Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 5488053, 13 pages
https://doi.org/10.1155/2022/5488053

Research Article

Natural Gas Price Forecasting by a New Hybrid Model Combining Quadratic Decomposition Technology and LSTM Model

Linjie Zhan[1] and Zhenpeng Tang[1,2]

[1]*School of Economics and Management, Fuzhou University, Fuzhou 350108, China*
[2]*School of Economics and Management, Fujian Agriculture and Forestry University, Fuzhou 350002, China*

Correspondence should be addressed to Zhenpeng Tang; zhenptang2020@126.com

Received 2 September 2022; Revised 16 October 2022; Accepted 21 November 2022; Published 5 December 2022

Academic Editor: Juan Frausto-Solis

Research on the price prediction of natural gas is of great significance to market participants of all kinds. In order to predict natural gas prices more reliably, this paper introduces a quadratic decomposition technology based on the combination of variational modal decomposition (VMD) and ensemble empirical modal decomposition (EEMD), which decomposes the residual term (Res) after VMD by EEMD; then, a new hybrid model called VMD-EEMD-Res.-LSTM is constructed in combination with the long short-term memory (LSTM) prediction model. The contribution of this new hybrid model is that, unlike existing application research that combines existing decomposition technology with the LSTM model, it does not ignore the important information contained in the residual after the VMD. In order to verify the predictive performance of the proposed new model, this paper uses the data of the spot price of natural gas in the United States to conduct a multistep-ahead empirical comparative analysis. The results show that the new hybrid model constructed in this paper has significant predictive advantages.

1. Introduction

As a clean energy, natural gas plays an important role in the transformation of the world's energy system, offering extensive advantages in dealing with global climate change. The development and utilization of natural gas have become important parts of the energy strategies of many countries in the world [1, 2]. In the natural gas market—an important component of the global energy market—the fluctuation of natural gas prices affects the behavior of natural gas suppliers and consumers, the development of the natural gas industry, and the production and operation of enterprises related to natural gas [3]. The reasonable and effective prediction of natural gas prices is helpful for researchers and decision makers in commodity trading and power production planning, allowing them to make better decisions and establish effective risk-avoidance mechanisms. Therefore, accurately predicting natural gas prices has become a concern of many scholars [4].

However, forecasting natural gas prices is a difficult task, largely due to the impact of uncertainties associated with extraction technology and economic factors. The existing research on energy price forecasting can be divided into multivariate forecasting and univariate forecasting according to the different variables involved in the price forecasting research model. In multivariate price forecasting research, the variables involved in the model not only include historical natural gas price data but also other market or policy factors that affect price changes [5, 6]. Univariate forecasting research only involves historical energy price data. This paper's research scope is aligned with univariate prediction.

Based on the review of research on univariate forecasting of energy prices, the existing forecasting models can be divided into three types. That is, traditional econometric models, single artificial intelligence forecasting methods, and hybrid model forecasting methods.

Regarding traditional econometric models, the classic models employed include autoregressive moving average (ARMA), autoregressive integrated moving average

(ARIMA), generalized autoregressive conditional heteroscedasticity (GARCH), random walk, vector autoregression, and error correction model, among others [7–10]. Lv and Shan [11] used the GARCH model with its characteristics of long memory and fat-tailed distributions to forecast the spot and futures prices of natural gas. Based on the generalized autoregressive scoring (GAS) model and the GARCH model, Xu and Lien [12] forecasted [7] the spot and futures prices of crude oil and natural gas in different developed markets.

In the second type of prediction research (i.e., research on the single artificial intelligence model), the typical models studied include the artificial neural network (ANN), support vector regression (SVR), inheritance calculation method, and random forest model, among other machine learning models [13–15]. Ceperic et al. [16] made a short-term prediction of the natural gas spot price for a distribution hub (Henry Hub) based on the improved support vector regression model. Mouchtaris et al. [17] used machine learning methods such as support vector machines (SVMs) and regression trees to predict the spot price of natural gas.

Various hybrid methods, the third kind of combination model prediction method, have been developed in recent years [18–20], of which the most typical method used is the TEI@I complex system research methodology [21]. Wang et al. [22] constructed a hybrid model for forecasting natural gas prices based on complete ensemble empirical mode decomposition with an adaptive noise sample entropy (CEEMDAN-SE) and a gated recurrent unit network that is optimized by the particle swarm optimization algorithm with an adaptive learning strategy (PSO-ALS-GRU). Liu et al. [23] combined variational pattern decomposition (VMD) and ANN to construct a VMD-ANN method to forecast the futures prices of various energy sources. This hybrid method combined decomposition technology and a prediction model to significantly improve the prediction performance of the whole model, showing better prediction accuracy than the singular model [24, 25].

Of the three methods described above, only the traditional econometric model requires the research object to meet certain statistical assumptions (e.g., data stationarity). Therefore, it has strong limitations when it comes to dealing with nonlinear and nonstationary time series and cannot effectively predict time series with irregular and random characteristics [26, 27]. Unlike the econometric model, the artificial intelligence model trains historical data by using. This machine learning technology [28], although it can better deal with nonstationary and nonlinear time series, it is sensitive to parameter setting and easily falls into local minima or over-fitting.

The hybrid TEI@I complex system methodology, which combines decomposition technology and a prediction model, can effectively overcome the limitations of traditional econometric models and single artificial intelligence models while improving the accuracy of prediction [29]. This approach combines decomposition preprocessing techniques with machine learning methods. The time series is decomposed with the signal processing method, and the decomposed components are predicted by the prediction model, which significantly improves the overall prediction accuracy of the model [30].

Based on the existing research on natural gas price prediction, it can be concluded that the precision of a singular forecasting model is limited. In contrast, the hybrid model based on decomposition technology and a forecasting model can effectively overcome the shortcomings of traditional econometric models and single artificial intelligence models.

However, this approach does feature multiple shortcomings. Firstly, in the existing research on forecasting models for natural gas prices, the decomposition of the original series by VMD technology usually only uses one-level decomposition technology; besides, there is no published research on forecasting natural gas prices by combining secondary decomposition technology with a deep learning forecasting model. In addition, the research on prediction models combining decomposition technology and LSTM models ignores the residual term after decomposition, failing to consider the complex information contained in the residual after modal decomposition and weakening the decomposition effect of data [31–33].

The motivation of this paper is that, in view of the shortcomings of the existing research, it proposes the quadratic decomposition technology for the residual term after applying VMD technology; further, it constructs the new VMD-EEMD-Res.-LSTM hybrid prediction model to predict natural gas prices.

Typical methods include wavelet decomposition (WD) [34], empirical mode decomposition (EMD) [35, 36], EEMD [37, 38], and VMD [39], among others. Among the aforementioned methods, the WD algorithm relies on the setting of the basis function, the EMD algorithm is prone to modal aliasing, and the EEMD algorithm has the problem of over-enveloping or under-enveloping due to its lack of mathematical basis and inability to separate components with similar frequencies, which limits the effects of data decomposition.

As an advanced forecasting model, the LSTM model performs excellently in solving prediction problems in the time series energy field [40–42]. However, notice that the analysis is a feature of the forecasting user. Besides, the tools for analysing the prediction results do not belong to LSTM.

Compared with the traditional artificial neural network model, recurrent neural networks (RNNs) can deal with the front-back dependency of time series data more effectively because of their internal self-circulating cells. However, RNN cannot effectively solve the problem of the long-term dependence on time series, and the gradient disappears in the process of propagation. As an advanced forecasting model, the LSTM model performs excellently in solving prediction problems in the time series energy field [40–42].

Compared to the previous research on forecasting natural gas prices with decomposition technology, this new hybrid model is novel for several reasons. After it uses VMD technology to decompose the original price series into several modal components and the residual term, it then uses EEMD technology to decompose the residual term. Furthermore, the LSTM model is used to predict the decomposed components. Finally, the prediction results of modal components and the residual term are linearly superimposed to form the final prediction results of the original sequence. In order to be consistent with the existing research and prove the superiority of the newly proposed VMD-EEMD-Res.-LSTM hybrid model, this paper constructs nine benchmark models: random forest, RNN, ANN, extreme learning machine (ELM), LSTM, EEMD-LSTM, VMD-LSTM, VMD-LSTM, and VMD-Res.-LSTM.

Unlike previous methods, the new model proposed in this paper can extract the nonlinear characteristics of natural gas prices more effectively, improve the prediction accuracy of natural gas prices, and maintain stable prediction abilities in all multistep ahead forecasting scenarios, providing a new perspective for univariate prediction research of natural gas prices.

The contribution of this paper is to fill the gap of the existing research on natural gas price prediction, which is achieved in three ways:

(1) The advanced decomposition technology is applied to the prediction of natural gas prices, reducing the complexity of natural gas price series. Compared with other decomposition algorithms, the VMD algorithm is an improved decomposition technology that adaptively decomposes the effective components corresponding to each center frequency in the frequency domain, giving it higher decomposition accuracy [33, 43].
(2) The LSTM model is introduced into the prediction research, which effectively improves the prediction accuracy. As a variant of RNN, the LSTM model uses long-distance time series information by introducing memory units into the network structure. This allows a more effective description of the long-term dependence between time series data, overcoming two problems: gradient disappearance and its lack of long-term memory ability [40, 41].
(3) To verify the accuracy and robustness of the model's prediction performance, different evaluation indicators (e_{RMSE}, e_{MAE}, and e_{MAPE}) are selected, and all models are applied to the multi-step-ahead forecasting based on the weekly frequency of natural gas prices from the natural gas spot price dataset of Henry Hub, a Louisiana distribution hub for a pipeline system. The paper then tests the effectiveness of different models in capturing the future price dynamics of natural gas under the one-step-ahead, three-step-ahead, and five-step-ahead forecasting scenarios.

The rest of this paper is arranged as follows: Section 2 describes the sub-models of the hybrid model and the specific modeling steps of the VMD-EEMD-Res.-LSTM hybrid model. Section 3 expounds on the empirical results and the corresponding discussion. Finally, Section 4 covers the main conclusions and future research directions of this study.

2. Methodology

The main goal of this paper is to develop a new hybrid model (namely, VMD-EEMD-Res.-LSTM) based on the quadratic decomposition algorithm and the LSTM model, which can be used to forecast the price data of natural gas more accurately and reliably. The following section describes the VMD algorithm, EEMD algorithm, and LSTM model, which constitute the VMD-EEMD-Res.-LSTM hybrid model; the section also explains the steps for constructing the VMD-EEMD-Res.-LSTM hybrid model. Tables 1 and 2 show the key abbreviations and variables (respectively) involved in the equations.

2.1. Variational Modal Decomposition. The variational modal decomposition algorithm determines the optimal solution for variational modes through iteration, and each modal function and the center frequency is constantly updated; then, a set of modal functions with a specific bandwidth is obtained. The VMD algorithm is an improvement of the EMD algorithm, more effectively avoiding problems such as mode aliasing and end-point effects. The VMD algorithm decomposes the signal into K, the predefined intrinsic modal component (IMF)—that is, the IMF is redefined as a modal function of an input signal $u_k(\mathrm{t})$:

$$u_k(\mathrm{t}) = A_k(t)\cos\left[\varphi_k(t)\right], \tag{1}$$

where $A_k(t)$ is the instantaneous amplitude and $A_k(t) \geq 0$; $\varphi_k(t)$ is the phase. The instantaneous frequency $\omega_k(t)$ is obtained by solving the first-order differential for t in $\varphi_k(t)$:

$$\omega_k(t) = \frac{d\varphi_k(t)}{dt}. \tag{2}$$

The specific steps are as follows:

Step 1: Perform HHT transformation on each $u_k(t)$ obtained by decomposition to calculate its related analytical signal and marginal spectrum:

$$\left(\delta(t) + \frac{j}{\pi t}\right)u_k(t), \tag{3}$$

where $\delta(t)$ is the unit impulse function.

Step 2: Transform the spectrum of each $u_k(t)$ to its corresponding baseband by adding an estimated center frequency $e^{-j\omega_k t}$:

$$\left[\left(\delta(t) + \frac{j}{\pi t}\right)u_k(t)\right]e^{-j\omega_k t}. \tag{4}$$

Step 3: The bandwidth of the corresponding mode is estimated by calculating the square of the L^2 norm of the gradient, and the following constrained variational problem is constructed:

Table 1: Information on key abbreviations.

Unabbreviated form	Abbreviation
Variational modal decomposition	VMD
Ensemble empirical mode decomposition	EEMD
Empirical mode decomposition	EMD
Intrinsic modal component	IMF
Long short-term memory	LSTM

Table 2: Information on key variables.

Name of the variable	Meaning
$A_k(t)$	The instantaneous amplitude
$\omega_k(t)$	The instantaneous frequency
$\delta(t)$	The unit impulse function
$N(t)$	The set of white noise
f_t	The memory cell state
h_{t-1}	The output of the previous layer

$$\min_{\{u_k\},\{w_k\}} \left\{ \sum_k \left\| \left[\partial_t \left(\delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] e^{-jw_k t} \right\|_2^2 \right\}, s.t = \sum_k u_k = f(t), \tag{5}$$

where $\{u_k\} = \{u_1, \cdots, u_k\}$ and $\{w_k\} = \{w_1, \cdots, w_k\}$ are the decomposed modal components and their corresponding center frequencies, K represents the number of decomposed modes, and $f(t)$ represents the original input signal.

Step 4: In order to solve equation (5), the quadratic penalty term α and the Lagrange multiplier $\lambda(t)$ are introduced, which are transformed into the following unconstrained variational problem:

$$L(\{u_k\}, \{w_k\}, \lambda) = \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \otimes u_k(t) \right] e^{-jw_k t} \right\|_2^2 + \left\| f(t) - \sum_k u_k(t) \right\|_2^2 + \langle \lambda(t), f(t) - \sum_k u_k(t) \rangle. \tag{6}$$

Combined with Fourier equidistant transformation, the alternate direction multiplier method is used to continuously optimize and update $u_k^{n+1}(w)$, w_k^{n+1}, and $\lambda^{n+1}(w)$. Then, the saddle point of (6) is found in the frequency domain-that is, the minimum value in (5). The updated formula is as follows:

$$\hat{u}_k^{n+1}(w) = \frac{\hat{f}(w) - \sum_{i<k} \hat{u}_i^{n+1}(w) - \sum_{i>k} \hat{u}_i^n(w) + \hat{\lambda}^n(w)/2}{1 + 2\alpha(w - w_k^n)^2}, \tag{7}$$

$$w_k^{n+1} = \frac{\int_0^\infty \omega \left| \hat{u}_k^{n+1}(w) \right|^2 dw}{\int_0^\infty \left| \hat{u}_k^{n+1}(w) \right|^2 dw}, \tag{8}$$

$$\hat{\lambda}^{n+1}(w) = \hat{\lambda}^n(w) + \tau \left[\hat{f}(w) - \sum_k \hat{u}_k^{n+1}(w) \right], \tag{9}$$

where n is the number of iterations.

Step 5: Specify the maximum number of iterations as N, so that a positive number n can satisfy the condition $\leq N$:

$$\sum_k \| \hat{u}_k^{n+1} - \frac{\hat{u}_k^n\|_2^2}{\| \hat{u}_k^n\|_2^2} < \varepsilon. \tag{10}$$

If there is any $\varepsilon > 0$ that can satisfy the (10), and the iteration is finished; otherwise, return to Step 4. Then, perform inverse Fourier transform on the frequency domain results of the K IMFs obtained by solving (7) to obtain time domain signals.

2.2. Ensemble Empirical Mode Decomposition.

The EEMD algorithm adds Gaussian white noise with limited amplitude to the signal in each decomposition, using the uniform distribution of the white noise spectrum to automatically distribute signals of different time scales to the appropriate reference scales. At the same time, by using the zero-mean characteristic of white noise, the noise can cancel each other out after being averaged many times, and the influence of noise can be suppressed or even completely eliminated. The specific steps of the EEMD algorithm are as follows:

Step 1: Obtain a population $x(t)$ by adding a set of white noise $N(t)$ to the analysis signal $X(t)$;

$$X(t) = x(t) + N(t). \tag{11}$$

Step 2: Apply the EMD algorithm to decompose the signal $X(t)$ to obtain a set of IMF components, namely,

$$X(t) = \sum_{j=1}^n C_j(t) + r_n(t). \tag{12}$$

Step 3: Add different white noises $N_i(t)$ to the analysis signal (t), and repeat the above two steps for a total of m times:

$$\begin{cases} X_i(t) = x(t) + N_i(t) \\ X_i(t) = \sum_{j=1}^n C_{ij}(t) + r_{\text{in}}(t) \end{cases}. \tag{13}$$

Step 4: Take the mean value of the corresponding m groups of IMF components as the final EEMD decomposition result, namely,

$$c_j(t) = \frac{1}{m}\sum_{i=1}^{m} C_{ij}(t). \quad (14)$$

2.3. Long Short-Term Memory. The core idea of the LSTM model is to introduce the memory unit and gating unit. The memory unit can keep the state over the passage of time, while the nonlinear gating unit can be used to adjust the inflow and outflow of information. Figure 1 shows the typical structure of the LSTM model, including the input gate, the forget gate, and the output gate. The LSTM model maintains and updates the state of memory cells through these three types of gate structures.

The information forgotten by the memory unit is determined by the sigmoid layer of the forget gate. The input is the input of the current layer x_t and the output of the previous layer h_{t-1}. The output of the memory cell state at time t is as follows:

$$f_t = \sigma\left(W^{(f)}x_t + U^{(f)}h_{t-1} + b^{(f)}\right). \quad (15)$$

The information stored in the memory cell state mainly consists of two parts. First, the result of the sigmoid layer of the input gate is used as the updated information. Secondly, the vector newly created by the $\tan h$ layer is added to the memory cell state. The old memory cell state f_t is multiplied by c_{t-1} to account for forgotten information, and an update of the memory cell state is generated together with the new candidate information $i_t \cdot u_t$.

$$\begin{aligned} i_t &= \sigma\left(W^{(i)}x_t + U^{(i)}h_{t-1} + b^{(i)}\right), \\ u_t &= \tan h\left(W^{(u)}x_t + U^{(u)}h_{t-1} + b^{(u)}\right), \\ c_t &= i_t \otimes u_t + f_t \otimes c_{t-1}. \end{aligned} \quad (16)$$

The output information is determined by the output gate. First, the sigmoid layer is used to determine the partial information of the memory cell state to be the output; then, the memory cell state is processed by the $\tan h$ layer. Finally, the output value is the product of these two pieces information.

$$\begin{aligned} c_t &= i_t \otimes u_t + f_t \otimes c_{t-1}, \\ h_t &= o_t \otimes \tan h(c_t). \end{aligned} \quad (17)$$

2.4. Construction of the Proposed VMD-EEMD-Res.-LSTM Hybrid Model. In order to address the deficiency of the existing research, this paper proposes a new VMD-EEMD-Res.-LSTM hybrid model by combining the secondary decomposition technology and the LSTM model. Figure 2 shows the construction flow of the VMD-EEMD-Res.-LSTM hybrid model. The detailed modeling steps of this model are as follows:

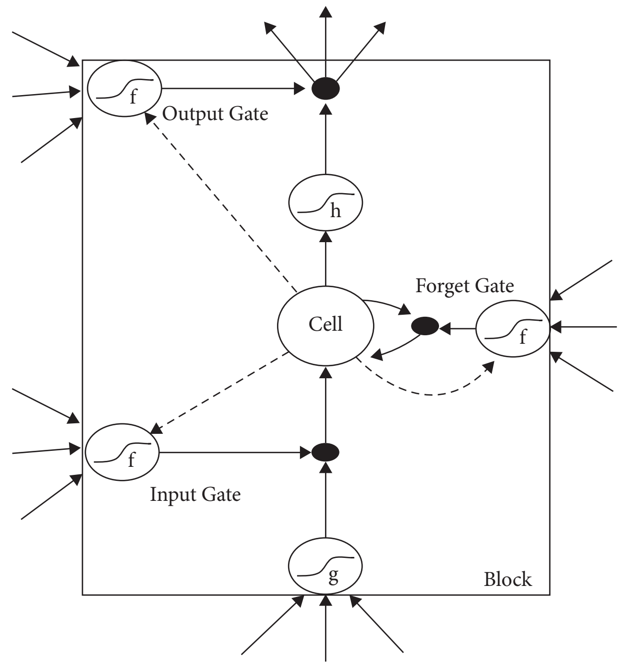

Figure 1: The neural network structure of the LSTM model.

Step 1: (VDM application and residual term): Firstly, use VMD to decompose the original natural gas price series to obtain each modal component VMF. Then subtract the sum of each VMF component's data from the original time series data to determine the residual term after applying VMD.

Step 2: Normalize the decomposed VMF components, then select an appropriate number of training samples and test samples. The LSTM model is used to train each VMF component, obtaining their prediction results.

Step 3: (EEMD and LSTM phase): First, apply EEMD to the residual term. Then, with LSTM predict each IMF component. Finally, superimpose the prediction results of each subsequence to obtain the final prediction result of the residual term.

Step 4: (Final prediction): Superimpose the prediction results of each component and the residual term after applying VMD technology to obtain the final prediction result of the original price series.

3. Experiments

3.1. Source of Data. This paper selects the weekly frequency spot price series from the Henry Hub's natural gas spot price dataset as its empirical research (https://www.eia.gov/dnav/ng/hist/rngwhhdw.htm). The date range of this sample series is from January 10th, 1997, to February 4th, 2022, which contains 1,308 pieces of trading data (among which the first 1,158 are training samples, and the last 150 are test samples). Figure 3 shows the trend of the sample data. Table 3 shows the results of descriptive statistics on the sample data. According to the results in Table 3, the skewness of the

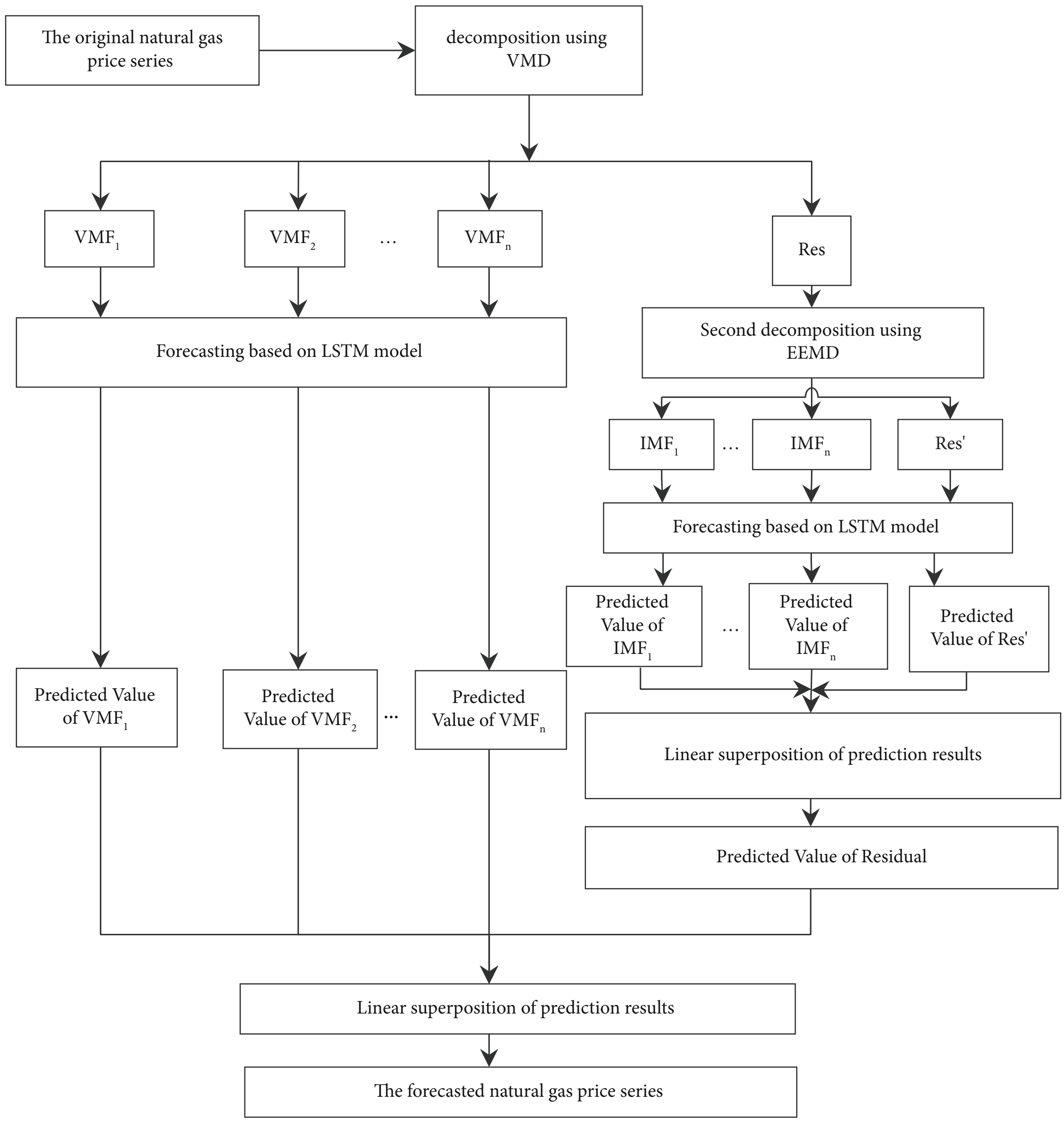

Figure 2: Flowchart of the VMD-EEMD-Res.-LSTM hybrid model.

sample data is greater than 0, and the kurtosis value is greater than 3, showing nonlinear and irregular distribution characteristics. In addition, the empirical research on all prediction models in this paper is all implemented in Matlab2019b.

3.2. Data Processing. In order to ensure the training effect of the LSTM model, it is necessary to normalize the subsequence data of each modal component. In this paper, the min-max dispersion standardization method is used to linearly process the data, and its expression is as follows:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \tag{18}$$

where x' is the standardized subsequence data, x is the original value, and $x_{\max}$ and $x_{\min}$ are the maximum and minimum values, respectively.

3.3. Benchmark Models. In order to verify the prediction ability of the VMD-EEMD-Res.-LSTM hybrid model proposed in this paper, this paper selects and builds nine benchmark models for comparative research in accordance with the existing literature's research path. These benchmark models include five single prediction models (random forest, RNN, ANN, ELM, and LSTM) and four hybrid models (EEMD-LSTM, EEMD-VMD-LSTM, VMD-LSTM, and VMD-Res.-LSTM).

Among these benchmark models, the random forest, RNN, ANN, ELM, and LSTM models are singular prediction

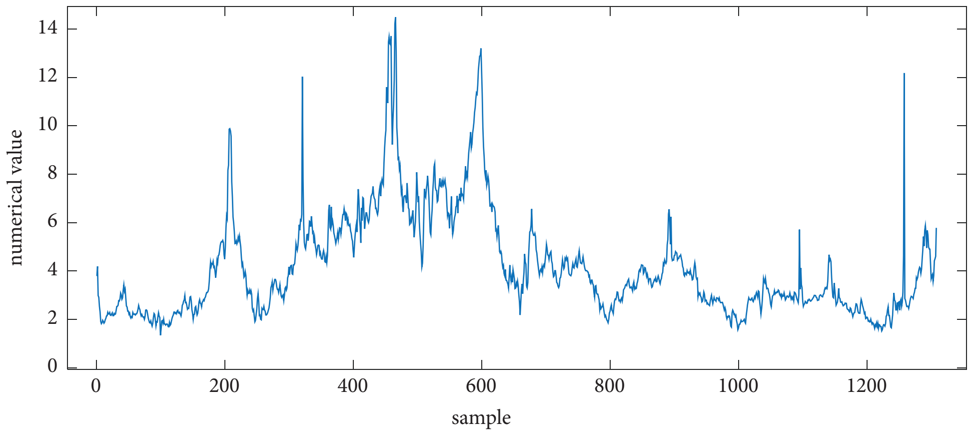

Figure 3: The trend chart of sample data.

Table 3: Descriptive statistics of sample data.

Mean value	Maximum	Minimum	Standard deviation	Skewness	Kurtosis
4.167	14.49	1.34	2.1682	1.6252	6.2509

models without decomposition technology. The EEMD-LSTM model means that, after applying EEMD technology to the natural gas price series, the residual term is discarded. Then, an LSTM model is used to predict a series of decomposed IMF components, and, finally, the prediction results of each component are superimposed. The EEMD-VMD-LSTM model means that, after applying EEMD technology to the natural gas price series, the residual term is discarded. Then, the first high-frequency component (IMF1) is decomposed by VMD. Next, the LSTM model is used to predict each component after decomposition. Finally, the prediction results of each component are superimposed. The VMD-LSTM model is used to apply the VMD technology to the natural gas price series, ignoring the residual term and using the LSTM model to predict a series of VMF components obtained by decomposition; finally, the prediction results of each component are superimposed.

Unlike the other constructed hybrid benchmark models, the VMD-Res.-LSTM model retains the residual term obtained after decomposition from applying VMD technology to natural gas price series, then uses the LSTM model to predict each component and the residual term after decomposition. Finally, it superimposes the prediction results of each component and the residual term. The hybrid model proposed in this paper and named VMD-EEMD-Res.-LSTM has the features shown in Table 4; as can be seen, it has additional good features that other models in the area do not have.

In addition, the research of this paper pertains to the quantitative prediction of a single price variable, adopting the rolling forecasting strategy to make one-step-, three-step-, and five-step- ahead forecasting for all models; that is, the data of the first six trading days are used to predict the data of the seventh, ninth, and 11th trading days.

3.4. Evaluation Indicators of Prediction Results. In this paper, three evaluation indexes; i.e., root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE), are selected to test the prediction effects of different models. The calculation formulas of each index are as follows:

$$e_{\mathrm{RMSE}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}, e_{\mathrm{MAE}} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right|, e_{\mathrm{MAPE}} = \frac{1}{n}\frac{\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right|}{y_i}, \tag{19}$$

where y_i and $\hat{y}_i$ are the true value and predicted value of the price series, respectively, n is the scale of the test sample, and i is the serial number of the test sample. The smaller the index values of e_{RMSE}, e_{MAE}, and e_{MAPE}, the higher the prediction accuracy of the model.

3.5. Experiment Results. To show the advantages of the proposed hybrid model in forecasting natural gas price series, this paper divides the benchmark model into two groups of methods for the empirical analysis: a single forecasting model and a hybrid model combined with

Table 4: Characteristics of different models.

	Decomposition technique	Including the residual after the first decomposition	Quadratic decomposition technology
Random forest (model 1)			
RNN (model 2)			
ANN (model 3)			
ELM (model 4)			
LSTM (model 5)			
EEMD-LSTM (model 6)	√		
EEMD-VMD-LSTM (model 7)	√		
VMD-LSTM (model 8)	√		
VMD-Res.-LSTM (model 9)	√	√	
VMD-EEMD-Res.-LSTM (model 10)	√	√	√

decomposition technology. After analyzing the prediction results of a single model, the proposed hybrid model is compared with the prediction results of other hybrid benchmark models.

3.5.1. Analysis Results of the Single Prediction Model. Table 5 shows the numerical results of the one-step-, three-step-, and five-step-ahead forecasting of natural gas price series using a single prediction model. Figures 4–5 compare the results of the one-step-, three-step-, and five-step-ahead forecasting of different single models, respectively.

Specifically, Table 2 shows that, compared with the random forest, RNN, ANN, and ELM models, the LSTM model has the best prediction performance in the multi-step prediction scenario, with lower e_{RMSE}, e_{MAE}, and e_{MAPE} values than other models. Taking the numerical results in the one-step-ahead forecasting scenario as an example, the values of e_{RMSE}, e_{MAE}, and e_{MAPE} for the random forest, RNN, ANN, ELM, and LSTM models are (0.9802, 0.3454, 0.1039), (1.0545,0.4440, 0.1545), (0.9468,0.4196, 0.1268), (1.1615, 0.3743, 0.1176), and (0.9107, 0.3401, 0.1029), respectively. Additionally, the improvement degree of the LSTM's model prediction results are (7.09%, 1.53%, 0.96%), (13.63%, 23.4%, 33.39%), (3.81%, 18.94%, 18.84%), and (21.59%, 9.13%, 12.5%), respectively. Based on the results of Figures 3–5, the prediction effect of the LSTM model in different scenarios is better than that of other models, which is closer to the real curve of natural gas prices.

This result shows that the LSTM model can effectively capture the data characteristics of natural gas price series, improve the accuracy and stability of natural gas price prediction, and result in better prediction performance than other single models. Therefore, the LSTM model is chosen as the prediction model of the hybrid model constructed in this paper.

3.5.2. Analysis Results of the Hybrid Prediction Model Combined with Decomposition Technology. Table 6 presents the numerical results of the hybrid model combined with the decomposition technique for the multistep-ahead forecasting of natural gas prices. Figures 7–9 compare the results of the one-step-, three-step-, and five-step-ahead forecasting of different hybrid models, respectively. Comparing the fitting effects of different models on the real values, it is evident that, with the increase of the number of prediction steps, the fitting ability of all models decreases; that is, the prediction error increases.

In addition, by comparing the numerical results in Tables 5 and 6, it can be inferred that, although the prediction ability of the LSTM model is better than that of other singular prediction models, its prediction performance is still inferior to any hybrid model combined with decomposition technology. Although singular artificial intelligence models have certain advantages in dealing with nonlinear regular sequences, the results show that it is still difficult to fully identify the complex features of natural gas price sequences. The hybrid model combined with decomposition technology can effectively identify the internal change patterns of natural gas price series; the model can also effectively decompose the natural gas price data into a series of different modal components, which reduces the complexity of natural gas price series and improves the prediction accuracy.

Firstly, by comparing the prediction results of the EEMD-LSTM and EEMD-VMD-LSTM models, it can be concluded that, in all multistep-ahead forecasting scenarios, the prediction performance of EEMD-VMD-LSTM combined with quadratic decomposition technology is better than that of the EEMD-LSTM model using single decomposition technology. Specifically, taking the numerical results in the one-step-ahead forecasting scenario as an example, the values of e_{RMSE}, e_{MAE}, and e_{MAPE} of the EEMD-LSTM and EEMD-VMD-LSTM models are (0.6746, 0.2958, 0.0906) and (0.4087, 0.2021, 0.0621), respectively. Compared with the prediction results of EEMD-LSTM, the improvement degree of the EEMD-VMD-LSTM model is (39.41%, 31.67%, and 31.45%). This result shows the superiority of quadratic decomposition technology in dealing with complex data.

Secondly, by comparing the prediction results of the VMD-LSTM, EEMD-LSTM, and EEMD-VMD-LSTM models, we can see that the prediction performance of VMD-LSTM with VMD technology is better than that of the EEMD-LSTM and EEMD-VMD-LSTM models. Specifically, taking the numerical results in the one-step-ahead

Table 5: Numerical results of multistep ahead forecasting with a single prediction model.

Prediction horizon	Error indicator	Model 1	Model 2	Model 3	Model 4	Model 5
One-step ahead	e_{RMSE}	0.9802	1.0545	0.9468	1.1615	0.9107
	e_{MAE}	0.3454	0.4440	0.4196	0.3743	0.3401
	e_{MAPE}	0.1039	0.1545	0.1268	0.1176	0.1029
Three-step ahead	e_{RMSE}	1.0411	1.1334	1.0308	1.3277	0.9986
	e_{MAE}	0.4933	0.5073	0.4923	0.4959	0.4805
	e_{MAPE}	0.1507	0.1634	0.1526	0.1602	0.1511
Five-step ahead	e_{RMSE}	1.0559	1.2153	1.1014	1.8051	1.0366
	e_{MAE}	0.5880	0.6451	0.5619	0.7097	0.5577
	e_{MAPE}	0.1910	0.2193	0.1736	0.2353	0.1829

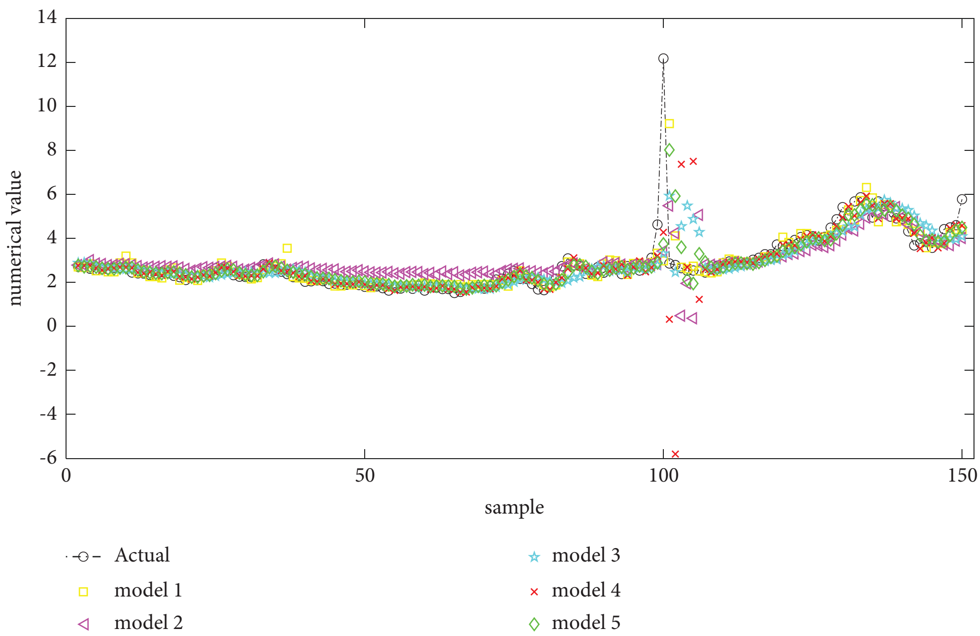

Figure 4: Comparison chart of one-step-ahead forecasting results of different single models.

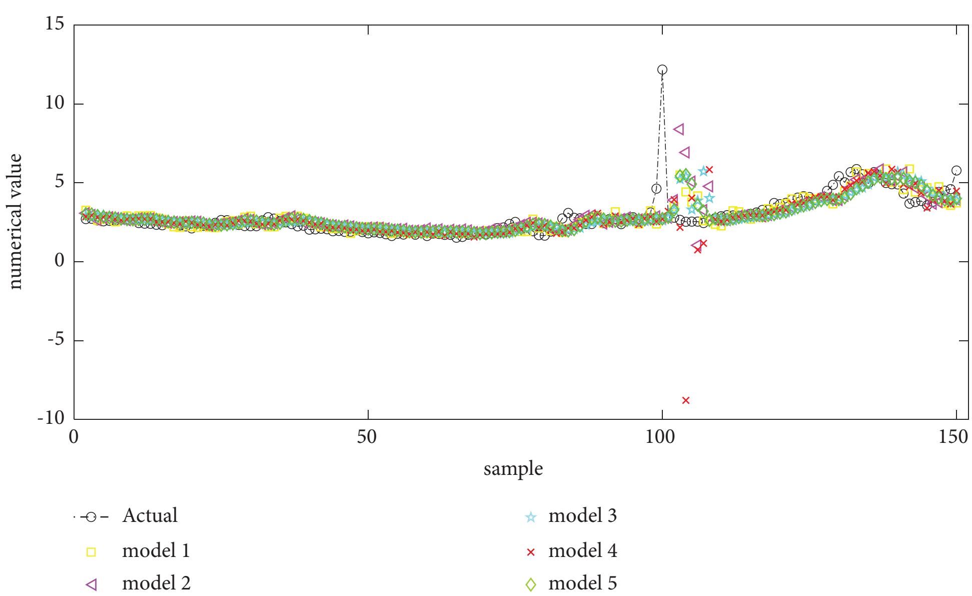

Figure 5: Comparison chart of three-step-ahead forecasting results of different single models.

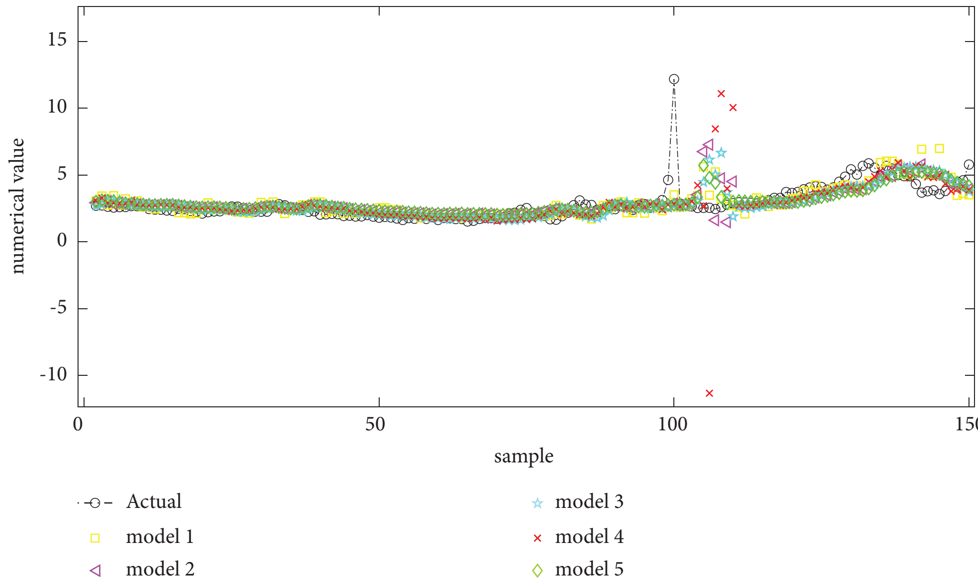

FIGURE 6: Comparison chart of five-step ahead forecasting results of different single models.

TABLE 6: Numerical results of multistep ahead forecasting with the hybrid prediction model.

Prediction horizon	Error indicator	Model 6	Model 7	Model 8	Model 9	Model 10
One-step ahead	e_{RMSE}	0.6746	0.4087	0.4397	0.4378	0.3647
	e_{MAE}	0.2958	0.2021	0.1948	0.1834	0.1793
	e_{MAPE}	0.0906	0.0621	0.0606	0.0559	0.0576
Three-step ahead	e_{RMSE}	0.7586	0.5520	0.5076	0.4998	0.4588
	e_{MAE}	0.3700	0.3107	0.2366	0.2266	0.2202
	e_{MAPE}	0.1069	0.0908	0.0744	0.0717	0.0703
Five-step ahead	e_{RMSE}	0.8900	0.8027	0.5684	0.5540	0.5263
	e_{MAE}	0.4475	0.4520	0.2716	0.2547	0.2470
	e_{MAPE}	0.1265	0.1262	0.0872	0.0805	0.0769

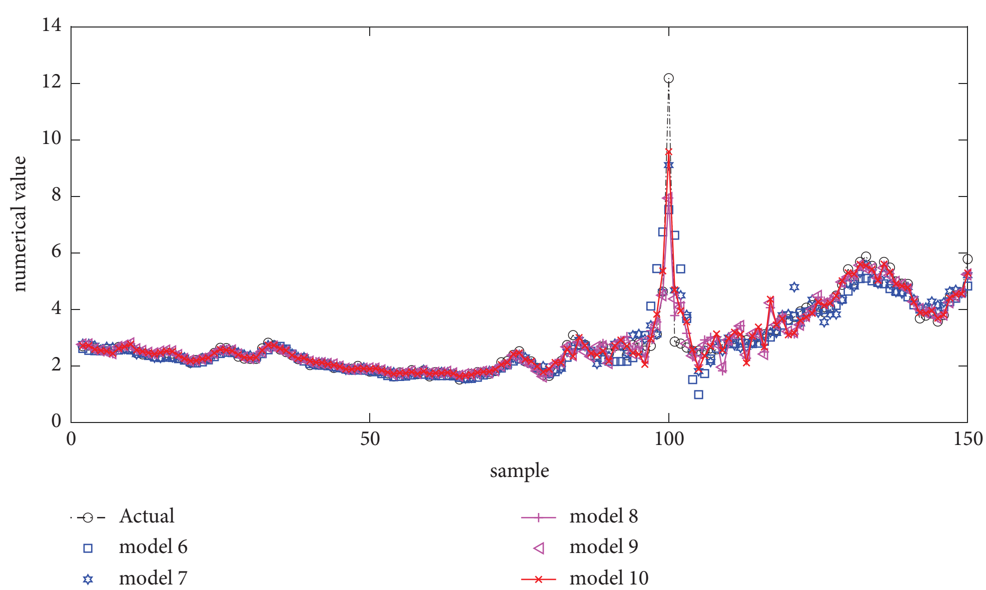

FIGURE 7: Comparison chart of one-step-ahead forecasting results of different hybrid models.

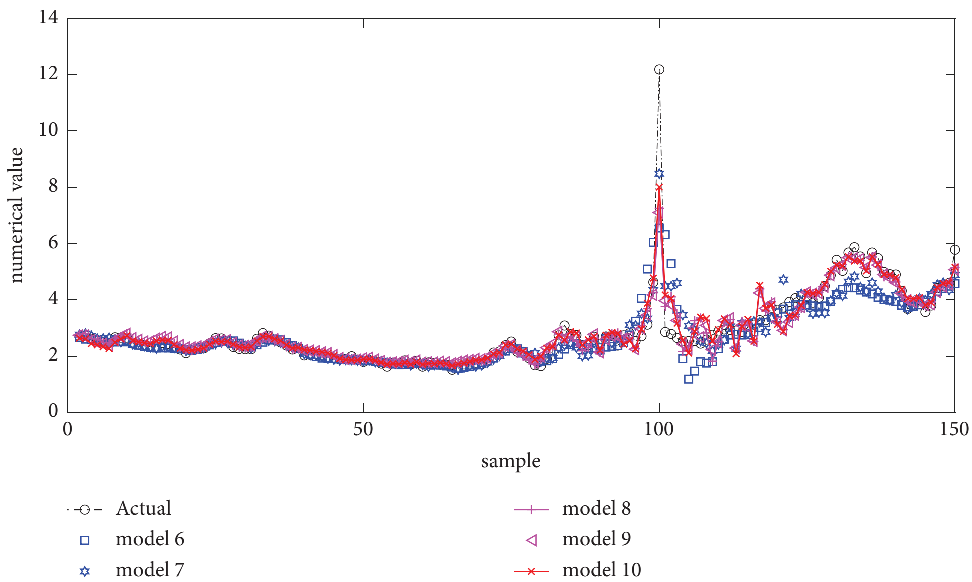

Figure 8: Comparison chart of three-step-ahead forecasting results of different hybrid models.

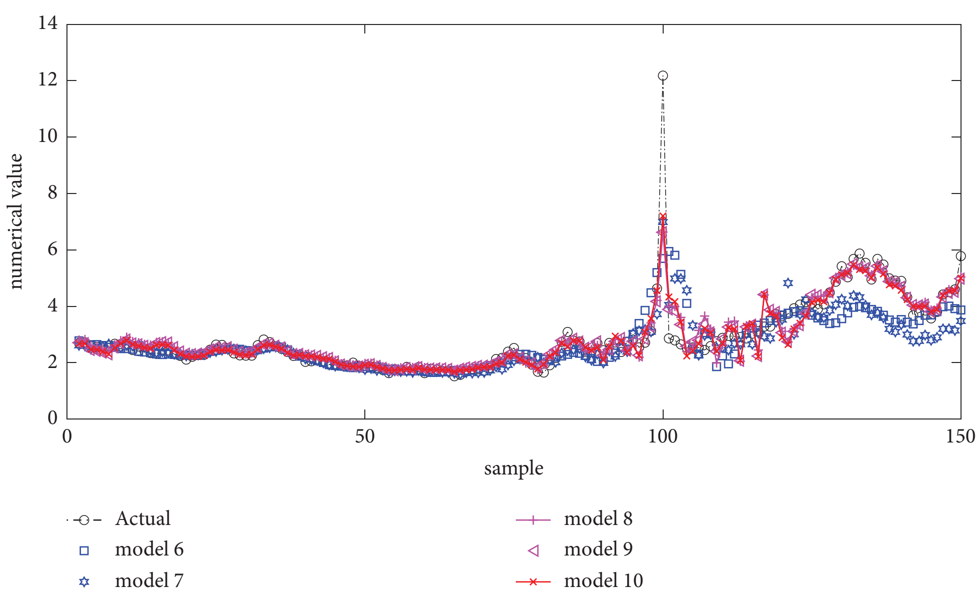

Figure 9: Comparison chart of five-step-ahead forecasting results of different hybrid models.

forecasting scenario as an example, the values of e_{RMSE}, e_{MAE}, and e_{MAPE} for the EEMD-LSTM, EEMD-VMD-LSTM, and VMD-LSTM models are (0.6746, 0.2958, 0.0906), (0.4087, 0.2021, 0.0621), and (0.4397, 0.1948, 0.0606), respectively. Compared with the EEMD-LSTM and EEMD-VMD-LSTM models, the predicted results of the VMD-LSTM model are improved by (34.82%, 34.14%, 33.11%) and (−0.75%, 3.61%, 2.41%), respectively. The result shows that in the prediction of natural gas prices, VMD technology has a stronger decomposition ability than EEMD when it pertains to complex price data and can better extract sequence data features and process complex signals.

Further, by comparing the prediction results of the VMD-LSTM and VMD-Res.-LSTM models, it can be concluded that the VMD-Res.-LSTM—which incorporates the residual term—can predict the natural gas price series more accurately. Specifically, taking the numerical results in the one-step-ahead forecasting scenario as an example, the values of e_{RMSE}, e_{MAE}, and e_{MAPE} of the VMD-LSTM and VMD-Res.-LSTM models are (0.4397, 0.1948, 0.0606) and (0.4378, 0.1834, 0.0555), respectively. Compared with the prediction results of the VMD-LSTM model, the VMD-Res.-LSTM model improved by (4.32%, 5.85%, and 7.75%), respectively. This result shows that, after applying the VMD technology to the natural gas price series, the obtained residual term contains rich information. After the residual term is decomposed again by EEMD technology, the dynamic changes of price data can be captured more completely; then, the prediction accuracy of the whole model can be improved.

Finally, by comparing the prediction results of the VMD-EEMD-Res.-LSTM hybrid model with those of other benchmark models, it can be concluded that the model's prediction performance—which combines the quadratic decomposition technology and LSTM model—is better than other hybrid models in the forward one-step-, three-step-, and five-step-ahead forecasting scenarios. This result shows that, in the hybrid model research that combines the LSTM model and the decomposition technology, the method of incorporating the residual term (which contains complex information) into the model for quadratic decomposition prediction analysis is not only suitable for short-term forecasting but also for long-term, multi-step forecasting. This further confirms the robustness and superiority of the prediction performance of the model constructed in this paper.

4. Conclusion

There are noises and redundant information present in natural gas price data. Considering the deficiencies of the existing hybrid forecasting models in the gas price area, we propose in this paper the VMD-EEMD-Res.-LSTM hybrid model for this area. This paper's goal is to fill the gaps in the existing research on natural gas price prediction. Specifically, the advanced decomposition algorithm and LSTM model are applied to predict natural gas prices, reducing the complexity of natural gas price series and effectively improving prediction accuracy. Additionally, to verify the accuracy, robustness, and the model's prediction performance, multistep-ahead forecasting is carried out for all models based on different evaluation indexes. The empirical analysis results show that, compared with other benchmark models, the hybrid model constructed in this paper effectively improves the prediction accuracy of the natural gas price series.

Although the model constructed in this paper has—to a certain extent—improved the prediction effect of natural gas price series, the research in this paper only involves a single price variable, which could be a limitation considering the trend of natural gas price series being influenced by multidimensional factors. Therefore, in future research, other influencing factors can be included to explore the effectiveness of different factors on price forecasting and improve the performance of the model forecasting.

Data Availability

The price data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This research was funded by the National Natural Science Foundation of China, No.71973028.

References

[1] J. Wang, M. Liu, Y. Bentley, L. Feng, and C. Zhang, "Water use for shale gas extraction in the Sichuan Basin, China," *Journal of Environmental Management*, vol. 226, pp. 13–21, 2018a.

[2] A. Duchowny, O. Mohnke, H. Thern et al., "Composition analysis of natural gas by combined benchtop NMR spectroscopy and mechanistical multivariate regression," *Energy Reports*, vol. 8, pp. 3661–3670, 2022.

[3] T. Wang and B. Lin, "China's natural gas consumption peak and factors analysis: a regional perspective," *Journal of Cleaner Production*, vol. 142, pp. 548–564, 2017.

[4] P. Jiang, Z. K. Liu, J. Z. Wang, and L. F. Zhang, "Decomposition-selection-ensemble forecasting system for energy futures price forecasting based on multi-objective version of chaos game optimization algorithm," *Resources Policy*, vol. 73, Article ID 102234, 2021.

[5] I. Ergen and I. Rizvanoghlu, "Asymmetric impacts of fundamentals on the natural gas futures volatility: an augmented GARCH approach," *Energy Economics*, vol. 56, pp. 64–74, 2016.

[6] F. Lu, F. Ma, P. Li, and D. S. Huang, "Natural gas volatility predictability in a data-rich world," *International Review of Financial Analysis*, vol. 83, Article ID 102218, 2022.

[7] D. P. Murphy, "The role of inventories and speculative trading in the global market for crude oil," *Journal of Applied Econometrics*, vol. 29, no. 3, pp. 454–478, 2014.

[8] B. Scholtens, "Modeling natural gas price volatility: the case of the UK gas market," *Energy*, vol. 72, pp. 126–134, 2014.

[9] Y. Xiao, "Forecasting extreme risk using regime-switching GARCH models: a case from an energy commodity," *International Journal of Emerging Markets*, vol. 16, no. 8, pp. 1556–1582, 2021.

[10] L. Wang, J. B. Wu, Y. Cao, and Y. R. Hong, "Forecasting renewable energy stock volatility using short and long-term Markov switching GARCH-MIDAS models: either, neither or both?" *Energy Economics*, vol. 111, Article ID 106056, 2022.

[11] X. D. Lv and X. Shan, "Modeling natural gas market volatility using GARCH with different distributions," *Physica A: Statistical Mechanics and Its Applications*, vol. 392, no. 22, pp. 5685–5699, 2013.

[12] Y. Y. Xu and D. Lien, "Forecasting volatilities of oil and gas assets: a comparison of GAS, GARCH, and EGARCH models," *Journal of Forecasting*, vol. 41, no. 2, pp. 259–278, 2022.

[13] Y. Chen, K. He, and G. K. Tso, "Forecasting crude oil prices: a deep learning-based model," *Procedia Computer Science*, vol. 122, pp. 300–307, 2017.

[14] S. M. Al-Fattah, "A new artificial intelligence GANNATS model predicts gasoline demand of Saudi Arabia," *Journal of Petroleum Science and Engineering*, vol. 194, pp. 107528–10, 2020.

[15] D. Kontogiannis, D. Bargiotas, A. Daskalopulu, A. I. Arvanitidis, and L. H. Tsoukalas, "Error compensation enhanced day-ahead electricity price forecasting," *Energies*, vol. 15, no. 4, p. 1466, 2022.

[16] E. Ceperic, S. Zikovic, and V. Ceperic, "Short-term forecasting of natural gas prices using machine learning and feature selection algorithms," *Energy*, vol. 140, pp. 893–900, 2017.

[17] D. Mouchtaris, E. Sofianos, P. Gogas, and T. Papadimitriou, "Forecasting natural gas spot prices with machine learning," *Energies*, vol. 14, no. 18, p. 5782, 2021.

[18] Q. Wu and H. Lin, "Short-term wind speed forecasting based on hybrid variational mode decomposition and least squares support vector machine optimized by bat algorithm model," *Sustainability*, vol. 11, no. 3, p. 652, 2019.

[19] W. Sun and C. A. Huang, "A carbon price prediction model based on secondary decomposition algorithm and optimized back propagation neural network," *Journal of Cleaner Production*, vol. 243, pp. 118671–11867113, 2020.

[20] C. J. Huang, Y. Shen, Y. H. Chen, and H. C. Chen, "A novel hybrid deep neural network model for short-term electricity price forecasting," *International Journal of Energy Research*, vol. 45, no. 2, pp. 2511–2532, 2021.

[21] S. Wang, L. Yu, and K. K. Lai, "Crude oil price forecasting with TEI@I methodology," *Journal of Systems Science and Complexity*, vol. 18, no. 2, pp. 145–166, 2005.

[22] J. Wang, J. X. Cao, S. Yuan, and M. Cheng, "Short-term forecasting of natural gas prices by using a novel hybrid method based on a combination of the CEEMDAN-SE-and the PSO-ALS-optimized GRU network," *Energy*, vol. 233, pp. 121082–12108215, 2021.

[23] W. P. Liu, C. Z. Wang, Y. G. Li, Y. S. Liu, and K. K. Huang, "Ensemble forecasting for product futures prices using variational mode decomposition and artificial neural networks," *Chaos, Solitons & Fractals*, vol. 146, Article ID 110822, 2021.

[24] M. Naderi, E. Khamehchi, and B. Karimi, "Novel statistical forecasting models for crude oil price, gas price, and interest rate based on meta-heuristic bat algorithm," *Journal of Petroleum Science and Engineering*, vol. 172, pp. 13–22, 2019.

[25] T. Zhang, Z. Tang, J. Wu, X. Du, and K. Chen, "Multi-step-ahead crude oil price forecasting based on two-layer decomposition technique and extreme learning machine optimized by the particle swarm optimization algorithm," *Energy*, vol. 229, pp. 120797–12079713, 2021.

[26] O. Taylan, "Modelling and analysis of ozone concentration by artificial intelligent techniques for estimating air quality," *Atmospheric Environment*, vol. 150, pp. 356–365, 2017.

[27] M. Wang, L. Zhao, R. Du et al., "A novel hybrid method of forecasting crude oil prices using complex network science and artificial intelligence algorithms," *Applied Energy*, vol. 220, pp. 480–495, 2018b.

[28] R. W. Liu, M. Liang, J. Nie et al., "STMGCN: mobile edge computing-empowered vessel trajectory prediction using spatio-temporal multigraph convolutional network," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 11, pp. 7977–7987, 2022a.

[29] Z. Qian, Y. Pei, H. Zareipour, and N. Chen, "A review and discussion of decomposition-based hybrid models for wind energy forecasting applications," *Applied Energy*, vol. 235, pp. 939–953, 2019.

[30] J. C. Li, Q. Q. Wu, Y. Tian, and L. G. Fan, "Monthly Henry Hub natural gas spot prices forecasting using variational mode decomposition and deep belief network," *Energy*, vol. 227, pp. 120478–12047815, 2021.

[31] H. D. Chacon, E. Kesici, and P. Najafirad, "Improving financial time series prediction accuracy using ensemble empirical mode decomposition and recurrent neural networks," *IEEE Access*, vol. 8, pp. 117133–117145, 2020.

[32] J. Wang and J. Wang, "A new hybrid forecasting model based on SW-LSTM and wavelet packet decomposition: a case study of oil futures prices," *Computational Intelligence and Neuroscience*, vol. 2021, pp. 7653091–22, 2021.

[33] Y. S. Huang and Y. Deng, "A new crude oil price forecasting model based on variational mode decomposition," *Knowledge-Based Systems*, vol. 213, pp. 106669–106711, 2021.

[34] B. Walczak and D. L. Massart, "Noise suppression and signal compression using the wavelet packet transform," *Chemometrics and Intelligent Laboratory Systems*, vol. 36, no. 2, pp. 81–94, 1997.

[35] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

[36] N. Bokde, A. Feijoo, D. Villanueva, and K. Kulat, "A review on hybrid empirical mode decomposition models for wind speed and wind power prediction," *Energies*, vol. 12, no. 2, p. 254, 2019.

[37] Z. Wu and N. E. Huang, "A study of the characteristics of white noise using the empirical mode decomposition method," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 460, no. 2046, pp. 1597–1611, 2004.

[38] P. Pandey, N. D. Bokde, S. Dongre, and R. Gupta, "Hybrid models for water demand forecasting," *Journal of Water Resources Planning and Management*, vol. 147, no. 2, pp. 40201061–402010613, 2021.

[39] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, 2014.

[40] J. H. Zhang, J. Yan, D. Infield, Y. Liu, and F. s. Lien, "Short-term forecasting and uncertainty analysis of wind turbine power based on long short-term memory network and Gaussian mixture model," *Applied Energy*, vol. 241, pp. 229–244, 2019.

[41] V. Gundu and S. P. Simon, "PSO-LSTM for short term forecast of heterogeneous time series electricity price signals," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 2, pp. 2375–2385, 2021.

[42] R. W. Liu, M. H. Liang, J. T. Nie, W. Y. B. Lim, Y. Zhang, and M. Guizani, "Deep learning-powered vessel trajectory prediction for improving smart traffic services in maritime internet of things," *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 5, pp. 3080–3094, 2022b.

[43] H. Hu, J. F. Zhang, and T. Li, "A comparative study of VMD-based hybrid forecasting model for nonstationary daily streamflow time series," *Complexity*, vol. 2020, pp. 4064851–4064921, 2020.