Artificial Intelligence in Geosciences 2 (2021) 165–170

Contents lists available at ScienceDirect

Artificial Intelligence in Geosciences

journal homepage: www.keaipublishing.com/en/journals/artificial-intelligence-in-geosciences

Data-driven approaches for time series prediction of daily production in the Sulige tight gas field, China

Qi Zhang [a], Ziwei Chen [b], Yuan Zeng [a], Hang Gao [a], Qiansheng Wei [a], Tiaoyu Luo [b], Zhiguo Wang [b,*]

[a] The Third Gas Production Plant of PetroChina Changqing Oilfield Branch, Wushenqi, Inner Mongolia, 017000, PR China
[b] School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, Shaanxi, 710049, PR China

ARTICLE INFO

Keywords:
Prediction of production time series
Long short-term memory neural network
Random forest
Support vector machine

ABSTRACT

The Sulige tight gas field is presently the largest gas field in China. Owing to the ultralow permeability and strong heterogeneity of the reservoirs in Sulige, the number of production wells has exceeded 3,000, keeping the stable gas supply in the decade. Thus, the daily production prediction of gas wells is significant for monitoring production and for implementing and evaluating stimulation measures. Therefore, on the basis of the three data-driven time series approaches, the daily production of 1692 wells over 10 years was mining for the daily production prediction of wells in Sulige. The jointed deep long short-term memory and fully connected neural network (DLSTM-FNN) model was proposed by introducing the recurrent neural network's sequential expression ability and was compared with random forest (RF) and support vector regression (SVR). After the daily production predictions of thousands of wells in Sulige, the proposed DLSTM-FNN model significantly improved the time series prediction accuracy and efficiency in the short training samples and had strong availability and practicability in the Sulige tight gas field.

1. Introduction

The Sulige gas field is located northwest of the Ordos Basin, China. Since 2011, it has become China's largest gas field (H. Yang et al., 2012). Owing to sedimentation, diagenesis, and accumulation processes (T. Yang et al., 2012; Zou et al., 2012; Jiang et al., 2007), the Sulige gas field reservoirs show ultralow permeability and strong heterogeneity, resulting in the low production of wells. Hence, in the last decade, the number of wells in the Sulige gas field had to be increased rapidly for the continuous expansion of productivity. Subsequently, liquid-accumulation, intermittent, and low production wells are gradually increasing, making the production management of the entire gas field more complicated. Therefore, it is crucial to explore the production rules of various gas wells in different stages, especially predicting production dynamics and monitoring production anomalies for the whole life cycle management of gas wells, which helps in maintaining long-term stable production and supply of natural gas.

The commonly used production prediction methods of gas wells can be divided into numerical simulation based on the physical model and production decline curve analysis based on the statistical law.

Reservoir numerical simulation utilizes computer models to estimate the fluids dynamics in porous media. The numerical simulation requires a large amount of static geological and dynamic production data to achieve accurate geological modeling and high-quality history matching. Therefore, the dynamic simulation takes a lot of time, and there are deviations between the predicted results and the actual gas field production results.

In reservoir engineering, many decline curve analysis models have been proposed to predict oil and gas production. As early as 1945, Arps proposed a classical exponential decline model, hyperbolic decline model, and harmonic decline model (Arps, 1945). Chen et al. proposed a generalized decline model, linear decline model, and pan-exponential decline model in China (Chen and Tang, 2016; Chen and Zhou, 2015; Chen and Fu, 2019). However, it is still challenging to select the parameters for tight sandstone gas in the abovementioned decline models, thus limiting its application in production dynamic prediction.

Recently, on the basis of the available production data, various machine learning methods, especially neural networks, have been successfully used to select the featured parameters that affect well production and then predict oil and gas well production under a supervised learning framework (Cao et al., 2016; Hou, 2019; Liu et al., 2019, 2020; Gu et al.,

* Corresponding author.
E-mail address: emailwzg@gmail.com (Z. Wang).

https://doi.org/10.1016/j.aiig.2022.02.005
Received 15 December 2021; Received in revised form 25 February 2022; Accepted 26 February 2022
Available online 3 March 2022

2019).

However, in the abovementioned oil and gas well production prediction methods, whether physical model-driven or data-driven algorithms, they focused more on medium-to-long-term production prediction in the unit of month or year and were less involved in time series prediction of a single well's daily production and abnormal monitoring in the whole life cycle of a tight gas field (Dong and Yang, 2009; Sagheer and Kotb, 2019; Jiang et al., 2021). Therefore, considering the current and past statuses of daily production time series, we proposed a deep network prediction method based on long short-term memory (LSTM) neural network and further compared it with traditional data-driven machine learning approaches such as random forest (RF) and support vector regression (SVR).

2. Data-driven approaches

2.1. Random forest and support vector regression

A decision tree (DT) is a supervised learning algorithm that selects the best decision scheme by constructing the decision process and calculating the corresponding mathematical expectation under the condition of known probabilities of all events. It can be expressed as a tree structure, in which the internal nodes represent different levels of decision schemes and the leaf nodes (endpoints) represent the final decision results. However, they are prone to overfitting. Therefore, ensemble learning methods are often used to replace a single DT, that is, RF (Breiman, 2001). Assuming that the prediction results of each DT are good, there is overfitting for different datasets. The RF notes that by constructing many DTs and averaging their prediction results, the robustness of the DTs can be improved while maintaining their prediction ability. The RF algorithm first conducts random sampling with replacement of the dataset to obtain several new datasets to create as many different DTs as possible. For these newly created datasets, each DT is constructed, and the weighted average is conducted according to the different importance of features to obtain the prediction results of the RF.

The model function of SVR is a linear function. The SVR constructs an interval band on both sides of the linear function. The loss is not calculated for the data samples falling into the interval band but for the samples located outside the interval band, and it is calculated according to a specific loss function. The SVR also introduces the kernel function method from the support vector machine to expand the algorithm to higher dimensional data space. There are many kinds of kernel functions. In this paper, the Gaussian kernel was used to map data samples x_1 and x_2 into high-dimensional space, which is defined as

$$k_{rbf}(x_1, x_2) = \exp\left(-\gamma ||x_1 - x_2||^2\right), \tag{1}$$

where γ is the parameter controlling the width of the Gaussian kernel.

2.2. Recurrent neural network

For time series such as the daily production of gas wells, the recurrent neural network (RNN) (Hopfield, 1982) was introduced to capture the correlation between the current and past moments. The RNN model has the following characteristics: the output at one moment will be a part of the input at the next moment; for the multilayer stacked RNN, the parameters of different layers can be shared, thus reducing the number of training parameters; and the length of the input data can be variable.

Fig. 1 shows the basic structure of RNN. Let the input data $\boldsymbol{X} = [\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N-1}]$. A represents a structural neuron unit, $\boldsymbol{x}_t$ represents the input of the t-th network, $\boldsymbol{o}_t$ represents the output of the t-th network, and $\boldsymbol{h}_t$ represents the hidden state of the output of the t-th network. The input–output relationship can be expressed as

$$\begin{cases} h_t = \sigma(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \\ o_t = W_{hy}h_{t-1} + b_o \end{cases}, \tag{2}$$

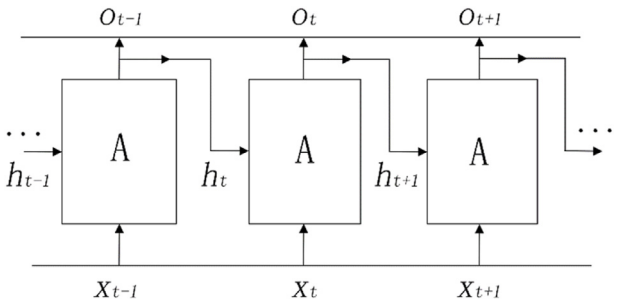

Fig. 1. The architecture of a recurrent neural network.

where σ is the activation function, $\boldsymbol{W}_{xh}$ is the weight matrix applied to $\boldsymbol{x}_t$ to compute the hidden state $\boldsymbol{h}_t$, $\boldsymbol{W}_{hh}$ is the weight matrix applied to the current hidden state $\boldsymbol{h}_t$ to calculate the next hidden state $\boldsymbol{h}_{t-1}$, $\boldsymbol{W}_{hy}$ is the weight matrix applied to the hidden state $\boldsymbol{h}_t$ to calculate the output $\boldsymbol{o}_t$, and $\boldsymbol{b}_h$ and $\boldsymbol{b}_o$ are offset items belonging to the hidden state and output, respectively. According to equation (2), $\boldsymbol{h}_t$ from any layer of the network can affect the subsequent output, which has strong practical significance for processing time series.

However, the RNN networks are challenging to achieve long-term transmission of the input. When the back-propagation mechanism is adopted to update the weights in the RNN, the gradient of the latter layer will be multiplied by that of the former network layer as the number of network layers increases gradually. Hence, the RNN is prone to gradient explosion/disappearance (Bengio et al, 1994). To solve the problem of long-term dependencies, many network structures have been proposed, such as LSTM neural networks and gated recurrent units.

2.3. LSTM neural network

LSTM is one of the most widely used RNN structures (Hochreiter and Schmidhuber, 1997). Compared with the traditional RNN structure, the LSTM adds memory and gating units on the original basis, which can effectively avoid the disappearance/explosion of the gradient and can extract the long-term correlation that may exist in the data. Fig. 2 depicts a schematic of an LSTM unit. The LSTM unit introduces three types of gating units, which are the input, forget, and output gates. The general equation of the input and output of the gating unit can be expressed as

$$\boldsymbol{y} = \sigma(\boldsymbol{W}\boldsymbol{x} + \boldsymbol{b}), \tag{3}$$

where $\boldsymbol{W}$ and $\boldsymbol{b}$ are the weight matrix and bias of the gating unit, respectively, and σ is the activation function. In the LSTM, the sigmoid function is generally used, and its expression is

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{4}$$

At time t, the input data of the LSTM are denoted by $\boldsymbol{x}_t$, the output data are denoted by $\boldsymbol{h}_t$, the memory unit is $\boldsymbol{c}_t$, ⊗ represents the

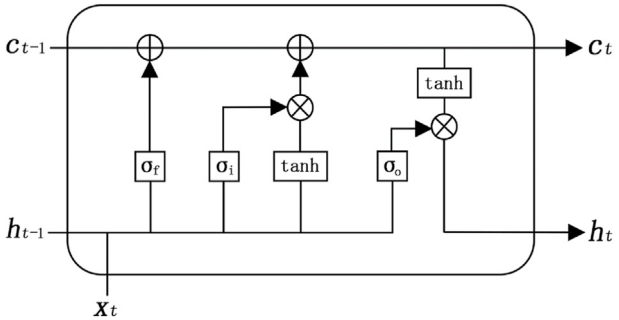

Fig. 2. A block of the long short-term memory.

multiplication of the corresponding bit elements, and the input gate $\boldsymbol{i}$ is used to determine how much input data $\boldsymbol{x}_t$ are allocated to the memory unit $\boldsymbol{c}_t$. The forward propagation process is

$$\boldsymbol{i}_t = \sigma\left(\boldsymbol{W}_{xi}\boldsymbol{x}_t + \boldsymbol{W}_{hf}\boldsymbol{h}_{t-1} + \boldsymbol{b}_i\right) \tag{5}$$

Considering some features in the time series may not be meaningful for the target task. Thus, the forgetting gate is introduced to reduce the influence of these features at subsequent time $t+n$. The forgetting gate $\boldsymbol{f}$ mainly selects the discarded sample features by changing the memory unit $\boldsymbol{c}_t$ as follows:

$$\boldsymbol{f}_t = \sigma\left(\boldsymbol{W}_{xf}\boldsymbol{x}_t + \boldsymbol{W}_{hi}\boldsymbol{h}_{t-1} + \boldsymbol{b}_f\right), \tag{6}$$

$$c_t = f_t \otimes c_{t-1} + i_t \otimes \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \tag{7}$$

The output gate $\boldsymbol{o}$ determines the output data $\boldsymbol{h}_t$ of the network through the memory unit $\boldsymbol{c}_t$ and the input data $\boldsymbol{x}_t$. At this time, the output $\boldsymbol{h}_t$ of the network takes into account both the current input $\boldsymbol{x}_t$ and the content learned by the previous network (that is, the memory unit $\boldsymbol{c}_{t-1}$), and its expression is

$$\boldsymbol{o}_t = \sigma(\boldsymbol{W}_{xo}\boldsymbol{x}_t + \boldsymbol{W}_{ho}\boldsymbol{h}_{t-1} + \boldsymbol{b}_o), \tag{8}$$

$$h_t = o_t \otimes \tanh(c_t) \tag{9}$$

In summary, for a sample s_i of the training set, its data and label are x_i and y_i, respectively. In this paper, the architecture of a deep network based on the deep LSTM and a fully connected network, that is, DLSTM-FNN, is finally proposed, as shown in Fig. 3. The forward propagation expression is

$$\begin{cases} f_1 = RNN(x_i) \\ f_2 = \mathrm{Re}lu(f_1) \\ f_3 = Dropout(f_2). \\ f_4 = Linear(f_3) \\ f_5 = Pool(f_4) \end{cases} \tag{10}$$

In equation (10), RNN represents the recurrent neural network. In this study, we selected the LSTM unit, where the depth of hidden layers was 30, which means the deep LSTM (DLSTM). *Relu* represents the linear rectification activation function, which is used to enhance the nonlinear expression capability of the network. *Dropout* represents the drop layer, and the inactivation rate was set to 0.5, which aims to alleviate the overfitting problem of the model. *Linear* and *Pool* represent the full connection layer (FNN) and the average pooling layer, respectively, to make the output f_5 have the same shape as the label y_i. The root mean square error was used as the loss function, and adaptive moment estimation (*Adam*) was used as the weight optimization algorithm.

3. Time series prediction

The reservoir indicators of gas wells generally include static geological and dynamic production data. Static geological data refer to wellhead position coordinates, depth, resistivity, sonic interval transit time, rock density, compensated neutrons, shale content, gas layer thickness, total porosity, permeability, and hydrocarbon saturation. Dynamic production data mainly refer to the daily gas production, gas pressure, casing pressure, and production time of a gas well under different production dates. To facilitate the comparison of different methods, this study ignores the remaining dynamic production characteristics and only considers the daily gas production time series. Next, the RF, SVR, and DLSTM-FNN models were applied to dynamically predict the daily gas production time series for 1692 tight gas wells in the Sulige gas field.

3.1. Sample set construction

In the daily production time series of 1692 tight gas wells in Sulige, the earliest production date started on November 5, 2006, and the longest production time was 5405 days as of August 28, 2021. Let the daily gas production of gas well v_i be $data_i$. Owing to the different production times and durations of each gas well, the length of $data_i$ in each gas well was not completely consistent. To enable large quantities of gas well data to be inputted into the DLSTM-FNN for training and prediction simultaneously, all $data_i$ must have the same length. One method is to fill in with zero values, that is, to make the value of $data_i$ at all nonmining times be zero. Another method is zero removal, that is, delete the zero values existing in $data_i$ and splice it. After splicing, the timing sequence length of the dataset is the minimum length of $data_i$ after zero removal. However, both methods have advantages and disadvantages. The zero-filling method retains the consistency of the daily gas production of each gas well in the production time, but it will result in a considerable number of zeros in the dataset, which reduces the convergence speed of the neural network. The zero removal method reduces the difficulty of neural network training because it retains effective production data, but it leads to inconsistencies in the corresponding production times of different gas wells. Additionally, because the effective production data of individual gas wells are relatively small, the time sequence length of the combined dataset will be very low, limiting the expressive ability of the neural network.

Therefore, if the zero-filling method is adopted to construct dataset $\boldsymbol{D}$ for all gas wells, its dimension $\boldsymbol{D}_{shape}$ is $(1692, 5405)$. In this study, it was unnecessary to consider all gas wells. Supposing the number of gas wells to be considered is n, for the zero-filling method, $\boldsymbol{D}_{shape}$ is equal to $(n, 5405)$; for the zero removal method, $\boldsymbol{D}_{shape}$ is equal to (n, l_{min}); and l_{min} is the minimum length after zero removal of $data_i$.

Let the dimension of the time series $\boldsymbol{D}$ be (n, l). First, $\boldsymbol{D}$ was divided

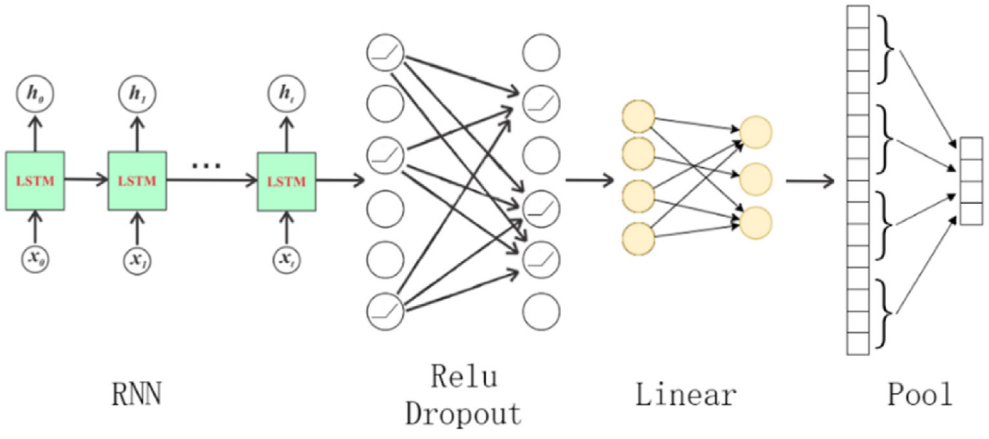

Fig. 3. The schematic of the DLSTM-FNN architecture.

into the training set $\boldsymbol{D}^{train}$ and the test set $\boldsymbol{D}^{test}$. That is, for a time series of length l, assign the first l_{train} information of $\boldsymbol{D}$ to $\boldsymbol{D}^{train}$ and the last l_{test} information to $\boldsymbol{D}^{test}$. Next, construct input data x and input label y for $\boldsymbol{D}^{train}$ and $\boldsymbol{D}^{test}$, respectively. At time t, we hope to use the sequence x for some time before t to predict the sequence y for some time after t. Assuming that the sequence length of x is l_x, the sequence length of y is l_y, and the sample formed by a pair of x and y at time t is s_t, the sample s_{t+1} can be generated at time $t+1$, as shown in Fig. 4. Following Fig. 4, for the training set $\boldsymbol{D}^{train}$, the sample dataset $\boldsymbol{S}^{train}$ can be constructed as $[s_0, s_1, \ldots, s_{m^{train}}]$, and the number of samples m^{train} is $l_{train} - l_x - l_y + 1$. Similarly, for the test set $\boldsymbol{D}^{test}$, $\boldsymbol{S}^{test}$ can be constructed as $[s_0, s_1, \ldots, s_{m^{test}}]$, and the number of test sets m^{test} is $l_{test} - l_x - l_y + 1$. If $\boldsymbol{S}^{train}$ and $\boldsymbol{S}^{test}$ are divided into training set data x^{train}, training set label y^{train}, test set data x^{test}, and test set label y^{test}, their dimensions are (m^{train}, l_x, n), (m^{train}, l_y, n), (m^{test}, l_x, n), and (m^{test}, l_y, n), respectively.

3.2. Result analysis

Using the information of 1692 gas wells in Sulige as the dataset, the RF, SVR, and DLSTM-FNN models were constructed. First, we reduced the dimensions of x and y to (m, l_x) and (m, l_y), respectively. Then, model training was performed on the basis of x^{train} and y^{train}. The length of the training sample was only 10% of the total sample data, which belonged to the short training sample set. On the basis of the data-driven approaches, we used the trained model to predict x^{train} and x^{test} to obtain

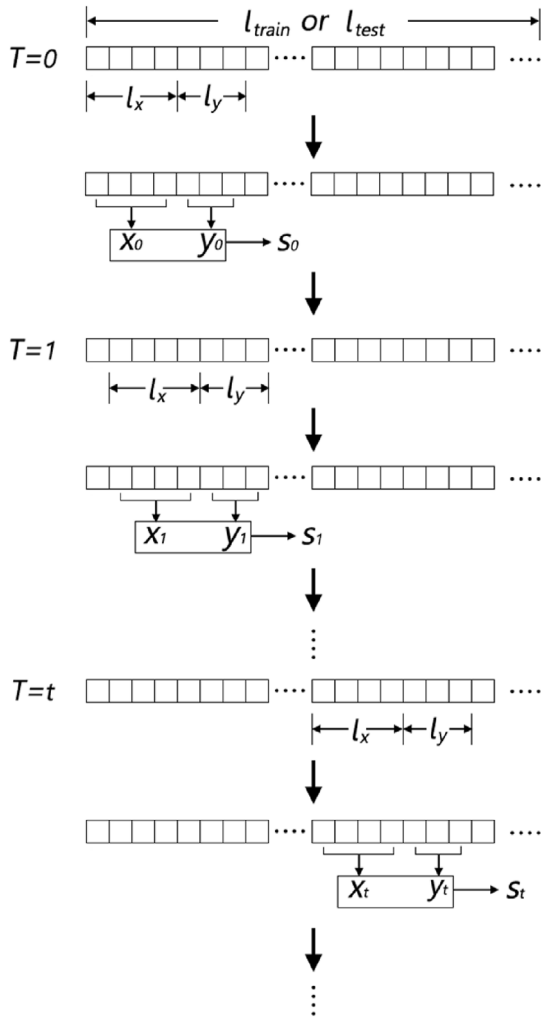

Fig. 4. The sample composition of the daily production time series.

y^{train}_{pred} and y^{test}_{pred}, and then, we compared them with the observed values y^{train} and y^{test} to analyze regression accuracy. As a deep learning model, the DLSTM-FNN needs to balance prediction accuracy and computational time. By constant optimizations of the network hyperparameters, finally, the shortest sequence length l_x was set to 1, the label sequence length l_y was 1, the learning rate was 0.0001, the weight attenuation was 0.0005, and the number of iterations was 150. Under optimal hyperparameters, the average predicting speed of the DLSTM-FNN model was 67.59 s per well, which can fulfill the real-time requirement of the actual production abnormality monitoring in Sulige. Here the root mean square error (RMSE) was used as an accuracy measurement, which can be given as follows

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - y^i_{pred}\right)^2}, \tag{11}$$

where y_i is the current observed value and y^i_{pred} is its predicted value. For the 1692 gas wells in the Sulige gas field, Table 1 compares the RMSE of the three gas well production dynamic prediction models: the RF, SVR, and DLSTM-FNN on the training set and testing set.

Although the RF had the smallest error in the short training sample set, which was 0.0884904, it had the largest error in the test set, which was 0.1878669. The DLSTM-FNN had the smallest error in the testing set, which was 0.1378165 and 0.03 less than the SVR prediction error. Additionally, its prediction was the best.

To further analyze the life cycle of the gas well, typical gas wells S1 and S2 that have produced for 3922 and 1145 days, respectively, were taken as examples. Fig. 5 shows the gas well production prediction results of the three data-driven models, and Table 2 shows their corresponding RMSE. For gas wells S1 and S2, DLSTM-FNN had the smallest RMSE on the testing set, which was 0.0722189 and 0.1025749, respectively. By comparing Fig. 5a, c, and 5e, the DLSTM-FNN showed more accurately predicted production trends and mutation locations in the measured continuous production period (1000–1500 days), the intermittent production period (1500–2000 days), and the economic inefficiency period (>2000 days) of gas well S1. During the 1145-day production cycle, the daily production of gas well S2 was zero due to multiple shutdowns, and the daily production fluctuated sharply because of multiple production stimulation measures. In the prediction results of gas well S2 (Fig. 5b, d, and 5f), the DLSTM-FNN model better predicted the locations of zero daily production abnormalities and severe fluctuations (Fig. 5f). Therefore, compared with the traditional RF and SVR, the data-driven DLSTM-FNN model had the advantages of short training sample sets, quick speed of large-scale predictions, and more accurate results. Furthermore, the prediction of DLSTM-FNN is expected to be used for monitoring production anomalies, for providing a specific time window, and for a quantitated production evaluation for the stimulation measures.

4. Conclusion

This paper studied the daily production prediction of the Sulige tight gas wells. Three data-driven approaches of time series prediction, including the RF, SVR, and LSTM, were applied. Considering the potential information of current and past statuses in the time series, we proposed an optimal set of the DLSTM-FNN hyperparameters suitable for the daily production prediction of tight gas wells.

The availability and practicability of the three data-driven models in

Table 1
The RMSE of the production prediction models for 1692 gas wells in Sulige.

Model	Training Set	Testing Set
RF	0.0884904	0.1878669
SVR	0.1409583	0.1636799
DLSTM-FNN	0.1663612	0.1378165

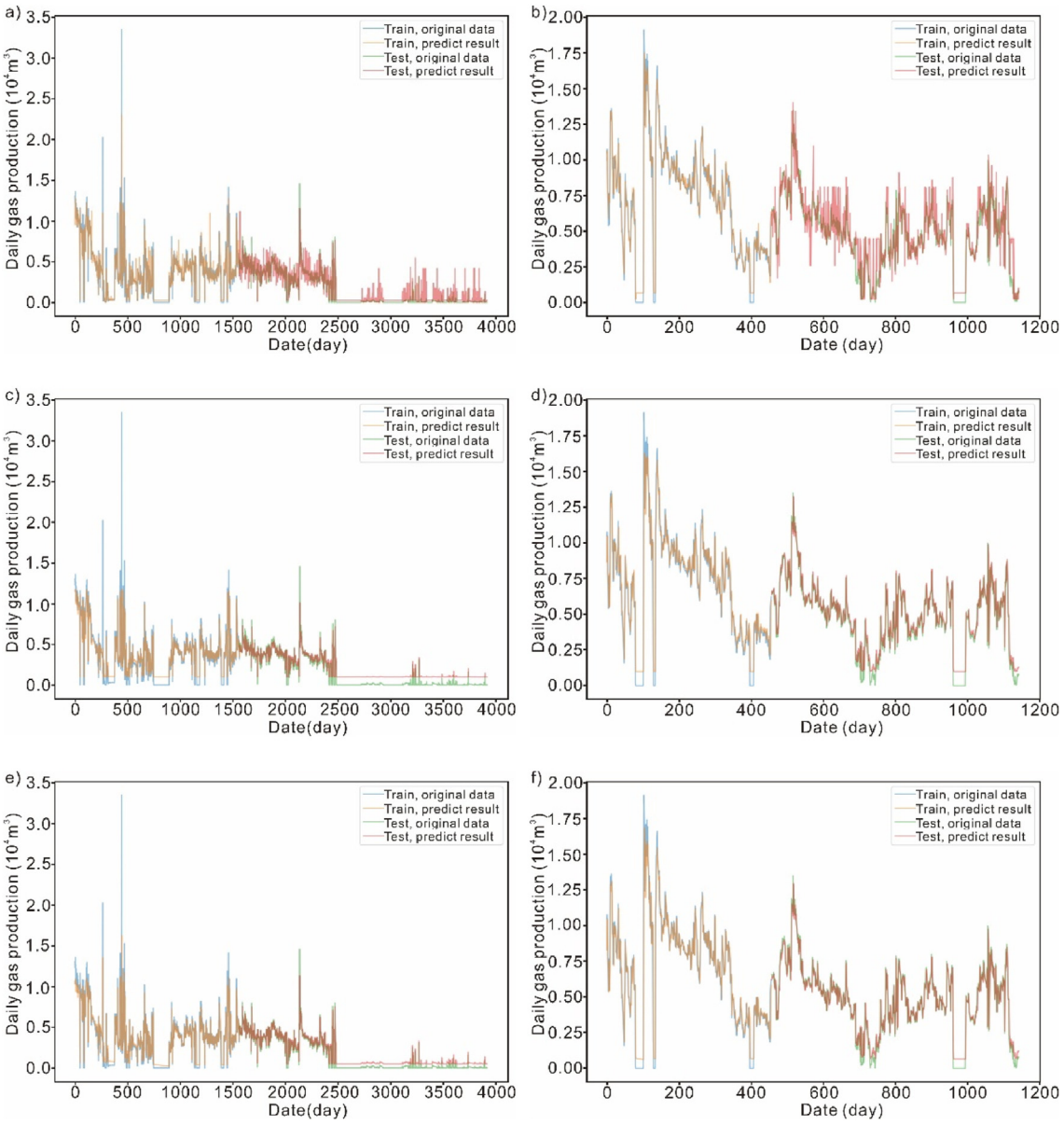

Fig. 5. Data-driven methods to predict daily productions of gas wells S1 and S2. (a): The prediction result of RF on S1. (b): The prediction result of RF on S2. (c): The prediction result of SVR on S1. (d): The prediction result of SVR on S2. (e): The prediction result of DLSTM-FNN on S1. (f): The prediction result of DLSTM-FNN on S2.

Table 2
The RMSE of the production prediction models for wells S1 and S2.

Model	S1 training set	S1 testing set	S2 training set	S2 testing set
RF	0.1090542	0.1285114	0.1269109	0.1587212
SVR	0.2011788	0.0957835	0.1945521	0.1073256
DLSTM-FNN	0.2042352	0.0722189	0.19246629	0.1025749

dynamic daily production prediction were verified using 1692 time series of wells in the Sulige gas field. Among them, the DLSTM-FNN deep learning model had a good ability in time series modeling. For the short training sample set and the rapid prediction of thousands of wells, the quantitated prediction performances revealed that the DLSTM-FNN model was more accurate and reliable and was more in line with the actual dynamic trend of the daily production of gas wells. Furthermore, the proposed DLSTM-FNN model can be applied to the time series prediction of other tight gas fields based on transfer learning.

Funding

This work was supported by the National Key R&D Program of China (2020YFA0713404).

Declaration of competing interest

The authors declare that they do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

References

Arps, J.J., 1945. Analysis of decline curves. Transact. Am. Inst. Min. Metal. Petrol. Eng. 160 (1), 228–247.
Bengio, Y., Simard, P., Frasconi, P., 1994. Learning long-term dependencies with gradient descent is difficult. IEEE Trans. Neural Network. 5 (2), 157–166.
Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.
Cao, Q., Banerjee, R., Gupta, S., et al., 2016. Data driven production forecasting using machine learning. In: SPE Argentina Exploration and Production of Unconventional Resources Symposium. SPE, Buenos Aires.

Chen, Y., Fu, L., 2019. Establishment, comparison and application of power function decline model. Petrol. Geol. Recov. Effic. 26 (6), 87–91.
Chen, Y., Tang, W., 2016. Establishment and application of generalized decline model. Acta Pet. Sin. 37 (11), 1410–1413.
Chen, Y., Zhou, C., 2015. Establishment, comparison and application of the linear decline type. Acta Pet. Sin. 36 (8), 983–987.
Dong, W., Yang, S., 2009. Factors affecting production decline performance and production forecasting. Nat. Gas Geosci. 20 (3), 411–415.
Gu, J., Zhou, M., Li, Z., et al., 2019. Oil well production forecast with long- short term memory network model based on data mining. Special Oil Gas Reservoirs 26 (2), 81–85, 135.
Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780.
Hopfield, J.J., 1982. Neural networks and physical systems with emergent collective computational abilities. In: Proceedings of the National Academy of Sciences of the United States of America, 79. National Academy of Sciences of the United States of America, pp. 2554–2558, 8.
Hou, C., 2019. New well oil production forecast method based on long-term and short-term memory neural network. Petrol. Geol. Recov. Effic. 26 (3), 105–110.
Jiang, F., Pang, X., Jiang, Z., et al., 2007. Physical simulation experiment of gas charging in tight sandstone. Geol. Rev. 53 (6), 844–849.
Jiang, Y., Chen, X., Bao, H., 2021. A new model for rapid prediction of horizontal well production decline in shale gas staged fracturing: case study of Fuling shale gas field. Nat. Gas Geosci. 32 (6), 845–850.
Liu, W., Liu, W.D., Gu, J., Shen, X., 2019. Predictive model for water absorption in sublayers using a machine learning method. J. Petrol. Sci. Eng. 182, 106367.
Liu, W., Liu, W., Gu, J., et al., 2020. Oil production prediction based on a machine learning method. Oil Drill. Product. Technol. 42 (1), 70–75.
Sagheer, A., Kotb, M., 2019. Time series forecasting of petroleum production using deep LSTM recurrent networks. Neurocomputing 323, 203–213.
Yang, H., Fu, J., Liu, X., Meng, P., 2012. Accumulation conditions and exploration and development of tight gas in the upper paleozoic of the Ordos basin. Petrol. Explor. Dev. 39 (3), 295–303.
Yang, T., Zhang, G., Liang, K., 2012. The exploration of global tight sandstone gas and forecast of the development tendency in China. Strat. Study CAE 14 (16), 64–68.
Zou, C., Yang, Z., Tao, S., et al., 2012. Nano-hydrocarbon and the accumulation in coexisting source and reservoir. Petrol. Explor. Dev. 39 (1), 13–26.