Research Article

EMD-GM-ARMA Model for Mining Safety Production Situation Prediction

Menglong Wu ¹, Yicheng Ye,^{1,2} Nanyan Hu ¹, Qihu Wang,¹ Huimin Jiang,¹ and Wen Li¹

¹School of Resource and Environmental Engineering, Wuhan University of Science and Technology, Wuhan, Hubei 430081, China

²Industrial Safety Engineering Technology Research Center of Hubei Province, Wuhan, Hubei 430081, China

Correspondence should be addressed to Nanyan Hu; hunanyan@wust.edu.cn

Received 8 January 2020; Revised 17 April 2020; Accepted 21 May 2020; Published 8 June 2020

Academic Editor: Alex Alexandridis

Copyright © 2020 Menglong Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to improve the prediction accuracy of mining safety production situation and remove the difficulty of model selection for nonstationary time series, a grey (GM) autoregressive moving average (ARMA) model based on the empirical mode decomposition (EMD) is proposed. First of all, according to the nonstationary characteristics of the mining safety accident time series, nonstationary original time series were decomposed into high- and low-frequency signals using the EMD algorithm, which represents the overall trend and random disturbances, respectively. Subsequently, the GM model was used to predict high-frequency signal sequence, while the ARMA model was used to predict low-frequency signal sequence. Finally, aiming to predict the mining safety production situation, the EMD-GM-ARMA model was constructed via superimposing the prediction results of each subsequence, thereby compared to the ARIMA model, wavelet neural network model, and PSO-SVM model. The results demonstrated that the EMD-GM-ARMA model and the PSO-SVM model hold the highest prediction accuracy in the short-term prediction, and the wavelet neural network has the lowest prediction accuracy. The PSO-SVM model's prediction accuracy decreases in medium- and long-term predictions while the EMD-GM-ARMA model still can maintain high prediction accuracy. Moreover, the relative error fluctuations of the EMD-GM-ARMA model are relatively stable in both short-term and medium-term predictions. This shows that the EMD-GM-ARMA model can provide high-precision predictions with high stability, proving the model to be feasible and effective in predicting the mining safety production situation.

1. Introduction

Mining is an important industry for the nation's economy, supporting the rapid development of society. However, it is also an industry afflicted with a high casualty rate. According to the statistics of mining accidents in recent years in China [1], 3,797 mining safety accidents causing 14,169 deaths occurred between 2005 and 2018, a figure higher than those in developed countries. Such security-related problems hinder the development of the economy and society. Therefore, accurate prediction of mining safety production situation is a prerequisite for providing mining safety decision-making and ensuring safe production. As a result, it is necessary to explore the development law of mining safety production situation and analyze its future changes.

The mining safety system is one of the complex circulations [2]. The issue of safe production in the mining industry has the characteristics of being nonlinear, nonstationary, and uncertain, and it is affected by several factors: natural conditions, social environment, laws and regulations, corporate environment, and operating environment. The occurrence of an accident presents a nonstationary random process with a certain trend over time. In 1970, Box and Jenkins published *Time Series Analysis: Forecasting and Control* [3], which laid the foundation for time series analysis. After half-a-century of development, time series analysis has made great progress and has been widely used in various fields. It makes inferences about the future changes of variables based on the constructed dynamic time series model. Until now,

studies on nonstationary random wave time series have yielded some results, mainly including the grey theory [4], the Markov chain theory [5], the time series analysis [6], the support vector machine (SVM) model [7], and the neural network [8]. Lan and Zhou [9] analyzed the causes of human error in mining accidents, combined the advantages of grey prediction and the Markov theory, and proposed an improved grey Markov model to simulate the development trend of mine safety accidents. Liu et al [10] used the GM (1, 1) for data trend prediction and the support vector machine to predict the residual sequence. Finally, the results of the two models were merged into a prediction model, namely, the grey support vector machine GM (1, 1)-SVM. Bao et al [11] combined the C²R model and the GM (1, 1) to predict the safety benefits of a mine, which can ensure the effective operation of a mine safety management system. Wang [12] proposed a Markov time series combination prediction model to fit the relative prediction error of the number of safety production accidents. Li [13] first established an environmental impact factor system by selecting the environmental factors that affected the safety production situation in the mining industry. He then determined the weight of each environmental factor using the analytic hierarchy process. Jang et al. [14] developed a prediction model using a long short-term memory (LSTM) recurrent neural network (RNN), which is a deep-learning algorithm. Wang and Jiang [15] used the rough sets (RS) method as a pre-processor to first reduce the feature parameters and then make classification modeling based on the SVM method. The mining safety production early warning index system was established on the basis of personnel, environment, equipment, and management. The genetic algorithm was used to reduce the RS attribute, and the rough set support vector machine combined forecasting model was used to predict the risks in coal mine production.

However, the development process of the Chinese mining safety accident database is imperfect, and limited accurate statistical data are available. Some data are greatly affected by random factors, and the distribution law cannot be identified correctly. Moreover, it is difficult to define the indicators of mining safety production accidents, and its data structure is single. The inaccuracy and incompleteness of the selection of the relative indicators of mining safety production and the quantization errors will affect the prediction results, leading to prediction errors.

In view of the various factors and their complicated interconnection that affects the safety production situation of the mining industry, the intricacies of various factors, and factors such as incomplete selection of indicators, the time series of 10,000 population mortality in mining production is selected as the research object. This can effectively solve the problems related to the uncertainty of selection and quantification of index. In 1998, the Chinese-American Norden E. Huang proposed a new signal processing method, i.e., the EMD method. Since then, domestic and foreign scholars have conducted a lot of research on the subject, and the EMD method has developed rapidly [16]. It is an adaptive binary filter bank, including two processes of

decomposition and reconstruction [17]. The signal is decomposed into a number of characteristic modal functions with different characteristic scales through the EMD, instead of decomposing the signal into a sine or cosine function like the usual Fourier transform [18]. The EMD can analyze both linear steady-state signals and nonlinear and nonstationary signals and has a good signal-to-noise ratio. Therefore, the EMD method of signal analysis and processing is considered to predict the situation of mining safety production. Through the EMD, the time series of mining safety production situation is decomposed into several stationary component sequences, and the characteristics of the original sequence are revealed by studying these components.

The stationary component sequence after EMD has strong time sequence, and it shows periodic fluctuations with time. Although machine learning and deep learning methods such as support vector machines and neural networks can better extract nonlinear information in stationary component sequences, they do not consider the timing of stationary sequences, and the prediction accuracy for medium and long-term is not high. The autoregressive moving average model is the most widely used model in time series analysis [19]. It only requires that the prediction sequence is a stationary sequence. Not only can the nonlinearity and timing of the stationary component be taken into account but also the amount of information required for modeling is small, which can achieve high-precision prediction of small samples. Good results can also be obtained in the mid- to long-term prediction [20]. Therefore, the ARMA model can be used to predict the stationary component sequence after EMD to achieve the purpose of solving complex practical problems with a simple method.

In the case of the EMD trend series, it has certain ambiguity and randomness. The ambiguity of the sequence is reflected in the statistical error of the information, and the grayness of the sequence is reflected in the incompleteness of the data information, and the trend sequence after EMD is a monotone sequence [21]. For this, the GM model can be used to predict the trend component series. It is a method for processing incomplete information. It has no special requirement nor is there any restriction on the data and its distribution. Even if there is only a small amount of historical data and an arbitrary random distribution, a high level of prediction accuracy can be obtained [22]. The sequence of trend components decomposed by the EMD method belongs to the “small sample” and “poor information” uncertain system where “some information is known and some information is unknown.” The GM can analyze the degree of disparity observed in development trends of system factors to correlate, analyze, and generate a series of observed numerical values. These reflect the characteristics of the prediction object and establish corresponding differential equations to predict the future development trend of things, including the two processes of establishing a GM model and using the GM model to make extrapolation predictions.

By processing the time series of the 10,000 population mortality rate in the mining production situation using the EMD method, decomposing a time series into several

components can decompose the hidden information of the original time series, make the characteristics of the original time series clearer, and preserve its dynamic characteristics. Then, these components are studied to reveal the characteristics of the original sequence. According to the characteristics of each component, the GM model is used to predict the trend sequence, and the ARMA model is used to predict the smooth detail sequence, and then the EMD-GM-ARMA model is established to predict the future development trend of mining safety production. It provides a new resolution for the study of mining safety production situation.

The rest of this paper is organized as follows: Section 2 provides a detailed introduction to the modeling principle and process of the EMD-GM-ARMA model. Section 3 explains the source of the data and validates the model. The comparison of prediction accuracy of the models is performed in Section 4. Section 5 discusses and elaborates the future research directions, while the conclusions drawn from this research are summarized in Section 6.

2. Establishment of the EMD-GM-ARMA Model

The time series of mining safety production situation shows a stochastic process with nonstationary characteristic, which makes it difficult to accurately identify its distribution law. The EMD algorithm can discretize the random nonstationary time series into a high-frequency and a low-frequency signals sequence, representing the overall change trend and random disturbance, respectively. It can completely retain the information of the original time series [23].

The GM model is suitable for the original dynamic sequence with obvious exponential law at equal intervals. It can describe the internal characteristics and development trend of the system and has higher prediction accuracy for variables with exponential increase or decrease trend [24]. Although the single sequence value constituting the time series in the ARMA model is uncertain, the change of the whole sequence has certain regularity, so it can be approximated using the corresponding mathematical model. The ARMA model is suitable for random and stable time series [25]. Therefore, the time series can be decomposed into low-frequency signals with random fluctuations and high-frequency signals with overall trend using the EMD method. The GM model is used to predict the high-frequency signals from the EMD, while the ARMA model is used to predict the low-frequency signals. The prediction results of the original time series can be obtained to establish the EMD-GM-ARMA model, which results from the superposition of abovementioned two model predictions. The modeling principle of the EMD-GM-ARMA model is shown in Figure 1. The modeling steps are as follows.

Step 1. Time series decomposition using the EMD method

Based on the local feature time scale of the signal, the EMD method is used to smoothen the original signal and decompose the complex signal into a finite number of data sequences with different characteristic scales, which is

named the intrinsic mode function (IMF). The IMF signal is based on the following three assumptions [26, 27]:

- (1) The original time series has at least one maximum point and one minimum point.
- (2) The characteristic time scale is defined as the time interval between adjacent extreme points.
- (3) If the original time series has only inflection points without extreme points, the original time series should be differentiated either once or multiple times to obtain extreme points before it is decomposed. Then, integration of the obtained components gives the corresponding results.

The EMD algorithm realizes the decomposition of the mining safety production situation time series through the process of “screening,” and the change trend of the original time series is clearly shown by transformation. The algorithm process is as follows [28]:

- (1) For the mining safety production situation time series $X(t)$, all extreme points of the time series are determined. The upper envelope $E^+(t)$ and the lower envelope $E^-(t)$ are obtained after the connection of all the maxima and minima points with a cubic spline curve. Here, h_t is defined as the following equation:

$$h_t = X(t) - \frac{[E^+(t) + E^-(t)]}{2}. \quad (1)$$

h_1 is considered as a new time series $X(t)$. The first-order IMF (recorded as C_1) is selected from the original mining safety production situation time series, once the h_t meets the two conditions of the IMF during the repeat process of the above-given steps.

- (2) A difference time sequence r_1 without the IMF1 component is obtained when the C_1 is separated from $X(t)$:

$$r_1 = X(t) - C_1. \quad (2)$$

r_1 is used as a new time series. The screening process of (1) is repeated until the residual sequence of the n_{th} order becomes a monotonic sequence.

$$r_n = r_{n-1} - C_n. \quad (3)$$

- (3) In mathematical terms, the original mining safety production situation time series $X(t)$ can be expressed as the sum of n IMF components and one residual term $r_n(t)$.

$$X(t) = \sum_{j=1}^n C_j(t) + r_n(t), \quad (4)$$

where $r_n(t)$ is the residual, representing the overall trend of the original time series.

FIGURE 1: The EMD-GM-ARMA modeling schematic diagram.

Each IMF component $C_j(t)$ represents the composition of the original time series from high-frequency segment to low-frequency segment.

The IMF component reflects the nature and real information of the original time series. The instantaneous frequency of time series has a certain physical meaning after it is decomposed by the EMD method [29]. The EMD algorithm is suitable for complex time series of nonstationary and nonlinear data sequences. Therefore, it is feasible to introduce the EMD algorithm into the time series prediction of the mining safety production situation.

Step 2. The prediction of the IMF sequence using the ARMA model

The ARMA model has high precision in stationary random sequence prediction, especially for the ARMA (p, q) model. It can be used to predict the IMF sequences [30]. The response c_t of the IMF sequence at time t is not only related to the response $c_{t-1}, c_{t-2}, \dots, c_{t-n}$ of its previous time $t-1, t-2, \dots, t-n$, but also affected by the disturbance $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-n}$, when entering the system at the time of $t-1, t-2, \dots, t-n$. This type of model is the ARMA model, recorded as ARMA (p, q) . Its mathematical model is as follows [31]:

$$\begin{cases} c_t = \varphi_0 + \varphi_1 c_{t-1} + \dots + \varphi_p c_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}; \\ \varphi_p \neq 0, \\ \theta_q \neq 0; \\ E(\varepsilon_t) = 0, \\ \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2, \\ E(\varepsilon_t \varepsilon_s) = 0, \\ s \neq t; \\ Ec_s \varepsilon_t = 0, \\ \forall s < t, \end{cases} \quad (5)$$

where c_t is the IMF sequence; ε_t is the white noise process; and p and q are the orders of the ARMA model.

The order of the model is determined by the AIC criteria proposed by Akaike in 1973, which defines the AIC function as

$$\begin{aligned} \text{AIC}(k, j) &= n \ln \hat{\sigma}^2(k, j) + 2(k + j), \quad k = 0, 1, \dots, P_{\max}, \\ j &= 0, 1, \dots, q_{\max}, \end{aligned} \quad (6)$$

where $\hat{\sigma}^2$ is an estimation of σ^2 in the ARMA (k, j) model. Generally, the values of $P_{\max}$ and $q_{\max}$ are empirically determined, and $\hat{p}$ and $\hat{q}$ are the orders obtained from the equation $\text{AIC}(\hat{p}, \hat{q}) = \min \text{AIC}(k, j)$.

Step 3. Residual sequence prediction using the GM model

The EMD method residuals have grey and fuzzy features, so the GM model is suitable for predicting the residuals. In order to construct the GM model, the original data is first treated by the association analysis, thereby transforming the irregular raw data into a generated sequence with strong correlation. Then, the relevant differential equation model is established. The GM prediction steps are as follows [32, 33].

The EMD method-created residual time series is $R^{(0)}(t)$,

$$R^{(0)}(t) = (r^{(0)}(1), r^{(0)}(2), \dots, r^{(0)}(n)). \quad (7)$$

An accumulation of $R^{(0)}(t)$:

$$R^{(1)}(t) = (r^{(1)}(1), r^{(1)}(2), \dots, r^{(1)}(n)). \quad (8)$$

$Z^{(1)}$ is the MEAN generation sequence of $R^{(1)}$:

$$Z^{(1)} = (z^{(1)}(2), z^{(1)}(3), \dots, z^{(1)}(n)). \quad (9)$$

Then,

$$x^{(0)}(k) + ax^{(1)}(k) = b. \quad (10)$$

It is called the grey model.

The basic form of the grey model corresponds to the whitenization equation:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b. \quad (11)$$

In this equation, the development coefficient (a) and the amount of grey action (b) can be obtained by the least squares method through statistical sequences.

The solution of the whitenization equation $dx^{(1)}/dt + ax^{(1)} = b$ (also called the time response function) is

$$\hat{r}(k+1) = \left[r^{(0)}(1) - \frac{b}{a} \right] e^{-ak} + \frac{b}{a}, \quad (k = 1, 2, \dots, n). \quad (12)$$

The time response sequence corresponding to the grey differential equation is

$$\hat{r}(k+1) = \left[r^{(1)}(0) - \frac{b}{a} \right] e^{-ak} + \frac{b}{a}, \quad (k = 1, 2, \dots, n). \quad (13)$$

From equation (13), $\hat{r}(0) = r^{(0)}(1)$. According to the $\hat{r}(k+1)$ reduction sequence, the prediction sequence of the residual sequence $\hat{R}^{(0)}$ can be obtained.

$$\begin{cases} \hat{r}(0) = r^{(0)}(1), \\ \hat{r}^{(0)}(k) = \hat{r}^{(1)}(k) - \hat{r}^{(1)}(k-1) = \left[r^{(0)}(1) - \frac{b}{a} \right] (1 - e^{-a}) e^{-a(k-1)}, \\ k = 1, 2, \dots, n. \end{cases} \quad (14)$$

The original mining safety production situation time series is decomposed into several IMF component sequences and one residual term sequence by the EMD method. The IMF component $C_j(t)$ is predicted by the GM model, and the residual $r_n(t)$ is predicted by the ARMA model [34]. The prediction results of the original sequence are used to establish the EMD-GM-ARMA model, which results from the superposition of the component prediction result.

3. Application of the EMD-GM-ARMA Model

Take every month's death rate per 10,000 people in the mining accidents from 2005 to 2019 as the research object. The data comes from the statistics of the domestic mining safety production accidents in the *Journal of Safety and Environment* and the statistical analysis of mining practitioners on the "National Data" website. The EMD-GM-ARMA model was established using 152 sets of data on the death rate per person from mining accidents from January 2005 to August 2017 as training samples. Twenty sets of data on the death rate per person from mining accidents from September 2017 to April 2019 were used as test sets to verify the reliability of the model. The statistics on the death rate of 10,000 people in the national mining safety accidents from 2005 to 2019 are shown in Figure 2.

3.1. Prediction of Time Series of Mining Safety Production Situation. The original time series is shown in Figure 2. The mining accident time series is a waveform with the change of time, so the method of time domain analysis can be adopted. In order to better describe the characteristics of the mining time series, the sequence can be expressed in frequency. It can be seen that the time series of mining accidents is unstable, but it shows a significant downward trend. Therefore, the deterministic information contained in the time series cannot be effectively extracted if it is analyzed directly. Furthermore, a single prediction model is difficult

FIGURE 2: Time series chart of death rate per 10,000 in mining safety accidents.

to predict the fluctuation trend of the mining accident time series. To solve this problem, the time series of mining accidents are decomposed into multiple intrinsic mode functions and one monotonic series via the EMD method. The sequence of mining accidents is decomposed by the EMD method as shown in Figure 3.

For the six IMF components, the ARMA method is used for prediction, and the order of each component model is determined by equation (6) using the AIC order method. The AIC result of the IMF component is represented by the heat map in Figure 4.

The grid with the smallest AIC value and the darkest color is selected as the order of the model. The order of each IMF component is shown in Table 1.

After the model is fixed, matrix programming is used to predict each IMF component. The residual (RES) trend component is predicted by the GM model. The prediction results of these two models are superimposed to obtain the total predicted results. The prediction results are shown in Figures 5–12.

The IMF components decomposed from the original stochastic volatility sequence using the EMD method are stable undulating sequences. The ARMA model has high prediction accuracy, and the RES is a monotonic sequence suitable for the grey model prediction. It can be seen from the prediction results that both the fitting precision of each component and total prediction results have high accuracy. Since January 2015, mining safety accidents have gradually shown an upward trend, leading to a lower fitting accuracy of the GM model from January 2015. The average relative error of the prediction of the EMD-GM-ARMA model is 11.45%.

3.2. Model Accuracy Test. In order to test whether the model's residual sequence obeys the normal distribution, the quantile-quantile (QQ) diagram for residual sequence of the EMD-GM-ARMA model prediction is shown in Figure 13.

It can be seen from the QQ diagram that the residuals, which are well-distributed around the straight line, fall in the

FIGURE 3: Mining safety accidents time series decomposition map by the EMD method.

95% acceptance interval. Therefore, it can be considered that the prediction residual of the EMD-GM-ARMA model is a comprehensive response independently influenced by many random phenomena, and the residual sequence satisfies a normal distribution.

Perform a Durbin–Watson (D-W) test on the model to test the first-order autocorrelation of the model and construct statistics [35, 36]:

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}, \quad (15)$$

where e_t is the residual at period t ; e_{t-1} is the residual at period $t-1$; and n is the sample size.

The D-W value of each component is shown in Table 2.

The D-W value of the prediction residuals of each IMF component is around 2.0. The D-W distribution table shows that there is no first-order autocorrelation for each IMF component, and the D-W value of the total prediction residual is 2.047751. This shows that there is no autocorrelation in the prediction results of the EMD-GM-ARMA model, indicating its better prediction effect.

4. Verification of the EMD-GM-ARMA Model

In the past, nonstationary time series often used the difference to smoothen the series so as to establish a differential autoregressive integrated moving average (ARIMA) model [37–39]. In order to compare the prediction accuracy of the EMD-GM-ARMA and ARIMA models, machine learning models (PSO-SVM) and deep learning models (wavelet neural network), all were used to train 152 sets of samples, and a total of 20 datasets from September 2017 to April 2019 were used to verify their reliability. The monthly death rate of 10,000 people in mining safety accidents is predicted. Among them, the selection order of the ARIMA model also assumes the thermal map order determination method. The selection of wavelet bases in the wavelet neural network is based on the correlation criterion and the entropy ratio criterion, which reduced the subjectivity of wavelet basis selection. The particle swarm optimization algorithm

parameters are selected by analysis of variance to analyze the search ability of the particles under different parameter conditions and optimization performance of the algorithm. On this basis, the two most important parameters of the particle swarm optimization performance inertia weight and acceleration factor are determined. The predicted values of each model are shown in Table 3.

The three absolute indicators of mean absolute error (MAE), mean relative error (MRE), and root mean square error (RMSE) were used to analyze the prediction results of each model. Among them, the MAE is used to measure the average absolute error between the predicted value and the real value on an experimental dataset. RE is the ratio of absolute error to the true value [40]. Generally, relative error can better reflect the credibility of the measurement. RMSE is more sensitive to outliers. The analysis of prediction error of each model is shown in Table 4.

The time series of mining safety accidents from September 2017 to April 2019 are predicted by the EMD-GM-ARMA model. The results showed that the MAE, MRE, and RMSE of the prediction are 0.3477%, 9.8935%, and 0.4029%, respectively, which were less than that of the contrast model. A comparative analysis of the errors of each prediction model is shown in Figure 14. The prediction effect of each model is shown in Figure 15.

It can be seen in Figure 14 that the relative error of the EMD-GM-ARMA model is lower than that of the ARIMA model, wavelet neural network model, and particle swarm optimization support vector machine model. Furthermore, the relative error fluctuation of the prediction results from the EMD-GM-ARMA model is more stable, indicating that the prediction stability of the EMD-GM-ARMA model is better. In addition, it can be seen from Figure 15 that the EMD-GM-ARMA model has a better fitting effect than other models.

5. Discussion and Future Work

From Figure 15, it can be clearly seen that the EMD-GM-ARMA model shows a better fitting effect than the ARIMA

FIGURE 4: Thermodynamic diagram. (a) IMF1AIC. (b) IMF2AIC. (c) IMF3AIC. (d) IMF4AIC. (e) IMF5AIC. (f) IMF6AIC.

TABLE 1: ARMA model parameters of IMF components.

Sequence	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6
ARMA parameter estimation	ARMA (2, 1)	ARMA (5, 4)	ARMA (5, 7)	ARMA (5, 3)	ARMA (5, 2)	ARMA (4, 2)

FIGURE 5: IMF1 predicted results.

FIGURE 7: IMF3 predicted results.

FIGURE 6: IMF2 predicted results.

FIGURE 8: IMF4 predicted results.

model. The average relative error of the EMD-GM-ARMA model is 9.8935%, the average relative error of the ARIMA model is 17.7318%, and the prediction accuracy of the EMD-GM-ARMA model is relatively high. Given that the ARIMA model uses the difference to smoothen the original sequence, it leads to the loss of the original sequence information. The more times the difference is lost, the more information is lost, and the original EMD can effectively retain the original time series information. Ahmadi et al. [41] also found a similar phenomenon in the research. Differential processing

will lose the internal information of the original time series. It is difficult to mine this part of the information during prediction, resulting in a decline in prediction accuracy. Therefore, the accuracy of the stationary time series after the EMD is significantly higher than the stationary series after difference processing, and the prediction stability of the model is also high.

Compared with the wavelet neural network, the average relative error of the EMD-GM-ARMA model is 9.8935%, which is smaller than that of the average relative error of the wavelet neural network model. Wavelet analysis is developed

FIGURE 9: IMF5 predicted results.

FIGURE 11: RES predicted results.

FIGURE 10: IMF6 predicted results.

FIGURE 12: Total prediction results.

for the insufficiency of Fourier transform, which improves the shortcoming that the Fourier transform has no resolution ability in the time domain. However, compared with empirical mode decomposition, the choice of wavelet base in wavelet neural network prediction has a strong subjectivity, which affects the prediction accuracy of the model, and similar phenomena have been found in the study. Also, neural networks often require a large number of training samples. Less training samples and uneven distribution of samples will lead to increased errors. However, the time series samples of the mining safety production situation are limited. When there are few data, the performance of deep learning algorithms is not precise. This is because deep learning algorithms require a lot of data to understand it perfectly, so the prediction effect of wavelet neural network is not ideal.

Compared with the PSO-SVM model, as can be seen from Figures 10 and 11, when the prediction step size is 8, the average relative error of the PSO-SVM model is closer to the EMD-GM-ARMA model, and the average relative errors of the EMD-GM-ARMA model and the particle swarm optimization support vector machine model are 9.1527% and 10.2834%, respectively. When the prediction step exceeds 8, the fitting effect of the particle swarm optimization support vector machine model is significantly lower than that of the EMD-GM-ARMA model. The PSO-SVM model has a good prediction effect in the short-term prediction of small samples and nonlinear sequences. But due to the lack of consideration of sequence timing, the prediction accuracy

FIGURE 13: QQ test chart.

TABLE 2: D-W test value.

Sequence	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	RES
D-W value	1.918757	1.947228	2.039361	2.025869	2.027511	1.911983	1.835048

TABLE 3: Comparison of model prediction results.

Serial number	Year	Actual value	Predicted value of EMD-GM-ARMA model	Predicted value of ARIMA model	Wavelet neural network prediction value	PSO-SVM prediction
1	September 2017	0.024155	0.026384	0.017070	0.017823	0.028374
2	October 2017	0.065876	0.070810	0.077484	0.046273	0.069373
3	November 2017	0.032938	0.035605	0.028937	0.041029	0.042813
4	December 2017	0.024155	0.026151	0.020423	0.032475	0.027364
5	January 2018	0.023524	0.025697	0.019736	0.035171	0.024485
6	February 2018	0.000000	0.008040	0.008586	0.030234	0.009230
7	March 2018	0.000000	0.008230	0.008434	0.012033	0.005392
8	April 2018	0.035286	0.039039	0.042802	0.034172	0.029305
9	May 2018	0.025876	0.028689	0.025535	0.042168	0.027293
10	June 2018	0.035286	0.037537	0.042490	0.048372	0.029484
11	July 2018	0.007057	0.007916	0.005843	0.012834	0.020293
12	August 2018	0.030581	0.035165	0.037057	0.021734	0.022388
13	September 2018	0.023524	0.025827	0.021576	0.032013	0.026344
14	October 2018	0.077629	0.082357	0.091440	0.042733	0.043823
15	November 2018	0.030581	0.034332	0.022855	0.045273	0.052723
16	December 2018	0.023524	0.026606	0.029769	0.039773	0.034785
17	January 2019	0.055000	0.058097	0.063285	0.034767	0.039374
18	February 2019	0.055000	0.060939	0.064205	0.049589	0.054989
19	March 2019	0.007500	0.008486	0.006482	0.007342	0.013546
20	April 2019	0.000000	0.0011312	0.004436	0.006273	0.008485

TABLE 4: Prediction error analysis.

Evaluation index	Contrast model			
	EMD-GM-ARMA (%)	ARIMA (%)	Wavelet neural network (%)	PSO-SVM (%)
MAE	0.3477	0.6133	1.2389	0.8560
MRE	9.8935	17.3718	29.7861	25.7764
RMSE	0.4029	0.7047	1.5085	1.1592

FIGURE 14: Relative error comparison chart of each model.

FIGURE 15: Comparative analysis of predicted values of models.

in medium long-term prediction is better than the EMD-GM-ARMA model, and in the short-term prediction, the average relative error is small, but the fluctuation is large, so the prediction stability in the short-term prediction is not so good as the EMD-GM-ARMA model.

Dhiman et al. [42] also found the subjectivity of wavelet transform and the phenomenon that SVM's accuracy becomes low and unstable in mid- and long-term prediction when using SVM built on a wavelet transform model for prediction. Through Figures 14 and 15, it can be seen intuitively that the mining safety production situation predicted by the EMD-GM-ARMA model is consistent with the real change trend. The EMD-GM-ARMA model not only considers the trend characteristics of mining safety production situation fluctuations but also considers the periodic characteristics of mining safety production situation fluctuations, that is, the long-term prediction of mining safety production situation timing and nonlinearity. The influence of factors, as in the long-term mining safety production situation prediction, achieved good prediction results. Therefore, it can be fully proved that the EMD-GM-ARMA model has good practical guiding significance in the prediction of mining safety production, which is useful to help safety decision makers accurately grasp changes in mining safety production and make correct safety decisions.

The relevant prediction results in this paper have clearly shown that the method proposed in this paper can obtain higher prediction accuracy in the time series prediction of mining safety production situation and can effectively solve the two core problems of nonstationarity and difficult to define indicators in time series prediction. However, in order to obtain higher prediction accuracy in future predictions, the next step will be to continue to carry out research in the following two aspects:

- (1) The accuracy of the time series prediction is inextricably linked to the inherent laws of the original time series. Due to the accident statistics, it is difficult to define the accident indicators, and there are phenomena such as concealment and failure to report. There will be certain statistical errors, which will lead to the weakening of inherent regularity of the time series, affecting the accuracy of the model. Therefore, the next step will be to study this. In future research, it is considered to blur the statistical accident information and to reflect the statistical error in the original time series of the mining safety production situation.
- (2) It can be seen from the relevant results in this paper that EMD can better retain the information of the original time series. But when the data are not pure white noise, some time scales in the decomposition will be lost, resulting in chaos of decomposition. Future research will focus on this aspect of the study, considering keeping the IMF mean within the normal dynamic filtering range, so that white noise cancels each other, to suppress modal aliasing and further improve prediction accuracy.

6. Conclusions

This paper presents an EMD-GM-ARMA model used for the prediction of mining safety production situation. By introducing the EMD algorithm of signal processing into the prediction of mining safety accident time series, the mining safety production situation time series is effectively decomposed into corresponding low-frequency fluctuation characteristic sequences and high-frequency trend sequences. That is, the smooth detailed sequence IMF components and the RES component of the overall change trend sequence explore the characteristics of the original sequence via studying the sequence of each component.

The EMD-GM-ARMA model was used to predict 152 sets of training samples, which perfectly fitted the original sequence. The prediction residuals were tested by the QQ test and D-W test, and the residual sequences had no autocorrelation. It showed that the EMD-GM-ARMA prediction model could extract the internal information in the mining safety production situation time series more completely. The EMD-GM-ARMA model, ARIMA model, wavelet neural network model, and PSO-SVM model were used to predict 20 sets of test samples. The relative error of wavelet neural network prediction results is the most volatile in the prediction for the small number of samples because the large number of training samples are necessary to guarantee the accuracy of deep learning methods. The relative error fluctuation of the prediction result of the EMD-GM-ARMA prediction model is the smallest, indicating that the prediction stability of the EMD-GM-ARMA model is better. It can provide relevant theoretical basis for mining safety decision-making. Besides, the average relative prediction error of the EMD-GM-ARMA model is 9.8935%, which is better than the ARIMA model of 17.3718% and the wavelet neural network of 29.7861%. When the prediction step is 8, the average relative error of the PSO-SVM is 10.2834%, which is close to the accuracy from the EMD-GM-ARMA model. When the prediction step exceeds 8, the prediction accuracy of the EMD-GM-ARMA model is significantly higher. The EMD-GM-ARMA model can be used to predict the situation of mining safety production, which has certain guiding significance for the decision-making in mining safety production.

Data Availability

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

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by the Central Guided Local Science and Technology Development Special Project of China (no. 2019ZYD060) and the National Natural Science Foundation of China (Grant no. 51704213).

References

- [1] S. C. Li and L. Xiao, "Statistical analysis of domestic production safety accidents from November to December 2018," *Journal of Safety and Environment*, vol. 19, no. 1, pp. 363–364, 2019.
- [2] M. Gul, M. F. Ak, and A. F. Guneri, "Pythagorean fuzzy VIKOR-based approach for safety risk assessment in mine industry," *Journal of Safety Research*, vol. 69, no. 69, pp. 135–153, 2019.
- [3] A.-S. Chen, H.-C. Chang, and L.-Y. Cheng, "Time-varying variance scaling: application of the fractionally integrated ARMA model," *IET Signal Processing*, vol. 47, no. 13, pp. 1–12, 2019.
- [4] J. Olamaee, M. Mohammadi, A. Noruzi, and S. M. H. Hosseini, "Day ahead price forecasting based on hybrid prediction model," *Complexity*, vol. 2, no. 5, pp. 156–164, 2016.
- [5] X. Yu, "Data fusion algorithm based on unbiased grey Markov forecasting WSN," *Journal of Transducer Technology*, vol. 31, no. 8, pp. 1266–1269, 2018.
- [6] F. Xie, "Time series model analysis of non-coal mine safety accidents," *Chemical Minerals and Processing*, vol. 43, no. 6, pp. 31–33, 2014.
- [7] N. Li, L. G. Wang, and M. T. Jia, "Rock burst prediction based on rough set theory and support vector machine," *Journal of Central South University*, vol. 48, no. 5, pp. 1268–1275, 2017.
- [8] Z. Wang and Y. Guo, "Non-coal underground mine safety evaluation model based on BP neural network," *Chinese Journal of Safety Science*, vol. 19, no. 2, pp. 124–128, 2009.
- [9] J.-Y. Lan and Y. Zhou, "Application of Gray Markov SCGM (1, 1) C model to prediction of accidents deaths in coal mining," *International Scholarly Research Notices*, vol. 2014, Article ID 632804, 7 pages, 2014.
- [10] S. B. Liu, Z. H. Zhao, and C. W. Wu, "Research on mine safety accident prediction based on grey support vector machine," *Journal of Yunnan University for Nationalities*, vol. 23, no. 4, pp. 270–272, 2014.
- [11] J. Bao, J. Zhang, and S.-P. Shi, "A projection analysis of C2R models applied in grey prediction on OHASA18001 safety investment and benefit of a mining industry," *Filomat*, vol. 3, no. 32, pp. 1581–1587, 2018.
- [12] Y. L. Wang, "Prediction of coal mine gas accidents based on time series-Markov model," *China Mining*, vol. 26, no. 12, pp. 179–183, 2017.
- [13] D. Q. Li, *Research on Safety Situation Prediction of Mining System Based on Environmental Impact*, Wuhan University of Science and Technology, Wuhan, China, 2014.
- [14] Y. Jang, I. Jeong, and Y. Cho, "Business failure prediction of construction contractors using a LSTM RNN with accounting," *Construction Market, and Macroeconomic Variables*, vol. 36, no. 2, 2020.
- [15] Y. Wang and G. Jiang, "Coal mine safety risk prediction based on RS-SVM combined model," *Journal of China University of Mining and Technology*, vol. 46, no. 2, pp. 423–429, 2017.
- [16] Y. Zhang, C. Zhang, J. Sun, and J. Guo, "Improved wind speed prediction using empirical mode decomposition," *Advances in Electrical and Computer Engineering*, vol. 18, no. 2, pp. 3–10, 2018.
- [17] X. Y. Li, "Research on ARMA-GM combined time series model for bridge structure state prediction," *Journal of Chongqing Jiaotong University*, vol. 35, no. 4, pp. 6–9, 2016.
- [18] N. A. Agana and A. Homaifar, "EMD-based predictive deep belief network for time series prediction: an application to drought forecasting," *Hydrology*, vol. 5, no. 1, p. 18, 2018.
- [19] B. N. Ghimire, "Application of ARIMA model for river discharges analysis," *Journal of Nepal Physical Society*, vol. 4, no. 1, pp. 27–32, 2017.
- [20] Z.-X. Wang and Q. Li, "Modelling the nonlinear relationship between CO₂ emissions and economic growth using a PSO algorithm-based grey verhulst model," *Journal of Cleaner Production*, vol. 207, pp. 214–224, 2019.
- [21] L. Tasci and M. Tuncez, "Monitoring of deformations in open-pit mines and prediction of deformations with the grey prediction model," *The Journal of Grey System*, vol. 4, no. 30, pp. 152–157, 2018.
- [22] K. Long, "Mine safety hidden dangers inspection platform based on network information security technology management," *Nonferrous Metals of the World*, vol. 3, no. 14, pp. 85–86, 2019.
- [23] B.-s. Oh, H. Zhuang, K.-A. Toh, and Z. Lin, "Augmented EMD for complex-valued univariate single," *North American Journal of Economics and Finance*, vol. 13, no. 4, pp. 424–433, 2019.
- [24] W.-B. Wang and Y.-C. Hu, "Multivariate grey prediction models for pattern classification irrespective of time series," *The Journal of Grey System*, vol. 3, no. 1, pp. 135–141, 2019.
- [25] C. Velasco and I. N. Lobato, "Frequency domain minimum distance inference for possibly noninvertible and noncausal arma models," *Annals of Statistics*, vol. 46, no. 2, pp. 555–579, 2018.
- [26] Z. M. Yaseen, M. Fu, C. Wang, W. H. M. W. Mohtar, R. C. Deo, and A. El-shafie, "Application of the hybrid artificial neural network coupled with rolling mechanism and grey model algorithms for streamflow forecasting over multiple time horizons," *Water Resources Management*, vol. 32, no. 5, pp. 1883–1899, 2018.
- [27] R. C. Bai, "Prediction of open source mine truck failure rate based on Mallat algorithm and ARMA model," *Journal of Safety Science and Technology*, vol. 28, no. 10, pp. 31–37, 2018.
- [28] F. Xu, X. Song, K.-L. Tsui, F. Yang, and Z. Huang, "Bearing performance degradation assessment based on ensemble empirical mode decomposition and affinity propagation clustering," *IEEE Access*, vol. 7, pp. 54623–54637, 2019.
- [29] J. L. Wescoat, A. Siddiqi, and A. Muhammad, "Socio-hydrology of channel flows in complex river basins: rivers, canals, and distributaries in Punjab, Pakistan," *Water Resources Research*, vol. 54, no. 1, pp. 464–479, 2018.
- [30] K. Li and X. Mu, "Containment control of stochastic multi-agent systems with semi-Markovian switching topologies," *International Journal of Robust and Nonlinear Control*, vol. 29, no. 14, pp. 4943–4955, 2019.
- [31] H. Lennon and J. Yuan, "Estimation of a digitised Gaussian ARMA model by Monte Carlo expectation maximisation," *Computational Statistics & Data Analysis*, vol. 133, no. 3, pp. 277–284, 2019.
- [32] L. Wu and H. Zhao, "Discrete grey model with the weighted accumulation," *Soft Computing*, vol. 23, no. 23, pp. 12873–12881, 2019.
- [33] C. L. Zheng, W. Z. Wu, and J. M. Jiang, "Forecasting natural gas consumption of china using a novel grey model," *Complexity*, vol. 2020, Article ID 3257328, 9 pages, 2020.
- [34] C. L. Zhao and Z. H. Yu, "Comparison of GM and ARMA model for survivability prediction in network information system," *Advanced Materials Research*, vol. 846–847, no. 46, pp. 1648–1651, 2014.

- [35] F. Proïa, "Testing for residual correlation of any order in the autoregressive process," *Communications in Statistics-Theory and Methods*, vol. 47, no. 3, pp. 628–654, 2018.
- [36] Z. Wang, J. Sun, G. Chen, and J. Chen, "Stability analysis of switched nonlinear delay systems with sampled-data inputs," *International Journal of Robust and Nonlinear Control*, vol. 29, no. 14, pp. 4700–4715, 2019.
- [37] L. Ma, C. Hu, R. Lin, and Y. Han, "ARIMA model forecast based on EVIEWS software," *IOP Conference Series: Earth and Environmental Science*, vol. 208, p. 012017, 2018.
- [38] C.-H. Hsiang, "Forecasting a point process with an ARIMA model," *Communications in Statistics-Theory And Methods*, vol. 45, no. 17, pp. 5066–5076, 2016.
- [39] Y. Wang, C. Wang, C. Shi, and B. Xiao, "Short-term cloud coverage prediction using the ARIMA time series model," *Remote Sensing Letters*, vol. 9, no. 3, pp. 274–283, 2018.
- [40] S. H. Xiao, X. P. Xiao, M. Y. Gao, and X. Wang, "Nonlinear fractional order grey model of urban traffic flow short term prediction," *The Journal of Grey System*, vol. 30, no. 4, pp. 1–17, 2018.
- [41] R. Ahmadi, J. Shahrabi, and B. Aminshahidy, "Forecasting multiple-well flow rates using a novel space-time modeling approach," *Journal of Petroleum Science and Engineering*, vol. 191, Article ID 107027, 2020.
- [42] H. S. Dhiman, D. Deb, and J. M. Guerrero, "Hybrid machine intelligent SVR variants for wind forecasting and ramp events," *Renewable and Sustainable Energy Reviews*, vol. 108, pp. 369–379, 2019.