Journal of Petroleum Science and Engineering 192 (2020) 107240

Contents lists available at ScienceDirect

Journal of Petroleum Science and Engineering

journal homepage: http://www.elsevier.com/locate/petrol

Daily natural gas price forecasting by a weighted hybrid data-driven model

Jianliang Wang [a,b], Changran Lei [c], Meiyu Guo [d,*]

[a] School of Economics and Management, China University of Petroleum, Beijing 102249, China
[b] Research Center for China's Oil and Gas Industry Development, China University of Petroleum, Beijing 102249, China
[c] College of Software, Beihang University, Beijing 100191, China
[d] Department of Geography, Hong Kong Baptist University, China

ARTICLE INFO

Keywords:
Natural gas price
Forecasting
Hybrid model

ABSTRACT

With the role of natural gas gaining increasing importance in the transition of the world energy system and addressing global climate change, accurate prediction of the price of natural gas becomes crucially important. This paper first introduces three widely used individual data-driven models, i.e., support vector regression (SVR) and long-term and short-term memory network (LSTM), and a modified data-driven model, i.e., the improved pattern sequence similarity search (IPSS). A new weighted hybrid data-driven model based on these three models is then proposed. To train the model, data regarding the daily natural gas spot price in the U.S. prior to June 2018 are used and the model's prediction ability is tested using data from June 2018 to May 2019. The results show that the new IPSS model can predict the daily price of natural gas accurately. In a comparison of prediction errors with other individual models, the proposed hybrid model demonstrated the highest prediction ability of all of the investigated models.

1. Introduction

Climate change is one of the main concerns regarding the sustainable development of human society (Wang et al., 2017). Usage of high-carbon energy sources such as coal and oil has been considered to be the most likely reason for the observed climate change (IPCC, 2013). The transition to a zero- or low-carbon energy system offers the most effective solution to the global climate problem (Wang et al., 2018a). Natural gas plays a key role and is considered to be a bridge energy for this transition (Vidic et al., 2013). Therefore, many countries around the world are promoting the switch to natural gas to reduce carbon emissions. In the gas market, price is a key factor influencing the consumption of natural gas (Wang and Lin, 2017; Salehnia et al., 2013). Furthermore, gas price also significantly affects the production behaviors of natural gas suppliers and the economic development of countries. Given the importance of gas price, it is necessary to get more accurate results.

A review of the existing literature shows that many studies have already addressed energy price forecasting, especially for oil prices. The most widely used methods are various types of econometrics models. For example, Lanza et al. (2005) used cointegration and error correction models to forecast daily oil prices from 01/2002 to 06/2002. Sadorsky (2006) used several generalized autoregressive conditional heteroskedasticity (GARCH) models to forecast daily volatility in oil price returns from 02/1998 to 01/2003. Baumeister and Kilian (2012) established a vector autoregressive (VAR) model to forecast monthly crude oil price from 01/1991 to 12/2010. Xiang and Zhuang (2013) used an autoregressive integrated moving average (ARIMA) model to forecast the Brent daily crude oil prices from 11/2012 to 04/2013. Klein and Walther (2016) used a mixture memory GARCH model to forecast the daily oil prices from 01/1995 to 12/2014. Hailemariam and Smyth (2019) applied a structural heterogeneous autoregressive VAR (SHVAR) model to analyze the volatility of monthly natural gas prices from 01/1978 to 07/2018.

Econometric models, however, usually assume data to be stationary, regular, and linear and cannot accurately forecast time series that are complex, irregular, nonlinear, or have dynamic and random features (Wang et al., 2018b). In reality, energy prices tend to be irregular and nonlinear. Recently, with the development of machine learning techniques, many new data-driven methods have been proposed for energy price forecasting, which can effectively distinguish random factors and capture hidden nonlinear features (Herrera et al., 2019). For example, Moshiri and Foroutan (2006) used a neural network (NN) model to forecast nonlinear crude oil futures prices. Xie et al. (2006) predicted

* Corresponding author.
E-mail address: meiyuguo@hkbu.edu.hk (M. Guo).

https://doi.org/10.1016/j.petrol.2020.107240
Received 28 December 2019; Received in revised form 16 March 2020; Accepted 25 March 2020
Available online 1 April 2020

crude oil prices using support vector machines (SVM). Guo et al. (2012) then used an improved SVM model to forecast crude oil prices. A semi-supervised learning model has been used to forecast the movement direction of crude oil prices (Shin et al., 2013). Gene expression programming is also applied to forecast crude oil prices (Mostafa and El-Masry, 2016). Some different types of deep learning models have also been proposed for crude oil price forecasting (Zhao et al., 2017; Chen et al., 2017). An integrated data fluctuation network (DFN) and several artificial intelligence (AI) algorithms have been established to predict crude oil prices (Wang et al., 2018b).

According to Armstrong (2001), no single forecasting method dominates any other, each one having pros and cons. Hybrid models have been proven to have better forecasting accuracies than single-type models (Zhang et al., 2015; Naderi et al., 2019), and many such hybrid methods have been proposed. One type of hybrid method, called the traditional hybrid model, combines several different types of traditional econometrics models or/and data-driven methods. For example, Zhang et al. (2015) proposed a new model that combines the least squares SVM together with particle swarm optimization (LSSVM–PSO) and the GARCH model to forecast crude oil prices. Kristjanpoller and Minutolo (2016) established a hybrid model that combines an artificial neural network (ANN) with a GARCH model to predict oil-price-return volatility. Dritsaki (2018) used a hybrid ARIMA-GARCH model to forecast oil price. The above hybrid methods could improve the forecast performance. However, Herrera et al. (2019) compared the long-horizon forecast performances of traditional hybrid models with those of data-driven models for the six main global energy commodities with respect to their monthly prices, and their results showed that data-driven methods outperform traditional hybrid methods. Therefore, another class of hybrid methods, named as the new hybrid models (i.e., that combine different types of data-driven methods), has been proposed. For example, Yu et al. (2014) established a machine learning technique that combines NN and support vector regression (SVR) for crude oil price forecasting. Chiroma et al. (2015) proposed a hybrid crude oil price forecasting model that combines an NN with a genetic algorithm. Compared with the first type of hybrid method, the second type has rapidly developed in recent years. The main focus of this paper is the second type of hybrid method. Among these data-driven methods, the most typical and commonly used are NN and SVM (Zhang et al., 2015; Papadimitriou et al., 2014), with many hybrid models including various types of NN and/or SVM, such as those proposed by Dai and Cao (2017), Zhao et al. (2017), and Su et al. (2019).

In addition to the econometrics and machine-learning models mentioned above, the sequence similarity search—a data-driven technique based on data mining—is also an important and popular method for the mining and forecasting of big time-series data (Sakurai et al., 2015). Similar to machine-learning techniques, the similarity search can achieve automatic mining and forecasting for complex and non-linear data series (Sakurai et al., 2015; Palpanas, 2015). The traditional application areas of the sequence similarity search are mainly medical analysis, financial analysis, and online text (Pearson, 1998; Sakurai et al., 2015; Lv and Hao, 2017). However, recently, some studies have attempted to introduce the sequence similarity search into energy price forecasting. For example, Yuan et al. (2003) used a sequence similarity search in data mining to predict the electricity price and found this method to be suitable for electricity price forecasting. Qiu (2017) also established a pattern-sequence-similarity-search model and used it to forecast short-term electricity prices. Wang and Lei (2020) used the sequence-similarity-search technique to forecast the US daily gas price and found the forecast performance to be excellent.

Hence, based on the above literature review, we can conclude that: 1) current studies of energy price forecasting are mainly focusing on crude oil. Only a few studies focus on natural gas price, such as Čeperić et al. (2017), Su et al. (2019a, 2019b), and Naderi et al. (2019). Given the growing importance of natural gas worldwide, it is essential to pay more attentions on natural gas price forecasting and fill the literature gap; 2) for complex and non-linear time-series data, data-driven methods tend to perform better than traditional econometrics models; 3) hybrid methods that combine different types of data-driven methods demonstrate better forecast performance; 4) the sequence-similarity-search technique, which is another type of data-driven method, has the capability to model the complex characteristics of big time-series data automatically, and could represent an effective emerging method for energy price forecasting. The application of this method will increase modeling accuracy.

Based on the above conclusions, this paper focuses on gas price and its efficient prediction by a hybrid method. The contribution of this paper is the application of a hybrid forecast model that represents an innovative combination of two widely used data-driven methods (i.e., NN and SVM) and one emerging data-driven method (i.e., the sequence similarity search) for energy price forecasting. The first two individual data-driven models have already been widely used in energy price forecasting areas, this paper conducted important tests on their suitability in gas price forecast. More importantly, this paper provided valuable modifications and improvements on third individual model (i.e., PSS model). Therefore, the novelty of this paper mainly focuses on the following three points: 1) By modifying the original sequence similarity search, an improved sequence similarity search model is proposed; 2) the empirical analysis shows that the improved sequence similarity search model has similar forecast capacity with current widely used models, which therefore can be used in the area of gas price forecasting. By doing so, we believe that the hybrid model by including already widely used models and the modified new emerging model could benefit the gas price forecast study.

The remainder of this paper is organized as follows. Section 2 introduces the methods. Section 3 describes the prediction results for the price of natural gas. Section 4 summarizes the main findings of this paper.

2. Methodology

2.1. Sequence-similarity-search model and its improvement

2.1.1. Sequence-similarity-search model

Sequence similarity searching is a method for searching pattern sequences to identify those that are similar to the target sequence (which is also named as pattern sequence similarity search, i.e., PSS model). As mentioned in Introduction, RSS model has been rarely used in energy price forecast, although this model has been widely used in other areas. One purpose of this paper is to introduce this model into energy price forecast area and evidence its usability. The procedure for PSS in a time series of natural gas prices is as follows: first, all of the pattern sequences are obtained using a data mining algorithm; second, the current to-be-searched sequence (i.e., target sequence) is intercepted from the entire sequence set; third, searching in the historical sequences is performed if a sequence similar to the target sequence is found, then the price of the next day of the sequence is included in the result set; lastly, the result set is processed to obtain the final prediction result. In the above analysis process, the acquisition of all pattern sequences and the definition of similarity are two key factors. Therefore, we present a detailed introduction of these two aspects.

1) Acquisition of all pattern sequences

Currently, the most popular method used to obtain all pattern sequences is the cluster analysis algorithm (Hu et al., 2007). Cluster analysis uses a classical unsupervised learning algorithm that can divide data into groups. There are many different algorithms used for cluster analysis, including the hierarchical clustering algorithm, segmentation clustering algorithm, constraint-based clustering algorithm, a clustering algorithm for machine learning, and a clustering algorithm for high-dimensional data (He et al., 2007). Of these, the K-means

algorithm, a type of segmentation clustering algorithm, is a common and suitable method for analyzing time series data and is widely used to acquire pattern sequences from time series data (Martínezálvarez et al., 2007).

2) Definition of similarity

After obtaining all pattern sequences by cluster analysis, the next step is to perform a similar-sequence search. This search work involves a data search and comparison process, which is generally implemented using a computer program. The key to achieving the best results is to ensure a reasonable judgment of similarity. This leads to the concepts of the difference degree $\boldsymbol{\theta}(\boldsymbol{\theta} \geq 0)$ and sequence similarity $\boldsymbol{D}(\boldsymbol{D} \geq 0)$. For a given degree of difference $\boldsymbol{\theta}$, if the similarity between sequence values $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ is $\boldsymbol{D} \leq \boldsymbol{\theta}$, it can be considered that sequences $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are similar.

Measuring similarity is a very important research topic in the field of time series similarity searching, and is also the basis for the implementation of sequence similarity searches. The Euclidean distance is a widely used metric for measuring the similarity between sequences (Zhang and Chen, 2011; Li et al., 2016). The core idea of Euclidean distance is the point distance in high-dimensional space. In high-dimensional space, the sequence S of length n is a point in the n-dimensional space. Calculating the Euclidean distance involves calculating the distance between two points in the n-dimensional space. The specific calculation formula is as follows:

$$D(\alpha,\beta) = \sqrt{\sum_{i=1}^{n} [\alpha(i) - \beta(i)]^2}, \tag{1}$$

where $\boldsymbol{\alpha}(\boldsymbol{i})$, $\boldsymbol{\beta}(\boldsymbol{i})$ represent the $\boldsymbol{i}^{th}$ elements of the sequences $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, respectively, and $\boldsymbol{n}$ represents the number of sequence elements.

2.1.2. Model improvement

The traditional PSS model described above has achieved good prediction results in many fields. However, the model still has some defects and shortcomings, which lead to discrepancies between the prediction results and the actual data. Therefore, in this paper, we introduce two improvements into the traditional method and propose the improved pattern sequence similarity search (IPSS) model.

1) Automatic adjustment mechanism for setting $\boldsymbol{\theta}$

When matching historical sequences, it is necessary to set an appropriate degree of difference, i.e., $\boldsymbol{\theta}$. In the traditional method, the degree of difference is generally set according to experience. If $\boldsymbol{\theta}$ is set too high, a large number of sequences are introduced into the result set (including many sequences that have low similarity with the target sequence), which results in errors in the prediction. If $\boldsymbol{\theta}$ is set too low, there will be cases in which no similar sequence can be found, that is, the result set is empty and no predicted price is obtained. To avoid the occurrence of error conditions and improve prediction accuracy, in this study, we designed an automatic adjustment mechanism, the principle of which is to check the result set after all of the retrievals have been completed. If the result set is found to be empty, the similarity requirement is lowered, and then the retrieval process is repeated. The specific procedure is as follows:

Step 1 Search the historical database; if the similarity meets the requirements, then include the price of the next-day data in the result set.
Step 2 After the search is completed, determine whether the result set is empty. If not, go to step four. If it is empty, go to step three.
Step 3 Start the adjustment mechanism, appropriately reduce the difference requirement, and repeat steps 1 and 2.
Step 4 Process the result set.

Obviously, the choice of $\boldsymbol{\theta}$ in the mechanism will affect the speed of data processing, and it can be adjusted according to the actual time limit.

2) Weighted average calculation for final result

Many searched results are included in the result set. The traditional way to obtain the final predicted result is to use a simple averaging method. Specifically, if the result set contains n searched prices, then the final prediction price, F, is:

$$F = \frac{\sum_{i=1}^{n} p_i}{n}. \tag{2}$$

This algorithm assumes that the matching sequences of different periods have the same effect on the final result. Obviously, this assumption has many limitations. Taking two matching sequences $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ as an example, if $\boldsymbol{\alpha}$ is closer to the target sequence in time than $\boldsymbol{\beta}$, which means that the difference between the external environment of $\boldsymbol{\alpha}$ and the target sequence tends to be smaller, then $\boldsymbol{\alpha}$ has a greater impact on the final result than $\boldsymbol{\beta}$. Therefore, it cannot be simply assumed that the matching sequences of different periods contribute in the same degree to the results. A weighted average method is proposed in this paper to deal with the above issue. After the historical database has been retrieved and the result set is determined to not be empty, the weight $\boldsymbol{\omega}$ is introduced for the processing of the result set to quantitatively reflect the particular contribution of a searched result to the final result. The definition of $\boldsymbol{\omega}$ is as follows:

$$\omega = \exp\left(-\frac{Len}{2\Omega^2}\right), \tag{3}$$

where $\boldsymbol{Len}$ is the time distance between the searched result and the target sequence. When $\boldsymbol{Len}$ is very small, that is, the distance is very small, the weight $\boldsymbol{\omega}$ is close to 1. Conversely, when $\boldsymbol{Len}$ is very large, the weight is a very small number that tends to zero. Applying this definition will result in higher weights for sequences that are closer to the target sequence. The role of the parameter $\boldsymbol{\Omega}$ is to control the proportion of weight increase or decrease.

By introducing the weight $\boldsymbol{\omega}$, the final predicted result F is:

$$F = \frac{\sum_{i=1}^{n} (p_i \cdot \omega_i)}{n}. \tag{4}$$

2.2. Support vector regression (SVR) model and its optimization

2.2.1. Support vector regression model

The SVM, a classic machine learning model originally proposed by Cortes and Vapnik (1995), was designed for solving classification problems. Therefore, it is more suitable to be used for classification instead of regression, although many studies use it to solve regression problems. To better deal with the regression problems, researchers expanded the SVM model and proposed SVR model (Drucker et al., 1996; Mohammadi et al., 2015), which provides a good solution for regression based on small-sample, nonlinear, and high-dimensional problems. Forecasting gas prices is a regression problem, therefore, SVR model is used in this paper.

The SVR prediction model maps the input parameters to high-dimensional spaces through the nonlinear mapping relationship $\phi(x)$, and discriminates between the nonlinear relationships in the original space by constructing linear functions in high-dimensional space. In the training sample set $\{x_i, y_i\} \subset R_i \times R, i = 1, 2, \ldots, n$, $\boldsymbol{x}$ is the input value and $\boldsymbol{y}$ is the output value. The nonlinear mapping equation maps the original spatial data into a high-dimensional space, and returns the data in the high-dimensional space by $f(x) = [w \cdot \phi(x)] + b$, where $\boldsymbol{b}$ is the threshold and $\boldsymbol{w}$ is the weight vector. According to the structural risk minimization principle of the SVR prediction model, the learning process of the model is transformed into the following convex optimization problem:

$$min\frac{1}{2}w^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \tag{5}$$

$$s.t.\begin{cases} y_i - w\cdot\phi(x_i) - b \le \varepsilon + \xi_i \\ w\cdot\phi(x_i) + b - y_i \le \varepsilon + \xi_i \\ \xi_i, \xi_i^* \ge 0 \end{cases},$$

where ξ_i, ξ_i^* are slack variables, C is a penalty factor, and ε is an insensitive loss function.

According to the dual theory and the kernel function theory (Smola and Schölkopf, 2004), the above problem is transformed into its dual problem:

$$max\left\{-\frac{1}{2}\sum_{i=1,j=1}^{n}(\beta_i-\beta_i^*)\left(\beta_j-\beta_j^*\right)K(x_i,x_j) + \sum_{i=1}^{n}\beta_i^*(y_i+\varepsilon) - \sum_{i=1}^{n}\beta_i(y_i-\varepsilon)\right\} \tag{6}$$

$$s.t.\sum_{i=1}^{n}(\beta_i-\beta_i^*) = 0\ 0 \le \beta_i, \beta_i^* \le C.$$

In the above, $K(x_i, x_j)$ is the kernel function of SVR and β_i, β_i^* are Lagrange parameters. The inner product operation in high-dimensional space can be realized by the function of the input space through the kernel function, which effectively avoids dimension disaster. In SVR applications, the Gaussian radial basis function (RBF) is often used as a kernel function. The RBF is defined as follows:

$$K(x_i, x_j) = \exp\left(\frac{-x_i - x_j^2}{2\sigma^2}\right). \tag{7}$$

In the above, x_i, x_j represents the feature vector that must perform the high-dimensional inner product operation and σ is a free parameter, also known as the width of the kernel function.

The final SVR regression function is:

$$f(x) = \sum_{i=1}^{n}(\beta_i - \beta_i^*)K(x_i, x_j) + b. \tag{8}$$

When $(\beta_i - \beta_i^*)$ is not zero, its corresponding training sample is called a support vector.

2.2.2. Model optimization method

In the SVR model, three parameters have an important impact on the results, i.e., ***C***, ***σ***, and ε. When ***C*** is too large, it will cause over-fitting of the model, and when ***C*** is too small, this will lead to increased training error. ***σ*** affects the distribution complexity of data samples in high-dimensional space. When ***σ*** is too small, the generalization ability of the model is poor, and when it is too large, it is difficult to achieve preset precision. ***ε*** reflects the sensitivity of the model to input noise. When ***ε*** is too small, the learning ability of the model is poor and when it is too large, it will affect the generalization ability of the model.

The accuracy and generalization of the SVR model largely depend on the above three parameters and their interactions. Therefore, to obtain an accurate result and to improve the model's generalization, it is important to find an algorithm or method that can obtain optimized parameter values. Common methods include cross-validation and grid verification (Bao et al., 2004; Espinoza et al., 2007). The main disadvantage of those two methods is their heavy requirement of time and computing resources (Bao et al., 2004). Given this drawback, to identify the optimal combination of SVR parameters, researchers have proposed the use of a particle swarm optimization (PSO) algorithm within a certain range, which has been widely used due to its high efficiency and good accuracy (Cao and Xu, 2007; Hong, 2009). Therefore, PSO algorithm is used in this paper to determine the optimal combination of parameters for the SVR model. In addition, the mean square error function for the real and predicted values is selected as the objective function of the SVR model.

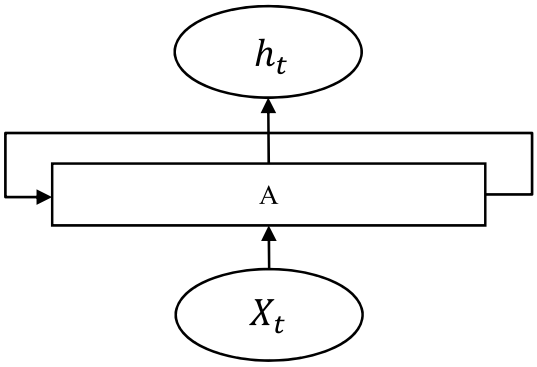

Fig. 1. Recurrent neural network structure.

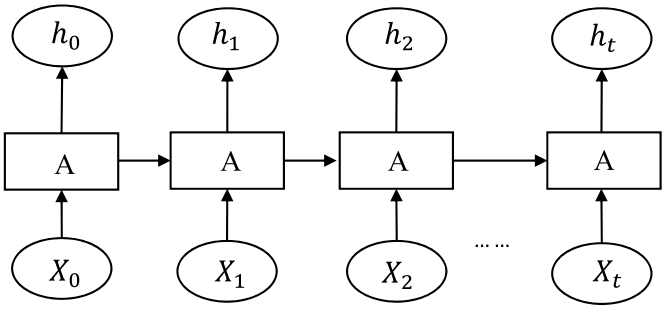

Fig. 2. Recurrent neural network expansion structure.

$$Obj = \frac{1}{n}\sum_{i=1}^{n}\left(y_i^* - y_i\right)^2. \tag{9}$$

In the above, n is the number of price data, y_i^* is the predicted value of the SVR prediction model and y_i is the actual value.

2.3. Long- and short-term memory network prediction model

NN is one of the widely used models in energy price forecast. The latest progress and future research directions for NN mainly focus on two areas, i.e., convolution neural networks (CNN) and recurrent neural networks (RNN) (Zhang et al., 2018). CNN is widely used in image processing areas, while RNN is mainly used in regression areas (Zhang et al., 2018). Gas price forecast is a regression problem, therefore, RNN is used in this paper. Furthermore, RNN often has the problem of gradient disappearance in the actual application process, that is, the perception of the previous time node decreases. To solve the above problems faced by RNN, scholars proposed LSTM (Zhang et al., 2019), which is an evolutionary model from RNN. Therefore, LSTM is used in this paper.

To better understand the LSTM, we first briefly introduce the RNN. The RNN contains a loop structure that senses information at each moment on the network, so the RNN output is affected not only by the current input, but also the historical input (Assaad et al., 2008). Because of this memory characteristic, RNN has an advantage in time-series prediction (Kumar et al., 2004). Fig. 1 shows the classical RNN structure.

In Fig. 1, module A of the RNN reads the input X_t to obtain the output value h_t. The loop facility can pass information from the current step to the next step. In fact, the RNN can be thought of as multiple copies of the same neural network, with each neural network module passing a message to the next module. If we expand this loop, we get the structure shown in Fig. 2, from which we can see that the output $h\_t = f(X\_t, h\_t\text{-}1)$ where h is the hidden state.

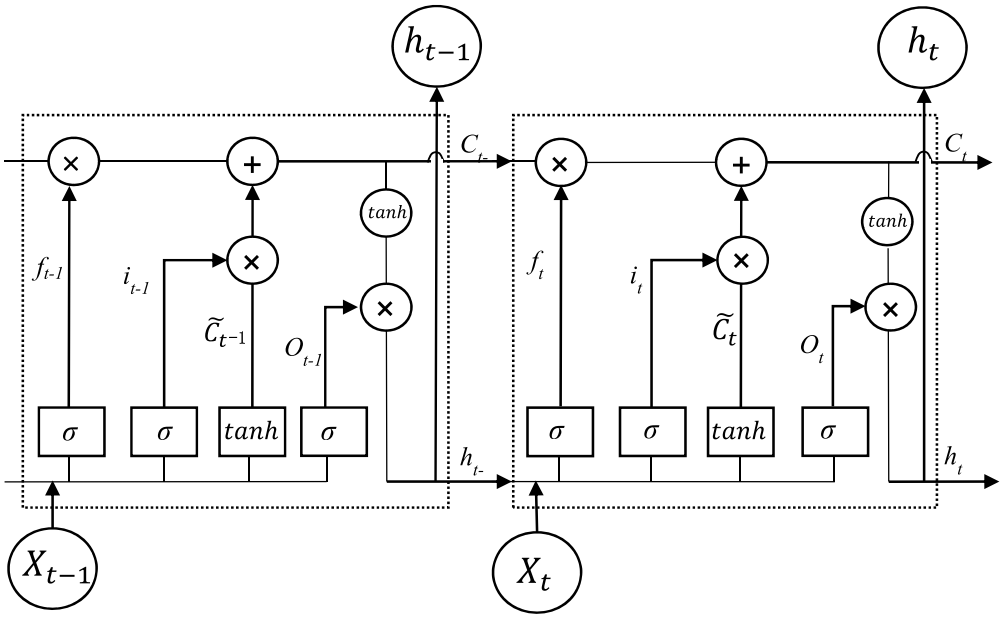

Fig. 3. LSTM structure.

LSTM is similar to a standard cyclic neural network with a hidden layer, but the nodes in each normal hidden layer are replaced by storage units. The LSTM model adds input gates, output gates, and "forget gates" to the neuron portion of the RNN. This structural design effectively solves the problem of gradient disappearance, which makes LSTM very suitable for dealing with long-term dependency problems. LSTM has been widely applied in the field of time series prediction, and has achieved good results (Peng et al., 2018; Kuo and Huang, 2018). Fig. 3 shows the specific structure of the LSTM model.

As can be seen from the above figure, the LSTM network system is similar to the general NN system in that it has an input and output structure. In addition to the activation function, there are many processing components in the core layer, which are described in detail below.

First, LSTM uses the forget gate to determine which information in the cell state must be discarded and which needs to be retained. The forget gate will read h_{t-1} and X_t, and use the *sigmoid* function to calculate the forget gate value f_t:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, X_t] + b_f\right). \tag{10}$$

Then, it determines which information is stored in the cell state. This step has two parts. First, the *sigmoid* layer, which is called the "input gate layer," determines which values will be updated. Second, a new vector $\widetilde{C}_t$ is created by the *tanh* layer and then added to the state. Next, we use the following two messages to update the status.

$$i_t = \sigma(W_i \cdot [h_{t-1}, X_t] + b_i) \tag{11}$$

$$\widetilde{C}_t = tanh(W_c \cdot [h_{t-1}, X_t] + b_c) \tag{12}$$

Next, we must determine how to update the information. We can do so by updating the old cell state C_{t-1} to C_t using Equation (13). This new candidate value will change based on the extent to which each state is updated.

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \widetilde{C}_t \tag{13}$$

Finally, the "output gate" is used to determine which information to output, based on the state of the cell. First, the *sigmoid* layer is used to determine the state of the cell to be output. Then, the cell state is processed through the *tanh* layer (obtaining a value between −1 and 1) and multiplied by the output of the *sigmoid* gate. Finally, only the portion that determines the output is output.

$$O_t = \sigma(W_O \cdot [h_{t-1}, X_t] + b_O) \tag{14}$$

$$h_t = O_t \cdot \tanh(C_t) \tag{15}$$

In the above, W_i, W_f, W_c, and W_O are weight vectors, and b_i,b_f,b_C, and b_O are corresponding deviation vectors.

2.4. Weighted hybrid model that combines three methods

In this study, we used a hybrid model that combines three basic models, i.e., IPSS, SVR, and the LSTM, to predict the short-term natural gas price. The basic principles of the hybrid model are as follows:

For a natural gas price series $\{x_t, t = 1, 2, \ldots, n\}$ of length n, the price prediction results of the three individual models, i.e., IPSS, SVR, and LSTM, are x_{1t}, x_{2t}, and x_{3t}, respectively. Let w_i be the weighting coefficient of the individual prediction model i, and $\sum_{i=1}^{3} w_i = 1$, then the prediction price of the hybrid model is:

$$\widehat{x}_{it} = \sum_{i=1}^{3} w_i x_{it}, \tag{16}$$

where $\widehat{x}_{it}$ is the final forecast price of the hybrid model at time t and x_{it} is the forecast price of the individual model i at time t.

From Equation (16), we can see that the weight is a crucial component in the hybrid model. To date, scholars have proposed a variety of weight determination methods, including the equal-weighted-average combination model (EWA), the variance–covariance preferred combination model (VCP), and the variance reciprocal method (Herrera et al., 2019; Zhu et al., 2011; Safari and Davallou, 2018). These three methods are provided as follows:

- EWA method assigns the same weights for each individual model. This is a very simple and easy calculated method. However, it does not reflect the principle of optimization, i.e., the good and bad models have same impacts on the final results. EWA method is suitable for the situation that the accuracy of each individual model is completely unknown (Gan, 2012).
- Compared to EWA, VCP method could provide a reasonable combination weight coefficient, which will improve the accuracy of prediction. However, the shortcoming of this method is that the weight

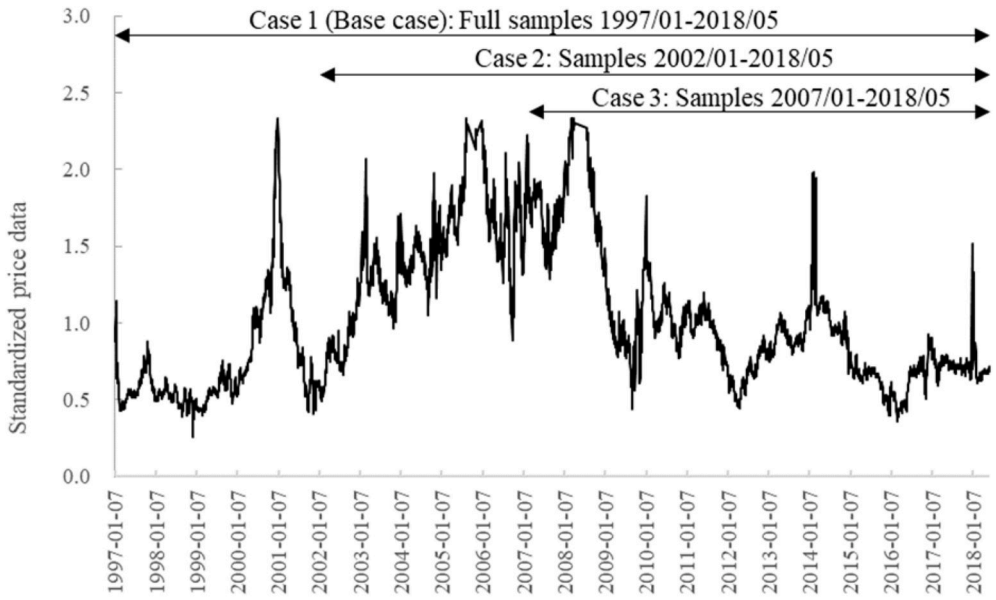

Fig. 4. Standardized training gas price data.

given by VCP is not stable. Furthermore, using this method will consume a lot of computing resources. Therefore, when the amount of data is large, it will affect the calculation efficiency (Jiang, 2002).

- The basic idea of the variance reciprocal method is that good model should have high impacts of the results, while bad models should have small impacts. The variance reciprocal method is based on the principle of the minimum sum of the squared error. It assigns a smaller weight to the larger-error square model and a larger weight to the model with the smaller square sum of errors. Compared to EWA and VCP, this method reflects the principle of optimization and has a stable weight (Chen, 2014; Liu et al., 2018). This method performs well in all respects and has received extensive and positive attention. Therefore, we use the variance reciprocal method to determine the weight.

To use the variance reciprocal method, we must first calculate the reciprocal of the sum of the squares of the individual model errors, and then divide the reciprocal sum by the respective reciprocals to obtain the respective weights. The specific calculation formula is as follows:

$$\omega_i = \frac{\varepsilon_i^{-1}}{\sum_{j=1}^{n} \varepsilon_j^{-1}} \quad i,j = 1,2,\ldots,n, \tag{17}$$

where ε_i is the squared errors of the ith individual model.

2.5. Performance evaluation criteria

To measure the prediction accuracy of the forecast models, two widely used indexes are applied in this paper (Wang et al., 2018b; Neupane et al., 2017), i.e., mean relative error (MRE) and mean absolute percentage error (MAPE), which are defined as:

$$MRE = \frac{1}{N}\sum_{j=1}^{N} \frac{\left|\widehat{x}_j - x_j\right|}{\bar{x}}, \tag{18}$$

$$MAPE = \frac{1}{N}\sum_{j=1}^{n} \frac{\left|\widehat{x}_j - x_j\right|}{x_j}, \tag{19}$$

where $\widehat{x}_j$ is the forecast price, x_j is the actual price, $\bar{x}$ is the mean price for the forecast period, and N is the number of test price data. Generally, the smaller are the MRE and MAPE, the greater is the degree of model accuracy.

3. Data processing and results

3.1. Data source and pre-processing

3.1.1. Data source

In this article, as an example, we use the prediction of the daily spot price of natural gas based on a data set from the Henry Hub Natural Gas Spot Price. This data set provides daily natural gas price data from the New York Mercantile Exchange (NYMEX) from 1997/01/07 to 2019/05/31 (see https://www.eia.gov/dnav/ng/ng_pri_fut_s1_d.htm, accessed on June 16, 2019). We divided the data into two parts, data from 1997/01/07-2018/05/31 as a training set; while data from 2018/06/01 to 2019/05/31 as test data.

3.1.2. Data pre-processing

In this paper, the pre-processing for original data includes three steps. The training data are used as an example to show how this pre-processing works.

1) Eliminating missing data samples

By downloading the data, we got 5385 data records. Data for 2018/1/5 is missing. Therefore, 5384 data points are left.

2) Processing outlier samples

The *Boxplot* method is used to identify outlier samples. The *Boxplot* is a graphical display where the outliers appear tagged. The first quartile and the third quartile of the sample size are labelled as Q_1 and Q_3. The *Interquartile Range* (IQR) is defined as "Q_3-Q_1". There are several methods for determining outliers in a sample. A very popular method is based on the following criterions are used in this paper, i.e. an observation data sample is declared as an outlier when it lies outside of the interval (Q_1-1.5IQR, Q_3+1.5IQR). By doing so, 177 outliers are identified from 5384 samples, which only accounting for 3.3% of total samples. Therefore, we eliminate these 177 outliers.

3) Standardizing data samples

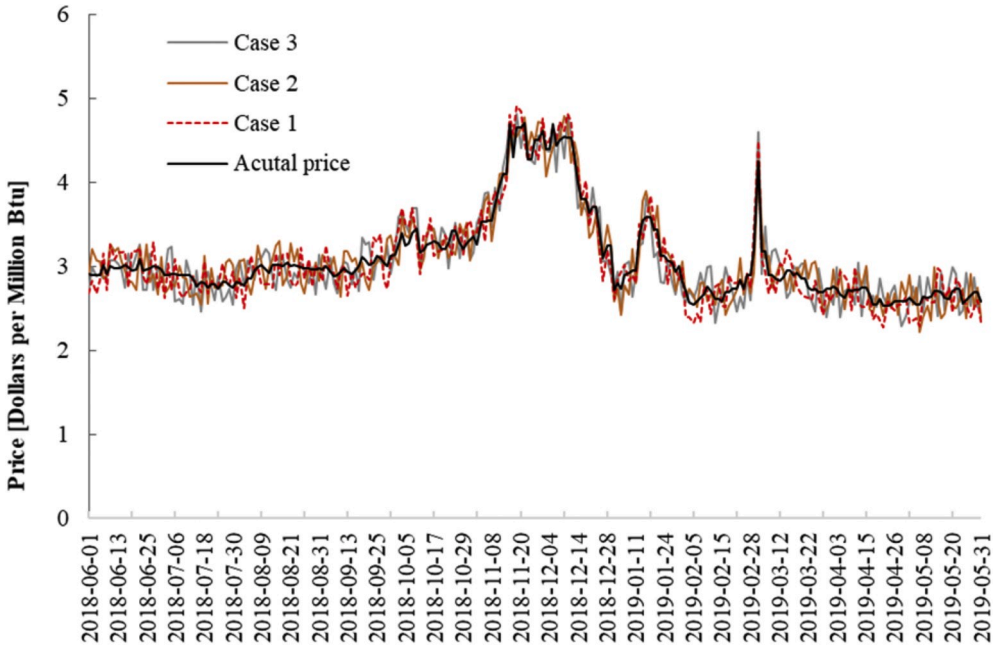

Fig. 5. Price forecast by using the hybrid model in three cases.

After eliminating missing data and outliers, 5207 data samples are left. To improve the efficiency and quality of the forecast, after obtaining the time series price of natural gas, it is generally necessary to first standardize the price data (Zhang and Zhu, 2010; Krizhevsky et al., 2012; He et al., 2016). In this study, we used the following commonly used proportional average method to standardize the historical data (Ke et al., 2017):

$$F_s = \frac{F_i}{\overline{F}}, \tag{20}$$

where F_i represents the raw data value of the natural gas price on the i-day and F_s represents the standardized data value. $\overline{F}$ represents the arithmetic mean of the data for the current month of the ith day, that is:

$$\overline{F} = \frac{1}{m}\sum_{i=1}^{m} F_i, \tag{21}$$

where m is the number of days in the month.

Standardized training price data for natural gas are shown in Fig. 4 (the detailed information for original data and pre-processing data is shown in *Supplement Information (SI)*). For a more intuitive display, the test data set was divided into 12 subperiods, that is, one period per month. We will use the proposed the model to conduct training tests on the data prior to June 2018 and then to forecast the prices from June 2018 to May 2019. The forecast prices were then compared with the actual prices to evaluate the forecast performances of the models. Furthermore, considering the impacts of the length of training data on prediction results, three cases are proposed here: 1) full samples (1997/01-2018/05) are used in case 1; 2) samples between 2002/01-2018/05 are used in case 2; 3) samples between 2007/01-2018/05 are used in case 3. It should be noted that the case 1 is seen as the base case in this paper.

In addition, to highlight the effective of the proposed hybrid model, several individual models are also used, i.e., PSS model, IPSS model, SVR model, LSTM model, GARCH model, ARIMA model and Gradient Boosting Machine (GBM) model. The first four models have already been introduced in this paper, and the last three models are widely used models by literature.

Table 1
Prediction errors in three cases.

Case	Training data period	Number of data points	MER	MAPE
Case 1	1997/01-2018/05	5207	4.99%	5.04%
Case 2	2002/01-2018/05	3970	5.19%	5.23%
Case 3	2007/01-2018/05	2811	5.49%	5.55%

3.2. Forecast performance of the hybrid model in the different cases

The overall forecast results and the prediction errors under three cases by using the hybrid model are shown in Fig. 5 and Table 1, respectively. We can see from the figure that the overall trends of the forecast prices in three cases are similar to those of the actual prices. However, the detailed forecast error and forecast performance could be different due to the effect of the length of training price data. From Table 1, we can see that with the increase of the length of training data, the prediction error reduces. For example, in case 3, 2811 data are used for training, then the values of MER and MAPE are 5.49% and 5.55%. While in case 2, 3970 data are used for training, and the values of MER and MAPE become 5.19% and 5.23%, which are lower than those in case 3. Furthermore, to considering effect degree of the time span of training data on the forecast results, we compare the change rate of the number of data points and the prediction errors under different cases. For example, comparing case 1 and case 2 shows that a reduction of nearly 24% in the number of training data points will result in a 4% increase in MER and 3.8% increase in MAPE. Therefore, the effect of the number of training data on forecast performance is not significant.

3.3. Forecast performance of different models in the base case

A comparison of forecast performance from our proposed hybrid model and other models in the base case is shown in Figs. 6 and 7. A very clear result is that for the overall performance, our proposed hybrid model has the lowest prediction error. Taking MAPE as an example, the MAPE of the hybrid model is 5.04% in the base case, while the MAPEs are 6.21% for IPSS, 6.47% for SVR, 7.35% for LSTM, 7.37% for GARCH, 7.99% for ARIMA, 8.17% for GBM and 8.54% for PSS. For the performance of each period, we can see that among 12 subperiods, the hybrid

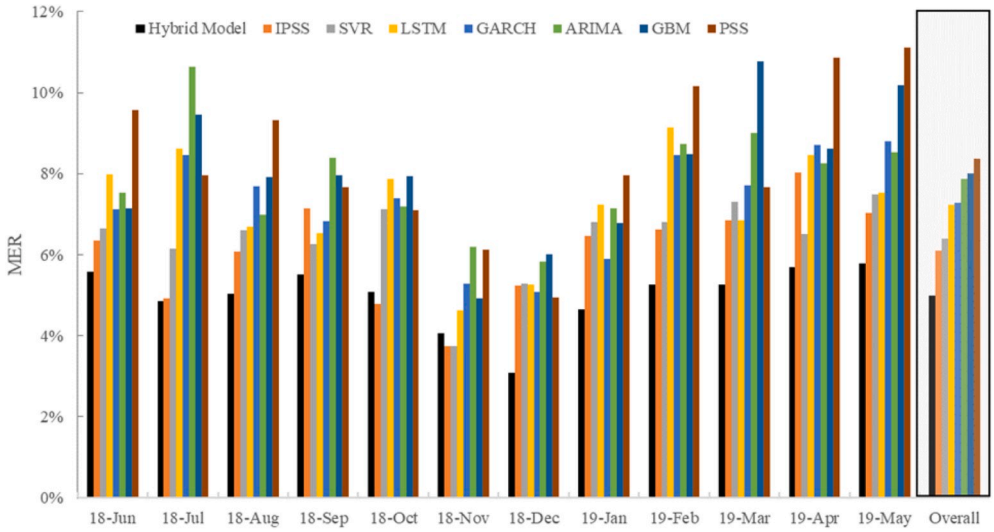

Fig. 6. MERs of different models in the base case.

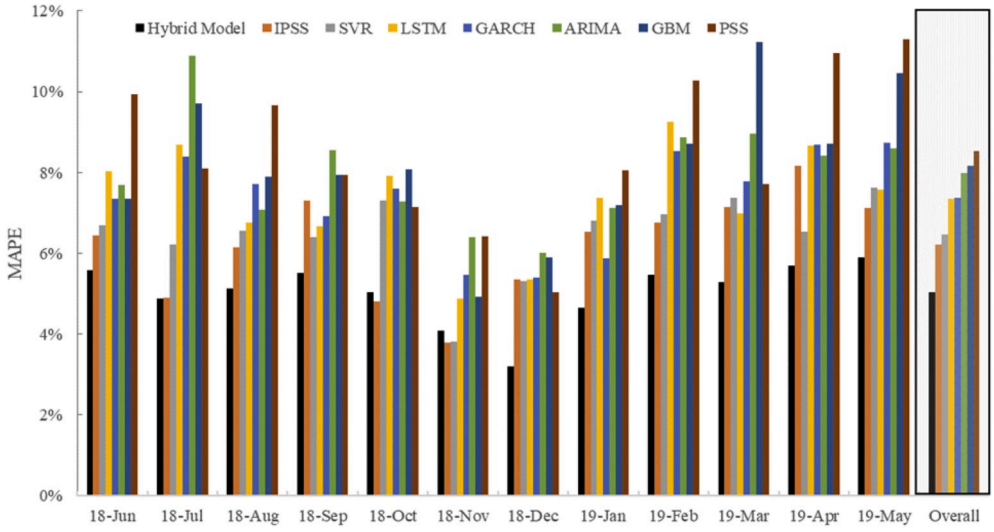

Fig. 7. MAPEs of different models in the base case.

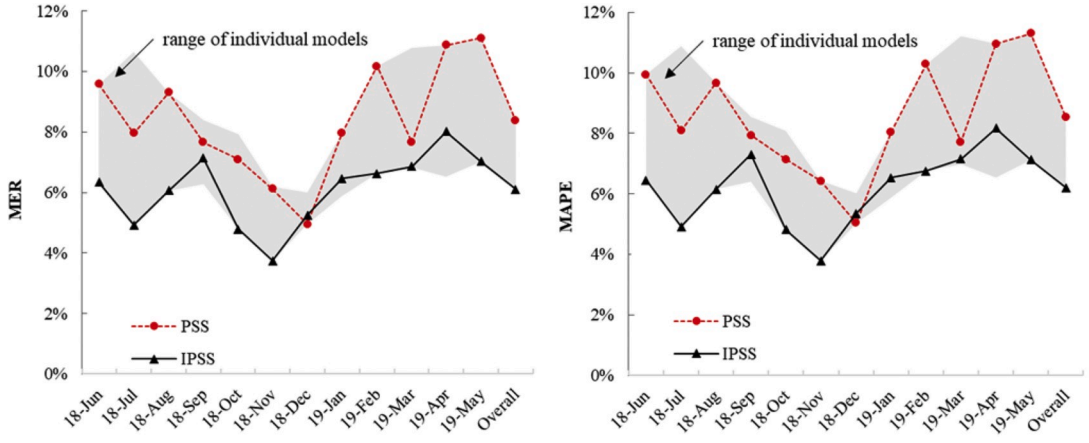

Fig. 8. Comparison of prediction errors of PSS and IPSS model in the base case.

Table 2
Prediction errors from current gas price forecast studies.

Sources	Type of model	Prediction error		Hybrid model in base case
Onour (2009)	Structure-based	RMSE	0.175*	0.174
Nguyen and Nabney (2010)	Structure-based	MAPE	1.8%*	5.04%
Viacaba et al. (2012)	Structure-based	RMSLE	0.0991	0.0188
Lin and Wesseh Jr (2013)	Data-based	MSE	0.0625–0.086	0.0301
Thakur et al. (2015)	Data-based	MSE	0.0727	0.0301
Čeperić et al. (2017)	Structure-based	MAPE	4.27%	5.04%
Naderi et al. (2019)	Data-based	MAPE	1.49%	5.04%
Siddiqui (2019)	Data-based	MSE	0.026–0.039	0.0301
Su et al. (2019a)	Structure-based	MAPE	11.2%–13.7%	5.04%
Su et al. (2019b)	Structure-based	MSE	0.4376	0.0301

Note: * is average value; MSE- Mean Squared Error; RMSE-Root Mean Squared Error; RMSLE- Root Mean Squared Logarithmic Error.

model has the best forecast capacity in 10 subperiods. In other two subperiods, i.e., 18-OCT and 18-NOV, the prediction error of the hybrid model is a little higher than IPSS and SVR, but is much lower than that of other models. Other two cases show similar results, and their results are not shown here to avoid redundancy. Therefore, we could conclude that our proposed model has the best forecast capacity during the investigated models.

3.4. Discussion on the effectiveness of the IPSS model

The IPSS model used here is a modified version of the PSS model. To demonstrate the strength of the improved model, a comparison of prediction errors of IPSS model and other individual models is shown in Fig. 8. We can see that original PSS model has the relatively highest prediction error among individual models, while the IPSS model has the relatively lowest prediction error among individual models. Other two cases show the similar results. Taking the overall MAPE as an example, the overall MAPE for PSS model is 8.54%, while this value for IPSS model is 6.21%, a 27.3% reduction compared to PSS model. Therefore, we could conclude that our modification on PSS model is effective.

3.5. Comparison of the prediction error with related studies

As we indicated in introduction, current studies on natural gas price forecast is relatively lacking compared to those oil price forecast studies. Furthermore, aims of some studies are actually to analyze the reasons of gas price fluctuations or the relationship between gas prices and other commodities' prices (Mishra, 2012). Therefore, there is no forecast for the future and no prediction errors (Buchanan et al., 2001; Yorucu and Bahramian, 2015). Some studies do provide the prediction errors and a comparison of their errors and the error from our proposed hybrid model in the base case is shown in Table 2. In Table 2, we define two types of models, i.e., structure-based model and data-based model. The first model forecasts gas price by considering many other factors, such as, oil price, gas production, gas consumption, gas imports, gas resources, temperature, gas rotary rigs, etc. While the second model forecasts gas price by using the historical gas price data only.

From Table 2, we can see that three studies do show lower prediction errors than our paper, i.e., Nguyen and Nabney (2010), Čeperić et al. (2017) and Naderi et al. (2019). The first two studies are structure-based models, the forecast performance largely depends on the factors and data they used. The choice of factors would be subjective, which may result in different factors and different prediction errors in different studies. The model in Naderi et al. (2019) is data-based model, the prediction error for gas price is only 1.49%, which is significant lower than our paper. However, this forecast performance is closely related to its data. In addition to gas price, Naderi et al. (2019) also forecast the oil price and interest rate by using the same method. MAPEs for oil price and interest rate are 5.45% and 8.06%, which are significant higher than gas price. The reason for this is that fluctuations for the data of oil price and interest rate are much significantly higher than that of gas price (see Fig. 9). A comparison of the gas price data in our paper (see Figs. 4 and 5) and those in Naderi et al. (2019) (see Fig. 9b) shows that the fluctuation in our paper is significantly higher than that in Naderi et al. (2019), which could be the main reason why this paper's MAPE is higher than Naderi et al. (2019). In addition to the above three studies, the prediction error in our paper is lower than others. In a word, we conclude that the prediction error of the hybrid model proposed in this paper are acceptable and could be used in gas price forecast.

4. Conclusions

4.1. Conclusions of this paper

Due to the complexity and nonlinearity of natural gas prices, as well as the limitation of individual models, in this paper, a hybrid model was proposed that combines three individual data-driven models (i.e., IPSS, SVR, and LSTM). The proposed model was then used to forecast the daily natural gas price in the U.S. Based on the results, the following conclusions can be made.

First, compared with the traditional PSS model and other individual models in this paper, the proposed IPSS model achieved higher forecast accuracy and therefore has higher forecast capacity. Second, the hybrid model achieved the lowest forecast error and highest prediction ability of all investigated models. Therefore, this weighted IPSS-SVR-LSTM hybrid model is suggested for use in gas price forecasting.

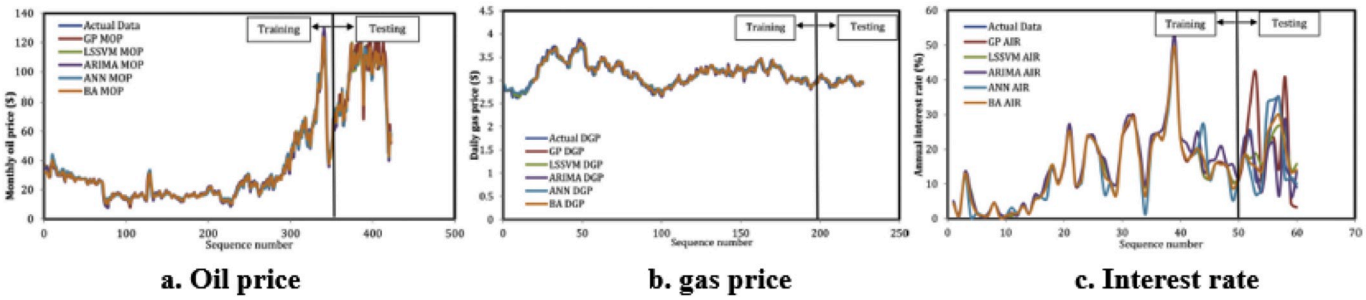

Fig. 9. Forecast results for oil price, gas price and interest rate in Naderi et al. (2019)
Source: this figure is from Naderi et al. (2019).

4.2. Limitations and future directions

Structure-based model and data-based model are two types of widely used model in price forecast. Each category has its own pros and cons. This proposed model in this paper belongs to the second category, which means some inherent shortcomings in this category are still existed. Furthermore, current model still can't capture the impacts of black swan events on forecast results. Therefore, in future, continuing improvement of data-based model is needed, such as considering the impacts of black swan events. In addition, we can see that the gas price in this paper has high variance. For these high variance data, precisely predicting future data points may not be realistic, especially when data is always influenced by multiple expected factors. Therefore, in future, converting the prediction target to be natural gas price range may be a feasible additive and should be study.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Jianliang Wang: Conceptualization, Methodology, Investigation, Writing - original draft, Supervision. **Changran Lei:** Data curation, Software, Validation, Formal analysis. **Meiyu Guo:** Conceptualization, Writing - review & editing.

Acknowledgements

This research has been supported by the National Natural Science Foundation of China (Grant Nos. 71874201, 71673297, 71503264 and 71874202), the Humanities and Social Sciences Youth Foundation of the Ministry of Education of China (Grant No. 19YJCZH106).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.petrol.2020.107240.

References

Armstrong, J.S., 2001. Principles of Forecasting: a Handbook for Researchers and Practitioners. Springer Science & Business Media.
Assaad, M., Romuald, B., Cardot, H., 2008. A new boosting algorithm for improved time-series forecasting with recurrent neural networks. Inf. Fusion 9 (1), 41–55.
Baumeister, C., Kilian, L., 2012. Real-time forecasts of the real price of oil. J. Bus. Econ. Stat. 30 (2), 326–336.
Bao, Y., Lu, Y., Zhang, J., 2004. Forecasting Stock Price by SVMs Regression. International Conference on Artificial Intelligence: Methodology, Systems, and Applications. Springer, Berlin, Heidelberg, pp. 295–303.
Buchanan, W.K., Hodges, P., Theis, J., 2001. Which way the natural gas price: an attempt to predict the direction of natural gas spot price movements using trader positions. Energy Econ. 23 (3), 279–293.
Chen, Y., He, K., Tso, G.K.F., 2017. Forecasting crude oil prices: a deep learning-based model. Proc. Comput. Sci. 122, 300–307.
Chen, Y.C., 2014. Wind Power Combination Forecasting Based on Wavelet Packet Transform. Guangdong University of Technology, China.
Chiroma, H., Abdulkareem, S., Herawan, T., 2015. Evolutionary Neural Network model for West Texas Intermediate crude oil price prediction. Appl. Energy 142, 266–273.
Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20 (3), 273–297.
Cao, C., Xu, J., 2007. Short-term traffic flow predication based on PSO-SVM. In: Peng, Q., Wang, K.C.P., Qiu, Y., Pu, Y., Luo, X., Shuai, B. (Eds.), Proceedings of the First International Conference on Transportation Engineering, vol. 246. American Society of Civil Engineers (ASCE), Chengdu, China, pp. 167–172.
Čeperić, E., Žiković, S., Čeperić, V., 2017. Short-term forecasting of natural gas prices using machine learning and feature selection algorithms. Energy 140, 893–900.
Drucker, H., Burges, C.J.C., Kaufman, L., Smola, A., Vapnik, V., 1996. Support vector regression machines. In: Moser, M., Jordan, M., Petsche, T. (Eds.), Advances in Neural Information Processing Systems, vol. 9. MIT Press, Cambridge, MA, pp. 155–161.
Dritsaki, C., 2018. The performance of hybrid ARIMA-GARCH modeling and forecasting oil price. Int. J. Energy Econ. Pol. 8 (3), 14–21.
Dai, H., Cao, Z., 2017. A wavelet support vector machine-based neural network metamodel for structural reliability assessment. Comput. Aided Civ. Infrastruct. Eng. 32 (2), 344–357.
Espinoza, M., Suykens, J.A.K., Belmans, R., et al., 2007. Electric load forecasting. IEEE Contr. Syst. Mag. 27 (5), 43–57.
Guo, X., Li, D.C., Zhang, A., 2012. Improved support vector machine oil price forecast model based on genetic algorithm optimization parameters. Aasri. Proc. 1, 525–530.
Gan, L., 2012. Application of Combination Forecast Model in the Medium and Long Term Power Load Forecast. Nanchang Hangkong University, Nanchang, China.
Hailemariam, A., Smyth, R., 2019. What drives volatility in natural gas prices? Energy Econ. 80, 731–742.
Herrera, G.P., Constantino, M., Tabak, B.M., et al., 2019. Long-term forecast of energy commodities price using machine learning. Energy 179, 214–221.
Hu, Q.L., Ye, N.Y., Zhu, M.F., 2007. Survey of cluster analysis in data mining. Comput. Digital Eng. 35 (2), 17–20.
He, L., Wu, L.D., Cai, Y.C., 2007. Survey of clustering algorithms in data mining. Appl. Res. Comput. 24 (1), 10–13.
Hong, W.C., 2009. Chaotic particle swarm optimization algorithm in a support vector regression electric load forecasting model. Energy Convers. Manag. 50 (1), 105–117.
He, K., Zhang, X., Ren, S., et al., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778.
Intergovernmental Panel on Climate Change (IPCC), 2013. In: Stocker, T.F.D., Qin, G.-K., Plattner, Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P. M. (Eds.), Climate Change 2013: the Physical Science Basis. Cambridge University Press, Cambridge, UK.
Jiang, L.K., 2002. A method of finding time-variant weight in combination forecasting model. J. Shanghai Marit. Univ. 23 (3), 79–81.
Klein, T., Walther, T., 2016. Oil price volatility forecast with mixture memory GARCH. Energy Econ. 58, 46–58.
Kristjanpoller, W., Minutolo, M.C., 2016. Forecasting volatility of oil price using an artificial neural network-GARCH model. Expert Syst. Appl. 65, 233–241.
Kumar, D.N., Raju, K.S., Sathish, T., 2004. River flow forecasting using recurrent neural networks. Water Resour. Manag. 18 (2), 143–161.
Kuo, P.H., Huang, C.J., 2018. An electricity price forecasting model by hybrid structured deep neural networks. Sustainability 10 (4), 1280.
Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet classification with deep convolutional neural networks. In: International Conference on Neural Information Processing Systems. Curran Associates Inc., pp. 1097–1105
Ke, G., Meng, Q., Finley, T., et al., 2017. Lightgbm: a highly efficient gradient boosting decision tree. Adv. Neural Inf. Process. Syst. 3146–3154.
Lanza, A., Manera, M., Giovannini, M., 2005. Modeling and forecasting cointegrated relationships among heavy oil and product prices. Energy Econ. 27 (6), 831–848.
Lv, T., Hao, Y.T., 2017. Stock price prediction based on K-line series similarity search. J. Comput. Appl. 32 (2), 229–235.
Li, Y., Song, T., Yang, J., et al., 2016. An alignment-free algorithm in comparing the similarity of protein sequences based on pseudo-markov transition probabilities among amino acids. PloS One 11 (12), e0167430.
Liu, Y.K., Wan, Z.S., Qiu, H.B., 2018. Research on methods of building deformation monitoring and forecasting. Mine Surv. 46 (4), 63–65.
Lin, B., Wesseh Jr., P.K., 2013. What causes price volatility and regime shifts in the natural gas market. Energy 55, 553–563.
Moshiri, S., Foroutan, F., 2006. Forecasting nonlinear crude oil futures prices. J. Energy 27, 81–95.
Mostafa, M.M., El-Masry, A.A., 2016. Oil price forecasting using gene expression programming and artificial neural networks. Econ. Modell. 54, 40–53.
Martínezálvarez, F., Troncoso, A., Riquelme, J.C., Riquelme, J.M., 2007. Partitioning-clustering techniques applied to the electricity price time series. Lect. Notes Comput. Sc. 4881, 990–999.
Mohammadi, K., Shamshirband, S., Anisi, M.H., et al., 2015. Support vector regression based prediction of global solar radiation on a horizontal surface. Energy Convers. Manag. 91, 433–441.
Mishra, P., 2012. Forecasting natural gas price-time series and nonparametric approach. In: Proceedings of the World Congress on Engineering, vol. 1. WCE, 2012, July 4 - 6, 2012, London, U.K.
Neupane, B., Woon, W., Aung, Z., 2017. Ensemble prediction model with expert selection for electricity price forecasting. Energies 10 (1), 77.
Naderi, M., Khamehchi, E., Karimi, B., 2019. Novel statistical forecasting models for crude oil price, gas price, and interest rate based on meta-heuristic bat algorithm. J. Petrol. Sci. Eng. 172, 13–22.
Nguyen, H.T., Nabney, I.T., 2010. Short-term electricity demand and gas price forecasts using wavelet transforms and adaptive models. Energy 35 (9), 3674–3685.
Onour, I.A., 2009. Natural gas markets: how sensitive are they to crude oil price changes? OPEC Energy Rev. 33 (2), 111–124.
Papadimitriou, T., Gogas, P., Stathakis, E., 2014. Forecasting energy markets using support vector machines. Energy Econ. 44, 135–142.
Palpanas, T., 2015. Data series management: the road to big sequence analytics. SIGMOD Rec. 44 (2), 47–52.
Pearson, W.R., 1998. Empirical statistical estimates for sequence similarity searches. J. Mol. Biol. 276 (1), 71–84.
Peng, L., Liu, S., Liu, R., et al., 2018. Effective long short-term memory with differential evolution algorithm for electricity price prediction. Energy 162, 1301–1314.
Qiu, H.Z., 2017. Research on Electricity Price Short-Term Forecasting. Harbin Institute of Technology.

Salehnia, N., Falahi, M.A., Seifi, A., et al., 2013. Forecasting natural gas spot prices with nonlinear modeling using Gamma test analysis. J. Nat. Gas Sci. Eng. 14, 238–249.
Sadorsky, P., 2006. Modeling and forecasting petroleum futures volatility. Energy Econ. 28 (4), 467–488.
Shin, H., Hou, T., Park, K., et al., 2013. Prediction of movement direction in crude oil prices based on semi-supervised learning. Decis. Support Syst. 55 (1), 348–358.
Su, M., Zhang, Z., Zhu, Y., et al., 2019. Data-driven natural gas spot price forecasting with least squares regression boosting algorithm. Energies 12 (6), 1094.
Sakurai, Y., Matsubara, Y., Faloutsos, C., 2015. Mining and forecasting of big time-series data. In: Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data. ACM, pp. 919–922.
Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. Stat. Comput. 14 (3), 199–222.
Safari, A., Davallou, M., 2018. Oil price forecasting using a hybrid model. Energy 148, 49–58.
Su, M., Zhang, Z., Zhu, Y., Zha, D., Wen, W., 2019a. Data driven natural gas spot price prediction models using machine learning methods. Energies 12 (9), 1680.
Su, M., Zhang, Z., Zhu, Y., Zha, D., 2019b. Data-driven natural gas spot price forecasting with least squares regression boosting algorithm. Energies 12 (6), 1094.
Siddiqui, A.W., 2019. Predicting natural gas spot prices using artificial neural network. In: 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS). https://doi.org/10.1109/cais.2019.8769586.
Thakur, A., Kumar, S., Tiwari, A., 2015. Hybrid model of gas price prediction using moving average and neural network. In: 2015 1st International Conference on Next Generation Computing Technologies. NGCT), Dehradun, pp. 735–737.
Vidic, R.D., Brantley, S.L., Vandenbossche, J.M., et al., 2013. Impact of shale gas development on regional water quality. Science 340 (6134), 1235009.
Viacaba, A., Poursaeidi, M.H., Kundakcioglu, O.E., 2012. Natural gas price forecasting via selective support vector regression. In: Proceedings of the 2012 Industrial and Systems Engineering Research Conference, pp. 1–9.
Wang, J., Feng, L., Tang, X., et al., 2017. The implications of fossil fuel supply constraints on climate change projections: a supply-side analysis. Futures 86, 58–72.
Wang, J., Liu, M., Bentley, Y., et al., 2018a. Water use for shale gas extraction in the Sichuan Basin, China. J. Environ. Manag. 226, 13–21.
Wang, M., Zhao, L., Du, R., et al., 2018b. A novel hybrid method of forecasting crude oil prices using complex network science and artificial intelligence algorithms. Appl. Energy 220, 480–495.
Wang, T., Lin, B., 2017. China's natural gas consumption peak and factors analysis: a regional perspective. J. Clean. Prod. 142, 548–564.
Wang, J.L., Lei, C.R., 2020. Research on the Forecasting Method for Natural Gas Price Based on the Data Mining Technique. China Mining Magazine, 2020; accepted.
Xiang, Y., Zhuang, X.H., 2013. Application of ARIMA model in short-term prediction of international crude oil price. Trans. Tech. Publ. 798, 979–982.
Xie, W., Yu, L., Xu, S., et al., 2006. A new method for crude oil price forecasting based on support vector machines. In: International Conference on Computational Science. Springer, Berlin, Heidelberg, pp. 444–451.
Yu, L., Zhao, Y., Tang, L., 2014. A compressed sensing-based AI learning paradigm for crude oil price forecasting. Energy Econ. 46, 236–245.
Yuan, G.C., Cheng, L., Wang, J.Q., 2003. Electric price forecasting using data mining. Proc. Chin. Soc. Univ. Electr. Power Syst. Autom. 15 (2), 19–23.
Yorucu, V., Bahramian, P., 2015. Price modelling of natural gas for the EU-12 countries: evidence from panel cointegration. J. Nat. Gas Sci. Eng. 24, 464–472.
Zhao, Y., Li, J., Yu, L., 2017. A deep learning ensemble approach for crude oil price forecasting. Energy Econ. 66, 9–16.
Zhang, J.L., Zhang, Y.J., Zhang, L., 2015. A novel hybrid method for crude oil price forecasting. Energy Econ. 49, 649–659.
Zhang, Y., Chen, W., 2011. A new measure for similarity searching in DNA sequences. MATCH Commun. Math. Comput. Chem. 65, 477–488.
Zhang, Z., Qin, H., Liu, Y., et al., 2019. Long Short-Term Memory Network based on Neighborhood Gates for processing complex causality in wind speed prediction. Energy Convers. Manag. 192, 37–51.
Zhu, S., Wang, J., Zhao, W., et al., 2011. A seasonal hybrid procedure for electricity demand forecasting in China. Appl. Energy 88 (11), 3807–3815.
Zhang, Z.F., Zhu, J.G., 2010. Application of SVR into quantitatively analyzing adverse selection contract model. J. Comput. Appl. 30 (3), 779–782.
Zhang, R., Li, W., Mo, T., 2018. Review of deep learning. Inf. Contr. 47 (4), 385–397.