IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 15, 2022, accepted May 9, 2022, date of publication May 13, 2022, date of current version May 19, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3174814

Variable Weights Combination MIDAS Model Based on ELM for Natural Gas Price Forecasting

LUE LI [1,2], CAIHONG HAN [1], SHENGWEI YAO [3], AND LIANGSHUO NING [1]

[1]School of Information and Statistics, Guangxi University of Finance and Economics, Nanning 530003, China
[2]Guangxi Key Laboratory of Cross-Border E-Commerce Intelligent Information Processing, Guangxi University of Finance and Economics, Nanning 530003, China
[3]Guangxi (ASEAN) Research Center of Finance and Economy, Guangxi University of Finance and Economics, Nanning 530003, China

Corresponding author: Caihong Han (hchong_math@163.com)

This work was supported in part by the Project of Guangxi Science and Technology under Grant GuikeAD21220082, Grant GuikeAD21220102, and Grant GuikeAD20238072; in part by the Basic Ability Improvement Project for Young and Middle-Aged Teachers in Guangxi Province under Grant 2020KY16016 and Grant 2022KY0652; in part by the Guangxi University of Finance and Economics for Doctoral Projects (BS201974); in part by the Guangxi Natural Science Foundation under Grant 2020GXNSFAA159014 and Grant 2021GXNSFAA220114; in part by the Guangxi Key Laboratory of Cross-Border E-Commerce Intelligent Information Processing; in part by the Guangxi Key Laboratory of Finance and Big Data; and in part by the Program for Innovative Team of the Guangxi University of Finance and Economics.

ABSTRACT Accurate and stable natural gas price forecasts are essential for effective management of energy systems. However, due to the mixed frequency of data and the inherent nonlinear fluctuation characteristics of natural gas price changes, it is difficult to achieve satisfactory forecasting performance. In order to effectively improve the forecasting results of mixed frequency data. In this study, the MIDAS regression model and the machine learning MIDAS models are successfully combined to form a novel combination forecasting model. Moreover, the extreme learning machine with the multi-objective grey-wolf algorithm is used to combine univariate MIDAS results and further improve the forecasting accuracy. In the empirical analysis, the weekly natural gas futures prices of the Intercontinental Exchange UK NBP are used to generate real-time forecasts to evaluate the forecasting performance of the proposed combination model. The experimental results show that the comprehensive forecasting accuracy of the novel combination MIDAS model is 26.35%, 8.82% and 12.91% higher than that of the benchmark MIDAS regression models, the combination MIDAS models and the multivariate MIDAS models, respectively. Based on the forecasting results, the non-linear, non-stationary and irregular natural gas futures prices can be effectively managed, which provides better investment and management tools.

INDEX TERMS Natural gas price forecasting, mixed data sampling model, combination-MIDAS-MOGWO-ELM model, forecasting accuracy.

I. INTRODUCTION

In this section, the background of natural gas energy and the mixed data sampling model (MIDAS), the literature review and our research innovation will be discussed in detail.

A. BACKGROUND

Over the past two decades, governments and scientists have been discussing alternative sources and global applications of clean and efficient energy. Global primary energy consumption is increasing at an alarming rate, and human production activities are becoming more urgent for energy needs, which has environmental implications. Natural gas has gained much attention. As the only low-carbon and clean energy among fossil fuels, natural gas is becoming a major force in promoting the global energy transition. In this endeavour, natural gas has become increasingly important as an alternative energy source throughout Europe.

Currently, the EU's main sources of energy are still mainly coal, natural gas and oil. In 2020, the EU's primary energy pattern will still be dominant, with fossil fuels such as oil, natural gas and coal accounting for about 75% [1]. Improving renewable energy and energy efficiency will promote a fair environmental transition to reduce associated CO2 emissions. To meet Paris Agreement in 2015, efforts have been made worldwide to significantly reduce greenhouse gas

The associate editor coordinating the review of this manuscript and approving it for publication was Mauro Tucci.

emissions [2]. The EU is addressing the "European Green Agreement", which aims to transform the EU into a modern, resource-efficient and competitive economy. In 2050, the net greenhouse gas emissions should stop increasing and decouple economic growth from resources [3].

In view of a future where renewable energy dominates in the EU, natural gas-to-gas technology play an important role. The green transition is promoted to curb the carbon emissions related to fossil fuels to be in line with the Paris Agreement and the European Green Agreement. As the earliest and most mature virtual natural gas trading point in Europe, the UK National Balancing Point (NBP) has become the benchmark for European natural gas market transactions in Europe. Through marketisation, policy makers have created incentives for companies to switch to green energy production mechanism. An important reference issue in the natural gas market is accurate price forecasts. First, the results of natural gas price forecasts and the understanding of determinants are important for market participants who use price information to make investment decisions. Second, information on changes in natural gas prices is conducive to establishing reasonable market orientation, enabling government managers to comprehensively evaluate green energy policies and make timely adjustments [4], [5]. Therefore, more accurate forecast of natural gas prices is essential for establishing an effective natural gas market and adjusting energy information policies. This article focuses on the natural gas price forecasting model, because researchers still need to solve many problems through experience, especially in terms of adjusting natural gas prices to the energy market and economic environment.

B. LITERATURE REVIEW

With the rapid changes in the natural gas price market, multi-source heterogeneous and mixed frequency data modeling has become desirable and challenging [6]. The forecasting of natural gas model formed by the discussion mainly focuses on two aspects. First, the natural gas forecasting model judges the trend of natural gas prices based on regression methods. Second, the natural gas forecasting model is evaluated by the machine learning model.

The statistical models for natural gas mainly include time series models and regression models. In the time series model, Erdogdu fitted the ARIMA model to forecast Turkey's natural gas demand and compared the results with the official forecasts to find that the elasticity of Turkey's natural gas consumption was low [7]. Dilaver established a structural time series technique to investigate the influence of income, current natural gas prices and potential energy demand trends on European natural gas consumption [8]. Khani proposed an online calibrated time series model from historical temperatures similar to autoregressive model to effectively improve the accuracy of natural gas forecasting [9]. In the regression model, Zhang proposed the combined method to forecast the regression model of natural gas consumption by Bayesian model averaging theory to improve the accuracy of prediction [10]. Chen proposed a novel natural gas demand forecasting model and developed exogenous variables into functional autoregressive model to produce high-resolution forecasts of natural gas flow [11]. Sen used socioeconomic indicators as descriptive variables and predicted future natural gas consumption through various multiple regression models [12]. Despite the popularity of statistical models, the model essentially assumes linear time correlation, which is inconsistent, especially for natural gas data with nonlinear properties.

In machine learning theory on natural gas forecasting, neural networks have been used to study the dependence of nonlinear features on energy information, including short-term natural gas forecasts. The multilayer perceptron (MLP), a special form of neural networks, is widely used and considered as a benchmark for various analyzes of energy forecasts. Szoplik proposed to use the MLP model for forecasting daily natural gas consumption in Szczecin (Poland) [13]. Dombayci developed an MLP model for forecasting the hourly heating energy consumption of a model house in Denizli (Turkey) and compared it with the recently developed Deep Neural Network (DNN) algorithm [14]. Afshin adapted the multilayer perceptron (MLP) and Radial Basis Function Neural Network (RBFNN) to predict the equilibrium dew point in a natural gas stream [15]. Qiao integrated the functions of BPNN, GRNN and ELM and forecasted the consumption of natural gas after integration to obtain good forecasting results [16]. It is worth mentioning that the artificial neural network can better deal with the problem of nonlinear characteristic data, but some shortcomings are still existed. For example, the generalization ability of MLP is insufficient, and difficult to effectively deal with multi-dimensional data. BPNN is easy to fall into local optimum, resulting in slow convergence speed.

In addition, support vector regression in machine learning models the principles of structural risk minimization and kernel function, effectively avoiding the shortcomings of artificial neural networks [17], [18]. Han considered the effects of temperature and holidays on consumption and adjusted SVM and LS-SVM to forecast daily natural gas consumption in Xi'an. The results show that the support vector machine forecasting method is better than the neural network method [19]. Zhu proposed an improved SVR to forecast the short-term consumption of British National Natural Gas Pipeline and proved that the error of the model is lower than that of ARIMA and ANN [20].

The above forecast model represents the forecast technology system for natural gas. Some forecasting technology models ignore relevant factors, which leads to the decrease of forecasting accuracy. The energy factor is the main determinant of natural gas prices due to energy and fuel competition in the market. The existing literature on the relationship between natural gas prices and energy factors mainly focuses on coal and oil factors. Coal and oil, as the main energy sources for consumption, have a great influence on natural gas price consumption. Sen believed that coal and oil prices have a greater impact on natural gas price [12]. Destek believed that the consumption of natural gas, coal and oil were related to each other and any shocks from natural gas, coal, and oil are long-lasting effects on most OECD countries [21]. Yang evaluated the relationship between natural gas, coal and oil by studying ethylene glycol and concluded that the technical and environmental performance of natural gas route is superior to oil and coal route [22]. Political factors also determine the price of natural gas. For example, the aforementioned ''Paris Agreement'' and ''European Green Deal'' call for continuous reduction of carbon emissions in the environment. Brehm believed in the negative correlation between natural gas consumption and carbon emissions [23]. In addition to energy factors, economic factors are important to natural gas consumption. When the economy is booming, the industry faces an increase in demand and provides more production. Bildirici compared the relationship between national economic change and the consumption of coal, oil and natural gas [24]. Wu studied that natural gas consumption provides the opportunity to achieve the double dividend of economic growth and CO2 emission reduction [25]. Magazzino developed the direction from dependency algorithm for natural gas consumption and economic growth, which used artificial neural network methods in Japan and Germany [26].

In view of the above, the importance of the related influences contained in energy factors and economic factors for the evaluation and forecast of natural gas has been pointed out. The above-mentioned forecasting model requires the identification of relevant driving factors when forecasting natural gas in order to improve the accuracy of the forecast. The multifactor forecasting model does consider the influence of exogenous variables. However, the above works are all based on the premise that the exogenous variables in forecasting natural gas are performed with the same frequency, which inevitably leads to the problem of error accumulation and inaccurate natural gas forecasting. In fact, relevant data are published with different frequencies and accompanied by non-linear and non-stationary characteristics. The current research shows that natural gas forecasting ignores high frequency data and non-linear information carried by the data. Therefore, new methods need to be explored to improve the accuracy of natural gas by using high-frequency exogenous variables.

Mixed data sampling (MIDAS) can forecast the corresponding low-frequency data by high-frequency exogenous variables. The MIDAS regression model has a wide range of applications in the financial market [27]–[31]. The univariate MIDAS model was proposed by Ghysels et al [32], [33]. For the MIDAS model, the volatility forecasting of the MIDAS model is performed by the GARCH class [34]–[39]. GARCH-MIDAS type models mostly study the influence of univariate mixed frequency data and multivariate MIDAS are processed by combination univariate to forecast. It is difficult to deal with the problems of multivariate MIDAS model. Xu proposed to apply the artificial neural network model to the univariate and multivariate MIDAS model, which proved that MIDAS-BPNN is an effective tool for processing nonlinear mixed data [40]. Xu proposed to use MIDAS-SVR to process the multivariate mixed data forecasting model with multiple emotional factors [41]. Pan constructed MIDAS and multi-output models and proposed a multi-output model MIDAS-MSVM to obtain the results of multiple consecutive points simultaneously [42]. Li combined the univariate and multivariate MIDAS with the Extreme Learning Machine to propose the MIDAS-ELM model and process incomplete data that demonstrated the effectiveness of the proposed MIDAS-ELM model [43]. The MIDAS-GRNN and MIDAS-ENN models used in this paper have not been proposed in the existing literature, but they can still be constructed according to the properties of the multivariate MIDAS theory and neural network structure.

Accounting for non-linearity, non-stationary and irregular data in the multivariate MIDAS model and actual natural gas market conditions. Different from the traditional univariate MIDAS regression model and the multivariate MIDAS regression model, this paper develops a novel combination-MIDAS forecasting model. Then, we utilize the novel combination-MIDAS-ELM model to conduct an empirical analysis of natural gas prices and compare them with three types of competitive models. (1) Benchmark types: ARIMA and GARCH; (2) Combination-MIDAS model: combination-MIDAS-regression model, combination-MIDAS-SVR model, combination-MIDAS-BPNN model, combination-MIDAS-GRNN model, combination-MIDAS-ELM model and combination-MIDAS-ENN model; (3) Multivariate MIDAS machine learning models: MMIDAS-SVR, MMIDAS-BPNN, MMIDAS-GRNN,MMIDAS-ELM and MMIDAS-ENN.

C. INNOVATIONS AND FRAMEWORK

In general, the setting of the fixed weight coefficient and the variable weight coefficient are two forms of the weight of the combination forecasting model [44], [45]. Considering the multivariate influence factors and the reparable advantages of the combination model, we have gradually proposed a novel variable weight forecasting model of the mixed frequency data, which can overcome the shortcomings of the estimated weight coefficients of the combination in the original forecasting of the mixed frequency data combination and achieve higher forecasting accuracy. In this research, a variable weight combination-MIDAS forecasting model is proposed to perform weekly natural gas price forecasting. The novel variable weight combination model proves to be effective in short-term weekly forecasting due to its excellent performance and improves the overall forecasting effect.

The contribution of the novel combination-MIDAS-ELM model is summarized as follows:

(1) The research method of the combination-MIDAS forecasting model was extended. Each univariate forecasting result of traditional MIDAS forecasting model is based on regression methods. This paper combined machine learning MIDAS models processing nonlinear data, the integration of regression MIDAS model and machine learning MIDAS models, which expanded the research scope of mixed

frequency data combination model and improved the forecasting results accuracy.

(2) The combination-MIDAS forecast model is combined with the extreme learning machine model to improve the accuracy. By constructing the combination-MIDAS forecasting model with novel variable weights, the combination forecasting results are improved. In particular, the combination-MIDAS forecasting model integrated by the extreme learning machine and multi-objective grey wolf optimization algorithm greatly improves the short-term forecasting performance.

(3) The extreme learning machine model is combined with the multi-objective grey wolf optimization algorithm. In the proposed combination-MIDAS forecasting model, the multi-objective grey wolf optimization algorithm is constructed to optimize the coefficients in the extreme learning machine, which overcomes the shortcomings of the single-objective optimization adjustment coefficient method. The proposed multi-objective optimization method effectively takes advantage of the component system while avoiding the low accuracy and instability of the single objective system.

(4) The combination-MIDAS-ELM model is constructed to forecast the weekly natural gas price in real time using the latest daily available factors. The forecast of weekly natural gas price is more informative for large traders in the energy market to optimize short-term trading strategies (e.g., natural gas companies).

The rest of the paper is organized as follows. In Section 2, the research on MIDAS regression and combination-MIDAS are reviewed and the newly proposed model is introduced in detail. In Section 3, the empirical and predictive value analysis is described, and the results of the empirical research are discussed and comparative analysis is performed. And summarize research in Section 4.

II. METHODOLOGY

A. THE MIDAS MODEL

MIDAS model is essentially a strictly parameterized form of regression that uses highly parsimonious distributed lag polynomials to model high frequency independent variables. Ensure simplified model specifications while allowing the use of data sampling at different frequencies. The influence of the independent variable on the dependent variable may have lag-order effect and the basic h-step-ahead regression MIDAS model to forecast the dependent variable y_t (low frequency) from the univariate variable (high frequency) can be expressed as MIDAS(m, K, h) :

$$y_t = \beta_0 + \beta_1 B\left(L^{1/m};\theta\right) x_{t-h}^{(m)} + \varepsilon_t, \tag{1}$$

where

$$B\left(L^{1/m};\theta\right) x_{t-h}^{(m)} = \sum_{k=0}^{K} \omega\left(k;\theta\right) L^{k/m} x_{t-h}^{(m)} = \sum_{k=0}^{K} \omega\left(k;\theta\right) x_{t-k/m}^{(m)}.$$

The parsimonious lagged coefficients of $\omega(k;\theta)$ is a key weight function. The MIDAS regression model depends on the polynomial weight $\omega(k;\theta)$ to connect the relationship of different frequency data. $L^{k/m}$ denotes the lag operator on $x_{t-h}^{(m)}$. m is the multiple of the high frequency variable data relative to the low frequency variable which appears in the same basic time unit and $k = 0, 1, 2, \ldots, K$ is the lag-order of the high-frequency variable. The appropriate parameter function structure of the MIDAS model is very critical and six parsimonious polynomial specifications $\omega(k;\theta)$ in this paper: beta density function with zero lag (Beta), beta density function with non-zero lag (BetaNN), exponential Almon lag polynomial (ExpAlmon), Almon lag polynomial (Almon), step function (Step) and unrestricted weight function (UMIDAS) have been considered [27]–[31].

In addition to building univariate MIDAS models with parsimonious parameter function $\omega(k;\theta)$, univariate MIDAS models can also be built by machine learning methods, such as MIDAS-SVR [41], MIDAS-BPNN [40], MIDAS-GRNN, MIDAS-ELM [43] and MIDAS-ENN. Machine learning learns and trains by constructing ordered number pairs. The following converts the basic MIDAS model formula (1) into ordered number pairs structure:

$$\begin{aligned} y_t &= \beta_0 + \beta_1 \times [b(0;\theta)x_{t-0/m}^m + b(1;\theta)x_{t-1/m}^m \\ &\quad + \ldots + b(K;\theta)x_{t-K/m}^m] + \varepsilon_t \\ &= \beta_0 + \beta_1 \times \left[(b(0;\theta), b(1;\theta), \ldots, b(K;\theta)) \begin{pmatrix} x_{t-0/m}^m \\ x_{t-1/m}^m \\ \vdots \\ x_{t-K/m}^m \end{pmatrix} \right] \\ &\quad + \varepsilon_t \end{aligned} \tag{2}$$

If let $\boldsymbol{b(K;\theta)} = (b(0;\theta), b(1;\theta), \ldots, b(K;\theta))$, $\boldsymbol{x_t} = (x_{t-0/m}^m, x_{t-1/m}^m, \ldots, x_{t-K/m}^m)^T$, Equation (2) turns

$$y_t = \alpha_0 + \alpha_1 \boldsymbol{b(K;\theta) x_t} + \varepsilon_t \tag{3}$$

In this way, the ordinal pair relationship of the MIDAS model for year t is established. Next, we establish the MIDAS model corresponding to different M years

$$\begin{cases} y_{t_1} = \alpha_{t_1,0} + \alpha_{t_1,1}\boldsymbol{b_{t_1}(K;\theta)x_{t_1}} + \varepsilon_{t_1} \\ y_{t_2} = \alpha_{t_2,0} + \alpha_{t_2,1}\boldsymbol{b_{t_2}(K;\theta)x_{t_2}} + \varepsilon_{t_2} \\ \cdots \\ y_{t_M} = \alpha_{t_M,0} + \alpha_{t_M,1}\boldsymbol{b_{t_M}(K;\theta)x_{t_M}} + \varepsilon_{t_M} \end{cases} \tag{4}$$

Further formula (4) can be written as

$$\boldsymbol{Y_t = \alpha_{t,0} + \alpha_{t,1} B_t(K;\theta) X_t + \varepsilon_t} \tag{5}$$

where $\boldsymbol{Y_t} = (y_{t_1}, y_{t_2}, \ldots, y_{t_M})^T$, $\boldsymbol{\alpha_{t,0}} = (\alpha_{t_1,0}, \alpha_{t_2,0}, \ldots, \alpha_{t_Q,0})^T$, $\boldsymbol{\alpha_{t,1}} = (\alpha_{t_1,1}, \alpha_{t_2,1}, \ldots, \alpha_{t_Q,1})^T$, $\boldsymbol{B_t(K;\theta)} = \boldsymbol{(b_{t_1}(K;\theta), b_{t_2}(K;\theta), \ldots, b_{t_Q}(K;\theta)x_{t_Q})}$, $\boldsymbol{X_t = (x_{t_1}, x_{t_2}, \ldots, x_{t_Q})^T}$, $\boldsymbol{\varepsilon_t} = (\varepsilon_{t_1}, \varepsilon_{t_2}, \ldots, \varepsilon_{t_Q})^T$. Equation (5) establishes an ordinal pair (X_t, Y_t) relation about time division for MIDAS model. Therefore, machine learning methods such as SVR, BPNN, GRNN, ELM and ENN can be used to learn and train on mixed frequency data.

The MIDAS regression model directly converts the mixed frequency data into the same frequency data through parsimonious parameter function $\omega(k;\theta)$ and further uses the regression model to forecast. The MIDAS machine learning model fully considers the nonlinear relationship between different frequency data. The following experiments show that the effect of the MIDAS machine learning model is better than that of the MIDAS regression model.

B. THE COMBINATION-MIDAS FORECASTING MODEL

Section A refers to the forecasting model as univariate MIDAS regression model and machine learning MIDAS model, so the best forecasting results are determined by the single independent variable. However, there is evidence that forecasting with multiple independent variables provides more accurate results and more stable performance [46], [47]. The combination forecasting method adjusts univariate MIDAS with different independent variables to construct the weighted average of the forecasting model, which can solve a variety of influencing factors and obtain more stable forecasts. The following formula gives the combination of N forecasting results from the univariate MIDAS model:

$$\hat{f}_{CM,T+h|T} = \sum_{j=1}^{N} \hat{w}_{j,T} \hat{y}_{j,T+h|T} \tag{6}$$

where $\hat{w}_{j,T}$ are the combination weight coefficients of the forecasting results of univariate MIDAS method, j represents the corresponding univariate forecasting index number, T represents the last observed value of univariate MIDAS model, $h = 1, 2, \ldots, H$ and $\hat{y}_{j,T+h|T}$ represents the Hth forecasting result obtained after univariate MIDAS model sample estimation. This article considers three types of $\hat{w}_{j,T}$. The mathematical expressions of three weight types are as follows:

1) FIXED EQUAL WEIGHT COEFFICIENT

The coefficients of the equal-weight combination model are considered as fixed weight coefficients. Considering the average comprehensive prediction results after each univariate prediction, the effects of the different univariate prediction results are averaged.

$$\hat{w}_{j,T} = 1/N \tag{7}$$

2) MSFE WEIGHT COEFFICIENT

The mean square forecast error (MSFE) combination method utilizes the mean square error to calculate the weight coefficient, and the formula to calculate the weight coefficient realizes the effect that the data contribute the near large part and the far small part in the weight distribution.

$$\hat{w}_{j,T} = m_{j,T}^{-1} / \sum_{j=1}^{n} m_{j,T}^{-1} \tag{8}$$

where $m_{j,T} = \sum_{i=T_0}^{t} (\delta^{i-T_0}(y_{j,T+s} - \hat{y}_{j,T+s|T}))^2/(t - T_0 + 1)$. If δ=1, $y_{j,T+s}$ represents the actual true value. $m_{j,T}$ is the MSFE of jth factor univariate MIDAS.

3) DMSFE WEIGHT COEFFICIENT

When δ=0.9 in the MSFE weighted forecast model discussed above, which becomes the DMSFE weight coefficient (Discounted Mean Square Forecast Error). δ weighs recent forecasts more heavily than distant ones through using a discounting factor [32], [33].

C. THE NOVEL COMBINATION-MIDAS-ELM MODEL

The combination forecasting model has become a prominent model in the field of forecasting, integrating the results of different forecasting models. In the result of combination-MIDAS forecasting obtained by the weight coefficient, the coefficient of combination weight is very important, and different combination coefficients obtain different combination forecasting results. As described above, in the three MIDAS combination forecast weight coefficient acquisition methods, each combination weight coefficient model has its own advantages and disadvantages. Generally, the combination forecast weight coefficients MIDAS are determined by the fixed weight coefficient structure and the variable weight coefficient structure, after obtaining different forecast results, the fixed weight coefficient and the variable weight coefficient structure of the different forecast results are directly used for the combination forecast. In this paper, when determining the weight coefficients of the combination-MIDAS forecast model, MOGWO-ELM is chosen to determine the variable combination weight coefficients.

Select different S high-frequency multiple independent variables influencing factors $x_1, x_2, \ldots, x_i, \ldots, x_S (i = 1, 2, \ldots, S)$, through formula (1) and the univariate MIDAS machine learning method, according to the different lag step $x_{i,t-k/m}^{(m)}$ of high-frequency variables to generate best M low-frequency forecasting results $y_{i1}^*, y_{i2}^*, \ldots, y_{iM}^*$ under different high-frequency multiple independent variables. The significance of this processing is that the best lag step $x_{i,t-k/m}^{*,(m)}$ and the corresponding MIDAS forecasting model can be obtained through different high-frequency multiple independent variables. Next, through S different high-frequency multiple independent variables influencing factors $x_1, x_2, \ldots, x_i, \ldots, x_S (i = 1, 2, \ldots, S)$, the best corresponding lag step $x_{i,t-k/m}^{*,(m)}$ and the best corresponding forecasting model, the scope of forecasting results are expanded. The new forecasting results $y_{i,1}^*, y_{i,2}^*, \ldots, y_{i,M}^*, y_{i,M+1}^*, \ldots, y_{i,M+N}^*$ is obtained, wherein the data length ratio of $y_{i,1}^*, y_{i,2}^*, \ldots, y_{i,M}^*$ and

$y^*_{i,M+1}, \ldots, y^*_{i,M+N}$ is 4 :1. Then, the combination-MIDAS-ELM model is constructed.

1) CONSTRUCTING THE COMBINATION-MIDAS-ELM MODEL

The extreme learning machine is the machine learning algorithm based on a single hidden layer feedforward neural network [48], [49]. The main feature of the extreme learning machine is to initialize the input weights randomly and the hidden layer to generate deviations. The learning process only needs to calculate the output weights. Extreme learning machine has the advantages of high learning efficiency and strong generalization ability. Given the mixed data for S forecasting results set containing $\left(\left\{Y_{i,t}\right\}_{t=1}^{M}, \left\{Y'_{i,t}\right\}_{t=M+1}^{M+N}\right), (i = 1, 2, \ldots, S)$ training samples, where $Y_{i,t} = (y^*_{i,1}, y^*_{i,2}, \ldots, y^*_{i,M})$, $Y'_{i,t} = (y^*_{i,M+1}, \ldots, y^*_{i,M+N})$, input $Y_{i,t} \in R^M$ and corresponding required output $Y_{i,t}{}' \in R^N$, L is the number of hidden nodes, and $g(Y)$ is the activation function. The Combination-MIDAS-ELM expression can be written as:

$$\sum_{j=1}^{L} \beta_j g(\boldsymbol{\upsilon}_j \cdot Y_i + b_j) = Y'_i, i = 1, 2, \ldots, S \tag{9}$$

where $\boldsymbol{\upsilon}_j = [\upsilon_{j1}, \upsilon_{j2}, \ldots, \upsilon_{jn}]^T$ is the input weight vector connecting the jth hidden node and the input node, b_j is the deviation of the jth hidden node randomly selected, β_j is the output weight connecting the jth hidden neuron and the output neuron, $\boldsymbol{\upsilon}_j \cdot Y_i$ is the vector inner product operation of $\boldsymbol{\upsilon}_j$ and Y_i that represents all the operations generated between the input weight vector and the input vector, $g_j(Y)$ is the activation function of the jth hidden node.

The above equation (9) corresponding to S equations can be written as:

$$\boldsymbol{H\beta} = \boldsymbol{Y'} \tag{10}$$

where,

$$\boldsymbol{H}(\upsilon_1, \ldots, \upsilon_L, Y_1, \ldots, Y_S, b_1, \ldots, b_L) = \begin{bmatrix} g(\upsilon_1 \cdot Y_1 + b_1) & \ldots & g(\upsilon_L \cdot Y_1 + b_L) \\ \vdots & \ldots & \vdots \\ g(\upsilon_1 \cdot Y_S + b_1) & \ldots & g(\upsilon_L \cdot Y_S + b_L) \end{bmatrix}_{S\times L}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1^T \\ \beta_2^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L\times N} \text{ and } \boldsymbol{Y'} = \begin{bmatrix} y_1'^T \\ y_2'^T \\ \vdots \\ y_S'^T \end{bmatrix}_{S\times N}. \tag{11}$$

The goal of Combination-MIDAS-ELM is to find the solution $\boldsymbol{\beta}$ of the output weight through the linear equation (10). However, in most cases, $\boldsymbol{H}$ is not necessarily guaranteed to be an invertible matrix, and $\boldsymbol{\beta}$ may not exist. The way to solve the problem is to use the Moore-Penrose pseudo-inverse matrix $\boldsymbol{H}^\dagger$, which requires $\boldsymbol{H}^\dagger$ to be invertible [50].

The Combination-MIDAS-ELM has the advantages of high learning efficiency and strong generalization ability, but the input weight vector $\boldsymbol{\upsilon_j}$ of the input node and the deviation $\boldsymbol{b_j}$ of the hidden node are randomly generated, which affects the forecasting accuracy and stability. The multi-objective grey wolf optimization algorithm is applied to determine the optimal input weight vector and the deviation of the hidden node of the Combination-MIDAS-ELM.

2) MULTI-OBJECTIVE GREY WOLF OPTIMIZATION

MOGWO is a multi-objective variant with respect to GWO proposed by Mirjalili et al [51]. In MOGWO, the main natural source of inspiration is the behavior of grey wolves in the population. Grey wolves belong to the group of canines that strictly adhere to the supremacy of social hierarchy [51], [52]. The first level of social hierarchy α-wolf: the α-wolf is defined as the dominant wolf because the other wolves must obey the commands of the α-wolf. The leading α-wolf is mainly responsible for making decisions about predators, habitat, working and resting hours, and other activities. The second level of the social hierarchy is the β-wolf, who obeys the α-wolf and assists him in making decisions. When the α-wolf dies or ages, the β-wolf becomes the most important candidate for the α-wolf. Although β-wolves obey α-wolves, β-wolves may dominate wolves at other social levels. At the third level of the social hierarchy, δ-wolves obey α- and β-wolves while dominating other wolves. According to the social dominance leadership hierarchy and the positional relationship of each wolf, when the grey wolf is searching for the prey target, it will gradually approach and surround the prey target.

$$\vec{D} = |\vec{C}\overrightarrow{X_p}(t) - \vec{X}(t)| \tag{12}$$

$$\vec{X}(t+1) = \overrightarrow{X_p}(t) - \vec{A}\vec{D} \tag{13}$$

$$\vec{A} = 2\vec{a}\vec{r}_1 - \vec{a} \tag{14}$$

$$\vec{C} = 2\overrightarrow{r_2} \tag{15}$$

where t represents the current iteration, $\vec{D}$ represents the distance between the individual grey wolf and the prey; $\overrightarrow{X_p}(t)$ and $\vec{X}(t)$ respectively represent the location vector of the search target and the individual grey wolf. $\vec{A}$ and $\vec{C}$ are the coefficient vectors for the distance adjustment, where elements of $\vec{a}$ decrease from 2 to 0 in the iterative process and $\vec{r}_1$, $\vec{r}_2$ are random vectors in [0,1]. The $\vec{C}$ parameter is intentionally designed to provide a random value that can guarantee that iteration continues when the local optimum stagnates.

Grey wolves have the ability to identify the location of potential prey (optimal solution). The search process is mainly performed by leading (α, β, δ) grey wolves. However, the properties of the solution space are unknown. Grey wolves cannot determine the exact location of their prey (optimal solution). To simulate the search behavior of grey wolves (proposed solutions), it is assumed that (α, β, δ) have a strong ability to identify the location of potential prey. Therefore, in each iteration, keep the best three grey wolves in the current population, then update the positions of the other search agents according to their position information.

$$\vec{D}_\alpha = |\vec{C}_1\vec{X}_\alpha - \vec{X}| \quad (16)$$
$$\vec{D}_\beta = |\vec{C}_2\vec{X}_\beta - \vec{X}| \quad (17)$$
$$\vec{D}_\delta = |\vec{C}_3\vec{X}_\delta - \vec{X}| \quad (18)$$
$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 \cdot (\vec{D}_\alpha) \quad (19)$$
$$\vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot (\vec{D}_\beta) \quad (20)$$
$$\vec{X}_3 = \vec{X}_\delta - \vec{A}_3 \cdot (\vec{D}_\delta) \quad (21)$$
$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \quad (22)$$

$\vec{X}_\alpha$, $\vec{X}_\beta$, $\vec{X}_\delta$ respectively represent the position vector of the current population, $\vec{X}$ represents the position vector of the grey wolf; $\vec{D}_\alpha$, $\vec{D}_\beta$, $\vec{D}_\delta$ respectively represent the distance between the current grey wolf candidate and the best three wolves. When $|A| > 1$, the grey wolf will spread as far as possible in different areas and search for prey. When $|A| < 1$, the grey wolf will focus on hunting for prey in a specific area or areas, forcing the search agent to deviate from the prey.

To use GWO for multi-objective optimization, the Pareto-optimal solution is loaded into the archive and the distribution of solutions in the MOGOA archive is adjusted by selecting the objective of each solution. Two important components need to be considered. The first is the archive library, which is responsible for storing the received non-dominated Pareto-optimal solutions. The second component is the leader selection strategy, which is used to select solutions from and in the archive as the leader of the search process.

Loading the Pareto-optimal solution into the archive and hope that the archive distribution of each Pareto solution is as broad as possible to avoid the MOGWO search process that concentrates each solution too much in the archive, and to achieve a global multi-objective optimization solution. Use the leader mechanism to obtain the best non-dominant solution by roulette method in the archives. The probabilities are as follows:

$$P_i = \frac{c}{N_i} \quad (23)$$

where c is a constant greater than 1, and N_i is the number of Pareto optimal solutions obtained in the ith segment. It can be seen that if the current memory is not too full, the possibility of selecting a new leader in the next iteration increases, making the solution distribution of the Pareto-optimal front in the archiver more extensive.

The limit of the archive controller is the maximum memory size of the non-dominant Pareto-optimal solutions. An oversized archive controller increases the computational cost. In the iterative process, the non-dominant solutions obtained so far are compared with the solutions in the archive and the archive update can be performed. Given three MOGWO principles:

- If the new non-dominant solution does not dominate any of the non-dominant solutions in the archive, the non-dominant solution must not be included in the archive.
- If the new non-dominant solution is greater than or equal to the non-dominant solution in an archive, the new non-dominant solution should be selected as dominant, the original non-dominant solution in the archive should be discarded and the new non-dominant solution is included in the archive.
- The new non-dominant solution and the non-dominant solution in the archive do not dominate each other and the new non-dominant solution should be included in the archive.

In short, MOGWO has the ability to search the Pareto-optimal solution and maintain the optimal solution, which greatly improves the distribution of the optimal multi-objective solution.

3) THE STEPS OF NOVEL COMBINATION-MIDAS-ELM FORECASTING MODEL

Step 1: Selecting the best mixed data forecasting model of univariate influencing factors by MIDAS regression model and machine learning MIDAS model.

The univariate MIDAS mixed frequency model is suitable for forecasting low frequency variables under the influence of high frequency univariate. Considering the nonlinear features and some linear features in the mixed frequency data, different types of models must be selected to deal with nonlinear and linear problems, and then satisfactory performance can be achieved in the subsequent combination forecasting. As mentioned in the literature review, MIDAS regression models have excellent functions in linear forecasting, while machine learning MIDAS models are good in nonlinear forecasting. Therefore, six MIDAS regression models and five machine learning MIDAS models are constructed as the basic technology for mixed frequency data forecasting. Moreover, previous studies have found that the univariate MIDAS regression models [47], [53]–[55] and the machine learning MIDAS models have good performance[40], [56], [57]. Six univariate MIDAS regression models are suitable for capturing the short-term correlation of mixed frequency time series and have been widely used for short-term forecasting of mixed frequency data. In addition, the machine learning MIDAS models are suitable for establishing the nonlinear relationship between input and output data, which can greatly reduce the forecasting errors and provide excellent functions for non-linear MIDAS forecasting. The specific method is select various S high-frequency independent variables influencing factors $x_1, x_2, \ldots, x_i, \ldots, x_S (i = 1, 2, \ldots, S)$, through the univariate regression and machine learning MIDAS models, according to the different lag step $x_{i,t-k/m}^{(m)}$ of high-frequency variables to generate best M low-frequency forecasting results $y_{i1}^*, y_{i2}^*, \ldots, y_{iM}^* (i = 1, 2, \ldots, S)$.

Step 2: Constructing the novel Combination-MIDAS-ELM forecasting model

Due to the complexity of the mixed frequency data and the advantages and disadvantages of the univariate MIDAS forecasting models, it is important to choose appropriate and optimal weighting coefficients for the combination-MIDAS forecasting model to compensate for the shortcomings of the various univariate MIDAS models. In this paper, ELM is adopted as an important aspect of regression forecasting for the final results after integrating the univariate mixed frequency components. The specific method is based on the first step, again through S different high-frequency independent variables influencing factors $x_1, x_2, \ldots, x_i, \ldots, x_S (i = 1, 2, \ldots, S)$, and the corresponding best forecasting models, expand the scope of the forecasting results, and the new result as $y^*_{i,1}, y^*_{i,2}, \ldots, y^*_{i,M}, y^*_{i,M+1}, \ldots, y^*_{i,M+N}$, where the ratio of the data lengths of $y^*_{i,1}, y^*_{i,2}, \ldots, y^*_{i,M}$ and $y^*_{i,M+1}, \ldots, y^*_{i,M+N}$ are 4:1. The forecasting length ratio of mixed frequency data is beneficial for extreme learning machine training. Therefore, a set of ordinal pairs of forecasting outcomes for mixed frequency data is constructed.

Furthermore, the combination-MIDAS-ELM coefficient $(\boldsymbol{v}, \boldsymbol{b})$ is used as the combination weight of the univariate mixed frequency components. Considering that the combination-MIDAS-ELM coefficients $(\boldsymbol{v}, \boldsymbol{b})$ vary with the data changes, the combination weight coefficients can be considered as one variable. More importantly, combination-MIDAS-ELM is good at determining the fitted variables in the training set. Therefore, it is recommended to use combination-MIDAS-ELM to complete the adaptive updating of the weight combination coefficients. In other words, the theory based on ELM to forecast weight combinations of mixed frequency variables can be better adapted to various factors of the changes of mixed frequency data, and the inherent law can be found based on the training of the previous high frequency data and low frequency data. Thus, the processing method can greatly improve the forecasting accuracy of the mixed frequency data.

Combination-MIDAS-ELM forecasting results are important, but it is difficult to determine parameters $(\boldsymbol{v}, \boldsymbol{b})$. Therefore, MOGWO is recommended to adjust parameters $(\boldsymbol{v}, \boldsymbol{b})$ to improve regression forecasting performance. In order to comprehensively consider the accuracy and stability of the MIDAS forecasting effect, two objectives are pursued to achieve the accuracy and stability of the mixed frequency data results. The first objective is the mean absolute error (MAE), which represents the forecasting accuracy of the model for mixed frequency data and is denoted by $fun_1(x)$. The second objective is the standard deviation (Std), which can be used to evaluate the stability of the mixed frequency data denoted as $fun_2(x)$. Therefore, it is recommended that the two objective function forecasting models that combine the accuracy and stability of mixed frequency data are as follows.

$$\min \begin{cases} fun_1(x) = MAE = \dfrac{1}{N}\displaystyle\sum_{j=1}^{N} \left|\widehat{y}_j - y_j\right| \\ fun_2(x) = Std(\widehat{y}_j - y_j)\ , j = 1, 2, \ldots, N \end{cases} \tag{24}$$

where N denotes the number of the data sample, y_j and $\widehat{y}_j$ represent the real value and forecasting value, respectively. Optimizing by multi-objective grey wolf optimization algorithm to obtain the best weights and deviations of the combination-MIDAS-ELM model. Thus, a novel variable weighting combination-MIDAS-ELM forecasting model is proposed by combining the integration of multivariate mixed frequency data and MOGWO-ELM theory.

Step 3: Forecasting the weekly natural gas prices consider h-step-ahead conditions

The novel combination-MIDAS-ELM model and h-step-ahead conditions are implemented in the forecasting framework. The principle is that the high-frequency independent variable can have a lagged effect on the low-frequency dependent variable. When $h = 0$, it means that the high frequency data has no effect on the low frequency data without lagged effect. When $h = 1$, it means that the high-frequency data lag by one unit has an impact on the low-frequency data, and so on. Take a positive integer $h = 0, 1, 2, \ldots, H$ as the lag time and consider the lag time effects to get the final forecasting value $\hat{f}_{CM,T+h|T}$. Figure 1 show the flow path of Combination-MIDAS-ELM model.

III. EMPIRICAL RESULTS AND ANALYSIS

In this part, we show that the proposed model is effective in forecasting mixed frequency data and recommend the novel Combination-MIDAS-ELM model to forecast the weekly natural gas futures price in UK NBP Intercontinental Exchange.

A. SOURCE OF MIXED FREQUENCY DATA

The source of mixed frequency data includes futures prices from UK Intercontinental Exchange retrieved at weekly frequency, and daily energy and economic indicators retrieved at higher frequency. Specifically, this paper considers the continuous futures contract settlement price of UK NBP natural gas traded in Intercontinental Exchange. NBP natural gas prices are referred to as British Intercontinental Exchange for natural gas futures delivery and represent the direction of changes in European natural gas prices. The daily data of NBP natural gas futures prices are got from extending the futures contract. The weekly natural gas futures price is the average daily price during the week. By selecting daily price data of carbon emissions, coal and crude oil as energy interaction factors. The price of carbon emissions is taken from the EU Emissions Trading Scheme. The European Emissions Trading Scheme (EU-ETS) is the world's largest carbon emissions trading market. By regulating corporate carbon emissions, the world has made a major contribution to reducing carbon emissions. The price of coal is taken from the price of the Newcastle coal futures contract. Newcastle is a major city in the coal industry in the UK as well as the largest coal port in the UK and a major coal exporter. The crude oil price is the settlement price of Brent crude oil futures and is the most commonly used reference price for crude oil products in the European market. The economic index used is the UK FTSE 100 Futures Index, which is the leading European futures contract index. All data used can be obtained from the Wind database.

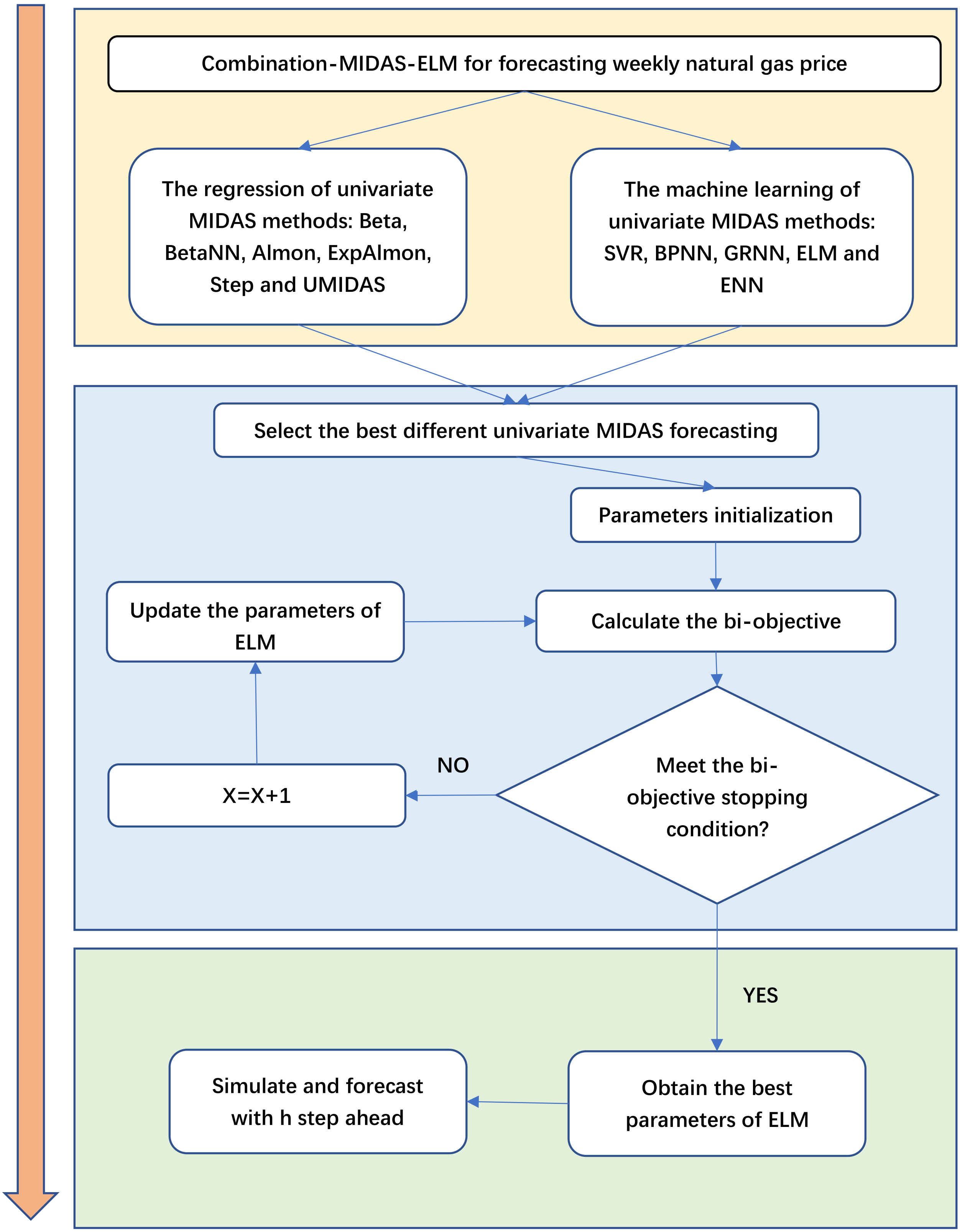

FIGURE 1. Flow path of combination-MIDAS-ELM model

In order to eliminate the heteroscedasticity in the subsequent empirical analysis, all the research objects are transformed into logarithmic rate of return growth rate variables, the expression is as follows:

$$return_{i,t} = \ln(value_{i,t}/value_{i,t-1}) \times 100\% \tag{25}$$

where $return_{i,t}$ represents the log return growth rate of the indicator in the period representing the prices of NBP natural gas, coal, crude oil and FTSE-100 respectively. In Figure 2, the rate of return growth of all variables in this paper is drawn by logarithmic transformation.

B. FORECASTING EVALUATION CRITERIA

A comprehensive forecast evaluation standard was proposed, including RMSE, MAE and MAPE. The detailed indicator expressions are shown in (26)-(29), and the optimal results are selected through comprehensive calculation expressions(CCE):

$$RMSE = \sqrt{\frac{1}{N} \times \sum_{j=1}^{N} (\widehat{y}_j - y_j)^2} \tag{26}$$

$$MAE = \frac{1}{N} \times \sum_{j=1}^{N} \left|(\widehat{y}_j - y_j)\right| \tag{27}$$

$$MAPE = \frac{1}{N} \times \sum_{j=1}^{N} \left|\frac{(\widehat{y}_j - y_j)}{y}\right| \times 100\% \tag{28}$$

$$CCE = \frac{1}{3} \times (RMSE + MAE + MAPE) \tag{29}$$

C. CHOOSING THE BEST UNIVARIATE MIDAS MODEL

The univariate model MIDAS uses the model shown in Section I and the machine learning models to forecast weekly natural gas prices. As the first step in constructing the combination-MIDAS-ELM model for weekly natural gas prices, selecting the best univariate MIDAS model directly affects the accuracy of the final forecasting results, which is an important step. Select the best univariate MIDAS model by comparing the CCE of different MIDAS regression models and machine learning MIDAS models. At the same time, the most sensitive natural gas forecast index is selected by the best univariate MIDAS model. The period from August 15, 2011 to July 6, 2018 is used to train the daily energy and economic data corresponding to the weekly natural gas price under the different univariate influencing factors, and the period from July 9, 2018 to March 29, 2020 is used to test the sample set. By comparing the CCE between the forecasting value and the true value, the best univariate MIDAS model is selected.

The optimal lag length of MIDAS regression and machine learning MIDAS parameters for daily indicators and weekly natural gas prices are obtained using the fixed-window regression method. In this paper, in order to reflect the change trend of CCE, the maximum lag order of carbon emission price, coal price, crude oil price and FTSE-100 index based

TABLE 1. The parameters setting of MIDAS regression models and machine learning MIDAS models.

Model	Parameter	Default value
MIDAS-Beta	Beta1	random
	Beta2	random
	Beta3	0
	Iterations number	10000
	Fitting requirements precision	0.0001
MIDAS-BetaNN	Beta1	1
	Beta2	random
	Beta3	random
	Iterations number	10000
	Fitting requirements precision	0.0001
MIDAS-ExpAlmon	ExpAlmon1	random
	ExpAlmon2	random
	ExpAlmonDegree	3
	Iterations number	10000
	Fitting requirements precision	0.0001
MIDAS-Almon	AlmonDegree	3
	Iterations number	10000
	Fitting requirements precision	0.0001
MIDAS-Step	Iterations number	10000
	Fitting requirements precision	0.0001
MIDAS-UMIDAS	Iterations number	10000
	Fitting requirements precision	0.0001
	Training requirements precision	0.0001
MIDAS-SVR	Input vector length	the X lag step of MIDAS
	C	1
	g	1
	Iterations number	10000
	Fitting requirements precision	0.0001
MIDAS-BPNN	Input nodes number	the X lag step of MIDAS
	Output nodes number	1
	Hidden nodes number	the X lag step of MIDAS
		lag step
	Learning rate	0.1
	Iterations number	10000
	Training requirements precision	0.0001
MIDAS-GRNN	Iterations Number	10000
	Input layer nodes number	the X lag step of MIDAS
	Pattern layer nodes number	the X lag step of MIDAS
	Summation layer nodes number	1
	Output layer nodes number	1
	Spread rate	0.6
MIDAS-ELM	Iterations Number	10000
	Input layer nodes number	the X lag step of MIDAS
	Output layer nodes number	1
	Hidden layer nodes number	the X lag step of MIDAS
MIDAS-ENN	Input nodes number	the X lag step of MIDAS
	Output nodes number	1
	Hidden nodes number	the X lag step of MIDAS
	Learning rate	0.1
	Iterations number	10000
	Training requirements precision	0.0001

on daily influence factors is set to 30. At the same time, the conditions of 0-step, 1-step, 2-step and 3-step ahead conditions are considered. The main parameter settings of MIDAS regression and machine learning MIDAS are listed in Table 1. It should be noted that the number of nodes of the hidden layer of machine learning MIDAS and the X lag step are set to be the same to maintain the consistency of the structure of the training set, and when the X lag step changes, the number of nodes of the hidden layer also changes accordingly.

In this part, we discuss the MIDAS regression models and machine learning MIDAS models of natural gas prices, and explain the mechanism for determining the best lag order and parameters. Table 2 show the CCE of the weekly natural gas forecast prices based on the univariate carbon

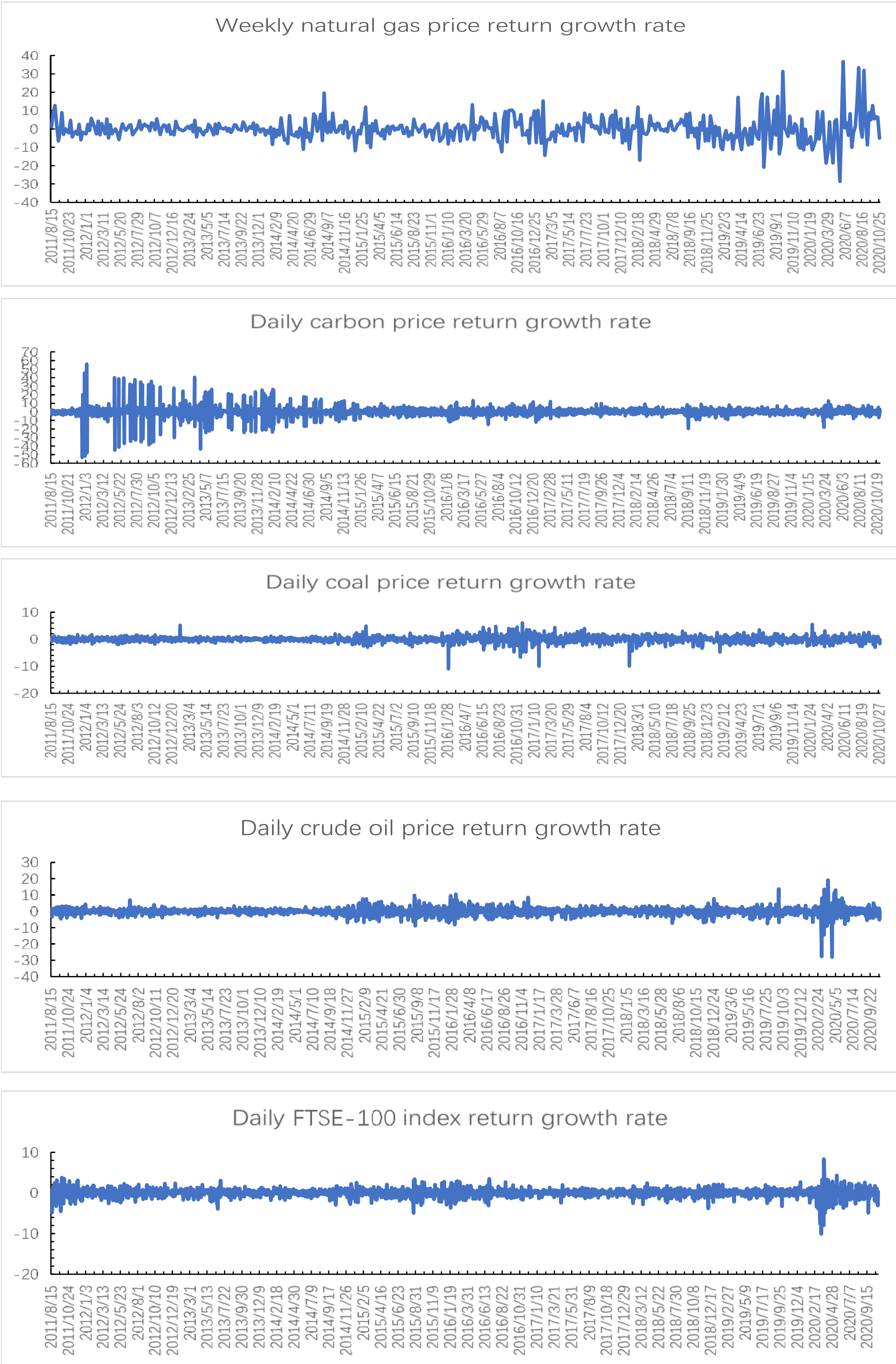

FIGURE 2. **Rate of return growth of weekly NBP natural gas and daily carbon prices, coal prices, crude oil prices and FTSE-100 index**

TABLE 2. The CCE results of daily carbon prices versus weekly natural gas prices when h=0.

Type	Xlag=5	Xlag=10	Xlag=15	Xlag=20	Xlag=25	Xlag=30
Beta	0.7088	0.7088	0.7088	0.7093	0.7093	0.7094
BetaNN	0.7093	0.7087	0.7109	0.7081	**0.7088**	0.7079
Almon	0.7097	0.7086	0.7104	0.7108	0.7112	0.7099
ExpAlmon	0.7097	0.7094	0.7102	0.7095	0.7097	0.7093
Step	0.7106	0.7106	0.7115	0.7119	0.7145	0.7145
UMIDAS	0.7091	0.7090	0.7098	0.7101	0.7122	0.7125
SVR	0.7122	0.7090	0.7151	0.7141	0.7095	0.7095
BPNN	0.7150	0.7114	**0.7060**	0.7095	0.7418	**0.7052**
RBFNN	0.7097	0.7095	0.7096	0.7095	0.7095	0.7096
ELM	**0.7068**	**0.7069**	0.7099	0.7073	0.7119	0.7080
ENN	0.7138	0.7232	0.7121	**0.7049**	0.7229	0.7069

price MIDAS regression model and the machine learning MIDAS models without step-ahead condition (h=0). When the lag order of the daily carbon price is 5 and 10, the best MIDAS models are ELM. It can be seen that the fitting and forecast effects of the Extreme Learning Machine are best when the high-frequency lag variables are short. When the lag order of daily carbon price is 15 and 30, the best MIDAS model is BPNN. When the lag of daily carbon price is 20, the best MIDAS model is ENN. If the lag order of the daily carbon price is 25, the best MIDAS polynomial is BetaNN. The empirical results show that the effect of machine learning MIDAS is generally better than the performance of MIDAS regression under all conditions of different lags of natural gas price, which indicates that the application of machine learning can better capture the relationship between natural gas prices and carbon prices when the lag of natural gas price changes. Moreover, the BPNN neural network achieves the best forecast performance when the lag order of carbon price is 30, namely BPNN-MIDAS (5, 30), which has the highest out-of-sample forecast accuracy. MIDAS (5, 30) represents the multiple of high frequency relative to low frequency and the specific representation can be found in the literature [56], [58].

Considering the different relationships between natural gas and carbon, coal, crude oil and FTSE-100, MIDAS regression and machine learning MIDAS models for natural gas prices and influencing factors are used to determine the best corresponding univariate MIDAS model. Select the best univariate MIDAS model (minimum CCE) under all conditions from the mechanism presented above. As shown in Table 3, the most accurate univariate MIDAS regression model for carbon price is the lag order of carbon price of 20 when the step forward of h is 1, BPNN-MIDAS (5,20). For coal, the most accurate univariate MIDAS regression model has the lag order of the coal price of 20 when the step ahead is 1, using the step polynomial weight Step-MIDAS (5,20). For crude oil, the univariate MIDAS regression model has the highest accuracy. The most accurate univariate MIDAS regression model is the lag order of crude oil at 10 when h is considered 2-step ahead, BPNN-MIDAS (5,10). For the FTSE-100, the most accurate univariate MIDAS regression model is the lag order of the FTSE-100 price at 20 when considered 1 step ahead, ELM-MIDAS (5,20). Of the four univariate factors affecting the analysis of natural gas price, natural gas is most sensitive to crude oil. When the step ahead of h is 0 to 3, the smallest CCE forecast effects are obtained because natural gas and crude oil are basically the same mining methods and then affect each other's prices, which determines the effect of crude oil on natural gas. From the results of the forecast indexs RMSE, MAE, MAPE and CCE, the corresponding machine learning model with a shorter lag step achieves better forecast results. In other words, this means that when the lag step size of the high-frequency independent variable is shorter, the effect on the low-frequency dependent variable is more obvious. This corresponds to the explanation that the previous maximum lag step size is selected as 30.

D. THE NOVEL COMBINATION-MIDAS-ELM MODELS RESULTS AND COMPARISON

The natural gas price driving factors considered have different information content, and the results of the combination forecast are usually better than the results of the single factor forecast [45], [56], [59]. In this paper, we consider the univariate MIDAS forecasting model, which is based on regression and machine learning. It not only uses a combination method to determine the weights of the individual forecast results of the best univariate MIDAS model, but also proposes the novel Combination-MIDAS-ELM model of the variable weighting scheme. Then select some competing models to evaluate the progress of the new model compared with the benchmark model.

1) FORECAST COMPARISON: THE NOVEL COMBINATION-MIDAS-ELM MODELS VS. ARIMA AND GARCH

In the first part of the comparison experiment, we compared the Combination-MIDAS-ELM with some other traditional benchmark combination models (such as ARIMA and GARCH). The detailed results of the forecast accuracy are shown in Table 4. The most satisfactory results are shown in bold in table 4.

Comparison of the experimental results showed that the novel Combination-MIDAS-ELM model has the most satisfactory evaluation standard. Specifically, the performance of ARIMA and GARCH forecast models are similar and inferior to that of novel Combination-MIDAS-ELM but not as good

TABLE 3. The best univariate MIDAS model under all conditions.

Indicator	h step ahead	BEST Univariate Model	RMSE	MAE	MAPE	CCE
Panel A: Carbon	0	BPNN-MIDAS(5,30)	1.8340	0.2406	0.0411	0.7052
	1	BPNN-MIDAS(5,20)	1.8376	0.2415	**0.0362**	**0.7051**
	2	SVR-MIDAS(5,5)	1.8598	**0.2364**	**0.0357**	0.7106
	3	SVR-MIDAS(5,5)	1.8574	**0.2353**	**0.0333**	0.7087
Panel B: Coal	0	ENN-MIDAS(5,10)	1.8263	0.2442	0.0454	0.7053
	1	Step-MIDAS(5,20)	1.8086	0.2391	0.0376	**0.6951**
	2	ELM-MIDAS(5,10)	1.8324	0.2437	0.0394	0.7052
	3	BPNN-MIDAS(5,10)	1.8416	0.2421	0.0378	0.7072
Panel C: oil	0	BPNN-MIDAS(5,5)	**1.8083**	**0.2358**	**0.0345**	**0.6929**
	1	ELM-MIDAS(5,20)	**1.7980**	0.2418	0.0386	**0.6928**
	2	BPNN-MIDAS(5,10)	**1.7666**	0.2391	0.0432	**0.6830**
	3	UMIDAS-MIDAS(5,10)	**1.8058**	0.2370	0.0355	**0.6927**
Panel D: FTSE-100	0	ELM-MIDAS(5,30)	1.8321	0.2444	0.0407	0.7058
	1	ELM-MIDAS(5,20)	1.8092	**0.2390**	0.0370	**0.6950**
	2	ELM-MIDAS(5,10)	1.8147	0.2382	0.0418	0.6982
	3	ELM-MIDAS(5,30)	1.8194	0.2428	0.0414	0.7012

TABLE 4. The novel combination-MIDAS-ELM models vs. ARIMA and GARCH.

Benchmark model	h step ahead	Forecast evaluation criteria			
		RMSE	MAE	MAPE	CCE
ARIMA	0	2.9434	0.4091	0.3413	1.2313
GARCH	0	14.8345	2.6610	2.0520	6.5158
Novel-Combination-MIDAS- ELM	0	**2.2204**	**0.3326**	**0.0371**	**0.8634**
ARIMA	1	3.2646	0.4691	0.3860	1.3732
GARCH	1	19.3309	3.3456	2.6080	8.4282
Novel-Combination-MIDAS- ELM	1	**2.5505**	**0.3317**	**0.0375**	**0.9732**
ARIMA	2	2.7562	0.3685	0.3086	1.1444
GARCH	2	18.6206	3.2285	2.5173	8.1221
Novel-Combination-MIDAS- ELM	2	**2.1034**	**0.2626**	**0.0373**	**0.8011**
ARIMA	3	2.7678	0.3717	0.3096	1.1497
GARCH	3	13.3448	2.3954	1.8475	5.8626
Novel-Combination-MIDAS- ELM	3	**2.5446**	**0.3308**	**0.0353**	**0.9702**

as novel Combination-MIDAS-ELM. When $h = 0, 1, 2, 3$, for ARIMA and GARCH models, among different forecast evaluation indicators RMSE, MAE, MAPE and CCE, the Combination-MIDAS-ELM proposed in this paper has the best effect, while the ARIMA model performs relatively better well. Notably, the mean MAPE values for ARIMA and GARCH are 33.64% and 225.62%, respectively, while the corresponding mean for Combination-MIDAS-ELM is 3.68%. It shows that the result of the ARIMA model is reasonable, the GARCH model is incorrect, and the forecasting result of the Combination-MIDAS-ELM model is the same excellent [60]. From the overall result CCE, the mean CCE values of ARIMA and GARCH are 1.2247 and 7.2322, respectively, while the corresponding mean of Combination-MIDAS-ELM is 0.9020.Compared with the best traditional ARIMA model, the average improvement is 26.35%. And it can be seen that when $h = 2$, the best forecast effect is obtained. This shows that in forecasting the weekly natural gas price, the daily influencing factors have the most obvious effect on the weekly natural gas price when the lag is two days. In other words, for the five working days of weekly data, the results of the third working day often have the most significant influence on that week. Therefore, we can conclude that the proposed Combination-MIDAS-ELM model outperforms the comparative benchmark model in terms of forecast accuracy. Therefore, Combination-MIDAS-ELM can improve the forecast performance of mixed frequency data to different degrees, with high forecast stability and accuracy.

2) FORECAST COMPARISON: THE NOVEL COMBINATION-MIDAS-ELM MODELS VS. THE REGRESSION AND MACHINE LEARNING COMBINATION-MIDAS MODELS

In the second part of the comparison experiment, we compared the novel Combination-MIDAS-ELM model with some other combination forecasting models (such as combination-MIDAS-Regression, combination-MIDAS-SVR, combinati on-MIDAS-BPNN, combination-MIDAS-GRNN, combinati on-MIDAS-ELM, combination-MIDAS-ENN). The detailed forecasting accuracy results are shown in Table 5 and 6. The most satisfactory results are also shown in bold.

The comparison of experimental results shows that the combination of machine learning MIDAS models are better than the combination regression MIDAS models. This is due to the nonlinear properties of the univariate MIDAS model data itself. The forecast effect of the univariate MIDAS is better, and the combination effect is also better. The model novel Combination-MIDAS-ELM developed

TABLE 5. The novel combination-MIDAS-ELM models vs. the regression and machine learning combination-MIDAS when H=0,1.

Benchmark model	combination weights	h step aheads	Forecast evaluation criteria			
			RMSE	MAE	MAPE	CCE
MIDAS-Regression	Equal Weights	0	2.6632	0.3492	0.0408	1.0177
	MSFE	0	2.6467	0.3478	0.0412	1.0119
	DMSFE	0	2.6278	0.3461	0.0416	1.0052
	Average	0	2.6459	0.3477	0.0412	1.0116
MIDAS-SVR	Equal Weights	0	2.6748	0.3582	0.0499	1.0276
	MSFE	0	2.6740	0.3580	0.0498	1.0273
	DMSFE	0	2.6748	0.3582	0.0499	1.0276
	Average	0	2.6745	0.3582	0.0498	1.0275
MIDAS-BPNN	Equal Weights	0	3.0058	0.4308	0.0809	1.1725
	MSFE	0	2.7981	0.3929	0.0724	1.0878
	DMSFE	0	2.8099	0.3958	0.0745	1.0934
	Average	0	2.8713	0.4065	0.0759	1.1179
MIDAS-GRNN	Equal Weights	0	2.5352	0.3661	0.0421	0.9811
	MSFE	0	2.5331	0.3657	0.0419	0.9802
	DMSFE	0	2.5334	0.3657	0.0419	0.9803
	Average	0	2.5339	0.3658	0.0420	0.9806
MIDAS-ELM	Equal Weights	0	6.4282	0.6148	0.1171	2.3867
	MSFE	0	2.5734	0.3349	0.0473	0.9852
	DMSFE	0	2.5415	0.3341	0.0458	0.9738
	Average	0	3.8477	0.4279	0.0700	1.4486
MIDAS-ENN	Equal Weights	0	3.0678	0.4263	0.0714	1.1885
	MSFE	0	2.8094	0.3760	0.0573	1.0809
	DMSFE	0	2.7465	0.3571	0.0528	1.0521
	Average	0	2.8746	0.3865	0.0605	1.1072
Novel Combination-MIDAS-	ELM	0	**2.2204**	**0.3326**	**0.0371**	**0.8634**
MIDAS-Regression	Equal Weights	1	2.6477	0.3394	0.0421	1.0098
	MSFE	1	2.6340	0.3382	0.0420	1.0047
	DMSFE	1	2.6214	0.3368	0.0420	1.0001
	Average	1	2.6344	0.3382	0.0420	1.0049
MIDAS-SVR	Equal Weights	1	2.6834	0.3598	0.0508	1.0313
	MSFE	1	2.6791	0.3591	0.0504	1.0295
	DMSFE	1	2.6805	0.3594	0.0506	1.0301
	Average	1	2.6810	0.3594	0.0506	1.0303
MIDAS-BPNN	Equal Weights	1	3.1517	0.4359	0.0798	1.2225
	MSFE	1	2.8690	0.3853	0.0652	1.1065
	DMSFE	1	2.8102	0.3719	0.0612	1.0811
	Average	1	2.9437	0.3977	0.0687	1.1367
MIDAS-GRNN	Equal Weights	1	2.5910	0.3480	0.0473	0.9954
	MSFE	1	2.5862	0.3476	0.0472	0.9937
	DMSFE	1	2.5853	0.3475	0.0472	0.9933
	Average	1	2.5875	0.3477	0.0472	0.9941
MIDAS-ELM	Equal Weights	1	3.0388	0.4129	0.0651	1.1723
	MSFE	1	2.7054	0.3598	0.0628	1.0427
	DMSFE	1	2.6096	0.3405	0.0613	1.0038
	Average	1	2.7846	0.3711	0.0631	1.0729
MIDAS-ENN	Equal Weights	1	3.7453	0.5170	0.0757	1.4460
	MSFE	1	3.5650	0.5193	0.0764	1.3869
	DMSFE	1	3.4857	0.5184	0.0821	1.3621
	Average	1	3.5987	0.5182	0.0781	1.3983
Novel Combination-MIDAS-	ELM	1	**2.5505**	**0.3317**	**0.0375**	**0.9732**

in this paper has the most satisfactory evaluation criteria. Similar to the above comparison experiment 1, the novel Combination-MIDAS-ELM model has the best forecasting effect compared with other combination model. It is worth noting that when $h = 0, 1, 2, 3$, the average MAPE values of MIDAS-Regression, MIDAS-SVR, MIDAS-BPNN, MIDAS-GRNN, MIDAS-ELM and MIDAS-ENN are 4.07%, 4.86%, 6.51%, 4.33%, 5.76% and 6.75% respectively, while the corresponding mean of Combination-MIDAS-ELM is 3.68%. It shows that the forecasting results of these seven combination models are all excellent [60]. From the overall CCE results, the average CCE values of MIDAS-Regression, MIDAS-SVR, MIDAS-BPNN, MIDAS-GRNN, MIDAS-ELM and MIDAS-ENN are 1.0031, 1.0222, 1.1066, 0.9892, 1.1703 and 1.1754, respectively, while the corresponding mean for combination-MIDAS-ELM is 0.9020. Compared with the best combination MIDAS-GRNN, the forecasting results have improved by 8.82% on average. It can be concluded that the novel Combination-MIDAS-ELM model outperforms the comparable combination MIDAS models in forecasting accuracy.

Remark: Compared with some combination MIDAS models, the evaluation criteria for accuracy and stability of the proposed model are the most satisfactory. In the Combination-MIDAS-ELM model proposed in our work, the univariate MIDAS regression models and machine learning

TABLE 6. The novel combination-MIDAS-ELM models vs. the regression and machine learning combination-MIDAS when H=2,3.

Benchmark model	combination weights	h step aheads	Forecast evaluation criteria			
			RMSE	MAE	MAPE	CCE
MIDAS-Regression	Equal Weights	2	2.6655	0.3333	0.0407	1.0132
	MSFE	2	2.6636	0.3331	0.0410	1.0126
	DMSFE	2	2.6619	0.3330	0.0426	1.0125
	Average	2	2.6637	0.3331	0.0415	1.0128
MIDAS-SVR	Equal Weights	2	2.6380	0.3487	0.0441	1.0103
	MSFE	2	2.6377	0.3486	0.0441	1.0102
	DMSFE	2	2.6381	0.3487	0.0441	1.0103
	Average	2	2.6379	0.3487	0.0441	1.0102
MIDAS-BPNN	Equal Weights	2	2.9485	0.3941	0.0563	1.1330
	MSFE	2	2.8340	0.3821	0.0520	1.0894
	DMSFE	2	2.7714	0.3807	0.0506	1.0676
	Average	2	2.8513	0.3856	0.0530	1.0966
MIDAS-GRNN	Equal Weights	2	2.5968	0.3406	0.0399	0.9924
	MSFE	2	2.5960	0.3406	0.0399	0.9922
	DMSFE	2	2.5960	0.3406	0.0399	0.9922
	Average	2	2.5963	0.3406	0.0399	0.9923
MIDAS-ELM	Equal Weights	2	2.9784	0.3958	0.0515	1.1419
	MSFE	2	2.5934	0.3365	0.0398	0.9899
	DMSFE	2	2.5179	0.3346	0.0391	0.9638
	Average	2	2.6966	0.3556	0.0434	1.0319
MIDAS-ENN	Equal Weights	2	2.9003	0.3917	0.0587	1.1169
	MSFE	2	2.7959	0.3800	0.0543	1.0767
	DMSFE	2	2.7778	0.3778	0.0563	1.0706
	Average	2	2.8246	0.3832	0.0564	1.0881
Novel Combination-MIDAS-	ELM	2	**2.1034**	**0.2626**	**0.0373**	**0.8011**
MIDAS-Regression	Equal Weights	3	2.5557	0.3610	0.0383	0.9850
	MSFE	3	2.5489	0.3594	0.0381	0.9821
	DMSFE	3	2.5498	0.3583	0.0380	0.9820
	Average	3	2.5514	0.3596	0.0381	0.9830
MIDAS-SVR	Equal Weights	3	2.6549	0.3593	0.0502	1.0215
	MSFE	3	2.6523	0.3590	0.0499	1.0204
	DMSFE	3	2.6530	0.3593	0.0501	1.0208
	Average	3	2.6534	0.3592	0.0501	1.0209
MIDAS-BPNN	Equal Weights	3	3.1577	0.4286	0.0683	1.2182
	MSFE	3	2.6377	0.3502	0.0565	1.0148
	DMSFE	3	2.5657	0.3486	0.0641	0.9928
	Average	3	2.7870	0.3758	0.0630	1.0753
MIDAS-GRNN	Equal Weights	3	2.5856	0.3422	0.0442	0.9906
	MSFE	3	2.5835	0.3420	0.0441	0.9899
	DMSFE	3	2.5828	0.3420	0.0442	0.9897
	Average	3	2.5840	0.3420	0.0441	0.9900
MIDAS-ELM	Equal Weights	3	3.0751	0.4228	0.0602	1.1860
	MSFE	3	2.8818	0.3888	0.0511	1.1072
	DMSFE	3	2.8407	0.3776	0.0509	1.0897
	Average	3	2.9325	0.3964	0.0541	1.1277
MIDAS-ENN	Equal Weights	3	3.1102	0.4453	0.1132	1.2229
	MSFE	3	2.7504	0.3844	0.0595	1.0648
	DMSFE	3	2.6884	0.3695	0.0523	1.0367
	Average	3	2.8497	0.3997	0.0750	1.1081
Novel Combination-MIDAS-	ELM	3	**2.5446**	**0.3308**	**0.0353**	**0.9702**

models MIDAS are combined, which makes the univariate forecast results more improved than before. Unlike the previous fixed weight and variable weight combination methods, the combination MIDAS model uses novel variable weight coefficient method based on ELM to determine the combination weights, which further improves the forecast accuracy of the combination MIDAS model.

3) FORECAST COMPARISON: THE NOVEL COMBINATION-MIDAS-ELM MODELS VS. THE MULTIVARIATE MIDAS MODELS

In the third part of the comparison experiment, we compared the novel Combination-MIDAS-ELM model with some other benchmark multivariate machine learning models MIDAS (such as MMIDAS-SVR, MMIDAS-BPNN, MMIDAS-GRNN, MMIDAS-ELM, and MMIDAS-ENN) that are directly multivariate. The detailed prediction accuracy results are shown in Table 7. The most satisfactory results are also highlighted in bold.

Comparison of the experimental results shows that the novel Combination-MIDAS-ELM model consistently exhibits the most satisfactory standard of evaluation. Similar to the comparison of experiment 1 and experiment 2 above, when compared with the multivariate machine learning MIDAS model, the novel Combination-MIDAS-ELM model can achieve the best forecast performance. It is worth

TABLE 7. The novel combination-MIDAS-ELM models vs. the multivariate MIDAS forecasting models.

Multivariate MIDAS benchmark model	h step aheads	Forecast evaluation criteria			
		RMSE	MAE	MAPE	CCE
MMIDAS-SVR	0	2.6481	0.3508	0.0453	1.0147
MMIDAS-BPNN	0	7.3445	1.0414	0.2337	2.8732
MMIDAS-GRNN	0	3.0254	0.4009	0.0593	1.1618
MMIDAS-ELM	0	3.2284	0.4480	0.0978	1.2581
MMIDAS-ENN	0	5.0308	0.7352	0.1643	1.9767
Novel-Combination-MIDAS-SVR	0	**2.2204**	**0.3326**	**0.0371**	**0.8634**
MMIDAS-SVR	1	2.6746	0.3569	0.0487	1.0267
MMIDAS-BPNN	1	9.9247	1.5368	0.3120	3.9245
MMIDAS-GRNN	1	2.9309	0.3846	0.0547	1.1234
MMIDAS-ELM	1	2.7731	0.3697	0.0672	1.0700
MMIDAS-ENN	1	6.3556	1.0134	0.2727	2.5472
Novel-Combination-MIDAS- ELM	1	**2.5505**	**0.3317**	**0.0375**	**0.9732**
MMIDAS-SVR	2	2.6481	0.3508	0.0453	1.0147
MMIDAS-BPNN	2	3.9802	0.5720	0.1346	1.5623
MMIDAS-GRNN	2	2.5345	0.3140	0.0672	0.9719
MMIDAS-ELM	2	3.0378	0.4020	0.0703	1.1700
MMIDAS-ENN	2	4.7934	0.7128	0.1635	1.8899
Novel-Combination-MIDAS- ELM	2	**2.1034**	**0.2626**	**0.0373**	**0.8011**
MMIDAS-SVR	3	2.8179	0.3783	0.0640	1.0867
MMIDAS-BPNN	3	5.3683	0.8012	0.2335	2.1343
MMIDAS-GRNN	3	2.9133	0.3989	0.0649	1.1257
MMIDAS-ELM	3	2.9644	0.3898	0.0786	1.1443
MMIDAS-ENN	3	6.7636	0.9409	0.1912	2.6319
Novel-Combination-MIDAS- ELM	3	**2.5446**	**0.3308**	**0.0353**	**0.9702**

noting that when $h = 0, 1, 2, 3$, the average MAPE values of MMIDAS-SVR, MMIDAS-BPNN, MMIDAS-GRNN, MMIDAS-ELM and MMIDAS-ENN are 8.08%, 22.84%, 6.15%, 7.85% and 19.79% respectively, while the corresponding mean of Combination-MIDAS-ELM is 3.68%. It can be explained that the results of the MMIDAS-SVR, MMIDAS-GRNN and MMIDAS-ELM models are excellent, the MMIDAS-BPNN and MMIDAS-ENN models are good, and the forecast results of the Combination-MIDAS-ELM model are also excellent [60]. From the overall result CCE, The mean CCE values for MMIDAS-SVR, MMIDAS-BPNN, MMIDAS-GRNN, MMIDAS-ELM and MMIDAS-ENN are 1.0357, 2.6236, 1.0957, 1.1606 and 2.2614, respectively, while the corresponding mean for Combination-MIDAS-ELM is 0.9020. Compared with the best multivariate MMIDAS-SVR, which is an average increase of 12.91%. Therefore, it can be concluded that the proposed Combination-MIDAS-ELM model outperforms the above models directly using the multivariate machine learning MIDAS for comparison in terms of forecasting accuracy.

Remark: Among the multivariate MIDAS benchmark models, the novel Combination-MIDAS-ELM model has the most satisfactory computational and evaluation criteria compared with some currently popular machine learning MIDAS models. Due to the direct use of the multivariate machine learning model MIDAS for forecasting, the increase of multivariate influencing factors leads to an increase in the degree of nonlinearity of the training dataset, which reduces the forecast accuracy and stability, especially for the MMIDAS-BPNN and MMIDAS-ENN models.

4) DM TESTS

This study uses the Diebold-Mariano test [61], an important hypothesis testing procedure, to compare the difference in predictive power between the proposed combination model and other comparison models. The theory of the DM test is described below. First, we state hypotheses based on current problems.

$$\begin{aligned} H_0 &: E[\psi(error^1)] = E[\psi(error^2)] \\ H_1 &: E[\psi(error^1)] \neq E[\psi(error^2)] \end{aligned} \tag{30}$$

where $error^1$ and $error^2$ are the difference between the real data and the predicted data of the different models, and ψ represents the loss function of the prediction error. We also select appropriate statistics to complete the statistical inference process. The prediction section of this paper belongs to the small sample test. Based on the DM test, the analysis of Harbor, Leybourne and Newbold [62] is introduced to find

$$k = \sqrt{n + 1 - 2h + n^{-1}h(h-1)} \tag{31}$$

n indicates the range of forecast data. h indicates that the length of the ahead step.

$$DM = \frac{\left[\sum_{i=1}^{n} \left(\psi\left(error^1\right) - \psi\left(error^2\right)\right)/n\right] * s^2}{\sqrt{s^2/n}} * k \tag{32}$$

The constructed DM test statistic, where s^2 is the consistent estimate of the variance and $\sum_{i=1}^{n}\left(\psi\left(error^1\right) - \psi\left(error^2\right)\right)/n$ is the average error of the loss functions of the two models. The next comparison with the given significance level determines whether the null hypothesis is acceptable. If the significance level α is given and the statistic DM is

TABLE 8. DM-test.

Model	h=0	h=1	h=2	h=3
ARIMA	-4.4694*	-12.6272*	-5.4142*	-9.4618*
GARCH	-42.1362*	-19.6202*	-18.9655*	-40.7115*
MIDAS	-3.0785*	-7.6989*	-4.8904*	-3.4404*
MIDAS-SVR	-2.9353*	-7.1642*	-4.6956*	-5.4047*
MIDAS-BPNN	-5.2620*	-6.8311*	-10.7365*	-8.6212*
MIDAS-GRNN	-2.1225**	-2.9661*	-4.2423*	-2.1060**
MIDAS-ELM	-3.8001*	-8.9105*	-9.0797*	-5.7218*
MIDAS-ENN	-4.4588*	-8.7412*	-9.1291*	-8.2888*
MMIDAS-SVR	-2.8019*	-4.3534*	-4.6939*	-5.2428*
MMIDAS-BPNN	-8.9402*	-15.3852*	-7.2843*	-10.4884*
MMIDAS-GRNN	-4.8147*	-4.8909*	-4.0330*	-5.4541*
MMIDAS-ELM	-5.2629*	-3.3270*	-5.0640*	-3.5731*
MMIDAS-ENN	-10.6995*	-14.8107*	-12.4333*	-10.4022*

*denote 1% significance level , **denote 5% significance level and ***denote 10% significance level

greater than the upper bound $Z_{\alpha/2}$ or less than the lower bound $-Z_{\alpha/2}$, the null hypothesis is rejected. This situation means that these combined and contrasting models have significant differences in predictive performance. Conversely, the alternative hypothesis is rejected as soon as α falls within the range of $[-Z_{\alpha/2}, Z_{\alpha/2}]$.

Table 8 lists the DM test results between the novel Combination-MIDAS-ELM models and all comparison methods. We can see that the novel Combination-MIDAS-ELM models are significantly different from the other forecast models at a confidence level of 1%, 5% or 10%. In particular, for the 0-step to 3-step ahead prediction, the test passed the 5% significance level, which means that the novel Combination-MIDAS-ELM models are significantly different from the other MIDAS models. As for the comparison between the novel Combination-MIDAS-ELM models and the traditional ARIMA, GARCH and multivariate MIDAS models, most of the DM values are greater than the critical value of the 1% confidence level. However, when h=0,3, the DM values comparing the novel Combination-MIDAS-ELM models and the MIDAS-GRNN based combination models are -2.1225 and -2.1060, respectively, which means that the forecast power of the novel Combination-MIDAS-ELM models is significantly different from the combination-MIDAS-GRNN models, with a significance level of 5%. Overall, the forecast ability of the proposed novel Combination-MIDAS-ELM models is significantly different from other comparative models.

IV. CONCLUSION

This work aims to make the MIDAS forecasting models more effective. Therefore, we propose a Combination-MIDAS-ELM model to overcome the weaknesses of natural gas price forecasting. First, the univariate MIDAS regression models and machine learning models MIDAS are analyzed based on carbon prices, coal, crude oil and FTSE-100, and the best univariate MIDAS model is selected based on CCE index. Second, the best univariate MIDAS model is combined by selecting the corresponding MIDAS model among the best influencing factors to expand the forecast range to construct a novel variable weight coefficient combination MIDAS model based on ELM and the optimal initial parameters are obtained by MOGWO. Third, the effectiveness and accuracy of the Combination-MIDAS-ELM proposed is demonstrated by comparison experiments with some relatively competitive models.

In general, we can draw some conclusions from the results of the empirical analysis. (1) The effect of the univariate MIDAS model based on machine learning is better than the univariate MIDAS model based on regression, which is due to the nonlinear properties of the data itself. (2) Natural gas prices are more responsive to crude oil prices than carbon emissions, coal and the FTSE-100. (3) In building the Combination-MIDAS-ELM model, the novel type of variable combination is more robust as a reference for determining the weighting of the forecast results of the combination MIDAS model. (4) The forecast accuracy of the Combination-MIDAS-ELM model is much higher than that of some other competing MIDAS models. Therefore, the proposed model can provide significant improvements in forecasting natural gas prices and is competitive in forecasting nonlinear and irregular natural gas prices.

In further research, we intend to choose other univariate MIDAS models to fit the weighting coefficients of the combination MIDAS model. Another challenge is to develop a more comprehensive system of factors for forecasting natural gas price, because there are other indicators that affect natural gas price, such as some unexpected information or political information.

REFERENCES

[1] A. Dominguez-Ramos and A. Irabien, "The role of power-to-gas in the European union," *Green Chem. Eng.*, vol. 1, no. 1, pp. 6–8, Sep. 2020.

[2] J. Rogelj, M. den Elzen, N. Höhne, T. Fransen, H. Fekete, H. Winkler, R. Schaeffer, F. Sha, K. Riahi, and M. Meinshausen, "Paris agreement climate proposals need a boost to keep warming well below 2°C," *Nature*, vol. 534, no. 7609, pp. 631–639, Jun. 2016.

[3] A. Sikora, "European green deal–legal and financial challenges of the climate change," *ERA Forum*, vol. 21, no. 4, pp. 681–697, Jan. 2021.

[4] P. Aatola, M. Ollikainen, and A. Toppinen, "Price determination in the EU ETS market: Theory and econometric analysis with market fundamentals," *Energy Econ.*, vol. 36, pp. 380–395, Mar. 2013.

[5] B. Zhu, S. Ye, D. Han, P. Wang, K. He, Y.-M. Wei, and R. Xie, "A multiscale analysis for carbon price drivers," *Energy Econ.*, vol. 78, pp. 202–216, Feb. 2019.

[6] B. Soldo, "Forecasting natural gas consumption," *Appl. Energy*, vol. 92, pp. 26–37, Apr. 2012.

[7] E. Erdogdu, "Natural gas demand in Turkey," *Appl. Energy*, vol. 87, no. 1, pp. 211–219, Jan. 2010.

[8] Ö. Dilaver, Z. Dilaver, and L. C. Hunt, "What drives natural gas consumption in Europe? Analysis and projections," *J. Natural Gas Sci. Eng.*, vol. 19, pp. 125–136, Jul. 2014.

[9] H. Khani and H. E. Z. Farag, "An online-calibrated time series based model for day-ahead natural gas demand forecasting," *IEEE Trans. Ind. Informat.*, vol. 15, no. 4, pp. 2112–2123, Apr. 2018.

[10] W. Zhang and J. Yang, "Forecasting natural gas consumption in China by Bayesian model averaging," *Energy Rep.*, vol. 1, pp. 216–220, Nov. 2015.

[11] Y. Chen, W. S. Chua, and T. Koch, "Forecasting day-ahead high-resolution natural-gas demand and supply in Germany," *Appl. Energy*, vol. 228, pp. 1091–1110, Oct. 2018.

[12] D. Sen, M. E. Günay, and K. M. M. Tunç, "Forecasting annual natural gas consumption using socio-economic indicators for making future policies," *Energy*, vol. 173, pp. 1106–1118, Apr. 2019.

[13] J. Szoplik, "Forecasting of natural gas consumption with artificial neural networks," *Energy*, vol. 85, pp. 208–220, Jun. 2015.
[14] Ö. A. Dombaycı, "The prediction of heating energy consumption in a model house by using artificial neural networks in Denizli–Turkey," *Adv. Eng. Softw.*, vol. 41, no. 2, pp. 141–147, Feb. 2010.
[15] T. Afshin, B.-H. Ali, M. Hossein, N. Saeid, B. Meysam, L. Moonyong, B. Alireza, and N.-M. Adel, "Prediction of water formation temperature in natural gas dehydrators using radial basis function (RBF) neural networks," *Natural Gas Ind. B*, vol. 3, no. 2, pp. 173–180, Mar. 2016.
[16] W. Qiao, Z. Yang, Z. Kang, and Z. Pan, "Short-term natural gas consumption prediction based on Volterra adaptive filter and improved whale optimization algorithm," *Eng. Appl. Artif. Intell.*, vol. 87, Jan. 2020, Art. no. 103323.
[17] Y. Bai, Z. Sun, B. Zeng, J. Long, L. Li, J. V. de Oliveira, and C. Li, "A comparison of dimension reduction techniques for support vector machine modeling of multi-parameter manufacturing quality prediction," *J. Intell. Manuf.*, vol. 30, no. 5, pp. 2245–2256, Jun. 2019.
[18] O. F. Beyca, B. C. Ervural, E. Tatoglu, P. G. Ozuyar, and S. Zaim, "Using machine learning tools for forecasting natural gas consumption in the province of Istanbul," *Energy Econ.*, vol. 80, pp. 937–949, May 2019.
[19] L. Han, D. Liu, G. Zheng, Y. Liang, and Y. Ni, "Research on natural gas load forecasting based on support vector regression," *Chin. J. Chem. Eng.*, vol. 12, pp. 732–736, Jun. 2004.
[20] L. Zhu, M. S. Li, Q. H. Wu, and L. Jiang, "Short-term natural gas demand prediction based on support vector regression with false neighbours filtered," *Energy*, vol. 80, pp. 428–436, Feb. 2015.
[21] M. A. Destek and S. A. Sarkodie, "Are fluctuations in coal, oil and natural gas consumption permanent or transitory? Evidence from OECD countries," *Heliyon*, vol. 6, no. 2, Feb. 2020, Art. no. e03391.
[22] Q. Yang, Q. Yang, S. Xu, S. Zhu, and D. Zhang, "Technoeconomic and environmental analysis of ethylene glycol production from coal and natural gas compared with oil-based production," *J. Cleaner Prod.*, vol. 273, Nov. 2020, Art. no. 123120.
[23] P. Brehm, "Natural gas prices, electric generation investment, and greenhouse gas emissions," *Resource Energy Econ.*, vol. 58, Nov. 2019, Art. no. 101106.
[24] M. E. Bildirici and T. Bakirtas, "The relationship among oil, natural gas and coal consumption and economic growth in BRICTS (Brazil, Russian, India, China, Turkey and South Africa) countries," *Energy*, vol. 65, pp. 134–144, Feb. 2014.
[25] D. Wu, Y. Geng, and H. Pan, "Whether natural gas consumption bring double dividends of economic growth and carbon dioxide emissions reduction in China?" *Renew. Sustain. Energy Rev.*, vol. 137, Mar. 2021, Art. no. 110635.
[26] C. Magazzino, M. Mele, and N. Schneider, "A D2C algorithm on the natural gas consumption and economic growth: Challenges faced by Germany and Japan," *Energy*, vol. 219, Mar. 2021, Art. no. 119586.
[27] Y. Wei, Q. Yu, J. Liu, and Y. Cao, "Hot money and China's stock market volatility: Further evidence using the GARCH–MIDAS model," *Phys. A, Stat. Mech. Appl.*, vol. 492, pp. 923–930, Feb. 2018.
[28] Z. Pan, Q. Wang, Y. Wang, and L. Yang, "Forecasting U.S. Real GDP using oil prices: A time-varying parameter MIDAS model," *Energy Econ.*, vol. 72, pp. 177–187, May 2018.
[29] X. Wu and H. Xie, "A realized EGARCH-MIDAS model with higher moments," *Finance Res. Lett.*, vol. 38, Jan. 2021, Art. no. 101392.
[30] A.-F. Allard, L. Iania, and K. Smedts, "Stock-bond return correlations: Moving away from 'one-frequency-fits-all' by extending the DCC-MIDAS approach," *Int. Rev. Financial Anal.*, vol. 71, Oct. 2020, Art. no. 101557.
[31] Z. Zhou, Z. Fu, Y. Jiang, X. Zeng, and L. Lin, "Can economic policy uncertainty predict exchange rate volatility? New evidence from the GARCH-MIDAS model," *Finance Res. Lett.*, vol. 34, May 2020, Art. no. 101258.
[32] E. Ghysels and R. Valkanov, "The MIDAS touch: Mixed data sampling regression models," *Cirano Work. Papers*, vol. 5, no. 1, pp. 512–517, 2004.
[33] E. Ghysels, P. Santa-Clara, and R. Valkanov, "Predicting volatility: Getting the most out of return data sampled at different frequencies," *J. Econometrics*, vol. 131, nos. 1–2, pp. 59–95, Mar. 2006.
[34] S. F. Razmi, B. R. Bajgiran, S. M. J. Razmi, and K. B. Oroumieh, "The effects of external uncertainties against monetary policy uncertainty on Iranian stock return volatility using GARCH-MIDAS approach," *Int. J. Energy Econ. Policy*, vol. 10, no. 4, pp. 278–281, May 2020.
[35] R. Demirer *et al.*, "Financial vulnerability and volatility in emerging stock markets: Evidence from GARCH-MIDAS models," Univ. Pretoria, South Africa, Work. Papers 202112, Dec. 2021.
[36] A. A. Salisu, R. Gupta, and R. Demirer, "Global financial cycle and the predictability of oil market volatility: Evidence from a GARCH-MIDAS model," *Energy Econ.*, vol. 108, Apr. 2022, Art. no. 105934.
[37] X. Yu and Y. Huang, "The impact of economic policy uncertainty on stock volatility: Evidence from GARCH–MIDAS approach," *Phys. A, Stat. Mech. Appl.*, vol. 570, May 2021, Art. no. 125794.
[38] C. Jiang, Y. Li, Q. Xu, and Y. Liu, "Measuring risk spillovers from multiple developed stock markets to China: A vine-copula-GARCH-MIDAS model," *Int. Rev. Econ. Finance*, vol. 75, pp. 386–398, Sep. 2021.
[39] X. Wu and H. Xie, "A realized EGARCH-MIDAS model with higher moments," *Finance Res. Lett.*, vol. 38, Jan. 2021, Art. no. 101392.
[40] Q. Xu, X. Zhuo, C. Jiang, and Y. Liu, "An artificial neural network for mixed frequency data," *Expert Syst. Appl.*, vol. 118, pp. 127–139, Mar. 2019.
[41] Q. Xu, L. Wang, C. Jiang, and Y. Liu, "A novel (U)MIDAS-SVR model with multi-source market sentiment for forecasting stock returns," *Neural Comput. Appl.*, vol. 32, no. 10, pp. 5875–5888, May 2020.
[42] Y. Pan, Z. Xiao, X. Wang, and D. Yang, "A multiple support vector machine approach to stock index forecasting with mixed frequency sampling," *Knowl.-Based Syst.*, vol. 122, pp. 90–102, Apr. 2017.
[43] L. Lue, "Research on forecasting model based on missing data with mixed frequency extreme learning machine," Ph.D. dissertation, Dept. Econ. Bus. Admin., Chongqing Univ., Chongqing, China, 2019.
[44] C. H. Aladag, E. Egrioglu, and U. Yolcu, "Forecast combination by using artificial neural networks," *Neural Process. Lett.*, vol. 32, no. 3, pp. 269–276, Dec. 2010.
[45] N. Tak, E. Egrioglu, E. Bas, and U. Yolcu, "An adaptive forecast combination approach based on meta intuitionistic fuzzy functions," *J. Intell. Fuzzy Syst.*, vol. 40, no. 5, pp. 9567–9581, Jan. 2021.
[46] E. Ghysels and N. Ozkan, "Real-time forecasting of the U.S. Federal government budget: A simple mixed frequency data regression approach," *Int. J. Forecasting*, vol. 31, no. 4, pp. 1009–1020, Oct. 2015.
[47] L. Ding, Z. Lv, M. Han, X. Zhao, and W. Wang, "Forecasting China's wastewater discharge using dynamic factors and mixed-frequency data," *Environ. Pollut.*, vol. 255, Dec. 2019, Art. no. 113148.
[48] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, nos. 1–3, pp. 489–501, May 2006.
[49] G.-B. Huang, X. Ding, and H. Zhou, "Optimization method based extreme learning machine for classification," *Neurocomputing*, vol. 74, nos. 1–3, pp. 155–163, Dec. 2010.
[50] A. E. Albert, *Regression and the Moore–Penrose Pseudo-Inverse*. New York, NY, USA: Academic, 1972.
[51] S. Mirjalili, S. Saremi, S. M. Mirjalili, and L. D. S. Coelho, "Multi-objective grey wolf optimizer: A novel algorithm for multi-criterion optimization," *Expert Syst. Appl.*, vol. 47, pp. 106–119, Apr. 2016.
[52] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, Mar. 2014.
[53] Q. Xu, X. Zhuo, C. Jiang, X. Liu, and Y. Liu, "Group penalized unrestricted mixed data sampling model with application to forecasting U.S. GDP growth," *Econ. Model.*, vol. 75, pp. 221–236, Nov. 2018.
[54] E. Ghysels, V. Kvedaras, and V. Zemlys-Balevicius, "Mixed data sampling (MIDAS) regression models," in *Handbook of Statistics*, vol. 42. Amsterdam, The Netherlands: Elsevier, 2020, ch. 4, pp. 117–153.
[55] E. Ghysels, A. Sinko, and R. Valkanov, "MIDAS regressions: Further results and new directions," *Econ. Rev.*, vol. 26, no. 1, pp. 53–90, Feb. 2007.
[56] M. Han, L. Ding, X. Zhao, and W. Kang, "Forecasting carbon prices in the Shenzhen market, China: The role of mixed-frequency factors," *Energy*, vol. 171, pp. 69–76, Mar. 2019.
[57] N. B. Karayiannis and G. W. Mi, "Growing radial basis neural networks: Merging supervised and unsupervised learning with network growth techniques," *IEEE Trans. Neural Netw.*, vol. 8, no. 6, pp. 1492–1506, Nov. 1997.
[58] Z. Xin, M. Han, L. Ding, and A. C. Calin, "Forecasting carbon dioxide emissions based on a hybrid of mixed data sampling regression model and back propagation neural network in the USA," *Environ. Sci. Pollut. Res.*, vol. 25, no. 3, pp. 1–12, 2017.
[59] Y. He and B. Lin, "Forecasting China's total energy demand and its structure using ADL-MIDAS model," *Energy*, vol. 151, pp. 420–429, May 2018.
[60] L.-C. Hsu and C.-H. Wang, "Forecasting integrated circuit output using multivariate grey model and grey relational analysis," *Expert Syst. Appl.*, vol. 36, no. 2, pp. 1403–1409, 2009.

[61] X. Zhao, M. Han, L. Ding, and W. Kang, ''Usefulness of economic and energy data at different frequencies for carbon price forecasting in the EU ETS,'' *Appl. Energy*, vol. 216, pp. 132–141, Apr. 2018.
[62] F. X. Diebold and R. S. Mariano, ''Comparing predictive accuracy,'' *J. Bus. Econ. Statist.*, vol. 20, no. 1, pp. 134–144, Jan. 2002.
[63] D. Harvey, S. Leybourne, and P. Newbold, ''Testing the equality of prediction mean squared errors,'' *Int. J. Forecasting*, vol. 13, no. 2, pp. 281–291, Jun. 1997.

LUE LI received the Ph.D. degree in management science and engineering from Chongqing University, Chongqing, China, in 2019. He is currently a Lecturer with the Guangxi Key Laboratory of Cross-Border E-Commerce Intelligent Information Processing, School of Information and Statistics, Guangxi University of Finance and Economics. His research interests include artificial intelligence and their application in prediction of clean energies.

CAIHONG HAN received the master's degree in basic mathematics from Guangxi University, in 2009. She is currently a Lecturer with the School of Information and Statistics, Guangxi University of Finance and Economics. Her research interests include difference equations, artificial intelligence, and machine learning.

SHENGWEI YAO received the M.S. degree in applied mathematics from Guangxi University, Nanning, China, in 2007, and the Ph.D. degree in science from the East China University of Science and Technology, in 2014. He is currently a Professor with the Guangxi University of Finance and Economics. His research interests include optimization and nonlinear analysis.

LIANGSHUO NING received the master's degree in applied mathematics from Hubei University, China, in 2010. He is currently a Lecturer with the School of Information and Statistics, Guangxi University of Finance and Economics. His research interests include pattern recognition, machine learning, and data mining.