

Article

Non-Ferrous Metal Price Point and Interval Prediction Based on Variational Mode Decomposition and Optimized LSTM Network

Yifei Zhao ^{1,2} , Jianhong Chen ^{1,*}, Hideki Shimada ²  and Takashi Sasaoka ²

¹ School of Resource and Safety Engineering, Central South University, Changsha 410083, China; zhaoyifei@csu.edu.cn

² Faculty of Engineering, Kyushu University, Fukuoka 819-0395, Japan; shimada@mine.kyushu-u.ac.jp (H.S.); sasaoka@mine.kyushu-u.ac.jp (T.S.)

* Correspondence: jhchen@csu.edu.cn

Abstract: The accurate forecasting of metal prices is of great importance to industrial producers as the supply of metal raw materials is a very important part of industrial production. The futures market is subject to many factors, and metal prices are highly volatile. In the past, most of the relevant research has focused only on deterministic point forecasting, with less research performed on interval uncertainty forecasting. Therefore, this paper proposes a novel forecasting model that combines point forecasting and interval forecasting. First, a novel hybrid price point forecasting model was established using Variational Modal Decomposition (VMD) and a Long Short-Term Memory Neural Network (LSTM) based on Sparrow Search Algorithm (SSA) optimization. Then, five distribution functions based on the optimization algorithm were used to fit the time series data patterns and analyze the metal price characteristics. Finally, based on the optimal distribution function and point forecasting results, the forecasting range and confidence level were set to determine the interval forecasting model. The interval forecasting model was validated by inputting the price data of copper and aluminum into the model and obtaining the interval forecasting results. The validation results show that the proposed hybrid forecasting model not only outperforms other comparative models in terms of forecasting accuracy, but also has a better performance in forecasting sharp fluctuations and data peaks, which can provide a more valuable reference for producers and investors.

Keywords: metal price forecasting; VMD; SSA; LSTM; point forecasting; interval forecasting

MSC: 91B84



Citation: Zhao, Y.; Chen, J.; Shimada, H.; Sasaoka, T. Non-Ferrous Metal Price Point and Interval Prediction Based on Variational Mode Decomposition and Optimized LSTM Network. *Mathematics* **2023**, *11*, 2738. <https://doi.org/10.3390/math11122738>

Academic Editor: Amir Mosavi

Received: 18 May 2023

Revised: 10 June 2023

Accepted: 12 June 2023

Published: 16 June 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Metal minerals are an important source of economic income for many metal-exporting countries, such as Chile and Zambia, and thus, fluctuations in metal prices have a significant impact on trade in these countries [1–3]. At the same time, price fluctuations also affect business management, raw material supply and investment risk, as metal raw materials are closely related to many industries and involved in industrial production [4]. This is why price forecasting has become an important part of the solution to price fluctuations. Forecasting is defined as the art and science of predicting future events [5]. Typically, the process of forecasting starts with collecting historical information and extrapolating it into the future using various mathematical models. Accurately predicting the future is essential for effective planning in all areas of business, making forecasting one of the most important tools managers can use to manage or organize.

In strategic mine planning, the price of mineral products is the most important and effective parameter in evaluating engineering projects, such as mines. The feasibility and economics of mining projects are typically assessed using the Net Present Value

(NPV) method, which calculates the present value of the economic value of a mining block [6]. Price forecasts are important in assessing the potential for economic extraction of reserves [7]. In addition, by forecasting changes in ore prices, mining project decision-makers can play a key role in making accurate project decisions, blocking development or limiting mining activities [8]. It is worth noting that the establishment of reliable and stable metal price forecasting models, even with small improvements in forecasting accuracy, can bring significant benefits to metal producers [9,10]. However, the uncertain nature of time series data is a significant barrier to forecasting accuracy. At the same time, metal price data have a high degree of non-linear complexity and are influenced by supply and demand and financial markets, and effective forecasting methods and high forecasting accuracy remain a challenge [1,9,11].

To address this issue, researchers have proposed a variety of forecasting methods to improve the accuracy of predictions. In terms of using statistical methods, a number of econometric models have been developed to forecast metal prices, such as the Autoregressive Integrated Moving Average (ARIMA) [12,13], Wavelet Analysis [14] and Generalized Autoregressive Conditional Heterogeneity (GARCH) models [13]. Kriechbaumer et al. proposed an improved model based on wavelet analysis and ARIMA to predict copper prices. The results showed that the performance of the ARIMA model in predicting metal prices was significantly improved compared to other methods [14]. However, the inherent limitations of these time series models in capturing the linear behavior of the time series and the specific non-linear patterns prevent them from further improving time series forecasting [11,15]. As technology advances, Artificial Intelligence (AI) models have been widely studied and used to identify and capture various features in the analysis of metal price time series [8,12,16–19]. In terms of using metaheuristic algorithms, Eguel et al. used simulated annealing (SA) and genetic algorithms (GAs), selected a 10-month price data set as a training set to predict copper prices, and proposed genetic algorithms as the best model for predicting copper prices [20]. Regarding the use of hybrid intelligence algorithms, Alameer et al. developed a novel hybrid artificial intelligence model (GA-ANFIS) based on a genetic algorithm and an adaptive neuro-fuzzy inference system (ANFIS) model for predicting copper prices. Several different models were also used for a comparison with the new model, such as GARCH, ARIMA, ANFIS (without optimization) and Support Vector Machine (SVM) models. In the end, they concluded that the GA-ANFIS model was better at predicting copper prices than the other models [21]. Hu et al. developed a novel intelligent model based on the ANN-LSTM-GARCH method for predicting copper price fluctuations using machine learning and deep learning methods. By optimizing an artificial neural network (ANN), the prediction accuracy was greatly improved [22]. However, these models also have their own drawbacks; for example, the widely used ANN can cause the algorithm to lack global search capability and easily fall into local optimization [23].

Although there is a large body of literature on metal price forecasting as described above, the uncertainty of metal price markets is not taken into account in forecasting, despite the prevalence of this reality. At the same time, in the face of highly volatile metal prices, deterministic forecasting can be difficult to achieve accurate predictions, which can increase the likelihood of mispricing and lead to irrational decisions or investments. In addition, most price forecasting focuses on single or multi-point forecasts at a deterministic time, while interval forecasting of metal prices has received little attention. The variability of forecasts due to uncertainty can be greatly reduced through interval quantification. At the same time, within the upper and lower limits of the significance level considered, interval forecasts can provide more reliable and stable references for mine managers and decision-makers. This study proposes a novel hybrid system consisting of point forecasting and interval forecasting models. The metal price point forecasts are used to support the price interval forecasts. For the point forecasting, an innovative and stable hybrid forecasting method was developed using Variational Modal Decomposition (VMD) and a Sparrow Search Algorithm Optimization (SSA)-based Long Short-Term Memory (LSTM) neural network to build the novel hybrid price point forecasting model. The VMD is used to

decompose the metal price time series data, and then, each decomposition unit is predicted using the SSA-LSTM model, and the prediction results of each decomposition unit are summed to obtain the point prediction results. Finally, based on the best distribution function and the point forecasting results, a range and confidence level are set and an interval forecasting model is established, with copper and aluminum price data fed into the model to obtain interval forecasting results. Three comparative models (LSTM, VMD-LSTM and SSA-LSTM) are also introduced for comparative experiments.

Further material is divided into several parts. Thus, Section 2 presents the modelling methodology; Section 3 describes the hybrid prediction model structure and the model prediction process; Section 4 presents the model data, parameter settings and evaluation metrics; Section 5 presents the analysis and discussion of the experimental prediction results; and finally, Section 6 presents the overall conclusions and recommendations.

2. Methodology

This paper introduces the VMD, SSA, LSTM and the Uncertainty Prediction Method.

2.1. Variational Mode Decomposition

In 1998, Huang et al. invented Empirical Mode Decomposition (EMD) to solve the characteristics of traditional decomposition methods (including wavelet transform and Fourier transform), which cannot satisfy the uncertainty in the face of nonlinear and non-smooth features, but this decomposition method has a serious modal confounding phenomenon. In order to solve this problem, Konstantin Dragomiretskiy et al. proposed its improved version of variational modal decomposition VMD in 2014. Unlike the EMD principle, the VMD decomposition method uses an iterative search for the optimal solution of the variational model to determine the center frequency and bandwidth of each decomposition component, which is a completely non-recursive model, with which the model searches for the modal components of the set of modal components and their respective center frequencies, and each mode is smoothed after demodulation into baseband, as demonstrated by experimental results: the method is more robust with respect to sampling and noise. This means that the non-periodic signal is analyzed in the frequency domain and the complex signal is decomposed into multiple harmonics [24]. It is currently used in many fields [25,26].

The goal of VMD is to decompose real-valued input signals $f(t)$ into discrete sub-signals (modes) u_k , assuming that each mode u_k is mostly compact in the vicinity of the frequency center ω_k . Use the VMD method to decompose $f(t)$ into k sub-sequences. The procedure is as follows.

Step 1: For each mode u_k , the associated resolved signal is calculated using the Hilbert transform and the spectrum is constructed.

Step 2: The spectrum of the modes is shifted to the baseband by the respective estimated center frequencies.

Step 3: The bandwidth is estimated from the Gaussian smoothness of the demodulated signal, i.e., the gradient L^2 of the signal to estimate the bandwidth. The resulting constrained variational problem is as follows:

$$\begin{aligned} \min_{\{u_k\}, \{\omega_k\}} & \left\{ \sum_k \|\partial_t[(\delta(t) + j/\pi t) * u_k(t)]e^{-j\omega_k t}\|_2^2 \right\} \\ \text{s.t.} & \sum_{k=1}^K u_k = f \end{aligned} \tag{1}$$

where, u_k is the first k modal component, ω_k is the frequency center of the k , which is the frequency center corresponding to the first mode and $\delta(t)$ is the unit pulse function.

Step 4: Introduce the quadratic penalty term and the Lagrange multiplier λ . The Lagrange expression is obtained by transforming the constrained variational problem into

an unconstrained variational problem, with the quadratic penalty term α . The effect is to reduce the interference of Gaussian noise.

$$L(\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_k \|\partial_t[(\delta(t) + j/\pi t) * u_k(t)]e^{-j\omega_k t}\|_2^2 + \|f(t) - \sum_k u_k(t)\|_2^2 + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle \tag{2}$$

Step 5: Finally, the solution is carried out using the alternating direction multiplier method (ADMM).

$$\begin{cases} \hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \hat{\lambda}(\omega)/2}{1 + 2\alpha(\omega - \omega_k)^2} \\ \omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k^{n+1}(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k^{n+1}(\omega)|^2 d\omega} \\ \hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + \gamma \left(\hat{f}(\omega) - \sum_k \hat{u}_k^{n+1}(\omega) \right) \end{cases} \tag{3}$$

2.2. Sparrow Search Algorithm

The Sparrow Search Algorithm SSA is a novel swarm intelligence optimization algorithm inspired by the foraging and anti-predation behavior of sparrows, proposed by Xue et al. in 2020 [27]. It has the characteristics of fast convergence and strong search ability, and thus, this paper uses the sparrow optimization algorithm to search for the optimization of LSTM parameters and improve its prediction accuracy.

In SSA, individuals are divided into three categories: discoverers, followers and vigilantes, and the position of each individual corresponds to a solution. The algorithm obtains the position of the optimal solution by continuously updating the positions of these three categories of individuals and calculating the fitness value of all individuals at each cycle, with the main update iteration steps shown below.

Step 1: Initialize the population, the proportion of predators and joiners and the number of iterations.

Step 2: Calculate the fitness values and sort them from largest to smallest.

Step 3: Update the finder position (1).

Step 4: Update the follower position (2).

Step 5: Update the vigilante position (3) (Aware of Danger Sparrow).

Step 6: Calculate the fitness value and update the sparrow position.

Step 7: If the requirements are met, output the result; otherwise, repeat steps 2–6.

(1) Finder position update.

The finder checks for predators in the foraging area and if not, searches extensively for food; if there are predators, it flies to a safe area. The expression is shown in Equation (10) below.

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t \cdot \exp\left(-\frac{i}{\alpha \cdot i_{max}}\right), R_2 < S_T \\ X_{i,j}^t + Q \cdot L, R_2 \geq S_T \end{cases} \tag{4}$$

where t is the number of current iterations; $X_{i,j}^t$ is the position of the i th sparrow in the j th dimension in generation t , i_{max} is the maximum number of iterations, which is a random number of $(0, 1]$, Q is a random number obeying a normal distribution and L is an all -1 matrix, R_2 is the alarm value and S_T is the safety threshold.

(2) Follower position update

When a follower joins, it is determined whether it is eligible to compete with the finder for food i.e., whether its location is better, if its location corresponds to a lower fitness level then it is not eligible to compete and it needs to fly to another area to forage; otherwise, the

joiner will forage in the vicinity of the best individual X_p ; otherwise, the joiner will forage near the best individual. The expressions are shown below.

$$X_{i,j}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{X_{worst}^t - X_{i,j}^t}{i^2}\right), i < \frac{n}{2} \\ X_p^{t+1} + |X_{i,j}^t - X_p^{t+1}| \cdot A^+ \cdot L, i \geq \frac{n}{2} \end{cases} \tag{5}$$

where X_{worst}^t denotes the position of the worst adapted individual in generation t , X_p^{t+1} denotes the position of the best adapted individual in generation $t + 1$ and A is a matrix of the same dimension as L with elements that are subsequently pre-defined as 1 or -1 and satisfy $A^+ = A^T (A A^T)^{-1}$.

(3) Vigilante position update

When individuals are at the periphery of the population, they need to adopt anti-predatory behavior to achieve a higher degree of adaptation; when they are at the center of the population, they need to move closer to their peers to stay away from the danger zone. The expressions are as follows.

$$X_{i,j}^{t+1} = \begin{cases} X_{best}^t + \beta \times |X_{i,j}^t - X_{best}^t|, f_i > f_g \\ X_{i,j}^t + K \times \frac{|X_{i,j}^t - X_{worst}^t|}{(f_i - f_w) + \epsilon}, f_i = f_g \end{cases} \tag{6}$$

where X_{best}^x is the global optimal position in the t generation, β is the control step size that follows a $(0, 1)$ normal distribution, K is a constant of $[-1, 1]$, ϵ is a constant that avoids the denominator being zero and f_g and f_w are the current best and worst fitness values, respectively.

2.3. Long Short-Term Memory Neural Network

Hidden variable models have long suffered from long-term information preservation and short-term input deficits. LSTM is a modified Recurrent Neural Network (RNN) proposed by Hochreiter et al. in 1997, which consists of an input layer, an implicit layer, a recurrent layer and an output layer [28]. Shown in Figure 1 is the Long Short-Term Memory Neural Network.

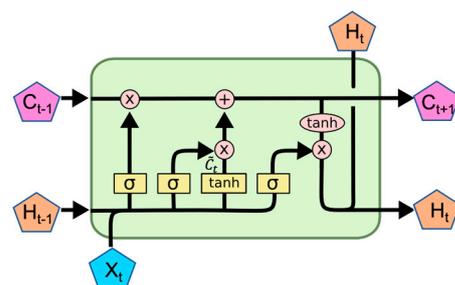


Figure 1. Long Short-Term Memory Neural Network.

In order to solve the long dependency problem of RNN based on historical data, the problem of long-term time series processing was enhanced by adding memory unit states in the implicit layer and also improving the problems of gradient disappearance and gradient explosion [29,30].

The control units are created in the implicit layer as forgetting gates f_t , input gates i_t and output gates o_t . The forgetting gate is to selectively forget the information in the cell; the input gate is to selectively record new information into the cell state. The forgetting gate selectively forgets the information in the cell; the input gate serves to selectively record new information into the cell state, and the input gate controls how much of the new data from

the candidate memory cell is used \tilde{C}_t . The output gate is to take the stored information to the next neuron. The expressions for the calculation of each variable are shown below:

$$\begin{cases} f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \\ i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \\ \tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \\ o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \\ C_t = f_t C_{t-1} + \tilde{C}_t \\ h_t = o_t \tanh(C_t) \end{cases} \quad (7)$$

where σ is the sigmoid activation function; W_f, W_i, W_c and W_o are the weight matrices of the respective gates; b_f, b_i, b_c and b_o are the bias parameters of the corresponding gates; and x_t, h_t and C_t are the cell inputs, outputs, and states at the current.

2.4. Interval Prediction Theory

In recent years, many researchers have started interval predictions of time series data based on point prediction results and the most appropriate distribution function to capture the trend characteristics of the time series objects they are studying, such as airborne particulate matter, wind speed, etc. [31–35].

This forecasting method first requires fitting the time series data with a distribution function, mining its relevant features and finding the most appropriate distribution function. This section introduces five metal price distribution models, Weibull, Logistic, Lognormal, Normal and Gamma five distribution functions DFs to represent metal price states. The associated probability density functions and cumulative distribution functions are shown in Table 1.

Table 1. Probability density functions (PDFs) and cumulative distribution functions (CDFs).

Distribution Functions	PDF/CDF	Parameters
Logistic	$f(x; \mu; \sigma) = \frac{\exp[-(x-\mu)/\sigma]}{\sigma(1+\exp[-\frac{x-\mu}{\sigma}])^2}$	$\infty > \mu > -\infty$ location parameter
	$F(x; \mu; \sigma) = \frac{1}{1+\exp[-\frac{x-\mu}{\sigma}]}$	$\sigma > 0$ scale parameter
Weibull	$f(x; \lambda; k) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} \exp\left[-\left(\frac{x}{\lambda}\right)^k\right], x > 0$	$k > 0$ shape parameter
	$F(x; \lambda; k) = 1 - \exp\left[-\left(\frac{x}{\lambda}\right)^k\right]$	$\lambda > 0$ shape parameter
Normal	$f(x; \mu; \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$	$\mu > 0$ location parameter
	$F(x; \mu; \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_0^x \exp\left[-\frac{(t-\mu)^2}{2\sigma^2}\right] dt$	$\sigma > 0$ scale parameter
Gamma	$f(x; k; \theta) = x^{k-1} \exp(-\frac{x}{\theta}) / [\theta^k \Gamma(k)], x > 0$	$k > 0$ shape parameter
	$F(x; k; \theta) = \frac{1}{\theta^k \Gamma(k)} \int_0^x t^{k-1} \exp[-\frac{t}{\theta}] dt, x > 0$	$\theta > 0$ scale parameter
Lognormal	$f(x; \mu; \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right], x \geq 0$	$\mu > 0$ location parameter
	$F(x; \mu; \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_0^x \frac{1}{t} \exp\left[-\frac{(\ln t - \mu)^2}{2\sigma^2}\right] dt$	$\sigma > 0$ scale parameter

Point forecast values and distribution functions are then used to provide uncertainty information about future values. A dynamic interval forecasting method is proposed to give uncertainty information about future prices by updating the expectation at the next point using the forecast value. For example, the forecast value of the next point is \hat{y} , and \hat{y} almost reaches the maximum of the historical data. If the forecasts are reliable and accurate, then the lower and upper limits of the interval at the next point are very large values. At the α significance level, the relationship between the upper and lower limits and the real value can be expressed by Equation (8):

$$P(I_{min} \leq Y_t \leq I_{max}) = 1 - 2\alpha \tag{8}$$

In this paper, price values are random variables for which estimates are considered as expectations of future points. Equation (8) can be written as Equation (9):

$$P\{I_{min} \leq Y_t \leq I_{max} \mid E(Y_t) = \hat{y}\} \cdot P\{E(Y_t) = \hat{y}\} = 1 - 2\alpha \tag{9}$$

In addition, we assume that the form of the predicted values is similar to the historical distribution f . Therefore, once the distribution function of the original time series has been determined, the historical variance S^2 can be used as the variance of the unknown quantity points. The conditional probability is then equal to Equation (10); thus, it is possible to calculate, at a certain confidence level, α . The values of the upper and lower bounds are calculated at a certain confidence level.

$$\left\{ (I_{min}, I_{max}) \mid I_{min} \leq Y_t \leq I_{max}, \int_{I_{min}}^{I_{max}} f(x|\hat{\Theta})dx = 1 - 2\alpha \right\} \tag{10}$$

Equation (10) can also be written as Equation (11):

$$\begin{cases} [\hat{I}_{min}, \hat{I}_{max}] = [\hat{I}_{min}, \hat{y}] [\hat{y}, \hat{I}_{max}] \\ \int_{\hat{I}_{min}}^{\hat{y}} f(x|\hat{\Theta})dx = F(\hat{y}) - \alpha \\ \int_{\hat{y}}^{\hat{I}_{max}} f(x|\hat{\Theta})dx = 1 - F(\hat{y}) - \alpha \end{cases} \tag{11}$$

3. Hybrid Forecasting Model

This section details the structure of the hybrid forecasting model developed and the model-forecasting process.

3.1. Model Structure

There are five main steps in the prediction model (Figure 2), described in detail below.

- (a) VMD decomposition: the experimental metal price dataset is decomposed into different modes through variational modal decomposition (VMD), as described in Section 2.1.
- (b) Optimization of prediction model parameters: the Sparrow search algorithm (SSA) was used to design an optimized long- and short-term memory neural network (LSTM), i.e., hybrid SSA-LSTM, and then, the established SSA-LSTM model was used to predict each mode.
- (c) Point prediction: the predictions of these models are summed according to the inverse decomposition model to obtain the final point prediction.
- (d) Distribution function (DF) estimation: five distribution functions, Weibull, Logistic, Lognormal, Normal and Gamma, are used to model the metal price time series. The ability to fit is strongly related to the estimated parameters of the DFs. Two different estimation methods, maximum likelihood estimation (MLE) and the optimization algorithm SSA, were used to determine the DF parameters. Finally, the distribution function with the best fitting ability (R^2) was used as the basis for the metal price characteristics in subsequent interval forecasts.

(e) Interval forecasting: interval metal price forecasting can be achieved based on point forecasts, the optimal distribution function for higher R^2 and the interval forecasting theory given in Section 2.4.

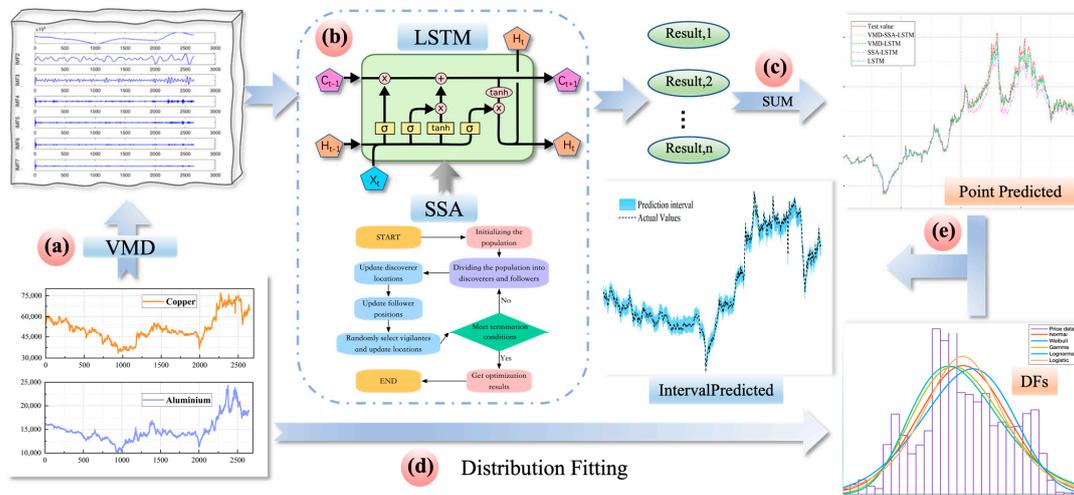


Figure 2. Metal price hybrid forecasting architecture.

3.2. Figures, Tables and Schemes

The VMD, SSA and LSTM are coupled to form a VMD–SSA–LSTM metal price forecasting model with the following forecasting steps.

- Step 1:** Metal price information for the previous n trading days is selected as input to the model.
- Step 2:** The original price time series is decomposed using the Variational Modal Decomposition (VMD) method to obtain k subcomponents.
- Step 3:** Set the SSA parameters of the sparrow search algorithm, set the population size N the maximum number of iterations T , use the SSA to find the optimal parameters of the LSTM and determine the optimal initial learning rate of the LSTM η and the number of neurons in the hidden layer H . Finally, a coupled model of the SSA-LSTM is developed.
- Step 4:** Input each of the k components into the SSA-LSTM prediction model to obtain the k prediction results.
- Step 5:** Finally, the point predictions are obtained by superimposing the predictions of the k components.
- Step 6:** Fitting the actual values of metal prices by means of a distribution function and selecting the distribution function with the largest determination coefficient determination coefficient.
- Step 7:** Based on the results of the optimal distribution function and the point forecast, the interval forecast is designed assuming uncertain information, such as the forecast range and confidence level.

4. Hybrid Predictive Model Applications

This section is divided into three parts: data introduction, parameter setting and model evaluation metrics, as described below.

4.1. Introduction to the Data

The data source for this study is the listed prices of copper and aluminum metal (in RMB) on the Shanghai Futures Exchange, and the daily prices of copper and aluminum metal are used to test the hybrid forecasting model developed. The time interval selected was from 5 January 2012 to 28 November 2022. Specifically, these datasets consist of two subsets: the training dataset (in-sample) ranging from 5 January 2012 to 21 August 2019

(70%), which is used to build the forecasting model; and the testing dataset (out-of-sample) ranging from 22 August 2019 to 28 November 2022 (30%), which is used to validate the performance of the designed model. Details of the total number of samples, the number of training samples, the number of test samples and the maximum, minimum, mean and standard deviation of the samples for the experimental dataset are shown in Table 2.

Table 2. The descriptive statistics of these two metal prices.

Metal Prices (RMB)	Total	Training	Test	Max	Min	Average	Standard Deviation
Copper	2647	1856	791	76,840	33,630	51,640.92	9519.910
Aluminum	2647	1856	791	24,330	9825	14,864.52	2551.76

4.2. Parameter Settings

When the VMD method is used to decompose the original metal price data, the value of the number of modes k has a great influence on the decomposition effect. When k is 7, the corresponding central frequency is more dispersed, and thus, the value is 7.

The neural network of the single LSTM model is constructed using a double hidden layer structure, with the number of neurons H contained in each hidden layer being 20, the training count E being 200 and the learning rate η being taken as 0.005.

The VMD–SSA–LSTM hybrid model was set to have a population size of 50 sparrows and a maximum number of iterations M of 10; 20% of the population was found, with the remainder being joiners. The safety threshold is 0.8, and when the warning value is less than 0.8, there is no predator; otherwise, there is a predator that is dangerous to the population and needs to go elsewhere to feed. The sparrow search algorithm optimization LSTM parameters number of neurons H , training number E and learning rate η are [10, 200], [10, 200] and [0.001, 0.02], respectively.

4.3. Evaluation Indicators

To evaluate the predictive capability of the proposed hybrid system, seven evaluation metrics are introduced in this paper. These performance metrics include the performance of the prediction model at five measurement points [36,37], the package Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Nash–Sutcliffe efficiency coefficient (NSE) and coefficient of determination R^2 , and two generic metrics reflecting interval prediction capability [38,39], Interval forecasts normalized average width (IFNAW) and Interval forecasts coverage probability (IFCP). The detailed formulas and definitions of the above model performance metrics are shown in Table 3.

Table 3. The considered evaluation indices.

Metrics	Definition	Equations
MAE	The mean absolute error	$MAE = \frac{1}{N} \sum_{t=1}^N y_t - \hat{y}_t $
RMSE	The root means square error	$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - \hat{y}_t)^2}$
MAPE	The mean absolute percent error	$MAPE = \frac{1}{N} \sum_{t=1}^N (y_t - \hat{y}_t)/y_t \times 100\%$
R^2	Coefficient of determination	$R^2 = 1 - \left(\frac{\sum_{t=1}^N (\hat{F}_t - F)}{\sum_{t=1}^N (F_t - \bar{F})} \right)^2$
NSE	Nash–Sutcliffe efficiency coefficient	$E = 1 - \frac{\sum_{t=1}^T (Q_t^o - Q_m^t)^2}{\sum_{t=1}^T (Q_t^o - \bar{Q}_o)^2}$
IFNAW	Interval forecasts normalized average width	$IFNAW = \frac{1}{N} \sum_{t=1}^N [(U_t - L_t) / (y_{max} - y_{min})]$
IFCP	Interval forecasts coverage probability	$IFCP = \frac{1}{N} \sum_{t=1}^N c_t, c_t = \begin{cases} 1, & \text{if } y_t \in [L_t, U_t] \\ 0, & \text{otherwise} \end{cases}$

5. Analysis of Experimental Results

In this section, two experiments are conducted: metal price point prediction and interval prediction. In order to demonstrate the better forecasting capability of this hybrid model, four benchmark models are introduced for a comparison, including LSTM, SSA-LSTM, VMD-LSTM and VMD-SSA-LSTM, thus illustrating the effectiveness and forecasting capability of the proposed hybrid system.

5.1. Point Forecasting of Metal Prices

In this section, an innovative hybrid model combining Variational Modal Decomposition (VMD) and SSA-LSTM, namely VMD-SSA-LSTM, is used for the point forecasting of prices of copper and aluminum metals. To demonstrate that the hybrid VMD-SSA-LSTM model proposed in this paper has good forecasting capability, three comparative models, LSTM, SSA-LSTM and VMD-LSTM, are also developed in this paper. A comparison of the price forecast and true value trends of the four models is shown in Figure 2, and a comparison of the forecast scatter is shown in Figure 3. In addition, five evaluation metrics were used to reflect the prediction level of the models, MAE, RMSE, MAPE, NSE and R^2 , and they are proposed. The detailed values of the errors for the proposed hybrid and comparative models are shown in Table 4, where the bolded labeled values are the best values for each evaluation indicator.

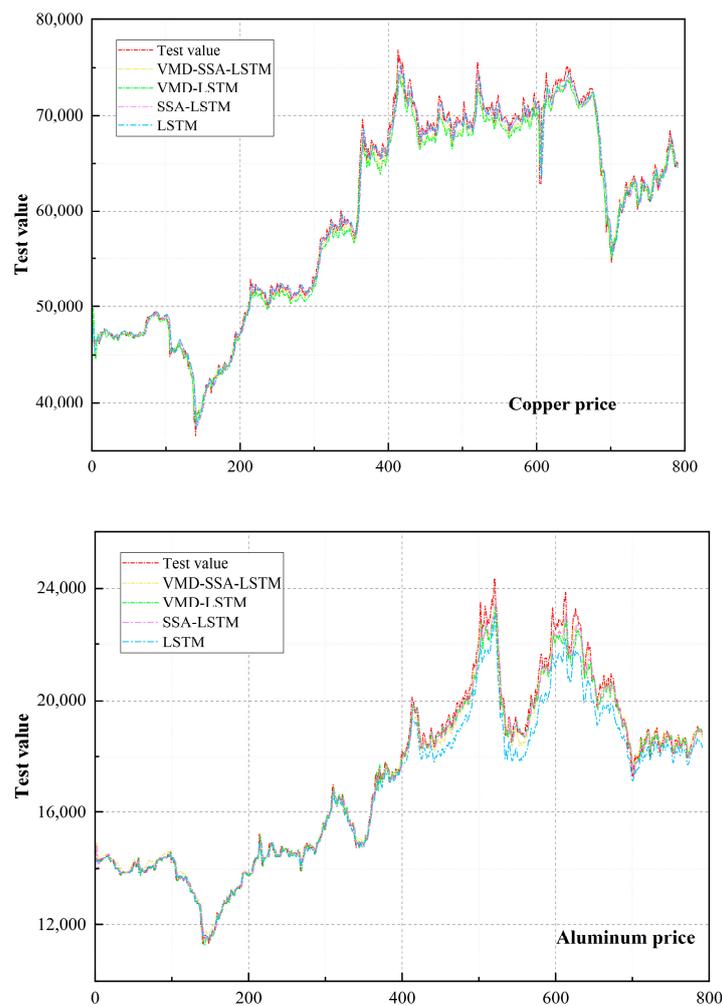


Figure 3. Comparison of the trend of the four models' price forecasts and true values.

Table 4. Forecast errors of different models for copper and aluminum metal prices.

Metal Prices	Methods	MAE	RMSE	MAPE	R ²	NSE
Copper price	LSTM	633.8126	995.2734	1.1216%	0.99012	0.98961
	SSA-LSTM	602.5823	874.5603	1.0183%	0.99546	0.99198
	VMD-LSTM	600.2264	821.1074	1.0841%	0.99608	0.99293
	VMD-SSA-LSTM	543.2012	763.0124	0.8789%	0.99827	0.99389
Aluminum price	LSTM	428.52	610.0284	2.1628%	0.99604	0.96192
	SSA-LSTM	212.084	316.8024	1.1392%	0.99289	0.98973
	VMD-LSTM	213.1229	303.9368	1.1715%	0.99221	0.99055
	VMD-SSA-LSTM	172.7794	233.5254	0.9464%	0.99733	0.99442

As can be seen from Table 4, the VMD-SSA-LSTM model has a more obvious advantage in predicting metal prices compared to the LSTM model and the VMD-LSTM model, where for copper metal prices, with a reduction of 14.30%, 9.85% and 9.50%, respectively, in the indicator MAE and 23.34%, 12.75% and 7.08%, respectively, in the RMSE, decreases in MAPE of 0.24%, 0.13% and 0.21%, respectively, increases in the R² indicator of 0.82%, 0.28% and 0.22% and increases in the NSE of 0.43%, 0.19% and 0.10%. For aluminum metal prices, there were decreases of 59.68%, 18.53% and 18.93%, respectively, in the indicator MAE, and the RMSE decreased by 61.72%, 26.29% and 23.17%, respectively, with decreases in MAPE of 1.22%, 0.19% and 0.23%, respectively. The indicator R² increased by 0.13%, 0.45% and 0.52%, respectively, and NSE increased by 3.38%, 0.47% and 0.39%, respectively.

Figure 3 reflects the comparison between the predicted and true price trends of the four models. It can be seen from Figure 2 that the LSTM model has poor prediction results due to its large deviation from the actual values, while the SSA model has the second best prediction results, and the VMD-SSA-LSTM and VMD-LSTM models have better prediction results due to their higher overlap with the measured values, but there is no significant difference between these two models; furthermore, according to the scatter plot in Figure 3, it can be seen that the VMD-SSA-LSTM model has the best convergence and slightly outperforms the VMD-LSTM and SSA-LSTM models in terms of peak prediction accuracy and fluctuation.

It can be seen from Figure 3 that the LSTM model has poor prediction results due to its large deviation from the actual values, while the SSA model has the second best prediction results, and the VMD-SSA-LSTM and VMD-LSTM models have better prediction results due to their higher overlap with the measured values, but there is no significant difference between these two models; furthermore, according to the scatter plot in Figure 4, it can be seen that the VMD-SSA-LSTM model has the best convergence and slightly outperforms the VMD-LSTM and SSA-LSTM models in terms of peak prediction accuracy and fluctuation.

In summary, it can be seen that the VMD-SSA-LSTM model has the highest prediction accuracy, and it can also be seen that after the data has been decomposed by VMD, the prediction accuracy is much higher than that of the data without VMD decomposition, which indicates that VMD decomposition is sufficient for noise reduction of the original time series and extracting the complex and effective information implied in the price data, and to a certain extent, it can better reflect the intrinsic mechanism of prices. Thus, it also shows the importance of VMD decomposition in time series forecasting. The SSA optimizes the model parameters of the LSTM and improves the efficiency of model parameter selection, which shows that the hybrid VMD-SSA-LSTM forecasting model proposed in this paper is effective.

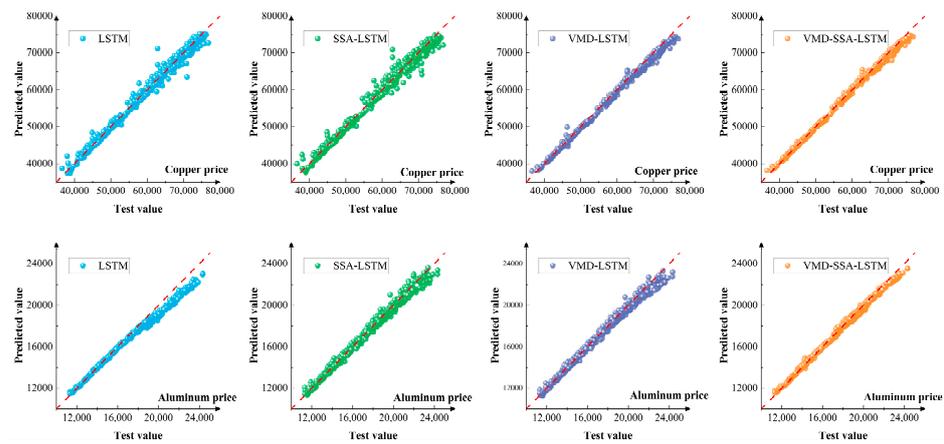


Figure 4. Scatter plot of price prediction values from the hybrid model and the three comparative models.

5.2. Interval Forecasting of Metal Prices

5.2.1. Distribution Fitting

Numerous scholars have introduced a variety of distribution functions, DFs, to describe time series in different fields [31,40]. The determination coefficient ($0 \leq R^2 \leq 1$) is used to determine the level of fit of these DFs. The larger the R^2 , the better the fit of the DFs.

In this section, five DFs—Weibull, Logistic, Lognormal, Normal and Gamma—are used to represent metal price states. The associated probability density functions and cumulative distribution functions are shown in Table 1. The calculation of the parameters of the DFs is an important step in this section. The traditional method for estimating the parameters of the DFs is to use the method of great likelihood estimation (MLE). However, parameters can also be calculated using intelligent optimization algorithms to maximize the value of the objective function of R^2 , ultimately improving the fit of the DFs.

In this paper, a Sparrow Search Algorithm (SSA) is used to optimize the parameters of interest, and MLE is used as a comparative method to illustrate the excellent optimization performance of SSA. Table 5 shows the parameters of the five DFs estimated using both MLE and SSA methods. These parameters can be used to describe the scales and translations of these DFs. Using the coefficient of determination R^2 as an evaluation indicator that reflects the ability to fit the distribution, the values are shown in Table 6. It can be seen that the values of R^2 calculated based on MLE and SSA are different, and the value of R^2 for SSA is higher than that of MLE, indicating that SSA is associated with a better estimation in all cases, which further confirms that SSA is more effective than MLE. We can also see that the Weibull distribution function is the most appropriate distribution function in most cases. It can also be seen from Table 5 that the values of these parameters are different for the DFs obtained using the two methods but are all within one order of magnitude.

Table 5. Parameter values of the different distribution functions determined using MLE and SSA.

Metal Prices	Methods	Logistic		Weibull		Normal		Gamma		Lognormal	
		μ	σ	λ	k	μ	σ	k	θ	μ	σ
Copper price	MLE	50,959.1	5312.24	55,680.8	5.5724	51,640.9	9519.91	30.33	1702.38	10.835	0.181
	SSA	48,940.5	5121.32	54,524.2	6.1856	50,237.3	9454.52	25.63	1759.68	10.152	0.192
Aluminum price	MLE	14518.1	1310.55	15,979.6	5.43863	14,864.5	2551.76	37.5543	395.814	9.59336	0.160201
	SSA	13125.6	1225.36	14,658.7	6.09564	13,582.9	2415.85	33.7821	409.254	8.87562	0.18.245

Table 6. R^2 values of the five distribution functions based on MLE and GWO.

Metal Prices	Methods	Logistic	Weibull	Normal	Gamma	Lognormal
Copper price	MLE	0.9853	0.9973	0.9927	0.9859	0.9796
	SSA	0.9958	0.9995	0.9986	0.9961	0.9938
Aluminum price	MLE	0.9856	0.9976	0.9956	0.9901	0.9852
	SSA	0.9989	0.9997	0.9996	0.9990	0.9984

5.2.2. Interval Forecasts

Unlike point forecasts, interval forecasts can provide upper and lower bounds on observations and construct interval forecasts at a given level of significance. It can provide investors in the metal’s financial markets with more uncertain information to help them analyze metal market conditions. Based on the point forecasting results of the hybrid model (VMD–SSA–LSTM), the metal price best-fit distribution function (normal distribution) and the interval prediction method [31–35], interval forecasts of metal prices can be made at a given significance α level. In addition, two evaluation metrics listed in Table 3 (IFNAW and IFCP) are presented in this section with the aim of evaluating the performance of interval prediction. It should be noted that the optimal interval prediction should satisfy the following conditions: the larger the IFCP ($0 \leq \text{IFCP} \leq 1$) and the smaller the IFNAW at the α significant level, the better the predictive performance of the interval prediction at the same time.

Table 7 presents the metal price interval forecasting results for five different significance levels. Based on the interval forecasting results in Table 7, it is evident that the interval forecasts vary in accuracy due to changes in width at different significance levels; for example, the interval forecasts for copper prices are 84.32% (IFCP), 0.0383 (IFNAW), 100% (IFCP) and 0.0454 (IFNAW) at significance levels of 0.1 and 0.05, respectively.

Table 7. The interval forecasting results of the metal price under different significance levels.

Level (α)		0.1	0.07	0.05	0.03	0.01
Copper price	IFCP	84.32%	96.59%	100%	100%	100%
	IFNAW	0.0383	0.0421	0.0454	0.0500	0.0589
Aluminum price	IFCP	85.34%	88.50%	90.52%	93.05%	99.12%
	IFNAW	0.0394	0.0433	0.0467	0.0514	0.0605

For a more visual presentation of the interval forecasts, Figure 5 shows the three metal price interval forecasts at three significant levels. As can be seen from the graph, the dots represent the actual values and the colored areas represent the forecast intervals. It is clear that the interval forecasts perform well, with a large number of actuals in the shaded areas. In particular, for point forecasts of a time series where the errors are often larger at the peaks and troughs and where there is greater volatility, interval forecasting can improve the accuracy of the forecasts in these areas, and it is also worth noting that the forecasting performance of interval forecasting depends, to a large extent, on the results of point forecasting.

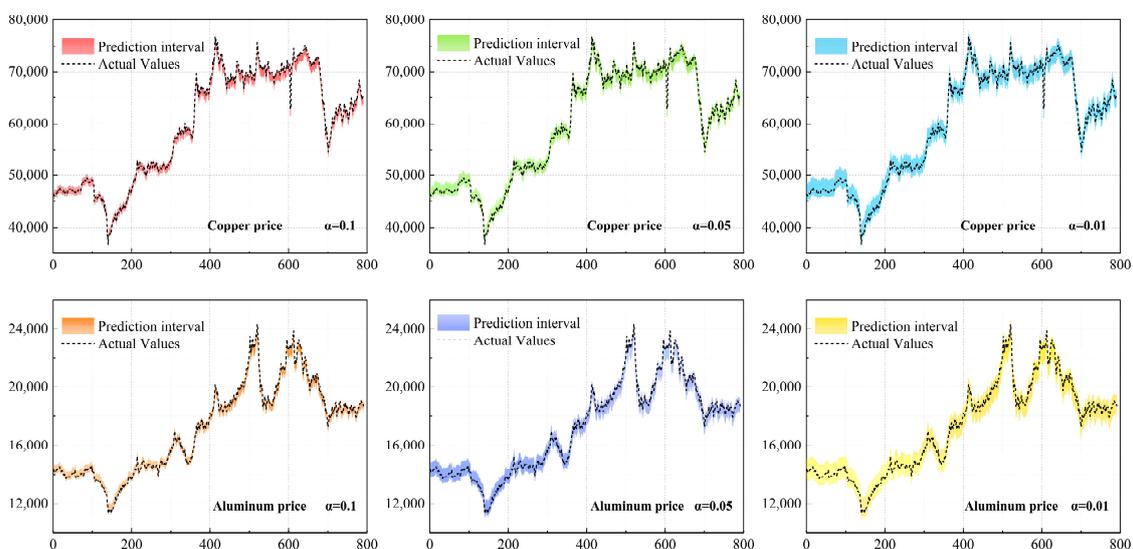


Figure 5. Interval forecasts for three metal prices with different significant levels.

6. Conclusions

This paper presents a hybrid forecasting system for a metal price time series. For point forecasting, a new hybrid forecasting model for metal price forecasting is built based on variational modal decomposition, a sparrow search algorithm and a long short-term memory neural network LSTM technique. Tests against other forecasting models showed a year-on-year reduction of around 10% in the Mean Absolute Error (MAE), indicating that the hybrid forecasting model outperformed other comparative models. In addition, the point forecasting results also fully illustrate that the forecasting model makes full use of the VMD decomposition technique for data noise reduction and the SSA for optimizing the LSTM neural network parameters to improve the accuracy of metal price forecasting. In terms of interval prediction, five distribution functions are first introduced, the distribution characteristics of metal prices are analyzed and the parameters of the distribution functions are optimized using the SSA, which improves the distribution fitting ability and shows strong optimization performance, with the best fitting of the positive-terrestrial distribution function and the highest coefficient of determination. The interval prediction results of metal prices were then obtained based on the optimal point prediction results and the optimal distribution function. The numerical results show that there is a good interval-prediction effect at different significance levels, with higher IFCP values and smaller IFNAW values, and the majority of the actual values of the data lie in the prediction interval, and the interval coverage can basically reach over 85%.

In conclusion, this study has developed a mixed point and interval forecasting model for copper and aluminum metal prices. The forecasting model has better accuracy and robustness in terms of forecasting results and forecasting capability, which can provide more reference value for futures investors and policy-makers. At the same time, the authors believe that the prediction results of this forecasting model need to be applied to engineering practice, and the next step in the research is to prepare the predicted price as dynamic parameters to be applied to the dynamic boundary grade measurement of metal mines to provide theoretical guidance for the future production planning of mining enterprises. In addition, the single-step forecasting approach has some limitations, and further research will be carried out in the future in the area of multi-step forecasting.

Author Contributions: Conceptualization, Methodology, Software, Validation, Writing—original draft, Writing—review & editing, Y.Z.; Visualization, H.S.; Supervision, T.S.; Funding acquisition, J.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation Project of China, grant number 52274163.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chen, Y.; He, K.; Zhang, C. A Novel Grey Wave Forecasting Method for Predicting Metal Prices. *Resour. Policy* **2016**, *49*, 323–331. [[CrossRef](#)]
2. He, K.; Lu, X.; Zou, Y.; Keung Lai, K. Forecasting Metal Prices with a Curvelet Based Multiscale Methodology. *Resour. Policy* **2015**, *45*, 144–150. [[CrossRef](#)]
3. Morales, L.; Andreosso-O'Callaghan, B. Comparative Analysis on the Effects of the Asian and Global Financial Crises on Precious Metal Markets. *Res. Int. Bus. Financ.* **2011**, *25*, 203–227. [[CrossRef](#)]
4. Bhatia, V.; Das, D.; Tiwari, A.K.; Shahbaz, M.; Hasim, H.M. Do Precious Metal Spot Prices Influence Each Other? Evidence from a Nonparametric Causality-in-Quantiles Approach. *Resour. Policy* **2018**, *55*, 244–252. [[CrossRef](#)]
5. Heizer, J.; Render, B. Operations Management. In *Social Sciences Books*; McGraw-Hill Education: New York, NY, USA, 2014.
6. Mai, N.L.; Topal, E.; Erten, O.; Sommerville, B. A New Risk-Based Optimisation Method for the Iron Ore Production Scheduling Using Stochastic Integer Programming. *Resour. Policy* **2019**, *62*, 571–579. [[CrossRef](#)]
7. Dooley, G.; Lenihan, H. An Assessment of Time Series Methods in Metal Price Forecasting. *Resour. Policy* **2005**, *30*, 208–217. [[CrossRef](#)]
8. Dehghani, H.; Bogdanovic, D. Copper Price Estimation Using Bat Algorithm. *Resour. Policy* **2018**, *55*, 55–61. [[CrossRef](#)]

9. Cheng, C.; Sa-Ngasoongsong, A.; Beyca, O.; Le, T.; Yang, H.; Kong, Z.; Bukkapatnam, S.T.S. Time Series Forecasting for Nonlinear and Non-Stationary Processes: A Review and Comparative Study. *IIE Trans.* **2015**, *47*, 1053–1071. [[CrossRef](#)]
10. Liu, D.; Li, Z. Gold Price Forecasting and Related Influence Factors Analysis Based on Random Forest. In *Proceedings of the 10th International Conference on Management Science and Engineering Management*; Xu, J., Hajiyev, A., Nickel, S., Gen, M., Eds.; Springer: Singapore, 2017; pp. 711–723.
11. Wang, C.; Zhang, X.; Wang, M.; Lim, M.K.; Ghadimi, P. Predictive Analytics of the Copper Spot Price by Utilizing Complex Network and Artificial Neural Network Techniques. *Resour. Policy* **2019**, *63*, 101414. [[CrossRef](#)]
12. Sánchez Lasheras, F.; de Cos Juez, F.J.; Suárez Sánchez, A.; Krzemień, A.; Riesgo Fernández, P. Forecasting the COMEX Copper Spot Price by Means of Neural Networks and ARIMA Models. *Resour. Policy* **2015**, *45*, 37–43. [[CrossRef](#)]
13. Buncic, D.; Moretto, C. Forecasting Copper Prices with Dynamic Averaging and Selection Models. *N. Am. J. Econ. Financ.* **2015**, *33*, 1–38. [[CrossRef](#)]
14. Kriechbaumer, T.; Angus, A.; Parsons, D.; Rivas Casado, M. An Improved Wavelet—ARIMA Approach for Forecasting Metal Prices. *Resour. Policy* **2014**, *39*, 32–41. [[CrossRef](#)]
15. De Gooijer, J.G.; Hyndman, R.J. 25 Years of Time Series Forecasting. *Int. J. Forecast.* **2006**, *22*, 443–473. [[CrossRef](#)]
16. Liu, C.; Hu, Z.; Li, Y.; Liu, S. Forecasting Copper Prices by Decision Tree Learning. *Resour. Policy* **2017**, *52*, 427–434. [[CrossRef](#)]
17. Dehghani, H. Forecasting Copper Price Using Gene Expression Programming. *J. Min. Environ.* **2018**, *9*, 349–360. [[CrossRef](#)]
18. Khoshalan, H.A.; Shakeri, J.; Najmoddini, I.; Asadizadeh, M. Forecasting Copper Price by Application of Robust Artificial Intelligence Techniques. *Resour. Policy* **2021**, *73*, 102239. [[CrossRef](#)]
19. Zhang, H.; Nguyen, H.; Bui, X.-N.; Pradhan, B.; Mai, N.-L.; Vu, D.-A. Proposing Two Novel Hybrid Intelligence Models for Forecasting Copper Price Based on Extreme Learning Machine and Meta-Heuristic Algorithms. *Resour. Policy* **2021**, *73*, 102195. [[CrossRef](#)]
20. Seguel, F.; Carrasco, R.; Adasme, P.; Alfaro, M.; Soto, I. A Meta-Heuristic Approach for Copper Price Forecasting. In *Proceedings of the Information and Knowledge Management in Complex Systems*; Liu, K., Nakata, K., Li, W., Galarreta, D., Eds.; Springer International Publishing: Cham, Switzerland, 2015; pp. 156–165.
21. Alameer, Z.; Elaziz, M.A.; Ewees, A.A.; Ye, H.; Jianhua, Z. Forecasting Copper Prices Using Hybrid Adaptive Neuro-Fuzzy Inference System and Genetic Algorithms. *Nat. Resour. Res.* **2019**, *28*, 1385–1401. [[CrossRef](#)]
22. Hu, Y.; Ni, J.; Wen, L. A Hybrid Deep Learning Approach by Integrating LSTM-ANN Networks with GARCH Model for Copper Price Volatility Prediction. *Phys. A Stat. Mech. Its Appl.* **2020**, *557*, 124907. [[CrossRef](#)]
23. Zhang, P.; Ci, B. Deep Belief Network for Gold Price Forecasting. *Resour. Policy* **2020**, *69*, 101806. [[CrossRef](#)]
24. Dragomiretskiy, K.; Zosso, D. Variational Mode Decomposition. *IEEE Trans. Signal Process.* **2014**, *62*, 531–544. [[CrossRef](#)]
25. Shi, G.; Qin, C.; Tao, J.; Liu, C. A VMD-EWT-LSTM-Based Multi-Step Prediction Approach for Shield Tunneling Machine Cutterhead Torque. *Knowl.-Based Syst.* **2021**, *228*, 107213. [[CrossRef](#)]
26. Zhang, Y.; Chen, B.; Pan, G.; Zhao, Y. A Novel Hybrid Model Based on VMD-WT and PCA-BP-RBF Neural Network for Short-Term Wind Speed Forecasting. *Energy Convers. Manag.* **2019**, *195*, 180–197. [[CrossRef](#)]
27. Xue, J.; Shen, B. A Novel Swarm Intelligence Optimization Approach: Sparrow Search Algorithm. *Syst. Sci. Control Eng.* **2020**, *8*, 22–34. [[CrossRef](#)]
28. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]
29. Dhruv, P.; Naskar, S. Image Classification Using Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN): A Review. In *Proceedings of the Machine Learning and Information Processing*; Swain, D., Pattnaik, P.K., Gupta, P.K., Eds.; Springer: Singapore, 2020; pp. 367–381.
30. Sherstinsky, A. Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) Network. *Phys. D Nonlinear Phenom.* **2020**, *404*, 132306. [[CrossRef](#)]
31. Du, P.; Wang, J.; Yang, W.; Niu, T. Point and Interval Forecasting for Metal Prices Based on Variational Mode Decomposition and an Optimized Outlier-Robust Extreme Learning Machine. *Resour. Policy* **2020**, *69*, 101881. [[CrossRef](#)]
32. Qin, S.; Liu, F.; Wang, J.; Song, Y. Interval Forecasts of a Novelty Hybrid Model for Wind Speeds. *Energy Rep.* **2015**, *1*, 8–16. [[CrossRef](#)]
33. Song, Y.; Qin, S.; Qu, J.; Liu, F. The Forecasting Research of Early Warning Systems for Atmospheric Pollutants: A Case in Yangtze River Delta Region. *Atmos. Environ.* **2015**, *118*, 58–69. [[CrossRef](#)]
34. Tian, C.; Hao, Y. Point and Interval Forecasting for Carbon Price Based on an Improved Analysis-Forecast System. *Appl. Math. Model.* **2020**, *79*, 126–144. [[CrossRef](#)]
35. Xu, Y.; Du, P.; Wang, J. Research and Application of a Hybrid Model Based on Dynamic Fuzzy Synthetic Evaluation for Establishing Air Quality Forecasting and Early Warning System: A Case Study in China. *Environ. Pollut.* **2017**, *223*, 435–448. [[CrossRef](#)] [[PubMed](#)]
36. Gupta, H.V.; Kling, H. On Typical Range, Sensitivity, and Normalization of Mean Squared Error and Nash-Sutcliffe Efficiency Type Metrics. *Water Resour. Res.* **2011**, *47*, W10601. [[CrossRef](#)]
37. McCuen, R.H.; Knight, Z.; Cutter, A.G. Evaluation of the Nash–Sutcliffe Efficiency Index. *J. Hydrol. Eng.* **2006**, *11*, 597–602. [[CrossRef](#)]
38. Khosravi, A.; Nahavandi, S.; Creighton, D.; Atiya, A.F. Lower Upper Bound Estimation Method for Construction of Neural Network-Based Prediction Intervals. *IEEE Trans. Neural Netw.* **2011**, *22*, 337–346. [[CrossRef](#)] [[PubMed](#)]

39. Khosravi, A.; Nahavandi, S.; Creighton, D. Prediction Intervals for Short-Term Wind Farm Power Generation Forecasts. *IEEE Trans. Sustain. Energy* **2013**, *4*, 602–610. [[CrossRef](#)]
40. Yang, H.; Zhu, Z.; Li, C.; Li, R. A Novel Combined Forecasting System for Air Pollutants Concentration Based on Fuzzy Theory and Optimization of Aggregation Weight. *Appl. Soft Comput.* **2020**, *87*, 105972. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.