that did not reflect the actual physics of the production mechanism effecting the behavior of the well. At the same time, the LSTM model predictions were compared with the numerical simulation results, which verified the long-term accurate prediction ability of the LSTM model. Fan et al. (2021) established a novel hybrid model that considers the advantages of linearity and nonlinearity, as well as the impact of manual operations. Werneck et al. (2022) used stacked recurrent neural networks (RNN) to make short-term predictions of fluid rates and bottom-hole pressures in oil and gas reservoirs for 30 days. Experiments on recurrent networks showed that long-term input data and designing specific key time data segments can effectively improve oil and gas production and pressure prediction. Li et al. (2022b) developed a deep learning model based on LSTM neural network that can consider human operations. This model can learn shutdown time, oil nozzle size and daily production time to predict oil well production under artificial operation conditions. Ning et al. (2022) presented a machine learning-based time series forecasting method, which considers the existing data as time series and extracts the salient characteristics of historical data to predict the values of a future time sequence. Li et al. (2022a) proposed a reservoir production prediction model based on a combined convolutional neural network (CNN) and a LSTM neural network model optimized by the particle swarm optimization (PSO) algorithm. Zha et al. (2022) proposed a CNN-LSTM model to predict production in a gas field in southwest China. Karasu and Altan (2022) proposed a model that can cope with uncertainty, non-stationarity, and nonlinearity in crude oil time series more effectively than other studies in the literature, thus exhibiting higher prediction performance in terms of both accuracy and robustness.

Despite of the advancement of neural network methods in predicting the production characteristics, these methods can only train the network through data-driven methods, and fail to consider the physical law governing the oil or gas flow through porous media. physics-informed neural network (PINN) is a machine learning technique that can embed the physics of the flow problem, i.e., the underlying partial differential equation (PDE), into the architecture of the neural network. Raissi et al. (2019) proposed the PINN method for solving nonlinear PDE, such as Schrödinger, Burgers and Allen-Cahn equations. This method can deal with the forward and inverse problems of estimating the solutions of governing equation and parameters from the observation data. It is a physics-informed learning approach, which allows the integration of physical laws in the form of PDEs into the loss function of the neural network (Karniadakis et al., 2021). The PINN can train a neural network to minimize a loss function, which includes both the terms with the initial, boundary conditions and the PDE residual along the domain and the terms reflecting the data mismatch based on automatic differentiation (Baydin et

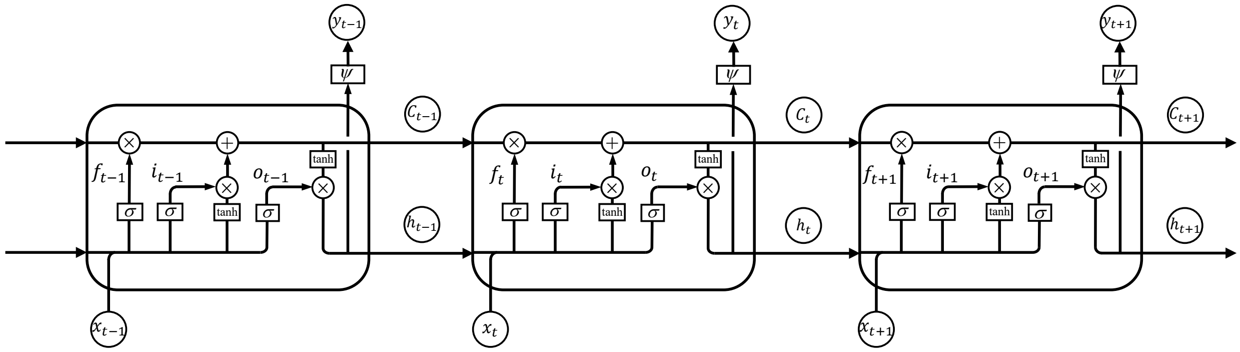


Fig. 1. Schematic diagram of LSTM.

al. 2018). Almajid and Abu-Al-Saud (2022) implemented the PINN to model the classical Buckley-Leverett problem. In addition, Fraces and Hamdi (2021) adopted the PINN to solve the two-phase immiscible flow problem and gained a physical solution by adding a diffusion term to the PDEs or an amount of observed data. Wang et al. (2020) proposed a theory-guided neural network model for subsurface flow using heterogeneous model parameters. More details of the research progress about PINN can be found in Cuomo et al. (2022) and Muther et al. (2023).

In the above-mentioned literatures, PINN-based approaches usually train the neural network through observation data with the constraint of the partial differential equation. However, many widely used empirical relationship models have been derived for gas well production performance prediction, such as those with the well-constructed characteristics of declining curve that governs the production declining process. In this work, enlightened by the strategy of PINN, a deep learning method driven jointly by DCA model and production data is proposed to predict the gas well production performance. The proposed method utilizes an appropriate long and short-term memory deep learning time series neural network architecture based on the characteristics of gas well production data, and embeds the conventional DCA model into the neural network architecture to further improve the accuracy of gas well performance prediction.

2. Methodology

2.1 LSTM neural network

Considering the temporal characteristics of gas well production, a time series model should be adopted for production forecasting. Time series models have achieved significant progress due to the natural language processing techniques in the artificial intelligence field, and have been successfully applied in oil and gas well performance prediction. These predictive models have improved generality because they only consider historical data. In this paper, based on the time-series relationship of gas reservoir production decline curve, the deep learning algorithms suitable for time series analysis are selected. Among these algorithms, the LSTM neural network is preferred, as this type of neural network algorithm has great flexibility and has made significant progress in various fields through years of research by many scholars (Chi, 2022; Chung et al., 2022; Rojc and Mlakar, 2022).

The LSTM neural network is an improved structure of RNN (Gers et al., 2000). Due to the sensitivity of traditional RNN to data gradients, the model prediction accuracy of RNN is largely controlled by the quality of data. To address the above-mentioned issues, researchers have introduced unit states and three control gates based on the RNN. The unit state is used to determine the information retention between different time steps, and the control gates are set to adjust the information transfer function between different positions. This modified neural network is called LSTM neural network, and its structure is shown in Fig. 1.

In the hidden layer of LSTM neural network, the input gate, output gate and forget gate are added to adjust the cell state. The forget gate can be expressed as:

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right) \tag{1}$$

where the f_t represents the forget gate information at time t; σ represents the sigmoid function; W_f and U_f are the weight parameters of the forget gate; b_f represents the bias of the forget gate; h_{t-1} is the output information of hidden layer.

The input gate is used to input the data x_t and the output data h_{t-1} from the previous time step to calculate the forget gate value of the memory cell. The input gate can be expressed as:

$$i_t = \sigma\left(W_i x_t + U_i h_{t-1} + b_i\right) \tag{2}$$

where i_t denotes the input information at time t; W_i and U_i are the weight parameters of the input gate; b_i represents the bias of the input gate. It outputs a value to transfer the unit state information and determines which information of the previous unit state should be retained or discarded. At this point, the candidate unit state C_t is calculated as:

$$C_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \tag{3}$$

where C_t represents the candidate cell status at time t; W_c and U_c are the weight parameters of the LSTM cell; b_c denotes the bias of the input gate; tanh represents the hyperbolic tangent activation function.

Finally, the output gate information is used in the output gate, and the output data of the output gate are calculated using

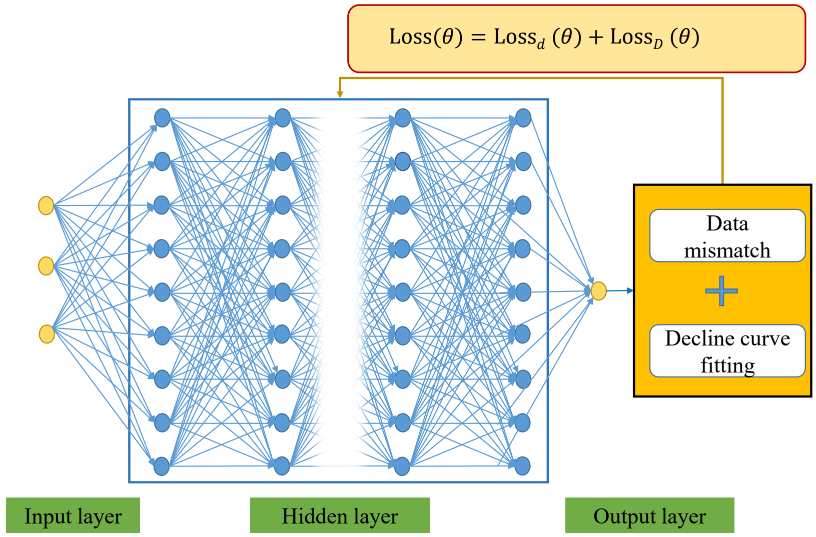


Fig. 2. Framwork of coupling neural network model.

Table 1. DCA models.

Name	Equation
Arps	$q(t)=q_i e^{-D_i t}$
	$q(t)=q_i(1+bD_i t)^{-1/b}$
	$q(t)=q_i(1+D_i t)^{-1}$
SEPD	$q=q_i\exp[-(t/\tau)^{n_{\mathrm{SEPD}}}]$
Duong	$q=q_i t^{-m_D}\exp[a_D(t^{1-m_D}-1)/(1-m_D)]$

the cell state:

$$o_t=\sigma\left(W_o x_t+U_o h_{t-1}+b_o\right) \tag{4}$$

$$h_t=o_t \tanh C_t \tag{5}$$

where o_t represents the output gate information; W_o and U_o are the weight parameters of the output gate; b_o is the bias of the output gate.

As demonstrated above, the finely designed operational structure of the LSTM neural network can well control the information transfer between the historical data and the input data, so that it has a higher prediction ability for time-series data. Therefore, the LSTM neural network is used to learn and extract data features from the time-series data, such that more accurate prediction models can be obtained. Importantly, the LSTM-based deep learning method requires the assumption that the historical production dataset is large and representative enough to sufficiently capture the production fluctuation tendency of the entire time series.

2.2 Production forecasting jointly driven by DCA model and data

The conventional neural network-based approaches can only train the network through data-driven methods, while they do not consider the characteristics of the decline curve governing the production decline process. Therefore, it is proposed to incorporate the well-constructed decline curve models into neural network methods, such as Arps and Duong decline curve models, to improve the training performance of the neural network. The decline curve model plays the role of an empirical verification for controlling the dynamics of gas production or a physically plausible law, and this prior information can serve as the driving condition and guide for the neural network training process. In this way, the training process can be converged rapidly and accurately to its optimal condition. Thus, an implicit coupling neural network driven jointly by both data and DCA models is innovatively used to predict the gas well performance, as shown in Fig. 2.

The construction of such method involves collecting and recording the production dynamic data of gas fields by using various monitoring devices. The production dynamic data can then be organized into a time series dataset, which is subsequently used by the neural network method to capture and extract potential relationships between the dataset segments, thereby endowing the model with predictive capability. The neural network model is established and trained based on production data. The DCA model represents a significant amount of prior knowledge of the empirical gas production decline analysis. By incorporating this knowledge into the data-driven neural network, the prediction performance of the neural network model can be effectively improved, and its interpretability can be enhanced.

Before incorporating the method of DCA into the neural network, it is necessary to select a suitable DCA model to ensure the effectiveness of the joint driving method. In this study, five types of DCA models are selected as the alternative options, which include exponential, hyperbolic, harmonic, stretched-exponential production decline (SEPD), and the Duong DCA models. The calculation formulae of the five DCA models are shown in Table 1. These five models are fitted to the training and testing datasets and based on the fitting performance of these five DCA models. Then, the optimal one is selected and then incorporated as a driving condition into the neural network.

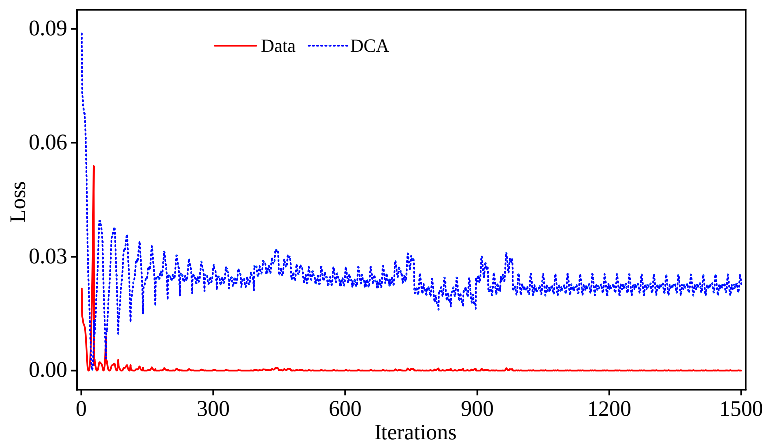


Fig. 3. Loss function values in the model training process.

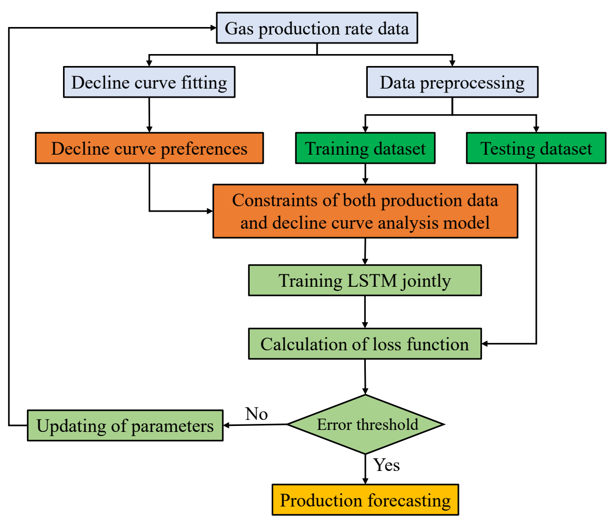


Fig. 4. Deep learning-based gas production prediction procedures.

Suppose that y denotes the real production rate data, y' denotes the output data of the neural network, y'' denotes the fitted value of the DCA model, and f_L denotes the loss function of the neural network. Then, the error Loss_d between the real production data and the output data of the neural network, and the error Loss_D between the real production data and the fitted value of the DCA model, can be calculated respectively.

$$\text{Loss}_d = f_L(y, y') \tag{6}$$

$$\text{Loss}_D = f_L(y, y'') \tag{7}$$

1) According to the error between the output value of the neural network and the target value Loss_d, the error between the output value of the neural network and the fitted value of the DCA model Loss_D, and the loss function of the neural network model by weighting these two parts, can be calculated. The weighting coefficients λ_d, λ_D can be determined by the trial and error method. As a rule of thumb, the weighting coefficients are adjusted to ensure that the values of data loss and DCA loss are on the same order of magnitude, and both losses are taken into account during the training process of neural network (in this case, λ_d is 0.7 and λ_D is 0.3). The total loss function can be expressed by:

$$\text{Loss} = \lambda_d \text{Loss}_d + \lambda_D \text{Loss}_D \tag{8}$$

2) Implement the back-propagation for the weight parameter W based on the loss function calculated by weighting the data loss and the DCA model loss;
3) Calculate the gradient from the back-propagated loss $\partial\text{Loss}/\partial W$;
4) Update the parameters of neural networks: $W_n = W_o - \alpha\partial\text{Loss}/\partial W$, where W_n and W_o represent updated and initial weight values.

In the neural network training process, the relationships between the loss function and the iterations are shown in Fig. 3. It can be inferred from the results that the data loss converges faster than DCA loss in this case. Furthermore, both data and DCA loss eventually decrease to levels that satisfy the convergence stopping criteria.

During the prediction process of gas production data, as shown in Fig. 4, the currently available gas production data are used as the input data. These data are fitted to the decline curve models to select the best-fitting one, which will be used to constrain the training process of the LSTM model. The input data are divided into two subsets: training dataset and testing dataset. The training data are used to construct the loss function of data mismatch, while the testing dataset is used to determine the proper hyper-parameters. The values of hyper-parameters are listed in Table 2. If the data and DCA models are both used to drive the training process, the selected DCA model fitting error and the data mismatch in the training dataset can be used to constrain the training process. The loss function are minimized by adjusting the neural network model parameters. The minimization iteration will be continued until the preset error thresholds are satisfied and the gas production rate can be predicted.

2.3 Data preprocessing

In order to the ensure the representativeness of production data from gas reservoirs over a certain period, an analysis of

Table 2. The values of hyper-parameters used for the LSTM model.

Hyper-parameter	Value
Number of hidden layers	2
Time window size	15
Number of neurons	32
Number of epochs	50
Activation function	Tanh
Optimizer	RMSprop
Learning rate	0.001

the data is conducted to select parameters that may have an impact on the production dynamics as an initial input data for the model. During the production process of gas wells, the dynamic parameters such as production rate and pressure may fluctuate due to various factors. Therefore, it is necessary to pre-process the raw data to remove the noise impact.

Initially, the Savitzky-Golay filter (S-G filter), proposed in 1964, is employed to process the data. This filter is a time-domain-based filtering method that uses the least-squares fitting approach (Press and Teukolsky, 1990). It is mainly used in noise reduction work and can effectively preserve the original features of the data while reducing noise. The basic formula for the S-G filter is as follows:

$$Y_i^* = \frac{\sum\limits_{-m}^{m} (C_i Y_i)}{N} \tag{9}$$

where Y_i and Y_i^* represent the original data and smoothed values, respectively; C_i denotes the S-G polynomial fit coefficient; m represents the filter half-length; N is the length of the filter and takes the value of $2\mathrm{m}+1$.

Firstly, after data denoising, it is necessary to normalize the production history data that contain the dynamic parameters of the gas reservoir. Normalization is a dimensionless data processing method that transforms the physical values of the gas reservoir production system into mathematical relationships; it is an effective way to simplify calculations and reduce the magnitude of values with features. Normalization processing is achieved by a linear transformation of the numerical values. The linear transformation method does not cause data to "fail" after processing but instead improves the performance of the data during the model training process. The calculation formula for normalization is as follows:

$$x_n = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{10}$$

where x_n, $x_{\min}$, $x_{\max}$, and x denote the normalized data, original data minimum, original data maximum, and original data, respectively.

After normalizing the production history data and converting it into sequential data, the production data are transformed into historical production data with a time window size and corresponding known parameter values, which serve as input parameters for the LSTM deep neural network model. The deep neural network algorithm is utilized to calculate the production dynamic data that need to be predicted, as shown in the left-hand side of Fig. 4.

The preprocessed data from the previous step is divided into three parts. The first 70% of the data are used as the training set to train the model. The next 15% of the data are taken as the test set to predict and validate the model's performance. Based on the test set error, an optimization algorithm is used to update the model parameters. The final 15% of the data are used as the prediction set for the model.

The objective of the neural network model established in this study is to utilize historical production dynamic data that contains production and production controlling parameters (such as choke size, bottom hole pressure and daily working time) to predict the future production indicator data (such as production volume).

The normalized production data are transformed, with production control parameters used as feature data and multiple time steps set as a time window. The production data from the current time step corresponding to the last time step in the window are taken as label data. Each time window corresponds to one sample, and different samples are constructed by moving the time window. The data are then transformed into a three-dimensional matrix of input features (sample size $\times$ time steps $\times$ number of features) and a corresponding two-dimensional matrix consisting of production volume (sample size $\times$ 1). These matrices are used as training data for the deep learning neural network model. A LSTM neural network algorithm is employed to predict the production data and the errors between the predictions and the measurements are calculated. Finally, an error back-propagation algorithm is employed in conjunction with a data-driven method based on a decreasing pattern to iteratively train and obtain the production prediction model.

2.4 Model settings and parameter optimization

In order to examine whether there is still room for improvement in the model being trained, the following steps are taken after completing the model training on the training set: First, the fitting effect of the model on the training set and the prediction error of the model on the test set is evaluated. If the performance of the model on both sets is unsatisfactory, the hyper-parameters of the model are adjusted to improve the prediction accuracy. These hyper-parameters include time step, number of neurons, regularization coefficient, etc. Finally, the trained and optimized model is applied to forecast the future production dynamics in the prediction set.

Traditionally, the approach to forecast one or several future time steps is to directly use the actual historical data as input data. However, this forecasting method only predicts a short period in the future and cannot reflect the long-term decreasing trend in production in the forecasting process. During production prediction in the testing and forecasting sets, the gas production data for each step is unknown. Thus, it is impossible to construct the input data for the forecasting set directly. Instead, the forecasting model is utilized to predict the value of the previous time step and construct the input data

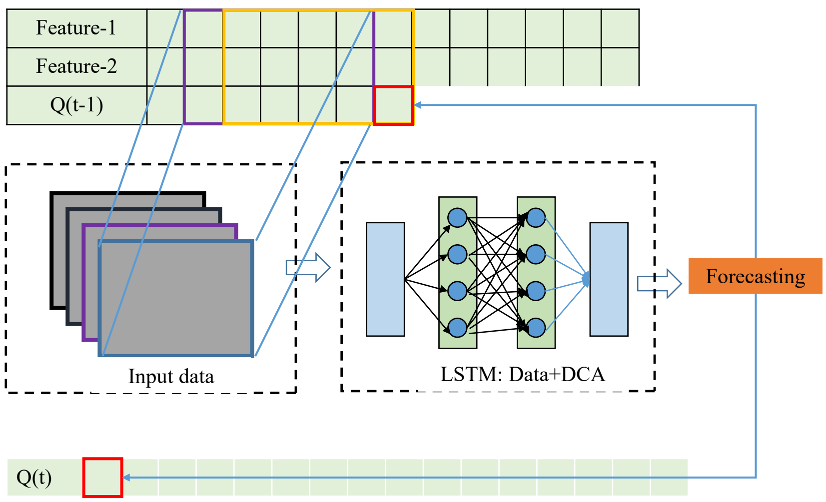


Fig. 5. Prediction dataset construction process and model prediction.

for the current time step. This involves adding the predicted production value of the previous time step to the production time series data and shifting the time window to achieve a cyclic prediction of future production. Finally, the predicted production results for the forecasting set are obtained. The process of cyclic prediction and data reconstruction of the model is illustrated in Fig. 5. During the prediction process, the controlling factors associated with the gas production data, such as production time and choke size setting, can be used as features. The historical feature data and the production data are used as input data for the deep learning LSTM model jointly driven by data and DCA model to predict the gas production rate in the next step. The input data window continues to slide towards the next time step, so that the time series prediction can be achieved, and the gas production rate fluctuation can be predicted as the time series.

3. Results and discussion

In order to demonstrate the predictive capability of the proposed innovative method, field data on gas production are collected from well X1 in a tight gas reservoir and well X2 in a carbonate reservoir in Southwest China. By using the five types of decline curve models mentioned above, the fitting process of the decline curves are performed on the training and testing datasets of well X1 and well X2. The results are illustrated in Fig. 6. Based on the fitting errors of the decline curves in the testing set, the optimal decline curve model is selected as the driving condition for the neural network models.

In Fig. 6, it can be observed that all five decline curve formulae exhibit satisfactory fitting results on well X1 and well X2. As the fitting results of the testing set have a significant impact on the prediction performance of the LSTM neural network model jointly driven by the DCA model and data on the prediction set, it is necessary to select the DCA model with the smallest prediction error on the testing set. To confirm which DCA model has the best fitting performance on the testing set, the mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square error (RMSE) of these five models are calculated as follows:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{11}$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{y_i} \times 100\% \tag{12}$$

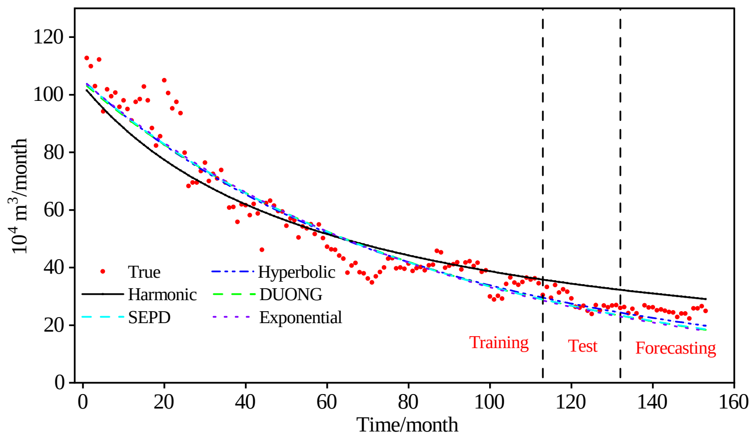
$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{13}$$

where y_i denotes the i^{th} true value, and $\hat{y}_i$ denotes the i^{th} prediction.

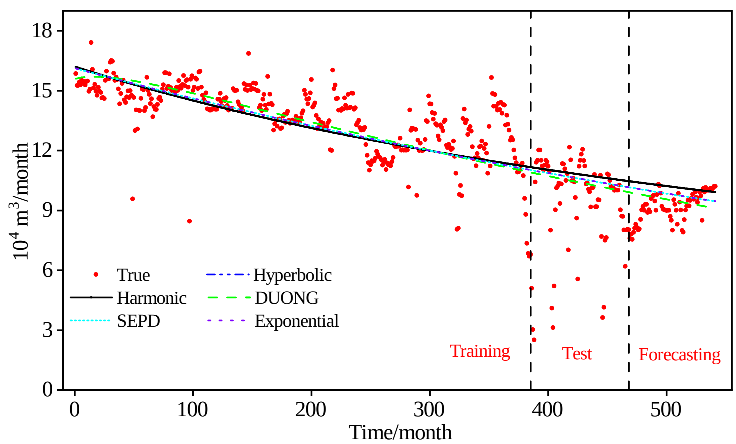
The results are shown in Table 3. As can be seen from the table, the hyperbolic DCA model best fits the actual production curve for well X1, and the Duong DCA model best fits the actual production curve for well X2. Therefore, the hyperbolic and Duong DCA models are incorporated into the neural network as the driving condition for the performance predictions of gas wells X1 and X2.

Next, the LSTM models driven jointly by production data and DCA model are trained and used for forecasting, and their performances are compared with that of the LSTM neural network models without the DCA models. The input data in the data-driven LSTM and jointly driven LSTM models are the historical gas production observation data. Fig. 7 illustrates the predictive performances for gas production rate in wells X1 and X2 when using the purely data-driven and the jointly driven LSTM models. As shown in Fig. 7, the DCA models alter the production trends predicted by the data-driven LSTM model on the testing set, making them more consistent with the production situation of wells X1 and X2.

Table 4 shows the performances of the two types of LSTM models with and without the constraints of DCA models. The accuracy evaluations are measured by the MAE, MAPE, and RMSE that occur in predicting the gas well production performance on the testing set. As seen in Table 4, after incorporating the DCA model into the LSTM model driven by data, the average absolute error decreases by 0.8845,



(a)



(b)

Fig. 6. Prediction results of five DCA models for wells X1 (a) and X2 (b).

Table 3. Fitting error of DCA models on the testing dataset.

Well	Predictive Models	MAE	MAPE (%)	RMSE
X1	Exponential	2.6761	9.01	3.2341
	Hyperbolic	1.9686	6.66	2.4313
	Harmonic	5.7199	21.16	6.1448
	Duong	2.3852	8.02	2.9368
	SEPD	2.3663	7.95	2.9154
X2	Exponential	1.6957	36.49	2.7937
	Hyperbolic	1.7828	38.17	2.9294
	Harmonic	1.7818	38.07	2.8994
	Duong	1.6461	35.37	2.7148
	SEPD	1.6957	36.49	2.7937

the relative percentage error decreases by 3.29%, and the root mean square error decreases by 0.9928 for well X1. From the perspective of the mean absolute error analysis, the LSTM model jointly driven by the production data and DCA model yields a 42.90% reduction for well X1 in the average absolute error compared to the LSTM model driven purely by production data. The average absolute error decreases by 0.093, the relative percentage error decreases by 1.44%, and the root mean square error decreases by 0.1138 for well X2. From the perspective of the mean absolute error analysis, the

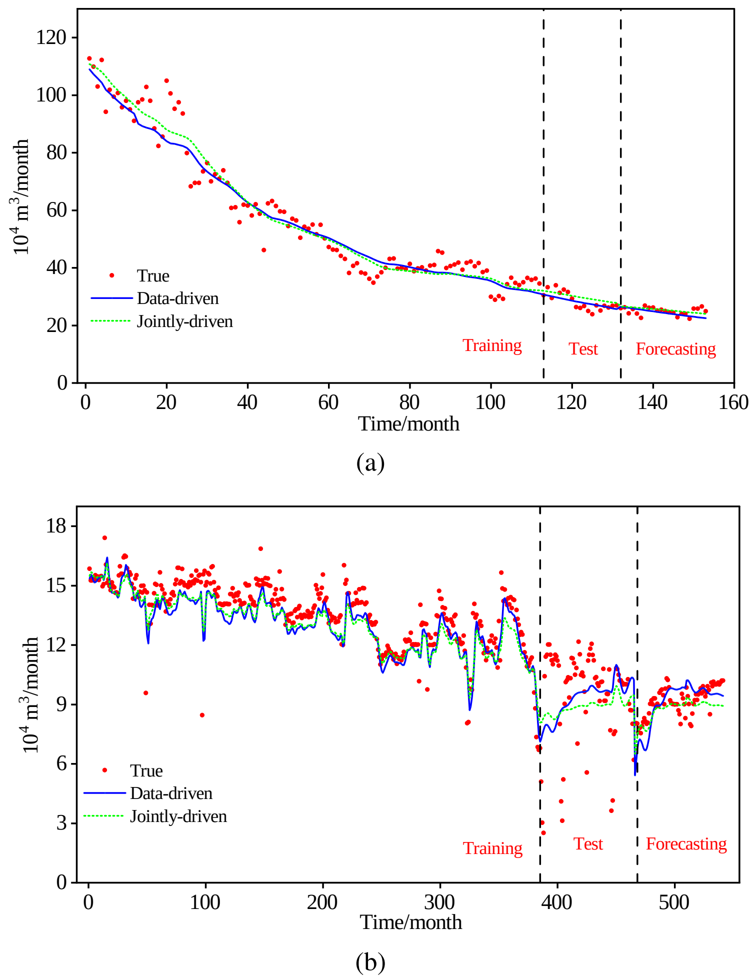


Fig. 7. Comparisons of prediction results based on data-driven and jointly driven LSTM models for wells X1 (a) and X2 (b).

Table 4. Errors of the data-driven and jointly-driven LSTM models in the prediction dataset.

Well	Models	MAE	MAPE (%)	RMSE
X1	Data-driven LSTM	2.0618	8.14	2.4720
	Jointly-driven LSTM	1.1773	4.85	1.4792
X2	Data-driven LSTM	0.6811	7.72	0.8258
	Jointly-driven LSTM	0.5881	6.28	0.7120

LSTM model jointly driven by the production data and DCA model results in a 13.65% reduction for well X2 in the average absolute error compared to the LSTM model driven purely by production data. Therefore, the proposed jointly driven approach significantly improves the ability of the LSTM model to predict the gas well production performance on the testing set and enhances its interpretability.

4. Conclusions

This study presents an interpretable and accurate time series neural network model jointly driven by production data and DCA model to forecast the gas well production performance. The following conclusions can be drawn:

1) The proposed method can incorporate prior knowledge related to gas production forecasting, including Arps, Duong and extended exponential DCA model, into the training process of the LSTM neural network. This knowledge is used as the driving force for neural network training, which greatly improves the conventional approach in which the neural network is trained only through data-driven methods. Consequently, a deep learning model jointly driven by both DCA model and production data can be innovatively established.
2) The forecasting performances of the purely data-driven deep learning model and the jointly driven deep learning model are compared with the conventional DCA

methods. The results show that the purely data-driven and jointly driven deep learning models have smaller prediction errors than the conventional decline curve methods, indicating that these deep learning models have better capability to extract the time dependence features from the gas well production data.

3) The methods jointly driven by DCA model and data are introduced as the driving conditions for the loss function of the LSTM neural network training, which improves the interpretability and prediction ability of gas well performance. In a practical example, deep learning models with both DCA model and data-driven training reduce the mean absolute error by 42.90% for the tight gas well and 13.65% for the carbonate gas well compared to data-driven deep learning models, which verifies the effectiveness of this training method in improving the accuracy of the deep learning-based prediction of gas well production performance.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 52274048), the Beijing Natural Science Foundation (No. 3222037), the Science Foundation of China University of Petroleum, Beijing (Nos. 2462022QNXZ002 and 2462021YXZZ010), and the CNPC 14th five-year perspective fundamental research project (No. 2021DJ2104).

Additional information: Author's email

felix.nanguka@tpdc.co.tz (F. Nanguka).

Conflict of interest

The authors declare no competing interest.

References

Almajid, M. M., Abu-Al-Saud, M. O. Prediction of porous media fluid flow using physics informed neural networks. Journal of Petroleum Science and Engineering, 2022, 208: 109205.

Arps, J. J. Analysis of decline curves. Transactions of the AIME, 1945, 160(1): 228-247.

Baydin, A. G., Pearlmutter, B. A., Radul, A. A., et al. Automatic differentiation in machine learning: A survey. Journal of Marchine Learning Research, 2018, 18: 1-43.

Blasingame, T. A., Poe, B. D. Semianalytic solutions for a well with a single finite-conductivity vertical fracture. Paper SPE 26424 Presented at the SPE Annual Technical Conference and Exhibition, Houston, Texas, 3-6 October, 1993.

Chi, D. Research on electricity consumption forecasting model based on wavelet transform and multi-layer LSTM model. Energy Reports, 2022, 8: 220-228.

Chung, W. H., Gu, Y. H., Yoo, S. J. District heater load forecasting based on machine learning and parallel CNN-LSTM attention. Energy, 2022, 246: 123350.

Cuomo, S., Di Cola, V. S., Giampaolo, F., et al. Scientific machine learning through physics-informed neural networks: Where we are and what's next. Journal of Scientific Computing, 2022, 92(3): 88.

Duong, A. N. Rate-decline analysis for fracture-dominated shale reservoirs. SPE Reservoir Evaluation & Engineering, 2011, 14(3): 377-387.

Fan, D., Sun, H., Yao, J., et al. Well production forecasting based on ARIMA-LSTM model considering manual operations. Energy, 2021, 220: 119708.

Fetkovich, M. J. Decline curve analysis using type curves. Paper SPE 4629 Presented at the Fall Meeting of the Society of Petroleum Engineers of AIME, Las Vegas, Nevada, 30 September-3 October, 1973.

Fraces, C. G., Hamdi, T. Physics informed deep learning for flow and transport in porous media. Paper SPE 203934 Presented at the SPE Reservoir Simulation Conference, On-Demand, 4-6 October, 2021.

Frooqnia, A., Torres-Verdín, C., Sepehrnoori, K. Numerical simulation and interpretation of production logging measurements using a new coupled wellbore-reservoir model. Paper SPWLA 2011 presented at the SPWLA 52nd Annual Logging Symposium, Colorado Springs, Colorado, 14-18 May, 2011.

Gers, F. A., Schmidhuber, J., Cummins, F. Learning to forget: Continual prediction with LSTM. Neural Computation, 2000, 12(10): 2451-2471.

Jia, X., Zhang, F. Applying data-driven method to production decline analysis and forecasting. Paper SPE 181616 presented at the SPE Annual Technical Conference and Exhibition, Dubai, UAE, 26-28 September, 2016.

Karasu, S., Altan, A. Crude oil time series prediction model based on LSTM network with chaotic Henry gas solubility optimization. Energy, 2022, 242: 122964.

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., et al. Physics-informed machine learning. Nature Reviews Physics, 2021, 3(6): 422-440.

Kuzma, H. A., Arora, N. S., Farid, K. Generative models for production forecasting in unconventional oil and gas plays. Paper URTEC 1928595 Presented at the SPE/AAPG/SEG Unconventional Resources Technology Conference, Denver, Colorado, 25-27 August, 2014.

Li, W., Wang, L., Dong, Z., et al. Reservoir production prediction with optimized artificial neural network and time series approaches. Journal of Petroleum Science and Engineering, 2022a, 215: 110586.

Li, X., Xiao, K., Li, X., et al. A well rate prediction method based on LSTM algorithm considering manual operations. Journal of Petroleum Science and Engineering, 2022b, 210: 110047.

Mollaiy, B. S., Shahbazian, M. An imperialist competitive algorithm artificial neural network method to predict oil flow rate of the wells. International Journal of Computer Applications, 2011, 26(10): 47-50.

Muther, T., Dahaghi, A. K., Syed, F. I., et al. Physical laws meet machine intelligence: Current developments and future directions. Artificial Intelligence Review, 2023, 56(7): 6947-7013.

Ning, Y., Kazemi, H., Tahmasebi, P. A comparative machine learning study for time series oil production forecasting: ARIMA, LSTM, and Prophet. Computers & Geosciences, 2022, 164: 105126.

Press, W. H., Teukolsky, S. A. Savitzky-Golay smoothing filters. Computers in Physics, 1990, 4(6): 669-672.

Raissi, M., Perdikaris, P., Karniadakis, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 2019, 378: 686-707.

Rojc, M., Mlakar, I. An LSTM-based model for the compression of acoustic inventories for corpus-based text-to-speech synthesis systems. Computers and Electrical Engineering, 2022, 100: 107942.

Sagheer, A., Kotb, M. Time series forecasting of petroleum production using deep LSTM recurrent networks. Neurocomputing, 2019, 323: 203-213.

Shabro, V., Torres-Verdin, C., Javadpour, F. Numerical simulation of shale-gas production: From pore-scale modeling of slip-flow, Knudsen diffusion, and Langmuir desorption to reservoir modeling of compressible fluid. Paper SPE 144355 Presented at the North American Unconventional Gas Conference and Exhibition, Woodlands, Texas, 14-16 June, 2011.

Song X., Liu, Y., Xue, L., et al. Time-series well performance prediction based on Long Short-Term Memory (LSTM) neural network model. Journal of Petroleum Science and Engineering, 2020, 186: 106682.

Temizel, C., Canbaz, C. H., Saracoglu, O., et al. Production forecasting in shale reservoirs through conventional DCA and machine/deep learning methods. Paper URTEC 20202878 Presented at the Unconventional Resources Technology Conference held in Denver, Colorado, 20-22 July, 2020.

Valkó, P. P., Lee, W. J. A better way to forecast production from unconventional gas wells. Paper SPE 134231 Presented at the SPE Annual Technical Conference and Exhibition, Florence, Italy, 20-22 September, 2010.

Wang, N., Zhang, D., Chang, H., et al. Deep learning of subsurface flow via theory-guided neural network. Journal of Hydrology, 2020, 584: 124700.

Werneck, R. O., Prates, R., Moura, R., et al. Data-driven deep-learning forecasting for oil production and pressure. Journal of Petroleum Science and Engineering, 2022, 210: 109937.

Zha, W., Liu, Y., Wan, Y., et al. Forecasting monthly gas field production based on the CNN-LSTM model. Energy, 2022, 260: 124889.