Reliable Prediction on Emerging Energy Supply for National Sustainability and Stability: A Case Study on Coal Bed Gas Supply in China Based on the Dual-LSTM Model

PING KANG¹, JUNMING LAO²,³, MINGXU YU³,⁴, HONGQING SONG²,³, AND CHENG WANG³,⁴

¹Beijing Engineering Consulting Company Ltd., Beijing 100025, China
²School of Civil and Resource Engineering, University of Science and Technology Beijing, Beijing 100083, China
³National and Local Joint Engineering Laboratory for Big-data Analysis and Computing Technology, Beijing 100190, China
⁴Beilong Zedata (Beijing Municipality) Company Ltd., Beijing 100055, China

Corresponding author: Hongqing Song (songhongqing@ustb.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51974357, and in part by the Fundamental Research Funds for the Central Universities of China under Grant FRF-TP-19-005B1.

ABSTRACT Aiming to prevent from the imbalance between supply and demand of energy in which the share of emerging type is rapidly increasing, to predict the supply of emerging energy reliably is significant. However, the expected distribution uncertain and high-noise characteristics of emerging energy supply impede the reliable prediction. The Dual-LSTM (Long Short-Term Memory) model was constructed for the characteristic extracting and effective prediction of the expected distribution uncertain high-noise emerging energy supply time series. A case study on coal bed gas supply in China was conducted. Results showed that the Dual-LSTM model effectively solved the problem of superfluous and non-quantifiable variables in the prediction of coal bed gas supply and extracted the statistical characteristics of expected distribution uncertain and high-noise data samples effectively with a relative error major less than 5% in short-term. Besides, the Dual-LSTM model has a significantly higher prediction accuracy while comparing with ARIMA model and original LSTM model. Ultimately, it is predicted that the year-on-year growth rates of coal bed gas supply of China from January to September, 2021, approximately maintains 75% in average based on the Dual-LSTM model. The Dual-LSTM model provides a reliable statistical model for policy decision to maintain national sustainability and stability.

INDEX TERMS Improved deep learning model, uncertain expected distribution high-noise sample, emerging energy, supply-demand stability, reliable prediction, time series analysis.

I. INTRODUCTION

The imbalance between supply and demand of energy is regarded as the potential impact on the national energy security and operation. Reliable prediction of energy supply is the key to avoid the imbalance of supply and demand [1]–[6]. Once if the predictive value of energy supply was inadequate, it would lead to surplus import. Meanwhile, the release of energy reserves would not only suppress the energy price which is not conducive to the survival of energy enterprises, but also result in a large outflow of national foreign exchange impacting on the national economy. If the predictive value of energy supply was excessive, it would lead to insufficient energy import and massive energy will be allocated and stored, endangering the national energy security and social stability [7]–[10].

In recent years, the supplies of emerging energy such as wind energy, solar energy, biomass energy and unconventional natural gas, etc have been rapidly boosting. For instance, the unconventional natural gas in China has increased rapidly. Figure 1 shows the natural gas supply from 2013 to 2020 in China and the structure of China’s natural gas supply of 2020. It can be found that the share of unconventional gas supply in China has been increasing, accounting for nearly 15% of the natural gas supply in the market [11].

The associate editor coordinating the review of this manuscript and approving it for publication was Jian Guo.
However, emerging energy is still in the development stage, whose supply is unstable and the randomness characteristic is obvious, leading to the prediction method of conventional energy inapplicable to be adopted [12]–[15]. First of all, diverse from conventional energy, emerging energy supply demonstrates the characteristics of non-linearity and high noise. In addition, variables affecting emerging energy supply are extensive while some of whom are non-quantifiable, limiting the application of prediction methods. Finally, distinguished from the widely studied problem, the single reservoir production prediction, emerging energy supply is the sum of the production capacity of the whole country in a specific period, which can not be described by a single capacity model.

Although there are few studies on emerging energy supply prediction, the existing reservoir production prediction models of conventional energy still remain certain reference value. Table 1 provides a comparison of various types of reservoir production prediction models. The conventional prediction models for reservoir production include generalized single- and multi-peak period models such as Weng’s model, cumulative growth models such as Hubbert model, Grey model, exponential smoothing model, linear regression analysis models such as differential auto-regressive moving average (ARIMA) model. The neural network models achieved a dominant position in alternative prediction models, such as back error propagation neural network (BPNN) model and long short term memory (LSTM) network model.

Weng’s model and Hubbert model are classical mathematical methods to describe the production and recoverable reserves of oil and gas fields, highly fitting the actual production capacity and energy storage curve of oil and gas field. However, the oil and gas production in the model is a single value function of time, lacking correlation between historical data and predictive values [16]–[18]. Weng’s model is based on the law of rise-growth-maturity-recession in natural process. In essence, it is a mathematical model belonging to continuous gamma distribution. Its mathematical expression is: $Q = at^b e^{-t}$, where, $Q$ is oil or gas production, $t$ is extraction cycle, $a, b$ are undetermined coefficients [8]–[10].

Although the Weng’s model has achieved high precision fitting of the actual oil field yield curve, it is limited to the prediction of oil field whose production corresponds to gamma distribution. The Hubbert model describes the relationship between the remaining recoverable reserves and the cumulative production of oil and gas reservoirs with the variation of mining time. The relationship follows the law of natural exponential distribution and its mathematical expression is: $(G_R - G_P)/G_P = ae^{-bt}$, where, $G_R$ is the residual recoverable reserves, $G_P$ is the cumulative production, $t$ is the extraction cycle, $a, b$ are undetermined coefficients [19], [20].

Grey model, exponential smoothing model and ARIMA model are typical time series analysis methods for oil and gas production. The Grey model weakens the random factors of the original data by accumulating the original data time series. It carries on the natural exponential regression to the data, and subsequently solves the parameters of the exponential function based on the least square method. Finally the regression model of the original data are obtained. The prediction accuracy of the grey model is limited by the distribution of the data. Once the accumulated data deviated from the natural exponential law, underfitting or overfitting would occur and the prediction effect of the Grey model would deviate from the actual value [21]–[24]. The exponential smoothing model applies the linear combination of time series to represent the prediction value at the next moment. The time complexity of the model and algorithm is relatively simple. However, the exponential smoothing model has strong linear characteristics and the predicted values lag behind the actual observations [25]. ARIMA model first obtains a time
TABLE 1. Comparison of prediction models of natural gas production or supply.

| Predictive model | Weng’s model | Hubbert model | Grey model | Exponential Smoothing model | ARIMA | BPNN | LSTM |
|------------------|-------------|--------------|-----------|----------------------------|-------|-----|------|
| Theoretical basis| Gamma distribution; Multi-peak period | Exponential distribution | Exponential curve fitting | Primary/secondary exponential smoothing | Auto-regression; Moving average; Information criterion; Auto-correlation coefficient | Gradient descent; Error back propagation | Input gate; Forgetting gate; Output gate; LSTM cell |
| Variable features| Single variable | Single variable | Single variable; Time series | Time series | Time series | Single/Multi-variable | Time series |
| Data type        | Obeys gamma distribution | Cumulatively obey exponential distribution | Cumulatively obey exponential distribution | Universal | Universal | Universal | Universal |
| Periodicity      | Applicable | Inapplicable | Inapplicable | Inapplicable | Applicable | Applicable | Applicable |
| Degree of freedom| $1 \times n$ | $1 \times n$ | $1 \times n + s$ | $s$ | $s$ | $m \times n$ | $s$ |
| Model time complexity | $O(n)$ | $O(n)$ | $O(n + s)$ | $O(s)$ | $O(s)$ | $O(m \times n)$ | $O(s \times (e \times h))$ |
| Scope of application | Medium and long term; Peak prediction | Medium and long term | Short term | Short term | Medium and long term | Short and medium term | Short, medium and long term |

series with certain data stationarity by data difference, and then expresses the predicted value as the sum of the linear combination of the $p$ order lag of the time series and the $q$ order lag of the prediction error, eliminating the influence of random factors (also called noise) in the original time series. Compared with Grey model and exponential smoothing model, ARIMA has better regression and prediction effect on expected distribution uncertain high noise time series. Nevertheless, the ARIMA model is still limited to extracting linear features of data [26]–[33].

The neural network model was proposed for effectively extracting the linear and nonlinear characteristics of the data. However, in the emerging energy supply prediction scenario, some influencing variables such as social policy, business operation dynamics and so on are failed to quantify accurately or included in the model input, affecting the accuracy of neural network prediction. The basic principle of BP neural network is to reduce the error between the predictive value and the actual value based on the gradient descent method. The hidden layer is used to process the input data with the weight and the specific activation function. The characteristics of the data are mined in the process of revising the weight value, and the model weight is optimized effectively to improve the fitting accuracy of the model [30], [31].

Long short term memory network (LSTM) is an improved recurrent neural network (RNN), which effectively retains the main information of the time series, discard the invalid information of the time series and be suitable for mining and predicting the nonlinear time series [32]. It is worth noting that the LSTM model is a member of the deep learning (DL) which is defined as a method of machine learning based on artificial neural networks with representation learning. In the deep learning framework, higher-level features from the raw input are extracted while compared with other algorithms. From the perspective of emerging energy supply which features expected distribution uncertain and high noise, deep learning model is desirable for whose regression analyse [33], [34]. Nevertheless, the application of deep learning algorithms on high noise or linear time series prediction were rarely studied due to that the deep learning algorithms have the capability to handle big data by capturing the inherent non-linear features through automatic feature extraction methods [35]–[37]. Moreover, overfitting frequently occurs while specifically predicting emerging energy supply [38]–[40].

For the exact purpose of reliable prediction on emerging energy supply, an improved LSTM model which integrate filtering and accurate regression capabilities is indispensable. In this study, a comprehensive introduction was conducted in Section I. The theory of expected distribution uncertain and high noise time series prediction and how to filter and effectively regress were discussed in Section II. The novel Dual-LSTM model was constructed after introducing the original LSTM model in Section III. The Dual-LSTM model performance and results of emerging energy supply prediction were presented in Section IV. Conclusions were drawn in Section V.

II. THEORY OF EXPECTED DISTRIBUTION UNCERTAIN HIGH NOISE TIME SERIES PREDICTION

From a statistical point of view, the observed value of a sample data set is equal to the sum of its expected distribution and random distribution (also known as statistical noise). Time series is a data set composed of the values under the same statistical index according to the order of its occurrence time. Time series analysis and prediction are based on the feature extraction of historical data. For a specific model, the closer the extracted data features are to the real value of the time
series, the higher the accuracy of fitting and prediction of the model. The characteristic of data sample refers to the statistical distribution of data sample, including data trend and data noise. Specifically, trend represents the concept of expected distribution of raw data set while noise represents the random features of raw data set. In general, the more obvious the statistical distribution of a data sample, the lower the noise and easier to realize the feature extraction and trend prediction of the data sample.

Figure 2 indicates the composition of a time series with expected distribution uncertain and high-noise characteristics, simply, the sum of trend and noise. Accordingly, to predict the trend and noise of a time series with expected distribution uncertain and high-noise characteristics separately would maintain the general features while avoiding significant deviations caused by over-fitting, which is the purpose of Dual-LSTM model. Deeper study of the Dual-LSTM model is carried forward in Section III.

Figure 3 shows the coal bed gas supply of total, Shanxi, Beijing municipality and Guizhou from March, 2013 to December, 2020 and respective characteristics. It could be found that the expected distribution uncertain characteristics and high-noise characteristics of coal bed gas supply are apparent, which brings more challenges to coal bed gas supply prediction.

A. TIME SERIES ANALYSIS AND PREDICTION

The time series contains the statistical characteristics of the trend, randomness and periodicity of the sample data in a certain period of time. The purpose of time series analysis is to extract and fit the above statistical characteristics of the sample by statistical method, construct a specific statistical model and predict the time series of the object basing on the model. Time series analysis is a functional analysis method, as shown in formula (1), where, $P_j$ is the value of studied subject at the j moment, $h_{j-1}$ to $h_{j-i}$ are the first to i historical values of the j moment, $x, y, z$ . . . are the impact variables on studied subject. The basic idea of time series analysis is that the value of the research object in a certain period of time is only related to its historical value. If the historical value of the research object was observable, it would be dispensable to consider the influence of other variables on the research object. Therefore, time series analysis effectively avoid the problem of superfluous variables and non-quantifiable variables, ensuring the reliability of prediction results. Time series analysis is an ideal analysis tool for unconventional natural gas supply with superfluous and non-quantifiable influencing variables.

$$P_j = P_j(h_{j-1}, h_{j-2}, \ldots, h_{j-i}), \quad i < j$$

B. EXPECTED DISTRIBUTION UNCERTAIN TIME SERIES

The characteristic of expected distribution uncertain time series is that the distribution of the raw time series is laborious to be described by typical and broadly utilized statistical model such as linear regression model, logistic regression model, exponential regression model, etc. According to Figure 3, the emerging energy supply time series have significant expected distribution uncertain characteristics since emerging energy, as a rapidly developing new energy resource, has increased its production year after year with the support of advanced developing technology.

To effectively predictive a expected distribution uncertain time series, two mainstream regression methods were proposed. On the one hand, the observation distribution of the raw data set is regarded as the actual distribution. Accordingly, the regression model work hard to fit the training set during the training process like common models such as artificial neural network including LSTM model usually do. However, once if the observation distribution of sample data significantly differed from the population, the model prediction performance would be inferior or called over-fitting. On the other hand, the observation distribution of the raw data set could be considered as the combination of expected distribution and random distribution, which was adopted in this research. Distinguished from general regression analyse, the expected distribution was not set but learned by Dual-LSTM model in this research, which is aiming to get closer to the actual distribution of data set.

C. HIGH NOISE TIME SERIES

In statistics, the definition of noise refers to the randomness found in a given data sample, the main form of which is

FIGURE 2. The sketch map of the composition of a time series with expected distribution uncertain and high-noise characteristics.
random error. The emerging energy supply is affected by considerable variables, a tiny change of each influencing variable may increase or decrease the coal bed gas supply significantly. Under the control of multiple variables, the emerging energy supply data has strong randomness, showing a high frequency fluctuation in curve. In the scene of statistical learning, the statistical model will learn the noise data together during the training process, which leads to the deviation between the features extracted by the model and the real law of the sample, finally affecting the prediction accuracy of the model.

According to the high noise time series prediction, two mainstream methods are optional. One is adopting the filter algorithm such as Kalman filter, Volterra adaptive filter, Bilateral filter, etc to the regression model which has been widely applied in the prediction on periodic data, stationary data and other data with clear expected distribution. It is worth noting that machine learning integrated with filter algorithm has been a powerful tool for prediction and optimal problems in high precision and efficiency [37]–[39]. The other is to regress the random distribution of data set after separating the expected distribution, which is applied by Dual-LSTM model in this research. Generally, the random distribution of a data set is assumed to follow the Gaussian distribution while the assumption is invalid mostly, however. In this research, the random distribution of the data set is regressed by the Dual-LSTM model either for higher regression accuracy. Consequently, the influence of noise in the time series is fully considered by the model and the prediction accuracy is improvable.

III. METHODOLOGY AND MODEL

A. RESEARCH METHODS

In this study, the Dual-LSTM model is constructed to fit and predict the coal bed gas supply sample time series. The mean square error (MSE) is introduced as a loss function to validate the fitting effect of the model. Formula (2) shows the hyperplane regression equation, which is the fundamental of neural network. Where, $Y_P$ is the forecast time series, $W$ is the weight of variable, $X$ is the independent variable which has been widely applied in the prediction on periodic data, stationary data and other data with clear expected distribution. It is worth noting that machine learning integrated with filter algorithm has been a powerful tool for prediction and optimal problems in high precision and efficiency [37]–[39]. The other is to regress the random distribution of data set after separating the expected distribution, which is applied by Dual-LSTM model in this research. Generally, the random distribution of a data set is assumed to follow the Gaussian distribution while the assumption is invalid mostly, however. In this research, the random distribution of the data set is regressed by the Dual-LSTM model either for higher regression accuracy. Consequently, the influence of noise in the time series is fully considered by the model and the prediction accuracy is improvable.

\[
Y_P = W^TX + b \tag{2}
\]

\[
Loss = MSE = \sqrt{\frac{||Y_P - Y||^2}{n}} \tag{3}
\]

\[
W^{(m+1)} = W^{(m)} - \alpha \frac{\partial Loss}{\partial W}
\tag{4}
\]

B. LSTM MODEL

Different from the general neural network, the input sample of LSTM is regarded as an independent and equally distributed data set. The recurrent neural network is based on the principle that the current data value is related to the
historical data which is retained while the input data are inputted to ensure the persistence of the information. LSTM is an improved recurrent neural network, which can solve the problem of long dependence (with the increase of time series data samples, the network will lose the early data information of time series) in general recurrent neural network.

Figure 4 shows the basic structure and working principle of the LSTM. Basic units of LSTM are called cells. Function realization depends on forgetting gate, input gate, output gate and cell status update. Each gate corresponds to a specific mathematical operation of the data in the cell. The model adopts Sigmoid() and tanh() functions as activation functions, whose expressions are shown in formula (5) and formula (6), respectively.

Formula (7) corresponds to the forgetting gate operation. The forgetting gate determines which information the cell state should retain or discard. Specifically, via coinciding the output value of the last LSTM cell, Yt-1, with the input value of the current LSTM cell, Xt, hyperplane regression and sigmoid() function processing, a vector, ft, consisting of numbers with values between 0 and 1 is obtained. 0 in the vector indicates that the corresponding original vector element is completely discarded and 1 indicates complete retention.

Formula (8)-(9) corresponds to input gate operation. The input gate determines what new information is added to LSTM cell. Specifically, via coinciding the output value of the last LSTM cell, Yt-1, with the input value of the current LSTM cell, Xt, hyperplane regressing and sigmoid() function processing, a vector, it, consisting of numbers with values between 0 and 1 is obtained. The 0 in the vector indicates that the corresponding original vector element does not change and 1 indicates a complete update. On the other hand, via tanh() function processing, a vector, ct, consisting of numbers with values between -1 and 1 is obtained for determining information updates.

The formula (10) is the cell state update operation. Specifically, the dot product operation of the last LSTM cell state Ct-1 and ft is summarized with the dot product operation of it and ct, obtaining the status Ct of current LSTM cells.

Formula (11)-(12) corresponds to the operation of output gate. In particular, the Yt-1 and Xt are combined and processed by hyperplane regression and Sigmoid(), and the Ct treated with tanh() is used for dot product operation to modify the state of the cell, obtaining the vector Yt as the current cell to output. The W, b of all formulas are distributed to represent the weight term and the bias term, while × represent the vector product operation.

\[
\text{Sigmoid} = \frac{1}{1 + e^{-x}} \\
\text{tanh} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \\
f_t = \text{Sigmoid}(W_f[Y_{t-1}, X_t] + b_f) \\
i_t = \text{Sigmoid}(W_i[Y_{t-1}, X_t] + b_i) \\
c_t = \text{tanh}(W_c[Y_{t-1}, X_t] + b_c) \\
C_t = f_t \times C_{t-1} + i_t \times c_t \\
o_t = \text{Sigmoid}(W_o[Y_{t-1}, X_t] + b_o) \\
Y_t = \text{tanh}(C_t) \times o_t
\]

where, [Yt-1,Xt] represents that the vector Yt-1 and Xt are merged by a tail prime connection. For instance, if the Yt-1 is a m dimension vector and the Xt is a n dimension vector, then [Yt-1,Xt] is a m + n dimension vector.

C. DUAL-LSTM MODEL

The detailed working flow of the Dual-LSTM model is shown in Figure 5. A better understanding of the Dual-LSTM model
could be achieved if comprehensively study both the Figure 4 and Figure 5.

According to the basic theory of regressing and predicting emerging energy supply adopted in this research, the specific operational structure and logical structure have been constructed as shown from Figure 5(a) to Figure 5(c). First the sample data, known as training set, will be processed mainly by interpolation to solve data missing. Next is followed by the normalization and input of the sample data. It should be noted that from the emerging energy supply perspective, feature of variables is one-dimensional. In a certain moment t, the sample data is input into the Dual-LSTM model and merged with the hidden state of last moment t-1. Then the procedure of expected distribution regression starts. After the data process done by input gate, forget gate and output gate the cell state and hidden state of expected distribution regression
are refreshed. When the epoch finishes, the expected distribution of sample is output. Meanwhile, the residual error is generated by subtracting the expected distribution from sample data which is also called the actual distribution. Further, the residual error is input and merged with the hidden state of random distribution regression process while the next procedures are consistent with that of expected distribution regression and the random distribution of sample is got. Consequently, the complete regression model is generated by summing the expected distribution and random distribution. Prediction is able to be conducted by inputting objective prediction steps.

IV. MODEL PERFORMANCE AND RESULTS

A. DATA SOURCES AND HYPER-PARAMETRIC SETTINGS

The data of this study are derived from the National Bureau of Statistics and the National Development and Reform Commission of China. For the supply of unconventional natural gas in China, the National Bureau of Statistics and the National Development and Reform Commission only track and disclose the supply of coal bed gas [41], [42]. Therefore, in this study, the monthly data of coal bed gas supply in China from March 2013 to December 2020 are selected as the observation data set, containing 94 samples. From March 2013 to June 2019, 76 samples were selected as training sets. From July 2019 to December 2020, 18 samples were selected as validation sets. Per time step represents one month. It is worth noting that data pre-processing was done to supply the missing values by linear interpolation method [41], [42]. Comparative analysis of regression and prediction performance of the ARIMA model, the LSTM model and Dual-LSTM model was performed. Finally, based on the Dual-LSTM model, the supply of coal bed gas in China from January to September, 2021 was predicted.

For the purpose of clarifying the reliability of the Dual-LSTM model, model comparison is performed in this research. First the ARIMA model is chosen since whom typically represents the extensively used linear model in conventional energy supply regression prediction. Besides, the original LSTM model is taken into account either for enhancing the diversity between original LSTM model and Dual-LSTM model and highlighting the advantages of the Dual-LSTM model.

Based on the Akaike information criterion (AIC) [43], [44], ARIMA model adopts the first order difference, third order autoregressive and third order moving average, namely $ARIMA (p, d, q) = ARIMA (1, 1, 1)$. The original
P. Kang et al.: Reliable Prediction on Emerging Energy Supply for National Sustainability and Stability

FIGURE 7. The noise prediction of coal bed gas supply. (a) The noise prediction of total coal bed gas supply in China. (b) The noise prediction of coal bed gas supply in Shanxi province. (c) The noise prediction of coal bed gas supply in Beijing municipality. (d) The noise prediction of coal bed gas supply in Guizhou province.

LSTM model is set without regularization and drop-out. After certain amount of experiments, the learning rate of the original LSTM model is set as 0.1, epoch steps are set as 1500, units of hidden states are set as 64, epoch size is set as 1 and window size is set as 94 (the full size of sample data). Parameters setting of Dual-LSTM model is supplied in Figure 5(c).

B. MODEL TRAINING AND GENERALIZATION VALIDATION

Figure 6 reveals the trend evaluations (from step 1 to step 76, green curves) and predictions (from step 77 to step 94, green curves) of coal bed gas supply of China total, Shanxi province, Beijing municipality and Guizhou province, respectively. Judging from the evaluation results, though the evaluation accuracy was relatively low caused by a small learning rate, the trend predictions of coal bed gas supply were appropriately fitted and predicted in the short term (no more than 9 months in this study).

Figure 7 reveals the noise prediction (gold curves) of coal bed gas supply of China total, Shanxi province, Beijing municipality and Guizhou province, respectively. Comparing with the trend predictions, the noise evaluations indicate a fitting accuracy reaching perfectly coincided according to a higher learning rate. From the noise prediction perspective, the noise prediction results basically follow the distribution law of the observation samples.

Figure 8 reveals the prediction performance of Dual-LSTM model on coal bed gas supply. According to the prediction performance of Dual-LSTM model, it is worth noting that the prediction results conducted by Dual-LSTM model valued similarly to the observation samples.

Figure 9 reveals the effective prediction of coal bed gas supply by Dual-LSTM model. Consistent with the trend prediction and noise prediction, the effective predictions of coal bed gas supply were with higher accuracy in the short term, especially for the predictions of coal bed gas supply of China total and Shanxi province.

C. MODEL COMPARISON AND ERROR ANALYSE

Figure 10 shows the comparison of prediction performance between ARIMA model, LSTM model and Dual-LSTM model. On the whole, the Dual-LSTM model conducted the best prediction on coal bed gas supply of either China total, Shanxi province, Beijing municipality or Guizhou province while comparing with the ARIMA model and original LSTM model.
FIGURE 8. The prediction performance of dual-LSTM model on coal bed gas supply, where $y$ represents the prediction value of Dual-LSTM model while $x$ represents the observation samples. (a) The prediction performance on total coal bed gas supply in China. (b) The prediction performance on coal bed gas supply in Shanxi province. (c) The prediction performance on coal bed gas supply in Beijing municipality. (d) The prediction performance on coal bed gas supply in Guizhou province.

model. In particular, though the ARIMA model effectively predict the trend of coal bed gas supply, the prediction results were only in linear form. From the original LSTM model perspective, the prediction results significantly indicate the over fitting situation since the training loss was controlled to a low value while the prediction results apparently deviate from observation values. It should be noted that for either ARIMA model, original LSTM model or Dual-LSTM model, the short-term predictions demonstrate a higher accuracy than the long-term predictions, indicating that as the prediction steps grow, the impact of training set on prediction results gradually fade. Especially for time series with expected distribution uncertain and high-noise characteristics, the stochastic noise impact significantly on the long-term predictions. Figure 11 reveals the relative error (RE) of prediction of Dual-LSTM model in diverse time steps. According to Figure 11 the prediction accuracy disparity is straightforward. It could be found that in the first 8 prediction steps relative errors are entirely within 8% while the majority is under 5%. From the ninth step on the prediction RE starts to rise prominently.

D. PREDICTION OF COAL BED GAS SUPPLY

Adopting the Dual-LSTM model and basing on the observation data set, the coal bed gas supply of China total, Shanxi province, Beijing municipality and Guizhou province from January to September, 2021 are predicted, as shown in Figure 12. The growth trends of coal bed gas supply in the future are also highlighted in Figure 12. Judging from the effective predictive results, the coal bed gas supply in China from January to September, 2021 would grow from 11.4 to 15.4 hundred million cubic metre. The year-on-year growth rate of coal bed gas supply in China from January to September, 2021 are about 46%, 60%, 69%, 75%, 77%, 101%, 90%, 88% and 83%, respectively, which still maintains a high growth rate. For Shanxi province, the coal bed gas supply January to September, 2021 would grow from 9.0 to 10.7 hundred million cubic metre while the year-on-year
growth rate of coal bed gas supply would be 63%, 73%, 76%, 79%, 76%, 104%, 90%, 91% and 85%, respectively, remaining the province with the most coal bed gas supply in China. Focusing on Beijing municipality, the coal bed gas supply maintains a slack and steady growth. The coal bed gas supply from January to September, 2021 would grow
from 1.6 to 1.8 hundred million cubic metre in fluctuation with year-on-year growth rate of 22%, 24%, 23%, 33%, 40%, 42%, 56%, 20% and 26%, respectively. From Guizhou province perspective, the future growth speed of coal bed gas supply indicates to slow down gradually. The coal bed gas supply from January to September, 2021 maintains
0.5 hundred million cubic metre firstly but remains 0.4 hundred million cubic metre finally with year-on-year growth rate of 69%, 61%, 57%, 10%, 7%, 16%, −23%, 18% and −20%, respectively.

V. CONCLUSION

Aiming to prevent from the imbalance between supply and demand of energy in which the share of unconventional type is rapidly increasing in China, to predict the supply of emerging energy reliably is significant. A Dual-LSTM model was constructed to extract and predict the emerging energy supply with expected distribution uncertain high noise characteristics. The fitting accuracy and generalization ability of the model were compared and analyzed. The prediction of coal bed gas supply in China was carried out.

The key conclusions are as follows: (1) The supply of unconventional natural gas in China continues to grow rapidly. In the whole 2020, unconventional gas supply accounted for nearly 15% of the total natural gas supply. (2) The Dual-LSTM model achieved solving the problem of superfluous and non-quantifiable variables. Meanwhile, the statistical features of expected distribution uncertain and high noise data samples were extracted effectively with a relative error major less than 5% in short-term. (3) In the study condition, the Dual-LSTM model has a significantly higher prediction accuracy while comparing with ARIMA model and original LSTM model. (4) Based on the Dual-LSTM model, the year-on-year growth rates of coal bed gas supply in China were predicted to be about 75% in average from January to September, 2021. The Dual-LSTM model effectively predicts the supply of unconventional gas in China in the short term. Moreover, it shows high robustness to the fitting and prediction of expected distribution uncertain high noise time series, providing a reliable statistical model for national sustainability and stability.

ACKNOWLEDGMENT

The authors would like to thank Beijing Engineering Consulting Company Ltd., for providing the monitoring data used in this study. They would also grateful to the University of Science and Technology Beijing, the National and Local Joint Engineering Laboratory for Big-Data Analysis and Computing Technology, and Beilong Zedata (Beijing Municipality) Company Ltd., either for offering technological supports.

REFERENCES

[1] A. Hu and X. Wang, “Safety of natural gas supply and countermeasure,” in (Chinese), Natural Gas Ind., vol. 10, no. 10, p. 125, 2008.
[2] A. Bittante and H. Saxén, “Design of small LNG supply chain by multi-period optimization,” Energies, vol. 13, no. 24, p. 6704, Dec. 2020, doi: 10.3390/en13246704.
[3] S.-H. Kim, Y.-Y. Lim, D.-W. Kim, and M.-K. Kim, “Swing suppliers and international natural gas market integration,” Energies, vol. 13, no. 18, p. 4661, Sep. 2020, doi: 10.3390/en13184661.
[4] K. Papageorghiou, E. I. Papageorghiou, K. Poczeza, D. Bochits, and G. Stamoulis, “Forecasting of day-ahead natural gas consumption demand in greece using adaptive neuro-fuzzy inference system,” Energies, vol. 13, no. 9, p. 2317, May 2020, doi: 10.3390/en13092317.
[5] Y. Liu, F. Li, and X. Yu, “Gas supply, pricing mechanism and the economics of power generation in China,” Energies, vol. 11, no. 5, p. 1058, Apr. 2018, doi: 10.3390/en11051058.
[6] X. Zhang, “Study on the safety assessments and countermeasures of natural gas supply in China,” Southeast Petroleum Univ., Chengdu, China, Tech. Rep., Apr. 2011. [Online]. Available: https://kns.cnki.net/kcms/detail/detail.aspx?FileName=10112252514 nh&DbName=CMFD2012
[7] J. Zhang, H. Song, W. Zhu, and J. Wang, “Liquid transport through nanoscale porous media with strong wettability,” Transp. Porous Media, vol. 190, pp. 1–5, Jan. 2021, doi: 10.1007/s11242-020-01519-5.
[8] Q. Xu, L. Liu, J. Feng, L. Qiao, C. Yu, W. Shi, C. Ding, Y. Zang, C. Chang, Y. Xiong, and Y. Ding, “A comparative investigation on the effect of different nanofluids on the thermal performance of two-phase closed thermosyphon,” Int. J. Heat Mass Transf., vol. 149, p. 119, Mar. 2020.
[9] Q. Xu, J. Feng, J. Zhou, L. Liu, Y. Zang, and H. Fan, “Study of a new type of radiant tube based on the traditional M-type structure,” Appl. Thermal Eng., vol. 150, pp. 849–857, Mar. 2019, doi: 10.1016/j.applthermaleng.2019.01.016.
[10] Q. Xu, J. Feng, L. Liu, J. Zhou, G. Ye, and C. Chang, “Analysis of mechanical-fluid-thermal performance of heat pipeline system with structural deformation effects,” Int. J. Heat Mass Transf., vol. 128, pp. 12–23, Jan. 2019, doi: 10.1016/j.ijheatmasstransfer.2018.08.122.
[11] Natural Gas Operation in the First Quarter of 2020, (in Chinese), Nat. Develop. Reform Commission People Republic China, Beijing, China, 2020.
[12] C. Wei, L. Wang, B. Li, L. Xiong, S. Liu, J. Zheng, S. Hu, and H. Song, “A study of nonlinear elasticity effects on permeability of stress sensitive shale rocks using an improved coupled flow and geomechanics model: A case study of the longmaxi shale in China,” Energies, vol. 11, no. 2, p. 329, Feb. 2018, doi: 10.3390/en11020329.
[13] C. Wei, H. Song, Y. Li, Q. Zhang, B. Song, and J. Wang, “Production characteristics with different superimposed modes using variogram: A case study of a super-giant carbonate reservoir in the Middle East,” Energies, vol. 10, no. 2, p. 250, Feb. 2017, doi: 10.3390/en10020250.
[14] J. Lao, H. Song, C. Wang, Y. Zhou, and J. Wang, “Reducing atmospheric pollutant and greenhouse gas emissions of heavy duty trucks by substituting diesel with hydrogen in Beijing-Tianjin-Hebei-Shandong region, China,” Int. J. Hydrogen Energy, vol. 46, no. 34, pp. 18137–18152, May 2021, doi: 10.1016/j.ijhydene.2020.09.132.
[15] H. Song, H. Guo, Y. Wang, J. Lao, H. Zhu, L. Tang, and X. Liu, “A novel hybrid energy system for hydrogen production and storage in a depleted oil reservoir,” Int. J. Hydrogen Energy, vol. 46, no. 34, pp. 18020–18031, May 2021, doi: 10.1016/j.ijhydene.2020.09.081.
[16] T. Wang, “The parameter estimation method of generalized Weng cycle model (Weng cycle model) and its applications in medicine,” in (Chinese), Chongqing Med. Univ., Chongqing, China, Tech. Rep., May 2013. [Online]. Available: https://kns.cnki.net/kcms/detail/detail.aspx?FileName=1011408806 nh&DbName=CMFD2014
[17] X. Zhang and G. Yu, “Research and application of solution method for generalized Weng model,” in (Chinese), Reservoir Eval. Develop., vol. 4, no. 6, p. 29, 2014.
[18] Y. Chen and C. Zo, “Model’s typical curve and its application for prediction production and recoverable reserves of oilfields,” in (Chinese), Acta Petrolei Sinica, vol. 35, no. 4, p. 749, 2014.
[19] B. Li, “Global peak oil prediction by application of multi-cycle Hubbert model,” Resour. Develop. Market, vol. 31, no. 4, pp. 427–430, 2015.
[20] D. Reynolds and M. Kolodziej, “North American natural gas supply forecast: The Hubbert method including the effects of institutions; ” Energies, vol. 2, no. 2, pp. 269–306, May 2009, doi: 10.3390/en20200269.
[21] H. Li and H. Wang, “China’s natural gas demand forecast based on the improved gray mode,” Henan Sci., vol. 32, no. 5, pp. 872–876, 2014.
[22] L. Chen, Z. Liu, and N. Ma, “Optimize production allocation for the oil-gas field basing on a novel grey model,” J. Natural Gas Geosci., vol. 4, no. 2, pp. 121–128, Apr. 2019.
[23] D. Wang, “Research on grey prediction models and their application in medium- and long-term power load prediction,” in (Chinese), Huazhong Univ. Sci. Technol., Wuhan, China, Tech. Rep., May 2013. [Online]. Available: https://kns.cnki.net/kcms/detail/detail.aspx?FileName=1013274351 nh&DbName=CFDF2014
[24] J. Wang, H. Jiang, Q. Zhou, J. Wu, and S. Qin, “China’s natural gas production and consumption analysis based on the multicycle Hubbert model and rolling grey model,” Renew. Sustain. Energy Rev., vol. 53, pp. 1149–1167, Jan. 2016.
[25] X. Shuai, “One natural gas production prediction grey model based on exponential smoothing,” (in Chinese), Comput. Technol. Develop., vol. 20, no. 3, pp. 243–245, 2010.

[26] P. Ji, “National coal bed gas production forecast based on ARIMA model,” (in Chinese), Coal Econ. Res., vol. 39, no. 11, pp. 29–32, 2019.

[27] J. Gao and X. Xu, “Forecast of China’s natural gas consumption based on ARIMA model,” Math. Pract. Theory, vol. 9, pp. 44–50, 2015.

[28] Q. Wang, S. Li, R. Li, and M. Ma, “Forecasting US shale gas monthly production using a hybrid ARIMA and metabolic nonlinear grey model,” Energy, vol. 160, pp. 378–387, Oct. 2018.

[29] Y. Wang, Z. Li, and C. Liu, “Prediction for coal bed gas output based on fractal and ARIMA,” Natural Gas Oil, vol. 29, no. 3, pp. 45–48, 2011.

[30] S. Du, R. Wang, C. Wei, Y. Wang, Y. Zhou, J. Wang, and H. Song, “The connectivity evaluation among wells in reservoir utilizing machine learning methods,” IEEE Access, vol. 8, pp. 47209–47219, 2020.

[31] H. Song, S. Du, R. Wang, J. Wang, Y. Wang, C. Wei, and Q. Liu, “Potential for vertical heterogeneity prediction in reservoir basing on machine learning methods,” Geofluids, vol. 2020, pp. 1–12, Aug. 2020.

[32] Q. Zhang, C. Wei, Y. Wang, S. Du, Y. Zhou, and H. Song, “Potential for prediction of water saturation distribution in reservoirs utilizing machine learning methods,” Energies, vol. 12, no. 19, p. 3597, Sep. 2019.

[33] A. Sagheer and M. Kohb, “Time series forecasting of petroleum production using deep LSTM recurrent networks,” Neurocomputing, vol. 323, pp. 203–213, Jan. 2019.

[34] O. Laib, M. T. Khadir, and L. Mihaylova, “Toward efficient energy systems based on natural gas consumption prediction with LSTM recurrent neural networks,” Energy, vol. 177, pp. 530–542, Jun. 2019.

[35] J. Devaraj, R. M. Elavarasan, G. Shafiuilah, T. Jamal, and I. Khan, “A holistic review on energy forecasting using big data and deep learning models,” Int. J. Energy Res., vol. 45, no. 9, pp. 13489–13530, Jul. 2021, doi: 10.1002/er.6679.

[36] K. Yudantaka, J.-S. Kim, and H. Song, “Dual deep learning networks based load forecasting with partial real-time information and its application to system marginal price prediction,” Energies, vol. 13, no. 1, p. 148, Dec. 2019, doi: 10.3390/en13010148.

[37] M. Godinho and R. Castro, “Comparative performance of AI methods for wind power forecast in portugual,” Wind Energy, vol. 24, no. 1, pp. 39–53, Jan. 2021, doi: 10.1002/we.2556.

[38] W. Qiao, K. Huang, M. Azimi, and S. Han, “A novel hybrid prediction model for hourly gas consumption in supply side based on improved whale optimization algorithm and relevance vector machine,” IEEE Access, vol. 7, pp. 88218–88230, 2019.

[39] W. Qiao and Z. Yang, “An improved dolphin swarm algorithm based on kernel fuzzy C-means in the application of solving the optimal problems of large-scale function,” IEEE Access, vol. 8, pp. 2073–2089, 2020.

[40] W. Qiao, Z. Yang, Z. Kang, and Z. Pan, “Short-term natural gas consumption prediction based on Volterra adaptive filter and improved whale optimization algorithm,” Eng. Appl. Artif. Intell., vol. 87, Jan. 2020, Art. no. 103323.

[41] China Statistics Yearbook-2020, (in Chinese), Nat. Bur. Statist., Beijing, China, 2020.

[42] Department Data: Coal, Electricity, Oil, Gas and Transportation, (in Chinese), Nat. Develop. Reform Commission People Republic China, Beijing, China, 2020.

[43] Y. Sakamoto, M. Ishiguro, and G. Kitagawa, Akaike Information Criterion Statistics, vol. 81, no. 10. Dordrecht, The Netherlands: D. Reidel, 1986, p. 26853.

[44] W. Pan, “Akaike’s information criterion in generalized estimating equations,” Biometrics, vol. 57, no. 1, pp. 120–125, Mar. 2001.