





Research Article

Improvement towards Prediction Accuracy of Principle Mineral Resources Using Threshold

Maria Qurban,¹ Mohammed M. A. Almazah ,^{2,3} Hafiza Mamona Nazir,⁴ Ijaz Hussain ,¹ Muhammad Ismail ,⁵ Faud S. Al-Duais ,^{6,7} Sana Amjad,⁸ and Mohammed N. Murshed⁹

¹Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan

²Department of Mathematics, College of Sciences and Arts (Muhiyl), King Khalid University, Muhiyl 61421, Saudi Arabia

³Department of Mathematics and Computer, College of Sciences, Ibb University, Ibb 70270, Yemen

⁴Department of Statistics, Government College Women University, Sialkot, Pakistan

⁵Department of Statistics, COMSATS University Islamabad, Lahore Campus, Lahore, Pakistan

⁶Mathematics Department, College of Humanities and Science, Prince Sattam Bin Abdulaziz University, Al Aflaj, Saudi Arabia

⁷Administration Department, Administrative Science College, Thamar University, Thamar, Yemen

⁸Department of Management Sciences, National University of Modern Languages, Lahore, Pakistan

⁹Department of Physics, College of Sciences and Arts (Muhiyl), King Khalid University, Muhiyl 61421, Saudi Arabia

Correspondence should be addressed to Mohammed M. A. Almazah; mmalmazah@kku.edu.sa and Ijaz Hussain; ijaz@qau.edu.pk

Received 11 November 2021; Revised 16 January 2022; Accepted 2 February 2022; Published 12 March 2022

Academic Editor: Firdous Khan

Copyright © 2022 Maria Qurban 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.

The production data of mineral resources are noisy, nonstationary, and nonlinear. Therefore, some techniques are required to address the problem of nonstationarity and complexity of noises in it. In this paper, two hybrid models (EMD-CEEMDAN-EBT-MM and WA-CEEMDAN-EBT-MM) flourish to improve mineral production prediction. First, we use empirical mode decomposition (EMD) and wavelet analysis (WA) to denoise the data. Second, ensemble empirical mode decomposition (EEMD) and complete ensemble empirical mode decomposition (CEEMDAN) are used for the decomposition of nonstationary data into intrinsic mode function (IMF). Then, empirical Bayesian threshold (EBT) is applied on noise dominant IMFs to consolidate noises, which are further used as input in the data-driven model. Next, other noise-free IMFs are used in the stochastic model as input for the prediction of minerals. At last, the predicted IMFs are ensemble for final prediction. The proposed strategy is exemplified using Pakistan's four major mineral resources. To measure the prediction performance of all the models, three methods, that is, mean relative error, mean square error, and mean absolute percentage error, are used. Our proposed framework WA-CEEMDAN-EBT-MM has shown improvement with minimum mean absolute percentage error value compared to other existing models in prediction accuracy for all four minerals. Therefore, our proposed strategy can predict the noisy and nonstationary time-series data with an efficient mechanism. Hence, it will be helpful to the policymakers for making policies and planning in mineral resource management.

1. Introduction

The industrial and economic development of any country is mainly based on mineral resources as those are among the most important natural resources of that country. Mineral resources provide the raw material to the different sectors of the country. It is said with considerable truth that the accessibility of essential minerals is one of the decisive factors in war. A strong positive correlation exists between

economic growth and natural mineral resources wealth reported by Dollar and Kraay [1]. So, accurate prediction of minerals is needed as it also plays a significant role in its economic development. For prediction purposes, many algorithms are available from the study of Aichouri et al. [2]; Ch et al. [3]; Cheng et al. [4]; Lapedes and Farber [5]; Solomatine and Ostfeld [6]. Many authors widely use the Box-Jenkins technique from the study of Tang et al. [7] in literature, that is, in autoregressive model (AR),

autoregressive moving average (ARMA) model, autoregressive integrated moving average (ARIMA) model, and many other models, but its drawback is that it considered the only linear and stationary behavior of the given process box [8]. On the other hand, data on the production of mineral resources are nonlinear, noisy, and nonstationary. The noisy and nonstationary characteristics of mineral resources make the prediction a challenging task.

The development of data-driven models makes it easy to deal with nonstationary and nonlinear time-series data from the study of Lapedes and Farber [5]. The data-driven models are further categorized as traditional statistical and machine learning (ML) models. The traditional statistical methods, that is, autoregressive integrated moving average (ARIMA), only consider the stationary and linear data. ARIMA model is successfully applied to predict the production of mineral resources [9]. Lapedes and Farber [5] compared neural network and conventional methods based on two time series having no noises and concluded that neural network performs better than traditional methods, many times by magnitude's order. Huang et al. [10] proposed the extreme learning machine (ELM) algorithm using a single hidden layer feed-forward neural network that randomly selects the input weights and then produces the output weights analytically [11]. They concluded that the ELM algorithm produced better results than traditional methods. Yaseen et al. [12] used an improved version of ELM in forecasting, exhibited its variability, and concluded that ELM with an enhanced version produced better results than traditional statistical methods. But the characteristics of varying time and noise of minerals production data are ignored by data-driven models. These drawbacks of data-driven models inhibit the researcher from predicting the data accurately.

To overcome the drawbacks of a data-driven model, hybrid models are introduced that capture the characteristics of varying times and reduce the noises, which eventually improves the accuracy of prediction from the study of Nourani et al. [13]; Pramanik et al. [14]; Yaseen et al. [15]. Hybrid models are the combination of some preprocessing techniques, that is, wavelet analysis (WA), empirical mode decomposition (EMD), and ensemble empirical mode decomposition (EEMD), with data-driven models. An advantage of hybrid models over data-driven models is that they decompose the data into the frequency components and remove the noises.

Several algorithms are developed, such as spectral analysis, WA, Fourier analysis, and EMD, to reduce these noises or stochastic volatiles from the data [16]. Fourier analysis and spectral analysis are used for those kinds of data that are stationary or linear. However, EMD and WA are the most commonly used preprocessing algorithms for nonlinear or nonstationary data and provide better results. The algorithms of WA decompose the nonlinear and nonstationary data of mineral resources into multiscale components [17]. These components are used as inputs at the prediction stage, and then, these predicted components are ensemble for final prediction. The present paper uses EMD and WA-based thresholds to reduce noises from the mineral production data.

The WA has developed a powerful tool for converting a signal into a stationary signal with specific effectiveness. There are many articles in the literature in which different hybrid models with wavelet decomposition are used to predict different kinds of nonlinear and nonstationary time-series data [18, 19]. Singh et al. [19] proposed hybrid models based on the discrete wavelet decomposition which can partition the data into series. The prediction indicates a sharp rise in death rate compared to the simple ARIMA model.

Wu et al. [20] used data preprocessing techniques coupled with data-driven models to predict the monthly streamflows and concluded that the hybrid techniques provide better prediction than traditional models. Azadeh et al. [21] described the effectiveness of the preprocessing techniques to enhance the precision of data-driven models by considering many data preprocessing techniques and concluded that simple statistical models and neural networks with preprocessing methods could efficiently forecast the nonlinear data. Asadi et al. [22] proposed a hybrid model by utilizing the preprocessing techniques coupled with neural networks to predict the runoff process and concluded that the proposed hybrid model predicted the runoff process better than artificial neural network (ANN) and neuro-fuzzy inference system (ANFIS) models. However, the performance of the WA depends upon the selection of the type of mother wavelet. Prior knowledge about the signal, which is to be analyzed, and prior knowledge about its frequency content, is needed for a suitable choice of the mother wavelet.

Huang et al. [23] proposed an EMD method to overcome the shortcomings of WA for scrutinizing the nonlinear data and nonstationary datasets. Complex time-series data can be decomposed into a small and finite number of IMFs by using EMD. The EMD strategy has the advantage of converting the nonstationary series into stationary series. Di et al. [24] proposed a four-stage hybrid model using EMD, EEMD, and WA techniques to remove the noise from the time series with soft and hard thresholds to find denoised time-varying information that decreases the complexity of hydrological series coupled with RBFNN for the prediction purpose. They concluded that the performance of their proposed hybrid model is better than conventional single-stage and other hybrid models without preprocessing techniques. Different studies exist in the literature that used EMD with different data-driven models such as EMD-ANN, EMD-radial basis function (EMD-RBF), EMD-support vector machines (EMD-SVM), EMD-relevant vector machine (EMD-RVM), EMD-ARIMA; these hybrid models improve the prediction accuracy from the study of Huang et al. [25]; Liu et al. [26]; Wang et al. [27]. The EMD is combined with ANN in many past studies, especially in hydrology from the study of Liu et al. [28], and also a novel model based on EMD and deep learning is used by Mi et al. [29] to reduce the noises and extract the information of trend of the original data of wind speed. Ruiz-Aguilar et al. [30] proposed a hybrid model by combining a preprocessing technique, EMD, an information-based method, the permutation entropy (PE), and a machine learning technique using an ensemble learning methodology for wind speed prediction.

These preprocessing techniques have their own shortcomings in the extraction of the optimal multiscale components, that is, the WA-based denoising technique depends upon the choice of a mother wavelet function, which may create problems and decrease its performance. EMD technique suffered from mode mixing problem although it is purely a data-driven technique. Because of the mode mixing problem, it provides bogus information about the time frequency. An improved version of EMD was introduced, which added the white Gaussian noise to solve the mode mixing problem. The EEMD method can separate the signals without inappropriate mode mixing. It uses white noise that helps to establish the dyadic reference frame on time-scale space. Many hybrid techniques based on EEMD are used for streamflow and wind speed prediction and in hydrology from the study of Niu et al. [31]; Santhosh et al. [32]. Although the EEMD technique of denoising has proved effective, it has its drawback as it may not be influential in the extraction of powerful IMFs through average. To tackle and improve the deficiency of the EEMD technique and to cope with the simple averaging problem of EEMD, complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN), Torres et al. [33] proposed CEEMDAN [34], which was used by Jun et al. [35] to decompose the given time-series data into frequency components which were further used for prediction purposes. Also, Jiang and Zhou [36] used CEEMDAN with Wigner-Ville distribution (WVD) to decompose and analyze the nonstationary signals of hydro-turbine. Dai et al. [37] used the CEEMDAN algorithm to predict the daily peak load in their proposed model. The proposed three-stage hybrid model comprises the CEEMDAN technique, which showed a robust decomposed ability of reliable prediction. Wang et al. [38] proposed a hybrid model based on CEEMDAN, detrended fluctuation analysis (DFA), and improved wavelet thresholding. In their proposed model, a denoising method of CEEMDAN-DFA-improved wavelet threshold function was presented to reduce the distortion of the noised signal. Qurban et al. [39] used CEEMDAN in combination with multimodels to bring improvements in the prediction of the production of mineral resources. They concluded that the CEEMDAN technique provides excellent performance. Johnstone and Silverman [40] used the empirical Bayesian threshold (EBT) with WA to denoise the multiscale components obtained from WA. They said that EBT could efficiently tackle the problem of noises by taking different priors for each level. Nazir et al. [41] used EBT with CEEMDAN to decompose the river flow time-series data. They concluded that their proposed model is efficient to predict the nonstationary and nonlinear noisy time-series data. It is hoped that such methods can be used to denoise the nonstationary and noisy data, enhancing the prediction accuracy.

This study aimed to develop a CEEMDAN-based hybrid model coupled with the EBT technique. Here, we consider the EBT as the threshold as it is a purely data-based technique, and it optimally reduces noises from IMFs, which are then used to predict mineral resources at the prediction stage. Furthermore, the current study explores its prediction

performance using a 33-step-ahead prediction by considering this emerging hybrid modeling technique. Based on the outlined above, this paper is focused on developing an improved model to improve the prediction of mineral resources production, hybrid models with CEEMDAN-EBT-based multimodels (EMD/WA-CEEMDAN-EBT-MM).

The remainder of this paper is organized as follows: Section 2 focuses on the motivations behind the proposed study to predict the production of mineral resources. A short review of models used for mineral data predictions and an introduction to EMD, EEMD, EBT, and their modified versions are also discussed in Section 2. Additionally, a short review of various approaches used in hybrid CEEMDAN-EBT models to select the appropriate prediction methods considering the characteristics of respective IMFs is also discussed in Section 2. The application of the proposed hybrid models, and data description is presented in Section 3. Section 4 describes details of the study area and data. Finally, Section 4 presents and discusses the case study results, while conclusions are made for this research in Section 5.

Accurate prediction of mineral resources has become a challenging task for researchers in recent years. Although Pakistan is blessed to have abundant mineral resources, it is still facing an alarming situation as its power generation is based on foreign exchange. Therefore, there is a need for accurate and improved prediction of the production of mineral resources and strong management and policymakers. The mineral sector of Pakistan is dominated by four principal minerals: gas, oil, coal, and gypsum. Therefore, there is a need to analyze and improve the accurate prediction of the production of these four major minerals to deal with emerging challenges.

The primary purpose of accurate prediction of mineral production data is to get the efficient and optimum utilization of natural resources in the development of the economy by dealing with the nonstationarity and nonlinearity of the data. Due to the nonlinearity and nonstationarity of mineral resources time-series data, statistical models usually cannot achieve satisfactory results by directly conducting predictions. Instead, there is a simple but also effective way called the divide and conquer rule, which means decomposing the complex data into simple components and extracting each component's relevant features for future prediction. Many researchers are working to deal with and improve different kinds of nonstationary data, having complex time-varying characteristics. These researchers work with different motivations, but they have the same goals.

Li et al. [42] improved CEEMDAN preprocessing technique to decompose the complex oil prices into different components. Each component is forecasted using ridge regression, and differential evolution is used to optimize the regularization item. They concluded that their experimental results showed that their proposed strategy achieved better results than other states of arts. Nazir et al. [41] used CEEMDAN with EBT to tackle the multiscale and noise complexity of hydrological time-series data, which decompose the nonstationary data into different noise

dominating and noise-free IMFs. They concluded that their proposed model provides efficient prediction results for nonstationary time-series data compared to HT, ST, and ITF by using different evaluation criteria.

2. Proposed Method

Here, an improved framework constitutes four stages, that is, decomposition, denoising using novel thresholds, prediction, and ensemble. For handling nonlinear and nonstationary data, EMD-CEEMDAN and WA-CEEMDAN techniques are used for decomposition to obtain IMF. After obtaining the IMFs, they are divided into two parts; one consists of the noisiest IMFs, and the other consists of noise-free IMFs. The noisiest IMFs contain sparsity and errors. Therefore, the noisiest IMFs are combined with EBT thresholds for removing the sparsity from the noisiest IMFs to denoise them. Then, for prediction purposes, the complex data-driven and simple stochastic models are used to predict the denoised and noise-free IMFs, respectively. Here, the denoised IMFs are predicted using data-driven models, and the remaining noise-free IMFs are predicted using stochastic methods.

At last, all the predicted IMFs are merged to obtain the eventual prediction. Here, we represent the improved novel framework of getting the multiscale IMFs by using the method of EMD-CEEMDAN-EBT and WA-CEEMDAN-EBT, which are combined with the optimal method of denoising. Here, in this study, EBT plays a crucial role in predicting the production data of minerals resources. For the convenience of the reader, the proposed technique is labeled as EMD/WA (denoised), CEEMDAN (decomposed), EBT (denoised using threshold), and MM (multimodels for prediction), that is, EMD-CEEMDAN-EBT-MM and WA-CEEMDAN-EBT-MM whose complete structure is described in Figure 1.

2.1. Denoising and Decomposition of IMFs. Empirical Mode Decomposition (EMD): Huang et al. [23] introduced the multiresolution technique for decomposing nonlinear and nonstationary series, called EMD.

From the study of Yang and Chen [43], the main steps of the EMD for an original time series $p(x)$, ($x = 1, 2, 3, \dots, N$) are as follows:

- (1) All the local extrema of the original time series $p(x)$ are identified
- (2) By using cubic spline interpolation, the upper and lower envelope is created as $U(x)$ and $G(x)$, respectively.
- (3) The mean value of upper envelope and lower envelope is estimated, that is, $m(x) = (U(x) + G(x))/2$.
- (4) In this step, the difference of the mean envelope from the original signal $x(p)$ is found. The difference $d(x)$ is calculated as $d(x) = p(x) - m(x)$.
- (5) The properties of the difference of mean envelope and signal $d(x)$, that is, are examined.

- (a) If (a) and (b) conditions are satisfied by the difference $d(p)$, then it is denoted as i^{th} IMF, and also, signal $x(p)$ is replaced by residue $e(x) = p(x) - d(x)$. The i^{th} IMF is symbolized with $C_i(x)$, and i is the order of the IMF.
- (b) Replace $p(x)$ with $d(x)$ if $d(x)$ is not an IMF.

- (6) Repeat 1–5 steps before the number of extrema are less than or equal to one so that no more IMF can be extracted from or residue $e(x)$ becomes a monotonic function.

At last, the signal can be shown as the sum of all the IMFs and residue $e(x)$, where n is the number of IMFs, $C_i(x)$ ($i = 1, 2, 3, \dots, n$) is the i^{th} IMF, and $e(x)$ is the residue (Tang et al., 1991). The way of denoised IMF is the same as mentioned in steps (1)–(5), except the last two because of the low frequencies IMFs which are entirely used without denoising from the study of Qurban et al. [39].

$$p(x) = \sum_i^n d_i + e(x), \quad (1)$$

where n shows the number of sifted IMF's as ($i = 1, 2, 3, \dots, n$), $e(x)$ is the trend of the signal, and $d_i(x)$ is the i^{th} IMF. Although it is a useful technique for noise reduction, it cannot accurately get the true information for the mode mixing problem. Mode mixing is defined as the phenomenon of the appearance of similar or different scales in the same mode. The closeness of the frequency of the noisy signal creates the problem of mutual energy penetration, and then, EMD cannot accurately separate the noise and useful signals.

2.2. Ensemble Empirical Mode Decomposition (EEMD). To improve the EMD and mitigate the mode mixing, EEMD is developed by N. E. Huang et al. [23]. EEMD is an effective tool for reducing noise and provides accurate results. The original signal includes the white noise and repeatedly calculating the mean of the IMFs. According to Zhang et al. [44] and Wei et al. [45], the procedure presented is as follows:

- (1) Initialization of the ensemble number Y
- (2) Set the amplitude of the added white noise, $i = 1$
- (3) Add the random white noise signal $wn_i(p)$ in the original signal $p(x)$

$$p_i(x) = p(x) + wn_i(x), \quad (2)$$

where $wn_i(x)$ is the i^{th} added series of white noise, and $p_i(x)$ denotes the i^{th} added noise signal $i = 1 \sim Y, Y > 1$.

- (4) By using EMD, decompose the noise signal $p_i(x)$ into N IMFs $C_{j,i}(x)$ ($j = 1, 2, 3, \dots, X$), where $C_{j,i}(x)$ shows the j^{th} IMF of the i^{th} noise signal and X is the total number of IMFs

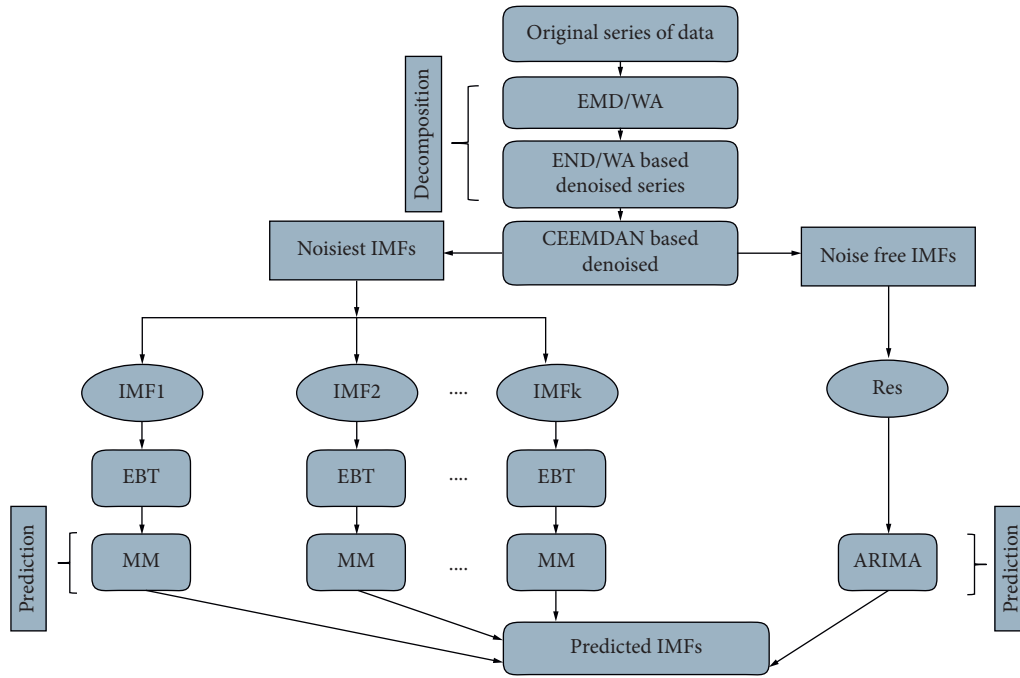


FIGURE 1: Proposed EMD/WA-CEEMDAN-EBT-MM structure for the prediction of mineral production.

- (a) Connects the maximal points of the signal with cubic splines after localizing them. The signal's upper envelope is denoted by $U(x)$
- (b) Connects the minimal points of the signal with cubic spline after localizing them. The signal's lower envelope is denoted by $G(x)$
- (c) To obtain the pro-IMF, $s_1(p)$ subtract the mean envelope $m_1(p)$ from the original $x_1(p)$, that is, $s_1(x) = p_1(x) - m_1(x)$ where $m_1(x) = (U(x) + G(x))/2$
- (d) Consider $s_1(x)$ as a new signal if the average of the lower and upper envelope becomes zero and if the number of zero-crossing and number of extrema are equal or almost one
- (e) Consider the points (a)-(d) of 4th step as far as the resulting signal as proper IMF $C_1(x)$
- (f) From the original signal, $p_1(x)$ subtract the resulting IMF $C_1(x)$. Consider the residue $e_1(x)$ as the new data and go back to Step 1.

$$p_1(x) = C_1(x) + e_1(x). \quad (3)$$

- (g) In the substep, if the residue becomes the monotonic function, complete the algorithm, and if $i < Y$, then go back to Step 3. The last residual is treated as the trend.

$$p_1(x) = \sum_{j=1}^x C_1(x) + e(x). \quad (4)$$

- (5) Estimate the ensemble mean $IMF_j(x)$ of all trials of each IMF.

$$IMF_j(x) = \frac{1}{Y} \sum_{i=1}^Y C_{j,i}(x) \quad j = 1, 2, 3, \dots, X \quad i = 1, 2, 3, \dots, Y. \quad (5)$$

- (6) Consider the mean $IMF_j(x)$ as the final mean of all the P IMFs.

In the final mean of the interrelated IMFs, the added white noise cancels each other apparently, and mean IMFs remain inside the natural dyadic filter windows, which significantly truncate the hazard of mode mixing and sustain the dyadic property. Therefore, this process can reduce the problem of mode mixing significantly and repeat the fundamental development in the original EMD. However, although EEMD can bring down the problem of mode mixing to a certain degree with added white noise sequence, an error cannot be eliminated after computation, the averaging to a finite number. Moreover, it affects the sequence of reconstruction.

2.3. Complete Ensemble Empirical Mode Decomposition (CEEMDAN). Consequently, Flandrin et al. [46] proposed CEEMDAN by considering the previous studies of EEMD. CEEMDAN adds the adaptive white noise to eliminate the mode mixing, smoothing pulse interference in decomposition. To make the decomposition of the data more complete, it uses the properties of the mean Gaussian white noise, whose mean is zero. The detailed procedure of the CEEMDAN from the study of Qian et al. [47] is as follows:

- (1) Persistent with EEMD, in the computation of CEEMDAN, P times decompose the original signal $p(x)$, that is, $p(x) + r_i w n_i(x)$, where r_i is the

parameter which controls the signal to noise ratio. The first component of IMF is as follows:

$$IMF_1(x) = \frac{1}{x} \sum_{i=1}^x c_{j,1}(x). \quad (6)$$

The residual of the signal is

$$e_1(x) = p(x) - IMF_1(x). \quad (7)$$

- (2) The $d(x)$ is defined as the l^{th} IMF component obtained by EMD. To get the second IMF component, the sequence $v_1(x) + r_1 d_1(x)(n_j(x))$ is decomposed as follows:

$$IMF_2(x) = \frac{1}{x} \sum_{j=1}^x d_1(e_1(x) + r_1 d_1(n_j(x))). \quad (8)$$

The second residual signal is

$$e_2(x) = e_1(x) - IMF_2(x). \quad (9)$$

Similarly, by following the above procedure, the expression of the l^{th} the residual signal is as follows:

$$e_l(x) = e_{l-1}(x) - IMF_l(x). \quad (10)$$

The expression of the $l + 1^{th}$ residual signal is

$$IMF_{l+1}(x) = \frac{1}{x} \sum_{j=1}^x d_l(e_l(x) + r_l d_l(n_j(x))). \quad (11)$$

Till the requirement and need of the lending criteria, the above procedure is repeated. The expression of the original sequence, if the number of IMF components is M , is as follows:

$$p(x) = \sum_{i=1}^M IMF_i(x) + e(x), \quad (12)$$

where $IMF_i(x)$ is the i^{th} IMF, $e(x)$ is the overall residual signal, and $p(x)$ is the signal obtained after decomposition. The selection of ensemble number and size of white noise is still an issue. Here, in this study, the ensemble size is set to be 900, and the standard deviation is set to be 0.2 from the study of Zhang et al. [48].

2.4. Identification of Noisy and Noise-free IMFs. After getting all IMFs through EMD, EEMD, and CEEMDAN, the screening of these IMFs started in the next step into noisy IMFs with high frequencies and noise-free IMFs which have low frequencies. The IMFs with high frequencies are noise-corrupted IMFs, and the low-frequency IMF is noise-free from the study of Wei et al. [45]. We calculated the crosscorrelation between all IMFs and original production data of mineral resources for the screening process. The low crosscorrelation value specifies that the high-frequency IMFs contain noises, and the high correlation shows that the low-frequency IMFs are noise-free. After identifying the noisy and noise-free IMFs, some thresholds are applied to the noisy IMFs to make them noise-free.

2.5. Denoising of Noisy IMFs through Thresholds. After decomposing and screening process of IMFs, the noisy IMFs are decomposed using appropriate thresholds. The main purpose of selecting the appropriate estimator is to find the optimal threshold value as the lowest, and maximum threshold values introduced the bias in prediction. For extracted IMF, the empirical Bayesian threshold estimator is used for denoising the noisy IMFs. In addition, latter soft and hard, and improved threshold functions are also used to denoise the noisy IMFs compared with EBT. The detailed procedure of EBT is described below.

2.6. Empirical Bayesian Threshold (EBT). EBT is used to estimate the noises and sparsity from the IMF from the study of Wei et al. [45] after decomposition. The idea of EBT was inspired by the wavelet denoising method. The first step of implementing the EBT is transformed to select the prior distribution for sparsities and noises. The data follow the normal distribution with mean ϑ_0 and variance one after scaled transformation. Then, for ϑ_0 , the mixture priors are selected as follows:

$$f_{\text{prior}}(\vartheta) = (1 - u)\delta_0(\vartheta) + u\gamma(\vartheta), \quad (13)$$

where $\delta_0(\vartheta)$ and $\gamma(\vartheta)$ are the zero part and density of the part other than zero of the scaled data, respectively. The prior density must be carefully selected; it should belong to a family of distributions whose parameters may be estimated using the maximum likelihood method for estimation of the parameters and weights of the mixture prior distributions. The main reason for using the maximum likelihood method to estimate the parameters is that the maximum likelihood method estimates the unknowns to be proportional to the maximum likelihood function from the study of Hossain et al. [49]. After the estimation procedure, the median of the posterior distribution $\tilde{\vartheta}_i(IMF, u)$ is estimated by using a mixture of the prior distribution. The posterior median is calculated as follows:

$$\tilde{F} * (\mu|IMF)d\mu. \quad (14)$$

The above-mentioned posterior median is used as a rule of a threshold for $\tilde{\mu}$. Generally, according to the estimation rule composed on $\zeta(IMF, x)$ specified for $u > 0$ is a thresholding condition $\Leftrightarrow \forall x > 0$, $\zeta(IMF, x)$ is an asymmetric and increasing kind of function of data and $\zeta(IMF, x) = 0 \Leftrightarrow \forall |IMF| \leq x$ where x is the median value which is estimated by using (14).

2.7. Existing Thresholds for Comparison. For comparison of EBT with other thresholding techniques, soft threshold (ST), hard threshold (HT), and improved threshold function (ITF) are used to decompose the nonlinear and noisy data from the study of Chang and Vetterli [50]; Jansen and Bultheel [51]; Jeng et al. [52]; Om and Biswas [53]. The mathematical expressions for soft, hard, and improved thresholds are given as follows:

$$\begin{aligned} \dot{c}_{x,l} &= \begin{cases} d_{x,l} & |c_{x,l}| \geq Th_l \\ 0 & |c_{x,l}| < Th_l \end{cases}, \\ \dot{c}_{x,l} &= \begin{cases} \text{sgn}(c_{x,l})(|c_{x,l}| - Th_l) & |c_{x,l}| \geq Th_l \\ 0 & |c_{x,l}| < Th_l \end{cases}, \end{aligned} \quad (15)$$

and

$$\dot{c}_{x,l} = \begin{cases} \text{sgn}(c_{x,l})|c_{x,l}| - \frac{Th_l}{[\alpha(|c_{x,l}| - Th_l)/Th_l]}, & |c_{x,l}| \geq Th_l \\ 0, & |c_{x,l}| < Th_l \end{cases}, \quad (16)$$

where Th_l is the threshold which is calculated as $Th_l = a\sqrt{2E_l \ln(N)}$, $l = 1, 2, 3, \dots, L$, where a is constant which considers the values between 0.4 and $\tilde{M}d_l$ is the median deviation, that is, $\tilde{M}d_l = \text{median}(\{c_{l-1,l} | l = 1, 2, 3, \dots, 2^{l-1} - 1\})/0.6745$.

2.8. Prediction and Ensemble. Because of the nonlinearity and nonstationarity of the production data of mineral resources, the traditional statistical techniques are not sufficient for capturing these characteristics of mineral production data. To overcome the shortcomings of the traditional techniques, data-driven models are used to predict the denoised and decomposed IMFs. In contrast, traditional techniques are used to predict the noise-free IMFs and residual terms. For the training model, 80% of the production data is used, and for the testing model, 20% of production data is used to test the accuracy of the proposed model. The description of the models which are used for prediction purposes is given below.

2.9. Prediction of Denoised IMFs Using Multilayer Perception Architecture (MLP). The neural network is a powerful technique to model complex nonlinear data. The multilayer perception architecture (MLP) model is an ANNs modeling approach widely used to model nonstationary time-series data. It belongs to a general structure called the feed-forward ANN model. This structure can deal with both continuous and integrable functions. The structure of MLP contains neurons that are grouped in layers.

In the MLP model, there is one layer for input nodes and one or more than one hidden layer. The structure of the MLP with a feed-forward network is exhibited in Figure 2.

The step-by-step process of the MLP network contains the following four parts:

- Step 1. Variable selection
- Step 2. Formulation of the training set, testing set, and validation set
- Step 3. Architecture
- Step 4. Verification of model and forecasting

The relationship between the output y_l layer and input $y_{l-1}, y_{l-2}, \dots, y_{l-v}$ layers has the following mathematical expression:

$$y_l = \varphi_o + \sum_{i=1}^u \varphi_i q \left(\sum_{j=1}^v \phi_{i,j} y_{l-1} + \phi_{oi} \right) + \varepsilon_l, \quad (17)$$

where φ_i ($i = 0, 1, 2, \dots, u$) and $\phi_{i,j}$ ($i = 0, 1, 2, \dots, u; j = 1, 2, 3, \dots, v$) are the model parameters, which are usually named as weights, u and v show the number of input and hidden nodes, respectively. Here, the MLP model executes the functional mapping of nonlinear types from past $y_{l-1}, y_{l-2}, \dots, y_{l-v}$ to future y_k observations.

$$y(l) = f(y_{l-1}, y_{l-2}, \dots, y_{l-v}, z) + \varepsilon_l. \quad (18)$$

The activation function which is usually used for activation is the logistic function, defined as follows:

$$h(y) = \frac{1}{1 + \exp(-\sum_{i=1}^u \varphi_i q (\sum_{j=1}^v \phi_{i,j} y_{l-1} + \phi_{oi}) + \varepsilon_l)}. \quad (19)$$

The back and forward propagation can be used by someone to optimize the neurons.

2.10. Prediction of Noise-free IMFs Using ARIMA Model. ARIMA model is selected to predict the residual term and noise-free IMFs whose description is given as follows:

$$\begin{aligned} IMF_x^l &= \psi_1 IMF_{x-1}^l + \dots + \psi_p IMF_{x-p}^l \\ &+ \varepsilon_p^l + \varphi_1 \varepsilon_{x-1}^l + \dots + \varphi_q \varepsilon_{x-q}^l, \end{aligned} \quad (20)$$

Here, IMF_x^l and ε_p^l that show the l^{th} IMF and residual term, which is obtained by CEEMDAN, p , and q , are the lag terms of the AMRA model. Suppose the time series is nonstationary and nonlinear. In that case, it is needed to make the difference to an appropriate degree, making the time-series stationary. If the difference is taken, then the model is called ARIMA (p, d, q), where d is used for the difference.

2.11. Multistep Ahead Prediction. Time-series prediction can be used for both single (one-step-ahead prediction) and multiple periods (multistep-ahead prediction). The multistep-ahead prediction has to deal with problems, such as accumulation error, uncertainty, and accuracy, unlike one-step-ahead prediction. However, accurate time-series prediction for long horizon has become challenging. A multistep-ahead time-series prediction consists of predicting the next H values of a time series consisting of N observations,

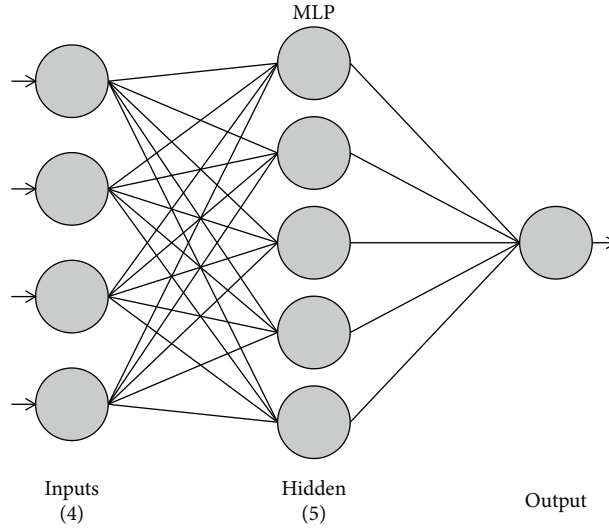


FIGURE 2: Structure of multilayer perceptron model.

where the forecasting horizon is denoted by $H > 1$ from the study of Sorjama et al. [54].

3. Application

In the current study, the production of mineral resources is considered to check the accuracy of the proposed model. Gas, oil, gypsum, and coal are among the principal mineral resources of Pakistan selected for the application of the proposed strategy. Pakistan is blessed with giant reserves of minerals covering 600,000 sq. Kms area. Out of the 92 minerals, 52 are exploited commercially, with 68.52 million metric tons per year. Our country has the second-largest coal deposits and billions of barrels of crude oil. However, the mineral sector's contribution towards the GDP of Pakistan is 3%, and exports are only 0.1% of the total of the world. Therefore, continuous development and planning towards the mineral sector are needed. Also, there is a need to analyze the problem and predict the accurate production of mineral resources of Pakistan.

3.1. Description of Data. Four major mineral resources are used in the current study to implement and investigate the proposed and improved framework. The production data comprises from 1st July to 30th June for the 2005–2019 period. The production of coal and gypsum is measured in metric tons, crude oil is measured in US barrels, and gas is measured in a million cubic feet. The used data are obtained from the Pakistan Bureau of Statistics. It consists of 168 monthly observations recorded from July 2005 to June 2019. The data are divided into training and testing data to observe the model performance. The data set contains 135 observations from July 2005 to April 2016, and the testing data set contains 33 observations from May 2016 to June 2019. The training data set consists of 80% observations of the observed series, and the testing data set includes 20% observations.

In a current study, several types of multistep-ahead prediction are conducted with L lag in the experiments. For

example, a kind of h -step-ahead prediction means predicting the production of mineral resources on the $(x + h)^{th}$ month with the h production samples before the x^{th} month but including the x^{th} month.

3.2. Evaluation Measures. After using the noise reduction techniques or thresholds, some distinctive approaches are needed to evaluate the denoised series of data performance. In the current study, the performance of the denoised and decomposed series, that is, EEMD-EBT and CEEMDAN-EBT, is observed based on three evaluation measures, that is, signal to noise reduction (SNR), mean absolute error (MAE), mean square error (MSE), and mean absolute percentage error (MAPE) from the study of Kim and Kim [55].

$$SNR = 10 \log_{10} \left(\frac{\sum_{x=1}^n (y(x)o)^2}{\sum_{x=1}^n (y(x)pred - y(x)o)^2} \right),$$

$$MAE = \frac{1}{n} \sum_{x=1}^n |y(x)_o - y(x)_{pred}|,$$

(21)

$$MSE = \frac{1}{n} \sum_{x=1}^n (y(x)_o - y(x)_{pred})^2,$$

$$MAPE = \frac{|y(x)_{m0} - y(x)_{m(pred)}|}{|y(x)_{m0}|} * 100,$$

where $y(x)_{pred}$ is the predicted series, $y(x)_o$ is the series of real data, and n is the data size also $y(x)_{m0}$ is the mean of the original series, $y(x)_{m(pred)}$ is the mean of the predicted series. Basically, MAE, MAPE, and MSE measure the deviation between the original values of the series and predicted values of the expected series. The performance of the proposed strategies, that is, EMD/WA-EEMD-EBT-MM and EMD/WA-CEEMDAN-EBT-MM, is evaluated using MAE, MSE, and MAPE described in equation (21).

4. Results and Discussion

This section presents the results of the proposed EMD/WA-EEMD-MM and EMD/WA-CEEMDAN-MM in comparison with other selected models.

4.1. Decomposition Stage. First, the augmented Dickey-Fuller (ADF) unit root test is used to confirm the non-stationarity of mineral resources data for all four minerals. The ADF unit root test results showed that the selected data of all four minerals are nonstationary with p values 0.2702, 0.0699, 1.4269, 1.1605 for gas, oil, and coal gypsum production, respectively. To decompose the data, EMD, WA, EEMD, and CEEMDAN decomposition methods are used to extract IMFs of production data of mineral resources. All four mineral production data are decomposed into six IMFs and one residual term. The first few IMFs showed high frequencies as compared to the last IMFs, and residual showed an overall trend. The WA-CEEMDAN-based decomposition results of gas and oil production are shown in Figure 3. The white noise magnitude is set to 0.2 from the study of Di et al. [24], and the number of ensemble members is fixed as 1000. According to Figure 3, it is noticed that IMF represents some fluctuations. Before proceeding to the next stage, the crosscorrelation method is used to find the IMFs with noises from all six IMFs. First, the decomposed IMFs are divided into two groups, that is, noisiest IMFs and noise-free IMFs, through crosscorrelation between IMFs and original data. The nature of correlation shows how much uncertainty exists in IMFs.

The low correlation indicates high uncertainty, and the high correlation shows less uncertainty in IMFs. The first five IMFs showed less correlation with original gas production data, which indicated the noisy IMFs. The graph of the crosscorrelation between the first IMF and fifth IMF and the original series of gas is shown in Figure 4. The cross correlation between these IMFs and gas production is very low at all lags indicating that these IMFs contain noises. The correlation graphs between both first IMF and sixth IMF with gas production are shown in part (A) and part (B) of Figure 4, respectively.

In plot (A), the starting IMF is full of noises with a very low correlation. In plot (B), it can be observed that it is noise-free, and there is a 0.60 correlation existing at lag zero and other lags for gas production. Therefore, for gas production data, the first five IMF components are labeled as noisiest, and the last two IMFs are characterized as noise-free IMFs and similarly for other minerals. For all four minerals, the first five IMFs are labeled as noisiest IMFs, and the last two IMFs are characterized as noise-free IMFs.

4.2. Denoising Stage. The noisiest IMFs obtained in the previous step are denoised in this step using different thresholds. The estimator of EBT is used to build the suggested model in which a mixture prior is assumed for each IMF as defined in (13) to remove noises from IMFs. The selection of mixture prior depends upon the nature of IMF. First, each IMF is transformed into normal distribution using scale transformation.

According to their nature, most of the coefficients are zero, and some are nonzero in IMFs, as shown in Figure 5; some are very high, and some are very low. By observing different parts of IMFs, the combination of probability at zero part of IMF and more than one distribution are contemplated for nonzero parts (Figure 5). After doing that, the Laplace distribution is considered as prior among all the priors of θ_i with a maximum value of SNR. Finally, the IMFs with the posterior median threshold estimator are selected to have the highest value of SNR and minimum MSE and MAE values. The values of SNR, MSE, and MAE for all minerals are presented in Table 1.

The ordinary denoising methods, that is, soft, hard thresholds, and ITF, are applied to all mineral production data for comparison. It is observed that the CEEMDAN-EBT-based decomposition and denoising method has the highest SNR value than CEEMDAN or EEMD-ST, CEEMDAN or EEMD-HT, and CEEMDAN or EEMD-ITF based denoising methods. For removing noises from data, the sparsity and noises are not considered by these methods separately except CEEMDAN-EBT based denoised method. The graphs of denoised methods CEEMDAN, EEMD-EBT, and other methods used for comparison are exhibited in Figures 6 and 7, respectively.

To understand the accomplishment of the suggested model, EEMD or CEEMDAN-EBT, the EEMD, and CEEMDAN occupied denoised and other denoising methods, that is, EEMD or CEEMDAN-HT, EEMD or CEEMDAN-ST, and EEMD or CEEMDAN-ITF models are used for comparison purposes. The evaluation of the suggested model and other models is measured using SNR, MSE, and MAE for the major four minerals resources. In Table 1, it can be observed clearly that from all decomposition and denoising methods, the performance of CEEMDAN-EBT is better than other models. Actually, CEEMDAN-EBT eliminates the noises efficiently by taking the mixture priors for IMFs, which attains the highest value of SNR and lowest values of MSE and MAE than CEEMDAN-HT, CEEMDAN-ST, and CEEMDAN-ITF techniques which have low values of SNR and high values of MSE. In addition, the decomposition and denoising technique based on the EEMD algorithm in combination with EBT, HT, ST, and ITF thresholds as compared to suggested CEEMDAN-EBT performed poorly for all four minerals. The poor performance of the EEMD-EBT is because of the mode mixing as shown in Figure 7, where the results of gas and oil production are plotted. Due to limitations in decomposition, the SNR values of all decomposition and denoising methods based on EEMD are low. The values of MSE and MAE are high for all four mineral production data, which shows that the CEEMDAN decomposition method has the capability of optimally extracting IMFs, which are denoised by optimal methods of denoising to obtain noise-free IMFs as presented in Figure 6; the results of denoising of all mineral production are plotted. From Table 1, it is concluded that the strategy of combining the CEEMDAN-based decomposition with the EBT denoising technique showed better results than other thresholding techniques. From Figures 6 and 7, it is shown that ST and HT combined with CEEMDAN overestimated the noises for gas and oil productions as these thresholds do not contemplate the

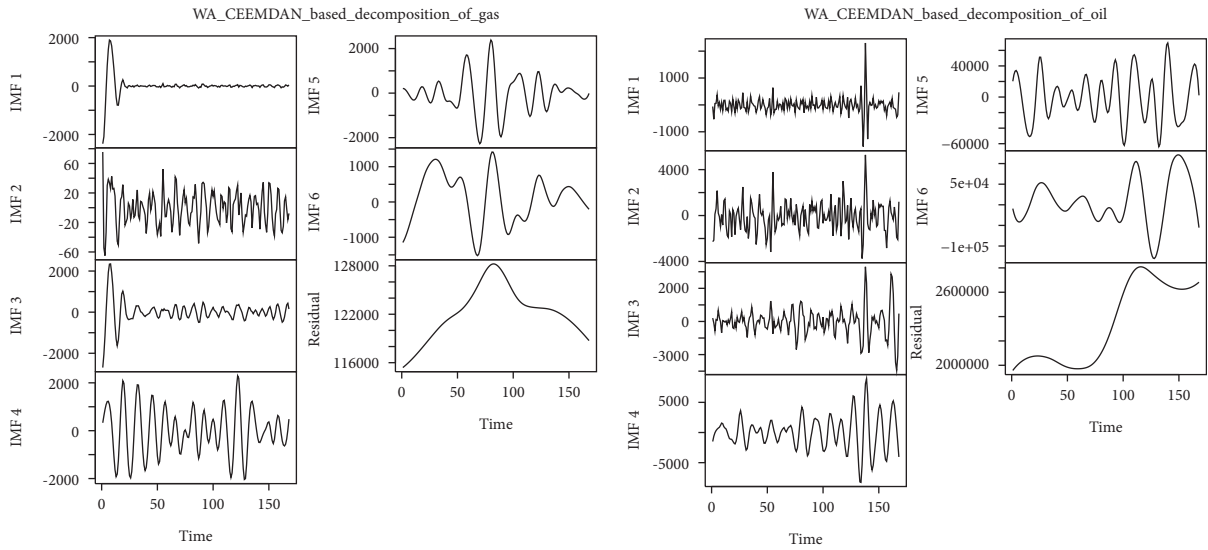


FIGURE 3: WA-CEEMDAN-based decomposition of gas and oil production. Both series are decomposed into six IMFs and one residual.

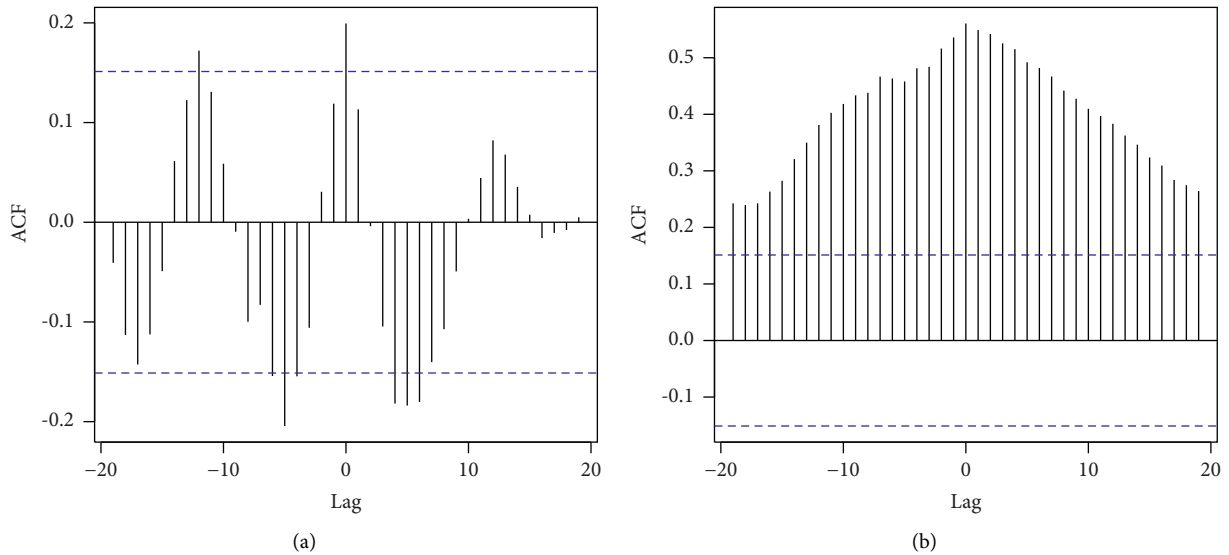


FIGURE 4: Crosscorrelation plot. Part (a) of the plot presents the correlation between the first IMF and gas production, and (b) part shows the correlation between the sixth IMF and gas production.

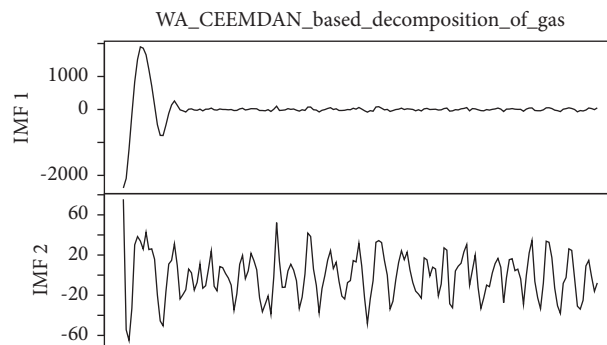


FIGURE 5: First two IMFs of WA-CEEMDAN for gas production.

TABLE 1: The numerical measures of suggested CEEMDAN-EBT and other denoising methods for four major minerals production.

| Mineral's production | Model | SNR | MSE | MAE |
|----------------------|-------------|--------------|--------------|-------------|
| Gas production | CEEMDAN-HT | -33.43 | 164.03 | 4.07 |
| | CEEMDAN-ST | -82.16 | 537323.50 | 629.80 |
| | CEEMDAN-ITF | -52.74 | 17922.28 | 112.60 |
| | CEEMDAN-EBT | 17.74 | 0.00 | 0.03 |
| | EEMD-HT | -39.30 | 1462.15 | 29.65 |
| | EEMD-ST | -54.38 | 18883.73 | 113.11 |
| | EEMD-ITF | -53.50 | 18873.04 | 113.19 |
| | EEMD-EBT | -61.81 | 206258.00 | 354.88 |
| Oil production | CEEMDAN-HT | 6.74 | 349.57 | 2.76 |
| | CEEMDAN-ST | -63.75 | 540139.50 | 591.92 |
| | CEEMDAN-ITF | -30.99 | 537953.40 | 588.10 |
| | CEEMDAN-EBT | 7.27 | 0.04 | 0.16 |
| | EEMD-HT | -79.00 | 12356570.00 | 2745.34 |
| | EEMD-ST | -86.93 | 86887096.00 | 7286.76 |
| | EEMD-ITF | -73.65 | 848831080.00 | 22770.34 |
| | EEMD-EBT | -97.94 | 859046156.00 | 22904.20 |
| Coal production | CEEMDAN-HT | -36.72 | 1651.05 | 9.87 |
| | CEEMDAN-ST | -70.71 | 324869.10 | 468.47 |
| | CEEMDAN-ITF | -64.89 | 316194.70 | 470.60 |
| | CEEMDAN-EBT | 18.00 | 0.00 | 0.03 |
| | EEMD-HT | -66.75 | 534782.20 | 570.28 |
| | EEMD-ST | -49.37 | 750473.30 | 677.51 |
| | EEMD-ITF | -53.98 | 484511.60 | 543.01 |
| Gypsum production | CEEMDAN-HT | -33.79 | 289.06 | 3.27 |
| | CEEMDAN-ST | -63.97 | 96132.27 | 251.26 |
| | CEEMDAN-ITF | -62.57 | 95933.80 | 251.12 |
| | CEEMDAN-EBT | 16.49 | 0.00 | 0.04 |
| | EEMD-HT | -78.28 | 21621484.00 | 3634.54 |
| | EEMD-ST | -76.52 | 21068132.00 | 3569.88 |
| | EEMD-ITF | -76.82 | 21071969.00 | 3570.38 |
| EEMD-EBT | -78.20 | 21623969.00 | 3633.90 | |

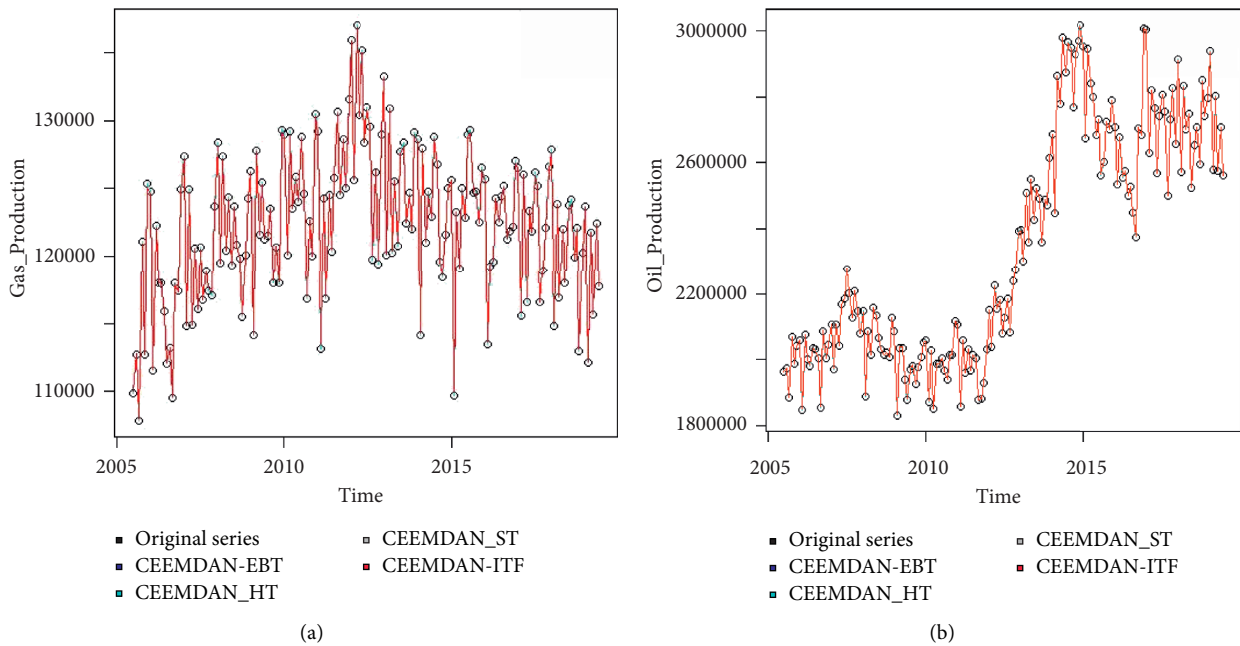


FIGURE 6: Decomposed and denoised series using CEEMDAN with suggested EBT and existing HT, ST, and ITF of (a) gas and (b) oil production.

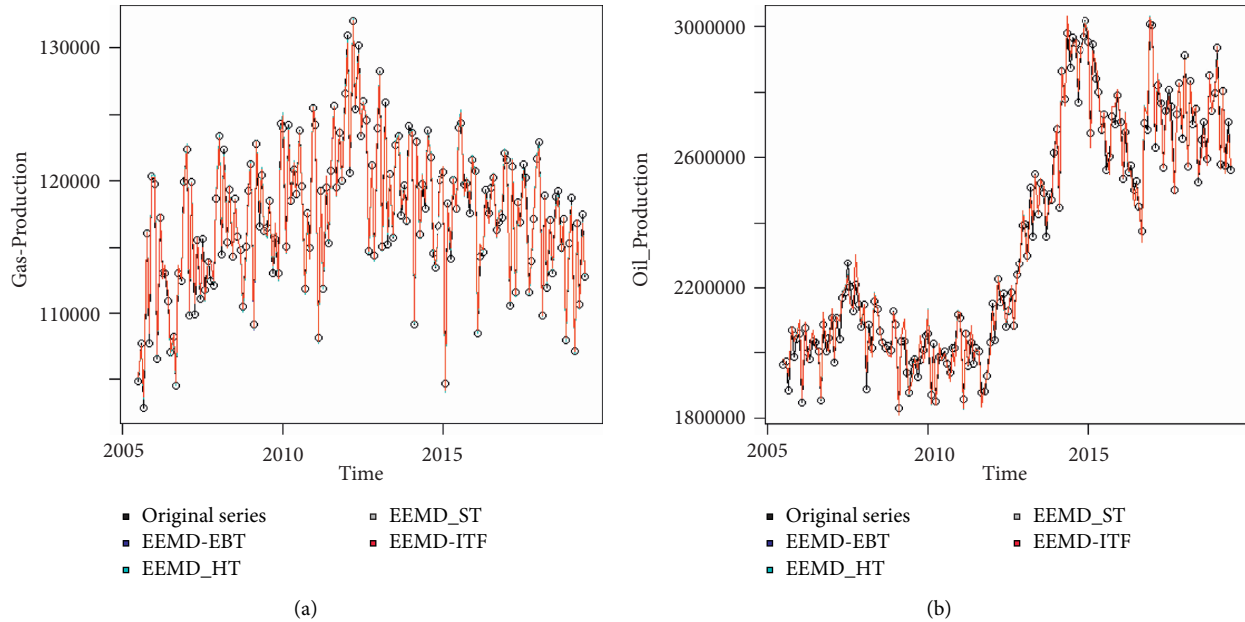


FIGURE 7: Decomposed and denoised series using EEMD with suggested EBT and existing HT, ST, and ITF of (a) gas and (b) oil production.

magnitude and sparsity of noises separately to eliminate the noises from data and ITF thresholds with CEEMDAN and EEMD worst performed for oil production data. However, the performance of the proposed CEEMDAN with EBT is optimal for the major four minerals.

4.3. Prediction Stage. At this stage, the IMFs obtained after denoising and decomposing are then predicted using traditional statistical and data-driven models. The prediction of denoised IMFs is obtained using the MLP-NN model. Training is carried out by utilizing forward propagation and back-propagation methods where the learning rate parameter is decided from 0.1 to 1. For the testing model, the appropriate learning rate of backpropagation is selected. The noise-free and denoised IMFs and residual are predicted using the ARIMA (p, d, q) model for all four major minerals. The mineral production data of all four minerals are divided into 70% for training and 30% for the testing set. The results of training errors after splitting the data are presented in Table 2. After estimating all IMFs and residuals, their accuracy is measured using MAE, MAPE, and MSE. The comparison of the proposed model with other existing models for four major mineral productions, that is, gas production, oil production, coal production, and gypsum production, is presented in Table 2. The proposed model, that is, WA-CEEMDAN-EBT-MM, illustrates its effectiveness by attaining the least values of MAE, MAPE, and MSE than other selected models for gas, oil, coal, and gypsum production. The proposed model demonstrated its productivity and efficiency for oil production by showing minimum values of MAPE and MSE compared to other models.

The predicted graphical view of WA-CEEMDAN-EBT-MM and WA-EEMD-EBT-MM with their benchmark models for gas, oil, and coal production is shown in Figures 8 and 9, respectively. For verification of the superiority of the proposed model by using the strategy WA-CEEMDAN-

EBT-MM for modeling mineral production data, we choose EMD-CEEMDAN-EBT-MM, WA or EMD-EEMD-EBT-MM, CEEMDAN or EEMD-EBT-MM, CEEMDAN or EEMD-HT-MM, CEEMDAN or EEMD-ST-MM, CEEMDAN or EEMD-ITF-MM model to find and analyze the results of prediction using nonlinear and noisy data.

Our proposed framework for prediction purposes based on the novel strategy of denoising and decomposition performs better than all other denoising and decomposition methods. It is evident from Table 2 and Figure 8 that the proposed model, that is, WA-CEEMDAN-EBT-MM, reveals good prediction results for oil, coal, and gypsum production with minimum values of MAE MSE and MAPE. The values of MAE and MAPE for gas production are also better than the selected models. The other selected models have no consistency in terms of efficiency except CEEMDAN-EBT-MM after WA-CEEMDAN-EBT-MM, as they vary in their behaviors during each mineral's prediction EMD-CEEMDAN-EBT-MM provides efficient results for oil and coal production but is not much efficient in predicting gas and gypsum production. However, the proposed model WA-CEEMDAN-EBT-MM provides excellent and consistent prediction results, illustrating that by using appropriate decomposition technique and novel denoising technique, one can get improvements in the performance of the data-driven model to handle the production of mineral resources.

To verify the dominance of our proposed strategy, that is, EMD and WA-CEEMDAN-EBT-MM, the testing data set of mineral production is also utilized. The results of the prediction error of the proposed and other benchmark models are exhibited in Table 3. It can be observed from Table 3 that the proposed methods performed better than other existing models, not only for training data set but also for testing data set because our proposed method attains the minimum value of MSE, MAD, and MAPE for all production of minerals. In

TABLE 2: Calculated index of estimating prediction error of proposed model (WA-CEEMDAN-EBT-MM) with all selected models for all minerals for training data set.

| Mineral production | Method | MAE | MAPE (%) | MSE |
|--------------------|--------------------|-----------------|-------------|--------------------|
| Gas production | CEEMDAN-HT-MM | 5304.52 | 0.26 | 12774831 |
| | CEEMDAN-ST-MM | 5317.78 | 0.24 | 11322218 |
| | CEEMDAN-ITF-MM | 5345.36 | 0.25 | 12102083 |
| | CEEMDAN-EBT-MM | 5032.89 | 0.19 | 7385481 |
| | EEMD-HT-MM | 5416.71 | 0.25 | 12017612 |
| | EEMD-ST-MM | 5485.98 | 0.24 | 11577010 |
| | EEMD-ITF-MM | 5300.14 | 0.24 | 11707157 |
| | EEMD-EBT-MM | 5544.93 | 0.23 | 10622160 |
| | EMD-CEEMDAN-EBT-MM | 5017.59 | 0.22 | 9093509 |
| | WA-CEEMDAN-EBT-MM | 4542.63 | 0.14 | 3985012 |
| Oil production | CEEMDAN-HT-MM | 113910.80 | 0.72 | 46194638775 |
| | CEEMDAN-ST-MM | 113765.70 | 0.74 | 48551504399 |
| | CEEMDAN-ITF-MM | 114119.50 | 0.83 | 61798819694 |
| | CEEMDAN-EBT-MM | 116116.80 | 0.71 | 44523217961 |
| | EEMD-HT-MM | 118991.10 | 0.73 | 47839442452 |
| | EEMD-ST-MM | 116942.50 | 0.73 | 46998880429 |
| | EEMD-ITF-MM | 103288.20 | 0.71 | 44903334225 |
| | EEMD-EBT-MM | 118982.80 | 0.73 | 47843257120 |
| | EMD-CEEMDAN-EBT-MM | 128277.10 | 0.66 | 39136609182 |
| | WA-CEEMDAN-EBT-MM | 89162.36 | 0.62 | 34516665678 |
| Coal production | EMD-EEMD-EBT-MM | 118733.20 | 0.73 | 47848016344 |
| | WA-EEMD-EBT-MM | 75571.16 | 0.71 | 44467200002 |
| | CEEMDAN-HT-MM | 79308.30 | 0.32 | 109961324 |
| | CEEMDAN-ST-MM | 76324.59 | 0.27 | 79147294 |
| | CEEMDAN-ITF-MM | 76124.60 | 0.28 | 85323250 |
| | CEEMDAN-EBT-MM | 75970.91 | 0.21 | 50089231 |
| | EEMD-HT-MM | 77767.61 | 0.34 | 127476519 |
| | EEMD-ST-MM | 77109.78 | 0.36 | 141105507 |
| | EEMD-ITF-MM | 77353.18 | 0.34 | 125934260 |
| | EEMD-EBT-MM | 79021.64 | 0.31 | 107708021 |
| | EMD-CEEMDAN-EBT-MM | 76677.20 | 0.16 | 28150475 |
| | WA-CEEMDAN-EBT-MM | 62097.61 | 0.02 | 631706 |
| | EMD-EEMD-EBT-MM | 78838.57 | 0.32 | 112101829 |
| | WA-EEMD-EBT-MM | 65352.05 | 0.20 | 44115854 |
| Gypsum production | CEEMDAN-HT-MM | 17597.64 | 4.66 | 2145766805 |
| | CEEMDAN-ST-MM | 17369.72 | 4.64 | 2127547391 |
| | CEEMDAN-ITF-MM | 17615.60 | 4.69 | 2171388807 |
| Mineral production | Method | MAE | MAPE | MSE |
| | CEEMDAN-EBT-MM | 15054.63 | 4.58 | 2070755942 |
| | EEMD-HT-MM | 18224.62 | 4.62 | 2111476775 |
| | EEMD-ST-MM | 17885.33 | 4.63 | 2116428672 |
| | EEMD-ITF-MM | 17754.59 | 4.65 | 2141721948 |
| | EEMD-EBT-MM | 19334.32 | 4.61 | 2130810674 |
| | EMD-CEEMDAN-EBT-MM | 15008.87 | 4.62 | 2104316567 |
| | WA-CEEMDAN-EBT-MM | 7359.75 | 4.43 | 1939268959 |
| | EMD-EEMD-EBT-MM | 17821.09 | 4.69 | 2175846815 |
| | WA-EEMD-EBT-MM | 12617.22 | 4.47 | 1976495272 |

Figure 10, all the models based on the EEMD strategy in the combination of ST, HT, ITF, and WA are exhibited. It can be inferred from Figure 10 that our proposed model (WA-CEEMDAN-EBT-MM) overcomes the problem of mode mixing, which provides better results of prediction for all minerals, that is, gas, oil, coal, and gypsum, as all the other models based on EEMD decomposition suffer from mode mixing problem.

4.4. Overall comparison of Proposed Model. Altogether, the proposed models behaved better than all other selected models as the proposed model attains the least values of MAE, MRE, and MSE for all four minerals. But WA-CEEMDAN-EBT-MM performs better among both proposed models by getting the lowest MAE, MRE, and MSE's compared to EMD-CEEMDAN-EBT-MM for all mineral productions listed in Tables 2 and 3. Furthermore, it is

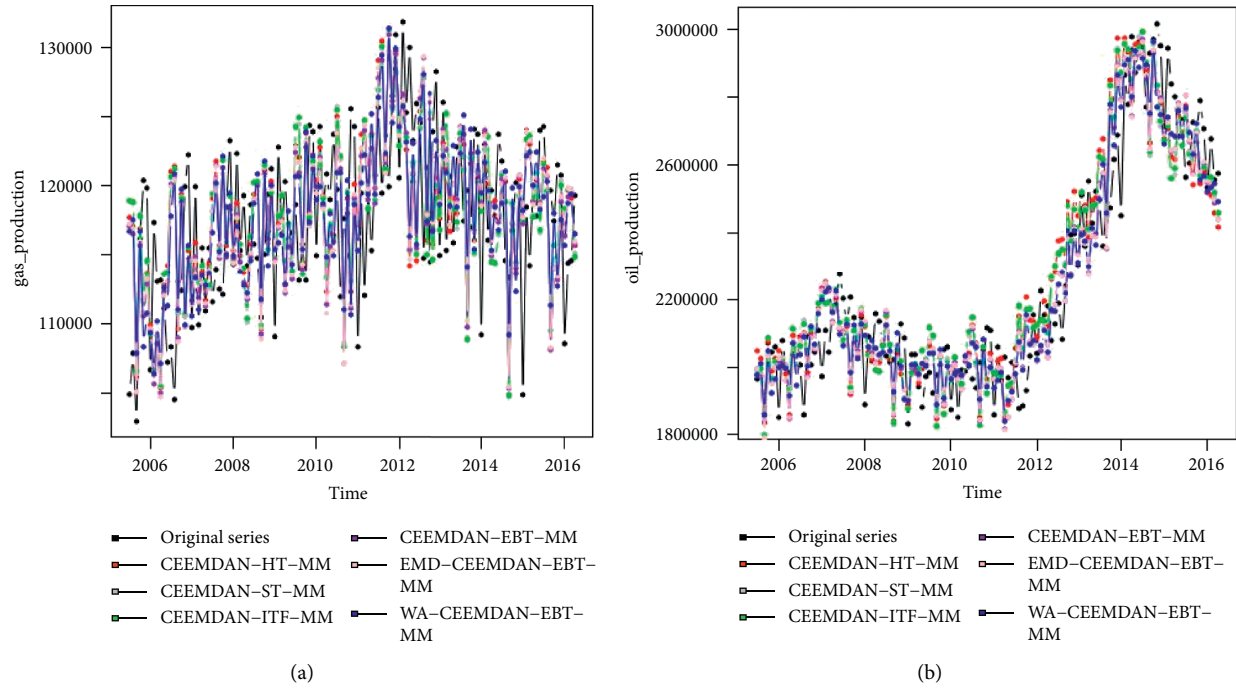


FIGURE 8: Prediction results of gas and oil production using proposed model EMD/WA-CEEMDAN-EBT-MM in comparison with other benchmark models.

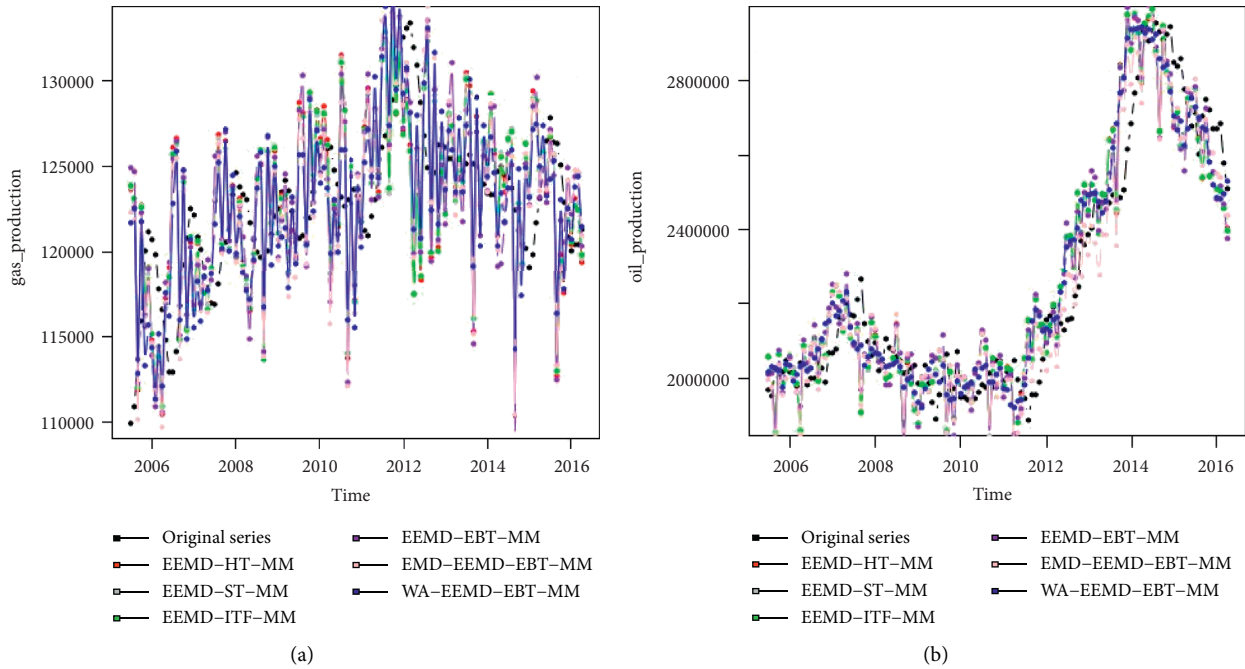


FIGURE 9: Prediction results of gas and oil production using EMD/WA-EEMD-EBT-MM in comparison with other benchmark models.

shown in these tables that in comparison with 1-stage, 2-stage, and 3-stage models, both proposed models perform well.

It is observed by considering the MAE from the quantitative analysis of the mineral resources data that WA-CEEMDAN-EBT-MM performed on average 32.5% better than CEEMDAN-HT-MM, 30.4% better than

CEEMDAN-ST-MM, 30.7% better than CEEMDAN-ITF-MM, 30% better than CEEMDAN-EBT-MM, 35.1% better than EEMD-HT-MM, 33.2% better than EEMD-ST-MM, 24.8% better than EEMD-ITF-MM, 36.6% better than EEMD-EBT-MM, 37.9% better than EMD-CEEMDAN-EBT-MM, 35.4% better than EMD-EEMD-EBT-MM, and 4.1% better than WA-EEMD-EBT-MM.

TABLE 3: Calculated index of the prediction error of proposed model (WA-CEEMDAN-EBT-MM) with all selected models for all minerals using testing data set.

| Mineral production | Method | MAE | MAPE (%) | MSE | |
|--------------------|--------------------|-----------------|-------------|-------------------|----------|
| Gas production | CEEMDAN-HT-MM | 6124.62 | 1.49 | 85776560 | |
| | CEEMDAN-ST-MM | 6067.94 | 1.49 | 86213206 | |
| | CEEMDAN-ITF-MM | 6060.91 | 1.49 | 85663915 | |
| | CEEMDAN-EBT-MM | 6497.42 | 1.45 | 80863903 | |
| | EEMD-HT-MM | 6299.46 | 1.52 | 89526223 | |
| | EEMD-ST-MM | 6162.96 | 1.49 | 85768731 | |
| | EEMD-ITF-MM | 6133.99 | 1.47 | 84174150 | |
| | EEMD-EBT-MM | 6196.67 | 1.47 | 83710228 | |
| | EMD-CEEMDAN-EBT-MM | 6087.81 | 1.47 | 83267636 | |
| | WA-CEEMDAN-EBT-MM | 5981.35 | 0.52 | 10644436 | |
| Oil production | EMD-EEMD-EBT-MM | 6498.35 | 1.39 | 75549936 | |
| | WA-EEMD-EBT-MM | 4656.62 | 1.32 | 67560768 | |
| | CEEMDAN-HT-MM | 94487.81 | 0.55 | 3574451945 | |
| | CEEMDAN-ST-MM | 89367.79 | 0.58 | 3950412921 | |
| | CEEMDAN-ITF-MM | 89112.90 | 0.59 | 4117443855 | |
| | CEEMDAN-EBT-MM | 95011.93 | 0.53 | 3331061007 | |
| | EEMD-HT-MM | 94687.77 | 0.55 | 3586516558 | |
| | EEMD-ST-MM | 94145.47 | 0.56 | 3718436914 | |
| | EEMD-ITF-MM | 94421.73 | 0.56 | 3763903042 | |
| | EEMD-EBT-MM | 97759.10 | 0.54 | 3526833910 | |
| Coal production | EMD-CEEMDAN-EBT-MM | 93675.74 | 0.54 | 3421733829 | |
| | WA-CEEMDAN-EBT-MM | 53672.63 | 0.01 | 1479480 | |
| | EMD-EEMD-EBT-MM | 97564.55 | 0.55 | 3582983333 | |
| | WA-EEMD-EBT-MM | 59518.42 | 0.59 | 4224621468 | |
| | CEEMDAN-HT-MM | 78961.34 | 4.74 | 6152379238 | |
| | CEEMDAN-ST-MM | 75085.79 | 4.78 | 6268716519 | |
| | CEEMDAN-ITF-MM | 75758.74 | 4.81 | 6351402799 | |
| | CEEMDAN-EBT-MM | 79648.01 | 4.58 | 5743522563 | |
| | EEMD-HT-MM | 79033.33 | 4.63 | 5881443482 | |
| | EEMD-ST-MM | 78581.31 | 4.65 | 5926195810 | |
| Gypsum production | EEMD-ITF-MM | 78326.25 | 4.63 | 5878805864 | |
| | EEMD-EBT-MM | 79137.63 | 4.65 | 6006845315 | |
| | EMD-CEEMDAN-EBT-MM | 78951.86 | 4.59 | 5782519245 | |
| | WA-CEEMDAN-EBT-MM | 46134.15 | 2.45 | 1627064349 | |
| | EMD-EEMD-EBT-MM | 78400.32 | 4.58 | 5862885160 | |
| | WA-EEMD-EBT-MM | 63512.19 | 4.86 | 6493517535 | |
| | Mineral production | Method | MAE | MAPE (%) | MSE |
| | Gypsum production | CEEMDAN-HT-MM | 12842.70 | 2.81 | 57920741 |
| | | CEEMDAN-ST-MM | 11996.57 | 3.51 | 90260488 |
| | | CEEMDAN-ITF-MM | 11809.15 | 3.46 | 87781588 |
| CEEMDAN-EBT-MM | | 12907.45 | 2.80 | 57304120 | |
| EEMD-HT-MM | | 13307.93 | 2.87 | 60794755 | |
| EEMD-ST-MM | | 13017.17 | 2.91 | 62552982 | |
| EEMD-ITF-MM | | 13016.39 | 2.89 | 61585436 | |
| EEMD-EBT-MM | | 13280.97 | 2.86 | 60331565 | |
| EMD-CEEMDAN-EBT-MM | | 12796.11 | 2.79 | 57183784 | |
| WA-CEEMDAN-EBT-MM | | 8215.01 | 0.02 | 4346.341 | |
| EMD-EEMD-EBT-MM | 13125.16 | 2.89 | 61419788 | | |
| WA-EEMD-EBT-MM | 9809.22 | 3.33 | 86638488 | | |

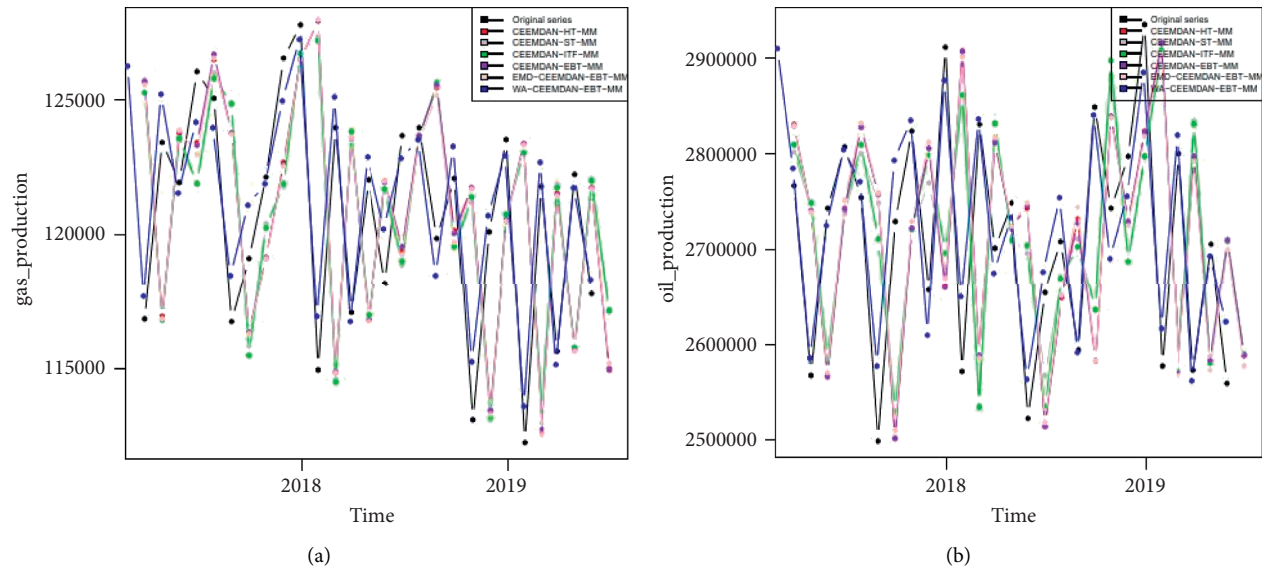


FIGURE 10: Prediction results of (a) gas and (b) oil production using EMD/WA-EEMD-EBT-MM in comparison with other benchmark models for testing data set.

Moreover, it is observed that many of the IMF components are predicted by using the ARIMA model, which shows that by using ARIMA in predicting IMFs, WA-EBT-CEEMDAN-MM provides more accurate prediction in mineral production.

5. Conclusion

For the optimal mineral supply of mineral resources, the accurate prediction of mineral resources is necessary. Here, some data processing methods are utilized to increase the prediction accuracy of such stochastic type data by using decomposition techniques to efficiently reduce the complexity of mineral production time-series data. Since the noises and nonlinearity of the mineral production data, a scheme is proposed here to improve the prediction accuracy of data-driven models with a suitable novel decomposition and denoising method. It is observed by considering the MAPE from the quantitative analysis of the mineral resources data that WA-CEEMDAN-EBT-MM performed on average 14.4% better than CEEMDAN-HT-MM, 13.1% better than CEEMDAN-ST-MM, 16.1% better than CEEMDAN-ITF-MM, 9.2% better than CEEMDAN-EBT-MM, 14% better than EEMD-HT-MM, 14.4% better than EEMD-ST-MM, 14% better than EEMD-ITF-MM, 12.9% better than EEMD-EBT-MM, 8.6% better than EMD-CEEMDAN-EBT-MM, 14.9% better than EMD-EEMD-EBT-MM, and 8.2% better than WA-EEMD-EBT-MM. Furthermore, our suggested method of denoising enhances the working of the decomposition method based on CEEMDAN through improvement in time-scale components which increases the prediction efficiency of the data-driven models. Our proposed method contained four stages: decomposition stage, denoising stage, prediction stage, and ensemble stage. The suggested model's performance, that is, WA-CEEMDAN-EBT-MM and EMD-CEEMDAN-EBT-

MM, is appraised using four major mineral production data. As a result, the WA-CEEMDAN-EBT-MM has the smallest value of MAPE for all four minerals compared to other models.

Data Availability

The secondary data set is used to validate the proposed models and will be provided on request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work under grant number (RGP.2/4/43), received by Mohammed M. Almazah (<http://www.kku.edu.sa>).

References

- [1] B. N. Birdsall and A. Hamoudi, vol. 7, 2002 W p n.
- [2] I. Aichouri, A. Hani, N. Bougherira, L. Djabri, H. Chaffai, and S. Lallahem, "river flow model using artificial neural networks," *Energy Procedia*, vol. 74, pp. 1007–1014, 2015.
- [3] S. Ch, N. Anand, B. K. Panigrahi, and S. Mathur, "Streamflow forecasting by SVM with quantum behaved particle swarm optimization," *Neurocomputing*, vol. 101, pp. 18–23, 2013.
- [4] C. Cheng, W. Niu, Z. Feng, J. Shen, and K. Chau, "Daily reservoir runoff forecasting method using artificial neural network based on quantum-behaved particle swarm optimization," *Water*, vol. 7, no. 8, pp. 4232–4246, 2015.
- [5] A. Lapedes and R. Farber, *Nonlinear Signal Processing Using Neural Networks: Prediction and System Modelling*, 1987.

- [6] D. P. Solomatine and A. Ostfeld, "Data-driven modelling: some past experiences and new approaches," *Journal of Hydroinformatics*, vol. 10, no. 1, pp. 3–22, 2008.
- [7] Z. Zaiyong Tang, C. De Almeida, and P. A. Fishwick, "Time series forecasting using neural networks vs. Box-Jenkins methodology," *SIMULATION*, vol. 57, no. 5, pp. 303–310, 1991.
- [8] G. Box, "Box and Jenkins: time series analysis, forecasting and control," in *A Very British Affair*, pp. 161–215, Palgrave Macmillan, London, UK, 2013.
- [9] S. Saadat, I. Hussain, and M. Faisal, "Modeling and forecasting of principal minerals production," *Arabian Journal of Geosciences*, vol. 14, no. 9, pp. 1–16, 2021.
- [10] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: a new learning scheme of feed-forward neural networks," in *Proceedings of the 2004 IEEE international joint conference on neural networks (IEEE Cat. No. 04CH37541)*, Budapest, Hungary, July 2004.
- [11] P. N. Suganthan and R. Katuwal, "Orig. randomization-based Feed. neural networks," *Applied Soft Computing*, vol. 105, Article ID 107239, 2021.
- [12] Z. M. Yaseen, S. O. Sulaiman, R. C. Deo, and K.-W. Chau, "An enhanced extreme learning machine model for river flow forecasting: state-of-the-art, practical applications in water resource engineering area and future research direction," *Journal of Hydrology*, vol. 569, pp. 387–408, 2019.
- [13] V. Nourani, A. Hosseini Baghanam, J. Adamowski, and O. Kisi, "Applications of hybrid wavelet-Artificial Intelligence models in hydrology: a review," *Journal of Hydrology*, vol. 514, pp. 358–377, 2014.
- [14] N. Pramanik, R. K. Panda, and A. Singh, "Daily river flow forecasting using wavelet ANN hybrid models," *Journal of Hydroinformatics*, vol. 13, no. 1, pp. 49–63, 2011.
- [15] 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.
- [16] K. L. Chong, S. H. Lai, and A. El-Shafie, "Wavelet transform based method for river stream flow time series frequency analysis and assessment in tropical environment," *Water Resources Management*, vol. 33, no. 6, pp. 2015–2032, 2019.
- [17] A. K. Tiwari, Z. Mukherjee, R. Gupta, and M. Balcilar, "A wavelet analysis of the relationship between oil and natural gas prices," *Resources Policy*, vol. 60, 2019.
- [18] V. Jayaswal, G. S. Gairola, and E. Chandrasekhar, "Identification of the typical frequency range associated with subsurface gas zones: a study using Hilbert-Huang transform and wavelet analysis," *Arabian Journal of Geosciences*, vol. 14, no. 5, pp. 1–10, 2021.
- [19] S. Singh, K. S. Parmar, J. Kumar, and S. J. S. Makkhan, "Development of new hybrid model of discrete wavelet decomposition and autoregressive integrated moving average (ARIMA) models in application to one month forecast the casualties cases of COVID-19," *Chaos, Solitons & Fractals*, vol. 135, Article ID 109866, 2020.
- [20] C. L. Wu, K. W. Chau, and Y. S. Li, "Predicting monthly stream-flow using data-driven models coupled with data-preprocessing techniques," *Water Resources Research*, vol. 45, no. 8, 2009.
- [21] A. Azadeh, M. Sheikhalishahi, M. Tabesh, and A. Negahban, "The effects of pre-processing methods on forecasting improvement of artificial neural networks," *Aust. J. Basic Appl. Sci.*, vol. 5, no. 6, pp. 570–580, 2011.
- [22] S. Asadi, J. Shahrabi, P. Abbaszadeh, and S. Tabanmehr, "A new hybrid artificial neural networks for rainfall-runoff process modeling," *Neurocomputing*, vol. 121, pp. 470–480, 2013.
- [23] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. London. Ser. A Math. Phys. Eng. Sci.*, vol. 454, 1998.
- [24] C. Di, X. Yang, and X. Wang, "A four-stage hybrid model for hydrological time series forecasting," *PLoS One*, 2014.
- [25] S. Huang, J. Chang, Q. Huang, and Y. Chen, "Monthly streamflow prediction using modified EMD-based support vector machine," *Journal of Hydrology*, vol. 511, pp. 764–775, 2014.
- [26] W. Liu, K. He, Q. Gao, and C. Liu, "Application of EMD-based SVD and SVM to coal-gangue interface detection," *Journal of Applied Mathematics*, vol. 2014, 2014.
- [27] Z.-Y. Wang, J. Qiu, and F.-F. Li, "Hybrid models combining EMD/EEMD and ARIMA for long-term streamflow forecasting," *Water*, vol. 10, no. 7, p. 853, 2018.
- [28] H. Liu, H.-Q. Tian, C. Chen, and Y.-f. Li, "A hybrid statistical method to predict wind speed and wind power," *Renewable Energy*, vol. 35, no. 8, pp. 1857–1861, 2010.
- [29] X. Mi, H. Liu, and Y. Li, "Wind speed prediction model using singular spectrum analysis, empirical mode decomposition and convolutional support vector machine," *Energy Conversion and Management*, vol. 180, pp. 196–205, 2019.
- [30] J. J. Ruiz-Aguilar, I. Turias, J. González-Enrique, D. Urda, and D. Elizondo, "A permutation entropy-based EMD-ANN forecasting ensemble approach for wind speed prediction," *Neural Computing & Applications*, vol. 33, no. 7, pp. 2369–2391, 2021.
- [31] W.-j. Niu, Z.-k. Feng, M. Zeng et al., "Forecasting reservoir monthly runoff via ensemble empirical mode decomposition and extreme learning machine optimized by an improved gravitational search algorithm," *Applied Soft Computing*, vol. 82, Article ID 105589, 2019.
- [32] M. Santhosh, C. Venkaiah, and D. M. Vinod Kumar, "Ensemble empirical mode decomposition based adaptive wavelet neural network method for wind speed prediction," *Energy Conversion and Management*, vol. 168, pp. 482–493, 2018.
- [33] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *Proceedings of the ICASSP, IEEE International Conference on Acoustics*, Prague, Czech, May 2011.
- [34] L. Chaabi, A. Lemzadmi, A. Djebala, M. L. Bouhalais, and N. Ouelaa, "Fault diagnosis of rolling bearings in non-stationary running conditions using improved CEEMDAN and multivariate denoising based on wavelet and principal component analyses," *International Journal of Advanced Manufacturing Technology*, vol. 107, no. 9, pp. 3859–3873, 2020.
- [35] W. Jun, L. Yuyan, T. Lingyu, and G. Peng, "A new weighted CEEMDAN-based prediction model: an experimental investigation of decomposition and non-decomposition approaches," *Knowledge-Based Systems*, vol. 160, pp. 188–199, 2018.
- [36] W. Jiang and J. Zhou, "Analysis of Hydro-Turbine Non-stationary Vibration Signals Based on CEEMDAN and Wigner-Ville Distribution," in *Proceedings of the 2017 5th International Conference on Frontiers of Manufacturing*

- Science and Measuring Technology (FMSMT 2017)*, Taiyuan, China, April 2017.
- [37] S. Dai, D. Niu, and Y. Li, "Daily peak load forecasting based on complete ensemble empirical mode decomposition with adaptive noise and support vector machine optimized by modified grey wolf optimization algorithm," *Energies*, vol. 11, no. 1, p. 163, 2018.
- [38] Y. Wang, C. Xu, Y. Wang, and X. Cheng, "A comprehensive diagnosis method of rolling bearing fault based on CEEM-DAN-DFA-improved wavelet threshold function and QPSO-MPE-SVM," *Entropy*, vol. 23, no. 9, p. 1142, 2021.
- [39] M. Qurban, X. Zhang, H. Mamona Nazir et al., "Development of Hybrid Methods for Prediction of Principal Mineral Resources," *Mathematical Problems in Engineering*, vol. 2021, Article ID 6362660, 2021.
- [40] I. M. Johnstone and B. W. Silverman, "Empirical Bayes selection of wavelet thresholds," *Annals of Statistics*, vol. 33, no. 4, pp. 1700–1752, 2005.
- [41] H. M. Nazir, I. Hussain, I. Ahmad, M. Faisal, and I. M. Almanjahie, "An improved framework to predict river flow time series data," *PeerJ*, vol. 2019, 2019.
- [42] T. Li, Y. Zhou, X. Li, J. Wu, and T. He, "Forecasting daily crude oil prices using improved CEEMDAN and ridge regression-based predictors," *Energies*, vol. 12, no. 19, 2019.
- [43] H.-F. Yang and Y.-P. P. Chen, "Hybrid deep learning and empirical mode decomposition model for time series applications," *Expert Systems with Applications*, vol. 120, pp. 128–138, 2019.
- [44] J. Zhang, R. Jiang, B. Li, and N. Xu, "An automatic recognition method of microseismic signals based on EEMD-SVD and ELM," *Computers & Geosciences*, vol. 133, Article ID 104318, 2019.
- [45] X. Wei, R. Lin, S. Liu, and C. Zhang, "Improved EEMD Denoising Method Based on Singular Value Decomposition for the Chaotic Signal," *The Use of Vibration Signals for Structural Health Monitoring, System Identification, Test Planning/Optimization, and Dynamic Model Validation/Updating*, vol. 2016, 2016.
- [46] P. Flandrin, E. Torres, and M. A. Colominas, *A COMPLETE ENSEMBLE EMPIRICAL MODE DECOMPOSITION WITH ADAPTIVE NOISE* "Laboratorio de Señales y Dinámicas no Lineales", pp. 4144–4147, Universidad Nacional de Entre Ríos, Laboratoire de Physique (UMR CNRS 5672), Ecole Normale Supérieure de Lyon, France, 2011.
- [47] Z. Qian, Y. Pei, H. Zareipour, and N. Chen, "A review and discussion of decomposition-based hybrid models for wind energy forecasting applications," *Applied Energy*, vol. 235, pp. 939–953, 2019.
- [48] X. Zhang, K. K. Lai, and S.-Y. Wang, "A new approach for crude oil price analysis based on empirical mode decomposition," *Energy Economics*, vol. 30, no. 3, pp. 905–918, 2008.
- [49] M. M. Hossain, T. J. Kozubowski, and K. Podgórski, "A novel weighted likelihood estimation with empirical Bayes flavor," *Communications in Statistics - Simulation and Computation*, vol. 47, no. 2, pp. 392–412, 2018.
- [50] C. Y. Sg and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Transactions on Image Processing*, vol. 9, no. 9, pp. 1532–1546, 2000.
- [51] M. Jansen and A. Bultheel, "Empirical bayes approach to improve wavelet thresholding for image noise reduction," *Journal of the American Statistical Association*, vol. 96, no. 454, pp. 629–639, 2001.
- [52] Y. Jeng, M.-J. Lin, C.-S. Chen, and Y.-H. Wang, "Noise reduction and data recovery for a VLF-EM survey using a nonlinear decomposition method," *Geophysics*, vol. 72, no. 5, pp. F223–F235, 2007.
- [53] H. Om and M. Biswas, "2012, An improved image denoising method based on wavelet thresholding," *Journal of Signal and Information Processing*, vol. 3, no. 1, Article ID 17686, 2012.
- [54] A. Sorjamaa, J. Hao, N. Reyhani, Y. Ji, and A. Lendasse, "Methodology for long-term prediction of time series," *Neurocomputing*, vol. 70, no. 16–18, pp. 2861–2869, 2007.
- [55] S. Kim and H. Kim, "A new metric of absolute percentage error for intermittent demand forecasts," *International Journal of Forecasting*, vol. 32, no. 3, pp. 669–679, 2016.